

Identification of Novel tRNA-Leu-CAA-Derived tsRNAs for the Diagnosis and Prognosis of Diffuse Gliomas

Bing Xu^{1,*}, Jian Liang^{1,*}, Hecun Zou^{2,*}, Jingwen Wang², Yubo Xiong^{1,3}, Jiao Pei¹

¹Xiangyang Central Hospital, Affiliated Hospital of Hubei University of Arts and Science, Xiangyang, Hubei, 441021, People's Republic of China;

²Institute of Life Sciences, Chongqing Medical University, Chongqing, 400016, People's Republic of China; ³Department of Neurosurgery, General Hospital of the Yangtze River Shipping, Wuhan Brain Hospital, Wuhan, Hubei, 430010, People's Republic of China

*These authors contributed equally to this work

Correspondence: Jiao Pei; Yubo Xiong, Email Ajiao2012@outlook.com; fengshu525@163.com

Purpose: tsRNA is a type of small non-coding RNA derived from tRNA. Diffuse gliomas are the most common brain tumors. This investigation focused on tsRNA identification and characterization within gliomas.

Methods: The sequences of human tRNA and tsRNAs were taken from GtRNAdb, tRFdb and tRFexplorer databases. Data processing and bioinformatic analysis were performed with R or Python software. The expression of tsRNAs in glioma tissues was analyzed by qRT-PCR assay.

Results: With computational approaches, we identified hundreds of tsRNAs with available expression abundance in the glioma datasets, most of them derived from the 3' end or 5' end of mature tRNA. Among the tsRNAs derived from tRNA-Leu-CAA, ts-26, tRFdb-3012a, and tRFdb-3012b (tRFdb-3012a/b) were significantly decreased in diffuse gliomas. The clinical survivals of glioma patients with low tsRNA (ts-26, tRFdb-3012a, and tRFdb-3012b) expression were remarkably worse than that of those with high expression. Expression of tRFdb-3012a/b was correlated with *IDH* mutant status and *MGMT* promoter mutation in gliomas, and tRFdb-3012a/b and ts-23 tended to be highly expressed in patients with the *IDH* mutant. The enrichment analysis showed that some tRFdb-3012a/b-related genes were enriched in RNA splicing and processing, the spliceosome pathway and astrocyte molecular signatures. Moreover, the 3' untranslated region of the *RBM43* gene was predicted to contain putative binding sites of tRFdb-3012a/b, ts-26 may directly bind to the 3' untranslated region of the *HOXA13* gene, and the expressions of both *RBM43* and *HOXA13* were up-regulated in diffuse gliomas. High *RBM43* and *HOXA13* expressions were significantly associated with poor survival outcome of glioma patients.

Conclusion: These results suggest that tRNA-Leu-CAA-derived tsRNAs (ts-26, tRFdb-3012a, and tRFdb-3012b) could be explored as diagnostic and prognostic biomarkers for diffuse gliomas, and tRFdb-3012a/b and ts-26 may play an important role in the progression of gliomas, through binding *RBM43* and *HOXA13*, respectively.

Keywords: tRNA-Leu-CAA, tRFdb-3012a/b, ts-26, diffuse gliomas, *RBM43*

Introduction

tsRNA (tRNA-derived small RNA) is a type of small non-coding RNA derived from transfer-RNA (tRNA), and also named tRNA-derived fragment (tRF).¹ tRNA fragments are short, ranging in length from 16 to 35 nucleotides, and are produced through the cleavage of mature or premature tRNAs at various positions.² A popular classification of tsRNAs is based on their position in accord with the parental tRNA sequence, which can be separated into five smaller categories, namely 5'-tRFs, 3'-tRFs, internal tRFs, 5'U-tRFs, and tRF-1. For example, in the fragments derived from tRNA-Leu-CAA, which is a leucine-transfer RNA, as shown in Figure 1A, the 5'-tRF (tRFdb-5023a) originated precisely from the 5'-end region of mature tRNA-Leu-CAA, two 3'-tRFs (tRFdb-3012a and tRFdb-3012b) were derived from the cleavage of the 3' end of mature tRNA-Leu-CAA, and the tRF-1 fragment (ts-26) was derived from the downstream 3' trailer sequence of the primary tRNA-Leu-CAA

A

	Acceptor stem	D-loop domain	Anticodon domain	T-loop domain	CCA tail
				tRFdb-3012a	AUCCACUUCUGACACCA
tRNA-Leu-CAA-1-1	GUCAGGAUGGCC	GAGUGGUCUA	AGGCGCCAGACUCAAGUUCUGGUCUC	AAUGGAGGCGUGGGUUCGAAU	CCCACUUCUGACA
tRNA-Leu-CAA-1-2	GUCAGGAUGGCC	GAGUGGUCUA	AGGCGCCAGACUCAAGUUCUGGUCUC	AAUGGAGGCGUGGGUUCGAAU	CCCACUUCUGACA
tRNA-Leu-CAA-2-1	GUCAGGAUGGCC	GAGUGGUCUA	AGGCGCCAGACUCAAGUUCUGGUCUC	GUAUGGAGGCGUGGGUUCGAAU	CCCACUUCUGACA
tRNA-Leu-CAA-3-1	GUCAGGAUGGCC	GAGUGGUCUA	AGGCGCCAGACUCAAGUUCUGGUCUC	GUAUGGAGGCGUGGGUUCGAAU	CCCACUUCUGACA
tRNA-Leu-CAA-4-1	GUCAGGAUGGCC	GAGUGGUCUA	AGGCGCCAGACUCAAGUUCUGGUCUC	GUAUGGAGGCGUGGGUUCGAAU	CCCACUUCUGACA
				tRFdb-3012b	UCGAAUCCCACUUCUGACACCA
tRNA-Leu-CAA-6-1	GUCAGGAUGGCC	GAGCAGUCUUA	AGGCGCUGCGUCAAUCGCACCCUCCGC	UGGAGGCGUGGGUUCGAAU	CCCACUUUUGACA
			tRFdb-5023a		GUCAGGAUGGCCGAGC

