Genome-Wide Perturbations of Alu Expression and Alu-Associated Post-Transcriptional Regulations Find a Uniqueness in Oligodendroglioma

Taeyoung Hwang
University of Colorado Boulder
Sojin Kim
Seoul National University Hospital Department of Neurosurgery
Tamrin Chowdhury
Seoul National University Hospital Department of Neurosurgery
Hyeon Jong Yu
Seoul National University Hospital Department of Neurosurgery
Kyung-Min Kim
Seoul National University Hospital Department of Neurosurgery
Ho Kang
Seoul National University Hospital Department of Neurosurgery
Jae Kyung Won
Seoul National University College of Medicine
Sung-Hye Park
Seoul National University College of Medicine
Joo Heon Shin
Lieber Institute for Brain Development
Chul-Kee Park (nsckpark@snu.ac.kr)
Seoul National University College of Medicine https://orcid.org/0000-0002-2350-9876

Research

Keywords: Glioma, Alu element, RNA editing, circular RNA

Posted Date: January 25th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-152468/v1

Abstract

Background

Alu is a primate-specific repeat element in the human genome and has been increasingly appreciated as a regulatory element in many biological processes. But the role of *Alu* has not been studied comprehensively in brain tumor because an evolutionary perspective has been the subject of little research in brain tumor. We aim to investigate the relevance of *Alu* to the gliomagenesis.

Methods

Using a total of 41 pairs of neurotypicial brain tissue samples and samples of diverse gliomas, we performed strand-specific RNA-seq and analyzed two *Alu*-associated post-transcriptional regulations, A-to-I editing and circular RNAs, and Alu expression in a genome-wide way.

Results

We found that while both A-to-I editing and circular RNA are decreased overall in gliomas, grade 2 oligodendrogliomas do not show this same pattern of global changes. Instead, in comparison with other gliomas, oligodendrogliomas showed a higher proportion of perturbed Alu RNA. Adenosine deaminase acting on RNA 2 (ADAR2) was down-regulated in gliomas other than grade 2 oligodendrogliomas, contributing to the observed *Alu*-associated perturbation.

Conclusions

Our results demonstrate that *Alu* is associated with glioma development and grade 2 oligodendroglioma exhibits a unique pattern of *Alu*-associated post-transcriptional regulations, which provides an insight to gliomagenesis from the perspective of an evolutionary genetic element.

Background

A large proportion of the human genome consists of repetitive elements. Transposable elements (TEs), a major repeat element, account for at least 45% of the human genome while coding sequences comprises less than 3% (1). Among TEs, Alu is the most abundant repeat element, consisting of about 10% of the human genome (1). Alu is primate-specific and expands through retrotransposition, an amplifying process of TE through an RNA intermediate. Alu RNA is about 300 nucleotide (nt)-long, a noncoding RNA (ncRNA) that is mainly transcribed by RNA polymerase III (2). As it has two monomers facing each other, it tends to form double-stranded RNA (dsRNA).

Although Alu was considered as “junk” DNA in the past, there has emerged increasing evidence that this element plays important regulatory roles in diverse cellular processes (3). Specifically, Alu’s regulatory roles are manifest at both DNA and RNA levels. At the DNA level, Alu insertion can generate regulatory elements including alternative splicing (4) and enhancer function (5). Effects of *Alu* at the RNA level can

be understood by both *cis*- and *trans*- mechanisms. As a *cis*-regulatory element, *Alu* mediates diverse post-transcriptional regulations including A-to-I editing (6, 7), an RNA modification changing RNA sequence at a single nucleotide from adenosine to inosine, and formation of circular RNA (8), a single-stranded RNA with a covalently closed loop structure. It has been shown that the dsRNA structure of *Alu* facilitates A-to-I editing and circular RNA formation. *Alu* also harbors a nuclear localization signal of long noncoding RNA (lncRNA) (9). As a *trans* factor of gene regulation, Alu RNA itself, a short ncRNA transcribed by RNA polymerase III, affects transcription by modulating RNA polymerase II (Pol II) (10, 11). In addition, Alu RNA has been reported to influence translation (12).

Alu also has implications in human diseases including cancer and neuropathological disorders. In cancer biology, Alu RNA has been studied in hepatocellular carcinoma (13) and in a metastatic colorectal cancer cell line (14), both showing that increased levels of Alu RNA are associated with tumor development. The importance of *Alu* has been cited in the context of neuropathological diseases as multiple studies reported *Alu*'s extensive effects on transcriptome in the nervous system (15, 16). *Alu*-associated A-to-I editing and circular RNA are abundant in neuronal tissues and involved in neuronal differentiation and potentially in developmental disorders (17, 18). Based on these findings, *Alu* has been hypothesized also to potentially play a role in neurodegenerative diseases (15, 19). However, the study of *Alu* in the context of brain tumor is lacking due at least in part to an underappreciation of the possible role of evolutionary and developmental processes on brain tumor pathogenesis and the difficulty in studying *Alu*'s noncoding functions in biological systems.

We have previously shown that the *Alu*-associated A-to-I editing pattern of glioblastoma (GBM) is similar to that of early stage neurodevelopment (17). In the present study, we focused on *Alu* at the RNA level as a dynamic factor affecting the progression of glioma. Using a spectrum of glioma samples with up-to-date molecular classification (20), we investigated expression levels of Alu RNA and the two *Alu*-associated post-transcriptional regulations that are abundant in human brain tissue, i.e. A-to-I editing and expression of circular RNA. We first showed that A-to-I editing sites and circular RNAs that were found in our brain tissue samples are significantly associated with *Alu* element. Then, by comparing genome-wide patterns between matched tumor and neurotypical brain samples, we observed that A-to-I editing levels and circular RNA expression are perturbed in gliomas and are globally decreased in high grade gliomas. In contrast, grade 2 oligodendroglioma, a glioma of favorable prognosis, does not present a global bias in either A-to-I editing levels or circular RNA expression. A unique pattern in Alu RNA expression also was found in grade 2 oligodendroglioma, specifically downregulation of Alu RNA in these tumors relative to matched normal brain. Our results demonstrate *Alu* is associated with gliomas through its own expression and associated post-transcriptional regulations, providing a potential insight into the molecular mechanisms of gliomas from the perspective of a primate-specific repetitive element.

Methods

Patient tissues Surgical specimens and clinical information were obtained from glioma patients who underwent surgery at Seoul National University Hospital. Informed consent was obtained from all

patients for the usage of samples. A total of 42 patients with a matched pair of tumor and normal samples were enrolled. Normal brain tissues were obtained when surgical approach to the tumor involves brain areas without evidence of microscopical involvement of tumor. The final diagnosis was rendered using the most recent update of cIMPACT-NOW guidelines (20). After performing stranded RNA-seq, we removed one patient for further analysis as its tumor sample has lower sequencing quality based on the number of counted reads for gene expression measurement (lower than 10% of total sequencing reads).

