scientific reports

Check for updates

OPEN

Prediction of neuroblastoma prognosis with a novel T-cell exhaustion-related gene signature

Jiangtao Chen, Xiao Zhai, Haiyang Kuang & Ai Zhang✉

Neuroblastoma (NB) is the most common type of pediatric extra-cranial tumor that arises in the sympathetic nervous system. The heterogeneity of T-cell exhaustion (TEX) has been linked to the determination of distinct clinical outcomes and the effectiveness of immunotherapy in numerous adult malignancies. Therefore, studying the heterogeneous TEX landscape in NB as well as its impact on clinical outcomes is meaningful. The gene expression and clinical datasets of the Sequencing Quality Control (SEQC), E-MTAB-8248, and Therapeutically Applicable Research to Generate Effective Treatments (TARGET) cohorts were downloaded from publicly available databases. Two TEX-related clusters for neuroblastoma were identified in the SEQC cohort. Patients in TEX-C1 exhibited superior overall survival (OS) and event-free survival (EFS) rates compared with those in TEX-C2. And TEX-C1 had more immune cells infiltrating, as well as higher expression of immune checkpoint genes. A total of 1984 genes were differentially expressed between these two clusters, of which 1712 were associated with OS. A gene signature consisting of ten TEX-related genes was developed, and a risk score was computed for each patient. Based on the risk score, SEQC patients were split into high- and low-risk groups with significantly different survival rates. The risk score was an independent risk factor predicting survival and showed superior prediction power for 3, 5, and 10-year survival compared to individual clinical parameters. The signature was further confirmed in the TARGET and E-MTAB-8248 cohorts. This study has successfully constructed a risk score model for NB prognosis, utilizing TEX as its foundation. The model provides risk classification and survival evaluation, which can further guide treatment.

Keywords Neuroblastoma, T-cell exhaustion, Signature, Overall survival, Event-free survival

Abbreviations

NB	Neuroblastoma
INSS	International Neuroblastoma Staging scale
TIME	tumor immune microenvironment
TEX	T-cell exhaustion
IFN-g	interferon-g
TNF	tumor necrosis factor
CTL	cytotoxic potential
IL-2	interleukin 2
SEQC	Sequencing Quality Control
TARGET	Therapeutically Applicable Research to Generate Effective Treatments
COG	Children's Oncology Group
EFS	event-free survival
LASSO	Least absolute shrinkage and selection operator
CD8 TCM	Central memory CD8 + T cells
CD4 TCM	Central memory CD4 + T cells
CD4 TEM	Effector memory CD4 + T cells
CD8 TEM	Effector memory CD8 + T cells

Cancer is the leading cause of death in children worldwide[1]. Neuroblastoma (NB), the most common type of pediatric extra-cranial tumor, is a heterogeneous tumor that arises in the sympathetic nervous system. It is

Department of Pediatrics, Tongji Hospital of Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, China. ✉email: Aizhang1109@163.com

responsible for 7–8% of childhood cancers and approximately 15% of all deaths of pediatric cancer[2,3]. Significant progress in therapeutic strategies has notably improved the survival rates of patients with NB, resulting in a 5-year survival rate of 82.1%[4]. NB patients are classified into low-, intermediate-, and high-risk categories based on their age at diagnosis, MYCN amplification status, International Neuroblastoma Staging scale (INSS) stage, histology, and tumor cell ploidy. This classification is done according to the 2000 COG risk scale. Although individuals with low- and intermediate-risk disease have a survival rate close to 100%, the 5-year survival rate for high-risk NB patients is below 50%. Heterogeneity remains present among high-risk NB patients. Patients who have a naturally favorable outlook but are categorized as high-risk according to the existing classification system will receive harmful treatment, which unnecessarily exposes them to the possibility of long-term negative effects[5]. Thus, it is crucial to identify more precise markers to predict prognosis and guide treatment in NB.

Immunotherapy has proven to be highly effective, resulting in the broad recognition of the tumor immune microenvironment (TIME) as a crucial factor in cancer progression, clinical prognosis, and therapeutic efficacy[6,7]. The survival rate of high-risk NB patients has risen in recent years following effective antiGD2 therapy, indicating the promising potential of immunotherapy in treating NB[8]. Although NB has low immunogenicity due to a low mutational burden and consequently low neoepitope expression, as well as low expression of *MHC-I*, the infiltration of T and NK cells into the tumor is nevertheless indicative of survival prognosis[9–11]. In addition to the intrinsic strategies employed by modified cells to escape immune detection during the development of NB, another cause contributing to NB growth is the impairment of immune cell function, which coincides with the formation of a suppressive tumor microenvironment[12]. T-cell exhaustion (TEX) is a type of immune cell dysfunction featured in many chronic infections and cancers[13]. T cells that are exhausted exhibit a gradual decline in their ability to perform effector activities, such as producing cytokines and carrying out killing functions. They also display the presence of many inhibitory receptors, disrupted metabolism, a weakened ability to recall memories, and homeostatic proliferation[14]. TEX is distinguished by the hierarchical impairment of several pathways, including interferon-g (*IFN-g*), tumor necrosis factor (*TNF*), cytotoxic potential (*CTL*), and interleukin 2 (*IL-2*) production[15,16]. The heterogeneity of TEX has been linked to the determination of distinct clinical outcomes and the effectiveness of immunotherapy in numerous adult malignancies[17,18]. However, the heterogeneous TEX landscape in pediatric cancer especially in NB as well as its impact on clinical outcomes are still not well comprehended.

In this study, we conducted a comprehensive analysis of TEX heterogeneity in NB using transcriptomic data from several datasets. Our study focused on investigating the presence of the TEX landscape in the context of NB and its impact on clinical prognosis. Furthermore, our objective was to discover a prognostic signature related to TEX to help improve the molecular classification and risk stratification of patients with NB.

Methods

Data acquisition and inclusion

This study used neuroblastoma dataset of the Sequencing Quality Control (SEQC), E-MTAB-8248, and Therapeutically Applicable Research to Generate Effective Treatments (TARGET) cohorts. Patients diagnosed as neuroblastoma who have available survival information were included in the study. Totally, 498 patients from SEQC, 150 patients from TARGET, and 223 patients from E-MTAB-8248 were included in the study. The gene expression and clinical datasets of SEQC, E-MTAB-8248, and TARGET cohorts were obtained from the Gene Expression Omnibus (GEO: GSE62564), ArrayExpress (E-MTAB-8248), and UCSC Xena (TARGET: NB) databases, respectively. The SEQC cohort was used as the training dataset, while the E-MTAB-8248 and TARGET datasets were used as the test datasets. The SEQC dataset also provided clinical information including age, gender, MYCN amplification, Children's Oncology Group (COG) classification, INSS phase, and event-free survival (EFS). The SEQC dataset utilized the INSS and the 2000 COG classification of risk systems.

Identifying TEX-related clusters

The molecular signature database (MSigDB, V7.2) provided information on TEX associated signaling pathways and marker genes. Similar to previous studies, we performed an unsupervised cluster analysis of neuroblastoma patients using IFN-γ, TNF, and IL-2 signaling pathways to represent the TEX pathway[17]. The method of "ssGSEA" was used to estimate the activity score of the TEX pathways in each patient using the 'GSVA' R package. Then Consensus clustering was performed to identify potential TEX clusters using the R package "ConsensusClusterPlus" based on the enrichment score of different TEX pathways. The Euclidean distance-based K-means (KM) clustering algorithm was applied in this analysis and was repeated for 100 iterations to guarantee dependability.

