Global RNA editome landscape discovers reduced RNA editing in glioma: Loss of editing of Gamma-Amino Butyric acid Receptor Alpha subunit 3 (GABRA3) favors glioma migration and invasion

Vikas Patil 1, 2, Jagriti Pal 1, Kulaandaivelu Mahalingam 2, Kumaravel Somasundaram Corresp. 1

1 Department of Microbiology and Cell Biology, Indian Institute of Science, Bangalore, India
2 Department of Bio-Medical Sciences, School of Biosciences and Technology, Vellore Institute of Technology, Vellore, India

Corresponding Author: Kumaravel Somasundaram
Email address: skumar1@iisc.ac.in

Background

Gliomas are the most common and lethal type of intracranial tumors. With the current treatment regime, the median survival of patients with grade IV glioma (glioblastoma/GBM) remains at 14-16 months. RNA editing modifies the function and regulation of transcripts. The development of glial tumors may be caused by altered RNA editing events.

Methods

In this study, we uncover the global RNA editome landscape of glioma patients from RNA-seq data of control, lower grade glioma (LGG) and GBM samples (n=1083).

Results

A-to-I editing events were found to comprise 80% of the total editing events of which 96% were located in the Alu regions. The total RNA editing events were found to be reduced in glioma compared to control samples. More specifically, we found Gamma-aminobutyric acid type A receptor alpha3 (GABRA3) to be edited (c.1026A-to-G; pI343M)) in 73% (editing ratio 0.8) of control samples compared to LGG (28.96%; 0.47) and GBM (5.2%; 0.53) samples. GABRA3 transcript level was found to be downregulated in glioma compared to control in a grade-specific manner with GBMs having the lowest level of the transcript. Further, GABRA3 transcripts were observed to be higher in edited compared to unedited glioma samples. The transcript and protein levels of exogenously expressed gene were found to be higher for edited compared to unedited GABRA3 in glioma cells. Further, exogenously expressed edited GABRA3 inhibited migration and invasion of glioma cells efficiently but not the unedited GABRA3.

Conclusion

Collectively, our study discovered a reduction in RNA editing during glioma development and demonstrated that editing of GABRA3 might have anti-migratory effect on the tumor cells. The reduced GABRA3 protein levels due to reduced stability of unedited RNA leads to loss of function in GBM thus resulting in an aggressive tumor.
Global RNA editome landscape discovers reduced RNA editing in glioma: Loss of editing of Gamma-Amino Butyric acid Receptor Alpha subunit 3 (GABRA3) favors glioma migration and invasion

Vikas Patil¹,², Jagriti Pal¹, Kulandaivelu Mahalingam² and Kumaravel Somasundaram*¹

¹ Department of Microbiology and Cell Biology, Indian Institute of Science, Bangalore, India.
² Department of Bio-Medical Sciences, School of Biosciences and Technology, Vellore Institute of Technology, Vellore, India

* Corresponding Author:
Kumaravel Somasundaram,
Department of Microbiology and Cell Biology,
Indian Institute of Science, Bangalore, Karnataka, 560012, India.
Email address: ksomasundaram1@gmail.com
Abstract

Background

Gliomas are the most common and lethal type of intracranial tumors. With the current treatment regime, the median survival of patients with grade IV glioma (glioblastoma/GBM) remains at 14-16 months. RNA editing modifies the function and regulation of transcripts. The development of glial tumors may be caused by altered RNA editing events.

Methods

In this study, we uncover the global RNA editome landscape of glioma patients from RNA-seq data of control, lower grade glioma (LGG) and GBM samples (n=1083).

Results

A-to-I editing events were found to comprise 80% of the total editing events of which 96% were located in the Alu regions. The total RNA editing events were found to be reduced in glioma compared to control samples. More specifically, we found Gamma-aminobutyric acid type A receptor alpha3 (GABRA3) to be edited (c.1026A-to-G; pI343M)) in 73% (editing ratio 0.8) of control samples compared to LGG (28.96%; 0.47) and GBM (5.2%; 0.53) samples. GABRA3 transcript level was found to be downregulated in glioma compared to control in a grade-specific manner with GBMs having the lowest level of the transcript. Further, GABRA3 transcripts were observed to be higher in edited compared to unedited glioma samples. The transcript and protein levels of exogenously expressed gene were found to be higher for edited compared to unedited GABRA3 in glioma cells. Further, exogenously expressed edited GABRA3 inhibited migration and invasion of glioma cells efficiently but not the unedited GABRA3.

Conclusion

Collectively, our study discovered a reduction in RNA editing during glioma development and demonstrated that editing of GABRA3 might have anti-migratory effect on the tumor cells. The reduced GABRA3 protein levels due to reduced stability of unedited RNA leads to loss of function in GBM thus resulting in an aggressive tumor.
Introduction

Thirty percent of all central nervous system (CNS) tumors and eighty percent of malignant brain tumors are composed of glial cells and are called gliomas (Goodenberger & Jenkins 2012). Astrocytomas are the most common and lethal type of glioma. They are divided into four categories as per WHO classification based on histopathology i.e., grade I, II, III and IV. Grade I, also known as, Pilocytic astrocytoma is benign, while grades II to IV are progressively more malignant. The most aggressive grade IV is called Glioblastoma (GBM). The median survival of GBM achieved till today through surgery, chemotherapy and radiotherapy is only 14.6 months (Stupp et al. 2009). Hence, further studies to understand the molecular pathways deregulated in GBM is important.

RNA editing is a molecular process by which RNA sequences are altered post-transcriptionally through base conversion or insertion/deletion. RNA editing increases proteomic diversity in cancer. In mammals, especially in humans, the most common type of editing changes include A-to-I and C-to-U base modifications (Gott & Emeson 2000). A-to-I occur in a large number of transcripts including miRNA and it is carried out by ADAR family of enzymes (Keegan et al. 2001; Nishikura 2010; Wulff et al. 2011). C-to-U editing events occur relatively rarely and are carried out by a particular deaminase family of enzymes called APOBEC (Rosenberg et al. 2011). ADAR enzyme binds to double-stranded RNAs and deaminates adenosine to inosine. Inosine in turn is recognized as guanosine by the cellular machinery. A-to-I editing events are most common in Alu repeats because of the double stranded RNA structures formed by inverted Alu repeats that spread across the genome (Levanon et al. 2004; O'Connell et al. 1995). RNA editing can lead to various changes in the mRNA including amino acid change, alteration in splice site, RNA stability, changes in secondary structures leading to alterations in proteins binding to it, alteration in miRNA binding etc.

In the present study, we propose to understand the role of RNA editing in glioma development. Until the last six years, various groups have identified RNA editing events by protein sequencing or cDNA sequencing of individual genes. Reduced A-to-I editing was observed in several human tumor types, including brain tumors (Maas et al. 2001; Paz et al. 2007). Also it was observed that restoration of the defective editing activity was able to inhibit proliferation of brain tumor cells (Cenci et al. 2008; Paz et al. 2007). Large scale mRNA sequencing was carried out to determine the regulation of RNA editing during brain development (Wahlstedt et al. 2009).
Subsequently, the advent of high-throughput sequencing technology led to the study of RNA editome landscape in different diseases including cancer. For example, increased A-to-I editing in hepatocellular carcinoma (HCC) has been identified in recent studies (Chan et al. 2014; Chen et al. 2013). Increased editing of AZIN1 transcript resulting in amino acid substitution was observed in HCC and the above change was seen to confer enhanced tumorigenicity (Chen et al. 2013). RNA sequencing and bioinformatic analysis have facilitated identification of novel RNA editing sites (Bahn et al. 2012; Choudhury et al. 2012; Ju et al. 2011; Li et al. 2011).

