ICELL JOURNAL | RESEARCH ARTICLE
Vol.1No.1(1)

Pan-Cancer Analysis for Identification of Tumor Antigens and Immune Subtypes in mRNA Vaccine Development

DIO:https://dx.doi.org/10.71373/FEHU5094
Submitted 1 October 2024
Accepted 26 October 2024
Published 1 November 2024

Yao Lin1,2#, Xinyue Xu2#, Ziyi Liu2#, Hubin Chen2, Yi Bao2, Jun Meng2, Zheng Li2, Sanqi An2,3, Yueqi Li2,3*, Xiaopeng Hu2*, Jiemei Chu2*

Purpose: Immunotherapy using vaccines proved effective for a variety of cancers, however, so far there is no vaccine against many types of cancers. The aim of this study was to identify potential tumor antigens for the development of antiall cancer mRNA vaccines and to map the immune landscape of pan-cancer to select suitable patients for vaccination.Methods: We collected normalized gene expression data and clinical data from the Chinese Glioma Genome Atlas (CGGA) and The Cancer Genome Atlas (TCGA). Gene Expression Profiling Interactive Analysis (GEPIA) version 2 was used to analyze the differential expression of cancer genes and to calculate the prognostic index of each selected antigen. We used the cBio Cancer Genomics Portal (cBioPortal) to visualize the gene alterations of potential tumor antigens in the TCGA data and the Tumor Immune Estimation Resource (TIMER) to analyze and visualize the association between immune infiltrates and expression level of the identified potent antigens. Consistency matrix analysis was used to identify corresponding gene modules and immune subtypes. Univariate and multivariate Cox regression as well as log-rank tests were used to assess the prognostic distinctions of the GBM immune subtypes. We used the reduce Dimension function of the Monocle package to visualize the distribution of immune subtypes across TCGA and CGGA patients.Results: Different cancer have different potential RNA Antigens. In GBM, seven tumor antigens (DIRAS3, EGFLAM, FUCA1, LILRB2, MPZL2, OSMR, and STC1) were identified as promising candidates for developing anti-GBM mRNA vaccines. Based on the immune gene expression profiles, GBM was classified into four immune subtypes (IS1–IS4). IS1 and IS4 were immunologically "hot," which means they are more likely to benefit from vaccines. In addition, six immune hub genes (CLCF1, CD54, SOCS3, PLAUR, LIF, and BCL3) were identified as potential biomarkers for GBM vaccine effectiveness.Conclusions: Different cancer types should have different potential RNA vaccine. For instance, DIRAS3, EGFLAM, FUCA1, LILRB2, MPZL2, OSMR and STC1 were identified to be related with poor survival and antigen presenting cell (APC) infiltration in GBM, and they are promising targets for mRNA vaccine development. GBM patients could be divided in to four immune subtypes based on the clustering of immune-related gene expression, and those in the IS1 and IS4 groups were candidates for vaccination.

Background

The concept of pan-cancer RNA vaccines aims to develop a universal cancer vaccine that can target multiple cancer types by exploiting common tumor-associated antigens (TAAs)shared among different tumors. This approach is based on the recognition that many cancers, despite arising from different tissues, often express similar sets of TAAs due to common oncogenic pathways and mutations. By identifying these shared antigens, pan-cancer RNA vaccines could potentially provide broad-spectrum protection and treatment across various cancer types.

For example, the prognosis of GBM patients is extremely poor, with median survival time of 14–15 months and < 10% probability of survival for 5 years [1, 2]. Conventional therapies for GBM include surgery, radiotherapy, and pharmacotherapy (usually by chemotherapy using temozolomide), but these treatments did not significantly improve survival outcomes [3, 4]. It is difficult to completely surgically resect a GBM tumor, because there is no clear boundary between tumor and normal brain tissue, and post-surgery radiotherapy will lead to radioactive spinal cord necrosis. In addition, most chemotherapeutic drugs cannot reach the lesion site because of the blood-brain barrier [5]. These issues limit the treatment of GBM, and novel therapies are urgently needed to improve the prognosis of GBM patients.

Immunotherapy has become an effective treatment for a number of cancers, as it has been shown to improve the prognosis of many patients with malignancies [6]. The ultimate goal of immunotherapy is to stimulate the patient's natural immune defense system through tumor-associated antigens and adjuvants. Simulating infection is the key to inducing defensive immunity. Cancer vaccines, as a kind of tumor immunotherapy, contribute to adaptive immune responses mainly by enhancing antigen presentation [7]. A recent Phase 1 study reported that a peptide-based vaccine targeting mutant isocitrate dehydrogenase (NADP(+)) 1 (IDH1) was successful in mitigating GBM patients [8]. The results are encouraging for exploring the potential of GBM-related vaccines.

There are four types of cancer vaccines: tumor or immune cell-based vaccines, peptide-based vaccines, viral vector-based vaccines, and nucleic acid-based vaccines. The la

[1] *College of Biomedical Engineering, Chongqing Medical University, Chongqing 400016, China.*
[2] *Life Sciences Institute, Biosafety Level-3 Laboratory, Guangxi Medical University, Nanning 530021, China.*
[3] *School of Basic Medical Sciences, Department of Biochemistry and Molecular Biology, Guangxi Medical University, Nanning 530021, China.*

#Equal contribution
***Corresponding author**

st one is a promising vaccine strategy for a variety of rea sons [9]. In particular, mRNA vaccines have recently eme rged as an attractive alternative to DNA vaccines for anti-cancer treatments. While mRNA vaccines share some bas ic features with DNA vaccines, they may also address so me of the concerns and limitations of the latter [10-13]. Two recent studies have shown the genome-wide antige ns of pancreatic adenocarcinoma (PAAD) and cholangioc arcinoma (CHOL) for the development of mRNA vaccines [14, 15], however, studies focused on identifying genome -wide antigens of GBM have not been reported to date. GBM could be divided into different subtypes based on tr anscription profiles, genetic alterations, and DNA methyl ation [16-18]. For instance, we previously found that gen etic alterations of the CDKN2A and TP53 genes play an es sential role in stratifying GBM [16]. The status of 1p19q c odeletion and MGMTp methylation are also widely used t o stratify GBM [18]. Immune subtype has a certain relati onship with disease progression and degree of malignanc y and can be used as an important factor in judging the p rognosis [14, 15]. Therefore, identification of immune su btypes will help guide patient screening and decisions ab out immunotherapy for GBM patients. However, most of the widely accepted GBM subtype classification systems are based on the status of driver genes rather than on co nsideration of the immune landscape. We propose that i mmune subtype of GBM is important for identifying suita ble patients for vaccination.

The aim of this study was to identify potential tumor anti gens for anti-GBM mRNA vaccine development and to ma p the immune landscape of GBM to select suitable patien ts for vaccination. In this study, seven candidate genes, D IRAS3, EGFLAM, FUCA1, LILRB2, MPZL2, OSMR and STC1, were identified to be related with poor survival and antig en presenting cell (APC) infiltration by screening for the overlap of altered and overexpressed genes in GBM. The n, GBM patients were divided in to four subtypes based o n the clustering of immune-related genes expression, the se results were validated in another independent cohort. Furthermore, we analyzed the relationship of the clinical features and immune subtypes. In addition, we demonst rated that the hub genes play an essential role in immun e landscape of GBM. Our findings provide theoretical sup port and candidate target genes for the development of mRNA vaccines against GBM.

2 Methods

2.1 GEPIA analysis

As described by Huang et al. [14, 15],we used Gene Expressio n Profiling Interactive Analysis (GEPIA) version 2 (GEPIA2, htt p://gepia2.cancer-pku.cn), which is an open-access online to ol that collects tumor and normal samples from The Cancer Genome Atlas (TCGA, (https://www.cancer.gov/tcga) and the Genotype-Tissue Expression databases, to perform the differ ential expression analysis of GBM genes and to calculate the prognostic index of each selected antigen. The GEPIA2 web s erver also provides extended gene expression quantification from the gene level to the transcript level and supports analy sis of a specific cancer subtype as well as comparison betwee n subtypes [19]. In this study, we used analysis of variance to identify the significantly differentially expressed genes based on |log2FC| values > 1 and q values < 0.01. We separated th e patients with a 50% (median) expression level cutoff and p erformed overall survival (OS) and relapse-free survival (RFS) analysis using the Kaplan-Meier method. We used the Cox pr oportional hazards regression model to calculate the hazard r atio. The parameter setting was consistent in each analysis w ithout adjustment for any *P* value. We considered *P* values < 0.05 to be statistically significant.

2.2cBioPortal analysis

We used the cBio Cancer Genomics Portal (cBioPortal, http:// cbioportal.org) to visualize the gene alteration of potential a ntigens against GBM in the TCGA. cBioPortal allows interactiv e exploration of multidimensional cancer genomics data sets [20]. Statistical significance was set at $P < 0.05$.

2.3TIMER analysis

The Tumor Immune Estimation Resource (TIMER, https://cist rome.shinyapps.io/timer/) is an open-access resource for co mprehensively investigate molecular characterization of tum or-immune interactions [21]. In this study, we used TIMER to analyze and visualize the association between immune infiltr ates and expression level of the potent antigens we identifie d. $P < 0.05$ was considered to be statistically significant.

2.4Data extraction of immune-related gene

The normalized gene expression data from 325 patients and clinical data from 306 patients were collected from the Chine se Glioma Genome Atlas (CGGA, http://www.cgga.org.cn/too ls.jsp), and RNA-seq data from 167 patients and clinical data f rom 600 patients were collected from the TCGA. We extracte d a total of 2006 immune-related genes from both the discov ery and validation cohorts, which included single immune cel l-specific, co-stimulating, and co-suppressing molecule-relate d genes, cytokines and cytokine receptor-related genes, antig en processing and presenting-related genes, and other immu ne-related genes. TCGA and CGGA tumor samples lacking clin ical information and normal tissue data were excluded, as we re genes with 0 transcripts per million (TPM) in > 50% of the samples. Ultimately, 103 immune cell-related genes with log 2 (TPM+1) fold-change were included in the subsequent anal ysis.