tRNA-Leu-CAA-1-2-downstream CAACTATCTTATTCTCCTTT
CAACUAUCUUUUCUCCUUU ts-26

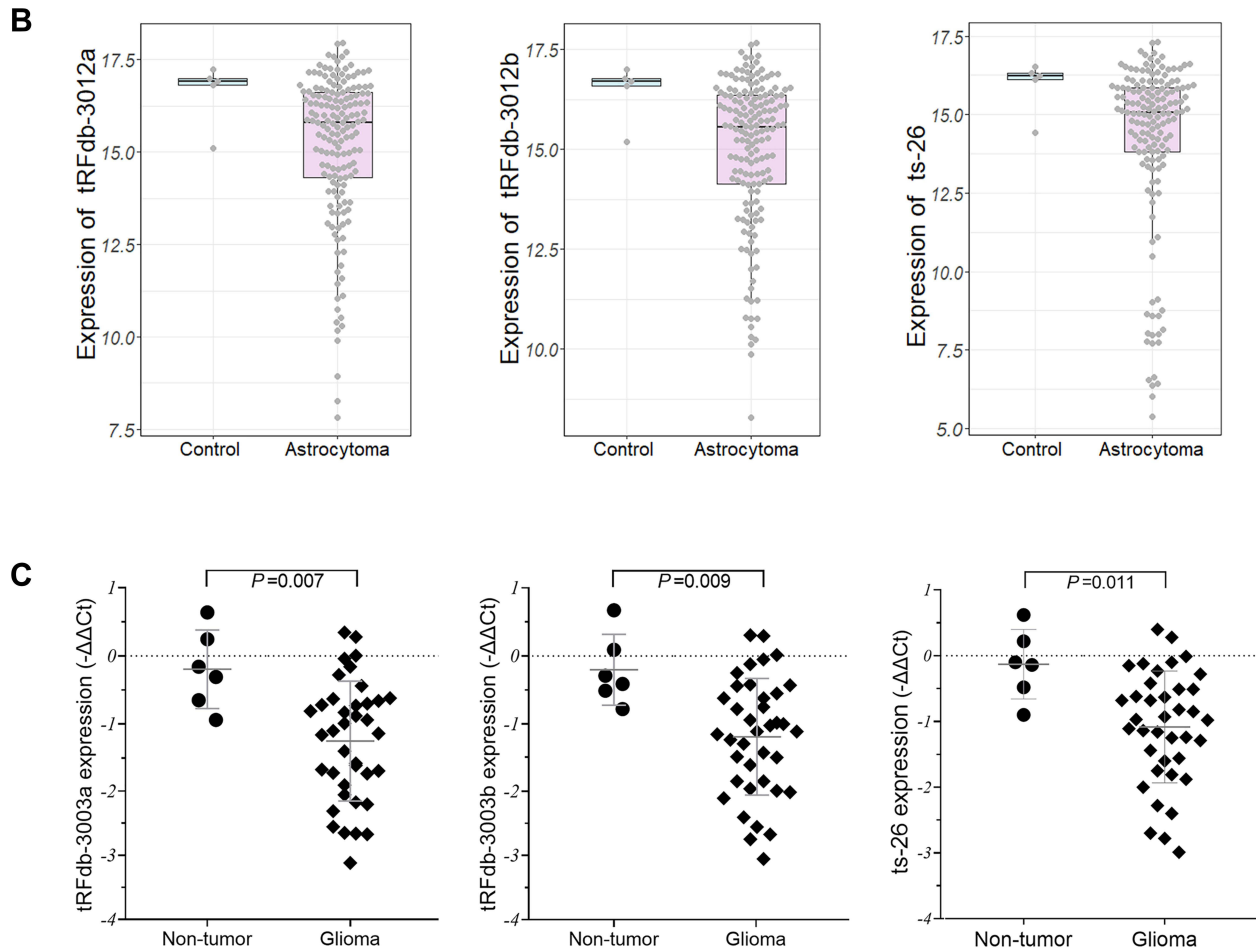


Figure 1 (A) Representative images of the sequence alignments between four tRNA fragments and candidate tRNA-Leu-CAA sources. (B) Expression patterns of three tRNA-Leu-CAA-derived tsRNA fragments (tRFdb-3012a, tRFdb-3012b, and ts-26) in glioma samples (astrocytoma vs non-tumor controls) of TCGA datasets. (C) qRT-PCR analysis of relative tRNA-Leu-CAA-derived tsRNA (tRFdb-3012a, tRFdb-3012b, and ts-26) expression between glioma tissues and non-tumor controls.

gene.²⁻⁴ The cleavage of tRNAs is context dependent, being affected by gender, different diseases, and tissue type. These dependencies give an increased urgency to the need to understand the regulatory roles of tsRNAs. Efforts in this area are gaining momentum, and include experimental and computational methods.³ Subsequent in-depth reanalysis attempted to classify and tabulate the tsRNAs that were being discovered in some higher organisms into a curated repository, and an exclusive database for tsRNAs was constructed.^{5,6}

A growing number of studies have demonstrated that many tsRNAs are differentially expressed in multiple tissues of cancer patients, and these dysregulated tsRNA fragments may serve as potential biomarker candidates for diagnosis and prognosis, and even contribute to the biological functions of tumor progression.^{7,8,9} Diffuse gliomas, especially astrocytoma, are the most common types of brain tumors originating from glial cells. The current standard of care includes surgical resection, followed by radiotherapy plus temozolomide chemotherapy, and adjuvant novel feasible target therapy.¹⁰⁻¹² According to the latest cancer statistics, the mortality due to diffuse gliomas is highest among brain and central nervous system (CNS) tumors.¹³ The pathogenesis of gliomas is extremely complicated, involving the aberrant activation of proto-oncogenes and inactivation of anti-oncogenes.¹⁴ As a critical functional molecule, tsRNA may also be involved in the development and progression of gliomas. However, it is not clear which tsRNA could serve as novel biomarkers for the diagnosis, prognosis, and target therapy of diffuse gliomas.

In this study, we aimed to identify the tsRNAs profiles within small non-coding RNA sequencing (sncRNA-seq) datasets of human glioma samples via a deterministic tsRNAs mining pipeline, and investigated the expression pattern and biological significance of tRNA-Leu-CAA-derived tsRNAs (ts-26, tRFdb-3012a, and tRFdb-3012b) in diffuse gliomas. The interactional relationships between three tsRNAs and their target genes were also explored in order to reveal the underlying regulatory mechanisms.

Materials and Methods

Datasets Collection

The sequences of human tsRNAs, including five fragment types: 5'-tRFs, 3'-tRFs, 5'U-tRFs, and tRF-1, were taken from the tRFdb database (<http://genome.bioch.virginia.edu/trfdb/>) and tRFexplorer database (<https://trfexplorer.cloud/>).^{6,15} The hg19 gene annotations and corresponding sequences of tRNAs in human genomes were taken from GtRNAdb (<http://gtrnadb.ucsc.edu/>).¹⁶ Then, a custom annotation of the reference human genome (hg19) containing only known tsRNAs was assembled. The sncRNA-seq datasets of glioma samples within TCGA, the RNA sequencing data of low-grade glioma (LGG) samples, and their corresponding patient clinical information were retrieved from GDC data portal (<https://portal.gdc.cancer.gov/>).¹⁷ The gene expression profiles of GSE4290 were downloaded from the Gene Expression Omnibus (GEO; <http://www.ncbi.nlm.nih.gov/geo/>) database¹⁸ and the gene expression microarrays are based on the Affymetrix Human Genome U133 Plus 2.0 Array platform (Affymetrix, USA).

Identification Approach

Data processing approach and flowchart for tsRNA identification are shown in [Figure 2](#). A conservative pipeline was implemented to obtain the accurate estimation of tsRNA expression, as described previously.¹⁵ In brief, the quality control and filtering of raw files on sncRNA-seq datasets were preprocessed using Trim Galore (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/), which is a wrapper tool for the Cutadapt and FastQC methods. Then, clean reads of small RNA sequencing were remapped to the reference human genome (hg19) and the sequences of our tsRNAs annotation files by applying the Bowtie tool.¹⁹ After alignment, only the mapped reads could be quantified to count the number of reads belonging to each of the candidate tsRNAs with HTSeq software,²⁰ while other ambiguous reads were removed to obtain a more accurate and conservative analysis. Finally, the expression values of tsRNAs were calculated and normalized as reads per million (RPM) of total raw counts,²¹ and the tsRNAs with an average expression level less than one log₂ RPM were filtered to exclude random degradative sequences.

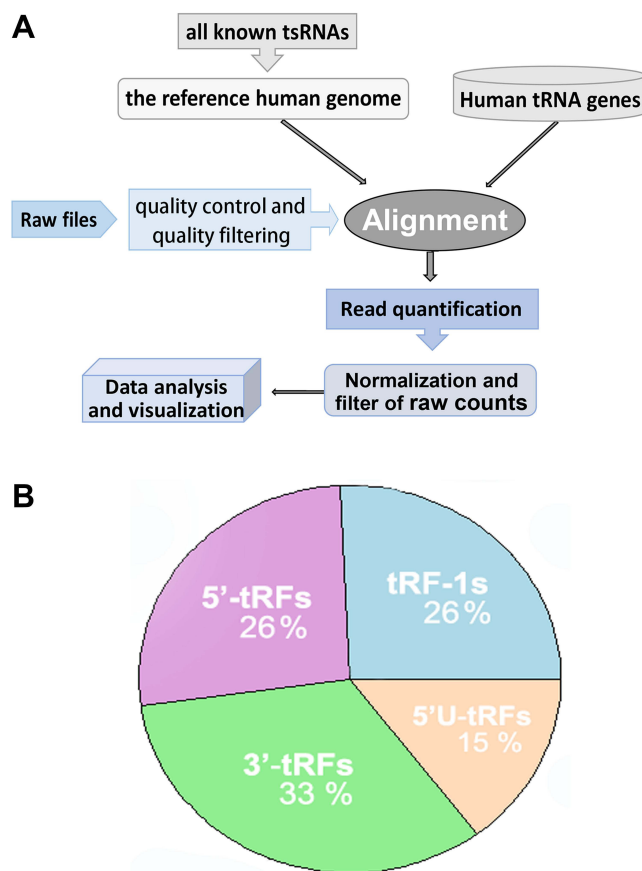


Figure 2 A summary of tsRNA identification and characterization. **(A)** Flowchart of data processing and tsRNA identification pipeline. **(B)** Characterization and distribution of tsRNA corresponding derived tRNA gene sources, including tRNA fragments derived from the 5' end (5'-tRFs) and 3' end (3'-tRFs) of mature tRNA, the 5' leader (5' U-tRFs) and 3' trailer (tRF-1s) sequence of primary tRNA genes.

