A long non-coding RNA signature predicts survival for glioblastoma as prognostic biomarkers

CURRENT STATUS: UNDER REVIEW

Zhenzhe Li
Harbin Medical University Cancer Hospital

Zhonghua Lv
Harbin Medical University Cancer Hospital

Lei Yu
Harbin Medical University Cancer Hospital

Sibin Zhang
Harbin Medical University Cancer Hospital

Yingjie Wang
Harbin Medical University Cancer Hospital

Siqi Ai
Harbin Medical University Cancer Hospital

Kaibin Zhu
Harbin Medical University Cancer Hospital

Qingsong Li
Second Affiliated Hospital of Harbin Medical University

Peng Liang
Harbin Medical University

Corresponding Author
liangpeng@hrbmu.edu.cn
ORCiD: https://orcid.org/0000-0002-1810-1163

SUBJECT AREAS
   Cancer Biology     Oncology

KEYWORDS
   long non-coding RNA, glioblastoma, prognosis, biomarkers
Abstract

**Background:** Glioblastoma (GBM) is one of the most fatal tumors in the central nervous system. Its prognosis is very poor. There is increasing evidence that long noncoding RNA (lncRNA) participates in the biological process of glioblastoma. Nevertheless, the role of lncRNA in predicting the prognosis of GBM is still uncertain.

**Methods:** In this study, using RNA-Seq and clinical follow-up data of GBM patients from The Cancer Genome Atlas (TCGA), we performed differential analysis of lncRNA, univariable and multivariable Cox regression analysis, Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis, and Gene Ontology (GO) analysis.

**Results:** We identified four lncRNAs closely interrelated with survival and prognosis of GBM patients. This lncRNA signature was effective in both the training set and the testing set, and it was independent to clinical factors.

**Conclusions:** Our data suggested that the four lncRNAs could be used as promising biomarkers for predicting prognosis in GBM patients.

**Background**

Glioma is the most common malignant form of primary tumors in central nervous system, accounting for nearly 80% of primary lethal brain tumors [1]. GBM is the most invasive type of glioma. The median survival time of patients diagnosed with GBM is less than 15 months following standard treatment [2]. Even with significant improvement of the therapeutic regimen in GBM, including surgery, radiotherapy as well as chemotherapy, the prognosis of patients is still extremely dismal. This failure of therapy is owing to the complex biological process of GBM, including the existence of refractory glioma stem cells, as well as chemotherapy protection caused by the blood–brain/tumor barrier [3]. Since there has no substantial breakthrough in treatment of GBM, it is essential to provide patients with valuable prognostic information. To date, widespread and effective prognostic biomarkers for GBM have not been available.

LncRNA is a kind of transcript without protein coding function, its length exceeds 200 nucleotides [4]. LncRNA can be assorted into five types according to their mode of action and functions: sense,
antisense, intronic, intergenic and bidirectional [5]. Last few years, increasing evidence reveals that IncRNA is involved in a variety of functional activities [6], including chromatin remodeling, transcriptional and post-transcriptional regulation, et al. Dysregulated IncRNAs exert tumor facilitator or suppressor function in various cancers [7, 8] Moreover, several studies indicate that numbers of IncRNAs are aberrantly expressed in glioma and correlate with glioma pathogenesis, such as initiation and progression [9, 10], which have significant clinical values in diagnosis and prognosis of glioma[11, 12]. Therefore, IncRNAs exhibit the potential to be promising survival predictors for patients with GBM.

In our study, analysis of GBM RNA-Seq data from TCGA was performed and differentially expressed IncRNAs were discerned. Then we identified a four-IncRNA signature which had important prognostic value for GBM patients based on differentially expressed IncRNAs Furthermore, we evaluated the performance of the four-IncRNA signature, and its effectiveness, generalization ability and independence were verified. Thus, this four-IncRNA signature could provide valuable prognostic information to GBM patients.

Methods

The GBM patient information

The GBM data and matching clinical information were downloaded from TCGA database. We excluded samples without overall survival information, and in total, 159 patients with GBM were analyzed in our study. 159 samples were randomly classified into training set(n = 80) and testing set(n = 79).

