



Prognostic implications of immune-related eight-gene signature in pediatric brain tumors

Yi Wang^{1*}, Chuan Zhou^{2*}, Huan Luo³, Jing Cao⁴, Chao Ma³, Lulu Cheng⁵, and Yang Yang⁵

¹Department of Neonatology and Neonatal Intensive Care, Zhumadian Central Hospital, Zhumadian, China

²Neonatal Intensive Care Unit, The Second Affiliated Hospital of Zhengzhou University, Zhengzhou, China

³Charité - Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and the Berlin Institute of Health, Berlin, Germany

⁴Department of Anatomy, College of Basic Medicine, Zhengzhou University, Zhengzhou, China

⁵Digital Medical Laboratory, Zhumadian Central Hospital, Zhumadian, China

Abstract

Genomic studies have provided insights into molecular subgroups and oncogenic drivers of pediatric brain tumors (PBT) that may lead to novel therapeutic strategies. Participants of the cohort Pediatric Brain Tumor Atlas: CBTTTC (CBTTTC cohort), were randomly divided into training and validation cohorts. In the training cohort, Kaplan-Meier analysis and univariate Cox regression model were applied to preliminary screening of prognostic genes. The LASSO Cox regression model was implemented to build a multi-gene signature, which was then validated in the validation and CBTTTC cohorts through Kaplan-Meier, Cox, and receiver operating characteristic curve (ROC) analyses. Also, gene set enrichment analysis (GSEA) and immune infiltrating analyses were conducted to understand function annotation and the role of the signature in the tumor microenvironment. An eight-gene signature was built, which was examined by Kaplan-Meier analysis, revealing that a significant overall survival difference was seen, either in the training or validation cohorts. The eight-gene signature was further proven to be independent of other clinic-pathologic parameters via the Cox regression analyses. Moreover, ROC analysis demonstrated that this signature owned a better predictive power of PBT prognosis. Furthermore, GSEA and immune infiltrating analyses showed that the signature had close interactions with immune-related pathways and was closely related to CD8 T cells and monocytes in the tumor environment. Identifying the eight-gene signature (*CBX7*, *JADE2*, *IGF2BP3*, *OR2W6P*, *PRAME*, *TICRR*, *KIF4A*, and *PIMREG*) could accurately identify patients' prognosis and the signature had close interactions with the immunodominant tumor environment, which may provide insight into personalized prognosis prediction and new therapies for PBT patients.

Key words: Pediatric brain tumor; Gene signature; Overall survival; Prognosis; Biomarkers

Introduction

Brain tumors are the most common solid tumor in pediatrics, accounting for 23.7% of new cancer diagnoses in children (1), and the second most common pediatric malignancy after leukemia (2,3). The life-saving treatments these children receive may result in impaired brain structure and function, leading to long-term major cognitive deficits (1,4–8). Current treatment options include surgical resection, cranial radiation, and chemotherapy. Survivors treated with high-load cranial radiation likely experience cognitive dysfunction, including difficulties in controlled attention, such as response inhibition and slower information processing (1,4–8). The distribution,

pathology, molecular characteristics, and treatment strategies for pediatric brain tumors (PBT) have essential differences compared to those of the adult population (9).

Although the cure rate of PBT has increased in the past two decades of the 20th century, which was primarily due to advances in imaging, neurosurgery, and radiation oncology technologies and the introduction of combined chemotherapy, unfortunately, the overall survival has remained static (10). The lack of advances in PBT treatment was hindered by our lack of knowledge about the molecular pathogenesis of brain tumors (10). Advanced genomic analysis of the entire spectrum of PBT heralds an

Correspondence: Chao Ma: <chao.ma@charite.de> | Lulu Cheng: <55526649@qq.com> | Yang Yang: <yangyang19800627@163.com>

*These authors contributed equally to this work.

Received September 3, 2020 | Accepted February 4, 2021

era in which this defect can be overcome by new technologies that will help us understand the genome pattern of PBT (10). Genomic studies have provided insights into molecular subgroups and oncogenic PBT drivers that may lead to novel therapeutic strategies (11).

The Gabriella Miller Kids First Data Resource Center (Kids First DR; <<https://kidsfirstdrc.org>>) is a new, collaborative, pediatric research effort to understand the genetic causes and links between childhood cancer and structural congenital disabilities. This is a brand-new public database that has only been launched in recent years. It is vital that this data is expressly set up for children's tumor research, and few researchers have begun to mine this database.

Here, we conducted a comprehensive mining of the Kids First database to determine the minimum number of potentially robust genes that can be used to predict PBT patients' prognosis. Importantly, we used the LASSO algorithm, which can effectively analyze high-dimensional sequencing data (12). Furthermore, we evaluated the accuracy of this eight-gene signature and validated it in a validation cohort. Moreover, gene set enrichment analysis (GSEA) and immune infiltrating analyses were conducted to explore the role of the signature in the tumor microenvironment.

Material and Methods

Data mining from the Kids First program

The gene expression profiles of PBT from 973 patients and their clinical and survival data were downloaded from Kids First Xena Hub (<<https://kidsfirst.xenahubs.net>>) with the cohort name: Pediatric Brain Tumor Atlas: CBTTTC (CBTTTC cohort). The study was conducted following the Declaration of Helsinki, and the Ethical Committee of Zhumadian Central Hospital approved the study.

Identification and validation of the prognostic gene signature

First, all patients in the CBTTTC cohort were randomly divided into training cohort (486 cases, 49.9%) and validation cohort (487 cases, 50.1%). In the training cohort, Kaplan-Meier analysis was used to screen the prognostic genes with a cutoff of P value <10.0E-15 in the log-rank test. Furthermore, univariate Cox regression analysis was performed on the training cohort to find prognostic genes with P values <10.0E-15. The intersected genes identified in Kaplan-Meier and univariate Cox analyses were then entered into the LASSO Cox regression model analysis, which was implemented in the training cohort utilizing R software (<www.r-project.org>) and the "glmnet" package (<cran.r-project.org>). Ten-time cross-validations were applied to detect the best penalty parameter lambda (12–15). According to the best lambda value, a list of prognostic genes with coefficients was obtained from the gene expression and patients' survival data.

Moreover, each patient's risk score was calculated based on the expression level of each prognostic gene and its corresponding coefficient. Using the median risk score as the cutoff point, the patients in the training cohort were distributed into high-risk or low-risk groups. Kaplan-Meier analysis was applied to evaluate the survival difference between the two groups. Cox and ROC analyses were conducted to further assess the gene signature's prognostic value in the training cohort. Furthermore, the prognostic gene signature was validated in the validation cohort. The same formula was conducted to compute risk scores like those in the training cohort. Kaplan-Meier, Cox, and ROC analyses were carried out as described earlier.

