

RESEARCH

Open Access



A signature based on five immune-related genes to predict the survival and immune characteristics of neuroblastoma

KeXin Ma¹, PeiPei Zhang², Yu Xia¹, Lin Dong¹, Ying Li¹, Liu Liu¹, YaJuan Liu¹ and YouJun Wang^{1*}

Abstract

Background: MYCN amplification (MNA) has been proved to be related to poor prognosis in neuroblastoma (NBL), but the MYCN-related immune signatures and genes remain unclear.

Methods: Enrichment analysis was used to identify the significant enrichment pathways of differentially expressed immune-related genes (DEIRGs). Weight gene coexpression network analysis (WGCNA) was applied to reveal the correlation between these DEIRGs and MYCN status. Univariate and multivariate Cox analyses were used to construct risk model. The relevant fractions of immune cells were evaluated by CIBERSORT and single-sample gene set enrichment analysis (ssGSEA).

Results: Five genes, including CHGA, PTGER1, SHC3, PLXNC1, and TRIM55 were enrolled into the risk model. Kaplan–Meier survival analysis and receiver operating characteristic (ROC) curve showed that our model performed well in predicting the outcomes of NBL (3-years AUC = 0.720, 5-year AUC = 0.775, 10-years AUC = 0.782), which has been validated in the GSE49711 dataset and the E-MTAB-8248 dataset. By comparing with the tumor immune dysfunction and exclusion (TIDE) and tumor inflammation signature (TIS), we further proved that our model is reliable. Univariate and multivariate Cox regression analyses indicated that the risk score, age, and MYCN can serve as independent prognostic factors in the E-MTAB-8248. Functional enrichment analysis showed the DEIRGs were enriched in leukocyte adhesion-related signaling pathways. Gene set enrichment analysis (GSEA) revealed the significantly enriched pathways of the five MYCN-related DEIRGs. The risk score was negatively correlated with the immune checkpoint CD274 (PD-L1) but no significant difference with the TMB. We also confirmed the prognostic value of our model in predicting immunotherapeutics.

Conclusion: We constructed and verified a signature based on DEIRG that related to MNA and predicted the survival of NBL based on relevant immune signatures. These findings could provide help for predicting prognosis and developing immunotherapy in NBL.

Keywords: Neuroblastoma, MYCN amplification, Immune-related gene, Prognosis

Introduction

Neuroblastoma (NBL) is the most common embryonal solid tumor of infancy, which accounts for 7–8% of malignancies and 15% of cancer-related mortality in children [1]. It is a heterogeneous tumor deriving from neural crest cells (NCCs) [2]. The genetic, morphological, and clinical heterogeneity have been described at

*Correspondence: wyjerke@outlook.com

¹ Department of Pediatrics, The Fifth Affiliated Hospital of Zhengzhou University, No. 3 Kangfuqian Street, Zhengzhou, China
Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

multiple levels, including anatomical localization, histology, genomics/molecular profile, and cellular and molecular levels. The considerable heterogeneity contributes to the clinical and prognostic diversity. Therefore, children with the same stage usually have different outcomes. Although treatments of NBL, for example, high-dose myeloablative chemotherapy, molecular-targeted therapy, autologous hematopoietic stem cell transplantation (AH SCT), and immunotherapy have significantly prolonged the survival of patients [3], NBL is still a life-threatening malignancy with 5-year survival rate less than 50% [4]. Therefore, it is urgently required to develop and validate novel prognostic predictors to guide the treatment of NBL.

MYCN amplification (MNA) accounts for about 20–25% of all primary tumors [5], which is an initiating event that drives the development of high-risk NBL and is strongly associated with high-risk disease and poor prognosis [6]. MYCN affects not only gene expression but also epigenetic factors [7]. It has been used as a biomarker of risk stratification, but the specific related immune features and genes remain unclear. Therefore, it is necessary to establish computational models of the immune-related genes in different MYCN status groups.

In recent years, treatments based on the tumor micro-environment (TME) have attracted attention. TME plays a significant role in tumor progression by providing a growth environment, reducing the efficacy of anti-tumor drugs, and helping tumor cells evade immune surveillance. Overall, tumor-infiltrating immune cells (TIICs) and immune-related genes (IRGs) are essential components of the TME [8, 9]. Studies have shown that the stromal and immune signatures are related to the survival of NBL. However, the specific mechanisms of TME related to tumor progression remain unclear. Therefore, we tried to analyze the relationship between the genes and TME alternations.

In this study, we developed and validated individualized prognostic characteristics of NBL based on the MYCN-related DEIRGs, and then compared the immune components between the two score groups. These results may provide insights for further prediction of the prognosis of NBL patients.

Material and methods

Public datasets

We downloaded the mRNA matrix, somatic mutation, and clinical data from the Therapeutically Applicable Research to Generate Effective Treatments (TARGET) database (www.ocg.cancer.gov/programs/target).

The validation cohorts were downloaded from the Gene Expression Omnibus (GEO) database (GSE49711, www.ncbi.nlm.nih.gov/geo/) and the

ArrayExpress database (E-MTAB-8248, [n=223](http://www.ebi.ac.uk/arrayexpress), www.ebi.ac.uk/arrayexpress). The immunotherapy dataset (IMvigor210, [n=298](http://research-pub.gene.com/IMvigor210CoreBiologies/packageVersions)) was downloaded from the “IMvigor210CoreBiologies” R package (<http://research-pub.gene.com/IMvigor210CoreBiologies/packageVersions>). 298 patients with metastatic urothelial cancer in the IMvigor210 dataset treated with an anti-PD-L1 agent (atezolizumab). IRGs list was obtained from the ImmPort database (www.immport.org/home) and the InnateDB database (www.innatedb.ca).

Analysis of DEIRGs and gene enrichment analysis

According to the MYCN status (MNA or non-MNA), we identified the differentially expressed genes (DEGs) with a false discovery rate (FDR) < 0.05 and $|\log_{2}FC| > 1$ between the two groups in the TARGET cohort using “limma” R package. Then we extracted differentially expressed immune-related genes (DEIRGs) from DEGs based on the IRGs list. These DEGs and DEIRGs were shown in the heatmap. In order to gain new insights into the mechanism and pathways related to these DEIRGs, we performed Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) [10–12] enrichment analyses.

WGCNA to identify key modules and hub genes

To screen hub genes associated with MYCN status, weight gene coexpression network analysis (WGCNA) was carried out based on the expression of these DEIRGs by “WGCNA” R package. We first calculated the degree of adjacency between every two genes and applied the standard scale-free network to evaluate the optimal soft threshold power. Topological overlap matrix (TOM) was used to reduce the effects of noise and spurious associations. TOM-based dissimilarity was used to form modules through dynamic tree cutting. The clustering dendrogram of genes was created through a dendrogram with colored assignments when cut height = 0.2, minModuleSize = 25. To select key modules associated with MNA, $p < 0.05$ was considered significant. Network screening was used to visualize hub genes in the key modules with Cytoscape 3.5.1 software.

Construction of a prognostic signature based on MYCN-related DEIRGs

Univariate Cox regression analysis was used to select prognostic genes located in the key modules. Multivariate Cox regression analysis was used to calculate the regression coefficients for each gene. We constructed a prognostic risk model based on these genes. The formula was as follows: risk score = $\beta_1 * X_1 + \beta_2 * X_2 + \dots + \beta_n * X_n$ (β , risk coefficient; X , the expression of a specific gene). Patients were assigned to the high- or low-risk

group based on the median risk score. Next, we performed univariate and multivariate analyses to identify whether the risk score can serve as an independent factor in NBL. The tumor immune dysfunction and exclusion (TIDE) and tumor inflammation signature (TIS) have been proved to be effective. By comparing with TIDE and TIS, we further proved that our model is reliable. Moreover, we evaluated the correlations between the risk model and clinical factors, including age (age ≤ 1.5 , age > 1.5), gender (female, male), INSS stage (stage 2, 3, 4), MYCN status (amplification, unamplification). We also calculated the proportions of patients in different age groups.

Gene set enrichment analysis

Gene set enrichment analysis (GSEA) was used to identify significantly enriched pathways in the two groups. The “c2.cp.kegg.v7.4.symbols.gmt” gene set collection in the MSigDB database was chosen as the reference. $p < 0.05$ was considered statistically significant.

Analysis of tumor mutation burden

We downloaded the somatic mutation data of 209 patients from the TARGET database and extracted relevant information through Perl script (www.perl.org/), then we estimated the tumor mutation burden (TMB) values. $TMB = (\text{total count of variants}) / (\text{the whole length of exons})$. The R package “maftools” was used to display specific mutation information for the top 20 genes in different groups.

The correlation between 5-DEIRGs and tumor-infiltrating immune cells

The R package “CIBERSORT” was used to quantify 22 kinds of human immune cell phenotypes in each sample and infer their relative proportions. Gene set LM22 was used as the reference. In each sample, the sum of the proportions of 22 immune cells was equal to 1. We set permutations = 1000 to improve the accuracy. Only samples with a CIBERSORT $p < 0.05$ were selected for the following analysis. TIICs infiltration was divided in low- or high- abundance. We also compared the expression difference of immune checkpoint gene CD274 (PD-L1) in two risk groups and the prognostic value of our model in predicting immunotherapeutics.

