m^{6}A function and tumor microenvironment infiltration characteristics in glioma

Aoran Yang
Department of Neurosurgery, Shengjing Hospital, China Medical University
Xinhuan Wang
Department of Neurosurgery, Shengjing Hospital, China Medical University
Chao Shang
Department of Neurobiology, School of Life Science, China Medical University
Yaofeng Hu
Department of Neurosurgery, Shengjing Hospital, China Medical University
Chen Pan
Department of Neurosurgery, Shengjing Hospital, China Medical University
Yang Hong (hongy@sj-hospital.org)
Department of Neurosurgery, Shengjing Hospital, China Medical University

Research Article

Keywords: m6A, TME, glioma, bioinformatics, PCA

Posted Date: December 13th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1142715/v1

Abstract

Background

Most studies related to N^6-methyladenosine (m^6A) in glioma have been conducted only at the single-gene level, while the m^6A modification patterns and correlation between m^6A and immune cells' tumor microenvironment infiltration have not been comprehensively reported.

Methods

Eighty-five clinical samples were obtained from the Gene Expression Omnibus; another 599 clinical samples and 174 transcriptome data were obtained from The Cancer Genome Atlas. Based on data analysis using the R program and principal component analysis, we established an m6Ascore quantitative scoring system and analyzed the modification of 22 m^6A-related genes with TME infiltration characteristics.

Results

Three m^6A modification patterns and the characteristics of immune cell infiltration were identified. Cluster A corresponded to the immune-desert phenotype, cluster B corresponded to the immune-excluded phenotype, and cluster C corresponded to the immune-inflamed phenotype. Cluster B showed the worst long-term prognosis, whereas cluster A had the best long-term prognosis. Anti-CD52/HE5 shows potential as an immune therapy for glioma.

Conclusions

This research provides a foundation for understanding m^6A modification patterns in glioma and a potential prognostic biomarker and immune therapy target for treat glioma.

1 Background

Four common internal modifications of mRNA have been identified: m^6A (N^6-methyladenosine), N^1-methyladenosine (m^1A), 5-methylcytosine (m^5C), and internal 7-methylguanosine (m^7G) [1]. m^6A is a common base modification that maintains mRNA and is performed by different type of enzymes. Methyltransferases of m^6A are known as readers that modify adenylates in mRNA by adding m^6A; genes such as METTL3 and METTL14 function as a complex in this catalysis [2]. Demethylases in m^6A are known as erasers and include ALKBH5 and FTO; these enzymes function by demethylating bases modified by m^6A. Studies have shown [3, 4] that FTO and ALKBH5 are closely associated with diverse types of tumors. Reader proteins specifically bind to the m^6A methylation region to affect the interaction

between proteins and RNA, which can alter the RNA secondary structure and weaken the binding of RNA-binding proteins [5–7]. In this research, 8 writers, 12 readers, and 2 erasers related to m^6A were comprehensively evaluated.

m^6A modification is related to numerous diseases. In tumors [8], the expression of enzymes that modify m^6A differs from that in normal cells. Additionally, environment changes and mutations can change the m^6A status.

Glioma is the most fatal intracranial tumor. Treatment strategies include excision, chemotherapy, and radiotherapy; patients with low-grade glioma subjected to aggressive treatment generally show a good prognosis [9, 10], whereas those with high-grade glioma do not. Although treatment prolong patient survival, there are also side effects. Surgery is often ineffective for removing tumor lesions in high-grade glioma, and normal brain tissue can be damaged during surgery [11]. Focal irradiation for glioma [12] can cause side effects such as radiation-induced necrosis. Chemotherapy using temozolomide, the most well-known specific anti-tumor drug, can only be used in patients in whom the promoter of O^6-methylguanine-DNA methyltransferase is methylated [13, 14].

Because of the development of drug tolerance, function of the blood-brain barrier, and high recurrence of glioma, treatment of patients with high-grade glioma typically involves palliative therapy, reducing the quality of life of patients. Numerous molecular-level studies have been performed to develop treatments for glioma. m^6A modification is closely associated with the progression and invasion of glioma [15]. Studies [16, 17] revealed that dysfunction of writers, readers, and erasers can regulate downstream proteins or target genes and thus affect glioma progression by altering the tumor microenvironment. Immunotherapy is a potential solution for treating glioma [18]. However, the m^6A-related glioma microenvironment and its correlation with immunity are unclear.

In this study, using bioinformatic tools, we focused on m^6A modification patterns in mRNA and comprehensively analyzed the function of m^6A biological in glioma. Transcriptome and clinical data were analyzed to determine the expression status of m^6A regulators and differences among immune cells. We also identified differentially expressed genes (DEGs) and predicted their biological functions.

Although studies [19, 20] have reported the curative effect of immune checkpoint blockade of CTLA4 and PD-1 in many types of tumors such as melanoma and prostate cancer, its curative effect in glioma is unknown. We investigated the value of immunotherapy for treating glioma and potential value of immune checkpoint blockade of CTLA4 and PD-1. We also identified a potential immunotherapy target in glioma. We determined the m^6A modification patterns, analyzed m^6A immune correlation, and predicted a potential immunotherapy target in glioma. This study provides insight into treating glioma and predicting the survival of patients with glioma.

2 Materials And Methods

2.1 Datum collecting and processing

Transcriptome data (n = 174) were obtained from the National Cancer Institute (NIH, Bethesda, MD, USA) GDC Data Portal. From the transcriptome profiles, we selected the gene expression quantification data with HTSeq-FPKM as the workflow type. We condensed the data using PERL and transformed the data into a list with the row showing the gene id and column showing the sample name. A total of 599 of clinical datasets was also downloaded from TCGA (The Cancer Genome Atlas) in “bcr xml” format. Clinical datasets were processed with PERL to build a fragments per kilobase million (FPKM) expression matrix with the rows showing the sample name and columns showing the clinical information. GEO (Gene Expression Omnibus) clinical data were obtained from dataset GSE4412 which was generated on the platform GPL96 [HG-U133A] Affymetrix Human Genome U133A Array (Santa Clara, CA, USA). The platform data were combined with GSE4412 data to transform the probe matrix into a gene matrix. Mutation data were downloaded from TCGA database with varscan2 variant aggregation and masking as the workflow type. To evaluate copy number variations in glioma, we downloaded data from UCSC Xena [21], which were processed with PERL to evaluate combinations of copy number data and m^6A-related gene data. Using the limma and sva packages in Bioconductor, we combined the TCGA and GEO data, which were transformed into m^6A expression data with the limma (42) package.