Gene set enrichment analysis

The Hallmark (v7.1) and C7 (v7.1) gene set collections were downloaded from the Molecular Signatures Database v7.1 download page (<https://www.gsea-msigdb.org/gsea/downloads.jsp>). GSEA was performed based on the downloaded gene set collections using GSEA software (v4.0.3, <https://www.gsea-msigdb.org/>). The entire CBTTTC cohort was taken to GSEA to reveal the functions and pathways in the differentially expressed genes between high-risk and low-risk groups. Only gene sets with family-wise error rate (FWER) P values <0.05 were considered significant.

Correlation of risk score with the proportion of tumor-infiltrating immune cells (TICs)

The CIBERSORT (16) and MCP-counter (17) methods were used to estimate all tumor samples' TIC abundance distribution in the CBTTTC cohort. The correlation was examined by the Spearman method.

Statistical analysis

All statistical calculations were performed in the R software. Kaplan-Meier analysis was conducted to determine the overall survival differences between groups. Univariate and multivariate Cox proportional hazard regression analyses were conducted to assess the association between risk score and overall survival. The ROC analysis was applied to examine the sensitivity and specificity of survival prediction using the gene signature risk score. An area under the ROC curve (AUC) served as a pointer of prognostic accuracy. The R package "pROC" was used for ROC analysis, and the "DeLong" method was used to study the significant differences among ROC curves. In all analyses, a P value <0.05 was considered statistically significant.

Results

Clinical characteristics

The flowchart of the present research is shown in Figure 1. A total of 973 cases in the CBTTTC cohort were

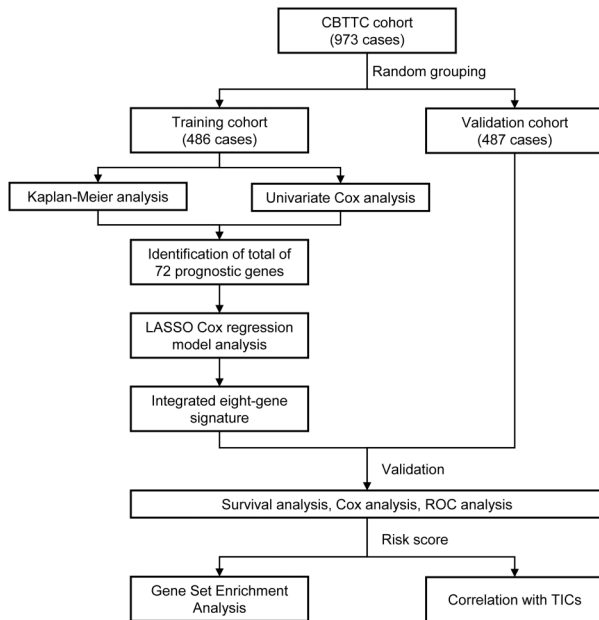


Figure 1. Flow chart of the study. The study was carried out in Pediatric Brain Tumor Atlas (CBTTC) cohort. The training cohort was used to identify prognostic genes. The LASSO regression model was used to establish a prognostic signature based on the prognostic genes. The prognosis analysis was validated in the validation cohort. LASSO: least absolute shrinkage and selection operator Cox regression model; ROC: receiver operating characteristic; TICs: tumor-infiltrating immune cells.

randomly distributed to a training cohort (N=486) and a validation cohort (N=487). The detailed clinical characteristics and grouping of these patients are summarized in Table 1.

Construction of prognostic signature from the training cohort

Ninety-two genes were extracted from the Kaplan-Meier analysis (Table S1), while 136 genes were identified as significant in the Cox regression analysis (Table S2). Taken together, 72 genes in the intersection of the two results are defined as prognostic genes for subsequent analyses (Table S3). These prognostic genes were then subjected to LASSO Cox regression analysis, and regression coefficients were calculated. The coefficient of each gene is illustrated in Figure 2A. When 8 genes were included, the model achieved the best performance (Figure 2B). These genes and corresponding coefficients are shown in Table 2.

Prognostic value of the eight-gene signature in the cohorts

The distribution of risk scores and survival status and the expression profiles of the eight-gene signature of the patients in the training cohort were plotted and are shown

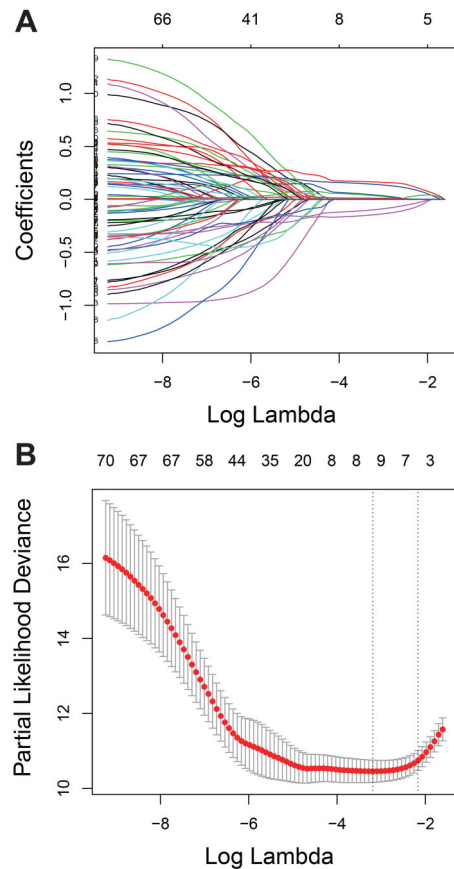


Figure 2. Establishment of prognostic gene signature by LASSO regression analysis. **A**, LASSO coefficient profiles of the 72 genes in the training cohort. **B**, Coefficient profile plot was generated against the log (lambda) sequence. Selection of the optimal parameter (lambda) in the LASSO model for training cohort. LASSO: least absolute shrinkage and selection operator Cox regression model.

in Figure 3A. As shown in the figure, there were more deaths in the high-risk patient group, and the survival time was shorter than that of the low-risk patient group. The heatmap indicates that *CBX7* and *JADE2* were under-expressed in high-risk patients, while *IGF2BP3*, *OR2W6P*, *PRAME*, *TICRR*, *KIF4A*, and *PIMREG* were highly expressed in high-risk patients. We also verified the performance of this eight-gene signature in the validation cohort. As shown in Figure 3B and C, the pattern was consistent with that in the training cohort. Furthermore, we examined this eight-gene signature in subtypes in the CBTTC cohort and the results were similar (Figure S1).

Kaplan-Meier survival analysis showed that patients in the high-risk group were associated with a poor prognosis trend in the training cohort (P value < 0.0001, Figure 4A). To confirm the efficacy of the eight-gene signature in predicting overall survival in PBT patients, it was examined in

Table 1. Clinical characteristics of 973 pediatric brain tumor patients involved in the study.