Processing of RNA-seq. Sequencing reads were aligned to the human reference genome (GRCh38 primary_assembly) with gene annotation from GENCODE (version 27). After the number of reads assigned to a gene was counted, statistical tests for differential expression between tumor and matched normal tissues were done with a regression model that has patient information as a covariate. See supplementary methods for detail.

RNA editing call. We developed a computational pipeline to call RNA editing sites from strand-specific RNA-seq. Briefly, mismatches relative to the reference genome were identified in the uniquely-mapped sequencing reads. Then, for every genomic site, the sequencing reads were counted separately for the reference sequence or mismatches. Finally, if a site has more than or equal to 3 sequencing reads for an alternative allele and does not overlap with SNP, it was called as a potential RNA editing site whose type is determined by considering strand information. See supplementary methods for detail.

Identification of differentially-edited sites between tumor and matched normal tissue. The sites shared by all the patients for a given pathology were statistically tested for whether they showed different A-to-I editing levels between tumor and matched normal tissues. Specifically, beta binomial regression was performed for a given site while controlling patient-specific effect by adding patient information as a covariate in the regression. The A-to-I editing level for a given sample was defined as the ratio of inosine-supporting read counts relative to the adenosine or inosine-supporting read counts. Multiple test correction was done by using false discovery rate (FDR). The sites whose FDR-adjusted p-value is less than 0.05 were called as differentially-edited sites.

Identification and analysis of circular RNA. Circular RNAs were identified using STAR and CIRCexplorer2 according to the recommended pipeline of CIRCexplorer2 (21). We focused on exonic circular RNAs based on the gene annotation from GENCODE (version 27). In order to compare the ratio of circular RNA expression levels relative to linear RNA expression levels across the samples, we performed a beta-binomal regression for a gene, where the number of RNA-seq junction reads supporting circular RNA was modelled to follow binomial distribution with a total number of junction reads as a parameter. Statistical significances of differences between tumor and matched normal samples were calculated while controlling patient-specific effect with a covariate in the regression if a gene has more than 3 junction reads in any sample.

Measurement of expression of Alu element. We measured expression levels of Alu elements using TEtranscript (22) with the gene annotation from GENCODE (version 27). Briefly, TEtranscript is a software package that utilizes both uniquely and ambiguously mapped reads to quantify RNA expression levels of

transposable elements from RNA-seq. Statistical tests for differential expression between tumor and matched normal tissues were done by using DESeq2 (23) with a regression model that has patient information as a covariate.

Results

Catalogue and pattern of RNA editing in glioma detected by strand-specific RNA-seq

We performed strand-specific RNA-seq for tumor and matched normal samples obtained from 41 patients across various pathologies of gliomas spanning oligodendroglioma *IDH* mutant and 1p/19q-codeleted grades 2 and 3 (O2 and O3), *IDH* mutant astrocytoma grades 2, 3 and 4 (A2, A3 and A4), and GBM (Supplementary table 1). In addition, we developed a computational pipeline to identify RNA-editing sites in a genome-wide context from strand-specific RNA-seq. While many previous methods identifying RNA-editing sites from RNA-seq depend on pre-compiled gene annotation to assign RNA editing types, our pipeline considers strand information embedded in stranded RNA-seq to determine RNA editing types, allowing unbiased identification of RNA-editing sites (see methods in detail). After we applied the computational pipeline to individual samples in our dataset, we additionally filtered potential DNA variants that have inconsistency across the samples in terms of editing type or strand. Also, we only chose sites that were found in at least two patients to minimize the contamination of rare DNA variants. Application of our computational pipeline with the additional filters identified a total of 708,902 RNA variant sites across samples in our dataset. The RNA variants were predominated by A-to-G which is a representation of A-to-I editing in RNA-seq (Figure 1A), comprising about 81% of RNA variant sites (number: 571,774). The second dominant type was C-to-U whose proportion is about 4%. This type of editing is known to be found in human cells. If we consider potential technical errors in strand-specific RNA-seq, A-to-I and C-to-U editing can be identified as their reverse complemented forms of C-to-T and G-to-A, respectively. These four types accounted for about 92% of all the identified sites, indicating our pipeline's higher specificity of identification.

As expected, most A-to-I editing sites in our list show their strong association with *Alu*: about 82% of the sites were identified in annotated *Alu* element regions. Also, most A-to-I editing sites were found in intronic regions (93%). The numbers of A-to-I editing sites vary across patients and tend to be smaller in tumor compared normal tissues (Figure 1B, paired t-test p-value: 0.002077). Many A-to-I editing sites were not found commonly across the patients. The proportion of A-to-I editing sites that were shared by patients decreases according to the increasing number of patients observing the sites (Supplementary Figure 1). But we also identified that some numbers (3,415) of A-to-I editing sites were found in all the patients, which implies that these A-to-I editing sites are regulated in human brain tissues, despite the variability of individual samples.

In order to identify A-to-I editing sites that show different editing levels between tumor and normal tissues, we performed regression-based statistical tests comparing tumor and normal tissues while controlling patient-specific differences (see methods). As pathology is a clear factor contributing to variation in our

dataset (Supplementary figure 2), we conducted the statistical tests per a pathology. We excluded the *IDH* mutant astrocytoma grade 4 from this analysis as the number of patients is too small (n=2). We found that about 0.4~7.5% of A-to-I editing sites showed differential editing between tumor and matched normal tissues (Table 1). Interestingly, low grade gliomas, O2 and A2 showed relatively lower number of differential editing sites (0.4% for O2 and 0.5% for A2) compared with high grade gliomas (5.9%, 3.4%, and 7.5% for O3, A3, and GBM, respectively). When we checked the overlap of the differentially-edited sites between two pathologies, we found that the degree of overlap between O2 and each of the others was much smaller than comparisons between other pairs of pathologies (Figure 1C and Supplementary table 2). We also compared the overall distribution of A-to-I editing level differences between tumor and matched normal tissues. In figure 1D, the changes of A-to-I editing levels are summarized by a cumulative distribution whose shift to the left to the zero indicates the overall decrease of A-to-I editing levels in tumor relative to matched normal tissues. We found that most A-to-I editing sites in higher grades gliomas or non-oligodendroglioma were decreased in general in tumors relative to the matched normal tissues. In contrast, grade 2 oligodendroglioma showed no shift or no bias in difference of A-to-I editing levels. On average, A-to-I editing levels do not show significant differences between tumors and normal tissues (average difference: 0.05, t-test p-value: 0.04) in O2 while the others have significant decreases: O3 (average difference: -0.08, t-test p-value<10^{-15}), A2 (average difference: -0.21, t-test p-value<10^{-15}), A3 (average difference: -0.19, t-test p-value<10^{-15}), GBM (average difference: -0.07, t-test p-value<10^{-15}). All together, these results suggest that both the perturbed sites and the direction of editing level changes are different between grade 2 oligodendroglioma and the other gliomas.