2.2 m6A-related gene phenotypes and mutation characteristics

Based on R, we transformed the processed data into a coordinate diagram. To distinguish differential of copy number variation frequencies, we used two colors to mark the 8 writers *METTLE3*, *METTLE14*, *METTLE16*, *WTAP*, *VIRMA*, *ZC3H13*, *RBM15*, and *RM15B*; 12 readers *YTHDC1*, *YTHDC2*, *YTHDF1*, *YTHDF2*, *YTHDF3*, *HNRNPC*, *FMR1*, *HNRNPA2B1*, *IGF2PBP1*, *IGF2PBP2*, *IGF2PBP3*, and *RBMX*; and 2 erasers *FTO* and *ALKBH5*. A red color represents increases in copy numbers of m^6A regulators, whereas the green color represents decreases in copy numbers. To identify the specific chromosome site containing m^6A regulators, we visualized the relationships among genes and their location in a gene loop diagram. RCircos [22] package in R was used to match regulators with their target chromosomes and generate the gene loop diagram. Based on TCGA datasets, we examined the expression differences in m^6A regulators between normal and tumor samples using R. The Limma package in Bioconductor, and reshape2 and ggpubr in R were used to calculate the p value; the results were visualized in a box plot. Mutation characteristic of m^6A regulators in glioma were further evaluated by establishing a waterfall plot. The mutation frequencies were determined as percentages. Based on the mutation frequency, each regulator was ranked and visualized using maftools in Bioconductor.

2.3 Association of m6A-related genes and overall survival and prognosis

TCGA and GEO data were analyzed to determine patient survival time and status. Survival-related data and merged data were analyzed using the packages limma, survival, and survminer. Genes were

considered as continuous variables and compared with the survival time and status by Cox regression analysis. Kaplan-Meier analysis was also conducted for survival analysis. Under best cut-off value, we separated each gene into two different expression status groups based on $p < 0.05$, each the relationship between each gene and survival value was confirmed.

2.4 Interactions of m6A-related genes and identification of risk-related genes

In this study, we identified cooperative relationships among these regulators and categorized based on their correlations. Based on Cox regression analysis of m^6A-related genes, we determined the risk of glioma for each m^6A regulator. We visualized the consequence by constructing a network diagram and profiled the interaction of m^6A-related gene using the packages igraph, psych, reshape2, and RColorBrewer in R.

2.5 m6A-related gene cluster analysis and immune correlation analysis

By using the ConsensusClusterPlus package in Bioconductor, we clustered m^6A-related genes and separated them into different groups based on their expression status; the groups were distinguished by labeling with different letters (A–C). Clusters were formed using three criteria as follows. The area under the curve of the cumulative distribution function did not increase significantly. In addition, the internal correlation of classification was closely related. Finally, samples in classifications should be abundant. To determine the survival differences among the three classifications, we performed overall survival analysis of m^6A-related gene clusters. The survival and survminer packages in R were used to draw survivorship curves. By constructing a heatmap with the pheatmap package in R, we visualized the expression status of genes in different classifications and clinical phenotypes including age, sex, and living status. Based on data dimension reduction, we applied PCA to evaluate obscure data while retaining information in the variables as much as possible and maintaining the number of variables as low as possible, helping to simplify the calculation. Principal component 1 and principal component 2 were used to score gene signatures. In the PCA diagram, the most relevant genes were casted onto the coordinate graph and interference by redundant genes was removed. The limma and ggplot2 packages were used to analyze and construct a scatter diagram.

To understand the immune infiltration patterns in different classifications, we analyzed the expression status of immune cells and calculated the differences in each immune cell along with the p value. Limma, GSEABase, and GSVA were used to visualize the differences by constructing a box plot.

2.6 GSVA and enrichment analysis of m6A-related genes

Because m^6A-related genes have different biological functions and are involved in different pathways, we conducted GSVA to evaluate variations in these factors in the different classifications. The three classifications were compared in pairs, with the results visualized in a heatmap. Gene sets of c2.cp.kegg.

v7.4. symbols were downloaded from the MSigDB [23] database. The limma, GSEABase, GSVA, and pheatmap packages were used for comparison and heatmap configuration. Based on the results of GSVA, we identified the intersection of the three classifications and abstracted m^{6}A-related DEGs. A Venn diagram showing the intersection was constructed using the limma and VennDiagram packages. To determine the specific biological functions and sites in which m^{6}A-related genes are enriched, we conducted GO enrichment analysis. The colorspace, stringi, and ggplot2 packages in R and org.Hs.eg.db, DOSE, clusterProfiler, and enrichplot packages in Bioconductor were used for analysis and to draw the bar plot and bubble diagram showing the results of GO enrichment analysis.

2.7 Evaluation of prognosis-related genes

Based on the identified DEGs, we examined the relationship among the DEGs, survival time, and survival status and identified prognosis-related genes. Cox regression analysis was conducted to identify correlations between the risk of glioma and prognosis-related genes. The limma and ConsensusClusterPlus packages were utilized for univariate Cox regression analysis and to show the results. Survival analysis was performed using the survival and survminer packages to detect differences in the survival status of the three groups. To illustrate the relationships among prognostic genes, clinical phenotypes, and their expression in m^{6}A clusters, we constructed a heatmap using the pheatmap package. Through variation analysis, we observed the expression status of m^{6}A-related genes in different groups of prognosis-related genes and drew a box plot with the limma, reshape, and ggpubr packages.

2.8 Quantification of m6Ascore and evaluation of TME infiltration characteristics

Based on PCA, we quantified the expression level of prognosis-related genes, and then defined differential expression levels based on the m6Ascore, which was determined as follows:

$$m6Ascore = \sum (PC1i + PC2i)$$

Survival analysis was performed to determine the differences between two groups busing the survival and survminer packages in R. The survival rate was determined for different clinical phenotypes such as sex and age. To evaluate the association between the m6Ascore and survival status, we compared the low and high m6Ascore groups using the ggpubr and ggplot2 packages and p-values. The results were shown in bar and box plots. Using the information on m^{6}A-related genes, prognosis-related genes, m6Ascore groups, and clinical phenotypes data, we used the ggplot2 and ggalluvial packages to draw a Sankey diagram to view data from all groups.

