Age-dependent Non-Silent Somatic Mutation with Transcriptomic Landscape and Prognosis of Lower Grade Glioma

YoungJoon Park1,2,*, JeongMan Park1,2,*, Ju Won Ahn1,2, JeongMin Sim2, Su Jung Kang1,2, Suwan Kim2, Sojung Hwang2,3, Kyoungh S. Sung4,*,‡ and Jaejoon Lim2,*,‡

Abstract: Glioma accounts for 80% of all malignant brain tumors and is the most common adult primary brain tumor. Age is an important factor affecting the development of cancer as somatic mutations accumulate with age. In this study, we aimed to analyze the significance of age-related non-silent somatic mutations in glioma prognosis. Histological tumor grade depends on age at diagnosis in patients with IDH1, TP53, ATRX, and EGFR mutations. The hierarchical clustering of patients was dominantly separated by IDH1 and EGFR mutations. Furthermore, patients with IDH1 mutation were dominantly separated by TP53 and ATRX double mutation and its double wildtype counterpart. Patients with the double mutation showed poorer prognosis than those with the double wild type genotype. In conclusion, among the many somatic mutations, those in IDH1, TP53, ATRX, and EGFR are important for glioma classification based on histological grade and related with onset age. Patients with EGFR mutation which showed late onset had the poorest prognosis, whereas those with only IDH1 mutation showed early onset had the best prognosis.

Keywords: age; glioma; mutation; TCGA; transcriptomic analysis

1. Introduction

Glioma, accounting for 80% of all malignant brain tumors, is the most common primary brain tumor in adults [1], and is usually treated using radiation therapy and using Temozolomide treatment after maximal safe resection [2]. Histologically, gliomas are classified as lower grade glioma (LGG) [World Health Organization (WHO) grade 2 (G2) and 3 (G3)] and glioblastoma multiforme (GBM) [WHO grade 4 (G4)]. LGG has a better prognosis than GBM, and the sensitivity of treatment depends on the molecular subtype of LGG in perspective of overall survival (OS) and progression free survival (PFI). In contrast, GBM has poor prognosis, and the development of new treatment methods based on its molecular characteristics is difficult because of infiltrative and integrative characteristics to normal brain tissues [3, 4].

Age is an important factor affecting the development of cancer. Most adult cancers develop exponentially with age [5], owing to the accumulation of somatic mutations [5-7]. However, all accumulated mutations do not contribute to cancer development. A mutation that contributes to the development of cancer is called a driver; a driver undergoes positive selection in the tissue microenvironment and leads to the development of cancer cell characteristics, such as cell growth [7, 8].
Current studies suggest signature mutation of glioma, but age-related mutation and related transcriptomic patterns are not reported yet. In this study, we investigated the effects of age to occurrence of non-silent somatic mutations on glioma prognosis and analyzed their transcriptomic significance using The Cancer Genome Atlas (TCGA) transcriptomic data. Finally, results of this study can be applied to molecular therapy with age and prediction of prognosis.

2. Results

2.1. Identification of Age-related Non-silent Mutations in LGG and GBM

Clinical data and non-silent somatic mutation data by multi-center mutation calling (MC3) from TCGA were used to identify somatic mutations associated with age. Using logistic regression analysis, we analyzed whether age at diagnosis affects somatic mutations in LGG (Table S1) and GBM (Table S2). Results showed that, in patients with LGG, IDH1, ATRX, and TP53 were less likely to mutate with aging, while EGFR acquired more mutations with aging (p-value < 0.05 after Bonferroni correction) (Figure 1a). Similar to LGG, ATRX was less likely to mutate with age in GBM (p-value < 0.05 after Bonferroni correction) (Figure 1b). Additional mutations information of four genes was included supplementary table S3.