Single-sample gene set enrichment analysis

We performed single-sample gene set enrichment analysis (ssGSEA) to evaluate the activities of 13 immune-related functions enrichment pathways and infiltration scores of 16 immune cells of each example using the R package “GSVA”.

Statistical analysis

Survival analysis was carried out by the K-M analysis, and the log-rank test was applied between groups. Univariate and multivariate Cox proportional-hazards models were utilized for model construction. The receiver operating characteristic (ROC) curve was used to estimate the predictive ability of our model. Wilcoxon test was used to analyze the correlation between the risk model and clinical features and immune characteristics. For all tests, $p < 0.05$ was considered to be significant. Data analysis in our study was performed by R language version 4.0.5 (www.r-project.org/) with the following packages: “glmnet,” “survminer,” “survival ROC,” and “clusterProfile.”

Results

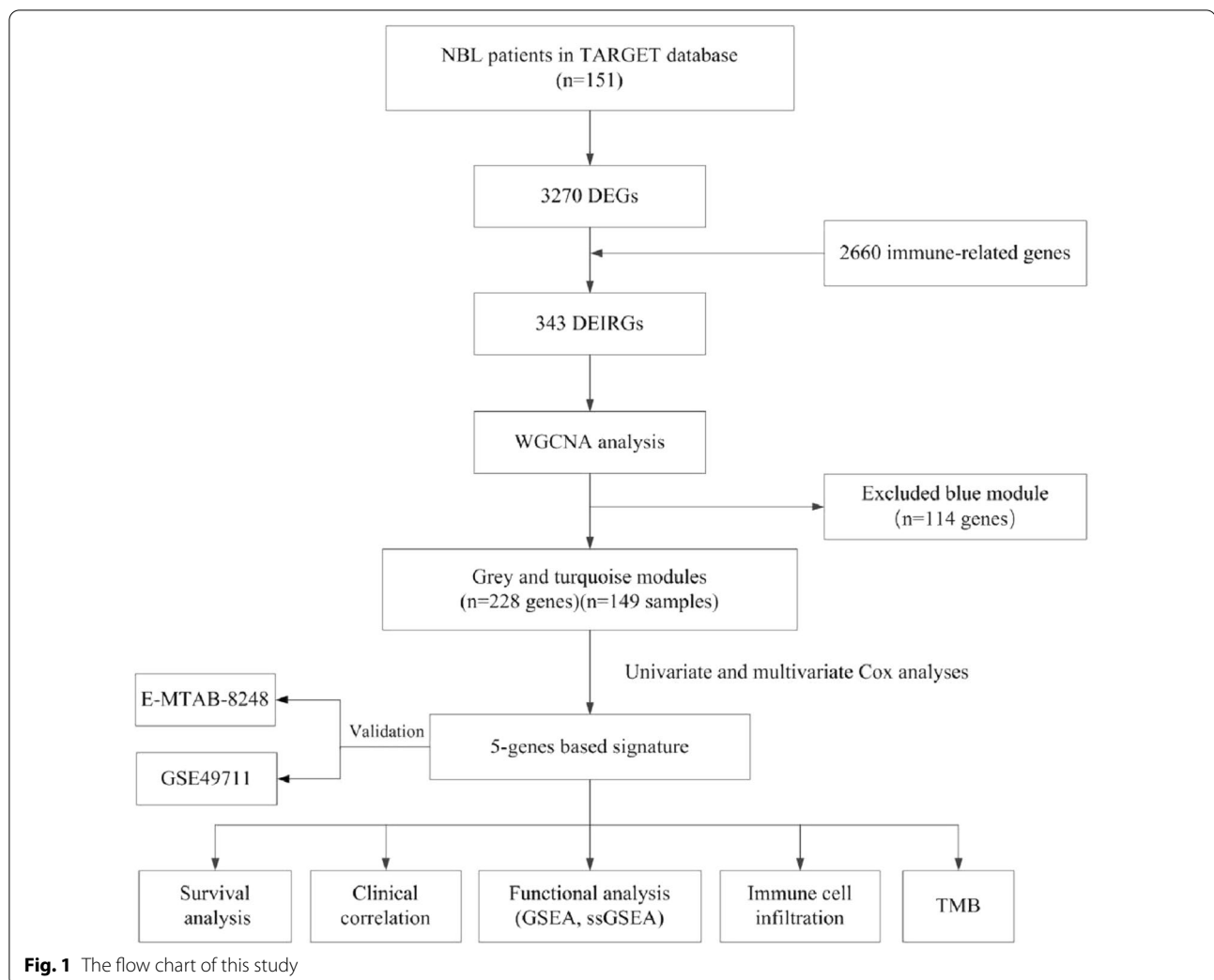
Identification of DEIRG

We downloaded transcriptome profiling and clinical databases of 151 NBL patients from the pediatric TARGET database. Transcriptome profiling of 498 NBL patients were downloaded from the GEO database and transcriptome profiling of 223 NBL patients were downloaded from the ArrayExpress, which served as the external validation groups. The corresponding flow chart is shown in Fig. 1A. We analyzed the DEGs between the MNA group ($n = 33$) and the non-MNA group ($n = 118$). A total of 3270 DEGs were identified, which were displayed in the heatmap (Fig. 2A). From ImmPort and InnateDB databases, we obtained 2660 immune-related genes (IRGs). We extracted 343 DEIRGs from the 3270 DEGs based on the IRGs list, among which 320 genes were down-regulated and 23 genes were up-regulated in the MNA group compared with non-MNA group (Fig. 2B).

Function enrichment analyses were performed using KEGG and GO analyses. By GO analysis (Fig. 2C), DEIRGs were significantly enriched in leukocyte-mediated immunity, T cell activation, signaling receptor activator activity, receptor-ligand activity, T cell receptor complex, and plasma membrane signaling receptor complex. KEGG analysis showed that these DEIRGs were mainly enriched in cytokine-cytokine receptor interaction, tuberculosis, and chemokine signaling pathway (Fig. 2D). In brief, these DEIRGs play immune roles in NBL mainly involved in transmitting various signaling pathways.

WGCNA to identify key modules and hub genes

To further identify DEIRGs related to MYCN status, we performed WGCNA to reveal the correlation between 343 DEIRGs and MYCN status. The scale-free network was constructed based on the expression profile of 343 DEIRGs with a soft threshold power of 8 (Fig. 3A). Next,

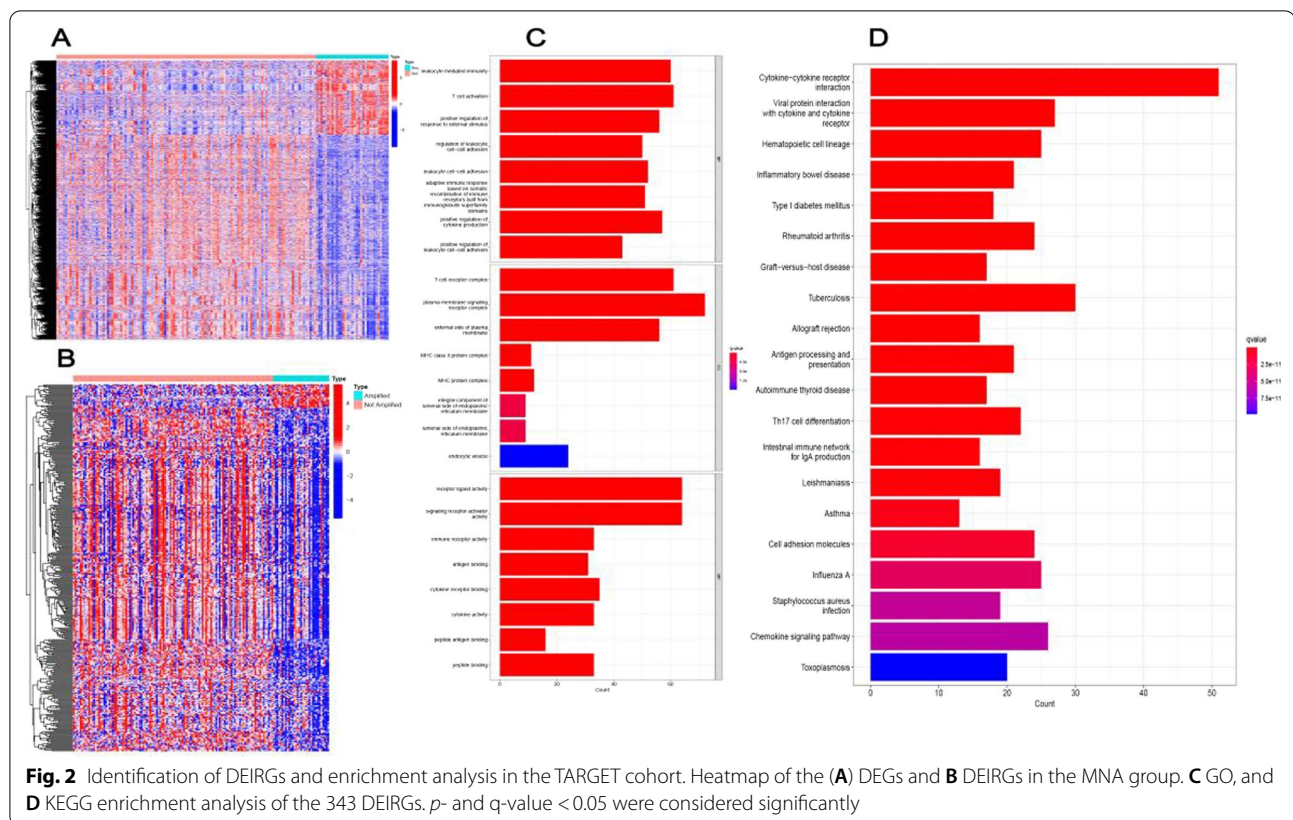


a total of 3 modules were identified by the dynamic cutting tree (minModuleSize=25, cut height=0.2) (Fig. 3B) and results showed that the turquoise ($R=0.44$, $p<0.001$) and grey ($R=0.75$, $p<0.001$) modules were closely correlated with MNA (Fig. 3C). 228 DEIRGs ($n=149$ samples) in the grey and turquoise modules were used to identify hub genes using the Cytoscape graph. The hub genes were located in the center of the modules: the larger the circle, the more likely it to be a core gene (Fig. 3D).

