scientific reports

OPEN

Molecular mechanisms and therapeutic targets in glioblastoma multiforme: network and single-cell analyses

Xiangyu Chen[1,5], Xiao Zhong[1,5], Feifei Zhang[2], Xiaomei Zhou[3], Xiaofeng Yue[4✉] & Xueru Li[1✉]

Glioblastoma multiforme (GBM) is a highly aggressive brain tumor associated with poor survival outcomes and is driven by a complex tumor microenvironment (TME) that promotes tumor progression and treatment resistance. To explore the role of the TME in GBM, we analyzed glioma-related microarray and single-cell RNA sequencing (scRNA-seq) datasets from the Gene Expression Omnibus (GEO). Functional enrichment and weighted gene coexpression network analyses revealed distinct immune profiles, metabolic alterations, and differences in chemotherapeutic drug sensitivity between the high-risk and low-risk patient groups. scRNA-seq data processed with the 'Seurat' package were used to identify differentially expressed genes in pericytes, endothelial cells, and glioma cells, particularly those involved in extracellular matrix (ECM) remodeling. A 17-gene prognostic signature developed through Cox regression and LASSO analyses revealed that key genes (*COL1A1*, *COL4A1*, and *VIM*) were significantly associated with survival outcomes in GBM patients. Drug sensitivity analyses using data from the Genomics of Drug Sensitivity in Cancer (GDSC) and Cancer Therapeutics Response Portal (CTRP) identified potential targeted therapies for GBM, including SB-505,124, staurosporine, and AZD8186. This integrative study underscores the critical roles of the ECM and synaptic remodeling in GBM and suggests novel therapeutic targets to improve personalized treatment strategies for GBM patients.

Keywords Glioblastoma multiforme, Weighted gene coexpression network analysis, Single-cell RNA sequencing, Molecular subtype, Prognosis, Tumor microenvironment, Extracellular matrix

Glioblastoma multiforme (GBM) is the most common and lethal brain tumor, with a median survival of only 16–18 months[1]. Conventional treatments, including surgery, radiotherapy, and chemotherapy[2,3], offer limited efficacy because of therapy resistance driven by glioma stem cells and the tumor microenvironment (TME)[4,5]. Recent studies have highlighted the potential of novel approaches, such as targeted therapies and combined chemoimmunotherapy, to improve outcomes[6,7]. However, successful treatment requires a deeper understanding of the intricate interactions within the TME.

Microenvironmental dependencies significantly modulate the behavior of cancer and stromal cells and regulate tumor progression, the chemotherapy response, and inflammatory cell infiltration[8,9]. A hallmark of tumor progression induced by these microenvironmental factors is aberrant deposition of the extracellular matrix (ECM)[10,11]. The ECM has various functions, including shaping stem cell niches[12], regulating ECM-dependent survival signaling cascades[8], and modulating the chemotherapy response[13,14]. Stromal cells, including fibroblasts, endothelial cells, and pericytes, are the major components of the ECM[15]. Previous studies have highlighted the importance of collagen, a key ECM protein widely expressed in various tissues, including the brain vasculature. Collagen is closely associated with basement membranes and influences tumor growth[16]. Additionally, collagen is reported to be upregulated in GBM and high-grade gliomas. Collagen upregulation in GBM and high-grade gliomas, but not in low-grade astrocytoma and healthy glial cells, is correlated with poor clinical outcomes[17–19].

[1]Key Laboratory of Major Brain Disease and Aging Research (Ministry of Education), Institute for Brain Science and Disease, Chongqing Medical University, Chongqing 400016, China. [2]Department of Blood Transfusion, School of Medicine, Sichuan Provincial People's Hospital, University of Electronic Science and Technology of China, Chengdu 610021, Sichuan, China. [3]Sichuan Provincial Chengdu Second People's Hospital, Chengdu 610021, Sichuan, China. [4]Department of Urology, The Third Affiliated Hospital of Chongqing Medical University, Chongqing 401120, China. [5]Xiangyu Chen and Xiao Zhong contributed equally to this work. ✉email: 652282@hospital.cqmu.edu.cn; 2020130299@stu.cqmu.edu.cn

In addition to ECM-related factors, neurons contribute to the glioma microenvironment by forming synaptic connections with tumor cells[20–22]. These glioma cells use neurotransmitter signals, such as glutamate, for their growth and invasive behavior[23]. This integration into neural circuits can lead to dynamic remodeling and lasting behavioral changes in tumors[24]. Understanding these interactions is crucial for identifying new therapeutic targets to disrupt protumorigenic networks.

Weighted gene coexpression network analysis (WGCNA), a robust statistical method for analyzing gene coexpression networks, is reported to be effective in identifying hub genes across various biological contexts[25]. For example, WGCNA of microarray and The Cancer Genome Atlas (TCGA) datasets revealed core genes associated with poor prognosis in GBM, providing useful insights into GBM prognosis and progression[26]. However, few studies have used WGCNA to examine the TME of GBM. Single-cell RNA sequencing (scRNA-seq) provides comprehensive insights into cellular transcriptomic heterogeneity and reveals the underlying gene expression patterns[27]. Integrating these methods enables precise characterization of the GBM microenvironment, offering insights into therapy resistance and patient prognosis.

This study systematically analyzed microarray and scRNA-seq data to construct a prognostic model for glioma patients. Key ECM-related genes, including *COL1A1*, *COL4A1*, and *VIM*, were upregulated in GBM tissues compared with low-grade tumors. Additionally, 16 small molecules that target ECM remodeling were predicted via molecular simulations. These findings highlight novel therapeutic opportunities to improve GBM management and patient outcomes.

Results
WGCNA and functional enrichment analysis of the hub genes in glioma

To investigate the correlation between gene expression and glioma progression, datasets GSE4290, GSE108476, and GSE15824 were analyzed using weighted gene co-expression network analysis (WGCNA). A soft-threshold power of 9 was selected based on the scale-free topology fit index ($R^2 = 0.9$), as determined via the pickSoftThreshold function in WGCNA (Supplementary Fig. S1A-B). Modules with similar expression profiles were merged using the dynamic tree cut method with a merge threshold of 0.25 (Fig. 1A). Subsequently, the adjacency matrix and topological overlap matrix (TOM) were constructed to define gene co-expression networks. The dynamic tree cut approach identified nine modules, including seven modules significantly associated with glioma progression, and a gray module representing non-coexpressed genes (Fig. 1B-C, Supplementary Fig. S1C-D). Among the seven glioma-associated modules, four modules exhibited positive correlations, whereas three modules displayed negative correlations with glioma progression (Fig. 1C). Functional enrichment analysis of hub genes within these modules was performed using the "Cluster Profiler" R package (Fig. 1D-E, Supplementary Fig. S2A). The red module was primarily enriched in biological processes such as glial cell differentiation, myelination, and neuron ensheathment (Supplementary Fig. S2B), indicating its role in neural tissue development and maintenance. Similarly, the brown module hub genes were significantly enriched in synapse organization and glutamatergic synapse processes (Supplementary Fig. S2C), reflecting their contributions to synaptic function and neurotransmission. In the turquoise module, enrichment was observed in processes such as the modulation of chemical synaptic transmission and calcium signaling pathways (Fig. 1F-G), further underscoring the module's involvement in neural communication and physiological brain function. Conversely, the hub genes of the black module were enriched in cell cycle-related processes, including chromosome segregation, nuclear division, and mitotic nuclear division. KEGG pathway analysis revealed significant enrichment in the p53 signaling pathway (Supplementary Fig. S2D-E), highlighting its regulatory role in glioma cell proliferation. Hub genes in the green-yellow module were associated with antiviral responses and innate immune regulation, with KEGG pathways enriched in Epstein-Barr virus infection, NOD-like receptor signaling, and measles (Supplementary Fig. S2F-G), suggesting involvement in immune system processes and pathogen response. The blue module was strongly associated with extracellular matrix (ECM) organization, and KEGG pathway analysis identified enrichment in ECM-receptor interactions (Fig. 1H-I), supporting its role in ECM remodeling during glioma progression. Finally, hub genes in the green module were enriched in immune activation processes, including Toll-like receptor signaling and NF-kappa B signaling pathways (Supplementary Fig. S2H-I), emphasizing its role in immune activation and inflammation in glioma.

Prognostic value of glioma module subnetworks: PPI network analysis and correlation with the response to chemoradiotherapy

The genes from different modules were imported into the STRING database to establish a PPI network, which was then visualized via Cytoscape. Next, the MCODE plugin was used to identify highly interconnected subnetworks (Supplementary Fig. S3A-G).To explore the correlation between module gene subnetworks and glioma, ssGSEA[28] was performed to calculate the enrichment scores for these seven module subnetworks among patients with LGG and those with GBM in the TCGA, CGGA325, and CGGA693 datasets. Patients with high enrichment scores in the red, brown, and turquoise module subnetworks presented increased survival rates and decreased HRs. Conversely, patients with high enrichment scores in the blue, black, green, and green-yellow module subnetworks presented poor survival rates and increased HRs (Supplementary Table 1).