![Figure 1. Frequencies of mutated and wild-type genes and logistic regression results showing the relationship of mutations with age in patients with lower grade glioma (a) and glioblastoma (b). Genes shown have mutation rates > 5%. OR, odds ratio; *, significant.](image-url)

**Table 1**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Wild-type(n)</th>
<th>Mutation(n)</th>
<th>OR</th>
</tr>
</thead>
<tbody>
<tr>
<td>IDH1</td>
<td>114</td>
<td>394</td>
<td>0.95*</td>
</tr>
<tr>
<td>TPS3</td>
<td>262</td>
<td>246</td>
<td>0.95*</td>
</tr>
<tr>
<td>ATRX</td>
<td>316</td>
<td>192</td>
<td>0.95*</td>
</tr>
<tr>
<td>CIC</td>
<td>398</td>
<td>110</td>
<td>1.01</td>
</tr>
<tr>
<td>TTN</td>
<td>448</td>
<td>60</td>
<td>1.03</td>
</tr>
<tr>
<td>FUBP1</td>
<td>463</td>
<td>45</td>
<td>1.03</td>
</tr>
<tr>
<td>PIK3CA</td>
<td>466</td>
<td>42</td>
<td>1.02</td>
</tr>
<tr>
<td>NOTCH1</td>
<td>472</td>
<td>36</td>
<td>1.02</td>
</tr>
<tr>
<td>MUC16</td>
<td>472</td>
<td>36</td>
<td>1</td>
</tr>
<tr>
<td>EGFR</td>
<td>473</td>
<td>35</td>
<td>1.1*</td>
</tr>
<tr>
<td>NF1</td>
<td>477</td>
<td>31</td>
<td>1.03</td>
</tr>
</tbody>
</table>

2.2. Histological Tumor Grade Depends on Age at Diagnosis in the IDH1 Wild-type group but not in the Mutant group

The patterns of age and IDH1 mutations were analyzed according to histological tumor grade. Results showed that, in patients with IDH1 wild-type, as the histologic tumor grade increased, the onset age of glioma was increased (Figure 2a). However, in IDH1
mutant patients, despite the rising histologic tumor grade of cancer, the onset age of glioma was constant (Figure 2a). In patients with G2 and G3 tumors, overall survival (OS) differed significantly between patients with and without IDH1 mutation (log-rank p-value in G2: 0.002; in G3: 1E-11; in G4: 0.02) (Figure 2b and table S4). In particular, the OS of patients with IDH1 wild-type and LGG did not differ significantly from that of patients with IDH1 mutation in GBM patients, but median survival was lower (median survival; IDH1 wild-type in LGG: 758 and IDH1 mutant in GBM: 1,024) (Figure 2b).

2.3. Transcriptomic Landscape of Tumors of Different Histological Grades Harboring IDH1 Mutations

To determine the transcriptomic landscape of tumors of different histological grades harboring IDH1 mutation, we performed a multi-label information gain (IG)-based feature selection of genes that possess discriminatory power according to combinations of histological tumor grades and IDH1 mutation statuses (Table S5). As a result, a hierarchical tree of patients dominantly separated by LGG and GBM was generated (Figure 3a). Interestingly, the gene expression pattern of IDH1 wild-type of G3 was similar to IDH1 wild-type of G4. In the IDH1 wild-type of G4 and G3-riched cluster, genes involved in the mitotic cell cycle and degradation of the extracellular matrix which are associated with malignancy of tumor in the Reactome 2020 database were more highly expressed than in LGG (Figure 3).
2.4. Distribution and Prognosis of Age-related Non-silent Somatic Mutations in LGG

Among several somatic mutations of four genes associated with pathological onset age in LGG, mutation pattern analysis revealed that ATRX and TP53 mutations occurred predominantly in patients with IDH1 mutations. The EGFR mutations dominantly occurred in patients with IDH1 wild-type (Figure 4a). The proportion of patients who divided by mutation status are suggested in supplementary (Table S6). The group with EGFR mutations had the worst prognosis, in terms of both OS and progression-free interval (PFI) (Figure 4b and Table S7). Among patients with IDH1 mutations, those with ATRX and TP53 mutations showed poorer PFI than those with the wild-type genes (Figure 4b).
Figure 4. Mutational patterns of IDH1, TP53, ATRX, and EGFR in the hierarchical tree (a). Survival analysis of patients with mutant or wild-type genes in terms of overall survival and progression-free survival (b). 0 indicates wild-type; 1 indicates mutation.