IncRNA expression profile

GBM RNA-seq data were obtained from TCGA data portal (https://tcga-data.nci.nih.gov/tcga/). After annotation using ensemble database, the expression levels of IncRNAs and mRNAs were calculated by Reads Per Kilobase per Million mapped reads (RPKM). We screened out IncRNAs according to the following three criteria [13]: 1) transcripts were not detected in any protein-coding domain; 2) transcript sequences had been annotated in GENCODE project [14]; 3) transcripts were expressed in at least half of the GBM samples. The IncRNA expression profiles were defined as those with an average RPKM ≥ 0.1 in 159 GBM samples. Finally, a total of 6315 IncRNAs were identified. Expression difference was analyzed by edgeR package [15] in R.
Statistical analysis
According to the training group, the interrelation between IncRNA expression level and patients’ overall survival was computed by univariate Cox regression analysis. Those IncRNAs with P-values less than 0.05 were regarded to be statistically significant. Then, the chosen IncRNAs were fitted in a multivariate Cox regression analysis in the training set. Risk scores were assessed by incorporating these chosen IncRNAs, which were weighted according to their estimated regression coefficients in the multivariable Cox regression model. The risk score can be computed for each GBM patient based on prognostic four-IncRNA signature. According to the median risk score, GBM patients could be classified into high-risk and low-risk groups, respectively. Survival differences between the two groups were assessed by the Kaplan-Meier survival analyses. To further test whether the risk score was independent of age and gender, multivariate Cox regression and data stratification analyses were performed. The time-dependent receiver operating characteristic (ROC) curve within 2–5 years were used to evaluate the sensitivity and specificity of the survival prediction in accordance with the risk score. All analyses were completed using R (version 3.6.1).

Functional enrichment analyses
In order to further understand the function of the four IncRNAs, we computed the expression relevance between the four IncRNAs and protein-coding genes by spearman correlation coefficients. GO and KEGG enrichment analyses for these co-expressed genes were executed with the cutoff value of P-value < 0.05 and enrichment score > 1.0.

Results
Identification of differentially expressed IncRNAs in GBM
After analysis of the GBM RNA-Seq data from TCGA, 631 differentially expressed IncRNAs were identified (|logFC|>2, false discovery rate(FDR) < 0.001), consisting of 259 upregulated genes and 372 downregulated genes between GBM tissues and normal tissues (Fig. 1).

Four prognostic IncRNAs were detected from the training set
The 159 GBM patients were allocated into two sets at random, a training set (n = 80) and a testing set (n = 79). Based on the training set, univariable Cox regression model of IncRNAs was carried out, and a set of four IncRNAs were identified to be considerably correlated with the patients’ overall survival (P-value < 0.05; Table 1). All of them had positive coefficients, indicating that their aberrant
expression was related to GBM tumorigenesis.

Expression level of the four lncRNAs
Here, we compared the expression levels of the four prognostic lncRNAs between normal brain tissue and GBM and found that AC096772.6 and UNC5B-AS1 were downregulated while HOTAIR and CTD-2263F21.1 were upregulated in GBM (Fig. 2). Except for HOTAIR, the other three lncRNAs were all firstly reported in GBM.

The four-lncRNA signature and patients’ survival in the training set
We designed a risk-score formula for prognostic prediction of GBM patients, based on the expression level of the four lncRNAs. The formula was as follows: Risk score= (0.91 × expression level of AC096772.6) + (0.14 × expression level of HOTAIR) + (0.36 × expression level of CTD-2263F21.1) + (0.31 × expression level of UNC5B-AS1). Afterwards, the four-lncRNA signature risk score was computed for each patient in the training set, and these patients were allocated into high-risk (n = 40) and low-risk groups (n = 40) by the median risk score value defined as the cutoff value. The Kaplan-Meier analysis indicated that patients in the high-risk group had lower overall survival than those in the low-risk group (P-value = 2.35e-05; Fig. 3A). We also performed time-dependent ROC analysis to assess the sensitivity and specificity of the four-lncRNA signature, and the area under the curve (AUC) score was 0.888 (2 years) (Fig. 3B). Moreover, ROC analysis from 3 to 5 years was performed (S-Fig. 1A-C). The result demonstrated that the four-lncRNA signature displayed a better prognostic predictive capability in the training set. Univariate Cox regression analysis showed that the four-lncRNA risk score was remarkably correlated with patients’ survival (P-value = 6.57E-08, HR = 1.409, 95% CI = 1.244-1.597; Table 2). The scatter plot of the four lncRNAs risk score and survival status in training set were showed in Fig. 3C. We found that the mortality rate of patients with high risk scores was higher than that of patients with low risk scores.

