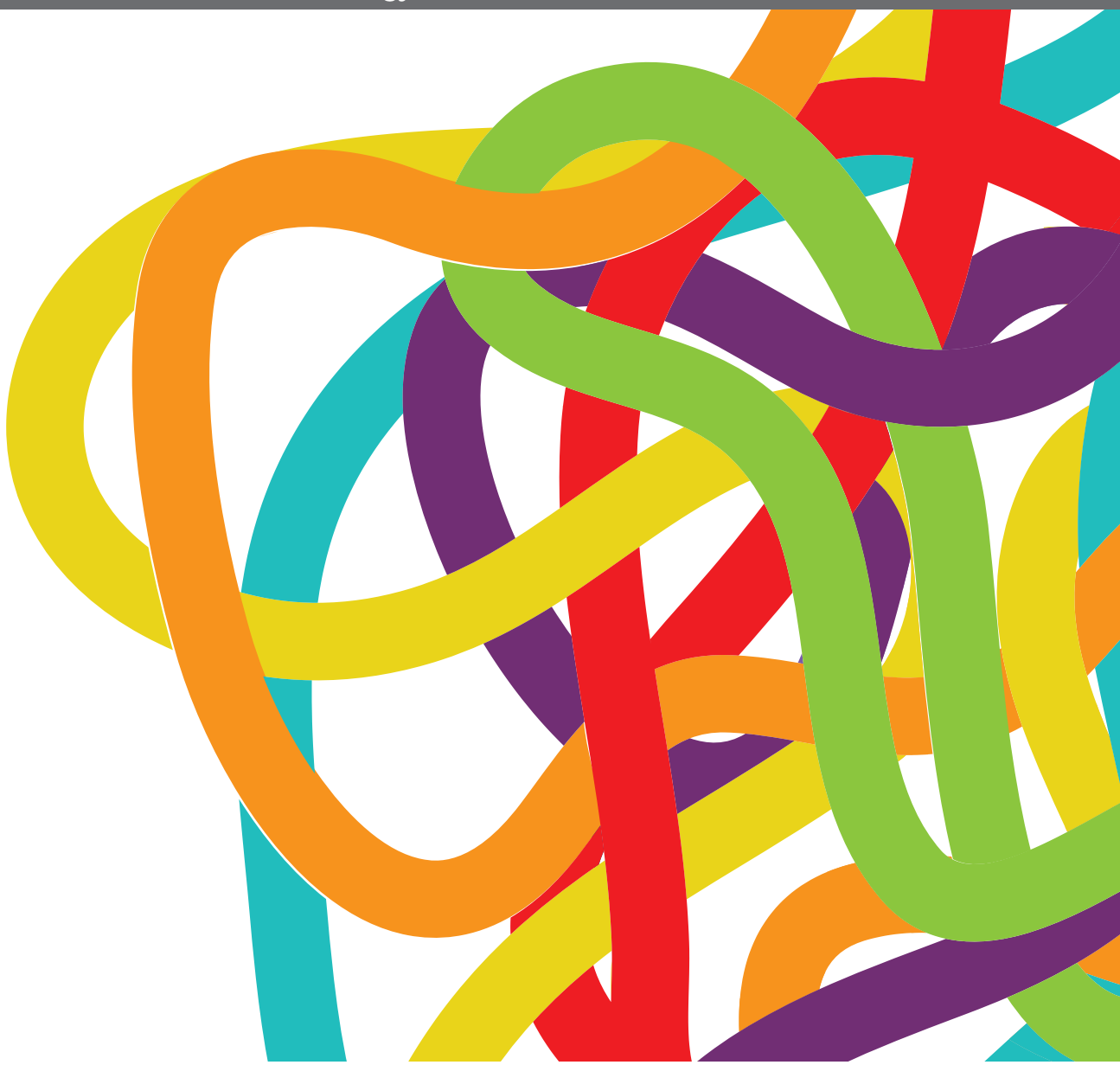
MOLECULAR DIAGNOSTICS OF PEDIATRIC CANCER

EDITED BY: Jing He, Yizhuo Zhang, Jinhong Zhu, Hua Tan and Jochen Rössler

PUBLISHED IN: Frontiers in Oncology and Frontiers in Pediatrics



frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-88974-300-1
DOI 10.3389/978-2-88974-300-1

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

MOLECULAR DIAGNOSTICS OF PEDIATRIC CANCER

Topic Editors:
Jing He, Guangzhou Medical University, China
Yizhuo Zhang, Sun Yat-sen University Cancer Center (SYSUCC), China
Jinhong Zhu, Harbin Medical University Cancer Hospital, China
Hua Tan, National Human Genome Research Institute (NHGRI), United States
Jochen Rössler, Insel Gruppe AG, Switzerland

Citation: He, J., Zhang, Y., Zhu, J., Tan, H., Rössler. J., eds. (2022). Molecular Diagnostics of Pediatric Cancer. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88974-300-1

Table of Contents

frontiers in Oncology

EDITORIAL
published: 11 October 2021
doi: 10.3389/fonc.2021.777662

Editorial: Molecular Diagnostics of Pediatric Cancer

Jing He[1], Yizhuo Zhang[2], Jinhong Zhu[3], Hua Tan[4] and Jochen Rössler[5]*

[1] Department of Pediatric Surgery, Guangzhou Institute of Pediatrics, Guangdong Provincial Key Laboratory of Research in Structural Birth Defect Disease, Guangzhou Women and Children's Medical Center, Guangzhou Medical University, Guangzhou, China, [2] Department of Pediatric Oncology, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Sun Yat-Sen University Cancer Center, Guangzhou, China, [3] Department of Clinical Laboratory, Biobank, Harbin Medical University Cancer Hospital, Harbin, China, [4] School of Biomedical Informatics, The University of Texas Health Science Center at Houston, Houston, TX, United States, [5] Division of Pediatric Hematology and Oncology, Department of Pediatrics, Inselspital, Bern University Hospital, University of Bern, Bern, Switzerland

Keywords: pediatric cancer, prognostic, biomarker, target, drug development

Editorial on the Research Topic
Molecular Diagnostics of Pediatric Cancer

Pediatric tumors are defined as tumors arising from the complex physiological growth process of embryonic stem cells (1). They differ from malignant adult tumors in cellular origin, epidemiology, genetic complexity, driver mutations, and potential mutational processes, and they are generally considered to be rare events (2). This Research Topic collects research related to molecular markers, signaling pathways, drug development and treatment, and emerging molecular technologies of pediatric tumors.

Chen et al. reviewed the progress of molecular epidemiology of hepatoblastoma (HB), focusing on the studies of single nucleotide polymorphisms (SNPs) related to the risk of HB. As treatment regimens for medulloblastoma (MB) are becoming subgroup-specific, methods are needed to discriminate its subgroups. Gershanov et al. used the SARC algorithm that reduces the set of 22 genes to only 6 genes, which could distinguish four MB subgroups reliably. The gene set identified is small enough to allow clinicians to easily obtain the qPCR-based classification of MB subtypes to better determine treatment options. Wang et al. found that the sensitivity of the NB5 method to detect neuroblastoma (NB) with micrometastases in bone marrow (BM) and peripheral blood (PB) was significantly higher than that of bone marrow biopsy (BMB). Liver and bone metastases are factors that affect the sensitivity of NB5 detection in the bone marrow and peripheral blood. Zhanghuang et al. illustrated that targeting the PI3K-AKT signaling pathway and microRNA-related proteins had high potential values for treating malignant rhabdoid tumors of the kidney (MRTK). Poot et al. described recent advances in the therapeutical development of pediatric cancer and illustrates how this methodology affects diagnosis and provides additional treatment options for these patients. These studies contribute to a better understanding, diagnosis, and treatment of pediatric cancer.

Pediatric cancers are characterized by high molecular heterogeneity. For instance, *CTNNB1*, *NFE2L2*, *AXIN1*, *APC*, *MYCN1*, and *IGF2* may be potential biomarkers for the diagnosis of HB. Hu et al. demonstrated that pediatric HB patients with causal genetic alterations had significantly lower complete remission (CR) rates than patients with wide-type gene counterparts ($P<0.05$). Moreover, regarding acute lymphoblastic leukemia (ALL), Liu et al. found that *METTL3* gene polymorphism was associated with an increased risk of ALL in children and suggested that *METTL3* gene polymorphism may be a potential biomarker for the selection of chemotherapy agents for pediatric ALL. Cai et al. proved that Prp19 regulates the expression of YAP through YAP pre-mRNA splice, thus

OPEN ACCESS

Edited and reviewed by:
Rimas J. Orentas, Seattle Children's Research Institute, United States

****Correspondence:***
Jing He
hejing198374@gmail.com

Specialty section:
This article was submitted to Pediatric Oncology, a section of the journal Frontiers in Oncology

Received: *15 September 2021*
Accepted: *24 September 2021*
Published: *11 October 2021*

Citation:
He J, Zhang Y, Zhu J, Tan H and Rössler J (2021) Editorial: Molecular Diagnostics of Pediatric Cancer. Front. Oncol. 11:777662. doi: 10.3389/fonc.2021.777662

affecting the invasion, migration, and EMT of NB cells. It was the first report to demonstrate that Prp19 is a potential therapeutic target and prognostic biomarker in patients with NB. Shi et al. showed that the high expression of CDC20 was involved in the tumorigenesis of Wilms tumor (WT), and inhibition of CDC20 could suppress the proliferation and migration of WT cells and arrest the cell cycle in the G2/M phase, suggesting that CDC20 could be a potential biomarker of WT. Liu et al. established a multinomial predictive survival model and a survival-associated ceRNA network, which provides a new potential biomarker for improving prognosis and treatment of WT patients. Taken together, these biomarkers may be able to predict clinical outcomes and hold great promise in clinical application of pediatric cancer.

The immune system is closely related to the occurrence and development of pediatric cancer, and understanding the immune microenvironment is helpful to the treatment of pediatric cancer (3). Li et al. used single-cell RNA sequences to reveal the characteristics of malignant cells and the immune microenvironment in subcutaneous panniculitis-like T-cell lymphoma (SPTCL), providing a better understanding of the transcriptional characteristics and immune microenvironment of this rare tumor. Feng et al. explored the immune microenvironment of Langerhans cell histiocytosis (LCH). They found that serum levels of immune indicators are somewhat representative of disease severity, and associated laboratory tests can be used to improve risk stratification and guide immunotherapy.

The rapid rise of gene sequencing and bioinformatics and the opening of relevant tumor databases provide opportunities to elucidate the molecular mechanisms of pediatric cancer and precise drug target therapy of pediatric cancer. Feng et al. applied artificial intelligence methods to improve the accuracy of gene express-based survival prediction for neuroblastoma. Ruan et al. showed that monitoring circulating tumor DNA (ctDNA) with next-generation sequencing-based analysis could provide more information about genetic mutations to guide the precise treatment of acute myeloid leukemia (AML) in children. Sun et al. established a random forest classifier and identified 10 HB core genes. These findings may help in the diagnosis, prediction, and targeted treatment of HB. Li et al. provided an overview of the techniques currently available *in vitro* and *in vivo* models of pediatric brain tumors and discussed the opportunities presented by new techniques such as 3D culture and organic-like compounds that can overcome the limitations of the simplicity of single-layer culture and the complexity of living models to accommodate greater precision in drug development for pediatric brain tumors. Wang et al. reported the first case of acute promyelocytic with *FIP1L1/RARA* identified by next-generation sequencing (NGS). NGS analysis is recommended as a routine test for patients with variant acute promyelocytic leukemia (APL). Cimmino et al. found that 9 out of 11 patients carried at least one pathogenic variant and developed a targeted NGS approach to identify tumor-specific alterations in ctDNA in NB patients. This information can be combined with clinical and pathological data at NB diagnosis. The goal of these molecular diagnostic studies for pediatric cancer is to translate them into the clinic to achieve more accurate diagnosis, more accurate risk stratification, and more effective and less toxic treatments.

In conclusion, the "Molecular Diagnostics of Pediatric Cancer" Research Topic highlights the most recent advance of diagnostic molecular biomarkers and novel therapeutic targets for pediatric cancer.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

FUNDING

This study was supported by a grant from the National Natural Science Foundation of China (No. 82173593).

REFERENCES

1. Jones DTW, Banito A, Grunewald TGP, Haber M, Jager N, Kool M, et al. Molecular Characteristics and Therapeutic Vulnerabilities Across Paediatric Solid Tumours. *Nat Rev Cancer* (2019) 19(8):420–38. doi: 10.1038/s41568-019-0169-x
2. Kattner P, Strobel H, Khoshnevis N, Grunert M, Bartholomae S, Pruss M, et al. Compare and Contrast: Pediatric Cancer *Versus* Adult Malignancies. *Cancer Metastasis Rev* (2019) 38(4):673–82. doi: 10.1007/s10555-019-09836-y
3. Miao L, Zhuo Z, Tang J, Huang X, Liu J, Wang HY, et al. FABP4 Deactivates NF-kappaB-IL1alpha Pathway by Ubiquitinating ATPB in Tumor-Associated Macrophages and Promotes Neuroblastoma Progression. *Clin Transl Med* (2021) 11(4):e395. doi: 10.1002/ctm2.395

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

ORIGINAL RESEARCH
published: 02 November 2020
doi: 10.3389/fonc.2020.575366

Prp19 Is an Independent Prognostic Marker and Promotes Neuroblastoma Metastasis by Regulating the Hippo-YAP Signaling Pathway

Yuanxia Cai[1,2†], *Kai Chen*[1,2†], *Cheng Cheng*[1,2], *Yonghu Xu*[3], *Qianqian Cheng*[1,2], *Guofeng Xu*[3], *Yeming Wu*[1,2,4*] *and Zhixiang Wu*[1,2,4*]

[1] *Department of Pediatric Surgery, Xinhua Hospital, School of Medicine, Shanghai Jiaotong University, Shanghai, China,* [2] *Division of Pediatric Oncology, Shanghai Institute of Pediatric Research, Shanghai, China,* [3] *Department of Pediatric Urology, Xinhua Hospital, National Key Clinical Specialty, Shanghai Top-Priority Clinical Center, School of Medicine, Shanghai Jiaotong University, Shanghai, China,* [4] *Department of Pediatric Surgery, Children's Hospital of Soochow University, Suzhou, China*

OPEN ACCESS

Edited by:
Jing He,
Guangzhou Medical University, China

Reviewed by:
Lisha Zhu,
University of Texas Health Science Center at Houston, United States
Ying Wang,
Jiangnan University, China

***Correspondence:**
Yeming Wu
wuyeming@xinhuamed.com.cn
Zhixiang Wu
wuzhixiang@xinhuamed.com.cn

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Pediatric Oncology, a section of the journal Frontiers in Oncology

Received: *23 June 2020*
Accepted: *19 August 2020*
Published: *02 November 2020*

Citation:
Cai Y, Chen K, Cheng C, Xu Y, Cheng Q, Xu G, Wu Y and Wu Z (2020) Prp19 Is an Independent Prognostic Marker and Promotes Neuroblastoma Metastasis by Regulating the Hippo-YAP Signaling Pathway. Front. Oncol. 10:575366. doi: 10.3389/fonc.2020.575366

Pre-mRNA processing factor 19 (Prp19) was previously reported to be involved in tumor progression. However, Prp19 expression and its functions remain elusive in neuroblastoma. Here, we aim to identify the functions and mechanisms of Prp19 in neuroblastoma. Neuroblastic tumor tissue microarrays and two independent validation data sets indicate that Prp19 is associated with high-risk markers and bone marrow metastasis and serves as a prognostic marker for worse clinical outcomes with neuroblastoma. Gain- and loss-of-expression assays reveal that Prp19 promotes invasion, migration, and epithelial–mesenchymal transition (EMT) of neuroblastoma cells *in vitro*. Bioinformatics analysis of RNA-seq data shows that the expressions of YAP and its downstream genes are significantly inhibited after downregulation of Prp19. Prp19 and YAP expression in metastatic lymph nodes is higher than *in situ* neuroblastoma tissue. Further experiments show that Prp19 regulates YAP expression and consequently affects cell invasion, migration, and EMT in neuroblastoma by pre-mRNA splicing of YAP. In conclusion, our findings provide the first evidence that Prp19 is a potential therapeutic target and prognostic biomarker for patients with neuroblastoma.

Keywords: Prp19, neuroblastoma, metastasis, YAP, RNA splicing

INTRODUCTION

Neuroblastoma, arising from neural crest progenitor cells of the sympathoadrenal lineage, is the most common extracranial solid tumor in children, accounting for 7.5% of all childhood cancers and 11–15% of all childhood cancer-related deaths (1, 2). The most prominent characteristic of neuroblastoma is extreme heterogeneity, which ranges from spontaneous regression in infants to metastasis and progression in older children despite intensive multimodality therapy. Children with limited lesions usually have a good prognosis, and the 5-year event-free survival (EFS) rate can reach 83% (3). However, the long-term prognosis for patients with distant metastasis is poor with a 5-year EFS of 35%, which has not improved in the last decades (3–5). A better understanding of the biological mechanism of metastatic neuroblastoma will likely refine treatment strategies and further improve the prognosis of metastatic patients.

Pre-mRNA processing factor 19 (Prp19)-associated complex (Prp19C) is a highly conserved multiprotein complex (6). As an important member of Prp19C (7), Prp19 has many vital biological functions, such as splicing of pre-mRNA, cell cycle regulation, DNA damage repair, protein degradation, and metastasis (8–13). Because these intracellular events are closely related to cell fate, aberrant Prp19 function may cause serious diseases, for example, cancer. In fact, studies have found Prp19 expression is higher in colon, larynx, and hepatocellular carcinoma compared with paracancerous tissues and is positively correlated with poor prognosis (13, 14). However, the function of Prp19 in tumors and its association with poor prognosis have not been well-demonstrated. Increasing evidence suggests that abnormalities in splicing events may involve tumorigenesis and development. For example, hnRNPK promotes gastric tumorigenesis through regulating CD44E alternative splicing (15). Alternative splicing of EZH2 pre-mRNA by SF3B3 contributes to the tumorigenic potential of renal cancer (16). As an important splicing factor, Prp19 may also participate in tumorigenesis and development by regulating alternative splicing of some key molecules.

The Hippo-YAP pathway is a highly conserved signaling pathway that functions in the regulation of organ size, cell proliferation, invasion, and metastasis (17, 18). When the Hippo-YAP signaling is off, YAP, the key factor in the Hippo pathway, translocates into the nucleus to drive transcription of downstream genes, promoting of cell proliferation, migration, and tumor growth. When Hippo-YAP is on, YAP is phosphorylated and retained in the cytoplasm, turning off downstream target gene expression (17). In addition to the classic Hippo-YAP signaling, YAP expression and stability can be regulated by a variety of molecules, such as CD44, HER3, and Fbxw7 (19–22). The mechanism governing YAP protein expression by alterative splicing remains poorly understood. This study presents a lot of overlap between the functions of Prp19 and YAP. On account of this, whether there is a regulatory effect between Prp19 and YAP is worth further exploration.

