Identification of a long noncoding RNA signature to predict outcomes of glioblastoma

DEPEI LI, JIE LU, HONG LI, SONGTAO QI and LEI YU

Department of Neurosurgery, Nanfang Hospital, Southern Medical University, Guangzhou, Guangdong 510515, P.R. China

Received August 27, 2018; Accepted March 15, 2019

DOI: 10.3892/mmr.2019.10184

Abstract. Long noncoding RNAs (lncRNAs) are a novel class of gene regulators involved in tumor biogenesis. Glioblastoma is the most common and malignant type of brain tumor. The function and prognostic significance of lncRNAs in glioblastoma remain unclear. In the present study, updated gene annotations were adopted to investigate lncRNA expression profiles in publicly available glioma microarray datasets from the Gene Expression Omnibus and the Repository for Molecular Brain Neoplasia Data. In a training set of 108 samples of glioblastoma, using univariate Cox regression analysis with a permutation P<0.005, four lncRNAs, including insulin-like growth factor binding protein 7-antisense 1 (IGFBP7-AS1), were significantly associated with patient overall survival. These four lncRNAs were integrated as an expression-based molecular signature to divide patients in the training set into high-risk and low-risk subgroups, with distinct survival rates (hazard ratio, 2.72; 95% CI, 1.71-4.31; P<0.001). The prognostic value of the lncRNA signature was confirmed in two additional datasets comprising a total of 147 samples from patients with glioblastoma. The prognostic value of this signature was independent of age and Karnofsky performance status. This signature was also able to predict different outcomes in cases of glioblastoma associated with an isocitrate dehydrogenase 1 mutation. Further bioinformatics analyses revealed that ‘epithelial-mesenchymal transition’ and ‘p53 pathway’ gene sets were enriched in glioblastoma samples with higher IGFBP7-AS1 expression. Furthermore, in vitro experiments demonstrated that knockdown of IGFBP7-AS1 inhibited the viability, migration and invasion of U87 and U251 glioma cells. In conclusion, the present study identified a lncRNA signature able to predict glioblastoma outcomes, and provided novel information regarding the role of IGFBP7-AS1 in glioma development.

Introduction

Glioma is the most common type of primary brain tumor, with an annual incidence of ~6.6/100,000 individuals in the USA (1). Approximately half of all glioma cases are classified as glioblastoma, which is the most malignant type of brain tumor (1). Despite current treatment standards involving maximal safe resection, followed by radiotherapy with concomitant and adjuvant temozolomide (TMZ) (2), the median survival of patients with glioblastoma is only 14-17 months (2,3). As a recognized disease with high heterogeneity, survival variability has been observed in cases of glioblastoma with similar clinical and histological features (4). Traditional clinicopathological factors, including histological grade, age and Karnofsky performance status (KPS), do not appear to be sufficient for precise outcome prediction (3,4). However, with the characterization of specific genetic alterations using advanced sequencing and microarray technologies, several genes and multi-gene expression-based molecular signatures have been discovered. These have since been used for the identification of novel glioblastoma subtypes, and may also have better prognostic significance (3-5).

Proteins are primary functional effectors in cells that represent the vast majority of prognostic biomarkers currently used for glioma (3). Over 68% of the human genome is transcribed to long noncoding RNAs (lncRNAs), a class of transcripts >200 nucleotides in length that lack protein-coding ability. These have emerged as important regulators of tissue physiology and disease processes, including tumorigenesis (6). lncRNAs can regulate the expression levels of oncogenes or tumor suppressors through various mechanisms, including chromatin modification, transcriptional control and post-transcriptional processing, and affect various aspects of cellular homeostasis (7). The expression pattern of lncRNAs is highly tissue- and cell type-specific, a characteristic that may be applied for accurate molecular cancer subclassification and outcome prediction (7). A successfully established example is the lncRNA prostate cancer antigen 3, which is currently used for early diagnosis and treatment response surveillance in prostate cancer (8).

