Research article

Prognostic value of anoikis-related genes revealed using multi-omics analysis and machine learning based on lower-grade glioma features and tumor immune microenvironment

Gu Linazi [a,1], Aierpati Maimaiti [b,1], Zulihuma Abulaiti [c,1], Hui Shi [a], Zexin Zhou [a], Mizhati Yimiti Aisa [a], Yali Kang [a], Ayguzaili Abulimiti [a], Xierzhati Dilimulati [a], Tiecheng Zhang [a], Patiman Wusiman [a], Zengliang Wang [b,**], Aimitaji Abulaiti [b,*]

[a] Department of Rehabilitation Medicine, The First Affiliated Hospital of Xinjiang Medical University, Urumqi, Xinjiang, 830054, China
[b] Department of Neurosurgery, Neurosurgery Centre, The First Affiliated Hospital of Xinjiang Medical University, Urumqi, Xinjiang, 830054, China
[c] Department of Obstetrics and Gynecology, The First Affiliated Hospital of Xinjiang Medical University, Urumqi, Xinjiang, 830054, China

A R T I C L E   I N F O

A B S T R A C T

Background: The investigation explores the involvement of anoikis-related genes (ARGs) in lower-grade glioma (LGG), seeking to provide fresh insights into the disease's underlying mechanisms and to identify potential targets for therapy.

Methods: We applied unsupervised clustering techniques to categorize LGG patients into distinct molecular subtypes based on ARGs with prognostic significance. Additionally, various machine learning algorithms were employed to pinpoint genes most strongly correlated with patient outcomes, which were then used to develop and assess risk profiles.

Results: Our analysis identified two distinct molecular subtypes of LGG, each with significantly different prognoses. Patients in Cluster 2 had a median survival of 2.036 years, markedly shorter than the 7.994 years observed in Cluster 1 (P < 0.001). We also constructed a six-gene ARG signature that efficiently classified patients into two risk categories, showing median survival durations of 4.084 years for the high-risk group and 10.304 years for the low-risk group (P < 0.001). Significantly, the immune profiles, tumor mutation characteristics, and drug sensitivity varied greatly among these risk groups. The high-risk group was characterized by a "cold" tumor microenvironment (TME), a lower IDH1 mutation rate (61.7 % vs. 91.4 %), a higher TP53 mutation rate (53.7 % vs. 38.9 %), and greater sensitivity to targeted therapies such as QS11 and PF-562271. Furthermore, our nomogram, integrating risk scores with clinicopathological features, demonstrated strong predictive accuracy for clinical outcomes in LGG patients, with an AUC of 0.903 for the first year. The robustness of this prognostic model was further validated through internal cross-validation and across three external cohorts.

Conclusions: The evidence from our research suggests that ARGs could potentially serve as reliable indicators for evaluating immunotherapy effectiveness and forecasting clinical results in patients with LGG.

---

* Corresponding author.

** Corresponding author.

E-mail addresses: wzl3ng@126.com (Z. Wang), aimi7973@163.com (A. Abulaiti).

[1] These authors have contributed equally to this work.

1. Introduction

Gliomas are the most common form of primary malignant brain tumors, comprising about 80 % of all cases. These tumors pose a significant public health challenge due to their high prevalence and associated morbidity and mortality. The age-adjusted annual mortality rate for gliomas stands at 4.19 per 100,000 people, reflecting the aggressive nature of these tumors and the limited effectiveness of current treatments in extending patient survival beyond a few years [1]. Within the broader category of gliomas, lower-grade gliomas (LGGs) represent a distinct subset. LGGs, while generally less aggressive than their higher-grade counterparts, still account for approximately 15 % of all gliomas and 5 % of primary brain tumors, making them a significant focus of neuro-oncological research[2,3], 2,3.

Characterized by a slower growth rate and a more protracted clinical course, LGGs are often associated with relatively better outcomes compared to higher-grade gliomas. Many patients with LGGs can survive for more than a decade following diagnosis, which contrasts sharply with the survival rates of patients with high-grade gliomas, who often face a median overall survival of only 12–15 months[4,5], 4,5. However, despite these more favorable outcomes, LGGs are not benign. They frequently recur after initial treatment and possess the potential to progress to more aggressive, higher-grade gliomas. This malignant transformation is particularly alarming, as it significantly decreases the median overall survival (OS) of patients to only 2.4 years after the onset of malignant progression, highlighting the critical need for continuous monitoring and intervention, even in cases that are initially diagnosed as lower-grade [6].

Current therapeutic strategies for managing LGG primarily involve maximal safe surgical resection to remove as much of the tumor as possible without compromising neurological function. This is usually followed by radiotherapy, chemotherapy, or a combination of both, tailored to the specific characteristics of the tumor and the individual patient. Treatment plans are highly individualized, taking into account a variety of factors including the extent of surgical resection, the location and size of the tumor, histopathological findings, molecular features such as IDH mutation status and 1p/19q codeletion, and patient-specific factors such as age and overall health status [7,8]. Despite these tailored approaches, the invasive and infiltrative nature of LGGs often limits the effectiveness of existing treatments. Tumor cells can invade surrounding brain tissue in a manner that makes complete surgical removal impossible, leading to recurrence and progression. Additionally, the molecular and genetic heterogeneity of LGGs further complicates treatment, as tumors with similar histological appearances can have vastly different biological behaviors and responses to therapy[9,10], 9,10.

Given these challenges, a deeper understanding of the mechanisms driving LGG infiltration, invasion, and progression is essential. In particular, there is a pressing need to identify reliable biomarkers that can accurately predict clinical outcomes and guide treatment decisions for LGG patients. Biomarkers that can differentiate between tumors likely to remain indolent and those prone to malignant transformation could significantly improve patient management, allowing for more aggressive treatment in high-risk cases while sparing low-risk patients from the side effects of intensive therapy.

Glioma infiltration is closely associated with abnormal alterations in the adhesion junctions of brain glial cells, which play a crucial role in maintaining the structural integrity and proper function of the brain's neural network [11]. This association points to a potential connection between gliomas and anoikis, a specialized form of programmed cell death that occurs when cells lose their adhesion to the extracellular matrix (ECM) or when cell adhesion is otherwise inadequate [12,13]. Anoikis plays a vital role in maintaining tissue homeostasis by preventing detached cells from surviving and proliferating in inappropriate locations, which could otherwise lead to abnormal growth and metastasis. The ECM, along with surrounding support cells, creates a microenvironment that is essential for normal cellular function, providing the necessary biochemical and mechanical signals for cell growth, differentiation, and the maintenance of physiological functions [14].

Anoikis is induced when ECM-mediated cell adhesion is disrupted, either during cell detachment or due to changes in the microenvironment. This disruption triggers signaling pathways that lead to the elimination of isolated cells, thereby preventing potential malignancies from developing in regions where they should not exist [15]. The regulation of anoikis primarily occurs through two major pathways: the intrinsic pathway, which is triggered by mitochondrial responses to cellular stress, and the extrinsic pathway, which is mediated by external factors such as tumor necrosis factor (TNF) and Fas ligands[16–18], 16–18, 16–18. Both pathways converge on the activation of caspases, a family of proteases that execute the cell death program by degrading key cellular components.

Although previous research has consistently underscored the critical importance of anoikis in cancer progression, especially in its role of inhibiting metastasis, the exploration of anoikis-related markers as potential prognostic tools in LGG has been limited. This gap in research leaves a significant opportunity to better understand and utilize these markers in the context of LGG prognosis[19–21], 19–21, 19–21. The ability of tumor cells to resist anoikis is a hallmark of metastatic potential, as it allows cancer cells to survive during detachment from the primary tumor and establish new growths in distant organs. Understanding the mechanisms by which LGG cells evade anoikis could provide valuable insights into the factors driving tumor progression and recurrence.

Therefore, this study aims to delve deeper into the role of anoikis in LGG by investigating the prognostic value of anoikis-related genes (ARGs). Given the critical role of anoikis in cancer progression and the limited research on its mechanisms in LGG, our study seeks to bridge this gap by exploring the expression profiles and prognostic significance of ARGs in LGG. We aim to evaluate the potential of ARGs as prognostic biomarkers that could predict patient outcomes and inform treatment strategies. By utilizing advanced bioinformatics tools and machine learning techniques, we intend to uncover the molecular mechanisms that drive LGG pathogenesis and identify novel therapeutic targets that could be exploited to develop more effective, targeted therapies for this challenging disease.

This research not only aims to enhance our understanding of the biological processes underlying LGG but also to contribute to the broader field of oncology by identifying potential biomarkers that could be used to stratify patients based on their risk of progression. By integrating multi-omics data and leveraging the power of machine learning, we hope to develop robust predictive models that can be used in clinical settings to guide treatment decisions and improve patient outcomes. Ultimately, our goal is to advance the

development of targeted therapeutic strategies that can effectively manage LGG and reduce the burden of this disease on patients and healthcare systems.

2. Methods

2.1. Data collection

RNA-seq gene expression data for both LGG and normal tissues were obtained from the UCSC Xena platform [22]. Corresponding clinical data for LGG patients were retrieved from the TCGA database. Samples lacking survival data, tumor grade information, or that were not primary LGG were excluded, leaving 105 normal cerebral cortex samples and 497 LGG samples for the analysis. For external validation, additional LGG samples from the CGGA325, CGGA693, and GSE16011 RNA-seq cohorts were used.

Mutation differences among risk groups were analyzed using the maftools R package, with Tumor Mutation Burden (TMB) calculated for each patient within the TCGA cohort. For protein expression analysis, data were sourced from the Human Protein Atlas (HPA) database (https://www.proteinatlas.org), which offers immunohistochemical (IHC) data. Additionally, IHC images of anoikis-related gene (ARG) model genes in LGG samples were obtained from the HPA.

Lastly, we analyzed tumor microenvironment (TME) heterogeneity using data from the TISCH database.

2.2. Identification of genes related to anoikis

We embarked on a thorough investigation to identify genes linked to anoikis by utilizing the GeneCards database (https://www.genecards.org/). GeneCards serves as an extensive, integrative database that provides detailed information on human genes, including their functions, related pathways, and involvement in various biological processes. By using "anoikis" as the keyword in our search, we were able to gather a list of genes that are potentially involved in this form of cell death, which occurs when cells lose adhesion to the extracellular matrix. The relevance of each gene was determined based on the GeneCards Inferred Function Score (GIFtS) [23]. For our study, we focused on genes with a GIFtS greater than 30, leading to the identification of 739 anoikis-associated genes for further investigation.

2.3. Identification of DEGs and genes associated with LGG

To identify genes that are differentially expressed between normal tissues and LGG tumors, we conducted a thorough differential expression analysis utilizing the R package limma. This package is a widely recognized and reliable tool for analyzing gene expression data, enabling us to accurately pinpoint genes with significant expression differences between the two tissue types. The analysis involved evaluating genes based on their fold change (FC), where the mean expression levels of each gene were used to calculate t-values. The resulting t-values were then converted into p-values, which were adjusted for multiple comparisons using the false discovery rate (FDR) method. This adjustment is crucial for reducing the likelihood of false positives, a common risk in analyses involving a large number of tests. Genes were classified as differentially expressed if they exhibited a |logFC| greater than 1 and an adjusted p-value of less than 0.05. The significant DEGs identified through this process were visualized using volcano plots, which effectively highlight the most pronounced differences in gene expression between the two groups.

To further explore the biological relevance of these DEGs in the context of LGG, we employed WGCNA [24], a powerful method for uncovering gene expression patterns across the entire dataset. This approach allowed us to identify clusters of genes, or "modules," that were highly co-expressed, meaning they had similar expression patterns across the samples. By applying the WGCNA package in R, we were able to detect these modules and assess their association with LGG. The strength of each module's association with LGG was quantified using correlation coefficients and corresponding p-values, providing a measure of the statistical significance of these relationships.

For more detailed analysis, we concentrated on the gene modules that demonstrated the strongest correlations with LGG. Within these key modules, we compared the identified genes with previously recognized ARGs. To assess the overlap between ARGs and the most relevant LGG-associated modules, we utilized an online Venn diagram tool (http://www.ehbio.com/test/venn/#/) [25]. This tool allowed us to visually and quantitatively determine the extent of shared genes between these categories. The overlapping genes were visually represented in Venn diagrams, which offered valuable insights into the shared genetic pathways between anoikis and LGG, potentially revealing critical targets for further investigation.

2.4. Comprehensive multivariate analysis of ARGs

We carried out an extensive analysis focusing on the expression levels of 77 ARGs to better understand their role in LGG. This analysis was visualized using the R package pheatmap, which allowed us to generate heatmaps that clearly display the expression patterns of these genes across different samples. By utilizing pheatmap, we could easily identify clusters of ARGs with similar expression profiles, offering insights into their potential co-regulation and involvement in common biological processes. To further investigate the relationships among these anoikis-related genes (ARGs), we constructed a protein-protein interaction (PPI) network using the R package STRINGdb. This approach enabled us to visualize and analyze the interactions between these proteins, providing deeper insights into their functional connections within the context of LGG. This network provided a detailed view of the potential physical and functional interactions between the proteins encoded by these ARGs, highlighting key nodes and pathways that may be

critical in the context of anoikis and its impact on LGG progression. Correlations between gene pairs were further analyzed and visualized with the R package corrplot.

To assess somatic mutations in these ARGs across patients, we utilized the R package maftools [26]. Additionally, we analyzed the copy number variation (CNV) status of each gene to determine whether there were gains or losses in specific chromosomal regions. This chromosomal information was then visualized using Circos diagrams, offering a comprehensive overview of the genetic alterations and interactions of the ARGs within the context of LGG.

2.5. Unveiling anoikis patterns through unsupervised clustering techniques

To identify distinct anoikis patterns, we began by conducting an initial screening of prognostic genes using univariate Cox regression analysis, which helped reduce the dimensionality for subsequent clustering. Only genes that demonstrated a strong association with anoikis and a significant relationship with patient prognosis were retained as key parameters for clustering.

We then performed unsupervised clustering using the ConsensusClusterPlus package [27]. The partitioning around medoid (PAM) algorithm was applied, utilizing Pearson correlation distances to form clusters. To ensure the robustness and stability of the classifications, the clustering process was repeated 1000 times, thereby solidifying the consistency of the cluster assignments.

2.6. Characterization of molecular subtypes

To correlate clinical data with molecular subtypes of LGG, we employed the heatmap package for effective visualization. To further explore the immune landscape associated with these subtypes, we analyzed the immune cell composition within LGG samples using the CIBERSORT method. To enhance the depth and breadth of our analysis, we integrated four additional algorithms to meticulously assess immune cell infiltration within the tumor samples. Each of these algorithms provides a unique perspective on the composition and abundance of various immune cell types within the tumor microenvironment. To further complement this analysis, we applied the ESTIMATE algorithm, which offered additional insights into the tumor microenvironment by evaluating both stromal and immune components. This combined approach enabled a more comprehensive understanding of the tumor's biological landscape. By combining these diverse computational approaches, we were able to perform a thorough evaluation of immune cell infiltration and delineate the immune profiles across the different molecular subtypes of lower-grade glioma (LGG). This multi-algorithm approach provided us with a more nuanced and detailed picture of the immune landscape associated with LGG, potentially uncovering critical insights into the tumor's behavior and response to therapy.

2.7. Utilizing machine learning to identify risk signatures and determine optimal multi-gene combinations for disease prognosis

In this study, we developed a predictive model by categorizing the TCGA cohort of LGG samples into distinct groups. This stratification allowed us to identify patterns and associations that could enhance the model's accuracy in predicting clinical outcomes. Specifically, 70 % of the samples were randomly selected to form the training set, which served as the basis for model development. The remaining 30 % of the samples were designated as the Test1 cohort, which was used for internal validation to assess the initial performance of the predictive model. This step was crucial for determining the model's reliability and effectiveness before further testing. Additionally, the entire TCGA cohort was utilized as a secondary internal validation group, which we referred to as the Test2 cohort, providing further assessment of the model's robustness. For external validation, we expanded our analysis to include LGG samples from three additional datasets: CGGA-325, CGGA-693, and GSE-16011, allowing us to comprehensively validate the model's generalizability across different data sources.

To identify key ARGs that could serve as potential prognostic markers, we applied a combination of random forest (RF) and lasso regression techniques. The RF model was implemented using the randomForestSRC package in R, with parameters configured to generate 1000 trees and a random split of 1. Concurrently, lasso regression was performed using the glmnet package in R, incorporating a 10-fold cross-validation procedure to ensure robust and reliable gene selection. This combination of methods aimed to enhance the predictive power and stability of the model. The overlap between the genes identified by both methods was used to finalize the gene signature for further analysis.