To understand the relationship between module subnetworks and glioma progression, predefined tumor progression gene sets were used[29]. The turquoise, red, and brown module subnetworks were negatively correlated with tumor progression, whereas the blue, black, green, and green-yellow module subnetworks were positively correlated with tumor progression (Fig. 2A, Supplementary Fig. S4A-C).This study also evaluated the expression of module subnetwork genes in patients with gliomas after chemoradiotherapy. The expression of genes in the turquoise, red, and brown module subnetworks was upregulated, whereas that of genes in the blue, black, green, and green-yellow module subnetworks was downregulated in patients who responded to treatment (Fig. 2B, Supplementary Fig. S5A-C). ROC curve analysis demonstrated that all module subnetworks had good predictive
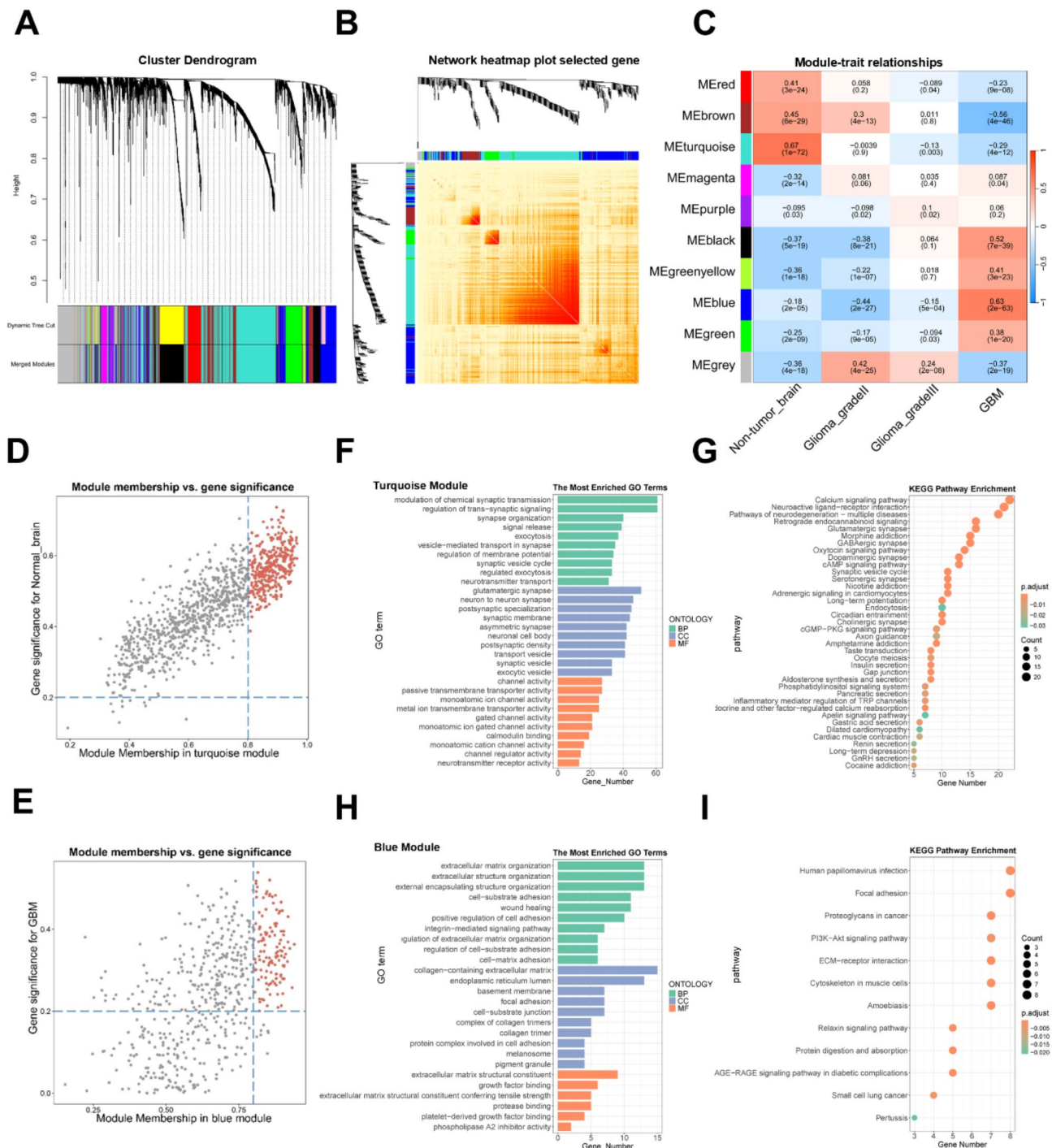
Fig. 1. Identification of module genes in glioma and non-tumorous samples using weighted gene coexpression network analysis (WGCNA). (A) Cluster dendrogram of the coexpression network modules. (B) A coexpression network was visualized via heatmaps. (C) Module-trait correlation heatmap. (D-E) Scatter plots of module membership (MM) values vs. gene significance (GS) values for the turquoise (D) and blue modules (E). Genes with MM values > 0.8 and GS values > 0.2 are colored. (F-I) Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses of the hub genes in the turquoise (F-G) and blue (H-I) modules. Hub genes are defined as those with MM values > 0.8 and GS values > 0.2.
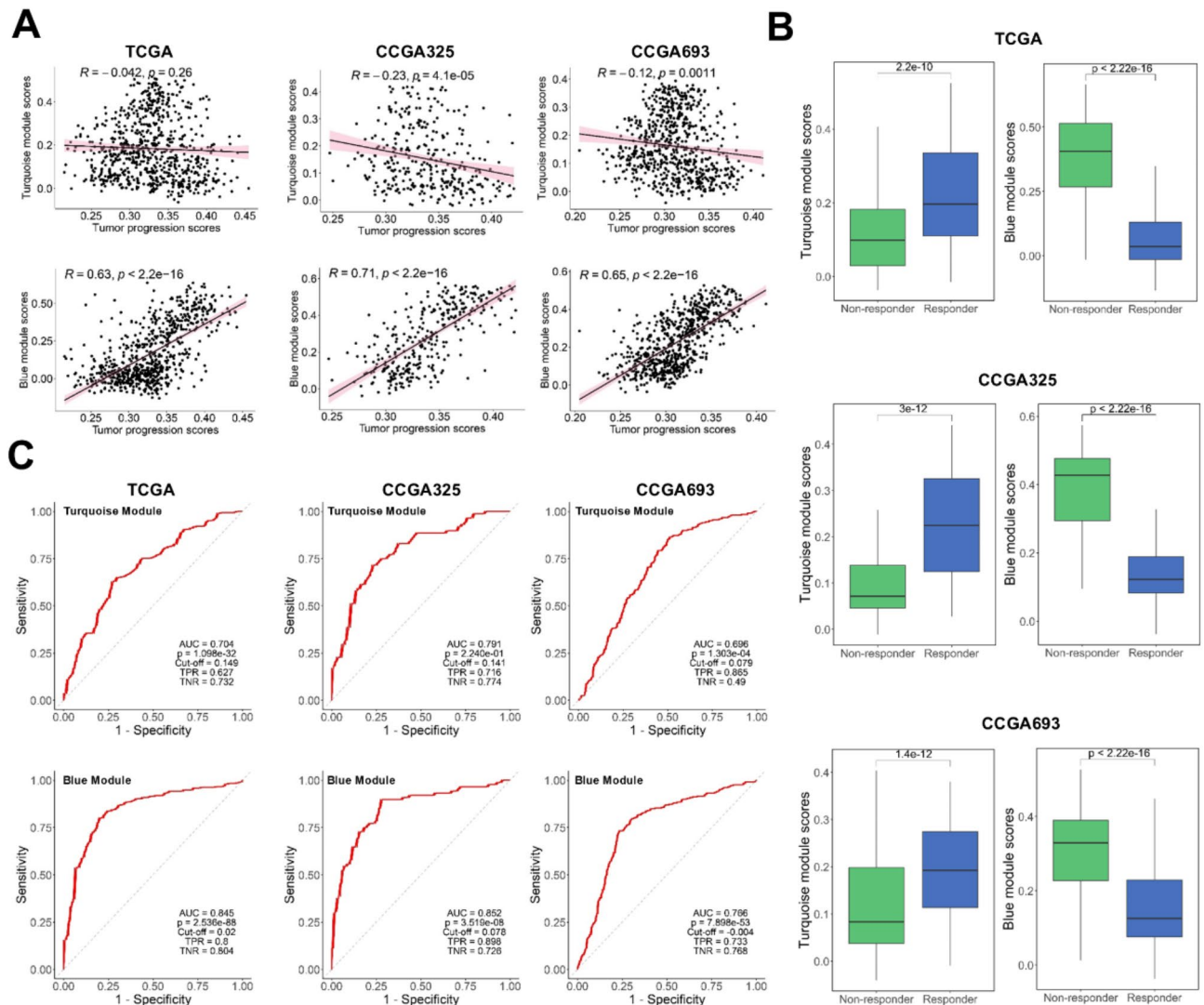
Fig. 2. Analysis of the correlation between module subnetworks and tumor progression. (**A**) Analysis of the correlation of the subnetwork scores of the turquoise and blue modules with tumor progression scores across The Cancer Genome Atlas (TCGA) and Chinese Glioma Genome Atlas (CGGA) datasets (CGGA325 and CGGA693). The top row presents the correlation between the turquoise module subnetwork scores and tumor progression scores. The bottom row shows the correlation between the blue module subnetwork scores and tumor progression scores. (**B**) Turquoise and blue module subnetwork scores in patients who responded (responders) and did not respond (non-responders) to chemoradiotherapy across the TCGA and CGGA datasets (CGGA325 and CGGA693). Patients who survived more than 14.2 months after treatment were classified as responders, whereas those who died within 14.2 months after treatment were classified as non-responders. (**C**) Receiving operating characteristic (ROC) curves for predicting the response to chemoradiotherapy via subnetwork scores of the turquoise and blue modules across the TCGA and CGGA datasets (CGGA325 and CGGA693).

power for chemoradiotherapy outcomes. The blue and black module subnetworks presented the highest AUC values and decreased p values in the TCGA and CGGA datasets (Fig. 2C, Supplementary Fig. S6A-C). Thus, the subnetwork genes selected from the seven modules effectively represent tumor progression and treatment prognosis. This comprehensive analysis highlights the distinct roles of each module in glioma progression and therapy response, providing valuable insights for future research and potential therapeutic targets.

Divergent roles of the blue and Turquoise module subnetworks in the GBM immune landscape and metabolism

The blue and turquoise modules, which were most strongly correlated with GBM, were selected for further analysis. Genes in the blue module subnetwork were upregulated in GBM samples, whereas those in the

turquoise module subnetwork were upregulated in non-tumorous samples (Fig. 3A, Supplementary Fig. S7A). GO and KEGG analyses revealed that blue module genes were enriched mainly in ECM remodeling and ECM-receptor interactions, highlighting the critical role of the ECM in glioma progression (Fig. 3B-C). In contrast, the turquoise module genes were enriched in the modulation of synaptic transmission and the synaptic vesicle cycle, suggesting their involvement in neuronal processes in LGG (Supplementary Fig. S7B-C).