To assess differences in gene mutations between the high and low m6Ascore groups, we used the maftools package in Bioconductor. We marked high m6Ascore group with red whereas low m6Ascore group was marked with blue. Gene mutation frequencies were determined as percentages and visualized by constructing a waterfall plot.

To assess relationships between immune cells and the m6Ascore, we performed correlation analysis with the corrplot package in R. We distinguished positive and negative correlations by marking them in red from blue, respectively.

2.9 Prediction of immunotherapy target and assessment of dual immunotherapy

Using the merged TCGA and GEO data, we identified potential immunotherapy targets for treating glioma. We selected CD52/HE5 as the target immune checkpoint. According to the m6Ascore, we analyzed expression variations between the high and low score groups using the limma and ggpubr packages, with the results shown in a box plot. From TCIA database [24], we obtained data correlated with immunotherapy in glioma. Based on the m6Ascore level, we grouped the data. Based on $p < 0.05$, we predicted the group with lower m6Ascore showed better curatives effects from immunotherapy. The result was visualized in a violin diagram drawn using the ggpubr package in R.

3 Results

3.1 Genetic variations in m^6A-related genes

To evaluate variations in the copy numbers of m^6A-related genes, we showed the results in a bar graph (Fig. 1A). The abscissa represents m^6A-regulated genes, whereas the ordinate represents the frequencies of copy number variations in each gene. Various colors represent the copy number of each gene. Green indicates the loss of copy numbers, whereas red represents a gain in copy numbers. This graph provides an overall perspective of glioma, in which most m^6A-related genes lost copy numbers and gained a few copy numbers of genes, suggesting that the expression of m6A-related genes is low in glioma.

To identify the locations of m^6A-related genes in human chromosomes, we drew a gene loop diagram (Fig. 1B). Each fraction of this gene loop represents the number of human chromosomes. Genes were distinguished based on their expression status. Red colors represent a high expression status whereas a blue color indicates a gene with low expression in glioma.

The results shown in the copy number variation frequency diagram reveal the differential expression status of each m^6A-related gene. To determine the differential expression of each gene, the p value was calculated and the results were shown in a box plot (Fig. 1C). We separated m^6A-related genes into two groups and evaluated their expression in both normal samples and tumor samples. The expression status in the tumor group is marked in red, whereas that in the normal group is marked blue. Most genes with asterisks showed higher expression in tumor tissues than in normal tissues.

We analyzed the downloaded mutation data from TCGA to determine the mutation status of m^6A-related genes in glioma by constructing a waterfall plot (Fig. 1D). m^6A-related genes showed a low mutation level

in glioma with a ratio of 0 to 0.02; the most common mutations were missense mutations. The mostly altered bases in glioma were C to T transitions.

3.2 Landscape of m^{6}A-related genes and relationship with survival and prognosis

Based on Cox regression analysis, we calculated the association between m^{6}A-related genes and glioma. The results of co-regression analysis are presented in Table 1, in which HR (hazard ratio) represents a gene's risk coefficient. If this risk coefficient was greater than one, the gene was considered as high-risk gene, whereas a risk coefficient lower than one indicated a low-risk gene.

Table 1
Univariate Cox regression analysis of m^{6}A-related genes based on HR value

Variables	**Univariate**				
	HR	HR95.L	HR95.H	p value	km
METTL3	0.784	0.606	1.013	0.063	0.002
WTAP	1.037	0.792	1.357	0.791	0.133
ZC3H13	0.818	0.656	1.020	0.075	0.0007
RBM15	0.955	0.685	1.332	0.788	0.092
RBM15B	0.877	0.682	1.129	0.311	0.103
YTHDC1	0.621	0.469	0.823	0.0009	0.0001
YTHDC2	0.782	0.552	1.108	0.167	0.002
YTHDF1	1.185	0.801	1.752	0.394	0.076
YTHDF2	1.284	0.859	1.919	0.221	0.048
YTHDF3	0.698	0.482	1.010	0.056	0.012
HNRNPC	0.494	0.335	0.729	0.0003	0.0007
FMR1	0.514	0.370	0.714	$7.161e^{-5}$	$2.95e^{-5}$
LRPPRC	0.582	0.396	0.855	0.005	0.0006
HNRNPA2B1	1.045	0.801	1.363	0.741	0.142
IGF2BP2	1.114	1.012	1.226	0.026	0.003
IGF2BP3	1.202	1.069	1.352	0.002	$6.82e^{-6}$
FTO	0.822	0.609	1.111	0.203	0.007

Prognostic evaluation was performed using the Kaplan-Meier method with cut-off criteria. High and low expression groups were evaluated, and the survival status was graphed. Graphs (Fig. 2) were drawn showing genes with $p < 0.05$, as these genes showed a prognostic difference between the high and low expression groups.

3.3 Phenotypes of m^6A-cluster and characteristics of immune cell infiltration

A previous study [25] revealed m^6A-related genes co-operate with each other to function as a complex. Based on univariate Cox regression analysis (Table 2), we constructed a correlation network of m^6A-related genes in R software to identify these correlation effects in glioma.

Table 2
Univariate Cox regression analysis of prognostic-related genes based on HR value

Variables	Univariate			
	HR	HR95.L	HR95.H	p value
IGF2BP2	1.114	1.012	1.226	0.026
IGFBP2	1.270	1.159	1.391	$2.577e^{-7}$
PLAT	1.328	1.180	1.494	$2.556e^{-6}$
MCAM	1.418	1.173	1.714	0.0003
NBR1	0.685	0.534	0.879	0.003
LMNB1	1.185	1.027	1.368	0.019
CENPF	1.157	1.021	1.312	0.021
ANKRD26	0.621	0.485	0.795	0.0001
WDFY3	0.692	0.532	0.901	0.006
FMR1	0.514	0.370	0.714	$7.161e^{-5}$
KAT6B	0.624	0.505	0.772	$1.314e^{-5}$
PDZD8	0.678	0.553	0.833	0.0002
HTATSF1	0.788	0.643	0.965	0.021
ALDH5A1	0.741	0.628	0.875	0.0004
GSE1	0.771	0.649	0.916	0.003
PBX1	0.633	0.504	0.794	$7.761e^{-5}$
LINC01278	0.666	0.484	0.916	0.012
ARHGAP5	0.691	0.552	0.865	0.001
SAT1	1.413	1.155	1.727	0.0007
SIRT1	0.546	0.414	0.718	$1.578e^{-5}$