Expression of circular RNA in glioma

Another *Alu*-associated post-transcriptional regulation that is abundant in brain tissue is a backsplicing process, where a 5' splice donor joins an upstream 3' splice acceptor, generating circular RNA. We first identified the genes that harbor circular RNA-supporting sequencing reads in our RNA-seq dataset (see methods). As expected, they are significantly associated with *Alu* (Figure 2A): the group of genes with circular RNA shows 99% association with *Alu* element while the group of genes without circular RNA expression only has 36% association (fisher exact test p-value<10^{-15}). Interestingly, the number of genes with circular RNA-supporting reads was significantly decreased in tumor compared to normal tissues in all gliomas except for O2 (Figure 2B: note that we used the relaxed p-value threshold (0.1) for statistical significance here and A4 is excluded for the detection of statistical significance due to lower number of samples). We further found genes showing different circular RNA expression levels between tumor and matched normal tissues while controlling patient-specific differences (see methods). O2 has the least unbiased distribution of differential expression while the others have variable degrees of shift in the distribution of decreased expression of circular RNA in tumors relative to normal tissues (Figure 2C): O2 (average difference: <0.01, t-test p-value= 0.005897), O3 (average difference: -0.02, t-test p-value<10^{-15}), A2 (average difference: -0.05, t-test p-value<10^{-15}), A3 (average difference: -0.06, t-test p-value<10^{-15}), GBM (average difference: -0.03, t-test p-value<10^{-15}).

Expression of Alu *RNA in glioma*

Alu itself can be expressed and affect cellular processes. We estimated the expression levels of known 47 *Alu* elements in our samples and compared their expression levels between tumor and the matched normal tissues using the computational pipeline that can handle repeat elements for differential expression analysis (see methods). In O2, a higher proportion of Alu RNAs (37 out of 47) was perturbed in tumors relative to matched normal tissues while the other gliomas only had one to three *Alu* RNAs as significantly-changed ones (Figure 3A). Alu RNAs were down-regulated in gliomas in general (Figure 3B) but, interestingly, O2 was identified as the most affected tumor in the overall expression change of Alu RNA, as shown by the largest shift of the distribution of expression changes toward negative in the figure 3C: average fold changes of tumor relative to normal tissues for O2, O3, A2, A3, A4 and GBM are -0.52, -0.28, -0.34, -0.23, -0.50 and -0.16, respectively in log2 scale (all of the pathologies showed statistically-significant decreases according to t-tests and the p-value threshold less than 0.01) .

Towards an integrative understanding of Alu-associated molecular processes in glioma

In order to understand the decreasing pattern of A-to-I editing and circular RNA expression in gliomas relative to normal tissues, we checked the expression levels of ADAR (Adenosine Deaminase Acting on Rna) families that are known to generate A-to-I editing and to affect the formation of circular RNA (Figure 4A). ADAR2 was down-regulated significantly in all gliomas except for O2, which is consistent with the identified decreasing patterns of A-to-I editing and circular RNA. In contrast, ADAR1 did not show any significant change across gliomas. ADAR3, known to antagonize ADAR1 and ADAR2 by competing with them also showed some decrease in high grade gliomas. Therefore, ADAR2 among ADAR families seems to be a *trans* factor underlying the decreased patterns of A-to-I editing and circular RNA expression that we observed. For Alu RNA expression, we checked the expression level of RNA polymerase III transcribing Alu RNA. We found that RNA polymerase III expressions were downregulated in all the pathologies in gliomas despite varying statistical significance (Supplementary figure 3). This downregulation of RNA polymerase III may in part contribute to lower expression of Alu RNA in gliomas compared to normal tissues.

We also sought to understand the relationship between A-to-I editing and circular RNA as they can compete or cooperate for Alu element-associated factors such as ADAR. We first found that A-to-I editing and circular RNA occurred in the same genes, but the overlap size is moderate (Figure 4B). Second, they were compared in terms of the direction of changes in tumor relative to normal tissue. The two processes showed consistency in that if the genes showed decreased circular RNA expression in tumor compared to normal, they also tended to have decreased A-to-I levels in tumor in general (Figure 4C). Finally, we looked into whether the two post-transcriptional processes affect similar pathways by performing gene ontology analyses for the genes harboring differentially-edited sites or the genes showing the different circular RNA expression. We found that the genes with perturbed A-to-I editing were over-represented in multiple gene ontology terms (Supplementary table 3). About 25% of the over-represented terms were shared by at least two pathologies and included the pathways known in neuronal tissues, for examples, glutamate ion

channels and the regulation of synapses (Figure 4D). For circular RNA, we found that different pathways are affected, including chromatin organization and regulation of metabolic process (Supplementary table 4). Therefore, our results demonstrated that A-to-I editing and circular RNA perturbations occurs concurrently at some genes, but they affect different genes in different pathways in general.

Discussion

There have been many attempts to understand neuropathological disorders from the perspective of genome evolution (19, 24). These efforts provide interesting hypotheses relating primate-specific genes or genetic elements to neuropsychiatric disorders as well as neurodegenerative disorders. However, brain cancer has not been appreciated well in terms of primate-specific elements as cancer is generally understood as diseases caused by genetic mutation associated with oncogenes or tumor suppressor genes and with environmental carcinogens. But increasing evidence suggests that evolutionary mechanisms affect brain tumors (25). In this study, we looked into a primate-specific *Alu* element to compare various pathologies in gliomas. We used the most recent version of classification criteria of gliomas and matched neurotypical samples for an elaborate comparison between tumor and normal brain tissues. We also performed extensive computational analyses with strand-specific RNA-seq in order to explore *Alu*'s dynamic effects on tumorigenesis in human brain tissue.

An obvious mode of *Alu* element's dynamic effect on brain pathologies is its expression as RNA. Although *Alu* is usually suppressed by epigenetic mechanisms, *Alu* RNA is transcribed by RNA polymerase III and dynamically regulated during development and in various diseases (2, 26). However, few studies have been conducted so far on Alu RNA in brain tumor, partly due to difficulty in measurements of repeated sequences in the human genome. We used a rigorous computational pipeline that considers repeat features in alignment of RNA-seq sequencing reads and controls patient-specific effects in a statistical comparison of Alu RNA expression between tumor and normal tissues. We found downregulation of Alu RNA in almost all pathologies of gliomas at varying degrees. This is in part explained by the downregulation of RNA polymerase III in glioma. But grade 2 oligodendroglioma is unique among the gliomas we studied in that downregulation of Alu RNA is accompanied by downregulation of other TEs (Figure 3B). In other pathologies, non *Alu* transposable elements do not show an overall bias of differential expression toward up- or down- regulation. This unique pattern of grade 2 oligodendroglioma is contrary to the previous studies showing increasing levels of RNA derived from TEs in progressive cancer (26) and may contribute to its better prognosis compared to other pathologies. One possible mechanism behind the downregulation of TE-derived RNA is histone modifications. A repressive histone mark of H3K9me3 was reported to be increased as a stress response, resulting in down-regulation of RNA of transposable elements in mouse brain (27). It will be interesting to see whether down-regulation of TE transcripts including Alu RNA observed in grade 2 oligodendroglioma is common in other cancer types of better prognosis.