Immune microenvironment and functional analysis between different clusters

To further explore the abundance of immunocytes in different clusters, tumor-infiltrating immune cells and tumor immune microenvironment were computationally estimated using the xCell and ESTIMATE methods. Differential expression analysis was conducted using the R package 'limma'. Those genes with $|\log2(\text{fold change})| > 1$ and false discovery rate (FDR) < 0.05 were considered significantly differentially expressed. GSVA is an unsupervised method that evaluates the variation in the gene set enrichment of each sample. Functional enrichment analysis was conducted using the GSVA method based on the R package 'clusterProfiler'. By using the GSVA analysis, it was possible to explore the differences between subgroups in signaling pathways for disease development.

Construction and validation of a TEX-related signature

Univariate Cox regression analysis was employed to assess the association between gene expression and overall survival (OS) on the SEQC dataset. The univariate Cox regression analysis identified prognosis related genes with

a P value < 0.05. The intersection between differentially expressed genes and prognosis related genes were used to construct the model. Then the Least absolute shrinkage and selection operator (LASSO) approach was employed to select genes from the intersection for the construction of the prediction model, leaving only 19 genes. Next, a stepwise multivariate Cox regression analysis comprising the 19 genes was used to develop the model, leaving only 10 genes. A risk score signature model for survival prediction in relation to TEX was developed based on gene expression and used the multivariate Cox regression coefficient as the weight. In addition, a multivariate Cox regression analysis was conducted to assess the independence of the signature, comprising the risk score as well as other clinical parameters such as age, gender, *MYCN* status, COG-risk, and INSS stage. The predictive power of the risk score was assessed using ROC curves and the AUC area for 3-, 5-, and 10-year survival. The risk score was validated in the E-MTAB-8248 and TARGET datasets.

Statistical analysis

All statistical analyses were conducted using the R software (V.3.4.4, https://www.r-project.org/). Wilcox test was used to compare the two groups. Kaplan–Meier method and log-rank tests were used to evaluate the differences in OS or EFS between different patient groups. A p-value less than 0.05 indicates statistical significance.

Results

Screening TEX-related clusters of neuroblastoma

In order to measure the extent of TEX in NB bulk tumor samples, we used a collection of TEX-specific pathways, which encompass tumor necrosis factor (*TNF*), interleukin (*IL*)-2, and interferon-g (*IFN-g*) pathway. Subsequently, we employed single-sample gene set enrichment analysis (ssGSEA) to analyze the active pathway profiles specific to TEX from the RNA-sequencing data of individual bulk tumor tissue samples obtained from the SEQC dataset. By performing clustering on the active TEX-specific pathway profiles, we identified two distinct clusters, which we refer to as TEX-C1 and TEX-C2 (Fig. 1A). The enriched scores of the *TNF*, *IL-2*, and *IFN-g* pathways were all greater in TEX-C1, as shown in Fig. 1B.

Difference between TEX-related clusters on neuroblastoma prognosis and clinical features

To examine the effects of TEX on clinical outcome, the prognostic significance of the TEX-related clusters was assessed in SEQC dataset. Survival analysis revealed that patients belonging to distinct TEX-related clusters exhibited considerably varied rates of OS. The patients belonging to the TEX-C2 cluster exhibited a lower OS rate compared to those in the TEX-C1 cluster, as seen in Fig. 1C. We assessed the event-free survival (EFS) rate in NB and observed that patients in TEX-C2 also exhibited a shorter EFS duration compared to patients in TEX-C1 (Fig. 1D). Next, we examined the distribution of clinical characteristics such as age, gender, INSS stage, COG risk, and MYCN amplification status in the two TEX clusters. We have detected distinct differences between these two groups in terms of age distribution, MYCN amplification, COG risk, and INSS stage. However, no disparities were discovered in terms of gender. Patients in the TEX-C2 group exhibited characteristics such as advanced age, higher MYCN amplification rate, a higher distribution of COG high-risk classification, and a higher INSS stage, in comparison to patients in the TEX-C1 cluster (Table 1).

TEX-related clusters show distinct immune microenvironments

To further characterize the difference of the tumor immune microenvironment of TEX clusters, we first estimated the abundance of infiltrating immune cells in the tumor immune microenvironment. The xCell method yielded results indicating higher levels of various immune cells, including *CD 8* + T cells, *CD 4* + T cells, Central memory *CD8* + T cells (*CD8* TCM), Central memory *CD4* + T cells (*CD4 TCM*), Effector memory *CD4* + T cells (*CD4 TEM*), Effector memory *CD8* + T cells (CD8 TEM) and dendritic cells (DC), within the TEX-C1 Cluster (Fig. 2A and S1). Further analysis revealed that the TEX-C1 group had significantly higher immunological score, immune microenvironment score, and stromal score compared to TEX-C2 group, as shown in Fig. 2B. The ESTIMATE approach yielded similar findings, with TEX-C1 exhibiting greater immune score, stromal score, and ESTIMATE score (Fig. S2). To evaluate the disparity in the expression of immune checkpoint genes, we conducted a comparison of the expression levels of immune checkpoint (ICP) genes between the two clusters. The findings indicated that the majority of ICP genes exhibited higher expression in TEX-C1, whereas only *CD276* and *VTCN1* showed higher expression in TEX-C2 group (Fig. 2C).

Differentially expressed genes between TEX-related clusters

We conducted a differential expression analysis comparing the two clusters, and found 1984 genes that exhibited differential expression (Fig. 3A). Among the whole, TEX-C1 exhibited increased expression in 1862 genes, whilst TEX-C2 showed increased expression in 162 genes. The functional enrichment analysis revealed that the genes up-expressed in the TEX-C1 cluster exhibited a tendency to be enriched in pathways related with immune response (Fig. 3B). In addition, we conducted univariate Cox analysis and identified an association between the expression of 1712 genes and the OS of NB patients. Among these genes, 122 were found to be more highly expressed in TEX-C2, whereas 1590 genes were found to be more highly expressed in TEX-C1.

A TEX-related gene signature for predicting survival and treatment efficacy

We performed LASSO and step COX analyses on the SEQC dataset, resulting in a prediction model consisting of ten genes (Table S1). The formula for the model is as follows: 0.4902*expression value of *SCNN1B* + 0.3224* expression value of *VGF*- 0.3364* expression value of *ECEL1–0*.2859* expression value of *CDH6*- 0.2663* expression value of *KLRK1* -0.31912* expression value of *CHD5* + 0.5664* expression value of *CENPF* + 0.3056* expression value of *HAPLN4* -0.2151* expression value of *TMC1* + 0.2169* expression value of *STAC2*. A risk score was computed for patients in the SEQC cohort using the ten-gene signature. The patients in the

