Cancer Immunology, Immunotherapy (2025) 74:175
https://doi.org/10.1007/s00262-025-04034-y

RESEARCH

Identification of SURF4 and RALGAPA1 as promising therapeutic targets in glioblastoma and pan-cancer through integrative multi-omics, CRISPR-Cas9 screening and prognostic meta-analysis

Fei Wang[1] · Yuxuan Chen[2] · Run Huang[2] · Dengfeng Lu[1] · Juyi Zhang[1] · Yanbo Yang[3] · Hanhan Dang[4] · Meirong Liu[5] · Zhouqing Chen[1] · Xiaoou Sun[1] · Zhong Wang[1]

Received: 15 January 2025 / Accepted: 24 March 2025

Abstract

Glioblastoma (GBM) is the most aggressive and malignant type of primary brain tumor, with a median survival time of less than two years and a uniformly poor prognosis, despite multimodal therapeutic approaches, which highlights an urgent need for novel therapeutic targets. In this study, by integrative multi-omics analysis from CPTAC database, DepMap database and seven independent GBM cohorts, four hub genes (CD44, SURF4, IGSF3 and RALGAPA1) were identified as essential genes regulated by cancer driver genes with robust prognostic value. GBM multi-omics data from public and in-house cohorts validated that CD44 and SURF4 might be synthetic lethal partners of loss-of-function tumor suppressor genes. Analysis for immune-related pathway activity revealed complex regulation relationships of the four hub genes in tumor microenvironment (TME). Further investigation on SURF4 in pathway activity, immune therapy response and drug sensitivity proposed that SURF4 emerged as a promising therapeutic target for GBM, even for pan-cancer. Pan-cancer multi-omics exploration suggested that RALGAPA1 may be a tumor suppressor gene. By screening the first-generation and second-generation DepMap database, four genes (CCDC106, GAL3ST1, GDI2 and HSF1) might be considered as synthetic targets after mutation of RALGAPA1 as a tumor suppressor gene.

Keywords SURF4 · RALGAPA1 · Glioblastoma · Multi-omics · Pan-cancer

Fei Wang, Yuxuan Chen and Run Huang contributed equally to this work.

✉ Zhouqing Chen
zqchen6@163.com

✉ Xiaoou Sun
sunxo76@163.com

✉ Zhong Wang
wangzhong761@163.com

Fei Wang
wangfeineu@163.com

Yuxuan Chen
1554567986@qq.com

Run Huang
2713315776@qq.com

Dengfeng Lu
544821174@qq.com

Juyi Zhang
zhangjuyi123456@163.com

Yanbo Yang
yangyb6@outlook.com

Hanhan Dang
nsdhh5124@pumc.edu.cn

Meirong Liu
meizai26@163.com

[1] Department of Neurosurgery, The First Affiliated Hospital of Soochow University, Suzhou 215006, Jiangsu Province, China

[2] Suzhou Medical College of Soochow University, Suzhou 215127, Jiangsu Province, China

[3] Department of Oncology, Division of Pediatric Oncology, Johns Hopkins School of Medicine, Baltimore, MD, USA

[4] China-Japan Friendship Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100000, China

[5] Department of Neurology, The First Affiliated Hospital of Soochow University, Suzhou 215000, Jiangsu Province, China

Published online: 18 April 2025

Introduction

Glioblastoma (GBM) is one of the most lethal and recurrence-prone brain tumors [1], accounting for 49% of primary central nervous system (CNS) malignancies and 57% of all gliomas. The median survival of GBM is less than two years [2], while the first-line treatment is surgery followed by a combination of radiotherapy and the alkylating chemotherapeutic agent temozolomide (TMZ) [3]. However, the standard treatment currently in GBM fails to achieve effective recovery, including unintended relapse, worse progression and poor survival. Due to intratumoral heterogeneity and therapeutic resistance, which contribute to the poor outcomes of numerous clinical trials, GBM remains non-curative to date and new therapeutic options are urgently required [4]. Advances in molecular targeted therapy and immunotherapy have led to more possibilities of cancer treatment while the vast majority of GBM clinical research has focused on therapeutic targeting of receptor tyrosine kinase (RTK) signaling pathways, other avenues for targeted therapy include epigenetics, metabolism and immune-targeted therapies, which have received less attention [5]. The limitations of existing targets such as drug resistance, low coverage, inability to eradicate cancer stem cells, the presence of the blood–brain barrier (BBB) and challenges in clinical practice such as transient efficacy, toxicity and lack of markers together constitute the "ceiling" of GBM treatment [6]. The exploration of new targets is not only to fill the existing gaps, but also to advance the treatment from "prolonging survival for several months" to "long-term control or even cure" through multi-mechanism collaboration, technological innovation and precision stratification. Therefore, the identification of predicting response to immunotherapy and screening for potentially intervenable targets is highly warranted.

The effect of specific genomic alterations on downstream pathway signaling is nonlinear and not always predictably consistent with genotype. Rapid advancement of next-generation sequencing has led to a bright insight into cancer research. Integrating and analyzing multi-omics data help to evaluate the flow of information from one omics level to the other and bridge the gap from genotype to phenotype. GBM was one of the earliest subjects of deep genomic, transcriptomic, epigenomic and proteomic analysis. These data possess the potential to discover therapeutic and diagnostic target candidates and advance the understanding of this lethal tumor. Apart from large-scale projects such as The Cancer Genome Atlas (TCGA) [7] and the Clinical Proteomic Tumor Analysis Consortium (CPTAC) [8], few studies have integrated multi-omics data to search for new therapeutic targets.

The identification of pan-cancer or cancer-specific dependency is a central goal of cancer research. CRISPR-Cas9 screening is capable of offering similar targeted loss-of-function outcomes in hundreds of cancer cell lines, which has identified essential genes in heterogeneous cancer cell models [9]. The combined efforts of the Sanger Institute and the Broad Institute have led to the Cancer Dependency Map (DepMap). The second-generation DepMap [10] announced recently integrated previous approaches, targeting a total of 17,647 genes in cancer cell lines derived from 27 cancer types, and included more cancer models, clinically relevant genomic, transcriptional, metabolic and proteomic data, providing new opportunities for screening novel targets for GBM. DepMap serves as a powerful community resource to accurately identify how essential one gene is across hundreds of different cancer cell lines and cancer types with the final goal of providing personalized treatments for patients based on their molecular characteristics [11].

By integrating multi-omics data from 99 GBM patients in CPTAC database, DepMap database with CRISPER-Cas9 screening on 67 GBM cell lines and prognostic meta-analysis on 890 GBM patients from 7 GBM cohorts, we identified four genes (CD44, IGSF3, RALGAPA1 and SURF4) as essential genes regulated by cancer driver genes with robust prognostic value. We then evaluated the four hub genes from aspects of GBM multi-omics data in public and in-house cohorts, immune-involved pathways and pan-cancer multi-omics. Taken together, SURF4 and RALGAPA1 were emerged as under-explored but promising therapeutic targets for GBM, even for pan-cancer.

Materials and methods

Data acquisition and processing

All the public-available data were listed in Table S1. A total of 890 GBM patients with RNA-seq data and complete survival information in our discover cohort were collected from 7 independent cohorts. These cohorts includes the Cancer Genome Atlas Program (TCGA, https://portal.gdc.cancer.gov), Chinese Glioma Genome Atlas (CGGA, http://www.cgga.org.cn/index.jsp)(CGGA693 and CGGA325) [12], Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/)(GSE121720 and GSE147352) [13], The Glioma Longitudinal AnalySiS (GLASS, http://www.synapse.org/glass) [14] and Clinical Proteomic Tumor Analysis Consortium (CPTAC, https://pdc.cancer.gov/pdc/). RNA-seq data of 1817 patients with diffuse glioma in our validation cohort were obtained from TCGA, CGGA (CGGA693 and CGGA325) and GLASS. Microarray data of 1151 patients with diffuse glioma in our validation cohort from GEO (Rembrandt and Gravendeel), ArrayExpress (https://www.

ebi.ac.uk/biostudies/arrayexpress) (Kamoun), and CGGA (CGGA301) and the Affymetrix® GPL8542 platform (Gravendeel cohort) were obtained from the GEO database.

The multi-Omics and clinical data for pan-cancer analyses were obtained from LinkedOmicsKB (https://kb.linkedomics.org), TCGA database and GTEx Portal (https://www.gtexportal.org/home/). RNA-seq data and response to ICIs of immunotherapy cohorts were accessed from GSE78220 [15] and PRJEB23709 [16], both of which focused on anti-PD-1 treatment in melanoma.

The raw read count data from RNA-seq experiments were converted into Transcripts Per Kilobase Million (TPM), followed by log-2 and z-score transformation. This transformation was performed to align the gene expression measurements with those obtained from microarrays and to improve the comparability between samples. Moreover, the raw microarray data from GEO were processed using the robust multiarray averaging (RMA) algorithm implemented in the Affy package. The "ComBat" function from the "sva" package was utilized to merge different datasets into a meta cohort and adjust for batch impacts caused by non-biological technical biases in each dataset using an empirical Bayes framework [17]. Additionally, we calculated the risk score for each patient in the meta cohort using the following formula.

$$Risk\ Score = \sum_{k=1}^{n} (Coef * x_k)$$

Characterization of GBM subtypes via multi-omics integration with MOVICS

In our research, we have identified the optimal number of clusters using the "getClustNum" function from the MOVICS package [18]. This function combines the Clustering prediction index (CPI) with Gaps-statistics to determine the most suitable quantity of subgroups within the data. The "getMOIC" function from the MOVICS package was utilized to perform clustering analysis. Nine distinct clustering algorithms (CIMLR, ConsensusClustering, SNF, iClusterBayes, PINSPlus, moCluster, NEMO, COCA and LRA) were employed by specifying them within the "methodslist" parameter. The default parameters provided by the MOVICS package were used for this analysis. Subsequently, we generated clustering outcomes for each individual method. Following the computation of clustering results using the nine methods, we employed the "getConsensusMOIC" function to integrate the outcomes derived from diverse algorithms. This integration approach, based on the concept of consensus clustering, aimed to enhance the stability and reliability of the clustering results.

CRISPRi screening for essential genes in GBM cell lines

The essential genes were identified through the second-generation DepMap database (https://cellmodelpassports.sanger.ac.uk/downloads). This database uses CRISPR-Cas9 technology to systematically knock out each gene in 930 cancer cell lines [10]. In this study, we obtained transcriptomic and proteomic data from the database. We used a strategy based on the fitness score of cell lines to identify essential genes. Specifically, we first defined cell lines with a fitness score of 1, which indicates that knockout of genes in these cell lines has a substantial negative impact on cellular survival and proliferation. We subsequently identified genes with elevated expression in cell lines with a fitness score of 1 compared to those with a fitness score of 0. We postulate that a gene whose expression is significantly upregulated in cell lines with a fitness score of 1 and whose deletion significantly impairs cell viability is essential for normal cellular function and survival. Accordingly, we define these genes, which exhibit both elevated expression levels and a fitness score of 1, as essential genes. Genomic and transcriptomic data from the first-generation DepMap database (https://depmap.org/portal/download/all/) were integrated to validate the essentiality of these genes.