The importance of RNA editing in glioma pathogenesis was reported in an earlier study where authors studied the role of ADAR2 in GBM growth (Galeano et al. 2013). In a more recent study, Silvestris et al. used the inosinome profile for patient stratification in GBM (Silvestris et al. 2019). In this study, we have carried out comprehensive analysis of RNA sequencing data of normal brain, lower grade glioma (LGG) and glioblastoma (GBM) samples from publicly available datasets. The data was carefully analysed to understand the distribution of RNA editing events across different regions of the genome in normal and glioma samples. Comparison of editing events between normal, LGG and GBM samples was performed to evaluate the global regulation patterns of RNA editing in normal versus diseased conditions. Regulation of ADAR enzymes in gliomas and the differential editing events between the different types of gliomas was queried to understand how the editing enzymes are regulated and how the regulation correlates with editing during glioma progression. Pathway analysis of differentially edited genes was performed to find out cellular processes regulated by editing in glioma. This revealed important genes involved in brain functions. Under-editing of glutamate ionotropic receptor AMPA type subunit 2 (GRIA2) has been reported to play a role in glioma aggressive phenotype (Maas et al. 2001; Oakes et al. 2017). Here, we further studied the effect of missense RNA editing in gamma-amino butyric acid receptor alpha subunit 3 (GABRA3), a gene involved in neuronal signaling in the brain. Experimental investigation revealed that loss of editing in this gene in glioma leads to a more aggressive tumor phenotype.

**Materials & Methods**

**Data acquisition**

We downloaded (downloaded in March 2014) 529 lower-grade glioma and 172 GBM RNA and exome sequencing samples from TCGA through the Cancer Genomics Hub (CGHub, dbGaP
accession number phs000178). RNA sequencing data of 45 glioma cell lines were also downloaded from CCLE through CGHub (dbGaP accession number phs000178)). Samples downloaded from TCGA and CCLE study were in BAM format, we converted BAM format to fastq format and used for further analysis. We also downloaded 174 LGG and 100 GBM RNA sequencing samples from Chinese Genome Glioma Atlas (CGGA) via SRA (SRP027383) (Bao et al. 2014). RNA sequencing data of control brain samples (SRP033725, SRP045638, SRP044668) (N=63) were downloaded from SRA (Akula et al. 2014; Gill et al. 2014; Jaffe et al. 2015). Data downloaded from SRA was in sra format; we converted sra format to fastq format using SRA Toolkit. Sample information is provided in Supplementary table 1.

**RNA editing pipeline**

BWA aligner (Li & Durbin 2009) was used to align the RNA-seq reads on hg19 and Ensemble164. The duplicate removal was carried out using picard (http://broadinstitute.github.io/picard/) followed by re-alignment and base re-calibration. GATK unifiedgenotyper tool (McKenna et al. 2010) was used in order to call the editing events from RNA-seq reads which was compared to hg19. Next, the total variants obtained were filtered and also the potential SNPs were removed by dbSNP (Sherry et al. 2001), 1000g (Genomes Project et al. 2015) and ESP6500 databases (http://evs.gs.washington.edu/EVS). For each read first six bases were discarded so as to remove the mismatches which were caused due to random-hexamer priming. Following steps were used to filter out the spurious changes: a) Editing events were removed which showed changes within the 4 bps of known splice junction. b) Editing events present in homopolymer runs of $\geq 5$ were removed. C) Editing events present in different locations with high similarity found in BLAT (Kent 2002) were also removed.

**Editing ratio calculation**

Coverage is very much important factor to predict accurate RNA editing events. If variant position was having 10 or more reads then only that position was considered for calculating editing ratio. Editing ratio was calculated by dividing alternate allelic depth by total depth for that base. If a position was having editing ratio greater than 0.2 then that position was considered as putative RNA editing events.

**Differential RNA editing**
We compared RNA editing ratios between control brain samples and tumor samples to find out differentially editing events. We calculated average editing ratios for both the conditions. Editing difference was calculated by subtracting average value of control brain samples from average value of tumor samples. Significant testing was carried out by using Mann-Whitney U-test. If RNA editing event follow the criteria - a) Benjamini/Hochberg FDR correction value is less than 0.05, b) absolute difference between average RNA editing of tumor samples versus control brain samples is greater than 0.2 and c) absolute difference in percentage of samples edited, tumor versus control brain, is more than 10% then that RNA editing event was considered as significantly differential RNA editing event.

Pathway analysis

Significant differentially edited events (glioma versus control brain sample) were used for gene ontology (GO) analysis using DAVID bioinformatics resources 6.7 (Database for Annotation, Visualization and Integrated Discovery) (Huang da et al. 2009). Missense editing events were identified as editing events that cause a change is protein sequence of the gene.

Cell lines, constructs and antibodies

Glioma cell lines were obtained from ECACC, Salisbury, UK. All cells were cultured in Dulbecco’s Modified Eagles’ Medium (DMEM) containing 10% Fetal Bovine Serum (FBS). The cells were grown at 37°C in 5% CO2.

Antibodies used include anti-GABRA3 (Sigma Aldrich, HPA000839-100UL), DDK (Origene, #TA-50011), Actin (Sigma Aldrich, # A3854). GABRA3 overexpression construct: (NM_000808) Human cDNA ORF Clone with catalog number RC206286.

Site Directed Mutagenesis (SDM)

pCMV-Entry-GABRA3 plasmid was subjected to SDM using QuikChange Multi Site Directed Mutagenesis Kit (Catalog no. 200515). 100 ng of plasmid was taken and mixed with buffer, dNTPs, mutation primer and Pfu polymerase enzyme mix. During PCR reaction, extension was performed at 65°C for 14 mins (2 mins/kb of plasmid length). The whole mix was then incubated with DpnI enzyme at 37°C for 1 hour. 2 μl of the reaction mix was transformed in E.coli
DH5α and plated on Kanamycin LB agar plate. Mutant colonies were picked and plasmid was isolated. The mutation status was verified by Sanger sequencing.

**Transfection of plasmid DNA in glioma cell lines**

0.5 x 10^6 cells were plated in 35 mm dishes. 7 μg of each of VC or GABRA3 plasmids were transfected using Lipofectamine 2000™ in OptiMEM™ medium. After 6 hours of transfection, the medium was changed to DMEM containing 10% FBS. Cells were plated for experiment or harvested after 24 hours of transfection for RNA or protein isolation.

**Boyden-chamber assay for cell migration**

For every experimental condition, 50000 cells were plated in triplicates in the upper-chamber (Boyden chamber from BD Biosciences) in serum-free medium. 10% FBS was added in the lower chamber. The cells were allowed to migrate for 6-8 hours. Next, cells were fixed in 100% chilled methanol and stained using 0.2% crystal violet. The experiment was carried out in two biological replicates and the quantification is shown for one experiment. Unpaired t-test was used for calculating statistical significance between two groups. *, ** and *** denote p value of <0.5, <0.1 and <0.001 respectively.

**Boyden-chamber matrigel assay for cell invasion**

Activation of the matrigel chambers (BD Biosciences) were performed by incomplete DMEM for half an hour. For each experimental condition, 75000 cells were plated in the upper chamber in serum-free medium. 10% FBS was added in the lower chamber. The cells were allowed to invade for 20-22 hours after which cells were fixed and stained using 0.2% crystal violet stain. The experiment was carried out in two biological replicates and the quantification is shown for one experiment. Unpaired t-test was used for calculating statistical significance between two groups. *, ** and *** denote p value of <0.5, <0.1 and <0.001 respectively.

