Am J Cancer Res 2023;13(3):1107-1117
www.ajcr.us /ISSN:2156-6976/ajcr0148661

Original Article

Circ8199 encodes a protein that inhibits the activity of *OGT* by JAK2-STAT3 pathway in esophageal squamous cell carcinoma

Chen-Jia Li[1,2,3*], Dong-Ge Li[1,2*], En-Jie Liu[1,2], Guo-Zhong Jiang[1,2]

[1]Department of Pathology, School of Basic Medicine, Zhengzhou University, Zhengzhou 450001, Henan, PR China; [2]Department of Pathology, The First Affiliated Hospital of Zhengzhou University, Zhengzhou 450052, Henan, PR China; [3]Luoyang Traditional Chinese Medicine Hospital, Luoyang 471099, Henan, PR China. [*]Equal contributors.

Received December 20, 2022; Accepted February 12, 2023; Epub March 15, 2023; Published March 30, 2023

Abstract: Esophageal squamous cell carcinoma (ESCC) is an invasive malignant tumor with a high incidence rate and mortality. It is imperative to study its tumorigenesis and development for better treatment. CircRNA has been proven to play an important role in various cancers. Our previous studies found that the *circ8199* gene is associated with tumor prognosis. To further clarify the role of *circ8199* in ESCC, we performed functional experiments and found that overexpression of *circ8199* significantly inhibited the proliferation of ESCC cells and the activity of O-linked N-acetylglucosamine transferase (*OGT*) simultaneously. Further experiments demonstrated that *circ8199* could interact with *OGT*, leading to a decrease in *OGT*'s activity. The reduction of *circ8199* expression stimulated the binding activity between *OGT* and its downstream gene JAK2, promoting the *O-GlcNAc* glycosylation modification of JAK2 and activating the JAK2-STAT3 pathway. Our study indicated that *circ8199* regulates the JAK2-STAT3 pathway through *OGT*, providing a candidate mechanism for drug discovery and development.

Keywords: *Circ8199*, ESCC, *OGT*, *O-GlcNAc* glycosylation, JAK2-STAT3 pathway

Introduction

Esophageal squamous cell carcinoma (ESCC), the fourth leading cause of cancer death in China, is a highly aggressive malignant tumor [1]. It has high morbidity and mortality with less than 50% of 5-year survival rate and increasing incidence in recent years. Elucidating the pathogenesis of ESCC and identifying potential molecular markers and drug targets have important scientific and clinical significance.

Non-coding RNA is considered an important regulator in the biological process [2]. Circular RNA (circRNA) is a class of evolutionarily highly conserved non-coding RNA produced by linear RNA splicing [3]. Unlike the canonical linear RNA (containing 5' and 3' ends), circRNA forms a 3'-5' covalently closed RNA loop, which is not affected by RNA exonuclease and is more stable [4]. Studies have shown that circRNAs are involved in various physiological and pathological processes and play important roles in tumorigenesis and development [5].

Previous studies have found that circRNA regulates the expression of target genes as a miRNA sponge by inhibiting miRNA activity [4, 6]. Multiple circRNAs encode proteins with important biological functions in tumors. For example, circ-E-Cad translates small protein C-E-Cad in glioblastoma, which activates EGFR/EGFR viii and promote the self-renewal of glioma stem cells [7]; circ-SHPRH participates in the progression of glioma by encoding the polypeptide SHPRH-146aa [8]; circ-AKT3 is involved in glioblastoma cell proliferation by encoding AKT3-174aa [9]; circFBXW7 inhibits the progression of malignant glioma and breast cancer by encoding FBXW7-185aa [10, 11]; β-catenin-370aa encoded by circRNA inhibits GSK-3β-induced β-catenin degradation, activates the Wnt/β-catenin pathway, and promotes the growth of hepatocellular carcinoma [12]. Stu-

dies in various types of cancer show that proteins encoded by circRNAs play important roles in tumorigenesis and development. However, there are only a few reports on whether there are similar circRNAs encoding proteins in ESCC.

ATXN10 is a protein that directly interacts with O-Linked β-N-acetylglucosamine (*O-GlcNAc*) transferase (*O-GlcNAc* transferase, *OGT*) and regulates its activity [13, 14]. ATXN10 transfers the *O-GlcNAc* glycosyl group from UDP-*GlcNAc* to the hydroxyl group of naked serine/threonine residues in intracellular proteins, which participate in biological processes such as cell signaling, recognition adhesion, and receptor activation by affecting the spatial structure and stability of the peptide chain. Abnormal *O-GlcNAc* glycosylation modification is associated with tumor proliferation and metastasis during multidrug resistance and epithelial-mesenchymal transition [15]. Therefore, as a protein involved in post-translational modification, *O-GlcNAc* glycosylation is expected to be a candidate therapeutic target for tumor. Studies have shown that *O-GlcNAc* glycosylation also plays an important role in tumorigenesis and development of ESCC [16], but its specific molecular mechanism remains to be elucidated.

In our previous study, we sequenced and analyzed 73 ESCC tumors and paired surrounding normal mucosal tissues and found that hsa_circ-00008199 (called *circ8199*) is significantly associated with prognosis. This study further explored the functional roles and specific regulatory molecular mechanisms of *circ8199* in ESCC.

Materials and methods

Cell culture and stable cell lines

The KYSE30, KYSE70, KYSE140, KYSE150, KYSE180, KYSE270, KYSE410, KYSE450, KYSE510, KYSE520 and HET-1A cells were purchased from the American Type Culture Collection (ATCC). KYSEs and HET-1A were cultured in RPMI 1640 (Hyclone) supplemented with 10% fetal bovine serum (FBS) (Gibco) at 37°C with 5% CO_2. According to the manufacturer's protocol, the plasmids were transfected into ESCC cells with Lipofectamine 3000 (Invitrogen, CA). Cells were harvested and analyzed 24-36 h after transient transfection. For the generation of stable cell lines, cells transfected post 24 h were digested with 0.25% trypsin and were transferred to the plates for further culture with DMEM medium containing 500 μg/l G418 and 10% FBS for 10 days till resistant cell clones was observed.

CCK8 assay

Cells were seeded into 96-wells plate at 1×10^3/well, and a Cell Counting Kit-8 kit was performed every 24 h to monitor cell proliferation according to the manufacturer's protocol. The light absorption value of each well at a wavelength of 450 nm was measured using a Microplate Reader (Thermo Fisher).

Migration assay

4×10^4 cells in 200 μL of serum-free RPMI-1640 medium were placed in the upper chamber of Transwell insert (BD Biosciences, Franklin, Laskes, NJ, USA) with 8 μm pore size filter. The well contained 600 μL RPMI-1640 with 20% FBS. After incubation for 48 h, the cells that migrated through the filter were stained with crystal violet and photographed under the microscope for statistics.