Meta-analysis for prognostic genes in glioma

We utilized the "meta" package [19] to conduct the meta-analysis. In order to mitigate heterogeneity among cohorts, we standardized gene expression by performing a log-2 transformation and converting it into z-scores across patients. Through univariate Cox regression analysis, we obtained the Benjamini-adjusted hazard ratio (HR) and 95% confidence interval (CI) for each gene within individual cohorts, based on which we performed the meta-analysis. The summary statistic HR and 95% CI were calculated using the random-effect model, and significance was determined by a p-value < 0.05.

Comprehensive analysis of immunogenomic molecular characterization

The IOBR package [20] was used to assess the infiltration level of immune cells in the TME. Various algorithms are used to quantify the level of immune cells, including quanTIseq, MCP-counter, CIBERSORT, xCell and EPIC. In addition, we used the Immune Cell Abundance Identifier portal (ImmuneCellAI, http://bioinfo.life.hust.edu.cn/ImmuCellAI#!/) [21] to estimate the frequencies of 18 T cell subtypes and six other immune cell types within the TME and used the Tracking Tumor Immunophenotype portal (TIP, http://biocc.hrbmu.edu.cn/TIP/) [22] to analyze the activity

of antitumor immune pathway of the Cancer-Immunity Cycle. From previously published articles, we have identified key pathways related to immune regulation [23–28]. Pre-calculated T cell dysfunction scores and the correlation between cytotoxic T lymphocyte levels and gene expression were obtained from the Tumor Immune Dysfunction and Exclusion portal (TIDE, http://tide.dfci.harvard.edu/) [29].

Exploration of drugability

A dataset of inhibitory concentration 50 (IC50) values for 746 small molecules tested across 1861 cancer cell lines, along with their corresponding mRNA gene expression profiles, was compiled. The dataset of the GSCA online website (http://bioinfo.life.hust.edu.cn/GSCA/#/drug) [30] was sourced from two significant databases: the Genomics of Drug Sensitivity in Cancer (GDSC) and the Genomics of Therapeutics Response Portal (CTRP), providing IC50 data for 265 and 481 small molecules in 860 and 1001 cell lines, respectively. By merging the drug sensitivity and mRNA expression data for each cell line, we conducted a comprehensive investigation into the correlation between gene expression and drug sensitivity.

Transcriptome- and proteome-based inference of pathway activity

At the transcriptomic level, we utilized the decoupleR package [31] to calculate the activity of 14 pathways, which uses a gene set-based approach to infer pathway activity from transcriptomic data. At the proteomic level, we obtained 50 pathway activity data from ProTrackPath (http://pancan.cptac-data-view.org/) [32] and 10 pathway activity data from The Cancer Proteome Atlas (TCPA, https://www.tcpaportal.org/tcpa/) [33]. We calculated the single-sample gene set enrichment analysis (ssGSEA) scores to quantify the activity of immune-related pathways at the protein level.

Single-cell transcriptome sequencing analysis

The expression profile of SURF4 at the single-cell level was obtained from the Tumor Immune Single-cell Hub 2 (TISCH2, http://tisch.comp-genomics.org/) [34], a scRNA-seq database dedicated to investigating the TME. TISCH2 provides comprehensive cell-type annotations at the single-cell level, allowing for the exploration of the TME across various cancer types. Specifically, the expression of SURF4 in single-cell level of glioma samples was extracted from 12 datasets, including GSE102130, GSE131928, GSE135437, GSE138794, GSE148842, GSE70630, GSE89567, GSE163108, GSE131928, GSE139448, GSE162631 and GSE84465.

Identification of synthetic lethal partner

Due to the absence of RALGAPA1-mutated glioblastoma cell lines in DepMap, we employed mRNA expression levels as a surrogate marker for RALGAPA1 mutation status, with low-expression indicative of mutation and high expression suggestive of normal function. Through the application of Wilcoxon rank-sum tests to the first-generation DepMap dataset, we identified 375 genes with significantly diminished chronos gene effect scores within the cell lines exhibiting low RALGAPA1 expression, thereby suggesting the potential for synthetic lethality. Similarly, in the second-generation DepMap dataset, Wilcoxon rank-sum test was conducted on the fitness scores and identified 71 genes with significantly more fitness scores of 1 in the low-expression group than in the high-expression group, denoting them as potential synthetic lethal partners for RALGAPA1.

Transcriptome sequencing

For RNA sequencing, we collected glioma tissues from the Gusu in-house dataset, with tissue samples processed to powder under liquid nitrogen, all under a protocol approved by the Ethics Committee of the First Affiliated Hospital of Soochow University. Total RNA was extracted using TRIzol reagent and quality assessed using a NanoPhotometer and the RNA Nano 6000 Assay System to ensure tissue integrity for sequencing. Library construction was performed using 1 μg of RNA per sample, including mRNA isolation, cDNA synthesis, end repair, dA-tailing and adaptor ligation. Indexed libraries were sequenced on an Illumina HiSeq/Novaseq/MGI2000 platform using a 2×150 PE configuration, and gene expression levels were determined using FPKM values.

Immunohistochemistry

From the Gusu in-house dataset, part of glioma samples initially utilized for RNA sequencing were selected for IHC staining. These samples were prepared by fixation in 4% paraformaldehyde (PFA) followed by paraffin embedding. Sections were obtained from the paraffin blocks and stained with hematoxylin, employing a reagent (cat. no. G1005; Servicebio). Immunohistochemistry staining of SURF4 (1:200; cat. no. 11599-1-AP; Proteintech) and GARNL1 (1:200; cat. no. 29847-1-AP; Proteintech) was executed following the manufacturer's protocols for the immunohistochemical assay kit (cat. no. PK10006; Proteintech).

Statistical analysis

All data processing, statistical analysis and visualization were conducted in R 4.3.2 software. Correlations between

two continuous variables were assessed via Spearman's correlation coefficients. The Wilcoxon rank-sum test was applied to compare continuous variables between different groups. For the univariate cox regression and multivariate cox regression, we utilized the coxph function from the survival package [35]. The ssGSEA algorithm of GSVA package [36] was used to characterize the activity of pathways within tumor cells by calculating the enrichment of collected gene lists. All statistical tests were two-sided and p value < 0.05 was regarded as statistically significant.

Results

Genes regulated by cancer driver genes using multi-data from CPTAC database

The overall design of this study was displayed in Fig. 1. Nine hub genes (CD44, FGFBP3, IGSF3, NUAK2, RALGAPA1, SLC29A4, SURF4, ZNF680 and RHOA) were identified as essential genes regulated by cancer driver

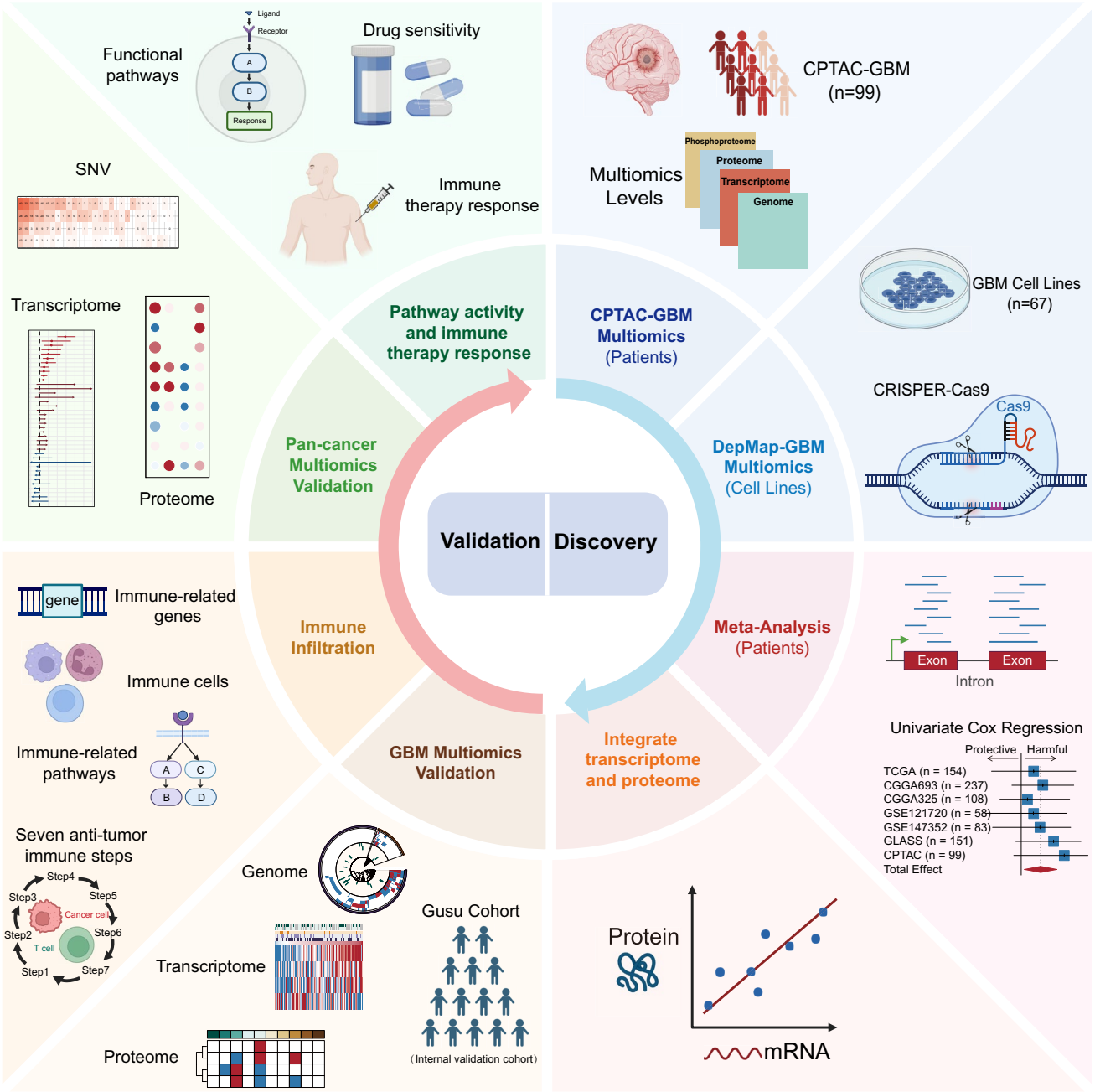


Fig. 1 Work flow of the study

genes with robust prognostic value. To achieve a more accurate unification between transcriptome and proteome, we performed correlation analysis between mRNA and protein expressions in CPTAC database. CD44, IGSF3, RALGAPA1 and SURF4 with high correlation ($p < 0.05$) were identified as potential targets for GBM treatment (Fig. S1).