**RNA isolation, cDNA conversion and Real-time qPCR**

Cells were harvested using Trizol™ and RNA was isolated immediately following phenol-chloroform extraction method. The RNA was quantified using nanodrop method and quality assessment was performed by running 1 μg of RNA on agarose-formaldehyde gel. 2 μg good
quality RNA was used per reaction for cDNA conversion. Applied Biosystems™ High Capacity cDNA Reverse Transcription kit (Part no. 4368813) was used. The cDNA strand synthesis was carried out in Biorad S1000™ Thermal Cycler. Thermo Scientific’s DyNAmo reagent was used for real time qPCR. Applied Biosystems™ 7900HT Fast Real-Time PCR system was used to measure RNA levels. GAPDH was used as internal control. The experiment was carried out in two biological replicates and the quantification is shown for one experiment.

The primer sequences used in the current study:

- **CMV promoter FP:** GCTCGTTTAGTGAACCGTCAG
- **GABRA3 5’ UTR FP:** CAGTCACACCACAGCGTCT
- **GABRA3 Common RP:** GGTTCCAGGGAGAATATTAATCAGG
- **GAPDH FP:** GTCTCCTCTGACTTCAACAGCG
- **GAPDH RP:** ACCACCCTGTTGCTGTAGCCAA

**Protein isolation and western blotting**

Cells were trypsinized and pelleted down at 4°C. RIPA lysis buffer containing 1mM NaF, 1mM PMSF and 1X SIGMA FAST™ protease inhibitor cocktail was used to lyse the cells. After 1 hr, the cell debris were pelleted at 14,000 rpm and 4°C for 30 minutes. The protein concentration of each sample was determined by Bradford assay. 50 µg of protein was used for SDS-PAGE at 12% concentration of the resolving gel and run at 80V for 3 hours. The proteins were then transferred onto PVDF membrane, blocked using 5% skimmed milk and subsequently probed using primary and secondary antibodies. The experiment was carried out in two biological replicates and the quantification is shown for one experiment.

**Results**

**Systematic identification of RNA editing events in glioma and control brain tissues**

The advent of Next Generation Sequencing (NGS) has facilitated identification of novel RNA editing events. The availability of RNA sequencing as well as, matched whole exome sequencing (WES) data provide us with the opportunity to study RNA editing events that are
present in the RNA but do not occur in the DNA. Hence, we aimed to unravel RNA editing events in glioma pathogenesis through analysis of RNA-seq and WES data. For this purpose, data of glioma tumor tissue samples from The Cancer Genome Atlas (TCGA) was analysed. Further, RNA-seq samples from Chinese Glioma Genome Atlas (CGGA), glioma derived cell lines from Cancer Cell Line Encyclopedia (CCLE) and control brain samples from Sequence Read Archive (SRA) were analyzed to find out potential RNA editing events.

A brief schematic of the overall workflow has been given in Figure 1. Whole RNA sequencing data from TCGA, CCLE, CGGA and SRA were downloaded and aligned to the human reference genome hg19. Post alignment, variants were called in comparison to hg19. These variants were filtered through various datasets - dbSNP, 1000 genome project and ESP6500 to remove variations that are essentially polymorphisms. Exome sequencing data for glioma samples was available only in TCGA dataset. Hence, the variants were compared with corresponding whole exome sequencing data only in TCGA glioma samples to eliminate individual-specific genetic changes. RNA editing events identified from glioma samples were compared with control brain RNA editing events. We validated differential RNA editing events in TCGA glioma compared to control brain samples in the independent CGGA cohort (Figure 1 and Supplementary Figure 1).

The distribution of RNA editing events and editing ratio

We found higher frequency of RNA editing events in Alu repeat region (94.81%) compared to non-Alu repeat (2.24%) and non-repeat (2.95%) regions in control brain samples (Figure 2A). We observed that majority of RNA editing were found in intronic (70.82%), followed by intergenic (17.91%), non-coding RNAs (6.66%), 5’flank (2.93%), UTR (1.5%) and exonic (0.19%) regions in the control brain samples (Figure 2B). We found majority of editing events were ADAR specific changes (A-to-I i.e. A-to-G or T-to-C) followed by APOBEC specific changes (C-to-U i.e. C-to-T or G-to-A) in control brain samples (Figure 2C). We found A-to-I editing events in higher frequency in the non-coding regions of the genome, such as, UTR, intronic, intergenic, non-coding RNAs and 5’flank regions; whereas we found higher frequency of C-to-U editing only in exonic region of the genome in control brain samples (Figure 2D). We also identified elevated levels of editing ratio in Alu repeat compared to non-Alu repeat and non-repeat regions in control brain suggesting that apart from number of editing events, editing ratio is also more in Alu repeat region (Figure 2E). We observed similar RNA editing distribution for TCGA LGG (Supplementary
Figure 2), TCGA GBM (Supplementary Figure 3), CGGA GBM (Supplementary Figure 4), CGGA LGG (Supplementary Figure 5) and CCLE glioma cell lines (Supplementary Figure 6 and Supplementary Figure 7). We observed no significant differences in the global RNA editing distribution patterns between the normal and the glioma samples. From this analysis, we conclude that majority of the editing events were ADAR-specific A to I changes that occur in the non-coding regions of the genome.

The distribution of RNA editing ratio by glioma phenotype

We identified large number of RNA editing sites in all the data sets. We chose those editing events that were covered by 10 or more reads and defined them as high confidence RNA editing events. We found higher number of high confidence RNA editing events (n=734,436) in control brains samples as compared to glioma samples in different data sets such as TCGA LGG (n=221,049), TCGA GBM (n=220,954), CGGA LGG (187,757), CGGA GBM (134,509), and CCLE glioma cell lines (140,952) (Supplementary Figure 8A and Supplementary Tables S2-S7). We calculated editing ratio (number of edited reads divided by total number of reads at a given site) for all the samples. A-to-I RNA editing, the most prevalent editing event found in our data, is carried out by the ADAR family of enzymes (Keegan et al. 2001; Nishikura 2010; Wulff et al. 2011). We found elevated levels of ADAR in glioma samples as compared to control brain samples, although overall editing was lesser in glioma (Supplementary Figure 8A and 8B). Interestingly, we found reduced levels of ADARB1 and ADARB2 in glioma samples, indicating that these two ADAR family enzymes may play an important role in A-to-I RNA editing events in glioma (Supplementary Figure 8A, 8C and 8D). We identified RNA editing ratios were higher in control brain as compared to LGG and GBM in TCGA (Figure 3A) and CGGA (Figure 3B) datasets. Similarly, we found RNA editing ratios were higher in control brain as compared to glioma cell lines (Figure 3C). From this, we can say that overall RNA editing events occur at a reduced frequency in glioma compared to normal and ADAR family enzymes ADARB1 and ADARB2 may be responsible for this decreased editing. Hence, RNA editing play a role in maintaining normal functions of the brain and decrease in editing leads to diseased condition.
Differential RNA editing events in glioma

We compared RNA editing ratio of control brain samples with LGG samples of TCGA cohort to find out differential RNA editing events in LGGs. We found increased level of editing ratio in 176 editing events in TCGA LGGs as compared to control brain samples and 4,842 editing events had a reduced level of editing ratio in TCGA LGGs as compared to control brain samples (Figure 4A and Supplementary Tables S8). Similarly, we found increased editing ratio of 185 editing events in CGGA LGGs as compared to control brain samples and 3,287 editing events exhibited reduced editing ratio in CGGA LGGs as compared to control brain samples (Supplementary Figure 9A and Supplementary Tables S9). Next, we compared RNA editing ratio of control brain samples with TCGA GBM samples. We identified higher editing ratio of 421 editing events in TCGA GBM as compared to control brain and reduced editing ratio of 3,054 editing events in TCGA GBM as compared to control brain samples (Figure 4B and Supplementary Tables S10). Similarly, we also observed that 237 editing events were having higher editing ratio in CGGA GBM samples as compared to control brain samples and 3,531 editing events were having reduced editing ratio in CGGA GBM samples as compared to control brain samples (Supplementary Figure 9B and Supplementary Tables S11). Next, we compared RNA editing ratio of glioma cell lines with control brain samples. We found 1,171 editing events had higher editing ratio in glioma cell lines as compared to control brain samples and 2,842 editing events had reduced levels of editing in glioma cell lines as compared to control brain samples (Supplementary Figure 9C and Supplementary Tables S12). We observed that majority of RNA editing events were reduced in LGGs and GBM samples as compared to control brain samples. Also, majority of significant downregulated editing events between LGG samples versus control brain samples and GBM samples versus control brain samples were overlapping (Figure 4C and 4D).

