

8-23-2024

## Expressions of the satellite repeat HSAT5 and transposable elements are implicated in disease progression and survival in glioma

SILA NAZ KÖSE

TUTKU YARAŞ

AHMET BURSALI

YAVUZ OKTAY

CİHANGİR YANDIM

*See next page for additional authors*

Follow this and additional works at: <https://journals.tubitak.gov.tr/biology>



Part of the [Biology Commons](#)

### Recommended Citation

KÖSE, SILA NAZ; YARAŞ, TUTKU; BURSALI, AHMET; OKTAY, YAVUZ; YANDIM, CİHANGİR; and KARAKÜLAH, GÖKHAN (2024) "Expressions of the satellite repeat HSAT5 and transposable elements are implicated in disease progression and survival in glioma," *Turkish Journal of Biology*. Vol. 48: No. 4, Article 3. <https://doi.org/10.55730/1300-0152.2700>

Available at: <https://journals.tubitak.gov.tr/biology/vol48/iss4/3>



This work is licensed under a [Creative Commons Attribution 4.0 International License](#).

This Research Article is brought to you for free and open access by TÜBİTAK Academic Journals. It has been accepted for inclusion in Turkish Journal of Biology by an authorized editor of TÜBİTAK Academic Journals. For more information, please contact [pinar.dundar@tubitak.gov.tr](mailto:pinar.dundar@tubitak.gov.tr).

---

## Expressions of the satellite repeat HSAT5 and transposable elements are implicated in disease progression and survival in glioma

### Authors

SILA NAZ KÖSE, TUTKU YARAŞ, AHMET BURSALI, YAVUZ OKTAY, CİHANGİR YANDIM, and GÖKHAN KARAKÜLAH

## Expressions of the satellite repeat HSAT5 and transposable elements are implicated in disease progression and survival in glioma

Sıla Naz KÖSE<sup>1,α</sup>, Tutku YARAŞ<sup>2,3,α</sup>, Ahmet BURSALI<sup>2</sup>, Yavuz OKTAY<sup>2,3</sup>,  
Cihangir YANDIM<sup>1,2,β</sup>, Gökhan KARAKÜLAH<sup>2,3,β,\*</sup>

<sup>1</sup>Department of Genetics and Bioengineering, Faculty of Engineering, İzmir University of Economics, İzmir, Türkiye

<sup>2</sup>İzmir Biomedicine and Genome Center (IBG), İzmir, Türkiye

<sup>3</sup>İzmir International Biomedicine and Genome Institute (IBG-İzmir), Dokuz Eylül University, İzmir, Türkiye

Received: 05.01.2024 • Accepted/Published Online: 01.07.2024 • Final Version: 23.08.2024

**Abstract:** The glioma genome encompasses a complex array of dysregulatory events, presenting a formidable challenge in managing this devastating disease. Despite the widespread distribution of repeat and transposable elements across the human genome, their involvement in glioma's molecular pathology and patient survival remains largely unexplored. In this study, we aimed to characterize the links between the expressions of repeat/transposable elements with disease progression and survival in glioma patients. Hence, we analyzed the expression levels of satellite repeats and transposons along with genes in low-grade glioma (LGG) and high-grade glioma (HGG). Endogenous transposable elements LTR5 and HERV\_a-int exhibited higher expression in HGG patients, along with immune response-related genes. Altogether, 16 transposable elements were associated with slower progression of disease in LGG patients. Conversely, 22 transposons and the HSAT5 satellite repeat were linked to a shorter event-free survival in HGG patients. Intriguingly, our weighted gene coexpression network analysis (WGCNA) disclosed that the HSAT5 satellite repeat resided in the same module network with genes implicated in chromosome segregation and nuclear division; potentially hinting at its contribution to disease pathogenesis. Collectively, we report for the first time that repeat and/or transposon expression could be related to disease progression and survival in glioma. The expressions of these elements seem to exert a protective effect during LGG-to-HGG progression, whereas they could have a detrimental impact once HGG is established. The results presented herein could serve as a foundation for further experimental work aimed at elucidating the molecular regulation of glioma genome.

**Key words:** Glioma, glioblastoma, transposon, satellite repeat, HSAT5, survival

### 1. Introduction

Gliomas are the most common type of central nervous system tumors representing approximately 80% of all malignant brain tumors (Ostrom et al., 2017). They can vary at genetic and epigenetic levels causing differences in tumor aggressiveness and hence a heterogeneous response to therapy (Noushmehr et al., 2010; Sturm et al., 2014; Sakthikumar et al., 2020; Garcia-Fabiani et al., 2021). Gliomas are histologically classified as astrocytic, oligodendroglial, or mixed (astrocytic and oligodendroglial) according to the cell type from which they originate (Louis et al., 2016). Importantly, they can be low-grade (LGG) or high-grade (HGG) according to their malignancy levels. LGGs typically arise de novo without any prior precursor or lower-grade lesions, and they have a better prognosis with approximately 75% five-year survival rates, while HGGs are grade III/IV gliomas generally differentiated from those of lower grades with a

much lower five-year survival rate of only 5% (Stupp et al., 2009; Crespo et al., 2015; Teng et al., 2022). Even though gliomas could arise from different cell types as mentioned above, they exhibit similar molecular characteristics based on their grade (i.e. LGG or HGG) and mutational statuses (Arcella et al., 2020).

Numerous studies have attempted to unravel histological and molecular differences between LGG and HGG to facilitate the identification of potential biomarkers and therapeutic targets. LGGs are usually characterized by IDH, ATRX, and TP53 alterations, as well as CDKN2A hypomethylation and 1p/19q codeletion (Parsons et al., 2008; Brat et al., 2015; Louis et al., 2016), whilst progression to HGG can be characterized by PTEN mutations, EGFR amplifications and MGMT promoter hypermethylation (Wang et al., 2016; Brito et al., 2019). It is noteworthy that some of these alterations may have distinct and sometimes opposite prognostic implications

\* Correspondence: gokhan.karakulah@ibg.edu.tr

<sup>α</sup>These authors contributed equally. <sup>β</sup>These authors contributed equally.

in LGG and HGG. Several mutations were found to be associated with better outcomes in LGG whilst presenting worse outcomes in HGG (Eckel-Passow et al., 2015; Juratli et al., 2018). Moreover, dysregulations in signaling pathways, mutations in oncogenes and tumor suppressors, and impaired epigenetic mechanisms may lead to genomic instability and prompt the chromatin of a glioma cell for further malignant progression (Louis et al., 2016; Mackay et al., 2017). Pivotal changes in chromatin architecture were observed for various types of gliomas and they were related to tumor aggressiveness, as well as patient survival (Mack et al., 2019; Wang et al., 2021; Garrett et al., 2022; Yang et al., 2022). Further insights into the dynamics of these chromatin aberrations would therefore pave the way for designing more effective treatment strategies.

There are a multitude of layers that influence chromatin architecture, one of which certainly is the content of DNA. At least 50% of the human genome is made up of repeat elements (International Human Genome Sequencing Consortium, 2001) and their transcription levels reflect the status of chromatin (Probst et al., 2010; Jagannathan and Yamashita, 2017). Even though their exact contribution to subnuclear regulation is yet to be characterized, expression of repeats was generally linked to vital biological processes such as heat stress response (Porokhovnik et al., 2021) and cellular senescence (Karakulah and Yandım, 2020). More importantly, tightly regulated expressions of these elements were particularly implicated in embryonic development (Ulitsky et al., 2011; Yandım and Karakulah, 2019a). Whereas tandemly repeated arrays such as the satellites could affect kinetochore formation and hence chromosomal stability (Vos et al., 2006; Zhu et al., 2011), interspersed elements such as DNA transposons or retrotransposons (i.e. long interspersed nuclear elements [LINEs], short interspersed nuclear elements [SINEs], long terminal repeats [LTRs], and human endogenous retroviruses [HERVs]) may act as a mutagenic force (Lee et al., 2012; Helman et al., 2014). Therefore, not surprisingly, dysregulated transcription levels of both satellites and transposons were reported to be potentially disease-causing in various cancers including but not limited to breast cancer (Zhu et al., 2011; Yandım and Karakulah, 2019b), hepatocellular carcinoma (Karakulah and Yandım, 2021), and pancreatic cancer (Ting et al., 2011; Yandım and Karakulah, 2022).

Despite reported changes in global chromatin and genome regulation, there have only been limited investigations into the contribution of repeat elements to the molecular pathogenesis and prognosis of gliomas. Ohka et al. reported that hypomethylation events in LINE elements were associated with shorter overall survival in glioma patients (Ohka et al., 2011). In addition, it was indicated that transcriptions of Alu elements were posttranscriptionally edited in oligodendroglioma (Hwang et al., 2022). Another recent study focusing on