Characteristics	CBTTC cohort (N=973)	Training cohort (N=486)	Validation cohort (N=487)
Age at diagnosis, years			
< 10	449 (46.14%)	209 (43.00%)	240 (49.28%)
≥ 10	381 (39.16%)	188 (38.69%)	193 (39.63%)
Unknown	143 (14.70%)	89 (18.31%)	54 (11.09%)
Sex			
Female	379 (38.95%)	189 (38.89%)	190 (39.01%)
Male	451 (46.35%)	208 (42.80%)	243 (49.90%)
Unknown	143 (14.70%)	89 (18.31%)	54 (11.09%)
Race			
American Indian or Alaskan Native	13 (1.34%)	5 (1.03%)	8 (1.64%)
Asian	28 (2.88%)	16 (3.29%)	12 (2.46%)
Black or African American	79 (8.12%)	32 (6.58%)	47 (9.65%)
More than one race	7 (0.71%)	3 (0.62%)	4 (0.82%)
Native Hawaiian or other Pacific Islander	2 (0.20%)	2 (0.41%)	0 (0%)
White	701 (72.05%)	339 (69.76%)	362 (74.34%)
Unknown	143 (14.70%)	89 (18.31%)	54 (11.09%)
Ethnicity			
Hispanic or Latino	28 (2.88%)	19 (3.91%)	9 (1.85%)
Not Hispanic or Latino	802 (82.42%)	378 (77.78%)	424 (87.06%)
Unknown	143 (14.70%)	89 (18.31%)	54 (11.09%)
Histological subtype			
Ependymoma	93 (9.56%)	44 (9.05%)	49 (10.06%)
Medulloblastoma	120 (12.33%)	57 (11.73%)	63 (12.94%)
Low-grade glioma/astrocytoma	252 (25.90%)	135 (27.78%)	117 (24.02%)
High-grade glioma/astrocytoma	109 (11.20%)	42 (8.64%)	67 (13.76%)
Others	399 (41.01%)	208 (42.80%)	191 (39.22%)

Data are reported as number and percent.

Table 2. Genes in the prognostic gene signatures.

Gene symbol	Full name	Risk coefficient
<i>CBX7</i>	Chromobox 7	-0.097248966
<i>IGF2BP3</i>	Insulin Like Growth Factor 2 mRNA Binding Protein 3	0.141763903
<i>JADE2</i>	Jade Family PHD Finger 2	-0.144454955
<i>KIF4A</i>	Kinesin Family Member 4A	0.041651464
<i>OR2W6P</i>	Olfactory Receptor Family 2 Subfamily W Member 6 Pseudogene	0.034386859
<i>PIMREG</i>	PICALM Interacting Mitotic Regulator	0.004196815
<i>PRAME</i>	Preferentially Expressed Antigen in Melanoma	0.010705036
<i>TICRR</i>	TOPBP1 Interacting Checkpoint and Replication Regulator	0.164511862

the validation cohort. According to the median risk score, patients were divided into high-risk and low-risk groups using the same classification method. Consistent with previous results, patients in the high-risk group showed significantly worse overall survival than patients in the low-risk group (P value <0.0001 , Figure 4B). In the entire CBTTC cohort, which was the sum of the training and validation cohorts, the eight-gene signature also had similar predictive ability (P value <0.0001 , Figure 4C). We tested the capacity of each of the eight genes via Kaplan-Meier analysis and found *CBX7* and *JADE2* predicted favorable outcomes, while the remaining genes

had poor effects on the prognosis (Figure S2). Moreover, the subtypes of PBT in the CBTTC cohort were also examined by Kaplan-Meier analysis, which showed that the eight-gene signature risk score predicted the survival of PBT subtypes (Figure S3).

Univariate and multivariate Cox analyses were performed in the training, validation, and CBTTC cohorts, using the available co-variables including risk score, age, gender, race, and ethnicity to confirm whether the prognostic capacity of our eight-gene signature was independent from the clinic-pathologic characteristics. In the training cohort, both univariate and multivariate Cox

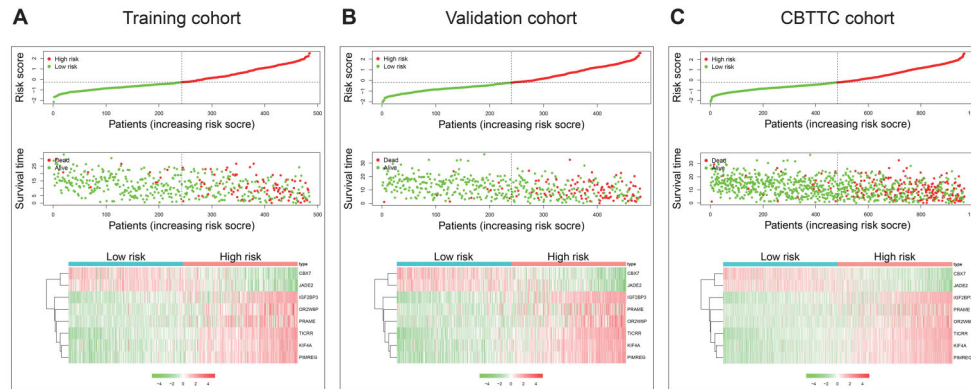


Figure 3. Characteristics of the eight-gene signature. Upper and middle, the distribution of risk score and patient’s survival time for training (A), validation (B), and Pediatric Brain Tumor Atlas (CBTTC) (C) cohorts. The black dotted line is the median cutoff dividing patients into low-risk and high-risk groups. Bottom, heatmap of the eight-gene expression profiles for prognostic signature for training, validation, and CBTTC cohorts.