Next we checked the concordance of RNA editing events in TCGA and CGGA data sets. We found 85.54 % of RNA editing events (both upregulated and downregulated) in CGGA LGGs were present in TCGA LGGs with downregulated editing events showing a higher overlap (Figure 5A and 5B). Similarly, we found that CGGA GBM RNA editing events had a concordance of 72.11 % with TCGA GBM RNA editing events (Figure 5C and 5D). Hence, majority of the differential editing events were found to be present in both TCGA and CGGA datasets.
Pathway analysis of genes exhibiting differential RNA editing in glioma

In depth analysis of the genes that underwent RNA editing identified several interesting facts. Two genes, MLLT6 and MPP2, that were most edited in control brain compared to GBM have been implicated in leukemogenesis and cell adhesion/cytoskeleton regulatory functions (Rademacher et al. 2016; Saha et al. 1995). We also carried out Gene Ontology (GO) analysis to find pathways that get regulated by editing to understand further the role of RNA editing in gliomagenesis. A total of 71, 50, 56, 44 and 28 GO biological processes were enriched for TCGA LGG, TCGA GBM, CGGA LGG, CGGA GBM and CCLE glioma cell lines respectively (Figure 6A). We observed 14 pathways to be common among all datasets which corresponded to neuronal function and protein translocation (Figure 6A and 6B). To identify proteins that are importantly regulated by editing in glioma, we identified genes from the above 14 pathways that get commonly altered in all the five datasets (Figure 6C). A total of 63 genes were found to be altered through editing that belong to the above pathways. Next, we evaluated the distribution of the above 63 edited genes across the genome which revealed only 0.78% of genes to be missense in nature (Figure 6D). A total of 5 out of the 63 genes that were enriched in the above pathways were found to be carrying missense type editing (Figure 6E) and these genes include GRIA2, GABRA3, GRIK2, NOVA1 and GRIK1.

Loss of editing in the coding region of GABRA3 cause aggressiveness of Glioma

It is evident that the majority of the aforementioned missense edited genes, that came up in our pathway enrichment analysis, play roles in neuronal signaling. These genes were found to be highly edited in normal brain and editing ratio and percentage of edited samples decreased significantly in glioma patients. We performed further investigation on GABRA3 because we found it to be highly edited (X: 151358319, A>G; pI342M) in normal brain (73%) with an average editing ratio of 0.8 while the editing percent and ratio reduced significantly in GBM samples (5.2% and 0.5 respectively). Indeed, GABRA3 editing reduced significantly in all datasets - TCGA LGG, CGGA LGG, TCGA GBM, and CGGA GBM compared to control brain (Figure 7A). Moreover, the transcript level of GABRA3 was found to be significantly lower in LGG and GBM samples of both datasets (Figure 7B). In fact, the transcript levels of unedited glioma samples (LGG + GBM) were found to be significantly lower compared to edited glioma samples (Figure 7C) and the RNA levels correlated significantly with the editing levels of the gene (R=0.51; p value <0.0001; Figure
Correlation of the overall editing events in all samples taken together with ADAR, ADARB1, and ADARB2 revealed that GABRA3 editing co-relate significantly with all three enzymes (Figure 7E). In case of normal brain, GABRA3 editing levels correlated significantly with ADAR and ADARB1 levels. For LGG samples (TCGA and CGGA), that exhibit moderate levels of editing in GABRA3, it was observed that the editing levels significantly correlate with all three enzymes. However, in GBM samples where GABRA3 editing is very low, the editing ratio did not correlate with any of the ADAR enzymes.

To understand the effect of loss of editing in glioma cells, we exogenously overexpressed edited and unedited forms of GABRA3 (Supplementary Figure 10A) in GABRA3 low cell lines LN229 and T98G which harbor no GABRA3 editing (Supplementary Figure 10B). Overexpression of edited GABRA3 in LN229 and T98G cells led to reduced migration and invasion potential of glioma cells compared to those expressing the unedited form (Figure 8A-F). Further, we predicted effect of amino acid change (I342M) as a result of RNA editing event on GABRA3 protein structure using PolyPhen-2 (Adzhubei et al. 2010), SIFT (Kumar et al. 2009), PROVEN (Choi et al. 2012) and MutationAssesor (Reva et al. 2011). Since all tools predicted neutral function of RNA editing events, we measured the transcript levels of overexpressed GABRA3 (normalized to endogenous GABRA3). This analysis revealed that loss of editing indeed leads to decreased transcript levels (Figure 8G; Supplementary Figure 10C). Consequently, the protein levels of GABRA3 were found to be reduced in unedited GABRA3 overexpressing cells compared to those expressing the edited form (Figure 8H, Supplementary Figure 11 and Supplementary Figure 12). In LN229 and T98G cells, we found 39% and 47% decrease in protein levels of unedited GABRA3 compared to edited GABRA3 overexpressing cells respectively. From this, we conclude that loss in editing of GABRA3 in glioma leads to reduced transcript and protein levels of GABRA3 which results in reduced inhibition of invasion and migration functions thus creating an aggressive GBM.

Discussion

Next generation sequencing techniques have helped to identify novel RNA editing events in many studies (Bahn et al. 2012; Chan et al. 2014; Chen et al. 2013; Choudhury et al. 2012; Ju et al. 2011; Li et al. 2011; Ramaswami et al. 2012; Ramaswami et al. 2013). In this study, we have
analysed RNA-seq data of 685 lower-grade glioma tumors, 272 glioblastoma tumors, 45 glioma cell lines, and 63 control brain samples from publicly available datasets.

A-to-I changes are the most common type of RNA editing events in mammals, especially in humans (Gott & Emeson 2000). A-to-I editing events, catalyzed by ADAR family of enzymes, are most prevalent in the Alu repeat regions because of the double stranded RNA structures formed by inverted Alu repeats that spread across the genome (Levanon et al. 2004; O’Connell et al. 1995). In this study, we found higher frequencies of RNA editing events in Alu repeat regions compared to non-Alu repeat and non-repeat regions in normal, LGG and GBM samples from all datasets. We also identified increased levels of editing ratio in Alu repeat compared to non-Alu repeat and non-repeat regions in all datasets suggesting that apart from the number of editing events, editing ratio is also more in Alu repeat region. We also observed that majority of editing events are ADAR specific changes (A-to-I i.e. A-to-G or T-to-C) followed by APOBEC specific changes (C-to-U i.e. C-to-T or G-to-A) similar to previously reported studies (Ramaswami et al. 2012; Ramaswami et al. 2013). A-to-I RNA editing events were found to be higher in the non-coding regions of the genome. This could be because the ADAR family of enzymes that bring about the A-to-I editing recognizes Alu repeat regions for binding and these repeat sequences are sparsely present in the coding regions of the genome (Grover et al. 2003). Conversely, we also observed that majority of the editing in the exonic region are of the C-to-U type catalyzed by the APOBEC family of enzymes (Rosenberg et al. 2011).

