Short Survival-Related Genes Harbor EMT Processes and Dendritic Cell Infiltration in Glioblastoma

Zhenkun Yang

Wuxi People's Hospital of Nanjing medical university <https://orcid.org/0000-0001-5763-7748>

Bo Zhang

Wuxi People's Hospital of Nanjing Medical University

Zhenhao Zhang

Wuxi People's Hospital of Nanjing Medical University

Jingjing Wang

Wuxi People's Hospital of Nanjing Medical University

Yaling Hu

Wuxi People's Hospital of Nanjing Medical University

Zhening Pu

Wuxi People's Hospital of Nanjing Medical University

Ying Yin

Wuxi People's Hospital of Nanjing Medical University

Jian Zou (✉ zoujan@njmu.edu.cn)

Wuxi People's Hospital of Nanjing Medical University

Research

Keywords: Immune cell infiltration, Glioblastoma, GEO, Bioinformatic analysis, Survival, Dendritic cell

Posted Date: November 13th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-104471/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Glioblastoma is an aggressive primary tumour with the lowest survival time among brain tumours. Tumour-infiltrating immune cells (TIICs) are involved in tumour progression and determine the prognosis, while the association of immune cell infiltration with glioblastoma is rarely unknown. This study aimed to screen survival-related (SR) genes and major biological processes through bioinformatic analysis and to identify the relationship between SR genes and TIICs.

Methods:

SR genes were screened by comparing the long-term (>36 months) and short-term (<12 months) survivors in the database GSE53733. Gene set enrichment analysis (GSEA) was applied to compare the differences in biological processes between long-term survivors and short-term survivors. The SR genes were identified using the limma package of R. Gene Ontology (GO) analysis was conducted through Metascape. The protein-protein interaction (PPI) network of the SR genes was established through the Search Tool for the Retrieval of Interacting Genes (STRING) website and further analysed by the Molecular Complex Detection (MCODE) algorithm. UALCAN and GlioVis were employed to analyse the expression levels and prognostic value of hub genes. The correlation of hub genes with immune cell infiltration was estimated by the Tumor Immune Estimation Resource (TIMER). The gene-drug interaction network was constructed using the Comparative Toxicogenomics Database (CTD).

Results: The functions of the detected genes were mainly enriched in epithelial mesenchymal transition (EMT) and oxidative phosphorylation. Of the detected genes, a total of 220 SR genes were identified, including 78 upregulated genes and 142 downregulated genes in long-term survivors. The upregulated genes were mainly related to neuron projection morphogenesis, extracellular matrix, and cation channel activity. The downregulated genes were mainly related to extracellular matrix organization and angiogenesis. The PPI network for SR genes was constructed with 65 edges and 195 nodes, and two significant modules were selected. The results indicated that COL1A2, COL6A2, COL8A1, and COL8A2 were hub SR genes. In addition, they were correlated with immune cell infiltration, especially dendritic cell infiltration.

Conclusions: These results revealed that collagens accounted for the progression and prognosis of glioblastoma. In addition, DC infiltration is a risk factor for glioblastoma patients. The expression of collagen protein COL6A2 was significantly correlated with the DC infiltration level and poor prognosis. Further, potential drugs that affect the function of COL6A2 could improve the outcomes of glioblastoma.

Background

Glioblastoma, sometimes called glioblastoma multiforme (GBM), is the most aggressive primary brain tumour, with a 5-year survival rate of 6.8%, and accounts for approximately 14.6% of all brain and other central nervous system tumours and 48.3% of malignant gliomas[1]. Despite the application of computed tomography (CT) and magnetic resonance imaging (MRI) for diagnosis and the advances in

multimodality therapy, the overall prognosis of GBM is still poor, and the long-term survival remains the lowest among brain and other central nervous system tumours, with a median survival of only 14 months[2].

The survival of cancer patients is mainly associated with molecular factors and clinical features including age, tumour site and therapies. In GBM, numerous studies have indicated that epidermal growth factor receptor activates the RTK/RAS/PI3K pathway, leading to increased proliferation and is associated with poorer survival[3–5]. Furthermore, methylation of O6-methylguanine-DNA methyltransferase (MGMT), mutation of isocitrate dehydrogenase 1/2 (IDH1/2), PTEN and immune cell infiltration also affect the prognosis of GBM[6, 7]. It seems that studying survival-related factors is meaningful for improving the outcomes of cancer patients.

In solid tumours, the tumour microenvironment (TME) is the soil for tumour progression and consists of cancer cells, endothelial cells, cancer-associated fibroblasts (CAFs), immune cells and noncancer stromal cells[8]. Immune cells are the guardians of the human body. Both innate and adaptive immune cells can interact with cancer cells via expressed antigens on the surface and affect the proliferation and invasion of carcinoma[9]. Firm evidence has indicated that tumours with high expression of programmed death ligand 1 (PD-L1) and indoleamine-2,3-dioxygenase (IDO) display high infiltration of regulatory T cells (Tregs) and have a poor prognosis[10]. Currently, a few drugs targeting PD-L1 have been approved for cancer immunotherapy to improve clinical outcomes[11, 12]. Interestingly, during the evolutionary trajectories of cancer, the remodelling of the TME by alteration of immune cell infiltration leads to diverse tumour behaviour, including immune escape and tolerance[13, 14]. Thus, it is important to clarify the relationship of tumour-related molecular expression with immune cell infiltration, especially in SR molecules.