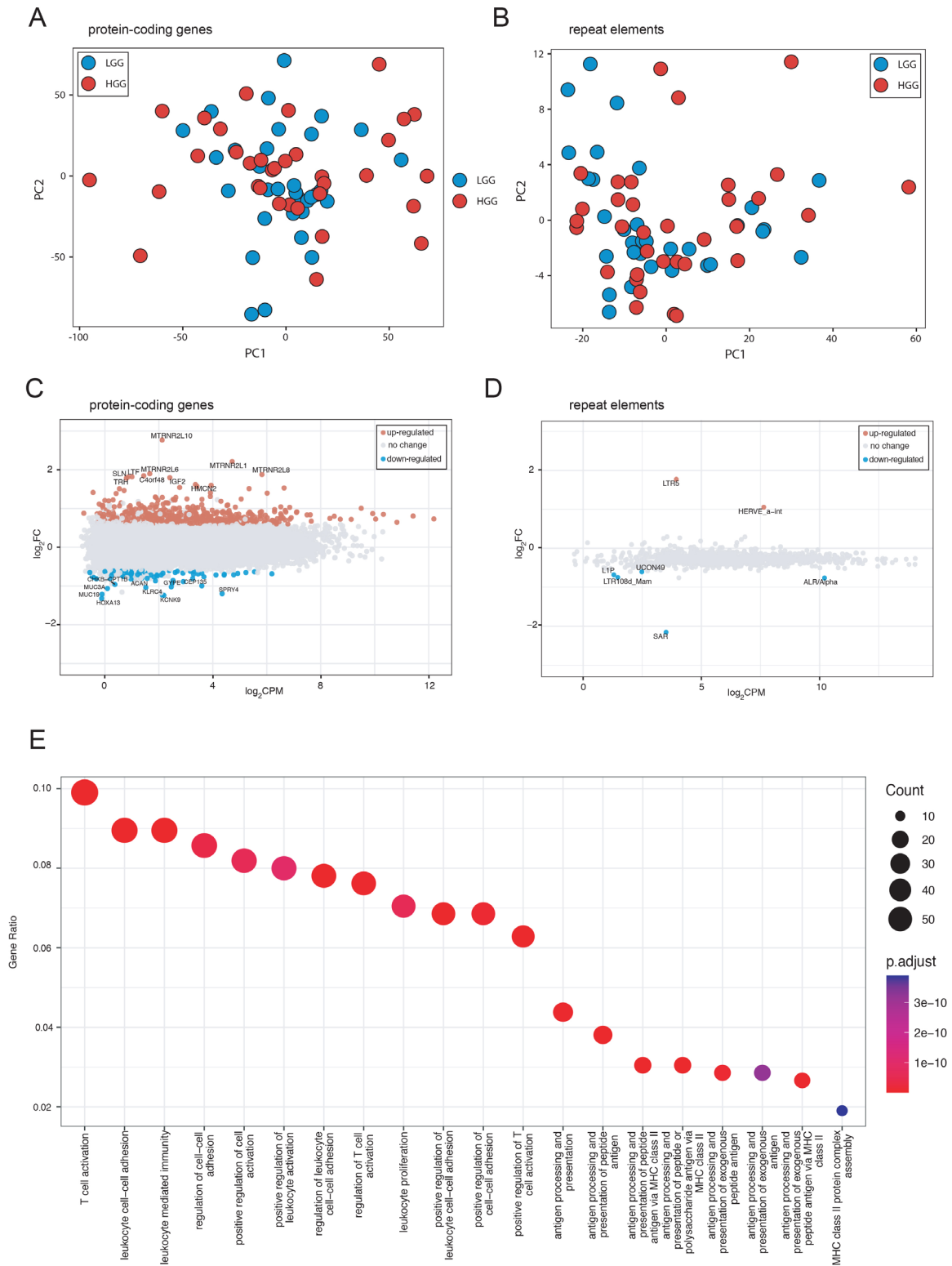
endogenous retroviruses showed that 43 HERV elements were upregulated in HGG compared to healthy brain tissue with implicated effects on proximal gene regulation (Koso et al., 2012; Yuan et al., 2021). On the other hand, the specific types of repeats/transposable elements that could be associated with disease progression and survival in glioma have not yet been characterized. This could be partly because most of the published RNA sequencing (RNA-seq) data is prepared with a poly(A) bias in library preparation, which renders the analysis of repeat-arisen transcripts inconclusive (Solovyov et al., 2018).

To elucidate the particular links between repeat/transposable element expression and disease progression and survival in glioma patients, we employed a publicly available dataset (GSE184941) for which total RNA-seq had been performed without an mRNA bias on the tumor resection samples of 32 LGG and 34 HGG patients with IDH mutations (Supplementary file 1: Table S1) (Xiao et al., 2022; Khan et al., 2023). This was the only available relevant dataset that allowed the quantification of total transcripts arisen from repeat elements and transposons whilst also presenting patient survival attributes. We identified differentially expressed satellite repeats and transposons between LGG and HGG samples. We also performed survival analyses considering the expressions of all human RepBase repeats (Bao et al., 2015) with Cox-regression and log-rank tests (Bewick et al., 2004), revealing a subset of repeats that are potentially important for LGG to HGG progression and survival.

## 2. Results

### 2.1. Identification of differentially expressed genes and repeat elements in HGG in comparison to LGG

To uncover the differences in gene and repeat/transposon expression, we quantified total RNA transcripts and performed a principal component analysis (PCA) (Figure 1 and Supplementary file 2: Figure S1). Even though a general distinction was observed in PCA analysis both for protein-coding gene expression and repeat arisen transcription levels between LGG and HGG samples (Figures 1A and 1B), the distribution of read percentages across patients were slightly different (Supplementary file 2: Figure S1A) with small changes in the abundance of repeat types giving rise to transcripts (Supplementary file 2: Figure S1B and S1C). When we applied a threshold of  $|\log_2(\text{FC})| > 0.6$  and false discovery rate (FDR)  $< 0.05$ , we identified 605 differentially expressed genes (Figure 1C and Supplementary file 3: Table S2) and 7 repeats (Figure 1D and Supplementary file 3: Table S3) in HGG samples when compared to LGG samples. A gene ontology analysis revealed that differentially expressed genes were mostly associated with biological processes implicating immunological response and cell-cell adhesion (Figure



**Figure 1.** Comparison of gene and repeat expressions between low-grade (LGG) and high-grade (HGG) gliomas. **A** Principal component analysis (PCA) of protein-coding genes. **B** PCA of repeat elements **C** MA plot of protein coding gene expressions **D** MA plot of repeat element expressions. **E** Gene ontology analysis results on biological processes for genes differentially expressed in HGG relative to LGG.

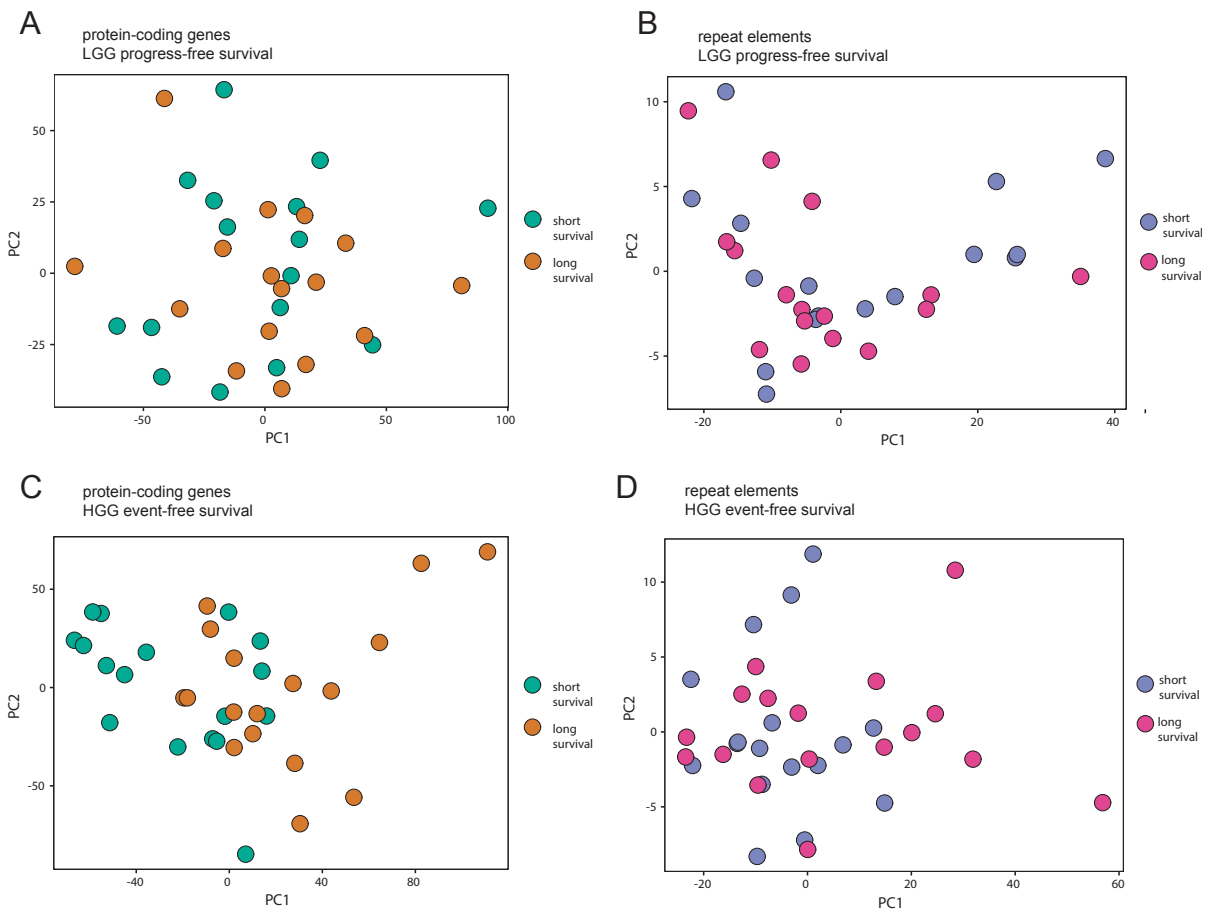
1E, Supplementary File 4: Table S4), aligning well with previous literature (Xu et al., 2019). Retrotransposons LTR5 and HERVE\_a-int were upregulated in HGG samples suggesting a possible role for them in the progression from LGG to HGG.

## 2.2. Identification of gene and repeat expressions implicated in disease progression and survival

Gene expression signatures that influence survival times of patients provide invaluable information on the carcinogenesis and metastasis process in cancer (Sahu et al., 2021). Therefore, they have the potential to open new treatment avenues. To provide a detailed picture of the survival-related expression events where repeat-arisen transcripts were also considered in glioma, we performed a Cox regression analysis followed by a log-rank test (with a median progression/survival time cut-off) not only for the survival of HGG patients but also for the progression of LGG patients (to HGG). A PCA analysis showed a

more pronounced separation between shorter and longer survival in the event-free survival analysis of HGG samples as opposed to progress-free survival analysis of LGG samples (Figure 2).