Colony formation assay

2×10^3 stably transfected cells were blended with pre-warmed 0.6% soft-agar containing RPMI-1640 medium and 10% FBS, and subsequently were poured on top of 1.2% agar in a 6-well plate. After 14 days, 50 cells were counted as a clone, and the number of colonies were photographed and counted for statistical analysis.

qRT-PCR

Total RNA was isolated by TRIzol™ Reagent (Thermo Fisher) and uses PrimeScript™ RT reagent Kit (Takara), which contains components needed for reverse transcription to facilitate cDNA synthesis. Then the relative expression of mRNA was further quantified by TB Green® Premix Ex Taq™ II kit (Takara) following the instructions. GAPDH is used as an internal reference, and $2^{-\Delta\Delta Ct}$ is the relative expression level of the gene. The primer sequences were designed as follows: *circ8199* divergent forward primer: 5'-ATGTGGCCAATGGGTTTAAG-3', and reverse forward primer: 5'-GCAAGCTCAACAGCATGAGA-3'; *circ8199* convergent forward primer: 5'-GAAACAGCACCCAGGACTA-3', and reverse forward primer: 5'-CTGATTGAATTCTGG-

TTCACAG-3'; GAPDH forward primer: 5'-GAGTCCACTGGCGTCTTCA-3' and reverse primer: 5'-TGATGATCTTGAGGCTGTTGTC-3'; ATXN10 forward primer: 5'-AGTCTCATCTCATTCGTCTG-3' and reverse primer: 5'-GGCTGTTGTCTTCGGTAA-3'; c-Myc forward primer: 5'-TGAGGAGACACCGCCCAC-3' and reverse primer: 5'-CAACATCGATTTCTTCCTCATCTTC-3'; Bcl-2 forward primer: 5'-GATGTGATGCCTCTGCGAAG-3' and reverse primer: 5'-CATGCTGATGTCTCTGGAATCT-3'; Survivin forward primer: 5'-AGAACTGGCCCTTCTTGGAGG-3' and reverse primer: 5'-CTTTTTATGTTCCTCTATGGGGTC-3'.

RNA fluorescence in situ hybridization

3×10^5 cells were cultured for 24 h on the 24-well plate round cell slide, washed twice in RNase-free PBS, 4% paraformaldehyde (Electron Microscopy Science, Hatfield, PA) was used to fix the cells at room temperature for 30 min, then permeabilized for 15 min at room temperature with 1% TritonX-100, and 5% RNase free BSA was used to seal the cells at room temperature for 30 min. Then used 50 μl of Hybridization Buffer (10% dextran sulfate, 10% formamide, 2×SSC, 0.2 mg/ml RNase free BSA) containing 5' labelled CY3 DNA probes (5 ng/sample) incubated at 37°C overnight. The FISH probe was designed to span the splice junction of *circ8199*: 5'-CTGGGTGCTGTTTCTCTTGTCTTGGT-3'. 4,6-diamidino-2-phenylindole (DAPI) staining was performed for 5 minutes and observed with fluorescence microscope.

Cell fractionations

All manipulations were at 4°C. Cells were fractionated using the Nuclear/Cytosol Fractionation kit (BioVision, Milpitas, CA) following the manufacturer's instructions. In Brief, cells were collected by trypsinization and washed with PBS, then pelleted and resuspended in 200 μl of CEB-A for incubation on ice for 10 min. Afterward, CEB-B was added to the lysates and incubated for 1 min on ice. The lysates were spun for 5 min in a microcentrifuge to yield a cytosolic supernatant and nuclear pellet. Both cytoplasmic and nuclear fractions were stored at -80°C until use.

Constructs

Circ8199 expression plasmid was generated by chemical gene synthesis the sequence of exon 2-9 of ATXN10. An additional cyclization promoter sequence from 400-547 nt Δ440-500 nt of ZKSCAN1 and sequence from 1437-1782 nt Δ1449-1735 nt of ZKSCAN1 were added upstream and downstream, respectively. The synthesized gene was cloned into pcDNA3.1 using the NheI- and HindIII-recognition sequences. 3xflag was placed before the stop codon, which was designated as *circ-8199*-3xflag. Furthermore, the start codon "ATG" in the *circ-8199*-3xflag was mutated to "ATC" in the *circ-8199*-3xflag-ATC. *circ-8199*-3xflag-NC was used as the corresponding negative control derived from *circ-8199*-3xflag, in which the cyclization sequence downstream was removed. The linear sequence of *circ8199* with Kozak sequence before ATG was directly cloned into pcDNA3.1, which was designated as *circ-8199*-3xflag-235aa, a positive control for 235aa. The potential IRES sequences of *circ8199* were synthesized and inserted in the middle of Rluc and Luc using KpnI and EcoRI introduced by primers. The lentivirus vector containing shRNA targeting *circ-8199* was generated by GenePhaema (Shanghai, China).

Luciferase reporter assay

Cells were seeded in 12-well plates and transfected with Luc2-IRES-Reporter vectors or the control. After 48 h post-transfection, the luminescent values for firefly luciferase reporter activities were detected and normalized by the Renilla internal control using the Dual luciferase reporter assay system (Promega). Results were determined as relative luciferase activities.

Immunoprecipitation

The cells were seeded in a 100 mm dish and transfected with indicated vectors. After 48 h, cells were washed three times with PBS (PH=7.4), and then were lysed in ice-cold IP lysis buffer (including RIPA, protease inhibitors, and phosphatase inhibitors). After lysis for 30 min, cells were collected and centrifuged at 12000 rpm for 10 min at 4°C. The soluble fractions in the supernatant was subjected to IP using the indicated primary antibodies at 4°C overnight by rotation. Beads were added and incubated at 4°C for 2 h, followed by washing with PBS (PH=7.4) three times. The eluent with 1×SDS loading buffer was subjected to immunoblotting analyses.

Western blot assay (WB)

Proteins were extracted by a lysis buffer and quantified using a bicinchoninic acid (BCA) Protein quantitative kit (KeyGen Biotech). Protein lysates were separated using 10% SDS-PAGE and transferred to a PVDF membrane (Millipore). The membrane was then incubated with the specific primary antibodies and second antibodies. The bands were developed by Pierce ECL Western Blotting Substrate (Thermo Scientific). Mouse monoclonal antibody recognizing JAK2 (AT53B7) were obtained from GeneTex. Rabbit polyclonal antibodies that recognize Phospho-Jak2 (Tyr1007/1008) (#3776) were purchased from Cell Signaling Technology. Rabbit polyclonal antibodies recognizing *OGT* (11576-2-AP), GAPDH (10494-1-AP), and mouse monoclonal antibodies recognizing Flag (66008-4-Ig) were from Proteintech. Mouse monoclonal antibody recognizing *O-GlcNAc* (PTM-952) were purchased from PTM Biolabs Inc. Rabbit polyclonal antibodies recognizing STAT3 (AF1492), Phospho-STAT3 (Tyr705) (AF-1276) were purchased from Beyotime Biotechnology. HRP Goat Anti-Rabbit IgG (AS014) was purchased from ABclonal Technology. HRP Goat Anti-Mouse IgG (CW0102) was purchased from CoWin Biotech (CWBIO).