To comprehensively analyze the correlation of the blue and turquoise module subnetworks with immune infiltration, ssGSEA was used to calculate enrichment scores for each sample. The CIBERSORT algorithm[30] was used to predict immune infiltration in groups with high and low enrichment scores for the turquoise and blue subnetworks. Notably, the proportions of 22 immune cell types varied significantly between these groups (Fig. 3D, Supplementary Fig. S7D).

In patients with high enrichment scores for the blue module subnetwork, the proportions of CD8 T cells, T follicular helper cells, regulatory T cells (Tregs), resting NK cells, M0 macrophages, M1 macrophages, M2 macrophages, and neutrophils were increased (Fig. 3E). The numbers of myeloid-derived suppressor cells (MDSCs) and exhausted T cells were also increased in these patients (Supplementary Fig. S7F-G). As shown in Supplementary Fig. S8A. The proportions of M2 macrophages, neutrophils, resting NK cells, and Tregs were negatively correlated with those of naïve B cells, plasma cells, naïve CD4 T cells, and activated NK cells. Correlation analysis also revealed that genes in the blue module subnetwork were positively correlated with immunosuppressive cells (Tregs, M2 macrophages, and neutrophils) and negatively correlated with immune-activating cells (plasma cells and activated NK cells) (Fig. 3F). These findings indicate that the blue module subnetwork promoted glioma progression through immune regulation.

Conversely, patients with high enrichment scores for the turquoise module subnetwork exhibited increased infiltration of naïve B cells, plasma cells, naïve CD4 T cells, activated natural killer (NK) cells, activated mast cells, and eosinophils (Supplementary Fig. S7E). These findings are supported by the data shown in Supplementary Fig. S8B. The proportions of naïve B cells were positively correlated with those of naïve CD4 T cells, activated NK cells, and plasma cells. Moreover, the proportions of CD8 T cells were positively correlated with those of resting dendritic cells and M1 macrophages. Correlation analysis revealed that 31 genes in the turquoise module subnetwork were significantly and positively correlated with the proportions of naïve B cells, plasma cells, naïve CD4 T cells, activated NK cells, activated mast cells, and eosinophils (Supplementary Fig. S7H). These findings indicate that the upregulation of these genes activates immune responses to inhibit glioma progression.

Further analysis via the "IBOR" R package revealed distinct metabolic pathway differences between the blue and turquoise subnetworks. In the blue subnetwork, pathways related to glycogen degradation, folate one-carbon metabolism, and oxidative phosphorylation were enriched, alongside the biosynthesis of homocysteine, pyrimidine, and purine (Supplementary Fig. S8C). In contrast, the turquoise subnetwork was enriched in pathways related to the urea cycle; ketone metabolism; glycerolipid metabolism; and the biosynthesis of fatty acids, cholesterol, and glycosaminoglycans (Supplementary Fig. S8D).

scRNA-seq analysis of blue module subnetwork genes revealed key roles in pericytes, endothelial cells, and glioma cells

Genes from the blue module subnetwork were further analyzed using scRNA-seq data (GSE182109) to identify the genes involved in glioma progression (GSE182109). After excluding recurrent samples and low-quality cells, data from 138,002 cells across 24 samples that passed all quality control steps were analyzed. Unsupervised clustering revealed 27 clusters with distinct gene expression patterns (Fig. 4A). These clusters were annotated using the "singleR" package, the CellMarker database, and references, resulting in the identification of 9 cell types: B cells, mast cells, endothelial cells, T cells, NK cells, myeloid cells, pericytes, glioma cells, and oligodendrocytes (Fig. 4B, Supplementary Fig. S9A-J). The enrichment scores of different cell subsets in the blue module subnetwork were calculated using the 'AddModuleScore' function. The expression of genes related to pericytes, endothelial cells, and glioma cells was upregulated in GBM patients compared with LGG patients (Fig. 4C). These findings indicate that several genes in the blue module are upregulated in these three cell subsets, especially in GBM samples (Fig. 4D-E). Next, these three cell subsets were extracted to analyze their gene expression profiles. As shown in Fig. 4F, most genes were highly expressed in pericytes, with a significant increase in expression in GBM samples. In endothelial cells, most of the blue module genes, including *COL1A2*, *COL4A1*, *COL4A2*, *COL5A2*, *COL6A2*, *COL18A1*, *AEBP1*, *ITGA5*, *LAMB1*, *NID2*, *SERPINH1*, and *VIM*, were highly expressed, with elevated expression levels in GBM samples compared with LGG samples (Fig. 4G). Within the glioma subset, *VIM* was the most highly expressed gene, maintaining elevated expression levels in GBM samples (Fig. 4H). GO and KEGG analyses revealed that in pericytes and endothelial cell subsets, DEGs between LGG and GBM were enriched mainly in the cellular component (CC) terms "collagen-containing ECM" and the KEGG pathway "ECM-receptor interaction" (Supplementary Fig. S10A-D). This highlights the critical role of the ECM in these cells. In the glioma subset, the most significantly enriched CC term was collagen-containing ECM (Supplementary Fig. S10E). These findings further indicate the critical role of the ECM in glioma progression. Additionally, KEGG pathway analysis revealed significant enrichment of the HIF-1 signaling pathway (Supplementary Fig. S10F), suggesting the potential involvement of this pathway in glioma progression and response to hypoxic conditions.

Validation and prognostic significance of COL1A1, COL4A1, and VIM expression in GBM tissues