In this study, we identified collagens as negative survival-related (NSR) genes through bioinformatic analysis of gene expression profiling and analysed the correlation of collagens with immune cell infiltration. Furthermore, we screened potential small molecular drugs that improve clinical outcomes by targeting NSR genes.

Methods

mRNA expression microarray data

The mRNA expression microarray dataset GSE53733 was downloaded from the Gene Expression Omnibus (GEO) (<http://www.ncbi.nlm.nih.gov/geo/>) database. The research subject of the microarray was *Homo sapiens*, and the research type was expression profiling by array. In dataset GSE53733, there were 70 samples, including 23 long-term survivors with > 36 months overall survival (OS), 16 short-term survivors with < 12 months OS, and 31 patients with intermediate OS. In this study, 23 long-term survivors and 16 short-term survivors were recruited. The former were assessed with GPL570 [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array.

Identification of SR genes

The limma package[15] in R was used to choose the SR genes in long-term survivors compared with those in short-term survivors. While a gene symbol may correspond to multiple probes (expression values), the max value is the final expression value for that mRNA. The cutoff criteria were p-value < 0.05 and |log fold change (FC)| > 0.5. The ggplot2 and pheatmap packages of R were applied for the visualization of SR genes in a volcano plot and a heatmap.

Functional enrichment analysis

Gene set enrichment analysis (GSEA, version 4.1.0) is a computational method that determines whether an a priori defined set of genes shows statistically significant, concordant differences between two biological states[16]. In this study, the hallmark gene set of Molecular Signatures Database v7.2 of GSEA was used to explore the enrichment of gene sets among the detected genes.

To obtain deep insight into the biological functions of critical genes, Gene Ontology (GO) annotation of SR genes was performed with Metascape[17] (<http://metascape.org/>). GO functional analysis contains biological processes (BP), molecular functions (MF), and cellular components (CC). The human genome (*Homo sapiens*) was selected as the background variable. Gene count ≥ 3 and p-value < 0.05 were set as the threshold. The results were visualized using Prism 8.

Protein-protein interaction network (PPI) construction

The PPI network of SR genes was constructed using the Search Tool for the Retrieval of Interacting Genes (STRING) database (<https://string-db.org/>). STRING is a biological database designed to construct a PPI network of SR genes based on known and predicted PPIs and then analyse the functional interactions between proteins[18]. The criterion was set at confidence score greater than 0.7. The Molecular Complex Detection (MCODE) plug-in of Cytoscape (version 3.9.0) was applied to screen the key modules of the PPI network[19]. The advanced options were set as degree cutoff = 3, K-Core = 2, and node score cutoff = 0.2.

Expression and survival analysis of collagen genes

The analysis of the relative expression of the four collagen genes was performed using UALCAN (<http://ualcan.path.uab.edu>), a user-friendly, interactive web resource for analysing cancer transcriptome data[20]. Furthermore, the ggplot2 and ggpvr packages of R were applied for the visualization of collagen gene expression between long-term survivors and short-term survivors in dataset GSE53733.

GlioVis[21] (<http://gliovis.bioinfo.cnio.es/>) is a user-friendly web application for data visualization and analysis used to explore brain tumour expression datasets, which is mainly based on the Chinese Glioma Genome Atlas (CGGA) and The Cancer Genome Atlas (TCGA) databases and other researcher's data. In this study, we analysed the relation of OS with the expression of collagen genes through Kaplan-Meier (KM) survival estimates in the TCGA-GBM dataset.

Correlation of collagens with immune cell infiltration

The Tumor Immune Estimation Resource (TIMER) database (<https://cistrome.shinyapps.io/timer/>) is an online tool for systematic analysis of immune cell infiltration across diverse cancer types from TCGA[22]. TIMER uses a deconvolution algorithm to estimate the abundance of tumour-infiltrating immune cells (TIICs) based on gene expression profiles[23]. First, we evaluated the association between collagens and tumour purity and the association between clinical outcome and the abundance of immune infiltrates in GBM. Then, we explored the correlation of collagen gene expression with the abundance of immune cell infiltration.

Gene-drug interaction network analysis

The gene-drug interaction network was constructed using the Comparative Toxicogenomics Database[24] (CTD) for chemotherapeutic drugs that could decrease or affect the mRNA, protein expression or methylation of genes.

Statistical analysis

Statistical analysis was performed with R (version 4.0.1). Student's t-tests were utilized for the comparison of two sample groups. Differences were considered statistically significant when $p < 0.05$.

