Bioinformatics identification of potential protein glycosylation genes associated with glioma stem cell signature

Tokumura Kazuya
Gifu Pharmaceutical University
Koki Sadamori
Gifu Pharmaceutical University
Makoto Yoshimoto
Gifu Pharmaceutical University
Jiajun Lyu
Gifu Pharmaceutical University
Yuki Tanaka
Gifu Pharmaceutical University
Kazuya Fukasawa
Gifu Pharmaceutical University
Eiichi Hinoi (hinoi-e@gifu-pu.ac.jp)
Gifu Pharmaceutical University

Research Article

Keywords: Glioma stem cells, glioblastoma, protein glycosylation, Gene Expression Omnibus, The Cancer Genome Atlas, Chinese Glioma Genome Atlas

Posted Date: February 7th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2485755/v1

Abstract

Glioma stem cells (GSCs) contribute to the pathogenesis of glioblastoma (GBM), which is the most malignant form of glioma. The implications and underlying mechanisms of protein glycosylation in GSC phenotypes and GBM malignancy are not fully understood. In this study, we aimed to investigate the implication of protein glycosylation and the corresponding candidate genes on the stem cell properties of GSCs and poor clinical outcomes in GBM, using single-cell RNA sequencing and bulk RNA sequencing datasets of clinical GBM specimens deposited in the Gene Expression Omnibus database, in addition to The Cancer Genome Atlas and the Chinese Glioma Genome Atlas databases of patients with glioma. We demonstrated that N-linked glycosylation was significantly associated with GSC properties and the prognosis of GBM by conducting integrated bioinformatics analyses of clinical specimens. N-linked glycosylation was associated with the glioma grade, molecular biomarkers, and molecular subtypes. The expression levels of the asparagine-linked glycosylation (ALG) enzyme family, which is essential for the early steps in the biosynthesis of N-glycans, were prominently associated with GSC properties and poor survival in patients with GBM with high stem-cell properties. Finally, the oxidative phosphorylation (OXPHOS) pathway was primarily enriched in GSCs with a high expression of the ALG enzyme family. Collectively, these findings uncover a pivotal role for N-linked glycosylation in the regulation of GSC phenotypes and GBM malignancy through, possibly in part, the ALG-OXPHOS axis, thereby revealing a potential target for GSC-directed therapy.

Introduction

The World Health Organization defines glioblastoma (GBM) as a grade IV cancer, and GBM is the most malignant form of glioma (1, 2). GBM is one of the most aggressive and fatal types of central nervous system cancer (3). Isocitrate dehydrogenase (*IDH*) status is a prognostic biomarker in patients with GBM (4). GBM is also classified by the following molecular subtypes: mesenchymal, classical, proneural, and neural (5). Patients with GBM have a significantly poor prognosis and rarely exhibit long-term survival, despite recent advances in multimodality therapy with a combination of surgical operation, radiation therapy, chemotherapy, and molecular targeted therapy (6). Glioma stem cells (GSCs) exhibit stem cell-like properties, such as self-renewal capacity, ability to differentiate into non-GSCs, and tumor-propagating potential (7). GSCs play important roles in several events, such as therapeutic resistance (radioresistance and chemoresistance), rapid recurrence, cancer invasion, and tumor angiogenesis, indicating that targeting GSCs is an effective strategy for improving GBM treatment (8).

Protein glycosylation, a typical posttranslational modification, is a complex and multistep process involving various glycan-modifying enzymes, including glycosyltransferases and glycosidases (9, 10). It regulates a diverse range of fundamental cellular and biological pathways, including protein trafficking, signal transduction, pluripotency, proliferation, differentiation, and survival (11). The most abundant and commonly occurring types of protein glycosylation include *N*-linked (asparagine-linked) and *O*-linked (serine/threonine-linked) glycosylation (12, 13). Among the multiple dysregulated posttranslational modifications driving tumorigenicity, aberrant protein glycosylation, such as *N*-glycan branching and *O*-

glycan truncation, is a well-known hallmark of cancer that contributes to tumor development and progression (14-16). Aberrant glycosylation increases the expression of sialylated and fucosylated glycans, which are associated with poor prognosis in glioma, and the *O*-linked glycan signature modulates immune suppression in GBM (17-19). Moreover, glycosyltransferase 8 domain containing 1 (GLT8D1) is involved in GBM pathogenesis and is associated with the regulation of glioma cell migration (20). GLT8D1 promotes GSC maintenance by inhibiting cluster of differentiation (CD) 133 degradation through *N*-linked glycosylation, exhibiting correlation with a higher grade of glioma and worse clinical outcomes (21). Moreover, alpha-1,6-mannosylglycoprotein 6-beta-*N*-acetylglucosaminyltransferase A (MGAT5), which catalyzes multibranched *N*-glycans, is a critical regulator of stiffness-driven invasion and GSC mechanotransduction (22).

Although extensive studies have been conducted to reveal the implication of protein glycosylation in the pathogenesis of glioma, limited evidence is available on the role of protein glycosylation alterations in stem-cell properties and aggressiveness of GSCs and GBM malignancy. We aimed to investigate the implication of protein glycosylation and the corresponding candidate genes on the stem cell properties of GSCs and poor clinical outcomes in GBM, using integrated bioinformatics analyses, single-cell RNA sequencing (scRNA-seq), and bulk RNA-seq datasets of clinical GBM specimens deposited in the Gene Expression Omnibus (GEO) database, in addition to The Cancer Genome Atlas (TCGA) and Chinese Glioma Genome Atlas (CGGA) databases of patients with glioma.

Materials And Methods

scRNA-seq data analysis

We used the gene expression data for patients with GBM (GSE84465) (23), which was downloaded from GEO. The downloaded data was analyzed using the Seurat package (ver. 4.2.1) on the R software (ver. 4.2.1). First, cells with more than 60 000, less than 600 expressed genes, or more than 100 000 counted genes were removed in advance. After performing sctransform normalization, dimensional compression was performed using principal component analysis (PCA) and uniform manifold approximation and projection (UMAP). Clustering was then performed using the Louvain algorithm, followed by identification of each cluster using known marker genes. Copy number variation (CNV) was performed using the inferCNV package (ver. 1.6.0). Wilcoxon's rank-sum test was performed to identify differentially expressed genes (DEGs) using the presto package (ver. 1.0.0). Single sample gene set enrichment analysis (ssGSEA) was performed using GSVA package (ver. 1.46.0). We defined the top 10 % of neoplastic tumor cells in the ssGSEA score of the "BEIER_GLIOMA_STEM_CELL_UP" gene set as cancer stem cell (CSC)-signature^{high} GBM cells and the rest as CSC-signature^{low} GBM cells. The gene sets were those deposited in MSigDB databases (http://www.gsea-msigdb.org/gsea/msigdb/index.jsp). Genes that were expressed in more than 10 % of cells and those that showed significant changes between the two groups (CSC-signature^{high} and CSC-signature^{low}) were identified as DEGs for genes related to the *N*-linked glycosylation pathway. GSEA was performed using the clusterProfiler package (ver. 4.6.0). The HALLMARK gene set is deposited in the H collection of MSigDB databases, the GOBP gene set in the C5

collection, and the KEGG gene set in the C2 collection. The KEGG-oxidative phosphorylation pathway was drawn using the pathview package (ver. 1.38.0).

Bulk RNA-seq data analysis

We performed gene expression analysis of GBM and non-tumor brain tissues from patients with GBM, using RNA-seq data (GSE33328, GSE48865, GSE59612, GSE62731, and GSE77530) (24-28), which were downloaded from the GEO using SRA Toolkit (ver. 2.10.4). Quality checks were performed with FastQC (ver. 0.11.8) and processed with Trimomatic (ver. 0.33) to exclude adapter sequences and low-quality bases. Clean reads were quantified at the transcript level against a human reference sequence (GRCh38 release 98) using Salmon (ver. 1.2.0). Conversion to gene expression and visualization were performed using the tximport (ver. 1.24.0) and ggplot2 (ver. 3.4.0) packages on the R software, respectively.

Additionally, we analyzed data from tissues of patients with glioma (e.g., grade, *IDH* mutation status, and subtype) in TCGA and CGGA. Statistical significance was determined using Wilcoxon's rank-sum test followed by Bonferroni's correction. For correlation analysis, we calculated Pearson's correlation coefficient.