Our investigation into the expression and potential role and mechanism of Prp19 in neuroblastoma reveals that the expression of Prp19 is positively correlated with bone marrow metastasis of neuroblastoma and that Prp19 promotes neuroblastoma invasion and migration *in vitro*. Also, YAP is identified as a candidate target of Prp19 because the mRNA maturity of it is regulated by Prp19. In metastatic lymph nodes, Prp19 and YAP show obvious higher expression than their paired primary tumor. Taken together, these results indicate that Prp19 promotes neuroblastoma metastasis via increasing pre-mRNA splicing to upregulate the level of YAP.

MATERIALS AND METHODS

Patients and Tissue Specimens

This study includes 62 pediatric patients with neuroblastoma who were diagnosed and treated in Xinhua Hospital, a subsidiary hospital of Shanghai Jiaotong University School of Medicine, from September 2012 to February 2015. The patient group includes 43 neuroblastoma/gneuroblastoma-N (NB/GNB-N) and 19 gneuroblastoma-I (GNB-I) cases. A total of 4 pairs of neuroblastoma *in situ* and their corresponding metastatic lymph node tissues were collected. The clinical and prognosis information of the patients was collected from recorded clinical data and follow-up through phone calls; data include gender, age at diagnosis, bone marrow metastasis, clinical stage, diagnostic category, Shimada pathologic type, and risk classification. All patients underwent surgery or biopsy in our hospital, and each tumor specimen was stored in liquid nitrogen until tissue microarray (TMA) analysis. The experimental protocols were approved by the Ethics Committee of the Xinhua Hospital affiliated to Shanghai Jiaotong University School of Medicine.

TMA Preparation and Immunohistochemistry (IHC)

IHC was performed as previously described (23). IHC staining was performed using a standard immunoperoxidase staining procedure, and the primary antibody included Prp19 (1:200, Abcam, ab126776) and YAP (ab52771, 1:50, abcam). Hematoxylin was used as a counterstain. The tissue sections were viewed independently by two pathologists in a double-blind fashion. IHC staining was graded on a specialized scale from 0 to 4, where 0 represents negative expression, 1 represents weakly positive expression (0–10% positive cells), 2 represents mildly positive expression (10–30% positive cells), 3 represents moderately positive expression (30–50% positive cells), and 4 represents strongly positive expression (50–100% positive cells). The scale was determined according to the average number of positive cells in 10 random fields on one slide. IHC staining grade 0–2 was defined as low expression, and IHC staining grade 3–4 was defined as high expression.

Cell Lines and Cell Culture

Human neuroblastoma cell lines SK-N-BE (2) and SK-N-AS were obtained from ATCC (Manassas, USA). All cell lines were cultured in a 1:1 mixture of Eagle's minimum essential medium and F12 medium (Gibco, USA) supplemented with 10% fetal bovine serum (Gemini, USA) in humid air at 37°C with 5% CO_2.

Knockdown and Overexpression of Prp19

Cells were plated in 6-, 12-, or 24-well-plates the day before transfection so that they achieved 30–50% confluence at the time of transfection. Cells were transfected with small interfering RNAs (siPrp19, siYAP, and negative control siNC) using RNAiMAX (ThermoFisher, USA) according to the manufacturer's instructions. siRNAs were synthetized by RiboBio (China); target sequences were as follows: Prp19-1 5′-GCCACTATCAGGATTTGGT-3′, Prp19-2 5′- GCCAAGTTCATCGCTTCAA-3′, and YAP 5′-GTAGCCAGTTACCAACACT-3′. Cells were incubated for 24 to 48 h at 37°C before harvesting cells for analyses.

A Prp19 overexpressing adenovirus with Flag and His tag was purchased from Vigenebio (China). Prp19 overexpression cell lines were constructed in SK-N-AS and SK-N-BE (2) cells.

Cell Invasion Assay

Matrigel was added to the top chamber of Transwell chambers before the chambers were placed in an incubator for 4 h. Then, siRNA-treated cells or Prp19 overexpressing cells (1×10^5 cells), having been washed with serum-free medium, were plated into chambers. After 36 h in 37°C, the cells were harvested and fixed with 4% paraformaldehyde for 20 min. Cells were then stained with 0.1% crystal violet staining for 10 min, and the Matrigel was wiped off with a cotton swab. Samples were observed under a microscope; 5–10 fields were randomly selected and photographed in each chamber.

Wound-Healing Assay

SK-N-BE (2) and SK-N-AS cells were seeded into 6-well-plates; after 24 h, the cells were transfected or infected with siNC, siPrp19, siYAP, or Prp19 or control adenovirus. Forty-eight hours later, the cells reached 100% confluence, and the cell monolayer was scraped with a 200-μl pipette tip. The well was washed twice with serum-free medium (50% MEM and 50% F12) and then replenished with fresh serum-free medium. At 0 h and every 12 h after incubation, images were captured with a light microscope (LeicaCTR6000 microscope system) at x50 magnification and analyzed quantitatively by ImageJ (X64, v. 2.1.4).

Protein Extraction and Western Blotting (WB)

WB was performed as previously described (23). Primary antibodies specific for Prp19 (ab126776, 1:1,000) and YAP (ab52771, 1:1,000) were purchased from Abcam; Primary antibodies specific for GAPDH (2118s, 1:2,000), cyclinD1 (2978s, 1:1,000), MMP9 (13667s, 1:1,000), E-cadherin (3195s, 1:1,000) were purchased from Cell Signaling Technology (Beverly, USA). A primary antibody specific for Actin was purchased from Yeasen (30101ES10, 1:5,000). Primary antibodies specific for CTGF (sc-101586, 1:200), FGF1 (sc-55520, 1:100) were purchased from Santa Cruz Biotechnology (USA). The results were quantitatively analyzed through Image J software (X64, v. 2.1.4).

Polymerase Chain Reaction (PCR) and Quantitative Real Time PCR (qPCR)

Total RNA was isolated with TRIzol (Invitrogen, USA) and subjected to DNase I treatment prior to reverse transcription (Promega, USA). Reverse transcription reactions were performed using the reverse transcription kit from TakaRa (Tokyo, Japan). PCR was performed using the MAX PCR Master Mix (TakaRa, Japan), and the PCR products were subjected to agarose gel electrophoresis. qPCR was conducted to measure the levels of mRNAs using SYBR Green reagent from Yeasen (China). Primer sequences were as follows: *Prp19* forward 5′-GTGCCAAGTTCCCAACCAAGTGTT-3′, reverse 5′-AG CACAGTGGCTTTGTCTTGAAGC-3′; *YAP-1*: forward 5′-CCCGACTCCTTCTTCAAGC-3′, reverse 5′-TGTCCCAGGAG AAACAGCTC-3′; *YAP-2*: forward 5′-TTGTGCCAACTTGAT TCAGC-3′, reverse 5′-TACATCCCGAGTGGGCTAAC-3′;*YAP-3*: forward 5′-CCTGCGTAGCCAGTTACCAA-3′, reverse 5′-CCATCTCATCCACACTGTTC-3′; *YAP-4*: forward 5′-TTGTCACCAAGCACAGAACC-3′, reverse 5′-TTCCTGTC CTGCAATGTCTG-3′; *GAPDH*: forward 5′-TCGACAGTCAGC CGCATCTTCTTT-3′, reverse 5′-GCCCAATACGACCAAA TCCGTTGA-3′. The transcript levels were calculated and analyzed by the $2^{-\Delta\Delta CT}$ method.

RNA Sequencing Assay and Data Analysis

siPrp19 and siNC small interference RNA were transfected to SK-N-BE (2) and SK-N-AS cells for 48 h, and then 3 μg of total RNA per sample (6 samples per cell line) were extracted by TOIzol® reagent (Thermo Fisher scientific, American). Library preparation for transcriptome sequencing and clustering and sequencing were performed by Novogene Bioinformatics Technology Co., Ltd. (Beijing, China). Per sample, biological replicates were performed three times. Differential expression analysis of the two groups was performed using the DESeq2R package (1.16.1), and genes with an adjusted $P < 0.05$ found byDESeq2 were assigned as differentially expressed. Gene ontology (GO) enrichment analysis of differentially expressed genes was implemented by the cluster Profiler R package, in which gene length bias was corrected. GO terms with corrected value <0.05 were considered significantly enriched by differentially expressed genes. KEGG is a database resource for understanding high-level functions and utilities of the biological system from molecular-level information, especially large-scale molecular data sets generated by genome sequencing and other high-throughput experimental technologies (http://www.genome.jp/kegg/). We used the cluster Profiler R package to test the statistical enrichment of differential expression genes in KEGG pathways. Sequencing data have been deposited in the Gene Expression Omnibus database, and the accession codes of SK-N-BE (2) and SK-N-AS are GSE153398 and GSE153432.

Neuroblastoma Data Set Analysis

In order to validate the conclusion we got from our clinical data, we analyzed the factors associated with the prognosis of neuroblastoma by the R2 Genomics Analysis and Visualization Platform (http://r2.amc.nl) using the following publicly available data sets: SEQC (GEO: GSE49710) and NRC (GEO: GSE85047). In addition, we also analyzed the differential expression of Prp19 among clinical stage, risk classification, *MYCN* status, and histology.

Statistical Analysis

Statistical analysis was performed using SPSS 20.0 software. The Mann-Whitney U rank sum test was used for the non-normally distributed variables. The different expression of Prp19 among genders, ages, bone marrow infiltration, clinical stage (INSS), tumor pathological diagnosis, Shimada pathological classification, and risk classification were analyzed by chi-square test. Multivariate Cox regression analysis was performed on the effects of Prp19, clinical stage, and age at diagnosis on survival in children. Survival curves between different expression groups of Prp19 were delineated using Kaplan–Meier curves, and differential analysis of survival curves was performed using log-rank test. Values of $p < 0.05$ were considered statistically significant.

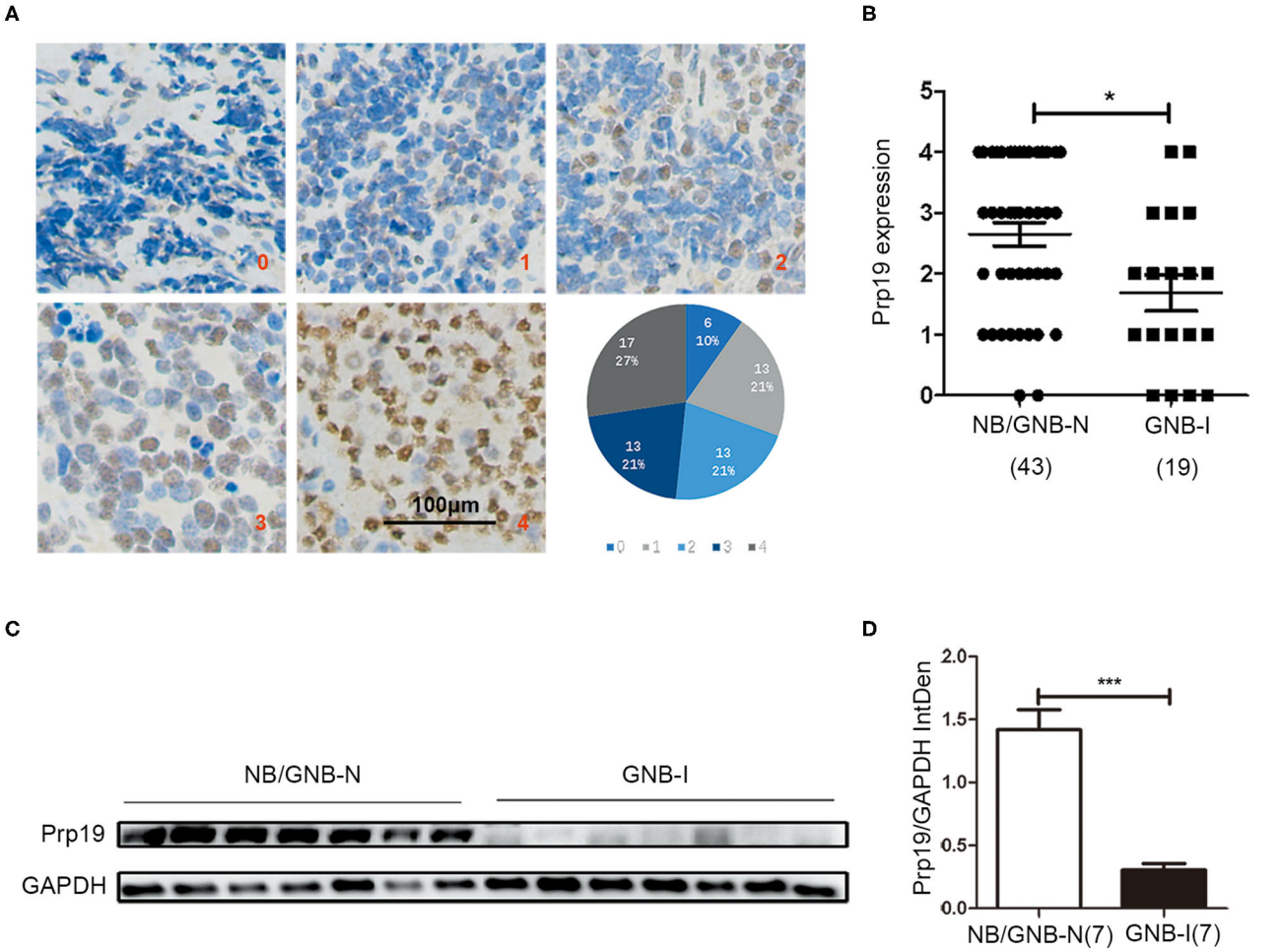


FIGURE 1 | Expression and roles of Prp19 in neuroblastoma. **(A)** Representative images of different levels (0–4) of IHC staining of Prp19 and the proportions of five levels. Scale bar: 100 μm. **(B)** Prp19 expression between NB/GNB-N and GNB-I in TMA. The data were analyzed by Mann-Whitney *U*-test ($n = 62$). **(C,D)** Proteins of 7 NB/GNB-N and 7 GNB-I tumor tissue proteins extracted for immunoblotting to detect Prp19 expression, and Image J was used to quantify protein bands. The data in **(D)** were analyzed by Mann-Whitney *U*-test ($n = 14$). $^{*}p < 0.05$ and $^{***}p < 0.001$.

RESULTS

Differential Expression of Prp19 Is Associated With Clinical Characteristics in Neuroblastoma

We first performed IHC analysis of Prp19 expression on a TMA consisting of 62 samples from patients diagnosed with neuroblastoma, including 43 NB/GNB-N and 19 GNB-I cases. Prp19 was mainly expressed in the cell nucleus and differentially expressed in neuroblastoma patients with different pathological types; the expression of Prp19 was as follows: grade 0 (10%), grade 1 (21%), grade 2 (21%), grade 3 (21%), and grade 4 (27%) (**Figure 1A**). Prp19 expression in patients with NB/GNB-N was significantly higher than that in patients with GNB-I ($p = 0.016$; **Figure 1B**). Western blot analysis of 14 frozen fresh samples further confirmed that the expression of Prp19 was significantly higher in patients with NB/GNB-N compared with patients with GNB-I ($p = 0.006$; **Figures 1C,D**). This indicates that Prp19 shows differential expression in different pathological types of neuroblastoma.

We next examined whether expression of Prp19 had clinical implications in our cohort. As shown in **Table 1**, high expression of Prp19 was significantly associated with bone marrow metastasis, NB/GNB-N, unfavorable histologic and high risk ($p = 0.009$, 0.021, 0.022, and 0.023, respectively); however, there were no associations with age at diagnosis and clinical stage in our TMA, possibly due to insufficient sample size. To further confirm the clinical implications of Prp19 in human neuroblastoma, we analyzed gene expression profiles from two cohorts of neuroblastoma primary tumors with larger sample sizes (SEQC: GSE62564 and NRC: GSE85047). The expression of Prp19 was positively correlated with clinical stage ($p < 0.001$; **Table 2**), and more interestingly, Prp19 expression increased with clinical stage from stage I to stage IV, but significantly decreased in stage

TABLE 1 | The association between Prp19 expression with clinical pathologic characteristics in the TMA cohort.

Clinical and pathological characteristics	Total	Prp19		χ^2	*P*
		High	Low		
Gender					
Male	35	14	21	2.264	0.132
Female	27	16	11		
Age at diagnosis					
≥18 months	42	21	21	0.136	0.713
<18 months	20	9	11		
Bone marrow metastasis					
Positive	21	15	6	6.751	0.009
Negative	41	15	26		
Clinical stages					
I–II IV–S	25	10	15	1.180	0.277
III–IV	37	20	17		
Diagnostic category					
NB and GNB-N	43	25	18	5.344	0.021
GNB-I	19	5	14		
Shimada pathologic type					
FH	30	10	20	5.274	0.022
UFH	32	20	12		
Risk classification					
Low and Intermediate	34	12	22	5.168	0.023
High	28	18	10		

NB, neuroblastoma; GNB-N, ganglioneuroblastoma-nodular; GNB-I, ganglioneuroblastoma-intermixed; FH, favorable histology; UFH, unfavorable histology.

TABLE 2 | Correlation analysis between clinical characteristics and expression of Prp19 in the SEQC data set.

Covariates	Total	Prp19		χ^2	*P*
		High	Low		
Age at diagnosis					
≥18 months	198	94	104		
<18 months	300	45	255	62.517	<0.001
Clinical stages					
I–II IV–S	252	18	234		
III–IV	246	121	125	109.362	<0.001
Risk					
Low risk	322	33	289		
High risk	176	106	70	498.000	<0.001
***MYCN* state**					
Amplification	92	50	42		
Non-amplification	401	87	314	39.758	<0.001

IVs (**Supplementary Figures 1A,C**). Furthermore, Prp19 was significantly overexpressed in children diagnosed at more than 18 months ($p < 0.001$) and high-risk cases ($p < 0.001$; **Table 2** and **Supplementary Table 1**; **Supplementary Figures 1B,D,E**). In addition, patients with *MYCN* amplification also tended to have higher Prp19 expression ($p < 0.001$; **Table 2** and **Supplementary Table 1**; **Supplementary Figures 1F,G**). Together, these results indicate that Prp19 shows differential expression in neuroblastoma tissues and higher Prp19 expression is associated with poor clinical characteristics in neuroblastoma.