Our study revealed RNA editing levels to be higher in control brain samples and it is significantly reduced in glioma samples which correlate with the RNA expression levels of the ADAR enzymes ADARB1 and ADARB2. Indeed, it was shown in a previous report that ADAR enzymes are downregulated in glioma and overexpression of the gene leads to tumor suppressive effects (Maas et al. 2001). However, out of the three ADAR family enzymes, ADAR was found to be expressed in glioma samples which could be the primary source of A-to-I editing in glioma. RNA editing events which are downregulated in glioma compared to control brain samples showed a significantly higher concordance between different datasets. This could be because in general, ADAR enzymes are downregulated in glioma leading to reduced editing while the upregulated editing events may be spurious.
The editing events of GRIA2 and GABRA3 as unearthed by this study have been previously reported to be essential for the functioning of normal mammalian brain (Ohlson et al. 2007; Sommer et al. 1991). Moreover, GRIA2 underediting was found to be responsible for aggressive phenotype in GBM (Oakes et al. 2017). RNA editing in GABRA3 was found to be present in the coding region of the gene which led to Isoleucine to Methionine conversion in the 342th amino acid position of the protein. ADAR and ADARB1 are the major ADAR enzymes in the brain. Hence, we observed a significant correlation between the RNA levels of GABRA3 with these two enzymes in the control brain and this correlation was lost in GBM samples where global editing levels decrease drastically. We further studied the effect of this missense RNA editing in GABRA3.

From literature, we see that GABRA3 editing increases during brain development in rats with a simultaneous decrease in the expression (Daniel et al., 2011). However, in our study, we reveal that loss of editing in GABRA3 led to reduced gene expression levels in LGG and GBM samples from both TCGA and CGGA datasets. Moreover, experimental studies revealed that loss of editing leads to reduced levels of GABRA3 RNA and protein accompanied by a more migratory and invasive phenotype of the glioma cell lines. Although we test experimentally in a small sample set (n=2), the results are confirmed by similar regulation the larger TCGA and CGGA datasets.

This opposing regulation of GABRA3 by editing could be cell type-specific and development-specific. In fact, it is possible that the regulation in adult brain is different. In the healthy adult human brain, editing of GABRA3 is found in different regions of the brain irrespective of enrichment of neuronal or glial populations (Additional figure xx; Frederico A C Azevedo et al., J. Comp. Neurol., 2009). Moreover, the editing of GABRA3 does not alter even with age (Holmes et al., 2013), suggesting the importance of the editing in non-cancerous brain tissue. But during early developmental stages, the effect of GABRA3 editing is probably specific to the neuronal population due to the requirement of a switch of GABA response from excitatory to inhibitory post-synaptic potentials (Ben-Ari et al., Nat. Rev. Neur., 2002).

In a recent study, it is shown that loss of editing in GABRA3 leads to invasive phenotype in breast cancer (Gumireddy et al. 2016). In this study, we reveal the tumor suppressive nature of editing of GABRA3 in glioma which stresses on the requirement for the gene to be edited in normal scenario. Moreover, loss of editing of GABRA3 in GBM was accompanied by a decrease in RNA
levels as seen in the patient tissue samples. In case of the patient data, we compared with control brain tissue which comprises of both neuronal and glial cells. In this study, we could not specifically evaluate the contribution of GABRA3 editing from glial cells alone. This is a limitation of the present study. However, from experimental data in GBM cell lines, we confirmed that loss of editing indeed led to a decrease of both RNA and protein level of GABRA3. Thus, GABRA3 editing in glioma may lead to decrease in the protein which might be responsible for the aggressive tumor phenotype but this requires further validation in future.

**Conclusion**

Thus, the present study unfolds the entire RNA editome landscape of glioma tissues which will help scientists in understanding the importance of post-transcriptional sequence alterations in diseased conditions. Furthermore, the effect of loss of editing of GABRA3 during glioma progression highlights the importance of RNA editing for the maintenance of tissue homeostasis. The study will open up novel avenues of research and therapeutic interventions for glioma.

**Acknowledgements**

The results published here are in whole or part based upon data generated by The Cancer Genome Atlas pilot project established by the NCI and NHGRI. Information about TCGA and the investigators and institutions that constitute the TCGA research network can be found at http://cancergenome.nih.gov/. KS acknowledges CSIR and DBT, Government of India for research grant. Infrastructure support by funding from DST-FIST, DBT grant-in-aid and UGC (Centre for Advanced Studies in Molecular Microbiology) to MCB is acknowledged. KS is a J. C. Bose Fellow of the Department of Science and Technology.

**References**

Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, and Sunyaev SR. 2010. A method and server for predicting damaging missense mutations. *Nat Methods* 7:248-249. 10.1038/nmeth0410-248

Akula N, Barb J, Jiang X, Wendland JR, Choi KH, Sen SK, Hou L, Chen DT, Laje G, Johnson K, Lipska BK, Kleinman JE, Corrada-Bravo H, Detera-Wadleigh S, Munson PJ, and McMahon FJ. 2014. RNA-sequencing of the brain transcriptome implicates dysregulation of neuroplasticity, circadian rhythms and GTPase binding in bipolar disorder. *Mol Psychiatry* 19:1179-1185. 10.1038/mp.2013.170
Bahn JH, Lee JH, Li G, Greer C, Peng G, and Xiao X. 2012. Accurate identification of A-to-I RNA editing in human by transcriptome sequencing. *Genome Res* 22:142-150. 10.1101/gr.124107.111

Bao ZS, Chen HM, Yang MY, Zhang CB, Yu K, Ye WL, Hu BQ, Yan W, Zhang W, Akers J, Ramakrishnan V, Li J, Carter B, Liu YW, Hu HM, Wang Z, Li MY, Yao K, Qiu XG, Kang CS, You YP, Fan XL, Song WS, Li RQ, Su XD, Chen CC, and Jiang T. 2014. RNA-seq of 272 gliomas revealed a novel, recurrent PTPRZ1-MET fusion transcript in secondary glioblastomas. *Genome Res* 24:1765-1773. 10.1101/gr.165126.113

Cenci C, Barzotti R, Galeano F, Corbelli S, Rota R, Massimi L, Di Rocco C, O'Connell MA, and Gallo A. 2008. Down-regulation of RNA editing in pediatric astrocytomas: ADAR2 editing activity inhibits cell migration and proliferation. *J Biol Chem* 283:7251-7260. 10.1074/jbc.M708316200

Chan TH, Lin CH, Qi L, Fei J, Li Y, Yong KJ, Liu M, Song Y, Chow RK, Ng VH, Yuan YF, Tenen DG, Guan XY, and Chen L. 2014. A disrupted RNA editing balance mediated by ADARs (Adenosine DeAminases that act on RNA) in human hepatocellular carcinoma. *Gut* 63:832-843. 10.1136/gutjnl-2012-304037

Choi Y, Sims GE, Murphy S, Miller JR, and Chan AP. 2012. Predicting the functional effect of amino acid substitutions and indels. *PLoS One* 7:e46688. 10.1371/journal.pone.0046688

Choudhury Y, Tay FC, Lam DH, Sandanraj E, Tang C, Ang BT, and Wang S. 2012. Attenuated adenosine-to-inosine editing of microRNA-376a* promotes invasiveness of glioblastoma cells. *J Clin Invest* 122:4059-4076. 10.1172/JCI62925

Chen L, Li Y, Lin CH, Chan TH, Chow RK, Song Y, Liu M, Yuan YF, Fu L, Kong KL, Qi L, Li Y, Zhang N, Tong AH, Kwong DL, Man K, Lo CM, Lok S, Tenen DG, and Guan XY. 2013. Recoding RNA editing of AZIN1 predisposes to hepatocellular carcinoma. *Nat Med* 19:209-216. 10.1038/nm.3043

Chen L, Li Y, Lin CH, Chan TH, Chow RK, Song Y, Liu M, Yuan YF, Fu L, Kong KL, Qi L, Li Y, Zhang N, Tong AH, Kwong DL, Man K, Lo CM, Lok S, Tenen DG, and Guan XY. 2014. A disrupted RNA editing balance mediated by ADARs (Adenosine DeAminases that act on RNA) in human hepatocellular carcinoma. *Gut* 63:832-843. 10.1136/gutjnl-2012-304037