Reports have revealed that lncRNAs are aberrantly expressed in glioma tissue relative to normal brain tissue (9,10).
Some IncRNAs have been implicated in glioma proliferation, apoptosis, motility and stemness maintenance (11,12). However, the detailed mechanisms and clinical application of most IncRNAs remain largely unknown. By virtue of the constant updating of gene annotation databases, such as GENCODE and RefSeq, more gene probes from previously published microarray studies are likely to be repurposed as IncRNAs, providing novel information for investigating the expression profiles and prognostic value of IncRNAs in cancer.

In the present study, the updated gene annotations were adopted to conduct IncRNA profiling on publicly available microarray datasets of glioma, and a four-IncRNA signature was identified to be able to predict the outcome of glioblastoma. Insulin-like growth factor binding protein 7-antisense 1 (IGFBP7-AS1), one of the survival-associated IncRNAs, was further investigated by bioinformatics analysis and functional experiments.

Materials and methods

Glioma datasets preparation. Gene expression profiles generated by the Affymetrix HG-U133 Plus 2.0 microarray platform (Affymetrix; Thermo Fisher Scientific, Inc.) and corresponding clinical data were downloaded from the publicly available Gene Expression Omnibus (GEO; www.ncbi.nlm.nih.gov/geo) and the Repository for Molecular Brain Neoplasia Data (Rembrandt) project (13). The Rembrandt database is available at the Georgetown Database of Cancer web portal (https://gdoc.georgetown.edu). Rembrandt database (13) includes 227 grade IV glioblastomas, 84 grade III gliomas, 99 grade II gliomas and 28 normal brain (NB) samples, according to the world health organization grading system (14). GEO dataset GSE4290 (15) includes 81 glioblastomas, 31 grade III gliomas, 45 grade II gliomas and 23 NB. All these samples were used for IncRNA expression analysis. To avoid interference from different treatment regimens on the analysis of survival-associated IncRNAs, only glioblastoma samples from patients who received post-operative radiation with or without chemotherapy were analyzed, including 108 samples from the GEO dataset GSE16011 (4), 69 from the GEO dataset GSE7696 (16) and 78 from Rembrandt database. The clinical characteristics of the 255 patients with glioblastoma used for survival analysis were described in Table S1.

Microarray data processing and analysis. The raw Affymetrix microarray expression data were processed and normalized using the Robust Multichip Average algorithm (17). The custom Chip Definition File was used to reorganize probes to Entrez Gene identifiers (version 21; National Center for Biotechnology Information, National Institutes of Health) and exclude inaccurate or wrongly annotated probes (18). To avoid systematic error across different experiments, each dataset was standardized independently by the Z-score method to transform the expression of each gene into having a mean of 0 and a standard deviation of 1 (19). Similar to previous methods (9), genes on microarray platforms were identified as IncRNAs according to their RefSeq transcript ID (National Center for Biotechnology Information, National Institutes of Health) and Ensembl gene ID (20) from the updated NetAffx annotation files (HG-U133_Plus_2 Annotations, Release 36; Affymetrix; Thermo Fisher Scientific, Inc.) and then filtered by removing pseudogenes, ribosomal RNAs, transfer RNAs, microRNAs (miRNAs) and other short noncoding RNAs. Finally, a gene list containing 1,895 IncRNAs was created for further analysis. Among these IncRNAs, 1,495 transcripts were annotated on RefSeq and Ensembl databases and 400 transcripts were annotated selectively on RefSeq (Table SII).

Hierarchical clustering analysis was performed to divide patients based on multi-IncRNA expression levels using centered correlation metrics and the average linkage method. Gene set enrichment analysis (GSEA) was implemented according to IncRNA median expression, with the gene sets from the Molecular Signatures Database (software.broadinstitute.org/gsea) (21).

Cell culture and transient transfection. The human glioma cell lines U87 (cat. no. HTB-14, glioblastoma from unknown origin) and U251 were purchased from the Chinese Academy of Sciences. The cells were cultured in Dulbecco's modified Eagle's medium (DMEM; Gibco; Thermo Fisher Scientific, Inc.) containing 10% fetal bovine serum (FBS; Gibco; Thermo Fisher Scientific, Inc.), 100 U/ml penicillin and 50 µg/ml streptomycin, and were incubated at 37°C in an atmosphere containing 5% CO₂.