The marker genes, blue module subnetwork genes, and DEGs of the three cell types identified from the scRNA-seq data were intersected via Venn diagrams. In total, 17 common intersecting genes were identified and defined as signature genes (Supplementary Fig. S11A). The training sets from the TCGA-LGG and TCGA-GBM cohorts were subjected to univariate Cox regression analysis[31]. All 17 genes were significantly associated with overall survival (OS) (Supplementary Fig. S11B). Next, genes were screened for model construction via the LASSO
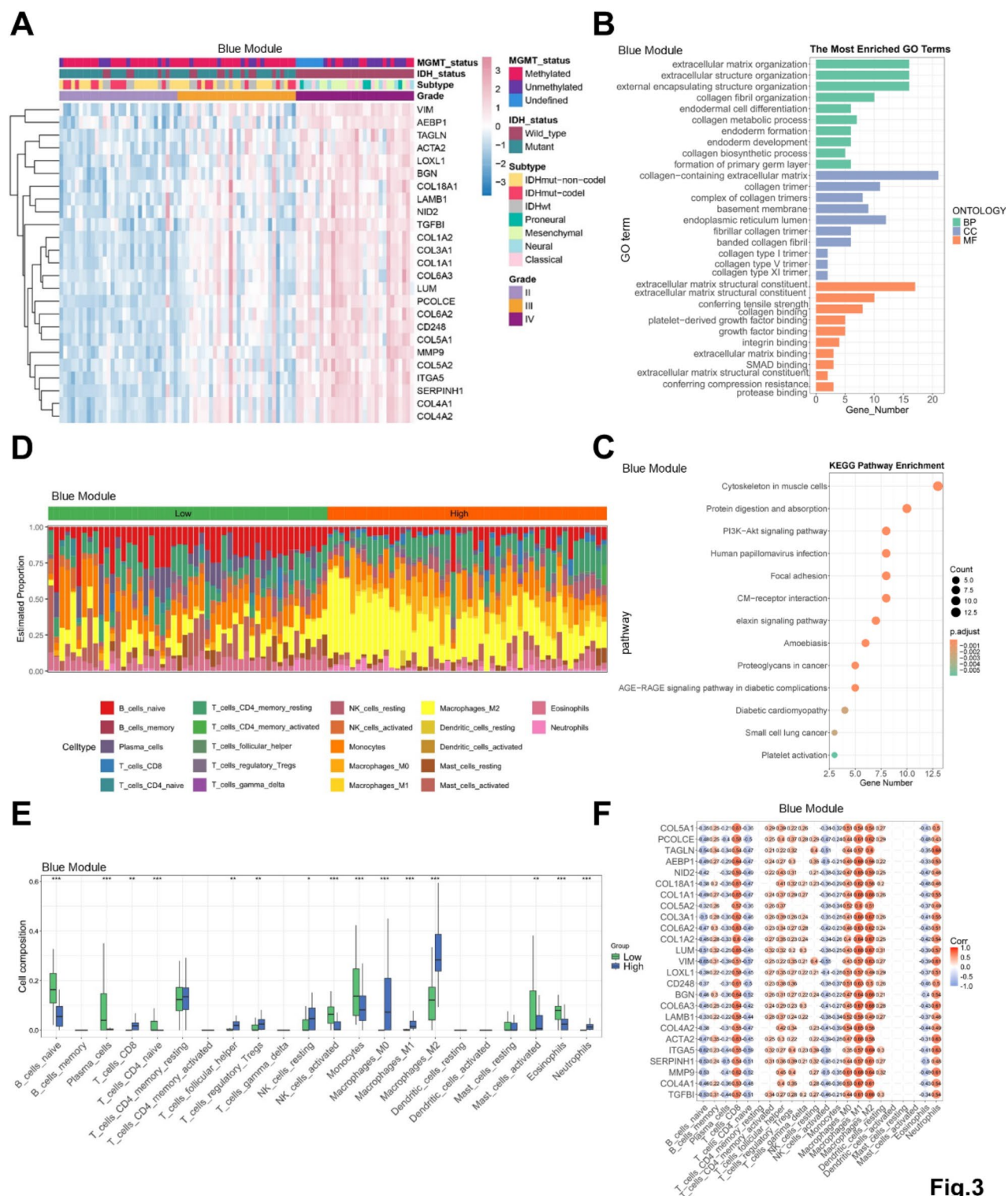
Fig. 3. Analysis of the blue module subnetwork genes in glioma and immune infiltration. (**A**) Heatmap of genes from the blue module subnetworks in The Cancer Genome Atlas-Low-Grade Glioma (TCGA-LGG) and TCGA-glioblastoma multiforme (GBM) datasets. (**B**–**C**) Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses of genes from the blue module subnetworks. (**D**) Differential proportions of 22 immune cell types between the blue module subnetwork low-score and blue module subnetwork high-score groups were determined via CIBERSORT. (**E**) Box plot displaying the differential infiltration levels of 22 immune cell types between the blue module subnetwork low-score and blue module subnetwork high-score groups. (**F**) Correlations of 22 immune cell types with blue module subnetwork genes.
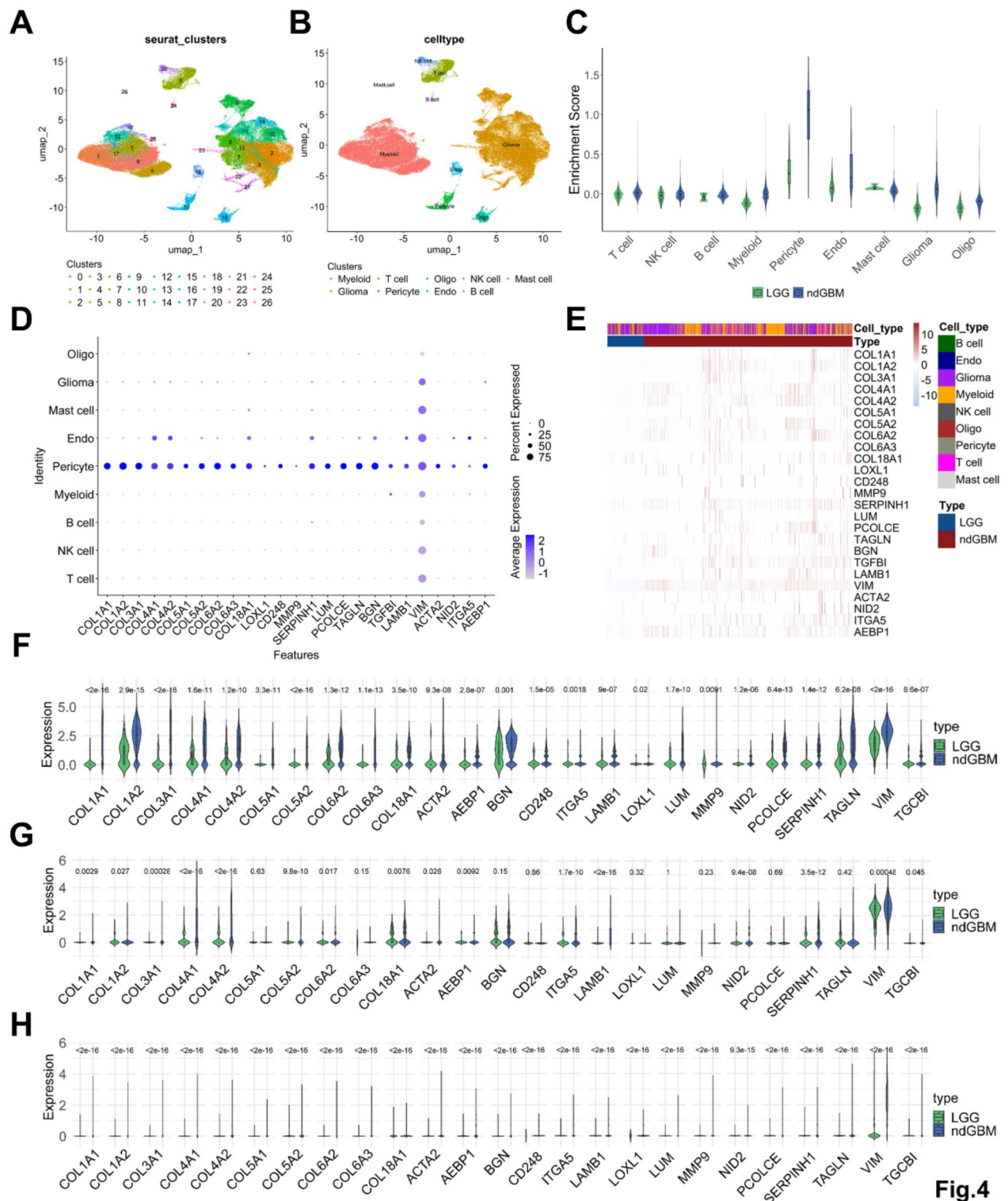
Fig. 4. Single-cell RNA sequencing (scRNA-seq) analysis of blue module subnetwork genes revealed their roles in pericytes, endothelial cells, and glioma cells. (**A**) Uniform manifold approximation and projection (UMAP) plots of 138,002 aggregate single cells from 24 samples. The clusters were identified via the UMAP algorithm. Each cluster is numbered and color-coded to represent different cell populations. (**B**) All nine cell clusters in glioma were annotated with "singleR" and CellMarker according to the composition of the marker genes. (**C**) Violin plot showing the expression of blue module subnetwork genes in nine cell clusters. (**D**) Dot plot presenting the expression of blue module subnetwork genes in nine cell clusters. (**E**) Heatmap displaying the expression of blue module subnetwork genes in low-grade glioma (LGG) and newly diagnosed glioblastoma multiforme (GBM) (ndGBM), as well as in nine distinct cell clusters. (**F**–**H**) Violin plot showing the expression of blue module subnetwork genes in pericytes (**F**), endothelial cells (**G**), and glioma cells (**H**) in patients with LGG or ndGBM.

algorithm[32]. The λmin with the highest accuracy was selected to construct the model. None of the 17 genes were eliminated (Supplementary Fig. S11C). Patients were then divided into high-risk and low-risk groups on the basis of their risk scores. K-M analysis revealed that the OS of patients with high-risk scores was significantly lower than that of patients with low-risk scores (Supplementary Fig. S11D-E). To further assess the validity of the risk model, an ROC curve for OS was generated. The AUC values at 1, 2, 3, 4, and 5 years were greater than 0.84, indicating enhanced efficacy of the risk model (Supplementary Fig. S11F). The functional validation of the model was performed with the internal validation set and the external validation set GSE184941. The validation analysis also revealed the high accuracy of the model (Supplementary Fig. S12A-F). These findings indicate that the prognostic model based on 17 genes has excellent prognostic value.

To identify independent prognostic factors, univariate and multivariate Cox regression analyses of clinical characteristics were performed. As shown in Fig. 5A-B, the risk score, PRS (progression status) type, histology, grade, and X1p19q codeletion are independent prognostic factors for patients with gliomas. These five independent factors, along with the chemotherapy response, were included in the nomogram model (Fig. 5C). The calibration curve demonstrated that the model had high predictive accuracy (Fig. 5D). Thus, the risk score is positively associated with glioma progression and can effectively predict the OS of patients with gliomas.

To further validate the expression of the previously identified signature genes in clinical samples, this study selected *COL1A1*, *COL4A1*, and *VIM* because they exhibited the highest fold changes in pericytes, endothelial cells, and glioma cells between LGG and GBM. qRT-PCR analysis revealed that the *COL1A1*, *COL4A1*, and *VIM* mRNA levels were significantly higher in the tumors than in the paracancerous tissues of patients with GBM (Fig. 5E).

Consistent with the results of the bioinformatics and qRT-PCR analyses, immunohistochemical analysis revealed that the COL1A1, COL4A1, and VIM protein levels are upregulated in the tumor core tissues of patients with GBM (Fig. 5F-G). Moreover, multiplex immunofluorescence staining revealed that the protein levels of COL1A1, COL4A1, and VIM were increased primarily in pericytes, endothelial cells, and glioma cells in tumor core tissues compared with those in paracancerous tissues, with lower expression levels of these proteins in macrophages (Fig. 5H-I, Supplementary Fig. S13A-C). Additionally, univariate and multivariate Cox regression analyses indicated that COL1A1, COL4A1, and VIM were independent risk factors. The ROC curve results revealed that the expression levels of COL1A1, COL4A1, and VIM can predict survival. The AUC values of the three genes for predicting 1-year, 3-year, and 5-year survival were greater than 0.72, indicating high diagnostic efficiency (Supplementary Fig. S14A-I).

Drug sensitivity analysis in high-risk and low-risk glioma groups: identification of potential therapeutics and pathway targets

The top 10 drugs that were negatively correlated with risk scores were identified in the GDSC dataset (Fig. 6A). The targets and pathways of these drugs are summarized in Fig. 6E. The sensitivity to these chemotherapeutic drugs was significantly different between the high-risk and low-risk groups (Fig. 6B). Additionally, analysis of the CTRP database revealed that birinapant, clofarabine, MLN2238, methotrexate, tanespimycin, and gemcitabine exhibited the strongest negative correlations with risk scores (Fig. 6C). The AUC values of these six drugs were significantly different between the high-risk and low-risk groups (Fig. 6D). Furthermore, the pathways targeted by these small-molecule drugs were identified (Fig. 6E). These findings suggest that the identified drugs are potential therapeutic agents for GBM on the basis of the differential sensitivities of the risk groups to these drugs.

Discussion

Microenvironmental factors regulate the phenotype and aggressiveness of GBM, which is the most aggressive and lethal brain tumor characterized by heterogeneity and resistance to conventional therapies[33,34]. Glioma cells express specific integrins and other receptors that interact with ECM components, facilitating their migration and invasion[35,36]. Altered protein synthesis in tumors contributes to ECM degradation, promotes invasion, and enhances the synthesis of ECM components in adjacent non-tumorous tissues[37,38]. Owing to the complexity of GBM progression and the critical role of the TME, advanced analytical methods must be used to elucidate the underlying molecular mechanisms.

To explore the GBM microenvironment, WGCNA was performed to identify gene modules from large mRNA datasets[39]. This approach enabled the identification of gene modules associated with GBM progression. In particular, the turquoise module was associated with physiological brain functions. Consistent with previous findings, GBM-induced remodeling of human neural circuits decreased survival, indicating the importance of neuron-tumor synaptic communication in the development of malignancy[23,40]. Additionally, increased enrichment scores in the red, brown, and turquoise module subnetworks were correlated with improved survival rates and decreased HRs. Conversely, patients with increased enrichment scores in the blue (*COL6A3*, *COL1A2*, *COL3A1*, *COL5A2*, *COL1A1*, *COL6A2*, *COL5A1*, *COL18A1*, and *COL4A2*), black, green, and green-yellow module subnetworks presented poor survival rates and high HRs, indicating an unfavorable outcome. Functional enrichment analysis revealed that the blue module hub genes were involved in critical processes and pathways, especially in ECM remodeling and ECM-receptor interactions. Consistent with previous findings, collagen VI (COL6), an ECM protein widely expressed in both healthy and pathological tissues, is distinctly expressed in GBM tissues[19,41]. Furthermore, ECM proteins, including various collagens, condense the tumor ECM and decrease its flexibility. This can impair the diffusion of neuroactive molecules or therapeutic agents[42], adversely affecting the treatment outcomes and survival rates of patients with GBM.