Glycosyltransferase assay

Glycosyltransferase activity of co-IP from KYSE70 transfected with *circ-8199*-3xflag-NC or *circ-8199*-3xflag-235aa was determined with the Glycosyltransferase Activity Kit (EA001, R&D system/USA). A glycosyltransferase reaction was carried out in 50 μl of reaction buffer containing 50 mM Tris HCl (PH=7.5), 100 μM UDP-glucose, the varying concentration of substrate (dissolved in DMSO), and 5 μg of precipitates at room temperature for 20 min, according to the manufacturer's instructions. The absorbance value was measured at 620 nm with a microplate reader.

Statistical analysis

Statistical analyses were performed using SPSS20.0 software (IBM). The differences between groups were examined by the two-tailed Student's t-test. Data were presented as the mean ± SD of at least 3 independent experiments. Differences were considered significant if $p < 0.05$; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

Results

Identification of circ8199 in ESCC

To explore the expression patterns of circRNAs in ESCC, we have performed circular RNA sequencing on 73 fresh ESCC and paired paracancerous tissue specimens and found that the overall survival in the *circ8199* high expression group was significantly better than that in the low expression group. *Circ8199* was highly expressed in cancer tissues, as shown by RT-qPCR in another 80 fresh ESCC tissue specimens [17]. This indicates that *circ8199* acts as a candidate prognostic biomarker in ESCC and plays an important role in the development of ESCC.

Circ8199 is derived from exons 2-9 of ATXN10 gene [18]. To further investigate the presence of *circ8199* in ESCC, the back-splice junction was confirmed by Sanger sequencing (**Figure 1A**, **1B**). In addition, we designed two sets of primers: divergent primers to amplify circular transcripts and convergent primers to detect linear transcripts. We then created two other sets of primers to amplify the circular and linear transcripts of *circ8199* in cDNA or gDNA of ESCC tissue or KYSE520 cells. Our results showed that the circular form could be amplified only in cDNA, whereas the linear transcripts were amplified in both cDNA and gDNA (**Figure 1C**). Furthermore, the linear transcripts of *circ8199* were degraded after RNase R treatment (**Figure 1D**).

The *circ8199* conjunction sequence was synthesized and labeled as a probe, and RNA fluorescence in situ hybridization (FISH) was carried out in KYSE520 and KYSE270 ESCC cells (**Figure 1E**), combined with the results of nucleocytoplasmic separation (**Figure 1F**), suggesting that *circ8199* is mainly localized in the nuclear.

Circ8199 inhibits cell proliferation and migration of ESCC in vitro

The fluorescence quantitative RT-PCR method was used to measure the expression levels of *circ8199* in 10 ESCC cell lines and one normal esophageal cell line (**Figure 2A**). In order to evaluate the biological functions of *circ8199* in ESCC, we constructed a *circ8199* expression vector on the base of the previous report [19]. The ZKSCAN1 gene exon2 front 400-547 nt

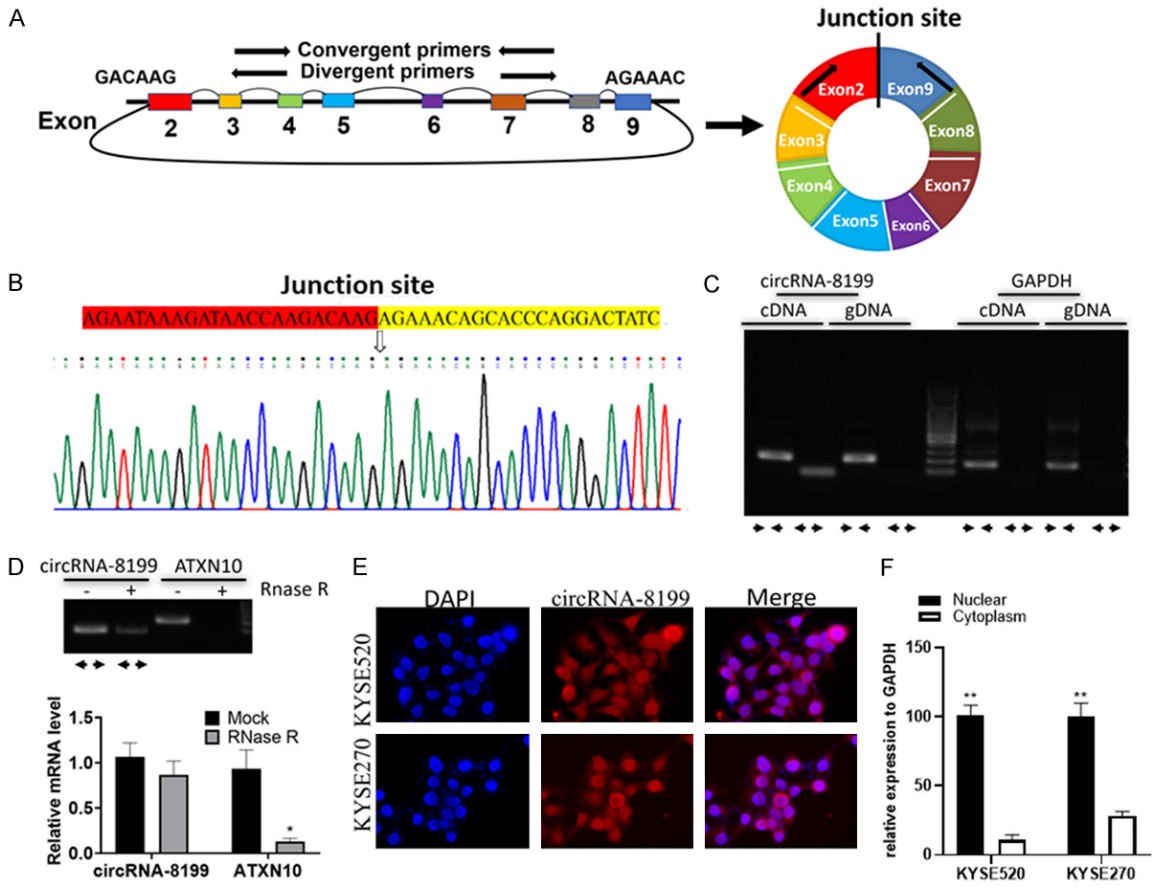


Figure 1. Identification and characteristics of *circ8199* in ESCC. (A) The schematic diagram of *circ8199* arose from exon 2-9 of ATXN10 gene. (B) Sanger sequencing was carried out to confirm head-to-tail splicing. (C) *Circ8199* was detected only in cDNA but not in gDNA. (D) RNase R degraded linear RNA but not *circ8199*. (E and F) Fish assays and nucleocytoplasmic separation assays indicated that *circ8199* is mainly located in the nuclear. The magnification is 40× in (E).