Construction and validation of the prognostic signature

A total of 228 DEIRGs ($n=149$ samples) originated from grey and turquoise modules that related to MNA were used for further analysis. As shown in the univariate Cox regression analysis, we screened 12 prognostic genes, including CRABP1, AMH, CHGA, PTGER1, SHC3, GAL, MAPT, PLXNC1, SCG2, LIFR, TRIM55, and OPTN (Fig. 4A). We applied the multivariate stepwise method (both forward and backward) to screen optimal

genes to construct the prognostic model. Finally, five genes, including CHGA, PTGER1, SHC3, PLXNC1, and TRIM55 were enrolled and the risk score of each patient was calculated based on the following formula: The risk score = $(-0.326 \times \text{the expression of CHGA}) + (0.384 \times \text{the expression of PTGER1}) + (0.199 \times \text{the expression of SHC3}) + (0.114 \times \text{the expression of PLXNC1}) + (-0.114 \times \text{the expression of TRIM55})$. Patients were assigned to the low- or high-risk group based on the median value. As shown in the K-M curves, the high-level expression of PLXNC1 and TRIM55 showed better survival outcomes in the high-risk group, while the high-level expression of CHGA, SHC3, and PTGER1 showed poorer survival outcomes in the high-risk group (Additional file 1: Fig. S1A–E). K–M analysis showed that the low-risk group patients had significantly better survival prognosis ($P<0.001$) (Fig. 4B), which was validated in the GSE49711 dataset ($P<0.001$) (Fig. 4C) and the E-MTAB-8248 dataset ($P<0.001$) (Fig. 4D). In addition, the univariate and



multivariate Cox regression analyses indicated that the risk score served as an independent prognostic factor in NBL patients in our model (Additional file 1: Fig. S2A, B). Subsequently, we identified independent prognostic factors in the E-MTAB-8248 and GSE49711 datasets. In the E-MTAB-8248 dataset, the univariate Cox regression analysis indicated that the age, stage, MYCN, and risk score were associated with the patient's prognosis (Fig. 4E). Multivariate Cox regression analysis indicated that the age, MYCN, and risk score were associated with the prognosis (Fig. 4F). The independent prognostic factors in the GSE49711 dataset are shown in the Additional file 1: Fig. S3A, B. Accordingly, the risk score, age, and MYCN all can serve as independent prognostic factors for NBL patients.

TIDE was used to compare the immunotherapy response between the two groups, and the higher the TIDE prediction score, the higher the probability of immune evasion. The low-risk group patients had a higher dysfunction ($*p < 0.05$), while the TIDE, MSI, and exclusion showed no significant differences in the two risk groups (Fig. 4G). Our model performed well in predicting the outcomes of NBL (3-years AUC = 0.720, 5-year AUC = 0.775, 10-years AUC = 0.782) (Fig. 4H), which has been validated in the GSE49711 dataset

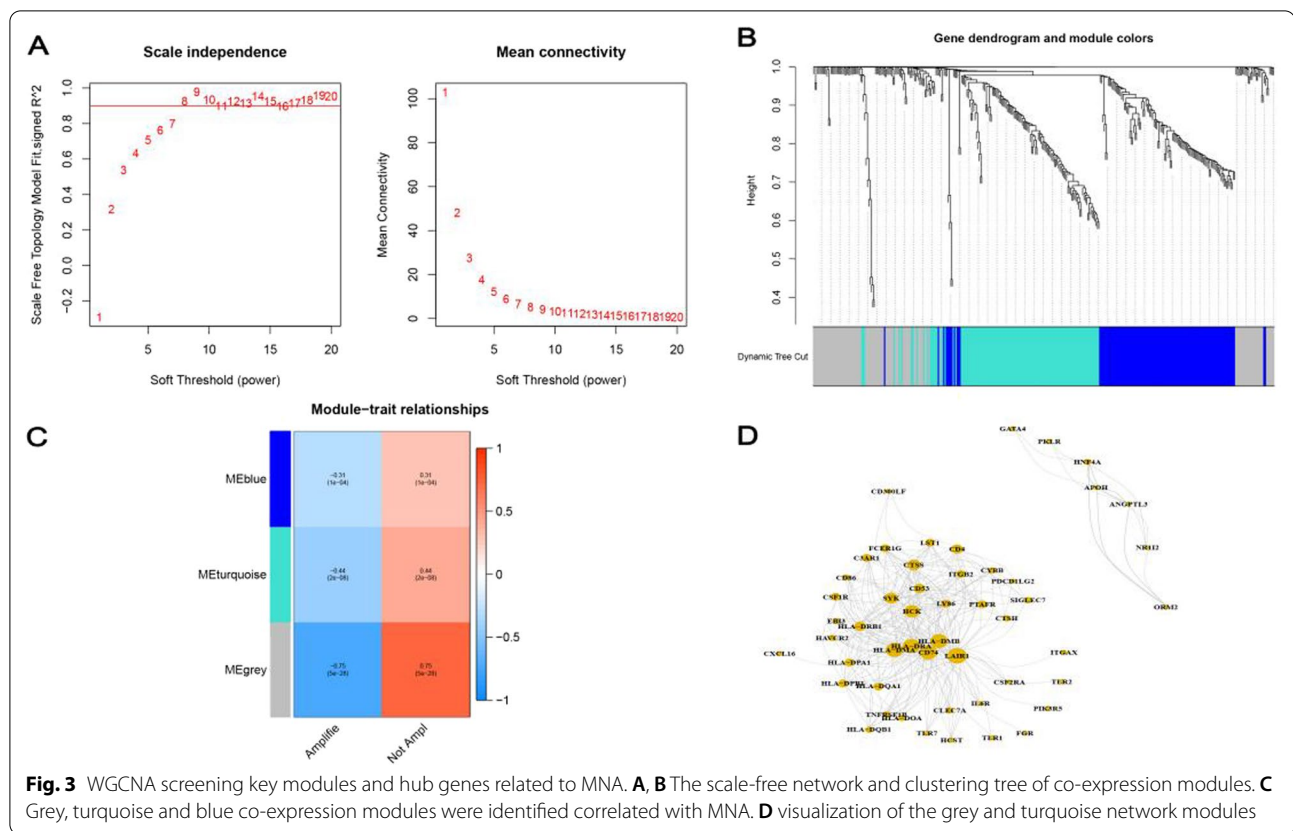
(3-years AUC = 0.784, 5-year AUC = 0.820, 10-years AUC = 0.825) (Fig. 4J) and the E-MTAB-8248 dataset (3-years AUC = 0.654, 5-year AUC = 0.603, 10-years AUC = 0.642) (Fig. 4I). The AUC values of risk model, TIDE, and TIS was 0.720, 0.493, and 0.458, respectively (Fig. 4K). Overall, the 5-MYCN-related DEIRGs signature has a better predictive ability.

Comparison of clinical features with the risk model

Age, histology, COG ($***P < 0.001$), and MKI ($*P < 0.05$) showed differences in the two risk groups (Fig. 5A). We calculated the proportions of patients in two age groups (age ≤ 1.5 , age > 1.5), and the results showed in Fig. 5B. All in all, some known clinical features of the prognosis [14] were associated with our constructed the 5-gene risk model.