In the next phase of our study, we constructed a Cox proportional hazards regression model, utilizing the expression data from the selected signature genes to evaluate their impact on patient survival. The model calculated a risk score for each patient using the following formula: ARG RiskScore = (Coef Gene1 * Exp Gene1) + (Coef Gene2 * Exp Gene2) + … + (Coef GeneN * Exp GeneN). Here, "Coef" represents the coefficient obtained from the Cox regression, reflecting the contribution of each gene to the overall risk, while "Exp" signifies the expression level of each gene in the patient. This approach allowed us to derive a comprehensive risk score that integrated the influence of multiple genes. Patients were then stratified into high-risk and low-risk groups, based on whether their risk score was above or below the median value. This stratification was crucial for identifying patients with differing levels of risk and potentially distinct clinical trajectories. To better illustrate the separation between the high-risk and low-risk groups, we plotted the distribution of risk scores. This visualization effectively highlighted the distinct division between the two groups, making the differences in risk levels more tangible and easier to interpret.

To assess the predictive power of the model, we generated receiver operating characteristic (ROC) curves using the Time ROC package in R. This approach allowed us to evaluate the model's accuracy over time, providing insights into its performance in distinguishing between different risk groups. These ROC curves were crucial for assessing the model's performance in predicting survival at specific time points—namely, 1 year, 3 years, and 5 years. The area under the ROC curve (AUC) provided a quantitative measure of

the model's accuracy, with higher AUC values indicating better predictive capability. In addition to ROC analysis, Kaplan-Meier survival curves were plotted to visualize the differences in OS between the high-risk and low-risk groups. These curves provided a clear comparison of survival probabilities over time, further highlighting the prognostic significance of the risk stratification. These survival curves were accompanied by log-rank tests to statistically compare survival outcomes between the two groups, allowing us to determine whether the risk stratification was significant in predicting patient prognosis. To further validate the model's consistency and reliability, we applied the same risk stratification procedure to all validation cohorts, from Test1 through Test5. This comprehensive validation process confirmed the model's robustness and applicability across different datasets, demonstrating that the model could reliably stratify patients into risk categories regardless of the cohort used.

2.8. GSEA and enrichment of risk subtype gene sets

To explore the biological differences between the tow groups, we undertook a detailed enrichment analysis of DEGs and conducted Gene Set Enrichment Analysis (GSEA) to illuminate the unique biological functions associated with these groups. Differential expression was evaluated using the R package **limma**, with a stringent threshold of $|logFC| > 1$ and a p-value <0.05 to identify significant DEGs between the two risk categories. This rigorous criterion ensured that only the most relevant DEGs were selected for further analysis. The resulting DEGs were then visualized with volcano plots, created using the ggplot2 package in R, offering a clear depiction of the genes that were significantly upregulated or downregulated in each group.

For functional enrichment, the identified DEGs were annotated with Gene Ontology (GO) terms using the org.Hs.eg.db package in R, which allowed us to categorize the DEGs according to their associated biological processes, cellular components, and molecular functions. Additionally, we retrieved the latest KEGG Pathway annotations via the KEGG REST API, facilitating a comprehensive enrichment analysis using the clusterProfiler package in R. This analysis helped in identifying specific pathways and functional categories that were overrepresented in the DEGs, shedding light on the molecular mechanisms underlying the differences between the high-risk and low-risk groups.

For the GSEA, we utilized version 4.0 of the GSEA software, which was downloaded from the official GSEA website [28]. We employed the c2.cp.kegg.v7.4.symbols dataset [29] to investigate the pathways and molecular mechanisms that were differentially enriched between the risk subtypes. This dataset was instrumental in highlighting the specific pathways that are more active in either the high- or low-risk groups, thus providing insights into the biological processes that might contribute to the progression of LGG. Overall, this systematic approach allowed us to uncover distinct biological processes that differentiate high-risk patients from those in the low-risk category, offering a deeper understanding of the molecular underpinnings of LGG progression.

2.9. Chemotherapeutic predictions and molecular docking

To predict the response of LGG samples to various antitumor drugs, we employed the OncoPredict R package [30], which offers a comprehensive analysis of drug responses and associated biomarkers. This tool allowed us to estimate the 50 % IC50 values for the LGG samples, providing insights into the sensitivity of these samples to different chemotherapeutic agents. The IC50 values were derived using data from the Cancer Treatment Response Portal (CTRP), enabling us to predict how effectively each drug could inhibit tumor cell growth at half-maximal concentration.

In our search for potential small molecule therapeutics, we utilized the Connectivity Map (CMap) platform, accessible at https://clue.io. By inputting DEGs associated with the risk subtypes into the CMap database, we aimed to identify small molecule drugs that could be particularly effective against LGG. The analysis produced a list of compounds with corresponding mean scores, where a negative mean score suggested potential therapeutic efficacy. Specifically, a negative score indicated that the drug might reverse disease-related abnormal biological states, potentially offering a therapeutic benefit. Statistical significance was defined as a p-value of less than 0.05, ensuring the reliability and robustness of the predictions. This threshold helped to minimize the likelihood of false positives in the analysis.

To further investigate the molecular interactions between the identified small molecule therapeutics and their target proteins, we conducted molecular docking studies using AutoDock Vina software. This analysis allowed us to calculate the free binding energy between the small molecules and their respective targets, providing an estimate of the strength and stability of the interaction. The docking results were then visualized using Pymol software, offering detailed structural insights into how these compounds might interact with target macromolecules. By examining the binding affinities and specific interactions, we gained a deeper understanding of the potential mechanisms through which these small molecules could act as therapeutic agents for LGG.

2.10. Development and validation of a prognostic nomogram

To develop a comprehensive prognostic tool for LGG patients, we gathered detailed clinical data from all training and validation cohorts. This data included key variables such as age, gender, tumor grade, IDH1 mutation status, and treatment history, encompassing both radiotherapy and chemotherapy. By integrating these clinical variables with the ARGs-based RiskScore, we utilized multivariate Cox regression analysis to construct prognostic nomograms. This method enabled us to create a comprehensive tool for predicting patient outcomes by combining genetic and clinical data into a single, easy-to-interpret model. These nomograms were designed to combine genetic factors with traditional clinical prognostic indicators, offering a more holistic approach to predicting patient outcomes.

The nomograms were visualized using the replot package in R, providing a clear and user-friendly representation of the integrated

prognostic model. To evaluate the predictive accuracy of these nomogram. This analysis assessed the model's ability to predict survival outcomes at 1, 3, and 5 years, providing a measure of its effectiveness in a clinical setting. Additionally, we calibrated the nomograms by generating calibration plots, which compared the predicted survival probabilities with the actual observed outcomes. These plots were crucial for assessing the precision and reliability of the model, ensuring that it could accurately guide treatment decisions and improve patient management.

3. Results

3.1. WGCNA-based analysis of LGG-associated gene module

The study's workflow is illustrated in Fig. 1. Through differential expression analysis, we identified 4311 DEGs, consisting of 2480 upregulated and 1831 downregulated genes when compared to normal tissues (Fig. 2A). These DEGs were further analyzed using WGCNA. A soft threshold of 10 was selected as the optimal parameter for network construction (Fig. 2B). By applying clustering criteria with minModuleSize = 30 and MEDissThres = 0.25, we identified eight distinct gene modules (Fig. 2C and D). The blue module exhibited the highest correlation among the modules with respect to differences between normal and tumor tissues (r = ±0.9, p < 0.0001) (as shown in Fig. 2E). For a more detailed analysis, a Venn diagram was employed to find the intersection between genes in the blue module with high correlation and the 77 ARGs associated with LGG, which were then selected for further study (Fig. 2F).

3.2. Characteristics of the ARGs

Our study revealed that several ARGs were notably upregulated in LGG, including genes such as SMARCB1, FYN, and ADGRG1. Conversely, certain genes like NOS3, XAF1, and LGALS8 were found to be downregulated in LGG tissues (Fig. 3A). To gain a clearer understanding of the interactions among these ARGs, we conducted a PPI network analysis using the STRING database. The resulting network, illustrating the connections between these proteins, is displayed in Fig. 3B. In addition, Pearson correlation analysis was conducted to examine the relationships between ARGs in LGG, which uncovered strong correlations between specific gene pairs, such as BRMS1 with GSTP1, NUDT1 with BRMS1, and SLC39A6 with CSNK2A1. These gene pairs exhibited correlation coefficients ranging from 0.7 to 0.769, with p-values below 0.001, indicating a strong likelihood that these genes may be involved in related biological functions (Fig. 3C).

The study also examined CNV and somatic mutation frequencies among ARGs in LGG. A significant occurrence of CNV-associated mutations was observed (Fig. 3D), with gene amplifications most commonly found in MYC, SMARCE1, and NOTCH1, while losses were frequently noted in NOS3, CASP3, and SMAD4 (Fig. 3E).

Furthermore, the somatic mutation rates of the 77 ARGs were investigated, revealing TP53 as the most frequently mutated gene, with a mutation rate of 45.3 %. Other genes, such as NOTCH1, PIK3CA, and EGFR, also exhibited higher mutation rates, ranging
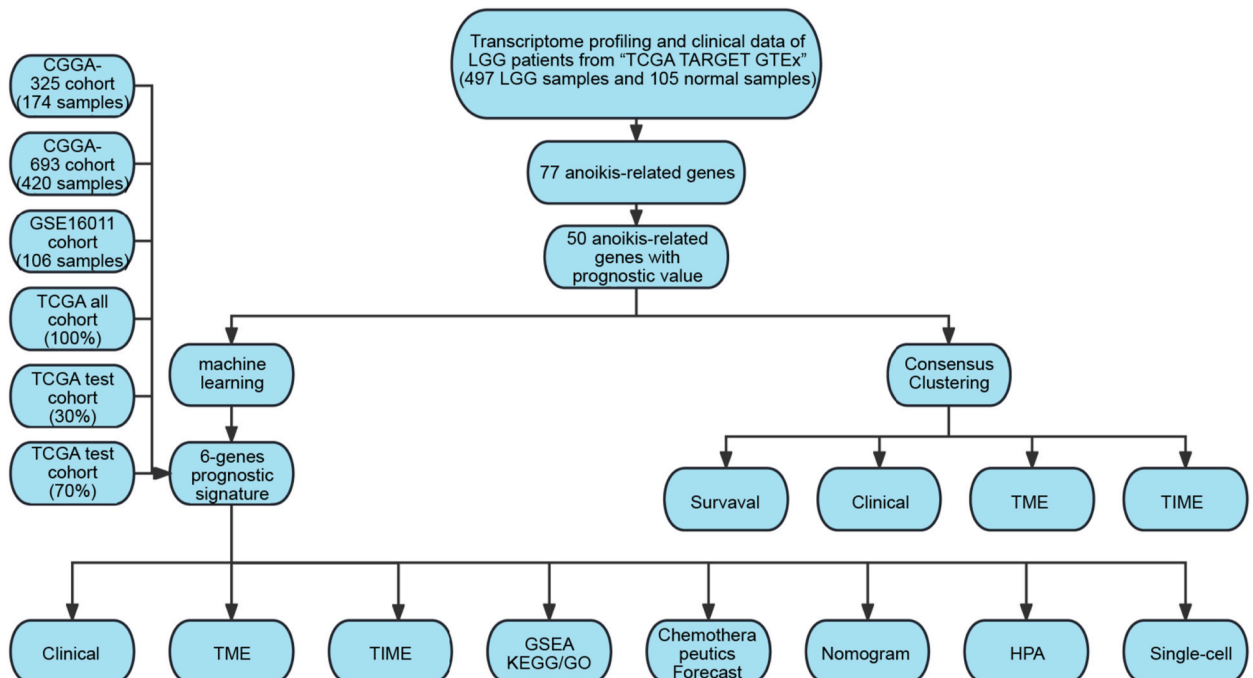

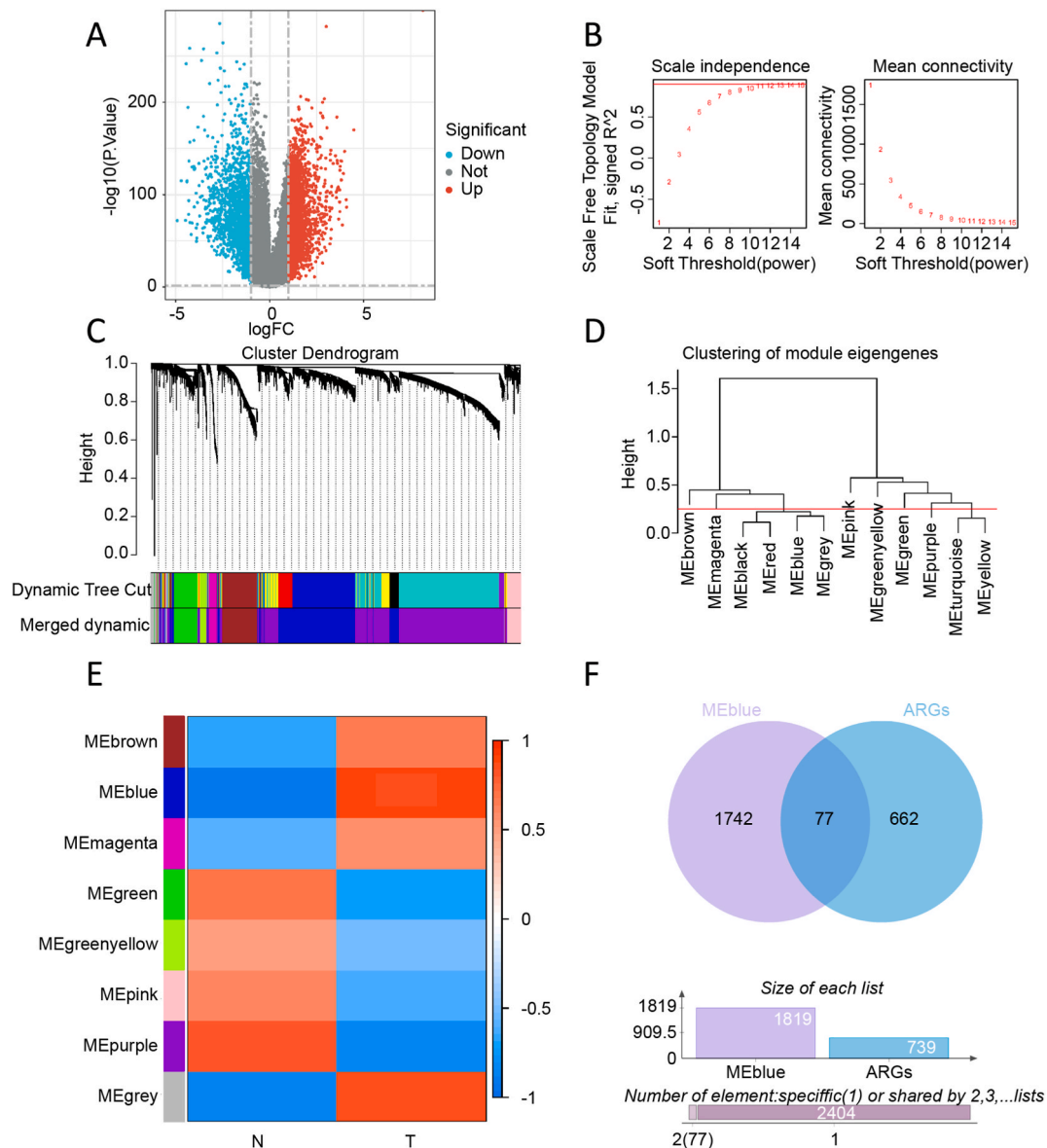**Fig. 1.** The graphic depicts the study's process.

Fig. 2. Anoikis-Relevant Gene Selection. (A) Volcano map display of DEGs, 2480 up-regulated genes, and 1831 down-regulated genes. (B) Scale-free fit indices for soft threshold power. (C) A tree diagram of differentially expressed gene clusters based on different metrics. (D) Clustering of the module eigengenes. (E) Correlations among gene modules are displayed on a heat map. The blue module contains 1819 of the best LGG-associated genes (r = ±0:9, p = 2e - 216). (F) Venn diagram for identifying 77 anoikis-related genes in LGG samples.

between 6.2 % and 7.4 %. The remaining ARGs demonstrated relatively lower mutation frequencies, as detailed in Fig. 3F.

