scientific reports

Check for updates

OPEN

CA3 bridges dietary restriction to glioblastoma suppression and tumor progression as a key downstream effector

Junxiang Mao[1,4], Zhibiao Cai[2,4], Dong Xie[3,4], Man Guo[3], Yu Gao[3], Guohui Zhao[3] & Jie Zhou[1]✉

Dietary restriction (DR) is recognized as a health-promoting, non-pharmacological intervention with demonstrated inhibitory effects on the initiation and progression of cancer. The molecular mechanisms underpinning DR's anticancer activity are pivotal, with documented evidence of its suppressive role across a spectrum of cancers. Glioblastoma multiforme (GBM) represents an aggressively malignant intracranial neoplasm, and despite incremental therapeutic and managerial advancements, the clinical outcomes remain suboptimal. Consequently, the discovery of novel molecular markers to augment diagnostic accuracy and therapeutic efficacy is imperative. Employing an array of bioinformatics strategies, we conducted an exhaustive analysis of molecules associated with DR, culminating in the identification of CA3 as a novel molecular marker for GBM. We evaluated its diagnostic and therapeutic potential within GBM. Our data indicate that the DR-associated molecule CA3 may exhibit correlations with multiple GBM phenotypes, including the immune contexture, with particular emphasis on the tumor's invasive and migratory capacities. Subsequent inquiries confirmed that modulating CA3 expression can effectively curb the genesis and progression of GBM. Our research substantiates that DR can mitigate the onset and development of GBM via the gene CA3, thereby validating a novel GBM marker and proposing a non-pharmacological interventional approach for this life-threatening condition.

Keywords Dietary restriction, Glioblastoma multiforme, CA3, Single-cell transcriptome, Biomarker

DR entails the reduction of caloric intake via dietary means, excluding malnourishment, yet ensuring sufficient provision of essential amino acids, vitamins, and nutrients—a regimen that has been demonstrated to enhance the well-being of rodents[1]. Neoplastic cells exhibit heightened sensitivity to deficiencies in nutrients and energy, and preclinical investigations have established that DR mitigates the toxic side effects of chemotherapy in animals, underscoring its potential as a non-pharmacological preventative measure against cancer[2,3]. The tumor-suppressive properties of DR are hypothesized to be associated with the modulation of growth factor signaling pathways, inflammatory cytokines, anabolic metabolic hormones, and attenuation of oxidative stress injuries[4]. Post-DR, there is a pronounced decrease in circulating levels of glucose, insulin-like growth factor 1 (IGF-1), and growth hormone, with the latter being implicated in tumorigenesis through the activation of the Ras/MAPK and PI3 K/AKT signaling cascades[5,6]. Additionally, DR has been shown to elicit the activation of AMP-activated protein kinase and to counteract the Warburg effect, thereby enhancing apoptosis, as confirmed in in vitro oncological models[2,7]. The pursuit of therapeutic interventions for cancer patients through DR to achieve oncological and preventative outcomes continues to be a focal point of scientific inquiry.

GBM, emerging from astrocytes, stands as one of the most invasive malignant neoplasms and is the predominant malignant primary brain tumor within the central nervous system. It constitutes the majority of malignant central nervous system tumors, representing roughly 48.6% of all malignant intracranial tumors[8]. Despite the current therapeutic regimens, the efficacy against GBM is still constrained, with the standard treatment protocol involving surgical resection followed by temozolomide-assisted chemoradiation and adjuvant

[1]The Second Hospital & Clinical Medical School, Lanzhou University, No.82 Cuiyingmen, Linxia Road, Chengguan District, Lanzhou City 730000, Gansu Province, China. [2]Department of Neurosurgery, The 940 th Hospital of Joint Logistics Support force of Chinese People's Liberation Army, No.333, Nanbinhe Road, Qilihe District, Lanzhou City 730000, Gansu Province, China. [3]Department of First Clinical College of Medicine, Gansu University of Traditional Chinese Medicine, No. 35 Dingxi East Road, Chengguan District, Lanzhou City 730000, Gansu Province, China. [4]Junxiang Mao, Zhibiao Cai and Dong Xie contributed equally. ✉email: ery_zhouj@lzu.edu.cn

chemotherapy[9]. Even with adherence to these standard treatments, the 5-year survival rate for GBM remains below 5%, with a median survival span of less than 15 months, and the adverse effects of chemotherapeutic agents also significantly affect patients' quality of life[10,11]. Consequently, the pursuit of novel biomarkers and therapeutic strategies for GBM is an imperative in contemporary medical research.

In the present study, leveraging multiple online databases, we performed a comprehensive multi-omics analysis at the molecular, cellular, and tissue levels to evaluate the influence of DR on the genesis and progression of GBM. We identified 14 genes exhibiting differential expression between neoplastic and normal tissues, correlating with DR, and delineated specific associations with genomic mutations and immune molecular signatures. Single-cell profiling from our investigation revealed heterogeneity in the molecular activity associated with DR within the tumor microenvironment, linked to cellular communication and paracrine interactions between neoplastic and immune cells. Moreover, our research established a significant correlation between the DR-related molecule CA3 and the invasiveness and migratory capacity of tumors, suggesting its potential as an early diagnostic biomarker for GBM. Significantly, our findings also indicate that CA3 is associated with enhanced overall survival (OS), progression-free interval (PFI), and disease-specific survival (DSS) in GBM patients. Collectively, our study underscores the role of DR-related molecule CA3 activity in the amelioration of GBM, uncovering a novel diagnostic biomarker and therapeutic target for this highly aggressive disease.

Results

Identification of DR-related genes in GBM

After merging GBM datasets from TCGA and GTEx databases, we ascertained 1474 DEGs, comprising 624 upregulated and 850 downregulated genes, which were illustrated via volcano plots and the heatmap of the top 30 genes based on the P value (Fig. 1a, b), followed by GO and KEGG enrichment analyses (Fig. 1c, d). The genes associated with drug resistance intersected with DEGs, culminating in 14 intersecting genes (KCNA1, SCRT1, GPR146, HIST2H4 A, HIST2H4B, NR1D1, IFIH1, G0S2, FZD1, RGS16, CA3, HSPA5, CTGF, EIF4EBP1) (Fig. 1e). These 14 genes were further validated for expression differences between normal and tumor groups using the RNA-seq dataset corrected within the PanCanAtlas database, revealing that only 8 genes exhibited differential expression (CA3, FZD1, G0S2, HSPA5, IFIH1, KCNA1, NR1D1, SCRT1) (Fig. 1f).

Survival analysis of multiple genes

ssGSEA was applied to the 8 genes of interest to ascertain enrichment scores, which subsequently informed KM curves for OS, DSS, and PFI in GBM (Fig. 2a, b, c). The findings suggested that elevated enrichment scores for these genes were correlated with a more adverse prognosis, a conclusion corroborated by the external Rembrandt dataset (Fig. 2d). A meta-analysis of survival data from TCGA and Rembrandt confirmed these observations (Fig. 2e).

Machine learning

In order to delineate the pivotal genes within the genomic landscape, a suite of six computational methods was implemented, encompassing the Boruta algorithm (Fig. 3a), Lasso regression (Fig. 3b), xgboost algorithm (Fig. 3c), the randomForest function from the randomForest package for constructing a random forest model (Fig. 3d), Multi-algorithmic approaches (Fig. 3e, f), and the support vector machine recursive feature elimination (SVM-RFE) algorithm (Fig. 3g). These methodologies were leveraged to discern the genes within the gene ensemble that exert a significant effect on GBM. The intersection of outcomes derived from these diverse algorithms culminated in the identification of the CA3 gene (see Supplementary Table S1 online), a constituent of the multi-gene family, predominantly functioning as a homodimeric enzyme that encodes carbonic anhydrase, as a crucial gene in the gene set's influence on GBM (Fig. 3h).

Discrepancy in expression and diagnostic utility

Upon comparison with normal tissues, we noted that CA3 is scarcely expressed in the brain (Fig. 4a, b). However, it is intriguing to observe the presence of CA3 in brain tumors, where its expression surpasses that of the majority of malignant neoplasms (Fig. 4c). The expression levels of CA3 in tumor tissues exceed those in normal tissues (Fig. 4d), and it demonstrates commendable diagnostic potential (Fig. 4e). These findings were substantiated through the integrated analysis of TCGA and GTEx datasets (Fig. 4f, g). Furthermore, we have detected a plethora of differentially expressed genes, particularly those that are highly expressed in tumors, which may have latent connections with the functionalities executed by CA3 (Fig. 4h). Concurrently, proteins that may interact with CA3 have been characterized within our study (Fig. 4i).

Survival analysis pertaining to CA3

Our observations have highlighted an inverse correlation between CA3 and DR, leading us to infer that elevated CA3 expression in GBM correlates with an adverse survival prognosis. To substantiate this hypothesis, a spectrum of survival-oriented analyses were conducted for CA3 across diverse databases. Analyses from the TCGA-GBM cohort demonstrate that individuals with heightened CA3 expression levels in GBM exhibit a markedly diminished OS rate relative to those with reduced expression (Fig. 5a), a finding corroborated by further assessments within the CGGA and Rembrandt databases (Fig. 5b, c, d). A synthesis of survival data through meta-analytical approaches underscores CA3's role as an independent predictor of poor prognosis in GBM (Fig. 5e).

Analysis of metabolic pathways and associated functionalities

In order to delineate the precise pathways through which CA3 mediates its biological effects, GSEA and KEGG enrichment analyses were utilized. The outcomes highlighted a significant enrichment of CA3 in pathways

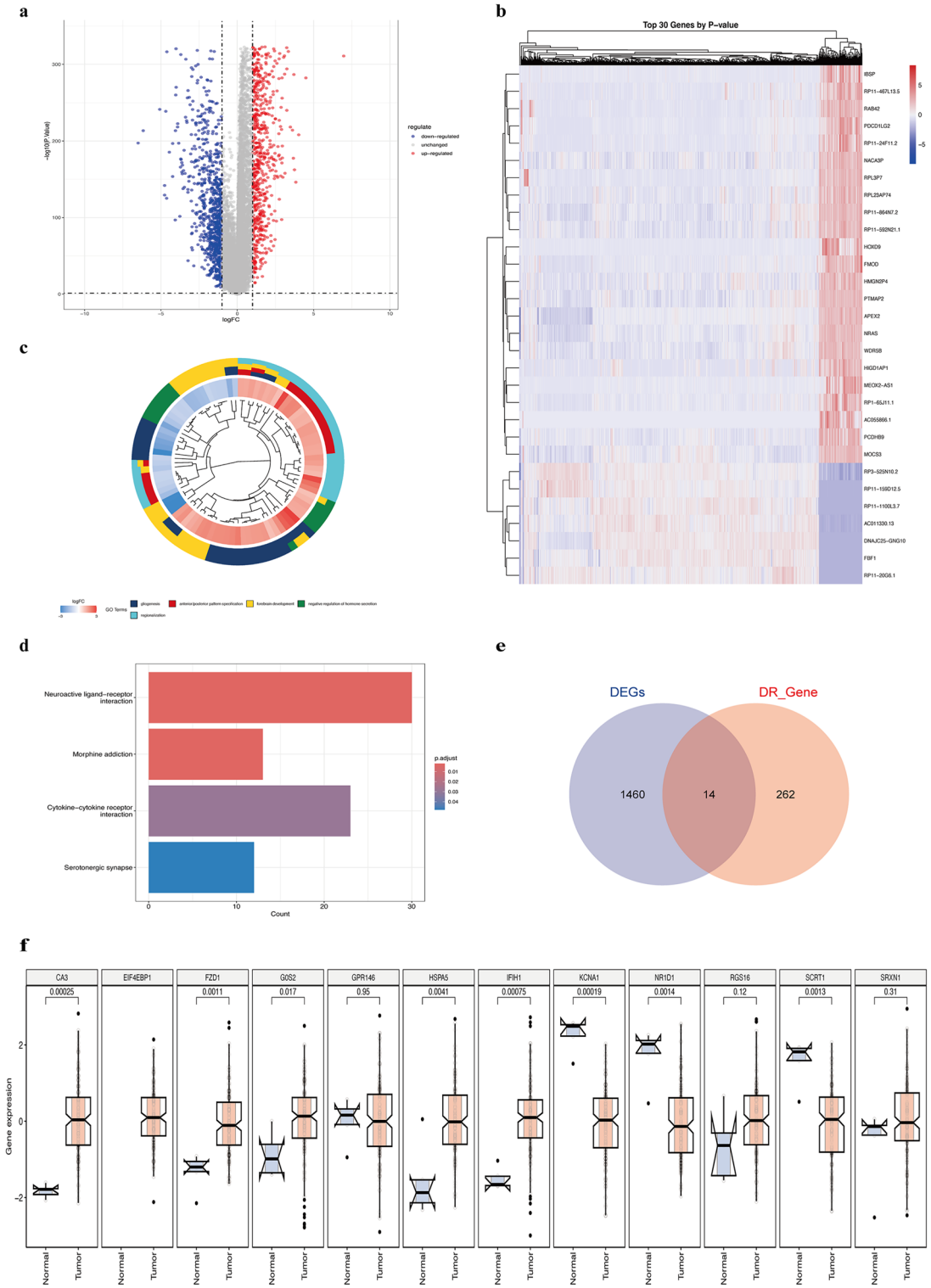


Fig. 1. Conjoint analysis of TCGA and GTEx databases identified differentially expressed genes and explored the misregulation of DR-associated genes within GBM. **(a)**The volcano plot delineating differential gene expression between TCGA and GTEx GBM datasets, with blue dots indicating downregulated genes and red dots signifying upregulated genes in tumor tissues compared to normal tissues. The threshold for selection was set at (log2|fold-change| > 1 and P value < 0.05). (b)The heatmap of the top 30 genes based on the P-value. **(c)**GO enrichment analysis for the DEGs. **(d)**KEGG enrichment analysis for the DEGs. **(e)**The intersection of DEGs with DR-associated genes identified 14 pertinent genes. **(f)**The expression patterns of the intersecting genes within GBM are depicted.