Confirmation of the four-lncRNA signature for the survival prediction in testing set and the entire TCGA data set
Subsequently, to validate our results, we verified the four-lncRNA signature in the testing set. Based on the same risk score method, the risk score of each patient in the testing set was computed, and GBM patients were divided into a high-risk (n = 32) and a low-risk group(n = 47) by the same
threshold. In line with the results in the training set, survival analysis suggested that the prognosis of the high-risk group was remarkably worse than that of the low-risk group (P-value = 0.0497; Fig. 4A). Similar result was yielded in the entire TCGA cohort (P-value = 5.87e-05; Fig. 4B). The ROC curves of the four-IncRNAs signature had AUC score of 0.782 in the testing set (Fig. 4C), and 0.746 in the entire set (2 years) (Fig. 4D). ROC analysis of the testing set and entire set from 3 to 5 years was also performed (S-Fig. 2A-C, S-Fig. 3A-C).

**Independence of the four-IncRNA signature and the other clinical variables**

Whether the survival prediction capability of the four-IncRNA signature was irrelevant to clinical factors? To answer this question, we executed multivariate Cox regression analysis based on IncRNA risk score and clinical characteristics such as age and gender (Table 2). The result suggested that when adjusted by clinical factors, four-IncRNA risk score remained to be closely related to overall survival in training set and entire set besides testing set. We considered that this may be attributed to small sample size. However, this did not affect the performance and predictive value of the four-IncRNA signature in the overall set. Moreover, we found that the age was also significantly related to prognosis. Then, stratification analysis was used, and all patients could be classified into younger stratum (age ≤ 60, n = 81) and older stratum (age > 60, n = 78). The analysis revealed that the four-IncRNA risk score could further subdivide GBM patients into high-risk and low-risk group within each age stratum (Fig. 5). These results indicated that prognostic value of the four-IncRNA signature was independent of age.

**Functional properties of the four prognostic IncRNAs**

Many IncRNAs involve in tumorigenesis via communicating with protein-coding genes. To deduce the biological function of the four IncRNAs in GBM, co-expression correlation between four IncRNAs and protein-coding genes were calculated. Protein-coding genes whose spearman correlation coefficients were more than 0.45 were chose as co-expressed genes of four prognostic IncRNAs. Finally, a total of 1621 protein-coding genes were notably associated with at least one of the four IncRNAs. Then we performed functional enrichment analysis, and the results showed that there were 479 GO terms and 11 KEGG pathways. Among the GO terms, Wnt signaling pathway, RNA catabolic process, chromatin
modification and histone modification were mainly enriched. The KEGG pathway included Ribosome, thermogenesis, regulation of actin cytoskeleton and so on (Fig. 6).

**Discussion**

Last several years, an extensive amount of lncRNAs have been discovered and become a novel class of gene regulators. Numerous studies have proved that lncRNAs act as vital parts in various biological functions by regulating gene expression at the transcriptional, post-transcriptional and epigenetic levels [16]. LncRNAs are also thought to serve as either tumor suppressors or oncogenes in different kinds of cancer [17]. In addition, a large number of lncRNAs are aberrantly expressed in glioma, which are closely concerning with diagnosis, prognosis, metastasis and tumorigenesis [9]. LncRNAs exhibit the potential to become ideal indicators for early diagnosis and prognostic prediction of cancers [18, 19]. Therefore, we speculated that differential expression of lncRNAs may serve as prognostic biomarkers for patients with GBM.