A gene was defined as exhibiting a strong cis effect if it was mutated in a way that significantly influenced the downstream products it corresponded to, including RNA, protein and phosphorylation. By associating genetic alterations—mutations, copy number variations (CNVs), fusions and structural variations (SVs) with transcriptome, proteome and phosphoproteome (Fig. 2A), we obtained six genes (TP53, RB1, PDGFRA, NF1, EGFR and ATRX) of strong cis effects as cancer driver genes of GBM. Among them, EGFR and PDGFRA showed significant increases in corresponding RNA, protein and phosphorylation at EGFR-S1166

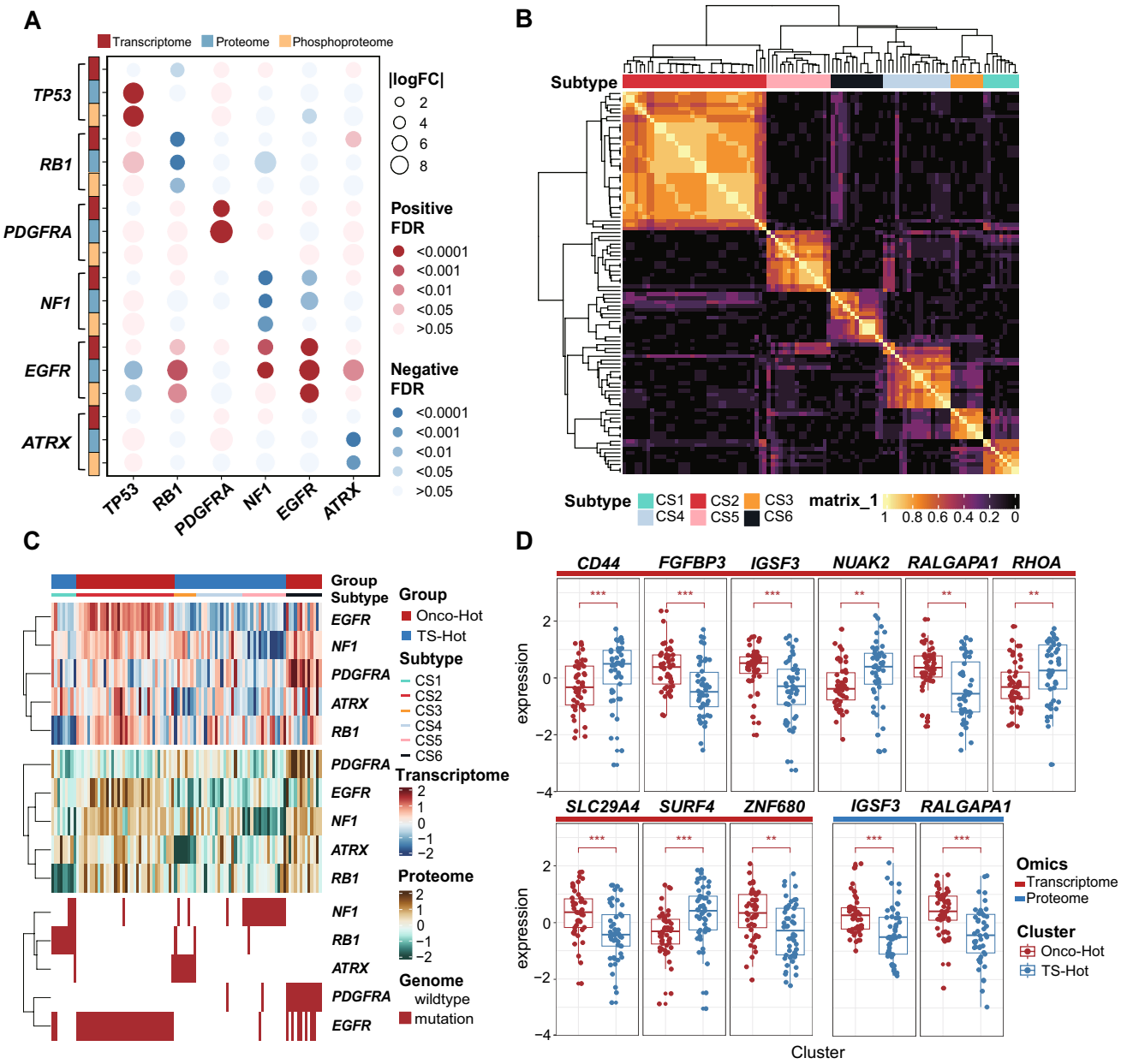


Fig. 2 Identification of potential targets for GBM by multi-omics data from CPTAC database. **A** Cis and trans effects of six cancer driver alterations on RNA expression, protein expression and phosphorylation level in CPTAC database. **B** Consensus clustering matrix for six subtypes based on the nine algorithms in CPTAC database. **C** Comprehensive heatmap of consensus ensemble subtypes by "MOVICS" based on genome, transcriptome and proteome of CPTAC database. **D** Distribution of mRNA and protein expression between proto-oncogene alteration hot group and tumor suppressor gene alteration hot group in CPTAC database

in patients with alterations. Tumor suppressors RB1, NF1 and ATRX demonstrated satisfying concordances between genetic alterations and decreased RNA, protein and phosphorylation levels of their respective gene products. We independently identified six subtypes from nine multi-omics ensemble clustering algorithms (Fig. 2B), and the clustering results were further combined through the consensus ensemble approach with distinctive molecular patterns of expression across genome, transcriptome and proteome from CPTAC database (Fig. 2C). Mutations in PDGFRA and EGFR were more frequent in cancer subtype 2 (CS2) and CS6 and were accompanied by higher levels of RNA and protein expression, while the other four CSs were the opposite. Therefore, patients were divided into proto-oncogene alteration hot (Onco-Hot) group and tumor suppressor gene alteration hot (TS-Hot) group. Subsequently, differential analysis between two groups identified 3976 differentially expressed genes (DEGs) (Table S2) and 760 differentially expressed proteins (DEPs) (Table S3). Among the nine hub genes, the expression of CD44, NUAK2, RHOA and SURF4 was significantly downregulated in Onco-Hot group, while FGFBP3, IGSF3, RALGAPA1, SLC29A4 and ZNF680 were upregulated in Onco-Hot group (Fig. 2D).

Essential genes that affected GBM cell survival in the DepMap database

The consequences of alterations in the DNA of cancer cells and subsequent vulnerabilities have not been fully understood. The second-generation DepMap, known as Sanger Project Score DepMap, aims to assign a dependency to every cancer cell in a patient which could be exploited to develop new therapies. Through CRISPR KO fitness score data, where a score < 0 indicated a statistically significant effect on cell fitness, we obtained a total of 144 fitness genes at the transcriptome level (Table S4) and 33 fitness genes at the proteome level (Table S5) by differential analysis between essential and nonessential groups (Fig. 3A). Among the nine hub genes, knockdown in GBM cell lines with high mRNA expression of CD44, FGFBP3, IGSF3, NUAK2, RALGAPA1, SLC29A4 and SURF4 and high protein expression of ZNF680 and RHOA could significantly inhibit cancer cell survival (Fig. 3B). The first-generation DepMap, known as Broad DepMap database, is a cancer dependency dataset composed of CRISPR-Cas9 loss-of-function screens performed in 1086 cancer cell lines to identify coessential gene. We further verified the four hub genes in the Broad DepMap database at the genomic level and transcriptomic level. The CNVs and mRNA expression of CD44 were negatively correlated with CRISPER effect in most cancers. For SURF4, mRNA expression had a greater impact on A variety of cancer cell survival (Fig. 3C). Unexpectedly but surprisingly, SURF4 played an essential role both in diffuse glioma (DG) and GBM cell lines separately, even more significant than CD44 (Fig. 3D).

Genes with robust prognostic value obtained from the meta cohort

Batch-size removal of seven cohorts was performed using log-transformed data (top half of Fig. 4A) and z-transformed data (bottom half of Fig. 4A) methods, respectively, where the latter was found to be superior, so we removed batch effects based on z-transformed transcriptomic data and clinical information before performing the prognostic meta-analysis. A total of 890 GBM patients in seven independent cohorts were included (Table S1), and 4073 genes were identified as genes with robust prognostic value (Table S6). Among the nine hub genes, FGFBP3 and RALGAPA1 were identified as favorable prognostic factors while CD44, IGSF3, NUCK2, RHOA, SLC29A4, SURF4 and ZNF680 were risk factors (Fig. 4B). We subsequently validated our findings in four DG RNA-seq cohorts (Fig. 4C) and four DG microarray cohorts (Fig. 4D), and most of the results were consistent with prior findings. Detailed HRs for CD44, IGSF3, RALGAPA1 and SURF4 are displayed in Fig. 4E.

GBM multi-omics data from public and in-house cohorts validated the four hub genes as promising targets

We explored whether the four hub genes we identified could be used as potential targets to treat GBM at the multi-omics level. At the genomic level, in the TCGA cohort and CPTAC database, the CNVs of SURF4 and IGSF3 in over half patients were amplified, which may be related to their upregulation in GBM and the high frequency of mutations in RALGAPA1 and IGSF3 suggested their potential roles in glioma driving (Fig. 5A). The findings of CNVs were also supported by GLASS cohort (Fig. 5B). CD44 was widely proven to be upregulated in GBM, but the result of CNV suggested that its DNA fragment was deleted. CNV could not necessarily explain the expression of mRNA and protein levels, a limitation of single Omics. There were multiple mechanisms that revealed the phenomenon behind it, including compensatory upregulation, posttranscriptional regulation, epigenetic changes, etc. At the transcriptomic level, we calculated the risk score for individual patients, and by associating mRNA expression of the four hub genes with clinical characteristics including gender, age, IDH mutation, 1p/19q codeletion and MGMT methylation, IDH mutation and 1p/19q co-codeletion were mainly distributed among patients with lower risk score (Fig. 5C). CD44, SURF4 and IGSF3 were unfavorable prognostic factors, while RALGAPA1 was identified as a favorable prognostic factor (Fig. 5D left panel). Risk score based on the four hub genes