Discovery and validation of the immune subtypes

Immune-related genes were clustered on the basis of their e xpression profiles, and we developed a consistency matrix to identify corresponding gene modules and immune subtypes. We applied the partition around medoids algorithm using th e “1-Pearson correlation” distance metric and performed 500 bootstraps, each involving 80% of the patients in the discove

ry cohort. We set cluster varied from 2 to 8 to determine the optimal number of clusters, and we defined the optimal partition by evaluating the consensus matrix and the consensus cumulative distribution function. We used an independent TCGA cohort to validate the immune subtypes with the same settings. The consistency of immune subtypes between discove The prognostic distinctions of the GBM immune subtypes were assessed using univariate and multivariate Cox regressiona nd log-rank tests with stage and grade as covariates. We used analysis of variance to assess the relationship between immune subtypes with different immune-related molecular and cellular characteristics. Gene Ontology analysis was performed using The Database for Annotation, Visualization and Integrated Discovery v6.8 (https://david.ncifcrf.gov/) with the genes expressed (TPM >1) as the background [22]. We used single-sample gene set enrichment analysis (ssGSEA), which is the measure of genes that are coordinately up- or down-regulated within a sample, to calculate immune enrichment scores for each sample. We screened the most frequently mutated genes using the chi-square test.

We performed dimensionality reduction analysis to visualize the distribution of immune subtypes across TCGA and CGGA patients using the reduce Dimension function of the Monocle package with a Gaussian distribution. According to the methods published earlier by Huang *et al*. [14, 15], the maximum number of components was set to 4. Finally, we used the function plot cell trajectory with color-coded immune subtypes in the Monocle package to visualize the immune landscape. Weighted correlation network analysis was used to constructed co-expression modules of immune genes according to the methods published earlier by Huang *et al*. [14, 15].

Results

Identification of potential tumor antigens of different cancer type

Using TCGA data, we analyzed the expression levels of natural receptors and ligands across different tumors and found that different tumors possess distinct natural receptors and ligands (Fig. 1). This indicates that different tumors should have different RNA vaccines. Among these, we found that gliomas

ry and validation datasets was quantified by calculating the Pearson correlation and the proportion of the intra-group in the centroid of the gene module scores.

2.6Prognostic evaluation of immune subtypes and construction of immune landscape

have unique receptor and ligand expression. Therefore, we will focus our subsequent research on the natural receptors and ligands of gliomas.

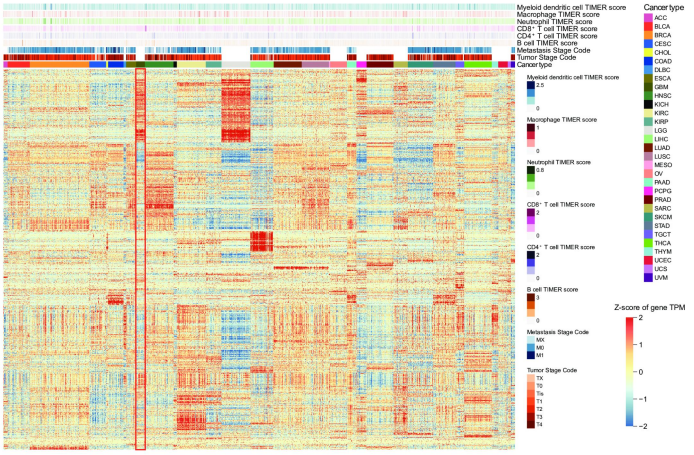

Figure 1. Heatmap showing the expression of receptor-ligand genes of 31 tumors in TCGA

Identification of potential tumor antigens of GBM

We screened the aberrantly expressed genes of GBM to identify potential antigens. Among the 7659 differentially expressed genes in GBM, we detected 5221 overexpressed genes that may encode tumor-associated antigens (Fig. 2A). Subsequently, 8182 mutated genes encoding tumor-specific antigens were screened by analyzing altered genome fraction and mutation counts in individual patients (Fig. 2B, C). Genes with the highest alteration frequency in the fraction genome-altered group were ASAH1, DES, DIS3, FAM86B2, PROSER1, TUBA4A, EGFR, and SEC61G (Fig. 2D). High mutation counts were observed in *CDKN2A-DT*, *SEC61G*, *MTAP*, *CDKN2A*, *CDKN2B*, *LANCL2*, *EGFR*, *VOPP1*, *PTEN*, and *DMD* (Fig. 2E). Overall, we identified 1633 overexpressed and frequently mutated tumor-specific genes.

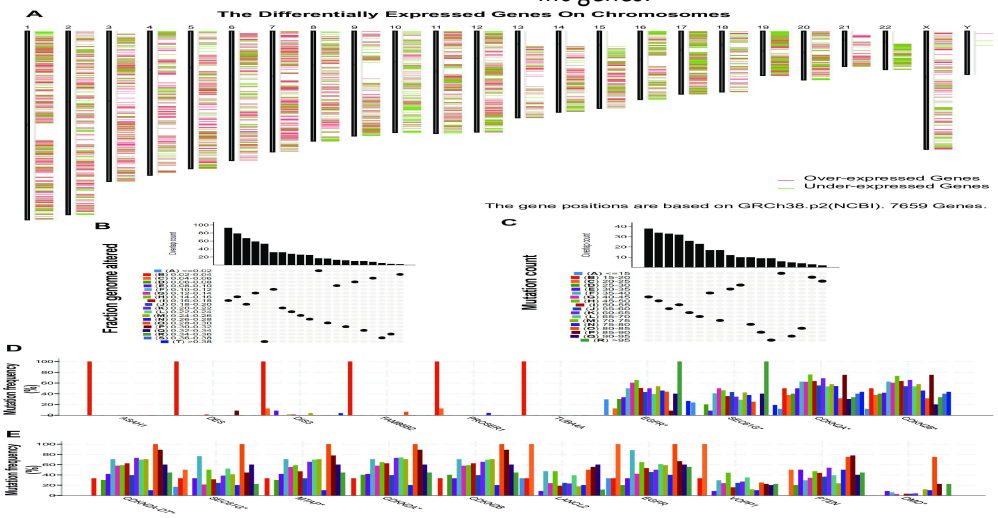

Figure 2. Identification of potential tumor antigens of GBM.(A)The differentially expressed genes on chromosomes of GBM. (B-E) Identification of potential tumor-specific antigens. Samples overlapping in(B)fraction genome altered and(C)mutation count. Genes with highest frequency in(D)fraction genome altered and(E)mutation count groups.

Identification of tumor antigens associated with GBM prognosis and antigen-presenting cells

We next screened the aforementioned genes for prognosis-related tumor antigens that may have immune stimulatory or inhibitory effects as candidates for developing mRNA vaccines. Among the 1633 overexpressed and frequently mutated tumor-specific genes, 35 genes were closely associated with the OS of GBM patients, and 7 of those genes showed significant correlation with RFS (Fig. 3A). Patients overexpressing the seven genes (DIRAS family GTPase 3 (*DIRAS3*), alpha-L-fucosidase 1 (*FUCA1*), leukocyte immunoglobulin like receptor B2 (*LILRB2*), myelin protein zero like 2 (*MPZL2*), oncostatin M receptor (*OSMR*), stanniocalcin 1 (*STC1*), and EGF like, fibronectin type III and laminin G domains (*EGFLAM*)) in the tumor tissues had significantly shorter survival time compared to the group with lower expression (Fig. 3B–H). Thus, these seven gene candidates were identified as being critical for GBM development and progression. What's more, higher expression levels of *DIRAS3*, *FUCA1*, and *LILRB2* were significantly associated with increased tumor infiltration of B cells, macrophages, and dendritic cells (Fig. 4A, C, D). *EGFLAM*, *MPZL2*, *OSMR*, and *STC1* expression levels also tended to be associated with increased infiltration of immune cells, but the patterns were more variable (Fig. 4B, 3E–G). All in all, these findings suggest that the seven identified tumor antigens may be directly processed and presented by antigen presenting cells (APCs) to T cells and recognized by B cells to trigger an immune response. Therefore, they are promising candidates for developing mRNA vaccines against GBM.

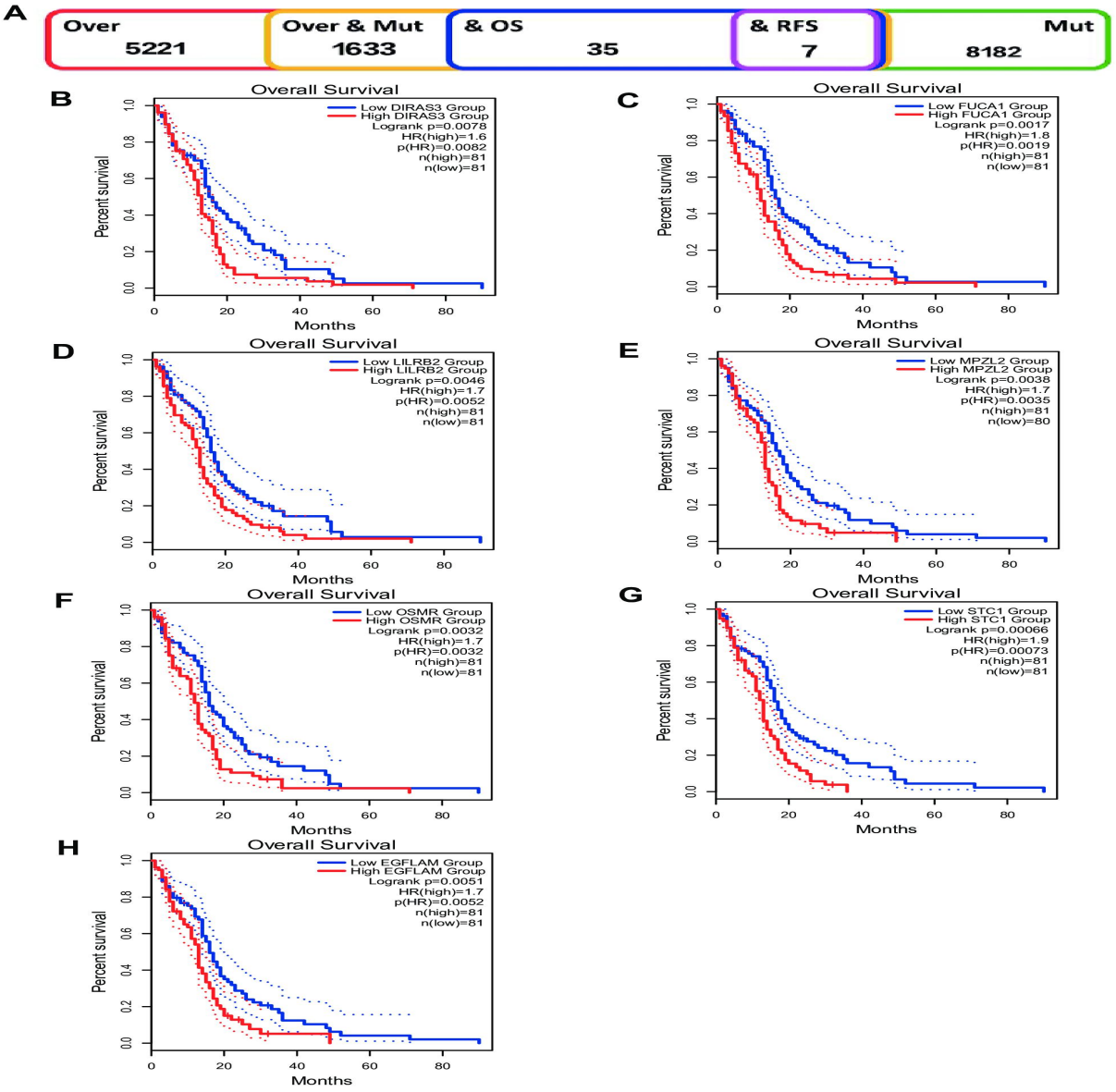

Figure 3. Identification of tumor antigens associated with the prognosis of GBM.(A)1633 potential tumor antigens with high expression and mutation in GBM, and 7 candidates significant associated with OS and RFS. (B-H) Kaplan-Meier curves showing OS of GBM patients stratified on the basis of(B) DIRAS3,(C)FUCA1,(D)LILRB2,(E)MPZL2,(F)OSMR,(G)STC1 and(H)EGFLAM expression levels.