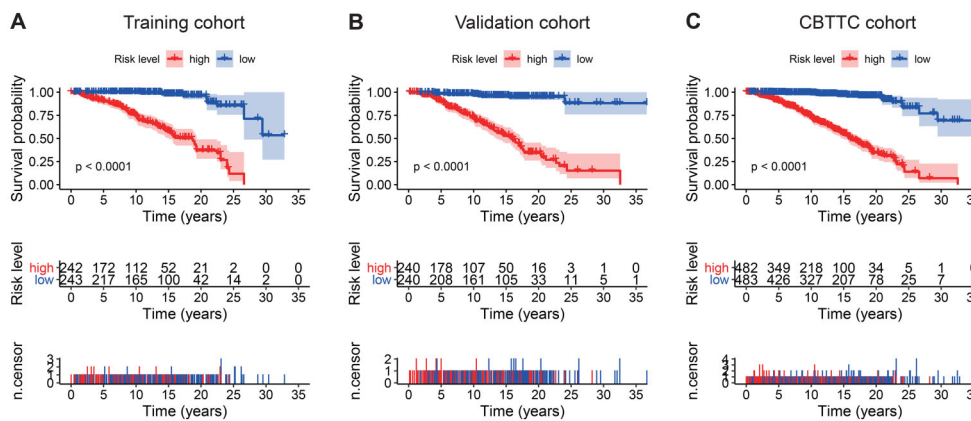


Figure 4. Kaplan-Meier analyses of the overall survival based on the eight-gene signature. **A**, Training cohort. **B**, Validation cohort. **C**, Pediatric Brain Tumor Atlas (CBTTC) cohort. The differences between the two curves were determined by the two-side log-rank test. The number of patients at risk are listed in the middle plot of each cohort.

regression analyses indicated that the eight-gene signature was a powerful variable associated with prognosis (HR=3.399, 95%CI: 2.713–4.258, P value < 0.001 and HR=3.135, 95%CI: 2.458–3.999, P value < 0.001, respectively; Figure 5A). Consistent with that in the training cohort, the eight-gene signature displayed a similar ability in the validation and CBTTC cohorts (Figure 5B and C). These results proved that the eight-gene signature was a strong and independent variable.

Subsequently, we conducted ROC analyses to assess how the eight-gene signature would behave in predicting prognosis. As shown in Figure 6A, the area under the ROC curve (AUC) of the eight-gene risk score model performed in the training cohort was 0.808, which was superior to those of age, gender, race, and ethnicity (0.492, 0.494, 0.482, and 0.514, respectively). This finding was also confirmed in the validation cohort (AUC=0.844, Figure 6B) and CBTTC cohort (AUC= 0.827, Figure 6C).

Gene set enrichment analysis with the eight-gene signature

Given the negative correlation between the eight-gene signature risk score and the overall survival of PBT patients, GSEA was conducted between the high- and low-risk groups. As displayed in Figure 7A and Table S4, genes in the high-risk group were mostly enriched in immune-related functions and pathways. They were relating to regulatory T cells, macrophages, CD4 T cells, TGF beta, IL6, and naive B cells. As to the low-risk score group, the genes were enriched in pathways involved in macrophages, CD4 T cells, T helper type 2 cells, and FOXP3+ regulatory T cells, which were also closely immune-related (Figure 7A and Table S5). For HALLMARK collection defined by the Molecular Signatures Database, multiple immune functional gene sets like HALLMARK_MITOTIC_SPINDLE, HALLMARK_G2M_CHECK POINT, and HALLMARK_E2F_TARGETS were significantly

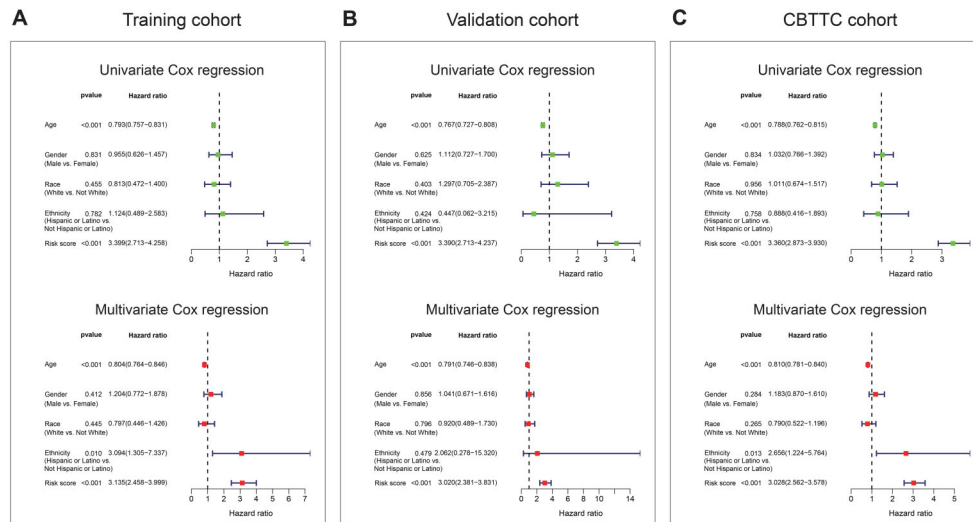


Figure 5. Forest plot summary of overall survival analyses. Univariate (upper) and multivariate (bottom) analyses based on the eight-gene signature and clinical covariates in training (A), validation (B), and Pediatric Brain Tumor Atlas (CBTTTC) (C) cohorts. The colored solid squares represent the hazard ratio (HR), and the transverse lines through HR represent 95% confidence intervals (CI). All P values were calculated using Cox hazards regression analysis.

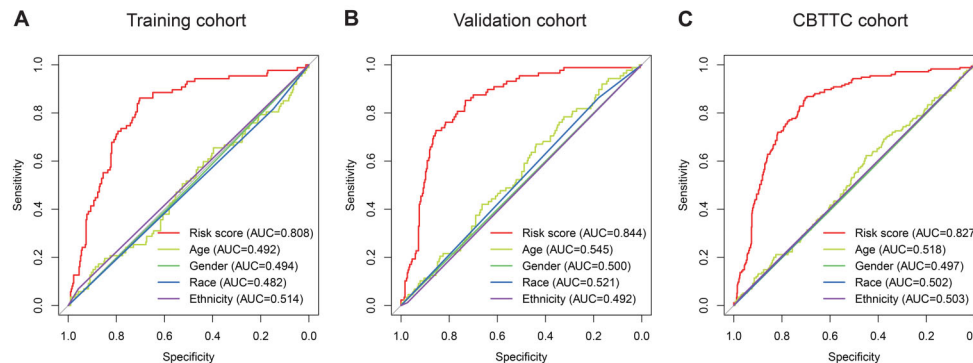


Figure 6. Receiver operating characteristic (ROC) analysis of the eight-gene signature risk score. ROC analysis of the sensitivity and specificity of the overall survival prediction by the eight-gene risk score, age, gender, race, and ethnicity in the training (A), validation (B), and Pediatric Brain Tumor Atlas (CBTTTC) (C) cohorts. AUC: area under the ROC curve.

enriched in the high-risk group (Figure 7B and Table S6); whereas only one gene set HALLMARK_ADIPOGENESIS was enriched in the low-risk group (Figure 7B and Table S7). These findings indicated that the risk score was potentially closely related to the status of tumor microenvironment.