Among *Alu*-associated regulatory processes, two post-transcriptional processes, A-to-I editing and backsplicing generating circular RNA are known to be abundant in brain tissue (17, 18). Many previous

studies have shown that these processes are affected in gliomas. For example, A-to-I editing is significantly altered, usually reduced in glioma (28). And there had been reports highlighting that alterations in specific RNA editing sites can contribute to tumor progression and classification of molecular signatures or grades in GBM (29). Moreover, a recent study comparing relative genome-wide RNA editing levels among genetic subgroups of glioma showed that the RNA editing signature can be used for prediction of *isocitrate dehydrogenase (IDH)* mutation and chromosome 1p/19q-codeletion status in gliomas (30). Regarding circular RNA, emerging evidence is accumulating on aberrant circRNAs expression and its oncological function in gliomas (31, 32). Studies have implicated circRNAs in the proliferation, invasion and angiogenesis of gliomas through cancer-associated signaling pathways (33). However, these studies mostly focused on individual sites or genes limiting their biological implications in terms of *Alu*-associated processes in brain cancer. Our results of global changes observed in gliomas suggest that there are perturbations affecting molecular mechanism of these processes, beyond individual sites or genes.

A-to-I editing and circular RNA are posited to mediate *Alu* element's contribution to the evolution of human brain. We and others have shown that A-to-I editing levels and circular RNA are abundant in neural genes that are enriched in the pathways of neurotransmission, neurogenesis, and synaptogenesis (8, 17, 18, 34, 35). *Alu* embedded in neural genes may serve as a mechanism to expand diversities in their functions and regulations that underlie remarkable complexities in the human nervous systems (36). The evolutionary benefits that *Alu* offers in human brain development, however, may turn into *Alu* element-specific adverse effects in pathological conditions in the nervous systems. Along with this idea, the Alu neurodegeneration hypothesis (19) proposes *Alu* as a "double-edged sword", whereby beneficial Alu-related processes also have the potential to disrupt mitochondrial homeostasis in neurodegenerative disorders including Alzheimer's disease. We propose that dysregulation of A-to-I editing and circular RNAs observed in higher grades of gliomas, affecting mainly glutamate signaling pathways and metabolic pathways, are also a manifestation of double-edged effect of *Alu* in glioma. From this evolutionary perspective, recent studies offer a clue as to how the beneficial effects of *Alu* might be turned into a tumor-promoting factor. For example, Venkataramani et al. reported that neuronal activities associated with glutamate synaptic connections contribute to progression of glioma (37). Here, *Alu* may enhance neuronal activity through A-to-I editing, an evolutionary advantage, but it also makes perturbation of A-to-I editing enzymes in glioma a favorable environment for glioma progression.

As *Alu* involves both A-to-I editing and the generation of circular RNA, it is intriguing to see whether these two *Alu*-associated post-transcriptional processes are influencing each other. The relation between circular RNA and A-to-I editing is unclear. Although some studies showed that circular RNA expression has a negative correlation with expression of ADAR1 (35, 38) in human cell lines, a recent study did not find a global correlation between the two processes (39) in mouse tissues. We observed decreased levels of circular RNA in genes showing decreased A-to-I editing levels, which implies that they are possibly co-regulated by other mechanisms, rather than ADAR1. In fact, our data did not show significant differences of ADAR1 expression in glioma compared with normal tissues. It is important to note, however, that the cellular pathways affected by A-to-I editing and circular RNA are different in our results. Although

previous studies showed that both A-to-I editing and circular RNAs in general occurred in genes encoding proteins of synaptic regulation, their perturbation may affect different pathways during gliomagenesis. Further studies are required to understand the observed loss of circular RNA in higher grades of gliomas and to uncover the relations between A-to-I editing and circular RNA.

Conclusions

Our study demonstrates that *Alu,* a primate-specific repeat element in the human genome, is associated with glioma development through post-transcriptional regulations such as A-to-I editing and circular RNAs, and its own RNA expression. In particular, grade 2 oligodendroglioma exhibits a unique global pattern compared to other gliomas, providing an insight to gliomagenesis from the perspective of an evolutionary genetic element.

Abbreviations

ADAR: Adenosine deaminase acting on RNA,

TE: Transposable elements,

ncRNA: noncoding RNA,

dsRNA: double-stranded RNA,

lncRNA: long noncoding RNA,

Pol II: RNA polymerase II,

GBM: glioblastoma.

Declarations

Ethics approval and consent to participate

This study was performed under the approval of the Institutional Review Board of Seoul National University Hospital (IRB approval No., H-1404-056-572).

Consent for publication

Not applicable

Availability of data and materials

The datasets are available through NCBI GEO (submission is under process but accession should be available before publication).

Competing interests

The authors declare that they have no competing interests.

Funding

This study was supported by Seoul National University Hospital Research Fund (2620160010) and the Basic Science Research Program (NRF-2019R1A2C2005144) through the National Research Foundation of Korea (NRF) funded by the Ministry of Science & ICT of Republic of Korea.

Authors' contributions

T.H. and C.K.P designed research; T.H. developed computational pipelines and performed statistical analyses; S.K., T.C., H.J.Y., K.M.K., and H.K. contributed to sample preparation and data generation; J.K.W. and S.H.P confirmed the genetic classification of patient samples; J.H.S. contributed analytic resources and discussed the results; T.H. and C.K.P. wrote the manuscript.

Acknowledgements

We thank Dr. Daniel R. Weinberger (Lieber Institute for Brain Development and Johns Hopkins University) for discussion and advising manuscript.

References

1. E. S. Lander, *et al.*, Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).
2. A. Orioli, C. Pascali, A. Pagano, M. Teichmann, G. Dieci, RNA polymerase III transcription control elements: Themes and variations. *Gene* **493**, 185–194 (2012).
3. R. A. Elbarbary, B. A. Lucas, L. E. Maquat, Retrotransposons as regulators of gene expression. *Science* **351** (2016).
4. G. Lev-Maor, R. Sorek, N. Shomron, G. Ast, The birth of an alternatively spliced exon: 3’ splice-site selection in Alu exons. *Science* **300**, 1288–1291 (2003).
5. M. Su, D. Han, J. Boyd-Kirkup, X. Yu, J.-D. J. Han, Evolution of Alu elements toward enhancers. *Cell Rep.* **7**, 376–385 (2014).
6. C. Daniel, G. Silberberg, M. Behm, M. Öhman, Alu elements shape the primate transcriptome by cis-regulation of RNA editing. *Genome Biol.* **15**, R28 (2014).
7. A. Athanasiadis, A. Rich, S. Maas, Widespread A-to-I RNA Editing of Alu-Containing mRNAs in the Human Transcriptome. *PLOS Biol.* **2**, e391 (2004).
8. W. R. Jeck, *et al.*, Circular RNAs are abundant, conserved, and associated with ALU repeats. *RNA N. Y. N* **19**, 141–157 (2013).