The small interfering RNA (siRNA) targeting IGFBP7-AS1 and nonspecific negative control (NC) sequence were synthesized by Shanghai GenePharma Co., Ltd. Cell transient transfection was performed using Lipofectamine® 2000 (Invitrogen; Thermo Fisher Scientific, Inc.) 24 h after seeding cells, at 70% confluency. Transfection complexes were prepared according to the manufacturer's protocol and added to the cells to a final oligonucleotide concentration of 100 nM. The transfection medium was replaced 6 h post-transfection. Cells were incubated for 2 days prior to subsequent experimentation. The siRNA sequences are presented in Table I.

Reverse transcription-quantitative (RT-q)PCR. Total RNA was extracted using RNAiso Plus® (Takara Bio, Inc.) according to the manufacturer's protocol. Total RNA (1 µg) was used to synthesize cDNA using the PrimeScript® RT reagent kit with gDNA Eraser (both from Takara Bio, Inc.) according to the manufacturer's protocol. Then, 1 µl cDNA was used to determine the quantification cycle (Cq) value of each sample using the SYBR® Premix Ex Taq II kit (Takara Bio, Inc.) in a LightCycler 480 system (Roche Applied Science). GAPDH served as the reference gene. Relative gene expression was calculated using the 2^(-ΔΔCq) method (22). The PCR primer sequences are in Table II.

Cell viability assay. Cells were seeded into 96-well plates at a density of 1,000-1,500 cells/well, for 1-5 days. The cell viability in each well was assessed at 1, 2, 3, 4 and 5 days using the Cell Counting kit-8 (CCK-8; Dojindo Molecular Technologies, Inc.) according to the manufacturer's protocol. Absorbance was measured at 450 nm to calculate the number of viable cells.

Wound healing and Transwell assays. A scratch was made on 80% confluent monolayer cell cultures, and culture medium was replaced with serum-free DMEM. The healing process was monitored for 24 h, and the percentage of wound closure was calculated using ImageJ software (version 1.51; National Institutes of Health).
in total, 5x10^4 cells in 200 µl serum-free DMEM medium were seeded onto the upper chamber of Transwell apparatus (Costar; Corning, Inc.) with a Matrigel-coated membrane. A total of 600 µl DMEM containing 10% FBS was added to the lower chamber as a chemoattractant. After a 12-h incubation at 37˚C in a 5% CO_2 atmosphere, the cells that invaded the lower chamber through the membrane were fixed with 4% methanol for 30 min at room temperature, stained with 0.1% crystal violet for 15 min at room temperature and counted under a light microscope (magnification, x100).

**Statistical analysis.** SPSS (version 20.0; IBM Corp.) and GraphPad Prism (version 7.0; GraphPad Software Inc.) were used for statistical analysis. ANOVA was performed for multiple comparisons followed by Dunnett’s post-hoc test. Data are presented as the mean ± SD, where appropriate. Student’s t-test was used to compare two groups. χ^2 test was used to compare qualitative variables. Survival-related lncRNAs were screened via univariate Cox regression analysis using Biometric Research Branch-Array Tools version 4.5.0 (National Center Institute, National Institutes of Health) and a permutation test method with 10,000 permutations (23). The selected lncRNAs were analyzed using a multivariate Cox regression analysis in the training set and a risk score formula was established including each identified lncRNA, and weighted using the estimated regression coefficients in the multivariate model (24,25). The risk score of each patient was calculated, and then patients were divided into a high-risk or low-risk group according to the median risk score (-0.0254) of the training set. Patients in the high-risk group exhibited significantly shorter survival times than those in the low-risk group (median 9.48 months vs. 13.68 months; log-rank P<0.001; Fig. 2B and E). The correlation of the risk score with overall survival was also significant when assessed in the univariate model (Table IV). A larger area under the curve in time-dependent ROC analysis was achieved using the risk scoring method, illustrating that this constituted a better algorithm for predicting outcomes, with good sensitivity and specificity compared with hierarchical clustering (Fig. 2C).