3.3. Prognostic impact of ARG-based immune phenotypes in LGG

To assess the prognostic impact of the 77 ARGs, we carried out an in-depth univariate Cox regression analysis. Through this analysis, we identified 50 ARGs that demonstrated a significant relationship with patient prognosis, underscoring their importance as predictive markers for survival outcomes. These findings are depicted in Fig. 4A, providing a visual overview of the genes that could be integral to understanding patient prognosis and guiding future research. To visualize the complex relationships among these genes and their implications for LGG prognosis, we created a network diagram (Fig. 4B). By employing consensus clustering techniques based on the expression patterns of these 50 ARGs, we were able to identify two distinct modification patterns, as illustrated in Fig. 4C and D. Subsequent survival analysis revealed a statistically significant difference in outcomes between the two clusters. Patients in Cluster 2 had a median survival time of 2.036 years, whereas those in Cluster 1 experienced a significantly longer median survival of 7.994 years (P < 0.001). This stark contrast underscores the prognostic relevance of these clusters. These findings, as depicted in Fig. 4E,
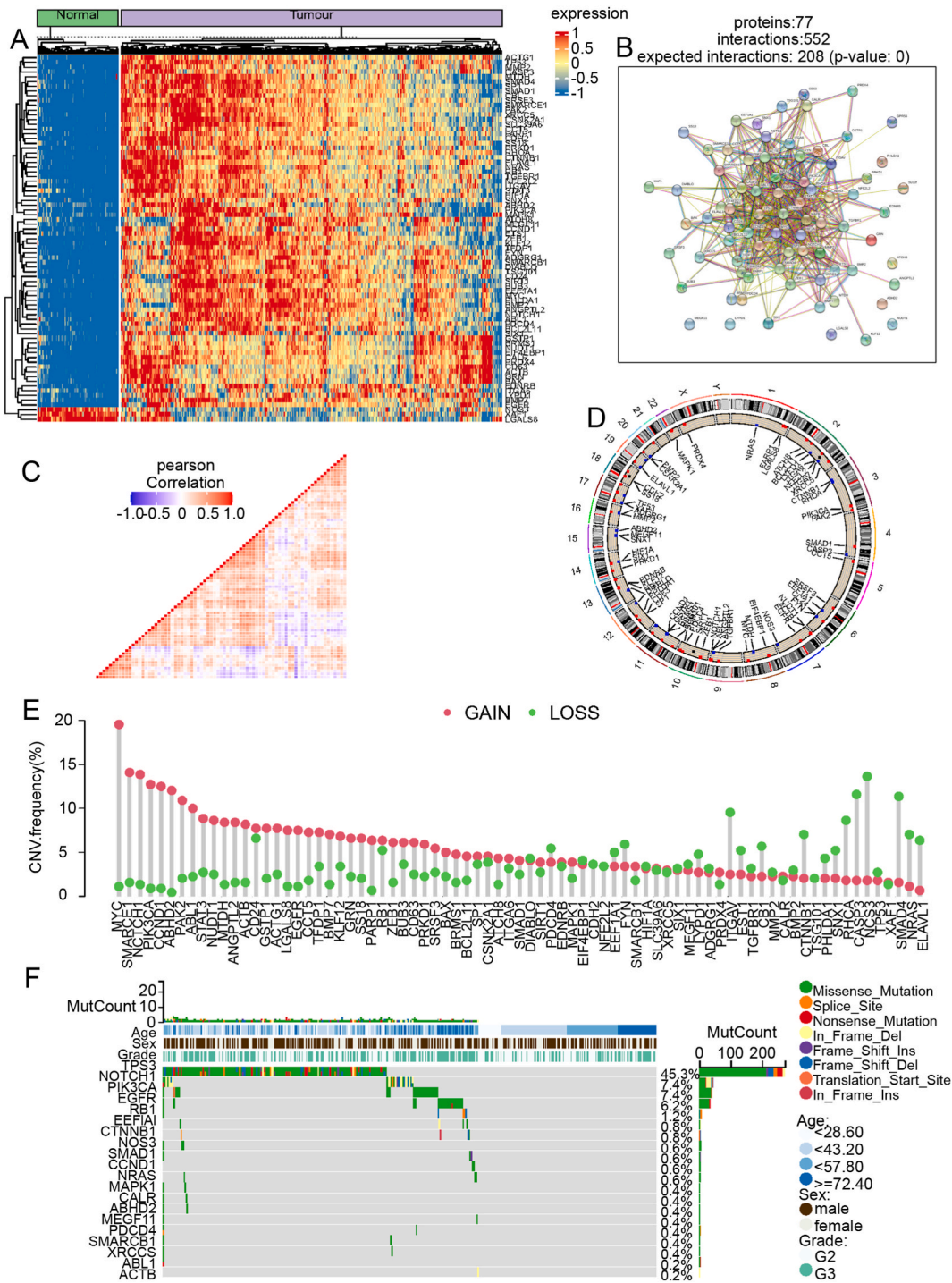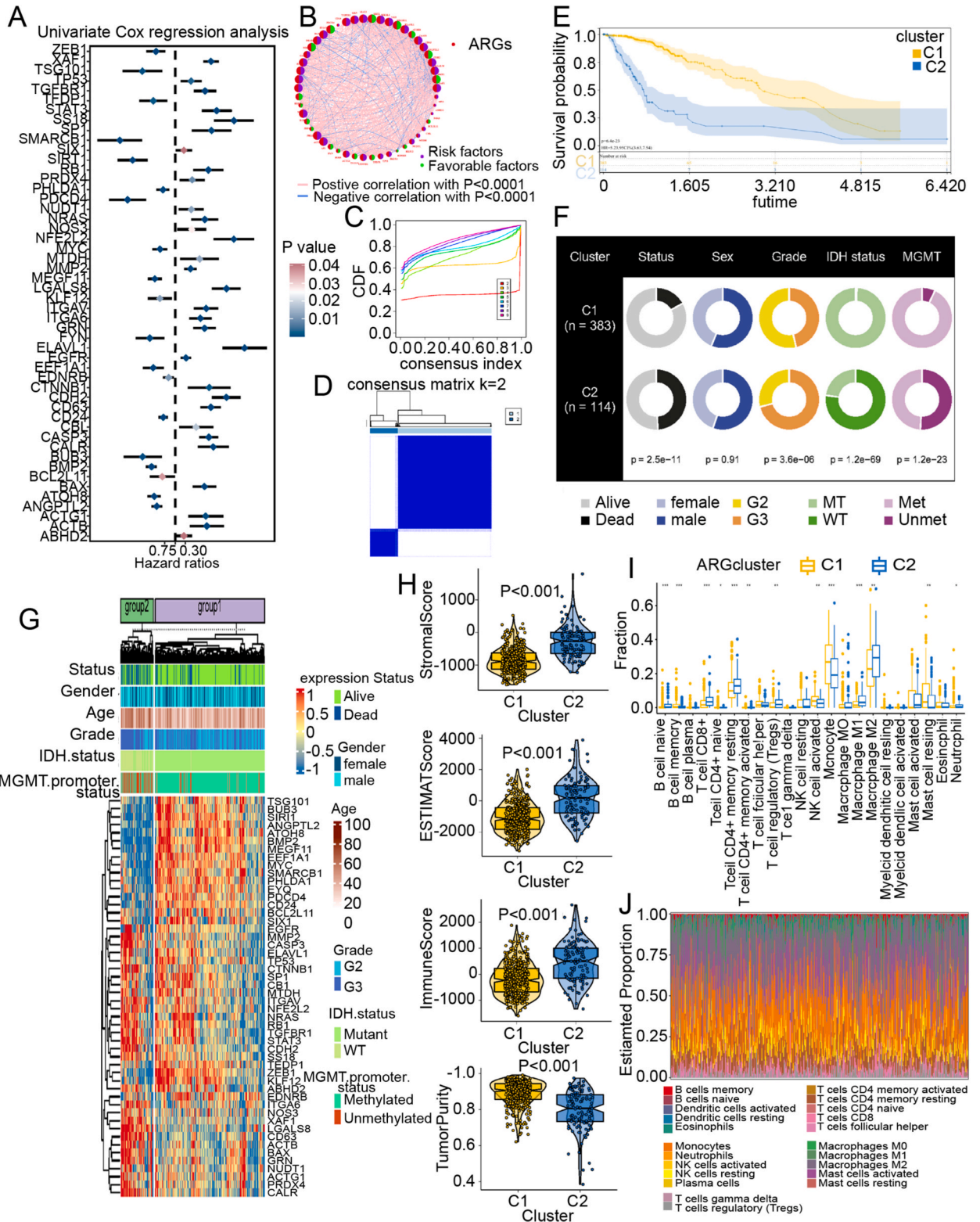
Fig. 3. Characterization of anoikis-related genes. (A) Heat map showing the gene expression profiles of 77 ARGs in normal and LGG brain tissue samples. (B) PPI among the ARGs.(C) Correlation of 77 anoikis-related genes. (D) Localization of nest loss-associated genes on LGG chromosome 23. (E) Frequency of CNV variants in anoikis-related genes in LGG samples. (F) The incidence of anoikis-related mutations in LGG samples shows the top 20 most frequent mutations.

underscore the prognostic value of these clusters in predicting patient survival.

Further examination of clinical characteristics between the clusters revealed notable distinctions. Cluster 2 had a higher proportion of grade G2 tumors (P < 0.001), a greater prevalence of IDH-WT status (P < 0.001), and a larger percentage of WGMT-unmet status compared to Cluster 1 (Fig. 4F). Heatmaps were used to illustrate the differential expression of ARGs across the two clusters (Fig. 4G).

(caption on next page)

Fig. 4. Characterization between molecular subtypes. (A) Forest plot of the association between 50 anoikis-related genes with prognostic value and OS. (B) 50 genes associated with anoikis and their interactions are depicted in a network diagram with predictive value. (C) Relative change in area under the cumulative distribution function (CDF) curve for consensus clustering when k = 2 to k = 10. (D) The heat map depicts the consensus clustering scheme for 50 ARGs out of 497 LGGs (k = 2). (E) Kaplan-Meier curves for OS of C1 and C2 subtypes. (F) Pie chart of the chi-square test for clinical information between C1 and C2 subtypes. (G) Heat map showing 50 ARGs expressed in various isoforms. Blue denotes low expression, whereas red denotes strong expression. (H) Compare the TME components of the two clustered subtypes. (I) The relative proportion of immune infiltrates in Cibersort. (J) Two clustering components of 22 immuno-infiltrative cells in the Cibersort algorithm.

We further analyzed the tumor microenvironment (TME) within each cluster, finding that Cluster 2 exhibited higher ImmuneScore, StromalScore, and EstimateScore, but a lower level of tumor purity compared to Cluster 1 (Fig. 4H). To gain deeper insights into the differences in immune cell infiltration between the two subtypes, we employed the CIBERSORT algorithm alongside the LM22 signature matrix. This approach enabled us to estimate the proportions of 22 distinct tumor-infiltrating immune cell types within each subtype (Fig. 4I and J). Consistent findings were observed across other analytical methods as well (Fig. S1). These results suggest a distinct immune landscape in each cluster, with Cluster 1 being characterized by a "cold" immune phenotype, indicative of a less active immune response, while Cluster 2 is associated with a "hot" immune phenotype, reflecting a more robust immune activity.

3.4. Development of a prognostic risk model

In this study, we systematically divided the dataset comprising 497 cases of LGG into two cohorts to facilitate robust analysis. The data was split in a 7:3 ratio, resulting in a training cohort of 348 cases and a validation cohort of 149 cases. This division ensured that the majority of the data was used to build and refine our models, while the remaining portion was reserved for validating the findings and confirming their generalizability. To identify genes significantly linked to overall survival, we employed both LASSO and RF regression techniques. During our analysis, the LASSO method identified a total of 15 genes that were significantly associated with clinical outcomes at a Lambda value of 0.02547996, as shown in Fig. 5A. In parallel, the RF approach was employed, which identified a broader set of 42 genes that also showed a strong association with the clinical outcomes of LGG patients, as depicted in Fig. 5B. These complementary methods provided a robust framework for pinpointing the key genetic factors influencing prognosis in LGG. By intersecting the results from both methods, we narrowed down the list to 12 genes common to both analyses (Fig. 5C).

To further refine our findings and concentrate on the most relevant genes, we conducted a multivariate Cox regression analysis on the 12 genes that were identified as common to both the LASSO and Random Forest analyses. This step allowed us to pinpoint the key genes most strongly associated with patient outcomes. This rigorous analysis further narrowed down the list to six critical genes that demonstrated the most significant associations with clinical outcomes, as illustrated in Fig. 5D. These six genes are likely to be key determinants in the prognosis of LGG patients and may serve as valuable targets for future research and therapeutic strategies. Using these six genes, we constructed a prognostic risk model with the following formula:

ARG RiskScore = 0.1069 * exp(EGFR) + 0.383904 * exp(SIX1) + 1.353163 * exp(SP1) - 0.22767 * exp(ANGPTL2) - 0.69252 * exp(BMP2) - 0.45003 * exp(PDCD4).

This six-gene signature represents a powerful and reliable tool for predicting patient outcomes in LGG.

3.5. Prognostic value of the signature genes

For each patient, we computed a risk score derived from the expression levels of the six selected genes. Using the median risk score as a threshold, we categorized the 348 patients in the training cohort into high-risk and low-risk groups. Upon analyzing the distribution of these risk scores, it became evident that higher scores were strongly associated with reduced OS. This finding suggests that patients with elevated RiskScores are likely to have a significantly worse prognosis compared to those with lower scores. The clear link between higher risk scores and poor survival outcomes highlights the potential utility of this risk model in predicting patient prognosis and guiding treatment strategies. More specifically, the increased expression of SP1, EGFR, and SIX1 was linked to a higher risk, classifying these genes as risk factors. On the other hand, the higher expression levels of ANGPTL2, BMP2, and PDCD4 were associated with a lower risk, thereby identifying them as protective factors (Fig. 5E–G).

Survival analysis further confirmed that patients in the high-risk group had a significantly shorter OS compared to those in the low-risk group (Fig. 5H–J). In the training cohort, the six-gene signature successfully distinguished between the two groups, with high-risk patients having a median survival of 4.084 years, while low-risk patients had a much longer median survival of 10.304 years (P < 0.001). This clear survival disparity underscores the prognostic power of the six-gene signature in identifying patients at higher risk of poor outcomes. The accuracy and predictive power of the risk score were confirmed through time-dependent ROC curve analysis, which showed a robust performance, with the Area Under the Curve (AUC) consistently exceeding 0.8, reaching up to 0.87 for the prediction of 1-year survival (Fig. 5K–M).