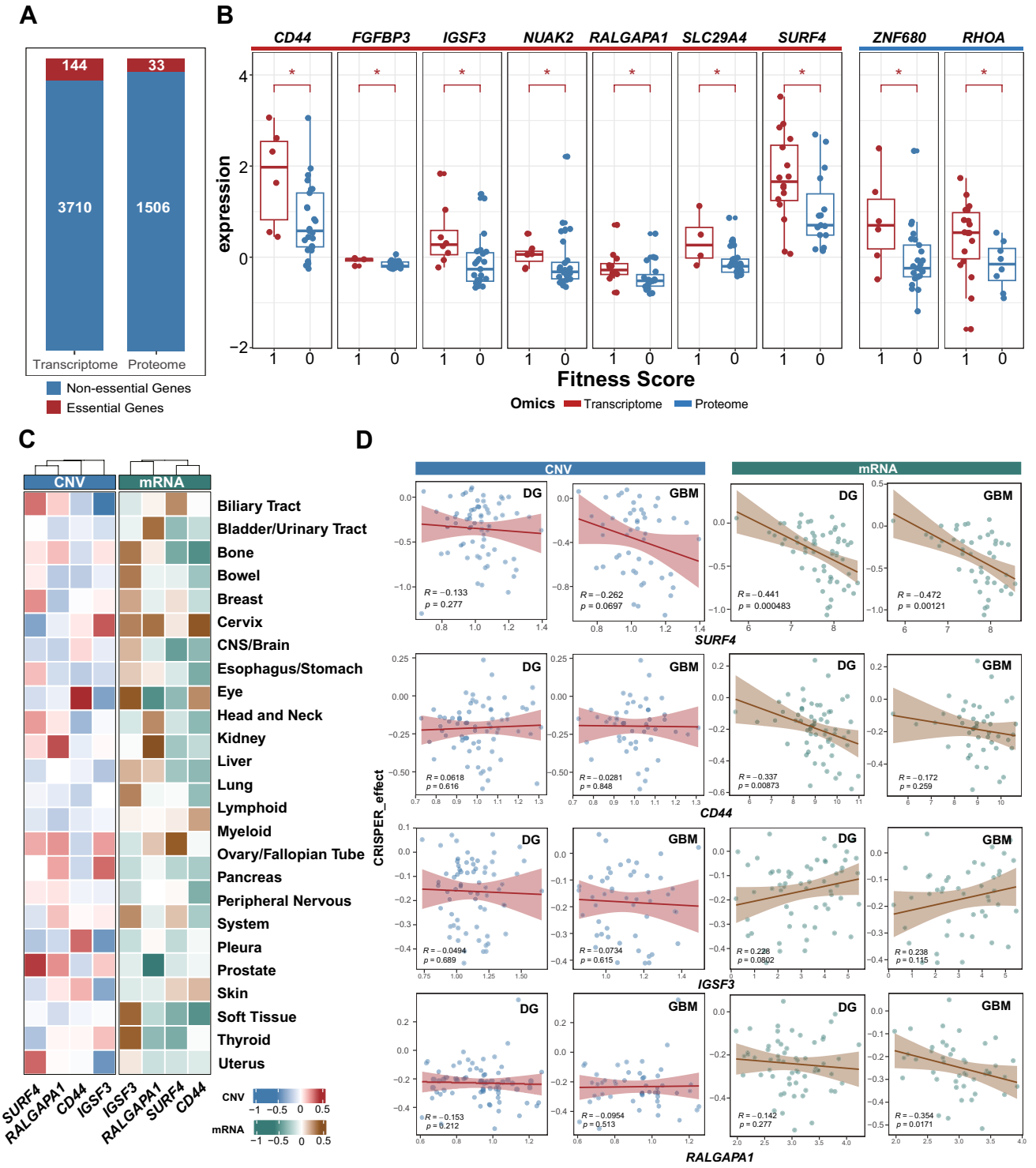


Fig. 3 Identification of potential targets for GBM in DepMap database. **A** The numbers of fitness genes identified from transcriptomic and proteomics data of 67 GBM cell lines in the Sanger Project Score DepMap database. **B** Differential analysis between the expression of nine hub genes and binary fitness scores of 67 GBM cell lines from the Sanger Project Score DepMap database. **C** Correlation analysis between CNV/mRNA expression of the four hub genes and the CRISPER effect in different cancer cell lines in the Broad DepMap database. **D** Correlation analysis between CNV/mRNA expression of the four hub genes and the CRISPER effect in DG/GBM cell lines in the Broad DepMap database

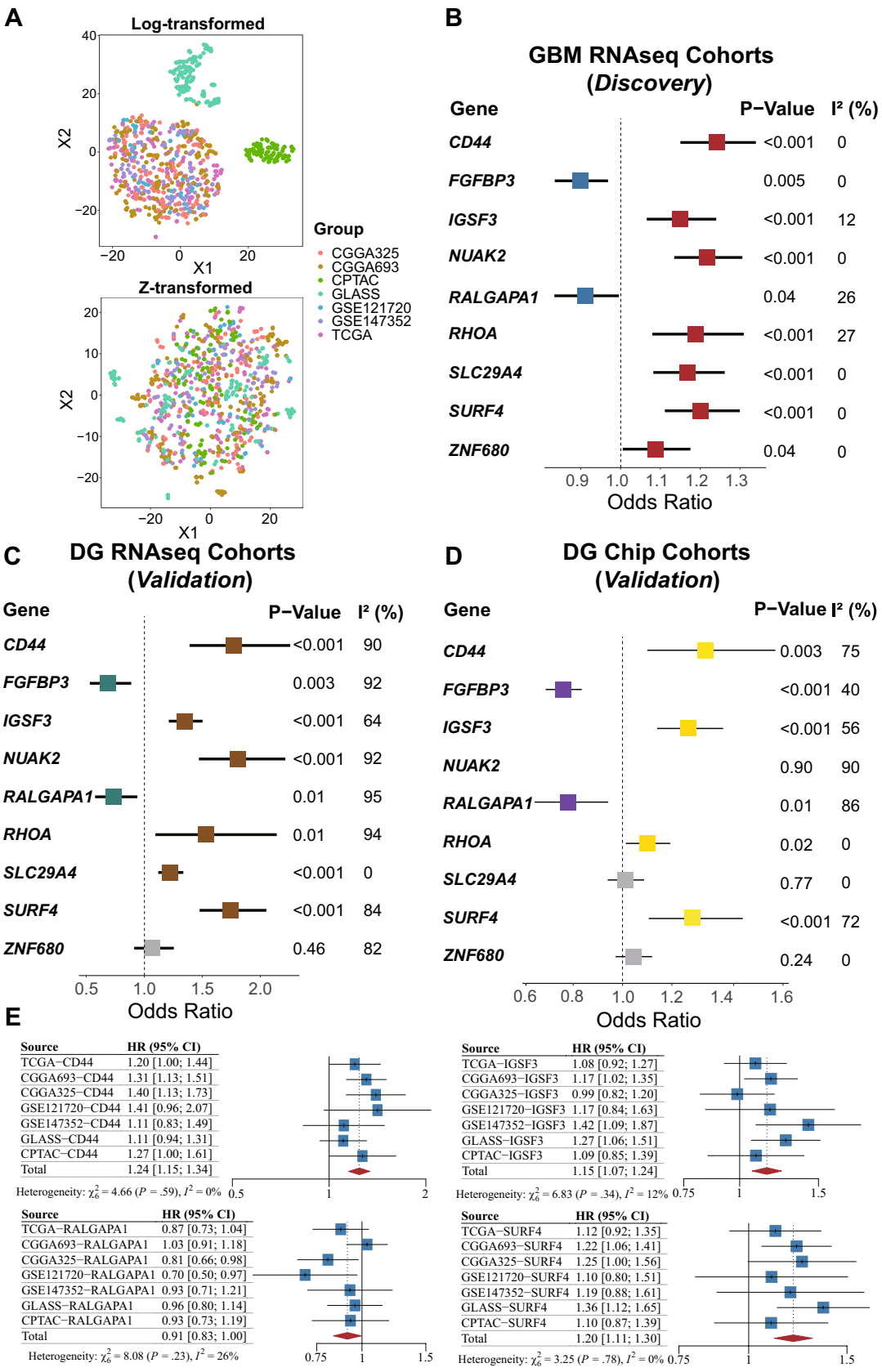
were more predictive than individual genes in multivariate Cox regression (Fig. 5D right panel). In transcriptomic and proteomic data of CPTAC database, the expression of CD44 and SURF4 was upregulated with NF1 and RB1 alteration, thus we proposed that CD44 and SURF4 might be synthetic lethal partners of loss-of-function tumor suppressor genes (Fig. 5E). In our in-house cohort (Table S7), there was a tendency for higher expression of SURF4 together with CD44 in Grade IV compared to Grade II/III, while RALGAPA1 exhibited a trend of lower expression in Grade IV than in Grade II/III (Fig. 5F). However, due to the limited sample size in our cohort, significant results could not be observed.

Expression of the four hub genes correlates with immune-involved pathways

The interaction between tumor cells and immune cells forms a complex TME, which affects not only the proliferation and migration of tumor cells, but also their response to chemotherapy and radiotherapy. There was a strong positive association trend between SURF4 and immune modulation in line with CD44, while the other two genes showed negative trend (Fig. 6A). ICAM-1 as an Ig-like cell adhesion molecule plays a role in inflammatory processes and in the T cell-mediated host defense system. Similarly, CD44 is a cell surface glycoprotein, which plays an important role in cell adhesion, migration, signal transduction and other processes, so their expressions were significantly synergistic. Similarly, the expressions of CD44 and SURF4 were also associated with high activity of immune hallmark pathway, and the other two genes were conversely (Fig. 6B), which is consistent with the trend observed in immune modulation (Fig. 6A). Among the negative correlation of RALGAPA1, the expression of RALGAPA1 was positively correlated with IFNγ pathway activity, so it is reasonable to speculate that increased expression of RALGAPA1 promotes the proliferation and differentiation of cells that can produce IFNγ. Given the association of the four hub genes identified with antitumor immunity, we next evaluated the infiltration levels of various immune cells using abundant deconvolution algorithms (Fig. 6C). Macrophage was enriched in patients with high CD44 expression, and stromal cell was enriched in patients with high SURF4 expression. IGSF3 and RALGAPA1 were positively correlated with the level of dendritic cell infiltration. By tracking tumor immunophenotype, we explored which step in antitumor immunity was more closely related to four genes (Fig. 6D). Among antitumor immune seven step activity, the relationship between the expression of SURF4 and RALGAPA1 with cancer-immunity cycle showed heterogeneity among individual cohorts. Eosinophil recruiting was thought to be positively associated with the expression of CD44 across individual cohorts and Step3 priming and activation was thought to be positively associated with the expression of IGSF3 across individual cohorts. In conclusion, our study revealed complex regulation relationships of the four hub genes in TME, especially the widely demonstrated CD44 and the under-characterized SURF4.