Based on progress-free survival analysis in LGG samples, we identified 71 genes and 16 repeats (Cox Hazard Ratio [HR] < 0.5, Cox p-value < 0.05 and log-rank p-value < 0.05) that were associated with slower LGG to HGG progression (Table 1, Supplementary file 5: Tables S5 and S7). On the other hand, 111 genes, but none of the repeats, (Cox HR > 1.5, Cox p-value < 0.05 and log-rank p-value < 0.05) were found to be related to faster LGG to HGG progression (Table 1, Supplementary file 5: Tables S6 and S8). In the samples where progression to HGG took place, event-free (event: death) survival analysis revealed 48 genes, but no repeats (Cox HR < 0.5, Cox p-value < 0.05 and Log-rank p-value < 0.05) were associated with longer survival (Table 1, Supplementary file 6: Tables S9 and



**Figure 2.** Principal component analyses (PCA) of the expressions of genes implicated in glioma progress and survival with a median cut-off. **A** PCA of samples associated with short and long progress-free survival in LGG considering the protein-coding gene expressions. **B** PCA of samples associated with short and long progress-free survival in LGG considering repeat element expressions. **C** PCA of samples associated with short and long event-free survival in HGG considering the protein-coding gene expressions. **D** PCA of samples associated with short and long progress-free survival in LGG considering repeat element expressions.

**Table 1.** Significant repeat elements linked to patient survival in LGG and HGG.

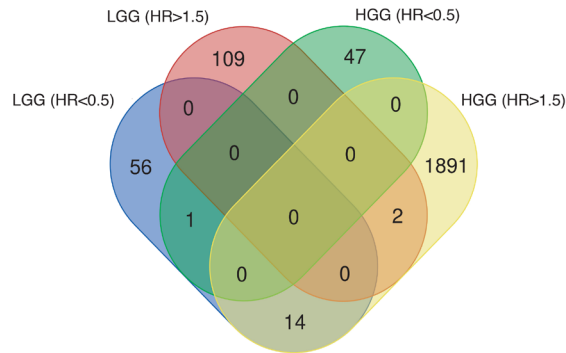
Low-grade glioma (LGG)				High-grade glioma (HGG)					
	Repeat	Cox HR (<0.5)	Repeat	Cox HR (>1.5)	Repeat	Cox HR (<0.5)	Repeat	Cox HR (>1.5)	
	Longer progress-free survival (Log-rank p<0.05)	LTR1F2		0.226566207		None		None	Longer event-free survival (Log-rank p < 0.05)
MER107		0.246705973	HERV1_LTRb	3.772926264					
MER1B		0.253616041	MLT1E3-int	3.693884586					
LTR57		0.342074886	L1M2a	3.661849157					
LTR1C		0.362824242	L1M3f	3.640586269					
MER130		0.379093156	CRI-11_Crp	3.092362862					
LTR40c		0.388071793	MER95	2.958200736					
LTR1		0.394057275	UCON97	2.91316852					
LTR1A2		0.394722746	MSTC-int	2.886032148					
LTR2752		0.398195734	MER70-int	2.688166503					
LTR18C		0.421946618	HERVI-int	2.664554289					
LTR27C		0.449130349	L1MDa	2.613437037					
LTR35A		0.458301493	HSAT5	2.598739703					
MER9B		0.465923328	LTR75	2.456202539					
LTR27		0.476049565	HERVK3-int	2.414286886					
LTR7C	0.491857327	L1MB2	2.396560074						
		L1M8	2.340619317						
		LTR70	2.3387068						
		L1M4a1	2.31165162						
		LTR18C	2.18835578						
		MER51-int	2.165050709						
		HERVKC4-int	1.908960599						
		Kanga1c	1.852908386						

S11). Interestingly, 1907 genes and 23 repeats (Cox HR >1.5, Cox p-value<0.05 and Log-rank p-value<0.05) were linked to shorter event-free survival in HGG (Table 1, Supplementary file 6: Tables S10 and S12). Only one gene (CASS4) was associated with longer progress-free survival in LGG, as well as longer event-free survival in HGG (Figure 3A). On the other hand, two genes (GPR156 and AGAP1), but no repeats, were associated with shorter survival both in progress-free analysis of LGG samples and event-free analysis of HGG samples at the same time. Intriguingly, 14 genes (IPO9, PLK1, DCAF4L1, ZNF345, LRRC75A, RAD18, BMP8A, LRR1, SETD6, AK2, MAP3K12, LMBR1L, ERCC6L, ZNF776) and one transposon (LTR18C) were related with longer time for LGG-to-HGG progression, yet in opposition, they were linked to shorter event-free survival in HGG samples. Gene ontology analyses revealed that LGG-to-HGG progression could be attributed to neurotransmitter release (Figure 3B, Supplementary file 7: Table S13), whereas survival in HGG could be related to various cell cycle events including spindle organization, chromosome segregation, and nuclear division (Figure 3C, Supplementary file 8: Table S14).

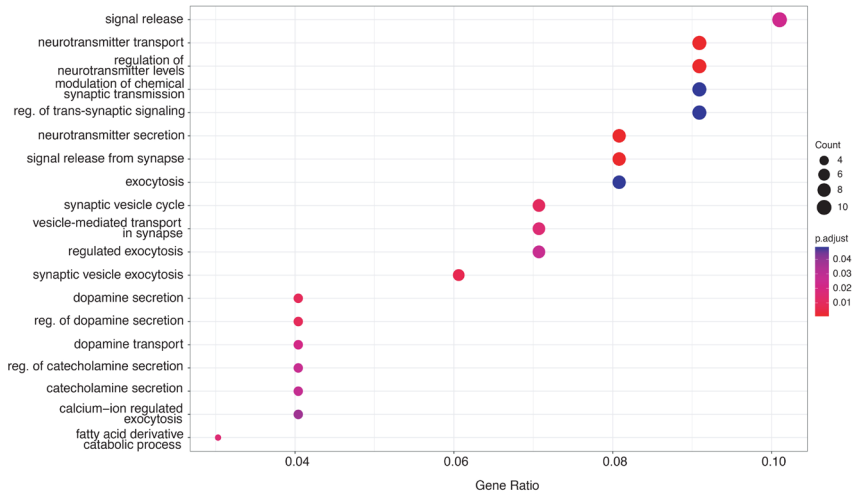
**2.3. Weighted gene co-expression network analysis (WGCNA) reveals the expression of HSAT5 satellite as a potential key player in HGG**

In order to further elucidate specific gene expression correlation patterns that could be distinctly related to LGG or HGG and to provide a better insight into the cooperation of repeat expression in these patterns, we performed weighted gene co-expression network analysis (WGCNA) (Zhang and Horvath, 2005; Cantone and Fisher, 2013). Our pipeline contained a pool approach where not only the transcript levels of genes but also those of repeats were analyzed altogether. Overall, 13 modules that contained genes and repeats were identified and these were represented with different colors (Figure 4A, Supplementary file 9: Table S15). Tan module was specifically emphasized in LGG samples, and the green module was particularly associated with HGG. Whereas the tan module contained no repeats but included 185 genes, the green module comprised only the satellite repeat HSAT5, no transposons, and 981 genes. The majority (987) of repeat elements/transposons fell into turquoise module, which was emphasized in LGG samples and where all 16 repeats linked to longer LGG progress-free survival