Bioinformatic Analyses

The RNA expression profiles of TCGA datasets were retrieved and prepared using TCGAbiolinks package of R.²² Survival analysis with Kaplan–Meier curves and a log-rank test were performed and visualized using the R survival and survminer packages.²³ Hierarchical cluster analysis was performed on tsRNA expression and primary pathology characteristics, and the corresponding heatmap was visualized by the pheatmap package (<https://cran.r-project.org/web/packages/pheatmap/>). Pearson correlations were estimated for each pair of tsRNA and candidate mRNA genes, and scatter diagrams were plotted by the ggplot2 and ggpvr package in R statistical software.²⁴ Enrichment analyses were performed to provide biological insights concerning the given tsRNA-related genes, and the molecular signatures database (MSigDB) was used as a functional gene set for gene set enrichment analysis (GSEA).²⁵ Gene ontology (GO), KEGG pathway, and GSEA analyses were performed and visualized using the enrichplot and clusterProfiler package of R (<https://yulab-smu.top/biomedical-knowledge-mining-book/index.html>). The computational binding relationships between a given tsRNA and its target genes were predicted via the tRFtarget database with RNAhybrid and IntaRNA methods.²⁶ The interactional relationships of the top 50 targets and tsRNAs were organized as a network using Cytoscape (v3.7.2) software.

Clinical Specimens

Thirty-eight glioma tissue specimens were obtained from patients diagnosed with diffuse gliomas undergoing surgical resection at the Department of Neurosurgery of Xiangya Hospital of Central South University from August 2015 to January 2019. Six non-tumor brain tissues were obtained from adult patients with craniocerebral injuries, which required partial resections of brain tissue as decompression treatment to reduce intracranial pressure. After excision, tissue specimens were immediately frozen in liquid nitrogen for subsequent use. This study was approved by the Ethics

Committees of Central South University and the patients provided written informed consent.²⁷ The study has been confirmed to comply with the Declaration of Helsinki.

Quantitative RT-PCR Analysis

Total RNA was extracted from tissues using the TRIzol reagent (Invitrogen, USA). One microgram of total RNA from each sample was reverse transcribed into cDNA under standard conditions. Quantitative real-time polymerase chain reaction (qRT-PCR) was performed to quantify the expression of tsRNAs, using with the Bulge-Loop miRNA qRT-PCR Starter Kit (Ribobio, China) with specific stem-loop RT primers, following the manufacturer's protocol. RNU6 (U6) was used as the internal control for tsRNA template normalization. Relative quantification of gene expression was calculated by the comparative cycle-threshold (CT) ($-\Delta\Delta CT$) method.¹¹ Stem-loop primers for tRFdb-3012a, tRFdb-3012b, ts-26, and RNU6 were designed and synthesized by Ribobio Biotech (Guangzhou, China).

Statistical Analysis

The statistical analyses were performed with R software (<https://www.r-project.org/>). GraphPad Prism version 8.0 (GraphPad Software, USA) was used for generating graphs and analysis. Data were shown as means \pm standard deviation (SD). Regarding the numerical variables, statistical significance of differences between two groups was assessed using the two-sided Student's *t*-test, and comparisons of multiple groups were made by one-way analysis of variance. All experiments were performed in triplicate and *P* values less than 0.05 were considered to indicate statistically significant differences.

Results

Identification of tsRNAs Within Diffuse Glioma Datasets

First of all, data on thousands of human tsRNA signatures were collected from tRFdb⁶ and tRFexplorer database;¹⁵ all known tsRNA sequences were remapped to the specific regions of tRNA. As statistical analysis with the collected tsRNA data, these fragments are mainly derived from the 5' end (5'-tRFs) and 3' end (3'-tRFs) of mature tRNA, as well as the 5' leader (5'U-tRFs) and 3' trailer (tRF-1s) sequences of primary tRNA genes.

Using computational identification approaches, we determined lots of tsRNAs and assessed their expression profile within the sncRNA-seq datasets of TCGA-LGG. A summary of the identification pipeline is shown in Figure 2A. After filtering and calculating, a total of 163 tsRNAs with available expression abundance were identified within glioma samples (Supplemental Material Table S1 and S2), of which about thirty-three percent were derived from the 3' end (3'-tRFs) and twenty-six percent from the 5' end (5'-tRFs) of mature tRNA, as well as twenty-six percent tsRNAs from the 3' trailer (tRF-1s) sequence of primary tRNA genes (Figure 2B).

Significance of tsRNAs Derived from tRNA-Leu-CAA Within Diffuse Gliomas

Among the identified tsRNAs, we were interested in the fragments derived from tRNA-Leu-CAA, which is a leucine-transfer RNA. As shown in Figure 1A, we identified four tsRNAs that derived from tRNA-Leu-CAA within the sncRNA-seq datasets of glioma samples. There are two 3'-tRF fragments (tRFdb-3012a and tRFdb-3012b) that remap to the cleavage of the 3' end of five different mature tRNA-Leu-CAA (Figures 1A and 3A), only one 5'-tRF fragment (tRFdb-5023a) that remaps to the 5' end of mature tRNA-Leu-CAA-6-1, and a tRF-1 fragment (ts-26) that remaps to the downstream regions (30 bases) of the primary *tRNA-Leu-CAA-1-2* gene.

To reveal the implications of these tsRNAs in diffuse gliomas, the expression patterns of the above tsRNAs were analyzed in the glioma samples of TCGA datasets. This showed that the expression levels of three tsRNAs (ts-26, tRFdb-3012a, and tRFdb-3012b) were decreased in astrocytoma samples compared to non-tumor controls (Figure 1B) (all $P < 0.05$); however, no significant difference for tRFdb-5023a was observed between astrocytoma and controls (both $P > 0.05$). As is well known, astrocytoma is the main subtype of diffuse gliomas. Our clinical specimen detections also demonstrated that expression levels of three tsRNAs (tRFdb-3012a, tRFdb-3012b, and ts-26) were significantly