The purpose of our study was to systematically identify lncRNAs that served as prognostic biomarkers for GBM by measuring the expression of lncRNAs in GBM. Thus, we obtained lncRNA expression profiles data and clinical data from TCGA database, and differential expression analysis of lncRNAs was performed. A four-lncRNA prognostic signature notably associated with the overall survival was identified on the basis of differentially expressed lncRNAs. Subsequently, the manifestation of this four-lncRNA signature was assessed by ROC analysis. And the results suggested that the four-lncRNA signature had significant predictive value for prognosis in two years, however, it performed unsatisfactorily in three to five years. [20]. Meanwhile, we validated the predictive capability of the four-lncRNA signature was independent of clinical factors. LncRNAs can exert its biological functions in various levels including interacting with protein-coding genes. A spearman correlation analysis that incorporated lncRNAs and protein-coding genes was performed to identify co-expressed genes. Functional enrichment analysis was carried out to expose the function of the four prognostic lncRNAs in GBM. And the results increased our understanding in how these lncRNAs affected prognosis of GBM. For example, a great deal of studies have shown that Wnt signaling pathway is abnormally activated in GBM, promoting the proliferation and invasion of GBM, and inducing resistance to temozolomide in
To our knowledge, in these four prognostic IncRNAs, HOTAIR and UNC5B-AS1 have been investigated in cancers, while AC096772.6 and CTD-2263F21.1 have not been reported in any cancers. HOTAIR was firstly discovered in 2007 by Chang et.al as a trans-acting intergenic IncRNA that is transcribed from the homeobox gene C cluster and inhibit the expression of distal HOXD locus [23]. And increasing evidences reveal that overexpressed HOTAIR is related with proliferation, invasion, metastasis and shorter overall survival in cancers via many mechanisms [24, 25]. UNC5B-AS1 transcribed from UNC5B is a newly confirmed IncRNA in thyroid cancer [26]. UNC5B-AS1 is 652 nucleotides in length, containing 2 exons and located in the genome 10q22.1 [27]. Until now, there are only several literatures correlated with UNC5B-AS1 in cancers can be searched. These studies reveal that UNC5B-AS1 is upregulated and relevant to proliferation, migration and invasion in ovarian cancer, breast cancer, thyroid cancer and colon cancer [28–31]. However, there have no reports of UNC5B-AS1 in GBM. Notably, contrary to reported results above, our data showed that UNC5B-AS1 was downregulated in GBM. We speculated that this may be induced by tissue specificity of IncRNA. Thus, it is a convincing speculation that the four prognostic IncRNAs may be implicated in pathological processes of GBM and their aberrant regulation may lead to occurrence and development of GBM. Nevertheless, except for HOTAIR, the relationship between the other three IncRNAs and GBM is reported for the first time, further biological and molecular experiments should be executed to help us to explore and understand their functional roles and underlying mechanisms in GBM.

Conclusion
A four-IncRNA signature which could predict prognosis of GBM patients was identified in our work. Additionally, we also validated the effect of the four-IncRNA signature. Our data greatly determined that the prognostic predictive value of the four-IncRNA signature was powerful, and GBM patients may benefit from it.

Abbreviations
GBM: glioblastoma; IncRNA: long noncoding RNA; TCGA: The Cancer Genome Atlas; KEGG: Kyoto Encyclopedia of Genes and Genomes; GO: Gene Ontology; RPKM: Reads Per Kilobase per Million
Declarations

Ethics approval and consent to participate

Not applicable.

Consent to publish

Not applicable.

Availability of data and materials

All datas are available. Please contact us to access if it is needed.

Competing interests

There are no conflicts of interest in this study.

Funding

The study did not accept any funding.

Authors' Contributions

PL and QL: conceive and design the research. ZL: analyze data and draft the manuscript. ZHL and LY: review and write guidance. SZ and KZ: download data and collate references. YW and SA: literature research. All authors read and approved the final manuscript.