As shown in the network diagram (Fig. 3A), high risk factors included *HNRNPA2B1, IGF2BP2, IGF2BP3, WTAP, YTHDF1, and YTHDF2, whereas favorable factors included METTL3, ZC3H13, RBM15, RBM15B, FTO, YTHDC1, YTHDC2, YTHDF3, HNRNPC, FMR1, and LRPPRC*. A negative correalation was observed between FTO and WTAP; other genes showed postive correlations with each other. We divided m^6A-related genes into three clusters based on the area under the curve of cumulative distribution and

samples' abundance. (Fig S1A–D). Subsequently, the survival status of each cluster was calculated. The result revealed that three different m^{6}A clusters exhibited distinct differences in the survival status, with cluster A showing better survival rates than clusters B and C ($p = 0.003$) (Fig. 3B). The principal component analysis (PCA) scatter diagram (Fig. 3C) clarified the differences among m^{6}A-related genes based on the clear boundaries among clusters A, B, and C. The heatmap (Fig. 3D) shows the expression status of m^{6}A-related genes in different clinical subgroups divided by age, sex, and living status. In cluster A, WTAP, IGF2BP2, and IGF2BP3 showed low expression, whereas in cluster C, WTAP and IGF2BP2 showed high expression. These results indicate that cluster A can better predict prognosis compared to cluster C because IGF2BP2 and IGF2BP3 expression is considered as a risk factor for glioma.

Differences in immune cell infiltration among m^{6}A clusters were identified. Immune cell infiltration was lower in cluster A than in clusters B and C. Immune cell infiltration was moderate in cluster B, whereas in cluster C, immune cell infiltration was high. This variation suggests that in cluster A, immunity is suppressed, whereas in cluster B, innate immunity is activated; in cluster C, adaptive immunity was activated (Fig. 3E).

3.4 Biological function and pathway analysis of m^{6}A-related genes

Based on Gene Ontology (GO) enrichment analysis and gene set variation analysis (GSVA), we predicted the molecular pathways affected by m^{6}A-related genes. The results of GSVA are shown as a heatmap (Fig. 4), whereas those of GO enrichment analysis are shown in supply files (Fig S2). In the heatmaps, we identified genes showing high and low expression. Cluster A was highly related to the calcium signaling pathway and metabolism of nutrients, whereas cluster B was highly related to the P53 signaling pathway and development of some tumors (Fig. 4A). Cluster B was activated in the Notch signaling pathway, cell cycle, basal transcription factor, spliceosome, RNA degeneration, and ubiquitin-mediated proteolysis; cluster C was highly related to the etiology of some diseases (Fig. 4B). Comparison between cluster A and C showed that the expression of most m^{6}A-related genes was low in cluster A, whereas expression was high in cluster C (Fig. 4C). GSVA was conducted and the output Venn diagram (Fig S2C) revealed 34 DEGs at the intersection of m^{6}A clusters. Subsequently, 34 DEGs were analyzed through GO analysis. The results of GO analysis were calculated as the p value are visualized in a bar plot and bubble plot (Fig S2A, S2B). In GO analysis, the biological characteristics of m^{6}A-related genes were determined in the biological process, cellular component, and molecular function categories. In the biological process category, m^{6}A-related genes were involved in acetylation of various amino acids, whereas in the cellular component category, the genes interacted with the nuclear envelope. In the molecular function section, m^{6}A-related genes were suggested to be related to histone binding.

3.5 Prognostic-related gene clinical correlation analysis based on DEGs

The DEGs indicated gene assemblies that are closely related to various m^6A clusters. Based on this assembly, we identified and divided m^6A prognostic-related genes into three clusters based on the area under the cumulative distribution function curve (Fig S3A-D). With the prognostic-related gene clusters, we predicted the biological roles of prognostic-related genes. The survival curve (Fig. 5A) revealed differences in survival among the three clusters ($p < 0.001$), suggesting that gene cluster A was associated with a better prognosis compared to gene clusters B and C. By analyzing each clinical phenotype, we found that most prognostic-related genes exhibited low expression in cluster C, whereas genes in cluster A showed high expression and cluster B genes showed moderate expression.

The analyzed the expression status of m^6A regulators by determining the p-values (Fig. 5C). As shown in Fig. 5C and Fig. 2, highly expressed genes in cluster A were related to good prognosis; prognosis-related genes in cluster C showed low expression, suggesting that cluster A was associated with better prognosis.

3.6 m6Ascore modification patterns in clinical traits and immnue cell TME inflitration

By quantifying m^6A prognostic-related genes by determining the m6Ascore, we analyzed m^6A modification characteristics and immune cell infiltration patterns in glioma (Fig. 6). Based on the level of the m6Ascore, we compared survival variation between groups with high and low m6Ascores. In the survival curve, a low m6Ascore was significantly associated with a better prognosis compared to a high m6Ascore ($p = 0.001$) (Fig. 6A). Analysis of clinical data supported these results. By analyzing patients' sex and age, we obtained four survival curves. The age data were evaluated for patients older and younger than or equal to 65 years. Survival analysis showed that patients <65 years old with a low m6Ascore (Fig S4A) had a better prognosis ($p = 0.011$) but no survival difference ($p = 0.074$) compared to patients older than 65 years (Fig S4B). Data on sex were also evaluated. The survival curve showed that females patients with low m6Ascores had a better prognosis ($p < 0.001$), whereas there was no survival difference in the male group between low m6Ascore group and high m6Ascore group (Fig S4C, S4D). A bar plot was drawn to show variations in the survival status of different m6Ascore groups (Fig. 6B). In the low m6Ascore group, the ratio of living to dead patients was approximately 3:7, whereas in the high m6Ascore group, this ratio was 1:9. The box plot showed that patient death was strongly associated with a high m6Ascore in glioma ($p = 0.0018$) (Fig. 6C). To determine the relationship among the m^6A gene clusters, prognostic gene clusters, m6Ascore, and survival status, we drew a Sankey diagram (Fig. 6D). This diagram showed that most m^6A- and prognosis-related genes in cluster A were closely correlated with a low m6Ascore and better prognosis. In the m^6A gene cluster and prognosis-related gene cluster, we determined the expression differences among clusters A–C and visualized the results in box plots. In the m^6A gene cluster (Fig. 6E), slight difference was detected between clusters A and B ($p = 0.76$), whereas prominent differences were observed between clusters A and C ($p < 1.9e^{-12}$) and clusters B and C ($p = 2.22e^{-16}$). In the prognosis-related gene cluster (Fig. 6F), obvious expression differences were observed between clusters A and B ($p < 0.0063$), clusters A and C ($p < 2.22e^{-16}$), and clusters B and C ($p <$

$2.22e^{-16}$). Both the m^6A gene cluster and prognosis-related cluster revealed that cluster A had a lower m6Ascore than the other clusters.