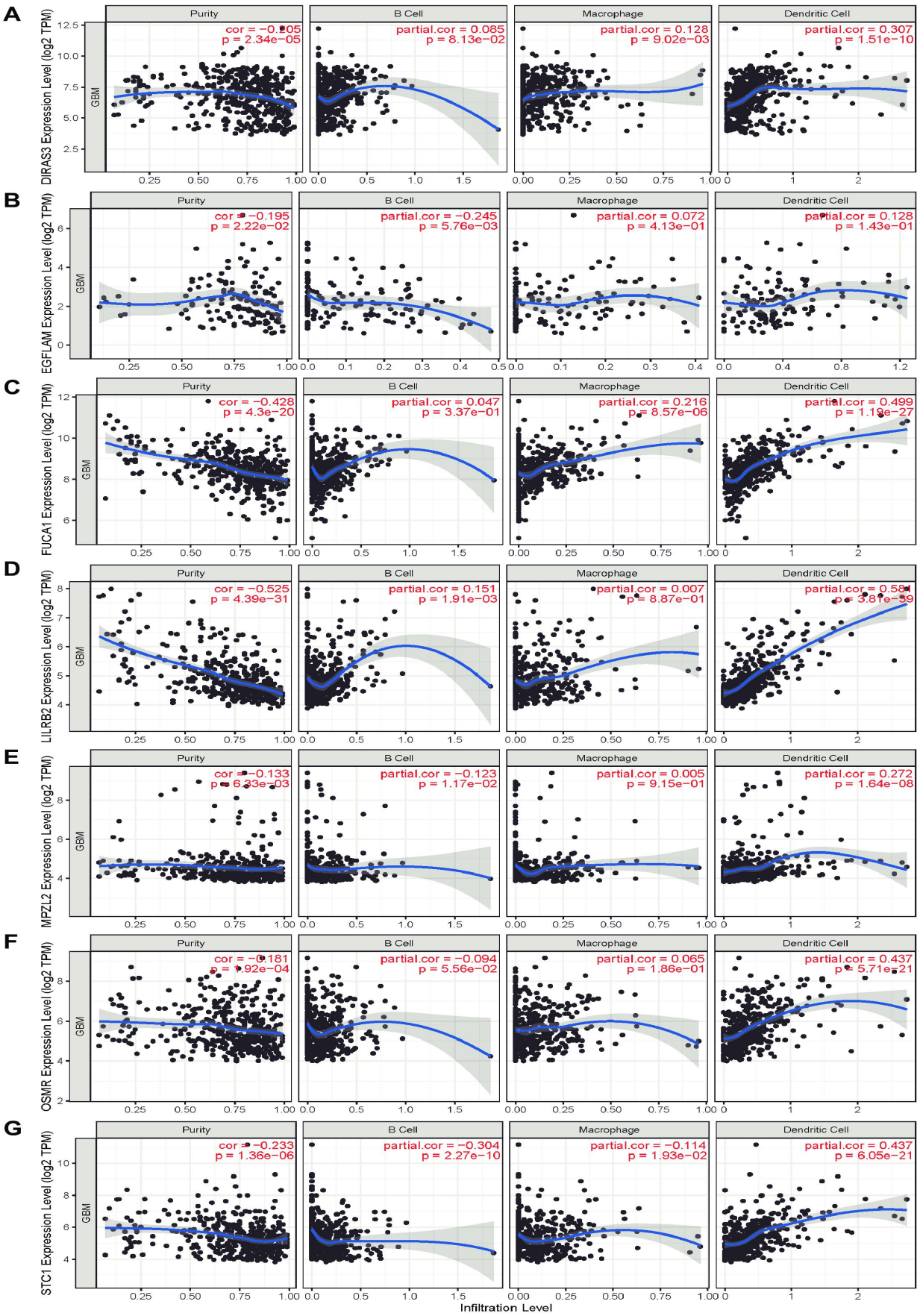

Figure 4. Correlation between the expression levels of.(A)DIRAS3,(B)EGFLAM,(C)FUCA1,(D)LILRB2,(E)MPZL2,(F)OSMR and(G)STC1 and infiltration of B cells, macrophages and dendritic cells in GBM.

Identification of potential immune subtypes of GBM

Immunophenotyping is helpful for identifying appropriate patients for vaccination since it can reflect the immune status of the tumor and its microenvironment. So, we analyzed the expression profiles of 402 immune-related genes in 325 GBM samples from the CGGA database to construct consensus clustering. We chose k = 4, where immune-related genes appeared to be stably clustered according to their cumulative distribution function and their function delta area (Fig. 5A, B), and we obtained four immune subtypes designated as IS1–IS4 (Fig. 5C). IS3 was associated with the best prognosis, whereas IS1 and IS2 had poor survival probability (Fig. 5D). We also analyzed the expression profiles of 396 immune-related genes in 167 GBM samples from the TCGA database to construct consensus clustering. Consistent with the results obtained with the CGGA cohort, four immune subtypes were obtained and related to prognosis in this cohort (Fig. 5E).

We investigated the subtype distribution in samples with different 1p19q codeletion status and MGMTp methylation status in the CGGA cohort and found that IS2 was only distributed in those with 1p19q non-code status (Fig. 5F), whereas IS1–IS4 were irregularly clustered in samples with both MGMTp methylated and un-methylated status (Fig. 5G). In the TCGA cohort, pan-*CDKN2A* and *TP53* were irregularly distributed in IS1–IS4 (Fig. 5H). Furthermore, we investigated the IS1–IS4 distribution across classical, mesenchymal, neural, and proneural subtypes, which are used for gene expression-based molecular classification of GBM. The classical subtype was only associated with IS2, the mesenchymal subtype was only associated with IS1, and the proneural subtype was only associated with IS3 (Fig. 5I). These results show that immunotyping can be used to predict prognosis of GBM patients and classical subtypes.

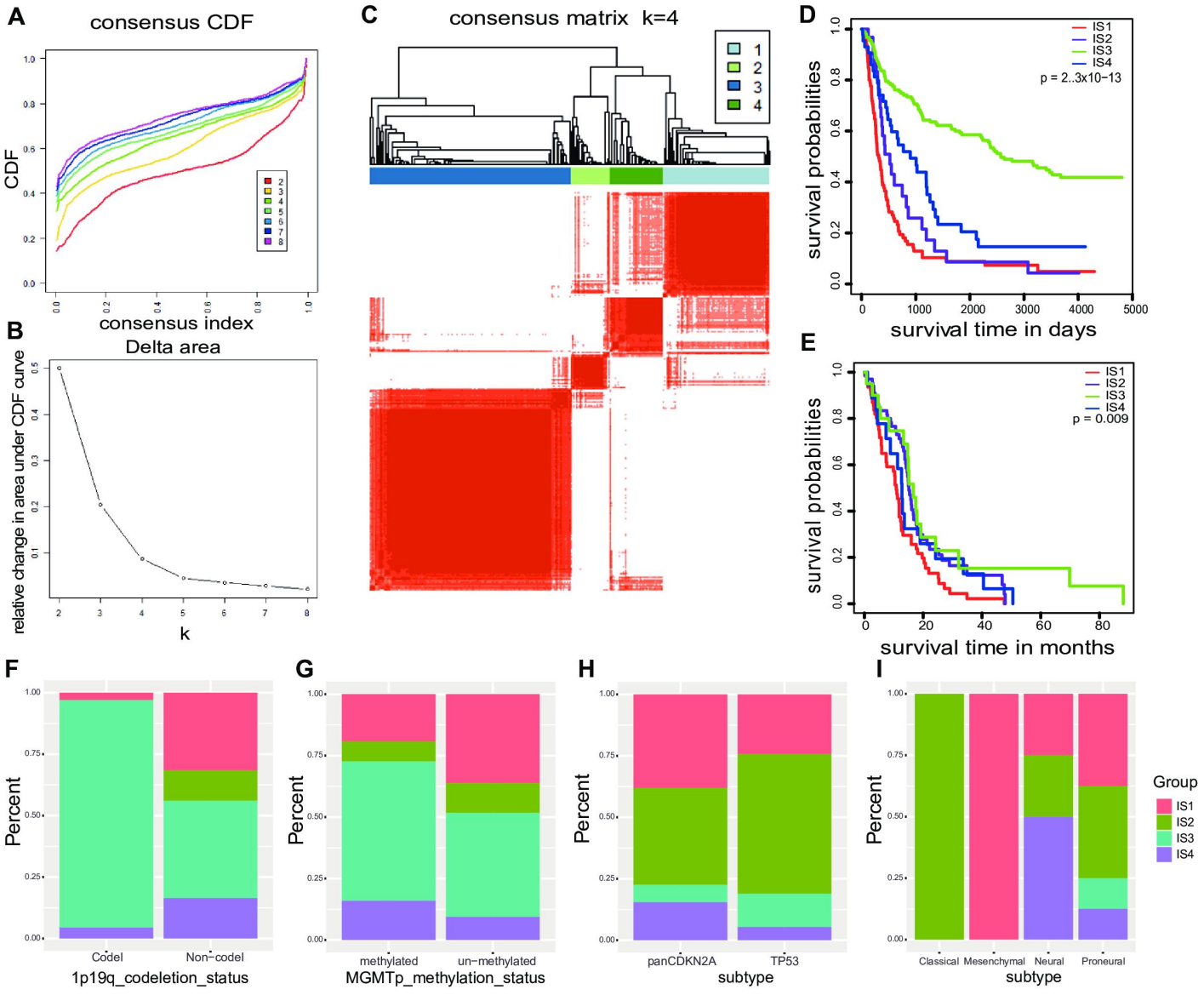

Figure 5. Identification of potential immune subtypes of GBM.(A)Cumulative distribution function curve and(B)delta area of immune-related genes in the CGGA cohort.(C)Sample clustering heat map.(D)Kaplan-Meier curves showing OS of GBM immune subtypes in the CGGA cohort.