2.5. ATRX and TP53 Mutation Occurred in Patients with Early Onset Age of G2 and G3 Glioma and IDH1 Mutation, and the Gene Expression Patterns of Mutant ATRX and TP53 in LGG Differed from those of their Wild-type counterpart

ATRX and TP53 mutations were common in patients with IDH1 mutation and were associated with the age of onset of LGG (Figure 4a). Therefore, we analyzed whether
ATRX and TP53 mutations were directly related to the early age of onset of patients with IDH1 mutation. Results showed that both ATRX and TP53 mutations occurred in patients with early onset age of G2 and G3 tumors (Figure 5a). In the transcriptomic analysis, the hierarchical tree of patients primarily clustered by ATRX and TP53 double mutations and their double wild-type versions rather than by histological grade (Figure 5b). In total, 121 genes passed the IG-based feature selection. Among these 121 genes, the telomere extension-related Reactome term was significantly enriched in the gene enrichment test. This term contained only TERT (Figure 5c, Table S8).

**Figure 5.** Comparison of patients harboring ATRX and TP53 wild-type and mutation with age of glioma onset at initial pathological diagnosis in the background of IDH1 mutation. p-values less than 0.001 and 0.0001 are shown by *** and ****, respectively. As the number of GBM patients with IDH1 mutation is few, GBM is excepted in analysis (a). Heatmap with the hierarchical tree of patients by gene expression pattern with information gain (IG) over 0.35. The top annotations show the grade and mutational status (mutant or wild-type) of ATRX and TP53. WT indicates wild-type; MT indicates mutation (b). Statistically significant results of pathway enrichment analysis of gene clusters in the ATRX and TP53 hierarchical group (c).

### 2.6. TERT Expression was Suppressed in LGG and GBM with ATRX Non-silent Mutations
We observed that TERT expression which is only significant in cluster 2 was related to ATRX and TP53 double mutation in the IDH1 wild-type subgroup in LGG (Figure 5b). In addition, TERT expression was highly suppressed in ATRX and TP53 double mutant patients in the IDH1 mutation subgroup in LGG (Figure 6a). Figure 4a shows that the ATRX and TP53 mutation patterns were similar. Therefore, we compared TERT expression with the ATRX or TP53 mutation status and tumor grade. TERT expression was highly suppressed in patients with ATRX mutation for all histological tumor grades, but not in those with TP53 mutation of G4 (Figure 6b). It indicated that the TERT expression was tightly regulated by ATRX but not TP53 mutation.

![Figure 6](image)

**Figure 6.** Variation in TERT expression level with the mutation statuses of ATRX and TP53 in tumors with IDH1 mutation (a). TERT expression level in tumors with IDH1 mutation with separated mutation statuses of ATRX and TP53 for each grade. WT indicates wild-type; MT indicates mutation (b).

2.7. EGFR Mutation Occurred at Late Onset Age in Patients with G3 tumor and was Associated with Poor Prognosis of GBM Patients but not was not Clustered in Transcriptomic Analysis

Unlike ATRX, TP53, and IDH1 mutations, EGFR mutations were more frequent in older patients. Age at initial pathologic diagnosis of patients with EGFR wild-type showed positive correlation with tumor grade. Interestingly, in G2 and G3 patients who have relatively an early onset age, the patients with EGFR mutation were occurred tumor at old age as well as G4. In addition, despite having LGG, patients with EGFR mutations in G3
showed poor prognosis similar to those of GBM patients (Figure 7b). In the transcriptomic analysis, the hierarchical tree of patients was dominantly divided into G2 and the rest, but not by EGFR mutation status (Figure 7c, Table S9). Most of the significant pathways in the pathway enrichment analysis were related to the regulation of the cell cycle and were downregulated in GBM (Figure 7d).

Figure 7. Comparison of ages at initial pathological diagnosis of tumors with EGFR wild-type and those with EGFR. p-values less than 0.01 and 0.0001 are shown as ** and **** respectively. ns, not significant (a). Survival analysis comparing EGFR mutant or wild-type and each grade with overall survival. 0 indicates wild-type; 1 indicates mutation (b). Heatmap with hierarchical tree of patients based on gene expression pattern with IG over 0.35. Top annotations show tumor grade and genotype (mutant or wild-type) of EGFR. WT indicates wild-type; MT indicates mutation (c). Statistically significant results of pathway enrichment analysis of gene clusters in the EGFR hierarchical group (d). Mutational pattern of GBM (e). Comparison of EGFR expression levels between tumors with EGFR mutant and those with EGFR wild-type against an IDH1 wild-type background (f).