Gene set enrichment analysis

We performed GSEA to explore the possible signaling pathways and mechanisms. The results showed that Cell cycle, mismatch repair, DNA replication, ribosome, and spliceosome pathway were highly enriched in the high-risk group (Fig. 5C). Complement and coagulation cascades, metabolism of xenobiotics by cytochrome p450, PPAR signaling pathway, drug metabolism cytochrome



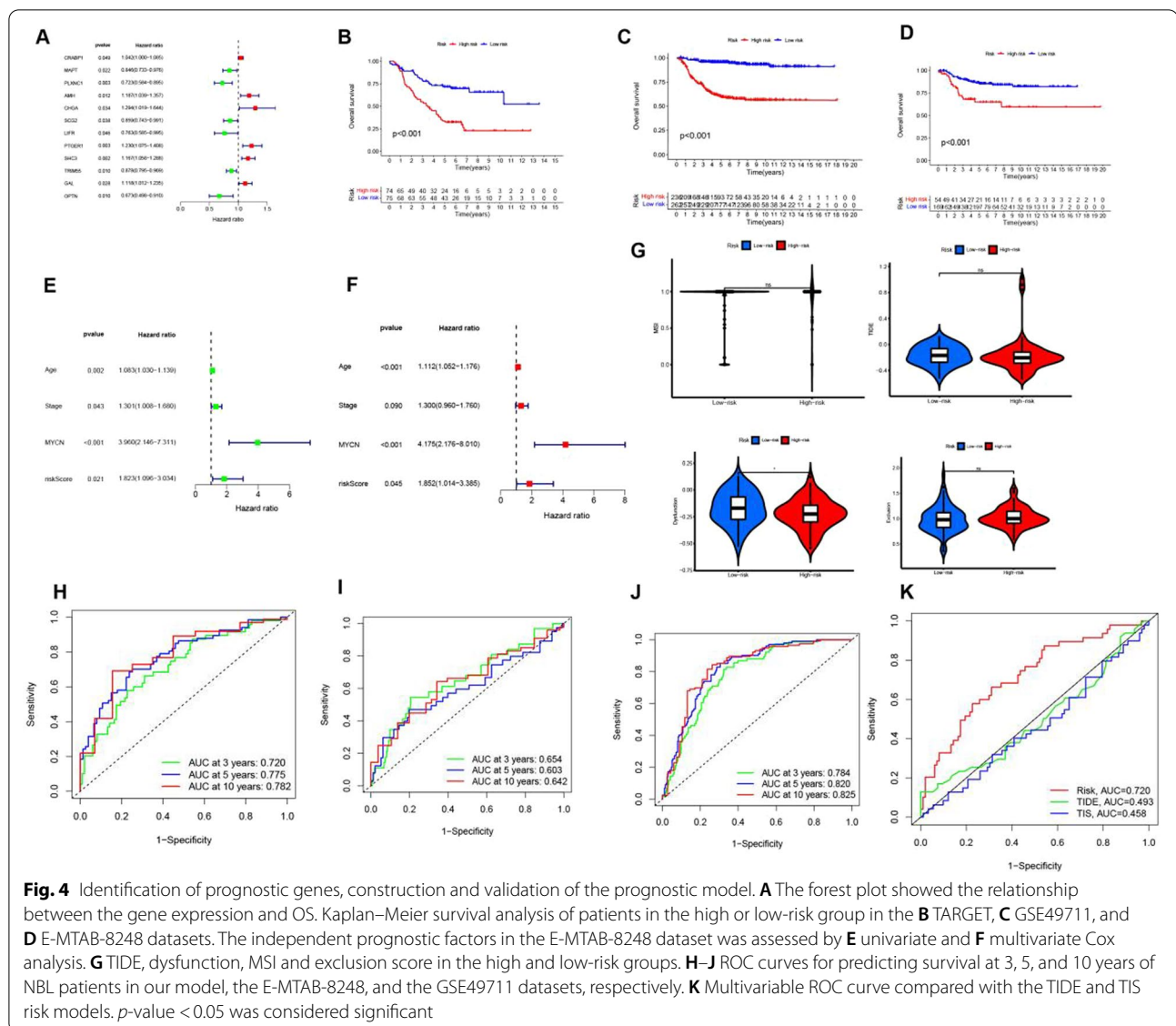
p450, and retinol metabolism were highly enriched in the low-risk group (Fig. 5D).

Correlation of tumor microenvironment with patients' prognosis

CIBERSORT was used to estimate the 22 kinds of TIICs infiltration abundance (Fig. 6A). 88 patients with $p < 0.05$ were selected for the subsequent analysis. There were no statistical differences of TIICs infiltration abundances in the two risk groups (Fig. 6B). According to the K-M analysis, the expression of four immune cells, including B cells naïve (cutpoint=0.007, $p=0.012$), macrophages M0 (cutpoint=0.384, $p=0.040$), mast cells resting (cutpoint=0.0003, $p=0.011$), and plasma cells (cutpoint=0.028, $p=0.031$) were associated with the prognosis (Fig. 6C-F). The high abundance of Macrophages M0 showed that patients had a good prognosis ($p < 0.05$) while B cells naïve, mast cells resting, and plasma cells showed a poor prognosis. Immune checkpoint gene CD274 showed a negative correlation with the risk score. ($R = -0.27$, $p = 0.001$) (Fig. 6H), and a higher expression in the low-risk group (Fig. 6G). To explore the prognostic value of the risk score for immune-checkpoint therapy, patients in the

IMvigor210 cohort were assigned to high- or low-risk groups. K-M analysis showed that the low-risk group patients had a significantly better prognosis ($P < 0.001$) (Additional file 1: Fig. S4A). We also compared the differences of immunosuppressive benefits between the two risk groups, the low-risk group patients had a higher complete response (CR)/partial response (PR) rate (Additional file 1: Fig. S4B).

4 kinds of immune functions, including APC-co-inhibition, neutrophils, T-helper cells and type-II-IFN-reponse, were suppressed in the high-risk group (Fig. 7A). According to the survival analysis, except for the immune function of CCR, the more active of the immune functions of aDCs (cutpoint=0.486, $p=0.012$), APC co stimulation (cutpoint=0.491, $p=0.035$), DCs (cutpoint=0.407, $p=0.015$), HLA (cutpoint=0.784, $p=0.013$), neutrophils (cutpoint=0.646, $p=0.006$), iDCs (cutpoint=0.414, $p=0.023$), NK cells (cutpoint=0.584, $p=0.009$), T cell co inhibition (cutpoint=0.495, $p=0.004$), parainflammation (cutpoint=0.743, $p=0.040$), T cell co stimulation (cutpoint=0.547, $p=0.028$), TIL (cutpoint=0.627, $p=0.043$), Th1 cells (cutpoint=0.455, $p=0.040$) and Treg (cutpoint=0.721, $p=0.021$), the better prognosis of NBL patients (Fig. 7B-O).



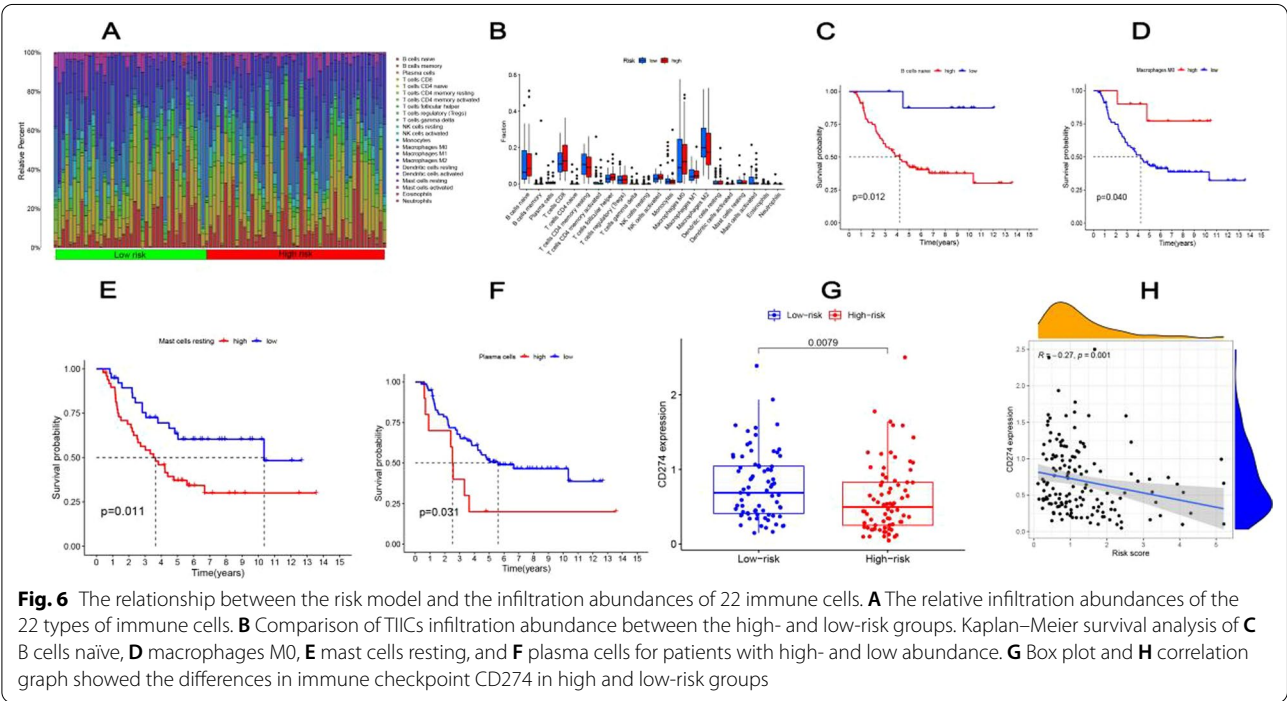
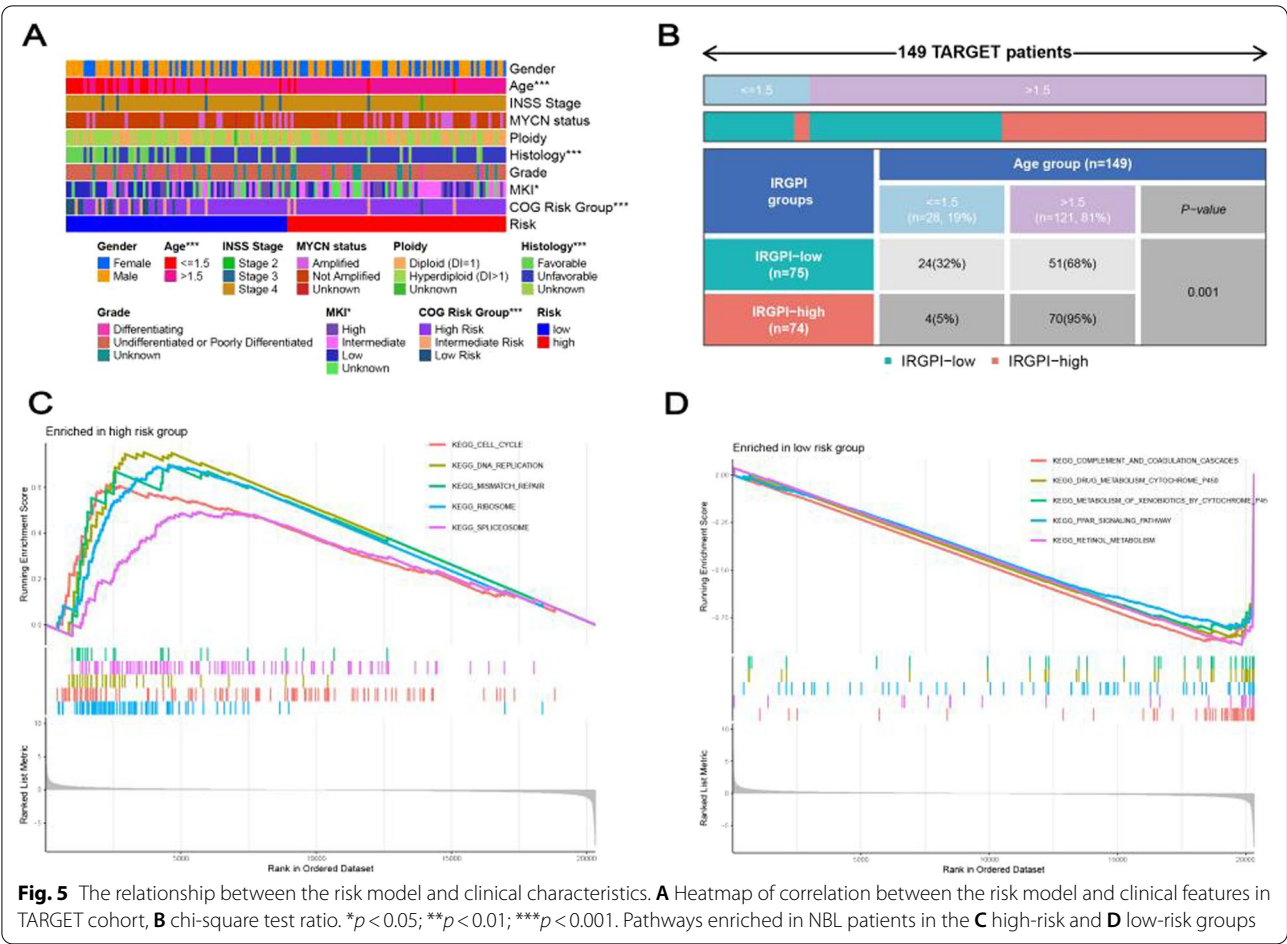
The tumor mutational burden of NBL samples