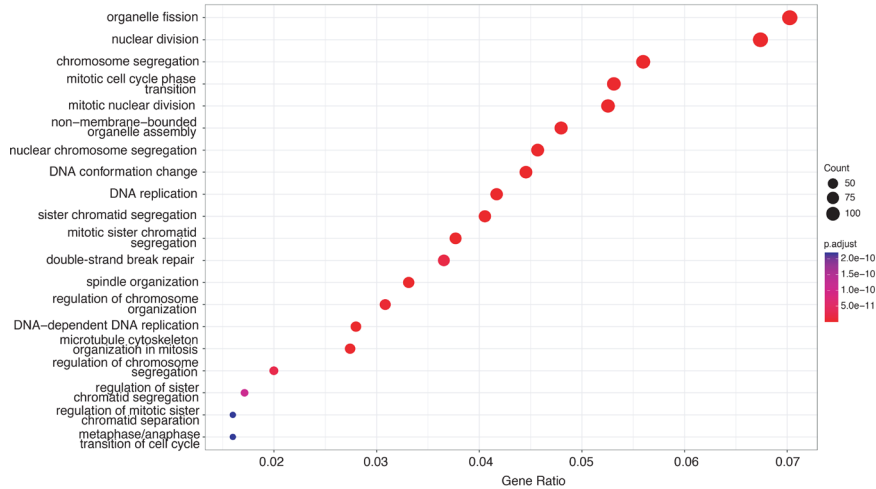
A



B



C



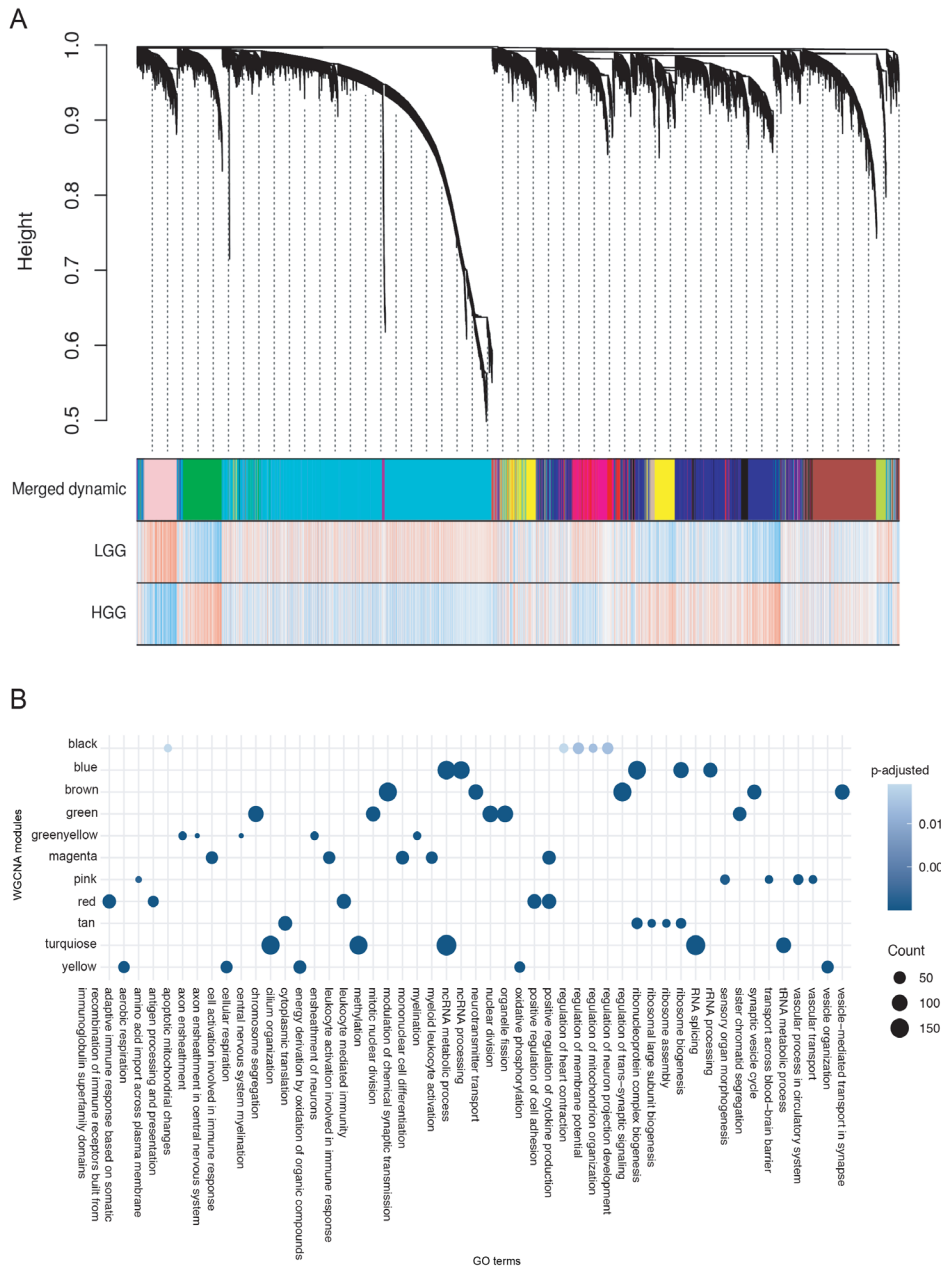
**Figure 3.** Progress-free survival genes in LGG and event-free survival genes in HGG. **A** Hazardous (HR>1.5, Log-rank p<0.05) and protective (HR<0.5, Log-rank p<0.05) genes based on Cox regression analyses and Log-rank test in LGG and HGG. **B** Gene ontology analysis of hazardous genes in LGG-to-HGG progression. **C** Gene ontology analysis of hazardous genes in HGG event-free survival.



(Table 1) were also present. The blue module, on the other hand, was pronounced in HGG and contained 5 repeats, among which only LIM2a was linked to event-free survival in HGG. Greenyellow module was slightly pointing to LGG, and it contained only one repeat.

When a gene ontology analysis was run for all of the WGCNA modules, genes related to ribosome biogenesis and protein translation seemed to be enriched in the tan module (Figure 4B, Supplementary file 10). Interestingly, the green module displayed an enrichment in genes

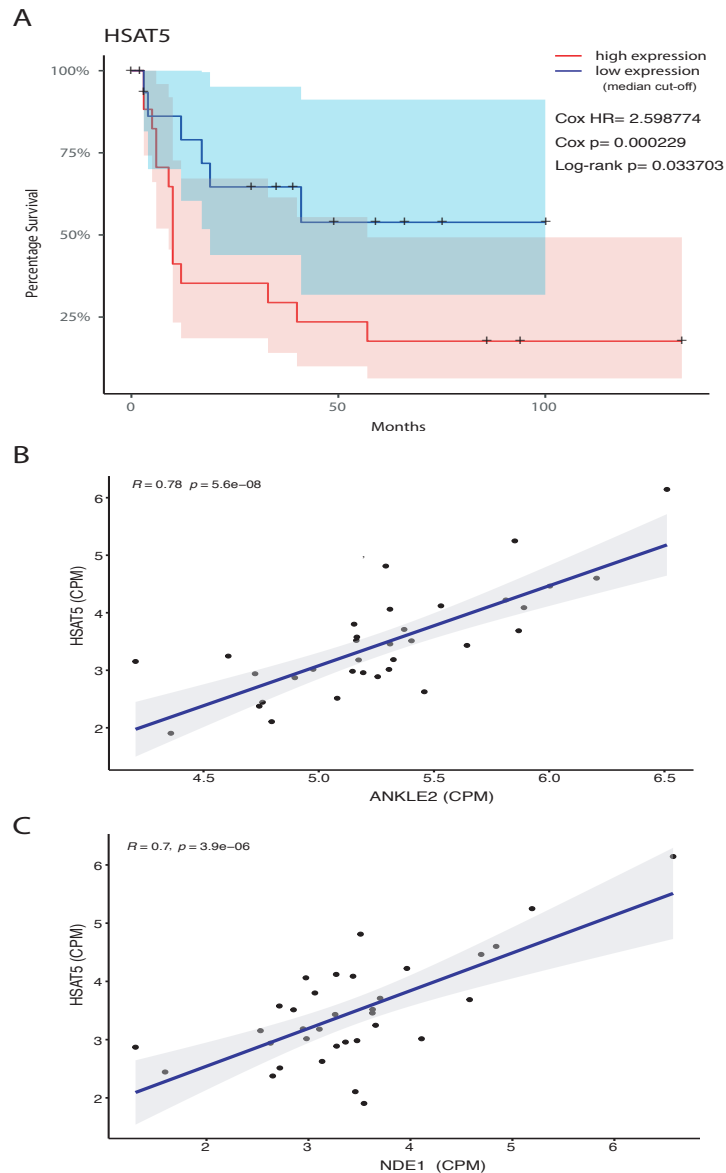
that were linked to nuclear division, chromosome segregation, and double-strand break repair. Turquoise module, where the majority of repeats reside, was enriched with genes linked to cilium organization, methylation and ncRNA metabolic process, splicing and tRNA metabolic process. The blue module was as well in an enrichment with ncRNA processing and ribosome biogenesis genes. On the other hand, the greenyellow module was related to myelination and neuron ensheathment.



**Figure 4.** Weighted Gene Co-expression Network Analysis (WGCNA) of gene and repeat expressions. A WGCNA modules B Gene ontology analysis of the genes in WGCNA modules.

Our interesting findings on HSAT5 satellite repeat which was linked to shorter event-free survival in HGG (Figure 5A) and the fact that it appeared in the same WGCNA module (green) with genes that were linked to similar biological phenomena (i.e. nuclear division and chromosome segregation) presented in Figure 3C for HGG event-free survival, prompted us to run a Pearson correlation analysis for the expression levels of HSAT5 and all protein-coding genes in the genome (Supplementary file 11). We found that 310 genes exhibited moderate-to-high level of correlation ( $r > 0.5$ )

with HSAT5 expression. Strikingly, 195 (63%) of these 310 genes were among those significantly linked to shorter HGG event-free survival as presented in Table S10 (Supplementary file 6). A thorough literature search also showed that most of these genes were reported to be influential in glioma (Table 2). Some of these HSAT5-correlated genes including but not limited to ANKLE2 (Figure 5B) and NDE1 (Figure 5C) could be involved in the regulation of HSAT5 expression due to their crucial functions on chromatin (Chomiak et al., 2022; Sebastian et al., 2022).



**Figure 5.** Survival analysis of HSAT5 expression and its correlation with key chromatin factors. A Kaplan-Meier event-free survival graph of HGG patients based on HSAT5 satellite repeat expression with a median cut-off. B,C Scatter plot representations showing the correlation of HSAT expression with the expressions of ANKLE2 and NDE1 genes in HGG patients.

**Table 2.** Genes, whose expressions significantly correlated with that of HSAT5 satellite repeat, and their known associations with gliomas. Gene names written in bold characters indicate a direct link with survival in HGG as represented in Figures 1A and 1C.

Gene name	p-values	Pearson's corr. coef. (r)	Direct association with glioma
<b>ANKRD20A1</b>	6.48E+02	0.723955	Membrane shaping (Wolf et al., 2019)
<b>QTRT2</b>	2.01E+02	0.712178	Ferroptosis (Sallam et al., 2022)
<b>ANKRD20A3P</b>	1.72E+04	0.688139	-
<b>IGF2BP2</b>	2.47E+04	0.683868	Chemoresistance (Han et al., 2022)
<b>RALY</b>	2.09E+04	0.656957	-
<b>CCDC150</b>	3.09E+05	0.651718	-
TRIO	3.79E+05	0.648938	Cell migration, invasion and disease outcome (Salhia et al., 2008)
ODF2	4.04E+04	0.648070	Hypoxia (Baghbani et al., 2013)
<b>ANKLE2</b>	6.30E+05	0.641924	Apoptosis and survival (Wang et al., 2018)
<b>SHC4</b>	7.97E+05	0.638604	EMT, disease progression and outcome (Tilak et al., 2021)
<b>TTF2</b>	1.27E+06	0.631903	Cell cycle and repair mechanisms, extracellular matrix organization (Chai et al., 2019)
<b>ANKRD11</b>	1.53E+06	0.629146	-
<b>GSDMB</b>	1.88E+06	0.626186	Immune cell infiltration (Lin J et al, 2022)
<b>MCM8</b>	1.91E+06	0.625905	Stemness (Wang et al., 2021)
<b>SGCA</b>	2.22E+06	0.623704	ECM signaling (Balvers et al., 2013)
<b>NDE1</b>	2.28E+06	0.623247	Cell proliferation, migration, and disease prognosis (Chen et al., 2017; Song et al., 2019)
<b>DIAPH3</b>	2.94E+06	0.619450	Chemoresistance (Chehade et al. 2022)
<b>PIAS2</b>	3.46E+06	0.616934	Posttranslational modification (Li and Meng, 2021)
<b>PLCG1</b>	3.74E+06	0.615770	Disease prognosis (Li et al., 2022)
FUS	3.76E+06	0.615690	Angiogenesis (He et al., 2019)
<b>QRICH2</b>	4.18E+06	0.614060	-
<b>CAPRIN2</b>	4.62E+06	0.612492	-
<b>WDR62</b>	4.98E+05	0.611313	-
<b>SFI1</b>	5.52E+06	0.609695	-
<b>GNL3L</b>	6.79E+06	0.606424	Survival (Sun et al., 2021)
<b>POLE</b>	7.04E+06	0.605863	Genomic stability and disease progression (Erson-Omay et al., 2015)
<b>ZBTB49</b>	7.97E+06	0.603872	-
<b>ANAPC7</b>	8.56E+06	0.602737	Disease progression (de Groot et al., 2011)
<b>PSMC3IP</b>	9.16E+06	0.601632	Angiogenesis (Zhang et al., 2017)

### 3. Discussion

The complexity of molecular pathogenesis underlying gliomas makes them refractory to current treatment strategies. Notably, most of the patients that die due to gliomas suffer from HGG, which in most cases stem from a preceding LGG (Bogdańska et al., 2017). Therefore, understanding the particular events that cause progression to HGG from LGG is important in terms of projecting more insightful approaches toward gliomas. In this study, we uncovered not only differentially expressed protein-coding genes but also transcripts arisen from satellite repeats and transposons in HGG as opposed to LGG, shedding further light into disease progress. We found that genes related to immune response were differentially expressed in HGG as opposed to LGG, agreeing well with previous reports (Gabrusiewicz et al., 2016; Lu et al., 2019; Xiao et al., 2022;

Khan et al., 2023). This could be a projected outcome due to immune cell infiltration as reported before (Domingues et al., 2016; Lin et al., 2021). Another interesting finding was the upregulated expression (in HGG) of HERV\_a-int and LTR5 transposon, the latter of which was previously linked to cancer (Lee et al., 2012) and human preimplantation development (Yandım and Karakülah, 2019a).