Acknowledgements

We thank Linyun Wan (Harbin Medical University College of Bioinformatics Science and Technology) and Jianyang Du (The second affiliated hospital of Harbin Medical University) for technical support.

Authors' Information

1 Department of Neurosurgery, Harbin Medical University Cancer Hospital, Harbin, Heilongjiang, People's Republic of China

2 Department of Neurosurgery, The second affiliated hospital of Harbin Medical University, Harbin, Heilongjiang, People's Republic of China

References

1. Wen, P. Y. & Kesari, S. Malignant gliomas in adults. *N Engl J Med* **359**, 492-507,
doi:10.1056/NEJMr0708126 (2008).

2 Tanaka, S., Louis, D. N., Curry, W. T., Batchelor, T. T. & Dietrich, J. Diagnostic and therapeutic avenues for glioblastoma: no longer a dead end? *Nat Rev Clin Oncol* **10**, 14-26, doi:10.1038/nrclinonc.2012.204 (2013).

3 Singh, S. K. *et al.* Identification of human brain tumour initiating cells. *Nature* **432**, 396-401, doi:10.1038/nature03128 (2004).

4 Koch, L. Functional genomics: Screening for IncRNA function. *Nat Rev Genet* **18**, 70, doi:10.1038/nrg.2016.168 (2017).

5 Ma, L., Bajic, V. B. & Zhang, Z. On the classification of long non-coding RNAs. *RNA Biol* **10**, 925-933, doi:10.4161/rna.24604 (2013).

6 Mercer, T. R., Dinger, M. E. & Mattick, J. S. Long non-coding RNAs: insights into functions. *Nat Rev Genet* **10**, 155-159, doi:10.1038/nrg2521 (2009).

7 Naidu, S. & Garofalo, M. microRNAs: An Emerging Paradigm in Lung Cancer Chemoresistance. *Front Med (Lausanne)* **2**, 77, doi:10.3389/fmed.2015.00077 (2015).

8 Patel, J. S. *et al.* Non-coding RNA as mediators in microenvironment-breast cancer cell communication. *Cancer Lett* **380**, 289-295, doi:10.1016/j.canlet.2015.11.016 (2016).

9 Bian, E. B. *et al.* LncRNAs: new players in gliomas, with special emphasis on the interaction of lncRNAs With EZH2. *J Cell Physiol* **230**, 496-503, doi:10.1002/jcp.24549 (2015).

10 Zhang, X. Q. & Leung, G. K. Long non-coding RNAs in glioma: functional roles and clinical perspectives. *Neurochem Int* **77**, 78-85, doi:10.1016/j.neuint.2014.05.008 (2014).

11 Wang, Q. *et al.* A novel cell cycle-associated lncRNA, HOXA11-AS, is transcribed from the 5-prime end of the HOXA transcript and is a biomarker of progression in glioma. *Cancer Lett* **373**, 251-259, doi:10.1016/j.canlet.2016.01.039 (2016).

12 Jing, S. Y. *et al.* Expression of long non-coding RNA CRNDE in glioma and its correlation with tumor progression and patient survival. *Eur Rev Med Pharmacol Sci* **20**, 3992-3996 (2016).

13 Shi, D. *et al.* A five-long non-coding RNA signature to improve prognosis prediction of clear cell renal cell carcinoma. *Oncotarget* **8**, 58699-58708, doi:10.18632/oncotarget.17506 (2017).
Derrien, T. et al. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res* **22**, 1775-1789, doi:10.1101/gr.132159.111 (2012).

Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140, doi:10.1093/bioinformatics/btp616 (2010).

Mercer, T. R. & Mattick, J. S. Structure and function of long noncoding RNAs in epigenetic regulation. *Nat Struct Mol Biol* **20**, 300-307, doi:10.1038/nsmb.2480 (2013).

Gibb, E. A., Brown, C. J. & Lam, W. L. The functional role of long non-coding RNA in human carcinomas. *Mol Cancer* **10**, 38, doi:10.1186/1476-4598-10-38 (2011).