Correlation of risk score with the proportion of TICs

To further confirm the correlation between the risk score and the immune microenvironment, as shown in Figure S4 and Figure S5, we used the CIBERSORT and MCP-counter algorithm to analyze the proportion of TICs subpopulations and constructed immune cell profiles in PBT samples. Combining the results of correlation

analysis (Figure 8A) and differential analysis (Figure 8B), a total of 11 TICs were associated with the eight-gene signature risk score (Figure 8C). In the result of the MCP-counter algorithm (Figure 9), a list containing 7 TICs identified closely with the signature. In summary, by adopting these methods, CD8 T cells and monocytes were overlapped in the two results and seen as critical cells that affected the eight-gene signature in the tumor environment of PBT.

Discussion

At present, diagnosis, prognosis, and treatment of PBT are highly dependent on the histopathological

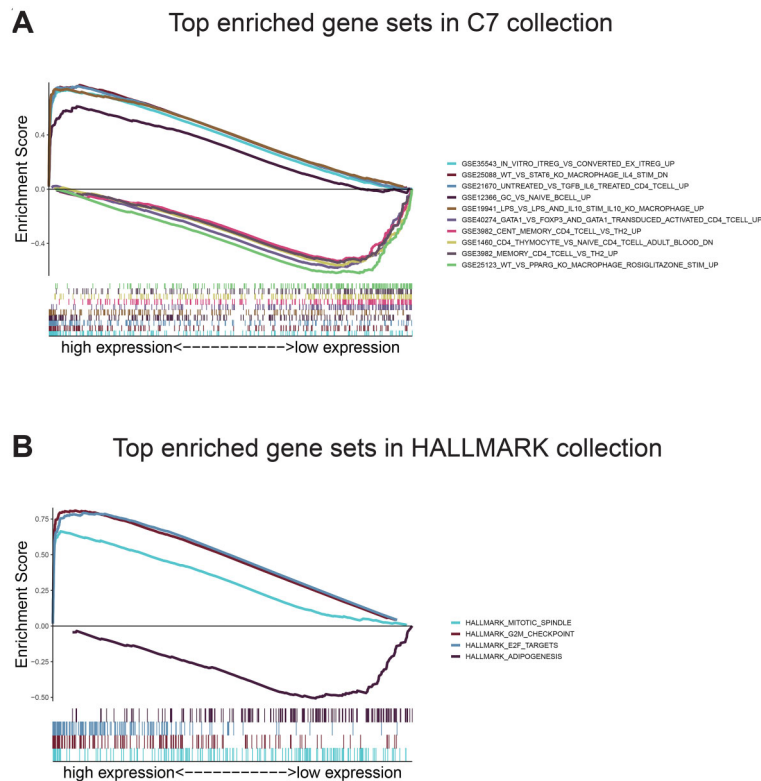


Figure 7. Gene Set Enrichment Analysis. **A**, Top ten enriched gene sets annotated by the C7 collection between the high- and low-risk groups in the CBTTTC cohort. **B**, Enriched gene sets annotated by the HALLMARK collection between the high- and low-risk groups in the CBTTTC cohort. Each line represents one specific gene set with unique color. Up-regulated genes are located at the left, approaching the origin of the coordinates, and the down-regulated are located at the right of the x-axis. Only gene sets with family-wise error rate (FWER) P values <0.05 were considered significant.

characteristics of the tumor (3,18). However, more importantly, given the current development of precision medicine and genetic research of tumors, in the past decade, significant changes have taken place in pediatric neuro-oncology, and exploring optimized tumor biomarkers will be the trend of future development (19,20). In this study, we built a PBT prognostic signature by comprehensively analyzing the Kids First database, designed to understand the genetic causes and connections of childhood cancer and congenital structural disabilities.

After we constructed the eight-gene signature, we firstly confirmed its capacity to distinguish the survival time and survival status of patients effectively. As shown in Figure 3A, the high-risk zone not only counted more deaths, but also the patients in it presented a shorter survival time than those in the low-risk zone. Moreover, the heatmap indicated that each of these eight genes had a differential expression pattern between the low- and high-risk groups. Importantly, this eight-gene signature had the same or similar performance in the validation cohort (Figure 3B) and the CBTTTC cohort (Figure 3C).

In addition, we examined the prognostic value of the eight-gene signature by Kaplan-Meier analysis in training, validation, and CBTTTC cohorts, finding its predicting ability in PBT patients significant (Figure 4). Furthermore, univariate and multivariate analyses were performed in the three cohorts to confirm whether our eight-gene signature could be independent from other variables in predicting PBT overall survival. As plotted in Figure 5, no matter in which cohort, whether it was univariate or multivariate Cox regression analysis, the variable of risk score was always statistically significant, and the results confirmed the predictive ability of the risk score, and its independence.

To further assess the predictive power of this eight-gene signature, ROC analysis was conducted. In diagnostic tests, AUC is used to check accuracy and determine the predictive capacity of biomarkers (21). ROC analysis indicated that the AUC of the eight-gene signature stayed above 0.8 in these three cohorts and was superior to age, gender, race, and ethnicity. These ROC results again suggested that our signature strengthened the predictive accuracy of prognosis in PBT.