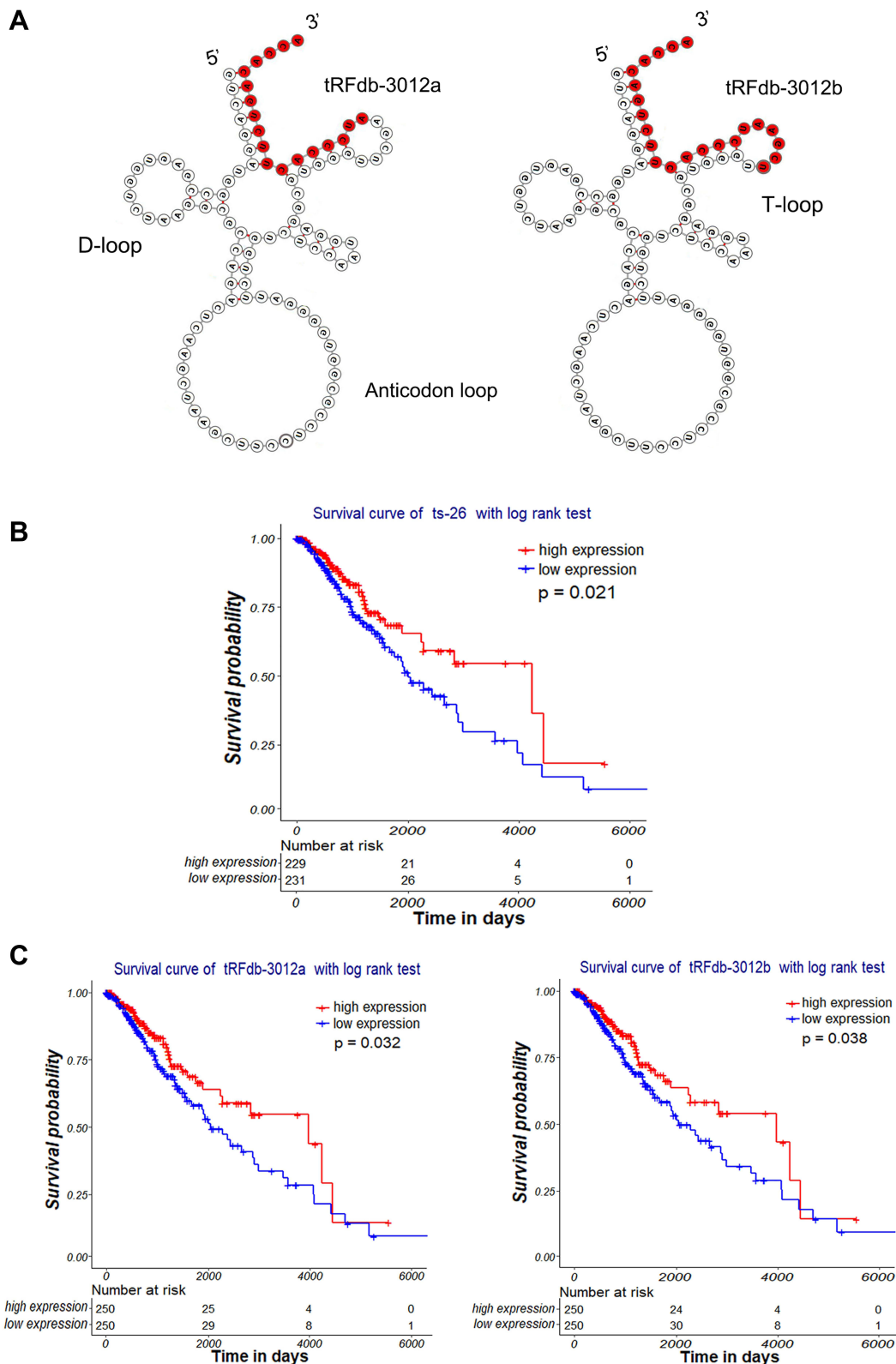


Figure 3 Clinical significance of tRNA-Leu-CAA-derived tsRNAs in gliomas. **(A)** Secondary structure of tRNA-Leu-CAA and its derived fragments (tRFdb-3012a and tRFdb-3012b). **(B and C)** Kaplan–Meier survival curve analysis with a log-rank test based on tRNA-Leu-CAA-derived tsRNA (tRFdb-3012a, tRFdb-3012b, and ts-26) expression within glioma samples of TCGA-LGG datasets.

decreased in glioma tissues compared to non-tumor controls (Figure 1C). These data imply that tRNA-Leu-CAA-derived tsRNAs may play an important role in diffuse gliomas, particularly in astrocytoma.

Next, we investigated the relationship between the expression of three tsRNAs and the overall survival of glioma samples using Kaplan–Meier curve analysis with a log-rank comparison. Based on the median of tsRNA expression, glioma samples were divided into a high expression group and a low expression group. As shown in Figure 3B and C, the clinical survival of glioma patients with low ts-26 expression ($P=0.021$) was significantly worse than that of the high-expression patients; the survival of glioma patients with low tRFdb-3012a expression ($P=0.032$) was worse than that of those with high expression; and the survival of patients with low tRFdb-3012b expression was also worse than that of those with high expression. These results suggest that down-regulated tsRNAs (ts-26, tRFdb-3012a, and tRFdb-3012b) were associated with poor survival outcome of glioma patients, and that tRNA-Leu-CAA-derived tsRNAs could be explored as diagnostic and prognostic biomarkers for diffuse gliomas.

Biological Insights into tRFdb-3012a/b and ts-26 on Gliomas

On the basis of the 2016 WHO CNS classification, the primary clinical and molecular pathology characteristics of glioma patients were retrieved and preprocessed.¹⁰ The correlations of the expression of three tsRNA with pathology characteristics were analyzed using chi-squared tests, and were visualized by hierarchical clustering heatmaps plus bee-swarm plots. As shown in Figure 4A and B, glioma samples with *IDH* wide-type status were clustered more in the left green block, with lower expression of the tsRNAs, while tRFdb-3012a, tRFdb-3012b, and ts-26 (right red block) tended to be up-regulated in patients with the *IDH* mutant (including *IDH* mutant codeletion and non-codeletion), and the expression of tRFdb-3012a/b and ts-26 was significantly correlated with *IDH* mutant status in glioma samples (both $P<0.05$). Moreover, high tRFdb-3012a/b and ts-26 expression was also correlated with *MGMT* promoter methylation status in glioma samples (both $P<0.05$). According to the transcriptome classification of TCGA, tRFdb-3012a, tRFdb-3012b, and ts-26 may be highly expressed in the pro-neural subtype of diffuse gliomas. However, no significant relationship was found between the expression of the given tsRNAs and other pathological characteristics, including gender, grade, BRAF status, 1p19q-codeletion, TERT, and ATRX status (all $P>0.05$). According to the 2016 WHO CNS classification, *IDH* mutant and *MGMT* promoter methylation have been used as specific molecular features for the diagnosis, prognosis, and therapy of diffuse gliomas.¹⁰ These findings imply that tRNA-Leu-CAA-derived tsRNAs may act as a key player in tumor progression of diffuse gliomas.

Moreover, enrichment analyses were performed to provide biological insights into the role of tRNA-Leu-CAA-derived tsRNAs (including tRFdb-3012a, tRFdb-3012b, and ts-26). TCGA-LGG mRNA expression profiling data were used to find some correlated genes of the given tsRNA by the Spearman correlation method. As shown in Figure 5A and B, for gene ontology analysis, some tRFdb-3012a-related genes were enriched in the specific GO terms, such as RNA splicing and processing (*LUC7L*, *OSGEP*, and *SNRNP70*), nuclear speck (*OSGEP* and *SNRNP70*), and catalytic activity (*HDAC10*). The dot plots showed that tRFdb-3012a-related genes were enriched in several KEGG pathways, such as spliceosome and RNA transport (*SNRNP70* and *HOOK2*). In the GSEA, two molecular signatures (astrocyte and STIM treatment response signature) were found to be associated with tRFdb-3012a, such as *LUC7L* and *ING5* were both the tRFdb-3012a correlated genes. Enrichment analyses of tRFdb-3012b and ts-26 were also conducted and results are presented in Supplemental Material Figure S1 and S2.

tRNA-Leu-CAA-Derived tsRNAs Regulate Expression of Its Targets in Gliomas

Using the bioinformatics database and tools, many potential target genes of tRNA-Leu-CAA-derived tsRNAs (including tRFdb-3012a, tRFdb-3012b, and ts-26) were predicted and screened out. The interactional relationships of tRFdb-3012a/b and their targets are presented in Figure 6A. Among these targets, we picked out a molecule with a strong binding score, *RBM43*, an RNA binding motif protein.²⁸ The 3' untranslated region of *RBM43* mRNA contained the putative binding site of tRFdb-3012a and tRFdb-3012b (Figure 6B). Moreover, the analysis of glioma samples within TCGA and GSE4290 datasets showed that *RBM43* was highly expressed in gliomas compared to non-tumor controls (Figure 7A), and tRFdb-3003a and tRFdb-3003b were decreased in diffuse gliomas (Figure 1B and C). Kaplan–Meier survival curve