Results

SR genes between long-term survivors and short-term survivors in GBM

First, we analysed the physiological processes involving the detected genes using GSEA. Epithelial mesenchymal transition (EMT), TNF- α signalling via the NF- κ B and the G2M checkpoint were enriched in short-term survivors, and oxidative phosphorylation was enriched in long-term survivors (Fig. 1a). According to the cutoff of a logFC value of 0.5, a total of 220 SR genes were identified, including 78 upregulated genes in long-term survivors, which were defined as positive SR genes (PSR genes), and 142 downregulated genes in long-term survivors, which were defined as NSR genes (Fig. 1b&c). Since EMT is a vital process associated with cancer prognosis[25] and had the highest score in GSEA, we analysed the genes involved in NSR and PSR genes. As shown in Fig. 1d, 11 NSR genes were enriched in the EMT gene set, while no PSR genes were enriched. The 11 enriched genes included COL1A2, COL6A2, COL8A2, FBN2, FMOD, LAMA1, PCOLCE, PLOD2, SERPINE1, TFPI2, and VEGFA.

Functional enrichment analysis of SR genes

To further determine the significance of SR genes in GBM, gene enrichment analysis was performed based on the PSR genes and NSR genes. The NSR genes were mainly enriched in embryonic organ development, embryonic morphogenesis, and skeletal system development in the BP terms; extracellular matrix (ECM), collagen-containing ECM, and basement membrane in the CC terms; and ECM structural constituent, proximal promoter sequence-specific DNA binding, and ECM structural constituent conferring tensile strength in the MF terms. Accordingly, the PSR genes were mainly associated with neuron

projection morphogenesis, plasma membrane bounded cell projection morphogenesis, cell projection morphogenesis in BP gene sets; extracellular matrix, axon, neuronal cell body in the CC gene sets; and potassium ion transmembrane transporter activity, voltage-gated potassium channel activity, cation channel activity in the MF gene sets (Fig. 2b). These results indicated that EMT-related processes are the core mechanism by which NSR genes affect GBM prognosis. To further confirm these findings, a protein-protein interaction (PPI) network derived from NSR genes and PSR genes was mapped using the STRING database (Fig. 2c). Interestingly, among the resulting nodes, two nodes with higher scores were screened out by MCODE (Fig. 2d, e). COL1A2, COL6A2, COL8A1, COL8A2 and PLOD2 comprised module 1 with a score of 5.0 (Fig. 2d). HOXB3, HOXA5, HOXB5, and HOXB7 comprised module 2 with a score of 4.0 (Fig. 2e). All the genes of module 1 and module 2 were derived from NSR genes.

Expression and survival analysis of collagen genes

Noting that COL1A2, COL6A2 and COL8A2 are EMT and significant PPI node participants, we further explored whether these genes predict poor outcomes in GBM. As shown in Fig. 3a, the expression of the three mentioned collagen proteins was markedly downregulated in short-term survivors ($p < 0.05$). Moreover, expression analysis based on TCGA-GBM RNAseq datasets shown that these genes are overexpressed in primary tumour tissues compared to normal tissues (Fig. 3b). Since it belongs to the collagen proteins, COL8A1 was also assessed. High expression of these genes was significantly associated with poor prognosis (Fig. 3c), which was consistent with their inclusion in the NSR gene set.

The expression of collagen genes is correlated with immune cell infiltration in GBM

The process of EMT is closely related to tumour heterogeneity, which contributes to resistance and treatment failure[26]. To explore whether the screened collagen genes are associated with the tumour heterogeneity of GBM, we examined the correlation of the tumour purity of GBM tissues and collagen proteins using TIMER. Although no significant difference in tumour purity was found between the short-term and long-term samples (Supplementary Fig. 1a), the expression of three screened collagen genes, COL1A2 ($r = -0.320$, $p < 0.05$), COL6A2 ($r = -0.261$, $p < 0.05$), and COL8A2 ($r = -0.369$, $p < 0.05$), was negatively correlated with tumour purity, suggesting that these three collagen genes are involved in tumour heterogeneity (Fig. 4a). The aggregation of TIICs is a key component of the tumour microenvironment. TIICs interact with tumour cells and domesticate each other, forming a community to maintain the malignant phenotype and heterogeneity[27]. We analysed the association between immune cells and survival in GBM. As shown in Fig. 4b, only dendritic cell (DC) infiltration was significantly associated with survival. Lower DC infiltration predicts better survival outcomes. Furthermore, the expression of COL1A2 ($r = 0.392$, $p < 0.05$), COL6A2 ($r = 0.461$, $p < 0.05$), and COL8A2 ($r = 0.350$, $p < 0.05$) was positively correlated with the infiltration level of DCs (Fig. 4c). To further confirm the findings, the correlation of somatic copy number of collagen proteins and DC infiltration was analysed. As shown in Fig. 4d, arm-level gain of COL1A2 and arm-level deletion of COL6A2 were associated with DC infiltration. The immune functions of tumour-infiltrating DCs are performed by activated DCs. Moreover, the expression of COL6A2 was negatively correlated with myeloid DC activation ($\rho = -0.182$, $p < 0.05$).