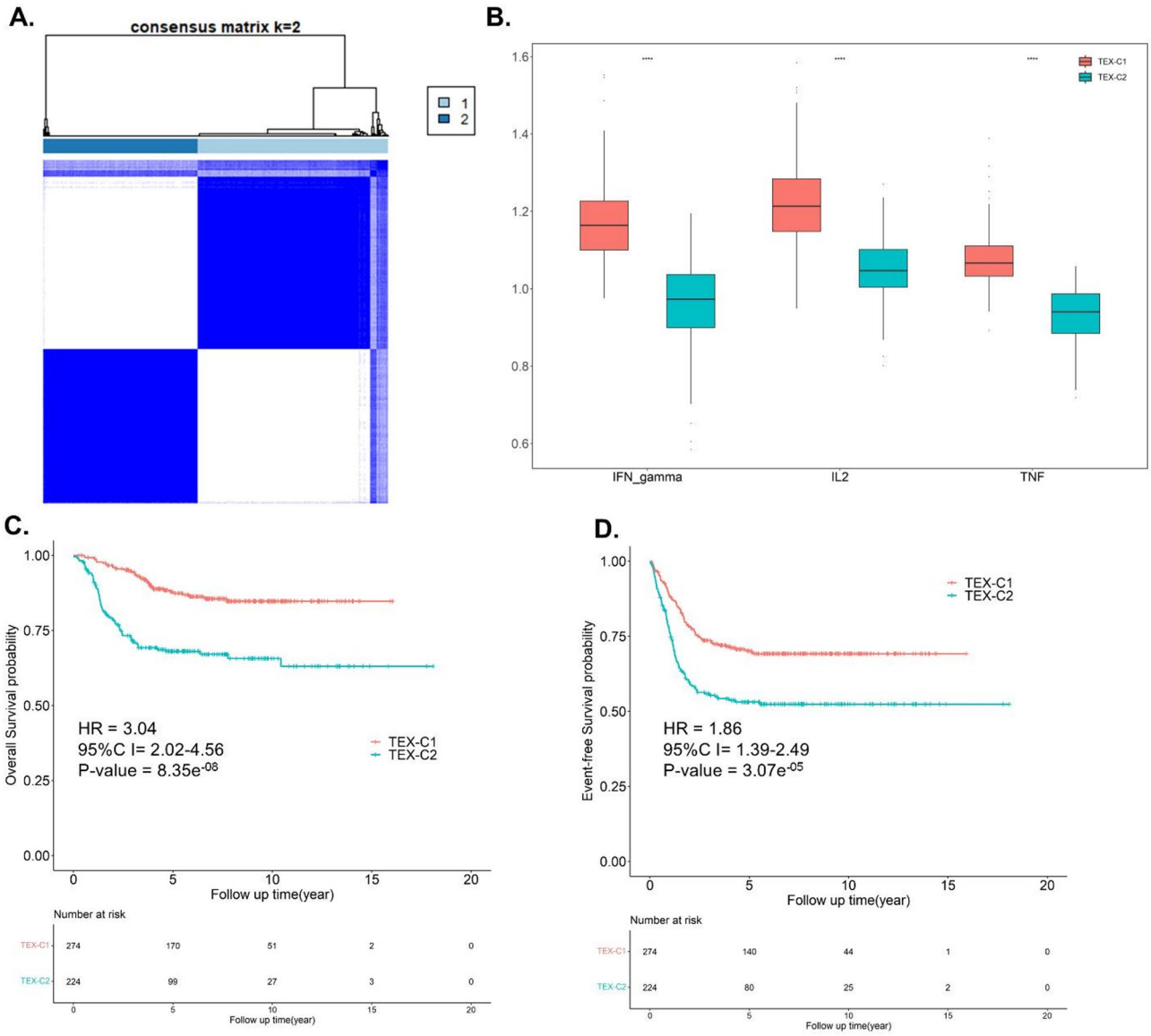

Fig. 1. Identifying heterogeneous clusters of neuroblastoma patients based on the T-cell exhaustion (TEX) associated pathways. (**A**) Consensus matrix identifying two clusters (k = 2), (**B**) the enrichment score of TEX associated pathways between the two clusters [* means significant difference between the two clusters ($P < 0.05$)], (**C**) Kaplan–Meier survival curves of overall survival between two distinctive TEX-related clusters. (**D**) Kaplan–Meier survival curves of event-free survival between two TEX-related clusters.

SEQC dataset were split into high-risk and low-risk groups based on the median value. Figure 4A displays the distribution of 10 gene signature risk scores, gene expression profiles, and survival status. The KM curves demonstrated that patients classified as high-risk had a worse OS rate compared to patients classified as low-risk (Fig. 4B). Subsequently, we conducted a multi-Cox regression analysis to determine the independence of the risk score from clinical features. Figure 4C demonstrates a significant association between the risk score and OS, even after adjusting for clinical variables such as age, gender, INSS stage, COG risk, and *MYCN* amplification status (HR = 36.39 [95%CI: 13.38–98.95], log-rank $P = 1.89E^{-12}$) (Fig. 4C). A 3-level risk classification was also done using the cut-off values from the x-tile software (https://medicine.yale.edu); the intermediate-risk group had different overall survival (OS) and event-free survival (EFS) compared to the high- and low-risk groups ($p < 0.0001$) (Fig. S3). These results from reclassification indicate the robustness of the T-cell exhaustion gene signature score.

Moreover, this risk score signature was utilized to assess the probability of survival. The ROC curves indicate that the AUC for OS at 3, 5, and 10 years was 0.954, 0.953, and 0.947 respectively in the SEQC cohort (Fig. 4D). In addition, we conducted a comparison of the prognostic capability of this risk score with other clinical characteristics. The ROC curves demonstrated that this risk score outperforms age, INSS stage, COG risk, and *MYCN* amplification status in accurately predicting 3-, 5-, and 10-year OS rates (Fig. 4D).

Characteristic		Total	TEX-C1	TEX-C2	Significance (*P*-value)*
Age (median, days)		444.5	379	518	1.40E-03
Gender	Male	287 (100.00%)	162 (56.45%)	125 (43.55%)	0.51
	Female	211 (100.00%)	112 (53.08%)	99 (46.92%)	
MYCN amplification	non-Amplified	401 (100.00%)	253 (63.09%)	148 (36.91%)	1.19E-12
	Amplified	92 (100.00%)	19 (20.65%)	73 (79.35%)	
	Unknown	5 (100.00%)	2 (40.00%)	3 (60.00%)	
COG-risk	High	176 (100.00%)	59 (33.52%)	117 (66.48%)	1.99E-12
	Low	322 (100.00%)	215 (66.77%)	107 (33.23%)	
INSS stage	1	121 (100.00%)	81 (66.94%)	40 (33.06%)	5.30E-06
	2	78 (100.00%)	54 (69.23%)	24 (30.77%)	
	3	63 (100.00%)	33 (52.38%)	30 (47.62%)	
	4	183 (100.00%)	74 (40.44%)	109 (59.56%)	
	4 S	53 (100.00%)	32 (60.38%)	21 (39.62%)	
Total		498 (100.00%)	274 (55.02%)	224 (44.98)	

Table 1. Distribution of clinical variables between T-cell exhaustion-related clusters. *The chi-test was used for comparing the distribution of categorical variables; the t-test was used for comparing continuous variables.

We additionally assessed the prognostic power of this risk score signature on EFS and observed that patients classified as high-risk exhibited significantly shorter EFS compared to patients classified as low-risk (Fig. 4E). According to the data presented in Fig. 4F, the risk score had a strong correlation with EFS even after taking into account several clinical factors such as age, gender, INSS stage, COG risk, and MYCN mutation status (HR = 6.76 [95%CI: 4.60–9.92], log-rank $P = 1.78E^{-22}$). The AUC values for the 3-, 5-, and 10-year EFS were 0.831, 0.827, and 0.833, respectively, for the risk score. These values were considerably higher than those obtained for other clinical variables such as age, INSS stage, COG risk, and *MYCN* amplification status (Fig. 4G).

Then we evaluated the heterogeneity of high- and low-risk patients. 72.7% of the low-risk patients were within the TEX-C1 cluster, while 62.7% of the high-risk patients were within the TEX-C2 cluster (chi-squared test $p = 4.624e\text{-}15$), which shows a higher risk score correlates with a worse TEX status (Fig. S4A). CD8 + T cell scores were significantly reduced in the high-risk group compared to the low-risk group [1.76-fold (mean value), $p = 0.002$] (Fig. S4B). Further, the MYCN amplification rate was 68% (92/249) in high-risk vs. 0% in low-risk (chi-squared test $p < 2.2e\text{-}16$). All these characteristics contribute to a better prognosis in low-risk patients.