Survival analysis

Gene expression and clinical data of patients with glioma were obtained from TCGA and CGGA database. We defined the top 50 % of patients with glioma in the ssGSEA score of the "BEIER_GLIOMA_STEM_CELL_UP" gene set as the CSC-signature^{high} group and the rest as the CSC-signature^{low} group. Survival analysis was performed using the log-rank test with the survival package, and Kaplan–Meier curves were plotted using ggplot2 and survminer (ver. 0.4.9) on the R software.

Data Availability Statement

The bioinformatics data used in this study are openly available in the Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/) and GlioVis (http://gliovis.bioinfo.cnio.es/) databases.

Results

N- and *C*-linked glycosylation pathways are associated with stem cell properties of GSCs

We first analyzed a scRNA-seq dataset of clinical GBM specimens deposited in the GEO database (GSE84465) to profile the properties of GSCs (Figure 1A). Seven clusters were successfully identified through UMAP analysis based on the genetic profile of the cell (Figure 1B), and canonical markers were used to annotate cell types: neoplastic cells (*EGFR*$^{+}$) and normal cells (myeloid [*PTPRC*$^{+}$], oligodendrocyte precursor cells [*GPR17*$^{+}$], oligodendrocytes [*MOG*$^{+}$], astrocytes [*AGXT2L1*$^{+}$], vascular cells [*DCN*$^{+}$], and neurons [*SYMN2*$^{+}$]) (Figure 1C). Subsequently, we distinguished malignant cells from non-malignant cells by CNV inference (Figure 1D).

We further divided neoplastic cells (GBM cell population) into two groups, CSC-signature^{high} GBM and CSC-signature^{low} GBM cells, on the basis of the gene set associated with “glioma stem cell” by ssGSEA (Figure 1E). We confirmed the enrichment of both gene sets involved in “stemness” and “stem cell” in CSC-signature^{high} GBM cells by ssGSEA and the shorter overall survival times in the CSC-signature^{high} patient group by Kaplan–Meier survival analysis of the TCGA and CGGA datasets, allowing us to define these cells as the GSC population (Figure 1F and 1G). Under these experimental conditions, among the gene sets associated with glycosylation, ssGSEA revealed a significant enrichment for gene sets related to the “glycosylation,” “*N*-linked glycosylation,” “deglycosylation,” and “*C*-linked glycosylation” in CSC-signature^{high} GBM cells (Figure 1H). In contrast, no significant enrichment was detected in gene sets involved in *O*-linked glycosylation in CSC-signature^{high} GBM cells (Figure 1H).

Collectively, these results indicate that the *N*- and *C*-linked glycosylation pathways are linked to the stem cell characteristics of GSCs.

N-linked glycosylation pathway in GSCs is linked to the aggressiveness and poor prognosis of GBM

Next, we analyzed five bulk-RNA-seq datasets of 107 patients with GBM (GSE33328, GSE48865, GSE59612, GSE62731, and GSE77530) and observed that gene sets involved in *N*- and *C*-linked glycosylation were more significantly upregulated in GBM tissues than in non-tumor brain tissues (Figure 2A). Consistent results were confirmed using the TCGA database (Figure 2B). Moreover, gene sets related to *N*- and *C*-linked glycosylation were positively associated with increased glioma grade (grades II, III, and IV) and were more significantly downregulated in patients with GBM harboring *IDH* mutation compared with that in patients exhibiting wildtype status, on the basis of the data from the TCGA and CGGA databases (Figure 2C–2F).Meanwhile, the *N*-linked glycosylation-related gene set was significantly higher in patients with GBM with the mesenchymal subtype, the most aggressive among the molecular GBM subtypes (5), compared with that in patients with classical and proneural subtypes, according to the data from the TCGA and CGGA databases (Figure 2G and 2H). However, expression of the *C*-linked glycosylation-related gene set was significantly increased in the classical subtype compared with that in the mesenchymal and proneural subtypes (Figure 2G and 2H).

Next, we assessed whether *N*- and *C*-linked glycosylation pathways in patients with glioma with higher stem cell properties were associated with poor prognosis. Kaplan–Meier survival analysis demonstrated that the CSC-signature^{high} patient group with elevated activity of the *N*-linked glycosylation pathway exhibited significantly shorter overall survival times than that with low glycosylation, according to the TCGA and CGGA databases (Figure 2I and 2J). We observed the same survival outcomes in patients with elevated activity of the *C*-linked glycosylation pathway (Figure 2I and 2J).

Collectively, these results suggest that the *N*-linked glycosylation pathway in GSCs is associated with malignancy, aggressiveness, and survival outcomes in patients with GBM.

Expression analysis of DEGs linked to *N*-linked glycosylation pathway in GSCs

Considering that a novel and valid association was found between *N*-linked glycosylation and molecular subtypes of GBM, in addition to the glioma grade, molecular biomarkers, and poor prognosis of GBM, we subsequently focused on the *N*-linked glycosylation pathway.

First, we identified DEGs related to the *N*-linked glycosylation pathway between CSC-signature^{high} GBM and CSC-signature^{low} GBM cells using the scRNA-seq dataset (GSE84465). Thirty-nine DEGs linked to the *N*-linked glycosylation pathway were screened, including five significantly upregulated genes and one significantly downregulated gene. Among the upregulated genes, *asparagine-linked glycosylation 1* (*ALG1*) and *ALG2*, essential enzymes for the biosynthesis of *N*-glycans at early stages (29), were the top two upregulated genes in CSC-signature^{high} GBM cells (Figure 3A). Moreover, the expression levels of *ALG6*, *ALG7*, and *ALG12* were significantly upregulated in CSC-signature^{high} GBM cells (Figure 3A). In contrast, the expression levels of other genes of the ALG family were not significantly changed in CSC-signature^{high} GBM cells (Figure 3B). Because all significantly upregulated genes related to the *N*-linked glycosylation pathway (*ALG1*, *ALG2*, *ALG6*, *ALG7*, and *ALG12*) belonged to the ALG enzyme family, we subsequently focused on the ALG enzyme family.

The expression levels of ALG enzyme family (*ALG1*, *ALG2*, *ALG6*, *ALG7*, and *ALG12*) were significantly more upregulated in GBM tissues compared to that in non-tumor brain tissues, according to analyses of five bulk-RNA-seq datasets and the TCGA database (Figure 3C and 3D). Moreover, the expression levels of ALG enzyme family were associated with increased glioma grade, and were significantly downregulated in patients with GBM harboring the *IDH* mutant, according to the data from the TCGA and CGGA databases (Figure 3E–3H). *ALG2* expression level was the highest in the mesenchymal subtype, according to the TCGA and CGGA databases (Figure 3I and 3J).

Expression of ALG enzyme family is associated with the stem cell properties of GSCs and poor prognosis of GBM

Next, we determined whether the expression levels of the ALG enzyme family were associated with the stem cell properties of GSCs using five bulk-RNA-seq datasets and the scRNA-seq dataset (GSE84465). A correlation analysis between the ALG family and stem cell markers revealed that the expression levels of *ALG1*, *ALG2*, *ALG6*, *ALG7*, and *ALG12* were positively correlated with that of *CD36*, *CD44*, *FUT4*, *MUC1*, *MYC* and *NES*in GBM specimens according to five bulk-RNA-seq datasets of patients with GBM (Figure 4A and Supplementary Figure 1A–C). In addition, gene sets involved in “stemness” and “stem cell” were significantly enriched in *ALG1*-, *ALG2*-, *ALG6*-, *ALG7*-, and *ALG12*-positive CSC-signature^{high} GBM cells, according to the scRNA-seq dataset (Figure 4C and 4D, and Supplementary Figure 1D–F).

Next, we assessed whether the expression levels of *ALG1*, *ALG2*, *ALG6*, *ALG7*, and *ALG12* in patients with glioma with higher stem cell properties were associated with poor prognosis. Kaplan–Meier survival analysis revealed that high expression levels of *ALG1*, *ALG2*, *ALG6*, *ALG7*, and *ALG12* were significantly

associated with poor prognosis in the CSC-signature^{high} patient group, according to the TCGA and CGGA databases (Figure 4E and 4F, and Supplementary Figure 1G and 1H).

Collectively, these results suggest that the ALG enzyme family in GSCs is prominently associated with the stem cell properties of GSCs and poor survival outcomes in patients with GBM.