**Results**

*Four lncRNAs, tumor protein P73-antisense 1 (TP73-AS1), IGFBP7-AS1, PAX interacting protein 1-antisense 2 (PAXIP1-AS2) and long intergenic non-protein coding RNA LINC00672 (LINC00672), are associated with glioblastoma outcomes.* Among the 159 glioblastoma patients in the GSE16011 dataset, 108 patients received post-operative radiation without chemotherapy. These 108 patients were used as a training set to screen survival-associated IncRNAs. A panel of four IncRNAs strongly associated with overall survival was identified via univariate Cox regression analysis, when the permutation was P<0.005, as listed in Table III. In addition, the upregulation of IncRNAs with a positive regression coefficient (TP73-AS1, IGFBP7-AS1 and PAXIP1-AS2), and the downregulation of IncRNAs with a negative coefficient (LINC00672) were observed in glioblastoma tissues compared with in normal brain tissues in the GSE4290 and Rembrandt datasets (Fig. 1).

As shown in Fig. 2A and D, hierarchical clustering with these four survival-associated IncRNAs subdivided patients in the training set into two clusters with distinct survival rates (median 8.76 months vs. 15.48 months). The four identified IncRNAs were integrated into a molecular signature using a risk score formula to weigh their prognostic effects. The risk score was calculated as follows: Risk score = (0.397x expression value of TP73-AS1) + (0.191x expression value of IGFBP7-AS1) + (0.468x expression value of PAXIP1-AS2) + (-0.129x expression value of LINC00672). The risk score of each patient was calculated, and then patients were divided into a high-risk or low-risk group according to the median risk score (-0.0254) of the training set. Patients in the high-risk group exhibited significantly shorter survival times than those in the low-risk group (median 9.48 months vs. 13.68 months; log-rank P<0.001; Fig. 2B and E). The correlation of the risk score with overall survival was also significant when assessed in the univariate model (Table IV). A larger area under the curve in time-dependent ROC analysis was achieved using the risk scoring method, illustrating that this constituted a better algorithm for predicting outcomes, with good sensitivity and specificity compared with hierarchical clustering (Fig. 2C).

Further validation of the four-lncRNA signature in two independent datasets. From the GSE7696 and Rembrandt...
datasets, 69 and 78 patients with glioblastoma, who had received postoperative radiation, respectively, were considered as independent datasets for validating the prognostic value of the lncRNA signature. Of these, a total of 57 patients accepted radiotherapy only (RT cohort), and 90 patients accepted radiotherapy combined with TMZ chemotherapy (RT-TMZ cohort). Using the same cutoff point for the lncRNA-risk score as for the training set, patients in the GSE4290 and Rembrandt datasets could be also divided into distinct survival subgroups (Fig. 3A and B). Risk score-based subclassification of the entire independent cohort (both the GSE7696 and Rembrandt patients) had similar results, whether the patients accepted TMZ treatment or not. Furthermore, for patients with a high-risk score, radiation combined with TMZ did not appear to significantly prolong the survival time relative to radiotherapy alone (Fig. 3C-E).

Prognostic value of the four-lncRNA signature is independent of conventional clinicopathological factors. Multivariate Cox regression analysis was adopted to estimate whether the prognostic value of the lncRNA signature was independent of the clinical characteristics of patients with glioblastoma. As shown in Table IV, the lncRNA-risk score remained significantly associated with overall survival when adjusted by age and KPS, when available, in every cohort. Data stratification analysis was also performed to assess the prognostic value of this signature with respect to the same clinical factors. For this, patients in the training set were first stratified into a younger group (age ≤50)

Table III. Long noncoding RNAs associated with overall survival of 108 cases of glioblastoma in the training set.

| Gene symbol | Permutation P-value | Hazard ratio | Coefficient | Ensembl ID | Entrez Gene ID | RefSeq ID       |
|-------------|---------------------|--------------|-------------|------------|----------------|----------------|
| TP73-AS1    | 0.001               | 3.229        | 1.172       | ENSG00000227372 | 57212          | NR_033708/    |
|             |                     |              |             |            |                | NR_033709/    |
|             |                     |              |             |            |                | NR_033710/    |
|             |                     |              |             |            |                | NR_033711/    |
|             |                     |              |             |            |                | NR_033712     |
| IGFBP7-AS1  | <0.001              | 1.447        | 0.369       | ENSG00000245067 | 255130         | NR_034081     |
| PAXIP1-AS2  | <0.001              | 2.537        | 0.931       | ENSG00000214106 | 100132707      | NR_024476/    |
|             |                     |              |             |            |                | NR_024477/    |
| LINC00672   | 0.003               | 0.605        | -0.502      | ENSG00000263874 | 100505576      | NR_038847     |