Choi Y, Sims GE, Murphy S, Miller JR, and Chan AP. 2012. Predicting the functional effect of amino acid substitutions and indels. *PLoS One* 7:e46688. 10.1371/journal.pone.0046688

Choudhury Y, Tay FC, Lam DH, Sandanraj E, Tang C, Ang BT, and Wang S. 2012. Attenuated adenosine-to-inosine editing of microRNA-376a* promotes invasiveness of glioblastoma cells. *J Clin Invest* 122:4059-4076. 10.1172/JCI62925

Galeano F, Rossetti C, Tomaselli S, Cifaldi L, Lezzerini M, Pezzullo M, Boldrini R, Massimi L, Di Rocco CM, Locatelli F, and Gallo A. 2013. ADAR2-editing activity inhibits glioblastoma growth through the modulation of the CDC14B/Spk2/p21/p27 axis. *Oncogene* 32:998-1009. 10.1038/onc.2012.125

Genomes Project C, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, Marchini JL, McCarthy S, McVean GA, and Abecasis GR. 2015. A global reference for human genetic variation. *Nature* 526:68-74. 10.1038/nature15393

Gill BJ, Pisapia DJ, Malone HR, Goldstein H, Lei L, Sonabend A, Yun J, Samanamud J, Sims JS, Banu M, Dovas A, Teich AF, Sheth SA, McKhann GM, Sisti MB, Bruce JN, Sims PA, and Canoll P. 2014. MRI-localized biopsies reveal subtype-specific differences in molecular and cellular composition at the margins of glioblastoma. *Proc Natl Acad Sci U S A* 111:12550-12555. 10.1073/pnas.1405839111

Goodenberger ML, and Jenkins RB. 2012. Genetics of adult glioma. *Cancer Genet* 205:613-621. 10.1016/j.cancergen.2012.10.009

Gott JM, and Emeson RB. 2000. Functions and mechanisms of RNA editing. *Annu Rev Genet* 34:499-531. 10.1146/annurev.genet.34.1.499

Grover D, Majumder PP, C BR, Brahmachari SK, and Mukerji M. 2003. Nonrandom distribution of alu elements in genes of various functional categories: insight from analysis of human chromosomes 21 and 22. *Mol Biol Evol* 20:1420-1424. 10.1093/molbev/msg153

Gumireddy K, Li A, Kossenkov AV, Sakurai M, Yan J, Li Y, Xu H, Wang J, Zhang PJ, Zhang L, Showe LC, Nishikura K, and Huang Q. 2016. The mRNA-edited form of GABRA3 suppresses
GABRA3-mediated Akt activation and breast cancer metastasis. *Nat Commun* 7:10715. 10.1038/ncomms10715

http://broadinstitute.github.io/picard/. Picard toolkit. Available at http://broadinstitute.github.io/picard/2014.

http://evs.gs.washington.edu/EVS. Exome Variant Server, NHLBI GO Exome Sequencing Project (ESP), Seattle, WA. Available at http://evs.gs.washington.edu/EVS2014.

Huang da W, Sherman BT, and Lempicki RA. 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4:44-57. 10.1038/nprot.2008.211

Jaffe AE, Shin J, Collado-Torres L, Leek JT, Tao R, Li C, Gao Y, Jia Y, Maher BJ, Hyde TM, Kleinman JE, and Weinberger DR. 2015. Developmental regulation of human cortex transcription and its clinical relevance at single base resolution. *Nat Neurosci* 18:154-161. 10.1038/nn.3898

Ju YS, Kim JI, Kim S, Hong D, Park H, Shin JY, Lee S, Lee WC, Kim S, Yu SB, Park SS, Seo SH, Yun JY, Kim HJ, Lee DS, Yavartanoo M, Kang HP, Gokcumen O, Govindaraju DR, Jung JH, Chong H, Yang KS, Kim H, Lee C, and Seo JS. 2011. Extensive genomic and transcriptional diversity identified through massively parallel DNA and RNA sequencing of eighteen Korean individuals. *Nat Genet* 43:745-752. 10.1038/ng.872

Keegan LP, Gallo A, and O'Connell MA. 2001. The many roles of an RNA editor. *Nat Rev Genet* 2:869-878. 10.1038/35098584

Kent WJ. 2002. BLAT--the BLAST-like alignment tool. *Genome Res* 12:656-664. 10.1101/gr.229202

Kumar P, Henikoff S, and Ng PC. 2009. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc* 4:1073-1081. 10.1038/nprot.2009.86

Levanon EY, Eisenberg E, Yelin R, Nemzer S, Hallegger M, Shemesh R, Fligelman ZY, Shoshan A, Pollock SR, Sztybel D, Olshansky M, Rechavi G, and Jantsch MF. 2004. Systematic identification of abundant A-to-I editing sites in the human transcriptome. *Nat Biotechnol* 22:1001-1005. 10.1038/nbt996

Li H, and Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25:1754-1760. 10.1093/bioinformatics/btp324

Li M, Wang IX, Li Y, Bruzel A, Richards AL, Toung JM, and Cheung VG. 2011. Widespread RNA and DNA sequence differences in the human transcriptome. *Science* 333:53-58. 10.1126/science.1207018

Maas S, Patt S, Schrey M, and Rich A. 2001. Underediting of glutamate receptor GluR-B mRNA in malignant gliomas. *Proc Natl Acad Sci U S A* 98:14687-14692. 10.1073/pnas.251531398

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, and DePristo MA. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 20:1297-1303. 10.1101/gr.107524.110

Nishikura K. 2010. Functions and regulation of RNA editing by ADAR deaminases. *Annu Rev Biochem* 79:321-349. 10.1146/annurev-biochem-060208-105251

O'Connell MA, Krause S, Higuchi M, Hsuan JJ, Totty NF, Jenny A, and Keller W. 1995. Cloning of cDNAs encoding mammalian double-stranded RNA-specific adenosine deaminase. *Mol Cell Biol* 15:1389-1397.

Oakes E, Anderson A, Cohen-Gadol A, and Hundley HA. 2017. Adenosine Deaminase That Acts on RNA 3 (ADAR3) Binding to Glutamate Receptor Subunit B Pre-mRNA Inhibits RNA Editing in Glioblastoma. *J Biol Chem* 292:4326-4335. 10.1074/jbc.M117.779868
Ohlson J, Pedersen JS, Haussler D, and Ohman M. 2007. Editing modifies the GABA(A) receptor subunit alpha3. *RNA* 13:698-703. 10.1261/rna.349107

Paz N, Levanon EY, Amariglio N, Heimberger AB, Ram Z, Constantini S, Barbash ZS, Adamsky K, Safran M, Hirschberg A, Krupsky M, Ben-Dov I, Cazacu S, Mikkelsen T, Brodie C, Eisenberg E, and Rechavi G. 2007. Altered adenosine-to-inosine RNA editing in human cancer. *Genome Res* 17:1586-1595. 10.1101/gr.6493107

Rademacher N, Schmerl B, Lardong JA, Wahl MC, and Shoichet SA. 2016. MPP2 is a postsynaptic MAGUK scaffold protein that links SynCAM1 cell adhesion molecules to core components of the postsynaptic density. *Sci Rep* 6:35283. 10.1038/srep35283

Ramaswami G, Lin W, Piskol R, Tan MH, Davis C, and Li JB. 2012. Accurate identification of human Alu and non-Alu RNA editing sites. *Nat Methods* 9:579-581. 10.1038/nmeth.1982

Ramaswami G, Zhang R, Piskol R, Keegan LP, Deng P, O'Connell MA, and Li JB. 2013. Identifying RNA editing sites using RNA sequencing data alone. *Nat Methods* 10:128-132. 10.1038/nmeth.2330