These conclusions were further supported by results from three independent external validation cohorts: CGGA-325 (Test3), CGGA-693 (Test4), and GSE-16011 (Test5). For each of these cohorts, risk scores were recalculated, and patients were stratified into high-risk and low-risk groups based on the cohort-specific median RiskScore. The results from these external validations aligned with those observed in the TCGA-LGG cohort, reinforcing the reliability and robustness of our six-gene prognostic signature (Fig. S2). This consistency across different cohorts highlights the potential of this signature as a dependable tool for predicting patient outcomes across various clinical settings.
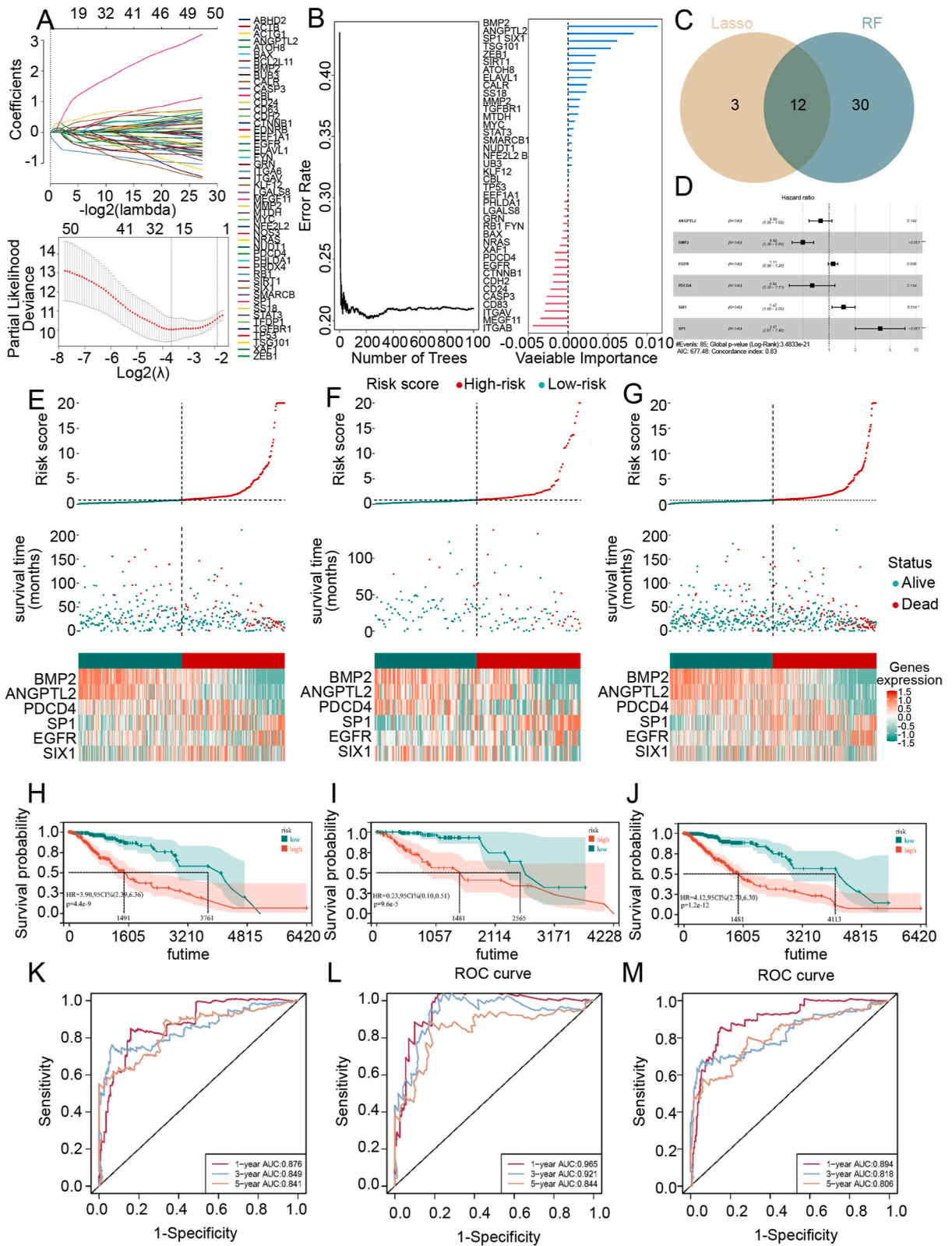
(caption on next page)

Fig. 5. Finding prognostic genes and predicting survival in TCGA cohort individuals with LGG. (A) Lasso regression of 50 anoikis-related genes with predictive value. (B) Random Forest (RF) algorithms of 50 anoikis-related genes with prognostic significance. (C) Gene selection by the Venn diagram is based on two algorithms. (D) Forest plot showing further screening for genes affecting prognosis using multifactorial COX analysis.Distribution of risk scores, survival times, and gene expression panels, TCGA-Train (E), TCGA-Test (F), TCGA-All (G).Kaplan-Meier curves for LGG overall survival (OS) based on risk scores, TCGA-Train (H), TCGA-Test (I), and TCGA-All (J). The ROC curves show the predictive efficiency of the risk scores for 1, 3, and 5 years, TCGA-Train (K), TCGA-Test (L), and TCGA-All (M).

3.6. Characterization of risk subtypes

In this research, we utilized the ESTIMATE algorithm to delve into the significance of ARGs within the TME and their correlation with high-risk subtypes among patients with LGG. The analysis revealed that individuals classified within the high-risk subgroup displayed several clinical and molecular characteristics associated with a poorer prognosis. Specifically, these patients exhibited a significantly higher incidence of mortality, a larger proportion of grade G3 tumors, a greater prevalence of IDH1 wild-type status, and a higher frequency of non-methylated MGMT promoters (all $P < 0.05$), which are all indicators of a more aggressive tumor phenotype (Fig. 6A).

Furthermore, the high-risk group was characterized by an elevated infiltration of immune cells, increased stromal content, and a higher overall estimated fraction, alongside a marked reduction in tumor purity (Fig. 6B). This suggests a more complex and reactive TME in the high-risk subgroup. Multiple immune infiltration analysis methods, including CIBERSORT, TIMER, QUANTISEQ, EPIC, and MCPCOUNTER, consistently identified higher levels of immune cell infiltrates in patients within the high-risk category (Fig. 6C–H). This likely reflects a compensatory response to a weaker local immune response, as the tumor microenvironment attempts to counteract the aggressive tumor behavior by mobilizing more immune cells.

Further examination of immune checkpoint genes and HLA genes revealed a pattern of upregulation in the high-risk subgroup and downregulation in the low-risk subgroup (Fig. 6H and I). These distinctions underscore the critical role that immune cell infiltration within the TME plays in determining the prognosis of LGG patients across different risk strata.

Additionally, the study explored the relationship between immune infiltration scores derived from the CIBERSORT algorithm and the expression levels of six prognostic genes associated with anoikis, further illuminating the interaction between these genes and the immune microenvironment (Fig. 7A). The analysis demonstrated that all six of these prognostic genes influenced the composition of the immune microenvironment. Notably, there were positive correlations between ARG-related risk scores and the abundance of certain immune cell types, including resting CD4 memory T cells, M1 and M0 macrophages, and regulatory T cells (Tregs) (Fig. 7C–F). Conversely, negative correlations were observed with activated NK cells, memory B cells, activated mast cells, and monocytes (Fig. 7G–J). These findings indicate that ARGs play a pivotal role in shaping immune cell infiltration within the TME, potentially affecting the stratification and prognostic evaluation of LGG patients.

The investigation also assessed the differences inTMB between the high- and low-risk LGG subgroups, along with their association with the TME. The results revealed significant disparities in TMB scores between the two groups ($P < 0.05$), with a positive correlation identified between TMB scores and ARG-related risk scores (Fig. 7K). This suggests that the risk scores employed in this study are effective in capturing variations in the TME and might be useful in evaluating the overall tumor landscape. Furthermore, the analysis identified the most frequently mutated genes within the high-risk subgroup, with IDH1, TPRB, and ATRX mutations being the most common. Notably, IDH1 mutations were significantly less frequent in the high-risk group (61.7 %) compared to the low-risk group (91.4 %), highlighting a potential molecular distinction that could inform prognosis and therapeutic strategies (Fig. 7L)

3.7. Biological functional analysis

This study sought to explore the molecular biological processes modulated by six key prognostic biomarkers associated with anoikis. To achieve this, we conducted a differential expression analysis, which led to the identification of 133 DEGs with a fold change threshold of 1 and a significance level of $P < 0.05$. Among these DEGs, 85 were found to be upregulated while 48 were downregulated (Fig. 8A). Our Gene Set Enrichment Analysis (GSEA) identified distinct pathways that were differentially enriched between the high-risk and low-risk patient groups. In the high-risk group, pathways such as Fc epsilon RI signaling, Fc gamma-mediated phagocytosis, and mismatch repair were significantly enriched, suggesting heightened immune activity and DNA repair mechanisms. On the other hand, the low-risk group showed enrichment in pathways related to ribosome function, selenoamino acid and taurine and hypotaurine metabolism, indicating a focus on cellular metabolism and protein synthesis (Fig. 8B). These findings highlight the biological differences underlying the risk stratification and provide insights into potential therapeutic targets.

Further functional characterization using the KEGG pathway analysis indicated that these DEGs were involved in a diverse array of biological processes, including cellular functions, environmental information processing and organismal systems (Fig. 8C–E, and G). In addition, GO analysis revealed that a significant portion of these DEGs were enriched in immune-related biological pathways (Fig. 8D–F, and H). These findings underscore the critical role that these molecular pathways play in regulating the anoikis process, which in turn impacts the biological characteristics and behavior of LGG across various risk subtypes.

3.8. Predicts chemotherapy regimens

To explore potential chemotherapeutic agents that could enhance the prognosis for LGG, we employed the oncoPredict package in
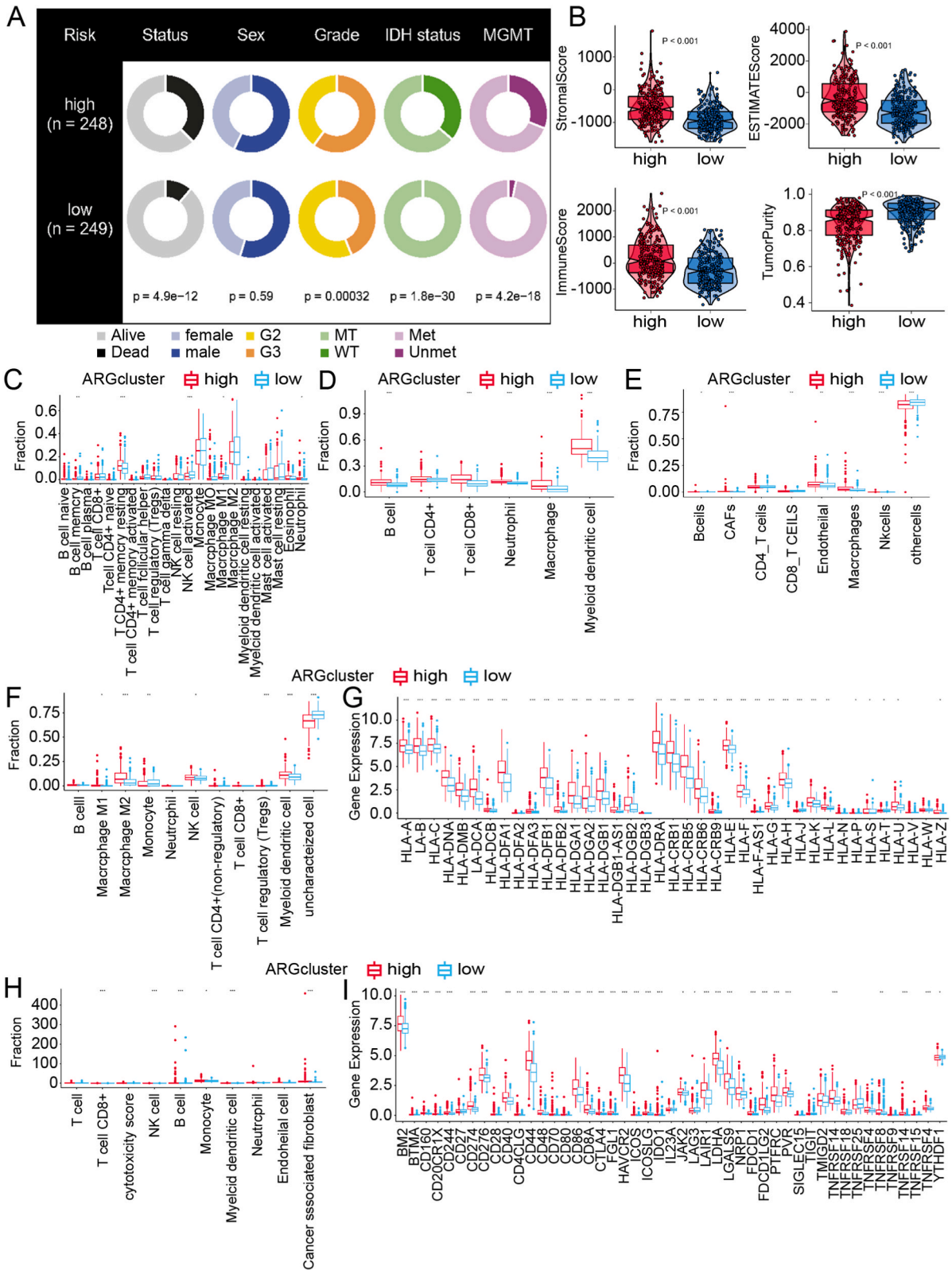
(caption on next page)

Fig. 6. Characterization between risk subtypes. (A) Pie chart of the chi-square test for clinical information between Risk-high and Risk-low subtypes. (B) Compare the TME components in the high-risk and low-risk groups. Immune infiltration in two risk subgroups based on the CIBERSORT algorithm(C), TIMER algorithm (D), QUANTISEQ algorithm (E), EPIC algorithm (F), and MCPCOUNTER algorithm (H).Differential expression of immune checkpoints(G) and HLA genes(I) between the two risk subtypes.
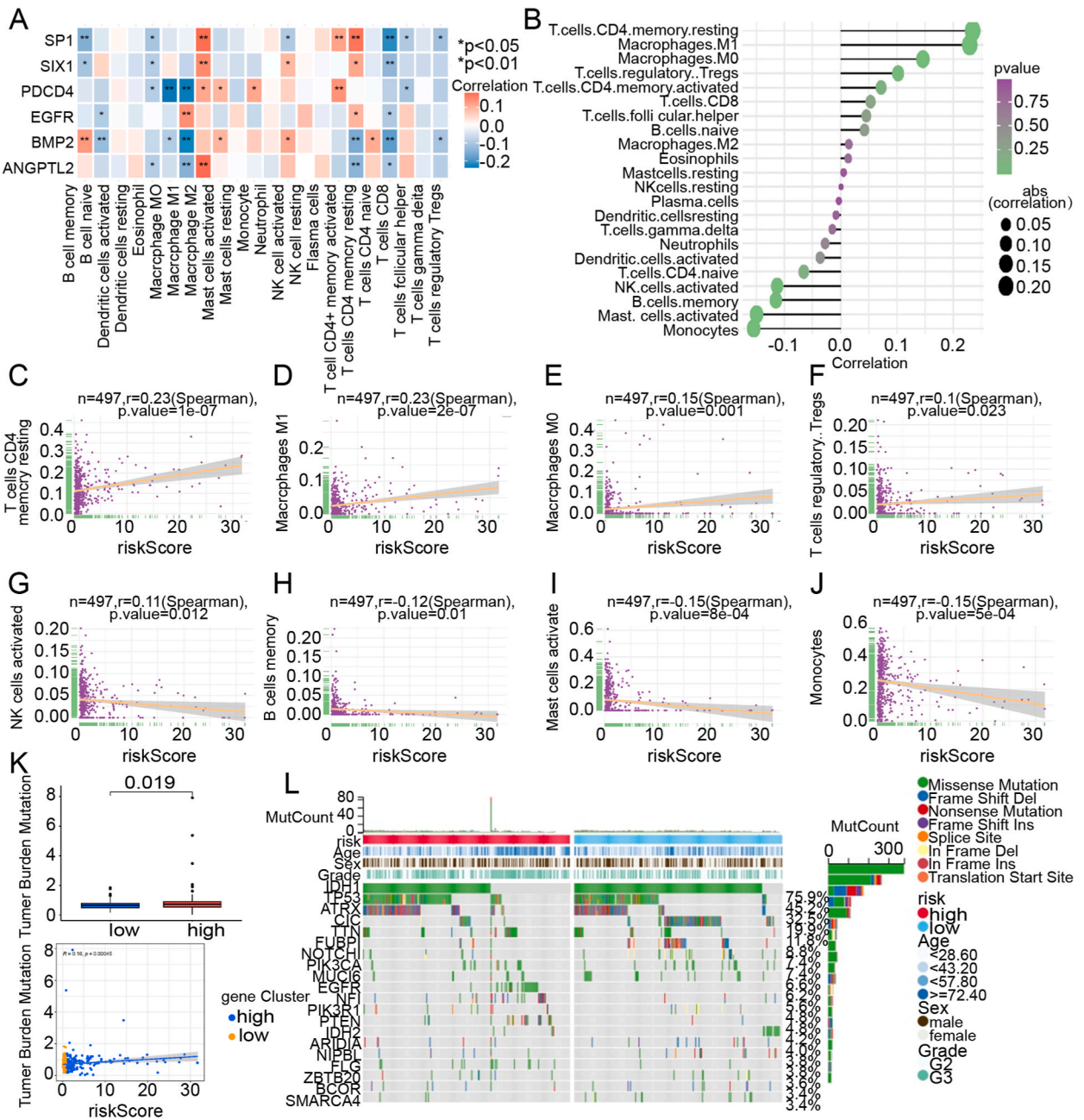