Risk stratification for patients with different clinical features

The prognosis for patients with *MYCN* amplification and COG high-risk classification was much poorer, necessitating prolonged treatment. Here, we assess the risk score based on different clinical variables. Our analysis revealed a significantly higher risk score in the *MYCN* amplification group compared to the *MYCN* non-amplification group (Fig. 5A). Patients with *MYCN* amplification were divided into high- and low-risk groups based on their median value. These groups showed significantly different OS and EFS rates, as shown in Fig. 5B. We additionally assessed the distribution of risk scores in patients with different COG-risk and INSS phases. The risk score for patients classified as COG high-risk is significantly greater than that of individuals classified as low-risk (Fig. 5C). The patients in the COG high-risk group were divided into distinct prognostic groups based on their median value (Fig. 5D). Comparable findings have been observed in patients across different INSS stages (Fig. S5 and S6). We also evaluate the effect of treatment on the risk stratification. As for 91 patients who died despite intensive chemotherapy in the SEQC cohort, 87/91 (95.60%) of them were divided into the high-risk group using the cut-off value of the whole cohort. And they were divided into two groups using the median score, which hold significant differences in overall survival [$P = 0.001$, HR = 2.03(1.32–3.12)] and event-free survival [$P = 0.0098$, HR = 1.75(1.14–2.67)] (Fig. S7).

Validation of the TEX-related gene signature in external datasets

In addition, we assessed the prognostic power of the signature in two additional cohorts of NB to ascertain its robustness. Using the same cutoff value, the signature divided 150 patients in the TARGET cohort into high-risk (93, 62%) and low-risk (57, 38%) groups. Figure 6A displays the distribution of 10 gene signature risk scores, gene expression profiles, and survival status. The OS of patients in the high-risk group was considerably lower compared to the low-risk group (HR = 2.32 (95%CI: 1.38–3.92), log-rank $P = 1.58E^{-3}$) (Fig. 6B). We calculated the risk score for patients in the E-MTAB-8248 cohort, which included 223 patients diagnosed with NB. The patients were categorized into high-risk (118, 53%) and low-risk (105, 47%) groups. Figure 6C displays the distribution of 10 gene signature risk scores, gene expression profiles, and survival status in the E-MTAB-8248 cohort. The survival rate of patients in the low-risk group was markedly superior to that of patients in the high-risk group (HR = 14.59 (95%CI: 4.51–47.24), log-rank $P = 7.74E^{-6}$) (Fig. 6D).

Discussion

The tumor microenvironment, which consists of cancer cells, stromal cells, blood vessels, nerve fibers, extracellular matrix, and associated acellular components, plays crucial and distinct roles in carcinogenesis,

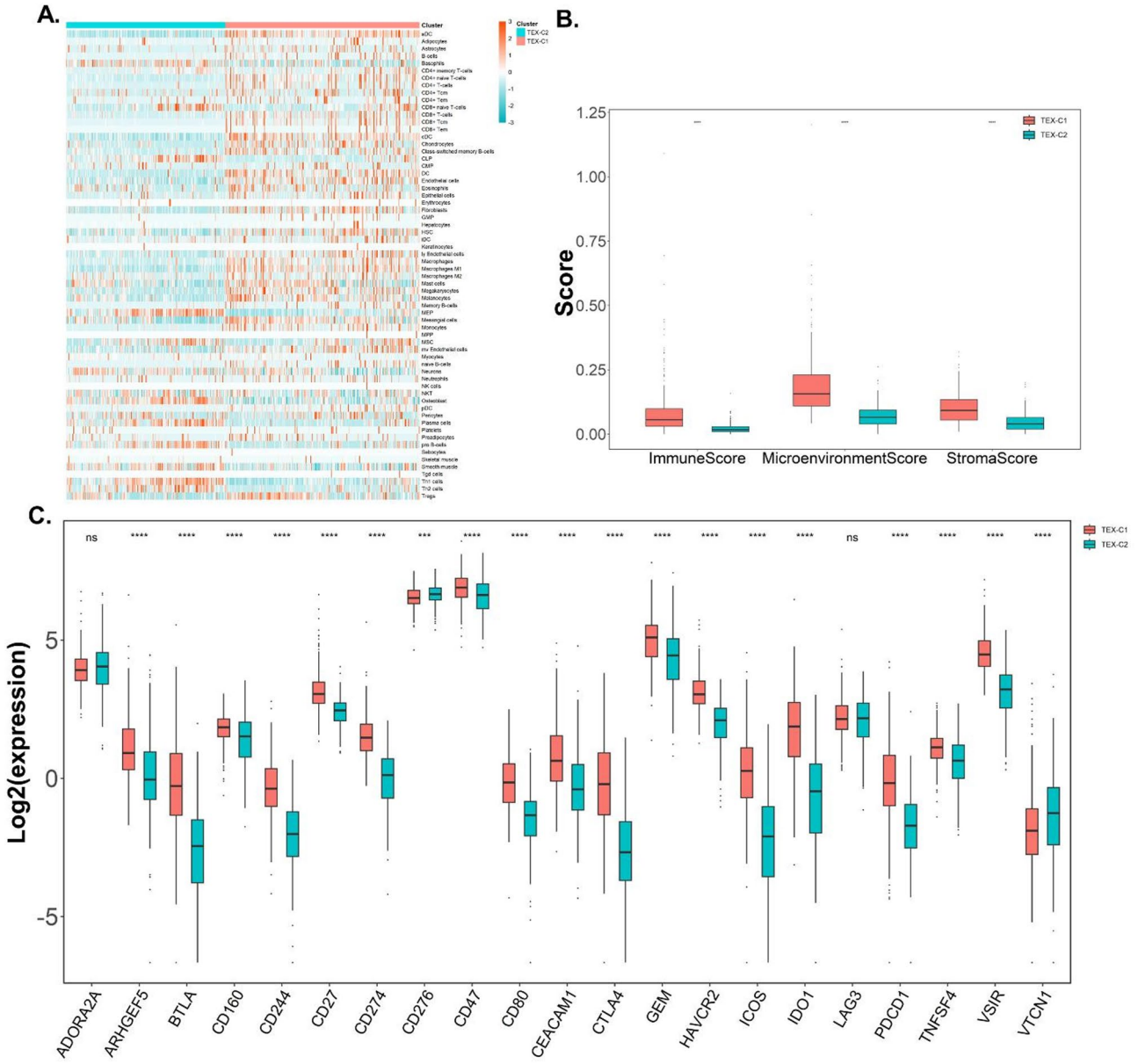

Fig. 2. Immune microenvironment characterization of T-cell exhaustion (TEX) -related clusters. (**A**) The abundance of infiltrating immune cells in the tumor immune microenvironment using the method of xcell, (**B**) difference in the immune score, immune microenvironment score and stromal score between TEX-related clusters [* means significant difference between the two clusters ($P<0.05$)], (**C**) difference in the expression of immune check-point genes between TEX-related clusters [* means significant difference between the two clusters ($P<0.05$)].

cancer progression, and therapeutic outcome[19,20]. TILs, being a critical component, exhibit variation in abundance and state across distinct forms of cancer[21]. Accumulating data indicates that TEX arises from a persistent condition of hierarchical malfunction in TILs, which involves a range of phenotypic and intermediate functional states. Hence, comprehending the dysregulation and exhaustion of T cells within the TIME is crucial for overcoming the tumor immune evasion barrier and enhancing the efficacy of immunotherapy in clinical practice. Nevertheless, there is a lack of comprehensive analysis of the heterogeneity of TEX within TIME of NB. In this study, we thoroughly analyzed the diverse TEX composition in the TIME of NB by utilizing TEX-associated pathway signatures obtained from databases. Due to its lack of definitiveness and quantitative nature, the single gene-defined indicator for TEX was inadequate in accurately representing the progressive and hierarchical processes of TEX. However, our multi-pathway-based results effectively displayed the heterogeneity in TEX in a progressive manner[17].