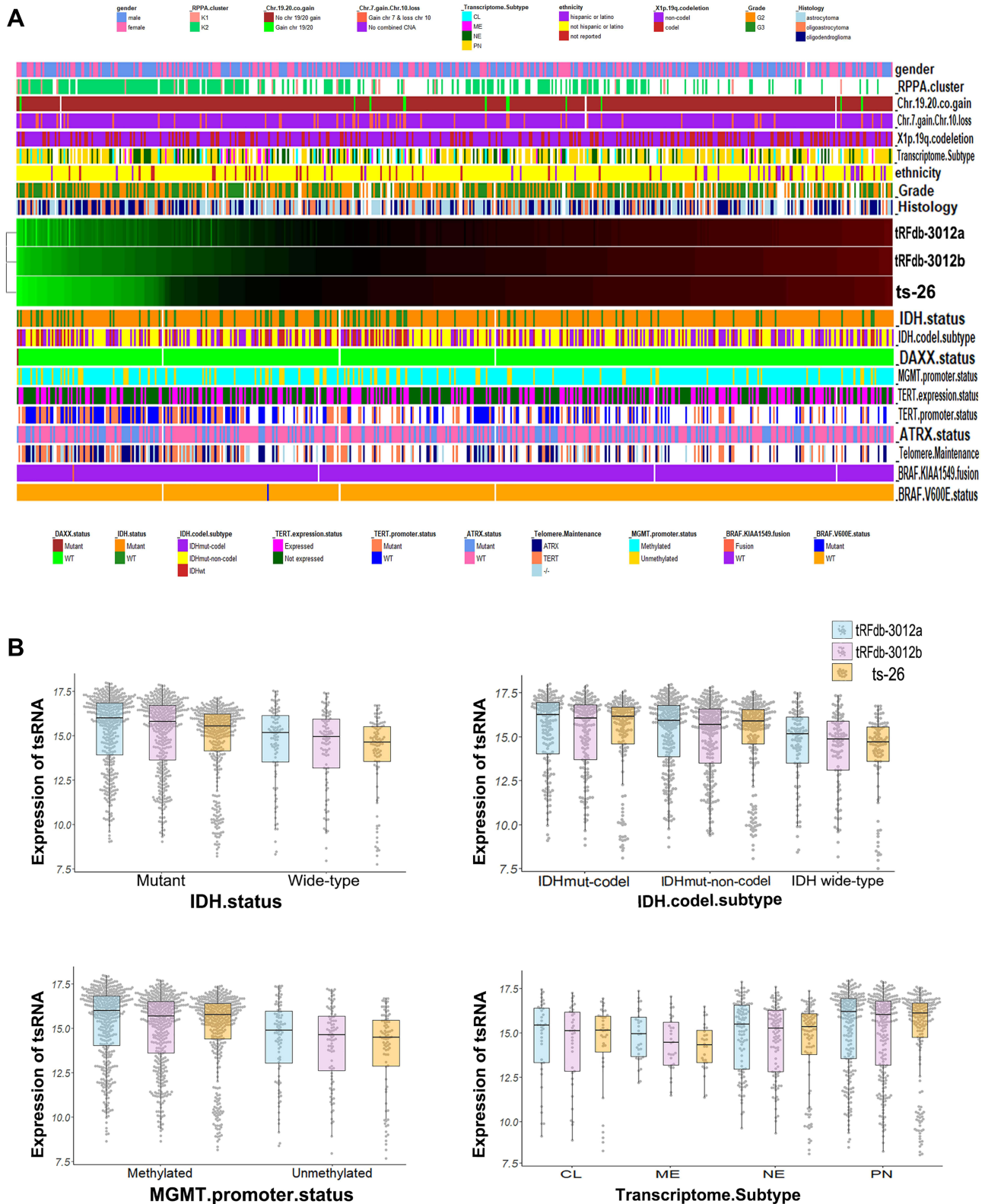


Figure 4 Correlation analysis of tRNA-Leu-CAA-derived tsRNA expression with the primary pathology characteristics. **(A)** Hierarchical clustering heatmap of tsRNA expression and the pathology characteristic parameters. **(B)** Bee-swarm plus box plots between tRNA-Leu-CAA-derived tsRNA (tRFdb-3012a, tRFdb-3012b, and ts-26) expression with four pathology parameters (*IDH* status, *IDH*-codeletion subtype, MGMT promoter status, and transcriptome subtype).

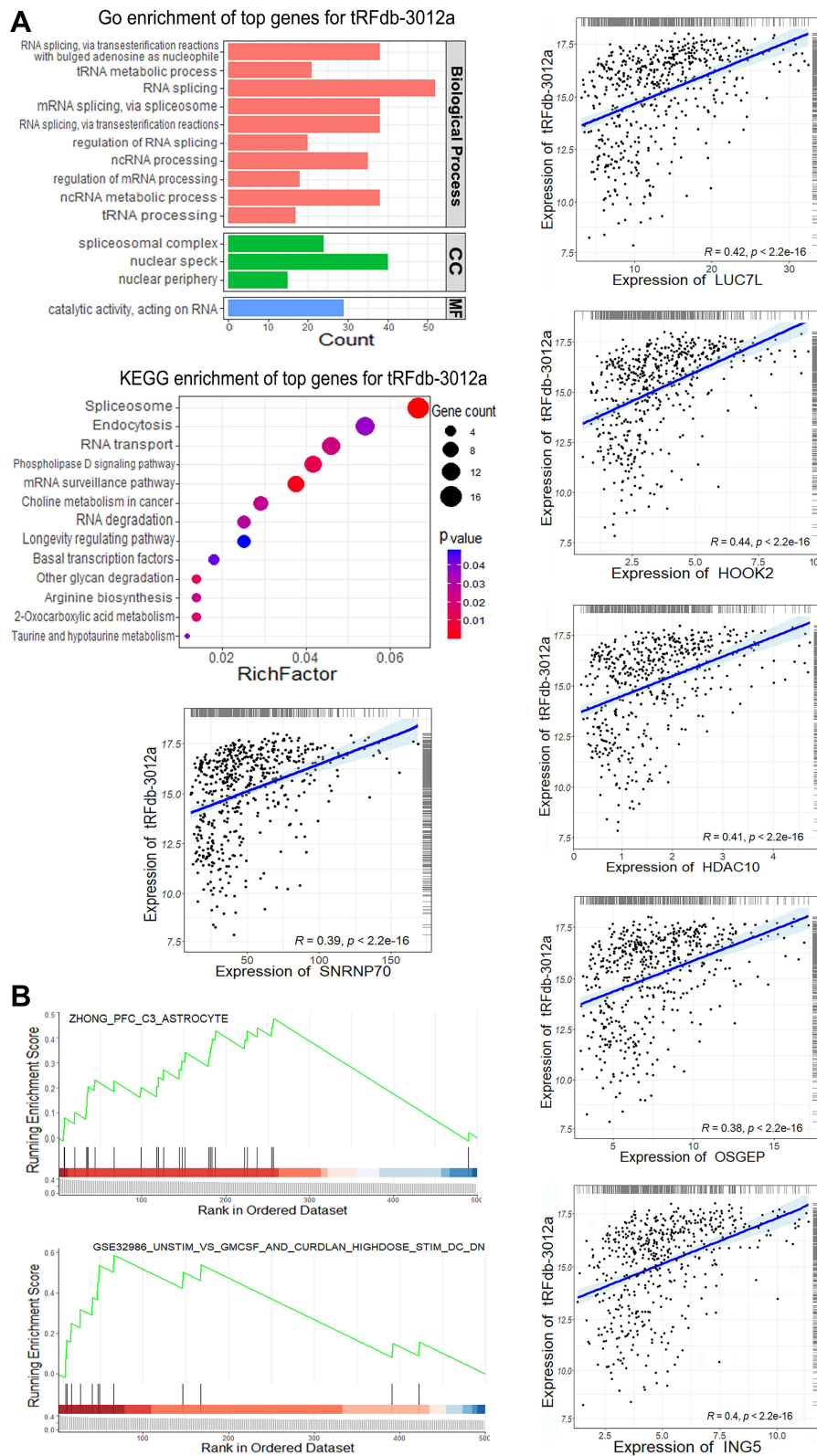


Figure 5 Enrichment analyses of tRFdb-3012a-related genes within TCGA-LGG datasets. **(A)** Top gene ontology (GO) terms, including biological process, cellular component (CC), and molecular function (MF), as well as the top KEGG pathway for tRFdb-3012a-related genes. **(A)** Correlation scatterplots of tRFdb-3012a and its correlated genes (*LUC7L*, *HOOK2*, *HDAC10*, *OSGEP*, *SNRNP70*, and *ING5*). **(B)** GSEA (gene set enrichment analysis) plots of two molecular signatures (astrocyte and STIM treatment response).