Oxidative phosphorylation pathway forms a link between GSC properties and ALG enzyme family

We next examined the molecular mechanisms by which the ALG enzyme family is involved in the control of GSC characteristics and GBM phenotypes. GSEA showed that the expression of the gene set involved in the oxidative phosphorylation (OXPHOS) pathway was ranked the highest in *ALG1*-, *ALG2*-, *ALG6*-, and *ALG7*-expression^{positive} CSCs and the third highest in *ALG12*-expression^{positive} CSCs (Figure 5A and 5B, and Supplementary Figure 2A–C). OXPHOS, the major source of adenosine triphosphate (ATP) in several cancer types, regardless of the increased aerobic glycolysis, plays a crucial role in tumorigenesis and tumor progression and addresses the energy demands of CSCs, including GSCs (30-32). Consistently, enrichment of the OXPHOS pathway in ALG-expression^{positive} CSCs was verified by GESA using the GOBP and KEGG gene sets (Figure 5C and 5D, and Supplementary Figure 2D–F). Moreover, KEGG analysis of the OXPHOS pathway revealed the significant upregulation of complex I and IV in ALG-expression^{positive} GSCs, along with the multiple dysregulated genes for each complex (Figure 5E and 5F, and Supplementary Figure 2G–I).

The mechanistic target of rapamycin (mTORC1), a serine/threonine kinase signaling complex, contributes to the properties of CSCs including GSCs, and it is dysregulated in GBM (33-37). c-MYC, a well-known stem cell transcription factor, is required for self-renewal and the tumorigenic potential of GSCs (38, 39). The gene sets associated with the mTORC1 and MYC pathways were consistently enriched in all ALG-expression^{positive} GSCs (Figure 5A and 5B, and Supplementary Figure 2A–C).

Collectively, these findings raise the possibility that the OXPHOS pathway is implicated in the regulation of GSC characteristics by the ALG enzyme family.

Discussion

Posttranslational modifications are frequently altered in GBM and are crucial for modulating the stemness and tumorigenicity of GSCs (40). Recently, we demonstrated that phosphorylation of the c-Myc axis by CDK8 and the extracellular signal regulated kinase 5/signal transducer and activator of transcription 3 axis by mitogen-activated protein kinase 5 controls the stemness and tumorigenicity of GSCs, contributing to GBM tumorigenesis (41). In addition, ubiquitination of the transforming growth factor-β receptor/R-Smad axis by Smad ubiquitin regulatory factor 2, an E3 ubiquitin ligase, controls the properties of GSCs and GBM malignancy (42). Moreover, methylation of regulator of chromosome condensation 1 by protein arginine methyltransferase 6 regulates tumorigenicity and radiation response of GSCs (43). Moreover, SUMOylation of promyelocytic leukemia protein by small ubiquitin-like modifier 1

regulates GSC maintenance and high malignancy (44). GSC properties are modulated by enzymes related to protein glycosylation, such as GLT8D1 and MGAT5 (20-22); however, the role of protein glycosylation alterations on the stemness and tumorigenicity of GSCs has not been investigated in detail. Although *in vitro* and *in vivo* analyses should be performed to validate our current results, to the best of our knowledge, this is the first study to show using integrated bioinformatics analyses that *N*-linked glycosylation is validly and strongly associated with the maintenance of GSC characteristics and GBM malignancy through, at least in part, the ALG-OXPHOS axis.

The assembly of *N*-linked oligosaccharides in eukaryotic cells is initiated by the successive addition of two *N*-acetylglucosamine, nine mannose, and three glucose molecules to the dolichol phosphate on the endoplasmic reticulum membrane by various glycosyltransferases, before transferring to an asparagine residue of a target protein by oligosaccharyltransferase (45-47). The ALG enzyme family participates in the sequential step of dolichol-linked oligosaccharide biosynthesis (45). *ALG1* encodes a β-1,4-mannosyltransferase that catalyzes the addition of the first β-1,4 mannose to Gn_2-dolichol-phosphate (48). *ALG2* encodes an α-1,3 mannosyltransferase that catalyzes the addition of both the second and third mannose residues to M_1Gn_2-dolichol phosphate (49). Although ALG mutations cause a rare autosomal-recessive disorder along with serious systemic diseases (50), the role of the ALG enzyme family in GBM development and progression remains unknown. Here, we demonstrated that the expression levels of the ALG enzyme family were significantly upregulated in CSC-signature^{high} GBM cells among the *N*-linked glycosylation-related genes, with *ALG1* and *ALG2* being the top two upregulated genes; however, the potential *N*-linked glycosylation-related genes associated with GSC properties need to be further characterized in future studies.

N-linked glycosylation is a crucial posttranslational modification that contributes to the stemness properties of GSCs and GBM prognosis. The ALG enzyme family may play a pivotal role in GBM pathogenesis by modulating the OXPHOS pathway in GSCs. OXPHOS is a major source of ATP in several cancer types; however, aerobic glycolysis contributes to tumorigenesis and tumor progression (30). Several independent lines have indicated that OXPHOS plays a crucial role in addressing the energy demands of CSCs, including GSCs (31, 32). Notably, GSEA showed that the gene sets involved in the mTORC1 and MYC pathways were enriched in all ALG enzyme family-expression^{positive} CSCs (Figure 5 and Supplementary Figure 2). Therefore, despite the additional molecular pathways that need to be explored, we reveal the possible crucial role of the ALG-OXPHOS axis in the regulation of GSC properties via the *N*-linked glycosylation pathway. Our findings improve our understanding of the molecular mechanism underlying the maintenance of stemness and tumorigenicity of GSCs and suggest that *N*-linked glycosylation status can represent a novel target for drug development in the treatment of GBM and malignant tumors associated with the stemness and aggressiveness of CSCs in humans.

Declarations

Conflict of Interest

E. Hinoi reports grants from Japan Society for the Promotion of Science during the study. No other disclosures are reported.

Author Contributions

K.T., K.S. and E.H. conceived the project. K.T., K.S., M.Y., J.L., Y.T., and K.F. performed the experiments and analysis. K.T. and E.H. wrote the manuscript.

Funding

This work was supported in part by the Japan Society for the Promotion of Science (20H03407 to E.H.).

List of non-standard abbreviations

ALG, asparagine-linked glycosylation; CNV, copy number variation; DEG, differentially expressed gene; GBM, glioblastoma; GOBP, Gene Ontology Biological Process; GSC, glioma stem cells; KEGG, Kyoto Encyclopedia of Genes and Genomes; OXPHOS, oxidative phosphorylation; scRNA-seq, single-cell RNA sequencing; ssGSEA, single sample gene set enrichment analysis; UMAP, uniform manifold approximation and projection

Acknowledgments

We wish to thank Dr. T. Horie (Kanazawa Medical University) for providing technical training in RNA-seq data analysis. Bioinformatic analyses were performed using the super-computing resource provided by Human Genome Center, the Institute of Medical Science, the University of Tokyo.

References

1. Louis DN, Perry A, Wesseling P, Brat DJ, Cree IA, Figarella-Branger D, et al. The 2021 WHO Classification of Tumors of the Central Nervous System: a summary. *Neuro-oncology* (2021) 23(8):1231-51. Epub 2021/06/30. doi: 10.1093/neuonc/noab106. PubMed PMID: 34185076; PubMed Central PMCID: PMCPMC8328013.
2. Wen PY, Packer RJ. The 2021 WHO Classification of Tumors of the Central Nervous System: clinical implications. *Neuro-oncology* (2021) 23(8):1215-7. Epub 2021/06/30. doi: 10.1093/neuonc/noab120. PubMed PMID: 34185090; PubMed Central PMCID: PMCPMC8328017.
3. van Solinge TS, Nieland L, Chiocca EA, Broekman MLD. Advances in local therapy for glioblastoma - taking the fight to the tumour. *Nature reviews Neurology* (2022) 18(4):221-36. Epub 2022/03/13. doi: 10.1038/s41582-022-00621-0. PubMed PMID: 35277681.
4. Melhem JM, Detsky J, Lim-Fat MJ, Perry JR. Updates in IDH-Wildtype Glioblastoma. *Neurotherapeutics : the journal of the American Society for Experimental NeuroTherapeutics* (2022) 19(6):1705-23. Epub 2022/06/01. doi: 10.1007/s13311-022-01251-6. PubMed PMID: 35641844; PubMed Central PMCID: PMCPMC9154038.