3. Discussion

Cancer is highly related with the occurrence of somatic mutations, which accumulate with age [9, 10]. In this study, we deeply investigated characteristics of molecular features and expression patterns of age-related somatic mutations in glioma.

IDH1 mutation is an important classifying point in both LGG and GBM [11]. In IDH1 mutant, conversion of α-ketoglutarate to 2-hydroxyglutarate (2-HG) is reduced, and increases the 2-HG level in patients by 100 folds [12]. Accumulation of 2-HG suppresses cell growth and also induces ferroptosis [13]. In the IDH1 mutation group, cell-cycle-related genes are down-regulated in contrast to IDH1 wild-type group (Figure 3a, b). However, hypermethylation of TP53 occurs in IDH1 mutant, owing to the accumulation of 2-HG [14]. In contrast, regulation of TP53 expression via methylation does not occur against the IDH1 wild-type background. Hypermethylation of DNA in IDH1 mutants results in the...
formation of the oncometabolite, 2-HG [15]. In short, most patients of grade 2 and 3 tumors with IDH1 mutation show suppression of cell growth and high methylation of TP53, resulting in the down-regulation of TP53 transcription in younger glioma patients. In contrast, promotion of cell growth and low methylation of TP53 are observed in most patients of grade 4 and a few of grade 3 tumors with IDH1 wild-type in older patients. Cell migration reactome terms were also included in the analysis. In the IDH1 wild-type background, degradation of keratan sulfate, which can interact with neuroregulatory ligands, dissolution of fibrin clot, and crosslinking of collagen fibrils, which are important for tumor cell and T cell migration, were higher than those in the IDH1 mutant [16-18]. Activation of NIMA kinases, which induce premature mitosis, is highly activated in tumors with IDH1 wild-type [19, 20]. Phosphorylation of early mitotic inhibitor (Emi1), which inhibits the anaphase-promoting complex/cyclosome (APC/C), is necessary for degradation of cyclins in mitosis [21]. APC/C with Cdc20 degrades type B cyclin and Nek2A and facilitates arrest of mitotic division [21-23]. Activity of Emi is lower for the IDH1 wild-type background; in contrast, activation of APC/C is higher for the IDH1 mutant background. In short, mitosis is suppressed more in tumors with IDH1 mutation than in tumors with IDH1 wild-type. Advanced glycosylation endproduct receptor signaling, which regulates cell proliferation, survival, differentiation, migration, and binding of formyl peptide receptors to formyl peptides and many other ligands, thereby regulating angiogenesis, cell proliferation and anti-apoptotic activities, is also higher in tumors with IDH1 wild-type [24, 25]. In contrast, BMP signaling associated with glioma cancer stem cell signaling by AKT-E17K, which may enhance signaling for cell-cycle arrest and regulation of FOXO by acetylation, thereby increasing cell death and decreasing cell proliferation, is higher in IDH1 mutant [26-28]. Negative regulation of mesenchymal epithelial transition and caspase-mediated cleavage of cytoskeleton, which are associated with morphological changes, are also higher in IDH1 mutants [29, 30].