Fig. 7. Characterization between risk subtypes. (A) Correlation of immune cells with 6 ARGs.(B) Correlation of immune cells with an ARG risk score. Correlation analysis of ARG risk scores with the proportion of T. cells.CD4.memory.resting(C), Macrophages.M1(D), Macrophages.M0(E), T. cells.regulatory.Tregs(F), NK.cells.activated(G), B.cells.memory(H), Mast.cells.activated(I) and Monocytes(J). (K) Differences in TMB scores between the two risk subtypes and Correlation of TMB scores with ARG risk scores. (L) Somatic mutations in LGG patients in high- and low-risk subgroups.
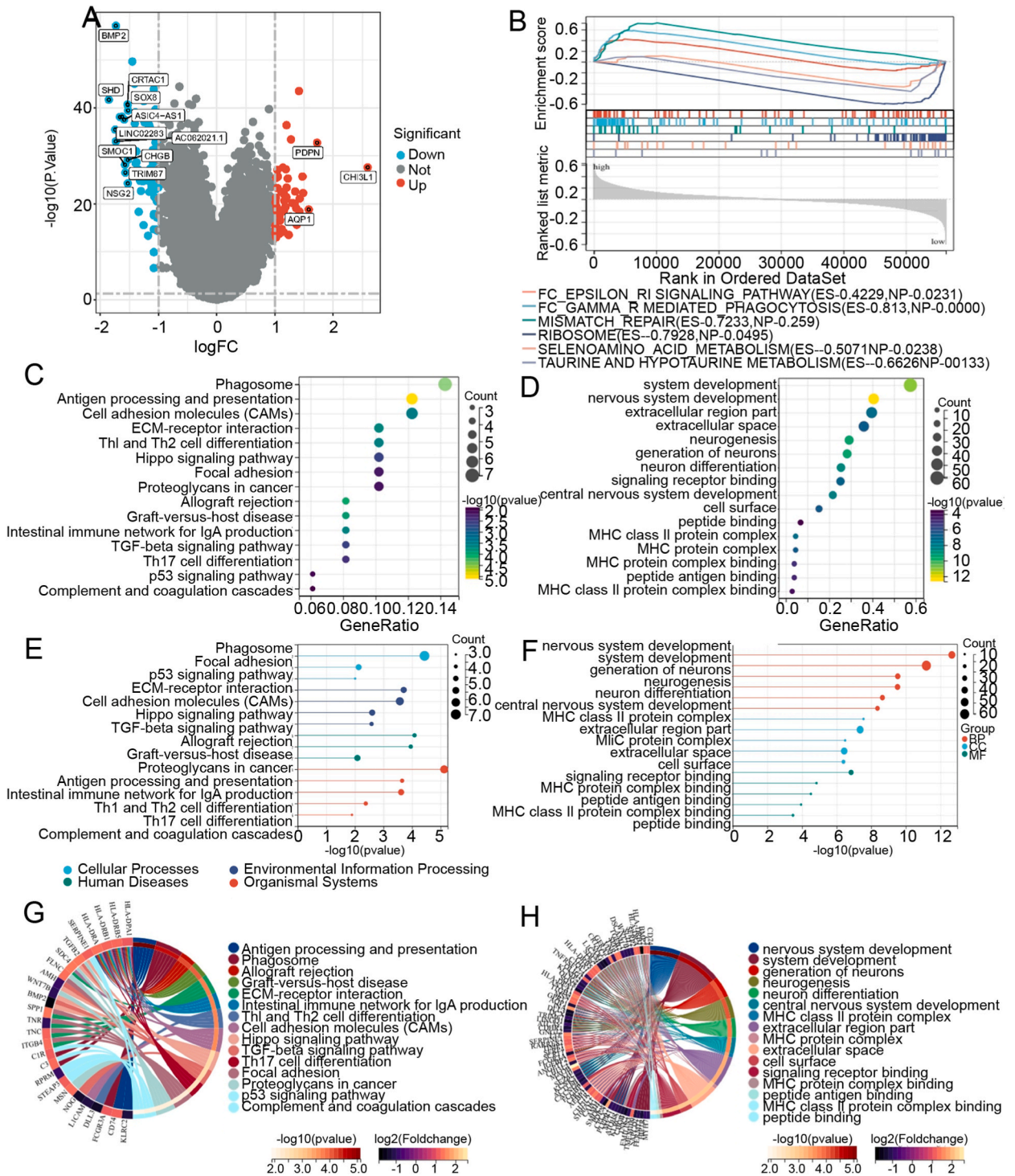
Fig. 8. Identification and biological functional analysis of DEGs between risk subtypes. (A) Volcano map analysis of differentially expressed genes, 85 up-regulated genes, and 48 down-regulated genes.Enrichment analysis of the KEGG (C, E, G) and GO (D, F, H) signaling pathways is presented.

R to evaluate the differential sensitivity of various chemotherapeutic drugs across the two identified risk subtypes. Through this analysis, we pinpointed eight drugs that demonstrated significant differences in efficacy between the high-risk and low-risk groups: LFM-A13, S-Trityl-L-cysteine, Rapamycin, Parthenolide, and MP470 (Fig. 9A).

Beyond the oncoPredict analysis, we also conducted a screening of DEGs between the high- and low-risk subtypes, which were

subsequently submitted to the CMap database. This screening led to the identification of four small molecule drugs—Risperidone, Pipamperone, and Erastin—that exhibited strong potential therapeutic effects (Fig. 9B). Importantly, a negative CMap score indicated that these drugs could potentially counteract the harmful biological processes associated with the high-risk subtype.

To further investigate the molecular interactions between these small molecules and their target proteins, we performed molecular docking studies using AutoDock Vina. The docking results, visualized using PyMOL, revealed significant binding affinities: DRD2 with Risperidone at −8.1 kcal/mol, DRD2 with Pipamperone at −6.9 kcal/mol, MAPK1 with FR-180204 at −7.4 kcal/mol, and VDAC2 with Erastin at −7.6 kcal/mol. A detailed analysis of these interactions showed the formation of three hydrogen bonds in the complexes of DRD2 with Risperidone (Fig. 9C), DRD2 with Pipamperone (Fig. 9D), and VDAC2 with Erastin (Fig. 9E), while four hydrogen bonds were identified in the binding of MAPK1 with FR-180204 (Fig. 9F). These findings highlight the strength and specificity of these molecular interactions.
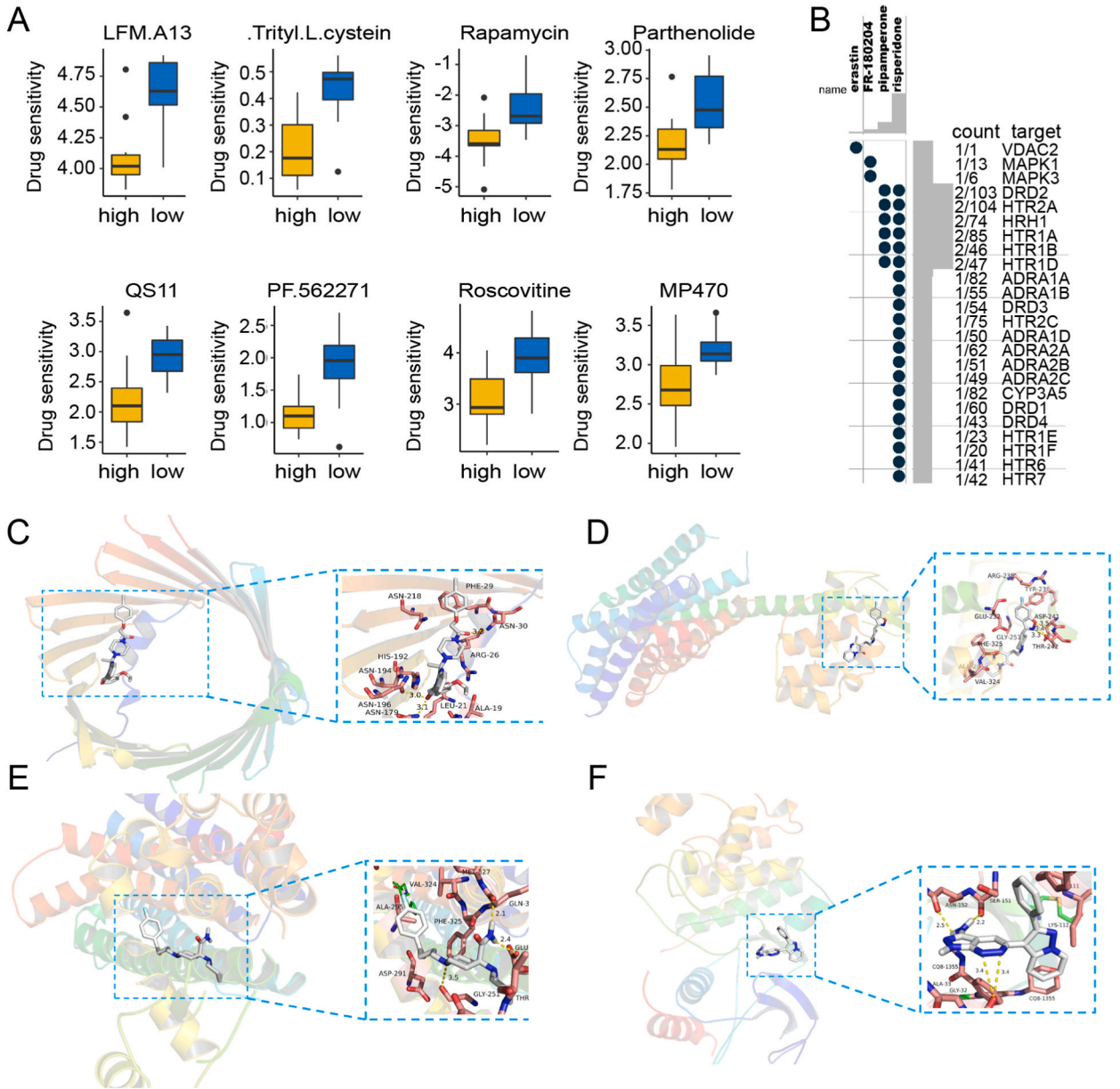

Fig. 9. Predicting new chemotherapy regimens. (A) To evaluate the sensitivity of chemotherapy, IC50 values for antineoplastic medicines were computed for high-risk and low-risk groups. (B) Predicting small molecule perturbation chemotherapeutics for high-risk patients based on DEGs.The binding sites and interaction results for the best-selected conformations showed that we found three hydrogen bonds in the binding of VDAC2 and Erastin (C), DRD2 and Risperidone (D), DRD2 and Pipamperone (E), and we found four hydrogen bonds in the binding of MAPK1 and FR-180204 (F).

These findings suggest that the identified small molecule drugs may offer therapeutic advantages for high-risk LGG patients by specifically targeting key molecular interactions. This highlights their potential for integration into new chemotherapy regimens aimed at improving patient outcomes.

3.9. Development and evaluation of survival models for nomograms

To assess the potential of the ARG-based RiskScore as an independent prognostic marker for LGG patients, we performed multivariate Cox regression analyses, leading to the development of a nomogram survival model. The analysis demonstrated that the RiskScore is a significant independent prognostic factor for patients in the TCGA training cohort, with a hazard ratio of 1.063, a 95 % confidence interval of 1.037–1.090, and a p-value of less than 0.001. This finding was consistently validated across multiple independent cohorts, underscoring the robustness of the RiskScore as a reliable predictor of patient outcomes (Fig. 10A).

Building on this foundation, we developed a robust and comprehensive nomogram model within the TCGA cohort, utilizing multivariate Cox regression analysis. This model was specifically designed to predict OS at 1, 3, and 5 years. It incorporated several key clinical variables, including patient age, gender, tumor grade (GII and GIII), IDH1 mutation status (mutated or wild-type), and the receipt of adjuvant therapies, such as radiation and/or chemotherapy. These adjuvant therapies were classified into three distinct groups: both radiation and chemotherapy, either radiation or chemotherapy, and no adjuvant therapy. Additionally, the ARG-based RiskScore was integrated into the model, enhancing its prognostic capability.

The performance of this nomogram was rigorously evaluated through ROC curve analysis, which confirmed its high diagnostic accuracy not only within the TCGA training cohort but also across five external validation cohorts. The model achieved a commendable C-index of 0.82, with a 95 % confidence interval ranging from 0.77 to 0.87 within the training cohort, demonstrating its reliability and robustness (Fig. 10B–D, F, H, J, L). Moreover, the predictive accuracy of the model for 1-, 3-, and 5-year survival was further validated through calibration curves, which showed a strong concordance between the predicted and actual survival outcomes for LGG patients (Fig. 10C–E, G, I, K, M).

The final nomogram, as illustrated in Fig. 11A, underscores the practical utility of this model in clinical settings. By offering a detailed and personalized survival prediction, the nomogram serves as a valuable tool for guiding treatment decisions, enabling clinicians to tailor therapeutic strategies based on individual patient risk profiles. This comprehensive approach not only enhances the precision of prognostication but also supports more informed and effective patient management.

3.10. Immunohistochemical analysis

In this study, we used IHC stained images from the HPA database to compare the protein expression levels of six ARGs between LGG tissues and normal brain tissues. Our objective was to determine whether these ARGs showed differential expression at the protein level, supporting their potential role in LGG. The analysis revealed that five out of the six ARGs were significantly more expressed in LGG tissues compared to normal brain tissues (Fig. 12A–E). This elevated protein expression aligns with previous findings at the RNA level, indicating that these genes might be actively involved in LGG progression. The overexpression of these proteins in tumor tissues suggests they could be key players in the tumor's survival and growth, making them potential targets for therapeutic interventions in LGG.

3.11. Single-cell analysi

To explore the expression of these six ARGs within the TME of LGG, we analyzed single-cell RNA sequencing data from the GSE89567 dataset, accessed through the TISCH database. This dataset provided a detailed view of gene expression across 17 different cell groups, including various cancer and immune cell types (Fig. 13A and B).

The analysis showed that hub genes were expressed within cancer cells, with PDCD4 being particularly noteworthy for its expression across all four intermediate cell types identified in the TME (Fig. 13C and D). This widespread expression suggests that PDCD4 and possibly other ARGs play a significant role not just in cancer cells but also in influencing other cell types within the tumor environment, potentially contributing to LGG progression.
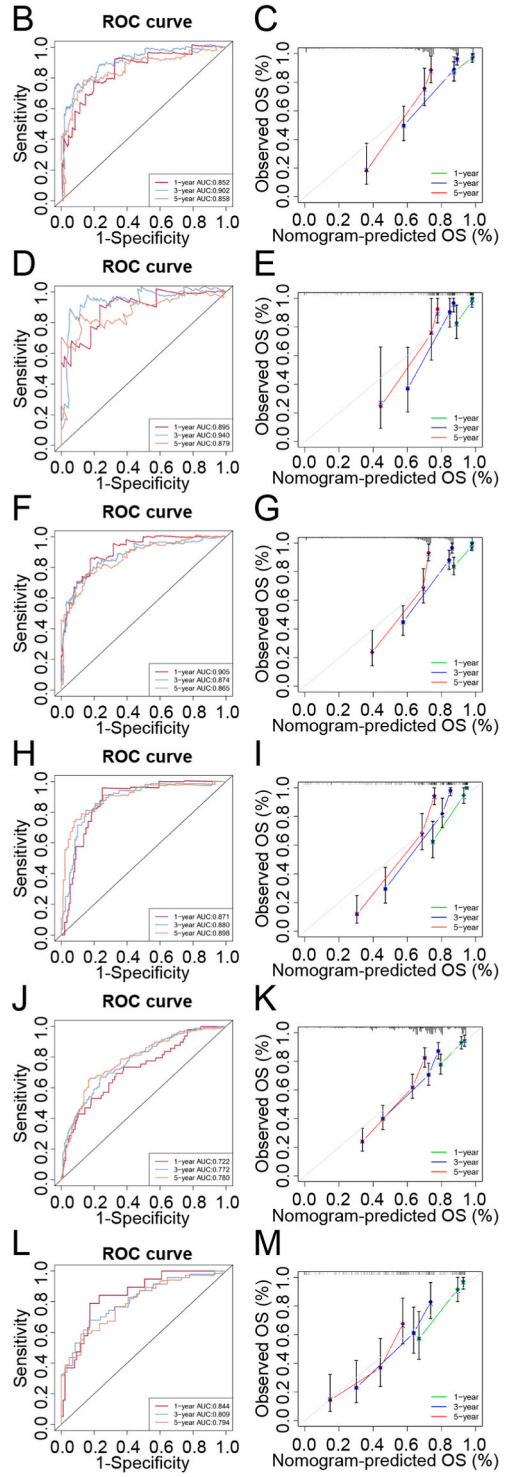
4. Discussion

LGG represent a common and aggressive type of primary brain tumor that originates from glial cells, the supportive cells of the central nervous system. These tumors pose significant challenges in treatment and understanding due to the complex and often elusive mechanisms underlying their pathogenesis. Unlike higher-grade gliomas, LGGs tend to progress more slowly, yet their potential for malignant transformation makes them a critical focus of neuro-oncological research. The relatively unpredictable nature of LGGs, coupled with their capacity for recurrence and progression to more aggressive forms, underscores the necessity for innovative therapeutic strategies and robust prognostic models to improve patient outcomes.