5. Behnan J, Finocchiaro G, Hanna G. The landscape of the mesenchymal signature in brain tumours. *Brain : a journal of neurology* (2019) 142(4):847-66. Epub 2019/04/05. doi: 10.1093/brain/awz044. PubMed PMID: 30946477; PubMed Central PMCID: PMCPMC6485274.
6. Sidaway P. Glioblastoma subtypes revisited. *Nature Reviews Clinical Oncology* (2017) 14(10):587-. doi: 10.1038/nrclinonc.2017.122.
7. Gimple RC, Bhargava S, Dixit D, Rich JN. Glioblastoma stem cells: lessons from the tumor hierarchy in a lethal cancer. *Genes & development* (2019) 33(11-12):591-609. Epub 2019/06/05. doi: 10.1101/gad.324301.119. PubMed PMID: 31160393; PubMed Central PMCID: PMCPMC6546059.
8. Prager BC, Bhargava S, Mahadev V, Hubert CG, Rich JN. Glioblastoma Stem Cells: Driving Resilience through Chaos. *Trends in cancer* (2020) 6(3):223-35. Epub 2020/02/27. doi: 10.1016/j.trecan.2020.01.009. PubMed PMID: 32101725; PubMed Central PMCID: PMCPMC8779821.
9. Schjoldager KT, Narimatsu Y, Joshi HJ, Clausen H. Global view of human protein glycosylation pathways and functions. *Nature Reviews Molecular Cell Biology* (2020) 21(12):729-49. doi: 10.1038/s41580-020-00294-x.
10. Reily C, Stewart TJ, Renfrow MB, Novak J. Glycosylation in health and disease. *Nature Reviews Nephrology* (2019) 15(6):346-66. doi: 10.1038/s41581-019-0129-4.
11. Varki A. Biological roles of glycans. *Glycobiology* (2017) 27(1):3-49. Epub 2016/08/26. doi: 10.1093/glycob/cww086. PubMed PMID: 27558841; PubMed Central PMCID: PMCPMC5884436.
12. Spiro RG. Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology* (2002) 12(4):43r-56r. Epub 2002/06/04. doi: 10.1093/glycob/12.4.43r. PubMed PMID: 12042244.
13. Breitling J, Aebi M. N-linked protein glycosylation in the endoplasmic reticulum. *Cold Spring Harbor perspectives in biology* (2013) 5(8):a013359. Epub 2013/06/12. doi: 10.1101/cshperspect.a013359. PubMed PMID: 23751184; PubMed Central PMCID: PMCPMC3721281.
14. Stowell SR, Ju T, Cummings RD. Protein glycosylation in cancer. *Annual review of pathology* (2015) 10:473-510. Epub 2015/01/27. doi: 10.1146/annurev-pathol-012414-040438. PubMed PMID: 25621663; PubMed Central PMCID: PMCPMC4396820.
15. Pinho SS, Reis CA. Glycosylation in cancer: mechanisms and clinical implications. *Nature Reviews Cancer* (2015) 15(9):540-55. doi: 10.1038/nrc3982.
16. Thomas D, Rathinavel AK, Radhakrishnan P. Altered glycosylation in cancer: A promising target for biomarkers and therapeutics. *Biochimica et biophysica acta Reviews on cancer* (2021) 1875(1):188464. Epub 2020/11/07. doi: 10.1016/j.bbcan.2020.188464. PubMed PMID: 33157161; PubMed Central PMCID: PMCPMC7855613.
17. Veillon L, Fakih C, Abou-El-Hassan H, Kobeissy F, Mechref Y. Glycosylation Changes in Brain Cancer. *ACS chemical neuroscience* (2018) 9(1):51-72. Epub 2017/10/06. doi: 10.1021/acschemneuro.7b00271. PubMed PMID: 28982002; PubMed Central PMCID: PMCPMC5771830.

18. Rosa-Fernandes L, Oba-Shinjo SM, Macedo-da-Silva J, Marie SKN, Palmisano G. Aberrant Protein Glycosylation in Brain Cancers, with Emphasis on Glioblastoma. *Advances in experimental medicine and biology* (2022) 1382:39-70. Epub 2022/08/28. doi: 10.1007/978-3-031-05460-0_4. PubMed PMID: 36029403.
19. Dusoswa SA, Verhoeff J, Abels E, Méndez-Huergo SP, Croci DO, Kuijper LH, et al. Glioblastomas exploit truncated O-linked glycans for local and distant immune modulation via the macrophage galactose-type lectin. *Proceedings of the National Academy of Sciences of the United States of America* (2020) 117(7):3693-703. Epub 2020/02/06. doi: 10.1073/pnas.1907921117. PubMed PMID: 32019882; PubMed Central PMCID: PMCPMC7035608.
20. Ilina EI, Cialini C, Gerloff DL, Duarte Garcia-Escudero M, Jeanty C, Thézénas ML, et al. Enzymatic activity of glycosyltransferase GLT8D1 promotes human glioblastoma cell migration. *iScience* (2022) 25(2):103842. Epub 2022/02/25. doi: 10.1016/j.isci.2022.103842. PubMed PMID: 35198895; PubMed Central PMCID: PMCPMC8850796.
21. Liu K, Jiang L, Shi Y, Liu B, He Y, Shen Q, et al. Hypoxia-induced GLT8D1 promotes glioma stem cell maintenance by inhibiting CD133 degradation through N-linked glycosylation. *Cell death and differentiation* (2022) 29(9):1834-49. Epub 2022/03/19. doi: 10.1038/s41418-022-00969-2. PubMed PMID: 35301431; PubMed Central PMCID: PMCPMC9433395.
22. Marhuenda E, Fabre C, Zhang C, Martin-Fernandez M, Iskratsch T, Saleh A, et al. Glioma stem cells invasive phenotype at optimal stiffness is driven by MGAT5 dependent mechanosensing. *Journal of experimental & clinical cancer research : CR* (2021) 40(1):139. Epub 2021/04/26. doi: 10.1186/s13046-021-01925-7. PubMed PMID: 33894774; PubMed Central PMCID: PMCPMC8067292.
23. Darmanis S, Sloan SA, Croote D, Mignardi M, Chernikova S, Samghababi P, et al. Single-Cell RNA-Seq Analysis of Infiltrating Neoplastic Cells at the Migrating Front of Human Glioblastoma. *Cell reports* (2017) 21(5):1399-410. Epub 2017/11/02. doi: 10.1016/j.celrep.2017.10.030. PubMed PMID: 29091775; PubMed Central PMCID: PMCPMC5810554.
24. Chen LY, Wei KC, Huang AC, Wang K, Huang CY, Yi D, et al. RNASEQR–a streamlined and accurate RNA-seq sequence analysis program. *Nucleic acids research* (2012) 40(6):e42. Epub 2011/12/27. doi: 10.1093/nar/gkr1248. PubMed PMID: 22199257; PubMed Central PMCID: PMCPMC3315322.
25. Bao ZS, Chen HM, Yang MY, Zhang CB, Yu K, Ye WL, et al. RNA-seq of 272 gliomas revealed a novel, recurrent PTPRZ1-MET fusion transcript in secondary glioblastomas. *Genome research* (2014) 24(11):1765-73. Epub 2014/08/20. doi: 10.1101/gr.165126.113. PubMed PMID: 25135958; PubMed Central PMCID: PMCPMC4216918.
26. Gill BJ, Pisapia DJ, Malone HR, Goldstein H, Lei L, Sonabend A, et al. MRI-localized biopsies reveal subtype-specific differences in molecular and cellular composition at the margins of glioblastoma. *Proceedings of the National Academy of Sciences of the United States of America* (2014) 111(34):12550-5. Epub 2014/08/13. doi: 10.1073/pnas.1405839111. PubMed PMID: 25114226; PubMed Central PMCID: PMCPMC4151734.