High Expression of Prp19 Is a Potential Poor Prognostic Factor in Children With Neuroblastoma

We next investigated if Prp19 had an influence on the prognosis of children with neuroblastoma. Out of the 62 patients from the TMA cohort, 50 cases were successfully followed up with a follow-up completion rate of 80.6%. The median follow-up time was 34 months, and the longest and shortest follow-up periods were 62 and 6 months, respectively. We examined OS and EFS by Kaplan-Meier curves in the TMA cohort and compared the survival curves by log-rank test. Patients with low Prp19 expression had a better OS ($p = 0.0040$) and EFS ($p = 0.0275$; **Figures 2A,D**). In addition, analyses of the SEQC and NRC data sets revealed similar results, and both the *p*-value of all OS and EFS rates were <0.0001 (**Figures 2B,C,E,F**). In the TMA cohort, univariate analysis found that high expression of Prp19 was a risk factor both for OS ($p = 0.010$; HR = 5.189) and EFS ($p = 0.045$; HR = 2.644) in children with neuroblastoma (**Table 3**). Multivariate analysis also found that patients with high Prp19 expression had a worse OS ($p = 0.027$; HR = 4.118) although there was no statistically significant difference in the EFS ($p = 0.174$; HR = 1.954). Clinical stage, as a recognized prognostic risk factor, was also associated with poor OS ($p = 0.016$; HR = 12.345) and EFS ($p = 0.003$; HR = 9.267). However, age at diagnosis did not show an effect on OS and EFS in children with neuroblastoma (**Table 4**). Age at diagnosis is a known risk factor affecting the prognosis of children with neuroblastoma (3). On account of this, we doubted that the negative result from **Table 4** was limited by the number of the TMA cohort. To further validate our results, we performed the analyses in the SEQC and NRC databases. Univariate and multivariate analyses show that high Prp19 expression, high clinical stage, and age over 18 months at diagnosis are associated with poor OS and EFS in the SEQC data set (**Supplementary Tables 2, 3**). In the NRC data set, OS and EFS are associated with Prp19 expression and clinical stage, but age at diagnosis, even with an impact on OS, has no effect on EFS (**Supplementary Tables 4, 5**). Taken together, our analyses indicate that Prp19 is a potential prognostic marker in neuroblastoma.

Prp19 Promotes Neuroblastoma Cell Invasion and Migration

Then, the function of Prp19 at the cellular level was explored. First, the expression level of Prp19 in different neuroblastoma cell

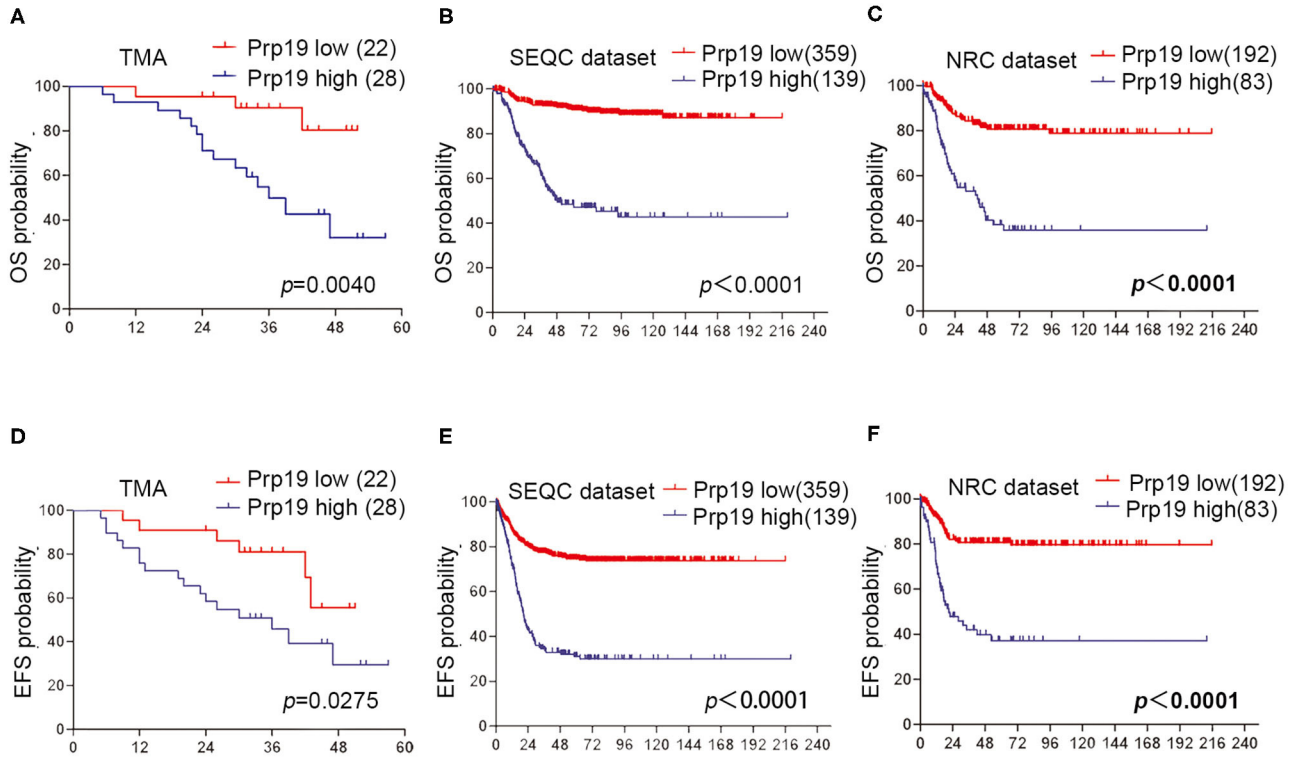


FIGURE 2 | Prp19's prognostic value in TMA cohort and data sets. **(A,D)** Kaplan-Meier analysis of OS and EFS for the TMA cohort based on Prp19 expression with the log-rank test *P*-value indicated ($n = 50$). **(B,E)** Kaplan-Meier analysis of OS and EFS for the SEQC data set based on Prp19 expression with the log-rank test *p* value indicated ($n = 498$). **(C,F)** Kaplan-Meier analysis of OS and EFS for the NRC data set based on Prp19 expression with the log-rank test *p*-value indicated ($n = 275$).

TABLE 3 | Univariate analyses in the TMA cohort.

Covariates	OS		EFS	
	HR (95%CI)	*P*	HR (95%CI)	*P*
Prp19 expression (high vs. low)	5.189 (1.487–18.109)	0.010	2.644 (1.023–6.836)	0.045

TABLE 4 | Multivariable analyses in the TMA cohort.

Covariates	OS		EFS	
	HR (95%CI)	*P*	HR (95%CI)	*P*
Prp19 expression (high vs. low)	4.118 (1.171–14.480)	0.027	1.954 (0.744–5.136)	0.174
Clinical stages (I–II IV–S vs. III–IV)	12.345 (1.613–94.489)	0.016	9.267 (2.114–40.624)	0.003
Age at diagnosis (< >18 months)	0.980 (0.351–2.740)	0.969	1.755 (0.667–4.616)	0.255

lines [SK-N-BE (2), SK-N-AS, SH-SY5Y, IMR32, and LAN1] was assessed, and the silencing efficiency of siRNA targeting Prp19 was examined (**Figures 3A,B**). Two cell lines [SK-N-BE (2) and SK-N-AS] with higher expression of Prp19 and the first siRNA were selected for the following assays. We next explored whether Prp19 influences migration or invasion of neuroblastoma cells. After 48 h silencing of Prp19 in SK-N-BE (2) and SK-N-AS cells, transwell assays were performed. As shown in **Figures 3C,D**, knockdown of Prp19 impaired the invasive ability of both SK-N-BE (2) and SK-N-AS cells, and upregulating Prp19 enhanced the invasive capacity of two cell lines. In addition, the wound-healing assay also showed increased or decreased migratory ability in Prp19 overexpressing cells or Prp19 silenced cells (**Figures 3E,F**). EMT-related molecules, such as E-cadherin and matrix metalloproteinase 9 (MMP9), were altered upon Prp19 expression; we observed increased E-cadherin and decreased MMP9 after Prp19 downregulation with opposite effects in Prp19 overexpression cells (**Figures 3G,H**). These results demonstrate the role of Prp19 in neuroblastoma cell invasion, migration, and EMT.

RNA-seq Identifies the Hippo/YAP Pathway as a Candidate Regulatory Target of Prp19

Given the change of invasion and migration observed in Prp19 knockdown and overexpressed cells, we speculated that Prp19 may affect specific cellular functions through regulating specific

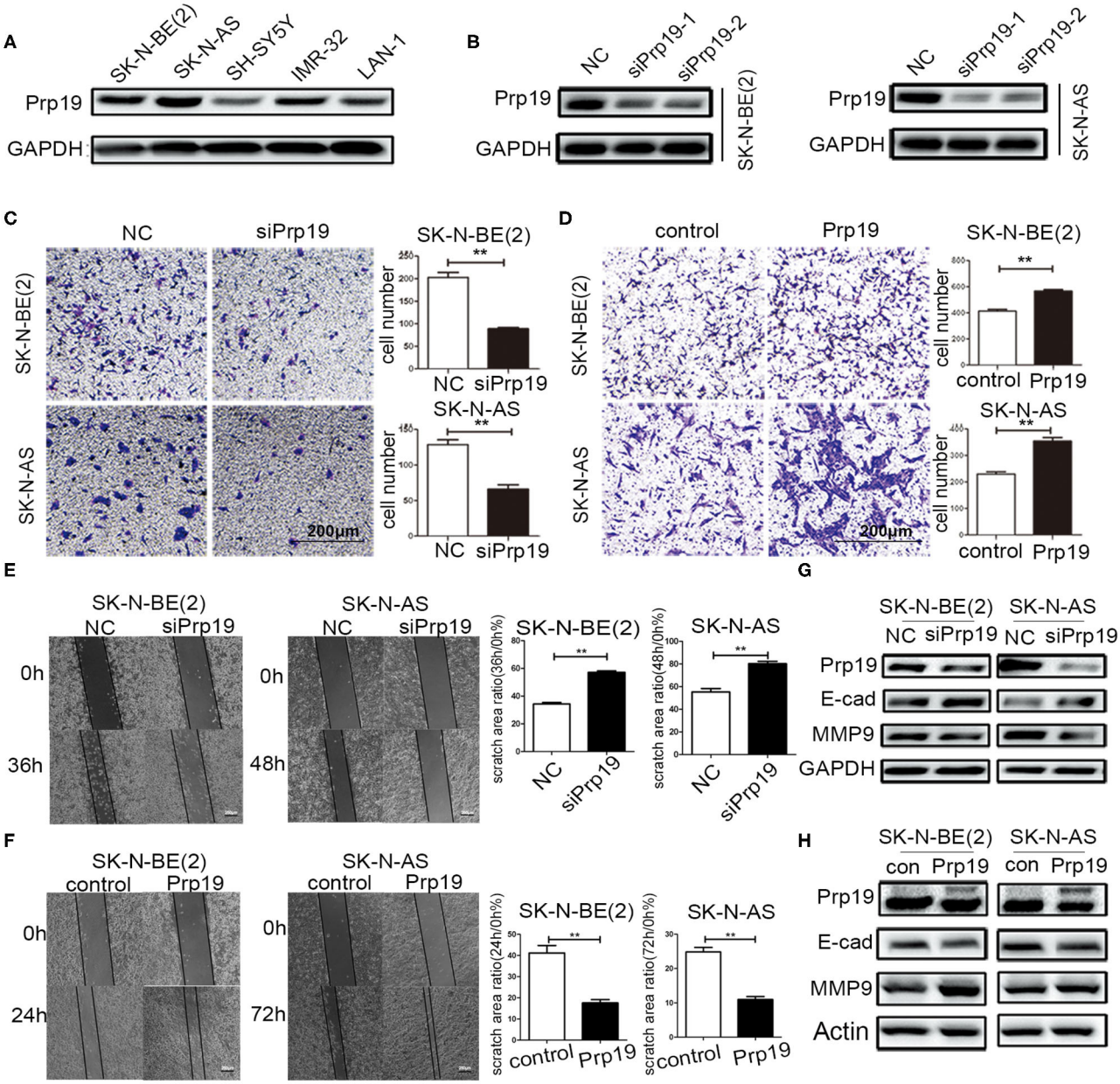


FIGURE 3 | Effects of Prp19 on cell biological behavior of neuroblastoma cells. **(A)** WB detected the expression levels of Prp19 in neuroblastoma cell lines SK-N-BE (2), SK-N-AS, SH-SY5Y, IMR-32, and LAN1, and two cell lines [SK-N-BE (2) and SK-N-AS] with higher expression of Prp19 were selected for subsequent experiments. **(B)** Prp19 was knocked down in SK-N-BE (2) and SK-N-AS using siRNA. The first siRNA was selected for subsequent experiments according to knockdown efficiency detected by WB. **(C,D)** Transwell array tested the invasive ability of SK-N-BE (2) and SK-N-AS between siPrp19 and siNC or between overexpression of Prp19 and control, and the number of cells crossing the bottom of the chamber were analyzed by GraphPad Prism 5. Scale bar: 200 μm $p = 0.0022$, respectively. **(E,F)** The cell migration ability was estimated using a wound-healing assay. The images were captured at indicated time after wounding (magnification: 50×; scale bar: 250 μm) $p = 0.0022$, respectively. **(G,H)** WB tested the expression of Prp19, E-cad, and MMP9 in Prp19 downregulation or overexpression cells. $^{*}p < 0.05$ and $^{**}p < 0.01$.

gene expression. To explore the mechanism by which Prp19 exerts its function in neuroblastoma, we performed RNA-seq analysis using SK-N-BE (2) and SK-N-AS cells transfected with siPrp19 compared with cells transfected with siNC. More than 400 genes were significantly regulated upon Prp19 downregulation (|fold change| >1.5, FDR <0.05; **Figure 4A** and **Supplementary Figure 2A**). Bioinformatic analysis shows that most of the enriched disease ontology (DO) catalogs and signaling pathways are suppressed upon Prp19 depletion. DO semantic and enrichment analysis reveal that the most highly enriched functional categories are related to cancer, including central nervous system cancer, autonomic nervous

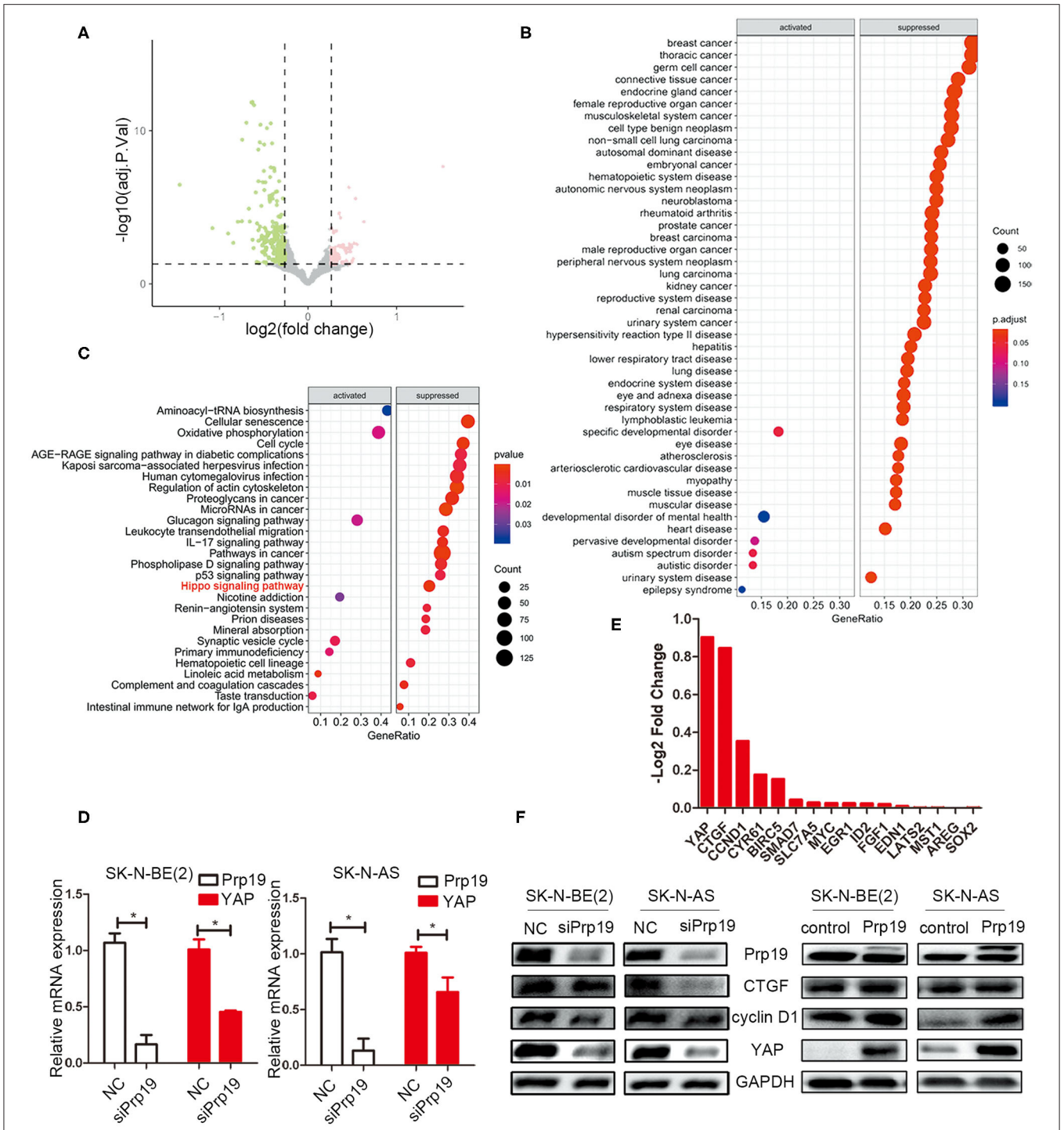


FIGURE 4 | Bioinformatics analysis of RNA-seq data. SK-N-BE (2) cells were transfected with siPrp19 for 48 h, and then subjected to RNA sequencing. **(A)** Volcano plot showing distribution of differential expression genes Prp19 downregulation (|Fold of changes| >1.5, FDR <0.05). **(B)** Disease ontology (DO) analysis enriched the affected diseases after Prp19 knockdown, which were mainly concentrated in nervous system cancer. **(C)** KEGG enriched the affected pathways after Prp19 knockdown, indicating that the Hippo signaling pathway was suppressed. **(D)** PCR analyzed the YAP mRNA level after Prp19 knockdown in SK-N-BE (2) and SK-N-AS. **(E)** RNA-seq analyzed the mRNA level of YAP and its downstream genes after Prp19 knockdown. It indicated that YAP and its downstream genes decreased following Prp19 downregulation. **(F)** The protein level of YAP, cyclin D1, and CTGF after downregulation or overexpression of Prp19. *$p < 0.05$.