IGFBP7-AS1, insulin-like growth factor binding protein 7-antisense 1; LINC00672, long intergenic non-protein coding RNA 672; PAXIP1-AS2, PAX interacting protein 1-antisense 2; TP73-AS1, tumor protein P73-antisense 1.
and an older group (age >50), then into a poor performance group (KPS ≤80) and a good performance group (KPS >80). The results indicated that the lncRNA signature remained a useful prognostic indicator within each age and KPS stratum (Fig. 4a-d).

The isocitrate dehydrogenase 1 (IDH1) mutation is one of the most valuable biomarkers for glioma diagnosis and prognosis. In the GSE16011 dataset, the lncRNA signature could predict different outcomes of glioblastoma carrying the IDH1 mutation. Patients exhibiting mutations in IDH1 gene and high-risk signature exhibited a poor survival, similarly to those with wild-type IDH1 (Fig. 4E).

Knockdown of IGFBP7-AS1 inhibits the viability and invasion of U87 and U251 glioma cells. As the oncogenic function of TP73-AS1 in glioma has been reported in recent studies (27-29), GSEA was performed to identify the biological processes and pathways of IGFBP7-AS1, PAXIP1-AS2, PAX interacting protein 1-antisense 2; TP73-AS1, tumor protein P73-antisense 1.
by dividing glioblastoma cases from the GSE16011 and Rembrandt datasets into high- and low-expression groups according to the median IncRNA expression. The results demonstrated that the ‘epithelial-mesenchymal transition’, ‘p53 pathway’ and ‘hypoxia’ gene sets were significantly enriched in glioblastoma samples with high IGFBP7-aS1 expression relative to those with low expression (Fig. 5a). IGFBP7-aS1 expression was also the most upregulated of the four prognostic IncRNAs in glioblastoma compared with NB (Fig. 1). Therefore, the role of IGFBP7-aS1 with regards to glioma viability and invasion was further assessed following the transfection of U87 and U251 glioma cell lines with a siRNA targeting IGFBP7-AS1 (Fig. 5B). IGFBP7-AS1 siRNA markedly reduced U87 and U251 cell viability compared with NC-transfected cells, as determined using a CCK-8 assay (Fig. 5C). After knockdown of IGFBP7-AS1, a wound-healing assay showed that the wound closure percentage was visibly reduced, and a Transwell assay revealed that the number of cells that invaded through the Matrigel matrix was significantly decreased (Fig. 5D and E). These results further demonstrated that IGFBP7-AS1 may serve a role in the regulation of glioma cell migration and invasion.

### Table IV. Univariate and multivariate Cox regression analyses for overall survival of patients with glioblastoma in public datasets.

#### A. GSE16011 (training set, n=108)

| Variables       | Univariate model | Multivariate model |
|-----------------|------------------|--------------------|
|                 | HR   | 95% CI       | P-value | HR   | 95% CI       | P-value |
| lnCRNA-risk score | 2.716 | 1.712-4.308 | <0.001<sup>a</sup> | 1.906 | 1.124-3.230 | 0.017<sup>a</sup> |
| Sex             | 0.856 | 0.565-1.298 | 0.465   | 0.841 | 0.552-1.282 | 0.421   |
| Age             | 1.033 | 1.016-1.051 | <0.001<sup>a</sup> | 1.019 | 0.999-1.039 | 0.057   |
| KPS             | 0.980 | 0.967-0.994 | 0.004<sup>a</sup> | 0.984 | 0.970-0.998 | 0.029<sup>a</sup> |

#### B. Entire independent dataset (n=147)

| Variables       | Univariate model | Multivariate model |
|-----------------|------------------|--------------------|
|                 | HR   | 95% CI       | P-value | HR   | 95% CI       | P-value |
| lnCRNA-risk score | 2.560 | 1.586-4.133 | <0.001<sup>a</sup> | 2.450 | 1.539-3.901 | <0.001<sup>a</sup> |
| Sex             | 0.893 | 0.610-1.307 | 0.560   | 1.030 | 1.011-1.050 | 0.658   |
| Age             | 1.029 | 1.011-1.048 | 0.002<sup>a</sup> | 0.917 | 0.625-1.346 | 0.002<sup>a</sup> |