Collectively, these data suggested that DC infiltration predicts poor outcomes in GBM patients and that the elevation of COL6A2 possibly regulates activated DC infiltration.

COL6A2-drug interaction network analysis

Based on the results of infiltration level analysis, COL6A2 is suggested to be the main regulator of DC-mediated immune energy in GBM. Overall, we screened 9 drugs that can reduce the expression or affect secretion of COL6A2 protein, including erianin, sodium fluoride and triclosan. The screening results also shown that 31 drugs can reduce the expression of COL6A2 mRNA, including cisplatin, clofibrate and cytarabine. In addition, there were 5 drugs that increase or affect the methylation of COL6A2, including valproic acid, sodium arsenite and aflatoxin B2 (Fig. 5), indicating that these drugs could potentially improve survival for GBM patients.

Discussion

In the present study, we identified GBM SR genes and uncovered that NSR genes regulate EMT processes in glioblastoma. Among the NSR genes, collagen genes, including COL1A2, COL6A2, and COL8A2, predict high-risk GBM with shorter survival time and are related to enhanced tumour heterogeneity and DC infiltration.

Tumour prognosis involves multiple mechanisms. EMT is an essential process associated with metastasis and drug resistance in cancer[28, 29]. In this study, we found that genes upregulated in short-term survivors were enriched in EMT, which had the highest normalized enrichment score (NES). EMT involves cell-cell and cell-extracellular matrix interactions[30]. Furthermore, we screened 220 SR genes and found that 11 EMT gene set members were integrated with NSR genes, including three matrix proteins. Among SR genes, the NSR genes were mainly associated with multicellular organism development and ECM organization. The ECM is a crucial structure for tumours that can promote the growth, survival, and invasion of tumours and modify fibroblast and immune cell behaviour[31, 32]. Moreover, the matrix and stromal cells can also modulate the efficacy of therapy, and methods that reshape the tumour matrix could improve the outcomes of patients[33, 34]. Thus, the ECM might be an interesting direction for exploring hub SR genes. In this study, the PPI network was constructed and modules were identified with the MCODE plug-in of Cytoscape. The results implied that the collagen proteins COL1A2, COL6A2, COL8A1, and COL8A2 were hub SR genes, and further investigations were carried out.

The collagen superfamily is the most important group of ECM proteins and is characterized by three signature features and comprises 28 members. Studies have revealed that collagens promote the proliferation, metastasis and invasion of cancers[35, 36]. In addition, overexpression of collagens can enhance the resistance of cancers to chemotherapy drugs, leading to poor prognosis[37, 38]. In this study, we found that COL1A2, COL6A2, COL8A1, and COL8A2 were overexpressed in GBM compared to normal tissue and that they were downregulated in long-term survivors compared to short-term survivors, which confirmed that the four mentioned collagen proteins are oncogenes and SR genes. Collagens interact

with cancer cells through receptors and play a crucial role[39]. Discoidin domain receptors (DDR) are a subfamily of tyrosine kinases that are activated by collagens. Research has shown that the activation of DDR2 by COL1 regulates SNAIL1 stability and promotes breast cancer cell invasion and migration[40]. Others further reported that the binding of COL11A1 to DDR2 activates Src-PI3K/Akt-NF- κ B signalling and inhibits cisplatin-induced apoptosis in ovarian cancer cells[41]. In addition to DDRs, integrin is also a receptor of collagens, and the binding of integrin to COL1 might enhance the proliferation and invasion of squamous cell carcinoma cells via the MEK/ERK signalling pathway[42]. Collectively, it seems that collagens promote the progression of tumours via multiple approaches.

Collagens bind to receptors that correlate with cancer behaviours and prognosis. Collagens combine with immune cells and comprise the main component of the TME. However, the correlation of collagen expression with immune cell infiltration in GBM is rarely known. Studies have shown that immune cell infiltration is highly relevant to antitumor responses and prognosis[43, 44]. In this study, we estimated the association of six immune infiltrates with patient survival. We found that tumour-infiltrating DCs (TIDCs) are significantly associated with clinical outcome in GBM, and a high level of DC infiltration predicts poor cumulative survival. In general, higher survival rates were found in patients harbouring larger populations of DCs and T lymphocytes[45, 46]. However, due to the presence of suppressive immune cells, especially Tregs, tumour cells are able to escape the immune system[47, 48]. In DC-mediated tumour immunity, the prognostic impact is related to the TIDC phenotype; mature DCs have been considered immune stimulatory, whereas immature DCs have been considered suppressive and tolerogenic[49]. In the current study, we found that the expression of COL6A2 was negatively correlated with the abundance of activated DCs. DC activation is suppressed by tumour-derived molecules, including PD-L1 and Tim3[50, 51]. A recent study shown that VEGF can impair the migration capacity and immune function of mature DCs and contribute to immunosuppression[52]. In our study, we found that the expression of COL1A2, COL6A2, and COL8A2 was positively correlated with the DC infiltration level. A previous study demonstrated that collagens could promote DC survival and promote the maturation of monocyte-derived DCs via osteoclast-associated receptors[53]. In this study, COL6A2 increased the level of DC infiltration but decreased the level of activated DC infiltration. Thus, it might be the main factor that leads immune escape and poor prognosis.