(E)Kaplan-Meier curves showing OS of GBM immune subtypes in the TCGA cohort. (F, G) Distribution of IS1–IS4 across GBM(F)1p19q codeletion status and(G) MGMTp methylation status in the CGGA cohort. (H, I) Distribution ratio of IS1–IS4 across GBM (H, I) subtypes in the TCGA cohort.

Association of immune subtypes with tumor mutational burden and mutational status

Higher tumor mutation burden (TMB) is correlated with stronger anti-cancer immunity. We calculated the TMB in each patient using the mutect2-processed mutation dataset of TCGA. However, there was no significant difference in TMB among the four immune subtypes (Fig. 6A). Eight genes (*TP53*, *EGFR*, *PTEN*, *TNN*, *SPTA1*, *RYR2*, *MUC16*, and *IDH1*) were most frequently mutated in these subtypes (Fig. 6B). These findings suggest that the tumor antigens do not differ significantly among the IS1–IS4 subtypes.

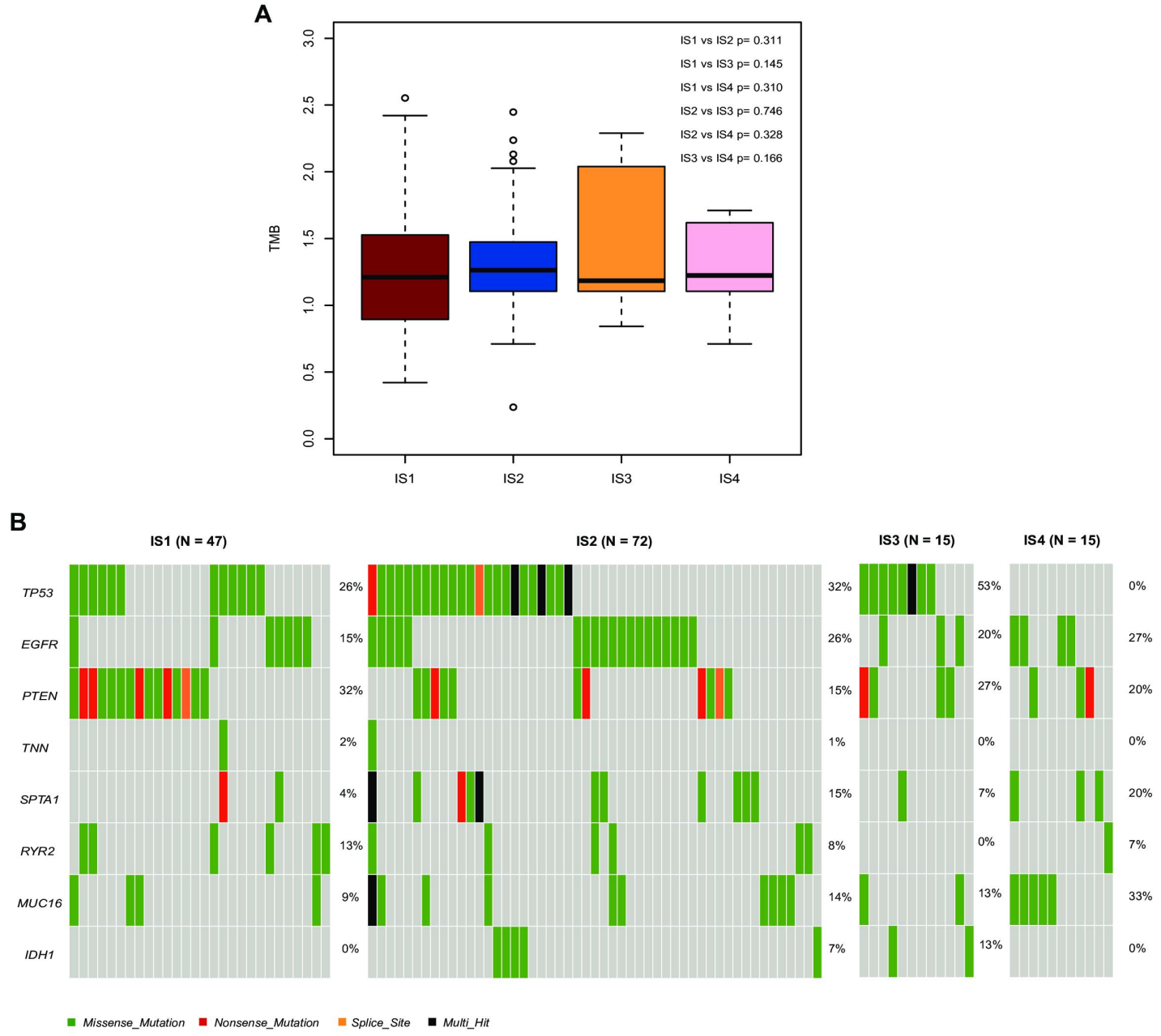

Figure 6. Association between immune subtypes and.(A)TMB in IS1–IS4 of GBM.(B)Eight highly mutated genes in GBM immune subtypes.

Association between immune subtypes of GBM and immune modulators

Both immune checkpoints (ICPs) and immunogenic cell death (ICD) modulators play critical roles in modulating host anti-tumor immunity, which could influence the efficacy of mRNA vaccines. Therefore, we analyzed their expression levels in the different subtypes. Forty-three ICP-related genes were detected in both the CGGA and TCGA cohorts, of which 40 (93%) genes in the CGGA cohort (Fig. 7A) and 39 (91%) genes in the TCGA cohort (Fig. 7B) were differentially expressed among the four immune subtypes. Furthermore, the overall expression level of ICPs in the TCGA cohort was higher than that in the CGGA cohort. Likewise, 21 ICD genes were detected in both the CGGA and TCGA cohorts, of which 20 (95%) genes in the

CGGA cohort (Fig. 7C) and 19 (90%) genes in the TCGA cohort (Fig. 7D) showed significant differences among the four immune subtypes. Therefore, immunotyping can reflect the expression levels of ICPs and ICD modulators, which may prove to be useful therapeutic biomarkers of the effectiveness of vaccines.

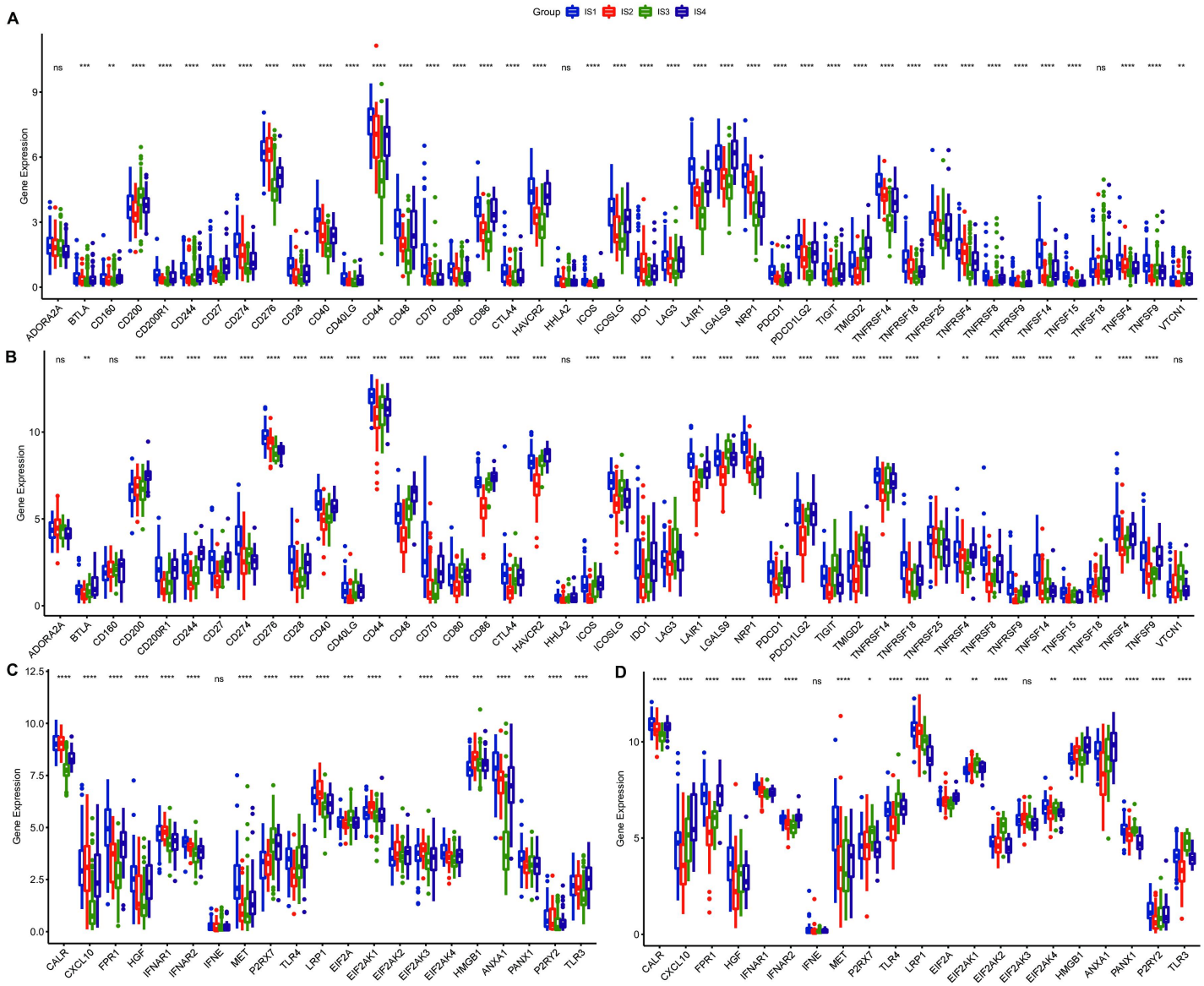

Figure 7. Differential expression of ICP genes among the immune subtypes of GBM in the.(A)CGGA and(B)TCGA cohorts. Differential expression of ICD modulator genes among the GBM immune subtypes in the (C) CGGA and (D) TCGA cohorts. * $p < 0.01$, ** $p < 0.001$, *** $p < 0.0001$, and ****$p < 0.00001$.