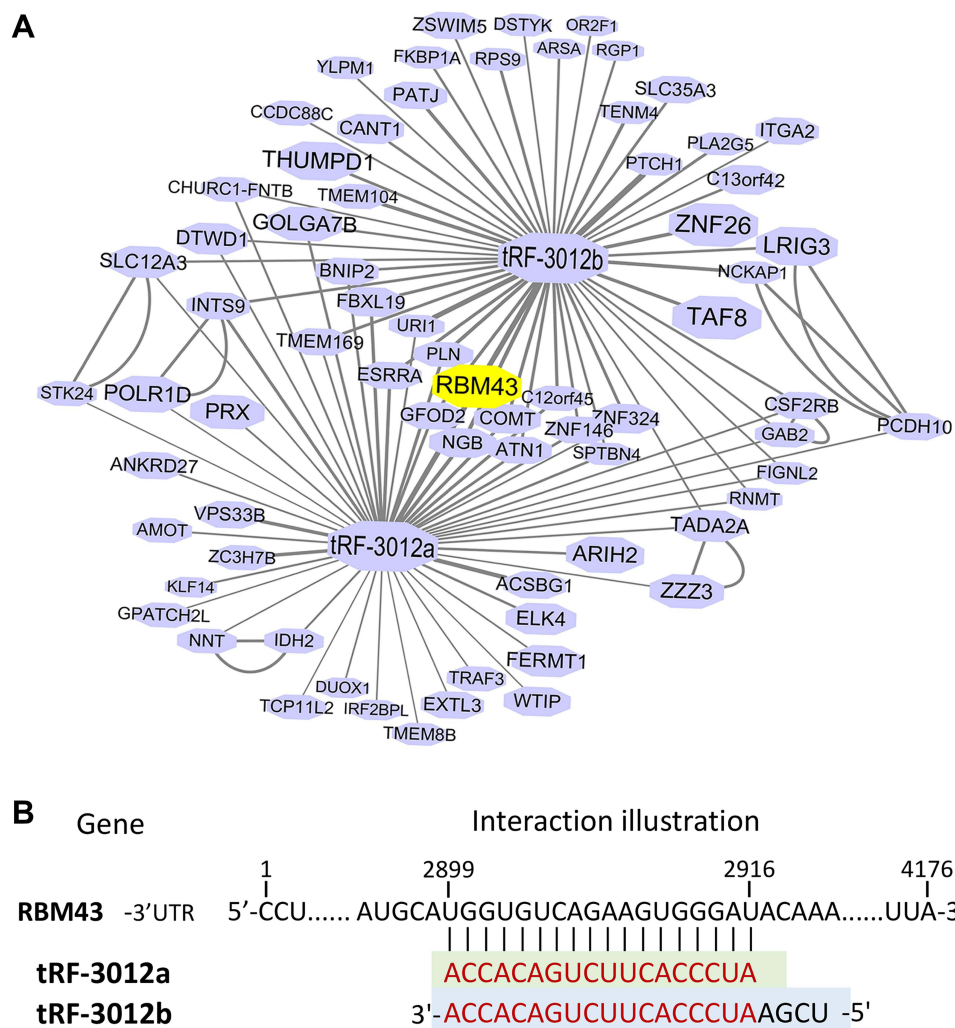


Figure 6 (A) Interaction networks that involved tRFdb-3012a/b and their target genes based on tRFtarget database. Nodes represent target genes and lines represent the interaction relationships. **(B)** Putative binding sites of tRFdb-3012a/b within the 3' untranslated regions of RBM43 based on bioinformatics prediction.

analysis showed that the overall survival of patients with gliomas with high RBM43 expression ($P < 0.0001$ and $P = 0.00013$) was significantly worse than that of those with low expression (Figure 7B).

The interactional relationships of ts-26 and its target genes are presented in Figure 8A. Among them, an oncogenic transcription factor of the homeobox family, *HOXA13*, was picked out for further analysis.²⁹ The 3' untranslated region of *HOXA13* mRNA contained a putative binding site of ts-26 (Figure 8B). The analysis of glioma samples within TCGA and GSE4290 datasets showed that *HOXA13* was highly expressed in gliomas compared to non-tumor controls (Figure 7C); however, ts-26 was decreased in diffuse gliomas. In addition, the clinical overall survival of glioma patients with high *HOXA13* expression ($P = 0.001$ and $P = 0.017$) was significantly worse than that of those with low expression (Figure 7D). These results suggest that tRFdb-3012a and tRFdb-3012b may directly target and regulate RBM43 expression, while ts-26 may directly target *HOXA13* and regulate *HOXA13* expression in gliomas.

Discussion

tsRNA is a type of small non-coding RNA derived from tRNA.¹ The identification and characterization of tsRNAs have been greatly accelerated over the past few years, owing to the emergence of next-generation sequencing (NGS) systems-level approaches to molecular genetics.² In this study, we first identified some tRNA fragments within the

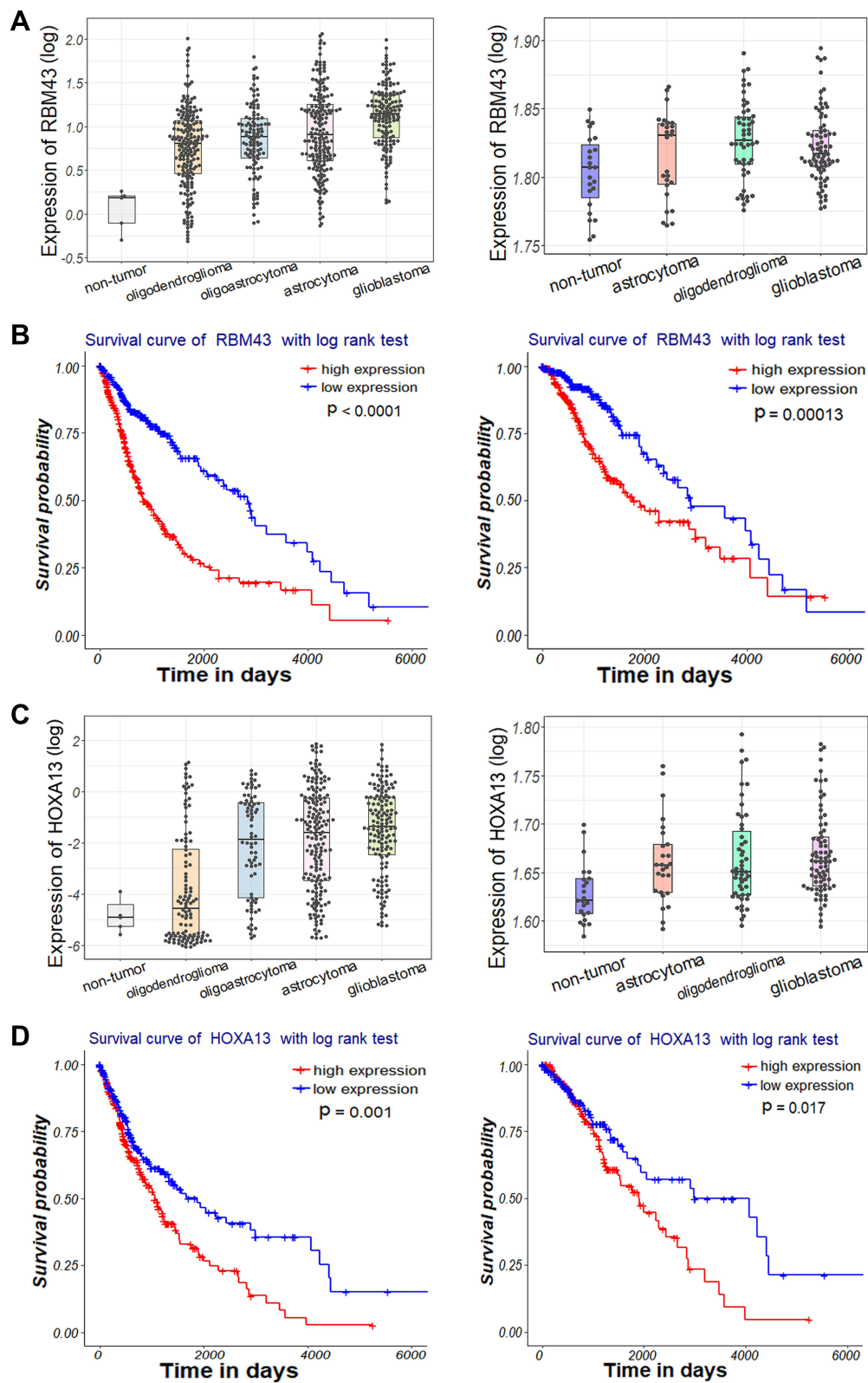


Figure 7 (A, C) Expression patterns of **(A)** RBM43 and **(C)** HOXA13 within diffuse glioma samples of TCGA (left) and GSE4290 (right) datasets. **(B, D)** Kaplan–Meier survival curve analysis with a log-rank test based on **(B)** RBM43 and **(D)** HOXA13 expression within glioma samples of TCGA dataset (left) and TCGA-LGG subtype dataset (right).