Δ440-500 nt was used as the upstream flanking sequence, which includes 40 nt of the AluSq2 element. The downstream flanking sequence spans from 1437-1782 nt Δ1449-1735 nt following the exon3 of the ZKSCAN1 gene, containing the AluSz and AluJr elements, which support circularization. The combined sequence was subcloned into the pcDNA3.1 (+) expression vector, in which CMV promoter was used (**Figure 2B**). As *circ8199*-NC only had a flanking ring-forming frame (IR) at one end, it could not form a circle correctly. Moreover, *circ8199*-OE is a plasmid vector for circular overexpression of *circ8199*. These constructed vectors were transfected into KYSE30 and KYSE70 ESCC cells which low expression of *circ8199*, respectively. Stable cell lines were screened for cell proliferation and clone formation experiments. Our results showed that the overexpression of *circ8199* significantly inhibited the growth (**Figure 2C**), migration (**Figure 2D** and **2E**) and clonogenicity (**Figure 2F**) of ESCC cells.

Circ8199 encodes a polypeptide that inhibits the proliferation of ESCC cells

The bioinformatics prediction website (http://reprod.njmu.edu.cn/circrnadb) shows that *circ8199* has a complete open reading frame of 708nt which encodes a protein/polypeptide of 235 amino acids similar to the AA sequence of ATXN10 from M179 to K391 (**Figure 3A**). After circularization, an extra polypeptide formed with sequence RNSTQDYLPKSSGYPKEIFSCC that is non-homologous to ATXN10. As cap structures are absent in circRNA, internal ribosome entry sites (IRESs) of RNA are required to

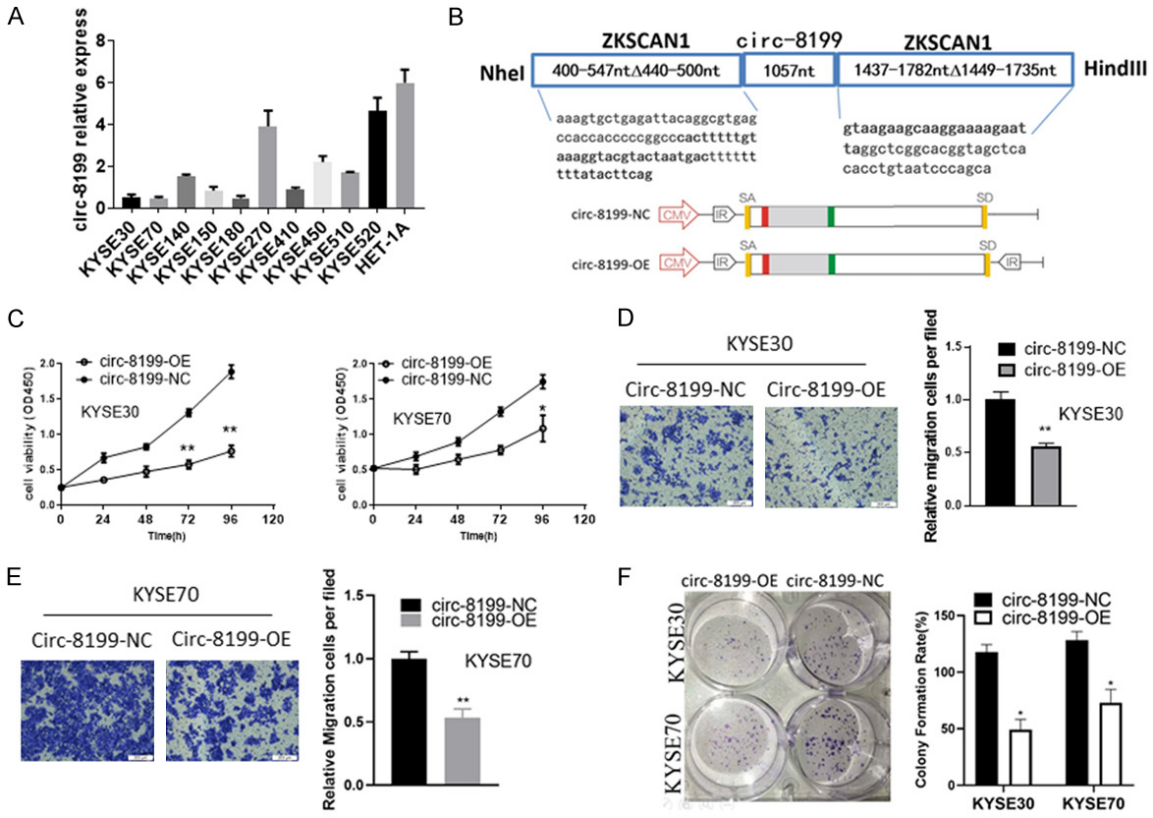


Figure 2. *Circ8199* promotes ESCC cell proliferation, immigration and colony formation in vitro. (A) *Circ8199* expression in a panel of ESCC cells was measured by qRT-PCR. (B) The schematic diagram of *circ8199* expression vector using the flanking sequences from ZKSCAN1 gene. IR, inverted repeats; SA, splice acceptor; SD, splice donor; CMV, cytomegalovirus promoter. The expression of *circ8199* inhibits proliferation (C), migration (D and E) and colony formation (F) of KYSE30 and KYSE70, respectively. All the results were shown as mean ± SD (n=3), $^{*}P < 0.05$; $^{**}P < 0.01$ (Student's test).

activate the translation process. We found a high-scoring IRES sequence in *circ8199* (**Figure 3B**). To verify whether the predicted IRES region has the ability to initiate translation, the HSV promoter between the two luciferase sequences in the Psicheck2 vector was replaced by the predicted IRES sequence with its downstream sequence (1-418, pre-initiation codon sequence) (named *circ8199*-IRES) (**Figure 3C**). The results of the dual-luciferase reporter experiment showed that the *circ8199*-IRES could initiate the translation of the downstream open reading frame (**Figure 3D**).