Association between immune subtypes and GBM-related tumor marker

Neuron-specific enolase (NSE) is the most commonly used prognostic tumor biomarker for neuroma, and a high value indicates cancer progression, poor prognosis, or cancer relapse [23]. In this study, both the CGGA and TCGA cohorts displayed significant differences in NSE expression levels across the immune subtypes (Fig. 8A, B). For example, IS4 showed lower NSE expression in the CGGA cohort (Fig. 8A), whereas IS3 showed lower NSE expression in the TCGA cohort (Fig. 8B). These outcomes demonstrate that immune subtype is a better indicator than NSE for anticipating GBM patient prognosis.

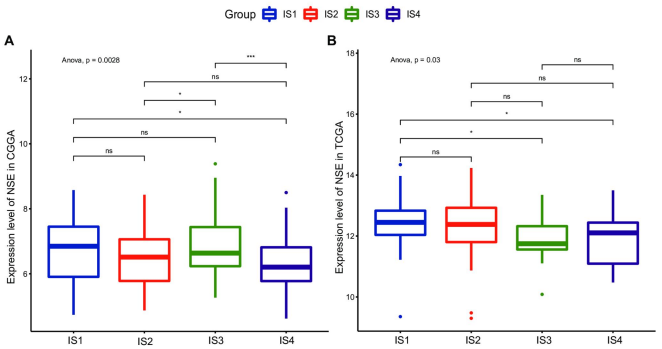

Figure 8. Association of immune subtypes with NSE. (A, B) NSE expression in GBM immune subtypes in the CGGA.(A)and TCGA (B)cohorts.

Cellular and molecular characteristics of immune subtypes

Because the tumor immune status to a great extent will impact the efficacy of mRNA vaccines, we further described the resistant cell segments in the four immune subtypes by scoring 28 marker genes in both the CGGA and TCGA cohorts utilizing ssGSEA. The immune cell composition differed significantly among the subtypes in the CGGA cohort (Fig. 9A). Most immune cells were more enriched in the IS1 and IS4 immune subtypes. For instance, the scores of activated CD8 T cells and immature B cells were significantly higher in IS1 and IS4 compared to IS2 and IS3 (Fig. 9A, C). Therefore, IS1 and IS4 are immunologically "hot" and IS2 and IS3 are immunologically "cold" phenotypes. Similar trends were seen in the TCGA cohort, as most immune cells were more enriched in the IS1 and IS4 immune subtypes (Fig. 9B, D). These results suggest that immune subtype reflects the GBM immune status and that it can be used to identify suitable patients for mRNA vaccination. The mRNA vaccines with these antigens should induce immune infiltration in patients with immunologically "cold" (IS2 and IS3) tumors.

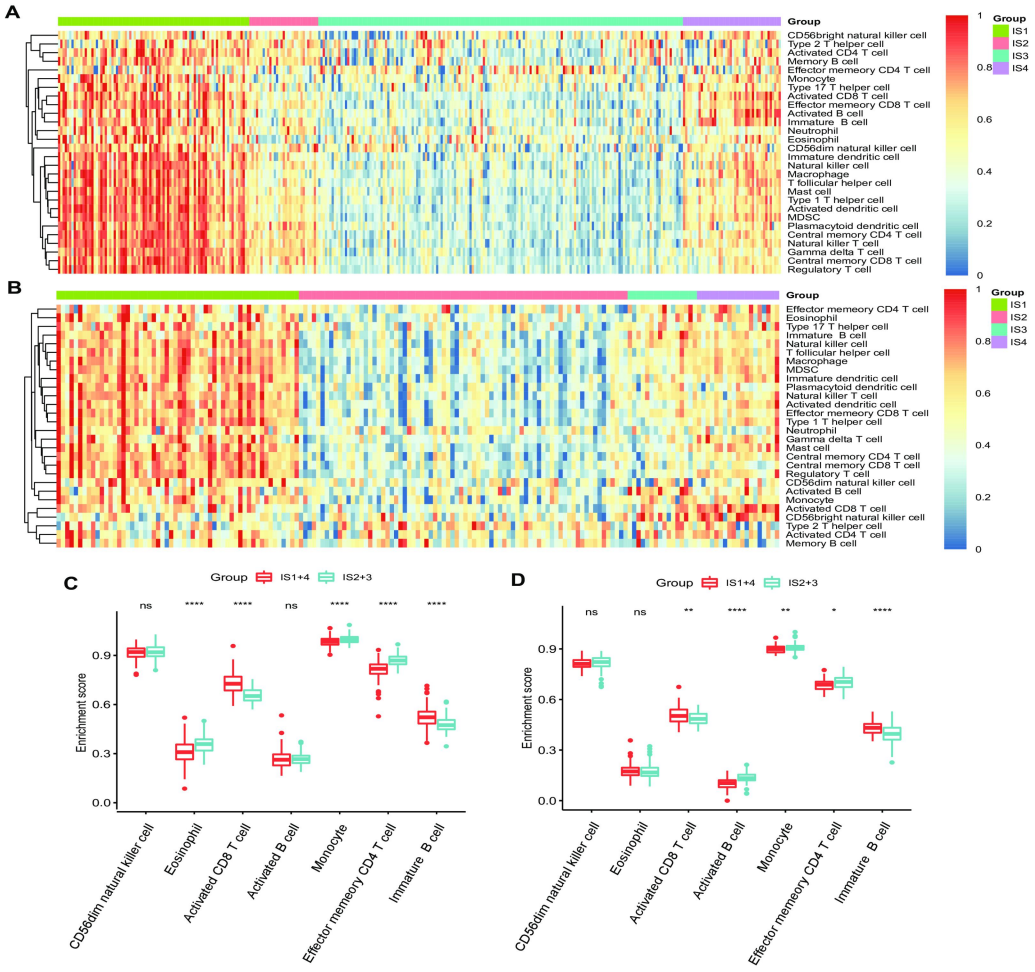

Figure 9. Differential enrichment scores of 28 immune cell signatures among GBM immune subtypes in the.(A)CGGA and (B)TCGA cohorts.(C)CGGA and(D)TCGA cohorts.

Immune landscape of GBM

The immune gene expression profiles of each patient were used to develop the immune landscape of GBM for mRNA vaccine application (Fig. 10A). The horizontal axis was correlated with various immune cells, of which effector memory CD8 T cells, type 1 T helper cells, natural killer T cells, mast cells, gamma delta T cells, central memory CD4 T cells, and activated CD8 T cell showed the most negative correlation, whereas the vertical coordinate was mostly negatively associated with central memory CD8 T cells (Fig. 10B). IS1, IS2, and IS4 were further divided into two subgroups according to the distribution location of the three immune subtypes in the immune landscape (Fig. 10C), and several immune cells differed significantly among these subsets (Fig. 10D). For instance, the enrichment scores of activated B cells, activated CD8 T cells, immature B cells, macrophages, MDSCs, natural killer cells, and T follicular helper cells in IS1B were lower than those in IS1A. Likewise, IS2A scored lower than IS2B in terms of activated B cells, activated CD4 T cells, activated CD8 T cells, CD56bright natural killer cells, CD56dim natural killer cells, effector memory CD4 T cells, memory B cells, monocytes, plasmacytoid dendritic cells, and type 2 T helper cells. IS4B scored lower in terms of effector memory CD8 T cells, macrophages, and MDSCs, while IS4A scored lower in terms of monocytes and type 2 T hel

per cells. Thus, mRNA vaccines may be relatively viable and more effective in IS1B and IS2A. Furthermore, patients were separated into six groups based on their location in the immune landscape and were subjected to prognostic comparison (Fig. 10E). Patients in group 5 showed the poorest survival probability (Fig. 10F), which is consistent with the aforementioned results. Overall, these findings suggest that the immune landscape based on immune subtypes can be used to define the immune components of each GBM patient and predict prognoses that help in selecting suitable patients and personalized therapeutics for mRNA vaccines.

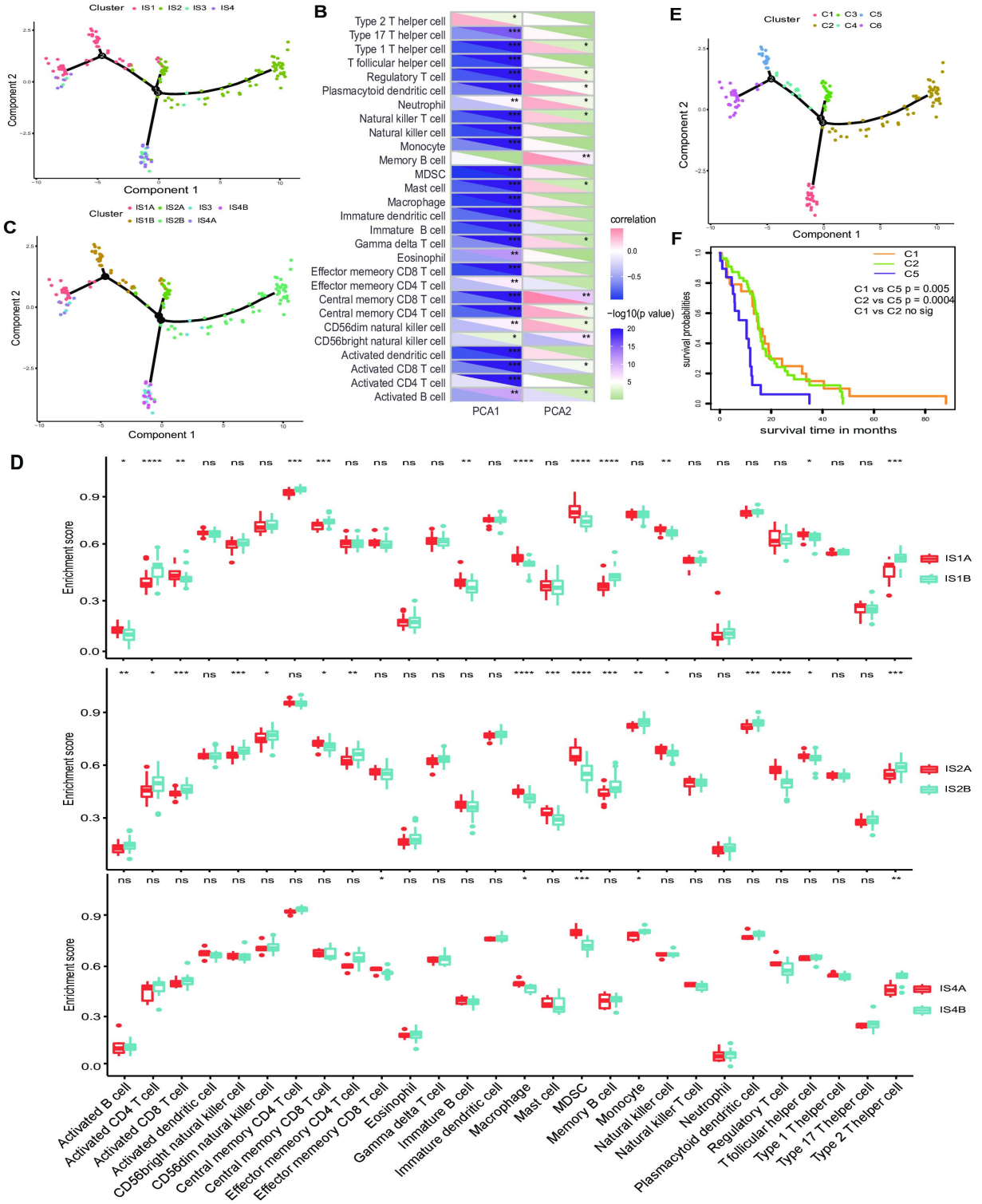

Figure 10. The immune landscape of GBM.(A)The horizontal axis represents the first principal component, and the vertical axis represents the second principal component.(B)Heat map of two principal components with 28 immune cell signatures.(C)Immune landscape of the subsets of GBM immune subtypes.(D)Differential enrichment scores of 28 immune cell signatures in the above subsets.(E)Immune landscape of samples from three extreme locations and(F)their prognostic status.