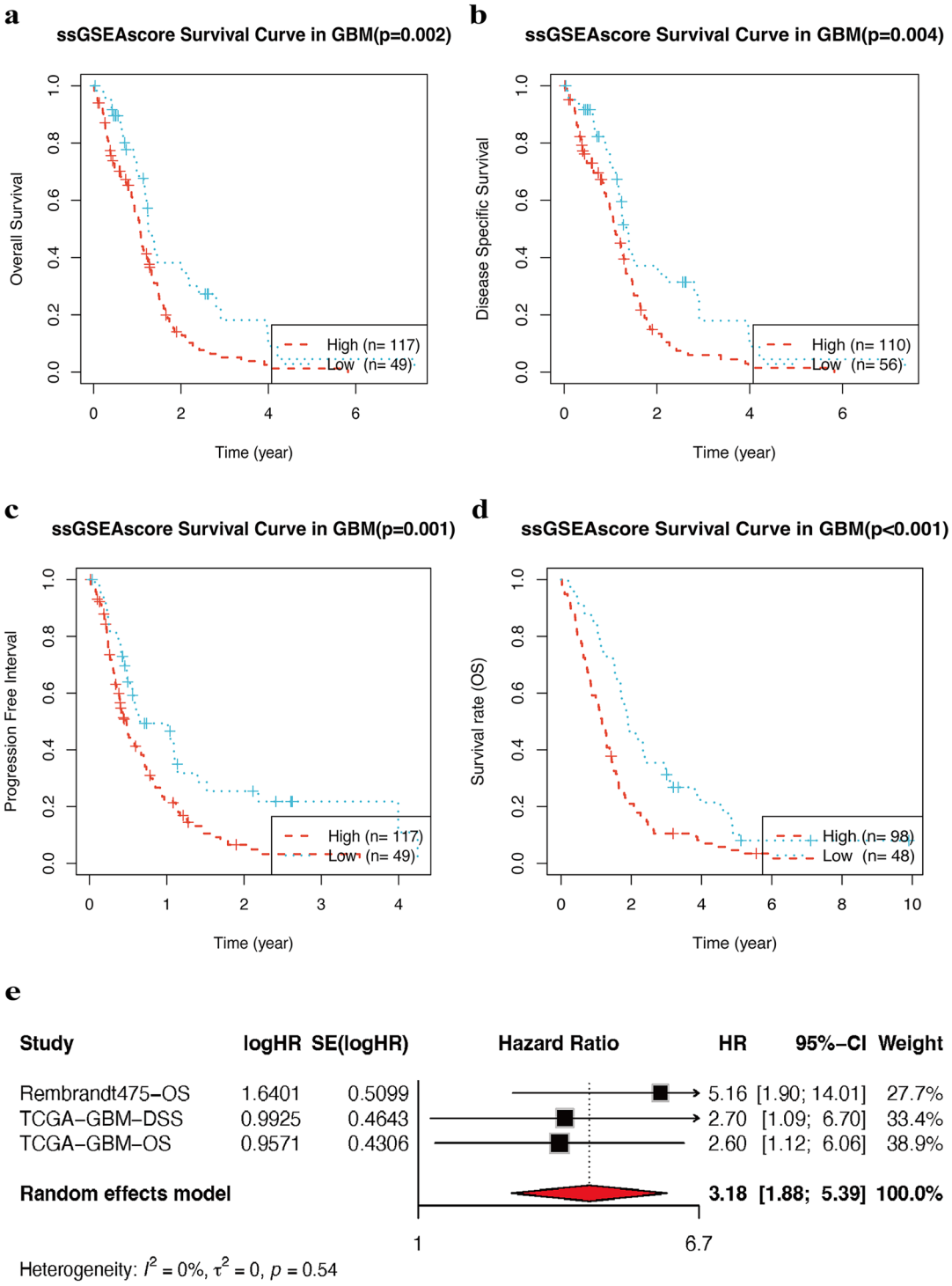


Fig. 2. ssGSEA of CA3, FZD1, G0S2, HSPA5, IFIH1, KCNA1, NR1D1, and SCRT1 in GBM. **(a-c)** KM survival curves from the survival analysis demonstrate that the high-expression cohort of these gene sets exhibits a more unfavorable prognosis in terms of OS, DSS, and PFI. **(d)** Examination of OS in the Rembrandt database for GBM similarly reveals that increased gene expression correlates with a poorer prognosis. **(e)** A meta-analysis of survival data from both the TCGA and Rembrandt databases validates that these eight genes are significant risk factors for GBM.

associated with immune responses, cellular proliferation, and metabolic processes (Fig. 6a, b)[12–14], the activation of which may be intricately linked to immune evasion in tumors, inflammatory milieus, metabolic reprogramming, and the formation of the tumor microenvironment. The HALLMARK_INTERFERON_ALPHA_RESPONSE pathway, which exhibited the highest normalized enrichment score (NES), was subjected

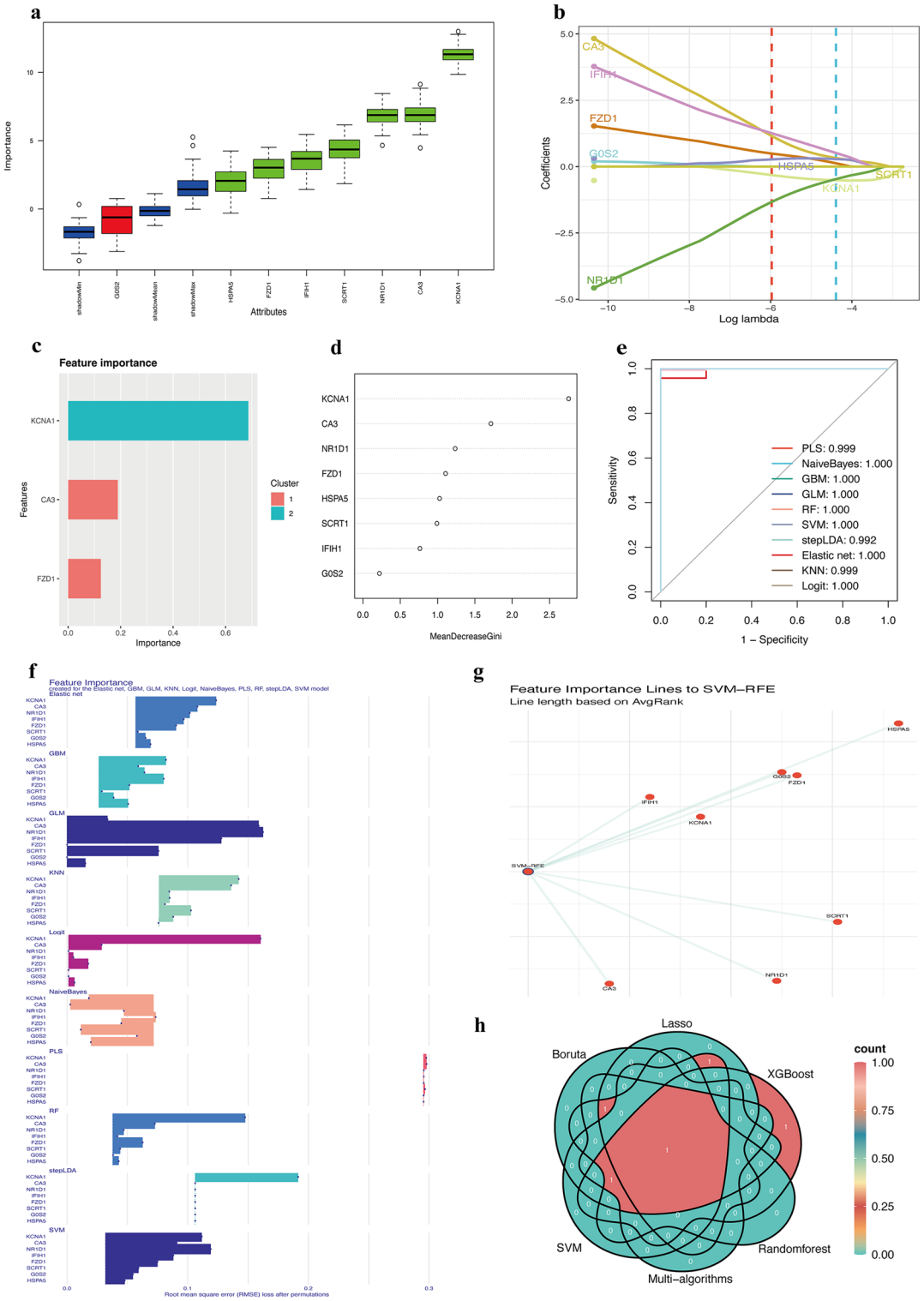
to further validation within the TCGA cohort, where it demonstrated pronounced enrichment in the group with elevated CA3 expression (Fig. 6c). These observations were substantiated in the Chinese Glioma Genome Atlas (CGGA) dataset across the 325 (Fig. 6d) and 693 (Fig. 6e) cohorts, as well as in the Rembrandt dataset (Fig. 6f). Employing GSVA, it was uncovered that numerous metabolic pathways, such as Nitrogen metabolism, Tryptophan metabolism, and Histidine metabolism, are activated in the GBM cohort with heightened CA3 expression (Fig. 6g, h). Concurrently, an exploration of the Pearson correlation between gene expression z-scores and GSVA scores, evaluated with 14 z-score parameters indicative of tumor states, uncovered a significant positive correlation between CA3 and attributes of Proliferation and Stemness within GBM (Fig. 6i).

Analysis of CA3 Co-expressed genes and associated proteins

Employing WGCNA, we investigated genes co-expressed with CA3, pinpointing 1561 genes within a green module that exhibit a high degree of correlation with CA3 (Fig. 7a, b, c, d). These genes demonstrate a marked positive correlation within the Q1 group, characterized by elevated CA3 expression, and a significant negative correlation within the Q4 group. An enrichment analysis was conducted on this gene set (Fig. 7d). GO analysis indicated that these genes are predominantly engaged in biological processes associated with cilia and molecular functions related to the activities of cellular skeletal motors. KEGG analysis delineated the enrichment of specific metabolic pathways, suggesting a potential involvement of these genes in cellular metabolic activities and functions (Fig. 7e, f)[12–14]. These findings may suggest that the 1561 genes highly correlated with CA3 are implicated in the invasion, migration, proliferation, and intercellular interactions within the tumor microenvironment of GBM, potentially exerting an influence on tumor cells through the modulation of cellular metabolism. Given the complexity of GBM, genetic modifications provide a limited understanding of disease etiology and progression. To further elucidate the impact of CA3 on the genesis and development of GBM, we utilized the TCPA database to identify the CD49B functional protein, which exhibits a significant positive correlation with CA3 (Fig. 7g). This finding further implies that elevated CA3 expression may enhance the invasive properties of GBM. Concurrently, a positive correlation between CA3 and the epithelial-mesenchymal transition (EMT) pathway within GBM was observed (Fig. 7h). In the cohort of patients afflicted with GBM, the epidermal growth factor receptor (EGFR) emerges as the predominant mutated oncogene[15]. The gene amplification of EGFR is recognized for its capacity to drive tumorigenesis, including proliferation, angiogenesis, and invasiveness, through activation of the RAS and PI3 K-AKT signaling axes, which correlates robustly with diminished OS in GBM[16]. The therapeutic targeting of EGFR has been validated as an efficacious strategy in GBM management. Our research delineates an upregulation of EGFR in GBM patients (see Supplementary Fig. S1a online), concomitant with a consistent positive correlation between EGFR and CA3 across diverse GBM dataset analyses (see Supplementary Fig. S1a online). Bcl-2-associated antiapoptotic protein 3 (BAG3) has been demonstrated to be intricately linked with CA3. It has been documented in the literature that GBM cells exhibit heightened sensitivity to apoptosis following the downregulation of BAG3[17]. Within the CGGA-693 cohort, a significant positive correlation is observed between CA3 and BAG3 (see Supplementary Fig. S1b online), implying a potential role for CA3 in the apoptotic mechanisms of GBM cancer cells. The PI3 K/AKT/mTOR signaling cascade, which orchestrates cellular growth, survival, and proliferation, is initiated by PI3 K, with the PIK3R1 gene encoding the principal regulatory subunit p85α of type I PI3 K[18]. Our research reveals an upregulation of PIK3R1 expression in GBM, which correlates positively with CA3 (see Supplementary Fig. S1c online). The phosphatase and tensin homolog (PTEN) acts as an inhibitor of the PI3 K/AKT/mTOR signaling pathway. Despite an upregulation of PTEN expression in GBM, a negative correlation with CA3 is observed (see Supplementary Fig. S1d online). Collectively, these findings suggest that CA3 may synergize with BAG3 and the PI3 K/AKT/mTOR signaling axis to foster tumorigenesis and proliferation. Although no studies have yet explored the nexus between CA3 and redox homeostasis in cancer, CA3 is known to exert antioxidant effects in intervertebral disc cells and osteoblasts of rats[19,20]. The protein encoded by thioredoxin reductase 1 (TXNRD1) is part of the pyridine nucleotide disulfide oxidoreductase family and constitutes a component of the thioredoxin system, which is implicated in tumor proliferation, invasion, and chemoresistance[21]. Our findings indicate that TXNRD1 expression is elevated in GBM and exhibits a positive correlation with CA3 (see Supplementary Fig. S1e online).