27. Stathias V, Pastori C, Griffin TZ, Komotar R, Clarke J, Zhang M, et al. Identifying glioblastoma gene networks based on hypergeometric test analysis. *PloS one* (2014) 9(12):e115842. Epub 2015/01/01. doi: 10.1371/journal.pone.0115842. PubMed PMID: 25551752; PubMed Central PMCID: PMCPMC4281219.
28. Gabrusiewicz K, Rodriguez B, Wei J, Hashimoto Y, Healy LM, Maiti SN, et al. Glioblastoma-infiltrated innate immune cells resemble M0 macrophage phenotype. *JCI insight* (2016) 1(2). Epub 2016/03/15. doi: 10.1172/jci.insight.85841. PubMed PMID: 26973881; PubMed Central PMCID: PMCPMC4784261.
29. Weerapana E, Imperiali B. Asparagine-linked protein glycosylation: from eukaryotic to prokaryotic systems. *Glycobiology* (2006) 16(6):91R-101R. doi: 10.1093/glycob/cwj099.
30. DeBerardinis RJ, Chandel NS. Fundamentals of cancer metabolism. *Science Advances* (2016) 2(5):e1600200. doi: doi:10.1126/sciadv.1600200.
31. Kaur J, Bhattacharyya S. Cancer Stem Cells: Metabolic Characterization for Targeted Cancer Therapy. *Frontiers in Oncology* (2021) 11. doi: 10.3389/fonc.2021.756888.
32. Sighel D, Notarangelo M, Aibara S, Re A, Ricci G, Guida M, et al. Inhibition of mitochondrial translation suppresses glioblastoma stem cell growth. *Cell reports* (2021) 35(4):109024. Epub 2021/04/29. doi: 10.1016/j.celrep.2021.109024. PubMed PMID: 33910005; PubMed Central PMCID: PMCPMC8097689.
33. Hay N. The Akt-mTOR tango and its relevance to cancer. *Cancer cell* (2005) 8(3):179-83. Epub 2005/09/20. doi: 10.1016/j.ccr.2005.08.008. PubMed PMID: 16169463.
34. Francipane MG, Lagasse E. Therapeutic potential of mTOR inhibitors for targeting cancer stem cells. *British journal of clinical pharmacology* (2016) 82(5):1180-8. Epub 2015/11/27. doi: 10.1111/bcp.12844. PubMed PMID: 26609914; PubMed Central PMCID: PMCPMC5061787.
35. Mecca C, Giambanco I, Donato R, Arcuri C. Targeting mTOR in Glioblastoma: Rationale and Preclinical/Clinical Evidence. *Disease markers* (2018) 2018:9230479. Epub 2019/01/22. doi: 10.1155/2018/9230479. PubMed PMID: 30662577; PubMed Central PMCID: PMCPMC6312595.
36. Sandoval JA, Tomilov A, Datta S, Allen S, O'Donnell R, Sears T, et al. Novel mTORC1 Inhibitors Kill Glioblastoma Stem Cells. *Pharmaceuticals (Basel, Switzerland)* (2020) 13(12). Epub 2020/12/02. doi: 10.3390/ph13120419. PubMed PMID: 33255358; PubMed Central PMCID: PMCPMC7761300.
37. Liu GY, Sabatini DM. mTOR at the nexus of nutrition, growth, ageing and disease. *Nature Reviews Molecular Cell Biology* (2020) 21(4):183-203. doi: 10.1038/s41580-019-0199-y.
38. Wang J, Wang H, Li Z, Wu Q, Lathia JD, McLendon RE, et al. c-Myc is required for maintenance of glioma cancer stem cells. *PloS one* (2008) 3(11):e3769. Epub 2008/11/21. doi: 10.1371/journal.pone.0003769. PubMed PMID: 19020659; PubMed Central PMCID: PMCPMC2582454.
39. Zheng H, Ying H, Yan H, Kimmelman AC, Hiller DJ, Chen AJ, et al. Pten and p53 converge on c-Myc to control differentiation, self-renewal, and transformation of normal and neoplastic stem cells in

glioblastoma. *Cold Spring Harbor symposia on quantitative biology* (2008) 73:427-37. Epub 2009/01/20. doi: 10.1101/sqb.2008.73.047. PubMed PMID: 19150964.

40. Lemjabbar-Alaoui H, McKinney A, Yang YW, Tran VM, Phillips JJ. Glycosylation alterations in lung and brain cancer. *Advances in cancer research* (2015) 126:305-44. Epub 2015/03/03. doi: 10.1016/bs.acr.2014.11.007. PubMed PMID: 25727152; PubMed Central PMCID: PMCPMC6080207.
41. Fukasawa K, Kadota T, Horie T, Tokumura K, Terada R, Kitaguchi Y, et al. CDK8 maintains stemness and tumorigenicity of glioma stem cells by regulating the c-MYC pathway. *Oncogene* (2021) 40(15):2803-15. Epub 2021/03/18. doi: 10.1038/s41388-021-01745-1. PubMed PMID: 33727660.
42. Hiraiwa M, Fukasawa K, Iezaki T, Sabit H, Horie T, Tokumura K, et al. SMURF2 phosphorylation at Thr249 modifies glioma stemness and tumorigenicity by regulating TGF-β receptor stability. *Commun Biol* (2022) 5(1):22. Epub 2022/01/13. doi: 10.1038/s42003-021-02950-0. PubMed PMID: 35017630; PubMed Central PMCID: PMCPMC8752672.
43. Huang T, Yang Y, Song X, Wan X, Wu B, Sastry N, et al. PRMT6 methylation of RCC1 regulates mitosis, tumorigenicity, and radiation response of glioblastoma stem cells. *Molecular cell* (2021) 81(6):1276-91.e9. Epub 2021/02/05. doi: 10.1016/j.molcel.2021.01.015. PubMed PMID: 33539787; PubMed Central PMCID: PMCPMC7979509.
44. Zhang A, Tao W, Zhai K, Fang X, Huang Z, Yu JS, et al. Protein sumoylation with SUMO1 promoted by Pin1 in glioma stem cells augments glioblastoma malignancy. *Neuro-oncology* (2020) 22(12):1809-21. Epub 2020/06/28. doi: 10.1093/neuonc/noaa150. PubMed PMID: 32592588; PubMed Central PMCID: PMCPMC7746948.
45. Bieberich E. Synthesis, Processing, and Function of N-glycans in N-glycoproteins. *Advances in neurobiology* (2014) 9:47-70. Epub 2014/08/26. doi: 10.1007/978-1-4939-1154-7_3. PubMed PMID: 25151374; PubMed Central PMCID: PMCPMC4236024.
46. Bieberich E. Synthesis, Processing, and Function of N-Glycans in N-Glycoproteins. In: Schengrund C-L, Yu RK, editors. *Glycobiology of the Nervous System*. Cham: Springer International Publishing (2023). p. 65-93.
47. Esmail S, Manolson MF. Advances in understanding N-glycosylation structure, function, and regulation in health and disease. *European journal of cell biology* (2021) 100(7-8):151186. Epub 2021/11/29. doi: 10.1016/j.ejcb.2021.151186. PubMed PMID: 34839178.
48. Xu X-X, Li S-T, Wang N, Kitajima T, Yoko-o T, Fujita M, et al. Structural and functional analysis of Alg1 beta-1,4 mannosyltransferase reveals the physiological importance of its membrane topology. *Glycobiology* (2018) 28(10):741-53. doi: 10.1093/glycob/cwy060.
49. Li ST, Wang N, Xu XX, Fujita M, Nakanishi H, Kitajima T, et al. Alternative routes for synthesis of N-linked glycans by Alg2 mannosyltransferase. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* (2018) 32(5):2492-506. Epub 2017/12/24. doi: 10.1096/fj.201701267R. PubMed PMID: 29273674.
50. Ng BG, Shiryaev SA, Rymen D, Eklund EA, Raymond K, Kircher M, et al. ALG1-CDG: Clinical and Molecular Characterization of 39 Unreported Patients. *Human mutation* (2016) 37(7):653-60. Epub

2016/03/05. doi: 10.1002/humu.22983. PubMed PMID: 26931382; PubMed Central PMCID: PMCPMC4907823.