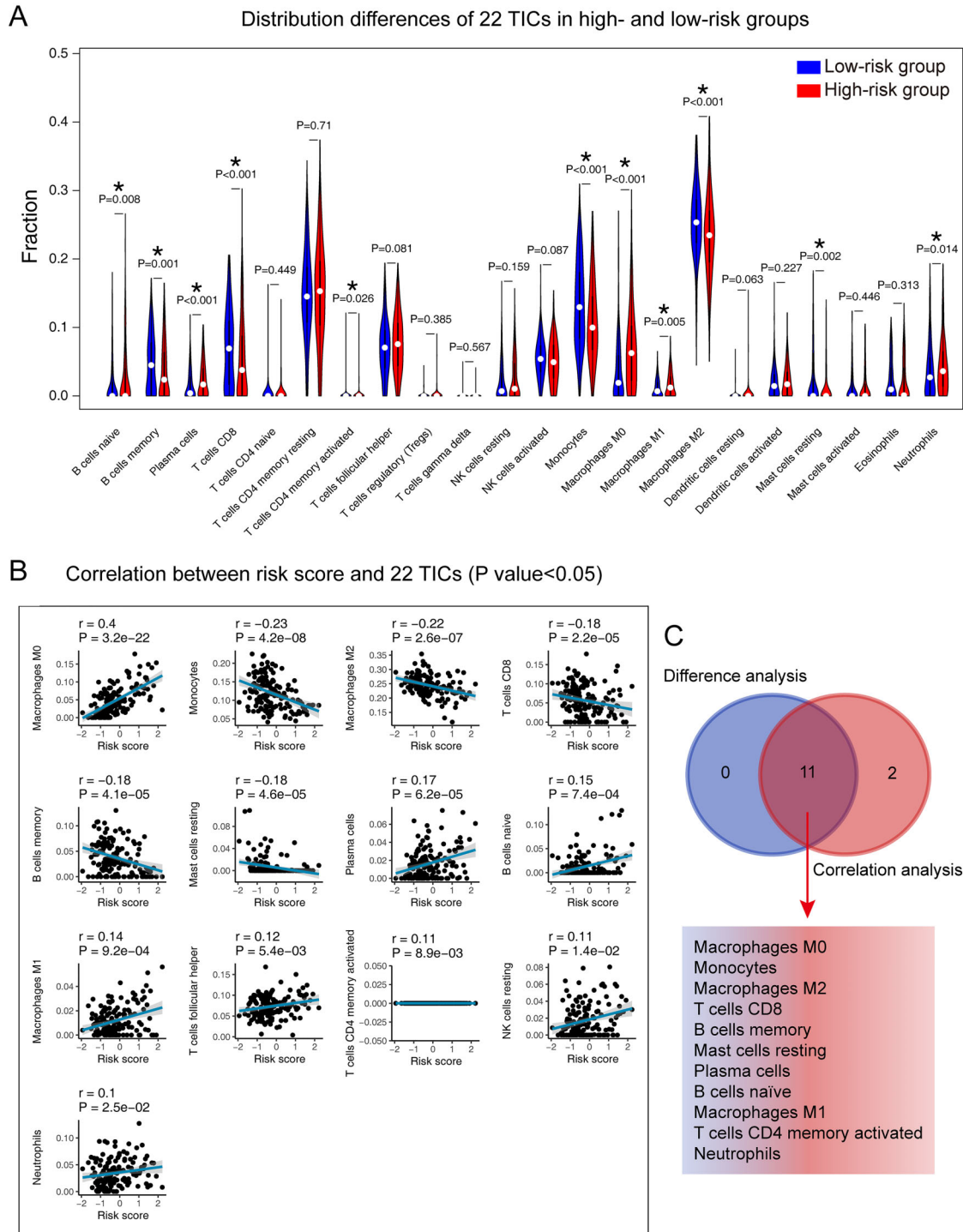


Figure 8. Correlation of TICs proportion with the eight-gene signature risk score in the Pediatric Brain Tumor Atlas (CBTTC) cohort (CIBERSORT method). **A**, Violin plot showing the ratio differentiation of 22 kinds of immune cells between PBT samples from low- and high-risk groups to the median risk score. Wilcoxon rank sum test was used to assess for significance. **B**, The blue line in each plot is the fitted linear model indicating the proportion tropism of the immune cell with risk score. The shade around the blue line represents the 95% confidence interval. Spearman coefficient was used for the correlation test. **C**, Venn plot displayed 11 TICs correlated with risk score co-determined by difference and correlation tests shown in violin and scatter plots, respectively. P value < 0.05 was the cutoff. TIC: tumor-infiltrating immune cell; PBT: pediatric brain tumor.

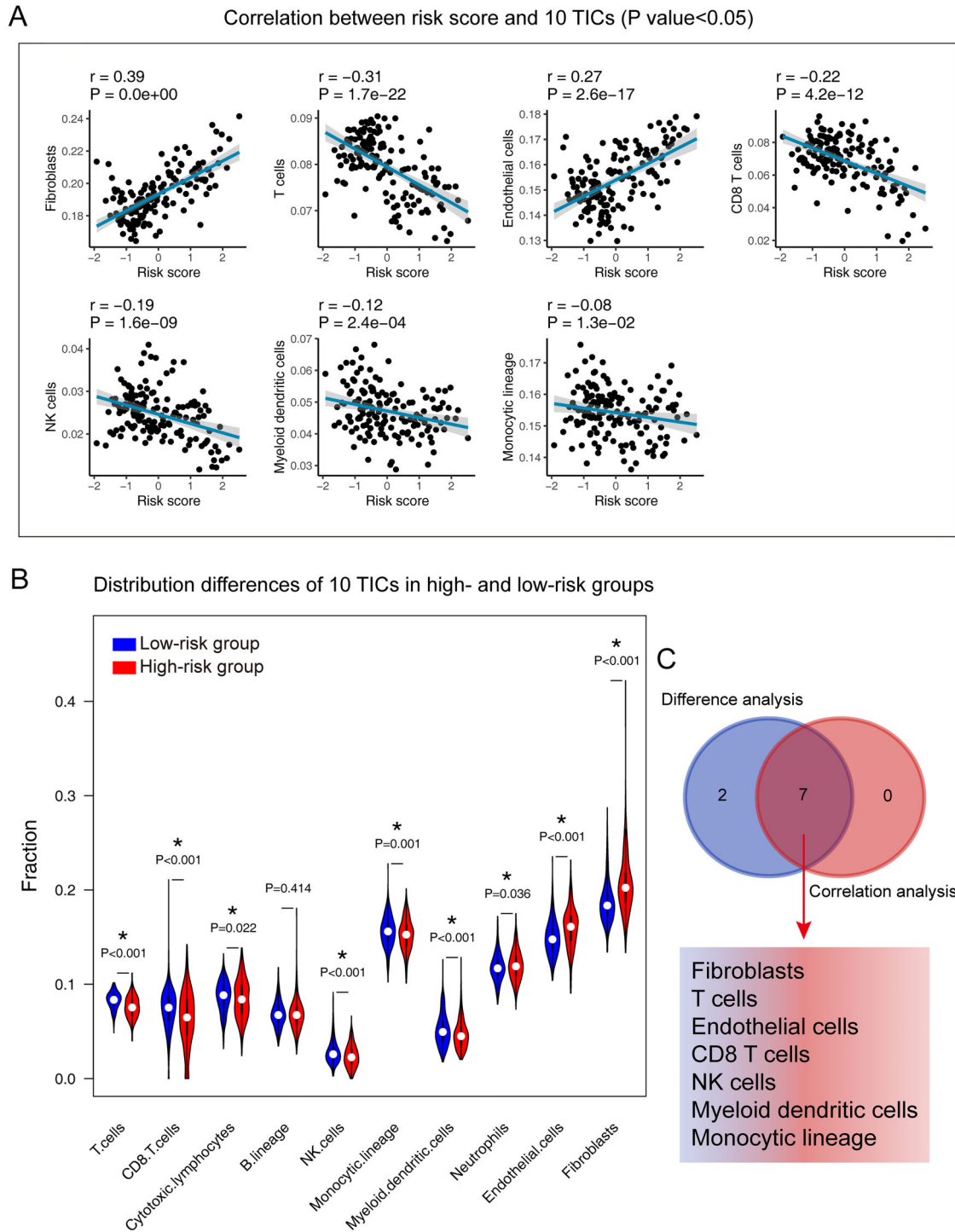


Figure 9. Correlation of TICs proportion with the eight-gene signature risk score in the Pediatric Brain Tumor Atlas (CBTTC) cohort (MCP-counter method). **A**, The blue line in each plot is the fitted linear model indicating the proportion tropism of the immune cell with risk score. The shade around the blue line represents the 95% confidence interval. Spearman coefficient was used for the correlation test. **B**, Violin plot showing the ratio differentiation of 10 types of immune cells between PBT samples from low- and high-risk groups to the median risk score. Wilcoxon rank sum test was used to assess for significance. **C**, Venn plot displays 7 TICs correlated with risk score co-determined by difference and correlation tests shown in violin and scatter plots, respectively. P value <0.05 was the cutoff. TIC: tumor-infiltrating immune cell; PBT: pediatric brain tumor.