Anoikis, a specific form of programmed cell death that is initiated when cells lose their adhesion to the extracellular matrix (ECM) and neighboring cells, has emerged as a promising therapeutic target in the battle against LGG. The concept of anoikis was first introduced by Steven M. Frisch in 1994, who discovered that epithelial cells undergo apoptosis upon detachment, a phenomenon they termed "anoikis," which translates to "homelessness" in Greek [13]. This mechanism is crucial for maintaining tissue homeostasis by ensuring that cells detached from their original location do not adhere to inappropriate sites, which could otherwise result in abnormal

A

Variable	HR	lower 95%CI	upper 95%CI	pvalue
TCGA-train				
Age	1.033	1.012	1.054	1.53e-03
Gender	0.763	0.482	1.208	2.49e-01
Grade	2.083	1.177	3.689	1.18e-02
IDH status	2.551	1.333	4.880	4.67e-03
ad_therapy	1.131	0.691	1.853	6.25e-01
riskScore	1.063	1.037	1.090	1.57e-06
TCGA-test				
Age	1.037	1.006	1.069	2.03e-02
Gender	0.868	0.409	1.842	7.12e-01
Grade	1.590	0.648	3.902	3.11e-01
IDH status	2.360	0.974	5.721	5.73e-02
ad_therapy	1.321	0.695	2.512	3.96e-01
riskScore	1.130	1.069	1.194	1.53e-05
TCGA-all				
Age	1.033	1.012	1.054	1.53e-03
Gender	0.763	0.482	1.208	2.49e-01
Grade	2.083	1.177	3.689	1.18e-02
IDH status	2.551	1.333	4.880	4.67e-03
ad_therapy	1.131	0.691	1.853	6.25e-01
riskScore	1.063	1.037	1.090	1.57e-06
CGGA325				
Age	1.011	0.989	1.033	3.31e-01
Gender	1.694	1.110	2.585	1.44e-02
Grade	3.053	1.888	4.938	5.34e-06
IDH status	1.411	0.832	2.393	2.02e-01
ad_therapy	1.355	0.905	2.028	1.40e-01
riskScore	1.220	1.112	1.338	2.65e-05
CGGA693				
Age	1.001	0.987	1.015	9.11e-01
Gender	0.865	0.651	1.150	3.17e-01
Grade	2.438	1.742	3.413	2.03e-07
IDH status	0.969	0.770	1.221	7.91e-01
ad_therapy	1.218	0.937	1.583	1.41e-01
riskScore	1.358	1.268	1.4552	48e-18
GSE16011				
Age	1.032	1.013	1.051	9.79e-04
Gender	0.792	0.492	1.275	3.37e-01
Grade	1.017	0.564	1.832	9.56e-01
IDH status	1.342	1.031	1.746	2.87e-02
ad_therapy	1.005	0.691	1.461	9.80e-01
riskScore	1.396	1.234	1.579	1.16e-07

$\log_2 HR$

B ROC curve — 1-year AUC:0.852, 3-year AUC:0.902, 5-year AUC:0.858

C Nomogram-predicted OS (%) — 1-year, 3-year, 5-year

D ROC curve — 1-year AUC:0.895, 3-year AUC:0.940, 5-year AUC:0.879

E Nomogram-predicted OS (%) — 1-year, 3-year, 5-year

F ROC curve — 1-year AUC:0.905, 3-year AUC:0.874, 5-year AUC:0.865

G Nomogram-predicted OS (%) — 1-year, 3-year, 5-year

H ROC curve — 1-year AUC:0.871, 3-year AUC:0.880, 5-year AUC:0.898

I Nomogram-predicted OS (%) — 1-year, 3-year, 5-year

J ROC curve — 1-year AUC:0.722, 3-year AUC:0.772, 5-year AUC:0.780

K Nomogram-predicted OS (%) — 1-year, 3-year, 5-year

L ROC curve — 1-year AUC:0.844, 3-year AUC:0.809, 5-year AUC:0.794

M Nomogram-predicted OS (%) — 1-year, 3-year, 5-year

(caption on next page)

Fig. 10. Independent predictive value of the ARG risk score. (A) Forest plots of multivariate Cox analysis of ARG hazard scores and clinical characteristics in six cohorts. The ROC curves show the predictive efficiency of the model at 1, 3, and 5 years, TCGA-Train (B), TCGA-Test (D), TCGA-All (F), CGGA-325(H), CGGA-693 (J), GSE-16011 (L). Calibration plots were used to predict 1-, 3- and 5-year survival rates in six cohorts, TCGA-Train (C), TCGA-Test (E), TCGA-All (G), CGGA-325(I), CGGA-693 (K), GSE-16011 (M).

growths and metastasis. In the context of cancer, resistance to anoikis is a defining feature of metastatic cells, enabling them to survive after detachment, travel through the circulatory system, and eventually colonize distant organs or tissues. Therefore, understanding and potentially manipulating anoikis pathways in LGG could offer novel approaches to thwart tumor progression and metastasis.

Anoikis plays a crucial role in preventing inappropriate cell attachment and migration, processes that are fundamental to tumor invasion and metastasis. By triggering cell death in detached cells, anoikis acts as a protective mechanism against the spread of cancerous cells to new areas of the body. The ability of tumor cells to resist anoikis is a key factor in their survival and dissemination, making it an attractive target for therapeutic intervention. The diverse prognoses associated with LGG, ranging from relatively indolent to highly aggressive, highlight the importance of identifying reliable biomarkers that can facilitate early intervention and targeted treatment for high-risk patients. The heterogeneity within LGG populations necessitates a more nuanced approach to treatment, where patients can be stratified based on molecular and genetic profiles to receive personalized therapies. Consequently, this study aimed to discover and validate a novel multigene biomarker that could stratify patients, predict prognosis, and guide treatment responses based on an ARG-related risk model in LGG.

Our study conducted a comprehensive analysis of 77 anoikis-associated genes, revealing distinct expression profiles in LGG samples. These genes were identified using variance analysis and WGCNA across datasets from TCGA and GTEx, providing a broad perspective on their roles in LGG. The use of WGCNA allowed us to identify gene modules that are highly correlated with clinical traits, thereby pinpointing potential key drivers of tumor behavior. It was observed that many ARGs were significantly overexpressed in LGG, with strong correlations detected among them. These findings suggest that anoikis resistance mechanisms are actively engaged in LGG and may contribute to the tumor's ability to evade programmed cell death, thus supporting its growth and invasiveness.

Additionally, a notable prevalence of CNV-associated mutations was identified among the ARGs, with copy number gains being the most frequent alteration. This underscores the potential role of genomic instability in driving the expression of genes that provide survival advantages to tumor cells, thereby contributing to tumor progression and resistance to cell death mechanisms like anoikis. Among the somatic mutations analyzed, TP53 showed the highest mutation frequency at 45.3 %, underscoring its critical role in glioma pathogenesis. TP53 is well-known as a tumor suppressor gene, and its mutation is closely linked to the loss of cell cycle control and resistance to apoptosis, which are critical factors in the progression of many cancers, including gliomas. The high frequency of TP53 mutations observed in our cohort suggests that this gene may play a pivotal role in the resistance to anoikis seen in lower-grade gliomas (LGG), potentially contributing to the tumor's ability to survive and proliferate despite detachment from the extracellular matrix.

The study identified 50 ARGs with significant prognostic value, which were subsequently used for unsupervised clustering and feature construction. This methodological approach enabled the classification of LGG patients into two molecular subgroups with distinct clinicopathological features, tumor immune microenvironments, and prognoses. These subgroups, identified through consensus clustering, provide a framework for understanding the heterogeneity within LGG and its implications for patient outcomes. The C1 cluster, primarily composed of patients with grade 2 tumors, IDH1 mutations, and MGMT promoter methylation, was associated with a better prognosis compared to the C2 cluster, which exhibited more aggressive characteristics. By calculating Immune-Score, StromalScore, EstimateScore, and tumor purity, we found that patients in the C1 cluster exhibited higher levels of stromal, immune, and extracellular components, whereas the C2 cluster was characterized by higher tumor purity. This observation is significant as it underscores the critical role that the composition of the tumor microenvironment (TME) plays in the progression and aggressiveness of lower-grade gliomas (LGG). The differing TME profiles between clusters suggest that interactions between tumor cells and their surrounding environment may influence tumor behavior and patient outcomes.

The TME, which is composed of various cellular and non-cellular components, is increasingly recognized as a critical factor in cancer development and progression. This dynamic environment includes immune cells, stromal cells, blood vessels, signaling molecules, and extracellular matrix components, all of which interact with tumor cells and influence their growth, survival, and metastatic potential. The interaction between tumor cells and their surrounding microenvironment influences not only the behavior of the tumor but also its response to therapy. In our study, the significant differences in TME composition between the C1 and C2 clusters suggest that targeting TME remodeling could be a promising strategy to inhibit tumor growth. By altering the cellular and extracellular components of the TME, it may be possible to disrupt the supportive environment that tumor cells rely on, thereby reducing their ability to proliferate, invade, and resist treatment. This approach highlights the potential of TME-focused therapies as a complementary strategy in the treatment of LGG [31]. Targeting the TME, particularly by modulating immune cell infiltration and stromal components, could enhance the efficacy of existing therapies and potentially lead to the development of new therapeutic approaches. Furthermore, the significant impact of the immune microenvironment on tumor behavior [32] prompted us to investigate immune cell infiltration within these subgroups. Our findings indicated that the C1 subtype is associated with a "cold" immune phenotype, characterized by lower levels of immune cell infiltration, suggesting a less active or suppressed immune environment within the tumor. In contrast, the C2 subtype is associated with a "hot" immune phenotype, reflecting a more active immune response with higher levels of immune cell infiltration. This distinction between the immune profiles of the two subtypes may have important implications for the effectiveness of immunotherapies and the overall prognosis of LGG patients. This distinction is crucial as it implies that patients in the C2 cluster, with higher immune cell infiltration, might benefit more from immunotherapy.
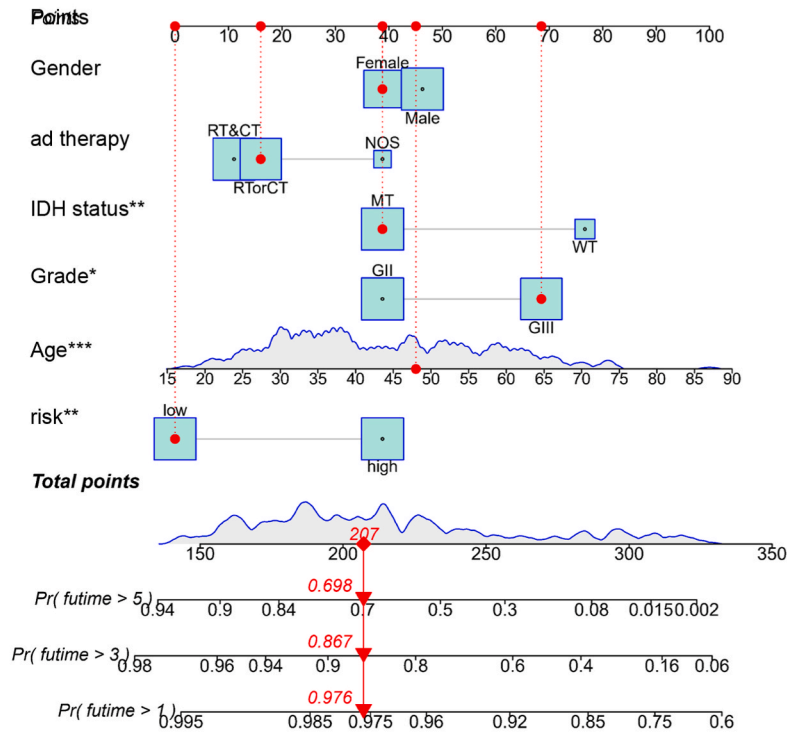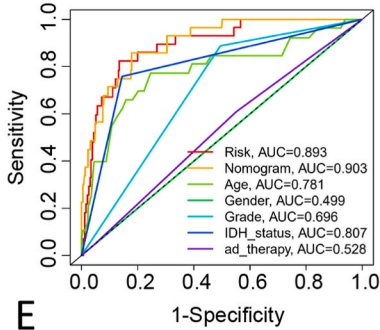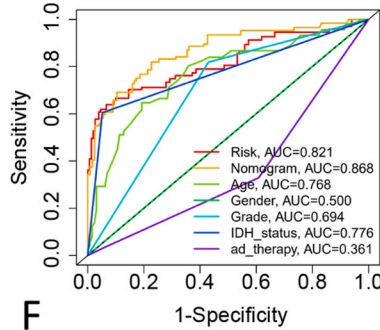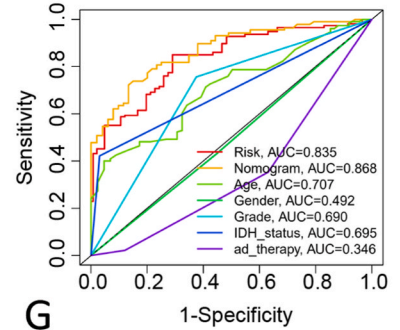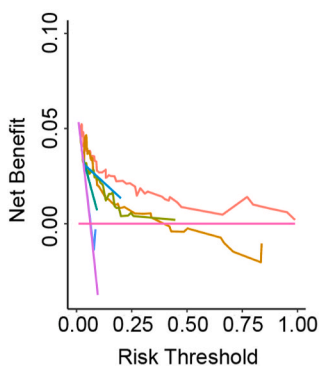
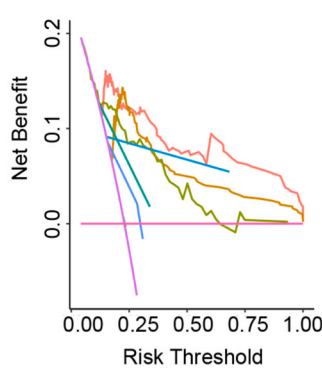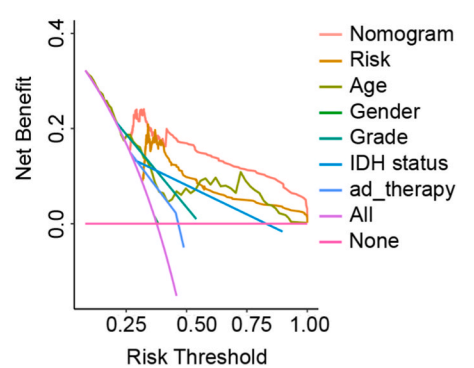**Fig. 11.** Independent predictive value of the ARG risk score. (A) Predicting overall survival using nomograms of the TCGA-LGG cohort. (B, C, D) ROC curves show risk score, age, gender, tumor grade, IDH mutation status, additional adjuvant therapy, and predictive efficiency of the Nomogram model at 1, 3, and 5 years. (E, F, G) DCA shows risk score, age, gender, tumor grade, IDH mutation status, additional adjuvant therapy, and predictive efficiency of the Nomogram model at 1, 3, and 5 years.
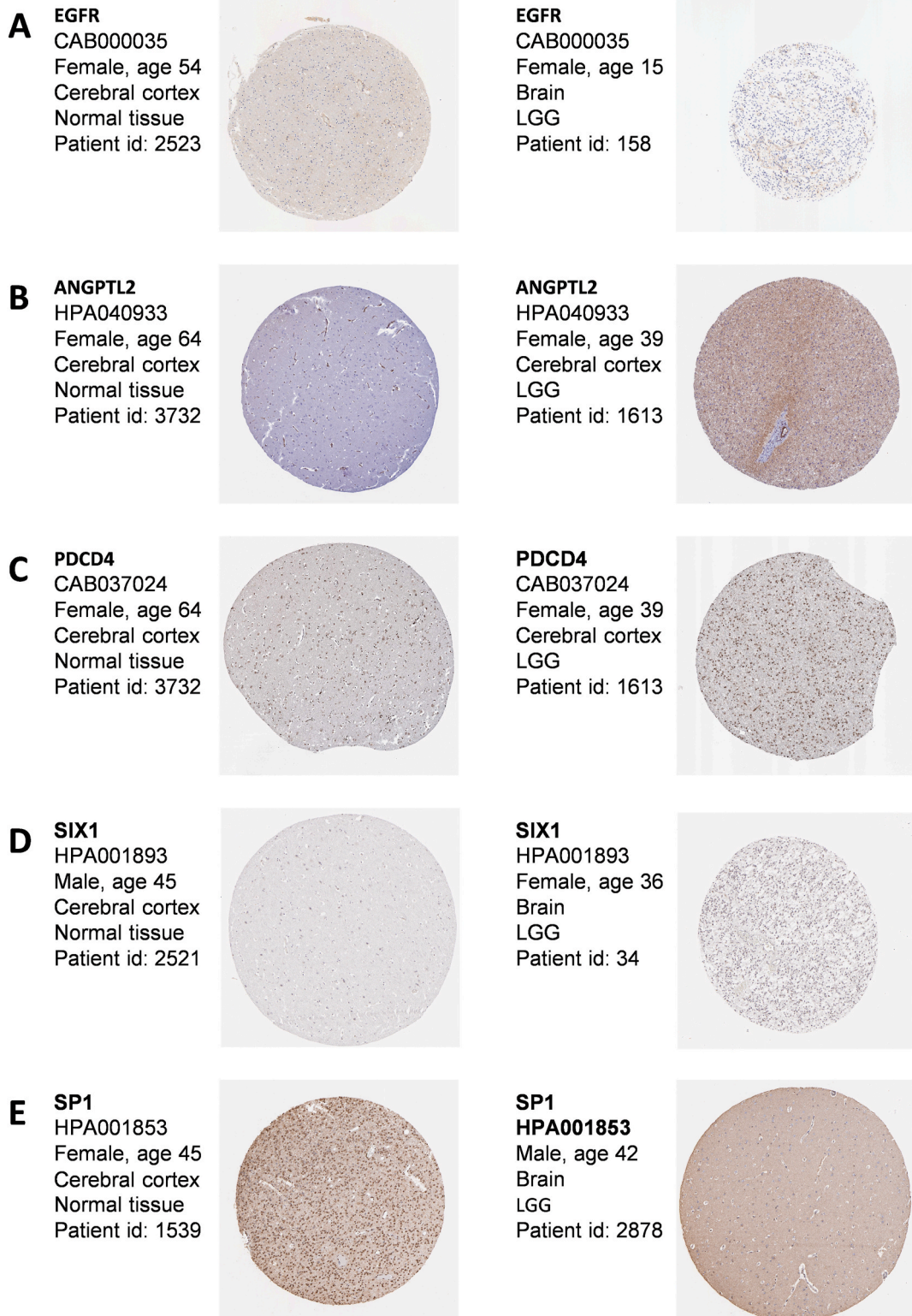
Fig. 12. Immunohistochemical analysis of six Anoikis-related model genes. Immunohistochemical analysis showed that the protein expression levels of EGFR(A), ANGPTL2(B), PDCD4(C), SIX1(D), and SP1(E) were significantly higher in LGG samples than in normal samples.
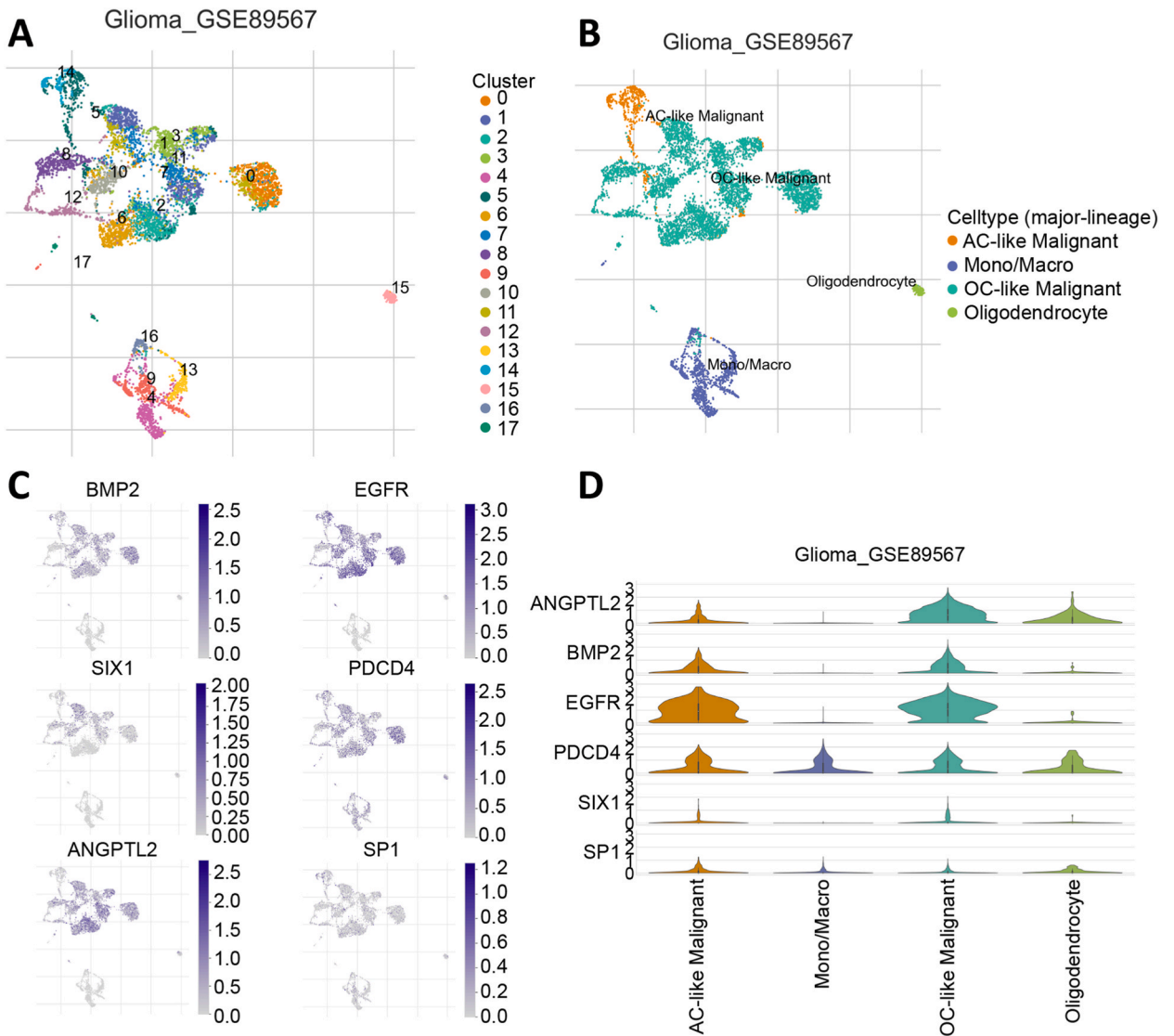
Fig. 13. Single-cell analysis of six Anoikis-related model genes. (A) We expressed notes for all cells in GSE89567.(B) Notes for all cell types in GSE89567.(C, D) EGFR, SIX1, SP1, ANGPTL2 PDCD4, and BMP2 in each cell type.