9. Y. Lubelsky, I. Ulitsky, Sequences enriched in Alu repeats drive nuclear localization of long RNAs in human cells. *Nature* **555**, 107–111 (2018).
10. P. Yakovchuk, J. A. Goodrich, J. F. Kugel, B2 RNA and Alu RNA repress transcription by disrupting contacts between RNA polymerase II and promoter DNA within assembled complexes. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 5569–5574 (2009).
11. P. D. Mariner, *et al.*, Human Alu RNA Is a Modular Transacting Repressor of mRNA Transcription during Heat Shock. *Mol. Cell* **29**, 499–509 (2008).
12. E. Ivanova, A. Berger, A. Scherrer, E. Alkalaeva, K. Strub, Alu RNA regulates the cellular pool of active ribosomes by targeted delivery of SRP9/14 to 40S subunits. *Nucleic Acids Res.* **43**, 2874–2887 (2015).
13. R.-B. Tang, *et al.*, Increased level of polymerase III transcribed Alu RNA in hepatocellular carcinoma tissue. *Mol. Carcinog.* **42**, 93–96 (2005).
14. F. Di Ruocco, *et al.*, Alu RNA accumulation induces epithelial-to-mesenchymal transition by modulating miR-566 and is associated with cancer progression. *Oncogene* **37**, 627–637 (2018).
15. P. A. Larsen, K. E. Hunnicutt, R. J. Larsen, A. D. Yoder, A. M. Saunders, Warning SINEs: Alu elements, evolution of the human brain, and the spectrum of neurological disease. *Chromosome Res.* **26**, 93–111 (2018).
16. W. Chen, E. Schuman, Circular RNAs in Brain and Other Tissues: A Functional Enigma. *Trends Neurosci.* **39**, 597–604 (2016).
17. T. Hwang, *et al.*, Dynamic regulation of RNA editing in human brain development and disease. *Nat. Neurosci.* **19**, 1093–1099 (2016).
18. X. You, *et al.*, Neural circular RNAs are derived from synaptic genes and regulated by development and plasticity. *Nat. Neurosci.* **18**, 603–610 (2015).
19. P. A. Larsen, *et al.*, The Alu neurodegeneration hypothesis: A primate-specific mechanism for neuronal transcription noise, mitochondrial dysfunction, and manifestation of neurodegenerative disease. *Alzheimers Dement.* **13**, 828–838 (2017).
20. D. N. Louis, *et al.*, cIMPACT-NOW update 6: new entity and diagnostic principle recommendations of the cIMPACT-Utrecht meeting on future CNS tumor classification and grading. *Brain Pathol. Zurich Switz.* **30**, 844–856 (2020).
21. X.-O. Zhang, *et al.*, Diverse alternative back-splicing and alternative splicing landscape of circular RNAs. *Genome Res.* **26**, 1277–1287 (2016).
22. Y. Jin, O. H. Tam, E. Paniagua, M. Hammell, TEtranscripts: a package for including transposable elements in differential expression analysis of RNA-seq datasets. *Bioinformatics* **31**, 3593–3599 (2015).
23. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

24. J. M. Sikela, V. B. Searles Quick, Genomic trade-offs: are autism and schizophrenia the steep price of the human brain? *Hum. Genet.* **137**, 1–13 (2018).
25. S. Hauptmann, W. D. Schmitt, Transposable elements – Is there a link between evolution and cancer? *Med. Hypotheses* **66**, 580–591 (2006).
26. Y. Luo, X. Lu, H. Xie, Dynamic Alu Methylation during Normal Development, Aging, and Tumorigenesis. *BioMed Res. Int.* **2014** (2014).
27. R. G. Hunter, *et al.*, Acute stress and hippocampal histone H3 lysine 9 trimethylation, a retrotransposon silencing response. *Proc. Natl. Acad. Sci.* **109**, 17657–17662 (2012).
28. N. Paz, *et al.*, Altered adenosine-to-inosine RNA editing in human cancer. *Genome Res.* **17**, 1586–1595 (2007).
29. X. Xu, Y. Wang, H. Liang, The Role of A-to-I RNA Editing in Cancer Development. *Curr. Opin. Genet. Dev.* **48**, 51–56 (2018).
30. S. C.-C. Chen, C.-M. Lo, S.-H. Wang, E. C.-Y. Su, RNA editing-based classification of diffuse gliomas: predicting isocitrate dehydrogenase mutation and chromosome 1p/19q codeletion. *BMC Bioinformatics* **20**, 659 (2019).
31. J. Zhu, *et al.*, Differential Expression of Circular RNAs in Glioblastoma Multiforme and Its Correlation with Prognosis. *Transl. Oncol.* **10**, 271–279 (2017).
32. Y. Yuan, *et al.*, Analyzing the interactions of mRNAs, miRNAs, lncRNAs and circRNAs to predict competing endogenous RNA networks in glioblastoma. *J. Neurooncol.* **137**, 493–502 (2018).
33. J. Sun, B. Li, C. Shu, Q. Ma, J. Wang, Functions and clinical significance of circular RNAs in glioma. *Mol. Cancer* **19**, 34 (2020).
34. J. O. Westholm, *et al.*, Genome-wide Analysis of Drosophila Circular RNAs Reveals Their Structural and Sequence Properties and Age-Dependent Neural Accumulation. *Cell Rep.* **9**, 1966–1980 (2014).
35. A. Rybak-Wolf, *et al.*, Circular RNAs in the Mammalian Brain Are Highly Abundant, Conserved, and Dynamically Expressed. *Mol. Cell* **58**, 870–885 (2015).
36. R. Dong, X.-K. Ma, L.-L. Chen, L. Yang, Increased complexity of circRNA expression during species evolution. *RNA Biol.* **14**, 1064–1074 (2017).
37. V. Venkataramani, *et al.*, Glutamatergic synaptic input to glioma cells drives brain tumour progression. *Nature* **573**, 532–538 (2019).
38. A. Ivanov, *et al.*, Analysis of intron sequences reveals hallmarks of circular RNA biogenesis in animals. *Cell Rep.* **10**, 170–177 (2015).
39. U. Kapoor, *et al.*, ADAR-deficiency perturbs the global splicing landscape in mouse tissues. *Genome Res.* **30**, 1107–1118 (2020).