#### C. GSE7696 dataset (n=69)

| Variables       | Univariate model | Multivariate model |
|-----------------|------------------|--------------------|
|                 | HR   | 95% CI       | P-value | HR   | 95% CI       | P-value |
| lnCRNA-risk score | 3.709 | 1.463-9.403 | 0.006<sup>a</sup> | 2.600 | 0.899-7.520 | 0.042<sup>a</sup> |
| Sex             | 0.887 | 0.491-1.602 | 0.690   | 0.540 | 0.457-1.507 | 0.540   |
| Age             | 1.041 | 1.008-1.075 | 0.015<sup>a</sup> | 1.026 | 0.989-1.063 | 0.167   |

#### D. Rembrandt dataset (n=78)

| Variables       | Univariate model | Multivariate model |
|-----------------|------------------|--------------------|
|                 | HR   | 95% CI       | P-value | HR   | 95% CI       | P-value |
| lnCRNA-risk score | 2.197 | 1.216-3.969 | 0.009<sup>a</sup> | 2.451 | 1.335-4.501 | 0.004<sup>a</sup> |
| Sex             | 0.950 | 0.575-1.571 | 0.842   | 1.189 | 0.706-2.002 | 0.516   |
| Age             | 1.031 | 1.007-1.055 | 0.011<sup>a</sup> | 1.028 | 1.003-1.053 | 0.027<sup>a</sup> |
| KPS             | 0.994 | 0.988-1.000 | 0.060   | 0.993 | 0.987-1.000 | 0.043<sup>a</sup> |

<sup>a</sup>P<0.05. KPS, Karnofsky performance status; CI, confidence interval; HR, hazard ratio.
Discussion

lncRNAs are a new class of noncoding gene regulators involved in cancer pathogenesis and prognosis (6,7). Although analyses of global expression patterns and functional characterization of lncRNAs have been performed, the molecular mechanisms and clinical applications of lncRNAs in the context of glioblastoma remain unclear (9-12). Previous studies demonstrated that lncRNA expression profiles can be obtained by mining existing microarray data, as thousands of probes on commonly used arrays are likely to be repurposed as lncRNAs (9,30). Moreover, the number of probes available to map lncRNAs is increasing rapidly due to the updating of lncRNA databases, such as Gencode, in order to provide a more comprehensive documentation of sequence information, functional annotation and expression profiles (31). By combining the microarray mining method and updated gene annotations, the present study identified 1,895 lncRNAs on the Affymetrix platform. To the best of our knowledge, more than half of these genes were not previously annotated as lncRNAs; therefore, the present study provided novel information for investigating lncRNA profiles in cancer, and a four-lncRNA signature was identified with potential for predicting outcomes in glioblastoma.

The current standard treatment for glioblastoma is maximal safe resection followed by radiotherapy and TMZ chemotherapy. Postoperative radio-chemotherapy confers better prognosis than surgery alone, or than radiation without chemotherapy (2). To avoid interference from different treatment regimens, only the patients who had received postoperative radiotherapy were chosen for survival analysis in the training set. The results revealed that four lncRNAs were significantly associated with overall survival in 108 cases of glioblastoma. Upregulation of TP73-aS1, IGFBP7-aS1 and PAXIP1-AS2, with a positive regression coefficient, and downregulation of LINC00672, with a negative coefficient, was observed in glioblastoma samples relative to normal brain tissue. The expression trends were also associated with the malignancy grades of glioma, indicating their potential use in glioma diagnosis and prognosis.