The above conclusions indicate that COL1A2, COL6A2, COL8A1, and COL8A2 are oncogenes and SR genes. Recent studies have shown that depleting collagens improves the therapeutic efficacy of antitumour drugs[54–56]. In this study, we found that cisplatin, clofibrate and cytarabine can decrease the expression of COL6A2 mRNA, while valproic acid can increase the methylation of COL6A2. Cisplatin is a first-line chemotherapy drug and is used in various tumours, and valproic acid is a selective inhibitor of histone deacetylase. Barne's research shown that valproic acid inhibits stromal cell function and exerts anticancer effects[57]. In GBM, valproic acid is considered a favourable SR drug that acts through multiple mechanisms[58–61].

In conclusion, our results highlight the key roles of collagens in GBM prognosis and immune cell infiltration. Furthermore, we reveal the potential mechanism through which valproic acid regulates GBM

progression via COL6A2 methylation. Future studies will aim to investigate additional mechanisms involved in collagen methylation and tumorigenesis.

Conclusions

In this study, we screened SR genes by comparing the long-term and short-term survivors in the database GSE53733. Through multiple bioinformatic analysis, it determined that short-survival genes were mainly associated with EMT associated process. We identified three collagen genes were most significantly associated with EMT and short survival regulation in GBM. Further analysis indicated that COL1A2, COL6A2 and COL8A2 negatively correlated to tumour purity and DCs infiltration. Survival analysis uncovered that DCs infiltration is a risk factor for GBM, and A novel finding is that COL6A2 was significant correlation with DCs infiltration and leading adverse prognosis.

Abbreviations

BP biological processes; CAFs:cancer-associated fibroblasts; CC:cellular components; CGGA:Chinese Glioma Genome Atlas; CT:computed tomography; CTD:Comparative Toxicogenomics Database; DC:dendritic cell; DDRs:Discoidin domain receptors; ECM:extracellular matrix; EMT:epithelial mesenchymal transition; GBM:glioblastoma multiforme; GEO:Gene Expression Omnibus; GO:Gene Ontology; GSEA:Gene set enrichment analysis; IDH1/2:isocitrate dehydrogenase 1/2; IDO:indoleamine-2,3-dioxygenase; KM:Kaplan-Meier; MCODE:Molecular Complex Detection; MF:molecular functions; MGMT:methylation of O6-methylguanine-DNA methyltransferase; MRI:magnetic resonance imaging; NSR:negative survival-related; OS:overall survival; PD-L1:programmed death ligand 1; PPI:protein-protein interaction; SR:survival-related; STRING:Search Tool for the Retrieval of Interacting Genes; TCGA:The Cancer Genome Atlas; TIICs:Tumour-infiltrating immune cells; TIMER:Tumor Immune Estimation Resource; TME:tumour microenvironment; Tregs:regulatory T cells; TIDCs:tumour-infiltrating DCs

Declarations

Acknowledgements

We thank AJE (<https://www.aje.com>)for assistance with language editing (verification code 57F7-7A7C-C411-57F6-C20P).

Authors' Contribution

Jian Zou, Ying Yin, and Zhenkun Yang conceived and designed the study. Zhenkun Yang, Bo Zhang performed the bioinformatics analysis. Zhenhao Zhang, Jingjing Wang and Yaling Hu prepared the figures. Zhenkun Yang wrote the manuscript. Zhening Pu, Ying Yin and Jian Zou revised the manuscript. All authors approved the final version of the manuscript.

Funding

This work was supported by Natural Science Foundation of China (NFSC) grants (numbers 81872056, 81802493), “333” Engineering Project Jiangsu province ((2016) III-0605), Medical Young Talents Program of Jiangsu Province (QNRC2016188).

Availability of data and materials

Not applicable

Ethics approval

Not applicable

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interests.

References

1. Ostrom QT, Cioffi G, Gittleman H, Patil N, Waite K, Kruchko C, et al. Cbtrus statistical report: Primary brain and other central nervous system tumors diagnosed in the united states in 2012–2016. *Neuro Oncol*. 2019;21:v1–100.
2. Louis DN, Perry A, Reifenberger G, von Deimling A, Figarella-Branger D, Cavenee WK, et al. The 2016 world health organization classification of tumors of the central nervous system: A summary. *Acta Neuropathol*. 2016;131:803–20.
3. Lee JH, Liu R, Li J, Wang Y, Tan L, Li XJ, et al. Egfr-phosphorylated platelet isoform of phosphofructokinase 1 promotes pi3k activation. *Mol Cell*. 2018;70:197–210 e7.
4. An Z, Aksoy O, Zheng T, Fan QW, Weiss WA. Epidermal growth factor receptor and egfrviii in glioblastoma: Signaling pathways and targeted therapies. *Oncogene*. 2018;37:1561–75.
5. Talasila KM, Soentgerath A, Euskirchen P, Rosland GV, Wang J, Huszthy PC, et al. Egfr wild-type amplification and activation promote invasion and development of glioblastoma independent of angiogenesis. *Acta Neuropathol*. 2013;125:683–98.
6. Delgado-Lopez PD, Corrales-Garcia EM. Survival in glioblastoma: A review on the impact of treatment modalities. *Clin Transl Oncol*. 2016;18:1062–71.
7. Klopfenstein Q, Truntzer C, Vincent J, Ghiringhelli F. Cell lines and immune classification of glioblastoma define patient's prognosis. *Br J Cancer*. 2019;120:806–14.
8. Lyssiotis CA, Kimmelman AC. Metabolic interactions in the tumor microenvironment. *Trends Cell Biol*. 2017;27:863–75.