Tables

Table 1. Summary of A-to-I editing sites according to pathologies in glioma Test sites mean the A-to-I editing sites found in all patients for a given pathology. Significant sites were determined by FDR-

adjusted p-value cutoff 0.05 in statistical comparison of A-to-I editing levels between tumor and matched normal tissues.

Pathology	Number of patients	Number of test sites	Number of significant sites	Proportion of significant sites (%)
O2	6	13852	53	0.4
O3	9	7647	453	5.9
A2	4	15855	73	0.5
A3	5	14148	476	3.4
GBM	15	6602	497	7.5

Figures

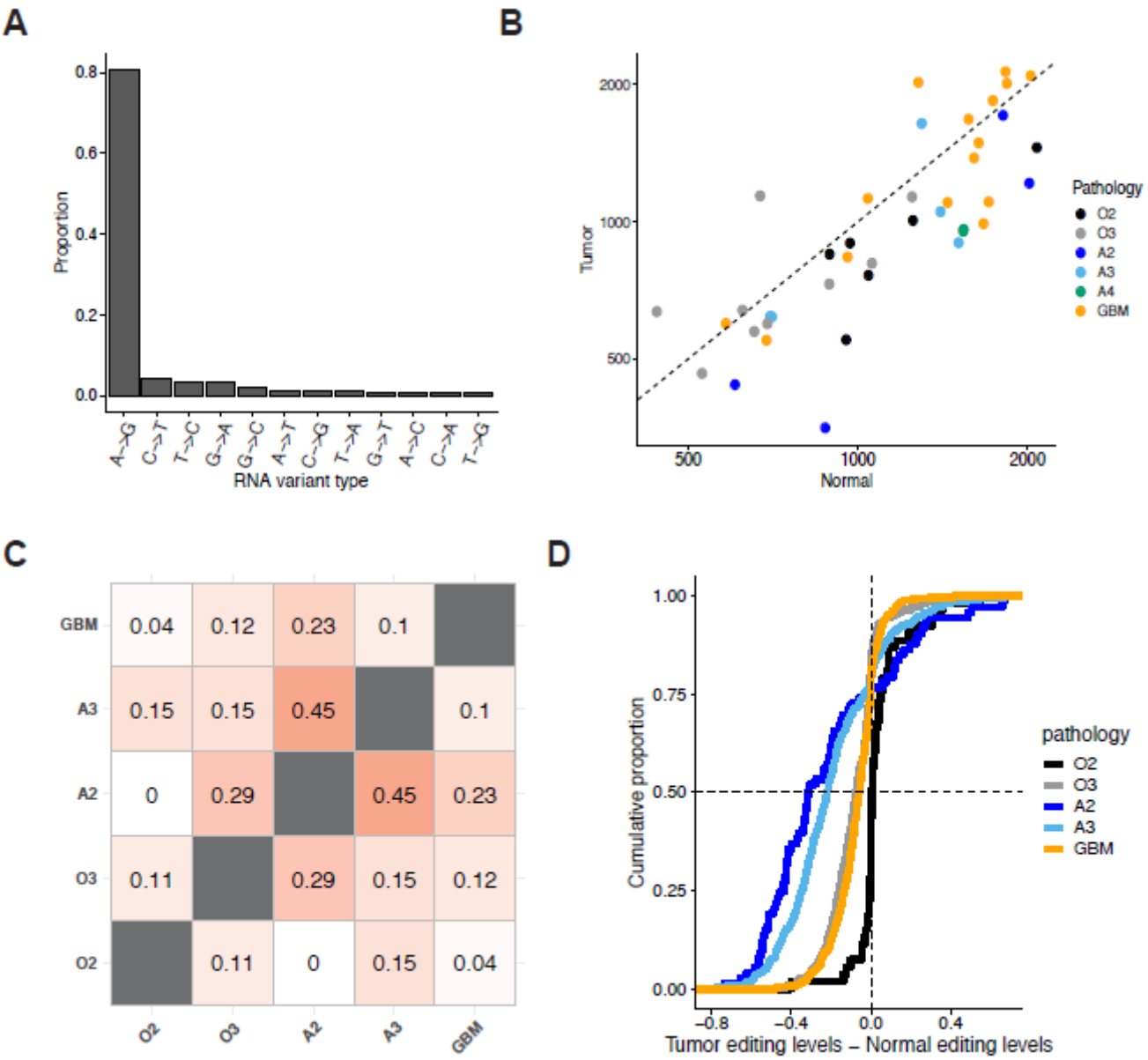

Figure 1

RNA editing sites and the A-to-I editing levels in glioma (A) Our unbiased computational pipeline detects RNA variants in 82 samples from patients of brain tumors. A-to-I or A-to-G type is most abundant among the identified RNA variants, indicating that most RNA variants are A-to-I editing sites. (B) Number of A-to-I editing sites per one million reads were depicted according to individual patients (dots) of different pathologies (colors) and tumor and normal tissues (y and x axis values respectively): grade 2 oligodendroglioma IDH mutant and 1p/19q-codeleted (O2), grade 3 oligodendroglioma IDH mutant and 1p/19q-codeleted (O3), grade 2 IDH mutant astrocytoma (A2), grade 3 IDH mutant astrocytoma (A3), grade 4 IDH mutant astrocytoma (A4), and glioblastoma (GBM). The dashed line indicates the line of y=x. (C) Overlaps of differentially-edited A-to-I editing sites between two pathologies. After differentially-edited A-to-I editing sites between tumor and the matched normal tissues were identified per pathology, overlaps of the sites were found and the proportions of the overlapped sites relative to the smaller numbers in the two comparing pathologies were described. Color gradation is proportional to the size of overlaps. (D) Distributions of A-to-I editing level changes in tumors relative to matched normal tissues were described in a cumulative way according to pathologies: the shift of curve to the left to 0 means overall decrease in tumor relative to normal tissues. Gliomas except for grade 2 oligodendroglioma (O2) has global decrease of A-to-I editing levels.