system neoplasm, and even neuroblastoma (adj. $p < 0.05$; **Figure 4B** and **Supplementary Figure 2B**). Results from KEGG analysis identified the Hippo-YAP pathway (**Figure 4C** and **Supplementary Figure 2C**), which has been shown to regulate multiple biological processes in various cancers but has not been indicated as downstream signaling pathway of Prp19. YAP, the key component of the Hippo-YAP signaling pathway, was significantly downregulated at both the mRNA and protein levels in cells silenced for Prp19 (**Figures 4D,F**) and was decreased in both the cytoplasm and nucleus

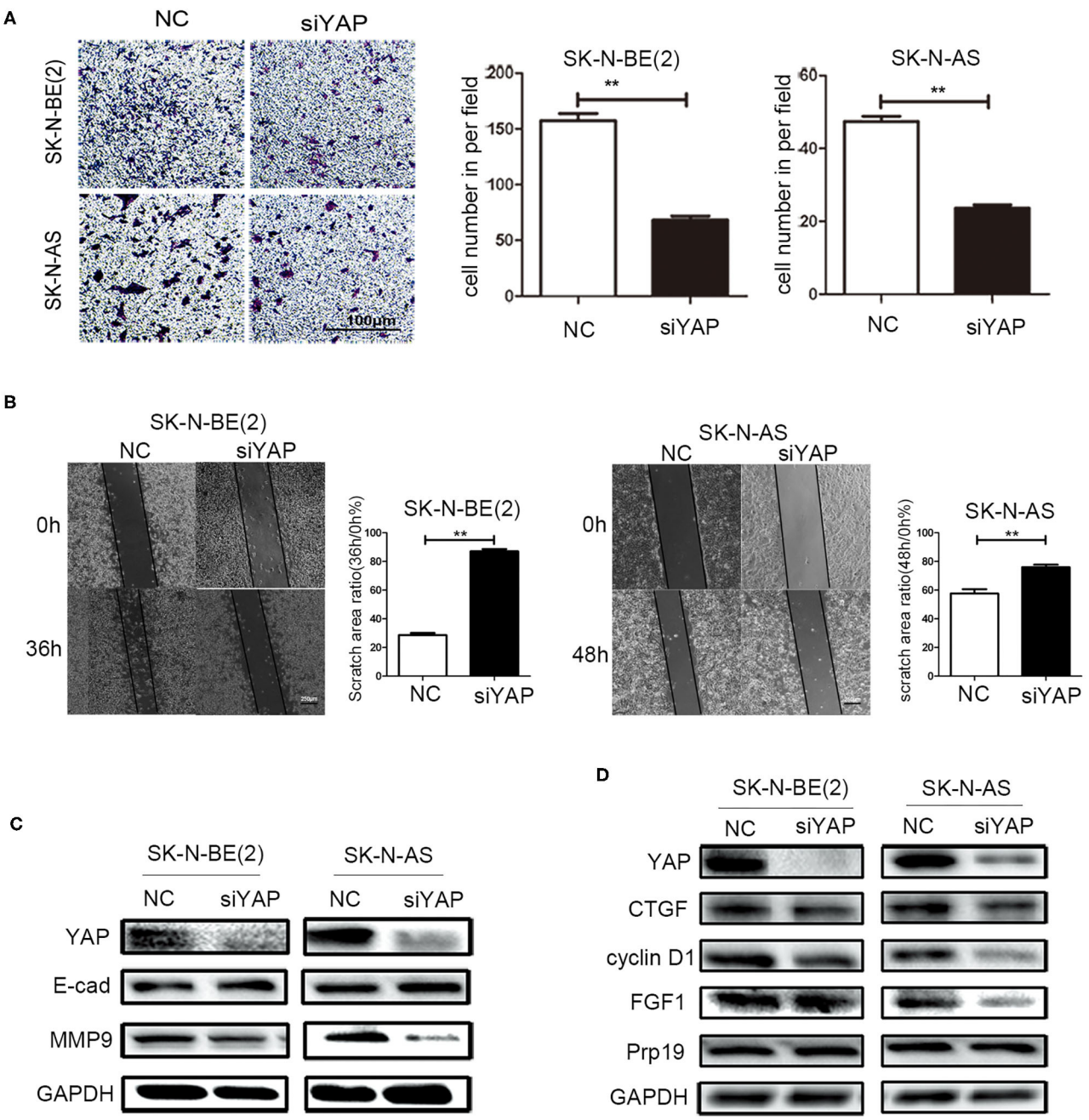


FIGURE 5 | Effects of YAP on cell biological behavior of neuroblastoma cells. **(A)** Transwell array tested the invasive ability of SK-N-BE (2) and SK-N-AS between siYAP and siNC, and the number of cells crossing the bottom of the chamber were analyzed by GraphPad Prism 5. Scale bar: 100 μm $p = 0.0079$, respectively. **(B)** The cell migration ability was estimated using the wound-healing assay after YAP downregulation. The images were captured at the time shown on the graph after wounding (magnification: 50×; scale bar: 250 μm) $p = 0.0022$, respectively. **(C)** WB tested the expression of YAP, E-cad, and MMP9 after downregulation of YAP. **(D)** WB tested the expression of YAP, CTGF, cyclin D1, FGF1, and Prp19 after downregulation of YAP. *$p < 0.05$ and **$p < 0.01$.

(**Supplementary Figure 3**). Furthermore, the downstream genes of the Hippo-YAP pathway, including *CTGF* and *CCND1*, were also significantly decreased with Prp19 knockdown (**Figure 4E** and **Supplementary Figure 2D**). In addition, the protein levels of YAP, CTGF, and cyclin D1 were also decreased or increased upon Prp19 downregulation or overexpression, respectively (**Figure 4F**). Together these data, thus, suggest that Prp19 may be involved in regulating the Hippo-YAP pathway by regulating YAP level.

Effect of Deletion of YAP on Cell Invasion and Migration

To examine whether Prp19 affects invasion and migration of neuroblastoma cells via the Hippo-YAP pathway, we first examined whether silencing YAP resulted in similar changes of invasion and migration of neuroblastoma cells as with Prp19 silencing. As a result, as we expected, invasion and migration abilities were impaired in YAP downregulated cells (**Figures 5A,B**). YAP knockdown was also accompanied with upregulation of E-cadherin and downregulation of MMP9 (**Figure 5C**). Expressions of downstream target genes of YAP were also downregulated after YAP knockdown, but Prp19 expression was not affected (**Figure 5D**). These results demonstrate that YAP knockdown causes almost same changes as Prp19 knockdown, implying that Prp19 may affect cell invasion and migration by regulating YAP expression.

Prp19 and YAP Are Involved in Neuroblastoma Metastasis

Because the enhancement of invasion and migration ability is closely related to metastasis, we speculated that Prp19 may play a role in tumor metastasis. Four pairs of neuroblastoma *in situ* tumor and metastatic lymph node tissues for the IHC experiment were collected. Among the 4 pairs of samples, Prp19 and YAP expression in two lymph node samples were higher than their corresponding tissues (**Figure 6**). The other 2 pairs of samples had received chemotherapy before we obtained them, so they did not show consistent results (data not shown). This result demonstrated Prp19 and YAP participate in tumor metastasis.

Prp19 Is Required for Efficient RNA Processing of YAP mRNA

Previous studies establish a critical role for Prp19 in regulating pre-RNA splicing. Our results show that Prp19 affects YAP gene levels, and we, therefore, examined the possibility that downregulation of YAP expression after Prp19 knockdown might be due to inefficient pre-mRNA splicing. To explore this possibility, we designed a series of PCR primer pairs targeting two neighboring constitutive exons or the interior of a single intron of the YAP gene as shown in **Figure 7A**. PCR agarose gel electrophoresis analyses found YAP intron retention (YAP-2 and YAP-4) and mature mRNA reduction (YAP-1 and YAP-3) in SK-N-BE (2) and SK-N-AS cells upon Prp19 knockdown, and opposite changes were found in Prp19 overexpression cells in addition to YAP-4 (**Figure 7B**). We verified the results with qPCR and obtained similar results. The relative RNA levels of YAP-1 and YAP-3, representing mature YAP RNA levels, decreased, and YAP-2 and YAP-4, representing immature YAP RNA levels, increased in Prp19 knockdown cells. In Prp19 overexpression cells, the relative RNA levels of YAP-1 and YAP-3 increased, and relative RNA levels of YAP-2 decreased while the relative RNA level of YAP-4 was unaffected (**Figure 7C**). These results demonstrated that Prp19 mediated YAP RNA splicing to regulate YAP abundance.

DISCUSSION

The International Neuroblastoma Pathology Classification (INPC) divides neuroblastoma into four subtypes: neuroblastoma, ganglioneuroblastoma-nodular (GNB-N), ganglioneuroblastoma-intermixed (GNB-I), and ganglioneuroma (GN) (24). Both GNB-I and GN have favorable clinical and biological characteristics, and the long-term survival rate exceeds 90% while neuroblastoma and GNB-N have relatively malignant clinical and biological characteristics with a long-term survival rate <80% (25). A recent study even indicated that GNB-I and GN patients only need regular follow-up rather than surgery first (26). Furthermore, in the current risk classification system, most GNB-I and GN are classified as low risk, whereas most neuroblastoma and GNB-N are classified as intermediate or high risk, which always accompany long-distance metastasis (3). Therefore, there may exist a different biological molecular distribution between neuroblastoma/GNB-N and GNB-I/GN that drive neuroblastoma/GNB-N possessing malignant biological behavior, for example, metastasis. In recent years, several studies have pointed out that molecular markers, for example, gene expression classification, can accurately affect the biological behavior, predicting outcomes for children with neuroblastoma (27–30). Revealing the molecular mechanisms of high-risk neuroblastoma may provide clues for target therapy, thereby improving the prognosis of children with high-risk neuroblastoma.

Our study reveals differential expression of Prp19 in different pathological types with higher Prp19 expression in neuroblastoma and GNB-N compared with GNB-I. This is consistent with several studies that show higher Prp19 expression in colon, laryngeal, and hepatocellular carcinoma tissues compared with normal tissues (13). In the SEQC and NRC databases, Prp19 expression was positively correlated with bone marrow metastasis, histological type, Shimada pathologic type and risk grade, age at diagnosis and clinical stage, and overexpression of Prp19 had a worse OS and EFS, which was also consistent with a previous study (13, 14). These data suggest that Prp19 may be one of the factors that results in differences in tumor biological behavior and prognosis in neuroblastoma patients.

The malignant biological behavior of tumors usually includes uncontrolled proliferation, metastasis, and multidrug resistance. In our clinical data, we noticed that Prp19 expression was

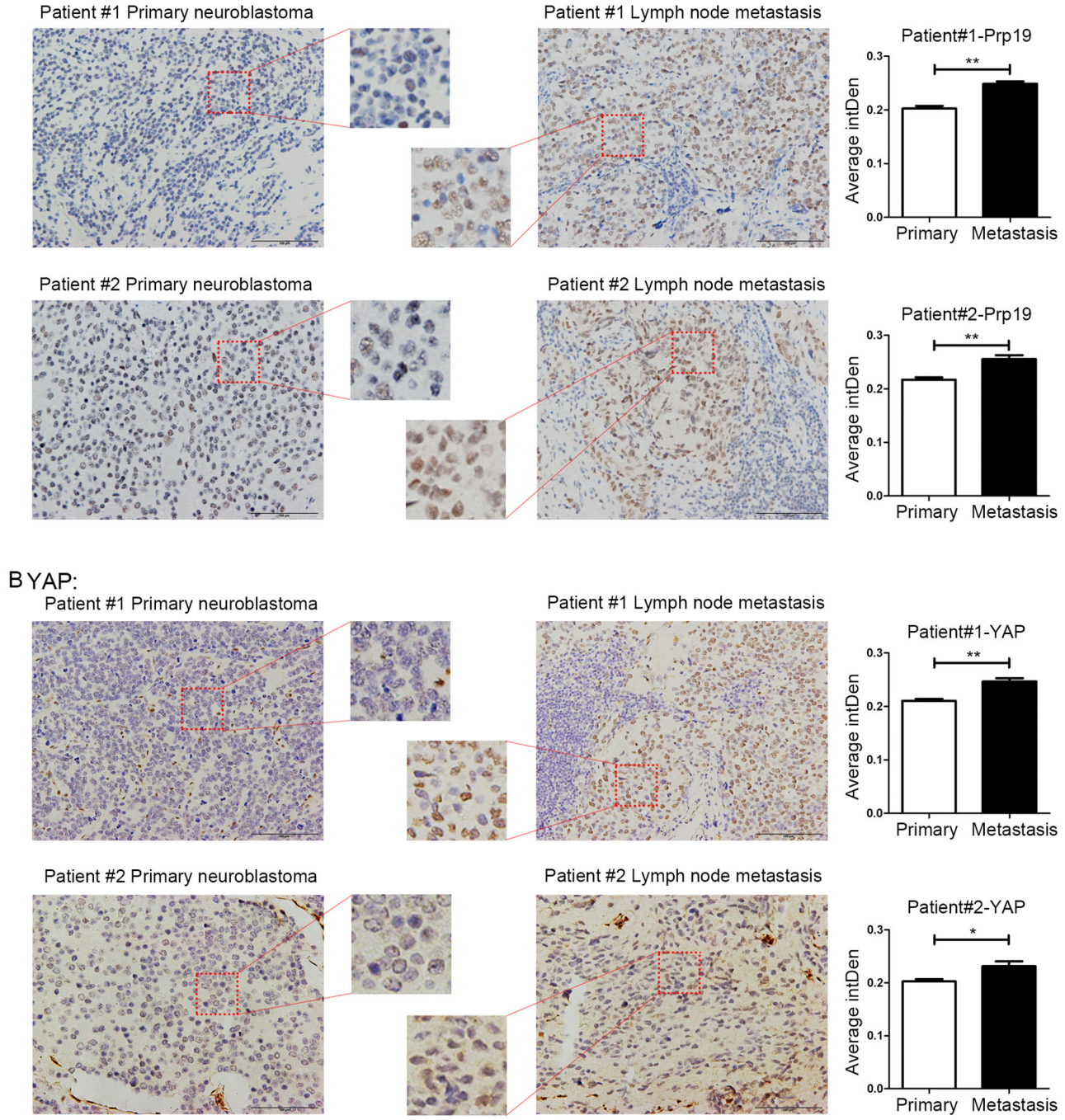


FIGURE 6 | Expression of Prp19 and YAP are higher in metastatic lymph nodes than primary sites. **(A)** The expression of Prp19 in metastatic lymph nodes of patients 1 and 2 is higher than that in primary tumor. **(B)** The expression of Prp19 in metastatic lymph nodes of patients 1 and 2 is higher than that in primary tumor. The average intDen in each group was analyzed by Mann-Whitney *U*-test ($n = 6$ per group). $^{*}p < 0.05$ and $^{**}p < 0.01$.

positively correlated with bone marrow metastasis. Therefore, metastasis may be one of the factors by which Prp19 causes poor prognosis of neuroblastoma. EMT has been associated with various tumor functions, including tumor initiation, tumor stemness, tumor cell migration, intravasation to the blood, metastasis, and resistance to therapy (31, 32). The characteristic of EMT is the cell conversion from the epithelial-like profile (marked with E-cadherin and β-catenin) to the mesenchymal

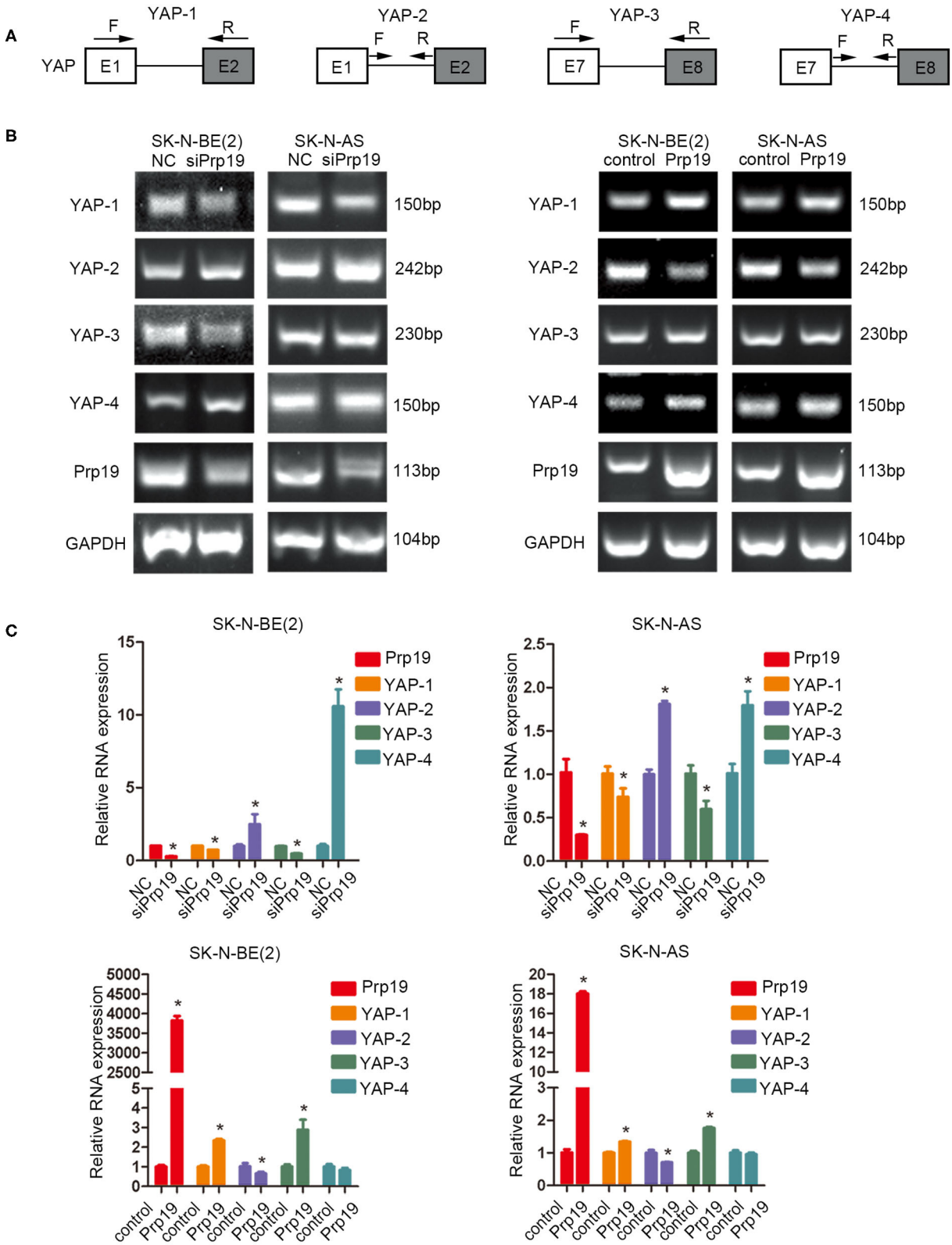


FIGURE 7 | Prp19 is required for efficient intron removal of YAP. **(A)** A schematic diagram of exons and introns of the YAP gene and the primer sets designed for PCR as well as the RT-PCR shown in **(B,C)**. **(B)** PCR analysis using primer sets shown in **(A)** to compare splicing efficiencies of YAP introns splicing in Prp19 siRNA-transfected and Prp19 overexpression cells. **(C)** qPCR analysis using primer sets shown in **(A)** to compare splicing efficiencies of YAP introns splicing in Prp19 siRNA-transfected and Prp19 overexpression cells. $^{*}p < 0.05$.

phenotype (marked with N-cadherin and MMP9) concomitant with the detachment of cells from intercellular adhesion and enhanced motility (32). A few studies have examined the involvement of Prp19 in EMT and metastasis of cancer. There is only one study showing that Prp19 promotes invasion, migration, and EMT in hepatocellular carcinoma cells (13). Here, we also found that downregulation of Prp19 inhibited invasion and migration of neuroblastoma cells and reversed EMT with an upregulation of E-cadherin and downregulation of MMP9, and overexpression of Prp19 had the opposite results. Abnormal activation of EMT not only results in the increased invasion and migration of cancer cells, but also triggers distant metastasis in diverse cancers (33, 34). Collectively, our findings identify the role of Prp19 in promoting cell invasion, migration, and EMT in neuroblastoma, and this means that Prp19 has the potential to promote tumor metastasis.