Cell subpopulations correlated with CA3 expression disruption in GBM

A pan-cancer screening of CA3 was conducted on single-cell data from the TISCH database, excluding datasets with null expression across all cells; datasets with uniformly low expression levels (with an average expression greater than 1) to mitigate false-positive outcomes; and datasets where expression was null in tumor cells (Fig. 8a, b, c). This process led to the identification of three datasets relevant to GBM: Glioma_GSE84465, Glioma_GSE131928_10X, and Glioma_GSE141460. We utilized Glioma_GSE131928_10X for further single-cell analysis (Fig. 8c). The UMAP result plots clearly delineate distinct cell populations separated based on their expression profiles (Fig. 8d). These segregated cell populations correspond to various cell types or states, including Mesenchymal Stem Cells (MES)-like Malignant, Neural Progenitor Cells (NPC)-like Malignant, Oligodendrocyte Progenitor Cells (OPC)-like Malignant, Astrocytes (AC)-like Malignant, Monocyte, M1, CD8 Tex cells, and others. The spatial distribution of each cell type in the UMAP plot mirrors their expression signatures and interactions with different cell types. Notably, we observed significant disparities in CA3 gene expression across different cell clusters, with a pronounced increase in malignant cell populations and markedly higher expression in MES-like Malignant and AC-like Malignant cells (Fig. 8e, f, g, h). Within the Glioma_GSE131928_10X dataset, the CA3 expression-positive cohort exhibited a substantially greater proportion of MES-like Malignant and AC-like Malignant cells compared to the CA3 expression-negative cohort (Fig. 8i).



Single-Cell pathway analysis of CA3

Our cellular-level analysis of pathways associated with CA3 revealed significant enrichment of CA3 in reactive oxygen species-related pathways within MES-like Malignant cells (Fig. 9a). Analysis of intercellular communication patterns shows that CA3 + Malignant cells are predominantly connected with MES-like Malignant cells, AC-like Malignant cells, OPC-like Malignant cells, and other malignant cell types (Fig. 9b, c, d). This implies that CA3 may act as a biomarker for GBM, with a strong correlation to the invasive and migratory behaviors of GBM. Furthermore, we observed that CA3^{+} Malignant cells modulate various cellular constituents within GBM through a spectrum of cellular signaling pathways, contributing to the initiation and progression of the tumor (Fig. 9e, f, g).

◀**Fig. 3**. Utilization of Diverse Machine Learning Algorithms for the Discovery of Critical Genes within Genomic Sets. **(a)**The Boruta algorithm elucidates the gene set's significance in GBM, highlighting cyan-colored features as those "confirmed" to be significantly correlated with the predictive variable by the Boruta algorithm. **(b)**The coefficient path plot from Lasso regression delineates the fluctuation in model coefficients (y-axis) across the spectrum of the regularization parameter λ (lambda), which is depicted on the x-axis in logarithmic scale. **(c)**The graphical representation of feature importance scores within a model is illustrated via a bar chart, where elevated scores denote a more significant contribution of the feature to the model's predictive capacity. The features are meticulously ordered from the most to the least influential according to their allocated importance scores. **(d)**The Mean Decrease in Gini index, or MeanDecreaseGini score, is delineated. **(e)**The top 10 genes of importance for each model are identified. **(f)**Proximity of the ROC curve to the top-left corner signifies model excellence. **(g)**Utilizing the SVM-RFE algorithm in conjunction with ten-fold cross-validation, the mean rank for each feature is ascertained, with each gene's mean rank represented by its proximity to the SVM-RFE scatter plot; an elevated mean rank is indicative of a position more adjacent to the SVM-RFE scatter plot.**(h)**Genes ascertained by the aforementioned algorithms culminate in the final identification of CA3 as the cross-gene.

Therapeutic potential of CA3 in GBM

Analyses have established CA3 as a diagnostic and prognostic marker in GBM and have hinted at its association with GBM's invasive and migratory behaviors. Consequently, we have forecast chemotherapeutic agents associated with CA3. Our findings indicate a significant inverse correlation between CA3 and a range of chemotherapeutic drugs, suggesting that higher CA3 expression correlates with increased GBM sensitivity to chemotherapy (Fig. 10a). Further refinement identified a pronounced negative correlation between CA3 expression and the IC50 of AZD6482, signifying that elevated CA3 expression confers heightened sensitivity to this drug (Fig. 10b). In contrast to normal brain tissues, CA3 is markedly overexpressed in GBM (Fig. 10c). Employing the X-Sum algorithm to deduce potential molecular and pharmacological interventions to counteract the biological repercussions of CA3 overexpression, we discovered that NU.1025 could potentially target the reversal of molecular signatures induced by CA3 expression anomalies, thus offsetting CA3-driven oncogenic activities (Fig. 10d). To authenticate our findings, we forecasted the chemotherapeutic efficacy of AZD6482 and NU.1025 within various GBM datasets, juxtaposing these with recognized CA activators (D-phenylalanine, histamine) and inhibitors (acetazolamide, topiramate, methazolamide). The data indicated that CA3 expression levels correlated positively with the AUC values of D-phenylalanine across multiple datasets from the CGGA, E-MTAB, GEO, and TCGA (see Supplementary Fig. S2a online), and with histamine across datasets from the CGGA, E-MTAB, and GEO (see Supplementary Fig. S2b online), suggesting that heightened CA3 expression is linked to augmented sensitivity to D-phenylalanine and histamine in GBM patients. Conversely, the AUC values for acetazolamide and topiramate revealed a negative correlation with CA3 expression (see Supplementary Fig. S2c, d online), implying that an escalation in CA3 expression is concomitant with increased resistance to these medications in GBM patients. Peculiarly, methazolamide's AUC values manifested a positive correlation with CA3 expression in the CGGA-693 and GSE3041 datasets, yet a negative correlation in the GSE7696, GSE43378, and E-MTAB-3892 datasets (see Supplementary Fig. S2e online). Significantly, the IC50 of AZD6482 exhibited a negative correlation with CA3 expression across various GBM datasets, including those from the CGGA, E-MTAB, and GEO databases (see Supplementary Fig. S3a online). Moreover, the expression levels of CA3 within multiple GBM datasets from the CGGA and GEO databases were robustly correlated with the AUC values of NU.1025, denoting that an upsurge in CA3 expression is predictive of heightened sensitivity to NU.1025 in GBM patients (see Supplementary Fig. S3b online). Collectively, these outcomes reinforce the candidacy of CA3 as a therapeutic target in GBM and underscore the promise of AZD6482 and NU.1025 as potential therapeutics. Additionally, to substantiate the role of CA3 in GBM, we conducted a comprehensive CRISPR-Cas9 screen across the DepMap database for cell lines with a pan- CERES. Our discovery that CA3 exhibits a negative CERES in several GBM cell lines, notably the M059 K cell line, implies that CA3 ablation elicits growth suppression and/or cell death (Fig. 10e). However, the CERES score did not plummet to −1 in GBM cell lines, indicating that CA3 is not an ubiquitously core essential gene, offering a basis for the selection of cells for future experimental validation of CA3's.

Discussion

GBM poses a formidable health challenge, highlighting the urgent need for novel biomarkers to facilitate early detection and intervention, thereby significantly enhancing the prospects for successful cancer treatment and improving patient outcomes in terms of cure rates, survival, and overall quality of life. A multitude of investigations has substantiated the efficacy of DR as a cancer therapeutic strategy, notable for its non-invasive nature that eschews pharmacological and surgical interventions. This approach offers a more cost-effective and accessible method for managing GBM, a condition that continues to challenge contemporary medical science[22]. Molecules implicated in DR are pivotal in orchestrating the anti-neoplastic effects of DR, yet a thorough evaluation of their attributes and their biological relevance within the context of GBM is an area ripe for further exploration. In this research, we leveraged the TCGA and GTEx databases to uncover differentially expressed genes correlated with the DR gene set in GBM. Employing a suite of analytical tools, including machine learning algorithms, WGCNA, GO, and KEGG enrichment analyses, we pinpointed the gene CA3 as a key element within the DR gene set pertinent to GBM. Our findings elucidate the interplay between CA3 and immune phenotypes, genomic traits, and clinical outcomes, offering insights that may catalyze advancements in GBM diagnostics and therapeutic strategies.

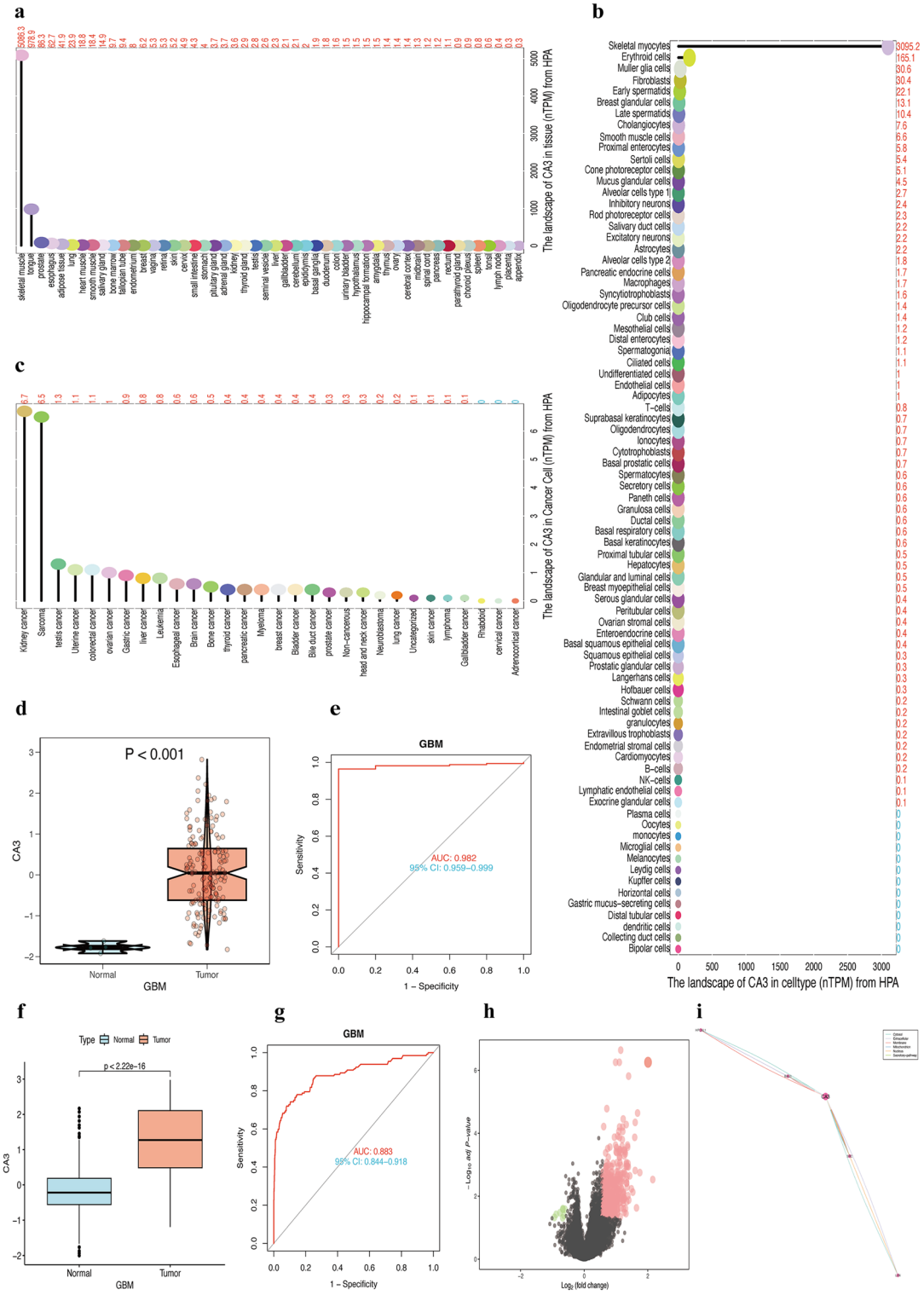


Fig. 4. Expressional Topography of CA3, Its Differential Expression in GBM, Diagnostic Utility, and CA3-Associated Genes within GBM. (**a**) The expression profile of CA3 within normal tissue contexts. (**b**)The expressional landscape of CA3 across cellular milieus. (**c**) The expressional landscape of CA3 within the sphere of cancer cells. (**d**) Discrepancies in CA3 expression specific to GBM. (**e**) Diagnostic efficacy attributed to CA3 for GBM as delineated by the TCGA database. (**f**) Varied expressional patterns of GBM as observed through the amalgamation of TCGA and GTEx datasets. (**g**) Diagnostic proficiency of CA3 for GBM, as ascertained by the integrated TCGA and GTEx databases. (**h**) A gradient of intensifying coloration and magnitude in the plotted points is observable with escalating logFC magnitudes and enhanced significance of the adjusted p values. **(i)** Proteins that may engage in potential interactive relationships with CA3.