This study also revealed significant variations in immune cell types and metabolic adaptations, offering insights into the complex interactions that drive GBM progression. The high-risk groups presented increased infiltration of immunosuppressive cells and alterations in metabolic pathways. Thus, novel biomarkers must be screened to aid in the development of personalized therapies and improve patient prognosis. To further dissect
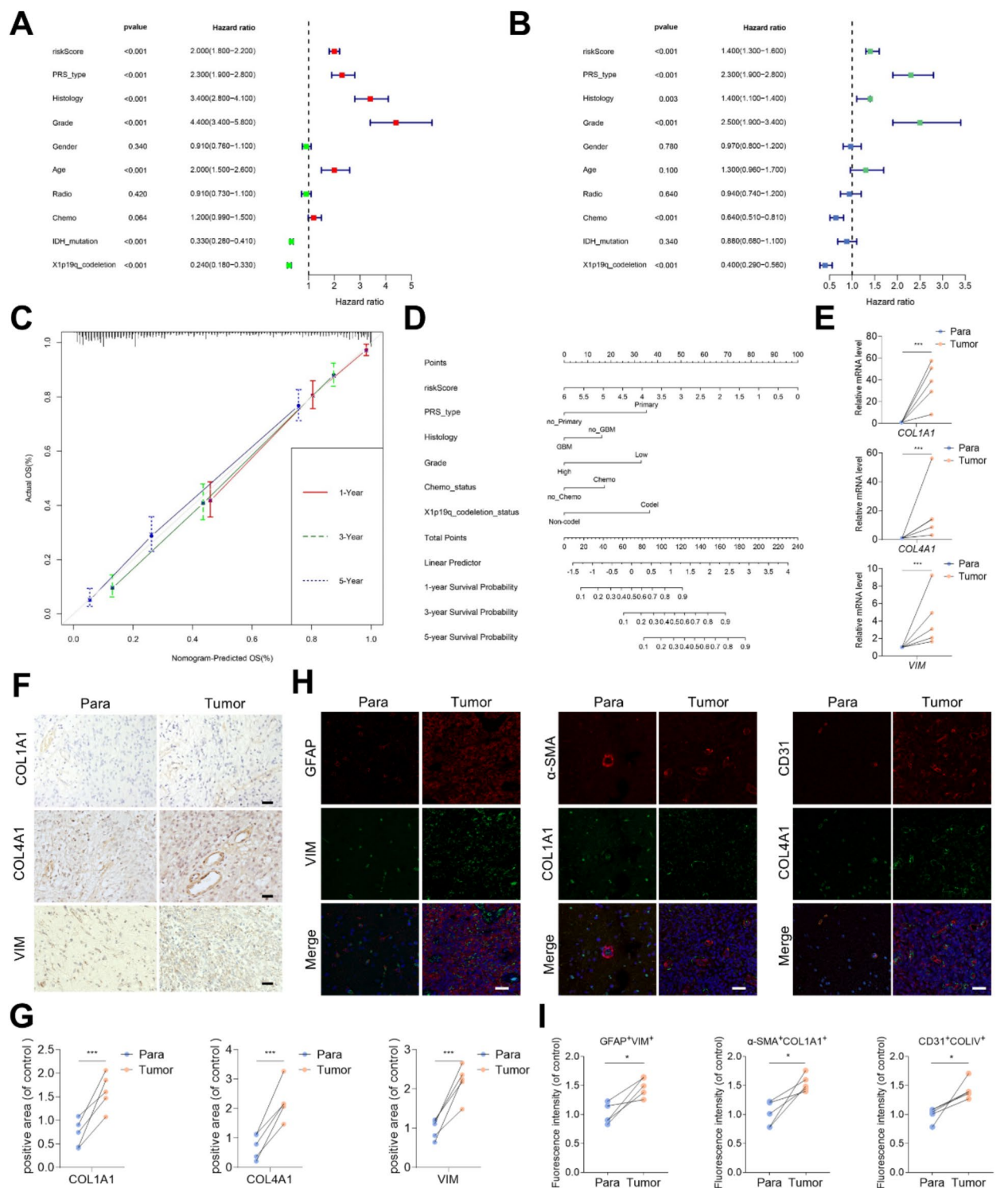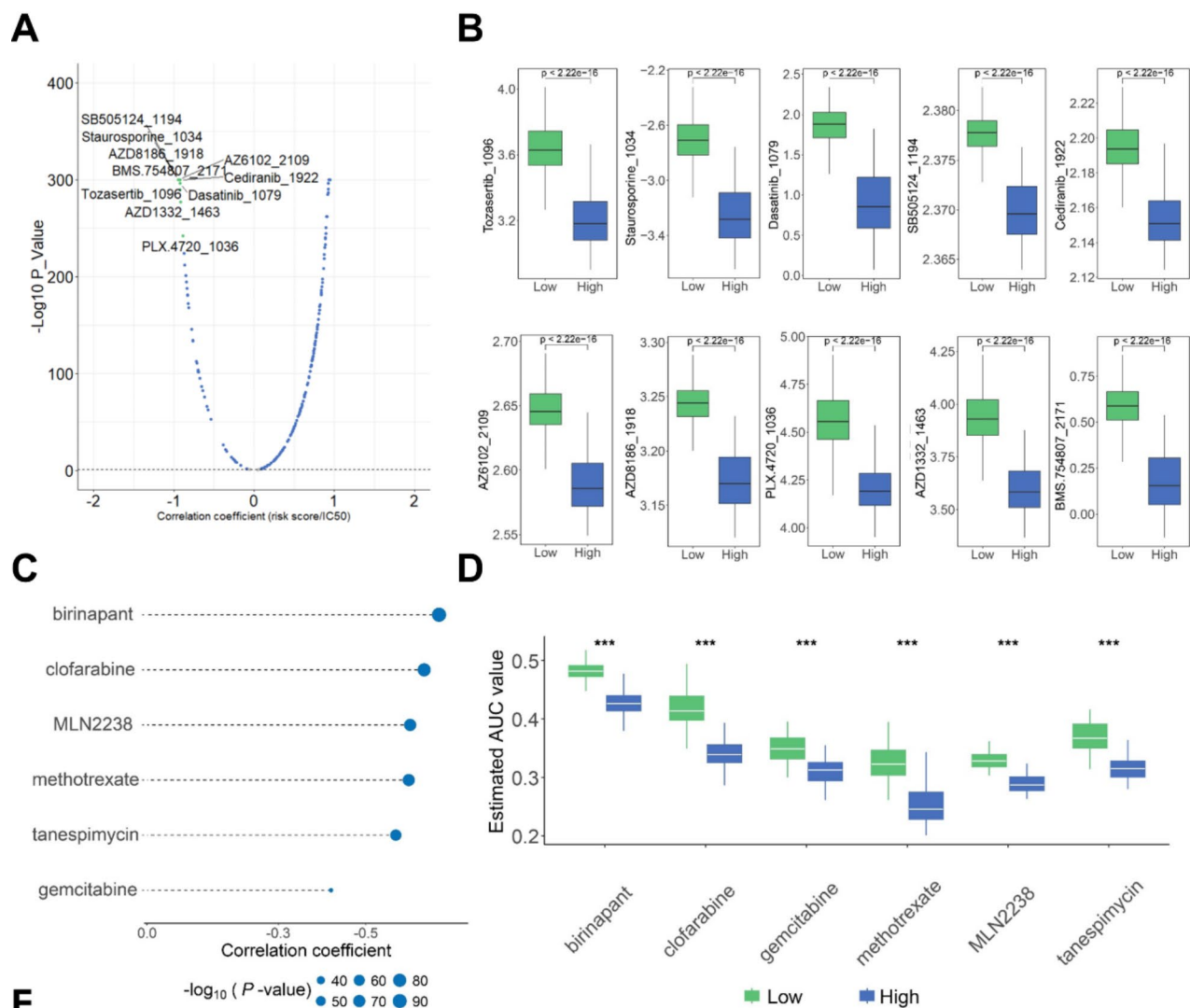
Fig. 5. Validation and prognostic significance of COL1A1, COL4A1, and VIM expression in GBM tissues, (**A**,**B**) Univariate (**A**) and multivariate (**B**) Cox regression analyses of risk scores and clinical characteristics. (**C**) Nomogram model construction. (**D**) Calibration curve of the nomogram. (**E**) mRNA expression levels of *COL1A1*, *COL4A1*, and *VIM* in paracancerous and tumor tissues of patients with glioblastoma multiforme (GBM). The data are from five replicates. (**F**–**G**) Immunohistochemical analysis and statistical analysis of COL1A1, COL4A1, and VIM expression levels in paracancerous and tumor tissues of patients with GBM. Scale bar, 50 μm. (**H**–**I**) Representative confocal images for GFAP, VIM, α-SMA, COL1A1, CD31, and COL4A1 of human tissues ($n = 3$) from paracancerous and tumor regions. A minimum of 5 randomly chosen fields of each tumor ($n = 5$) were evaluated. Scale bars, 50 μm.