We investigated gene mutations and immune cell infiltration based on the m6Ascore by drawing heatmaps in R for the high m6Ascore group (Fig. 6G) and low m6Ascore group (Fig. 6H). Comparison both heatmaps revealed higher gene mutation frequencies in the high m6Ascore group; however, this group showed fewer mutation types compared to the low m6Ascore group. Immune cells exhibit different infiltration characteristics among different m^6A gene clusters. To estimate immune cell infiltration patterns, we used the m6Ascore to quantify immune cells. Using R, we analyzed the relationship between the m6Ascore and immune cells (Fig. 6I), which showed that most immune cells were positively correlated with a high m6Ascore.

3.7 Potential of immune checkpoint and immunotherapy in glioma

Immune checkpoint blockade is a new therapy that has been shown to be effective for treating tumors [26]. We found that CD52/HE5 earned higher m6Ascore in glioma ($p = 3.7e^{-14}$) and a high m6Ascore was associated with worse prognosis, as shown in the box plot in Fig. 7A. Thus, anti-CD52/HE5 may be useful for treating glioma. Subsequently, we analyzed the effects of dual immunotherapy based on R and the m6Ascore. The violin diagram revealed better curative effects of anti-PD-1 immunotherapy in the low m6Ascore group compared to in the high m6Ascore group ($p = 0.044$) (Fig. 7B).

4 Discussion

Glioma has been considered as the most dangerous tumor among intracranial tumors. Because curative treatments for glioma are lacking, numerous molecular-level studies have been carried out to understand this disease, revealing the molecular mechanisms of glioma from different perspectives. Studies of regulation of the blood-brain barrier [27, 28], competing endogenous RNA mechanism [29, 30], and glioma stem cell mechanisms [31, 32] have provided feasible methods for treating glioma and improving patient prognosis. m^6A modification has also been reported to interfere with glioma development and invasion. Various studies [33, 34] showed that m^6A is associated with glioma; however, most of these studies determined the relationship between a single m^6A regulator and glioma, and comprehensive studies of the relationship between m^6A-related genes and glioma are lacking.

It was previously thought that the central nervous system (CNS) has an immune-privileged status. However, researchers' studies revealed lymphatic vessels in the CNS [35, 36], suggesting that the CNS is under surveillance by the immune system, with the mechanism determined in 2021 [37]. CNS-derived antigens in the cerebrospinal fluid flowing through lymphatic vessels enter and accumulate in the dural sinus. Antigen-presenting cells capture these antigens and present them to T cells. Thus, immune surveillance does not occur within the brain tissue directly but rather at the margin of the brain.

Immune cells can lead to intracranial disease; however, the relationship of m^6A and immune cell infiltration and how these factors regulate the tumor microenvironment are not well-understood. Here, we evaluated three modification patterns of m^6A and analyzed immune cell infiltration. We found that in glioma, most m^6A regulators showed reduced copy numbers, which may have led to their low expression. The mutation frequency of m^6A regulators was low, suggesting that in the brain, most m^6A regulators are in a steady state, whereas in other cancers such as gastric cancer, prominent gene [38] mutations are observed. However, even low-frequency gene mutations can lead to heterogeneity. Based on survival analysis, the different expression status of m^6A regulators can lead to different prognoses, and thus m^6A regulators can be divided into oncogenes and tumor suppressor genes. Studies [39, 40] showed that changes in the expression level of oncogenes and gene suppressor genes can affect the growth and invasion of glioma cells.

In this research, we constructed three different m^6A modification patterns, clusters A–C, which were used to analyze immune cell infiltration. Unexpectedly, most immune cells including B cells, CD4 T cells, CD8 T cells, dendritic cells, CD56 bright natural killer cells, CD56 dim natural killer cells, eosinophils, gamma delta T cells, immature B cells, immature dendritic cells, myeloid-derived suppressor cells, macrophages, mast cells, monocytes, natural kill T cells, neutrophils, plasmacytoid dendritic cells, regulatory T cells, T follicular helper cells, type 17 T helper cells, and type 1 T helper cells were highly infiltrated in cluster C. Cluster A was considered to suppress immunity and to be related to an immune-desert phenotype. Cluster B can be considered to activate innate immunity, which is related to an immune-excluded phenotype, and cluster C can be considered to activate adaptive immunity, which is related to an immune-inflamed phenotype. However, unexpectedly, cluster C was not associated with a better prognosis corresponding to its high immune cell infiltration, whereas cluster A, which lacked immune cell infiltration, was associated with the best prognosis.

Immune cells can kill tumor cells via their cytotoxic effects or phagocytosis. Traditional analysis also showed that high immune infiltration is effective for combating glioma[41]; however, immune infiltration may act as a double-edged sword in glioma. Cluster C was highly related to autoimmune diseases such as lupus erythematosus, allograft rejection, and autoimmune thyroid disease. This result may be explained the high immune infiltration in cluster C, and clusters B and C were found to closely interact with extracellular matrix (ECM) receptors, whereas cluster A did not. Thus, to kill tumor cells, immune cells must pass through the collagen network of the ECM, the killing effect of immune cells is highly related to their migration speed; a dense collagen network may weaken immune cell function [42]. Research [43] has shown that remodeling and deposition of the ECM can create a stiffer environment in which tumor cells reside. Tumor cells are protected by fibrocytes and collagen cells; to eliminate these structures, immune cells must gather and attack (high infiltration status), during which the ECM is consistently compressed, causing nucleus and membrane damage and immune cell depletion [44]. This progress may damage the normal brain tissue and accelerate the invasion of glioma cells, explaining the poor prognosis in cases with high immune cell infiltration. Furthermore, the unique microenvironment may be regulated by the blood-brain barrier, and ECM stiffness in glioma may be higher than that for other tumors

in the body. Thus, high levels of immune infiltration may contribute to glioma growth and invasion by interfering with the ECM.