Identification of immune gene co-expression modules and hub genes of GBM

Co-expression modules of immune genes were constructed by clustering the samples using weighted correlation network analysis (Fig. 11A). We set the soft threshold at 4 for the scale-free network (Fig. 11B). We calculated eigengenes of each module and merged the close modules into a new one (Fig. 11C). We further analyzed the distribution of the eigengenes in the four immune subtypes of these modules and detected significantly different distributions in every module (Fig. 11D). IS3 had the lowest number of eigengenes in the black, blue, pink, brown, and red modules, and IS2 had the lowest number of eigengenes in the magenta, turquoise, yellow, and green modules. IS1 had the highest number of eigengenes in the black, blue, pink, and turquoise modules, and IS4 had the highest number of eigengenes in the magenta, brown, red, and yellow modules. Thus, IS3 and IS2 corresponded to immunologically cold and IS1 and IS4 to inflamed tumors.

Further prognostic correlation analysis showed that the blue module was significantly associated with the prognosis of GBM (Fig. 12A). Moreover, genes in the blue module were enriched with cytokine-cytokine receptor interaction, chemotaxis, interleukin-related, and immune-related pathways and biological processes (Fig. 12B). The blue module also was negatively related to component 1 of the immune landscape (Fig. 12C). Analysis of the prognostically relevant genes of the blue module showed that higher expression scores correlated with better prognosis in the CGGA and TCGA cohorts, which is consistent with the abovementioned prognosis findings (Fig. 12D). The immune-related genes *CLCF1*, *CD54*, *SOCS3*, *PLAUR*, *LIF*, and *BCL3* were associated with the blue module, and we conclude that these six immune hub genes are potential biomarkers for effectiveness of mRNA vaccines in GBM.

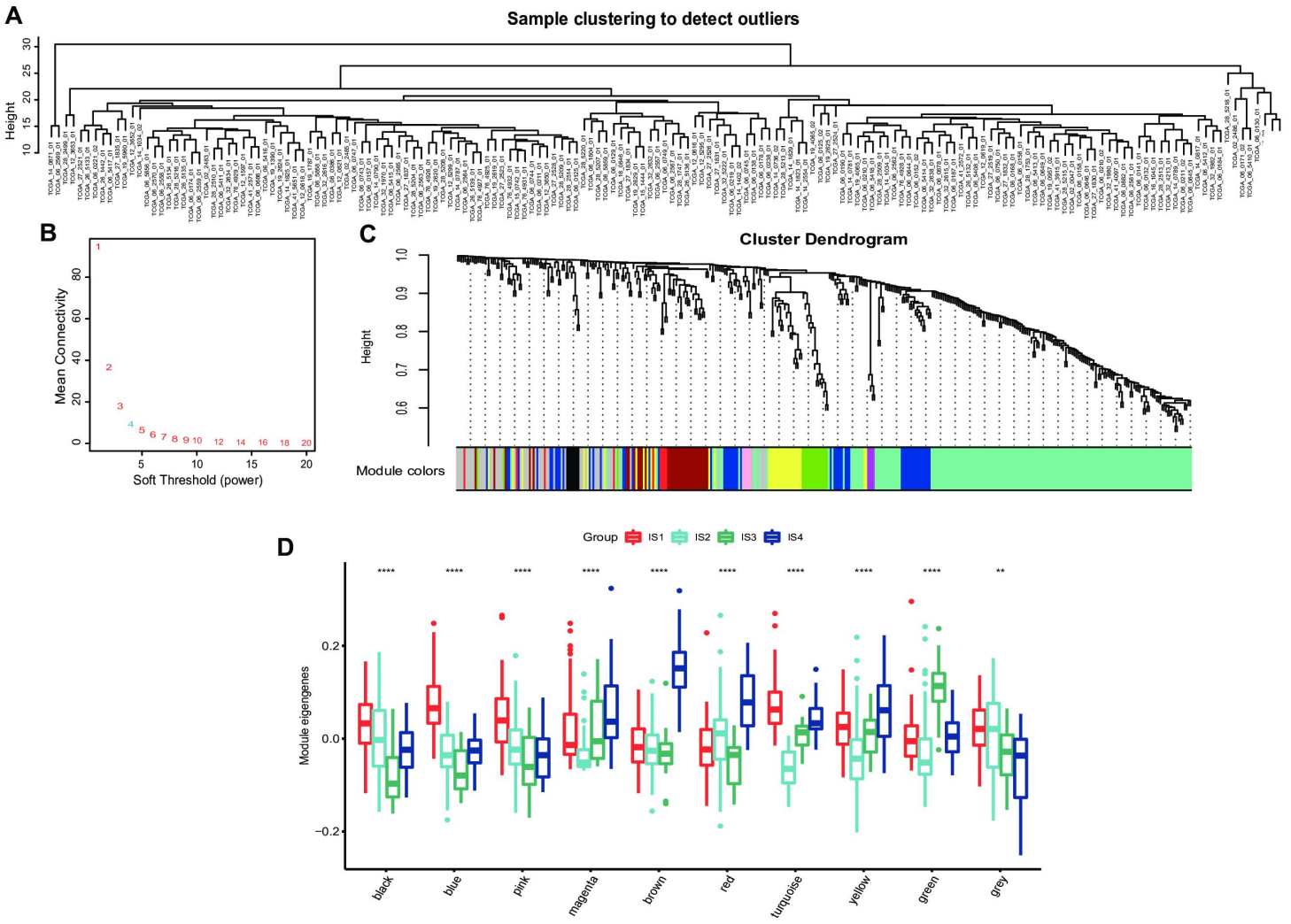

Figure 11. Identification of immune gene co-expression modules of GBM.(A)Sample clustering.(B)Mean connectivity for various soft-thresholding powers.(C)Dendrogram of all differentially expressed genes clustered based on a dissimilarity measure (1- topological overlap measure).(D)Differential distribution of feature vectors of each module in GBM immune subtypes.

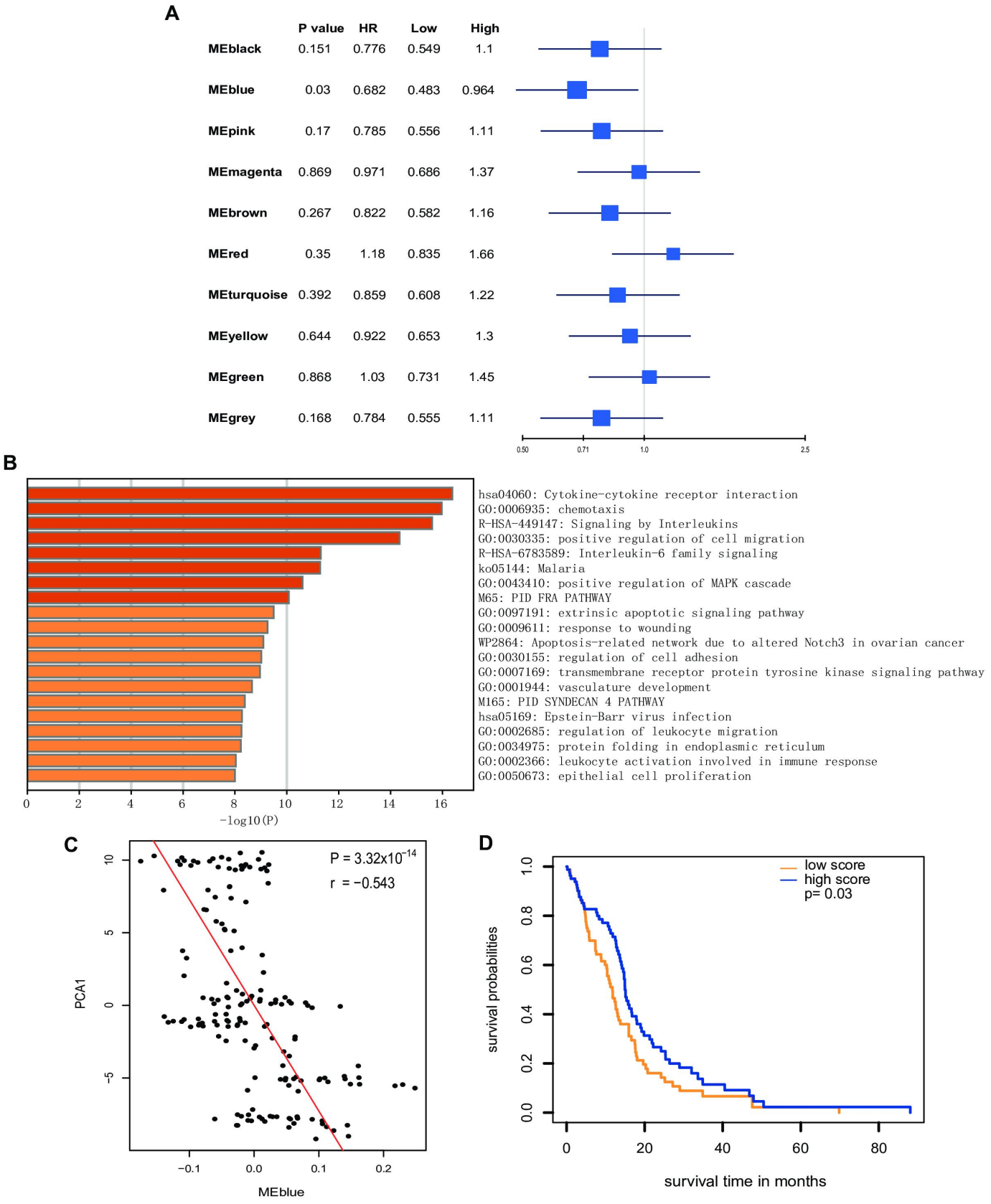

Figure 12. Identification of immune hub genes of GBM.(A)Forest maps of single factor survival analysis of 10 modules in GBM.(B)Gene Ontology enri chment analyses of genes in the blue module.(C)Correlation between the blue module feature vector and the second principal component in the im mune landscape.(D)Survival analysis of high and low mean group in the blue module.