Assisted validation of the four hub genes from pan-cancer multi-omics

We further evaluated the four hub genes as potential targets at the pan-cancer and multi-omics level. At the genomic level, the waterfall plot presented 496 samples that at least had one single nucleotide variant (SNV) (Fig. 7A). RALGAPA1 showed SNVs in 50% of all patients, while SURF4 only showed SNVs in 9% of all patients, which was consistent with the results in GBM, and variations in all four genes were dominated by Missense mutations. A genome that was not significantly altered indicated that SURF4 probably was not involved in driving tumor and may be in the downstream of tumor development and progression. UCEC, SKCM, COAD, STAD and LUAD showed the highest deleterious mutation frequencies among 29 cancer types, while SNV frequencies of the four hub genes in GBM and LGG were general (Fig. 7B). At the transcriptomic level, CD44 and IGSF3 showed a tendency to act as prognostic risk factors in about a half cancer types and as prognostic protective factors in the other half. RALGAPA1 was a protective factor in 5 cancer types, and its lower expression led to a trend toward poorer survival in more than 50% of cancer types. SURF4 was a risky factor in 11 cancer types and its higher expression led to a trend toward poorer survival in more than two-thirds cancer types (Fig. S2A). mRNA expression of the four genes also demonstrated significant differences between tumor and normal tissues in pan-cancer, among which CD44, IGSF3 and SURF4 showed higher expression of tumor tissue in most cancer types (Fig. S2B). At the proteomic level, protein expression of all four genes consistently differed between tumor and peritumor tissues, which was nearly upregulated within tumor in CD44, IGSF3 and SURF4, and was downregulated in RALGAPA1 (Fig. 7C). Given the high mutation frequency in pan-cancer and the low level of expression in tumor center, we inferred that RALGAPA1 might be a tumor suppressor gene. Based on TIDE algorithm, large heterogeneity among cancer types was observed in the correlation with cytotoxic T lymphocyte (CTL) and contribution to T dysfunction, which illustrated the complexity of four genes involvement in immune regulation, and the deeper relationship was still worth exploring (Fig. 7D). Cancer-related pathways, including apoptosis, cell cycle and endothelial mesenchymal transition (EMT) pathways exhibited discernible correlations with the expression of SURF4, CD44 and IGSF3, all of which were associated with malignant degree of tumor (Fig. 7E). The higher


A
Log-transformed
X2
X1
Group
CGGA325
CGGA693
CPTAC
GLASS
GSE121720
GSE147352
TCGA
Z-transformed
B
GBM RNAseq Cohorts
(Discovery)
Gene
P−Value
I² (%)
CD44 <0.001 0
FGFBP3 0.005 0
IGSF3 <0.001 12
NUAK2 <0.001 0
RALGAPA1 0.04 26
RHOA <0.001 27
SLC29A4 <0.001 0
SURF4 <0.001 0
ZNF680 0.04 0
Odds Ratio
C
DG RNAseq Cohorts
(Validation)
CD44 <0.001 90
FGFBP3 0.003 92
IGSF3 <0.001 64
NUAK2 <0.001 92
RALGAPA1 0.01 95
RHOA 0.01 94
SLC29A4 <0.001 0
SURF4 <0.001 84
ZNF680 0.46 82
Odds Ratio
D
DG Chip Cohorts
(Validation)
CD44 0.003 75
FGFBP3 <0.001 40
IGSF3 <0.001 56
NUAK2 0.90 90
RALGAPA1 0.01 86
RHOA 0.02 0
SLC29A4 0.77 0
SURF4 <0.001 72
ZNF680 0.24 0
Odds Ratio
E
Source HR (95% CI)
TCGA−CD44 1.20 [1.00; 1.44]
CGGA693−CD44 1.31 [1.13; 1.51]
CGGA325−CD44 1.40 [1.13; 1.73]
GSE121720−CD44 1.41 [0.96; 2.07]
GSE147352−CD44 1.11 [0.83; 1.49]
GLASS−CD44 1.11 [0.94; 1.31]
CPTAC−CD44 1.27 [1.00; 1.61]
Total 1.24 [1.15; 1.34]
Heterogeneity: $\chi^2_6 = 4.66$ ($P = .59$), $I^2 = 0\%$
Source HR (95% CI)
TCGA−IGSF3 1.08 [0.92; 1.27]
CGGA693−IGSF3 1.17 [1.02; 1.35]
CGGA325−IGSF3 0.99 [0.82; 1.20]
GSE121720−IGSF3 1.17 [0.84; 1.63]
GSE147352−IGSF3 1.42 [1.09; 1.87]
GLASS−IGSF3 1.27 [1.06; 1.51]
CPTAC−IGSF3 1.09 [0.85; 1.39]
Total 1.15 [1.07; 1.24]
Heterogeneity: $\chi^2_6 = 6.83$ ($P = .34$), $I^2 = 12\%$
Source HR (95% CI)
TCGA−RALGAPA1 0.87 [0.73; 1.04]
CGGA693−RALGAPA1 1.03 [0.91; 1.18]
CGGA325−RALGAPA1 0.81 [0.66; 0.98]
GSE121720−RALGAPA1 0.70 [0.50; 0.97]
GSE147352−RALGAPA1 0.93 [0.71; 1.21]
GLASS−RALGAPA1 0.96 [0.80; 1.14]
CPTAC−RALGAPA1 0.93 [0.73; 1.19]
Total 0.91 [0.83; 1.00]
Heterogeneity: $\chi^2_6 = 8.08$ ($P = .23$), $I^2 = 26\%$
Source HR (95% CI)
TCGA−SURF4 1.12 [0.92; 1.35]
CGGA693−SURF4 1.22 [1.06; 1.41]
CGGA325−SURF4 1.25 [1.00; 1.56]
GSE121720−SURF4 1.10 [0.80; 1.51]
GSE147352−SURF4 1.19 [0.88; 1.61]
GLASS−SURF4 1.36 [1.12; 1.65]
CPTAC−SURF4 1.10 [0.87; 1.39]
Total 1.20 [1.11; 1.30]
Heterogeneity: $\chi^2_6 = 3.25$ ($P = .78$), $I^2 = 0\%$

◂**Fig. 4** Identification of potential targets for GBM by prognostic meta-analysis. **A** Comparison for batch effects using log-transformed data and z-transformed data from seven independent cohorts. **B** Robust prognostic value of nine hub genes using meta-analysis of seven GBM RNA-seq cohorts. **C** Robust prognostic value of nine hub genes using meta-analysis in four DG RNA-seq cohorts. **D** Robust prognostic value of nine hub genes using meta-analysis in four DG microarray cohorts. **E** The HRs of four hub genes in seven independent GBM RNA-seq cohorts and meta cohort

RALGAPA1 expressed, the more EMT pathway was inhibited. The EMT pathway is closely associated with the pathological processes of cancer metastasis and inflammatory response, which further suggested that RALGAPA1 may be a tumor suppressor gene.

SURF4 was emerged as an under-explored but promising therapeutic target for GBM, even for pan-cancer

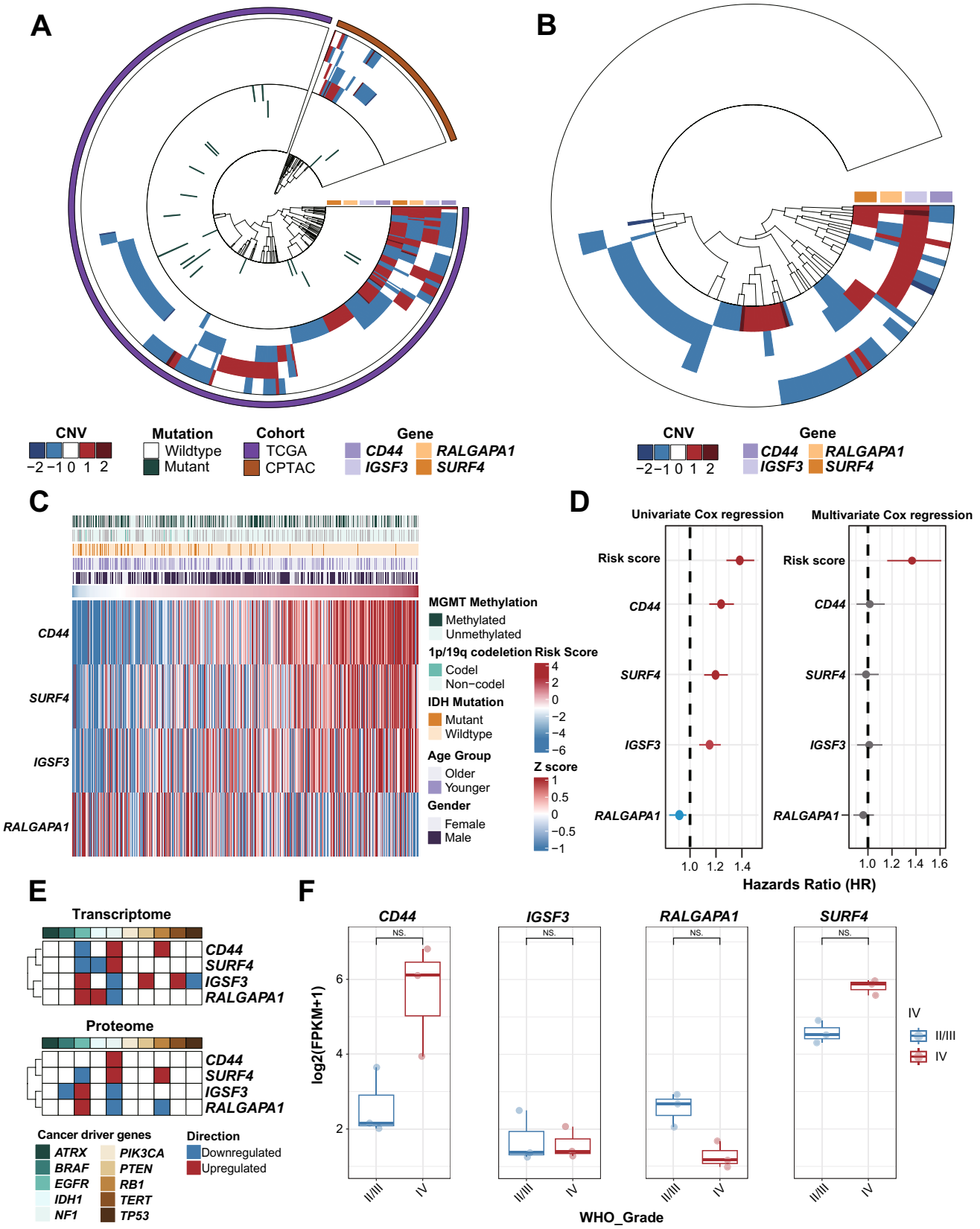
In the previous assessment of the four hub genes, we found unique advantages of SURF4. SURF4, which was consistently upregulated in glioma and pan-cancer, was identified as an essential gene regulated by cancer driver genes, predicting a poor prognosis, and showed amplified CNVs but rare mutation. Further investigation into the tumor immune microenvironment (TIME) showed that high expression of SURF4 correlated with a hot phenotype, which supported the potential of SURF4 to overcome the immunosuppressive microenvironment of GBM. Therefore, the next step was to carry out further studies on SURF4 from pathway activity, immune therapy response and drugability. At the transcriptomic level, we found that SURF4 was mainly involved in the activation of VEGF, TNFα and TGFβ, which were closely related to the proliferation, differentiation, growth and metastasis of tumor cells (Fig. 8A). At the proteomic level, we consistently observed the activation of TNFα and TGFβ pathways as the expression of SURF4 increased (Fig. 8B). We explored the potential of SURF4 as an immunotherapy response marker in several cohorts treated with ICIs. Patients with lower expression of SURF4 were more likely to respond to anti-PD-L1 immunotherapy in GSE78220 and PRJEB23709 (Fig. 8C), which contradicted our previous finding that high SURF4 expression was associated with stronger antitumor immunity. We speculated that this was because these two cohorts were in melanoma, whereas our findings were primarily found in glioma, where the same gene may have different functions in different cancers. The result of immunohistochemistry (IHC) showed that the expression of SURF4 protein in high-grade glioma was significantly increased compared with low-grade glioma (Fig. 8D). By integrating the relationship between Genomics of Drug Sensitivity in Cancer (GDSC)/The Cancer Therapeutics Response Portal (CTRP) drug sensitivity and mRNA expression, SURF4 showed a higher correlation compared with other genes, which supported the higher drugability of SURF4 (Fig. S3A and S3B). We further explored the expression of SURF4 at a single-cell resolution. The expression was consistently high in all types of cells in the TIME, which confirmed that SURF4 had a complex function in TIME (Fig. S4). In conclusion, SURF4 was associated with cancer signaling pathways like VEGF, TGFβ and TNFα, can predict the response to immunotherapy, and had strong drugable capacity. In conclusion, all results provide a rationale for SURF4 as a promising therapeutic target for GBM, even for pan-cancer.