Figures

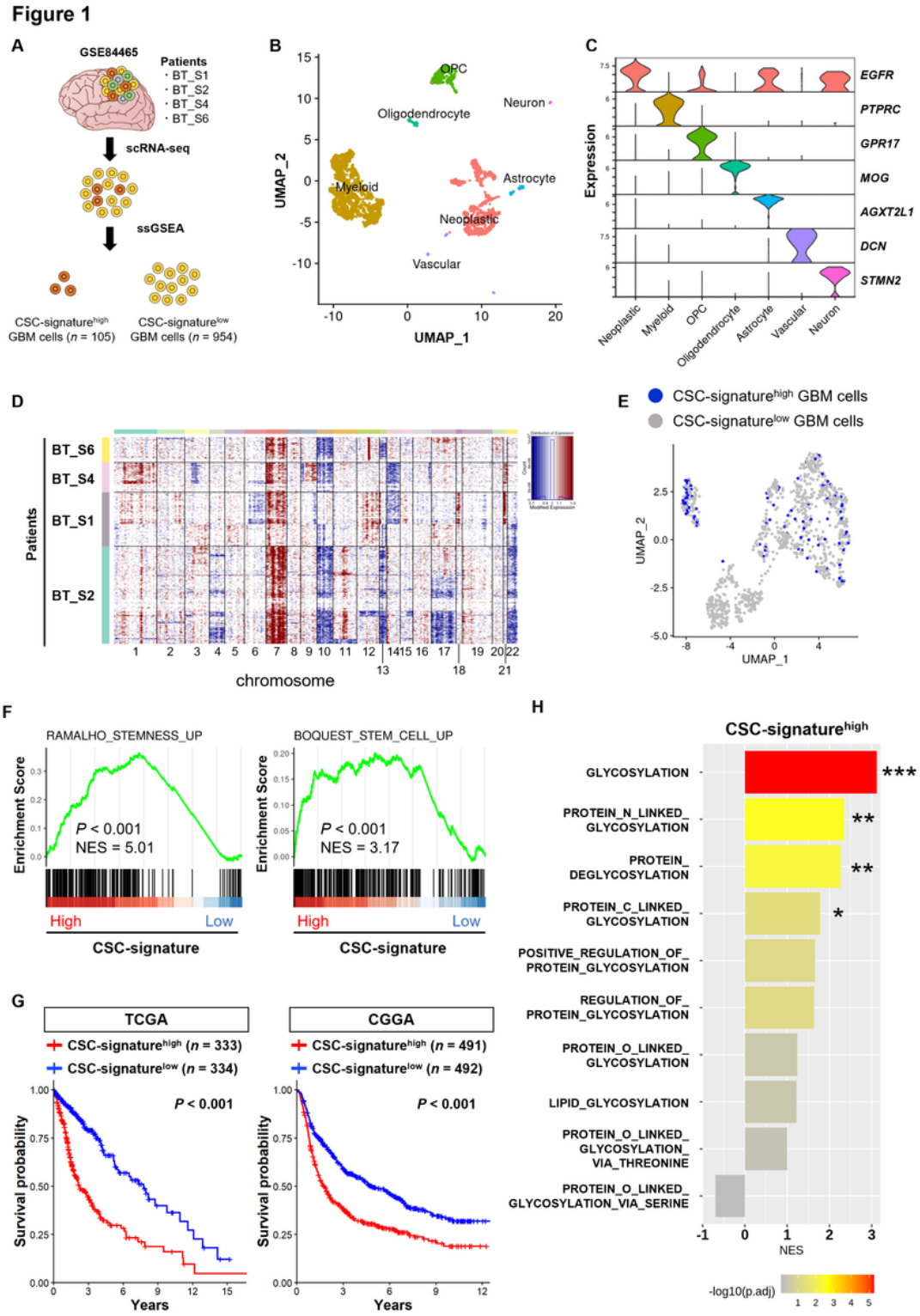

Figure 1

***N*- and *C*-linked glycosylation pathways are enhanced in GSCs.** (A) Schematic diagram of scRNA-seq and ssGSEA of GBM cells in GSE84465. (B) Uniform manifold approximation and projection (UMAP) plot of the seven identified clusters in GBM tissue. (C) Violin plot of canonical marker gene in each cell type. (D) Copy number variation (CNV) profile of cells from four patients. Red indicates amplifications, and blue indicates deletions. (E) UMAP plot showing the CSC-signature^{high} GBM (n = 105) and CSC-signature^{low} GBM (n = 954) cells. (F) Gene set enrichment analysis (GSEA) plot of stemness- and stem cell-related gene sets in CSC-signature^{high} GBM cells. (G) Kaplan–Meier survival curve for the CSC-signature^{high} patient group (n = 333) and CSC-signature^{low} patient group (n = 334) in the TCGA database (left) and the CSC-signature^{high} patient group (n = 491) and CSC-signature^{low} patient group (n = 492) in the CGGA database (right). (H) GSEA of glycosylation-related gene sets in Gene Ontology Biological Process (GOBP) ($^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$).

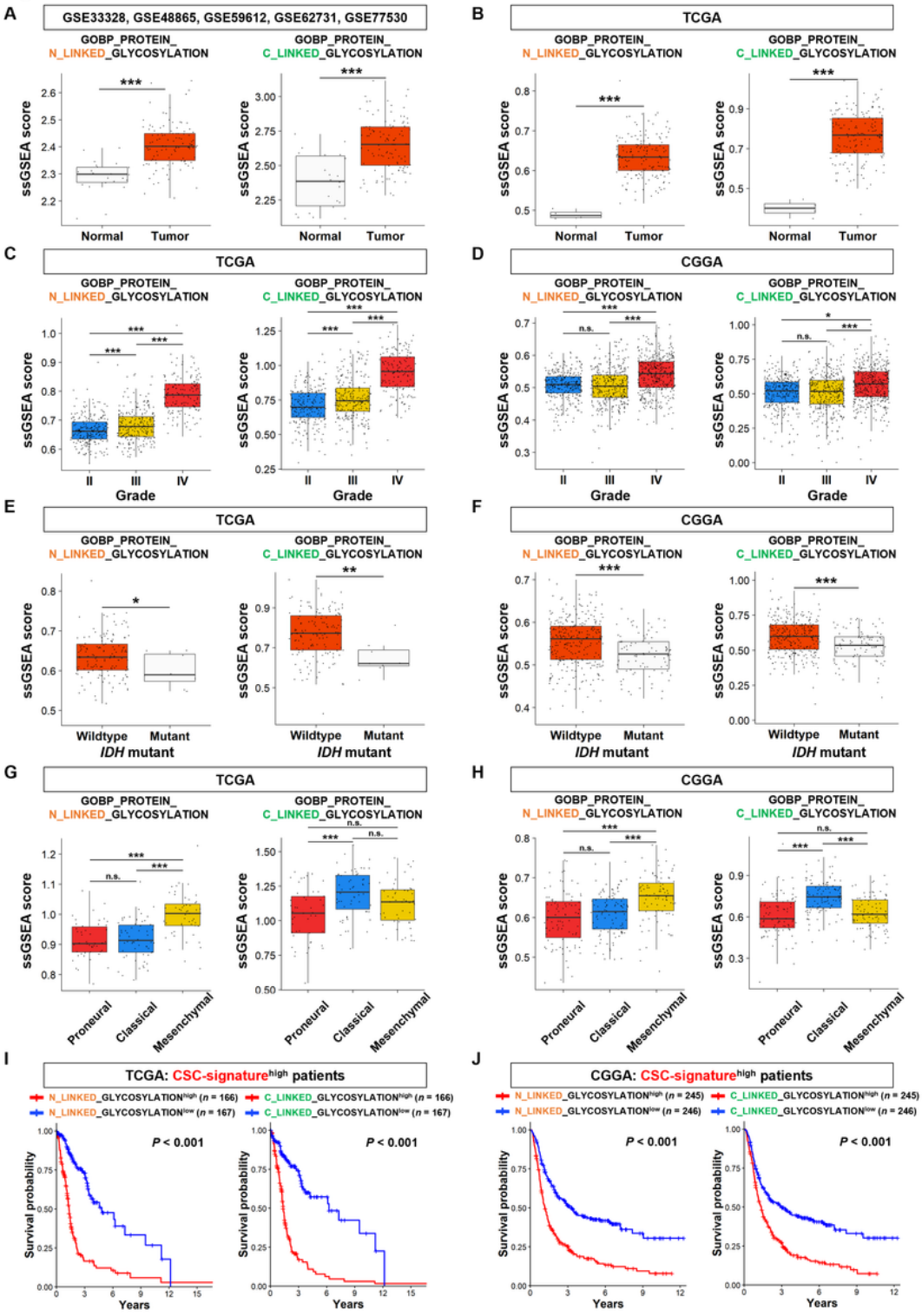

Figure 2

***N*-linked glycosylation pathways are associated with poor prognosis of patients with glioma.** (A–H) Comparison of ssGSEA scores of *N*- and *C*-linked glycosylation pathways. (A) Normal tissues (n = 22) and GBM tissues (n = 107) in GSE33328, GSE48865, GSE59612, GSE62731, and GSE77530 ($^{***}P <$ 0.001). (B) Normal tissues (n = 4) and GBM tissues (n = 156) in TCGA ($^{***}P < 0.001$). (C) Tissues exhibiting each grade of glioma in TCGA (Grade II, n = 226; Grade III, n = 244; Grade IV, n = 150) ($^{***}P <$