Our progress-free survival analysis of LGG samples on gene expressions revealed that genes related to the regulation of neurotransmitter release, dopamine secretion and exocytosis were linked to faster progression to HGG. Even though genes within this context were not reported in some of the studies that analyzed clinical samples (Wang et al., 2020; Zhang et al., 2020), in vitro and animal studies previously suggested dopamine release as a potential factor in glioma (Li et al., 2014); however, its effect was not

entirely understood (Lan et al., 2017). Another interesting finding in our study was that all repeat elements associated with LGG-to-HGG progression were those that had a hazard ratio of less than 0.5. In other words, transposon expression could potentially slow down the progression to HGG. This could be due to increased antigenicity and better immune response caused by transposon expression as suggested earlier (Kong et al., 2019; Yuan et al., 2021). In opposition, event-free survival analysis in HGG revealed that none of the repeat elements were associated with longer survival. However, it identified one satellite (i.e. HSAT5) and 23 transposons linked to shorter survival. Though this result suggests that repeat element expression may lead to opposite consequences in LGG and HGG, only one repeat (i.e. LTR18C) appeared in the survival analyses of both LGG (HR<0.5) and HGG (HR>1.5) groups. In line with this, the genes linked to shorter survival in HGG were also pointing out a different context as opposed to those in LGG. Genes with HR>1.5 were mostly linked to chromosome segregation and nuclear division, generally agreeing well with previous literature (Bredel et al., 2005; Phillips et al., 2006; Murat et al., 2008; Q. Wang et al., 2017).

In an effort to dissect the impact of repeat element expression on the molecular pathogenesis and survival in glioma, we not only presented LGG progress-free survival related repeats and HGG event-free survival linked repeats but also differentially expressed repeats in HGG samples in comparison to LGG. Yuan et al. (2021) reported differentially expressed repeats in tissue samples of HGG (GBM) patients as opposed to normal brain samples; however, the data they employed was prepared with an mRNA bias only detecting poly(A)+ transcripts. Our study utilizes sequencing data obtained from total RNA samples without the poly (A) bias (Khan et al., 2023) as previously recommended (Solovyov et al., 2018). Notably, LTR5 was reported to be upregulated both in a previous study (Yuan et al., 2021), which compares the HGG tissue with normal tissue, and in our study that compares with the HGG tissue with LGG, even though no survival related effect was detected. Another element we detected in HGG event-free survival analysis was HERVK3-int, which was reported to be upregulated in HGG samples with a potential effect on disease progression (Shah et al., 2022). Our holistic approach identified many other repeats/transposons previously unmentioned in the glioma literature. Perhaps, the most interesting one among these is the satellite repeat HSAT5, which predominantly resides in pericentromeric regions of chromosomes (Hillier et al., 2005; Hubley et al., 2016) and was also reported to be upregulated in cellular senescence (Karakulah and Yandim, 2020). Though no studies so far reported HSAT5 within the context of cancer, other satellites such as HSATI (Zhu et al.,

2011), HSATII (Ting et al., 2011), and GSATII (Karakulah and Yandim, 2021) were linked to the molecular pathologies of solid cancers. Satellite repeats are known to play pivotal roles in maintaining heterochromatin architecture and kinetochore formation (Rošić et al., 2014; Nishibuchi and Déjardin, 2017) and their abnormal expressions were linked to genomic catastrophes in cell division (Zhu et al., 2011; Bersani et al., 2015; Kishikawa et al., 2016). Their usage as biomarkers is also being valuably considered (Özgür et al., 2021). Our results point out HSAT5 satellite repeat as a potential prognostic biomarker in HGG survival. More importantly, its cooperative expression levels with HGG-specific genes that are known to involve in chromosome segregation and nuclear division could implicate its potential in the molecular pathology of this devastating disease. The relationship of HSAT5 DNA or RNA with chromatin-related factors such as NDE1 and ANKLE2 could illuminate the mechanism of HSAT5 expression further.

#### 4. Conclusion

In this study, we performed a comprehensive analysis on the expressions of repeat elements covering both satellite repeats and transposons to figure out those implicated in glioma progression and survival. Our results interestingly revealed that the expressions of transposable elements could be linked to slower LGG to HGG progression, yet this beneficial effect is overturned once the progress to HGG is manifested. One explanation for the beneficial effect in LGG could be the improved antigenicity caused by repeat elements, which provides better immune response. On the other hand, once the progression to HGG takes place, repeats could be contributing to dysregulated events in chromosome segregation and cell cycle as mostly the results on HSAT5 satellite repeat here suggested. Results presented here could serve as a guide for further experimental work in understanding the molecular basis of gliomas further.

#### 5. Methods

##### 5.1. Data acquisition and preprocessing

We obtained raw sequencing data from the Sequence Read Archive (SRA) database (Leinonen et al., 2011) for 66 glioblastoma patients (SRA accession: SRP339173; GEO accession: GSE184941) using the “fastq-dump -gzip -skip-technical -readids -dumpbase -clip -split-3” command of the SRA Tool Kit v.2.9.0. We collected the GRCh38 human genome assembly and comprehensive gene annotation (release 34) in gene transfer format from the GENCODE<sup>1</sup> website (Frankish et al., 2021). We obtained genomic coordinates of repeat instances across the genome from the RepeatMasker<sup>2</sup> website. We aligned paired-end reads from

<sup>1</sup><https://www.genecodegenes.org>

<sup>2</sup><http://www.repeatmasker.org/>

each tumor sample to the human reference genome using the Rsubread package v1.34.7 (Liao et al., 2014) in the R statistical computation environment with default parameters. We used SAMtools suite v1.3.1 (Li et al., 2009) to sort and index the binary alignment map files. We counted reads using the featureCounts function of the Rsubread package (Liao et al., 2014), which allowed us to differentiate between exonic and repetitive regions. We merged repeat and gene counts into a single expression matrix, and calculated counts per million (CPM) values for each gene and repeat element across all tumor samples. We excluded any gene or repeat element that was not expressed above 1 CPM in at least 25% of the tumor samples from further analysis. Besides, when filtering for HGG and LGG samples, IDH mutants and percentage of total mapped reads greater than 70 were included in both samples. To improve the accuracy of estimated repeat element expressions, we only considered sequencing reads that overlapped nonexonic regions and uniquely mapped to DNA, LINE, SINE, LTR, and satellite repeat regions.

## 5.2. Coexpression network analysis of protein-coding genes and repeat elements

To perform coexpression network analysis, we used the R package WGCNA v1.47 (Langfelder and Horvath, 2008). Adjacencies between all protein-coding genes and repeats across samples were calculated. In this step, CPM values of genomic features were used as input. We designated the soft threshold value as 12 and the average linkage hierarchical clustering method was selected for grouping the genomic features with similar expression patterns. Additionally, the dynamic tree cut algorithm (Langfelder et al., 2008) was utilized to determine network modules and minimum network module size was set to 30. Gene ontology (GO) enrichment analysis of each network module was performed with the R package clusterProfiler v3.18.0 (Yu et al., 2012).

<sup>3</sup>v3.6.0; <https://www.r-project.org/>

## 5.3. Statistical analysis and graphical representation

Statistical analysis and graphical representation were performed using the R environment<sup>3</sup> for the LGG and HGG samples. Differential expression (DE) analysis of genomic features was performed using the edgeR package (version 3.28.2) (Robinson et al., 2010). We obtained a PCA plot of samples to visualize clusters of LGG and HGG samples based on their similarities (quantile median = 0.50). We used the log rank test (logrank\_test function of the coin package, version 1.4.1) (Royston et al., 2019) to test the statistical significance of survival time differences between groups and the Cox regression analysis (coxph function of the survival package, version 3.2.11) to estimate hazard ratios (HR). The Kaplan-Meier curves were drawn using the survfit function to show the probabilities of survival for a certain time interval. We used the ggplot2 package (version 3.3.6) to create other graphical representations.

## Ethical approval and consent to participate

Not applicable.

## Consent for publication

Not applicable.

## Availability of data and material

The dataset analyzed during the current study is available in the GEO repository, with accession number GSE184941.

## Competing interests

The authors declare that there are no competing interests.

## Funding

Not applicable.