Four genes were proposed to act as synthetic targets after mutation of RALGAPA1 as a tumor suppressor gene

We hypothesized that RALGAPA1 is a potential tumor suppressor, its mutation might lead to the downregulation of its mRNA and protein expressions, and the expression of synthetic lethal partnership will be upregulated correspondingly at this point. If we take action to intervene oncogenes, then it can contribute to killing tumor cells. Because there were no GBM cell lines with RALGAPA1 alterations in the DepMap, the mRNA expression of transcriptome was used to find the synthetic lethal partners of RALGAPA1, with low RALGAPA1 expression representing RALGAPA1 mutation, and high expression representing normal gene function. In the first-generation DepMap, CRISPER effect of all genes was compared between RALGAPA1-high group and RALGAPA1-low group, and those with significantly lower CRISPER effect in RALGAPA1-low group were considered as synthetic lethal partners of RALGAPA1. In the second-generation DepMap, RALGAPA1-high group and RALGAPA1-low group compared binary classification of fitness score of all genes, and those in RALGAPA1-low group with more significant 1 were considered as synthetic lethal partners of RALGAPA1 (Fig. 8E). Therefore, we collected 375 synthetic lethal partners in the first-generation DepMap and 71 synthetic lethal partners in the second-generation DepMap (Table S8). After taking the intersection of the two, four genes (CCDC106, GAL3ST1, GDI2 and HSF1) were obtained (Fig. 8F and G). In addition, IHC showed that the expression of RALGAPA1 protein decreased with increasing grade (Fig. 8H). They were proposed to act as synthetic targets after mutation of RALGAPA1 as a tumor suppressor gene.

Discussion

Previous glioblastoma research has predominantly relied on single-omics data, such as gene expression profiles or protein levels. While these approaches provide a snapshot


A
B
CNV
−2 −1 0 1 2
Mutation
Wildtype
Mutant
Cohort
TCGA
CPTAC
Gene
CD44
RALGAPA1
IGSF3
SURF4
C
CD44
SURF4
IGSF3
RALGAPA1
MGMT Methylation
Methylated
Unmethylated
1p/19q codeletion
Codel
Non−codel
IDH Mutation
Mutant
Wildtype
Age Group
Older
Younger
Gender
Female
Male
Risk Score
4
2
0
−2
−4
−6
Z score
1
0.5
0
−0.5
−1
D
Univariate Cox regression
Multivariate Cox regression
Risk score
CD44
SURF4
IGSF3
RALGAPA1
1.0 1.2 1.4
1.0 1.2 1.4 1.6
Hazards Ratio (HR)
E
Transcriptome
CD44
SURF4
IGSF3
RALGAPA1
Proteome
CD44
SURF4
IGSF3
RALGAPA1
Cancer driver genes
ATRX
BRAF
EGFR
IDH1
NF1
PIK3CA
PTEN
RB1
TERT
TP53
Direction
Downregulated
Upregulated
F
CD44
IGSF3
RALGAPA1
SURF4
NS.
log2(FPKM+1)
6
4
2
II/III
IV
WHO_Grade
IV
II/III
IV

◂**Fig. 5** Evaluation of the four hub genes as potential targets from perspective of GBM multi-omics. **A** SNVs and CNVs of the four hub genes in genomic data of TCGA cohort and CPTAC database. **B** CNVs of the four hub genes in genomic data of GLASS cohort. **C** The relationships between clinicopathologic characteristics including gender, age, IDH mutation, 1p/19q codeletion and MGMT methylation ranked in increasing order of risk score in the meta cohort. **D** Prognostic value of the four hub genes in univariate Cox regression analysis (left panel) and multivariable Cox regression analysis of the four hub genes based on clinicopathological features (right panel) in the meta cohort. **E** Differential expression of mRNA and protein of the four hub genes in 10 cancer driver alterations compared to unaltered patients in the CPTAC database. **F** mRNA expression of the four hub genes between Grade II/III and Grade IV in Gusu Cohort (in-house data)

of molecular changes, they often overlook the complex interplay between different molecular layers, limiting a comprehensive understanding of tumor heterogeneity. In contrast, patients in our study were divided into two groups based on six cancer driver genes with strong cis effects (TP53, RB1, PDGFRA, NF1, EGFR and ATRX) in glioblastomas using a multi-omics consensus combinatorial approach in the CPTAC database. This approach provides a holistic understanding of the molecular landscape and an enhanced understanding of molecular heterogeneity. Secondly, previous studies on the necessity of genes have relied solely on the fitness score [37], which is straightforward but may not comprehensively reflect the role of genes in different cellular states. Therefore, we integrated the fitness score of second-generation DepMap data to refine the identification of essential genes in cancer. The second-generation DepMap, which is the most comprehensive analysis of cancer dependencies to date, integrates a larger number of cancer models, including clinically relevant genomic, transcriptomic and proteomic data. This expanded dataset provides a more comprehensive perspective on the importance of genes. By correlating the fitness score with comprehensive transcriptomic and proteomic profiles, we overcame the limitations of single-omics analysis. The combination of multi-omics analysis and the fitness score not only increases the specificity of gene screening, but also reduces the sensitivity to false positives, allowing for a more precise and reliable assessment of gene significance, ultimately strengthening the validity of our findings across a wide range of cancer models. Additionally, in terms of prognostic analysis, most previous studies have been based on single or limited cohorts, which may lead to different interpretations of the prognostic significance of the same gene. We integrated expression profiles and survival information from multiple central GBM cohorts and used meta-analysis to systematically review and statistically integrate the data. This approach not only enhances the stability of the prognostic analysis, but also amplifies the statistical power of our findings.

Our study found that CD44, IGSF3, RALGAPA1 and SURF4 might be promising novel targets for gliomas. The protein encoded by the CD44 gene is a transmembrane glycoprotein that acts as a key receptor for hyaluronan (HA), a glycosaminoglycan that is present in abundance in the central nervous system. The interaction between CD44 and HA is of critical importance for the growth and progression of glioblastoma multiforme cells, as well as influencing their response to chemotherapy [38]. Previous studies have identified CD44 as a cancer stem cell marker in a range of malignancies, including pancreatic, prostate, colorectal, breast and GBM. The persistence of CD44-marked tumor stem cells represents a significant factor in tumor recurrence, even in the context of initial treatment regimens that have successfully reduced tumor size [39–41]. This serves to confirm the essentiality of CD44 as a target for the further understanding and combating of cancer resilience and recurrence. Furthermore, IGSF3 expression was observed to be specifically increased in human gliomas, and was found to correlate with a poor prognosis for the patient [42]. To date, the role of RALGAPA1 in glioblastoma or even pan-cancer has not been investigated in the literature [43]. Previous researches have shown that SURF4 may be a potential target for stem cell-based cancer therapy [44]. The expression of BIRC3 in OCSCs was decreased after SURF4 downregulation. As a member of the inhibitor of apoptosis protein (IAP) family, BIRC3 not only regulates apoptosis, but is also closely associated with immune regulation [45]. Analyses that are comprehensive, multidimensional and supported by the existing literature lend greater credibility to our findings and help us to identify promising targets that are potentially clinically relevant in the treatment of glioblastoma.

Our further analysis revealed that SURF4 and RALGAPA1 might be novel targets for GBM and even pan-cancer. SURF4, which was consistently upregulated in glioblastoma and pan-cancer, was identified as an essential gene regulated by cancer driver genes, predicting a poor prognosis. Meanwhile, SURF4 was associated with immune-related pathways like neoantigens generation and IFNγ activation as well as cancer signaling pathways like VEGF, TGFβ and TNFα, and can predict the response to immunotherapy. All results provide a rationale for SURF4 as a promising therapeutic target for GBM, even for pan-cancer. Our study also reveals that RALGAPA1 is frequently mutated in gliomas and across a spectrum of cancers, indicating its potential as a tumor suppressor gene. Synthetic lethality offers an important strategy to address targets that are traditionally difficult to drug, in particular loss-of-function mutations in tumor suppressor genes that are critical for cancer initiation. In this context, our research has identified the genes CCDC106, GAL3ST1, GDI2 and HSF1 as potential synthetic lethal