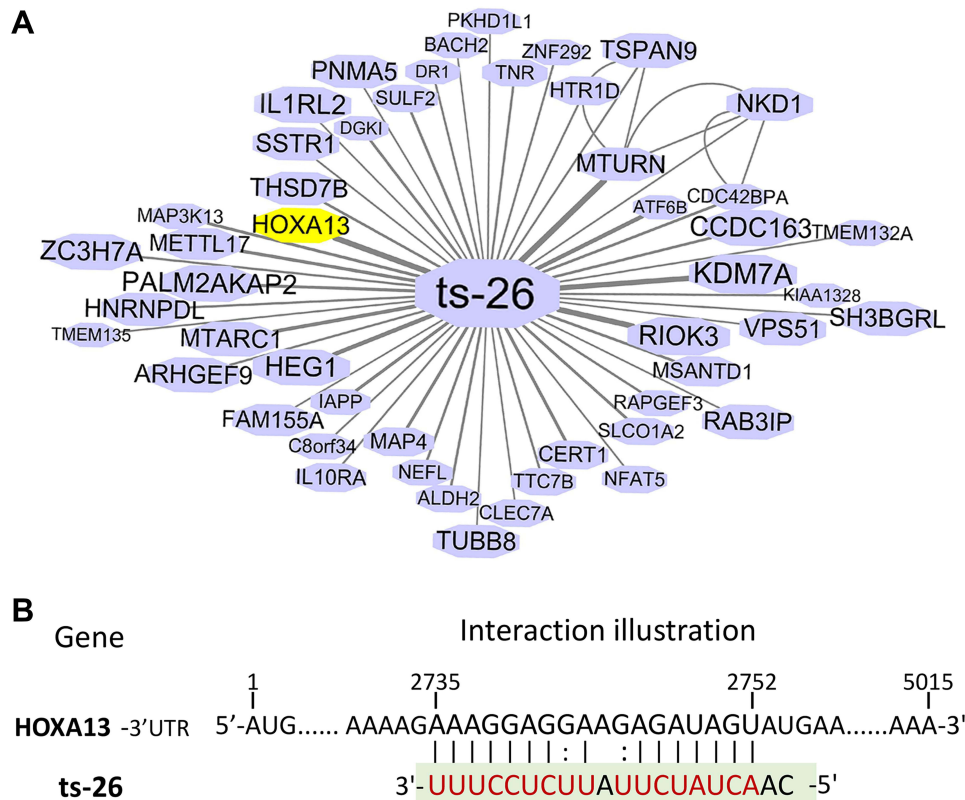


Figure 8 (A) Interaction networks that involved ts-26 and their target genes based on tRFtarget database. Nodes represent target genes and lines represent the interaction relationships. (B) Putative binding sites of ts-26 within the 3' untranslated regions of HOXA13 based on bioinformatics prediction.

sncRNA-seq data of glioma samples; most of fragments were derived from the 3' end (3'-tRFs) or 5' end (5'-tRFs) of mature tRNA. Among them, we picked out four tRNA fragments derived from tRNA-Leu-CAA, which is a leucine-transfer RNA. There are two 3'-tRF fragments (tRFdb-3012a and tRFdb-3012b) that can remap to the cleavage of the 3' end of mature tRNA-Leu-CAA, and a tRF-1 fragment (ts-26) that remaps to the downstream regions of primary tRNA-Leu-CAA genes. Further analyses showed that the expression of three tsRNAs (ts-26, tRFdb-3012a, and tRFdb-3012b) was significantly down-regulated in gliomas compared to non-tumor controls. Kaplan–Meier curve analyses demonstrated that overall survival of glioma patients with low ts-26 expression was significantly worse than that of those with high expression, and clinical survival of patients with gliomas with low tRFdb-3012a/b expression was also worse than that of those with high expression; this suggests that down-regulated tsRNAs (ts-26, tRFdb-3012a and tRFdb-3012b) were associated with poor survival outcome in glioma patients. Moreover, tRFdb-3012a/b expression was related to *IDH* mutation status and *MGMT* promoter mutation in gliomas, and tRFdb-3012a and tRFdb-3012b may tend to be up-regulated in patients with *IDH* mutation. As might be expected, heterozygous mutations in the catalytic arginine residues of isocitrate dehydrogenase gene (*IDH1/2*) are common in diffuse gliomas and acute myeloid leukemia, and contribute to the pathogenesis of several tumors.³⁰ According to the 2016 WHO CNS classification, *IDH* mutant and *MGMT* promoter methylation have been used as specific molecular features for the diagnosis, prognosis, and therapy of diffuse gliomas.¹⁰ In addition, our enrichment analysis showed that tRFdb-3012a was associated with specific GO terms, such as RNA splicing and processing (*LUC7L*, *OSGEP*, and *SNRNP70*), nuclear speck (*OSGEP* and *SNRNP70*), and catalytic activity (*HDAC10*); tRFdb-3012a-related genes may refer to RNA transport and spliceosome pathway (*SNRNP70* and *HOOK2*), as well as two GSEA molecular signatures (astrocyte and STIM treatment response signature). These results indicate that tRNA-Leu-CAA-derived tsRNAs (ts-26, tRFdb-3012a, and tRFdb-3012b) could be explored as

potential biomarkers for the diagnosis and prognosis of gliomas, and down-regulated tRFdb-3012a/b may play a tumor-suppressor role in glioma progression through a specific signaling pathway.

A growing number of studies have linked the tRNA fragments to some cancers, such as lung, colorectal,⁷ and breast cancer.⁸ They demonstrated that some tsRNAs were differentially expressed in the tissues of neoplasm patients. These dysregulated tsRNAs may serve as novel candidates for diagnosis and prognosis biomarkers.⁹ For instance, Tao et al determined the expression profile of tsRNAs in colorectal cancers by a sequencing approach, and found an internal tsRNA derived from mature histidine tRNA, 5'tiRNA-His-GTG, which is up-regulated in cancer tissues.⁷ Their experiments showed that 5'tiRNA-His-GTG promoted cancer cell pro-proliferation and tumor growth in nude mice, and it may be a carcinogenic factor in the development and progression of colorectal cancers.⁷ Although the biological functions of tsRNAs are complex and need further elaboration, current knowledge of their function has been a key focus in cancer research. Several studies have described that tRNA fragments can bind to argonaute proteins in a manner similar to miRNAs, and play an important role in expression silencing by targeting the 3' untranslated region of specific genes.² For instance, Goodarzi et al identified some novel tsRNAs derived from tRNA (Glu), tRNA (Asp), tRNA (Gly), and tRNA (Tyr) through hypoxic stress induction. These tRNA fragments could bind and displace the 3' untranslated region of the *YBX1* gene, which encoded a RNA-binding protein, and then suppressed the stability of multiple oncogenic transcripts such as *EIF4EBP1* and *AKT1*, leading to tumor-suppressive and metastasis-suppressive activity in breast cancer cells.⁸ Our bioinformatics analysis showed the interactional relationship of tRNA-Leu-CAA-derived tsRNAs (such as tRFdb-3012a, tRFdb-3012b, and ts-26) and their predicted target genes. Among them, tRFdb-3012a/b may directly bind to the 3' untranslated region of *RBM43*, and ts-26 may directly target the 3' untranslated region of *HOXA13* mRNA, and the two genes were both highly expressed in diffuse gliomas. Survival curve analyses showed that up-regulated *RBM43* and *HOXA13* were significantly associated with poor survival outcome of glioma patients. Previous research demonstrated that *RBM43* is an RNA binding motif protein and has been found to be critical to cancer risk.²⁸ *HOXA13* is an oncogenic transcription factor of the homeobox families, and functions as a key role in cancer development and progression.²⁹

In summary, our investigation identified four tRNA fragments derived from tRNA-Leu-CAA within the sncRNA-seq datasets of glioma samples. In particular, three tsRNAs (ts-26, tRFdb-3012a, and tRFdb-3012b) were significantly down-regulated in gliomas compared to non-tumor controls, and the expression of tRFdb-3012ab and ts-26 was associated with clinical survival outcome and *IDH* mutant status of diffuse glioma patients. Down-regulated tRFdb-3012a/b may have a tumor-suppressive role in glioma progression via a specific signaling pathway. In addition, tRFdb-3012a/b may directly target the 3' untranslated region of *RBM43*, while ts-26 may directly target the 3' untranslated region of *HOXA13* and regulate its expression in gliomas. These results imply that tRNA-Leu-CAA-derived tsRNAs (ts-26, tRFdb-3012a, and tRFdb-3012b) may regulate tumor progression of diffuse glioma, and could be explored as novel diagnostic and prognostic biomarkers for diffuse gliomas.