Unsupervised clustering revealed that mutations in IDH1, TP53, and ATRX occur simultaneously (Figure 4a). Onset of mutations in TP53 and ATRX occur earlier in life, similar to onset of the IDH1 mutation (Figure 5a, b). Mutation in TP53, a cell-cycle related oncogene, is prevalent in various types of cancer and occurs frequently in various glioma subtypes and glioblastoma [31, 32]. The main function of normal TP53 involves the induction of cell-cycle arrest in the presence of DNA damage via binding to DNA and prevention of DNA replications [32]. Mutations in ATRX, which is mainly related to alternative lengthening of telomere (ALT) and regulation of epigenetic activity via methylation, are widely detected in glioma [33, 34]. Modulation of TP53 by ATRX regulates the upstream events of cancer [35-37]. Using unsupervised clustering of gene set extracted via multi-label IG, patients were divided on the basis of TP53 and ATRX double mutation and TP53 and ATRX wild-type with IDH1 mutation. (Figure 5b). PTEN regulation, negative control of transcription by E2F, and regulation of TP53 via acetylation were upregulated in the double mutation group. However, telomere extension by telomerase was promoted in the double wild-type group. Interestingly, the level of TERT expression in patients with double mutation was almost zero (Figure 6a). After separating the mutations of TP53 and ATRX, every patient with ATRX mutation showed a low level of TERT expression; however, TERT expression in patients with TP53 mutation and a grade 4 tumor did not differ from that in the wild-type patients. TERT encodes telomerase reverse transcriptase, which preserves the length of telomeress [38]. Loss of function of TERT induces telomere shortening and cell senescence, a sign of permanent cell-cycle arrest [38]. In short, stabilization of telomeres by telomerase or ALT is associated with various primary cancer types and malignant tumors [39-41]. The results of this study showed the relationship between TERT expression and mutations in ATRX and TP53. Recent studies have reported that TERT expression prevents extension of telomeres by ALT, which increases loss of ATRX; conversely, decrease in TERT expression, which is linked to the hypomethylation of TERT promoter, precedes ALT in zebrafish [42, 43]. Another study reported that TERT expression is downregulated in cells with ALT [44]. In short, ATRX loss resulting from ATRX mutation induces an increase in ALT and reduces TERT expression in hypomethylation
of the TERT promoter, which is related to the positive control of TERT expression in human brain cancer [34, 45]. Regarding the relationship between TP53 mutation and TERT expression, TP53 is normally activated by DNA damage and induces p21 and cell-cycle arrest [46]. ATM-Chk2-p53 may also participate in the feedback regulation of TERT expression in lung cancer [46]. Conversely, the lack of TP53 induces an increase in TERT expression [46]. ATRX and TP53 mutations differentially affect TERT expression, although we can infer that the impact of ATRX mutation is more than that of the TP53 mutation on prognosis of patients (Figure 6b). Regarding the relationship between TERT expression and ATRX mutation, we inferred that patients with ALT, which can occur due to loss of ATRX, have poorer prognosis than patients who have higher TERT expression.

EGFR signaling induces proliferation, which is one of the common events in various types of cancer [47]. Interestingly, the occurrence of EGFR mutation increased with aging and was independent of its occurrence with IDH1, TP53, and ATRX mutations (Figure 1a, 4a, and 7c). Unlike in patients with GBM, EGFR mutation occurred frequently in older patients with LGG. Patients with LGG harboring EGFR mutation showed poor prognosis for both grade 2 and 3; however, EGFR mutation in GBM did not affect prognosis (Figure 7a). After conducting multi-label IG and pathway enrichment analysis, most of the related pathways were found to be involved in the regulation of the cell cycle and cell signaling. Cell-cycle related pathways were downregulated and transcriptional regulation by VENTX was upregulated in grade 3 glioma and GBM. Detoxification of reactive oxygen species and interleukin-12 (IL-12) signaling, which exert an antitumor effect on tumors, including glioma and glioblastoma, were downregulated in grade 3 glioma and GBM [48-51].

In conclusion, mutations of IDH1, TP53, and ATRX occur in early-onset glioma, whereas EGFR mutation is associated with late-onset glioma. Interestingly, these mutations are associated with the WHO central nervous system tumor classifications. We believe that our results regarding mutations in IDH1, TP53, ATRX and EGFR can be applied to classify glioma according to age and precisely predict the prognosis of glioma. Furthermore, the results of transcriptomic analysis can suggest molecular therapeutic targets of age dependent glioma.