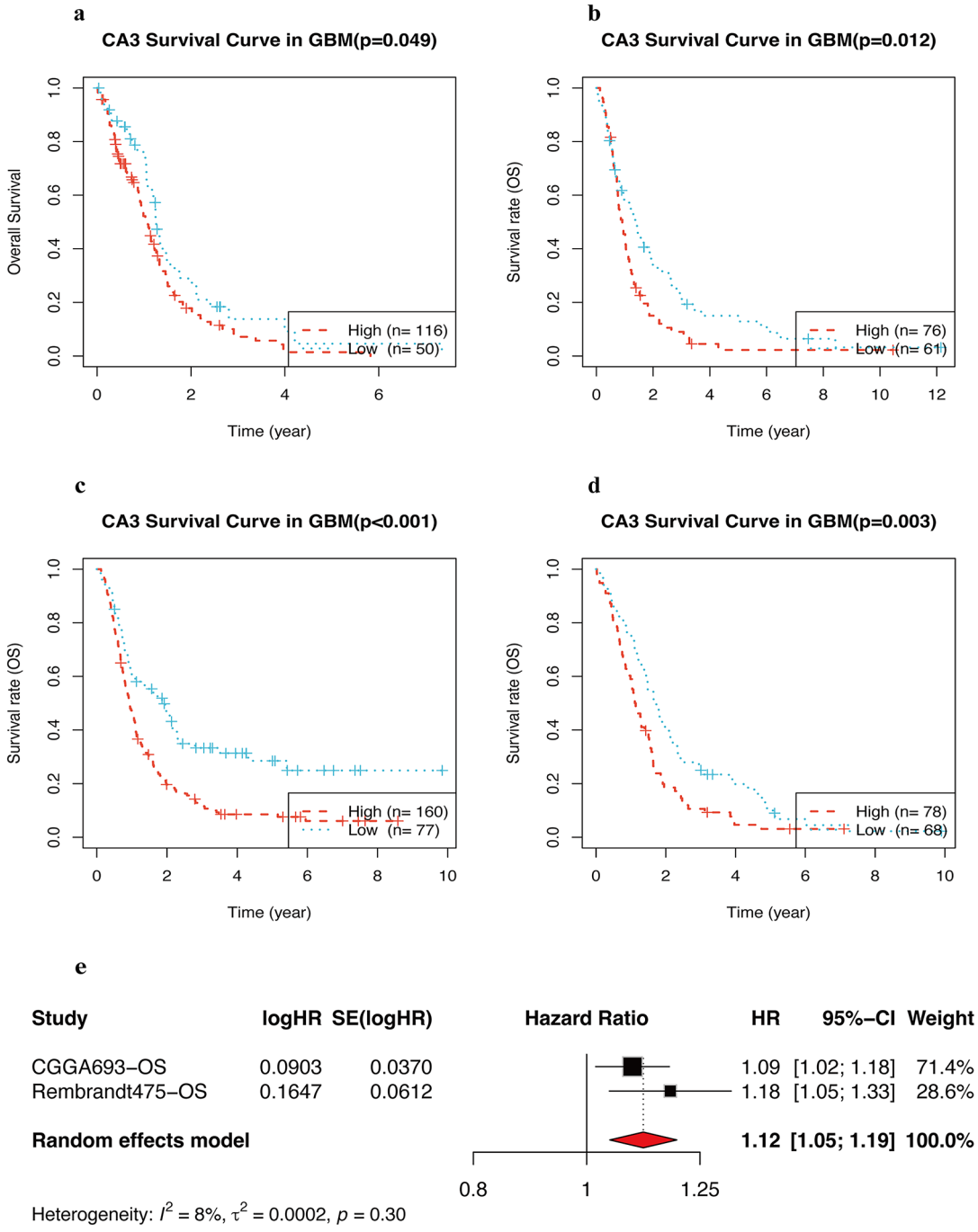
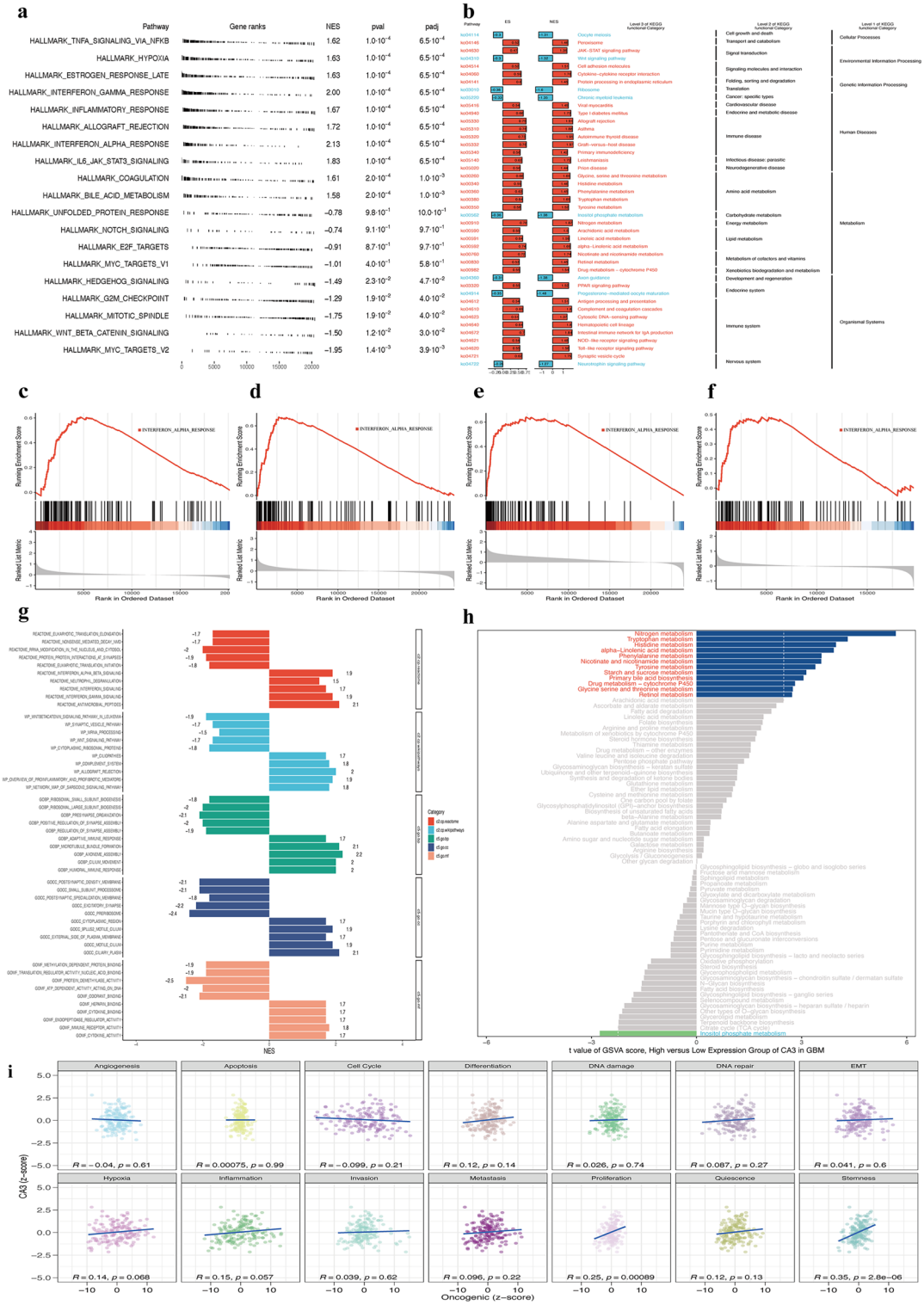


Fig. 5. Prognostic Evaluation of CA3 within the Context of GBM. **(a)** Within the TCGA dataset, an elevated expression of CA3 correlates with a diminished OS in GBM. **(b-d)** Validation through external datasets confirms the association between CA3 expression and GBM prognosis, as evidenced by a positive correlation between heightened CA3 expression and reduced OS rates across the CGGA-325 (**b**), CGGA-693 (**c**), and Rembrandt-475 (**d**) cohorts. (**e**) A meta-analysis of hazard ratios for survival underscores the datasets where CA3 exerts a significant predictive influence on GBM prognosis.

Significantly, prior research has established that CA3 holds a pivotal role in malignancies such as lung, pancreatic, and bladder cancers[23,24]。 However, the diagnostic potential of CA3 in GBM remains unexplored. Our study ascertains that CA3 offers precise diagnostic efficacy for GBM, suggesting that its heightened expression could be harnessed as a precious biomarker for the early diagnosis of GBM. Moreover, the expression of CA3 could be targeted through DR to curb the invasive and migratory behaviors of GBM. Preliminary studies on CA3 have indicated its significance in lung cancer and, in bladder cancer, identified CA3 as a novel gene associated with tumor-, chromosome-, and pathway-specific methylation differences, marking it as a potential cancer biomarker for early detection[25]. Our findings reveal that GBM patients with elevated CA3 expression



levels are associated with a reduced likelihood of survival. Additionally, our ROC analysis underscores CA3's substantial diagnostic merit, evident in analyses confined to the TCGA database or expanded to include the GTEx database. Mesenchymal cells, categorized as non-tumorigenic stromal cells, are detected within the GBM microenvironment, at the tumor's core and periphery, where they facilitate tumor cell proliferation, invasion, and angiogenesis, thereby exerting a critical influence on GBM pathogenesis[26]. The prevailing hypothesis posits that GBM-associated MES stem from the migration of MES from the spinal cord or other tissues[26]. Although the differentiation mechanisms remain obscure, scant evidence hints at a potential origin from GBM cells themselves[26]. CA3, predominantly expressed in skeletal muscle, adipose tissue, and hepatocytes, is also found to be chiefly expressed in MES-like malignant cells within GBM, as revealed by our single-cell data analysis, aligning with anticipated expression patterns. Our findings suggest that MES-like malignant cells within the GBM microenvironment may have their origins in the spinal cord or other tissues. A number of therapeutic

◀**Fig. 6**. Influence of CA3 Gene Expression on Metabolic Pathways in GBM. (**a**) A negative ES or NES implies that the central molecules of the gene set are concentrated in the low-expression group on the right, indicating that the target pathway is significantly enriched in the low-expression group; on the contrary, it suggests that these molecules are predominantly in the high-expression group on the left, signifying that the pathway is notably enriched in the high-expression group. (**b**) KEGG enrichment analysis delineates the principal pathways engaged, where red denotes significant enrichment in the high-expression group, and blue indicates enrichment in the low-expression group. (**c-f**) GSEA confirms that the HALLMARK_INTERFERON_ALPHA_RESPONSE pathway is markedly enriched in the high-expression group. **G.** For divergent gene sets, a bar chart oriented to the left denotes significant enrichment in the low-expression group; otherwise, it suggests enrichment in the high-expression group. (**h**) GSVA uncovers that various metabolic pathways, including Nitrogen metabolism, Tryptophan metabolism, and Histidine metabolism, are activated in the high CA3 expression group within GBM. (**i**) The Pearson correlation between CA3 expression z-scores and GSVA scores, as determined by 14 z-score parameters reflective of tumor states, is examined.

strategies have been devised to target the interactions between GBM tumor cells and MES-like malignant cells, albeit currently confined to the preclinical realm. Our research elucidates the advancements in GBM research by targeting CA3 and its intimate association with MES-like malignant cells, potentially heralding new avenues for the development of therapies that address the interplay between tumor cells and MES-like malignant cells in the future.

Research pertaining to hepatocellular carcinoma has identified a downregulation of CA3, which is implicated in cellular demise and the execution of programmed cell death pathways[27]. Our GSVA utilizing datasets from TCGA, CGGA, and Rembrandt has indicated a potential link between CA3 and the modulation of immune responses, as well as the metabolic handling of nutrients within GBM. Additionally, a significant positive correlation has been observed between CA3 expression levels and the proliferative and stem-like characteristics of tumors. Through WGCNA, we have pinpointed genes that exhibit a high degree of correlation with CA3 expression in GBM. Subsequent KEGG and GO enrichment analyses of these genes intimate that they are primarily engaged in the regulation of tumor cell survival and motility through pathways associated with cellular metabolism and migration. CD49B, a cell surface glycoprotein, acts as a marker for Tr1 cells, which are known to modulate immune responses within tumors, and downregulation of CD49B has been demonstrated to impede tumor progression[28]. This study has uncovered a pronounced co-expression pattern between CA3 and the functional protein CD49B in GBM, hinting that the suppression of CA3 expression may retard the advancement of GBM. EMT is a fundamental biological process that plays a pivotal role in embryonic development and tissue repair. Nonetheless, the dysregulated reactivation of EMT is associated with the acquisition of aggressive traits by tumor cells during oncogenic progression and metastasis. This encompasses the facilitation of cell migration and invasive capabilities, an increase in tumorigenic potential, and a heightened resistance to chemotherapeutic and immunotherapeutic interventions[29,30]. Our data demonstrate a significant positive correlation between CA3 and the EMT phenotype in GBM tumor cells. These observations may imply that CA3 modulates the metabolic and migratory behavior of GBM malignant cells, thereby impacting their viability. Furthermore, CA3 could potentially influence the course of GBM and its resistance to chemotherapy by engaging in the EMT program of these cells. CA3 is typically involved in various biological processes in the body, such as redox balance, pH balance, glycolysis, and apoptosis[31,32]. Previous literature has reported that metabolic genes related to glycolysis are positively correlated with EGFR in GBM, and the expression of CA3 increases with the amplification of EGFR copy number, which aligns with our findings[33]. Additionally, our study highlights the positive correlation between CA3 and TXNRD1, suggesting that CA3 may cooperate with TXNRD1 to promote tumor development. While CAs are generally considered intracellular pH regulators, mainly through the hydration/dehydration reaction of CO_2/HCO_3^- CA3 exhibits significantly lower CO_2 hydration activity compared to CA9 and CA12, indicating that CA3 does not significantly contribute to intracellular pH regulation[34–36]. Our research confirms the association of CA3 with apoptosis, proliferation, metastasis, invasion, redox balance, and glycolysis in GBM cells, implying that CA3 might interact with these biological processes to promote the occurrence and development of GBM. This suggests that CA3 is indeed a key target for DR inhibition of GBM progression, although further experimental validation is required.