9. Giraldo NA, Sanchez-Salas R, Peske JD, Vano Y, Becht E, Petitprez F, et al. The clinical role of the tme in solid cancer. *Br J Cancer*. 2019;120:45–53.
10. Gajewski TF. Failure at the effector phase: Immune barriers at the level of the melanoma tumor microenvironment. *Clin Cancer Res*. 2007;13:5256–61.
11. Overman MJ, Lonardi S, Wong KYM, Lenz HJ, Gelsomino F, Aglietta M, et al. Durable clinical benefit with nivolumab plus ipilimumab in DNA mismatch repair-deficient/microsatellite instability-high metastatic colorectal cancer. *J Clin Oncol*. 2018;36:773–9.
12. Hellmann MD, Callahan MK, Awad MM, Calvo E, Ascierto PA, Atmaca A, et al. Tumor mutational burden and efficacy of nivolumab monotherapy and in combination with ipilimumab in small-cell lung cancer. *Cancer Cell* 2018; 33:853–61 e4.
13. Mascaux C, Angelova M, Vasaturo A, Beane J, Hijazi K, Anthoine G, et al. Immune evasion before tumour invasion in early lung squamous carcinogenesis. *Nature*. 2019;571:570–5.
14. Tanaka A, Sakaguchi S. Regulatory t cells in cancer immunotherapy. *Cell Res*. 2017;27:109–18.
15. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential expression analyses for rna-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43:e47.
16. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005;102:15545–50.
17. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun*. 2019;10:1523.
18. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, et al. String v10: Protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res*. 2015;43:D447-52.
19. Bader GD, Hogue CW. An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics*. 2003;4:2.
20. Chandrashekar DS, Bashel B, Balasubramanya SAH, Creighton CJ, Ponce-Rodriguez I, Chakravarthi B, et al. Ualcan: A portal for facilitating tumor subgroup gene expression and survival analyses. *Neoplasia*. 2017;19:649–58.
21. Bowman RL, Wang Q, Carro A, Verhaak RG, Squatrito M. Gliovis data portal for visualization and analysis of brain tumor expression datasets. *Neuro Oncol*. 2017;19:139–41.
22. Li T, Fan J, Wang B, Traugh N, Chen Q, Liu JS, et al. Timer: A web server for comprehensive analysis of tumor-infiltrating immune cells. *Cancer Res*. 2017;77:e108–10.
23. Li B, Severson E, Pignon JC, Zhao H, Li T, Novak J, et al. Comprehensive analyses of tumor immunity: Implications for cancer immunotherapy. *Genome Biol*. 2016;17:174.
24. Davis AP, Wieggers TC, Johnson RJ, Lay JM, Lennon-Hopkins K, Saraceni-Richards C, et al. Text mining effectively scores and ranks the literature for improving chemical-gene-disease curation at the comparative toxicogenomics database. *PLoS One*. 2013;8:e58201.