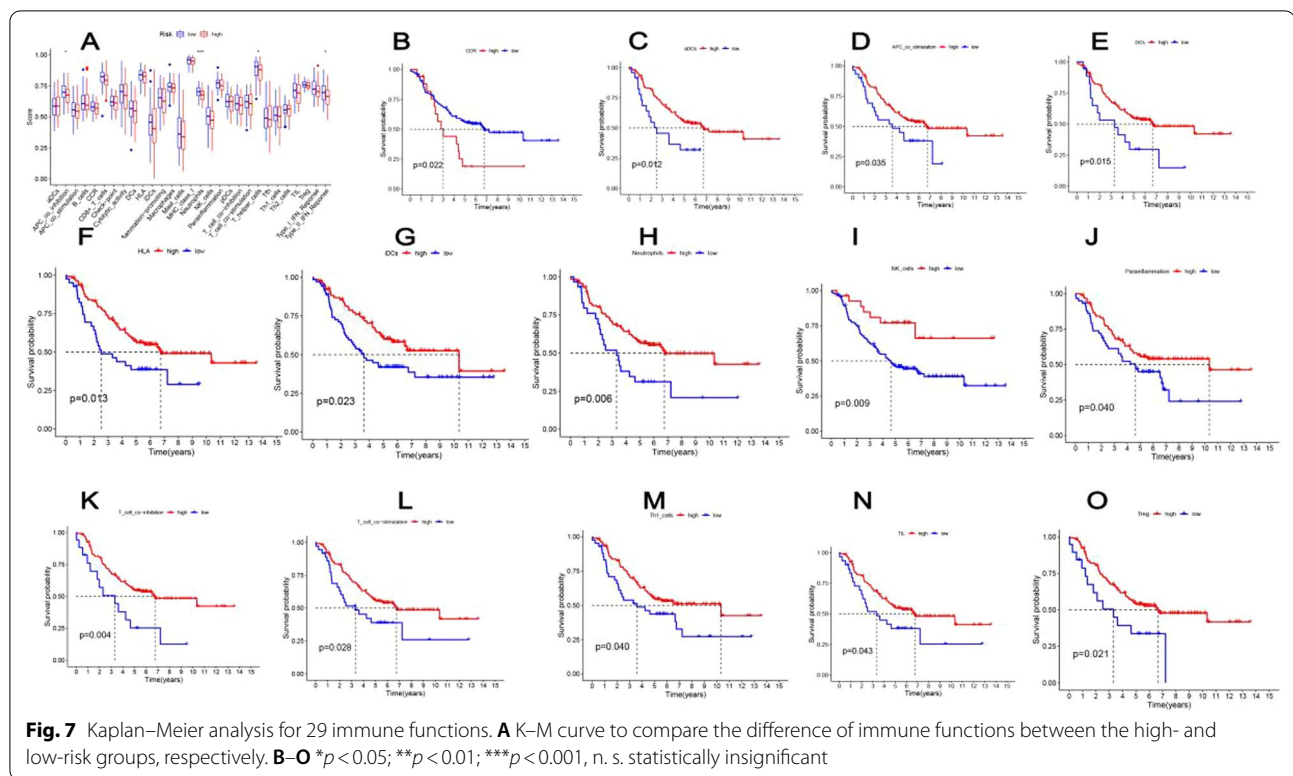
We showed the top 20 most mutated genes in somatic mutation profiles patients. The waterfall plot was used to distinguish different mutation types and show the relationship between gender and MYCN status (Fig. 8A). Missense mutations, C>A mutation, and single-nucleotide polymorphism (SNP) accounted for the majority of different classifications. In addition, counting each sample separately, the maximum mutations were 186. The box plot showed the number of variant classifications in the different samples. The top 3 mutated genes were ALK (10%), MUC16 (8%), FLG (4%) (Fig. 8B). To identify the differences of the TMB, we compared the top 20 most common mutated genes between the two risk groups. ALK, MUC16, FLG, FAT2, and ATP10B were the

commonly mutated genes in the low-risk group (Fig. 8C), while ALK, MUC16, FLG were commonly mutated genes in the high-risk group (Fig. 8D). Accordingly, ALK (10%) was the most common mutant gene in the 149 NBL patients. The boxplot and correlation graphs showed that the TMB had no significant difference in the two risk groups (Fig. 8E, F).

Discussion

NBL is an embryonal malignancy originating from neural crest cells, which is frequently characterized by the amplification and overexpression of MYCN oncogene [13, 14]. MNA is associated with high-risk disease, advanced-stage disease, metastatic behavior, rapid progression, poor survival, and unfavorable prognosis [15,