Prior studies have delineated six unique malignant cellular subtypes within GBM, with the MES-like Malignant subtype demonstrating a robust association with immune cell infiltration and chronic hypoxic microenvironments. These MES-like GBM cells exhibit an elevated expression of M2-polarizing ligands relative to other cellular states, and all subtypes are implicated in the M2 polarization and immunosuppressive functions of tumor-associated macrophages/microglia (TAMs) through a repertoire of 10 ligand-receptor signaling axes, which are correlated with an unfavorable prognosis in GBM[37]. In this investigation, single-cell transcriptome analysis has unveiled the intratumoral heterogeneity of GBM, revealing its composition of diverse cellular elements, including MES-like Malignant, NPC-like Malignant, OPC-like Malignant, AC-like Malignant, and immune cells. Notably, CA3 exhibits significantly heightened expression in MES-like Malignant cells compared to other subtypes, hinting that its overexpression may serve as a valuable biomarker indicative of aggressive tumor behavior. Analyses of intercellular interactions and signal transduction pathways have further established that intense intercellular communication exists between CA3 + Malignant and other subtypes such as MES-like Malignant, OPC-like Malignant, and AC-like Malignant, influencing the genesis and progression of GBM through diverse signaling mechanisms. Consequently, this research posits that CA3 could be harnessed as a predictive biomarker for the development and metastatic spread of GBM.

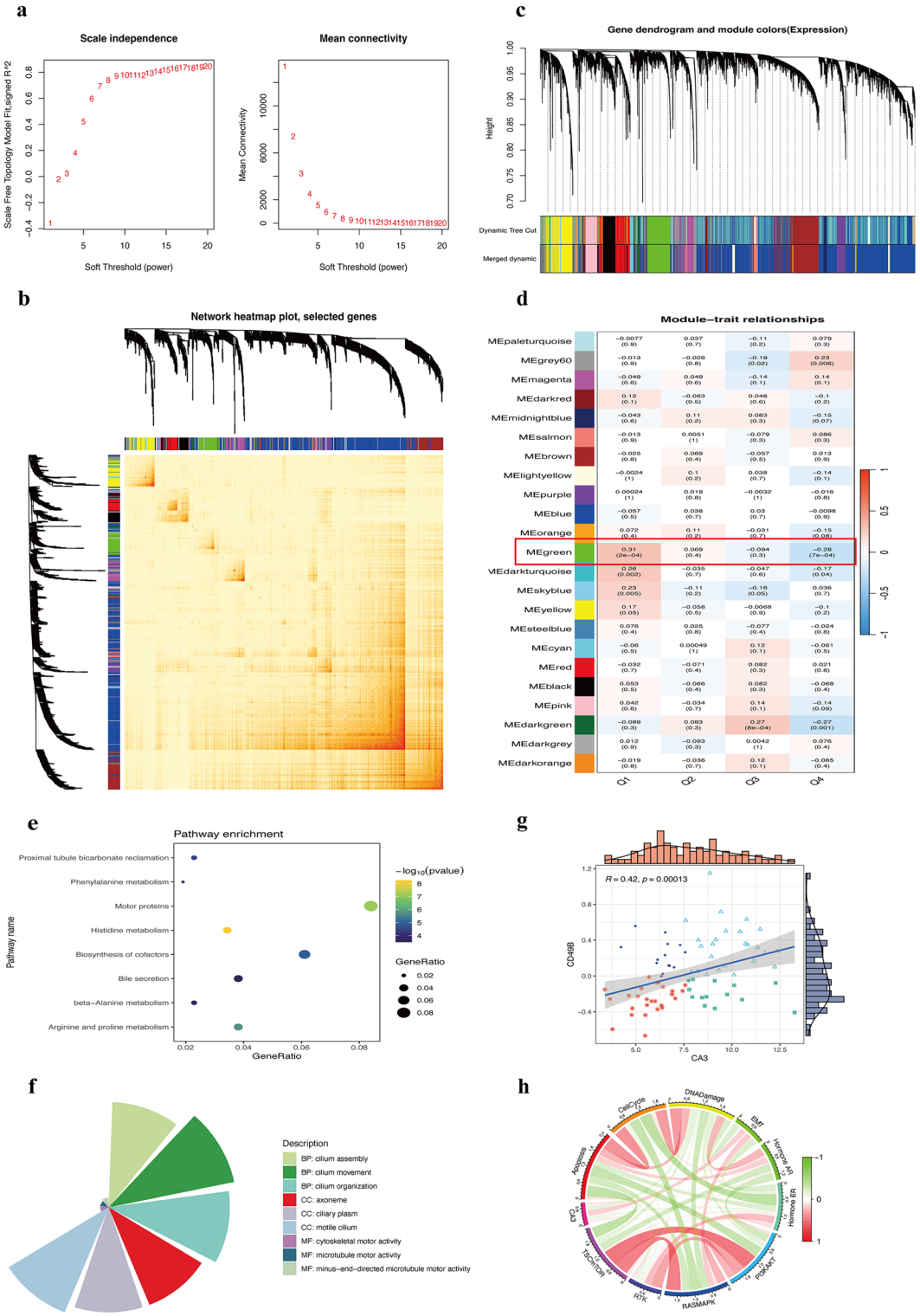


Fig. 7. WGCNA Elucidates CA3-Associated Genes, with TCPA Database Highlighting CD49B as a Significantly Positively Correlated Functional Protein of CA3, Delving into the Influence of CA3 on GBM. (**a**) Attaining the state of a scale-free network, with the optimal soft-thresholding value determined automatically. (**b**) Visualization of the weighted co-expression network, employing a heatmap to depict the gene network. (**c**) Genes are grouped into distinct modules based on their co-expression profiles, modules with low diversity are consolidated. (**d**) Quartiles of gene expression are used to categorize samples into four phenotypes: Q1, Q2, Q3, Q4 (Q1 encompasses the top 25% of samples with the highest expression levels, whereas Q4 comprises the bottom 25% with the lowest expression levels). Heatmap visualizations of the association between each module and the four phenotypes are presented, with gene modules exhibiting. (**e**, **f**) KEGG and GO analysis of the MEGreen module gene. (**g**) Correlation analysis of CA3 with functional protein of CD49B. (**h**) Correlation analysis between CA3 and multiple pathways.

The inverse relationship between DR and CA3 potentially implies that CA3 could be leveraged as an innovative therapeutic target to curb the progression of GBM. Current research underscores that AZD6482 elicits antitumor activity by suppressing cell proliferation and promoting apoptosis in human glioma cells. As such, AZD6482 could be integrated as an adjunct therapy to bolster the efficacy of GBM treatment regimens[38,39]. Our findings suggest that, relative to established CA activators and inhibitors, AZD6482 and NU.1025 exhibit favorable chemotherapeutic predictive treatment efficacy when CA3 is utilized as a therapeutic target for GBM. Nonetheless, the predictive outcomes for methazolamide, an additional CA inhibitor, presented inverse results, highlighting the need for further investigation to substantiate these observations. The pan-cancer cell line screening outcomes derived from the DepMap database are poised to further inform and enhance the design of subsequent biological assays.

In this research, employing a multi-omics approach through various databases, we have elucidated that DR can suppress the progression of GBM, and have pinpointed CA3 as a crucial gene mediating the inhibitory effects of DR. However, the precise mechanisms underlying DR's regulation of CA3 necessitate further investigation. Additionally, the role of CA3 in GBM awaits validation through direct evidence from both in vitro and in vivo studies, and the prognostic significance of CA3 in GBM treatment requires further substantiation with molecular biological and clinical data. Regarding clinical applications, while our findings suggest that CA3 possesses diagnostic potential for GBM, the absence of research into CA3's expression in blood and cerebrospinal fluid constrains its use as a non-invasive diagnostic biomarker. Moreover, the question of how to utilize CA3 expression in GBM patients to inform the clinical application of DR warrants deeper consideration and investigation.

Conclusions

To conclude, differential expression analysis in our study has shown that the suppression of GBM progression by DR is associated with the genes CA3, FZD1, G0S2, HSPA5, IFIH1, KCNA1, NR1D1, and SCRT1. Further integrated multi-omics analysis and functional studies have identified CA3, which is negatively correlated with DR, as the most critical gene and target for DR's inhibition of GBM progression. Additionally, our investigation into the mechanisms by which DR inhibits GBM development has revealed that CA3 is connected to immune evasion, invasion, and migration in GBM. Our findings suggest that DR may inhibit GBM progression through the downregulation of CA3. While our research has delved into the associations between the DR, CA3, and GBM, additional experimental validation and investigative efforts are imperative to substantiate our findings and dissect the intricate mechanisms underlying the interplay of DR, CA3, and GBM. Such advancements are anticipated to pave the way for innovative therapeutic strategies against this highly malignant disease.

Materials and methods

Expression and transcription analysis

Clinical and RNA-sequencing datasets from patients and normal controls were procured from The Cancer Genome Atlas (TCGA) database through the UCSC Xena platform (https://xena.ucsc.edu/), while normal control samples were derived from the Genotype-Tissue Expression (GTEx) project databases (https://www.gtexportal.org/home/). To address batch effects between TCGA and GTEx samples, we employed the limma package for normalization and differential expression analysis. This process entailed applying a stringent filter of $|\log \text{fold change (logFC)}| > 1$ and a P-value threshold of < 0.05 to discern differentially expressed genes (DEGs). Subsequently, the DEGs were visually represented using the dplyr, tidyverse, and heatmap packages, facilitating an intuitive interpretation of the genomic alterations associated with the disease under investigation.

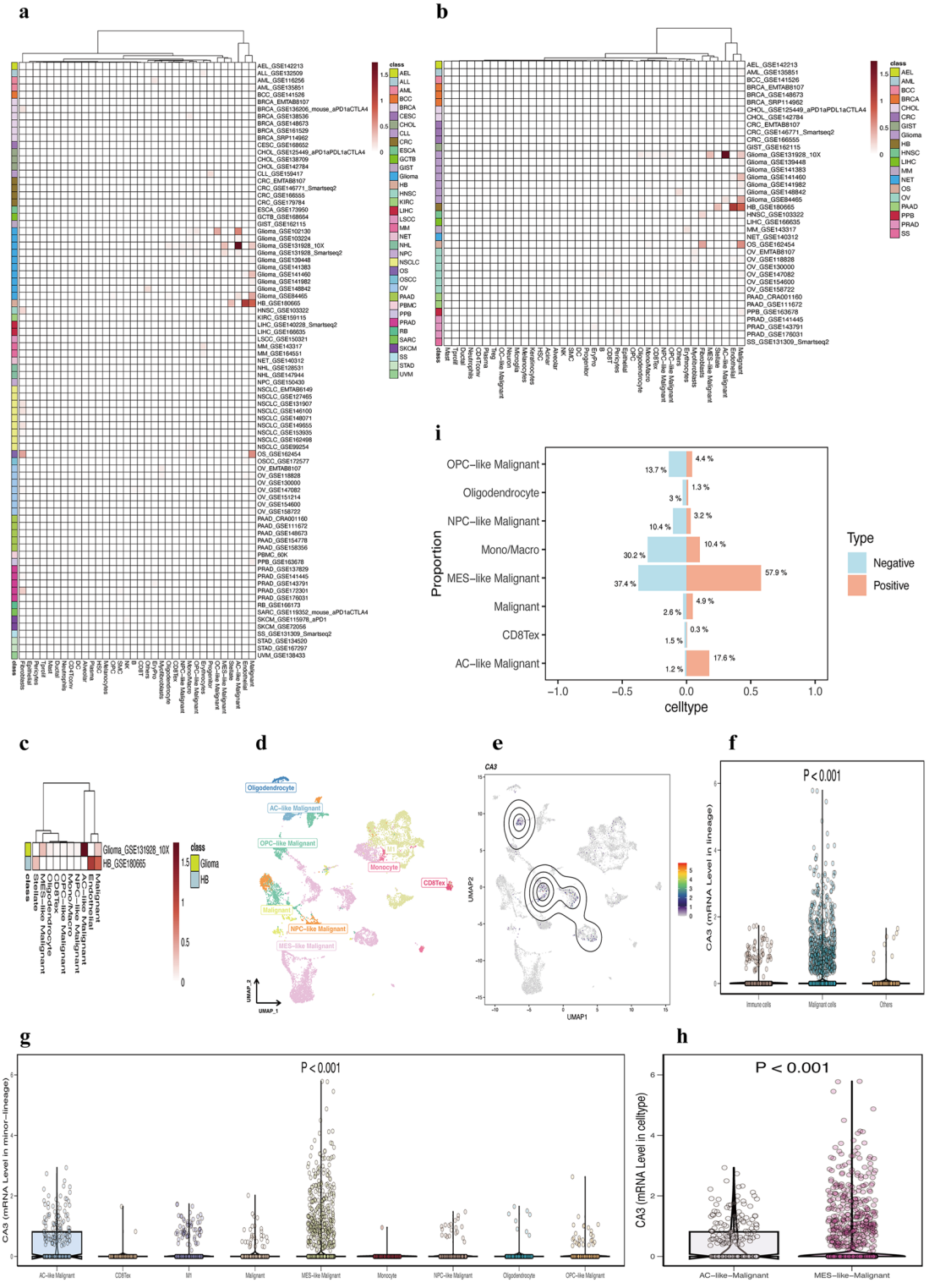
DR-Related gene acquisition and selection

A cohort of 276 genes implicated in the DR mechanism was discerned from the Dietary Restriction Gene Database (GenDR) and extant scholarly articles (Table 1.)[3]. By computationally intersecting this DR gene set with our DEGs, we have pinpointed 14 genes that are common to both datasets. The Veen package was utilized to graphically represent the intersection of these gene sets in a Veen diagram.

Gene expression data for the intersecting genes were extracted from the PanCanAtlas, leveraging the EBPlusPlusAdjustPANCAN_IlluminaHiSeq_RNASeqV2.geneExp.tsv file. This dataset, generated through the Firehose pipeline employing MapSplice for alignment and RSEM for quantification, was normalized by setting the upper quartile to a threshold of 1000. We then standardized the data across tumors by converting it into Z-Scores, calculated as, and subsequently filtered out outliers with Z-Scores exceeding 3 or falling below −3 to enhance the reliability of our subsequent analyses. To quantitatively evaluate the differential gene expression between tumor and normal tissues within the GBM dataset, we conducted Wilcoxon Rank Sum Tests. These non-parametric statistical tests were selected to compare the gene expression profiles of the two groups, considering the potential non-normal distribution of the data. This approach allows for a comprehensive assessment of the gene expression patterns associated with DR in the context of GBM, providing valuable insights into the molecular mechanisms that may underlie the therapeutic potential of DR interventions.