For further verification, we constructed serial vectors for cell experiments: *circ8199*-3xflag is a *circ8199*-235aa overexpression vector containing 3xFlag tag; *circ8199*-3xflag-ATC contains 3xflag tag with ATG mutated to ATC of *circ8199*-235aa which could be circularized but not translated; *circ8199*-3xflag-NC is a vector containing the 3xflag tag and with the downstream IR deleted (negative control that cannot be circularized); *circ8199*-3xflag-235aa is a linear expression vector only for the predicted ORF, as a positive control for 235aa (**Figure 3E**). The above vectors were transfected into KYSE30 respectively. Western blot results showed that the protein expression of *circ8199*-3xflag was detected. When the start codon of *circ8199* was mutated from ATG to ATC, the expression of *circ8199*-3xflag could not be detected (**Figure 3F**), indicating that *circ8199* encodes polypeptides.

Circ8199-3xflag-NC, *circ8199*-3xflag, *circ8199*-3xflag-ATC and *circ8199*-3xflag-235aa were transferred into KYSE30 and KYSE70 ESCC cells, respectively. The results showed that the overexpression of *circ8199*-3xflag-235aa or *circ8199*-3xflag significantly inhibited the proliferation of ESCC cells, and *circ8199*-3xflag-

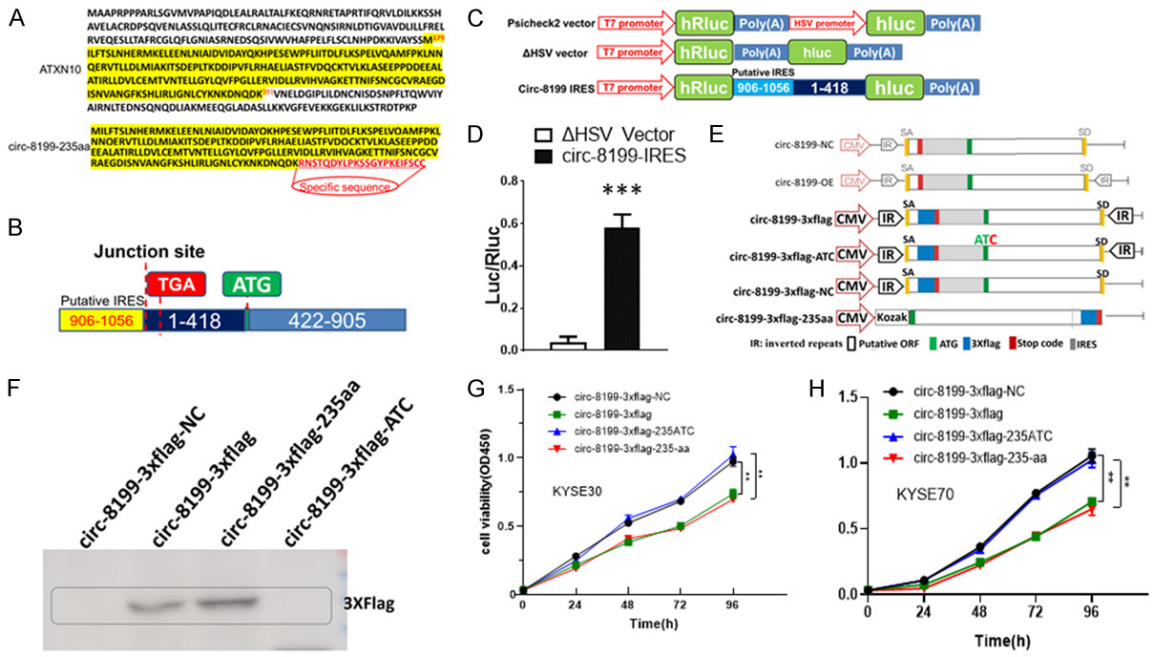


Figure 3. *Circ8199* encodes a predicted 235aa protein which inhibits cell proliferation. A. The amino acid sequences compared between ATXN10 and *circ8199*-235aa, and the extra sequence was indicated in red. B. The sequence of *circ8199* with putative IRES was cloned into pcDNA3.1, in which ATG of *circ8199*-235aa was mutated to TGA. C. The schematic diagram of luciferase expression vector. The HSV promoter between two luciferase genes in pSicheck2 vector was replaced with *circ8199* IRES and its downstream. D. Luciferase activity was measured using a luminometer. Rluc normalized Fluc activity is represented in *circ8199*-IRES and control infected cells. Error bars indicate the standard deviations (SD) of at least three repeats. Student's t-test was used to determine *P*-values; $^{***}P < 0.001$. E. The expression vectors of 3xflagged putative ORF sequence of circRNA8199: *circ8199*-3xflag; *circ8199*-3xflag-ATC; *circ8199*-3xflag-NC *circ8199*-3xflag-235aa. F. *circ8199*-235aa was detected by immunoblotting using anti-flag antibody in KYSE30. G and H. Upregulation of *circ8199*-235aa inhibits the cell proliferation in KYSE30 and KYSE70, respectively (n=5). $^{**}P < 0.01$.

ATC mutant almost abolished *circ8199*-3xflag-induced inhibited the proliferation of ESCC cells (**Figure 3G** and **3H**).

Circ8199-235aa directly binds OGT and inhibits its enzyme activity

We measured the total protein *O-GlcNAc* glycosylation levels in ESCC fresh tissue specimens by WB. The protein *O-GlcNAc* glycosylation levels in the *circ8199* overexpression group were significantly lower than those in the low *circ8199* expression group, indicating that the level of *O-GlcNAc* glycosylation in ESCC was negatively correlated with the expression of *circ8199* (**Figure 4A** and **4B**). Co-immunoprecipitation experiment (CO-IP) showed that *OGT* interacted with *circ8199*-3xflag-235aa (**Figure 4C**). CO-IP experiment with flag antibody further confirmed this finding that *OGT* and *circ8199*-3xflag-235aa could interact with each other (**Figure 4D**).

The KYSE70 cell line was transfected with *circ8199*-3xflag-NC and *circ8199*-3xflag-235aa plasmids, respectively. The CO-IP experiment was performed with *OGT* protein, and then the activity of *OGT* was detected according to a commercial kit (R&D, #EA001). The results showed that overexpression of *circ8199*-3xflag-235aa in KYSE70 cells could significantly inhibit the enzymatic activity of *OGT* (**Figure 4E**).