Discussion

Previous studies have reported that IDH1, 1p/19q, and TERT play key roles in GBM and that IDH1 mutation is associated with prognosis of GBM [24]. The recently published results of the Phase 1 trial revealed that the experimental peptide-based vaccine focused of IDH1 in GBM is safe and without serious side effects [8]. This vaccine should help a patient's immune system better target brain tumors. Arginine mutation to histidine (R132H) is the most common type of IDH1 mutation. Researchers have found that IDH1 mutations in glioblastoma mainly occur in high-grade glioblastoma and that IDH1 mutations in primary glioblastoma are almost always wild type [25]. We previously reported that IDH1 defines a molecular subtype of glioma [26]. Glioblastoma patients with IDH1 mutation had longer overall survival of 31 months compared with the 15-month survival of patients with wild-type IDH1. Antibodies and CD4+ Th1 cells recognize IDH1(R132H) specifically, and the IDH1(R132H)-specific peptide vaccine (IDH1-vac) induces specific therapeutic T helper cell responses that are effective against IDH1(R132H)+ tumors. In fact, IDH1 vaccine-induced immune responses were observed in 93.3% of patients [8]. Thus, the IDH1 vaccine is a breakthrough in patients with IDH1 (R132H), as some patients benefitted from it. However, many genes are associated with GBM, and potent mRNA vaccines against GBM are still not fully defined. Compared with DNA, mRNA offers strong advantages as a therapeutic or more specifically as a vaccine [27].

Using clinical data to analyze the overexpressed and mutational landscape of GBM, we identified IRAS3, EGFLAM, FUCA1, LILRB2, MPZL2, OSMR, and STC1 to be targetable antigens and promising mRNA vaccine candidates for GBM. We also analyzed the survival outcomes and relationship with immunity of those targetable antigens. This is the report to systematically identify potent antigens in GBM for mRNA vaccine development. Although further clinical evaluation of these candidate antigens is required, the potential of these tumor antigens for use in an anti-GBM mRNA vaccine is supported in previous reports. For instance, Ahmad-Sharanek et al. reported that the OSMR gene controls glioma stem cell respiration and confers resistance of glioblastoma to ionizing radiation (IR) and that suppression of OSMR improves glioblastoma response to IR and prolongs lifespan [28]. In another study, upregulation of DIRAS3 was found to promote glioma cell proliferation and invasion by activating EGFR-AKT signaling [29]. EGFLAM silencing inhibited the proliferation, migration, and invasion of U87 cells, which was regulated through repression of the PI3K/AKT pathway [30]. Overexpressed FUCA1 induced autophagy in glioma, and lower levels of tumor-infiltrating macrophages were identified in FUCA1-downregulated glioma tissues [31]. LILRB2, which regulates synaptic plasticity in an Alzheimer's model, was highly expressed in the immune compartment at the mRNA level [32]. MPZL2 was shown to have a key function in the proliferation and tumorigenesis of glioblastoma-initiating cells [33]. STC1 regulates glioblastoma migration and invasion by activating the TGF-β/SMAD4 signaling pathway [34]. Due to the high heterogeneity of GBM and the lack of common specific antigen expression, it is difficult to kill all tumor cells with a single mRNA vaccine, thus the combination of multiple antigens may produce better results.

Because GBM is a malignant tumor with strong heterogeneity, we classified GBM into four immune subtypes according to their immune gene expression profiles. GBM classification based on immune genes is helpful for selecting the appropriate patient population for vaccination. Most molecular classifications of gliomas are based on genetic variation and gene expression profiles, which group gliomas into four subtypes (classical, neural, proneural, and mesenchymal) based on transcriptional features [17]. In our previous research, we proposed using DNA methylation classification and IDH1, TP53, and CDKN2A mutation status as a basis for GBM classification [26, 35]. In addition, distribution and infiltration of immune components are reported to be associated with the commonly described GBM subgroups [36]. In the current study, we found that all mesenchymal GBMs belonged to IS2 and all classical GBMs belonged to IS1. Most of the GBMs were immunologically "cold" phenotypes, as reported before [3, 36]. However, IS1 and IS4 showed significantly higher scores of activated CD8 T cells, eosinophils, activated B cells, monocytes, and effector memory CD4 T cells compared to the others, which suggests that IS1 and IS4 are immunologically "hot." Because patients with different immune subtypes would likely respond differently to the vaccine, IS1 and IS4 should benefit more from the vaccine. Immunotype also can be a prognostic biomarker for GBM, which indicates that our new immunotyping method is reliable. However, the potential vaccine antigens identified still require further exploration.

Conclusion

We found that tumor antigens IRAS3, EGFLAM, FUCA1, LILRB2, MPZL2, OSMR, and STC1 were related to poor survival and that infiltration of APCs are potent antigens that could be targeted to treat GBM. GBM patients were divided in to four subtypes based on the clustering of immune-related gene expression, and patients classified as IS1 and IS4 should be suitable for vaccination. We are the first to identify potent antigens and immune subtypes in GBM, and these data will be useful for future mRNA vaccine development and selection of suitable vaccine recipients.

References

1.Ohgaki H and Kleihues P (2005) Epidemiology and etiology of gliomas. Acta Neuropathol 109:93-108. doi: 10.1007/s00401-005-0991-y

2.Sasmita AO, Wong YP and Ling APK (2018) Biomarkers and therapeutic advances in glioblastoma multiforme. Asia Pac J Clin Oncol 14:40-51. doi: 10.1111/ajco.12756

3.Lim M, Xia Y, Bettegowda C and Weller M (2018) Current state of immunotherapy for glioblastoma. Nat Rev Clin Oncol 15:422-442. doi: 10.1038/s41571-018-0003-5

4.Parlato C, Barbarisi M, Moraci M and Moraci A (2006) Surgery, radiotherapy and temozolomide in treating high-grade gliomas. Frontiers in Bioscience 11:1280-1283. doi: 10.2741/1881