Given the pleiotropic changes observed in Prp19 knockdown cells, we hypothesized that Prp19 might take part in the regulation of certain genes, thereby regulating the expression of multiple genes. In order to seek a downstream target or pathway of Prp19, RNA-seq analysis was performed on neuroblastoma cells transiently transfected with Prp19 siRNA. Among the downregulated genes, it is noteworthy that the YAP gene changed by more than 0.9-log2fold when Prp19 was lowered by 1.0-log2fold. To the best of our knowledge, this is the first report suggesting that Prp19 may regulate the Hippo-YAP pathway in neuroblastoma cells. The Hippo-YAP pathway is a highly conserved signaling pathway regulating the regulation of organ size, cell proliferation, invasion, and apoptosis (17, 35). Many studies have shown that YAP, one of the key downstream terminal effectors of the Hippo pathway (36), promotes proliferation and metastasis of cancer cells (18, 37–39), and can also be regulated by other protein (21). Further analysis of changes in disease types and cellular pathways associated with downregulation of Prp19 found that the most highly enriched functional categories were related to cancer and that the Hippo-YAP pathway was inhibited. Results also imply that many protein and mRNA expressions were downregulated downstream of YAP. In addition, silencing YAP led to decreased invasive and migratory ability and reversal of EMT in neuroblastoma cells, similar to the effects of siPrp19. These results indicate that Prp19 might mediate tumor cell invasion and migration by regulating YAP.

The above results suggest that both Prp19 and YAP have the potential to promote tumor metastasis because they can promote tumor migration, invasion, and EMT transformation. Recent literature also has shown that YAP can promote tumor metastasis to lymph nodes and is highly expressed in metastatic lymph nodes (18). On this basis, our results reveal that the expression of Prp19 and YAP in neuroblastoma metastatic lymph nodes is significantly higher than that in *in situ* tumors. From the above results, we can see that Prp19 can affect the expression of YAP, but not the opposite. Therefore, it can be speculated that Prp19 promotes YAP expression and further promotes lymph node metastasis of the tumor though further experiments need to be done to verify.

The next point is how Prp19 regulates YAP expression. As far as we know, the best-described function of Prp19 is its role in splicing (40–42), a critical step in gene expression that involves the removal of introns with specificity and precision. As a pivotal step in eukaryotic gene regulation, splicing enables excision of introns from pre-mRNA and the generation of protein-coding transcripts, which takes place on the spliceosome, a dynamic macromolecular complex. Prp19 associates with the spliceosome concomitantly with or immediately after dissociation of U4, finally leading to spliceosomal activation and pre-mRNA splicing (8). We designed a series of PCR primer pairs targeting two neighboring constitutive exons, reflecting the mature mRNA, and the interior of a single intron of the YAP gene, reflecting the immature pre-mRNA. Results show that the mature mRNA of YAP is decreased with the decrease of Prp19 and increased with the increase of Prp19 except for YAP-4 in the Prp19 overexpression group. As for the YAP-4 group, it may be hypothesized that once Prp19 is overexpressed, the expression of YAP increased dramatically, so there was also a higher immature pre-mRNA of YAP. Taken together, these findings demonstrate that Prp19 controls YAP abundance by coordinated regulation of RNA splicing.

In conclusion, our study showed that Prp19 is positively correlated with various adverse clinical pathological parameters and could serve as a potential prognostic marker indicating a poor clinical prognosis in neuroblastoma. Moreover, Prp19 promotes metastasis of neuroblastoma cells through controlling the level of YAP by RNA splicing.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. This data can be found in Bioproject GSE153398 and GSE153432.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics Committee of Xin Hua Hospital Affiliated to Shanghai Jiao Tong University School of Medicine. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

YC carried out the studies, participated in the study design, statistical analysis, and drafted the manuscript. KC performed cell culture and participated in all *in vitro* experiments. KC, CC, and QC performed proteomics analysis and analyzed the RNA-Seq data. YX and GX participated in the histological examination of tissue samples and shRNA interference assay. ZW and YW conceived the study, participated in its design and coordination, and helped to

draft the manuscript. All authors read and approved the final manuscript.

FUNDING

This work was supported by the Suzhou Clinical Medicine Innovation Team Introduction Project (SZYJTD201706) to YW, and Natural Science Foundation of China (No. 81874234), Shanghai Rising-Star Program (16QA1402900) to ZW and Natural Science Foundation of China (No. 81672488) to GX.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2020.575366/full#supplementary-material

REFERENCES

1. Malcolm AS, Sean FA, Peter CA, Gregory HR, Nita LS. Declining childhood and adolescent cancer mortality. *Cancer.* (2014) 120:2497–506. doi: 10.1002/cncr.28748
2. Matthay KK, Maris JM, Schleiermacher G, Nakagawara A, Mackall CL, Diller L, et al. Neuroblastoma. *Nat Rev Dis Primers.* (2016) 2:16078. doi: 10.1038/nrdp.2016.78
3. Cohn SL, Pearson AD, London WB, Monclair T, Ambros PF, Brodeur GM, et al. The international neuroblastoma risk group (INRG) classification system: an INRG task force report. *J Clin Oncol.* (2009) 27:289–97. doi: 10.1200/JCO.2008.16.6785
4. Tas ML, Reedijk AMJ, Karim-Kos HE, Kremer LCM, van de Ven CP, Dierselhuis MP, et al. Neuroblastoma between 1990 and 2014 in the Netherlands: increased incidence and improved survival of high-risk neuroblastoma. *Eur J Cancer.* (2019) 124:47–55. doi: 10.1016/j.ejca.2019.09.025
5. Hara J. Development of treatment strategies for advanced neuroblastoma. *Int J Clin Oncol.* (2012) 17:196–203. doi: 10.1007/s10147-012-0417-5
6. Chanarat S, Strasser K. Splicing and beyond: the many faces of the Prp19 complex. *Biochim Biophys Acta.* (2013) 1833:2126–34. doi: 10.1016/j.bbamcr.2013.05.023
7. Grote M, Wolf E, Will CL, Lemm I, Agafonov DE, Schomburg A, et al. Molecular architecture of the human Prp19/CDC5L complex. *Mol Cell Biol.* (2010) 30:2105–19. doi: 10.1128/MCB.01505-09
8. Yin J, Zhu JM, Shen XZ. New insights into pre-mRNA processing factor 19: a multi-faceted protein in humans. *Biol Cell.* (2012) 104:695–705. doi: 10.1111/boc.201200011
9. Huang R, Xue R, Qu D, Yin J, Shen XZ. Prp19 arrests cell cycle via Cdc5L in hepatocellular carcinoma cells. *Int J Mol Sci.* (2017) 18:778. doi: 10.3390/ijms18040778
10. Mahajan K. hPso4/hPrp19: a critical component of DNA repair and DNA damage checkpoint complexes. *Oncogene.* (2016) 35:2279–86. doi: 10.1038/onc.2015.321
11. Garschall K, Dellago H, Galikova M, Schosserer M, Flatt T, Grillari J. Ubiquitous overexpression of the DNA repair factor dPrp19 reduces DNA damage and extends drosophila life span. *NPJ Aging Mech Dis.* (2017) 3:5. doi: 10.1038/s41514-017-0008-9
12. Minocha R, Popova V, Kopytova D, Misiak D, Hüttelmaier S, Georgieva S, et al. Mud2 functions in transcription by recruiting the Prp19 and TREX complexes to transcribed genes. *Nucleic Acids Res.* (2018) 46:9749–63. doi: 10.1093/nar/gky640
13. Jie Y, Lan W, Zhu JM, Qian Y, Xue RY, Ying F, et al. Prp19 facilitates invasion of hepatocellular carcinoma via p38 mitogen-activated protein kinase/Twist1 pathway. *Oncotarget.* (2016) 7:21939–51. doi: 10.18632/oncotarget.7877
14. Confalonieri S, Quarto M, Goisis G, Nuciforo P, Donzelli M, Jodice G, et al. Alterations of ubiquitin ligases in human cancer and their association with the natural history of the tumor. *Oncogene.* (2009) 28:2959–68. doi: 10.1038/onc.2009.156
15. Peng WZ, Liu JX, Li CF, Ma R, Jie JZ. hnRNPK promotes gastric tumorigenesis through regulating CD44E alternative splicing. *Cancer Cell Int.* (2019) 19:335. doi: 10.1186/s12935-019-1020-x
16. Chen K, Xiao H, Zeng J, Yu G, Zhou H, Huang C, et al. Alternative splicing of EZH2 pre-mRNA by SF3B3 contributes to the tumorigenic potential of renal cancer. *Clin Cancer Res.* (2017) 23:3428–41. doi: 10.1158/1078-0432.CCR-16-2020
17. Liu H, Du S, Lei T, Wang H, He X, Tong R, Wang Y. Multifaceted regulation and functions of YAP/TAZ in tumors (Review). *Oncol Rep.* (2018) 40:16–28. doi: 10.3892/or.2018.6423
18. Lee CK, Jeong SH, Jang C, Bae H, Kim YH, Kim SK et al. Tumor metastasis to lymph nodesrequires YAP-dependent metabolic adaptation. *Cancer.* (2019) 363:644–9. doi: 10.1126/science.aav0173
19. Lai CJ, Lin CY, Liao WY, Hour TC, Wang HD, Chuu CP. CD44 promotes migration and invasion of docetaxel-resistant prostate cancer cells likely via induction of hippo-yap signaling. *Cells.* (2019) 8:295. doi: 10.3390/cells8040295
20. Li C, Wang S, Xing Z, Lin A, Liang K, Song J, et al. A ROR1-HER3-lncRNA signalling axis modulates the Hippo-YAP pathway to regulate bone metastasis. *Nat Cell Biol.* (2017) 19:106–19. doi: 10.1038/ncb3464
21. Tu K, Yang W, Li C, Zheng X, Lu Z, Guo C, et al. Fbxw7 is an independent prognostic marker and induces apoptosis and growth arrest by regulating YAP abundance in hepatocellular carcinoma. *Mol Cancer.* (2014) 13:110. doi: 10.1186/1476-4598-13-110
22. Shi Y, Cao T, Sun Y, Xia J, Wang P, Ma J. Nitidine chloride inhibits cell proliferation and invasion via downregulation of YAP expression in prostate cancer cells. *Am J Transl Res.* (2019) 11:709–20.
23. Wang Y, Chen K, Cai Y, Cai Y, Yuan X, Wang L, et al. Annexin A2 could enhance multidrug resistance by regulating NF-kappaB signaling pathway in pediatric neuroblastoma. *J Exp Clin Cancer Res.* (2017) 36:111. doi: 10.1186/s13046-017-0581-6
24. Shimada H, Ambros IM, Dehner LP, Hata J, Joshi VV, Roald B. Terminology and morphologic criteria of neuroblastic tumors: recommendations by the international neuroblastoma pathology committee. *Cancer.* (1999) 86:349–63. doi: 10.1002/(SICI)1097-0142(19990715)86:2<349::AID-CNCR20>3.0.CO;2-Y
25. Decarolis B, Simon T, Krug B, Leuschner I, Vokuhl C, Kaatsch P, et al. Treatment and outcome of ganglioneuroma and ganglioneuroblastoma intermixed. *BMC Cancer.* (2016) 16:542. doi: 10.1186/s12885-016-2513-9
26. Alexander N, Sullivan K, Shaikh F, Irwin MS. Characteristics and management of ganglioneuroma and ganglioneuroblastoma-intermixed in children and adolescents. *Pediatr Blood Cancer.* (2018) 65:e26964. doi: 10.1002/pbc.26964
27. Rosswog C, Schmidt R, Oberthuer A, Juraeva D, Brors B, Engesser A, et al. Molecular classification substitutes for the prognostic variables stage, age, and MYCN status in neuroblastoma risk assessment. *Neoplasia.* (2017) 19:982–90. doi: 10.1016/j.neo.2017.09.006
28. Oberthuer A, Juraeva D, Hero B, Volland R, Sterz C, Schmidt R, et al. Revised risk estimation and treatment stratification of low- and intermediate-risk neuroblastoma patients by integrating clinical and molecular prognostic markers. *Clin Cancer Res.* (2014) 21:1904–15. doi: 10.1158/1078-0432.CCR-14-0817
29. Oberthuer A, Hero B, Berthold F, Juraeva D, Faldum A, Kahlert Y, et al. Prognostic impact of gene expression–based classification for neuroblastoma. *J Clin Oncol.* (2010) 28:3506–15. doi: 10.1200/JCO.2009.27.3367
30. Vermeulen J, De Preter K, Naranjo A, Vercruysse L, Van Roy N, Hellemans J, et al. Predicting outcomes for children with neuroblastoma using a multigene-expression signature: a retrospective SIOPEN/COG/GPOH study. *Lancet Oncol.* (2009) 10:663–71. doi: 10.1016/S1470-2045(09)70154-8
31. Nieto MA, Huang RY, Jackson RA, Thiery JP. Emt: 2016. *Cell.* (2016) 166:21–45. doi: 10.1016/j.cell.2016.06.028

32. Pastushenko I, Blanpain C. EMT transition states during tumor progression and metastasis. *Trends Cell Biol.* (2019) 29:212–26. doi: 10.1016/j.tcb.2018.12.001
33. Krebs AM, Mitschke J, Lasierra Losada M, Schmalhofer O, Boerries M, Busch H, et al. The EMT-activator Zeb1 is a key factor for cell plasticity and promotes metastasis in pancreatic cancer. *Nat Cell Biol.* (2017) 19:518–29. doi: 10.1038/ncb3513
34. Zheng X, Carstens JL, Kim J, Scheible M, Kaye J, Sugimoto H, et al. Epithelial-to-mesenchymal transition is dispensable for metastasis but induces chemoresistance in pancreatic cancer. *Nature.* (2015) 527:525–30. doi: 10.1038/nature16064
35. Guan K-L. Abstract SY29-03: regulation and function of the Hippo-YAP pathway in organ size, tumorigenesis, and metastasis. *Cancer Res.* (2012) 72:SY29–03. doi: 10.1158/1538-7445.AM2012-SY29-03
36. Hong W, Guan KL. The YAP and TAZ transcription co-activators: key downstream effectors of the mammalian hippo pathway. *Semin Cell Dev Biol.* (2012) 23:785–93. doi: 10.1016/j.semcdb.2012.05.004
37. Hansen CG, Ng YL, Lam WL, Plouffe SW, Guan KL. The hippo pathway effectors YAP and TAZ promote cell growth by modulating amino acid signaling to mTORC1. *Cell Res.* (2015) 25:1299–313. doi: 10.1038/cr.2015.140
38. Zhang S, Liu Z, Wu L, Wang Y. MiR-361 targets yes-associated protein (YAP) mRNA to suppress cell proliferation in lung cancer. *Biochem Biophys Res Commun.* (2017) 492:468–73. doi: 10.1016/j.bbrc.2017.08.072
39. Nallet-Staub F, Marsaud V, Li L, Gilbert C, Dodier S, Bataille V, et al. Pro-invasive activity of the Hippo pathway effectors YAP and TAZ in cutaneous melanoma. *J Invest Dermatol.* (2014) 134:123–32. doi: 10.1038/jid.2013.319
40. Hogg R, McGrail JC, O'Keefe RT. The function of the nineteen complex (NTC) in regulating spliceosome conformations and fidelity during pre-mRNA splicing. *Biochem Soc Trans.* (2010) 38:1110–15. doi: 10.1042/BST0381110
41. Chen SP, Kao DI, Tsai WY, Cheng SC. The Prp19p-associated complex in spliceosome activation. *Science.* (2003) 302:279–82. doi: 10.1126/science.1086602
42. Chan SP, Cheng SC. The Prp19-associated complex is required for specifying interactions of U5 and U6 with pre-mRNA during spliceosome activation. *J Biol Chem.* (2005) 280:31190–9. doi: 10.1074/jbc.M505060200

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

MINI REVIEW
published: 19 November 2020
doi: 10.3389/fonc.2020.578286

The Current Status and Future Potential of Theranostics to Diagnose and Treat Childhood Cancer

Alex J. Poot[1,2], *Marnix G. E. H. Lam*[1] *and Max M. van Noesel*[2*]

[1] Department of Radiology and Nuclear Medicine, University Medical Center Utrecht, Utrecht, Netherlands, [2] Department of Solid Tumors, Princess Maxima Center for Pediatric Oncology, Utrecht, Netherlands

OPEN ACCESS

Edited by:
Hua Tan,
University of Texas Health Science Center at Houston, United States

Reviewed by:
Jaume Mora,
Hospital Sant Joan de Déu Barcelona, Spain
Yong-mi Kim,
Children's Hospital of Los Angeles, United States

***Correspondence:**
Max M. van Noesel
m.m.vannoesel@prinsesmaximacentrum.nl

Specialty section:
This article was submitted to Pediatric Oncology, a section of the journal Frontiers in Oncology

Received: *30 June 2020*
Accepted: *09 October 2020*
Published: *19 November 2020*

Citation:
Poot AJ, Lam MGEH and van Noesel MM (2020) The Current Status and Future Potential of Theranostics to Diagnose and Treat Childhood Cancer. Front. Oncol. 10:578286. doi: 10.3389/fonc.2020.578286

In theranostics (i.e., therapy and diagnostics) radiopharmaceuticals are used for both therapeutic and diagnostic purposes by targeting one specific tumor receptor. Biologically relevant compounds, e.g., receptor ligands or drugs, are labeled with radionuclides to form radiopharmaceuticals. The possible applications are multifold: visualization of biological processes or tumor biology *in vivo*, diagnosis and tumor staging, therapy planning, and treatment of specific tumors. Theranostics research is multidisciplinary and allows for the rapid translation of potential tumor targets from preclinical research to "first-in-man" clinical studies. In the last decade, the use of theranostics has seen an unprecedented value for adult cancer patients. Several radiopharmaceuticals are routinely used in clinical practice (e.g., [^{68}Ga/^{177}Lu]DOTATATE), and dozens are under (pre)clinical development. In contrast to these successes in adult oncology, theranostics have scarcely been developed to diagnose and treat pediatric cancers. To date, [$^{123/131}$I]meta-iodobenzylguanidine ([$^{123/131}$I]mIBG) is the only available and approved theranostic in pediatric oncology. mIBG targets the norepinephrine transporter, expressed by neuroblastoma tumors. For most pediatric tumors, including neuroblastoma, there is a clear need for novel and improved radiopharmaceuticals for imaging and therapy. The strategy of theranostics for pediatric oncology can be divided in (1) the improvement of existing theranostics, (2) the translation of theranostics developed in adult oncology for pediatric purposes, and (3) the development of novel theranostics for pediatric tumor-specific targets. Here, we describe the recent advances in theranostics development in pediatric oncology and shed a light on how this methodology can affect diagnosis and provide additional treatment options for these patients.