0.001). (D) Tissues exhibiting each grade of glioma in CGGA (Grade II, *n* = 291; Grade III, *n* = 334; Grade IV, *n* = 388) ($^{*}P < 0.05$, $^{***}P < 0.001$, n.s.: not significant). (E) Wildtype (*n* = 142) and *IDH* mutant (*n* = 8) GBM tissues in TCGA ($^{*}P < 0.05$, $^{**}P < 0.01$). (F) Wildtype (*n* = 288) and *IDH* mutant (*n* = 90) GBM tissues in CGGA ($^{***}P < 0.001$). (G) Tissues exhibiting each subtype of glioma in TCGA (Proneural, *n* = 163; Classical, *n* = 199; Mesenchymal, *n* = 166) ($^{***}P < 0.001$, n.s.: not significant). (H) Tissues exhibiting each subtype of glioma in CGGA (Proneural, *n* = 93; Classical, *n* = 106; Mesenchymal, *n* = 89) ($^{***}P < 0.001$, n.s.: not significant). (I) Kaplan–Meier survival curve for the glycosylation pathways^{high} patient group (*n* =166) and glycosylation pathways^{low} patient group (*n* =167), in the CSC-signature^{high} patient group (*n* = 333) in the TCGA database. (J) Kaplan–Meier survival curve for the glycosylation pathways^{high} patient group (*n* = 245) and glycosylation pathways^{low} patient group (*n* = 246), in the CSC-signature^{high} patient group (*n* = 491) in the CGGA database.

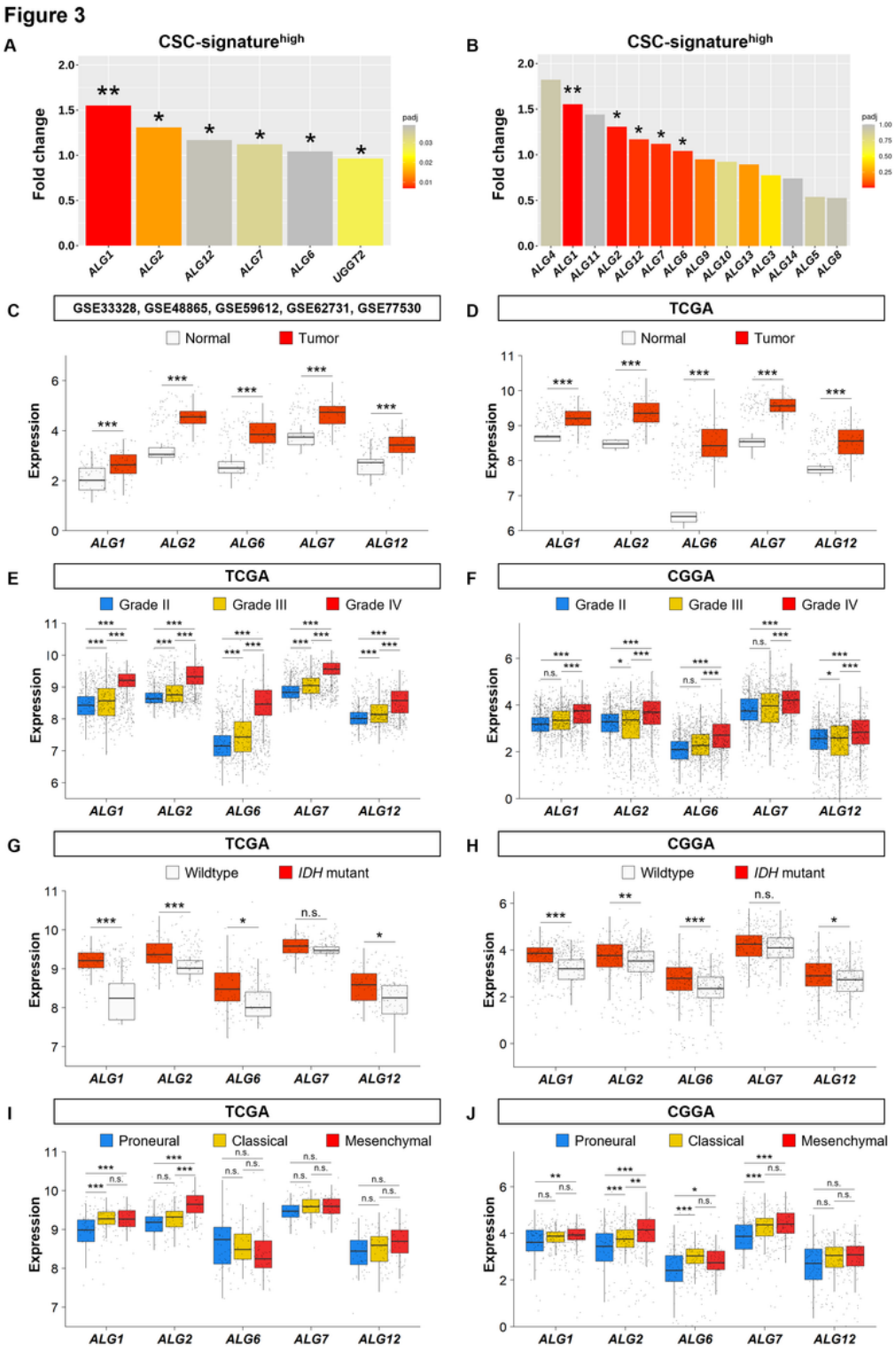

Figure 3

The ALG enzyme family is upregulated in GSCs. (A) Differentially expressed genes (DEGs) related to the *N*-linked glycosylation pathway between CSC-signature^{high} GBM cells (n = 105) and CSC-signature^{low} GBM cells (n = 954) ($^{*}P < 0.05$, $^{**}P < 0.01$). (B) DEGs in the ALG enzyme family between CSC-signature^{high} GBM cells (n = 105) and CSC-signature^{low} GBM cells (n = 954) ($^{*}P < 0.05$, $^{**}P < 0.01$). (C–J) Comparison of *ALG1*, *ALG2*, *ALG6*, *ALG7*, or *ALG12* expression. (C) Normal tissues (n = 22) and GBM tissues (n = 107)

in GSE33328, GSE48865, GSE59612, GSE62731, and GSE77530 (***$P < 0.001$). (D) Normal tissues ($n = 4$) and GBM tissues ($n = 156$) in TCGA (***$P < 0.001$). (E) Tissues exhibiting each grade of glioma in TCGA (Grade II, $n = 226$; Grade III, $n = 244$; Grade IV, $n = 150$) (***$P < 0.001$). (F) Tissues exhibiting each grade of glioma in CGGA (Grade II, $n = 291$; Grade III, $n = 334$; Grade IV, $n = 388$) (*$P < 0.05$, ***$P < 0.001$, n.s.: not significant). (G) Wildtype ($n = 142$) and *IDH* mutant ($n = 8$) GBM tissues in TCGA (*$P < 0.05$, ***$P < 0.001$, n.s.: not significant). (H) Wildtype ($n = 288$) and *IDH* mutant ($n = 90$) GBM tissues in CGGA (*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, n.s.: not significant). (I) Tissues exhibiting each subtype of glioma tissues in TCGA (Proneural, $n = 163$; Classical, $n = 199$; Mesenchymal, $n = 166$) (***$P < 0.001$, n.s.: not significant). (J) Tissues exhibiting each subtype of glioma in CGGA (Proneural, $n = 93$; Classical, $n = 106$; Mesenchymal, $n = 89$) (*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, n.s.: not significant).