5.Birk HS, Han SJ and Butowski NA (2017) Treatment options for recurrent high-grade gliomas. CNS Oncology 6:61-70. doi: 10.2217/cns-2016-0013
6.Kruger S, Ilmer M, Kobold S, Cadilha BL, Endres S, Ormanns S, Schuebbe G, Renz BW, D'Haese JG, Schloesser H, Heinemann V, Subklewe M, Boeck S, Werner J and von Bergwelt-Baildon M (2019) Advances in cancer immunotherapy 2019 - latest trends. J Exp Clin Cancer Res 38:268. doi: 10.1186/s13046-019-1266-0
7.Velcheti V and Schalper K (2016) Basic Overview of Current Immunotherapy Approaches in Cancer. American Society of Clinical Oncology 35:298-308. doi: 10.1200/EDBK_156572
8.Platten M, Bunse L, Wick A, Bunse T, Le Cornet L, Harting I, Sahm F, Sanghvi K, Tan CL, Poschke I, Green E, Justesen S, Behrens GA, Breckwoldt MO, Freitag A, Rother LM, Schmitt A, Schnell O, Hense J, Misch M, Krex D, Stevanovic S, Tabatabai G, Steinbach JP, Bendszus M, von Deimling A, Schmitt M and Wick W (2021) A vaccine targeting mutant IDH1 in newly diagnosed glioma. Nature 592:463-468. doi: 10.1038/s41586-021-03363-z
9.Faghfuri E, Pourfarzi F, Faghfouri AH, Abdoli Shadbad M, Hajiasgharzadeh K and Baradaran B (2021) Recent developments of RNA-based vaccines in cancer immunotherapy. Expert Opin Biol Ther 21:201-218. doi: 10.1080/14712598.2020.1815704
10.Miao L, Zhang Y and Huang L (2021) mRNA vaccine for cancer immunotherapy. Mol Cancer 20:41. doi: 10.1186/s12943-021-01335-5
11.Tombacz I, Weissman D and Pardi N (2021) Vaccination with Messenger RNA: A Promising Alternative to DNA Vaccination. Methods Mol Biol 2197:13-31. doi: 10.1007/978-1-0716-0872-2_2
12.Pardi N, Hogan MJ, Porter FW and Weissman D (2018) mRNA vaccines - a new era in vaccinology. Nat Rev Drug Discov 17:261-279. doi: 10.1038/nrd.2017.243
13.Pardi N, Hogan MJ and Weissman D (2020) Recent advances in mRNA vaccine technology. Curr Opin Immunol 65:14-20. doi: 10.1016/j.coi.2020.01.008
14.Huang X, Tang T, Zhang G and Liang T (2021) Identification of tumor antigens and immune subtypes of cholangiocarcinoma for mRNA vaccine development. Mol Cancer 20:50. doi: 10.1186/s12943-021-01342-6
15.Huang X, Zhang G, Tang T and Liang T (2021) Identification of tumor antigens and immune subtypes of pancreatic adenocarcinoma for mRNA vaccine development. Mol Cancer 20:44. doi: 10.1186/s12943-021-01310-0
16.Xia Zhou, Gonghua Li, Sanqi An, Wenxing Li, Huihui Yang, Yicheng Guo, Zhi Dai, Shaoxing Dai, Junjuan Zheng, Jingfei Huang, Antonio Iavarone and Zhao. X (2018) A new method of identifying glioblastoma subtypes and creation of corresponding animal models. Oncogene 37:4781-4791. doi: 10.1038/s41388-018-0305-1
17.Verhaak RG, Hoadley KA, Purdom E, Wang V, Qi Y, Wilkerson MD, Miller CR, Ding L, Golub T, Mesirov JP, Alexe G, Lawrence M, O'Kelly M, Tamayo P, Weir BA, Gabriel S, Winckler W, Gupta S, Jakkula L, Feiler HS, Hodgson JG, James CD, Sarkaria JN, Brennan C, Kahn A, Spellman PT, Wilson RK, Speed TP, Gray JW, Meyerson M, Getz G, Perou CM, Hayes DN and Cancer Genome Atlas Research N (2010) Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. Cancer Cell 17:98-110. doi: 10.1016/j.ccr.2009.12.020
18.Ceccarelli M, Barthel FP, Malta TM, Sabedot TS, Salama SR, Murray BA, Morozova O, Newton Y, Radenbaugh A, Pagnotta SM, Anjum S, Wang J, Manyam G, Zoppoli P, Ling S, Rao AA, Grifford M, Cherniack AD, Zhang H, Poisson L, Carlotti CG, Jr., Tirapelli DP, Rao A, Mikkelsen T, Lau CC, Yung WK, Rabadan R, Huse J, Brat DJ, Lehman NL, Barnholtz-Sloan JS, Zheng S, Hess K, Rao G, Meyerson M, Beroukhim R, Cooper L, Akbani R, Wrensch M, Haussler D, Aldape KD, Laird PW, Gutmann DH, Network TR, Noushmehr H, Iavarone A and Verhaak RG (2016) Molecular Profiling Reveals Biologically Discrete Subsets and Pathways of Progression in Diffuse Glioma. Cell 164:550-63. doi: 10.1016/j.cell.2015.12.028
19.Tang Z, Kang B, Li C, Chen T and Zhang Z (2019) GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. Nucleic Acids Res 47:W556-W560. doi: 10.1093/nar/gkz430
20.Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, Antipin Y, Reva B, Goldberg AP, Sander C and Schultz N (2012) The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov 2:401-4. doi: 10.1158/2159-8290.CD-12-0095
21.Li T, Fan J, Wang B, Traugh N, Chen Q, Liu JS, Li B and Liu XS (2017) TIMER: A Web Server for Comprehensive Analysis of Tumor-Infiltrating Immune Cells. Cancer Res 77:e108-e110. doi: 10.1158/0008-5472.CAN-17-0307
22.Huang da W, Sherman BT and Lempicki RA (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res 37:1-13. doi: 10.1093/nar/gkn923
23.Yan T, Skaftnesmo KO, Leiss L, Sleire L, Wang J, Li X and Enger PO (2011) Neuronal markers are expressed in human gliomas and NSE knockdown sensitizes glioblastoma cells to radiotherapy and temozolomide. BMC Cancer 11:524. doi: 10.1186/1471-2407-11-524
24.Eckel-Passow JE, Lachance DH, Molinaro AM, Walsh KM, Decker PA, Sicotte H, Pekmezci M, Rice T, Kosel ML, Smirnov IV, Sarkar G, Caron AA, Kollmeyer TM, Praska CE, Chada AR, Halder C, Hansen HM, McCoy LS, Bracci PM, Marshall R, Zheng S, Reis GF, Pico AR, O'Neill BP, Buckner JC, Giannini C, Huse JT, Perry A, Tihan T, Berger MS, Chang SM, Prados MD, Wiemels J, Wiencke JK, Wrensch MR and Jenkins RB (2015) Glioma Groups Based on 1p/19q, IDH, and TERT Promoter Mutations in Tumors. N Engl J Med 372:2499-508. doi: 10.1056/NEJMoa1407279
25.Ducray F, Marie Y and Sanson M (2009) IDH1 and IDH2 mutations in gliomas. N Engl J Med 360:2248-9; author reply 2249.
26.Zhou X, Li G, An S, Li WX, Yang H, Guo Y, Dai Z, Dai S, Zheng J, Huang J, Iavarone A and Zhao X (2018) A new method of identifying glioblastoma subtypes and creation of corresponding animal models. Oncogene 37:4781-4791. doi: 10.1038/s41388-018-0305-1
27.RNA Vaccines: Methods and Protocols.
28.Sharanek A, Burban A, Laaper M, Heckel E, Joyal JS, Soleimani VD and Jahani-Asl A (2020) OSMR controls glioma stem cell respiration and confers resistance of glioblastoma to ionizing radiation. Nat Commun 11:4116. doi: 10.1038/s41467-020-17885-z
29.Peng Y, Jia J, Jiang Z, Huang D, Jiang Y and Li Y (2018) Oncogenic DIRAS3 promotes malignant phenotypes of glioma by activating EGFR-AKT signaling. Biochem Biophys Res Commun 505:413-418. doi: 10.1016/j.bbrc.2018.09.119
30.Chen J, Zhang J, Hong L and Zhou Y (2019) EGFLAM correlates with cell proliferation, migration, invasion and poor prognosis in glioblastoma. Cancer Biomark 24:343-350. doi: 10.3233/CBM-181740
31.Xu L, Li Z, Song S, Chen Q, Mo L, Wang C, Fan W, Yan Y, Tong X and Yan H (2020) Downregulation of alpha-l-fucosidase 1 suppresses glioma progression by enhancing autophagy and inhibiting macrophage infiltration. Cancer Sci 111:2284-2296. doi: 10.1111/cas.14427
32.Close HJ, Stead LF, Nsengimana J, Reilly KA, Droop A, Wurdak H, Mathew RK, Corns R, Newton-Bishop J, Melcher AA, Short SC, Cook GP and Wilson EB (2020) Expression profiling of single cells and patient cohorts identifies multiple immunosuppressive pathways and an altered NK cell phenotype in glioblastoma. Clin Exp Immunol 200:33-44. doi: 10.1111/cei.13403
33.Wesdorp M, Murillo-Cuesta S, Peters T, Celaya AM, Oonk A, Schraders M, Oostrik J, Gomez-Rosas E, Beynon AJ, Hartel BP, Okkersen K, Koenen H, Weeda J, Lelieveld S, Voermans NC, Joosten I, Hoyng CB, Lichtner P, Kunst HPM, Feenstra I, de Bruijn SE, Admiraal RJC, Yntema HG, van Wijk E, Del Castillo I, Serra P, Varela-Nieto I, Pennings RJE and Kremer H (2018) MPZL2, Encoding the Epithelial Junctional Protein Myelin Protein Zero-like 2, Is Essential for Hearing in Man and Mouse. Am J Hum Genet 103:74-88. doi: 10.1016/j.ajhg.2018.05.011
34.Xiong Y and Wang Q (2019) STC1 regulates glioblastoma migration and invasion via the TGF‐ β/SMAD4 signaling pathway. Mol Med Rep 20:3055-3064. doi: 10.3892/mmr.2019.10579

35.Ceccarelli M, Barthel FP, Malta TM, Sabedot TS, Salama SR, Murray BA, Morozova O, Newton Y, Radenbaugh A, Pagnotta SM, Anjum S, Wang J, Manyam G, Zoppoli P, Ling S, Rao AA, Grifford M, Cherniack AD, Zhang H, Poisson L, Carlotti CG, Jr., Tirapelli DP, Rao A, Mikkelsen T, Lau CC, Yung WK, Rabadan R, Huse J, Brat DJ, Lehman NL, Barnholtz-Sloan JS, Zheng S, Hess K, Rao G, Meyerson M, Beroukhim R, Cooper L, Akbani R, Wrensch M, Haussler D, Aldape KD, Laird PW, Gutmann DH, Noushmehr H, Iavarone A and Verhaak RG (2016) Molecular Profiling Reveals Biologically Discrete Subsets and Pathways of Progression in Diffuse Glioma. Cell 164:550-63. doi: 10.1016/j.cell.2015.12.028

36.Martinez-Lage M, Lynch TM, Bi Y, Cocito C, Way GP, Pal S, Haller J, Yan RE, Ziober A, Nguyen A, Kandpal M, O'Rourke DM, Greenfield JP, Greene CS, Davuluri RV and Dahmane N (2019) Immune landscapes associated with different glioblastoma molecular subtypes. Acta Neuropathol Commun 7:203. doi: 10.1186/s40478-019-0803-6

Abbreviations: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.**Acknowledgement**s:GBM: Glioblastoma Multiforme; IDH1: Isocitrate Dehydrogenase (NADP(+)) 1; PAAD: Pancreatic Adenocarcinoma; CHOL: Cholangiocarcinoma; APC: Antigen Presenting Cell; GEPIA2: Gene Expression Profiling Interactive Analysis version 2; TCGA: The Cancer Genome Atlas; GTEx: Genotype-Tissue Expression; ANOVA: Analysis of Variance; OS: Overall survival; RFS: Relapse-free survival; cBioPortal: cBio Cancer Genomics Portal; TIMER: Tumor Immune Estimation Resource; CGGA: Chinese Glioma Genome Atlas; DAVID: Database for Annotation, Visualization and Integrated Discovery; ssGSEA: single-sample GSEA; DIRAS3: DIRAS family GTPase 3; FUCA1: alpha-L-fucosidase 1; LILRB2: Leukocyte Immunoglobulin Like Receptor B2; MPZL2: Myelin Protein Zero Like 2; OSMR: Oncostatin M Receptor; STC1: Stanniocalcin 1; EGFLAM: EGF like, fibronectin type III and laminin G domains; TMB: Tumor mutation burden; ICPs: Immune checkpoints; ICD: Immunogenic cell death; NSE: Neuron-specific enolase.**Acknowledgments:**The authors would like to thank CGGA (https://www.cgga.org.cn/tools.jsp) and TCGA (https://www.cancer.gov/tcga) for data collection, as well as GEPIA2 (https://gepia2.cancer-pku.cn), cBioPortal (https://www.cbioportal.org) and TIMER (https://cistrome.shinyapps.io/timer) for the provision of data processing and customizable functions. We also thank International Science Editing (https://www.internationalscienceediting.com) for editing this manuscript.**Authors' contributions:**J.M.C. and X.P.H. conceptualized and designed the project. J.M.C., X.P.H. and Y.Q.L. supervised the project. Y.L., and Z.Y.L. collected the omics data. Y.L., X.Y.X., Z.Y.L. and Y.B. conducted the analysis. Y.L., H.B.C., J.M., Z.L. and X.P.H. produced the diagrams. Y.L., X.Y.X., Z.Y.L., Y.Q.L., S.Q.A. and J.M.C. interpreted the data and results. Y.L., X.Y.X., Z.Y.L., Y.Q.L., S.Q.A. and J.M.C. wrote the draft of the manuscript. Y.L., X.Y.X. and Z.Y.L. contributed equally to this work. Y.Q.L., S.Q.A. and J.M.C. share the senior authorship of this study. All authors reviewed and approved the final manuscript.**Funding:**This research was supported by the National Natural Science Foundation of China (Grant No.: 82160389 to Sanqi An), Guangxi Medical University Training Program for Distinguished Young Scholars to Sanqi An, Guangxi Science and Technology Base and Talent Project (Grant No.: 2022AC19006 to Sanqi An).**Availability of data and materials:**All data generated and described in this article are available from the corresponding web servers, and are freely available to any scientist wishing to use them for noncommercial purposes, without breaching participant confidentiality. Further information is available from the corresponding author on reasonable request.**Ethics approval and consent to participate:**Not applicable.**Consent for publication:**Not applicable.**Competing interests:**The authors declare no competing financial interest.