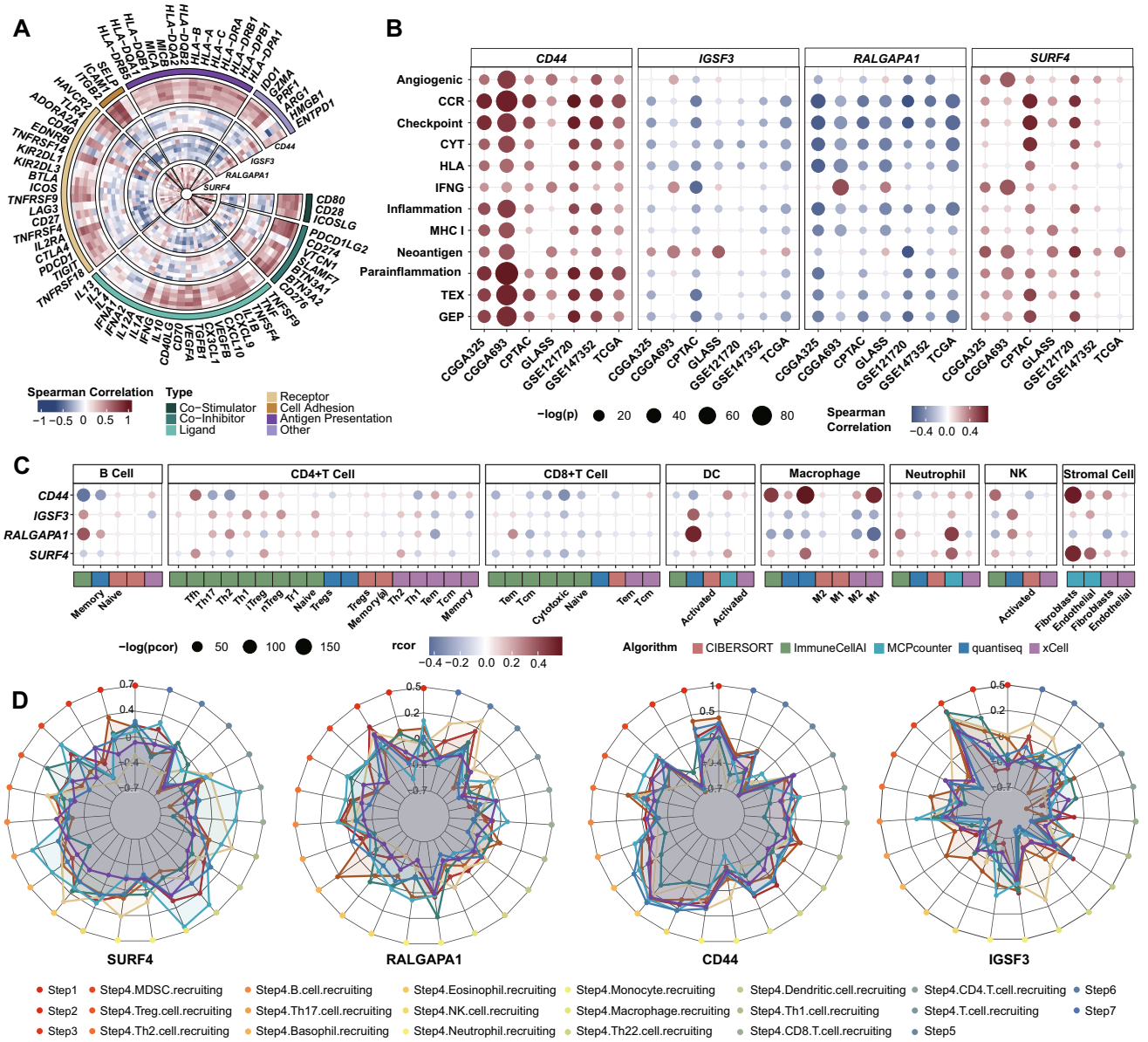


Fig. 6 Correlation analysis between mRNA expression and immune-involved pathways. **A** Correlation analysis between mRNA expression of the four hub genes and 73 key immune modulators in the meta cohort. **B** Correlation analysis between mRNA expression of the four hub genes and the ssGSEA score of immune-related pathways in seven GBM cohorts. **C** Correlation analysis between mRNA expression of the four hub genes and the enrichment of infiltrating immune cells in the meta cohort. **D** Correlation analysis between mRNA expression of the four hub genes and steps of antitumor immunity in the meta cohort

targets. These interactions, if validated, could offer new avenues for targeted cancer therapies.

Despite some of the advantages of our study, there are inherent limitations that need to be considered. Firstly, bulk RNA-seq captures aggregate gene expression levels across a heterogeneous cell population, which may dilute the signal from rare or minority cell types, making it difficult to discern their expression patterns. And bulk RNA-seq does not provide information on the precise cellular location or distribution of gene expression. This impedes comprehension of the intricate dynamics within the tumor microenvironment and the interactions between disparate cell types in the context of cancer progression. Secondly, our four-gene target identification approaches simply taking the intersection of different biological significance, which may filter out

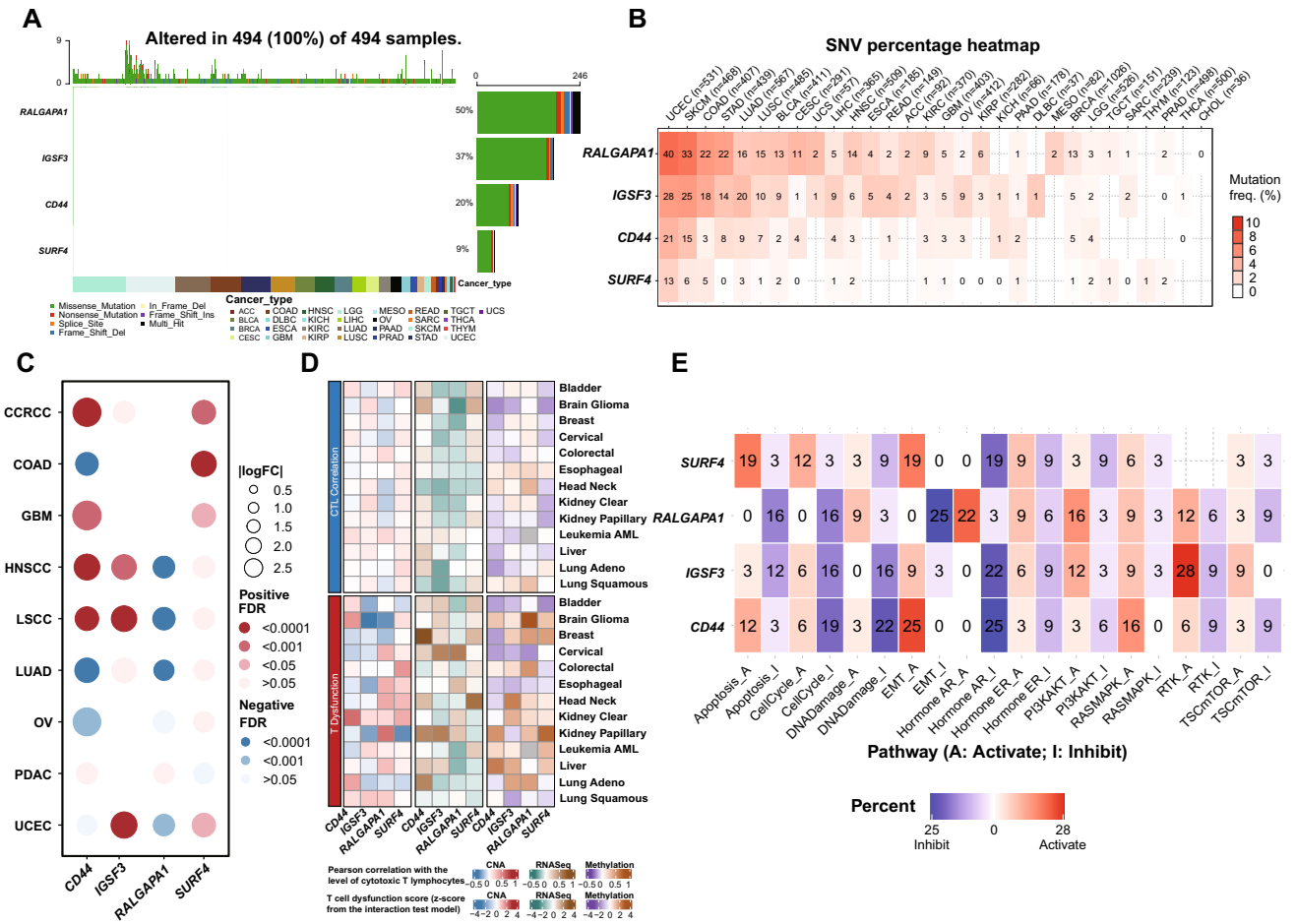


Fig. 7 Evaluation of the four hub genes as potential targets from perspective of pan-cancer multi-omics. **A** Count of deleterious mutations (Missense_Mutaton, Nonsense_Mutation, Frame_Shift_Ins, Splice_Site, Frame_Shift_Del, In_Frame_Del, Multi_Hit) of the four genes in 29 cancer types. **B** Frequency of deleterious mutations in 29 cancer types. **C** Differential protein expression of the four hub genes between tumor and peritumor tissues in CPTAC database and GTEx database. **D** Correlation analysis between the level of cytotoxic T lymphocytes and gene expression and the pre-calculated T cell dysfunction score using the interaction test model in TCGA pan-cancer cohorts. **E** Percentage of cancers between protein expression and caner related pathway activity in TCPA pan-cancer cohorts

many other potential targets. Thirdly, our analysis identified potential correlations between gene mutations and their implications in glioblastoma and pan-cancer. It is important to note that these correlations do not necessarily imply a causal regulatory relationship. Nevertheless, the relatively limited sample size of our own cohort may potentially limit the generalizability of our findings. Additionally, although SURF4 and RALGAPA1 have demonstrated potential immunogenic properties in our data, these require experimental validation to fully elucidate its role in the context of cancer immunotherapy.