Figure 4

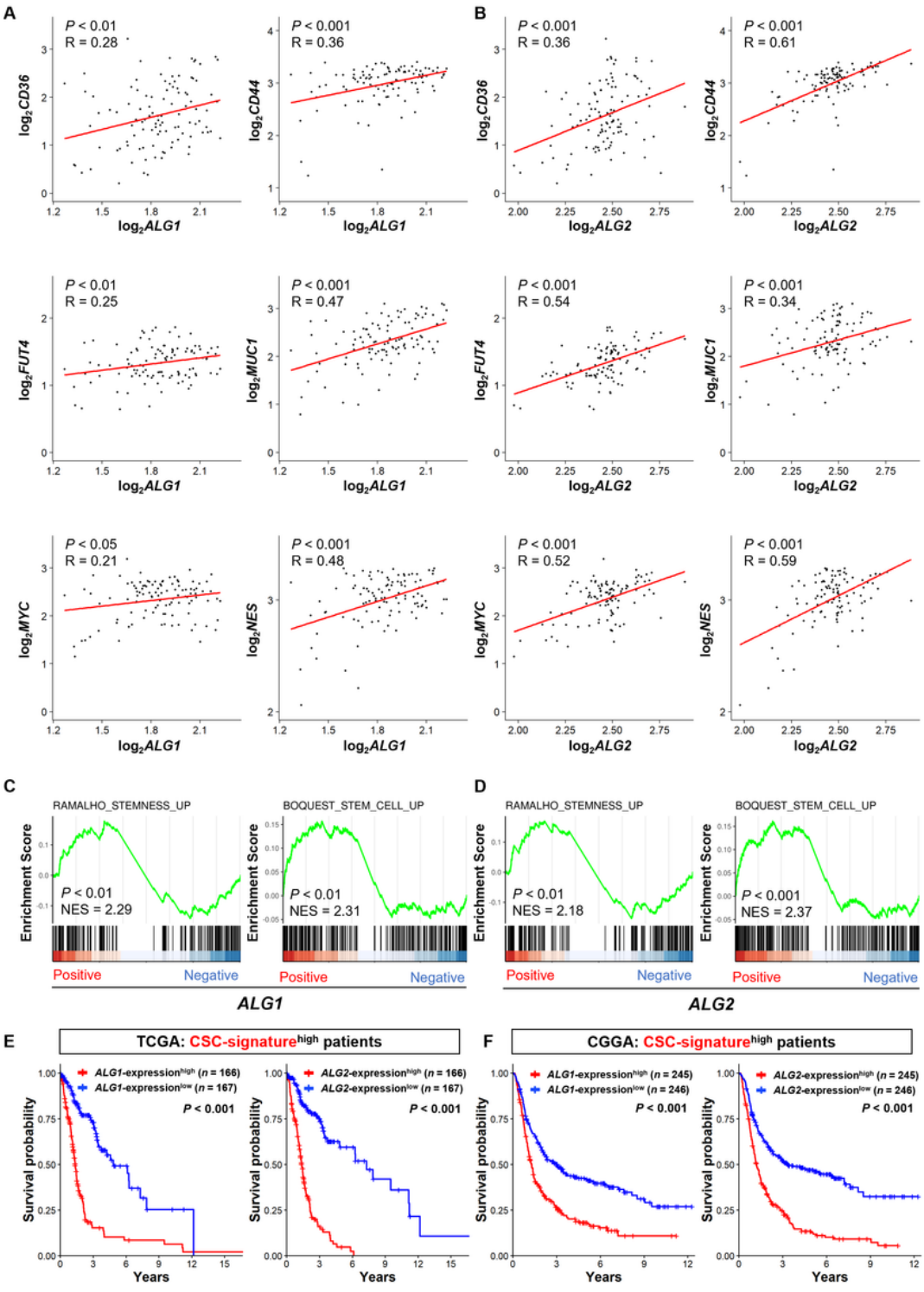

Figure 4

ALG1 and 2 are associated with the stem cell properties of GSCs and poor prognosis of GBM. (A and B) Scatter plot of correlation analysis between *ALG1* or *ALG2* and GSC marker genes in GBM tissues (*n* = 107) in GSE33328, GSE48865, GSE59612, GSE62731, and GSE77530. (C and D) GSEA plot of stemness- and stem cell-related gene sets in *ALG1*- or *ALG2*-positive CSC-signature^{high} GBM cells. (E) Kaplan–Meier survival curve of *ALG1*- or *ALG2*-expression^{high} patient group (*n* =166) and *ALG1*- or *ALG2*-expression^{low}

patient group (n =167) in the CSC-signature^{high} patient group (n =333) in the TCGA database. (F) Kaplan–Meier survival curve for *ALG1*- or *ALG2*-expression^{high} patient group (n =245) and *ALG1*- or *ALG2*-expression^{low} patient group (n =246) in the CSC-signature^{high} patient group (n =491) in the CGGA database.

Figure 5

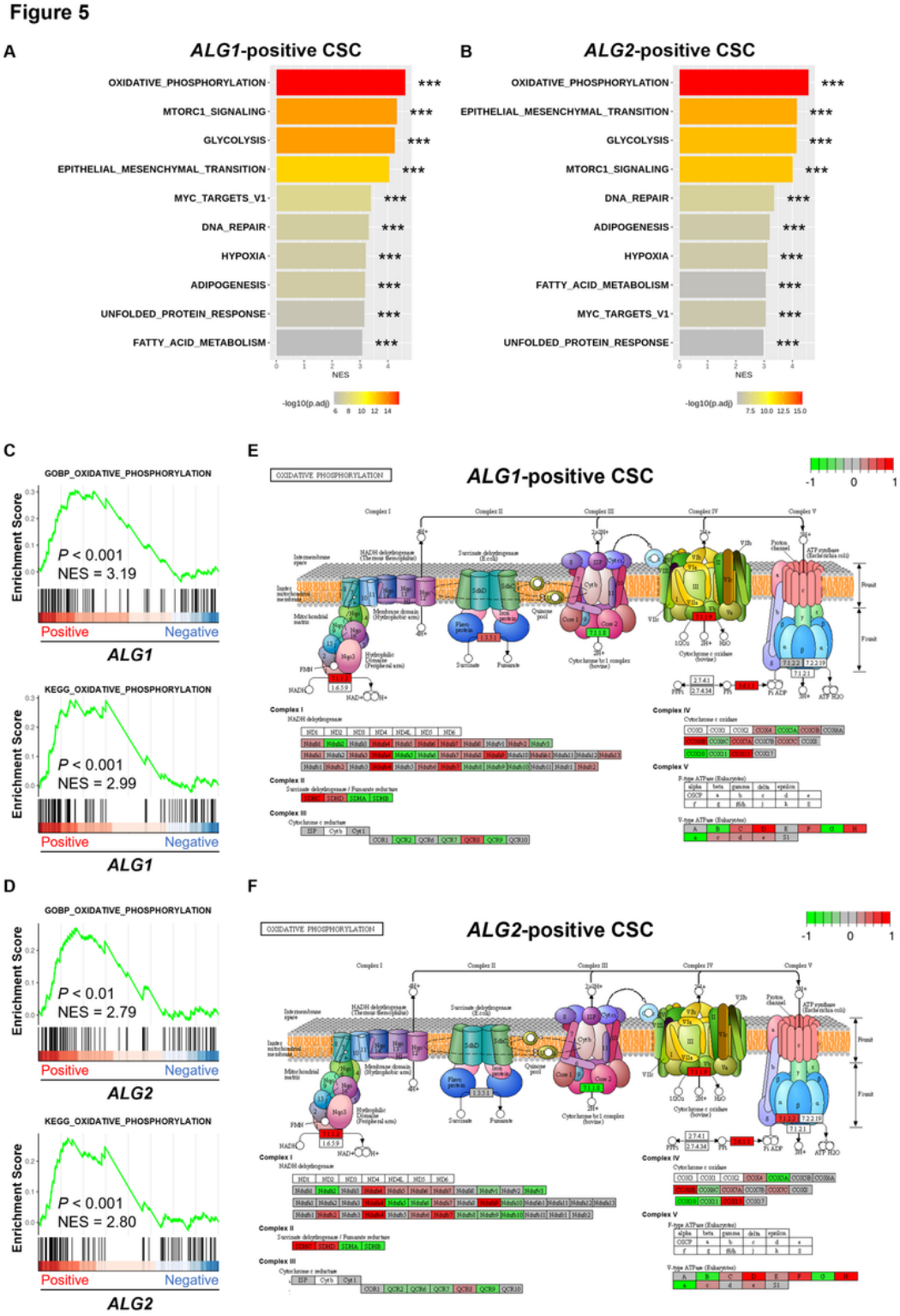

Figure 5

Oxidative phosphorylation pathway forms a link between GSC properties and ALG1 and 2. (A and B) GSEA of the Hallmark gene sets in *ALG1*- or *ALG2*-positive CSC-signature^{high} GBM cells ($^{***}P < 0.001$). (C and D) GSEA plot of GOBP- and Kyoto Encyclopedia of Genes and Genomes (KEGG)-oxidative phosphorylation pathway gene sets in *ALG1*- or *ALG2*-positive CSC-signature^{high} GBM cells. (E and F) KEGG-oxidative phosphorylation pathway for each respiratory chain complex of *ALG1*- or *ALG2*-positive CSC-signature^{high} GBM cells. Green represents downregulation, and red represents upregulation.

Supplementary Files

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

- SupplinformationGSCglycosylationfinal.pdf