Reva B, Antipin Y, and Sander C. 2011. Predicting the functional impact of protein mutations: application to cancer genomics. *Nucleic Acids Res* 39:e118. 10.1093/nar/gkr407

Rosenberg BR, Hamilton CE, Mwangi MM, Dewell S, and Papavasiliou FN. 2011. Transcriptome-wide sequencing reveals numerous APOBEC1 mRNA-editing targets in transcript 3' UTRs. *Nat Struct Mol Biol* 18:230-236. 10.1038/nsmb.1975

Saha V, Chaplin T, Gregorini A, Ayton P, and Young BD. 1995. The leukemia-associated-protein (LAP) domain, a cysteine-rich motif, is present in a wide range of proteins, including MLL, AF10, and MLLT6 proteins. *Proc Natl Acad Sci U S A* 92:9737-9741. 10.1073/pnas.92.21.9737

Silvestris DA, Picardi E, Cesarini V, Fosso B, Mangraviti N, Massimi L, Martini M, Pesole G, Locatelli F, and Gallo A. 2019. Dynamic inosinome profiles reveal novel patient stratification and gender-specific differences in glioblastoma. *Genome Biol* 20:33. 10.1186/s13059-019-1647-x

Sommers B, Kohler M, Sprengel R, and Seeburg PH. 1991. RNA editing in brain controls a determinant of ion flow in glutamate-gated channels. *Cell* 67:11-19.

Stupp R, Hegi ME, Mason WP, van den Bent MJ, Taphoorn MJ, Janzer RC, Ludwin SK, Allgeier A, Fisher B, Belanger K, Hau P, Brandes AA, Gijtenbeek J, Marosi C, Vecht CJ, Mokhtari K, Wesseling P, Villa S, Eisenhauer E, Gorlia T, Weller M, Lacombe D, Cairncross JG, Mirimanoff RO, European Organisation for R, Treatment of Cancer Brain T, Radiation Oncology G, and National Cancer Institute of Canada Clinical Trials G. 2009. Effects of radiotherapy with concomitant and adjuvant temozolomide versus radiotherapy alone on survival in glioblastoma in a randomised phase III study: 5-year analysis of the EORTC-NCIC trial. *Lancet Oncol* 10:459-466. 10.1016/S1470-2045(09)70025-7

Wahlstedt H, Daniel C, Enstroen M, and Ohman M. 2009. Large-scale mRNA sequencing determines global regulation of RNA editing during brain development. *Genome Res* 19:978-986. 10.1101/gr.089409.108

Wulff BE, Sakurai M, and Nishikura K. 2011. Elucidating the inosinome: global approaches to adenosine-to-inosine RNA editing. *Nat Rev Genet* 12:81-85. 10.1038/nrg2915
Figure legends

Figure 1: Scheme of the computational analyses carried out in the study. The data used for this study from TCGA glioma samples (n=683), CGGA glioma samples (n=274), CCLE glioma cell lines (n=45) and SRA control brain samples (n=63).

Figure 2: Distribution of editing events across the genome for control brain samples. A. Distribution of total editing events in different regions of the genome: in Alu repeat, non-Alu repeat and non-repeat regions. B. Distribution of total editing events in different regions of the genome: in exonic, UTR, intergenic and intronic regions. C. Distribution of different types of editing events in different regions of the genome (Alu repeat, non-Alu repeat and non-repeat). D. Distribution of types of RNA editing events across different portions of the genome (exonic, intronic, intergenic and UTR). E. Cumulative distribution function of editing levels of Alu repeat, non-Alu repeat and non-repeat regions for control brain samples. Significance testing was performed using Mann–Whitney U-test. Black color p value is between alu repeat regions versus non-alu repeat regions; blue color p value is between alu repeat regions versus non-repeat regions.

Figure 3: The distribution of RNA editing ratio by phenotype. A. Cumulative distribution function of editing ratios comparison for control brain (n=63) samples, TCGA GBM (n=173) samples and TCGA LGG (n=511) samples. Red color p value is between control brain samples versus TCGA GBM samples; blue color p value is between control brain samples versus TCGA LGG samples. B. Cumulative distribution function of editing ratios comparison for control brain (n=63) samples, CGGA GBM (n=100) samples and CGGA LGG (n=174) samples. Red color p value is between control brain samples versus CGGA GBM samples; blue color p value is between control brain samples versus CGGA LGG samples. C. Cumulative distribution function of editing ratios comparison for control brain (n=63) samples and CGGA glioma cell lines (n=45). Significance testing was performed using Mann–Whitney U-test.

Figure 4: Differential RNA editing events in glioma. A. Heat map representing two-way hierarchical clustering of significant differential RNA editing events in TCGA LGG (n=511) versus control brain (n=63) samples. Samples are shown vertically, magenta colour represents control brain and orange colour represents TCGA LGG tumors. RNA editing are shown in rows, red colour represents upregulated editing events in TCGA LGG tumors as compared to control
brain samples and green colour represents downregulated editing events in TCGA LGG tumors as compared to control brain samples. B. Heat map representing two-way hierarchical clustering of significant differential RNA editing events in TCGA GBM (n=172) versus Control brain samples (n=63). Samples are shown vertically, magenta colour represents control brain and orange colour represents TCGA GBM tumors. RNA editing are shown in rows, red colour represents upregulated editing events in TCGA GBM tumors as compared to control brain samples and green colour represents downregulated editing events in TCGA GBM tumors as compared to control brain samples. C and D. Venn diagram representing the common number of differentially editing events in control brain samples versus TCGA LGG tumors and control brain samples versus TCGA GBM tumors.

Figure 5: Validation of differential RNA editing events in independent dataset. A and B. Venn diagram representing the common number of differentially editing events in TCGA LGG tumors and CGGA LGG tumors. C and D. Venn diagram representing the common number of differentially editing events in TCGA GBM tumors and CGGA GBM tumors.

Figure 6: Pathway analysis of edited events. A. Venn diagram represents Gene Ontology (GO) pathways of genes altered by RNA editing in TCGA GBM, CGGA GBM, TCGA LGG, CGGA LGG and CCLE samples. B. The 14 pathways common between the five datasets - TCGA GBM, CGGA GBM, TCGA LGG, CGGA LGG and CCLE. C. Venn diagram represents genes from the above 14 pathways that are common between the five datasets. D. Distribution of the editing events across the different regions of the genome. E. Venn diagram represents overlap between the missense edited genes and the genes enriched in our pathway analysis that are common among the five datasets.

Figure 7: Regulation of GABRA3 gene by RNA editing event. A. Scatter plot representing RNA editing levels of GABRA3 in control brain (n=63) samples, TCGA LGG (n=511), TCGA GBM (n=172), CGGA LGG (n=174), CGGA GBM (n=100) and CCLE glioma cell lines (n=45). B. Scatter plot representing mRNA expression of GABRA3 in control brain (n=63) samples, TCGA LGG (N=511), TCGA GBM (n=172), CGGA LGG (n=174), CGGA GBM (n=100) and CCLE glioma cell lines (n=45). C. Scatter plot showing mRNA expression of GABRA3 in GABRA3 edited samples (n=217) versus GABRA3 unedited samples (n=755). D. Scatter plot of correlation between RNA editing level of GABRA3 (n=972) and mRNA expression of GABRA3 (n=972). E.
Correlation between mRNA expression of ADAR family of enzymes and RNA editing levels of GABRA3. Significance testing was performed using Mann–Whitney U-test.