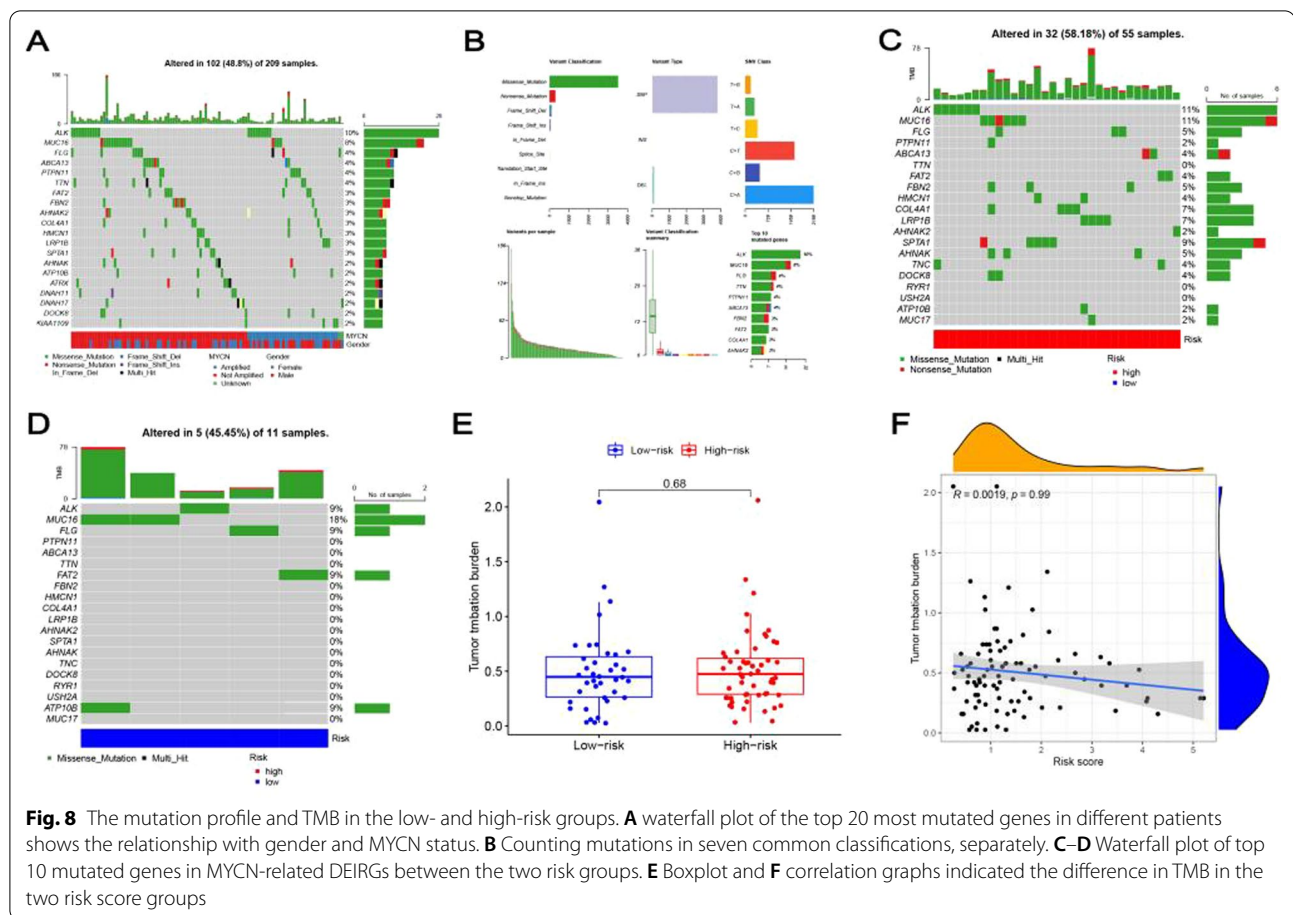


16]. The MYCN-amplified subtype constitutes the most aggressive and least treatable form of NBL. Despite intensive treatment models, the long-term survival of high-risk NBL is still around 50% [14]. Tumors with MNA have a unique gene expression profile, and the altered gene expression is considered to be the primary oncogenic function of MYCN [17]. However, targeting this oncogene remains challenging in clinical treatment. In this study, we analyzed EDGs between the two MYCN status groups. Based on the ImmPort and InnateDB databases, 343 DEIRGs, including 320 down-regulated and 23 up-regulated genes in the MNA group were extracted from DEGs based on the IRGs list. Functional enrichment analysis showed that these DEIRGs participated in leukocyte adhesion-related signaling pathways, including the regulation of leukocyte cell–cell adhesion, leukocyte cell–cell adhesion, leukocyte mediated immunity, and positive regulation of leukocyte cell–cell adhesion. Several cell adhesion molecules (CAMs) are involved in leukocyte adhesion. These CAMs are associated with the adhesion and metastatic behavior of NBL cells [18, 19]. These enriched pathways might mean that leukocyte adhesion is partially responsible for the malignant progression of NBL.

Based on the WGCNA, we screened 3 gene co-expression modules, of which the turquoise and grey modules were closely correlated with MNA ($p < 0.01$), and a total

of 228 DEIRGs were selected for the following research. Through univariate and multivariate regression analyses, we finally selected five genes (PLXNC1, CHGA, PTGER1, SHC3, and TRIM55) to construct the risk model. The survival curves showed a poorer prognosis in the high-risk group which was validated in the GSE49711 and the E-MTAB-8248 datasets. Two kinds of tumor cells immune escape mechanisms have been founded at present. TIDE was used to identify the potential factors of tumor cells immune evasion mechanisms. However, except for dysfunction, the other three TIDE, MSI, and exclusion showed no significant differences in the two risk groups. TIDE was calculated in different cancer databases. The number of our samples is limited and all of them were from the TARGET database, which may lead to bias. Larger research is needed to explore the specific mechanisms in the future.

Here, the prediction accuracy of the risk AUC was better than the TIS and TIDE. Therefore, the signature based on the five MYCN-related DEIRGs had an excellent predictive effect. The signature may be helpful to evaluate the prognosis and update treatment for NBL patients. Heterogeneous diseases have a large number of clinicopathological characteristics and risk factors. So, we should analyze whether the risk model can be an independent factor. Through univariate and multivariate analyses, we concluded that the risk score was



an independent prognosis factor in our model. However, former studies have proved that MYCN is related to poor prognosis in NBL patients. Subsequently, we performed univariate and multivariate analyses in the E-MTAB-8248 and GSE49711 datasets, the results showed that the risk score, age, and MYCN all can serve as independent prognostic factors for NBL patients. After our analysis, data collections in the TARGET database can lead to bias in our model. Thus, more datasets are needed to identify the result.

When it comes to clinical characteristics, the risk model was strongly correlated with age, histology, MKI, and COG. International Neuroblastoma Risk Group [20] uses histologic category, age, grade, MYCN status, DNA ploidy, etc., as a risk group for NBL patients, which helps to develop individualized treatment options according to different risk characteristics.

Through univariate Cox regression analysis, we selected 12 genes, including CRABP1, AMH, CHGA, PTGER1, SHC3, GAL, MAPT, PLXNC1, SCG2, LIFR, TRIM55, and OPTN. These 12 genes are associated with the progression and prognosis of cancer. For example, the expression of CRABP1 of cancers from various origins

are significantly different. High levels of CRABP1 were detected in NBL but not in NSCLC, ovarian cancer, and glioblastoma. By DNA methylation analysis, CRABP1 was identified as a hypermethylated target gene of ovarian cancer [21]. The decreased expression of CRABP1 is associated with the poor prognosis of serous and clear cell ovarian adenocarcinoma [22]. MAPT is aberrantly expression in some cancers. The aberrantly expression of MAPT is an independent prognostic factor in prostate cancer, and its knockdown can reduce cell growth [23]. Marachelian et al. [24], first reported that the mRNA quantification of 5-related genes (CHGA, DCX, DDC, PHOX2B, and TH) could be used as biomarkers for relapsed/refractory NBL. In a pan-cancer study, the overall expression of PLXNC1 is up-regulated in primary and metastatic tumors and may be associated with a more aggressive cancer phenotype [25]. SCG2 is an independent prognostic factor in colorectal cancer, which is associated with macrophage polarization and immune infiltration [26]. The expression of gene LIF/LIFR is overexpressed in solid tumors, and plays an essential role in cancer metastasis, progression, and invasion [27]. PTGER1, one of the prostaglandin receptors,

conjugates G-proteins to activate the protein kinase C. DNA methylation is a critical link in tumor transformation, PTGER1 is closely related to DNA methylation in non-functioning adrenocortical adenoma [28]. In addition, the high-level expression of PTGER1 is correlated with a poor prognosis in clear cell renal cell carcinoma [29]. SHC3 is ectopically overexpressed in various cancers and associated with their progression. The interaction between SHC3 and HIF-1 α may prevent NBL cells from hypoxia. 8 IRGs, including TRIM55, were associated with the survival and clinical features of lung squamous carcinoma (LUSC) [30]. By augmenting protein degradation of Snail1 via promoting the ubiquitination pathway, TRIM55 inhibits the malignant behavior of lung adenocarcinoma [31]. The expression of OPTN is down-regulated in urothelial carcinoma [32]. Overexpression of OPTN increases mitophagy and is related to worse prognosis in hepatocellular carcinoma (HCC) patients [33]. The expression of GAL up-regulated in neuroblastic cancers. The expression of Galanin and galanin receptors are correlated with abnormal differentiation of neuroblastic tumors [34]. However, except for MAPT, the other 11 genes have not been reported to be associated with the prognosis of NBL. MAPT microarray values are associated with MNA. High level expression of MAPT can improve the survival of NBL, which is consistent with the apoptosis-effector and proliferation genes [35]. ADAM22, GAL, KLHL13, and TWLST1 are up-regulated in ultra-high risk NBL, and they are overexpressed in patients with MNA [36]. Furthermore, GAL can be an independent prognosis factor for high-risk NBL. These twelve genes were closely associated with MNA. Some genes have been reported to be associated with cancer progression. So, further researches are needed to explore the relationship between the 12 genes and MNA and their roles in NBL prognosis.