Based on the expression levels of the four lncRNAs, patients with glioblastoma in the training set could be divided into two groups with distinct survival by both hierarchical clustering and a risk scoring method. Risk scoring is a widely used and effective method to establish a multi-gene expression-based molecular signature to predict outcomes of various types of cancer (24,25). By performing ROC analysis, risk scoring was determined as a better estimator of overall survival, with better sensitivity and specificity compared with hierarchical clustering. Overall, patients in the training set with high-risk lncRNA signatures exhibited poorer survival. The reproducibility of this signature was validated in additional independent datasets, including patients who received radiotherapy alone, and those treated with radiation combined with TMZ. The results indicated that the prognostic value of the lncRNA signature in cases of glioblastoma was not restricted by different treatment options. In addition, the prognostic value of this signature was independent of well-known clinical prognostic factors, including age and KPS, further attesting to its potential usefulness in clinical practice. It is worth noting that, compared with radiotherapy alone, postoperative radiation combined with TMZ may have prolonged the survival of patients with glioblastoma, who exhibited a low-risk lncRNA signature.
signature. For high-risk patients, TMZ failed to improve prognosis. This observation implies that the roles of these lncRNAs in glioma progression may be associated with treatment resistance.

IDH1 and IDH2 mutations were previously identified as key biomarkers for molecular pathology and outcome prediction in glioblastoma. The R132H mutation in the IDH1 gene is the most frequent type of IDH mutation (14,32). Mutations in IDH1 mainly occur in lower grade glioma and recurrent glioblastoma, which have improved survival rates compared with primary glioblastoma. In primary glioblastoma, patients exhibiting IDH mutations also possess better survival rates than patients without mutation (32). IDH mutation can induce global DNA hypermethylation and deregulate genes, including

Figure 4. Stratification analysis for the four-lncRNA signature in the training set of glioblastoma cases. (A) Kaplan-Meier survival analysis for younger patients (age ≤50) and (B) older patients (age >50), as well as (C) poor performance patients (KPS ≤80) and (D) good performance patients (KPS >80). (E) Kaplan-Meier survival analysis for patients with IDH1 mutation subdivided into (a) low-risk and (b) high-risk compared with (c) patients with wt IDH1. lncRNA, long noncoding RNA; IDH1, isocitrate dehydrogenase 1; wt, wild type.
Figure 5. Knockdown of IGFBP7-AS1 reduces the viability and invasion of U87 and U251 glioma cells. (A) Gene set enrichment analysis of IGFBP7-AS1 in glioblastoma samples from GSE16011 and Rembrandt datasets. (B) Interference efficiency of IGFBP7-AS1 siRNA in glioma cells analyzed via reverse transcription-quantitative PCR. (C) Cell viability of glioma cell lines following transfection with siRNA or NC measured by Cell Counting kit-8 assay. (D) Migration of glioma cells following transfection with siRNA or NC measured by wound-healing assay. Scale bar, 200 µm. (E) Invasion of glioma cells following transfection with siRNA or NC detected by Transwell assay. Scale bar, 100 µm. Each bar represents the mean ± SD of three independent experiments. *P<0.05 and **P<0.01. IGFBP7-AS1, insulin-like growth factor binding protein 7-antisense 1; NC, negative control; NES, Normalized enrichment score; OD, optical density; siRNA, small interfering RNA.
IncRNAs, inhibiting cellular differentiation, which is involved in tumor initiation and progression (32,33). These genetic findings promoted the emergence of therapeutic approaches aimed to target the mutant form of IDH in glioma and other types of cancer (32). In the GSE16001 dataset, the expression levels of the four prognostic IncRNAs were significantly different between glioblastoma cases with and without the IDH1 mutation. A correlation between these IncRNA levels and the IDH mutation has also been reported in a RNA-sequencing report (10). Furthermore, in the present study, IDH1 mutant glioblastoma with a high-risk IncRNA signature exhibited poor survival, which was comparable with IDH wild type glioblastoma. This may have clinical implications for identifying at-risk patients, among those with IDH mutant glioblastoma, who may therefore require more intensive treatment, such as IDH-targeted therapy in clinical trials (32). These results also indicated the possibility that these IncRNAs are involved in IDH mutation-associated glioma tumorigenesis. Further studies are needed to elucidate the relationship between the IDH mutation and IncRNA deregulation.