Li, D. et al. LncRNA MAGI2-AS3 Is Regulated by BRD4 and Promotes Gastric Cancer Progression via Maintaining ZEB1 Overexpression by Sponging miR-141/200a. *Mol Ther Nucleic Acids* **19**, 109-123, doi:10.1016/j.omtn.2019.11.003 (2019).

Zhou, M. et al. Identification and validation of potential prognostic IncRNA biomarkers for predicting survival in patients with multiple myeloma. *J Exp Clin Cancer Res* **34**, 102, doi:10.1186/s13046-015-0219-5 (2015).

Li, Z. et al. Modulating IncRNA SNHG15/CDK6/miR-627 circuit by palbociclib, overcomes temozolomide resistance and reduces M2-polarization of glioma associated microglia in glioblastoma multiforme. *J Exp Clin Cancer Res* **38**, 380, doi:10.1186/s13046-019-1371-0 (2019).

Han, M. et al. Interfering with long non-coding RNA MIR22HG processing inhibits glioblastoma progression through suppression of Wnt/beta-catenin signalling. *Brain*, doi:10.1093/brain/awz406 (2019).

Paw, I., Carpenter, R. C., Watabe, K., Debinski, W. & Lo, H. W. Mechanisms regulating glioma invasion. *Cancer Lett* **362**, 1-7, doi:10.1016/j.canlet.2015.03.015 (2015).

Rinn, J. L. et al. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* **129**, 1311-1323, doi:10.1016/j.cell.2007.05.022 (2007).

Tan, S. K. et al. Serum long noncoding RNA HOTAIR as a novel diagnostic and prognostic
biomarker in glioblastoma multiforme. *Mol Cancer* **17**, 74, doi:10.1186/s12943-018-0822-0 (2018).

25 Li, L. *et al.* HOX cluster-embedded antisense long non-coding RNAs in lung cancer. *Cancer Lett* **450**, 14-21, doi:10.1016/j.canlet.2019.02.036 (2019).

26 Lu, W., Xu, Y., Xu, J., Wang, Z. & Ye, G. Identification of differential expressed IncRNAs in human thyroid cancer by a genome-wide analyses. *Cancer Med* **7**, 3935-3944, doi:10.1002/cam4.1627 (2018).

27 Venter, J. C. *et al.* The sequence of the human genome. *Science* **291**, 1304-1351, doi:10.1126/science.1058040 (2001).

28 Wang, H., Su, H. & Tan, Y. UNC5B-AS1 promoted ovarian cancer progression by regulating the H3K27me on NDRG2 via EZH2. *Cell Biol Int*, doi:10.1002/cbin.11300 (2020).

29 Cao, Z. *et al.* Long non-coding RNA UASR1 promotes proliferation and migration of breast cancer cells through the AKT/mTOR pathway. *J Cancer* **10**, 2025-2034, doi:10.7150/jca.29457 (2019).

30 Wang, Y. *et al.* The lncRNA UNC5B-AS1 promotes proliferation, migration, and invasion in papillary thyroid cancer cell lines. *Hum Cell* **32**, 334-342, doi:10.1007/s13577-019-00242-8 (2019).

31 Zhang, Y., Li, Z. & Lan, Z. Silencing UNC5B antisense IncRNA 1 represses growth and metastasis of human Colon cancer cells via raising miR-622. *Artif Cells Nanomed Biotechnol* **48**, 60-67, doi:10.1080/21691401.2019.1699809 (2020).