TMB can be used as a new immunotherapy biomarker. We analyzed the mutation profiles of NBL patients to explore the potential mechanisms. The C>A mutations accounted for the majority. The three commonly mutated genes were ALK (10%), MUC16 (8%), and FLG (4%). MUC16, encoding cancer antigen 125 (CA125), promotes tumor proliferation and metastasis and can play an immunosuppressive role. ALK is the most common single-gene alteration in family NBL. Bresler et al. [37] found that ALK amplification and MNA concomitantly occurred in almost all cases, and both of them indicated that NBL had a poor prognosis. TMB is associated with the overall survival of malignant tumors. By analyzing 459 NBL patients, a higher somatic mutational burden is associated with lower survival in NBL [38]. According to the waterfall plot, there was no significant difference in TMB between the two groups, and there was no

significant correlation between TMB and the risk score. However, the top 20 most common mutated genes were different. Overall, NBL patients have low mutational frequencies. Therefore, larger data and specific researches are needed to analyze the relationship between TMB and the prognosis.

By GSEA functional enrichment analysis, the 12 MYCN-related DEIRGs may be involved in the pathogenesis of NBL. The high-risk group was closely associated with DNA replication, cell cycle, mismatch repair, ribosome, and spliceosome. In contrast, the low-risk group was correlated with complement and coagulation cascades, PPAR signaling pathway, drug metabolism of xenobiotics by cytochrome P450, drug metabolism cytochrome P450, retinol metabolism, and cell cycle. DNA replication is a potential pathway for the initiation and treatment of NBL. Studies have shown that MYC proto-oncogene participates in DNA replication through transcriptional or nontranscriptional mechanisms and can control the initiation of DNA replication [39]. Alterations of defects-related genes associated with DNA damage response (DDR) are frequently observed in the high-risk NBL [40]. MYCN is involved in regulating gene expressions related to ribosomal biogenesis. MYC transcription factors induced ribosome hyperactivity is related to the poor prognosis of NBL [41]. By RNA sequencing analysis, the high level expression of spliceosome factors is associated with high-risk NBL [42]. PPAR agonists and target the retinol signaling pathway may be involved in the treatment of NBL [43]. Ethanol Induce Cytochrome P450 protects against MPP+ induced NBL toxicity [44]. These studies have proved that these pathways can serve as potential therapeutic targets for NBL.

Growing evidence shows that the efficacy of immunotherapy is associated with TME. The dynamic balance of anti-tumor and pro-tumor immune cells lead to the malignant progression of tumors. The TIICs infiltration may affect immunotherapy. Several studies have analyzed the relationship between TIICs infiltration abundance and MNA [45]. Therefore, we explored the immune signatures of NBL. We used ssGSEA, CIBERSORT, and immune checkpoint to analyze the TME. The TIICs infiltration showed no significant differences in the two risk groups. According to the survival curves, four TIICs infiltration abundance, including the B cells naïve, macrophages M0, mast cells resting, and plasma cells were screened associated with NBL prognosis. The high abundance of B cells naïve, plasma cells, and mast cells resting were tightly related to a poor prognosis. APC-co-inhibition, neutrophils, T-helper cells and type-II-IFN-response were suppressed in the high-risk group. Except for CCR, the other 13 showed a better prognosis in a more active of immune functions. Several kinds of studies are similar

to the conclusion of this study [46, 47]. However, some reports are different [48]. Therefore, we still require further analysis to realize the relationship between immune functions and the prognosis, which may be helpful for immunotherapy. The high level expression of the immune checkpoint is related to immune cell infiltration and poor prognosis of children with solid tumors. Immunocheck-point inhibitor is a novel method of immunotherapy. In this study, the immune checkpoint gene CD274 was negatively correlated with the risk score, and the level expression of CD274 was higher in the low-risk group. The conclusion is consistent with other studies [49]. The result suggests that the low-risk group has a better therapeutic effect on anti-PD-L1 therapy. Some reports have indicated that anti-PD-L1 immunotherapy can enhance the response of T-cell to control tumor progression [50]. According to the IMvigor210 cohort, we concluded that patients in the low-risk group had a higher CR/PR rate. Thus, this immune checkpoint can be considered as a predictive marker and a potential therapeutic target.

Conclusion

Our study has several limitations. First, although we used the GEO database for external verification, it is still a retrospective study, lacking long-term follow-up. Second, the role of these five genes in NBL progression and prognosis should be confirmed using *in vivo* and *in vitro* experiments. Therefore, we still need corresponding clinical trials. In conclusion, we selected and analyzed DEIRGs between the two MYCN status and then constructed a well-performed five immune-related gene signature to predict the prognosis of NBL patients. A risk AUC indicated an excellent prediction result combined with the independent prognostic factor. GSEA indicated the potential pathways of NBL, which may help identify the underlying mechanisms. Furthermore, we explored the value of immune signatures and immune checkpoints in different risk score groups, which will help identify the prognostic indicator and immune treatment target. Collectively, these results provide new treatment strategies for NBL patients.

Abbreviations

AHSCT: Autologous hematopoietic stem cell transplantation; CA125: Cancer antigen 125; CAMs: Cell adhesion molecules; CR: Complete response; DDR: DNA damage response; DEGs: Differentially expressed genes; DEIRGs: Differentially expressed immune-related genes; FDR: False discovery rate; GEO: Gene Expression Omnibus; GO: Gene Ontology; GSEA: Gene set enrichment analysis; HCC: Hepatocellular carcinoma; IRGs: Immune-related genes; KEGG: Kyoto Encyclopedia of Genes and Genomes; K-M: Kaplan–Meier; LUSC: Lung squamous Carcinoma; MNA: MYCN amplification; NBL: Neuroblastoma; NCCs: Neural crest cells; OS: Overall survival; PR: Partial response; ROC: Receiver operating characteristic; ssGSEA: Single-sample gene set enrichment analysis; TARGET: Therapeutically Applicable Research to Generate Effective Treatments; TIDE: Tumor Immune Dysfunction and Exclusion; TILs: Tumor-infiltrating immune cells; TIS: Tumor inflammation signature; TMB: Tumor mutation

burden; TME: Tumor microenvironment; WGCNA: Weighted gene coexpression network analysis.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12920-022-01400-y>.

Additional file 1. Fig. S1. The correlation between the prognostic genes and overall survival. **Fig. S2.** Identification of the independent prognosis factor in the risk model. **Fig. S3.** Identification of the independent prognosis factor in the GSE49711 dataset. **Fig. S4.** Immune signature predicts immunotherapy benefit.

Acknowledgements

We wish to express thanks for sharing the data from TARGET, GEO and InnateDB, and the ImmPort database.

Author contributions

KM and PZ contributed equally to this work. KM and PZ contributed to the conception and methodology, validation, and formal analysis. KM contributed to the writing of the original draft. LL, YJL, and YL contributed to the review of the manuscript. LD and YX contributed to the study visualization. YW contributed to the study supervision and project administration. All authors read and approved the final manuscript.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Availability of data and materials

Our data can be found in the Therapeutically Applicable Research to Generate Effective Treatments (TARGET) database (www.ocg.cancer.gov/programs/target). The Gene Expression Omnibus (GEO, www.ncbi.nlm.nih.gov/geo/, GSE49711) database. The ArrayExpress (www.ebi.ac.uk/arrayexpress, E-MTAB-8248) database. The IMvigor210 cohort (<http://research-pub.gene.com/IMvigor210CoreBiologies>). The ImmPort database (www.immport.org/home) and the InnateDB database (www.innatedb.ca).

Declarations

Ethics approval and consent to participate

Our study did not require an ethical board approval because all data are from public databases, including TARGET, GEO, ArrayExpress and ImmPort, and the InnateDB database. The patients involved in the database have obtained ethical approval. Users can download relevant data for free for research and publish relevant articles. Our study is based on open source data, so there are no ethical issues and other conflicts of interest.

Consent for publication

Not applicable.

Competing interests

The authors confirm no conflict of interest related to the manuscript.

Author details

¹Department of Pediatrics, The Fifth Affiliated Hospital of Zhengzhou University, No. 3 Kangfuqian Street, Zhengzhou, China. ²Department of Neonatology, Children's Hospital, Capital Institute of Pediatrics, Beijing, China.

Received: 15 May 2022 Accepted: 18 November 2022

Published online: 23 November 2022

References

1. Mlakar V, Jurkovic Mlakar S, Lopez G, Maris JM, Ansari M, Gumy-Pause F. 11q deletion in neuroblastoma: a review of biological and clinical implications. *Mol Cancer*. 2017;16(1):114.