25. Pastushenko I, Blanpain C. Emt transition states during tumor progression and metastasis. *Trends Cell Biol.* 2019;29:212–26.
26. Teeuwssen M, Fodde R. Cell heterogeneity and phenotypic plasticity in metastasis formation: The case of colon cancer. *Cancers (Basel)* 2019; 11.
27. Balatoni T, Mohos A, Papp E, Sebestyen T, Liskay G, Olah J, et al. Tumor-infiltrating immune cells as potential biomarkers predicting response to treatment and survival in patients with metastatic melanoma receiving ipilimumab therapy. *Cancer Immunol Immunother.* 2018;67:141–51.
28. Gugnoni M, Sancisi V, Gandolfi G, Manzotti G, Ragazzi M, Giordano D, et al. Cadherin-6 promotes emt and cancer metastasis by restraining autophagy. *Oncogene.* 2017;36:667–77.
29. Du B, Shim JS. Targeting epithelial-mesenchymal transition (emt) to overcome drug resistance in cancer. *Molecules* 2016; 21.
30. Dongre A, Weinberg RA. New insights into the mechanisms of epithelial-mesenchymal transition and implications for cancer. *Nat Rev Mol Cell Biol.* 2019;20:69–84.
31. Kai F, Drain AP, Weaver VM. The extracellular matrix modulates the metastatic journey. *Dev Cell.* 2019;49:332–46.
32. Mierke CT. The matrix environmental and cell mechanical properties regulate cell migration and contribute to the invasive phenotype of cancer cells. *Rep Prog Phys.* 2019;82:064602.
33. Hirata E, Sahai E. Tumor microenvironment and differential responses to therapy. *Cold Spring Harb Perspect Med* 2017; 7.
34. Vennin C, Murphy KJ, Morton JP, Cox TR, Pajic M, Timpson P. Reshaping the tumor stroma for treatment of pancreatic cancer. *Gastroenterology.* 2018;154:820–38.
35. Liu J, Shen JX, Wu HT, Li XL, Wen XF, Du CW, et al. Collagen 1a1 (col1a1) promotes metastasis of breast cancer and is a potential therapeutic target. *Discov Med.* 2018;25:211–23.
36. Kumagai Y, Nio-Kobayashi J, Ishida-Ishihara S, Tachibana H, Omori R, Enomoto A, et al. The intercellular expression of type-xvii collagen, laminin-332, and integrin-beta1 promote contact following during the collective invasion of a cancer cell population. *Biochem Biophys Res Commun.* 2019;514:1115–21.
37. Zhang H, Fredericks T, Xiong G, Qi Y, Rychahou PG, Li JD, et al. Membrane associated collagen xiii promotes cancer metastasis and enhances anoikis resistance. *Breast Cancer Res.* 2018;20:116.
38. Walsh AJ, Cook RS, Lee JH, Arteaga CL, Skala MC. Collagen density and alignment in responsive and resistant trastuzumab-treated breast cancer xenografts. *J Biomed Opt.* 2015;20:26004.
39. Orgel J, Madhurapantula RS. A structural prospective for collagen receptors such as ddr and their binding of the collagen fibril. *Biochim Biophys Acta Mol Cell Res.* 2019;1866:118478.
40. Zhang K, Corsa CA, Ponik SM, Prior JL, Piwnica-Worms D, Eliceiri KW, et al. The collagen receptor discoidin domain receptor 2 stabilizes snail1 to facilitate breast cancer metastasis. *Nat Cell Biol.* 2013;15:677–87.

41. Rada M, Nallanthighal S, Cha J, Ryan K, Sage J, Eldred C, et al. Inhibitor of apoptosis proteins (iaps) mediate collagen type xi alpha 1-driven cisplatin resistance in ovarian cancer. *Oncogene*. 2018;37:4809–20.
42. Hayashido Y, Kitano H, Sakaue T, Fujii T, Suematsu M, Sakurai S, et al. Overexpression of integrin alphav facilitates proliferation and invasion of oral squamous cell carcinoma cells via mek/erk signaling pathway that is activated by interaction of integrin alphavbeta8 with type collagen. *Int J Oncol*. 2014;45:1875–82.
43. Rohr-Udilova N, Klinglmüller F, Schulte-Hermann R, Stift J, Herac M, Salzmann M, et al. Deviations of the immune cell landscape between healthy liver and hepatocellular carcinoma. *Sci Rep*. 2018;8:6220.
44. Law AM, Lim E, Ormandy CJ, Gallego-Ortega D. The innate and adaptive infiltrating immune systems as targets for breast cancer immunotherapy. *Endocr Relat Cancer*. 2017;24:R123–44.
45. Zhang Z, Huang J, Zhang C, Yang H, Qiu H, Li J, et al. Infiltration of dendritic cells and t lymphocytes predicts favorable outcome in epithelial ovarian cancer. *Cancer Gene Ther*. 2015;22:198–206.
46. Muppa P, Parrilha Terra SBS, Sharma A, Mansfield AS, Aubry MC, Bhinge K, et al. Immune cell infiltration may be a key determinant of long-term survival in small cell lung cancer. *J Thorac Oncol*. 2019;14:1286–95.
47. Wing JB, Tanaka A, Sakaguchi S. Human foxp3(+) regulatory t cell heterogeneity and function in autoimmunity and cancer. *Immunity*. 2019;50:302–16.
48. Wolf D, Soppe S, Pircher A, Gastl G, Wolf AM. Treg(s) in cancer: Friends or foe? *J Cell Physiol*. 2015;230:2598–605.
49. Tran Janco JM, Lamichhane P, Karyampudi L, Knutson KL. Tumor-infiltrating dendritic cells in cancer pathogenesis. *J Immunol*. 2015;194:2985–91.
50. Fu C, Jiang A. Dendritic cells and cd8 t cell immunity in tumor microenvironment. *Front Immunol*. 2018;9:3059.
51. Koelblinger P, Emberger M, Drach M, Cheng PF, Lang R, Levesque MP, et al. Increased tumour cell pd-l1 expression, macrophage and dendritic cell infiltration characterise the tumour microenvironment of ulcerated primary melanomas. *J Eur Acad Dermatol Venereol*. 2019;33:667–75.
52. Long J, Hu Z, Xue H, Wang Y, Chen J, Tang F, et al. Vascular endothelial growth factor (vegf) impairs the motility and immune function of human mature dendritic cells through the vegf receptor 2-rhoa-cofilin1 pathway. *Cancer Sci*. 2019;110:2357–67.
53. Schultz HS, Nitze LM, Zeuthen LH, Keller P, Gruhler A, Pass J, et al. Collagen induces maturation of human monocyte-derived dendritic cells by signaling through osteoclast-associated receptor. *J Immunol*. 2015;194:3169–79.
54. Fang L, Kong SS, Zhong LK, Wang CM, Liu YJ, Ding HY, et al. Asiatic acid enhances intratumor delivery and the antitumor effect of pegylated liposomal doxorubicin by reducing tumor-stroma collagen. *Acta Pharmacol Sin*. 2019;40:539–45.