The study identified two distinct TEX clusters that could be utilized for immunological subtyping. The two groupings exhibited diverse TEX patterns and distinct survival outcomes. The prognosis of patients in the TEX-C1 cluster, characterized by a higher infiltration of immune cells, was significantly more favorable than that

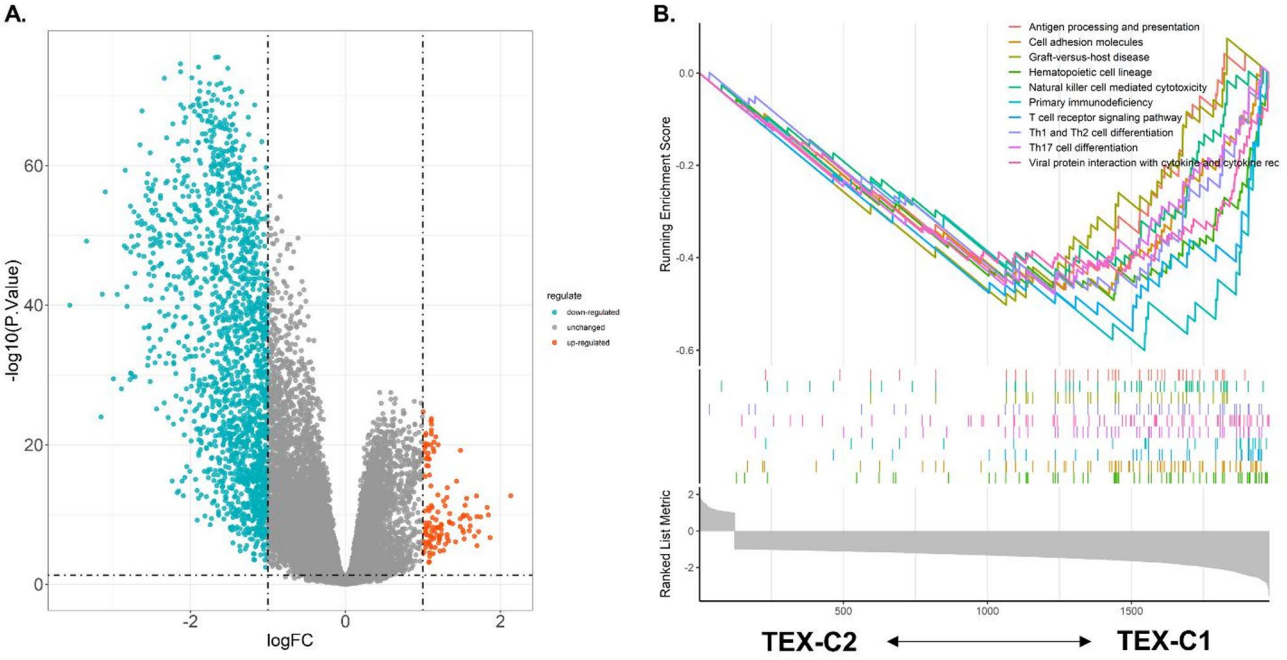

Fig. 3. Differentially expressed genes and enriched pathways between T-cell exhaustion (TEX) -related clusters. (**A**) Volcano plot showing differentially expressed genes between TEX-related clusters, (**B**) Enriched pathways in the two TEX-related clusters.

of patients in the TEX-C2 subset, characterized by a worse TEX status. This highlights the significant impact of TEX status on prognosis and demonstrates the importance of evaluating TEX status in order to improve prognosis. Our findings align with a prior investigation on NB, which suggests that the increased expression of genes associated with immune cell exhaustion in tumors with large levels of cytotoxic immune cells is indicative of a more favorable prognosis[22]. In addition, TEX is shown to be accompanied by the presence of inhibitory cell surface receptors, which serve as effective targets for immune checkpoint blockade therapies to reactivate cytotoxic immune cells[23,24].

The INSS stage is currently used to determine treatment strategy and assessment of NB prognosis. NB patients are also commonly classified as at high-, intermediate-, or low- risk of death according to the Children's Oncology Group (COG) risk classification[25]. However, both the INSS and COG categorization fail to fully elucidate the highly disparate clinical prognosis of NB. In this study, we identified genes that exhibited differential expression between TEX clusters and employed a machine learning technique to develop a TEX-related risk score signature for predicting the prognosis of NB patients. When applying this risk score model to the training dataset from SEQC dataset and the test datasets from TARGET and E-MATB datasets, it successfully divided patients into two distinct groups with significantly different survival outcomes. This demonstrates the model's strength and reliability. The risk score outperforms both INSS stage and COG risk in predicting survival. Moreover, the risk score model has been demonstrated to be an autonomous and robust tool for assessing NB EFS. Notably, when applying this risk score to patients with *MYCN* amplification or COG high-risk, the patients were also categorized into high and low risk groups. This categorization will aid in assessing the survival of these high-risk patients prior to administering hazardous medication and preventing long-term complications.

T cell exhaustion-related models have been constructed for numerous malignancies to predict prognosis[26–28]. But there were no comparable T cell exhaustion-related neuroblastoma models available prior to our study, which was the first to generate a T cell exhaustion-related gene signature. We compared our model to two models generated based on the SEQC dataset to determine whether our T-cell exhaustion-related model outperformed these models. Hu et al. discovered a ferroptosis-related gene signature, and their ROC areas for predicting 3-, 5-, and 10-year survival were 0.902, 0.912, and 0.924, respectively, which were all lower than our model[29]. Tan et al. developed a *ALKBH5*-related 5-gene signature, and the ROC areas for predicting 3- and 5-year survival were 0.902 and 0.916, respectively, which were all lower than our model[30]. These findings demonstrate our 10-gene model's outstanding performance. As for the clinical use, the 10-gene signature could be designed as a NanoString nCounter gene panel in the future[31]. The NanoString nCounter only requires 20 ng RNA, and its hybridization-based technology achieves higher sensitivity for low-abundance transcripts and better inter-laboratory reproducibility compared to RNA-seq, particularly critical for degraded FFPE samples common in neuroblastoma prognosis evaluation.