Keywords: theranostics, childhood cancer, radiopharmaceuticals, nuclear imaging (e.g. PET, SPECT), nuclear therapy

INTRODUCTION

Theranostics in nuclear medicine includes the use and application of two identical or very closely related radiopharmaceuticals for *ther*apy and diag*nos*is. In oncology, tumor-specific substrates, receptor ligands, or drugs can serve as lead for theranostic development when labeled with specific radionuclides for imaging or therapy (**Figure 1A**). As the molecular structure of both the diagnostic and therapeutic radiopharmaceuticals are identical, diagnostic images can become predictive for therapeutic response because the biological characteristics and binding potential of both are similar, irrespective of the radionuclide (2–4).

For diagnosis, positron emission tomography (PET) is a nuclear imaging technique that enables the visualization and quantification of molecules equipped with positron emitting radionuclides. The most used radionuclide for imaging is Fluorine-18 (^{18}F) in the form of [^{18}F]FDG. [^{18}F]FDG PET visualizes increased carbohydrate uptake in tissue, e.g., tumor tissue, and is important for diagnosis, staging and treatment monitoring. For PET tracer development, any molecule that displays tumor-specific targeting can be used, including small molecules, peptides or biologicals. Radionuclides used for PET tracer development are, among others, Carbon-11 (^{11}C) and Fluorine-18 (^{18}F) facilitating small molecule labeling, Gallium-68 (^{68}Ga) for peptide radiolabeling, and Copper-64 (^{64}Cu) or Zirconium-89 (^{89}Zr) for the labeling of monoclonal antibodies (mAbs) and other biologicals. PET imaging enables studying the distribution and kinetics of labeled molecules and the biochemical and physiological processes. Molecular imaging by means of PET can, thus, facilitate and guide cancer treatment in many ways (5, 6). Currently, PET is the most sensitive technique for nuclear imaging; it requires nanomolar amounts of the radiopharmaceutical for imaging. These nanomolar amounts will not induce pharmacological effects, hold minimal risks for toxicity, and are described as the micro-dosing concept. Micro-dosing allows for fast translation of novel PET tracers into clinical trials in small "first-in-man" or phase 0 studies, when produced under good manufacturing practice (GMP). Single photon emission computed tomography (SPECT) is an alternative nuclear imaging technique and enables the visualization of γ-emitting radionuclides and was the basis for

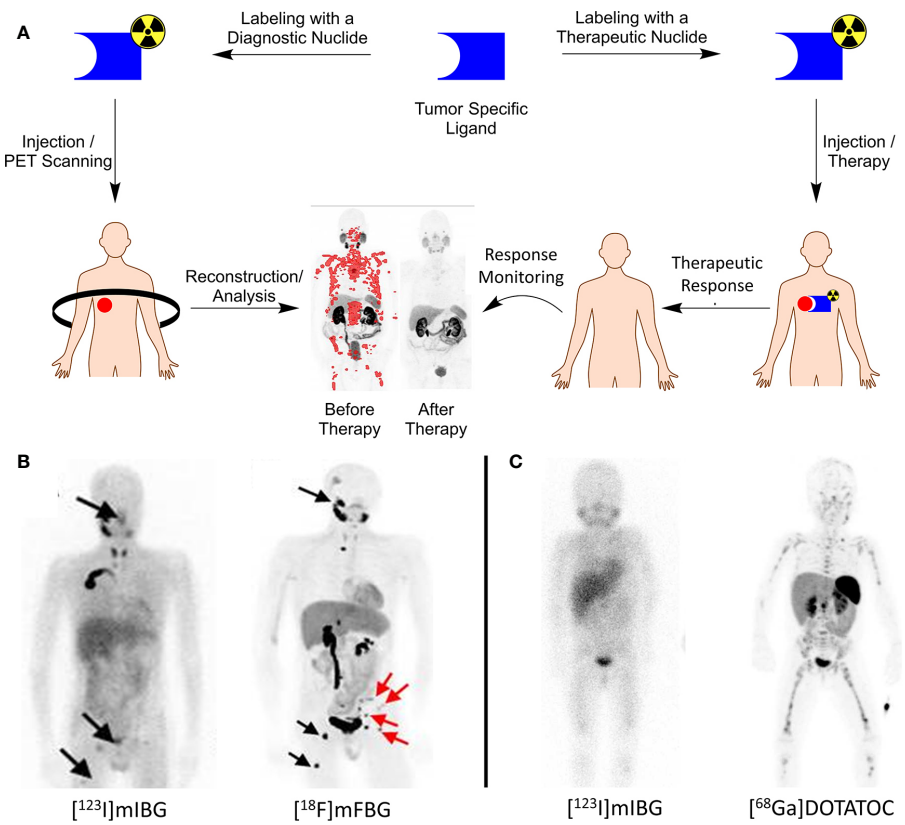


FIGURE 1 | **(A)** Theranostics concept explained. A tumor-specific ligand can be used for both imaging and therapy, dependent on the nuclide of choice. PET images before/after therapy in a prostate cancer patient diagnosed and treated with [^{68}Ga/^{177}Lu]PSMA. PET image adapted from SNMMI image of the year 2018 by Hofman et al.; **(B)** left, [^{123}I]mIBG SPECT image of a neuroblastoma patient with lesions indicated with black arrows; right, [^{18}F]mFBG PET image of the same patient showing greater contrast and additional lesions that were not observed with [^{123}I]mIBG. Image adapted from Pandit-Taskar et al. (1); **(C)** left, [^{123}I]mIBG SPECT image of a neuroblastoma patient with only vague tumor uptake; right, [^{68}Ga]DOTATOC PET image of the same patient showing SSTR-2A expression, greater contrast and additional lesions. Patient is treated with [^{177}Lu]DOTATATE with an additional survival of 24 months (unpublished data, UMC Utrecht).

early theranostics development, where, among others, the different radionuclides of iodine were used for imaging (e.g., Iodine-123 (^{123}I) and Iodine-131 (^{131}I)).

Therapeutic radiopharmaceuticals for treatment of cancer are predominantly labeled with β-emitting radionuclides. The radionuclide ^{131}I, Lutetium-177 (^{177}Lu) and Yttrium-90 (^{90}Y) are frequently used for this purpose. The emitted β-particles travel 1–12 mm through tissue upon decay while losing energy and causing cytotoxic damage to the cell to induce apoptosis. Alternatively and more recently, α-emitting radionuclides, e.g., Astatine-211 (^{211}At) or Actinium-225 (^{225}Ac) were explored for therapy (7–9). The high energy deposition and a limited range of the α-particles in tissue (0.005–0.11 mm) result in very strong cytotoxic and therapeutic effects. Nowadays, α-emitting radionuclides become more widely available, research toward the development of therapeutic radiopharmaceuticals with these radionuclides is emerging, and first-in-man studies are expected in the near future.

Successful theranostics have been developed for somatostatin receptor positive neuroendocrine tumors with [^{68}Ga/^{177}Lu]DOTATATE and prostate-specific membrane antigen (PSMA) positive prostate cancer patients as prime examples (10–13). Currently, for childhood cancers and more specifically norepinephrine transporter (NET) positive neuroblastoma tumors, [$^{123/131}$I]meta-iodobenzylguanidine ([$^{123/131}$I]mIBG) is the only available theranostic to date (14–16). Despite the proven value of theranostics in adult oncology, its potential was minimally explored for childhood cancer and is still at its infant stage. However, many opportunities and applications present themselves. In this review, we discuss different strategies for theranostics development for childhood cancer and divided these into (1) the existing theranostics and improvement thereof, (2) theranostics developed for adult oncology and translation thereof for childhood cancer, and (3) the development of novel theranostics for specific pediatric tumor targets. By describing the recent advances in theranostics research we discuss how it can affect diagnosis and therapy for childhood cancer in the future.

CURRENT THERANOSTICS IN PEDIATRIC ONCOLOGY

[$^{123/131}$I]mIBG is the only theranostic currently available for routine clinical use to image and treat neuroblastoma tumors that express the norepinephrine transporter (NET). mIBG is a structural analog of the neurotransmitter norepinephrine and is actively transported into the tumor by NET. Inside the cell, mIBG is stored in the cytoplasm, mitochondria, and in vesicular monoamine transporter (VMAT)-coated and neurosecretory vesicles (17–21). [^{123}I]mIBG SPECT imaging is currently the standard of care to diagnose primary tumors and distant metastases in neuroblastoma and for staging and disease response evaluation after treatment. In total, approximately 95% of neuroblastoma tumors are [^{123}I]MIBG avid. The remaining 5% of tumors are either well-differentiated ganglioblastoma or very undifferentiated neuroblastoma with little or no NET transporter expression. Although [^{123}I]mIBG SPECT has a high specificity and sensitivity, it also has disadvantages being poor image resolution, long scanning times, and iodine-driven thyroid toxicity. Accompanied by imaging, [^{131}I]mIBG initially showed therapeutic effectiveness in bulky tumors (22). Subsequently, it was shown that [^{131}I]mIBG was feasible and effective in the first treatment of high-risk neuroblastoma patients (23). However, two systematic reviews failed to show a survival advantage for [^{131}I]mIBG treated patients (24, 25). In two studies [^{131}I]mIBG was combined with busulfan and melphalan followed by autologous stem cell rescue. For both, acceptable toxicity in highly pretreated patients and encouraging responses were observed. This has led to the implementation of this combination for ultra-high-risk patients who failed to respond adequately during induction treatment for high-risk neuroblastoma. The current European SIOPEN VERITAS study explores the role of [^{131}I]mIBG in combination with topotecan and stem cell rescue followed by another high-dose consolidation with Buslfan and Melphalan and a second stem cell rescue. The aim is to increase the survival of these ultra-high-risk patients. In conclusion, [^{131}I]mIBG treatment is still under investigation and its definitive role has not been determined. In addition to the discussion on therapeutic response, patients receiving [^{131}I]mIBG also suffer from iodine uptake in the thyroid and increased risk for long-term thyroid dysfunction or secundary thyroid cancer. Last, after [^{131}I]mIBG administration, patients need to live in isolation for 5–7 days and strict precepts for 2–3 weeks. Despite the value of [$^{123/131}$I]mIBG as a theranostic, both imaging and therapy have serious disadvantages and limitations that steer the research toward novel approaches.

An ^{18}F-labeled analog of [^{123}I]mIBG, [^{18}F]meta-fluorobenzylguanidine ([^{18}F]mFBG]) has long been proposed as a possible PET alternative for the imaging of NET-positive neuroblastoma tumors (26). The radionuclide ^{18}F is a cyclotron produced β^+-emitter with a short range *in vivo*, resulting in a high image quality. Furthermore, PET-CT (or PET-MRI) images can be analyzed quantitatively for tracer distribution. ^{18}F-labeled radiopharmaceuticals are, therefore, ideal for high-resolution diagnosis, faster acquisition, and low radiation burden. Until recently, however, the production of [^{18}F]mFBG has been challenging. It requires a nucleophilic aromatic substitution of an electron-rich molecule (27, 28). Recent advances and novel radiofluorination reactions now give access to the production and clinical translation of [^{18}F]mFBG (1, 29).

Pandit-Taskar et al. reported the first clinical results with [^{18}F]mFBG, described a biodistribution and dosimetry study in neuroblastoma patients, and compared the results with [^{123}I]mIBG. In all five neuroblastoma patients, [^{18}F]mFBG scored better than [^{123}I]mIBG with respect to lesion counts, improved image quality, and the absence of any thyroid uptake (**Figure 1B**). These encouraging results gave rise to additional and more extensive clinical testing of [^{18}F]mFBG as an alternative to [^{123}I]mIBG as the current gold standard (**Table 1**) (30, 31).

TABLE 1 | Theranostics under preclinical development and in clinical trials for pediatric cancers.

Molecular Target	Pediatric Cancer	Theranostic	Development Phase	Pediatric Clinical Trial Number	Refs
Norepinephrine	Neuroblastoma	[$^{123/131}$I]mIBG,	Routine Care	Multiple Trials	(14–16, 22–25, 78),
Transporter		[^{18}F]mFBG,	Phase I/II	NCT02348749	(29, 30)
		[^{211}At]mABG	Preclinical		(34)
Somatostatin	Neuroblastoma	[^{68}Ga]DOTATATE,	Phase I/II	NCT04040088	(39–42),
Receptor 2A		[^{177}Lu]DOTATATE,	Phase I/II		
		[^{68}Ga]OPS202,	Phase I/II		
		[^{177}Lu]OPS201,	Phase I/II		
C-X-C Chemokine	Neuroblastoma	[^{68}Ga]Pentixafor,	Early Phase I		(48, 50),
Receptor 4	Rhabdomyosarcoma	[^{177}Lu]Pentixather	Early Phase I		
	Glioblastoma				
	ALL &AML				
Fibroblast	Glioblastoma	[^{68}Ga]FAPI,	Phase I/II		(56, 57),
Activation Protein		[^{177}Lu]FAPI	Preclinical		
Ganglioside D2	Neuroblastoma	[^{89}Zr]Dinutuximab,	Preclinical		(68–70)
	Osteosarcoma	[^{68}Ga]WHWRLPS	Preclinical		
	Glioblastoma				
B7-H3 (CD276)	Pontine Glioma	[$^{124/131}$I]8H9	Phase I/II	NCT03275402/	(75, 76),
	Neuroblastoma	(omburtamab)		NCT01502917/	
				NCT04022213	

In addition to improved imaging, research is now focused on the development of an improved alternative for [^{131}I]mIBG therapy. ^{131}I is a β^--emitter with a $t_{1/2}$ of 8.04 days. Furthermore, when [^{131}I]mIBG is metabolized and ^{131}I is released, it will accumulate in the thyroid. Therefore, the thyroid is blocked as a preventive action by administration of excess iodine to avoid undesired effects. As an alternative for ^{131}I, ^{211}At has been explored. ^{211}At is an α-emitter with a $t_{1/2}$ of 7.2 h and a range of 0.005–0.11 mm in tissue. These physical properties cause very strong cytotoxic and therapeutic effects. Furthermore, ^{211}At does not accumulate in the thyroid and potentially will not cause any undesired damage (32, 33), As ^{211}At has benefits over ^{131}I, [^{211}At]meta-astatobenzylguanidine ([^{211}At]mABG) was reported as an alternative for [^{131}I]mIBG for the treatment of NET positive tumors. To date, [^{211}At]mABG has only been evaluated in preclinical models on PC12 xenografted mice (**Table 1**). [^{211}At]mABG showed a dose-dependent tumor regression and increased survival compared to the control animals. It should, however, be noted that a high dose of [^{211}At]mABG caused the death of the animals. Therefore, the toxicity profile and maximum tolerated dose of [^{211}At]mABG needs to be assessed and compared to [^{131}I]mIBG. An important additional note is the availability of ^{211}At to produce [^{211}At]mABG, which may become a practical concern. ^{211}At can only be produced by high-energy cyclotrons, of which a few are installed worldwide, and thereby the access is limited (34).

FROM ADULT ONCOLOGY TO PEDIATRIC ONCOLOGY

Theranostics available in routine clinical care are a rich source of potential theranostic candidates in pediatric oncology.

The somatostatin receptor (SSTR) family is one of the first discovered and most successful targets identified for which theranostics were developed. To date, 5 subtypes of SSTR (i.e., SSTR-1, 2A, 3, 4, and 5) are characterized. In particular, SSTR-2A is important with high expression levels for neuroendocrine tumors. It is involved in secretion, proliferation, and the induction of apoptosis (35). For pediatric cancers, SSTR-2A expression was reported for neuroblastoma tumors by Alexander et al. as well as for neuro-oncological malignancies (e.g., glioblastomas and medulloblastomas) (36–38). Analogs of somatostatin, the natural ligand of SSTR-2A, have successfully been developed to inhibit neuroendocrine tumor growth. Radiolabeling of these compounds led to the development of [^{68}Ga]DOTATATE as a PET tracer and received FDA approval in 2016 (**Table 1**). In 2018, [^{177}Lu]DOTATATE (Lutathera, AAA/Novartis) was approved as a therapeutic agent to treat SSTR-2A positive tumors. As DOTATATE is an SSTR-2A agonist, it stimulates the receptors, potentially causing undesired tumor growth. To circumvent these agonistic effects, the theranostics pair [^{68}Ga]OPS202/[^{177}Lu]OPS201 (Ipsen) was developed as an SSTR-2A antagonist and is currently in Phase I/II trials (**Table 1**) (10, 11, 39, 40). Because SSTR-2A expression was also validated for neuroblastomas and neuro-oncological malignancies and with several theranostics available, a straightforward translation to pediatric oncology is feasible. Small-scale experimental pilot studies were reported for these pediatric cancers with [^{68}Ga]DOTATATE, and results are encouraging (41). This warrants further clinical studies on imaging and treating SSTR-2A positive pediatric cancers with these theranostics in the near future (**Figure 1C**) (42).