2. Delloye-Bourgeois C, Castellani V. Hijacking of embryonic programs by neural crest-derived neuroblastoma: from physiological migration to metastatic dissemination. *Front Mol Neurosci*. 2019;12:52.
3. Thompson D, Vo KT, London WB, et al. Identification of patient subgroups with markedly disparate rates of MYCN amplification in neuroblastoma: a report from the International Neuroblastoma Risk Group project: Predictors of MNA in Neuroblastoma. *Cancer*. 2016;122(6):935–45.
4. Matthay KK, Maris JM, Schleiermacher G, et al. Neuroblastoma. *Nat Rev Dis Primers*. 2016;2:16078.
5. Maris JM, Hogarty MD, Bagatell R, Cohn SL. Neuroblastoma. *Lancet*. 2007;369(9579):2106–20.
6. Brodeur GM, Seeger RC, Schwab M, Varmus HE, Bishop JM. Amplification of N-myc in untreated human neuroblastomas correlates with advanced disease stage. *Science*. 1984;224(4653):1121–4.
7. Huang M, Weiss WA. Neuroblastoma and MYCN. *Cold Spring Harbor Perspect Med*. 2013;3(10):a014415–a014415.
8. Li B, Cui Y, Diehn M, Li R. Development and validation of an individualized immune prognostic signature in early-stage nonsquamous non-small cell lung cancer. *JAMA Oncol*. 2017;3(11):1529.
9. Cao J, Yang X, Li J, et al. Screening and identifying immune-related cells and genes in the tumor microenvironment of bladder urothelial carcinoma: based on TCGA database and bioinformatics. *Front Oncol*. 2020;9:1533.
10. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. 2000;28(1):27–30.
11. Kanehisa M. Toward understanding the origin and evolution of cellular organisms. *Protein Sci*. 2019;28(11):1947–51.
12. Kanehisa M, Furumichi M, Sato Y, Ishiguro-Watanabe M, Tanabe M. KEGG: integrating viruses and cellular organisms. *Nucleic Acids Res*. 2021;49(D1):D545–51.
13. Maris JM. Recent advances in neuroblastoma. *N Engl J Med*. Published online 2010:10.
14. Pinto NR, Applebaum MA, Volchenboum SL, et al. Advances in risk classification and treatment strategies for neuroblastoma. *JCO*. 2015;33(27):3008–17.
15. Association of multiple copies of the N-myc oncogene with rapid progression of neuroblastomas - PubMed. Accessed 27 Jan 2022. <https://pubmed.ncbi.nlm.nih.gov/4047115/>.
16. Schwab M, Westermann F, Hero B, Berthold F. Neuroblastoma: biology and molecular and chromosomal pathology. *Lancet Oncol*. 2003;4(8):472–80.
17. Westermann F, Muth D, Benner A, et al. Distinct transcriptional MYCN/c-MYC activities are associated with spontaneous regression or malignant progression in neuroblastomas. *Genome Biol*. 2008;9(10):R150.
18. Yoon KJ, Danks MK. Cell adhesion molecules as targets for therapy of neuroblastoma. *Cancer Biol Ther*. 2009;8(4):306–11.
19. Schwankhaus N, Gathmann C, Wicklein D, Riecken K, Schumacher U, Valentiner U. Cell adhesion molecules in metastatic neuroblastoma models. *Clin Exp Metastasis*. 2014;31(4):483–96.
20. Cohn SL, Pearson ADJ, London WB, et al. The international neuroblastoma risk group (INRG) classification system: an INRG Task Force Report. *JCO*. 2009;27(2):289–97.
21. Wu Q, Lothe RA, Ahlquist T, et al. DNA methylation profiling of ovarian carcinomas and their in vitro models identifies HOXA9, HOXB5, SCGB3A1, and CRABP1 as novel targets. *Mol Cancer*. 2007;6(1):45.
22. Miyake T, Ueda Y, Matsuzaki S, et al. CRABP1-reduced expression is associated with poorer prognosis in serous and clear cell ovarian adenocarcinoma. *J Cancer Res Clin Oncol*. 2011;137(4):715–22.
23. Yang J, Yu Y, Liu W, Li Z, Wei Z, Jiang R. Microtubule-associated protein tau is associated with the resistance to docetaxel in prostate cancer cell lines. *RRU*. 2017;9:71–7.
24. Marachelian A, Villablanca JG, Liu CW, et al. Expression of five neuroblastoma genes in bone marrow or blood of patients with relapsed/refractory neuroblastoma provides a new biomarker for disease and prognosis. *Clin Cancer Res*. 2017;23(18):5374–83.
25. Zhang X, Shao S, Li L. Characterization of class-3 semaphorin receptors, neuropilins and plexins, as therapeutic targets in a pan-cancer study. *Cancers*. 2020;12(7):1816.
26. Wang H, Yin J, Hong Y, et al. SCG2 is a prognostic biomarker associated with immune infiltration and macrophage polarization in colorectal cancer. *Front Cell Dev Biol*. 2022;9:795133.
27. Zhang C, Liu J, Wang J, Hu W, Feng Z. The emerging role of leukemia inhibitory factor in cancer and therapy. *Pharmacol Ther*. 2021;221:107754.
28. Itcho K, Oki K, Kobuke K, et al. Aberrant G protein-receptor expression is associated with DNA methylation in aldosterone-producing adenoma. *Mol Cell Endocrinol*. 2018;461:100–4.
29. Liao Z, Yao H, Wei J, et al. Development and validation of the prognostic value of the immune-related genes in clear cell renal cell carcinoma. *Transl Androl Urol*. 2021;10(4):1607–19.
30. Li N, Wang J, Zhan X. Identification of immune-related gene signatures in lung adenocarcinoma and lung squamous cell carcinoma. *Front Immunol*. 2021;12:752643.
31. Guo T, Zhang Z, Zhu L, et al. TRIM55 suppresses malignant biological behavior of lung adenocarcinoma cells by increasing protein degradation of Snail1. *Cancer Biol Ther*. 2022;23(1):17–26.
32. Li J, Abraham S, Cheng L, et al. Proteomic-based approach for biomarkers discovery in early detection of invasive urothelial carcinoma. *Prot Clin Appl*. 2008;2(1):78–89.
33. Inokuchi S, Yoshizumi T, Toshima T, et al. Suppression of optineurin impairs the progression of hepatocellular carcinoma through regulating mitophagy. *Cancer Med*. 2021;10(5):1501–14.
34. Perel Y, Amrein L, Dobremez E, Rivel J, Daniel JY, Landry M. Galanin and galanin receptor expression in neuroblastic tumours: correlation with their differentiation status. *Br J Cancer*. 2002;86(1):117–22.
35. Zaman S, Chobrutskiy BI, Blanck G. MAPT (Tau) expression is a biomarker for an increased rate of survival in pediatric neuroblastoma. *Cell Cycle*. 2018;17(21–22):2474–83.
36. Z L, Cn G, L S, Ba M, Vs S, Hg W. Expression patterns of immune genes reveal heterogeneous subtypes of high-risk neuroblastoma. *Cancers*. 2020;12(7).
37. Bresler SC, Weiser DA, Huwe PJ, et al. ALK mutations confer differential oncogenic activation and sensitivity to ALK inhibition therapy in neuroblastoma. *Cancer Cell*. 2014;26(5):682–94.
38. Hwang WL, Wolfson RL, Niemierko A, Marcus KJ, DuBois SG, Haas-Kogan D. Clinical impact of tumor mutational burden in neuroblastoma. *JNCI: J Natl Cancer Inst*. 2019;111(7):695–9.
39. Dominguez-Sola D, Gautier J. MYC and the control of DNA replication. *Cold Spring Harb Perspect Med*. 2014;4(6):a014423–a014423.
40. Southgate HED, Chen L, Curtin NJ, Tweddle DA. Targeting the DNA damage response for the treatment of high risk neuroblastoma. *Front Oncol*. 2020;10:371.
41. Hald ØH, Olsen L, Gallo-Oller G, et al. Inhibitors of ribosome biogenesis repress the growth of MYCN-amplified neuroblastoma. *Oncogene*. 2019;38(15):2800–13.
42. Shi Y, Yuan J, Rrakli V, et al. Aberrant splicing in neuroblastoma generates RNA-fusion transcripts and provides vulnerability to spliceosome inhibitors. *Nucleic Acids Res*. 2021;49(5):2509–21.
43. Kim EJ, Park KS, Chung SY, et al. Peroxisome proliferator-activated receptor-γ activator 15-deoxy-Δ^{12,14}-prostaglandin J₂ inhibits neuroblastoma cell growth through induction of apoptosis: association with extracellular signal-regulated kinase signal pathway. *J Pharmacol Exp Ther*. 2003;307(2):505–17.
44. Fernandez-Abascal J, Ripullone M, Valeri A, Leone C, Valoti M. β-Naphthoflavone and ethanol induce cytochrome P450 and protect towards MPP⁺ toxicity in human neuroblastoma SH-SY5Y cells. *IJMS*. 2018;19(11):3369.
45. Wienke J, Dierselhuis MP, Tytgat GAM, Künkele A, Nierkens S, Molenaar JJ. The immune landscape of neuroblastoma: challenges and opportunities for novel therapeutic strategies in pediatric oncology. *Eur J Cancer*. 2021;144:123–50.
46. Yang Y, He W, Wang ZR, et al. Immune cell landscape in gastric cancer. Zhang L, ed. *BioMed Research International*. 2021;2021:1–12.
47. Melaiu O, Chierici M, Lucarini V, et al. Cellular and gene signatures of tumor-infiltrating dendritic cells and natural-killer cells predict prognosis of neuroblastoma. *Nat Commun*. 2020;11(1):5992.
48. Schaafsma E, Jiang C, Cheng C. B cell infiltration is highly associated with prognosis and an immune-infiltrated tumor microenvironment in neuroblastoma. *J Cancer Metastasis Treat*. 2021;7(34).
49. Kang W, Hu J, Zhao Q, Song F. Identification of an autophagy-related risk signature correlates with immunophenotype and predicts immune checkpoint blockade efficacy of neuroblastoma. *Front Cell Dev Biol*. 2021;9:731380.

50. Cha JH, Chan LC, Li CW, Hsu JL, Hung MC. Mechanisms controlling PD-L1 expression in cancer. *Mol Cell*. 2019;76(3):359–70.

Publisher's Note

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

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