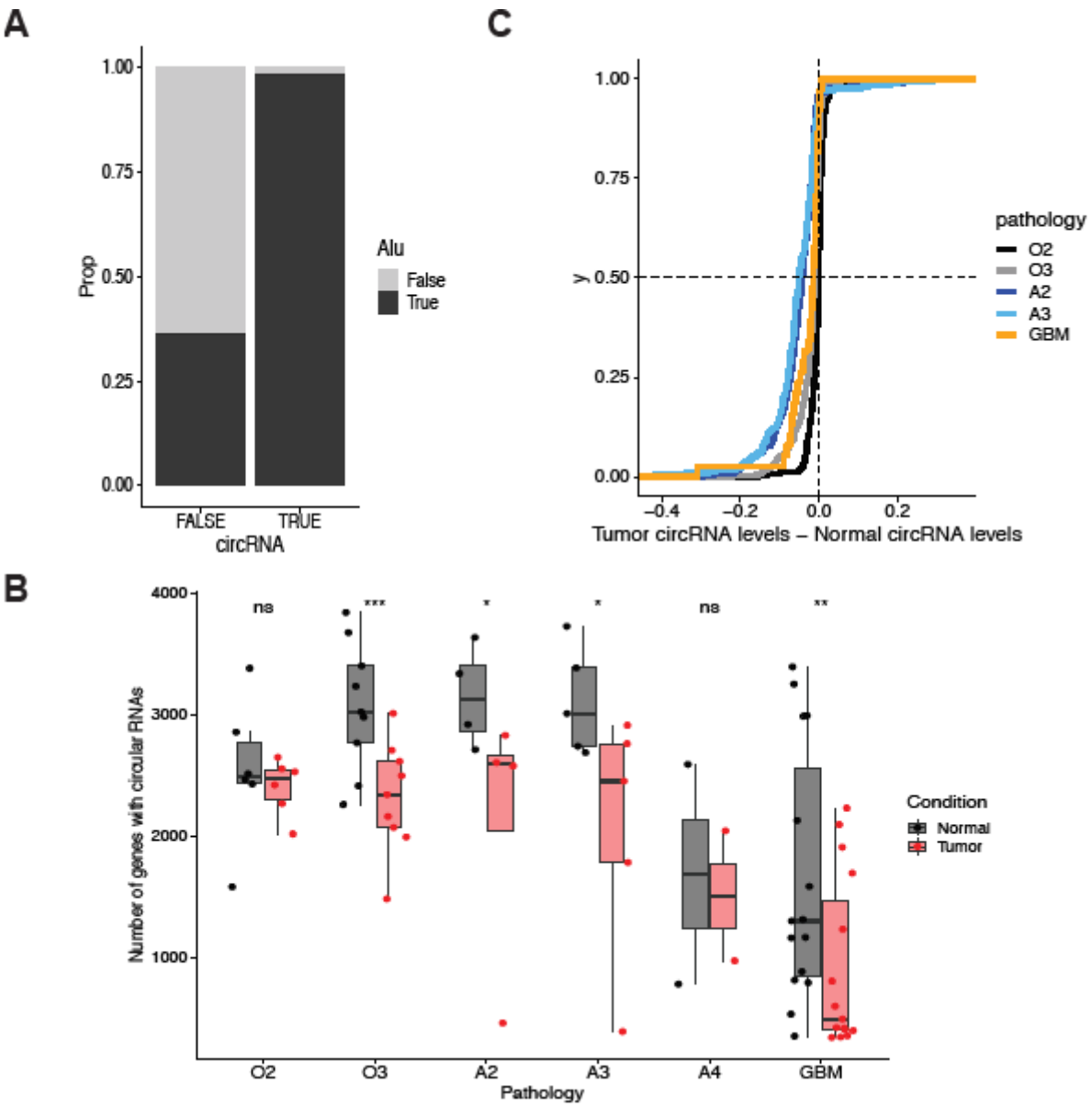

Figure 2

Circular RNA expression in glioma (A) Association of genes with Alu elements in the two different gene groups depending on the detection of circular RNA in gene bodies. (B) Number of genes containing circular RNA in gene bodies according to individual samples (dots) and pathologies: grade 2 oligodendroglioma IDH mutant and 1p/19q-codeleted (O2), grade 3 oligodendroglioma IDH mutant and 1p/19q-codeleted (O3), grade 2 IDH mutant astrocytoma (A2), grade 3 IDH mutant astrocytoma (A3), and glioblastoma (GBM). Tumors and normal tissues were indicated by different colors (red: tumor, gray: normal) and compared in terms of the average numbers of detected genes by Wilcoxon rank sum tests whose resulting statistical significances are indicated at the top of the plot: *** (p-value<0.01), ** (p-value<0.05), * (p-value<0.1) and ns (p-value>=0.1). (C) Distributions of circular RNA level changes in tumors relative to matched normal tissues were described in a cumulative way according to pathologies: the shift of curve to the left to 0 means overall decrease in tumor relative to normal tissues.