4. Materials and Methods

4.1. Description of TCGA Data and Entire Analysis Process

Transcriptome sequencing and clinical data of GBM and low-grade glioma were downloaded from the UCSC Xena database. The tumor statuses were LGG (n = 508) and GBMLGG (n = 657), and all the samples of the GBMLGG dataset were primary tumors (n = 657). All samples used in this study have clinical data, survival data, non-silent somatic mutation data, and transcriptomic data. The description of entire analysis process are visualized in flow chart (Figure 8).
4.2. Extraction of Candidate Somatic Mutations Related to Age

Candidate somatic mutations related to ages were selected based on a somatic mutation rate of at least 5% from somatic mutation data from multi-center mutation calling (MC3). Somatic mutation rate is calculated by individual who has mutation/all individuals. Logistic regression analysis was performed to identify somatic mutations related to age. Bonferroni correction was conducted to correct the p-value of logistic regression. Logistic regression analysis was conducted by binomial and corrected p-value which is 0.05 and 95 percent confidence intervals. The somatic mutation rate and results of logistic regression were visualized using forest plot v 1.10 (http://gforge.se/packages/)

4.3. Comparing Age with Mutation Status and Grade

The relationship between non-silent somatic mutations selected by logistic regression analysis and onset age are validated by comparing age at initial pathologic diagnosis with each mutations and grade. These results are visualized by ggpubr (https://rpkgs.datanovia.com/ggpubr/reference/ggboxplot.html).

4.4. Survival Analysis

Pairwise log-rank tests with Benjamini-Hochberg corrections were performed to compare prognoses between the groups, which were divided via agglomerative clustering with whole genes; Kaplan-Meier plots were generated in all survival analyses. Survival analysis was conducted using survival (https://CRAN.R-project.org/package=survival), and the survival plot was visualized using survminer (http://www.sthda.com/english/rpkgs/survminer/), both of which are R packages.
4.5. Multi-label IG based feature selection

Multi-labels, based on the histological grade and presence of somatic mutations in specific genes, were allotted to each individual. IG using multi-label individual and gene-level expression data was performed using FSlector v 0.31 [52, 53]. IG more than 0.35 was used as the feature selection threshold for the grade-mutation statuses of IDH1, ATRX, TP53, and EGFR.

4.6. Hierarchical Clustering and Heatmap with Annotations of Information

Hierarchical clustering of patients was performed by the gene expression levels of the gene set with an IG over 0.35. Visualization of hierarchical clustering data was performed using a Complex heatmap with annotation, which has information regarding tumor grade and genotype (mutant or wild-type) [54].

4.7. Pathway Enrichment Analysis

Reactome pathway enrichment analysis was performed using the ClueGo software, and significant pathways with Benjamini-Hochberg correction < 0.05 were extracted in Cytoscape [55]. p-values of the enriched pathways are presented as -log10 (p-value).

Supplementary Materials: The following are available online at www.mdpi.com/xxx/s1.

Author Contributions: Y.J. Park, J.M. Park and J.J. Lim contributed to the study conceptualization, study analysis, manuscript writing with the feedback from K.S. Sung. Y.J. Park, J.M. Park, J.W. Ahn, S.W. Kim and J.M. Shim processed genome data curation for public data base and performed bioinformatic analyses. S.J. Kang assisted in visualization and genome data curation. S.J. Hwang, K.S. Sung and J.J. Lim contributed to clinical data interpretation. J.J. Lim and K.S. Sung supervised the entire project and approved the final manuscript submission.

Funding: Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Korea Government (MSIT) (Grant No. 2018R1C1B5086460).

Acknowledgments: The authors thank the National Research Foundation of Korea (NRF).

Conflicts of Interest: The authors declare no conflict of interest.

Data Availability Statement

Abbreviations

<table>
<thead>
<tr>
<th>ALT</th>
<th>Alternative lengthening of telomere</th>
</tr>
</thead>
<tbody>
<tr>
<td>APC/C</td>
<td>Anaphase-promoting complex/cyclosome</td>
</tr>
<tr>
<td>GBM</td>
<td>Glioblastoma multiforme</td>
</tr>
<tr>
<td>IG</td>
<td>Information gain</td>
</tr>
<tr>
<td>LGG</td>
<td>Lower grade glioma</td>
</tr>
<tr>
<td>OS</td>
<td>Overall survival</td>
</tr>
<tr>
<td>PFI</td>
<td>Progression free interval</td>
</tr>
</tbody>
</table>

References

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 February 2021


33 Amorim, J.P.; Santos, G.; Vinagre, J.; Soares, P. The role of ATRX in the alternative lengthening of telomeres (ALT) phenotype. *Genes* 2016, 7, 66.