Tables

Table 1: Four lncRNA significantly associated with overall survival

| Gene symbol       | Coefficient | HR   | 95%CI of HR | P-value       |
|-------------------|-------------|------|-------------|---------------|
| AC096772.6        | 0.91        | 2.49 | 1.35-4.59   | 3.46e-03      |
| HOTAIR            | 0.14        | 1.15 | 1.01-1.31   | 2.92e-02      |
| CTD-2263F21.1     | 0.36        | 1.44 | 1.12-1.84   | 3.97e-03      |
| UNC5B-AS1         | 0.31        | 1.37 | 1.03-1.83   | 3.27e-02      |

Table 2: Univariable and multivariable Cox regression analyses in each data set.

| Variables | Univariable model | Multivariable model |
|-----------|-------------------|---------------------|
|           |                   |                     |
|                       | Coefficient | HR    | 95%CI of HR | P-value | Coefficient | HR    | 95%CI of HR | P-value |
|-----------------------|-------------|-------|-------------|---------|-------------|-------|-------------|---------|
| Training set          |             |       |             |         |             |       |             |         |
| Risk score            | 0.343       | 1.409 | 1.244-1.597 | 6.57e-08 | 0.349       | 1.419 | 1.248-1.613 | 8.51e-08 |
| Gender                | 0.026       | 1.026 | 0.548-1.920 | 0.936   | 0.006       | 1.006 | 0.531-1.908 | 0.984   |
| Age                   | 0.028       | 1.029 | 0.998-1.059 | 0.066   | 0.027       | 1.028 | 0.997-1.059 | 0.074   |
| Testing set           |             |       |             |         |             |       |             |         |
| Risk score            | 0.138       | 1.148 | 0.919-1.434 | 0.221   | 0.141       | 1.151 | 0.901-1.471 | 0.261   |
| Gender                | 0.339       | 1.404 | 0.842-2.342 | 0.193   | 0.338       | 1.402 | 0.826-2.381 | 0.211   |
| Age                   | 0.037       | 1.038 | 1.017-1.059 | 3.99e-04| 0.037       | 1.037 | 1.016-1.059 | 6.01e-04|
| Entire set            |             |       |             |         |             |       |             |         |
| Risk score            | 0.241       | 1.273 | 1.155-1.404 | 1.19e-06| 0.251       | 1.286 | 1.159-1.426 | 1.91e-06|
| Gender                | 0.181       | 1.199 | 0.811-1.772 | 0.363   | 0.210       | 1.234 | 0.829-1.835 | 0.299   |
| Age                   | 0.035       | 1.036 | 1.018-1.053 | 5.47e-05| 0.034       | 1.035 | 1.017-1.053 | 1.04e-04|

P<0.05 was considered statistically significant in all analyses.

Figures
Figure 1

Differentially expressed IncRNAs in GBM. (A) Volcano plot of differentially expressed IncRNAs. The red dot represents up-regulated IncRNAs, and green dot represents down-regulated IncRNAs. (B) The heatmap of hierarchical clustering of differentially expressed IncRNAs between glioma samples and normal brain tissues.
The expression level of the four prognostic IncRNAs: (A-D) AC096772.6 and UNC5B-AS1 were significantly downregulated, HOTAIR and CTD-2263F21.1 were significantly upregulated in GBM.
Figure 3

The performance of the four-IncRNA signature in the training set: (A) Kaplan–Meier estimates of the overall survival of GBM patients between high-risk group and low-risk group (n=80). (B) The Receiver operating characteristic (ROC) analysis of the overall survival prediction based on the four-IncRNA risk score. (C) The distribution of the four-IncRNA risk score and patients’ survival status.
Figure 4

The performance of the four-lncRNA signature in the testing set and entire set: (A-B) Kaplan–Meier estimates of the overall survival of GBM patients between high-risk group and low-risk group in the testing set (n=79) and entire set (n=159). (C-D) ROC analysis of the overall survival prediction based on the four-lncRNA risk score.
Stratification analyses of all patients adjusted to age using the five-lncRNA signature. (A) The Kaplan-Meier curves for younger GBM patients (age ≤ 60, n=81). (B) The Kaplan-Meier curves for elder GBM patients (age > 60, n=78). (C) The Kaplan-Meier curves for the entire GBM patients (n=159).

Functional enrichment analysis of the four lncRNAs co-expressed protein-coding genes. (A) The significant enriched GO terms of co-expressed protein-coding genes. (B) The significant enriched KEGG pathways of co-expressed protein-coding genes.

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

supplementaryfile2.pdf
supplementaryfile1.pdf
supplementaryfile3.pdf