## References

- Apridita Sebastian W, Shiraishi H, Shimizu N, Umeda R, Lai Set al. (2022). Ankle2 deficiency-associated microcephaly and spermatogenesis defects in zebrafish are alleviated by heterozygous deletion of vrk1. *Biochemical and Biophysical Research Communications*, 624: 95–101. <https://doi.org/10.1016/j.bbrc.2022.07.070>
- Arcella A, Limanaqi F, Ferese R, Biagioni F, Oliva MA et al. (2020). Dissecting molecular features of gliomas: genetic loci and validated biomarkers. *International Journal of Molecular Sciences*, 21 (2). <https://doi.org/10.3390/ijms21020685>
- Baghbani F, Raoofian R, Hasanzadeh NM, Hamzehloei T, Soukhtanloo M et al. (2013). Identification of novel hypoxia response genes in human glioma cell line a172. *Iranian Journal of Basic Medical Sciences*, 16 (5): 675–82.
- Bao W, Kojima KK, Kohany O (2015). Repbase Update, a database of repetitive elements in eukaryotic genomes. *Mobile DNA*, 6: 11. <https://doi.org/10.1186/s13100-015-0041-9>
- Balvers RK, Kleijn A, Kloezeman JJ, French PJ, Kremer A et al. (2013). Serum-free culture success of glial tumors is related to specific molecular profiles and expression of extracellular matrix-associated gene modules. *Neuro-Oncology*, 15 (12): 1684–95. <https://doi.org/10.1093/neuonc/not116>
- Bersani F, Lee E, Kharchenko PV, Xu AW, Liu M et al. (2015). Pericentromeric satellite repeat expansions through RNA-derived DNA intermediates in cancer. *Proceedings of the National Academy of Sciences of the United States of America*, 112 (49): 15148–15153. <https://doi.org/10.1073/pnas.1518008112>

- Bewick V, Cheek L, Ball J (2004). Statistics review 12: survival analysis. *Critical Care* (London, England), 8 (5): 389–394. <https://doi.org/10.1186/cc2955>
- Bogdańska MU, Bodnar M, Piotrowska MJ, Murek M, Schucht P et al. (2017). A mathematical model describes the malignant transformation of low grade gliomas: prognostic implications. *PLoS One*, 12 (8): e0179999. <https://doi.org/10.1371/journal.pone.0179999>
- Brat DJ, Verhaak RGW, Aldape KD, Yung WKA, Salama SR et al. (2015). Comprehensive, integrative genomic analysis of diffuse lower-grade gliomas. *The New England Journal of Medicine*, 372 (26): 2481–2498. <https://doi.org/10.1056/NEJMoa1402121>
- Bredel M, Bredel C, Juric D, Harsh GR, Vogel H et al. (2005). Functional network analysis reveals extended gliomagenesis pathway maps and three novel MYC-interacting genes in human gliomas. *Cancer Research*, 65 (19): 8679–8689. <https://doi.org/10.1158/0008-5472.CAN-05-1204>
- Brito C, Azevedo A, Esteves S, Marques AR, Martins C et al. (2019). Clinical insights gained by refining the 2016 WHO classification of diffuse gliomas with: *EGFR* amplification, *TERT* mutations, *PTEN* deletion and *MGMT* methylation. *BMC Cancer*, 19 (1): 968. <https://doi.org/10.1186/s12885-019-6177-0>
- Cantone I, Fisher AG (2013). Epigenetic programming and reprogramming during development. *Nature Structural & Molecular Biology*, 20 (3): 282–289. <https://doi.org/10.1038/nsmb.2489>
- Chai RC, Li YM, Zhang KN, Chang YZ, Liu YQ et al. (2019). RNA processing genes characterize RNA splicing and further stratify lower-grade glioma. *JCI Insight*, 5 (17): e130591. <https://doi.org/10.1172/jci.insight.130591>
- Chehade G, Ruiz-Reig N, Raftopoulos C, Tajeddine N, Tissir F (2022). P10.02.B The fragile balance between proliferation and neoplasia in neural stem cells: role of DIAPH3 in glioblastoma. *Neuro-Oncology*, 24: (Issue Supplement\_2), ii48. <https://doi.org/10.1093/neuonc/noac174.167>
- Chen HY, Lin LT, Wang ML, Laurent B, Hsu CH et al. (2017). Musashi-1 enhances glioblastoma cell migration and cytoskeletal dynamics through translational inhibition of Tensin3. *Scientific Reports*, 7 (1): 8710. <https://doi.org/10.1038/s41598-017-09504-7>
- Chomiak AA, Guo Y, Kopsidas CA, McDaniel DP, Lowe CC et al. (2022). Nde1 is required for heterochromatin compaction and stability in neocortical neurons. *iScience*, 25 (6): 104354. <https://doi.org/10.1016/j.isci.2022.104354>
- Crespo I, Vital AL, Gonzalez-Tablas M, Patino MDC, Otero A et al. (2015). Molecular and Genomic Alterations in Glioblastoma Multiforme. *The American Journal of Pathology*, 185 (7): 1820–1833. <https://doi.org/10.1016/j.ajpath.2015.02.023>
- Domingues P, González-Tablas M, Otero Á, Pascual D, Miranda D et al. (2016). Tumor infiltrating immune cells in gliomas and meningiomas. *Brain, Behavior, and Immunity*, 53: 1–15. <https://doi.org/10.1016/j.bbi.2015.07.019>
- Eckel-Passow JE, Lachance DH, Molinaro AM, Walsh KM, Decker PA et al. (2015). Glioma Groups Based on 1p/19q, IDH, and TERT Promoter Mutations in Tumors. *The New England Journal of Medicine*, 372 (26): 2499–2508. <https://doi.org/10.1056/NEJMoa1407279>
- Erson-Omay EZ, Çağlayan AO, Schultz N, Weinhold N, Omay SB et al. (2015). Somatic *POLE* mutations cause an ultramutated giant cell high-grade glioma subtype with better prognosis. *Neuro-Oncology*, 17 (10): 1356–64. <https://doi.org/10.1093/neuonc/nov027>
- Frankish A, Diekhans M, Jungreis I, Lagarde J, Loveland JE et al. (2021). GENCODE 2021. *Nucleic Acids Research*, 49 (D1): D916–D923. <https://doi.org/10.1093/nar/gkaa1087>
- Gabrusiewicz K, Rodriguez B, Wei J, Hashimoto Y, Healy LM et al. (2016). Glioblastoma-infiltrated innate immune cells resemble M0 macrophage phenotype. *JCI Insight*, 1 (2). <https://doi.org/10.1172/jci.insight.85841>
- Garcia-Fabiani MB, Haase S, Comba A, Carney S, McClellan B et al. (2021). Genetic Alterations in Gliomas Remodel the Tumor Immune Microenvironment and Impact Immune-Mediated Therapies. *Frontiers in Oncology*, 11: 631037. <https://doi.org/10.3389/fonc.2021.631037>
- Garrett MC, Albano R, Carnwath T, Shah S, Woo D et al. (2022). Chromatin structure predicts survival in glioma patients. *Scientific Reports*, 12 (1): 8221. <https://doi.org/10.1038/s41598-022-11019-9>
- de Groot JF, Lamborn KR, Chang SM, Gilbert MR, Cloughesy TF et al. (2011). Phase II study of aflibercept in recurrent malignant glioma: a North American Brain Tumor Consortium study. *Journal of Clinical Oncology*, 29 (19): 2689–95. <https://doi.org/10.1200/JCO.2010.34.1636>
- Han J, Yu X, Wang S, Wang Y, Liu Q et al. (2022). IGF2BP2 induces U251 glioblastoma cell chemoresistance by inhibiting FOXO1-mediated PID1 expression through stabilizing lncRNA DANCR. *Frontiers in Cell and Developmental Biology*, 9: 659228. <https://doi.org/10.3389/fcell.2021.659228>
- He Z, Ruan X, Liu X, Zheng J, Liu Y et al. (2019). FUS/circ\_002136/miR-138-5p/SOX13 feedback loop regulates angiogenesis in Glioma. *Journal of Experimental & Clinical Cancer Research*, 38 (1): 65. <https://doi.org/10.1186/s13046-019-1065-7>
- Helman E, Lawrence MS, Stewart C, Sougnez C, Getz G et al. (2014). Somatic retrotransposition in human cancer revealed by whole-genome and exome sequencing. *Genome Research*, 24 (7): 1053–1063. <https://doi.org/10.1101/gr.163659.113>
- Hillier LW, Graves TA, Fulton RS, Fulton LA, Pepin KH et al. (2005). Generation and annotation of the DNA sequences of human chromosomes 2 and 4. *Nature*, 434 (7034): 724–731. <https://doi.org/10.1038/nature03466>
- Hubley R, Finn RD, Clements J, Eddy SR, Jones TA et al. (2016). The Dfam database of repetitive DNA families. *Nucleic Acids Research*, 44 (D1): D81–9. <https://doi.org/10.1093/nar/gkv1272>