Survival analysis validation

By employing the gsva function within the GSVA package for single-sample Gene Set Enrichment Analysis (ssGSEA), we have conducted an in-depth analysis to ascertain the relative enrichment of gene sets within individual samples. This methodology involves the ranking of gene expression profiles across samples and the calculation of the cumulative distribution function for genes within a gene set of interest, culminating in the determination of a gene set enrichment score.To substantiate the prognostic relevance of the eight genes identified through our selection criteria, we have leveraged external validation using 475 cohorts from the



Rembrandt database. Application of a random effects model has facilitated a comprehensive assessment of the correlation between the gene set enrichment scores and patient outcomes across these cohorts, thereby reinforcing the clinical significance of our findings.

Machine learning positioning core genes

In our research, we have strategically implemented a spectrum of machine learning algorithms, encompassing Bruute, Lasso, SVM-RFE, RandomForest, and XGBoost, to discern and delineate the core genes. These methodologies were chosen for their exceptional capacity to manage high-dimensional datasets and to execute feature selection with precision. Leveraging the train function from the caret package in R, we have meticulously trained models for these algorithms, optimizing the parameters through rigorous cross-validation to bolster the models' ability to generalize to new data.

◀**Fig. 8**. Differential Expression of CA3 at the Single-Cell Level. (**a**) Single-cell expression within datasets from the TISCH database where CA3 expression is uniformly low. (**b**) Datasets from the TISCH database in which CA3 expression is nullified in malignant cells. (**c**) Single-cell expression profiles of CA3 within GBM. (**d**) The distribution of single-cell data post UMAP dimensionality reduction is depicted, where each dot represents an individual cell, positioned according to its expression profile similarity in high-dimensional space, and color-coded based on cell type or the expression levels of specific genes. (**e**) The color legend on the right of the Figure represents gene expression levels, with red indicating high expression and gray denoting low expression, each dot symbolizing a cell. (**f**) Discrepancies in CA3 expression are observed across immune, stromal, and malignant lineages. (**g**) Post-cell classification, variations in CA3 expression are noted. (**h**) Distinct expression patterns of CA3 are evident in MES-like Malignant and AC-like Malignant cells exhibiting heightened expression. (**i**) The proportion of various cell types within CA3 expression-positive and -negative groups is delineated.

To augment the models' interpretability, we have harnessed the explain function from the DALEX package, which provides a clear and comprehensive understanding of the underlying decision-making mechanisms within each model. Additionally, we have rigorously evaluated the models' predictive accuracy through the predict function and have graphically represented the models' efficacy via Receiver Operating Characteristic (ROC) curves. The variable_importance function from the DALEX package has been instrumental in quantifying the significance of each variable within the models, thereby facilitating a deeper comprehension of the genes that exert the most considerable influence on the predictive outcomes of our models. This holistic approach has been pivotal in unraveling the genetic determinants that are critical to the biological phenomena under investigation.

The expression of CA3

In this study, we thoroughly investigated the expression patterns of the CA3 gene in both normal tissues and various cancers. We commenced by extracting RNA sequencing data from the Human Protein Atlas (HPA) and GTEx databases. These data included transcript expression levels across 50 different tissues and 81 cell types. We then calculated the consensus normalized transcript per million (nTPM) value for each gene, which is defined as the maximum nTPM value from either of the data sources. For tissues with multiple sub-tissues, we adopted the highest value among all sub-tissues to represent the expression level of the tissue type. A lollipop plot was utilized to visualize these data.

Expression of CA3 in GBM

In our work, we investigated the expression differences of the CA3 gene in GBM. We used RNA-seq data from the corrected TCGA dataset, which was processed through the Firehose pipeline (MapSplice + RSEM) and normalized according to the upper quartile. We further transformed the data into unitless Z-Score values to standardize the expression levels across different samples.

To expand the normal sample size, we used the normal sample TPM expression from the GTEx database to pair with the TCGA tumor TPM expression for analysis. We also transformed these data into Z-Score values and removed outliers greater than 3.0 or less than − 3.0 to ensure the accuracy of the analysis. CA3 protein expression data comes from the corrected CPTAC database; we used mass spectrometry proteomics data and transformed it into Z-Score values. Through Wilcoxon Rank Sum Tests, we compared the statistical differences in expression levels between tumor and normal tissues in these GBM datasets.

Receiver operating characteristicanalysis

Utilizing the pROC package, we performed ROC analysis to ascertain the discriminative capacity of CA3 gene expression in delineating neoplastic from normal tissues. The ROC analysis, a benchmark for evaluating the diagnostic efficacy of binary classification models, plots an ROC curve by determining the true positive rate (sensitivity) and the false positive rate (1 - specificity). We derived the area under the curve (AUC), a pivotal parameter for gauging diagnostic accuracy, with an AUC approaching unity signifying superior diagnostic proficiency. Furthermore, we ascertained the 95% confidence interval to appraise the statistical significance of the AUC. To bolster the precision and dependability of our analysis, we augmented the cohort of normal samples, persisting in our integration of normative sample data from the GTEx database.

Survival analysis and status

Employing the survival package within the R environment, we executed Kaplan-Meier (KM) survival analysis to evaluate the prognostic significance of CA3 in GBM. The "survminer" package was engaged to ascertain the optimal cut-off thresholds for groups with high and low expression, maintaining a minimum proportion of 0.3 for each to mitigate bias. Utilizing the survfit function, we performed the log-rank test to assess the statistical significance of survival disparities between cohorts. CA3 expression was stratified into quartiles (Q1, Q2, Q3, Q4) to thoroughly evaluate its prognostic implications and to scrutinize the prognostic roles of these categories in GBM patient survival. A meta-analysis was conducted on univariate Cox survival analyses from various datasets using the inverse variance technique, with the natural logarithm of the hazard ratio (log HR) as the principal metric, to determine the association between CA3 expression levels and the survival duration of GBM patients.

Gene set enrichment analysis enrichment analysis

Initially, we divided the samples into high-expression and low-expression groups based on gene expression levels, with each group comprising 30% of the samples. We employed the limma package for differential expression

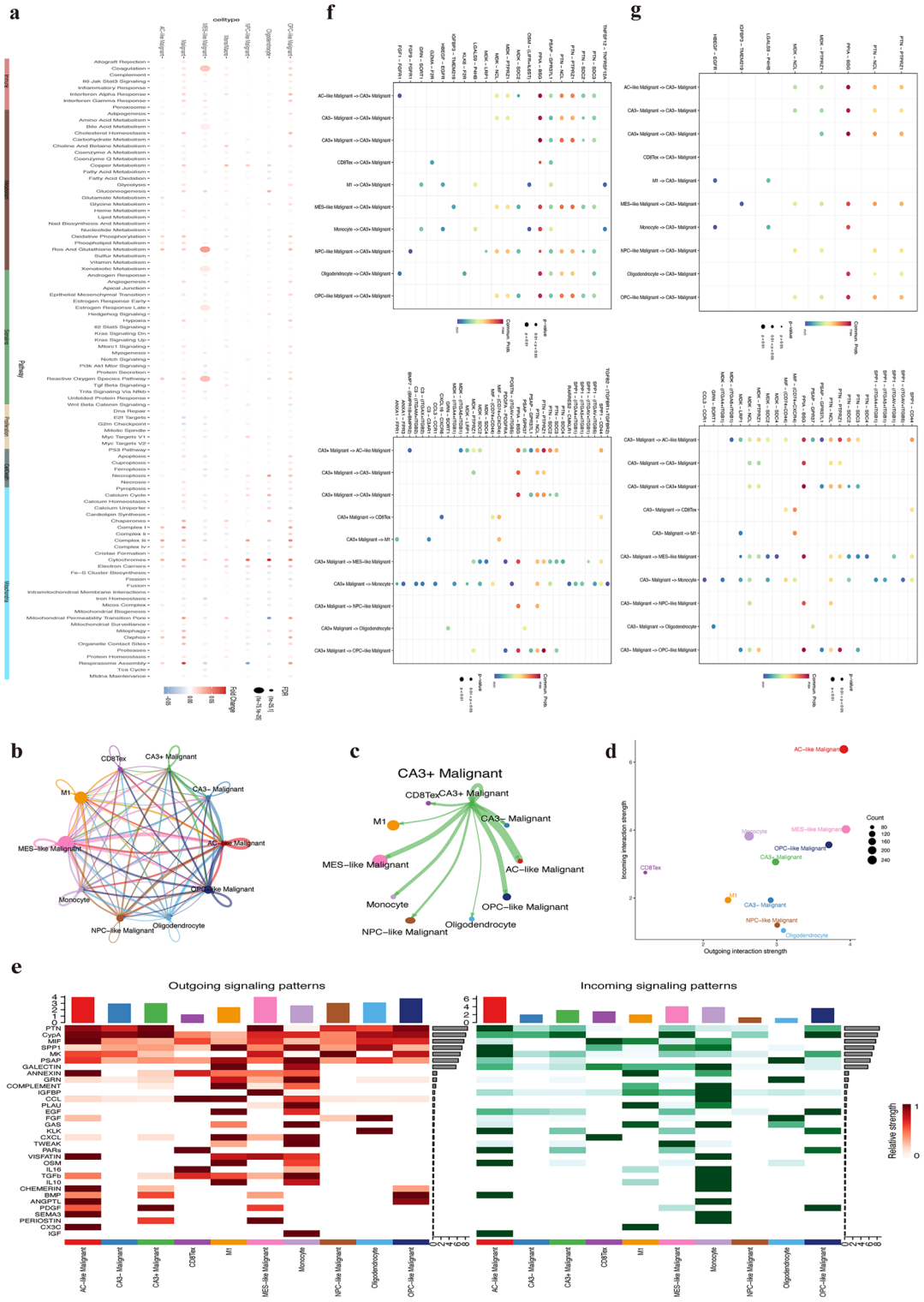


Fig. 9. Crosstalk and Signaling Pathways Among Multiple Cell Types. (**a**) Distinct pathway profiles between CA3-expressing positive and negative groups across various cell types. (**b**) Intercellular communication networks within GBM. (**c**) Interactions between CA3-positive Malignant cells and a spectrum of other cell types. (**d**) Intensity of signaling interactions among different cell types. (**e**) Cellular populations characterized as signal originators (Outgoing) or recipients (Incoming). (**f**, **g**) Visualization of specific ligand-receptor interactions at the cellular level.

analysis, calculated the log2 FC of each gene, and sorted the genes according to these values. Subsequently, we utilized the fgsea function in the fgsea package to conduct gene set enrichment analysis based on the hallmark gene set, computed the gene set enrichment score (ES) value, and performed significance tests and multiple hypothesis tests. The clusterProfiler package was adopted to perform enrichment analysis on the KEGG gene set, calculate the ES value, and conduct significance tests and multiple hypothesis tests. We regarded gene set enrichment results with a p value less than 0.05 and an adjust p value less than 0.25 as significant and visualized them. To further investigate the enrichment of CA3 in HALLMARK_INTERFERON_ALPHA_RESPONSE, we divided the samples into high-expression and low-expression groups according to the median value of genes and used the limma package for differential analysis. The clusterProfiler package was used for gene set enrichment analysis based on a custom gene set, calculated the ES value, and conducted significance tests and multiple hypothesis tests. Additionally, we verified the enrichment results of CA3 through an external dataset.

Gene set variation analysis

The CancerSEA website is employed to process single-cell data from datasets such as HCMDB, Cyclebase, and StemMapper for functional analysis and to redefine a total of 14 functional states[40]. We utilized the z-score algorithm in the R package GSVA to compute these functional state gene sets and converted the value of each gene set into a z-score to reflect the activity of a given pathway[41]. Additionally, we employed Pearson correlation analysis to calculate the statistical correlation between genes and the z-score of each gene set. To further investigate the relationship between gene expression and metabolic pathway activity, we used the gsva parameter algorithm in the R package GSVA to score and calculate the metabolic gene sets in 73 KEGG databases. We divided the samples into high-expression and low-expression groups based on gene expression levels and used the limma package to compare the differences in metabolic GSVA scores between these two groups.

Weighted correlation network analysis (WGCNA)

WGCNA is a systems biology approach that clusters genes into co-expression modules by constructing a correlation-based weighted gene co-expression network. We optimize the topological structure of the network by adjusting the soft threshold parameter for network construction to ensure its biological significance. Initially, samples are divided into four phenotypes based on the quartiles of gene expression levels: Q1, Q2, Q3, and Q4. Q1 represents the 25% of samples with the highest expression, while Q4 represents the 25% of samples with the lowest expression. After network construction is completed, we perform module identification and identify the module most closely associated with CA3 through correlation analysis between module characteristic genes and CA3 expression levels.