Our signature was composed of eight genes, which are *CBX7*, *JADE2*, *IGF2BP3*, *OR2W6P*, *PRAME*, *TICRR*, *KIF4A*, and *PIMREG*. In the signature model, *CBX7* and *JADE2* were protective genes, whereas other genes were unfavorable on the overall survival of PBT patients. *IGF2BP3* is an oncofetal protein that binds RNA, thereby influencing the fate of target transcripts, and it is up-expressed in a variety of malignant tumors and represents a promising cancer biomarker (22). *PRAME* is a tumor-associated antigen that was first identified through analysis of the specificity of tumor-reactive T-cell clones derived from a patient with metastatic cutaneous melanoma (23). Subsequently, it was found that *PRAME* is not only expressed in cutaneous melanoma, but also in ocular melanoma and various non-melanoma cell malignancies (24). *PRAME* expression can be detected in 82% of medulloblastoma samples, regardless of molecular and histopathological subgroups. The high expression of *PRAME* is also related to poor outcomes of patients with medulloblastoma. Studies have shown that adoptive immunotherapy that redirects T cells to *PRAME* antigen may represent an innovative treatment for medulloblastoma (25). *TICRR* is one of the important replication initiation factors. The knockout of *TICRR* significantly inhibits tumor cell growth, migration, and colony formation *in vitro*, and inhibits tumor growth in xenograft models (26). A recent study demonstrated that *TICRR* is a major carcinogen, which can accelerate the proliferation of cancer cells by promoting the initiation and progression of DNA replication (27). *KIF4A* was found to be implicated in the regulation of chromosome condensation and segregation during mitotic cell division, which is essential for eukaryotic cell proliferation (28). *KIF4A* is aberrantly expressed in a variety of cancers, and it is overexpressed in most tumors but also low-expressed in a few tumors, suggesting distinct functions and mechanisms for different tumors (29,30). *PIMREG* plays a key role in regulating cell proliferation and is induced by mitogens, and its protein level is related to the cell cycle (31). Jiang et al. found that *PIMREG* played a crucial role in the promotion of breast cancer aggressiveness via constitutive activation of the NF- κ B signaling pathway (32). There are relatively few studies related to these genes and PBT. However, the eight-gene signature had a significant role in predicting and diagnosing PBT in our research. The eight-gene signature or each of them may indicate specific directions for future research on PBT.

The findings of the GSEA analysis demonstrated that the eight-gene signature might potentially participate in the immune-dominant tumor microenvironment. The analysis based on CIBERSORT algorithm for the proportion of TICs showed that half of TICs (11/22) were correlated with the

eight-gene signature risk score in PBT patients, further supporting that the signature interacted closely with the tumor environment. Combining the CIBERSORT and MCP-counter methods, we found CD8 T cells and monocytes were in a close relationship with the eight-gene signature in the tumor environment of PBT. Strategies targeting the tumor microenvironment of pediatric brain cancers have the potential to improve the efficacy of standard and genome-based molecular therapeutics and to help resolve many of the challenges associated with developing new drugs and running clinical trials for a relatively small PBT population (33). The specific pathways and TICs revealed in our analysis have potential for tumor microenvironment targets in further studies.

Our research also had some limitations. For the study of PBT, currently available public databases are very limited. The datasets in GEO and TCGA are not eligible for validation purposes because of the population's age distribution. Our eight-gene signature came from retrospective data, and more prospective data is needed for proving the clinical utility of it. In addition, due to the very limited clinical characteristics of patients included in the CBTTTC cohort, we could not perform certain clinical subgroup analyses. Furthermore, there are currently no wet experimental data explaining the relationship between these eight genes and their mechanism in PBT samples. Therefore, more research is needed to clarify the potential relationship.

In conclusion, our research defined a robust eight-gene signature in PBT. It was a comprehensive analysis of the new Kids First database. This signature was related to PBT's overall survival and accurately identified the prognostic risk of patients. Notably, we assessed the reliability and accuracy of the signature in a validation cohort. In addition, the functions and immune infiltrating analyses showed that the signature had close interactions with CD8 T cells and monocytes in the tumor environment, which may advance the development of new therapies for PBT treatment.

Supplementary Material

[Click to view \[pdf\].](#)

Acknowledgments

Chao Ma and Huan Luo thank Zhengzhou University Overseas Virtual Research Institute. Chao Ma thanks China Scholarship Council (No. 201708410121). We acknowledge support from the German Research Foundation (DFG) and the Open Access Publication Fund of Charité - Universitätsmedizin Berlin.