- Hwang T, Kim S, Chowdhury T, Yu HJ, Kim KM et al. (2022). Genome-wide perturbations of Alu expression and Alu-associated post-transcriptional regulations distinguish oligodendroglioma from other gliomas. *Communications Biology*, 5 (1): 62. <https://doi.org/10.1038/s42003-022-03011-w>
- International Human Genome Sequencing Consortium (2001). Initial sequencing and analysis of the human genome. *Nature*, 409 (6822): 860–921. <https://doi.org/10.1038/35057062>
- Jagannathan M, Yamashita YM (2017). Function of junk: pericentromeric satellite DNA in chromosome maintenance. *Cold Spring Harbor Symposia on Quantitative Biology*, 82: 319–327. <https://doi.org/10.1101/sqb.2017.82.034504>
- Juratli TA, Stasik S, Zolal A, Schuster C, Richter S et al. (2018). *TERT* promoter mutation detection in cell-free tumor-derived DNA in patients with *IDH* wild-type glioblastomas: a pilot prospective study. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*, 24 (21): 5282–5291. <https://doi.org/10.1158/1078-0432.CCR-17-3717>
- Karakülah G, Yandım C (2020). Signature changes in the expressions of protein-coding genes, lncRNAs, and repeat elements in early and late cellular senescence. *Turkish Journal of Biology* 44 (6): 356–370. <https://doi.org/10.3906/biy-2005-21>
- Karakülah G, Yandım C (2021). Identification of differentially expressed genomic repeats in primary hepatocellular carcinoma and their potential links to biological processes and survival. *Turkish Journal of Biology* 45 (5): 599–612. <https://doi.org/10.3906/biy-2104-13>
- Khan AB, Lee S, Harmanci AS, Patel R, Latha K et al. (2023). CXCR4 expression is associated with proneural-to-mesenchymal transition in glioblastoma. *International Journal of Cancer*, 152 (4): 713–724. <https://doi.org/10.1002/ijc.34329>
- Kishikawa T, Otsuka M, Yoshikawa T, Ohno M, Ijichi H et al. (2016). Satellite RNAs promote pancreatic oncogenic processes via the dysfunction of YBX1. *Nature Communications*, 7, 13006. <https://doi.org/10.1038/ncomms13006>
- Kong, Y, Rose CM, Cass AA, Williams AG, Darwish M et al. (2019). Transposable element expression in tumors is associated with immune infiltration and increased antigenicity. *Nature Communications*, 10 (1): 5228. <https://doi.org/10.1038/s41467-019-13035-2>
- Koso H, Takeda H, Yew CCK, Ward JM, Nariai N et al. (2012). Transposon mutagenesis identifies genes that transform neural stem cells into glioma-initiating cells. *Proceedings of the National Academy of Sciences*, 109 (44): E2998–E3007. <https://doi.org/10.1073/pnas.1215899109>
- Lan YL, Wang X, Xing JS, Lou JC, Ma XC et al. (2017). The potential roles of dopamine in malignant glioma. *Acta Neurologica Belgica*, 117 (3): 613–621. <https://doi.org/10.1007/s13760-016-0730-2>
- Langfelder P, Horvath S (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*, 9: 559. <https://doi.org/10.1186/1471-2105-9-559>
- Langfelder P, Zhang B, Horvath S (2008). Defining clusters from a hierarchical cluster tree: the Dynamic Tree Cut package for R. *Bioinformatics* 24 (5): 719–720. <https://doi.org/10.1093/bioinformatics/btm563>
- Lee E, Iskow R, Yang L, Gokcumen O, Haseley P et al. (2012). Landscape of somatic retrotransposition in human cancers. *Science*, 337 (6097): 967–971. <https://doi.org/10.1126/science.1222077>
- Leinonen R, Sugawara H, Shumway M on behalf of the International Nucleotide Sequence Database Collaboration (2011). The sequence read archive. *Nucleic Acids Research*, 39 (suppl\_1) D19–D21. <https://doi.org/10.1093/nar/gkq1019>
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, 25 (16): 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>
- Li J, Zhu S, Kozono D, Ng K, Futalan D et al. (2014). Genome-wide shRNA screen revealed integrated mitogenic signaling between dopamine receptor D2 (DRD2) and epidermal growth factor receptor (EGFR) in glioblastoma. *Oncotarget*, 5 (4): 882–893. <https://doi.org/10.18632/oncotarget.1801>
- Li T, Yang Z, Li H, Zhu J, Wang Y et al. (2022). Phospholipase C $\gamma$ 1 (PLCG1) overexpression is associated with tumor growth and poor survival in *IDH* wild-type lower-grade gliomas in adult patients. *Laboratory Investigation*, 102 (2): 143–153. <https://doi.org/10.1038/s41374-021-00682-7>
- Li X, Meng Y (2021). SUMOylation regulator-related molecules can be used as prognostic biomarkers for glioblastoma. *Frontiers in Cell and Developmental Biology*, 9: 658856. <https://doi.org/10.3389/fcell.2021.658856>
- Liao Y, Smyth GK, Shi W (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30 (7): 923–930. <https://doi.org/10.1093/bioinformatics/btt656>
- Lin J, Lai X, Liu X, Yan H, Wu C. (2022). Pyroptosis in glioblastoma: a crucial regulator of the tumour immune microenvironment and a predictor of prognosis. *Journal of Cellular and Molecular Medicine*, 26 (5): 1579–1593. <https://doi.org/10.1111/jcmm.17200>
- Lin WW, Ou GY, Zhao WJ (2021). Mutational profiling of low-grade gliomas identifies prognosis and immunotherapy-related biomarkers and tumour immune microenvironment characteristics. *Journal of Cellular and Molecular Medicine*, 25 (21): 10111–10125. <https://doi.org/10.1111/jcmm.16947>
- Louis DN, Perry A, Reifenberger G, von Deimling A, Figurearella-Branger D et al. (2016). The 2016 World Health Organization classification of tumors of the central nervous system: a summary. *Acta Neuropathologica*, 131 (6): 803–820. <https://doi.org/10.1007/s00401-016-1545-1>
- Lu J, Li H, Chen Z, Fan L, Feng S et al. (2019). Identification of 3 subpopulations of tumor-infiltrating immune cells for malignant transformation of low-grade glioma. *Cancer Cell International*, 19: 265. <https://doi.org/10.1186/s12935-019-0972-1>

- Mack SC, Singh I, Wang X, Hirsch R, Wu Q et al. (2019). Chromatin landscapes reveal developmentally encoded transcriptional states that define human glioblastoma. *The Journal of Experimental Medicine*, 216 (5): 1071–1090. <https://doi.org/10.1084/jem.20190196>
- Mackay A, Burford A, Carvalho D, Izquierdo E, Fazal-Salom J et al. (2017). Integrated molecular meta-analysis of 1,000 pediatric high-grade and diffuse intrinsic pontine glioma. *Cancer Cell*, 32 (4): 520–537.e5. <https://doi.org/10.1016/j.ccell.2017.08.017>
- Murat A, Migliaiavacca E, Gorlia T, Lambiv WL, Shay T et al. (2008). Stem cell-related 'self-renewal' signature and high epidermal growth factor receptor expression associated with resistance to concomitant chemoradiotherapy in glioblastoma. *Journal of Clinical Oncology*, 26 (18): 3015–3024. <https://doi.org/10.1200/JCO.2007.15.7164>
- Nishibuchi G, Déjardin J (2017). The molecular basis of the organization of repetitive DNA-containing constitutive heterochromatin in mammals. *Chromosome Research*, 25 (1): 77–87. <https://doi.org/10.1007/s10577-016-9547-3>
- Noushmehr H, Weisenberger DJ, Diefes K, Phillips HS, Pujara K et al. (2010). Identification of a CpG island methylator phenotype that defines a distinct subgroup of glioma. *Cancer Cell*, 17 (5): 510–522. <https://doi.org/10.1016/j.ccr.2010.03.017>
- Ohka F, Natsume A, Motomura K, Kishida Y, Kondo Y et al. (2011). The global DNA methylation surrogate LINE-1 methylation is correlated with *MGMT* promoter methylation and is a better prognostic factor for glioma. *PLoS One*, 6 (8): e23332. <https://doi.org/10.1371/journal.pone.0023332>
- Ostrom QT, Gittleman H, Liao P, Vecchione-Koval T, Wolinsky Y et al. (2017). CBTRUS Statistical Report: Primary brain and other central nervous system tumors diagnosed in the United States in 2010–2014. *Neuro-Oncology*, 19 (suppl\_5): v1–v88. <https://doi.org/10.1093/neuonc/nox158>
- Özgür E, Mayer Z, Keskin M, Yörüker EE, Holdenrieder S et al. (2021). Satellite 2 repeat DNA in blood plasma as a candidate biomarker for the detection of cancer. *Clinica Chimica Acta*, 514: 74–79. <https://doi.org/10.1016/j.cca.2020.12.008>
- Parsons DW, Jones S, Zhang X, Lin JCH, Leary RJ et al. (2008). An integrated genomic analysis of human glioblastoma multiforme. *Science*, 321 (5897): 1807–1812. <https://doi.org/10.1126/science.1164382>
- Phillips HS, Kharbanda S, Chen R, Forrest WF, Soriano RH et al. (2006). Molecular subclasses of high-grade glioma predict prognosis, delineate a pattern of disease progression, and resemble stages in neurogenesis. *Cancer Cell*, 9 (3): 157–173. <https://doi.org/10.1016/j.ccr.2006.02.019>
- Porokhovnik LN, Veiko NN, Ershova ES, Kostyuk SV (2021). The role of human satellite III (1q12) copy number variation in the adaptive response during aging, stress, and pathology: a pendulum model. *Genes* 12 (10). <https://doi.org/10.3390/genes12101524>
- Probst AV, Okamoto I, Casanova M, El Marjou F, Le Baccon P et al. (2010). A strand-specific burst in transcription of pericentric satellites is required for chromocenter formation and early mouse development. *Developmental Cell*, 19 (4): 625–638. <https://doi.org/10.1016/j.devcel.2010.09.002>
- Robinson MD, McCarthy DJ, Smyth GK (2010). edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, 26 (1): 139–140. <https://doi.org/10.1093/bioinformatics/btp616>
- Rošić S, Köhler F, Erhardt S (2014). Repetitive centromeric satellite RNA is essential for kinetochore formation and cell division. *The Journal of Cell Biology*, 207 (3): 335–349. <https://doi.org/10.1083/jcb.201404097>
- Royston P, Choodari-Oskooei B, Parmar MKB, Rogers JK (2019). Combined test versus logrank/Cox test in 50 randomised trials. *Trials*, 20 (1): 172. <https://doi.org/10.1186/s13063-019-3251-5>
- Sahu D, Chang YL, Lin YC, Lin CC (2021). Characterization of the survival influential genes in carcinogenesis. *International Journal of Molecular Sciences*, 22 (9). <https://doi.org/10.3390/ijms22094384>
- Sakthikumar S, Sakthikumar S, Roy A, Haseeb L, Pettersson ME et al. (2020). Whole-genome sequencing of glioblastoma reveals enrichment of non-coding constraint mutations in known and novel genes. *Genome Biology*, 21 (1): 127. <https://doi.org/10.1186/s13059-020-02035-x>
- Salhia B, Tran NL, Chan A, Wolf A, Nakada M et al. (2008). The guanine nucleotide exchange factors trio, Ect2, and Vav3 mediate the invasive behavior of glioblastoma. *The American Journal of Pathology*, 173 (6): 1828–38. <https://doi.org/10.2353/ajpath.2008.080043>
- Sallam M, Mysara M, Baatout S, Guns P-J, Ramadan R et al. (2022). Meta-analysis of RNA-seq datasets identifies novel players in glioblastoma. *Cancers*, 14 (23): 5788. <https://doi.org/10.3390/cancers14235788>
- Shah AH, Govindarajan V, Doucet-O'Hare TT, Rivas S, Ampie L et al. (2022). Differential expression of an endogenous retroviral element [HERV-K(HML-6)] is associated with reduced survival in glioblastoma patients. *Scientific Reports*, 12 (1): 6902. <https://doi.org/10.1038/s41598-022-10914-5>
- Solovoyov A, Vabret N, Arora KS, Snyder A, Funt SA et al. (2018). Global Cancer Transcriptome Quantifies Repeat Element Polarization between Immunotherapy Responsive and T Cell Suppressive Classes. *Cell Reports*, 23 (2): 512–521. <https://doi.org/10.1016/j.celrep.2018.03.042>
- Song X, Wan X, Huang T, Zeng C, Sastry N et al. (2019). SRSF3-regulated RNA alternative splicing promotes glioblastoma tumorigenicity by affecting multiple cellular processes. *Cancer Research*, 79 (20): 5288–5301. <https://doi.org/10.1158/0008-5472.CAN-19-1504>
- Stupp R, Hegi ME, Mason WP, van den Bent MJ, Taphoorn MJB et al. (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. *The Lancet Oncology*, 10 (5): 459–466. [https://doi.org/10.1016/S1470-2045\(09\)70025-7](https://doi.org/10.1016/S1470-2045(09)70025-7)