Down-regulation of circ8199 activates OGT, thereby promoting the level of JAK2 O-GlcNAc glycosylation and activating the JAK2-STAT3 pathway

In order to explore the signal pathways of the *circ8199* protein involved, we performed KEGG pathway enrichment analysis on two groups of tumor samples, the high-expression group and low-expression group, based on the expression level of *circ8199* from the high-throughput

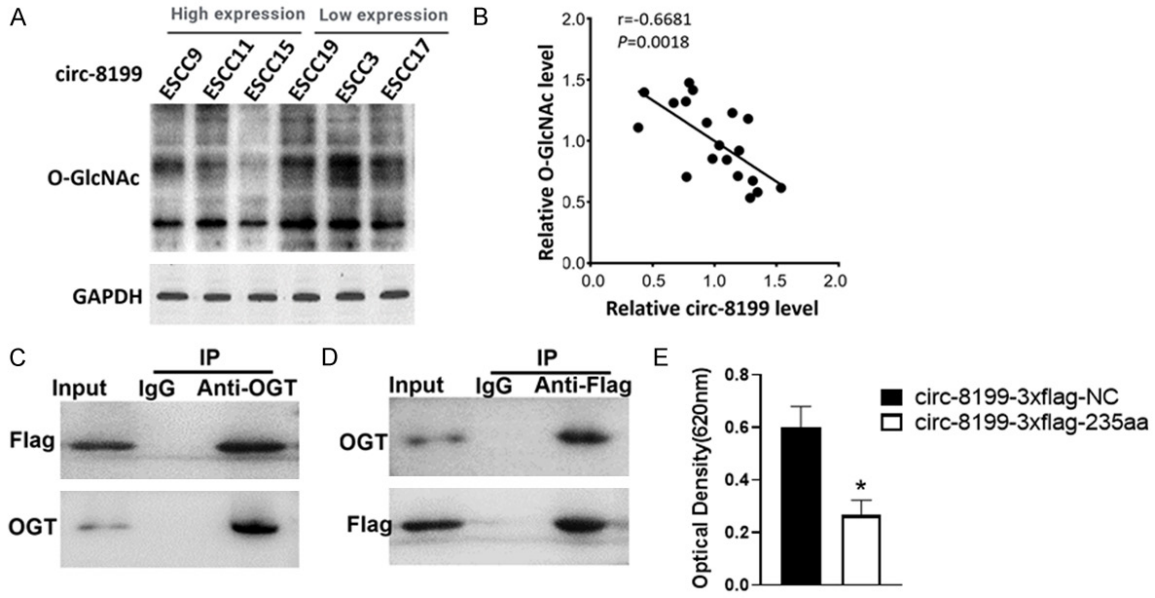


Figure 4. *Circ8199*-235aa binds *OGT* and inhibits its activity. A and B. The results showed that the expression of *circ8199* were negative correlated with *O-GlcNAc* glycosylation of protein in ESCC tissues. C and D. Immunoprecipitation and western blot showed *circ8199*-235aa interacts with *OGT* in KYSE70. E. *circ8199*-235aa inhibits *OGT* activity in KYSE70. The results were shown as mean ± SD (n=3), $^{*}P < 0.05$ (Student's test).

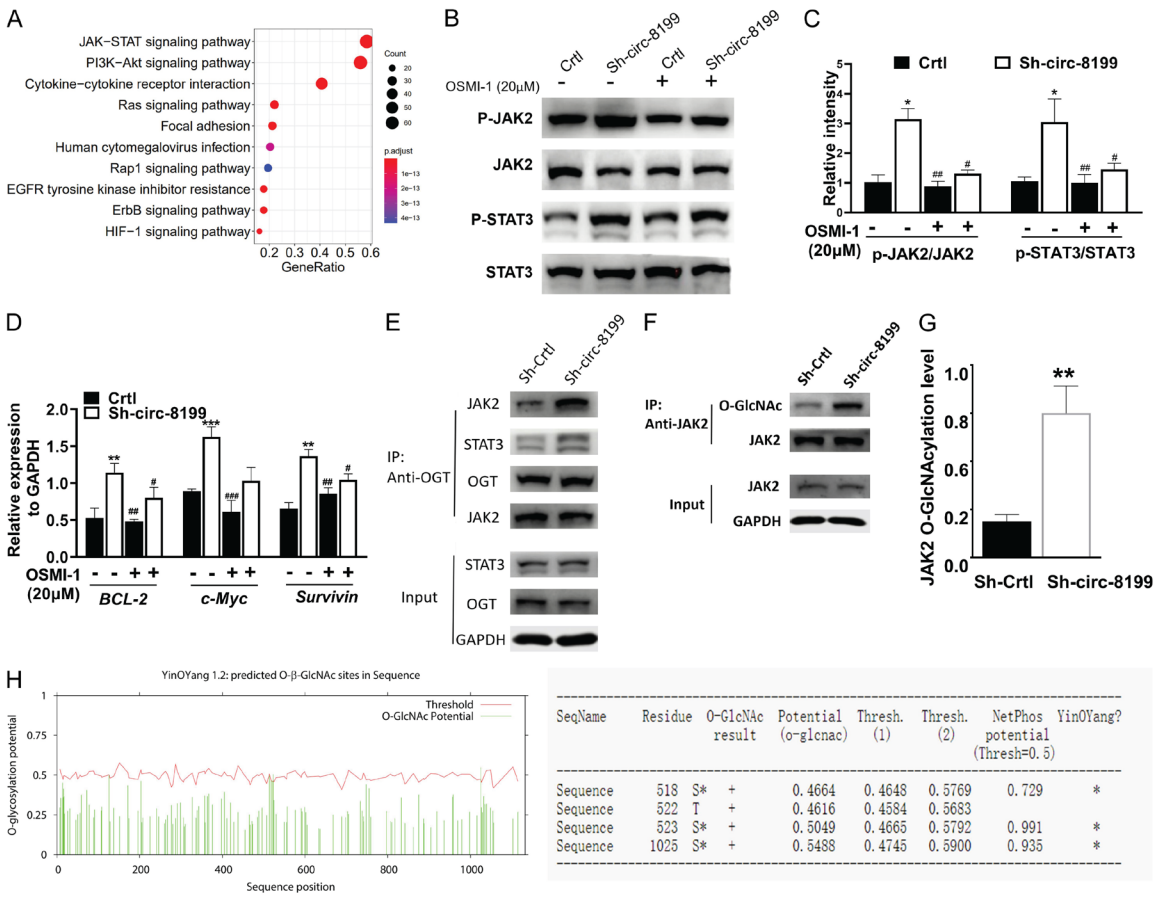
sequencing analysis of mRNA and circRNA of ESCC in a previous large cohort [17]. KEGG pathway enrichment analysis showed that lower expression of *circ8199* in ESCC was associated with JAK-STAT signaling pathway (**Figure 5A**).