Moreover, quantifying the m6Ascore helped us evaluated patient prognosis, gene mutation frequencies, and immune cell correlations and can be considered as an independent prognostic biomarker for glioma. Additionally, based on the m6Ascore, we predicted that anti-CD52/HE5 can be used to treat glioma.

There were some limitations to this study. In m6Ascore-related survival analysis, because patients grouped by age and sex were imbalanced, survival differences differed between the male and female groups and in patients older and younger than 65 years. Moreover, samples from Gene Expression Omnibus (GEO) were astrocytoma-related, and thus the description of relationship between m^6A and glioma may be incomplete. Data regarding microsatellite instability in the glioma tumor microenvironment were also lacking and require further analysis. The process of the application of our results in the clinic to improve patient survival needs to also be examined.

5 Conclusion

We identified the m^6A modification patterns in glioma and immune cell infiltration in the tumor microenvironment. We determined the prognosis status of each m^6A regulator and showed that the low m6Ascore group has a better long-term prognosis. We also showed that the m6Ascore is a reliable biomarker for evaluating the prognosis of patients with glioma. This comprehensive analysis provides a foundation for further understanding glioma.

6 List Of Abbreviations

m^6A
N^6-methyladenosine
m^1A
N^1-methyladenosine
m^5C
5-methylcytosine
m^7G
7-methylguanosine
DEGs
differentially expressed genes
TCGA
The Cancer Genome Atlas
GEO
Gene Expression Omnibus
PCA

principal component analysis
GO
Gene Ontology
GSVA
gene set variation analysis
TME
tumor microenvironment
HR
hazard ratio
CNS
central nervous system

Declarations

7. Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

8. Author contributions

YA conceived the idea and drafted the manuscript. WX collected and preprocessed data from TCGA and GEO database. PC and HYF processed the format of images and reviewed the data and manuscript. SC and HY verified the experimental dataset and revised the manuscript.

9. Funding

This study was funded by the National Nature Science Foundation of China (No. 81872067) and Natural Science Foundation of Liaoning Province (No. 20180530024).

10. Data Availability Statement

The GSE4412 dataset analyzed in this study can be found in the Gene Expression Omnibus [https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi].

Transcriptome data and 599 clinical sample datasets analyzed in this study can be found in The Cancer Genome Atlas Genomic Data Commons Data Portal [https://portal.gdc.cancer.gov/].

11. Consent for publication

All the authors have read and approved the final version of manuscript and give consent.

References

1. Barbieri, I. & Kouzarides, T. Role of RNA modifications in cancer. *Nat Rev Cancer*, **20** (6), 303–22 (2020).
2. Wang, X. *et al.* Structural basis of N(6)-adenosine methylation by the METTL3-METTL14 complex., **534** (7608), 575–8 (2016).
3. Chen, J. & Du, B. Novel positioning from obesity to cancer: FTO, an mA RNA demethylase, regulates tumour progression. *J Cancer Res Clin Oncol*, **145** (1), 19–29 (2019).
4. Hao, L. *et al.* ALKBH5-mediated mA demethylation of FOXM1 mRNA promotes progression of uveal melanoma. *Aging (Albany NY)*, **13** (3), 4045–62 (2021).
5. Xu, Y. *et al.* YTH Domain Proteins: A Family of mA Readers in Cancer Progression. *Front Oncol*, **11**, 629560 (2021).
6. Zheng, G. *et al.* ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. *Mol Cell*, **49** (1), 18–29 (2013).
7. Liao, S., Sun, H. & Xu, C. YTH Domain: A Family of N-methyladenosine (mA) Readers.Genomics Proteomics Bioinformatics. 2018;16(2)
8. Huang, H., Weng, H. & Chen, J. mA Modification in Coding and Non-coding RNAs: Roles and Therapeutic Implications in Cancer., **37** (3), 270–88 (2020).
9. Buckner, J. C. *et al.* Radiation plus Procarbazine, CCNU, and Vincristine in Low-Grade Glioma. *N Engl J Med*, **374** (14), 1344–55 (2016).
10. Jakola, A. S. *et al.* Surgical resection versus watchful waiting in low-grade gliomas. *Ann Oncol*, **28** (8), 1942–8 (2017).
11. Krivosheya, D., Prabhu, S. S., Weinberg, J. S. & Sawaya, R. Technical principles in glioma surgery and preoperative considerations. *J Neurooncol*, **130** (2), 243–52 (2016).
12. Yang, I., Huh, N. G., Smith, Z. A., Han, S. J. & Parsa, A. T. Distinguishing glioma recurrence from treatment effect after radiochemotherapy and immunotherapy. *Neurosurg Clin N Am*, **21** (1), 181–6 (2010).
13. Whitelaw, B. C. How and when to use temozolomide to treat aggressive pituitary tumours. *Endocr Relat Cancer*, **26** (9), R545–R52 (2019).
14. Bell, E. H. *et al.* Association of MGMT Promoter Methylation Status With Survival Outcomes in Patients With High-Risk Glioma Treated With Radiotherapy and Temozolomide: An Analysis From the NRG Oncology/RTOG 0424 Trial. *JAMA Oncol*, **4** (10), 1405–9 (2018).
15. Lin, S. *et al.* Prognosis Analysis and Validation of mA Signature and Tumor Immune Microenvironment in Glioma. *Front Oncol*, **10**, 541401 (2020).
16. Chang, Y-Z. *et al.* METTL3 enhances the stability of MALAT1 with the assistance of HuR via m6A modification and activates NF-κB to promote the malignant progression of IDH-wildtype glioma. *Cancer Lett*, **511**, 36–46 (2021).
17. Li, F., Zhang, C. & Zhang, G. m6A RNA Methylation Controls Proliferation of Human Glioma Cells by Influencing Cell Apoptosis. *Cytogenet Genome Res*, **159** (3), 119–25 (2019).