- Sturm D, Bender S, Jones DTW, Lichter P, Grill J et al. (2014). Paediatric and adult glioblastoma: multiform (epi)genomic culprits emerge. *Nature Reviews Cancer*, 14 (2): 92–107. <https://doi.org/10.1038/nrc3655>
- Sun R, Pan Y, Mu L, Ma Y, Shen H et al. (2021). Development of a 3 RNA binding protein signature for predicting prognosis and treatment response for glioblastoma multiforme. *Frontiers in Genetics*, 12: 768930. <https://doi.org/10.3389/fgene.2021.768930>
- Teng C, Zhu Y, Li Y, Dai L, Pan Z et al. (2022). Recurrence- and malignant progression-associated biomarkers in low-grade gliomas and their roles in immunotherapy. *Frontiers in Immunology* (Vol. 13). <https://doi.org/10.3389/fimmu.2022.899710>
- Tilak M, Alural B, Wismer SE, Brasher MI, New LA et al. (2021). Adaptor protein ShcD/*SHC4* interacts with Tie2 receptor to synergistically promote glioma cell invasion. *Molecular Cancer Research*, 19 (5): 757–770. <https://doi.org/10.1158/1541-7786.MCR-20-0188>
- Ting DT, Lipson D, Paul S, Brannigan BW, Akhavanfard S et al. (2011). Aberrant overexpression of satellite repeats in pancreatic and other epithelial cancers. *Science*, 331 (6017): 593–596. <https://doi.org/10.1126/science.1200801>
- Ulitsky I, Shkumatava A, Jan CH, Sive H, Bartel DP (2011). Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. *Cell*, 147 (7): 1537–1550. <https://doi.org/10.1016/j.cell.2011.11.055>
- Vos LJ, Famulski JK, Chan GKT (2006). How to build a centromere: from centromeric and pericentromeric chromatin to kinetochore assembly. *Biochemistry and Cell Biology*, 84 (4): 619–639. <https://doi.org/10.1139/o06-078>
- Wang H, Wang X, Xu L, Zhang J, Cao H (2020). Prognostic significance of age related genes in patients with lower grade glioma. *Journal of Cancer*, 11 (13): 3986–3999. <https://doi.org/10.7150/jca.41123>
- Wang L, Shang Z, Zhou Y, Hu X, Chen Y et al. (2018). Autophagy mediates glucose starvation-induced glioblastoma cell quiescence and chemoresistance through coordinating cell metabolism, cell cycle, and survival. *Cell Death & Disease*, 9 (2): 213. <https://doi.org/10.1038/s41419-017-0242-x>
- Wang Q, Hu B, Hu X, Kim H, Squatrito M et al. (2017). Tumor evolution of glioma-intrinsic gene expression subtypes associates with immunological changes in the microenvironment. *Cancer Cell*, 32 (1): 42–56.e6. <https://doi.org/10.1016/j.ccell.2017.06.003>
- Wang X, Zhang L, Song Y, Jiang Y, Zhang D et al. (2021). MCM8 is regulated by EGFR signaling and promotes the growth of glioma stem cells through its interaction with DNA-replication-initiating factors. *Oncogene*, 40 (27): 4615–4624. <https://doi.org/10.1038/s41388-021-01888-1>
- Wang J, Cazzato E, Ladewig E, Frattini V, Rosenbloom DIS et al. (2016). Clonal evolution of glioblastoma under therapy. *Nature Genetics*, 48 (7): 768–776. <https://doi.org/10.1038/ng.3590>
- Wang J, Huang TYT, Hou Y, Bartom E, Lu X et al. (2021). Epigenomic landscape and 3D genome structure in pediatric high-grade glioma. *Science Advances*, 7 (23). <https://doi.org/10.1126/sciadv.abg4126>
- Wolf D, Hofbrucker-MacKenzie SA, Izadi M, Seemann E, Steiniger F et al. (2019). Ankyrin repeat-containing N-Ank proteins shape cellular membranes. *Nature Cell Biology*, 21 (10): 1191–1205. <https://doi.org/10.1038/s41556-019-0381-7>
- Xiao F, Shen J, Zhou L, Fang Z, Weng Y et al. (2022). ZNF395 facilitates macrophage polarization and impacts the prognosis of glioma. *American Journal of Cancer Research*, 12 (9): 4312–4325. <https://pubmed.ncbi.nlm.nih.gov/36225626/>
- Xu Y, Geng R, Yuan F, Sun Q, Liu B et al. (2019). Identification of differentially expressed key genes between glioblastoma and low-grade glioma by bioinformatics analysis. *PeerJ*, 7: e6560.
- Yandım C, Karakülah G (2019a). Expression dynamics of repetitive DNA in early human embryonic development. *BMC Genomics*, 20 (1): 439. <https://doi.org/10.1186/s12864-019-5803-1>
- Yandım C, Karakülah G (2019b). Dysregulated expression of repetitive DNA in ER+/HER2- breast cancer. *Cancer Genetics*, 239: 36–45. <https://doi.org/10.1016/j.cancergen.2019.09.002>
- Yandım C, Karakülah G (2022). Repeat expression is linked to patient survival and exhibits single nucleotide variation in pancreatic cancer revealing LTR70:r.879A>G. *Gene*, 822: 146344. <https://doi.org/10.1016/j.gene.2022.146344>
- Yang Q, Jiang N, Zou H, Fan X, Liu T et al. (2022). Alterations in 3D chromatin organization contribute to tumorigenesis of EGFR-amplified glioblastoma. *Computational and Structural Biotechnology Journal*, 20: 1967–1978. <https://doi.org/10.1016/j.csbj.2022.04.007>
- Yu G, Wang LG, Han Y, He QY (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OmicS: A Journal of Integrative Biology*, 16 (5): 284–287. <https://doi.org/10.1089/omi.2011.0118>
- Yuan Z, Yang Y, Zhang N, Soto C, Jiang X et al. (2021). Human endogenous retroviruses in glioblastoma multiforme. *Microorganisms* 9 (4). <https://doi.org/10.3390/microorganisms9040764>
- Zhang B, Horvath S (2005). A general framework for weighted gene co-expression network analysis. *Statistical Applications in Genetics and Molecular Biology*, 4, Article17. <https://doi.org/10.2202/1544-6115.1128>
- Zhang D, Zhao J, Han C, Liu X, Liu J et al. (2020). Identification of hub genes related to prognosis in glioma. *Bioscience Reports*, 40 (5). <https://doi.org/10.1042/BSR20193377>
- Zhang L, Wang Y, Rashid MH, Liu M, Angara K et al. (2017). Malignant pericytes expressing GT198 give rise to tumor cells through angiogenesis. *Oncotarget*, 8 (31): 51591–51607. <https://doi.org/10.18632/oncotarget.18196>
- Zhu Q, Pao GM, Huynh AM, Suh H, Tonnu N et al. (2011). BRCA1 tumour suppression occurs via heterochromatin-mediated silencing. *Nature*, 477 (7363): 179–184. <https://doi.org/10.1038/nature10371>

**Supplementary files**

Supplementary files can be downloaded from the following link:

<https://aperta.ulakbim.gov.tr/record/273725>

**Supplementary File 1:**

**Table S1.** SRR numbers of data used for the study.

**Supplementary File 2:**

**Figure S1.** Distribution of read percentages of LGG and HGG samples. **A.** Violin plot representing the distribution of read percentages of protein-coding genes and repeat elements in LGG and HGG samples. **B.** Pie chart representing the percentages of repeat classes in LGG samples. **C.** Pie chart representing the percentages of repeat classes in HGG samples.

**Supplementary File 3:**

**Table S2.** Differentially expressed genes in HGG samples in comparison to LGG samples.

**Table S3.** Differentially expressed repeats in HGG samples in comparison to LGG samples.

**Supplementary File 4:**

**Table S4.** Gene ontology analysis result list on biological processes of differentially expressed genes in HGG relative to LGG.

**Supplementary File 5:**

**Table S5.** Protective genes ( $HR < 0.5$ ) obtained in LGG progress-free survival analysis.

**Table S6.** Hazardous genes ( $HR > 1.5$ ) obtained in LGG progress-free survival analysis.

**Table S7.** Protective repeats ( $HR < 0.5$ ) obtained in LGG progress-free survival analysis.

**Table S8.** Hazardous repeats ( $HR > 1.5$ ) obtained in LGG progress-free survival analysis.

**Supplementary File 6:**

**Table S9.** Protective genes ( $HR < 0.5$ ) obtained in HGG event-free survival analysis.

**Table S10.** Hazardous genes ( $HR > 1.5$ ) obtained in HGG event-free survival analysis.

**Table S11.** Protective repeats ( $HR < 0.5$ ) obtained in HGG event-free survival analysis.

**Table S12.** Hazardous repeats ( $HR > 1.5$ ) obtained in HGG event-free survival analysis.

**Supplementary File 7:**

**Table S13.** Gene ontology analysis result list on biological processes of hazardous genes identified in LGG progress-free analysis.

**Supplementary File 8:**

**Table S14.** Gene ontology analysis result list on biological processes of hazardous genes identified in HGG event-free analysis.

**Supplementary File 9:**

**Table S15.** Genes listed in WGCNA modules.

**Supplementary 10:** Gene ontology analysis of genes identified in WGCNA modules pertaining to biological processes.

**Supplementary File 11:**

**Table S16.** Correlations of the expressions of protein-coding genes with HSAT5 satellite repeat expression.