The functions of IncRNAs are closely associated with their expression level, as they do not encode proteins (30). The expression of TP73-AS1 is increased in a variety of malignancies, including hepatocellular carcinoma, breast cancer and glioblastoma, and the upregulation of TP73-AS1 also predicts the poor prognosis of cancer (28,34,35). TP73-AS1 may combine with miRNAs to regulate glioma growth, apoptosis and invasion (27-29). To the best of our knowledge, the remaining three IncRNAs were identified as survival predictors in glioma for the first time in the present study. However, in endometrial cancer, lncRNAs were identified as survival predictors in glioma for the first time in the present study. In conclusion, a combination of microarray mining and gene re-annotation may be considered an effective method to investigate the expression profiles of IncRNAs in cancer. A novel four-IncRNA signature was identified as a composite biomarker for glioblastoma outcome prediction. The role of IGFBP7-AS1 in influencing glioma cell viability, migration and invasion was demonstrated with further in vitro experiments. The present study provided a feasible method and novel information for understanding the functions of IncRNAs in glioma pathogenesis and prognosis.

Acknowledgements

Not applicable.

Funding

The present study was funded by The National Natural Science Foundation of China (grant no. 81372692) and National Key Clinical Specialist Construction Program of China.

Availability of data and materials

All data generated or analyzed during this study are included in this published article.

Authors' contributions

DL and LY designed this study and wrote the manuscript. DL analyzed and interpreted the datasets. JL and HL performed the cell experiments. SQ analyzed the data. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Ostrom QT, Gittleman H, Xu J, Kromer C, Wolinsky Y, Kruchko C and Barnholtz-Sloan JS: CBTRUS statistical report: Primary brain and other central nervous system tumors diagnosed in the united states in 2009-2013. Neuro Oncol 18: (Suppl 5): v1-v75, 2016.
2. Stupp R, Mason WP, van den Bent MJ, Weller M, Fisher B, Taphoorn MJ, Belanger K, Brandes AA, Marosi C, Bogdahn U, et al.: Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. N Engl J Med 352: 987-996, 2010.

3. Reifenberger G, Wirsching HG, Knobbe-Thomsen CB and Weller M: Advances in the molecular genetics of gliomas-implications for classification and therapy. Nat Rev Clin Oncol 14: 434-452, 2017.

4. Gravendeel JS, Kouwenhoven MC, Gevaert O, de Rooij JJ, Huarte M: The emerging role of lncRNAs in cancer. Nat Med 21: 107-115, 2015.

7. Reon BJ, Anaya J, Zhang Y, Mandell J, Purow B, Abounader R, Madhavan S, Zenklusen JC, Kotliarov Y, Sahni H, Fine HA and Sun LX, Hui AM, Su Q, Vortmeyer A, Kotliarov Y, Pastorino S, Dai M, Wang P, Boyd AD, Kostov G, They B, Jones GE, Cheadle C, Vawter MP, Freed WJ and Becker KG: Analysis of long noncoding RNAs in the human transcriptome. Nat Genet 47: 199-208, 2015.

10. Verhaak RGW, Hoadley KA, Purdom E, Wang V, Qi Y, Willenbrink CR, Ding L, Golub T,Mesirov JP, et al.: Integrative genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. Cancer Cell 17: 98-110, 2010.

13. Iyer MK, Niknafs YS, Malik R, Singhal U, Sahu A, Hosono Y, Barrette TR, Prendergast J, Evans JR, Zhao S, et al.: The landscape of long noncoding RNAs in the human transcriptome. Nat Genet 39: 9065-9072, 2009.

15. Weller M: Advances in the molecular genetics of gliomas -implications for classification and therapy. Nat Rev Clin Oncol 14: 379-392, 2017.

18. Stubbs PA, Duijm JE, Damen A, Bleeker FE, Bralten LB, Kloosterhof NK, et al.: Intrinsic gene expression profiles of gliomas are a better predictor of survival than histology. Cancer Res 69: 9065-9072, 2009.

19. Madhavan S, Zenklusen JC, Kotliarov Y, Sahni H, Fine HA and Sun LX, Hui AM, Su Q, Vortmeyer A, Kotliarov Y, Pastorino S, Dai M, Wang P, Boyd AD, Kostov G, They B, Jones GE, Cheadle C, Vawter MP, Freed WJ and Becker KG: Analysis of long noncoding RNAs in the human transcriptome. Nat Genet 47: 199-208, 2015.

22. Livak KJ and Schmittgen TD: Analysis of relative gene expression data using real-time quantitative PCR and the 2(-delta delta C(T)) method. Methods 25: 402-408, 2001.