Another theranostic candidate target that was extensively explored in adult oncology is the C-X-C chemokine receptor 4 (CXCR4). The expression levels of CXCR4 and its natural ligand, CXCL12, are correlated to tumor development and metastasis

and were validated for breast cancer, prostate cancer, lung cancer, colorectal cancer, and primary brain tumors (43). By immunohistochemical staining, CXCR4 expression was also demonstrated for neuroblastomas, rhabdomyosarcomas, glioblastomas, and hematological malignancies (44–47). To date, several CXCR4-targeting drugs are under (pre)clinical development, e.g., Ulocuplumab, PRX177561, AMD3100, and Plerixafor, which demonstrates that CXCR4 targeting is clinically feasible and relevant. For theranostic development, the PET tracer and cyclic-pentapeptide [^{68}Ga]Pentixafor (Scintomics) is currently the most advanced and under investigation in multiple Phase I clinical trials (**Table 1**) (48). Labeling of pentixafor with ^{177}Lu or ^{90}Y to obtain the therapeutic counterpart of the diagnostic led to a strongly decreased affinity for the target receptor. This affinity could be restored after small molecular adaptations to the pentixafor scaffold and resulted in the successful development of [^{177}Lu] Pentixather (**Table 1**) (43, 49, 50). [^{68}Ga]Pentixafor and [^{177}Lu] Pentixather are candidates for clinical trials in pediatric patients as well as CXCR4 is reported for these tumors.

A target that recently received much attention is the fibroblast activation protein α (FAP) (51). FAP is a serine protease that is selectively expressed in the stromal fibroblasts of the tumor, which is often observed for breast cancer, colon cancer, and pancreatic cancers (52). FAP expression is observed in glioblastomas and can be a valuable theranostic target for pediatric cancers. FAP-specific inhibitors (FAPI) have been developed based on quinoline scaffolds. For diagnostic purposes, promising results were obtain after radiolabeling with ^{68}Ga (53, 54). In particular, [^{68}Ga]FAPI-04, -21, and -46 resulted in high-contrast images, and as a proof-of-concept, 28 different tumor types were visualized with [^{68}Ga]FAPI-04 (55, 56). All FAPI compounds allow radiolabeling with ^{177}Lu too to obtain the corresponding therapeutic radiopharmaceutical. Preclinical studies with [^{177}Lu]FAPI-21 and -46 in tumor-bearing mice gave promising results (**Table 1**) (57). As FAP is also expressed by glioblastomas, these theranostics have potential for the diagnosis and treatment of pediatric cancers.

Monoclonal antibodies (mAbs) and mAb-fragments had unprecedented impact on the treatment of cancer patients. However, clinical benefit is usually only achieved in a percentage of the patient population. The application of ^{89}Zr-labeled mAbs as ImmunoPET tracers has become increasingly important to visualize these compounds *in vivo* and assess the distribution, kinetics, and the biochemical and physiological behaviour (4, 58). Nowadays, more than 75 clinical trials are ongoing with ^{89}Zr-labeled mAbs and the radiolabeling can be achieved *via* generic methods (59, 60). Despite the clinical impact of ImmunoPET with [^{89}Zr]mAbs for adult oncology and other indications, ImmunoPET with available [^{89}Zr]mAbs has barely been explored for pediatric cancers. The only reported application of ImmunoPET was [^{89}Zr] bevacizumab in diffuse intrinsic pontine glioma to study vascular endothelial growth factor (VEGF) excretion and the potential to treat these patients with bevacizumab (61). Though ImmunoPET in pediatric cancer patients is not common, it should be anticipated that this methodology can also have an impact for these patients in the future.

SPECIFIC THERANOSTIC TARGETS IN PEDIATRIC ONCOLOGY

Pediatric cancers have a distinct biological profile with unique molecular targets that are not expressed in adult cancers. These targets embody unique opportunities for the diagnosis and treatment of pediatric cancers, but due to small patient populations, it remains a challenge to identify them and develop theranostics against these targets.

A target of interest for theranostic development is ganglioside D2 (GD2). GD2 is a glycosphingolipid and selective cellular marker that is expressed by neuroblastomas, osteosarcomas, and glioblastomas (62, 63). Though its exact function is still not fully understood, it is assumed that it plays a crucial role in cell adhesion, migration, and tumor metastasis. Dinutuximab (Unituxin®, United Therapeutics) is FDA approved, and Dinutuximab beta (Qarziba®, EUSA Pharma) is EMA approved for the treatment of GD2-positive neuroblastoma tumors (64, 65). Despite increased survival rates from 46% to 66%, for high-risk neuroblastoma patients, 30% of the patients will relapse independent of the GD2 expression levels (**Table 1**) (66). Several radiopharmaceuticals have been developed to image GD2-positive tumors. ^{64}Cu-labeled hu14.18K322A showed clear accumulation and retention in preclinical osteosarcoma models, and [^{89}Zr]dinutuximab was mentioned as a PET tracer in meeting abstracts (67–69). In addition to radiolabeled mAbs, Müller et al. reported on the development of [^{68}Ga]DOTA- WHWRLPS heptapeptide and demonstrated its accumulation in neuroblastoma xenografted mice (70). Though encouraging, clinical translation of these radiopharmaceuticals has yet to be achieved.

More recently, B7-H3 (CD276) has become a validated pediatric cancer target for immunotherapy in pontine gliomas and neuroblastomas (71, 72). To date, two mAbs were developed, Enoblituzumab (MacroGenics) and Omburtamab (Y-mAbs), to treat B7-H3 positive tumors (73, 74) Based on these immunotherapeutics, attempts at the development of theranostics are reported. Especially with 8H9 (i.e., Omburtamab) multiple clinical trials are ongoing. The theranostics pair [$^{124/131}$I]8H9 is investigated for B7-H3 positive pontine glioma tumors and a modest survival benefit was reported (**Table 1**) (75, 76). As specific brain tumors (e.g., gliomas) express B7-H3, it is important that passage and delivery of the radiopharmaceutical across the blood–brain barrier is achieved. As such, radiolabeled [$^{124/131}$I]Omburtamab is ideal to investigate drug targeting in these patients. As B7-H3 is acknowledged as a pan-tumor target, theranostics targeting B7-H3 might become of general importance for childhood cancer.

CONSIDERATIONS AND REQUIREMENTS FOR NUCLEAR MEDICINE IN CHILDHOOD CANCER

The application of theranostics for the diagnosis and treatment of childhood cancers is in its infancy. With the availability of

radiopharmaceuticals and theranostics for the various adult cancers, there is a lot of potential to translate and directly apply these for childhood cancers. Clinically available SPECT/PET tracers and therapeutic radiopharmaceuticals can directly be applied for pediatric cancers when the target is present and validated for the respective tumor type. Examples include SSTR-2A, CXCR4, and FAP-positive tumors. As childhood cancers have unique target expression profiles, with GD2 and B7-H3 as examples, novel theranostics can be developed for these yet unexplored targets. Unique target-finding programs are in place to unveil novel childhood cancer-specific biological features for which theranostics can be developed. A critical note and challenge is that target expression of cancers in general cannot always be directly correlated to positive imaging and treatment results. Preclinical research programs are, therefore, required to validate target expression and the potential of the target against which to developed theranostics.

A successfully developed theranostic that shows potential in preclinical studies warrants clinical translation. To achieve that, the theranostic needs to be produced under GMP to guarantee product quality and patient safety (77). Clinical translation of developed theranostics is relatively straightforward as procedures and GMP production facilities are widely available.

CONCLUSION

Theranostics have unprecedented value to diagnose and treat cancers. Many novel theranostics are under development and expected to enter clinical trials and care in the near future. For the diagnosis and treatment of childhood cancers, theranostics research is still in its infancy, but following the path of adult oncology, its value is promising. They are expected to become additional and valuable tools to diagnose and treat childhood cancers.

AUTHOR CONTRIBUTIONS

AP drafted the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

Sponsored by UMC Utrecht and Princess Máxima Center for Pediatric Oncology, Netherlands.

REFERENCES

1. Rotstein B, Wang L, Liu R, Patteson J, Kwan EE, Vasdev N, et al. Mechanistic Studies and Radiofluorination of Structurally Diverse Pharmaceuticals with Spirocyclic Iodonium(III) Ylides. *Chem Sci* (2016) 7:4407–17. doi: 10.1039/C6SC00197A
2. Langbein T, Weber W, Eiber M. Future of Theranostics: An Outlook on Precision Oncology in Nuclear Medicine. *J Nucl Med* (2019) 60:13S–9S. doi: 10.2967/jnumed.118.220566
3. Turner JH. Recent advances in theranostics and challenges for the future. *Br J Radiol* (2018) 91:1091. doi: 10.1259/bjr.20170893
4. Rösch F, Herzog H, Qaim SM. The Beginning and Development of the Theranostic Approach in Nuclear Medicine, as Exemplified by the Radionuclide Pair ^{86}Y and ^{90}Y. *Pharmaceuticals (Basel)* (2017) 10:56. doi: 10.3390/ph10020056
5. Van Dongen GAMS, Poot AJ, Vugts DJ. PET imaging with radiolabeled antibodies and tyrosine kinase inhibitors: immuno-PET and TKI-PET. *Tumour Biol* (2012) 33:607–15. doi: 10.1007/s13277-012-0316-4
6. Slobbe P, Poot AJ, Windhorst AD, Van Dongen GAMS. PET imaging with small-molecule tyrosine kinase inhibitors: TKI-PET. *Drug Discovery Today* (2012) 17:1175–87. doi: 10.1016/j.drudis.2012.06.016
7. Navalkissoor S, Grossman A. Targeted Alpha Particle Therapy for Neuroendocrine Tumours: The Next Generation of Peptide Receptor Radionuclide Therapy. *Neuroendocrinology* (2019) 108:256–64. doi: 10.1159/000494760
8. Guérard F, Gestin J-F, Brechbiel MW. Production of [^{211}At]-Astatinated Radiopharmaceuticals and Applications in Targeted α-Particle Therapy. *Cancer Biother Radiopharm* (2013) 28:1–20. doi: 10.1089/cbr.2012.1292
9. Morgenstern A, Apostolidis C, Kratochwil C, Sathekge M, Krolicki L, Bruchertseifer F. An Overview of Targeted Alpha Therapy with ^{225}Actinium and ^{213}Bismuth. *Curr Radiopharm* (2018) 11:200–8. doi: 10.2174/1874471011666180502104524
10. Sanli Y, Garg I, Kandathil A, Kendi T, Baladron Zanetti MJ, Kuyumcu S, et al. Neuroendocrine Tumor Diagnosis and Management: ^{68}Ga-DOTATATE PET/CT. *Am J Roentgenol* (2018) 2:267–77. doi: 10.2214/AJR.18.19881
11. Mittra ES. Neuroendocrine Tumor Therapy: ^{177}Lu-DOTATATE. *Am J Roentgenol* (2018) 2:278–85. doi: 10.2214/AJR.18.19953
12. Rahbar K, Afshar-Oromieh A, Jadvar H, Ahmadzadehfar H. PSMA Theranostics: Current Status and Future Directions. *Mol Imaging* (2018) 17. doi: 10.1177/1536012118776068
13. Ahmadzadehfar H, Rahbar K, Essler M, Biersack HJ. PSMA-Based Theranostics: A Step-by-Step Practical Approach to Diagnosis and Therapy for mCRPC Patients. *Semin Nucl Med* (2019) 50:98–109. doi: 10.1053/j.semnuclmed.2019.07.003
14. Matthay KK, Shulkin B, Laderstein R, Michon J, Giammarile F, Lewington V, et al. Criteria for evaluation of disease extent by ^{123}I-metaiodobenzylguanidine scans in neuroblastoma: a report for the International Neuroblastoma Risk Group (INRG) Task Force. *Br J Cancer* (2010) 102:1319–26. doi: 10.1038/sj.bjc.6605621
15. Matthay KK, Panina C, Huberty J, Price D, Glidden DV, Tang HR, et al. Correlation of tumor and whole-body dosimetry with tumor response and toxicity in refractory neuroblastoma treated with ^{131}I-MIBG. *J Nucl Med* (2001) 42:1713–21.
16. Kraal KCJM, Tytgat GAM, Van Eck-Smit BLF, Kam B, Caron HN, Van Noesel MM. Upfront treatment of high-risk neuroblastoma with a combination of ^{131}I-MIBG and Topotecan. *Pediatr Blood Cancer* (2015) 62:1886–91. doi: 10.1002/pbc.25580
17. Vaidyanathan G. Meta-iodobenzylguanidine and analogues: chemistry and biology. *Q J Nucl Med Mol Imaging* (2008) 52:351–68.
18. Glowniak JV, Kilty JE, Amara SG, Hoffman BJ, Turner FE. Evaluation of metaiodobenzylguanidine uptake by the norepinephrine, dopamine and serotonin transporters. *J Nucl Med* (1993) 34:1140–6.
19. Gaze MN, Huxham IM, Mairs RJ, Barrett A. Intracellular localization of metaiodobenzyl guanidine in human neuroblastoma cells by electron spectroscopic imaging. *Int J Cancer* (1991) 47:875–80. doi: 10.1002/ijc.2910470615
20. Smets LA, Janssen M, Metwally E, Loesberg C. Extragranular storage of the neuron blocking agent meta-iodobenzylguanidine (MIBG) in human neuroblastoma cells. *Biochem Pharmacol* (1990) 39:1959–64. doi: 10.1016/0006-2952(90)90615-R
21. Smets LA, Janssen M, Rutgers M, Ritzen K, Buttenhuis C. Pharmacokinetics and intracellular distribution of the tumor-targeted radiopharmaceutical m-