Among the set of 10 genes in the model, *SCNN1B* is located on chromosome 16p12-p13. Prior research has investigated the function of *SCNN1B* in gastrointestinal cancers, where it is typically identified as a tumor suppressor gene. However, other studies have demonstrated that reducing the expression of *SCNN1B* resulted in a decrease in the migration of glioma cells[32–34]. The neuroendocrine regulatory polypeptide *VGF* is a protein

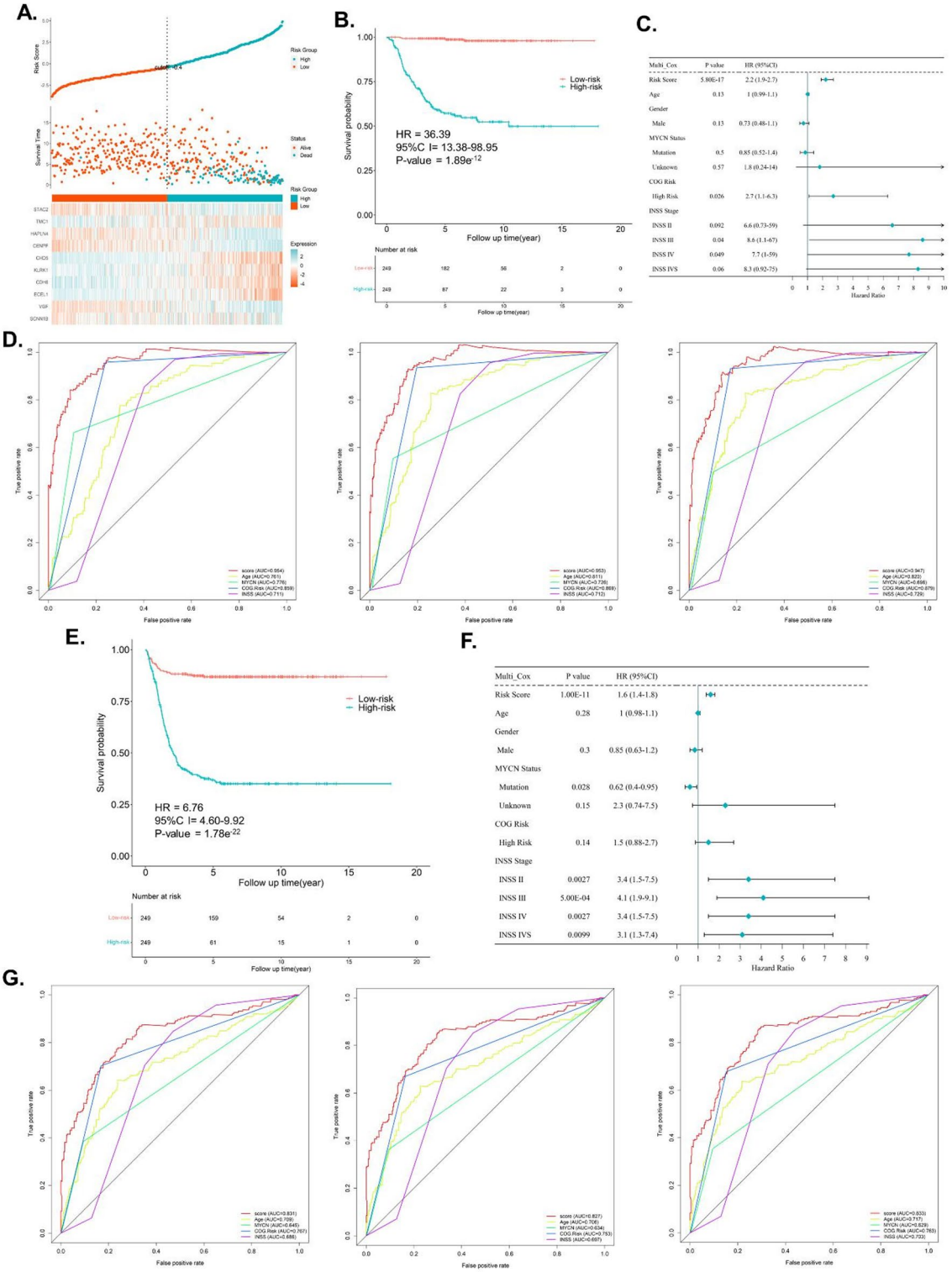

Fig. 4. Development of the ten gene T-cell exhaustion-related signature to predict prognosis. (**A**) The distribution of risk scores, gene expression profiles and survival status, (**B**) Kaplan–Meier survival curves of overall survival between two risk groups according to the risk score, (**C**) Forest plot showing the results of multivariate analysis on overall survival comprising the risk score and clinical characteristics, (**D**) ROC curves for predicting 3-, 5-, and 10- year overall survival, (**E**) Kaplan–Meier survival curves of event-free survival between two risk groups according to the risk score, (**F**) Forest plot showing the results of multivariate analysis on event-free survival comprising the risk score and clinical characteristics, (**G**) ROC curves for predicting 3-, 5-, and 10- year event-free survival.

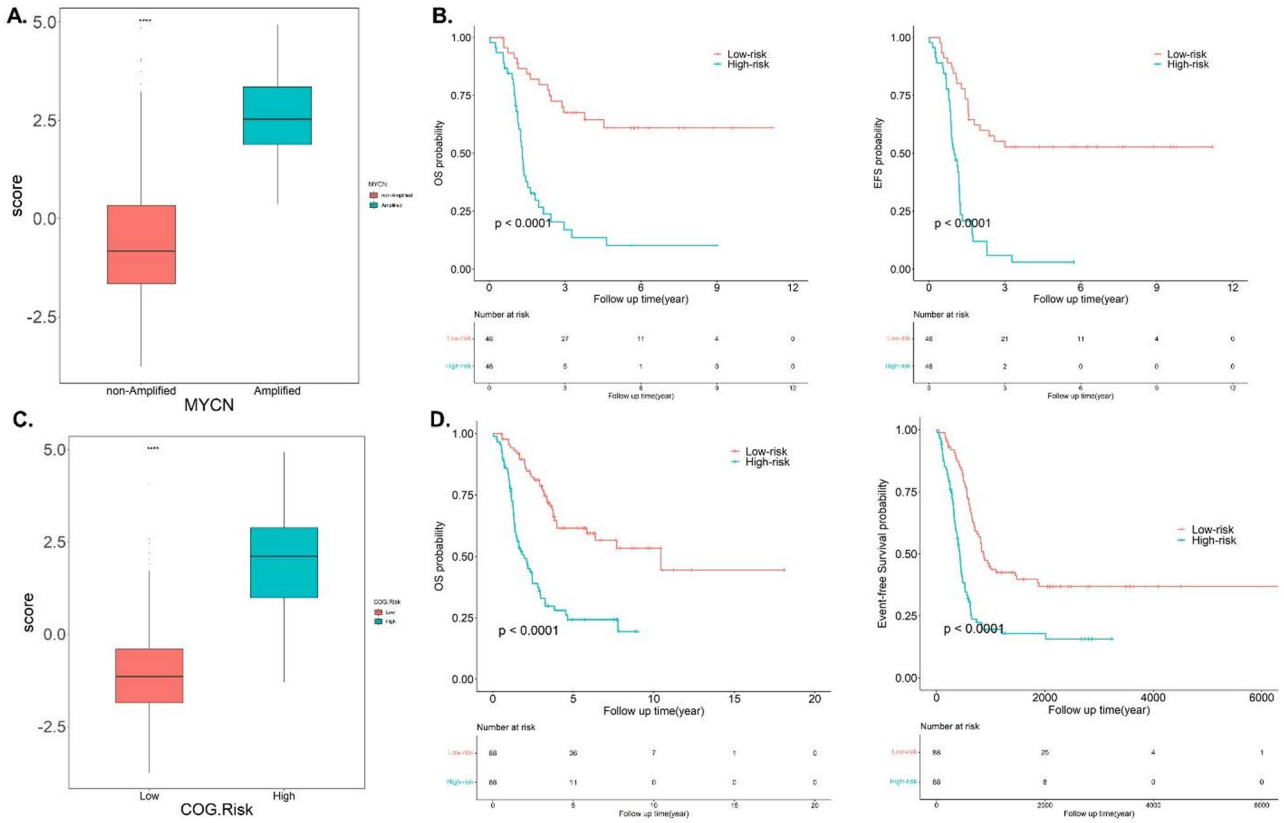

Fig. 5. Evaluation of the risk score under different clinical characteristics. (**A**) Difference in risk score between *MYCN* non-amplified and amplified patients [* means significant difference between the two groups ($P < 0.05$)], (**B**) the difference in overall survival and event-free survival between high and low risk groups in the *MYCN* amplified patients, (**C**) difference in risk score between COG high- and low- risk patients [* means significant difference between the two groups ($P < 0.05$)], (**D**) the difference in overall survival and event-free survival between high and low risk groups in the COG high-risk patients. * means significant difference between the two clusters ($P < 0.05$).