55. Chakravarthy D, Munoz AR, Su A, Hwang RF, Keppler BR, Chan DE, et al. Palmatine suppresses glutamine-mediated interaction between pancreatic cancer and stellate cells through simultaneous inhibition of survivin and col1a1. *Cancer Lett.* 2018;419:103–15.
56. Tang Y, Liu Y, Wang S, Tian Y, Li Y, Teng Z, et al. Depletion of collagen by losartan to improve tumor accumulation and therapeutic efficacy of photodynamic nanoplateforms. *Drug Deliv Transl Res.* 2019;9:615–24.
57. Barneh F, Salimi M, Goshadrou F, Ashtiani M, Mirzaie M, Zali H, et al. Valproic acid inhibits the protective effects of stromal cells against chemotherapy in breast cancer: Insights from proteomics and systems biology. *J Cell Biochem.* 2018;119:9270–83.
58. Proske J, Walter L, Bumès E, Hutterer M, Vollmann-Zwerenz A, Eyupoglu IY, et al. Adaptive immune response to and survival effect of temozolomide- and valproic acid-induced autophagy in glioblastoma. *Anticancer Res.* 2016;36:899–905.
59. Garcia CG, Kahn SA, Geraldo LHM, Romano I, Domith I, Silva D, et al. Combination therapy with sulfasalazine and valproic acid promotes human glioblastoma cell death through imbalance of the intracellular oxidative response. *Mol Neurobiol.* 2018;55:6816–33.
60. Tran LNK, Kichenadasse G, Sykes PJ. Combination therapies using metformin and/or valproic acid in prostate cancer: Possible mechanistic interactions. *Curr Cancer Drug Targets.* 2019;19:368–81.
61. Riva G, Cilibrasi C, Bazzoni R, Cadamuro M, Negroni C, Butta V, et al. Valproic acid inhibits proliferation and reduces invasiveness in glioma stem cells through wnt/beta catenin signalling activation. *Genes (Basel)* 2018; 9.

Figures

Figure 1

Identification of SR genes between long-term survivors and short-term survivors. (a) Top 5 gene sets that were significantly enriched for detected genes by GSEA. (b) Volcano plot of SR genes in GSE53733. (c) Heatmap showing the expression of all NSR genes and PSR genes. (d) Enrichment plot of EMT and the integrated genes between the gene set and NSR genes.

Figure 2

GO enrichment analysis and PPI network construction for NSR and PSR genes by Metascape and STRING. (a) NSR gene enrichment analysis included BPs, CCs and MFs. (b) PSR gene enrichment analysis included BPs, CCs and MFs. The top 10 terms were chosen. (c) PPI network constructed with proteins encoded by all NSR genes and PSR genes. (d) Module 1 with an MCODE score of 5.0. (e) Module 2 with an MCODE score of 4.0.

Figure 3

Collagen gene over expression is associated with poor prognosis in patients with GBM. (a) Comparison of collagen gene transcript levels between short-term and long-term survivors in the GSE53733 dataset. (b) Comparison of collagen gene transcript levels between tumour samples and normal tissues in the TCGA-GBM RNAseq dataset. The results were derived from the Gene-Cloud of Biotechnology Information

(GCB) based on RNAseqV2 datasets. (c) Prognostic analysis of collagen genes by KM survival estimates in GlioVis.

Figure 4

Collagen genes are associated with infiltration levels of immune cells in GBM. (a) The gene expression of collagens, beside COL8A1, is negatively correlated with tumour purity. (b) High-level infiltration of DCs is associated with poor prognosis in GBM. (c) The gene expression of collagens, including COL8A1, is

positively correlated with the infiltration levels of DCs. (d) Correlation of collagen copy number in glioblastoma tissues with infiltration levels of DCs.(e)Expression of COL6A2 is negatively correlated with myeloid DC activation.

Figure 5

Gene-drug interaction network constructed with COL6A2 and chemotherapeutic drugs. Overview of drugs that decrease protein and mRNA expression and affect methylation of COL6A2. Drugs that decrease the mRNA expression were exhibited for 15 terms.

Supplementary Files

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

- [supplementfiguresNI.pdf](#)