Fig. 6. Drug sensitivity analysis of the high-risk and low-risk groups. (**A**) A glioma dataset from the Genomics of Drug Sensitivity of Cancer (GDSC) database was subjected to Pearson correlation analysis to examine the correlation between glioma and estimated half-maximal inhibitory concentration (IC50) values. The top 10 negatively correlated candidate compounds were identified. (**B**) Box plot of the IC50 values of the top 10 negatively correlated candidate compounds in the high-risk and low-risk groups. (**C**) The area under the curve (AUC) values of compounds from the Cancer Therapeutics Response Portal (CTRP) database were estimated for each patient with glioma. The significant correlation between glioma and AUC values was determined via Pearson correlation analysis on the basis of the following criteria: $R > 0.4$ and $p < 0.05$. (**D**) Box plot of the AUC values of negatively correlated compounds from the CTRP database in the high-risk and low-risk groups. (**E**) Table showing the names, IDs, targets, and pathways of the 10 drugs.

the TME, scRNA-seq was used for the transcriptional stratification of cell subpopulations and the identification of specific biomarkers and heterogeneity among different cell types in LGG and GBM. A comprehensive scRNA-seq analysis provided a detailed view of the cellular heterogeneity in GBM, emphasizing the roles of pericytes, endothelial cells, and glioma cells in ECM remodeling. Compared with LGG patients, GBM patients presented upregulated expression of ECM-related genes in these cell types. Pericytes envelop endothelial cells, forming a sheath around blood vessels, and contribute to the formation and stabilization of the vascular basement membrane[43]. Blood vessels play crucial roles in gliomas by providing nutrition, enabling immune cell infiltration, influencing drug penetration, and contributing to tumor growth and treatment response[44]. Previous studies have demonstrated that altering the mechanical properties of the collagen matrix regulates angiogenesis and promotes a tumor vasculature phenotype. In particular, increased collagen density elevates the energy state of endothelial cells, leading to increased rates of tip-stalk cell switching[45,46]. This process is mediated by mechanosensors, such as integrins, the actin cytoskeleton, and ion channels, which convert mechanical signals into biochemical responses[47,48]. Furthermore, increased MMP activity in a stiff TME promotes angiogenesis, invasion, and new vessel branching, enhancing the overall tumor vasculature[49,50]. Thus, these mechanisms must be elucidated for the development of effective therapeutic strategies targeting tumor angiogenesis.

Therapies that target the ECM use various methods, such as antibodies, RNA interference, physical or enzymatic modifications, and TME modulation[13,51,52]. The main challenge in cancer treatment is the lack of effective and specific ECM targets. Next, this study aimed to understand GBM progression and identify therapeutic targets involved in ECM remodeling. The upregulation of ECM-related genes, such as *COL1A1*, *COL4A1*, and *VIM*, was significantly associated with GBM aggressiveness. These findings indicate the critical roles of these genes in disease pathology. Consequently, a prognostic model was developed on the basis of these genes. The prognostic model enabled risk stratification among patients with gliomas. Validation of the prognostic model across multiple datasets and the correlation of the risk score with OS indicated the clinical relevance of the prognostic model. Additionally, drug sensitivity analysis revealed potential therapeutic agents that can modulate the expression of COL1A1, COL4A1, and VIM. For example, SB505124 inhibits the TGF-β pathway to downregulate COL1A1 and COL4A1 synthesis. Staurosporine downregulates VIM by modulating cell adhesion and ECM assembly. AZD8186 targets the PI3K/AKT pathway to downregulate collagen protein expression. AZ6102 inhibits the Wnt/β-catenin pathway to suppress the transcriptional regulation of COL1A1 and COL4A1.

The ECM serves as a therapeutic target and diagnostic marker. The second harmonic generation properties of ECM components have been previously used to achieve high-contrast imaging of deep tumors in vivo[53]. Additionally, radiomics combined with machine learning algorithms is effective in identifying early malignant tumors[54]. Noninvasive liquid biopsies, which can detect ECM components or fragments, are potential strategies for early diagnosis. Thus, ECM components, including COL1A1, COL4A1, and VIM, can serve as early-stage diagnostic markers and therapeutic targets.

This study has several limitations. The regulatory mechanisms of the signature genes in GBM have not been elucidated. Additionally, the clinical sample size was small. Furthermore, the number of LGG samples in the single-cell dataset was relatively small, which may affect the robustness of the study conclusions. Therefore, future studies must address these limitations by expanding clinical sample sizes and elucidating the regulatory networks of these key genes to enhance our understanding of GBM pathogenesis and develop novel therapeutic strategies for GBM.

Conclusions

This study improved our understanding of the molecular landscape of GBM, contributing to ongoing efforts to develop effective treatments for this lethal malignancy. The integration of multiomics data and innovative bioinformatics approaches can enable the development of precision medicine and consequently enhance the clinical outcomes of patients with GBM.

Methods
Clinical samples

Human paracancerous and cancer tissues were obtained from the First Affiliated Hospital of Chongqing Medical University. Approval for tissue collection and experimentation was granted by the Institutional Review Board of the First Affiliated Hospital of Chongqing Medical University. All methods were performed in accordance with the relevant guidelines and regulations, ensuring compliance with ethical standards.

Dataset collection and preprocessing

The glioma microarray datasets GSE4290, GSE108476, and GSE15824 were sourced from the Gene Expression Omnibus (GEO) database. WGCNA was performed with the data of 553 samples (176, 29, and 348 samples from the GSE4290, GSE15824, and GSE108476 datasets, respectively). These samples were strategically classified into the following four cohorts: Non-tumor, Grade II, Grade III, and GBM groups. The expression profiles and clinical data of patients with GBM and low-grade glioma (LGG) were extracted from the TCGA and the Chinese Glioma Genome Atlas (CGGA) databases. The matrisome gene sets were obtained from the Molecular Signature Database (MSigDB v5.0)[55]. Additionally, the scRNA-seq data were retrieved from the GSE182109 dataset, and recurrent samples were excluded. The raw data (GSE4290, GSE108476, and GSE15824) were processed using R (version 4.4.0). The "sva" R package was used to perform batch normalization across combined datasets to precisely mitigate batch effects.

WGCNA

WGCNA was performed via the "WGCNA" R package to construct a gene coexpression network. The analysis initially focused on the top 20% of the most variable genes, paring down the noise and computational load to identify 4020 genes from the original 20,099 genes. An optimal soft threshold was selected to ensure the maintenance of scale-free properties. The weighted adjacency matrix was transformed into a topological overlap matrix (TOM) to assess the intricacies of network connectivity. A dendrogram was constructed on the basis of TOM via average linkage hierarchical clustering. A minimum module size of 30 was set to ensure the identification of modules, and a threshold of 0.25 was used to merge similar modules. Modules associated with clinical traits were identified. Hub genes in each module that are significantly linked to clinical traits were identified on the basis of the module membership (MM) and gene significance (GS) values. These hub genes were subjected to functional enrichment analysis to identify specific biological processes and signaling pathways.

Functional enrichment analysis

Hub genes of each module were subjected to Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses. The Benjamini-Hochberg (BH) correction method was used to adjust P-values for multiple comparisons in both GO and KEGG pathway enrichment analyses. Significantly enriched terms and pathways were identified on the basis of the following criteria: $p < 0.01$ and a q value (false discovery rate) < 0.05.

Protein-protein interaction (PPI) network construction

The genes of each module were exported via the 'exportNetworkToCytoscape' function and imported into the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database (https://string-db.org/), an extensive resource for known and predicted PPIs. These interactions were then visualized using Cytoscape. The MCODE plugin was used to identify highly interconnected subnetworks to indicate the functional relevance of each module. These subnetworks, which were selected on the basis of the clustering coefficient and density, were primed for further analysis.

TCGA and CGGA dataset analysis

The TCGA and CGGA datasets were used to evaluate module subnetworks related to glioma progression, survival, and therapy response. The enrichment scores for the module subnetworks and tumor progression gene sets were computed via single-sample gene set enrichment analysis (ssGSEA). The "ggpubr" package was used to analyze the correlation between module subnetwork enrichment scores and glioma progression. Survival analysis was performed to assess the prognostic value of module subnetwork enrichment scores by calculating hazard ratios (HRs), confidence intervals (CIs), and p values. Patients were classified as responders or non-responders on the basis of a 14.2-month posttreatment survival threshold. Module subnetwork enrichment scores were then compared between the responders and non-responders. The "multipleROC" R package was used to evaluate the differential characteristics between responders and non-responders. This provided a robust method for validating the prognostic significance of the module enrichment scores. Additionally, patients were categorized on the basis of their enrichment scores. The data of patients in the top 0.75 quantile and the bottom 0.25 quantile were extracted for further analysis.

Immune cell infiltration and metabolism analysis

Using the TCGA dataset, ssGSEA was performed on genes from MCODE-identified subnetworks via the MCODE plugin in Cytoscape. The infiltration levels of 22 immune cell types in 50 patients with the highest ssGSEA scores and 50 patients with the lowest ssGSEA scores were determined via cell type identification via the estimation of relative subsets of RNA transcripts (CIBERSORT) algorithm with the "CIBERSORT" R package. Metabolic pathway analysis was performed using the "IOBR" R package.

scRNA-seq data analysis

Analysis of the scRNA-seq data (GSE182109) was performed using Seurat 4.4.0, with recurrent GBM samples excluded. Initial quality control filtered out cells expressing fewer than 500 genes. Additional quality checks involved filtering cells with more than 5000 genes, more than 50% ribosomal gene counts, or more than 20% mitochondrial gene counts. After low-quality cells were removed, the remaining data were normalized via the "SCTransform". PCA was then performed on the most variable genes, followed by batch correction using "Harmony". Unsupervised clustering was conducted using Louvain-based graph clustering, with resolution parameters ranging from 0.1 to 1, adjusted according to the dataset's characteristics. Marker genes specific to each cluster were identified using the 'FindMarkers' function. The cell types were annotated using the "singleR" package, the CellMarker database, and relevant references. The expression profiles of distinct cell types were visualized via dot plots to characterize their unique features.