that is secreted by brain cells. Increased levels of *VGF* expression have been observed in certain malignancies that have a neuroendocrine origin or have neuroendocrine characteristics[35–37]. *CHD5* was reported to inhibit the metastasis of NB, but its role in tumor immunology needs further study[38]. *CENPF* has been involved in the malignant advancement of various malignancies, including breast cancer, prostate cancer, and gastric cancer[39–41]. It was also observed to be overexpressed in NB compared with ganglioneuroblastoma[42].

There were certain limitations in our study. At first, the present study lacks in fundamental experiments, and further extensive-scale molecular experiments are required to authenticate the findings. In addition, all the datasets utilized in the present study were available from the public. The following study should be conducted using a prospective dataset to assess the performance of this signature. Furthermore, the absence of NB treatment data prevented the assessment of the efficacy of chemotherapy and immunotherapy using the signature.

Conclusion

In summary, our study discovered the heterogeneity of TEX in NB and developed a ten-gene signature associated with TEX that can predict a more precise prognosis. This signature has the potential to provide risk stratification and prognostic assessment, hence optimizing treatment decision-making.

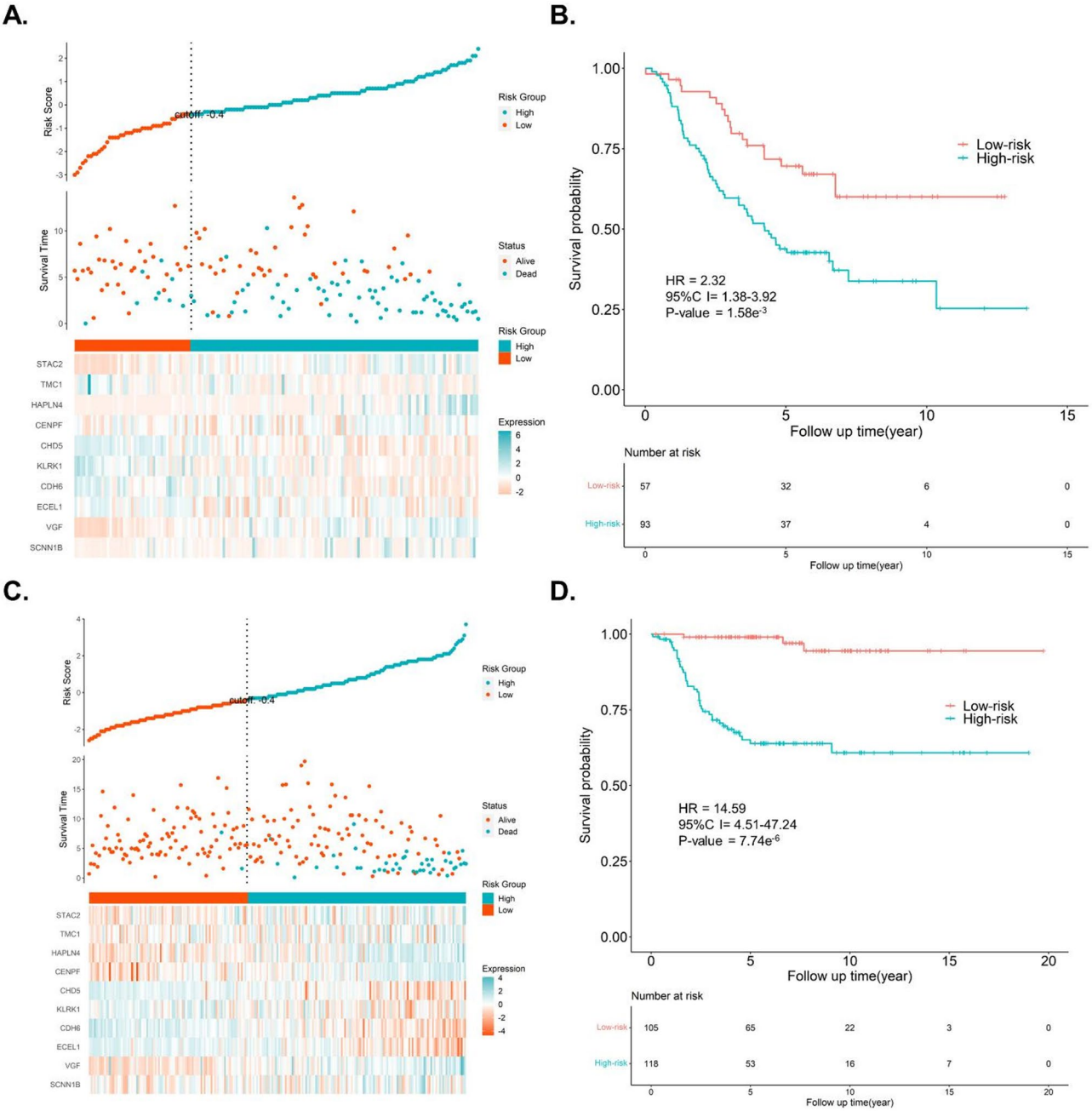

Fig. 6. Validation the risk score signature in external cohorts. (**A**) The distribution of risk scores, gene expression profiles and survival status in TARGET cohort, (**B**) Kaplan–Meier survival curves of overall survival between two risk groups in TARGET cohort, (**C**) The distribution of risk scores, gene expression profiles and survival status in E-MTAB-8248 cohort, (**D**) Kaplan–Meier survival curves of overall survival between two risk groups in E-MTAB-8248 cohort.

Data availability

The datasets used in this study can be found in the GEO database (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE62564), ArrayExpress database (https://www.ebi.ac.uk/biostudies/arrayexpress/studies/E-MTAB-8248?query=E-MTAB-8248) and UCSC Xena database (https://xenabrowser.net/datapages/?cohort=TARGET-neuroblastoma&removeHub=https://xena.treehouse.gi.ucsc.edu:443).