To validate this finding, shRNA-*circ8199* was transfected into KYSE520 cells, and cells were lysed 48 hours later for WB and RT-qPCR tests. The phosphorylation levels of JAK2 and STAT3 were significantly increased by shRNA-*circ8199* compared with the control counterparts, which was reversed by the *OGT* inhibitor OSMI-1 (**Figure 5B-D**). Accordingly, the binding of *OGT* and JAK2 was enhanced when *circ8199* was knocked down (**Figure 5E**). And *OGT* had a lower binding activity to STAT3 than that of JAK2 (**Figure 5E**). Furthermore, the knockdown of *circ8199* could significantly enhance the level of *O-GlcNAc* glycosylation of JAK2 (**Figure 5F** and **5G**). To investigate the *O-GlcNAc* glycosylation modification site of JAK2, the amino acid sequence of JAK2 was analyzed using the online *O-GlcNAc* glycosylation prediction software http://www.cbs.dtu.dk/services/YinOYang/. The result showed that JAK2 may undergo *O-GlcNAc* glycosylation at Ser518, Thr522, Ser523 and Ser1025 (**Figure 5H**).

Discussion

Our previous study identified the circRNA *circ8199* by sequencing, analysis and screening of circRNA in paired ESCC tissues. The expression of *circ8199* was relatively downregulated in ESCC. Survival analysis showed that the high expression of *circ8199* was associated with prolonged survival of ESCC patients. In this study, we found that overexpression of *circ8199* could inhibit the proliferation and migration of tumor cells and encode a 235aa polypeptide mainly located in the cytoplasm, which could bind *OGT* and inhibit its activity through the JAK2-STAT3 pathway.

Circ8199 locates on chromosome 22 and comes from ATXN10, involving neuron survival, neuron differentiation, and neuritogenesis [20] via activation of the mitogen-activated protein kinase cascade [21]. Diseases associated with ATXN10 include Spinocerebellar Ataxia 10 and Spinocerebellar Ataxia 4 [20, 22], related to Akt signaling pathways. *Circ8199* is formed by reverse cleavage of exons 2 to 9 of ATXN10, which is 1057 bp in length. There was no significant change in the parental gene expression. In contrast, *circ8199* expression was interfered with by shRNA with the junction



SeqName	Residue	O-GlcNAc result	Potential (o-glcnac)	Thresh. (1)	Thresh. (2)	NetPhos potential (Thresh=0.5)	YinOYang?
Sequence	518	S*	+	0.4664	0.4648	0.5769	0.729	*
Sequence	522	T	+	0.4616	0.4584	0.5683		
Sequence	523	S*	+	0.5049	0.4665	0.5792	0.991	*
Sequence	1025	S*	+	0.5488	0.4745	0.5900	0.935	*

Figure 5. Inhibition of *circ8199*-235aa activated JAK2-STAT3 pathway by JAK2 *O-GlcNAc* glycosylation. A. KEGG pathways enriched among differentially expressed coding genes. B. Western blot were performed in *circ8199* shRNAs or scramble shRNA stably transfected KYSE520 cells with or without OSMI-1. C and D. The mRNA expressions of indicated genes were measured by qRT-PCR in *circ8199* shRNAs or scramble shRNA stably transfected KYSE520 cells with or without OSMI-1. E. Interaction between JAK2/*OGT* was determined mutually by immunoprecipitation in KYSE520. F and G. The *O-GlcNAc* glycosylation of JAK2 was detected following sh-*circ8199* stably transfection in KYSE520 cells. H. The putative sites of *O-GlcNAc* glycosylation on JAK2. Error bars indicate the standard deviations (SD) of at least three repeats. Student's t-test was used to determine *P*-values; $^{*}P < 0.05$, $^{**}P < 0.01$.

point, suggesting that the expression of *circ8199* in ESCC is independent of the parental gene. Therefore, *circ8199* can be used as a potential biomarker for ESCC molecular diagnosis.

Our study showed an ORF with a length of 708 nt in the *circ8199* transcript, which can encode a 235aa polypeptide (called *circ8199*-235aa in this study). A high-scoring IRES sequence analyzed by software was found upstream of the ORF. This IRES sequence could initiate the translation of circRNA-encoded proteins, as demonstrated by the dual-luciferase reporter gene experiment. The results of Western blotting confirmed that *circ8199* could encode a protein via *circ8199*-related expression vectors transfected into ESCC cells. Further experiments showed that *circ8199* lost its ability to encode a protein when its start codon ATG was mutated to ATC. In addition, overexpression of the predicted protein encoded by *circ8199* can inhibit the proliferation of ESCC cells in vitro.

Investigations have shown that many circRNAs encode proteins with essential biological functions related to cell proliferation, differentiation, metastasis and other processes in tumors and raise challenge to their traditional role only as gene transcription regulator through "sponge" adsorption of miRNAs. Meanwhile, there is not only high homology between the sequences of circRNA and linear RNA produced by the same transcripts of a gene but also high

homology between their encoded proteins. Some studies have reported that circRNA-encoded proteins can enhance the function of parental proteins. On the contrary, there are also circRNA-derived proteins that antagonize parental genes. To explore the regulating mechanism of *circ8199*-235aa, co-immunoprecipitation was performed using *circ8199*-235aa as the bait, and *OGT* was found to be the most abundant in the precipitates. *Circ8199* can recognize the same sequence on *OGT* as ATXN10, which may be necessary for binding. A previous study showed that ATXN10 directly regulates the activity of *OGT*. In this study, overexpression of *circ8199*-235aa could significantly inhibit the enzymatic activity of *OGT*, suggesting that *circ8199*-235aa may competitively impede the binding between ATXN10 and *OGT* through specific protein domain and inhibit *OGT* activity. Meanwhile, the down-regulation of *circ8199* in ESCC weakened this inhibitory effect and subsequently increased *OGT* activity, which promotes the progression of ESCC.

Recent studies have found that *O-GlcNAc* glycosylation on STAT3 inhibits its phosphorylation and the transcription of its target genes, which is inconsistent with the experimental results we observed. The potential reason for this inconsistency is that the down-regulation of *circ8199* could enhance the binding of *OGT* and JAK2 in ESCC cells, while *OGT* to STAT3 was weak. As expected, the level of *O-GlcNAc* glycosylation modification on JAK2 increased when *circ8199* was down-regulated. The above results suggest that the *O-GlcNAc* glycosylation on STAT3 may not play a dominant role in ESCC under the low expression of *circ8199*. On the contrary, the *O-GlcNAc* glycosylation of JAK2 may be the main factor for activating the JAK2-STAT3 pathway. Studies have shown that *OGT* inhibits the ubiquitination of its substrate proteins through *O-GlcNAc* glycosylation, thereby increasing the stability of its substrate proteins. In addition, *O-GlcNAc* glycosylation mainly occurs on the serine and threonine of the protein by competitively inhibiting the phosphorylation of these sites, and JAK2 has multiple phosphorylation sites that negatively regulate its kinase activity. Therefore, we concluded that *OGT* promotes *O-GlcNAc* glycosylation of JAK2 with low expression of *circ8199* in ESCC. Inhibiting its protein degradation or regulating its inhibitory phosphorylation site will finally activate the JAK2-STAT3 pathway. *Circ8199* regulates the JAK2-STAT3 pathway through *OGT*, providing a candidate mechanism for drug discovery and development. In addition, we found that JAK2 can undergo *O-GlcNAc* glycosylation at serine 518 (Ser518), threonine 522 (Thr522), serine 523 (Ser523) and serine 1025 (Ser1025) by using bioinformatics prediction, which needs to be validated in further studies.