18. Lim, M., Xia, Y., Bettegowda, C. & Weller, M. Current state of immunotherapy for glioblastoma. *Nat Rev Clin Oncol*, **15** (7), 422–42 (2018).
19. Anti-PD-. 1-CTLA4 Combo Hits Prostate Cancer. *Cancer Discov*, **9** (5), 569–70 (2019).
20. LAG3-PD-1 Inhibitor Combo Impresses in Melanoma.Cancer Discov. 2021
21. Goldman, M. J. *et al.* Visualizing and interpreting cancer genomics data via the Xena platform. *Nat Biotechnol*, **38** (6), 675–8 (2020).
22. Zhang, H., Meltzer, P. & Davis, S. RCircos: an R package for Circos 2D track plots. *BMC Bioinformatics*, **14**, 244 (2013).
23. Liberzon, A. *et al.* The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst*, **1** (6), 417–25 (2015).
24. Prior, F. W. *et al.* TCIA: An information resource to enable open science. *Annu Int Conf IEEE Eng Med Biol Soc*, **2013**, 1282–5 (2013).
25. Schöller, E. *et al.* Interactions, localization, and phosphorylation of the mA generating METTL3-METTL14-WTAP complex. *RNA*, **24** (4), 499–512 (2018).
26. Topalian, S. L., Taube, J. M., Anders, R. A. & Pardoll, D. M. Mechanism-driven biomarkers to guide immune checkpoint blockade in cancer therapy. *Nat Rev Cancer*, **16** (5), 275–87 (2016).
27. Carpentier, A. *et al.* Clinical trial of blood-brain barrier disruption by pulsed ultrasound. *Sci Transl Med*, **8** (343), 3432 (2016).
28. Wang, Y. *et al.* Remodelling and Treatment of the Blood-Brain Barrier in Glioma. *Cancer Manag Res*, **13**, 4217–32 (2021).
29. Chen, Y. *et al.* HOXD-AS1/miR-130a sponge regulates glioma development by targeting E2F8. *Int J Cancer*, **142** (11), 2313–22 (2018).
30. Shang, C., Tang, W., Pan, C., Hu, X. & Hong, Y. Long non-coding RNA TUSC7 inhibits temozolomide resistance by targeting miR-10a in glioblastoma. *Cancer Chemother Pharmacol*, **81** (4), 671–8 (2018).
31. Tabu, K. *et al.* Glioma stem cell (GSC)-derived autoschizis-like products confer GSC niche properties involving M1-like tumor-associated macrophages., **38** (8), 921–35 (2020).
32. Sun, Z. *et al.* Glioblastoma Stem Cell-Derived Exosomes Enhance Stemness and Tumorigenicity of Glioma Cells by Transferring Notch1 Protein. *Cell Mol Neurobiol*, **40** (5), 767–84 (2020).
33. R-2HG Targets FTO to Increase mALevels and Suppress Tumor Growth. Cancer Discov. 2018; 8(2):137
34. Zhang, S. *et al.* mA Demethylase ALKBH5 Maintains Tumorigenicity of Glioblastoma Stem-like Cells by Sustaining FOXM1 Expression and Cell Proliferation Program.Cancer Cell. 2017;31(4)
35. Louveau, A. *et al.* Structural and functional features of central nervous system lymphatic vessels., **523** (7560), 337–41 (2015).
36. Louveau, A. *et al.* Corrigendum: Structural and functional features of central nervous system lymphatic vessels., **533** (7602), 278 (2016).

37. Rustenhoven, J. *et al.* Functional characterization of the dural sinuses as a neuroimmune interface.Cell. 2021;184(4)
38. Zhang, B. *et al.* mA regulator-mediated methylation modification patterns and tumor microenvironment infiltration characterization in gastric cancer. *Mol Cancer*, **19** (1), 53 (2020).
39. Kowalski-Chauvel, A. *et al.* The m6A RNA Demethylase ALKBH5 Promotes Radioresistance and Invasion Capability of Glioma Stem Cells.Cancers (Basel). 2020;13(1)
40. Visvanathan, A., Patil, V., Abdulla, S., Hoheisel, J. D. & Somasundaram, K. Nâ¶-Methyladenosine Landscape of Glioma Stem-Like Cells: METTL3 Is Essential for the Expression of Actively Transcribed Genes and Sustenance of the Oncogenic Signaling.Genes (Basel). 2019;10(2)
41. Yang, J. *et al.* MiR-15a/16 deficiency enhances anti-tumor immunity of glioma-infiltrating CD8+ T cells through targeting mTOR. *Int J Cancer*, **141** (10), 2082–92 (2017).
42. Salmon, H. *et al.* Matrix architecture defines the preferential localization and migration of T cells into the stroma of human lung tumors. *J Clin Invest*, **122** (3), 899–910 (2012).
43. Di Modugno, F. *et al.* 3D models in the new era of immune oncology: focus on T cells, CAF and ECM. *J Exp Clin Cancer Res*, **38** (1), 117 (2019).
44. Moreau, J-F. *et al.* The emerging role of ECM crosslinking in T cell mobility as a hallmark of immunosenescence in humans. *Ageing Res Rev*, **35**, 322–35 (2017).

Figures

Figure 1

m6A-related gene phenotypes and mutation status. (A) Copy number variation frequency of each gene. IGF2BP2, HNRNPA2B1, FMR1, YTHDF1, RBMX, and VIRMA gained more copy numbers than were lost. IGF2BP3, YTHDF2, HNRNPC, ZC3H13, FTO, IGF2BP1, RBM15, METTLE14, WTAP, METTLE13, ALKBH5, METTLE16, YTHDC1, YTHDC2, YTHDC3, RBN15B, and LRPPRC lost more copy numbers than were gained. (B) Genes loop diagram. Each m6A-related gene chromosome site is marked on the loop. (C) m6A-related gene expression box plot. Significant differences in m6A-related gene expression between normal and tumor tissue identified based on the p-value. (D) Waterfall plot. In the upper graph, five gene mutations were detected, including missense mutation, frameshift mutation, splice site mutation, multisite mutation, and nonsense mutation. The grey grids represent non-mutation samples. In the lower graph, transformed bases are marked with different colors. Red represents a C to T transformation, green represents a T to A transformation, navy blue represents a C to G transformation, yellow represents a T to C transformation, blue represents a C to A transformation, and orange represents a T to G transformation. To identify the locations of m6A-related genes in human chromosomes, we drew a gene loop diagram (Fig 1B). Each fraction of this gene loop represents the number of human chromosomes. Genes were

distinguished based on their expression status. Red colors represent a high expression status whereas a blue color indicates a gene with low expression in glioma.