Chemotherapy

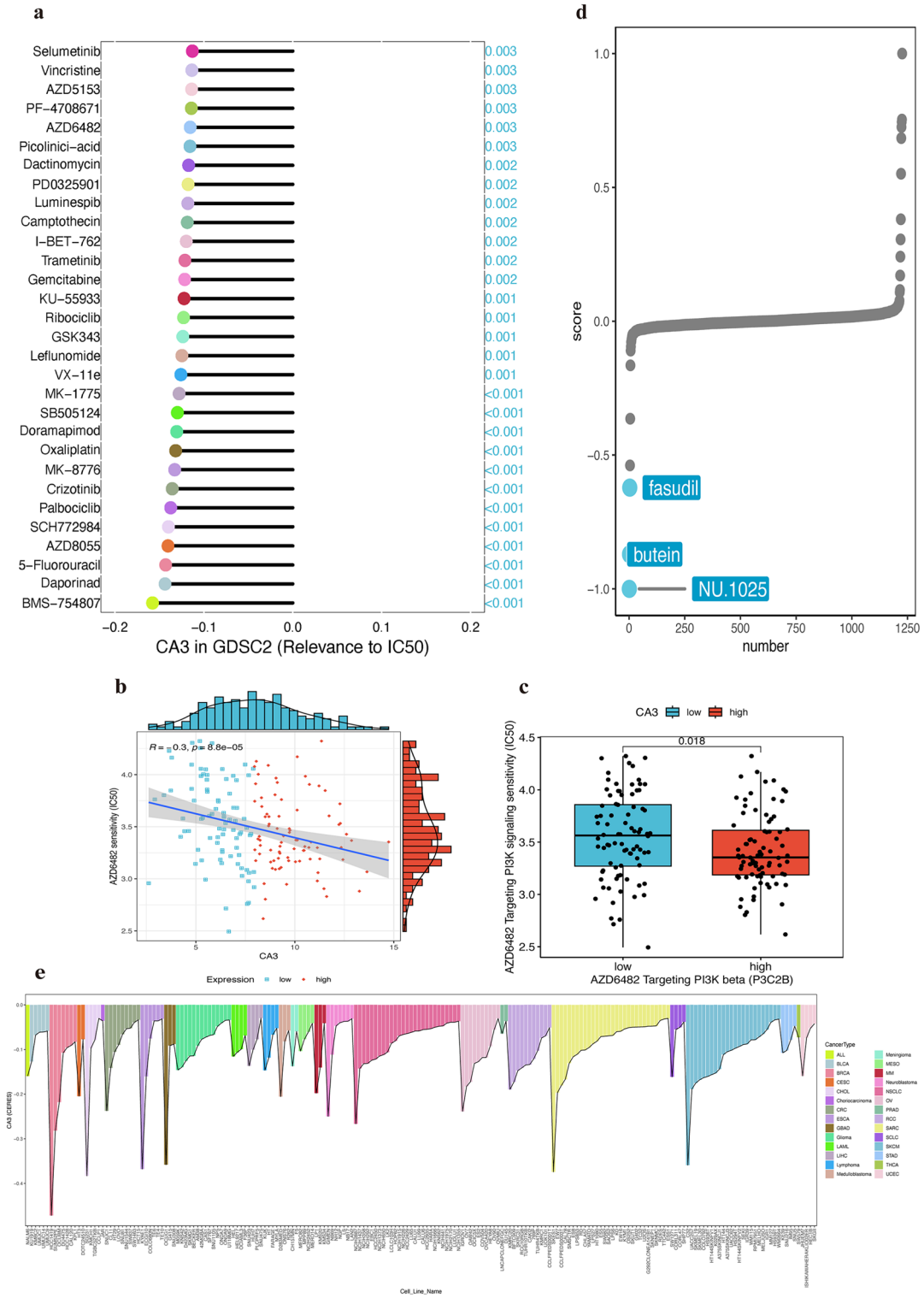
The Genomics of Drug Sensitivity in Cancer (GDSC) dataset is derived from the R package oncoPredict, which has meticulously organized rdata files for these databases (https://osf.io/c6tfx/). We commenced our analysis by employing Spearman's rank correlation to correlate gene expression data with the IC50 values of quantitative antagonists within the GDSC database (https://www.cancerrxgene.org/), which facilitated the identification of chemotherapeutic agents sensitive to CA3. Following this, we harnessed the pRRophetic R package to prognosticate gene expression levels and anticipate clinical responses to chemotherapy. Within the GDSC1 database, we estimated the IC50 values for GBM patients across various datasets in response to AZD6482, thereby predicting chemotherapy responses. Furthermore, within the PRISM database (https://www.theprismlab.org), we projected the AUC values under the dose-response curve for multiple GBM datasets, encompassing CA activators (D-phenylalanine, histamine), CA inhibitors (acetazolamide, topiramate, methazolamide), and NU.1025, to mirror the responses to chemotherapy treatments.

Targeted therapy

In an effort to explore potential treatment regimens capable of counteracting gene-mediated tumor-promoting effects, we resorted to Connectivity Map (cMAP) analysis. cMAP, a public database, identifies small molecules that might influence specific biological processes by comparing gene expression patterns. We initially constructed a gene signature consisting of 150 most significantly upregulated and 150 most significantly downregulated genes. This was accomplished by comparing patients with high and low gene expression in tumors. The CMAP_gene_signatures.RData file, containing 1288 compound-related features, offers us a rich compound library for further analysis. Subsequently, the optimal feature matching method XSum was employed to compare gene-related features with cMAP gene features, yielding similarity scores for 1288 compounds[42,43]. Compounds with lower scores are potentially capable of inhibiting gene-mediated cancer-promoting effects.

Common pathway activity differences

In the easier package, scores for 14 pathways are computed using PROGENy. These pathways comprise androgen, EGFR, estrogen, hypoxia, JAK-STAT, MAPK, NF-κB, p53, PI3 K, TGF-β, TNF-α, Trail, VEGF, and WNT. Pathway-specific signatures are obtained by examining changes in gene expression when pathways are perturbed. A linear regression model is employed to fit genes influenced by pathway perturbations, and then pathway-specific signatures and gene expression data are utilized to infer pathway signaling activity. The easier package directly calculates pathway scores based on version 1.10.0 of the PROGENy R software package. Since these scores are a linear transformation of gene expression data, the easier package removes 448 genes used to calculate immune response proxies that have an extremely high correlation with the original pathway activity (average pan-cancer Pearson correlation = 0.99, $p < 10\text{^}{-}16$). After obtaining the scores of 14 pathways, the data is transformed into unitless Z-Score values via the formula $(x - \mu)/\sigma$ to standardize the scores.



Functional proteomics

First, obtain the protein expression data of reverse phase protein array (RPPA) from the The Cancer Proteome Atlas (TCPA) database and preprocess the data, encompassing standardization and removal of outliers. Subsequently, employ multiple correlation analysis approaches, such as spearman correlation analysis and feature selection methods in machine learning, to batch-analyze the correlation between the target gene and each protein. For instances where the absolute value of the correlation coefficient is greater than 0.3 and the p-value is less than 0.05, draw scatter plots based on the magnitude of the correlation coefficient and the significance of the p-value for classification. Next, in accordance with published research results, elaborate on the calculation method of pathway activity score and conduct verification[44]. Finally, utilize the cor.test function and other appropriate methods to compute the correlation and p-value between the target gene and the pathway activity score, and conduct comprehensive analysis and discussion of the results.

◀**Fig. 10**. Investigating CA3 as a Therapeutic Agent for GBM. (**a**) The x-axis denotes the Spearman correlation coefficient relating drug IC50 or AUC values to gene expression levels, while the y-axis highlights the top 30 drugs with the most significant p-values. Drugs are distinctly colored, with red signifying a positive correlation and blue a negative correlation with the gene. The correlation coefficient's magnitude is reflected in the length of the lollipop stick. (**b**) Each dot corresponds to a sample, with axes indicating CA3 expression levels and IC50 values for AZD6482. Samples are bifurcated into high (red) and low (blue) expression groups based on the gene's median expression. (**c**) TCGA-GBM dataset analysis delineating IC50 disparities between high and low CA3-expressing groups. The box's upper and lower bounds represent the interquartile range, with the median depicted by the central line. Pathways targeted by the drug are annotated on the left, and specific targets are noted at the bottom. (**d**) Each plotted point corresponds to a distinct compound, with the y-axis illustrating the similarity scores for 1288 compounds derived from the optimal feature matching method XSum, which compares gene-related characteristics to cMAP gene profiles. Compounds that achieve lower scores are potentially capable of suppressing oncogenic activities mediated by genes. (**e**) A visual representation of the top 200 cancer cell lines from the DepMap database, ranked by their CERES as determined by whole-gene CRISPR-Cas9 screens. The y-axis denotes the CERES score, the x-axis signifies various cell lines, and the color coding differentiates between distinct cancer types.

CRISPR-Cas9 screening of CA3-sensitive cell lines

Download genome-wide CRISPR screening data from the DepMap website and employ the Correcting Essential gene Ranking Evaluation based on Single-cell Expression Patterns(CERES) algorithm with stringent quality control and standardized processing to calculate the dependency scores of approximately 17,000 candidate genes[45]. Conduct repeated experiments and validate with independent datasets for the conclusion that negative scores imply cell growth inhibition and/or death after gene knockout. Scores of 0 and − 1 respectively represent the median effect of non-essential genes and common core essential genes. Display the top 200 negatively scored cell lines via various visualization means such as heatmaps, 3D charts and traditional bar charts. Simultaneously, provide detailed method descriptions and code examples so that other researchers can reproduce and expand the study.

Single-cell data analysis

Acquire gene expression profiles at pan-cancer single-cell resolution from the Tumor Immune Single-cell Hub (TISCH) database and employ the pheatmap package to generate heatmaps visualizing the pan-cancer single-cell expression landscape of genes. Utilizing Euclidean distance as the metric and applying Ward's minimum variance method for hierarchical clustering facilitates the recognition of patterns and trends within the data, assisting in the identification of conserved gene expression signatures. In this study, Uniform Manifold Approximation and Projection (UMAP) dimensionality reduction technique is implemented to render high-dimensional data into two-dimensional heatmaps. Additionally, a visual analysis of the expression data for the CA3 gene is conducted, which aids in comprehending gene expression patterns more intuitively and discerning potential biological disparities.

Utilizing the Kruskal-Walli's rank sum test, we evaluated the expression variances of specific genes within diverse cellular contexts. As a non-parametric statistical approach, the Kruskal-Walli's test is adept at handling scenarios where datasets deviate from normality, robustly discerning significant disparities among several independent sample cohorts.

Positive-related	Negative-related
ADH1C	SLC6A6
PER2	CA3
POR	DHCR7
INMT	ARNTL
DBP	ZFP64
NAT8	SREBF1
EHHADH	GCK
CYP2J2	COL15A1
ABCG5	G0S2
FAM107A	INSIG1
KLF15	C9
SDS	PHLDA1
FKBP5	HSPA5
ZBTB16	DPP9
ANGPTL4	ALAS2
USP2	TMEM132D
COBLL1	IRF7
FMO3	FABP5
CYP7A1	TNFSF10
ABLIM3	ACLY
NR1I3	SCLY
HERPUD1	CASC5
CTGF	SERPINH1
SLC37A4	IFIH1
TENC1	C11orf49
WEE1	TTLL12
KLF9	AQP8
PPARA	CLDN1
TP53I13	NR1D1
IRS2	GHR
FAM195A	R3HDM2
ACOT4	HIPK2
NTF3	RSC1A1
TMEM218	CYP2F1
ALDH1A1	HSD3B1
PIM3	EXTL1
AQP6	SC5D
DECR2	G6PD
CRY1	SCRT1
TSC22D3	PTPRJ
CBR1	PSMB8
RGS16	SLC10A2
HACL1	ACTG1
SULT1C2	NTN3
GYS2	STAC3
CYP2E1	MMP15
PLIN5	GTF2IRD1
CPT1 A	PHF19
IGFBP2	INHBE
ARRDC2	COL3A1
ST3GAL5	CDC42EP2
SLC25A25	DNASE1L2
LPIN1	LITAF
GPR146	PDIA3
ADCY1	LY6E
IFRD1	HSPB7
Continued	

Positive-related	Negative-related
MAT1A	
ACOT12	
NFKBIA	
EPB41	
HSD17B2	
SUN2	
MGP	
HCAR2	
RDH16	
DUSP1	
KLF10	
RHBDD2	
DECR1	
CD163	
PLCXD3	
BNIP3	
FZD1	
PER1	
ENPEP	
SALL1	
SLC25A42	
ZNF354A	
PLA2G12A	
MAP3K6	
RBP7	
RHOBTB1	
CRYM	
PLIN4	
SMOC1	
TOB1	
WWP2	
WWP1	
ITCH	
MTOR	
PIK3C3	
HSF2	
ERN2	
KCNA2	
KCNA1	
KCNA3	
KCNA4	
KCNA6	
KCNA5	
KCNA10	
KCNA7	
PIK3CA	
PIK3CB	
PIK3CD	
PIK3CG	
CCKAR	
CCKBR	
IGF1R	
INSR	
INSRR	
COQ7	
UBE2L3	
Continued	

Positive-related	Negative-related
RHEB	
RHEBL1	
EP300	
CREBBP	
NFE2L1	
NFE2L3	
NFE2L2	
NFE2	
BECN1	
FAAH	
CRYAB	
CRYAA	
HSPB6	
HSPB2.1	
SIRT1	
ATG7	
SMEK1	
SMEK2	
ULK2	
ULK1	
MAT2 A	
GATA3	
GATA2	
PRKAA2	
PRKAA1	
TXN	
TXNDC8	
RAB10	
CBS	
EIF4EBP2	
EIF4EBP1	
COX5B	
HADHA	
HDAC2	
HDAC1	
SLC13 A2	
SLC13 A5	
SLC13 A1	
SLC13 A3	
TSC2	
ETFB	
NDUFB5	
IRS1	
IRS4	
UCP2	
SC5DL	
METTL21C	
ELOVL4	
ELOVL1	
VPS8	
ADH5	
RPL31	
HIST2H4B	
HIST2H4A	
HIST1H4E	
HIST1H4I	
Continued	

Positive-related	Negative-related
HIST1H4L	
HIST1H4D	
HIST1H4F	
HIST1H4B	
HIST1H4K	
HIST1H4H	
HIST4H4	
HIST1H4C	
HIST1H4J	
HK2	
HKDC1	
HHAT	
GCLC	
SRXN1	
MDH2	
SP4	
SP2	
SP5	
MAST1	
MAST2	
MAST4	
MAST3	
CYC1	
RAB5A	
RAB5C	
RAB5B	
SPHK2	
GPD2	
FOXC1	
FOXC2	
FOXD2	
FOXD3	
FOXA2	
FOXE1	
FOXA1	
FOXD4L1	
FOXI2	
RRAGA	
RRAGB	
PRDX2	
PRDX1	
PRDX4	
RAB7A	
CHMP6	
PPAT	
DLAT	
SIRT3	
Continued	

Positive-related	Negative-related
SIRT2	
DPP4	
FAP	
DPP10	
FAM120B	

Table 1.. DR-related Gene.