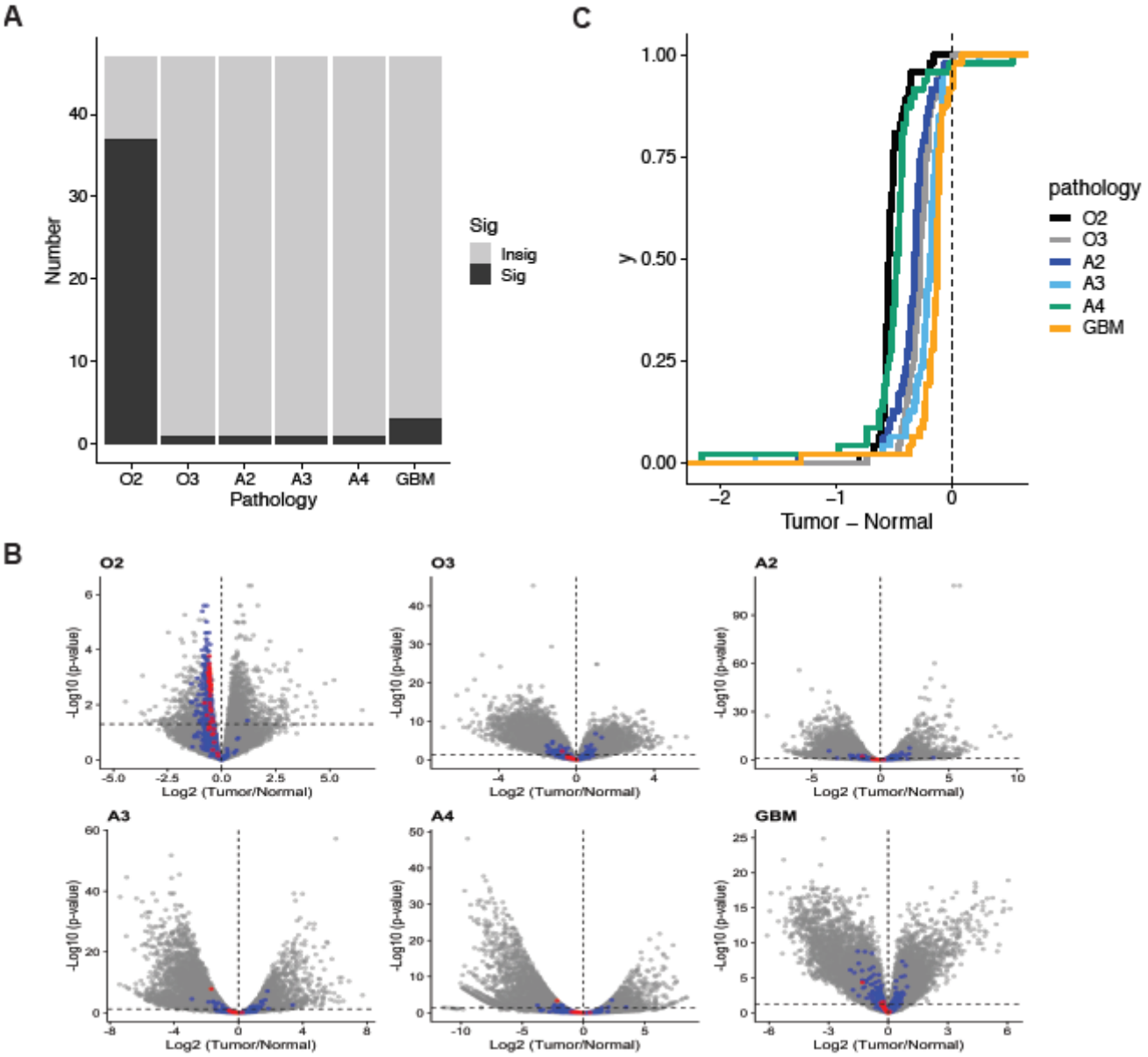

Figure 3

Alu RNA expression in glioma (A) The number of significantly-dysregulated Alu RNAs among total 47 Alu RNAs annotated in GENCODE (v27). (B) Volcano plots of differential gene expression analyses per brain tumor pathologies. Alu elements (red), non-Alu Transposable elements (blue) and the annotated genes (gray) were described according to the results of differential gene expression analysis. X-axis is fold change of RNA levels in log2 scale (tumor relative to normal) and y-axis is p-value of negative log scale. (C) Distribution of Alu RNA expression changes in tumors relative to matched normal tissues were described in a cumulative way according to pathologies. Label of pathologies: grade 2 oligodendroglioma (O2), grade 3 oligodendroglioma (O3), grade 2 astrocytoma (A2), grade 3 astrocytoma (A3), grade 4 astrocytoma (A4) and glioblastoma (GBM).

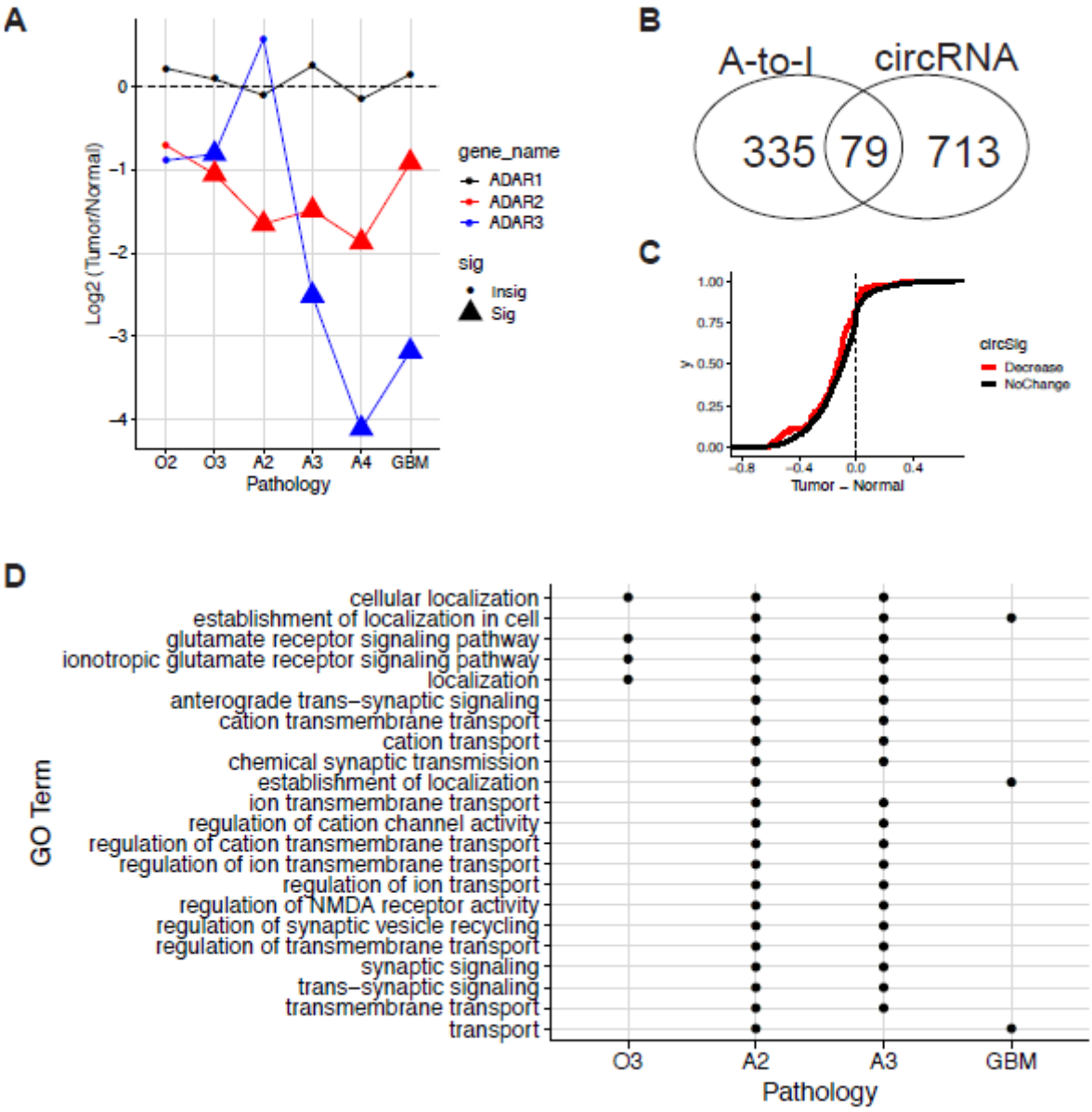

Figure 4

Integrative understanding of Alu-associated molecular processes in glioma (A) Average fold change of mRNA expression levels of A-to-I editing enzyme ADAR families (ADAR1, ADAR2 and ADAR3). Statistical significance of differential expression was determined by a R package DESeq2: significant if FDR-adjusted p-value<0.05. (B) Ven diagram of the genes with differentially-edited A-to-I editing sites and the genes showing differential expression of circular RNA (C) The distributions of A-to-I editing level changes in tumors relative to matched normal tissues were compared between the two groups made by whether genes show the decreased circular RNA expression levels in tumors relative to matched normal tissues (Kolmogorov-Smirnov test p-value: 0.06919). (D) Gene ontology terms that were enriched with the genes whose gene bodies harbor the differentially-edited A-to-I editing sites between tumor and matched normal tissues. The terms found in at least two pathologies of glioma were shown.

Supplementary Files

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

- Supplementarysubmitted.docx