References

- Cox E, Bells S, Timmons BW, Laughlin S, Bouffet E, de Medeiros C, et al. A controlled clinical crossover trial of exercise training to improve cognition and neural communication in pediatric brain tumor survivors. *Clin Neurophysiol* 2020; 131: 1533–1547, doi: 10.1016/j.clinph.2020.03.027.
- Cacciotti C, Fleming A, Ramaswamy V. Advances in the molecular classification of pediatric brain tumors: a guide to the galaxy. *J Pathol* 2020; 251: 249–261, doi: 10.1002/path.5457.
- AlRayahi J, Zapotocky M, Ramaswamy V, Hanagandi P, Branson H, Mubarak W, et al. Pediatric brain tumor genetics: what radiologists need to know. *Radiographics* 2018; 38: 2102–2122, doi: 10.1148/rg.2018180109.
- Dockstader C, Wang F, Bouffet E, Mabbott DJ. Gamma deficits as a neural signature of cognitive impairment in children treated for brain tumors. *J Neurosci* 2014; 34: 8813–8824, doi: 10.1523/JNEUROSCI.5220-13.2014.
- Dockstader C, Gaetz W, Bouffet E, Tabori U, Wang F, Bostan SR, et al. Neural correlates of delayed visual-motor performance in children treated for brain tumours. *Cortex* 2013; 49: 2140–2150, doi: 10.1016/j.cortex.2012.09.004.
- Wolfe KR, Madan-Swain A, Kana RK. Executive dysfunction in pediatric posterior fossa tumor survivors: a systematic literature review of neurocognitive deficits and interventions. *Dev Neuropsychol* 2012; 37: 153–175, doi: 10.1080/87565641.2011.632462.
- Mabbott DJ, Snyder JJ, Penkman L, Witol A. The effects of treatment for posterior fossa brain tumors on selective attention. *J Int Neuropsychol Soc* 2009; 15: 205–216, doi: 10.1017/S1355617709090249.
- Palmer SL, Reddick WE, Gajjar A. Understanding the cognitive impact on children who are treated for medulloblastoma. *J Pediatr Psychol* 2007; 32: 1040–1049, doi: 10.1093/jpepsy/jsl056.
- Collins KL, Pollack IF. Pediatric low-grade gliomas. *Cancers (Basel)* 2020; 12: 1152, doi: 10.3390/cancers12051152.
- Gajjar A, Bowers DC, Karajannis MA, Leary S, Witt H, Gottardo NG. Pediatric brain tumors: innovative genomic information is transforming the diagnostic and clinical landscape. *J Clin Oncol* 2015; 33: 2986–2998, doi: 10.1200/JCO.2014.59.9217.
- Brabetz S, Leary SES, Grobner SN, Nakamoto MW, Seker-Cin H, Girard EJ, et al. A biobank of patient-derived pediatric brain tumor models. *Nat Med* 2018; 24: 1752–1761, doi: 10.1038/s41591-018-0207-3.
- Tibshirani R. The lasso method for variable selection in the Cox model. *Stat Med* 1997; 16: 385–395, doi: 10.1002/(SICI)1097-0258(19970228)16:4<385::AID-SIM380>3.0.CO;2-3.
- Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw* 2010; 33: 1–22, doi: 10.18637/jss.v033.i01.
- Goeman JJ. L1 penalized estimation in the Cox proportional hazards model. *Biom J* 2010; 52: 70–84, doi: 10.1002/bimj.200900028.
- Sauerbrei W, Royston P, Binder H. Selection of important variables and determination of functional form for continuous predictors in multivariable model building. *Stat Med* 2007; 26: 5512–5528, doi: 10.1002/sim.3148.
- Newman AM, Steen CB, Liu CL, Gentles AJ, Chaudhuri AA, Scherer F, et al. Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat Biotechnol* 2019; 37: 773–782, doi: 10.1038/s41587-019-0114-2.
- Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol* 2016; 17: 218, doi: 10.1186/s13059-016-1070-5.
- Sturm D, Pfister SM, Jones DTW. Pediatric gliomas: current concepts on diagnosis, biology, and clinical management. *J Clin Oncol* 2017; 35: 2370–2377, doi: 10.1200/JCO.2017.73.0242.
- Johnson KJ, Cullen J, Barnholtz-Sloan JS, Ostrom QT, Langer CE, Turner MC, et al. Childhood brain tumor epidemiology: a brain tumor epidemiology consortium review. *Cancer Epidemiol Biomarkers Prev* 2014; 23: 2716–2736, doi: 10.1158/1055-9965.EPI-14-0207.
- Levy JMM, Towers CG, Thorburn A. Targeting autophagy in cancer. *Nat Rev Cancer* 2017; 17: 528–542, doi: 10.1038/nrc.2017.53.
- Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982; 143: 29–36, doi: 10.1148/radiology.143.1.7063747.
- Wang Z, Tong D, Han C, Zhao Z, Wang X, Jiang T, et al. Blockade of miR-3614 maturation by IGF2BP3 increases TRIM25 expression and promotes breast cancer cell proliferation. *EBioMedicine* 2019; 41: 357–369, doi: 10.1016/j.ebiom.2018.12.061.
- Ikeda H, Lethe B, Lehmann F, van Baren N, Baurain JF, de Smet C, et al. Characterization of an antigen that is recognized on a melanoma showing partial HLA loss by CTL expressing an NK inhibitory receptor. *Immunity* 1997; 6: 199–208, doi: 10.1016/S1074-7613(00)80426-4.
- Lezcano C, Jungbluth AA, Nehal KS, Hollmann TJ, Busam KJ. PRAME expression in melanocytic tumors. *Am J Surg Pathol* 2018; 42: 1456–1465, doi: 10.1097/PAS.0000000000001134.
- Orlando D, Miele E, De Angelis B, Guercio M, Boffa I, Sinibaldi M, et al. Adoptive immunotherapy using PRAME-specific T cells in medulloblastoma. *Cancer Res* 2018; 78: 3337–3349, doi: 10.1158/0008-5472.CAN-17-3140.
- Itou H, Muramatsu S, Shirakihara Y, Araki H. Crystal structure of the homology domain of the eukaryotic DNA replication proteins Sld3/Treslin. *Structure* 2014; 22: 1341–1347, doi: 10.1016/j.str.2014.07.001.
- Yu Q, Pu SY, Wu H, Chen XQ, Jiang JJ, Gu KS, et al. TICRR contributes to tumorigenesis through accelerating DNA replication in cancers. *Front Oncol* 2019; 9: 516, doi: 10.3389/fonc.2019.00516.
- Hou PF, Jiang T, Chen F, Shi PC, Li HQ, Bai J, et al. KIF4A facilitates cell proliferation via induction of p21-mediated cell

- cycle progression and promotes metastasis in colorectal cancer. *Cell Death Dis* 2018; 9: 477, doi: 10.1038/s41419-018-0550-9.
29. Hu G, Yan Z, Zhang C, Cheng M, Yan Y, Wang Y, et al. FOXM1 promotes hepatocellular carcinoma progression by regulating KIF4A expression. *J Exp Clin Cancer Res* 2019; 38: 188, doi: 10.1186/s13046-019-1202-3.
 30. Rath O, Kozielski F. Kinesins and cancer. *Nat Rev Cancer* 2012; 12: 527–539, doi: 10.1038/nrc3310.
 31. Archangelo LF, Greif PA, Holzel M, Harasim T, Kremmer E, Przemeck GK, et al. The CALM and CALM/AF10 interactor CATS is a marker for proliferation. *Mol Oncol* 2008; 2: 356–367, doi: 10.1016/j.molonc.2008.08.001.
 32. Jiang L, Ren L, Zhang X, Chen H, Chen X, Lin C, et al. Overexpression of PIMREG promotes breast cancer aggressiveness via constitutive activation of NF-kappaB signaling. *EBioMedicine* 2019; 43: 188–200, doi: 10.1016/j.ebiom.2019.04.001.
 33. Batista A, Riedemann L, Vardam T, Jain RK. Targeting the tumor microenvironment to enhance pediatric brain cancer treatment. *Cancer J* 2015; 21: 307–313, doi: 10.1097/PPO.000000000000125.