Acknowledgements

This work was supported by a grant from the Natural Science Foundation of China (No. 82172941) and the Henan Province Medical Science and Technology Research Project (No. 212102310134).

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Guo-Zhong Jiang, Department of Pathology, The First Affiliated Hospital of Zhengzhou University, Jianshe Road, Zhengzhou 450052, Henan, PR China. E-mail: guozhongjiang@zzu.edu.cn

References

[1] Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, Jemal A, Yu XQ and He J. Cancer statistics in China, 2015. CA Cancer J Clin 2016; 66: 115-132.

[2] Zhang Y, Qian J, Gu C and Yang Y. Alternative splicing and cancer: a systematic review. Signal Transduct Target Ther 2021; 6: 78.

[3] Jeck WR and Sharpless NE. Detecting and characterizing circular RNAs. Nat Biotechnol 2014; 32: 453-461.

[4] Chen LL. The biogenesis and emerging roles of circular RNAs. Nat Rev Mol Cell Biol 2016; 17: 205-211.

[5] Yang Q, Li F, He AT and Yang BB. Circular RNAs: expression, localization, and therapeutic potentials. Mol Ther 2021; 29: 1683-1702.

[6] Zang J, Lu D and Xu A. The interaction of circRNAs and RNA binding proteins: an important part of circRNA maintenance and function. J Neurosci Res 2020; 98: 87-97.

[7] Gao X, Xia X, Li F, Zhang M, Zhou H, Wu X, Zhong J, Zhao Z, Zhao K, Liu D, Xiao F, Xu Q, Jiang T, Li B, Cheng SY and Zhang N. Circular RNA-encoded oncogenic E-cadherin variant promotes glioblastoma tumorigenicity through activation of EGFR-STAT3 signalling. Nat Cell Biol 2021; 23: 278-291.

[8] Zhang M, Huang N, Yang X, Luo J, Yan S, Xiao F, Chen W, Gao X, Zhao K, Zhou H, Li Z, Ming L, Xie B and Zhang N. A novel protein encoded by the circular form of the SHPRH gene suppresses glioma tumorigenesis. Oncogene 2018; 37: 1805-1814.

[9] Xia X, Li X, Li F, Wu X, Zhang M, Zhou H, Huang N, Yang X, Xiao F, Liu D, Yang L and Zhang N. A novel tumor suppressor protein encoded by circular AKT3 RNA inhibits glioblastoma tumorigenicity by competing with active phosphoinositide-dependent Kinase-1. Mol Cancer 2019; 18: 131.

[10] Yang Y, Gao X, Zhang M, Yan S, Sun C, Xiao F, Huang N, Yang X, Zhao K, Zhou H, Huang S, Xie B and Zhang N. Novel role of FBXW7 circular RNA in repressing glioma tumorigenesis. J Natl Cancer Inst 2018; 110: 304-315.

[11] Ye F, Gao G, Zou Y, Zheng S, Zhang L, Ou X, Xie X and Tang H. circFBXW7 inhibits malignant progression by sponging miR-197-3p and encoding a 185-aa protein in triple-negative breast cancer. Mol Ther Nucleic Acids 2019; 18: 88-98.

[12] Liang WC, Wong CW, Liang PP, Shi M, Cao Y, Rao ST, Tsui SK, Waye MM, Zhang Q, Fu WM and Zhang JF. Translation of the circular RNA circbeta-catenin promotes liver cancer cell growth through activation of the Wnt pathway. Genome Biol 2019; 20: 84.

[13] Marz P, Stetefeld J, Bendfeldt K, Nitsch C, Reinstein J, Shoeman RL, Dimitriades-Schmutz B, Schwager M, Leiser D, Ozcan S, Otten U and Ozbek S. Ataxin-10 interacts with O-linked beta-N-acetylglucosamine transferase in the brain. J Biol Chem 2006; 281: 20263-20270.

[14] Andrali SS, Marz P and Ozcan S. Ataxin-10 interacts with *O-GlcNAc* transferase *OGT* in pancreatic beta cells. Biochem Biophys Res Commun 2005; 337: 149-153.

[15] Pinho SS and Reis CA. Glycosylation in cancer: mechanisms and clinical implications. Nat Rev Cancer 2015; 15: 540-555.

[16] Yuan Y, Wang L, Ge D, Tan L, Cao B, Fan H and Xue L. Exosomal *O-GlcNAc* transferase from esophageal carcinoma stem cell promotes cancer immunosuppression through up-regulation of PD-1 in $CD8^+$ T cells. Cancer Lett 2021; 500: 98-106.

[17] Wang W, Zhu D, Zhao Z, Sun M, Wang F, Li W, Zhang J and Jiang G. RNA sequencing reveals the expression profiles of circRNA and identifies a four-circRNA signature acts as a prognostic marker in esophageal squamous cell carcinoma. Cancer Cell Int 2021; 21: 151.

[18] Marz P, Probst A, Lang S, Schwager M, Rose-John S, Otten U and Ozbek S. Ataxin-10, the spinocerebellar ataxia type 10 neurodegenerative disorder protein, is essential for survival of cerebellar neurons. J Biol Chem 2004; 279: 35542-35550.

[19] Liang D and Wilusz JE. Short intronic repeat sequences facilitate circular RNA production. Genes Dev 2014; 28: 2233-2247.

[20] Hernandez-Castillo CR, Diaz R, Vaca-Palomares I, Torres DL, Chirino A, Campos-Romo A, Ochoa A, Rasmussen A and Fernandez-Ruiz J. Extensive cerebellar and thalamic degeneration in spinocerebellar ataxia type 10. Parkinsonism Relat Disord 2019; 66: 182-188.

[21] Waragai M, Nagamitsu S, Xu W, Li YJ, Lin X and Ashizawa T. Ataxin 10 induces neuritogenesis via interaction with G-protein beta2 subunit. J Neurosci Res 2006; 83: 1170-1178.

[22] Veliz-Otani D, Inca-Martinez M, Bampi GB, Ortega O, Jardim LB, Saraiva-Pereira ML, Mazzetti P and Cornejo-Olivas M. ATXN10 microsatellite distribution in a Peruvian Amerindian population. Cerebellum 2019; 18: 841-848.