Received: 6 August 2024; Accepted: 15 May 2025
Published online: 23 May 2025

References

1. Steliarova-Foucher, E. et al. International incidence of childhood cancer, 2001-10: a population-based registry study. *Lancet Oncol.* **18**, 719–731 (2017).
2. Whittle, S. B. et al. Overview and recent advances in the treatment of neuroblastoma. *Expert Rev. Anticancer Ther.* **17**, 369–386 (2017).
3. Cone, E. B. et al. Biomarkers for Wilms tumor: A systematic review. *J. Urol.* **196**, 1530–1535 (2016).
4. Fu, X. L. et al. Incidence of suicide mortality among childhood cancer survivors: A population-based retrospective study. *Psychiatry Res.* **304**, 114119 (2021).
5. Vermeulen, J. et al. Predicting outcomes for children with neuroblastoma using a multigene-expression signature: a retrospective SIOPEN/COG/GPOH study. *Lancet Oncol.* **10**, 663–671 (2009).
6. Binnewies, M. et al. Understanding the tumor immune microenvironment (TIME) for effective therapy. *Nat. Med.* **24**, 541–550 (2018).
7. Bagaev, A. et al. Conserved pan-cancer microenvironment subtypes predict response to immunotherapy. *Cancer Cell.* **39**, 845–865e7 (2021).
8. Yu, A. L. et al. Anti-GD2 antibody with GM-CSF, interleukin-2, and isotretinoin for neuroblastoma. *N Engl. J. Med.* **363**, 1324–1334 (2010).
9. Gröbner, S. N. et al. The landscape of genomic alterations across childhood cancers. *Nature* **555**, 321–327 (2018).
10. Wölfl, M. et al. Expression of MHC class I, MHC class II, and cancer germline antigens in neuroblastoma. *Cancer Immunol. Immunother CII.* **54**, 400–406 (2005).
11. Wienke, J. et al. The immune landscape of neuroblastoma: challenges and opportunities for novel therapeutic strategies in pediatric oncology. *Eur. J. Cancer Oxf. Engl. 1990.* **144**, 123–150 (2021).
12. Pei, M. et al. The transcription factor TOX is involved in the regulation of T-cell exhaustion in neuroblastoma. *Immunol. Lett.* **248**, 16–25 (2022).
13. Wherry, E. J. T cell exhaustion. *Nat. Immunol.* **12**, 492–499 (2011).
14. Kurachi, M. CD8 + T cell exhaustion. *Semin Immunopathol.* **41**, 327–337 (2019).
15. Blank, C. U. et al. Defining 'T cell exhaustion'. *Nat. Rev. Immunol.* **19**, 665–674 (2019).
16. Wherry, E. J. & Kurachi, M. Molecular and cellular insights into T cell exhaustion. *Nat. Rev. Immunol.* **15**, 486–499 (2015).
17. Zhang, Z. et al. Pan-cancer landscape of T-cell exhaustion heterogeneity within the tumor microenvironment revealed a progressive roadmap of hierarchical dysfunction associated with prognosis and therapeutic efficacy. *EBioMedicine* **83**, 104207 (2022).
18. Wang, Q., Qin, Y. & Li, B. CD8 + T cell exhaustion and cancer immunotherapy. *Cancer Lett.* **559**, 216043 (2023).
19. Bejarano, L., Jordão, M. J. C. & Joyce, J. A. Therapeutic targeting of the tumor microenvironment. *Cancer Discov.* **11**, 933–959 (2021).
20. Jin, M. Z. & Jin, W. L. The updated landscape of tumor microenvironment and drug repurposing. *Signal. Transduct. Target. Ther.* **5**, 166 (2020).
21. Zheng, L. et al. Pan-cancer single-cell landscape of tumor-infiltrating T cells. *Science* **374**, abe6474 (2021).
22. Wei, J. S. et al. Clinically relevant cytotoxic immune cell signatures and clonal expansion of T-Cell receptors in High-Risk MYCN-Not-Amplified human neuroblastoma. *Clin. Cancer Res. Off J. Am. Assoc. Cancer Res.* **24**, 5673–5684 (2018).
23. Crespo, J., Sun, H., Welling, T. H., Tian, Z. & Zou, W. T cell anergy, exhaustion, senescence, and stemness in the tumor microenvironment. *Curr. Opin. Immunol.* **25**, 214–221 (2013).
24. Brahmer, J. et al. Nivolumab versus docetaxel in advanced Squamous-Cell Non-Small-Cell lung Cancer. *N Engl. J. Med.* **373**, 123–135 (2015).
25. Irwin, M. S. et al. Revised neuroblastoma risk classification system: A report from the children's oncology group. *J. Clin. Oncol. Off J. Am. Soc. Clin. Oncol.* **39**, 3229–3241 (2021).
26. Lian, C. et al. Identification of T-cell exhaustion-related genes and prediction of their immunotherapeutic role in lung adenocarcinoma. *J. Cancer*. **15**, 2160–2178 (2024).
27. Wu, J., Li, L. & Cheng, Z. System analysis based on T-cell exhaustion-related genes identifies PTPRT as a promising diagnostic and prognostic biomarker for gastric cancer. *Sci. Rep.* **14**, 21049 (2024).
28. Qi, Q. et al. Targeted drug screening leveraging Senescence-Induced T-Cell exhaustion signatures in hepatocellular carcinoma. *Int. J. Mol. Sci.* **25**, 11232 (2024).
29. Hu, J. et al. Integrative analysis of multi-omics data for discovery of ferroptosis-related gene signature predicting immune activity in neuroblastoma. *Front. Pharmacol.* **14**, 1162563 (2023).
30. Tan, K., Wu, W., Zhu, K., Lu, L. & Lv, Z. Identification and characterization of a glucometabolic prognostic gene signature in neuroblastoma based on N6-methyladenosine eraser ALKBH5. *J. Cancer*. **13**, 2105–2125 (2022).
31. Geiss, G. K. et al. Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nat. Biotechnol.* **26**, 317–325 (2008).
32. Qian, Y. et al. Sodium channel subunit SCNN1B suppresses gastric Cancer growth and metastasis via GRP78 degradation. *Cancer Res.* **77**, 1968–1982 (2017).
33. Qian, Y. et al. The sodium channel subunit SCNN1B suppresses colorectal cancer via suppression of active c-Raf and MAPK signaling cascade. *Oncogene* **42**, 601–612 (2023).
34. Li, J. et al. Identification of heterogeneous subtypes and a prognostic model for gliomas based on mitochondrial dysfunction and oxidative stress-related genes. *Front. Immunol.* **14**, 1183475 (2023).
35. Rindi, G. et al. Peptide products of the neurotrophin-inducible gene Vgf are produced in human neuroendocrine cells from early development and increase in hyperplasia and neoplasia. *J. Clin. Endocrinol. Metab.* **92**, 2811–2815 (2007).
36. Matsumoto, T. et al. A new possible lung cancer marker: VGF detection from the conditioned medium of pulmonary large cell neuroendocrine carcinoma-derived cells using secretome analysis. *Int. J. Biol. Markers*. **24**, 282–285 (2009).
37. Annaratone, L. et al. Search for neuro-endocrine markers (chromogranin A, synaptophysin and VGF) in breast cancers. An integrated approach using immunohistochemistry and gene expression profiling. *Endocr. Pathol.* **25**, 219–228 (2014).
38. Laut, A. K. et al. CHD5 inhibits metastasis of neuroblastoma. *Oncogene* **41**, 622–633 (2022).
39. Sun, J. et al. Overexpression of CENPF correlates with poor prognosis and tumor bone metastasis in breast cancer. *Cancer Cell. Int.* **19**, 264 (2019).
40. Chen, E. B. et al. HnRNPR-CCNB1/CENPF axis contributes to gastric cancer proliferation and metastasis. *Aging* **11**, 7473–7491 (2019).
41. Shahid, M. et al. Downregulation of CENPF remodels prostate Cancer cells and alters cellular metabolism. *Proteomics* **19**, e1900038 (2019).
42. Albino, D. et al. Identification of low intratumoral gene expression heterogeneity in neuroblastic tumors by genome-wide expression analysis and game theory. *Cancer* **113**, 1412–1422 (2008).

Author contributions

AZ designed the study, JTC and XZ collected and prepared the data, JTC and HYK performed the data analysis, JTC, XZ, and HYK prepared the table and figures, JTC and AZ wrote and revised the manuscript. All authors

have read and approved the submitted version.

Declarations

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-02661-0.

Correspondence and requests for materials should be addressed to A.Z.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.