Data Sharing Statement

All of the data are described within the manuscript. The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Consent for Publication

All authors have agreed with the content of the manuscript for publication.

Acknowledgments

This work was supported by grants from the National Natural Science Foundation of China (no. 82003854), and sponsored by the Natural Science Foundation of Chongqing, China (cstc2021jcyj-msxmX0302).

Disclosure

The authors declare that they have no conflicts of interest.

References

1. Kim HK, Fuchs G, Wang S, et al. A transfer-RNA-derived small RNA regulates ribosome biogenesis. *Nature*. 2017;552(7683):57–62. doi:10.1038/nature25005
2. Shi J, Zhang Y, Zhou T, et al. tsRNAs: the Swiss army knife for translational regulation. *Trends Biochem Sci*. 2019;44(3):185–189. doi:10.1016/j.tibs.2018.09.007
3. Xie Y, Yao L, Yu X, et al. Action mechanisms and research methods of tRNA-derived small RNAs. *Signal Transduct Target Ther*. 2020;5(1):109. doi:10.1038/s41392-020-00217-4
4. Su Z, Wilson B, Kumar P, et al. Noncanonical roles of tRNAs: tRNA fragments and beyond. *Annu Rev Genet*. 2020;54:47–69. doi:10.1146/annurev-genet-022620-101840
5. Telonis AG, Magee R, Loher P, et al. Knowledge about the presence or absence of miRNA isoforms (isomiRs) can successfully discriminate amongst 32 TCGA cancer types. *Nucleic Acids Res*. 2017;45(6):2973–2985. doi:10.1093/nar/gkx082
6. Kumar P, Mudunuri SB, Anaya J, et al. tRFdb: a database for transfer RNA fragments. *Nucleic Acids Res*. 2015;43(Database issue):D141–5. doi:10.1093/nar/gku1138
7. Tao EW, Wang H-L, Cheng WY, et al. A specific tRNA half, 5'tiRNA-His-GTG, responds to hypoxia via the HIF1alpha/ANG axis and promotes colorectal cancer progression by regulating LATS2. *J Exp Clin Cancer Res*. 2021;40(1):67. doi:10.1186/s13046-021-01836-7
8. Goodarzi H, Liu X, Nguyen HB, et al. Endogenous tRNA-derived fragments suppress breast cancer progression via YBX1 displacement. *Cell*. 2015;161(4):790–802. doi:10.1016/j.cell.2015.02.053
9. Yu M, Lu B, Zhang J, et al. tRNA-derived RNA fragments in cancer: current status and future perspectives. *J Hematol Oncol*. 2020;13(1):121. doi:10.1186/s13045-020-00955-6
10. Louis DN, Perry A, Reifenberger G, et al. The 2016 World Health Organization classification of tumors of the central nervous system: a summary. *Acta Neuropathol*. 2016;131(6):803–820. doi:10.1007/s00401-016-1545-1
11. Zou H, Wu L-X, Yang Y, et al. lncRNAs PVT1 and HARI1 are prognosis biomarkers and indicate therapy outcome for diffuse glioma patients. *Oncotarget*. 2017;8(45):78767–78780. doi:10.18632/oncotarget.20226
12. Yang K, Wu Z, Zhang H, et al. Glioma targeted therapy: insight into future of molecular approaches. *Mol Cancer*. 2022;21(1):39. doi:10.1186/s12943-022-01513-z
13. Ostrom QT, Patil N, Cioffi G, et al. CBTRUS statistical report: primary brain and other central nervous system tumors diagnosed in the United States in 2013–2017. *Neuro Oncol*. 2020;22(12 Suppl 2):iv1–iv96. doi:10.1093/neuonc/noaa200
14. Zou H, Wu L-X, Tan L, et al. Significance of single-nucleotide variants in long intergenic non-protein coding RNAs. *Front Cell Dev Biol*. 2020;8:347. doi:10.3389/fcell.2020.00347
15. La Ferlita A, Alaimo S, Veneziano D, et al. Identification of tRNA-derived ncRNAs in TCGA and NCI-60 panel cell lines and development of the public database tRFExplorer. *Database*. 2019;2019. doi:10.1093/database/baz115
16. Chan PP, Lowe TM. GtRNAdb 2.0: an expanded database of transfer RNA genes identified in complete and draft genomes. *Nucleic Acids Res*. 2016;44(D1):D184–9. doi:10.1093/nar/gkv1309
17. Hutter C, Zenklusen JC. The cancer genome atlas: creating lasting value beyond its data. *Cell*. 2018;173(2):283–285. doi:10.1016/j.cell.2018.03.042
18. Barrett T, Wilhite SE, Ledoux P, et al. NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Res*. 2013;41(Database issue):D991–5. doi:10.1093/nar/gks1193
19. Langmead B, Trapnell C, Pop M, et al. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol*. 2009;10(3):R25. doi:10.1186/gb-2009-10-3-r25
20. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics*. 2015;31(2):166–169. doi:10.1093/bioinformatics/btu638
21. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*. 2014;30(7):923–930. doi:10.1093/bioinformatics/btt656
22. Colaprico A, Silva TC, Olsen C, et al. TCGAAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res*. 2016;44(8):e71. doi:10.1093/nar/gkv1507
23. Peng Z, Chen Y, Cao H, et al. Protein disulfide isomerases are promising targets for predicting the survival and tumor progression in glioma patients. *Aging*. 2020;12(3):2347–2372. doi:10.18632/aging.102748
24. Wickham H. *ggplot2: Elegant Graphics for Data Analysis, in Use R!* Cham: Springer International Publishing; Imprint: Springer; 2016:1. online resource XVI, 260 pages 232 illustrations, 140 illustrations in color.
25. Liberzon A, Birger C, Thorvaldsdóttir H, et al. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst*. 2015;1(6):417–425. doi:10.1016/j.cels.2015.12.004
26. Li N, Shan N, Lu L, et al. tRFtarget: a database for transfer RNA-derived fragment targets. *Nucleic Acids Res*. 2021;49(D1):D254–D260. doi:10.1093/nar/gkaa831
27. Zou H, Wen C, Peng Z, et al. P4HB and PDIA3 are associated with tumor progression and therapeutic outcome of diffuse gliomas. *Oncol Rep*. 2018;39(2):501–510. doi:10.3892/or.2017.6134
28. Feng H, Liu J, Qiu Y, et al. RNA-binding motif protein 43 (RBM43) suppresses hepatocellular carcinoma progression through modulation of cyclin B1 expression. *Oncogene*. 2020;39(33):5495–5506. doi:10.1038/s41388-020-1380-7
29. Janmaat VT, Nesteruk K, Spaander MCW, et al. HOXA13 in etiology and oncogenic potential of Barrett's esophagus. *Nat Commun*. 2021;12(1):3354. doi:10.1038/s41467-021-23641-8
30. Flavahan WA, Drier Y, Liau BB, et al. Insulator dysfunction and oncogene activation in IDH mutant gliomas. *Nature*. 2016;529(7584):110–114. doi:10.1038/nature16490

Cancer Management and Research

Dovepress

Publish your work in this journal

Cancer Management and Research is an international, peer-reviewed open access journal focusing on cancer research and the optimal use of preventative and integrated treatment interventions to achieve improved outcomes, enhanced survival and quality of life for the cancer patient. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/cancer-management-and-research-journal>