Figure 2

m6A-related gene survival differences. Better prognosis in glioma was suggested by (A) High expression of ZC3H13 ($p < 0.001$), (B) Low expression of YTHDF2 ($p = 0.048$), (C) High expression of YTFDH3 ($p = 0.012$), (D) High expression of LRPPRC ($p < 0.001$), (E) High expression of YTHDC1 ($p < 0.001$), (F) High expression of YTHDC2 ($p = 0.003$), (G) Low expression of IGF2BP2 ($p = 0.003$), (H) Low expression of IGF2BP3 ($p < 0.001$), (I) High expression of HNRNPC ($p < 0.001$), (J) Low expression of FMR1 ($p < 0.001$), (K) High expression of FTO ($p = 0.008$), and (L) High expression of METTL3 ($p = 0.003$).

Figure 3

m6A-related gene co-expression network and clinical and immune cell information for different m6A-related gene clusters. (A) Each m6A-related gene was presented as a circle, with the circle size based on the p-value determined in Cox regression analysis. Each left semicircle represents the subtype of m6A regulators, whereas the right semicircle represents the risk level of m6A regulators. Erasers are marked in red, reads are marked in orange, and writers are marked in gray. Risk factors are marked in purple, whereas favorable factors are marked in green. Postive co-expression relationships are shown in pink, whereas negative co-expression relationships are shown in blue. (B) Survival status of different m6A clusters. Survival status of cluster A was significantly better than that of clusters B and C ($p < 0.001$). (C) PCA scatter diagram. Samples in m6A clusters were included in the coordinate graph and m6A clusters were distinguished as A–C with clear boundaries. (D) Heatmap of m6A cluster clinical information. Data on sex, age, and living status were compared within different m6A clusters. (E) Box plot of immune cell differences within different m6A clusters based on the p-value.

Figure 4

Heatmaps showing comparison of m6A clusters. The blue section represents low gene expression and red represents high gene expression. Modification patterns of m6A-related genes were distinguished in each biological function of molecular pathway. (A) Comparison of m6A clusters A and B. (B) Comparison of m6A clusters B and C. (C) Comparsion of m6A clusters A and C.

Figure 5

m6A prognostic-related gene clinical correlation analysis. (A) Survival status of each gene cluster. Cluster A was associated with better long-term prognosis compared to clusters B and C ($p < 0.001$). (B)

Comparison of gene clusters in different clinical phenotypes. Blue represents lower expression and red represents higher expression. (C) m6A regulator differences among each gene cluster. Differences are shown in the box plot based on the p-value.

Figure 6

m6Ascore modification patterns and immune cells infiltration. (A) Survival comparison versus two groups. (B) Bar plot showing the relationship between m6Ascore status and prognosis. Abscissa represents m6Ascore and ordinate represents percent weight. Blue area represents survival and red area represents death. (C) Box plot. Abscissa indicates survival and ordinate represents the m6Ascore. (D) Sankey diagram. Four columns showing relationships among m6A cluster, gene cluster, m6Ascore, and survival status. (E) Box plot showing the relationship among m6A-related genes clusters. Abscissa represents m6A cluster and ordinate represents m6Ascore. (F) Box plot showing the relationship among prognosis-related gene clusters. Abscissa represents prognosis-related genes and ordinate represents m6Ascore. (G) Heatmap of gene mutation frequency in high m6Ascore group. Abscissca represents m6Ascore and ordinate represents glioma samples. Different color modules represent different mutation types and grey modules represent non-mutated samples. (H) Heatmap of gene mutation frequency in low m6Ascore group. Abscissca represents m6Ascore and ordinate represents glioma samples. Different color modules represent different mutation types and grey modules represent non-mutated samples. (I) Immune cells' relationship with m6Ascore based on p value. Abscissca represents m6Ascore and diagonal represents immune cells.

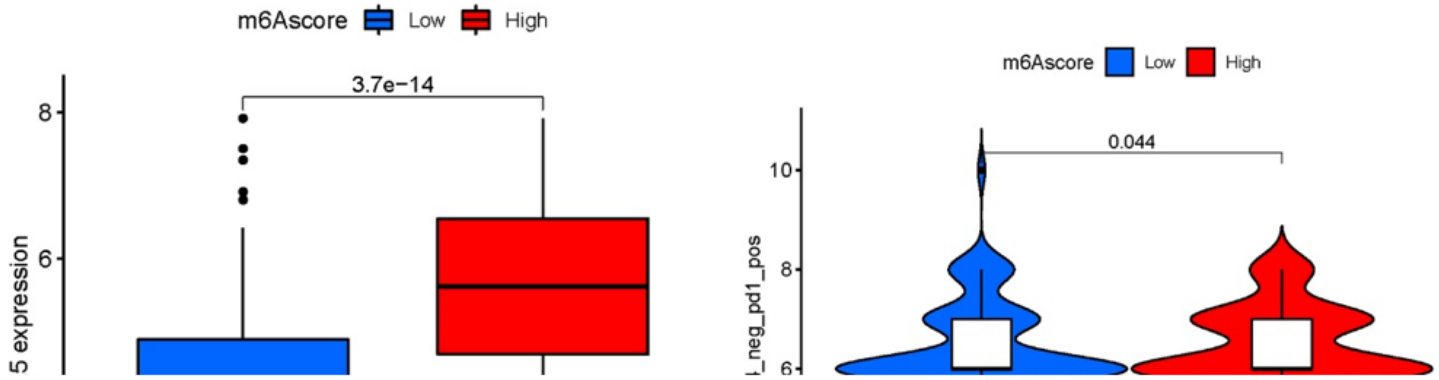

Figure 7

Potential immunotherapy for glioma patients. (A) Box plot of target gene expression based on m6Ascore. Abscissca represents m6Ascore level and ordinate represents target gene. Blue box represents low

expression of m6Ascore and red box represents high expression of m6Ascore. (B) Violin diagram. Abscissca represents m6Ascore level and ordinate represents dual immunotherapy score.

Supplementary Files

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

- supplyment.docx