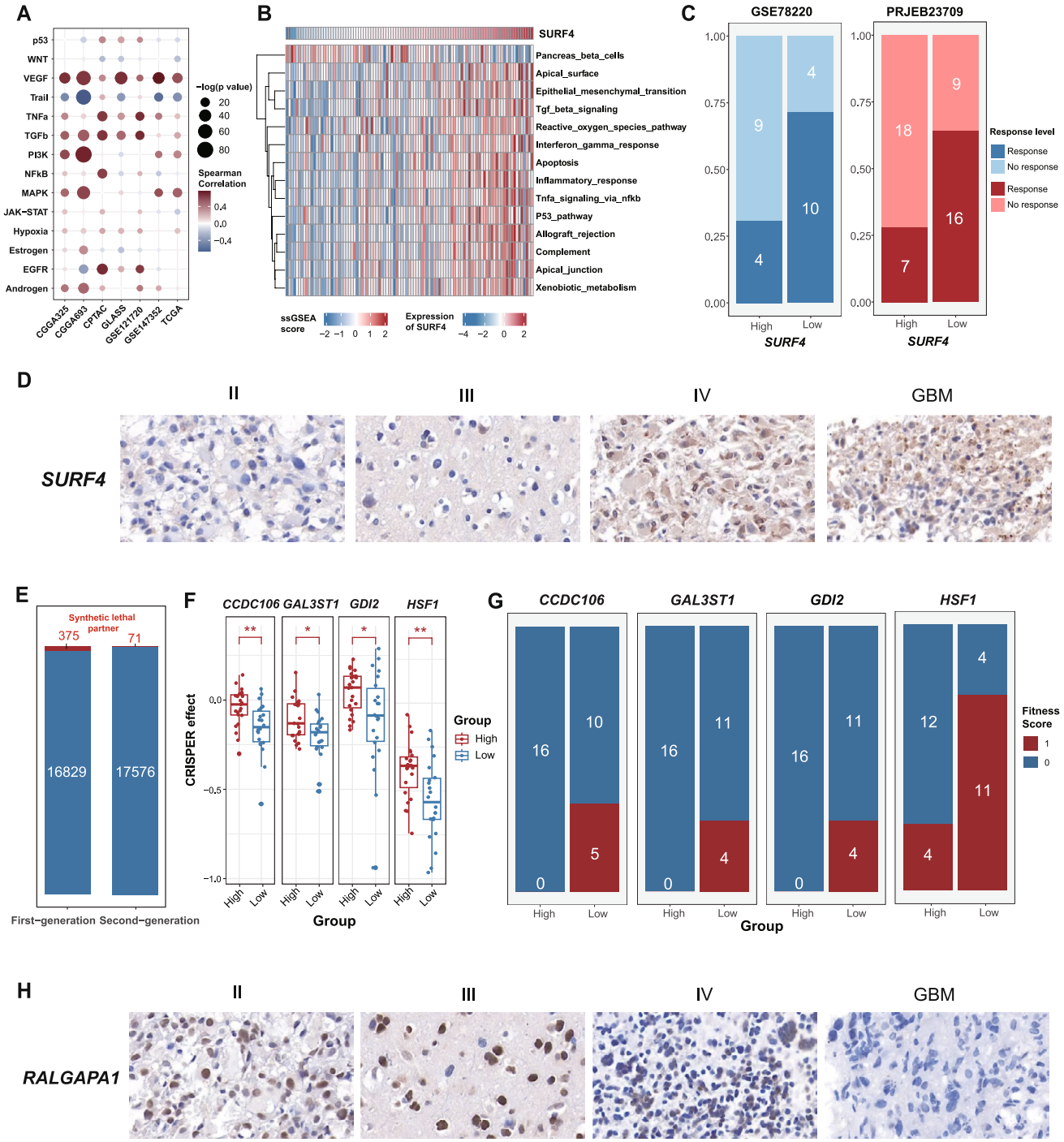


Fig. 8 Evaluation of SURF4 as a potential target from perspectives of pathway activity and immune therapy response, discovery of synthetic lethal partners of RALGAPA1, and immunohistochemical of the two genes. **A** Correlation analysis between SURF4 and 14 pathway activity inferred by "DecoupleR" algorithm at the transcriptomic level in the meta cohort. **B** Pathway activity inferred by ssGSEA algorithm with significant differences by Wilcoxon test between SURF4-high group and SURF4-low group at the proteomic level in CPTAC database. **C** Proportion of patients in GSE78220 and PRJEB23709 cohorts with immune response in the SURF4-high group and SURF4-low group. **D** Immunohistochemical validation of SURF4 protein expression between different grades of glioma. **E** The numbers of synthetic lethal partners identified from transcriptomic data in the first-generation and second-generation DepMap. **F** Differential analysis between high and low RALGAPA1 expression groups with CRISPER effect of the four genes in the first-generation DepMap. **G** Proportion of the four genes with fitness score in the RAPGAPA1-high group and RALGAPA1-low group from the second-generation DepMap. **H** Immunohistochemical validation of RALGAPA1 protein expression between different grades of glioma

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00262-025-04034-y) contains supplementary material, which is available to authorized users.

Acknowledgements We appreciate the valuable and constructive suggestions and assistance from our Team of Neurosurgical Study.

Author Contributions FW presented conceptualization, data curation and visualization. YXC conducted formal analysis, investigation and writing. RH performed investigation, resources and Writing. DFL carried out methodology and validation. JYZ provided resources and methodology. YBY approved project administration and reviewing. HHD analyzed investigation and validation. MRL developed project administration and supervision. ZQC contributed funding acquisition and reviewing. XOS involved in methodology and visualization. ZW prepared conceptualization and funding acquisition. All authors reviewed the manuscript.

Funding This study was supported by the National Natural Science Foundation of China (No. 82171309, 82201445).

Data availability No datasets were generated or analyzed during the current study.

Declarations

Conflict of interests The authors declare no competing interests.

Ethical approval and consent to participate This study was approved by the Ethics Committee of The First Affiliated Hospital of Soochow University. The informed consent was signed by all the patients. According to ethical and legal standards, every specimen was made and handled anonymously.

Consent for publication Not applicable.

References

1. Tan AC et al (2020) Management of glioblastoma: state of the art and future directions. CA Cancer J Clin 70(4):299–312
2. McKinnon C et al (2021) Glioblastoma: clinical presentation, diagnosis, and management. BMJ 374:n1560
3. Schaff LR, Mellinghoff IK (2023) Glioblastoma and other primary brain malignancies in adults: a review. JAMA 329(7):574–587
4. Bedard PL et al (2020) Small molecules, big impact: 20 years of targeted therapy in oncology. Lancet 395(10229):1078–1088
5. Dewdney B et al (2023) From signalling pathways to targeted therapies: unravelling glioblastoma's secrets and harnessing two decades of progress. Signal Transduct Target Ther 8(1):400
6. Eskilsson E et al (2018) EGFR heterogeneity and implications for therapeutic intervention in glioblastoma. Neuro Oncol 20(6):743–752
7. (2008) Comprehensive genomic characterization defines human glioblastoma genes and core pathways. Nature 455(7216):1061–1068
8. Wang LB et al (2021) Proteogenomic and metabolomic characterization of human glioblastoma. Cancer Cell 39(4):509-528.e20
9. Behan FM et al (2019) Prioritization of cancer therapeutic targets using CRISPR-Cas9 screens. Nature 568(7753):511–516
10. Pacini C et al (2024) A comprehensive clinically informed map of dependencies in cancer cells and framework for target prioritization. Cancer Cell 42(2):301-316.e9
11. Katti A et al (2022) CRISPR in cancer biology and therapy. Nat Rev Cancer 22(5):259–279
12. Zhao Z et al (2021) Chinese Glioma Genome Atlas (CGGA): a comprehensive resource with functional genomic data from Chinese glioma patients. Genomics Proteomics Bioinform 19(1):1–12
13. Barrett T et al (2013) NCBI GEO: archive for functional genomics data sets–update. Nucleic Acids Res 41(Database issue):D991–D995
14. (2018) Glioma through the looking GLASS: molecular evolution of diffuse gliomas and the Glioma Longitudinal Analysis Consortium. Neuro Oncol 20(7):873–884
15. Hugo W et al (2017) Genomic and transcriptomic features of response to anti-PD-1 therapy in metastatic melanoma. Cell 168(3):542
16. Gide TN et al (2019) Distinct immune cell populations define response to anti-PD-1 monotherapy and anti-PD-1/anti-CTLA-4 combined therapy. Cancer Cell 35(2):238-255.e6
17. Leek JT et al (2012) The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics 28(6):882–883
18. Lu X et al (2021) MOVICS: an R package for multi-omics integration and visualization in cancer subtyping. Bioinformatics 36(22–23):5539–5541
19. Balduzzi S, Rücker G, Schwarzer G (2019) How to perform a meta-analysis with R: a practical tutorial. Evid Based Ment Health 22(4):153–160
20. Zeng D et al (2021) IOBR: multi-omics immuno-oncology biological research to decode tumor microenvironment and signatures. Front Immunol 12:687975
21. Miao YR et al (2020) ImmuCellAI: a unique method for comprehensive T-Cell subsets abundance prediction and its application in cancer immunotherapy. Adv Sci (Weinh) 7(7):1902880
22. Xu L et al (2018) TIP: a web server for resolving tumor immunophenotype profiling. Cancer Res 78(23):6575–6580
23. Cristescu R et al (2018) Pan-tumor genomic biomarkers for PD-1 checkpoint blockade-based immunotherapy. Science. https://doi.org/10.1126/science.aar3593
24. Martinez-Lage M et al (2019) Immune landscapes associated with different glioblastoma molecular subtypes. Acta Neuropathol Commun 7(1):203
25. Rooney MS et al (2015) Molecular and genetic properties of tumors associated with local immune cytolytic activity. Cell 160(1–2):48–61
26. Ayers M et al (2017) IFN-γ-related mRNA profile predicts clinical response to PD-1 blockade. J Clin Invest 127(8):2930–2940
27. White K et al (2023) Identification, validation and biological characterisation of novel glioblastoma tumour microenvironment subtypes: implications for precision immunotherapy. Ann Oncol 34(3):300–314

28. He Y et al (2018) Classification of triple-negative breast cancers based on Immunogenomic profiling. J Exp Clin Cancer Res 37(1):327
29. Jiang P et al (2018) Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. Nat Med 24(10):1550–1558
30. Liu CJ et al (2023) GSCA: an integrated platform for gene set cancer analysis at genomic, pharmacogenomic and immunogenomic levels. Brief Bioinform 24(1):bbac558
31. Badia IMP et al (2022) decoupleR: ensemble of computational methods to infer biological activities from omics data. Bioinform Adv 2(1):vbac016
32. Calinawan AP et al (2020) ProTrack: an interactive multi-omics data browser for proteogenomic studies. Proteomics 20(21–22):e1900359
33. Li J et al (2013) TCPA: a resource for cancer functional proteomics data. Nat Methods 10(11):1046–1047
34. Han Y et al (2023) TISCH2: expanded datasets and new tools for single-cell transcriptome analyses of the tumor microenvironment. Nucleic Acids Res 51(D1):D1425-d1431
35. Therneau TM (2023) A package for survival analysis in R
36. Hänzelmann S, Castelo R, Guinney J (2013) GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinform 14:7
37. Kim E, Hart T (2021) Improved analysis of CRISPR fitness screens and reduced off-target effects with the BAGEL2 gene essentiality classifier. Genome Med 13(1):2
38. Xu Y, Stamenkovic I, Yu Q (2010) CD44 attenuates activation of the hippo signaling pathway and is a prime therapeutic target for glioblastoma. Cancer Res 70(6):2455–2464
39. Jijiwa M et al (2011) CD44v6 regulates growth of brain tumor stem cells partially through the AKT-mediated pathway. PLoS ONE 6(9):e24217
40. Carvalho AM, Reis RL, Pashkuleva I (2023) Hyaluronan receptors as mediators and modulators of the tumor microenvironment. Adv Healthc Mater 12(5):e2202118
41. Xu Y et al (2024) CD44 and its implication in neoplastic diseases. MedComm. 5(6):e554
42. Curry RN et al (2023) Glioma epileptiform activity and progression are driven by IGSF3-mediated potassium dysregulation. Neuron 111(5):682-695.e9
43. Wagner M et al (2020) Bi-allelic variants in RALGAPA1 cause profound neurodevelopmental disability, muscular hypotonia, infantile spasms, and feeding abnormalities. Am J Hum Genet 106(2):246–255
44. Zhai J et al (2022) High SURF4 expression is associated with poor prognosis of breast cancer. Aging (Albany NY) 14(22):9317–9337
45. Yue Y et al (2020) SURF4 maintains stem-like properties via BIRC3 in ovarian cancer cells. J Gynecol Oncol 31(4):e46