Data availability

The datasets generated during and/or analysed during the current study are available in the TCGA repository, [https://www.cancer.gov/ccg/research/genome-sequencing/tcga], UCSC Xena platform [https://xena.ucsc.edu/], GTEx repository, [https://www.gtexportal.org/home], CGGA, [http://www.cgga.org.cn], and GEO repository [https://www.ncbi.nlm.nih.gov/geo/].

Received: 10 January 2025; Accepted: 9 May 2025
Published online: 28 May 2025

References

1. Green, C. L., Lamming, D. W. & Fontana, L. Molecular mechanisms of dietary restriction promoting health and longevity. *Nat. Rev. Mol. Cell. Biol.* **23**, 56–73. https://doi.org/10.1038/s41580-021-00411-4 (2022).
2. Ibrahim, E. M., Al-Foheidi, M. H. & Al-Mansour, M. M. Energy and caloric restriction, and fasting and cancer: a narrative review. *Support Care Cancer*. **29**, 2299–2304. https://doi.org/10.1007/s00520-020-05879-y (2021).
3. Song, X. et al. An integrative pan-cancer analysis of the molecular characteristics of dietary restriction in tumour microenvironment. *EBioMedicine* **102**, 105078. https://doi.org/10.1016/j.ebiom.2024.105078 (2024).
4. Hursting, S. D., Dunlap, S. M., Ford, N. A., Hursting, M. J. & Lashinger, L. M. Calorie restriction and cancer prevention: a mechanistic perspective. *Cancer Metab.* **1**, 10. https://doi.org/10.1186/2049-3002-1-10 (2013).
5. Longo, V. D. & Mattson, M. P. Fasting: molecular mechanisms and clinical applications. *Cell. Metab.* **19**, 181–192. https://doi.org/10.1016/j.cmet.2013.12.008 (2014).
6. Berrigan, D. et al. Phenotypic effects of calorie restriction and insulin-like growth factor-1 treatment on body composition and bone mineral density of C57BL/6 mice: implications for cancer prevention. *Vivo* **19**, 667–674 (2005).
7. Bowers, L. W., Rossi, E. L., O'Flanagan, C. H., deGraffenried, L. A. & Hursting, S. D. The role of the Insulin/IGF system in cancer: lessons learned from clinical trials and the energy Balance-Cancer link. *Front. Endocrinol. (Lausanne)*. **6**, 77. https://doi.org/10.3389/fendo.2015.00077 (2015).
8. Mukherjee, P., Abate, L. E. & Seyfried, T. N. Antiangiogenic and proapoptotic effects of dietary restriction on experimental mouse and human brain tumors. *Clin. Cancer Res.* **10**, 5622–5629. https://doi.org/10.1158/1078-0432.CCR-04-0308 (2004).
9. Grochans, S. et al. Epidemiology of glioblastoma Multiforme-Literature review. *Cancers (Basel)*. **14** https://doi.org/10.3390/cancers14102412 (2022).
10. Lin, H. et al. Understanding the immunosuppressive microenvironment of glioma: mechanistic insights and clinical perspectives. *J. Hematol. Oncol.* **17** https://doi.org/10.1186/s13045-024-01544-7 (2024).
11. Young, J. S., Al-Adli, N., Scotford, K., Cha, S. & Berger, M. S. Pseudoprogression versus true progression in glioblastoma: what neurosurgeons need to know. *J. Neurosurg.* **139**, 748–759. https://doi.org/10.3171/2022.12.JNS222173 (2023).
12. Kanehisa, M., Furumichi, M., Sato, Y., Matsuura, Y. & Ishiguro-Watanabe, M. KEGG: biological systems database as a model of the real world. *Nucleic Acids Res.* **53**, D672–D677. https://doi.org/10.1093/nar/gkae909 (2025).
13. Kanehisa, M. Toward Understanding the origin and evolution of cellular organisms. *Protein Sci.* **28**, 1947–1951. https://doi.org/10.1002/pro.3715 (2019).
14. Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30. https://doi.org/10.1093/nar/28.1.27 (2000).
15. Sigismund, S., Avanzato, D. & Lanzetti, L. Emerging functions of the EGFR in cancer. *Mol. Oncol.* **12**, 3–20. https://doi.org/10.1002/1878-0261.12155 (2018).
16. Wang, H. et al. Clinical roles of EGFR amplification in diffuse gliomas: a real-world study using the 2021 WHO classification of CNS tumors. *Front. Neurosci.* **18**, 1308627. https://doi.org/10.3389/fnins.2024.1308627 (2024).
17. Karpel-Massler, G. et al. Combined Inhibition of Bcl-2/Bcl-xL and Usp9X/Bag3 overcomes apoptotic resistance in glioblastoma in vitro and in vivo. *Oncotarget* **6**, 14507–14521. https://doi.org/10.18632/oncotarget.3993 (2015).
18. Liu, Y. et al. Pan-cancer analysis on the role of PIK3R1 and PIK3R2 in human tumors. *Sci. Rep.* **12**, 5924. https://doi.org/10.1038/s41598-022-09889-0 (2022).
19. Shi, C. et al. Carbonic anhydrase III protects osteocytes from oxidative stress. *FASEB J.* **32**, 440–452. https://doi.org/10.1096/fj.201700485RR (2018).
20. Silagi, E. S., Batista, P., Shapiro, I. M. & Risbud, M. V. Expression of carbonic anhydrase III, a nucleus pulposus phenotypic marker, is Hypoxia-responsive and confers protection from oxidative Stress-induced cell death. *Sci. Rep.* **8**, 4856. https://doi.org/10.1038/s41598-018-23196-7 (2018).
21. Nguyen, P., Awwad, R. T., Smart, D. D., Spitz, D. R. & Gius, D. Thioredoxin reductase as a novel molecular target for cancer therapy. *Cancer Lett.* **236**, 164–174. https://doi.org/10.1016/j.canlet.2005.04.028 (2006).
22. Ji, S. et al. Cellular rejuvenation: molecular mechanisms and potential therapeutic interventions for diseases. *Signal. Transduct. Target. Ther.* **8**, 116. https://doi.org/10.1038/s41392-023-01343-5 (2023).
23. Ning, Y. et al. YAP1 synergize with YY1 transcriptional co-repress DUSP1 to induce osimertinib resistant by activating the EGFR/MAPK pathway and abrogating autophagy in non-small cell lung cancer. *Int. J. Biol. Sci.* **19**, 2458–2474. https://doi.org/10.7150/ijbs.79965 (2023).
24. Nimmakayala, R. K. et al. PAF1 cooperates with YAP1 in metaplastic ducts to promote pancreatic cancer. *Cell. Death Dis.* **13**, 839. https://doi.org/10.1038/s41419-022-05258-x (2022).
25. Reinert, T. et al. Comprehensive genome methylation analysis in bladder cancer: identification and validation of novel methylated genes and application of these as urinary tumor markers. *Clin. Cancer Res.* **17**, 5582–5592. https://doi.org/10.1158/1078-0432.CCR-10-2659 (2011).

26. Kang, Q. M., Wang, J., Chen, S. M., Song, S. R. & Yu, S. C. Glioma-associated mesenchymal stem cells. *Brain* **147**, 755–765. https://doi.org/10.1093/brain/awad360 (2024).
27. Kuo, W. H. et al. The differential expression of cytosolic carbonic anhydrase in human hepatocellular carcinoma. *Life Sci.* **73**, 2211–2223. https://doi.org/10.1016/s0024-3205(03)00597-6 (2003).
28. Contreras-Kallens, P. et al. CD49b targeting inhibits tumor growth and boosts Anti-tumor immunity. *Front. Oncol.* **12**, 928498. https://doi.org/10.3389/fonc.2022.928498 (2022).
29. Huang, Y., Hong, W. & Wei, X. The molecular mechanisms and therapeutic strategies of EMT in tumor progression and metastasis. *J. Hematol. Oncol.* **15**, 129. https://doi.org/10.1186/s13045-022-01347-8 (2022).
30. Debaugnies, M. et al. RHOJ controls EMT-associated resistance to chemotherapy. *Nature* **616**, 168–175. https://doi.org/10.1038/s41586-023-05838-7 (2023).
31. Armignacco, R. et al. The adipose stem cell as a novel metabolic actor in adrenocortical carcinoma progression: evidence from an in vitro tumor microenvironment crosstalk model. *Cancers (Basel)*. **11** https://doi.org/10.3390/cancers11121931 (2019).
32. Cantini, G. et al. The role of metabolic changes in shaping the fate of Cancer-Associated adipose stem cells. *Front. Cell. Dev. Biol.* **8**, 332. https://doi.org/10.3389/fcell.2020.00332 (2020).
33. Beckner, M. E., Pollack, I. F., Nordberg, M. L. & Hamilton, R. L. Glioblastomas with copy number gains in EGFR and RNF139 show increased expressions of carbonic anhydrase genes transformed by ENO1. *BBA Clin.* **5**, 1–15. https://doi.org/10.1016/j.bbacli.2015.11.001 (2016).
34. Yu, Y., Poulsen, S. A., Di Trapani, G. & Tonissen, K. F. Exploring the redox and pH dimension of carbonic anhydrases in cancer: A focus on carbonic anhydrase 3. *Antioxid. Redox Signal.* **41**, 957–975. https://doi.org/10.1089/ars.2024.0693 (2024).
35. Supuran, C. T. & Scozzafava, A. Carbonic anhydrases as targets for medicinal chemistry. *Bioorg. Med. Chem.* **15**, 4336–4350. https://doi.org/10.1016/j.bmc.2007.04.020 (2007).
36. Duda, D. M. et al. Human carbonic anhydrase III: structural and kinetic study of catalysis and proton transfer. *Biochemistry* **44**, 10046–10053. https://doi.org/10.1021/bi050610h (2005).
37. Xiao, Y. et al. Single-Cell transcriptomics revealed Subtype-Specific tumor immune microenvironments in human glioblastomas. *Front. Immunol.* **13**, 914236. https://doi.org/10.3389/fimmu.2022.914236 (2022).
38. Xu, P. F. et al. PI3Kbeta inhibitor AZD6482 exerts antiproliferative activity and induces apoptosis in human glioblastoma cells. *Oncol. Rep.* **41**, 125–132. https://doi.org/10.3892/or.2018.6845 (2019).
39. Zhao, H. F. et al. Synergism between the phosphatidylinositol 3-kinase p110beta isoform inhibitor AZD6482 and the mixed lineage kinase 3 inhibitor URMC-099 on the Blockade of glioblastoma cell motility and focal adhesion formation. *Cancer Cell. Int.* **21**, 24. https://doi.org/10.1186/s12935-020-01728-4 (2021).
40. Yuan, H. et al. CancerSEA: a cancer single-cell state atlas. *Nucleic Acids Res.* **47**, D900–D908. https://doi.org/10.1093/nar/gky939 (2019).
41. Lee, E., Chuang, H. Y., Kim, J. W., Ideker, T. & Lee, D. Inferring pathway activity toward precise disease classification. *PLoS Comput. Biol.* **4**, e1000217. https://doi.org/10.1371/journal.pcbi.1000217 (2008).
42. Malta, T. M. et al. Machine Learning Identifies Stemness Features Associated with Oncogenic Dedifferentiation. *Cell* 173, 338–354 e315, (2018). https://doi.org/10.1016/j.cell.2018.03.034
43. Yang, C. et al. A survey of optimal strategy for signature-based drug repositioning and an application to liver cancer. *Elife* **11** https://doi.org/10.7554/eLife.71880 (2022).
44. Liu, C. J. et al. GSCA: an integrated platform for gene set cancer analysis at genomic, Pharmacogenomic and Immunogenomic levels. *Brief. Bioinform.* **24** https://doi.org/10.1093/bib/bbac558 (2023).
45. Meyers, R. M. et al. Computational correction of copy number effect improves specificity of CRISPR-Cas9 essentiality screens in cancer cells. *Nat. Genet.* **49**, 1779–1784. https://doi.org/10.1038/ng.3984 (2017).

Acknowledgements
Not applicable.

Author contributions
JXM， DX, and JZ designed and developed the experiments. JXM and DX prepared the draft of the manuscript. MG, YG, and GHZ participated in all the experiments. JXM analyzed the data and drafted the manuscript. JZ supervised all research and revised the manuscript. ZBC has supplemented the revised manuscript with additional data and made key modifications. All authors have read and approved the final manuscript.

Funding
This work supported by Natural Science Foundation of Gansu Province (Project No: 23 JRRA532), Lanzhou Municipal Bureau of Science and Technology Science and Technology Program Projects (Project No: 2023-ZD-167) for the funding support, Natural Science Foundation of Gansu Province (Project No: 21 JR7RA008), The 940 th Hospital of Joint Logistics Support force of Chinese People's Liberation Army G Program (Project No: 2024-G3-3).

Declarations

Competing interests
The authors declare no competing interests.

Consent for publication
Not applicable.

Ethics approval and consent to participate
Not applicable.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-01986-0.

Correspondence and requests for materials should be addressed to J.Z.