**Figure 8: Effect of GABRA3 editing on glioma cell migration and invasion.** A. Migration of LN229 and T98G cells overexpressing vector control (VC), edited and unedited GABRA3 gene. B and C. Quantification of migration of LN229 and T98G cells overexpressing vector control (VC), edited and unedited GABRA3 gene. D. Invasion potential of LN229 and T98G cells overexpressing vector control (VC), edited and unedited GABRA3 gene. E and F. Quantification of invasion potential of LN229 and T98G cells overexpressing vector control (VC), edited and unedited GABRA3 gene. ** represents p-value of <0.01. Significance testing was performed using Mann–Whitney U-test. G. RNA levels of GABRA3 gene in LN229 glioma cells having edited/unedited GABRA3 overexpressed ectopically. The RNA levels were normalised with respect to vector control (VC). H. Protein levels of GABRA3 gene in LN229 glioma cells having edited/unedited GABRA3 overexpressed ectopically in comparison to VC. The quantification has been provided underneath each blot with the levels of edited GABRA3 normalised to 1.0.
Figure 1

Scheme of the computational analyses carried out in the study.

The data used for this study from TCGA glioma samples (n=683), CGGA glioma samples (n=274), CCLE glioma cell lines (n=45) and SRA control brain samples (n=63).
Figure 2

Distribution of editing events across the genome for control brain samples.

**A.** Distribution of total editing events in different regions of the genome: in Alu repeat, non-Alu repeat and non-repeat regions. **B.** Distribution of total editing events in different regions of the genome: in exonic, UTR, intergenic and intronic regions. **C.** Distribution of different types of editing events in different regions of the genome (Alu repeat, non-Alu repeat and non-repeat). **D.** Distribution of types of RNA editing events across different portions of the genome (exonic, intronic, intergenic and UTR). **E.** Cumulative distribution function of editing levels of Alu repeat, non-Alu repeat and non-repeat regions for control brain samples. Significance testing was performed using Mann–Whitney U-test. Black color p value is between alu repeat regions versus non-alu repeat regions; red color p value is between alu repeat regions versus non-repeat regions.
Figure 3

The distribution of RNA editing ratio by phenotype.

**A.** Cumulative distribution function of editing ratios comparison for control brain (n=63) samples, TCGA GBM (n=173) samples and TCGA LGG (n=511) samples. Red color p value is between control brain samples versus TCGA GBM samples; blue color p value is between control brain samples versus TCGA LGG samples. **B.** Cumulative distribution function of editing ratios comparison for control brain (n=63) samples, CGGA GBM (n=100) samples and CGGA LGG (n=174) samples. Red color p value is between control brain samples versus CGGA GBM samples; blue color p value is between control brain samples versus CGGA LGG samples. **C.** Cumulative distribution function of editing ratios comparison for control brain (n=63) samples and CGGA glioma cell lines (n=45). Significance testing was performed using Mann–Whitney U-test.
Figure 1: Cumulative fraction plots for three datasets: TCGA, CGGA, and CCLE glioma cell lines. For each dataset, there are three groups: Control brain, TCGA GBM, and TCGA LGG. The plots show the editing ratio against the cumulative fraction. The significance levels are indicated as follows:

- A: TCGA
  - p < 0.0001
  - p < 0.0001

- B: CGGA
  - p < 0.0001

- C: CCLE glioma cell lines
  - p < 0.0001
Figure 4

Differential RNA editing events in glioma.

A. Heat map representing two-way hierarchical clustering of significant differential RNA editing events in TCGA LGG (n=511) versus control brain (n=63) samples. Samples are shown vertically, magenta colour represents control brain and orange colour represents TCGA LGG tumors. RNA editing are shown in rows, magenta colour represents upregulated editing events in TCGA LGG tumors as compared to control brain samples and orange colour represents downregulated editing events in TCGA LGG tumors as compared to control brain samples. B. Heat map representing two-way hierarchical clustering of significant differential RNA editing events in TCGA GBM (n=172) versus Control brain samples (n=63). Samples are shown vertically, magenta colour represents control brain and orange colour represents TCGA GBM tumors. RNA editing are shown in rows, magenta colour represents upregulated editing events in TCGA GBM tumors as compared to control brain samples and orange colour represents downregulated editing events in TCGA GBM tumors as compared to control brain samples. C and D. Venn diagram representing the common number of differentially editing events in control brain samples versus TCGA LGG tumors and control brain samples versus TCGA GBM tumors.
Figure 5

Validation of differential RNA editing events in independent dataset.

**A and B.** Venn diagram representing the common number of differentially editing events in TCGA LGG tumors and CGGA LGG tumors. **C and D.** Venn diagram representing the common number of differentially editing events in TCGA GBM tumors and CGGA GBM tumors.
Figure 6

Pathway analysis of edited events.

A. Venn diagram represents Gene Ontology (GO) pathways of genes altered by RNA editing in TCGA GBM, CGGA GBM, TCGA LGG, CGGA LGG and CCLE samples. B. The 14 pathways common between the five datasets - TCGA GBM, CGGA GBM, TCGA LGG, CGGA LGG and CCLE. C. Venn diagram represents genes from the above 14 pathways that are common between the five datasets. D. Distribution of the editing events across the different regions of the genome. E. Venn diagram represents overlap between the missense edited genes and the genes enriched in our pathway analysis that are common among the five datasets.
Figure 7

Regulation of GABRA3 gene by RNA editing event.

A. Scatter plot representing RNA editing levels of GABRA3 in control brain (n=63) samples, TCGA LGG (n=511), TCGA GBM (n=172), CGGA LGG (n=174), CGGA GBM (n=100) and CCLE glioma cell lines (n=45). B. Scatter plot representing mRNA expression of GABRA3 in control brain (n=63) samples, TCGA LGG (n=511), TCGA GBM (n=172), CGGA LGG (n=174), CGGA GBM (n=100) and CCLE glioma cell lines (n=45). C. Scatter plot showing mRNA expression of GABRA3 in GABRA3 edited samples (n=217) versus GABRA3 unedited samples (n=755). D. Scatter plot of correlation between RNA editing level of GABRA3 (n=972) and mRNA expression of GABRA3 (n=972). E. Correlation between mRNA expression of ADAR family of enzymes and RNA editing levels of GABRA3. Significance testing was performed using Mann–Whitney U-test.
**Table E**

| Gene          | ADAR  | ADARB1 | ADARB2 |
|---------------|-------|--------|--------|
| r² p value    | r²    | p value| r²    | p value|
| All           | 0.0223| < 0.0001| 0.2687 | < 0.0001| 0.1556 | < 0.0001|
| Control Brain | 0.2332| < 0.0001| 0.4003 | < 0.0001| 0.0002 | 0.9189 |
| TCGA LGG      | 0.0739| < 0.0001| 0.3398 | < 0.0001| 0.0555 | < 0.0001|
| TCGA GBM      | 0.0029| 0.4807 | 0.0366 | 0.0119 | 0.0713 | 0.0004 |
| CGGA LGG      | 0.0484| 0.0036 | 0.2339 | < 0.0001| 0.1171 | < 0.0001|
| CGGA GBM      | 0.0119| 0.2791 | 0.0352 | 0.0616 | 0.112  | 0.0007 |
Figure 8

Effect of GABRA3 editing on glioma cell migration and invasion.

A. Migration of LN229 and T98G cells overexpressing vector control (VC), edited and unedited GABRA3 gene. B and C. Quantification of migration of LN229 and T98G cells overexpressing vector control (VC), edited and unedited GABRA3 gene. D. Invasion potential of LN229 and T98G cells overexpressing vector control (VC), edited and unedited GABRA3 gene. E and F. Quantification of invasion potential of LN229 and T98G cells overexpressing vector control (VC), edited and unedited GABRA3 gene. ** represents p-value of <0.01. Significance testing was performed using Mann–Whitney U-test. G. RNA levels of GABRA3 gene in LN229 glioma cells having edited/unedited GABRA3 overexpressed ectopically. The RNA levels were normalised with respect to vector control (VC). H. Protein levels of GABRA3 gene in LN229 glioma cells having edited/unedited GABRA3 overexpressed ectopically in comparison to VC. The quantification has been provided underneath each blot with the levels of edited GABRA3 normalised to 1.0.