To construct a reliable prognostic model, we utilized advanced machine learning techniques, including Lasso regression and random forest (RF) algorithms. These methods enabled us to identify the most relevant features and predict patient outcomes with greater accuracy by efficiently handling high-dimensional data and reducing the risk of overfitting. These models were meticulously validated across multiple internal and external cohorts, including those from TCGA, CGGA, and GEO, to ensure robustness and predictive accuracy. The rigorous validation process underscores the reliability of the findings and their potential applicability in clinical settings. The ARG-based signature identified includes six crucial genes, creating a robust risk-scoring system. Notably, EGFR, a critical member of the tyrosine kinase family within the epidermal growth factor receptor group, is essential in cancer development [33,34]. The EGFR signaling pathway, which interacts with ligands like AREG, BTC, EGF, EREG, HBEGF, and TGF, controls glioma resistance to anoikis by activating downstream pathways such as PI3K/AKT, MAPK, and PLCγ/PKC[21,35–37], 35, 36, 21,35–37. This involvement underscores the pivotal role of EGFR in ensuring tumor survival and resisting programmed cell death.

In normal cells, downregulation of EGFR leads to the upregulation of Bim, a BH3-only protein, which initiates apoptosis via the intrinsic pathway [38]. However, in cancer cells, EGFR overexpression has been shown to inhibit Bim through the MAPK pathway, thereby inducing resistance to anoikis [39]. This resistance allows tumor cells to survive under conditions where they would normally undergo programmed cell death, contributing to the persistence and spread of the tumor. The PI3K/AKT signaling pathway is intricately involved in various cellular processes, one of which is the regulation of anoikis resistance. NF-kB, a critical transcription factor and a downstream target within this pathway, is specifically influenced by EGFR activity, thereby playing a significant role in promoting resistance to anoikis in cancer cells, particularly in gliomas [40]. This resistance is crucial for tumor cells, as it allows them to

survive despite the loss of cell adhesion, a key step in metastasis. Additionally, the amplification of EGFR, which is frequently observed in aggressive gliomas, can lead to the activation of NF-kB via the transmembrane protein TMEM43/LUMA. This mechanism further strengthens the ability of glioma cells to resist anoikis, contributing to their survival, proliferation, and potential to invade other tissues [41]. This complex network of interactions underscores the importance of EGFR as a therapeutic target in LGG and highlights the potential of targeting this pathway to overcome resistance to anoikis.

SIX1, one of the genes identified in our study, is a transcription factor that plays a critical role in embryonic development. It is involved in the regulation of several developmental processes, including cell proliferation, differentiation, and survival. Beyond its essential functions during early development, SIX1 has been increasingly recognized for its role in oncogenesis. Aberrant expression of SIX1 has been implicated in a variety of malignancies, where it contributes to tumor progression and metastasis. In particular, in the context of gliomas, SIX1 is associated with the promotion of aggressive tumor behavior, aiding in the proliferation and survival of glioma cells, which ultimately contributes to the malignancy of this type of brain cancer[42–44], 42–44, 42–44. SIX1 promotes the expression of α5β1 integrin, enhancing tumor cell adhesion to ECM molecules, which in turn supports invasion and resistance to anoikis, facilitating metastasis [45]. The ability of SIX1 to promote cell adhesion and migration makes it a key player in the metastatic process and a potential target for therapies aimed at preventing tumor spread. Similarly, SP1, another transcription factor critical for embryonic development, upregulates survivin, an inhibitor of apoptosis, thereby protecting cancer cells from anoikis by suppressing intrinsic apoptotic pathways [46–48]. The role of SP1 in regulating apoptosis suggests that it could be a valuable target for therapies designed to induce cell death in tumors that have become resistant to other forms of treatment.

ANGPTL2, a gene identified in our study, is well-documented for its role in exacerbating metastasis across various types of cancers. This protein has been shown to facilitate the spread of tumor cells to distant sites, significantly contributing to the aggressiveness of malignant diseases. Moreover, ANGPTL2 is also involved in the inhibition of anoikis resistance, a process that typically prevents detached cells from surviving and establishing secondary tumors [49]. In the context of lower-grade gliomas (LGG), the presence of high levels of ANGPTL2 is particularly concerning, as it suggests a greater propensity for the tumor cells to survive after detachment and potentially metastasize. This dual role of ANGPTL2 in both promoting metastasis and influencing anoikis resistance underscores its importance as a potential therapeutic target. Targeting ANGPTL2 could be especially beneficial for LGG patients who exhibit elevated levels of this protein, as it may help reduce tumor cell survival and limit the spread of the disease.

On the other hand, PDCD4 is a tumor suppressor gene that plays a critical role in regulating apoptosis, the programmed cell death process that serves as a natural defense mechanism against cancer. PDCD4 functions by promoting apoptosis, thereby eliminating potentially dangerous cells before they can proliferate uncontrollably. Additionally, PDCD4 inhibits tumor cell proliferation and metastasis, making it a crucial factor in preventing the spread of cancer [50]. The loss or downregulation of PDCD4 has been associated with poor outcomes in various cancers, including gliomas, as it removes a key barrier to unchecked tumor growth and dissemination. Restoring or enhancing PDCD4 function could therefore be a promising strategy in cancer therapy, aiming to suppress tumor progression and improve patient prognosis. Its expression, regulated by miR-21, is downregulated in conditions that increase anoikis resistance [51]. The loss of PDCD4 in tumors suggests that restoring its function could enhance the effectiveness of treatments that rely on the induction of apoptosis. Lastly, BMP2, a member of the TGF-β superfamily, plays a role in embryonic development and tissue homeostasis. Both BMP2 and BMP9 are known to inhibit anchorage-independent survival, thereby promoting anoikis[52,53], 52,53. The role of these proteins in maintaining normal cellular architecture and preventing the survival of detached cells makes them important targets for therapeutic intervention in LGG.

The analysis of our prognostic model revealed a clear distinction in outcomes between patients classified into different risk groups. Specifically, those categorized as high-risk exhibited significantly poorer prognoses, as evidenced by a combination of factors including increased infiltration of immune cells within the tumor microenvironment and elevated expression levels of immune checkpoint genes. These findings suggest that the high-risk patients are likely to benefit from therapies targeting immune checkpoints, as their tumors may be more susceptible to this type of immunotherapy. This observation is consistent with the growing body of research that highlights the crucial role of the immune system in cancer treatment. In particular, the activation of immune checkpoints has been shown to effectively unleash the body's natural immune response against cancer cells, providing a powerful therapeutic strategy to combat the disease. The correlation between high-risk status and immune checkpoint activity underscores the potential for checkpoint inhibitor therapies to improve outcomes in these patients, making it a promising avenue for targeted cancer treatment. The model included several independent prognostic factors, such as age, tumor grade, IDH1 status, history of adjuvant treatment, and the ARG-based RiskScore. The inclusion of these variables in the model highlights the multifaceted nature of LGG and the need for a comprehensive approach to treatment and prognosis. The development of a nomogram integrating these factors demonstrated strong predictive accuracy, as indicated by high AUC values. However, the retrospective nature of this study underscores the need for prospective validation and further research to confirm the role of the ARG signature in LGG pathogenesis.

Future research should prioritize conducting prospective studies to confirm the prognostic value of the ARG signature across various patient populations. Additionally, there should be an exploration into its potential role as a predictive marker for determining patient responses to therapies aimed at targeting anoikis pathways. Moreover, delving into the molecular mechanisms that drive anoikis resistance in LGG could uncover new therapeutic targets, paving the way for the development of more effective treatment strategies. Given the complexity of LGG and its interactions with the TME, a multidisciplinary approach involving genomics, proteomics, and immunology will be essential to unravel the full potential of anoikis-related therapies. The integration of ARG signatures into clinical practice could pave the way for personalized treatment strategies, ultimately improving outcomes for patients with this challenging disease.

5. Conclusion

In conclusion, our research highlights the critical role of anoikis in the progression of LGG. Through the examination of ARGs, we have identified pivotal molecular mechanisms that impact both the development of LGG and patient outcomes. The ARG-based risk model we developed serves as a valuable tool for predicting prognosis and guiding treatment decisions, underlining the significance of targeting anoikis pathways in therapeutic strategies. These findings enhance our understanding of LGG and may lead to the development of more effective, targeted therapies for patients facing this challenging disease.

Ethics approval

Not required.

Funding

No Funding.

Consent for publication

Not applicable.

Data availability statemen

The data supporting the findings of this study are derived from publicly accessible databases, promoting transparency and reproducibility. For further details or requests regarding the datasets and research materials utilized, please reach out to the corresponding author.

CRediT authorship contribution statement

Gu Linazi: Writing – original draft. **Aierpati Maimaiti:** Writing – original draft. **Zulihuma Abulaiti:** Formal analysis, Writing – original draft. **Hui Shi:** Data curation. **Zexin Zhou:** Data curation. **Mizhati Yimiti Aisa:** Methodology. **Yali Kang:** Methodology. **Ayguzaili Abulimiti:** Methodology. **Xierzhati Dilimulati:** Validation. **Tiecheng Zhang:** Validation. **Patiman Wusiman:** Validation. **Zengliang Wang:** Writing – review & editing. **Aimitaji Abulaiti:** Writing – review & editing, Writing – original draft.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We are appreciative of everyone who contributed to the open datasets utilized for this investigation.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e36989.