Construction and validation of a prognostic model

Differentially expressed genes (DEGs) in pericyte, endothelial cell, and tumor cell subpopulations between LGG and GBM were identified from scRNA-seq data. These DEGs were compared with genes from the blue module subnetwork identified using WGCNA to obtain common genes. The TCGA glioma dataset was randomly divided into a training set (70%) and a validation set (30%). The common genes were subjected to univariate Cox regression analysis using the training set to identify significant candidate prognostic genes ($p < 0.05$). The identified prognostic genes were refined via least absolute shrinkage and selection operator (LASSO) regression with the "glmnet" R package to construct a prognostic model. The risk score was calculated as follows:

risk score = (gene 1 expression × β1) + (gene 2 expression × β2) + … + (gene n expression × βn), where gene expression denotes the gene expression level and β denotes the corresponding LASSO regression coefficient. A risk score for each patient was calculated using the coefficients obtained from the LASSO regression. This risk score was used to perform survival analysis with the validation set to assess its prognostic value. The prognostic model was validated using an external dataset (GSE184941). The predictive power of the prognostic model was evaluated using receiver operating characteristic (ROC) curves generated using the "survivalROC" package.

Prognostic analysis and nomogram model construction

Univariate and multivariate Cox analyses were performed to evaluate the prognostic significance of clinical characteristics and gene expression, including risk score, PRS type, histology, grade, sex, age, radiotherapy, chemotherapy, *IDH* mutation, and X1p19q codeletion. The 'coxph' function in the "survival" R package was used to perform univariate analysis and calculate the HR and 95% CI. Differences were considered significant at $p < 0.05$. Significant variables were included in a multivariate Cox regression model to identify independent prognostic factors. A nomogram model was constructed via the 'cph' and 'nomogram' functions with the R package "rms" to predict the 1-year, 3-year, and 5-year survival probabilities. Calibration curves were generated with the calibrate function to validate the predictive accuracy of the nomogram.

Chemotherapy drug sensitivity analysis

To further investigate the potential of risk scores to guide chemotherapy, drug sensitivity data from the Genomics of Drug Sensitivity in Cancer (GDSC) and the Cancer Therapeutics Response Portal (CTRP) databases were analyzed. Drug sensitivity values were obtained using the R package "oncoPredict." In the GDSC dataset, the top 10 drugs that are most strongly correlated with risk scores were identified on the basis of their half-maximal inhibitory concentration (IC50) values. Moreover, the area under the dose-response curve (AUC) values were retrieved from the CTRP dataset. Pearson correlation analysis was performed to examine the correlation between drug sensitivity (IC50 and AUC values for the GDSC and CTRP datasets, respectively) and the computed risk scores. In the CTRP dataset, drugs with an absolute Pearson correlation coefficient (R) > 0.4 were identified. Next, the sensitivity values (IC50 and AUC values for the GDSC and CTRP datasets, respectively) were compared between the high-risk and low-risk groups and were defined by the median risk score. Patients with risk scores above the median were categorized into the high-risk group, whereas those with scores below the median were categorized into the low-risk group. The results were visualized using box plots and lollipop plots with the R package "ggplot2".

Quantitative real-time polymerase chain reaction (qRT-PCR)

Total RNA was isolated using TRIzol (catalog no. 15596026; Invitrogen). GBM tissues from paracancerous and tumor regions were lysed with 1 mL of TRIzol using ultrasonication. The supernatant obtained after ultrasonication was mixed with chloroform (catalog no. C2432-500ML, Sigma-Aldrich), vigorously mixed for 90 s, and incubated at room temperature for 3–5 min. Next, the mixture was centrifuged at 12,000 ×g and 4 °C for 15 min to achieve phase separation. The aqueous phase containing RNA was carefully transferred to a new tube, mixed with isopropanol (catalog no. 59300-4×4 L-R; Sigma-Aldrich), and incubated at 4 °C for 30 min to precipitate the RNA. The sample was subsequently centrifuged at 12,000 ×g and 4 °C for 10 min, after which the pellet was washed with 75% ethanol. After the supernatant was discarded, the RNA pellet was air-dried for 5–10 min and resuspended in RNase-free water. The RNA concentration was quantified using spectrophotometry. The RNA (1 μg) sample was reverse-transcribed into complementary DNA using HiScript II Q RT SuperMix for QPCR (catalog no. R223-01, Vazyme). qRT-PCR analysis was performed using ChamQ Universal SYBR QPCR Master Mix (catalog no. Q711-02, Vazyme).

Immunohistochemical staining

Paraffin-embedded tissue sections were deparaffinized with xylene, rehydrated in an ethanol gradient, and washed with phosphate-buffered saline (pH 7.4). To retrieve antigens, the sections were boiled with citrate buffer for 20 min. Peroxidase activity was quenched by incubating the sections with 3% hydrogen peroxide for 10 min. The sections were blocked with 5% goat serum for 1 h at room temperature to suppress nonspecific binding. Next, the sections were incubated overnight at 4 °C with anti-COL1A1, anti-COL4A1, and anti-VIM primary antibodies (Proteintech; all 1:200). The sections were then incubated with horseradish peroxidase-conjugated secondary antibodies (Proteintech; all 1:100), followed by staining with 3,3′-diaminobenzidine. The fluorescence and colocalization signals were quantified using image analysis software (Fiji).

Multiplex Immunofluorescence

For multiplex immunofluorescence, tissue samples were deparaffinized, rehydrated, subjected to antigen retrieval, and blocked. The sections were then incubated with the primary antibody for the first target, followed by incubation with an HRP-conjugated secondary antibody (catalog no. AFIHC024, AiFang Biological). Fluorescently labeled tyramide was then applied to visualize the target. After staining, antigen retrieval and endogenous peroxidase activity were quenched again, and blocking was performed before proceeding with subsequent rounds of staining. Each subsequent round involved incubation with a different primary antibody, followed by incubation with an HRP-conjugated secondary antibody and fluorophore labeling, with repeated blocking and quenching steps between each cycle. After all the targets were labeled, the sections were counterstained with DAPI to visualize the nuclei and then mounted with Fluoromount-G.

Cox regression validation

Univariate and multivariate Cox regression analyses were performed to evaluate the significance of *COL1A1*, *COL4A1*, and *VIM*. The data for these analyses were sourced from the TCGA database. The "timeROC" package was used to predict the 1-year, 3-year, and 5-year survival probabilities on the basis of the Cox regression results, offering a comprehensive assessment of the prognostic value of these biomarkers.

Statistical analysis

Data analysis and graphical visualization were performed using R (version 4.4.0). Continuous variables were compared using the Wilcoxon rank-sum test or Student's t test, whereas categorical variables were compared using the chi-square test or Fisher's exact test. Differences were considered significant at $p < 0.05$ ($^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, ns: nonsignificant).

Data availability

Data is provided within the manuscript or supplementary information files.

References

1. Cela, I., Capone, E., Trevisi, G. & Sala, G. Extracellular vesicles in glioblastoma: biomarkers and therapeutic tools. *Sem. Cancer Biol.* **101**, 25–43. https://doi.org/10.1016/j.semcancer.2024.04.003 (2024).
2. Bikfalvi, A. et al. Challenges in glioblastoma research: focus on the tumor microenvironment. *Trends cancer.* **9**, 9–27. https://doi.org/10.1016/j.trecan.2022.09.005 (2023).
3. van Tellingen, O. et al. Overcoming the blood-brain tumor barrier for effective glioblastoma treatment. *Drug Resist. Updates: Reviews Commentaries Antimicrob. Anticancer Chemother.* **19**, 1–12. https://doi.org/10.1016/j.drup.2015.02.002 (2015).
4. Yang, F. et al. An immunosuppressive vascular niche drives macrophage polarization and immunotherapy resistance in glioblastoma. *Sci. Adv.* **10**, eadj4678. https://doi.org/10.1126/sciadv.adj4678 (2024).
5. Akiyama, T. et al. Stromal reprogramming through dual PDGFRα/β Blockade boosts the efficacy of Anti-PD-1 immunotherapy in fibrotic tumors. *Cancer Res.* **83**, 753–770. https://doi.org/10.1158/0008-5472.can-22-1890 (2023).
6. Wu, W. et al. Glioblastoma multiforme (GBM): an overview of current therapies and mechanisms of resistance. *Pharmacol. Res.* **171**, 105780. https://doi.org/10.1016/j.phrs.2021.105780 (2021).
7. Li, Q., Aishwarya, S., Li, J. P., Pan, D. X. & Shi, J. P. Gene expression profiling of glioblastoma to recognize potential biomarker candidates. *Front. Genet.* **13**, 832742. https://doi.org/10.3389/fgene.2022.832742 (2022).
8. P Cruz, S. et al. Dampened regulatory circuitry of TEAD1/ITGA1/ITGA2 promotes TGFβ1 signaling to orchestrate prostate Cancer progression. *Adv. Sci. (Weinheim Baden-Wurttemberg Germany).* **11**, e2305547. https://doi.org/10.1002/advs.202305547 (2024).
9. Mao, X. et al. Crosstalk between cancer-associated fibroblasts and immune cells in the tumor microenvironment: new findings and future perspectives. *Mol. Cancer.* **20**, 131. https://doi.org/10.1186/s12943-021-01428-1 (2021).
10. Winkler, J., Abisoye-Ogunniyan, A., Metcalf, K. J. & Werb, Z. Concepts of extracellular matrix remodelling in tumour progression and metastasis. *Nat. Commun.* **11**, 5120. https://doi.org/10.1038/s41467-020-18794-x (2020).
11. Fuller, A. M. et al. Oncogene-induced matrix reorganization controls CD8 + T cell function in the soft-tissue sarcoma microenvironment. *J. Clin. Invest.* **134** https://doi.org/10.1172/jci167826 (2024).
12. Hendriks, D. et al. Human fetal brain self-organizes into long-term expanding organoids. *Cell* **187**, 712–732. .e738 (2024).
13. Jiang, Y. et al. Targeting extracellular matrix stiffness and mechanotransducers to improve cancer therapy. *J. Hematol. Oncol.* **15** https://doi.org/10.1186/s13045-022-01252-0 (2022).
14. Bertero, T. et al. Tumor-Stroma mechanics coordinate amino acid availability to sustain tumor growth and malignancy. *Cell Metabol.* **29**, 124–140e110. https://doi.org/10.1016/j.cmet.2018.09.012 (2019).
15. Denton, A. E., Roberts, E. W. & Fearon, D. T. Stromal cells in the tumor microenvironment. *Adv. Exp. Med. Biol.* **1060**, 99–114. https://doi.org/10.1007/978-3-319-78127-3_6 (2018).
16. Karamanos, N. K. et al. Extracellular matrix-based cancer targeting. *Trends Mol. Med.* **27**, 1000–1013. https://doi.org/10.1016/j.molmed.2021.07.009 (2021).
17. Cui, X. et al. RUNX1/NPM1/H3K4me3 complex contributes to extracellular matrix remodeling via enhancing FOSL2 transcriptional activation in glioblastoma. *Cell. Death Dis.* **15**, 98. https://doi.org/10.1038/s41419-024-06481-4 (2024).
18. Sun, L., Jiang, Y., Tan, H. & Liang, R. Collagen and derivatives-based materials as substrates for the establishment of glioblastoma organoids. *Int. J. Biol. Macromol.* **254**, 128018. https://doi.org/10.1016/j.ijbiomac.2023.128018 (2024).
19. Cescon, M. et al. Collagen VI sustains cell stemness and chemotherapy resistance in glioblastoma. *Cell. Mol. Life Sci.* **80**, 233. https://doi.org/10.1007/s00018-023-04887-5 (2023).
20. Mancusi, R. & Monje, M. The neuroscience of cancer. *Nature* **618**, 467–479. https://doi.org/10.1038/s41586-023-05968-y (2023).
21. Zhang, X. et al. CHD2 regulates Neuron-glioma interactions in pediatric glioma. *Cancer Discov.* https://doi.org/10.1158/2159-8290.Cd-23-0012 (2024).
22. Drexler, R. et al. A prognostic neural epigenetic signature in high-grade glioma. *Nat. Med.* **30**, 1622–1635. https://doi.org/10.1038/s41591-024-02969-w (2024).
23. Venkatesh, H. S. et al. Electrical and synaptic integration of glioma into neural circuits. *Nature* **573**, 539–545. https://doi.org/10.1038/s41586-019-1563-y (2019).
24. Bailey, C. P. et al. Pharmacologic Inhibition of lysine-specific demethylase 1 as a therapeutic and immune-sensitization strategy in pediatric high-grade glioma. *Neuro Oncol.* **22**, 1302–1314. https://doi.org/10.1093/neuonc/noaa058 (2020).
25. Wang, Y. et al. Unveiling the key genes, environmental toxins, and drug exposures in modulating the severity of ulcerative colitis: a comprehensive analysis. *Front. Immunol.* **14** https://doi.org/10.3389/fimmu.2023.1162458 (2023).
26. Heiland, D. H. et al. Molecular differences between cerebral blood volume and vessel size in glioblastoma multiforme. *Oncotarget* **8**, 11083–11093. https://doi.org/10.18632/oncotarget.11522 (2017).
27. Ni, L. et al. Transcriptome and single-cell analysis reveal the contribution of immunosuppressive microenvironment for promoting glioblastoma progression. *Front. Immunol.* **13**, 1051701. https://doi.org/10.3389/fimmu.2022.1051701 (2022).
28. Hänzelmann, S., Castelo, R. & Guinney, J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinform.* **14** https://doi.org/10.1186/1471-2105-14-7 (2013).
29. Han, M. H. et al. Identification of genes from ten oncogenic pathways associated with mortality and disease progression in glioblastoma. *Front. Oncol.* **12**, 965638. https://doi.org/10.3389/fonc.2022.965638 (2022).