References

[1] Q.T. Ostrom, N. Patil, G. Cioffi, K. Waite, C. Kruchko, J.S. Barnholtz-Sloan, CBTRUS statistical report: primary brain and other central nervous system tumors diagnosed in the United States in 2013-2017, Neuro Oncol. 22 (2020) iv1–iv96.
[2] R.S. Nunna, S. Khalid, J.S. Ryoo, A. Sethi, R.W. Byrne, A.I. Mehta, Radiotherapy in adult low-grade glioma: nationwide trends in treatment and outcomes, Clin. Transl. Oncol. : official publication of the Federation of Spanish Oncology Societies and of the National Cancer Institute of Mexico 23 (2021) 628–637.
[3] D.N. Louis, A. Perry, P. Wesseling, D.J. Brat, I.A. Cree, D. Figarella-Branger, C. Hawkins, H.K. Ng, S.M. Pfister, G. Reifenberger, R. Soffietti, A. von Deimling, D. W. Ellison, The 2021 WHO classification of tumors of the central nervous system: a summary, Neuro Oncol. 23 (2021) 1231–1251.
[4] S. Tabrizi, H.A. Shih, The path forward for radiation therapy in the management of low-grade gliomas, Neuro Oncol. 22 (2020) 748–749.
[5] R. Jooma, M. Waqas, I. Khan, Diffuse low-grade glioma - changing concepts in diagnosis and management: a review, Asian journal of neurosurgery 14 (2019) 356–363.
[6] M.C. Tom, D.Y.J. Park, K. Yang, C.M. Leyrer, W. Wei, X. Jia, V. Varra, J.S. Yu, S.T. Chao, E.H. Balagamwala, J.H. Suh, M.A. Vogelbaum, G.H. Barnett, R. A. Prayson, G.H.J. Stevens, D.M. Peereboom, M.S. Ahluwalia, E.S. Murphy, Malignant transformation of molecularly classified adult low-grade glioma, Int. J. Radiat. Oncol. Biol. Phys. 105 (2019) 1106–1112.
[7] D.A. Forst, B.V. Nahed, J.S. Loeffler, T.T. Batchelor, Low-grade gliomas, Oncol. 19 (2014) 403–413.

[8] L.B. Nabors, J. Portnow, M. Ahluwalia, J. Baehring, H. Brem, S. Brem, N. Butowski, J.L. Campian, S.W. Clark, A.J. Fabiano, P. Forsyth, J. Hattangadi-Gluth, M. Holdhoff, C. Horbinski, L. Junck, T. Kaley, P. Kumthekar, J.S. Loeffler, M.M. Mrugala, S. Nagpal, M. Pandey, I. Parney, K. Peters, V.K. Puduvalli, I. Robins, J. Rockhill, C. Rusthoven, N. Shonka, D.C. Shrieve, L.J. Swinnen, S. Weiss, P.Y. Wen, N.E. Willmarth, M.A. Bergman, S.D. Darlow, Central nervous system cancers, version 3.2020, NCCN clinical practice guidelines in oncology, J. Natl. Compr. Cancer Netw. : J. Natl. Compr. Cancer Netw. 18 (2020) 1537–1570.

[9] R. Stupp, W.P. Mason, M.J. van den Bent, M. Weller, B. Fisher, M.J. Taphoorn, K. Belanger, A.A. Brandes, C. Marosi, U. Bogdahn, J. Curschmann, R.C. Janzer, S. K. Ludwin, T. Gorlia, A. Allgeier, D. Lacombe, J.G. Cairncross, E. Eisenhauer, R.O. Mirimanoff, Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma, N. Engl. J. Med. 352 (2005) 987–996.

[10] T. Jiang, D.H. Nam, Z. Ram, W.S. Poon, J. Wang, D. Boldbaatar, Y. Mao, W. Ma, Q. Mao, Y. You, C. Jiang, X. Yang, C. Kang, X. Qiu, W. Li, S. Li, L. Chen, X. Li, Z. Liu, W. Wang, H. Bai, Y. Yao, S. Li, A. Wu, K. Sai, G. Li, K. Yao, X. Wei, X. Liu, Z. Zhang, Y. Dai, S. Lv, L. Wang, Z. Lin, J. Dong, G. Xu, X. Ma, W. Zhang, C. Zhang, B. Chen, G. You, Y. Wang, Y. Wang, Z. Bao, P. Yang, X. Fan, X. Liu, Z. Zhao, Z. Wang, Y. Li, Z. Wang, G. Li, S. Fang, L. Li, Y. Liu, S. Liu, X. Shan, Y. Liu, R. Chai, H. Hu, J. Chen, W. Yan, J. Cai, H. Wang, L. Chen, Y. Yang, Y. Wang, L. Han, Q. Wang, Clinical practice guidelines for the management of adult diffuse gliomas, Cancer Lett. 499 (2021) 60–72.

[11] P.G. Gritsenko, N. Atlasy, C.E.J. Dieteren, A.C. Navis, J.H. Venhuizen, C. Veelken, D. Schubert, A. Acker-Palmer, B.A. Westerman, T. Wurdinger, W. Leenders, P. Wesseling, H.G. Stunnenberg, P. Friedl, p120-catenin-dependent collective brain infiltration by glioma cell networks, Nat. Cell Biol. 22 (2020) 97–107.

[12] P. Chiarugi, E. Giannoni, Anoikis: a necessary death program for anchorage-dependent cells, Biochem. Pharmacol. 76 (2008) 1352–1364.

[13] S.M. Frisch, H. Francis, Disruption of epithelial cell-matrix interactions induces apoptosis, J. Cell Biol. 124 (1994) 619–626.

[14] C. Frantz, K.M. Stewart, V.M. Weaver, The extracellular matrix at a glance, J. Cell Sci. 123 (2010) 4195–4200.

[15] H.J. Han, J.Y. Sung, S.H. Kim, U.J. Yun, H. Kim, E.J. Jang, H.E. Yoo, S.H. Goh, A. Moon, J.S. Lee, S.K. Ye, J. Shim, Y.N. Kim, Fibronectin regulates anoikis resistance via cell aggregate formation, Cancer Lett. 508 (2021) 59–72.

[16] C.D. Simpson, K. Anyiwe, A.D. Schimmer, Anoikis resistance and tumor metastasis, Cancer Lett. 272 (2008) 177–185.

[17] N.D. Amoedo, M.F. Rodrigues, F.D. Rumjanek, Mitochondria: are mitochondria accessory to metastasis? Int. J. Biochem. Cell Biol. 51 (2014) 53–57.

[18] P. Paoli, E. Giannoni, P. Chiarugi, Anoikis molecular pathways and its role in cancer progression, Biochim. Biophys. Acta 1833 (2013) 3481–3498.

[19] L. Jin, J. Chun, C. Pan, A. Kumar, G. Zhang, Y. Ha, D. Li, G.N. Alesi, Y. Wang, I.L. Zhou, W.M. Yu, K.R. Magliocca, F.R. Khuri, C.K. Qu, C. Metallo, T.K. Owonikoko, S. Kang, The PLAG1-GDH1 Axis promotes anoikis resistance and tumor metastasis through CamKK2-AMPK signaling in LKB1-deficient lung cancer, Mol. Cell 69 (2018) 87–99.e7.

[20] J. Wang, Z. Luo, L. Lin, X. Sui, L. Yu, C. Xu, R. Zhang, Z. Zhao, Q. Zhu, B. An, Q. Wang, B. Chen, E.L. Leung, Q. Wu, Anoikis-associated lung cancer metastasis: mechanisms and therapies, Cancers 14 (2022).

[21] Z. Zhu, C. Fang, H. Xu, L. Yuan, Y. Du, Y. Ni, Y. Xu, A. Shao, A. Zhang, M. Lou, Anoikis resistance in diffuse glioma: the potential therapeutic targets in the future, Front. Oncol. 12 (2022) 976557.

[22] J. Vivian, A.A. Rao, F.A. Nothaft, C. Ketchum, J. Armstrong, A. Novak, J. Pfeil, J. Narkizian, A.D. Deran, A. Musselman-Brown, H. Schmidt, P. Amstutz, B. Craft, M. Goldman, K. Rosenbloom, M. Cline, B. O'Connor, M. Hanna, C. Birger, W.J. Kent, D.A. Patterson, A.D. Joseph, J. Zhu, S. Zaranek, G. Getz, D. Haussler, B. Paten, Toil enables reproducible, open source, big biomedical data analyses, Nat. Biotechnol. 35 (2017) 314–316.

[23] F. Yang, T. Wang, P. Yan, W. Li, J. Kong, Y. Zong, X. Chao, W. Li, X. Zhao, J. Wang, Identification of pyroptosis-related subtypes and establishment of prognostic model and immune characteristics in asthma, Front. Immunol. 13 (2022) 937832.

[24] P. Langfelder, S. Horvath, WGCNA: an R package for weighted correlation network analysis, BMC Bioinf. 9 (2008) 559.

[25] T. Chen, H. Zhang, Y. Liu, Y.X. Liu, L. Huang, EVenn: easy to create repeatable and editable Venn diagrams and Venn networks online, Journal of genetics and genomics = Yi chuan xue bao 48 (2021) 863–866.

[26] A. Mayakonda, D.C. Lin, Y. Assenov, C. Plass, H.P. Koeffler, Maftools: efficient and comprehensive analysis of somatic variants in cancer, Genome Res. 28 (2018) 1747–1756.

[27] M.D. Wilkerson, D.N. Hayes, ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking, Bioinformatics 26 (2010) 1572–1573.

[28] A. Subramanian, P. Tamayo, V.K. Mootha, S. Mukherjee, B.L. Ebert, M.A. Gillette, A. Paulovich, S.L. Pomeroy, T.R. Golub, E.S. Lander, J.P. Mesirov, Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles, Proc. Natl. Acad. Sci. U.S.A. 102 (2005) 15545–15550.

[29] A. Liberzon, A. Subramanian, R. Pinchback, H. Thorvaldsdóttir, P. Tamayo, J.P. Mesirov, Molecular signatures database (MSigDB) 3.0, Bioinformatics 27 (2011) 1739–1740.

[30] D. Maeser, R.F. Gruener, R.S. Huang, oncoPredict: an R package for predicting in vivo or cancer patient drug response and biomarkers from cell line screening data, Briefings Bioinf. 22 (2021).

[31] A.R. Lim, W.K. Rathmell, J.C. Rathmell, The tumor microenvironment as a metabolic barrier to effector T cells and immunotherapy, Elife 9 (2020).

[32] M. Angelova, B. Mlecnik, A. Vasaturo, G. Bindea, T. Fredriksen, L. Lafontaine, B. Buttard, E. Morgand, D. Bruni, A. Jouret-Mourin, C. Hubert, A. Kartheuser, Y. Humblet, M. Ceccarelli, N. Syed, F.M. Marincola, D. Bedognetti, M. Van den Eynde, J. Galon, Evolution of metastases in space and time under immune selection, Cell 175 (2018) 751–765.e16.

[33] D.M. Kelly, L. Li, A.I. Burgess, D.L. Poole, J.M. Duerden, P.M. Rothwell, Associations of blood biomarkers with glomerular filtration rate in patients with TIA and stroke: population-based study, Stroke and vascular neurology 6 (2021) 48–56.

[34] D.A. Sabbah, R. Hajjo, K. Sweidan, Review on epidermal growth factor receptor (EGFR) structure, signaling pathways, interactions, and recent updates of EGFR inhibitors, Curr. Top. Med. Chem. 20 (2020) 815–834.

[35] S. Kumagai, S. Koyama, H. Nishikawa, Antitumour immunity regulated by aberrant ERBB family signalling, Nat. Rev. Cancer 21 (2021) 181–197.

[36] K.B. Runkle, A. Kharbanda, E. Stypulkowski, X.J. Cao, W. Wang, B.A. Garcia, E.S. Witze, Inhibition of DHHC20-mediated EGFR palmitoylation creates a dependence on EGFR signaling, Mol. Cell 62 (2016) 385–396.

[37] Y. Yarden, B.Z. Shilo, SnapShot: EGFR signaling pathway, Cell 131 (2007) 1018.

[38] M.R. Quadros, S. Connelly, C. Kari, M.T. Abrams, E. Wickstrom, U. Rodeck, EGFR-dependent downregulation of Bim in epithelial cells requires MAPK and PKC-delta activities, Cancer Biol. Ther. 5 (2006) 498–504.

[39] C.L. Buchheit, B.L. Angarola, A. Steiner, K.J. Weigel, Z.T. Schafer, Anoikis evasion in inflammatory breast cancer cells is mediated by Bim-EL sequestration, Cell Death Differ. 22 (2015) 1275–1286.

[40] N. Shao, Z. Lu, Y. Zhang, M. Wang, W. Li, Z. Hu, S. Wang, Y. Lin, Interleukin-8 upregulates integrin β3 expression and promotes estrogen receptor-negative breast cancer cell invasion by activating the PI3K/Akt/NF-κB pathway, Cancer Lett. 364 (2015) 165–172.

[41] C. Jiang, Y. Zhu, Z. Zhou, J. Gumin, L. Bengtsson, W. Wu, Z. Songyang, F.F. Lang, X. Lin, TMEM43/LUMA is a key signaling component mediating EGFR-induced NF-κB activation and tumor progression, Oncogene 36 (2017) 2813–2823.

[42] J.P. Kumar, The sine oculis homeobox (SIX) family of transcription factors as regulators of development and disease, Cell. Mol. Life Sci. : CM 66 (2009) 565–583.

[43] G. Chen, Z. Chen, H. Zhao, MicroRNA-155-3p promotes glioma progression and temozolomide resistance by targeting Six1, J. Cell Mol. Med. 24 (2020) 5363–5374.

[44] Z.X. Fang, C.L. Li, Z. Wu, Y.Y. Hou, H.T. Wu, J. Liu, Comprehensive analysis of the potential role and prognostic value of sine oculis homeobox homolog family in colorectal cancer, World J. Gastrointest. Oncol. 14 (2022) 2138–2156.

[45] D. Liu, X.X. Zhang, D.Y. Wan, B.X. Xi, D. Ma, H. Wang, Q.L. Gao, Sine oculis homeobox homolog 1 promotes α5β1-mediated invasive migration and metastasis of cervical cancer cells, Biochem. Biophys. Res. Commun. 446 (2014) 549–554.

[46] C. Vizcaíno, S. Mansilla, J. Portugal, Sp1 transcription factor: a long-standing target in cancer chemotherapy, Pharmacol. Therapeut. 152 (2015) 111–124.

[47] K.A. Ivanenko, V.S. Prassolov, E.R. Khabusheva, [Transcription factor Sp1 in the expression of genes encoding components of MAPK, JAK/STAT, and PI3K/akt signaling pathways], Mol. Biol. 56 (2022) 832–847.

[48] C.S. Mak, M.M. Yung, L.M. Hui, L.L. Leung, R. Liang, K. Chen, S.S. Liu, Y. Qin, T.H. Leung, K.F. Lee, K.K. Chan, H.Y. Ngan, D.W. Chan, MicroRNA-141 enhances anoikis resistance in metastatic progression of ovarian cancer through targeting KLF12/Sp1/survivin axis, Mol. Cancer 16 (2017) 11.

[49] Y. Takeshita, T. Motohara, T. Kadomatsu, T. Doi, K. Obayashi, Y. Oike, H. Katabuchi, M. Endo, Angiopoietin-like protein 2 decreases peritoneal metastasis of ovarian cancer cells by suppressing anoikis resistance, Biochem. Biophys. Res. Commun. 561 (2021) 26–32.

[50] K. Lu, Q. Chen, M. Li, L. He, F. Riaz, T. Zhang, D. Li, Programmed cell death factor 4 (PDCD4), a novel therapy target for metabolic diseases besides cancer, Free Radic. Biol. Med. 159 (2020) 150–163.

[51] M.Y. Zhao, L.M. Wang, J. Liu, X. Huang, J. Liu, Y.F. Zhang, MiR-21 suppresses anoikis through targeting PDCD4 and PTEN in human esophageal adenocarcinoma, Current medical science 38 (2018) 245–251.

[52] T.T. Li, Y.W. Lai, X. Han, X. Niu, P.X. Zhang, BMP2 as a promising anticancer approach: functions and molecular mechanisms, Invest. N. Drugs 40 (2022) 1322–1332.

[53] Z. Shonibare, M. Monavarian, K. O'Connell, D. Altomare, A. Shelton, S. Mehta, R. Jaskula-Sztul, R. Phaeton, M.D. Starr, R. Whitaker, A. Berchuck, A.B. Nixon, R. C. Arend, N.Y. Lee, C.R. Miller, N. Hempel, K. Mythreye, Reciprocal SOX2 regulation by SMAD1-SMAD3 is critical for anoikis resistance and metastasis in cancer, Cell Rep. 40 (2022) 111066.