30. Newman, A. M. et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods*. **12**, 453–457. https://doi.org/10.1038/nmeth.3337 (2015).
31. Liu, C. H. et al. Development and validation of a nomogram for esophagogastric variceal bleeding in liver cirrhosis: A cohort study in 1099 cases. *J. Dig. Dis.* **23**, 597–609. https://doi.org/10.1111/1751-2980.13145 (2022).
32. Deng, H. et al. Dynamic nomogram for predicting thrombocytopenia in adults with acute pancreatitis. *J. Inflamm. Res.* **14**, 6657–6667. https://doi.org/10.2147/jir.S339981 (2021).
33. DeCordova, S. et al. Molecular heterogeneity and immunosuppressive microenvironment in glioblastoma. *Front. Immunol.* **11** https://doi.org/10.3389/fimmu.2020.01402 (2020).
34. Yabo, Y. A., Niclou, S. P. & Golebiewska, A. Cancer cell heterogeneity and plasticity: A paradigm shift in glioblastoma. *Neuro-oncology* **24**, 669–682. https://doi.org/10.1093/neuonc/noab269 (2022).
35. Safarians, G. et al. Glioblastoma spheroid invasion through soft, Brain-Like matrices depends on hyaluronic Acid-CD44 interactions. *Adv. Healthc. Mater.* **12**, e2203143. https://doi.org/10.1002/adhm.202203143 (2023).
36. Bellail, A. C., Hunter, S. B., Brat, D. J., Tan, C. & Van Meir, E. G. Microregional extracellular matrix heterogeneity in brain modulates glioma cell invasion. *Int. J. Biochem. Cell Biol.* **36**, 1046–1069. https://doi.org/10.1016/j.biocel.2004.01.013 (2004).
37. Wu, Z., Yang, Y., Chen, M. & Zha, Y. Matrix metalloproteinase 9 expression and glioblastoma survival prediction using machine learning on digital pathological images. *Sci. Rep.* **14**, 15065. https://doi.org/10.1038/s41598-024-66105-x (2024).
38. Quesnel, A., Karagiannis, G. S. & Filippou, P. S. Extracellular proteolysis in glioblastoma progression and therapeutics. *Biochim. Et Biophys. Acta Reviews cancer.* **1874**, 188428. https://doi.org/10.1016/j.bbcan.2020.188428 (2020).
39. Xiao, F., Zhao, Y., Wang, X., Jian, X. & Zhou, H. Analysis of differential mRNA and MiRNA expression induced by heterogeneous grafting in Gleditsia sinensis. *Int. J. Biol. Macromol.* **270**, 132235. https://doi.org/10.1016/j.ijbiomac.2024.132235 (2024).
40. Venkataramani, V. et al. Glutamatergic synaptic input to glioma cells drives brain tumour progression. *Nature* **573**, 532–538. https://doi.org/10.1038/s41586-019-1564-x (2019).
41. Mierke, C. T. The matrix environmental and cell mechanical properties regulate cell migration and contribute to the invasive phenotype of cancer cells. *Rep. Progress Phys. Phys. Soc. (Great Britain)*. **82**, 064602. https://doi.org/10.1088/1361-6633/ab1628 (2019).
42. Nia, H. T., Munn, L. L. & Jain, R. K. Physical traits of cancer. *Sci. (New York N Y)*. **370** https://doi.org/10.1126/science.aaz0868 (2020).
43. Ahmed, T. A., El-Badri, N. & Pericytes The role of multipotent stem cells in vascular maintenance and regenerative medicine. *Adv. Exp. Med. Biol.* **1079**, 69–86. https://doi.org/10.1007/5584_2017_138 (2018).
44. Wälchli, T. et al. Single-cell atlas of the human brain vasculature across development, adulthood and disease. *Nature* **632**, 603–613. https://doi.org/10.1038/s41586-024-07493-y (2024).
45. Wang, W. et al. Collagen density regulates tip-stalk cell rearrangement during angiogenesis via cellular bioenergetics. *APL Bioeng.* **8**, 026120. https://doi.org/10.1063/5.0195249 (2024).
46. Gross, S. J. et al. Correction to: Notch regulates vascular collagen IV basement membrane through modulation of Lysyl hydroxylase 3 trafficking. *Angiogenesis* **24**, 807. https://doi.org/10.1007/s10456-021-09801-w (2021).
47. Albig, A. R., Roy, T. G., Becenti, D. J. & Schiemann, W. P. Transcriptome analysis of endothelial cell gene expression induced by growth on matrigel matrices: identification and characterization of MAGP-2 and lumican as novel regulators of angiogenesis. *Angiogenesis* **10**, 197–216. https://doi.org/10.1007/s10456-007-9075-z (2007).
48. Fang, Y., Wu, D. & Birukov, K. G. Mechanosensing and mechanoregulation of endothelial cell functions. *Compr. Physiol.* **9**, 873–904. https://doi.org/10.1002/cphy.c180020 (2019).
49. Heath, E. I. & Grochow, L. B. Clinical potential of matrix metalloprotease inhibitors in cancer therapy. *Drugs* **59**, 1043–1055. https://doi.org/10.2165/00003495-200059050-00002 (2000).
50. Das, A., Monteiro, M., Barai, A., Kumar, S. & Sen, S. MMP proteolytic activity regulates cancer invasiveness by modulating integrins. *Sci. Rep.* **7**, 14219. https://doi.org/10.1038/s41598-017-14340-w (2017).
51. Huang, J. et al. Extracellular matrix and its therapeutic potential for cancer treatment. *Signal. Transduct. Target. Therapy.* **6** https://doi.org/10.1038/s41392-021-00544-0 (2021).
52. Baldari, S., Di Modugno, F., Nisticò, P. & Toietta, G. Strategies for efficient targeting of tumor collagen for Cancer therapy. *Cancers (Basel).* 14. https://doi.org/10.3390/cancers14194706 (2022).
53. Mostaço-Guidolin, L. B. et al. Defective fibrillar collagen organization by fibroblasts contributes to airway remodeling in asthma. *Am. J. Respir. Crit. Care Med.* **200**, 431–443. https://doi.org/10.1164/rccm.201810-1855OC (2019).
54. Floca, R. et al. Radiomics workflow definition & challenges - German priority program 2177 consensus statement on clinically applied radiomics. *Insights into Imaging.* **15** https://doi.org/10.1186/s13244-024-01704-w (2024).
55. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 15545–15550. https://doi.org/10.1073/pnas.0506580102 (2005).

Acknowledgements

Author contributions

YXF, ZFF, and ZXM conceived this project. LXR, CXY, and ZX were involved in project administration. LXR and CXY performed experiments and data analysis and interpretation. CXY, LXR, and ZX prepared the original draft and critically reviewed and edited the manuscript. All authors have read and approved the final version of the manuscript.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-92867-z.

Correspondence and requests for materials should be addressed to X.Y. or X.L.