iodo-benzylguanidine in SK-N-SH neuroblastoma and PC-12 pheochromocytoma cells. *Int J Cancer* (1991) 48:609–15. doi: 10.1002/ijc.2910480421
22. Hoefnagel CA, De Kraker J, Valdes Olmos RA, Voûte PA. 131I-MIBG as a first-line treatment in high-risk neuroblastoma patients. *Nucl Med Commun* (1994) 9:712–7. doi: 10.1097/00006231-199409000-00008
23. Kraal KCJM, Bleeker GM, Van Eck-Smit BLF, van Eijkelenburg NKA, Berthold F, Van Noesel MM, et al. Feasibility, toxicity and response of upfront metaiodobenzylguanidine therapy therapy followed by German Pediatric Oncology Group Neuroblastoma 2004 protocol in newly diagnosed stage 4 neuroblastoma patients. *Eur J Cancer* (2017) 76:188–96. doi: 10.1016/j.ejca.2016.12.013
24. Kraal KCJM, Van Dalen EC, Tytgat GAM, Van Eck-Smit BLF. Iodine-131-meta-iodobenzylguanidine therapy for patients with newly diagnosed high-risk neuroblastoma. *Cochrane Database Syst Rev* (2017) 4:CD010349. doi: 10.1002/14651858.CD010349.pub2
25. Wilson JS, Gains JE, Moroz V, Wheatley K, Gaze MN. A systematic review of 131I-meta iodobenzylguanidine molecular radiotherapy for neuroblastoma. *Eur J Cancer* (2014) 50:801–5. doi: 10.1016/j.ejca.2013.11.016
26. Garg PK, Garg S, Zalutsky MR. Synthesis and preliminary evaluation of para- and meta-[18F]fluorobenzylguanidine. *Nucl Med Biol* (1994) 21:97–103. doi: 10.1016/0969-8051(94)90135-x
27. Cole EL, Stewart MN, Littich R, Hoareau R, Scott PJ. Radiosyntheses using fluorine-18: the art and science of late stage fluorination. *Curr Top Med Chem* (2014) 14:875–900. doi: 10.2174/1568026614666140202205035
28. Van der Born D, Pees A, Poot AJ, Orru RVA, Windhorst AD, Vugts DJ. Fluorine-18 labelled building blocks for PET tracer synthesis. *Chem Soc Rev* (2017) 46:4709–73. doi: 10.1039/c6cs00492j
29. Tredwell M, Preshlock SM, Taylor NJ, Gruber S, Huiban M, Passchier J, et al. A General Copper Mediated Nucleophilic ^{18}F Fluorination of Arenes. *Angew Chem Int Ed* (2014) 53:7751–5. doi: 10.1002/anie.201404436
30. Pandit-Taskar N, Zanzonico P, Staton KD, Carrasquillo JA, Reidy-Lagunes D, Lyashchenko S, et al. Biodistribution and dosimetry of ^{18}F-Meta Fluorobenzyl Guanidine (MFBG): A first-in-human PET-CT imaging study of patients with neuroendocrine malignancies. *J Nucl Med* (2018) 59:147–53. doi: 10.2967/jnumed.117.193169
31. Zhang H, Huang R, Cheung N-KV, Guo H, Zanzonico PB, Thaler HT, et al. Imaging the norepinephrine transporter in neuroblastoma: a comparison of [^{18}F]-MFBG and ^{123}I-MIBG. *Clin Cancer Res* (2014) 20:2182–91. doi: 10.1158/1078-0432.CCR-13-1153
32. Zalutsky MR, Pruszynski M. Astatine-211: Production and Availability. *Curr Radiopharm* (2011) 4:177–85. doi: 10.2174/1874471011104030177
33. Vaidyanethan G, Zalutsky MR. Applications of ^{211}At and ^{223}Ra in Targeted Alpha-Particle Radiotherapy. *Curr Radiopharm* (2011) 4:283–94. doi: 10.2174/1874471011104040283
34. Ohshima Y, Sudo H, Watanabe S, Nagatsu K, Tsuji AB, Sakashita T, et al. Antitumor effects of radionuclide treatment using α-emitting meta-^{211}At-astato-benzylguanidine in a PC12 pheochromocytoma model. *Eur J Nucl Med Mol Imaging* (2018) 45:999–1010. doi: 10.1007/s00259-017-3919-6
35. Mizutani G, Nakahishi Y, Watanabe N, Honma T, Obana Y, Seki T, et al. Expression of Somatostatin Receptor (SSTR) Subtypes (SSTR-1, 2A, 3, 4 and 5) in Neuroendocrine Tumors Using Real-time RT-PCR Method and Immunohistochemistry. *Acta Histochem Cytomchem* (2012) l45:167–76. doi: 10.1267/ahc.12006
36. Alexander N, Marrano P, Thorner P, Naranjo A, Van Ryn C, Martinez D, et al. Prevalence and clinical correlations of somatostatin receptor-2 (SSTR2) expression in neuroblastoma. *J Pediatr Hematol Oncol* (2019) 41:222–7. doi: 10.1097/MPH.0000000000001326
37. Kiviniemi A, Gardberg M, Kivinen K, Posti JP, Vourinen V, Sipilä J, et al. Somatostatin receptor 2A in gliomas: Association with oligodendrogliomas and favourable outcome. *Oncotarget* (2017) 8:49123–32. doi: 10.18632/oncotarget.17097
38. Guyotat J, Champier J, Pierre GS, Jouvet A, Bret P, Brisson C, et al. Differential Expression of Somatostatin Receptors in Medulloblastoma. *J Neuro-Oncol* (2001) 51:93–103. doi: 10.1023/a:1010624702443
39. Chen S, Cheung W, Leung YL, Cheng K, Wong KN, Wong YH, et al. ^{177}Lu-DOTATATE radionuclide therapy for pediatric patients with relapsed high-risk neuroblastoma negative on ^{131}I-MIBG imaging - a pilot study. *J Nucl Med* (2018) 59:Suppl 1–307.
40. Nicolas GP, Beykan S, Bouterfa H, Kaufmann J, Bauman A, Lassmann M, et al. Safety, Biodistribution, and Radiation Dosimetry of ^{68}Ga-OPS202 in Patients with Gastroenteropancreatic Neuroendocrine Tumors: A Prospective Phase I Imaging Study. *J Nucl Med* (2018) 59:909–14. doi: 10.2967/jnumed.117.199737
41. Nicolas GP, Mansi R, McDougall L, Kaufman J, Bouterfa H, Wild D, et al. Biodistribution, Pharmacokinetics, and Dosimetry of ^{177}Lu-, ^{90}Y-, and ^{111}In-Labeled Somatostatin Receptor Antagonist OPS201 in Comparison to the Agonist ^{177}Lu-DOTATATE: The Mass Effect. *J Nucl Med* (2017) 58:1435–41. doi: 10.2967/jnumed.117.191684
42. *search term: Neuroblastoma, SSTR2, DOTATATE* (2020). Available at: www.clinicaltrials.gov (Accessed May 25, 2020).
43. Kircher M, Herhaus P, Schottelius M, Buck AK, Werner RA, Wester H-J, et al. CXCR4-directed theranostics in oncology and inflammation. *Ann Nucl Med* (2018) 32:503–11. doi: 10.1007/s12149-018-1290-8
44. Russell HV, Hicks J, Okcu MF, Nuchtern JG. CXCR4 expression in neuroblastoma primary tumors is associated with clinical presentation of bone and bone marrow metastases. *J Pediatr Surg* (2004) 39:1506–11. doi: 10.1016/j.jpedsurg.2004.06.019
45. Miyoshi K, Kohashi K, Fushimi F, Yamamoto H, Kishimoto J, Taguchi T, et al. Close correlation between CXCR4 and VEGF expression and frequent CXCR7 expression in rhabdomyosarcoma. *Hum Pathol* (2014) 45:1900–9. doi: 10.1016/j.humpath.2014.05.012
46. Eckert F, Schilbach K, Klumpp L, Bardoscia L, Sezgin EC, Schwab M, et al. Potential Role of CXCR4 Targeting in the Context of Radiotherapy and Immunotherapy of Cancer. *Front Immunol* (2018) 9:3018:3018. doi: 10.3389/fimmu.2018.03018
47. Peled A, Klein S, Beider K, Burger JA, Abraham M. Role of CXCL12 and CXCR4 in the pathogenesis of hematological malignancies. *Cytokine* (2018) 109:11–6. doi: 10.1016/j.cyto.2018.02.020
48. *search term: Pentixafor, CXCR4* (2020). Available at: www.clinicaltrials.gov (Accessed June 8, 2020).
49. Cooper TM, Sison EAR, Baker SD, Li L, Ahmed A, Trippett T, et al. A phase 1 study of the CXCR4 antagonist plerixafor in combination with high-dose cytarabine and etoposide in children with relapsed or refractory acute leukemias or myelodysplastic syndrome: A Pediatric Oncology Experimental Therapeutics Investigators' Consortium study (POE 10-03). *Pediatr Blood Cancer* (2017) 64. doi: 10.1002/pbc.26414
50. Herrmann K, Schottelius M, Lapa C, Osl T, Poschenrieder A, Hänscheid H, et al. First-in-Human Experience of CXCR4-Directed Endoradiotherapy with ^{177}Lu- and ^{90}Y-Labeled Pentixather in Advanced-Stage Multiple Myeloma with Extensive Intra- and Extramedullary Disease. *J Nucl Med* (2016) 57:248–51. doi: 10.2967/jnumed.115.167361
51. Busek P, Mateu R, Zubal M, Kotackova L, Sedo A. Targeting Fibroblast Activation Protein in Cancer - Prospects and Caveats. *Front Biosci (Landmark Ed)* (2018) 23:1933–68. doi: 10.2741/4682
52. Puré E, Blomberg R. Pro-tumorigenic roles of fibroblast activation protein in cancer: back to the basics. *Oncogene* (2018) 37:4343–57. doi: 10.1038/s41388-018-0275-3
53. Busek P, Balaziova E, Matrasova I, Hilser M, Tomas R, Syrucek M, et al. Fibroblast Activation Protein Alpha Is Expressed by Transformed and Stromal Cells and Is Associated With Mesenchymal Features in Glioblastoma. *Tumour Biol* (2016) 37:13961–71. doi: 10.1007/s13277-016-5274-9
54. Jansen K, Heirbaur L, Cheng JD, Joossens J, Ryabtsova O, Cos P, et al. Selective Inhibitors of Fibroblast Activation Protein (FAP) with a (4-Quinolinoyl)-glycyl-2-cyanopyrrolidine Scaffold. *ACS Med Chem Lett* (2013) 4:491–6. doi: 10.1021/ml300410d
55. Lindner T, Loktev A, Altmann A, Giesel F, Kratochwil C, Debus J, et al. Development of Quinoline-Based Theranostic Ligands for the Targeting of Fibroblast Activation Protein. *J Nucl Med* (2018) 59:1415–22. doi: 10.2967/jnumed.118.210443
56. Giesel FL, Kratochwil C, Lindner T, Marschalek MM, Lehnert W, Debus J, et al. ^{68}Ga-FAPI PET/CT: Biodistribution and Preliminary Dosimetry Estimate of 2 DOTA-Containing FAP-Targeting Agents in Patients with

Various Cancers. *J Nucl Med* (2018) 60:386–92. doi: 10.2967/jnumed.118.215913
57. Kratochwil C, Flechsig P, Lindner T, Abderrahim L, Altmann A, Mier W, et al. ^{68}Ga-FAPI PET/CT: Tracer Uptake in 28 Different Kinds of Cancer. *J Nucl Med* (2019) 60:801–5. doi: 10.2967/jnumed.119.227967
58. Loktev A, Lindner T, Burger A-M, Altmann A, Giesel F, Kratochwil C, et al. Development of Fibroblast Activation Protein-Targeted Radiotracers With Improved Tumor Retention. *J Nucl Med* (2019) 60:1421–9. doi: 10.2967/jnumed.118.224469
59. Jauw YWS, Menke-van der Houven van Oordt CW, Hoekstra OS, Hendrikse NH, Vugts DJ, Zijlstra JM, et al. Immuno-Positron Emission Tomography with Zirconium-89-Labeled Monoclonal Antibodies in Oncology: What Can We Learn from Initial Clinical Trials? *Front Pharmacol* (2016) 7:131. doi: 10.3389/fphar.2016.00131
60. Heskamp S, Raavé R, Boerman O, Rijpkema M, Goncalves V, Denat F. ^{89}Zr-Immuno-Positron Emission Tomography in Oncology: State-of-the-Art ^{89}Zr Radiochemistry. *Bioconjug Chem* (2017) 28:2211–23. doi: 10.1021/acs.bioconjchem.7b00325
61. *89Zr, Zirconium-89* (2020). Available at: www.clinicaltrials.gov (Accessed June 8, 2020).
62. Jansen MH, Veldhuijzen van Zanten SEM, Van Vuurden DG, Huisman MC, Vugts DJ, Hoekstra O, et al. Molecular Drug Imaging: ^{89}Zr-Bevacizumab PET in Children with Diffuse Intrinsic Pontine Glioma. *J Nucl Med* (2017) 58:711–6. doi: 10.2967/jnumed.116.180216
63. Sait S, Modak S. Anti-GD2 immunotherapy for neuroblastoma. *Expert Rev Anticanc* (2017) 17:889–904. doi: 10.1080/14737140.2017.1364995
64. Hung J-T, Yu A. Chapter-4: GD2-targeted immunotherapy of neuroblastoma. In: *Neuroblastoma, Molecular Mechanisms and Therapeutic Interventions*. Cambridge, United States: Academic Press (2019). p. 63–78. doi: 10.1016/B978-0-12-812005-7.00004-7
65. Hoy S. Dinutuximab, a review in high-risk neuroblastoma. *Target Oncol* (2016) 11:247–53. doi: 10.1007/s11523-016-0420-2
66. Yu AL, Gilman AL, Ozkaynak MF, London WB, Kreissman SG, Chen HX, et al. Anti-GD2 antibody with GM-CSF, interleukin-2, and isotretinoin for neuroblastoma. *N Engl J Med* (2010) 363:1324–34. doi: 10.1056/NEJMoa0911123
67. Terzic T, Cordeau M, Herblot S, Teira P, Cournoyer S, Beaunoyer M, et al. Expression of Disialoganglioside (GD2) in Neuroblastic Tumors: A Prognostic Value for Patients Treated With Anti-GD2 Immunotherapy. *Ped Dev Path* (2018) 21:355–62. doi: 10.1177/1093526617723972
68. Butch ER, Mead PE, Diaz VA, Tillman H, Stewart E, Mishra JK, et al. Positron Emission Tomography Detects In Vivo Expression of Disialoganglioside GD2 in Mouse Models of Primary and Metastatic Osteosarcoma. *Cancer Res* (2019) 79:3112–24. doi: 10.1158/0008-5472.CAN-18-3340
69. Butch E, Mishra J, Diaz VA, Vavere A, Snyder S. Selective detection of GD2-positive pediatric solid tumors using ^{89}Zr-Dinutuximab PET to facilitate anti-GD2 immunotherapy. *J Nucl Med* (2018) 59:suppl 1–170.
70. Müller J, Reichel R, Vogt S, Sauerwein W, Brandau W, Eggert A, et al. Identification and Tumour-Binding Properties of a Peptide with High Affinity to the Disialoganglioside GD2. *PloS One* (2016) 11:e0163648. doi: 10.1371/journal.pone.0163648
71. Castellanos JR, Purvis IJ, Labak CM, Guda MR, Tsung AJ, Velpula KK, et al. B7-H3 role in the immune landscape of cancer. *Am J Clin Exp Immunol* (2017) 6:66–75.
72. Flem-Karlsen K, Fodstad O, Tan M, Nunes-Xavier CE. B7-H3 in Cancer – Beyond Immune Regulation. *Trends Cancer* (2018) 4:401–4. doi: 10.1016/j.trecan.2018.03.010
73. Powderly J, Cote G, Flaherty K, Szmulewitz RZ, Ribas A, Weber J, et al. Interim results of an ongoing Phase I, dose escalation study of MGA271 (Fc-optimized humanized anti-B7-H3 monoclonal antibody) in patients with refractory B7-H3-expressing neoplasms or neoplasms whose vasculature expresses B7-H3. *J Immunother Cancer* (2015) 3:Suppl 2-O8. doi: 10.1186/2051-1426-3-S2-O8
74. Modak S, Kramer K, Gultekin SH, Guo HF, Cheung N-KV. Monoclonal Antibody 8H9 Targets a Novel Cell Surface Antigen Expressed by a Wide Spectrum of Human Solid Tumors. *Cancer Res* (2001) 61:4048–54.
75. Kramer K, Kushner BH, Modak S, Pandit-Taskar N, Smith-Jones P, Zanzonico P, et al. Compartmental intrathecal radioimmunotherapy: results for treatment for metastatic CNS neuroblastoma. *J Neurooncol* (2010) 97:409–18. doi: 10.1007/s11060-009-0038-7
76. Luther N, Zhou Z, Zanzonico P, Cheung NK, Humm J, Edgar MA, et al. The potential of theragnostic ^{123}I-8H9 convection-enhanced delivery in diffuse intrinsic pontine glioma. *Neuro Oncol* (2014) 6:800–6. doi: 10.1093/neuonc/not298
77. *European Pharmacopoea 10th edition* (2020). Available at: https://www.edqm.eu/en/european-pharmacopoeia-ph-eur-10th-edition (Accessed June 8, 2020).
78. Weiss B, Vora A, Huberty J, Hawkins RA, Matthay KK. Secondary myelodysplastic syndrome and leukemia following 131I-metaiodobenzylguanidine therapy for relapsed neuroblastoma. *J Pediatr Hematol Oncol* (2003) 7:543–7. doi: 10.1097/00043426-200307000-00009

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

ORIGINAL RESEARCH
published: 14 December 2020
doi: 10.3389/fonc.2020.596191

A Targeted Gene Panel for Circulating Tumor DNA Sequencing in Neuroblastoma

Flora Cimmino[1†], *Vito Alessandro Lasorsa*[1,2†], *Simona Vetrella*[3], *Achille Iolascon*[1,2] *and Mario Capasso*[1,2*]

[1] CEINGE Biotecnologie Avanzate, Napoli, Italy, [2] Dipartimento di Medicina Molecolare e Biotecnologie Mediche, Università degli Studi di Napoli Federico II, Napoli, Italy, [3] Department of Pediatric Oncology, Santobono-Pausilipon Children's Hospital, Naples, Italy

OPEN ACCESS

Edited by:
Jing He,
Guangzhou Medical University, China

Reviewed by:
Jiani Yin,
University of California, Los Angeles, United States
Andrea Di Cataldo,
University of Catania, Italy

***Correspondence:**
Mario Capasso
mario.capasso@unina.it

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Pediatric Oncology, a section of the journal Frontiers in Oncology

Received: 18 August 2020
Accepted: 27 October 2020
Published: 14 December 2020

Citation:
Cimmino F, Lasorsa VA, Vetrella S, Iolascon A and Capasso M (2020) A Targeted Gene Panel for Circulating Tumor DNA Sequencing in Neuroblastoma. Front. Oncol. 10:596191. doi: 10.3389/fonc.2020.596191

Background: Liquid biopsies do not reflect the complete mutation profile of the tumor but have the potential to identify actionable mutations when tumor biopsies are not available as well as variants with low allele frequency. Most retrospective studies conducted in small cohorts of pediatric cancers have illustrated that the technology yield substantial potential in neuroblastoma.

Aim: The molecular landscape of neuroblastoma harbors potentially actionable genomic alterations. We aimed to study the utility of liquid biopsy to characterize the mutational landscape of primary neuroblastoma using a custom gene panel for ctDNA targeted sequencing.

Methods: Targeted next-generation sequencing (NGS) was performed on ctDNA of 11 patients with primary neuroblastoma stage 4. To avoid the detection of false variants, we used UMIs (unique molecular identifiers) for the library construction, increased the sequencing depth and developed *ad hoc* bioinformatic analyses including the hard filtering of the variant calls.

Results: We identified 9/11 (81.8%) patients who carry at least one pathogenic variation. The most frequently mutated genes were *KMT2C* (five cases), *NOTCH1/2* (four cases), *CREBBP* (three cases), *ARID1A/B* (three cases), *ALK* (two cases), *FGFR1* (two cases), *FAT4* (two cases) and *CARD11* (two cases).

Conclusions: We developed a targeted NGS approach to identify tumor-specific alterations in ctDNA of neuroblastoma patients. Our results show the reliability of our approach to generate genomic information which can be integrated with clinical and pathological data at diagnosis.

Keywords: genetic mutation, bioinformactics analysis, next generation sequencing, neuroblastoma, liquid biopsy and circulating tumor DNA

INTRODUCTION

Circulating tumor DNA (ctDNA), a subfraction of cell-free DNA (cfDNA), is fragmented genomic DNA poured in the blood flow and other biological fluids as the result of apoptosis and necrosis of tumor cells (1, 2). The isolation and sequencing of ctDNA from biological fluids is called Liquid Biopsy (LB). LB is a low-cost and safe non-surgical procedure to access tumor's genetic information. It is a valid alternative for tumors that are not easy to tissue biopsy and represents a complementary tool for more accurate diagnoses. Furthermore, the ctDNA is representative of intra-tumoral and metastasis heterogeneity. The analysis of ctDNA can be useful for monitoring tumor clonal evolution and for the detection of therapy-relevant novel mutations arising during treatment (3, 4).

The development of ctDNA assays relies on the identification and quantification of somatic mutations. A variety of technical approaches have been optimized to detect diagnostic/prognostic/therapeutic markers in ctDNA with great sensitivity, including BEAMing technologies (5), panel sequencing of cancer-associated genes (6, 7), targeted amplicon sequencing (8, 9) and droplet digital PCR technologies (10). Currently, the FDA-approved qPCR test to identify Epidermal Growth Factor Receptor (EGFR) mutations, is a diagnostic test that replaces the tissue biopsy in patients with metastatic non-small cell lung cancer who would be eligible for treatment with EGFR-targeted therapy (erlotinib) (11). Nevertheless, comprehensive genomic analysis such as Whole Exome Sequencing (WES) or Whole Genome Sequencing (WGS) of ctDNA, cannot be considered in diagnostic routines because of the detection of large numbers of variants with uncertain significance. This, makes it difficult to interpret the data in a clinical setting and, in addition, often requires the profiling of a tumor tissue biopsy as a reference. However, ctDNA assays designed to target selected genes in customized panels could map tumor heterogeneity and may facilitate the identification of druggable mutations. Several commercially available ctDNA sequencing panels, designed to target specific exons or mutational hotspots, have shown their validity in clinical settings in adult cancers as in lung cancers (12).

The completely different patterns of genetic mutations between pediatric and adult cancers emphasize the need to develop specific approaches for ctDNA profiling of pediatric cancers. Recent comprehensive sequencing efforts show that only 45% of driver genes in pediatric cancers correspond to those found in adults, as demonstrated through pan-cancer studies. In children, these genes are mainly involved in biological processes belonging to epigenetic/chromatin remodelling pathways (25%). By contrast, the most relevant biological processes affected in adult cancers belong to PIK3 pathway (31%) which is altered only in the 3% of pediatric cancers (13, 14).

In the last year, feasibility of detecting, quantifying, and profiling ctDNA has been established in patients with five of the most common pediatric solid tumors: Ewing sarcoma, osteosarcoma, alveolar rhabdomyosarcoma, Wilms tumor and neuroblastoma (15, 16).

Clinical stages of neuroblastoma, according to the International Neuroblastoma Staging System (INSS), have been divided as the following: localized stages 1, 2, or 3; disseminated stage 4; and disseminated stage 4S occurring in patients younger than 1 year of age. Moreover, the International Neuroblastoma Risk Group classification system (INRG) classifies neuroblastoma patients in four categories: high-, intermediate-, low- and very low-risk (17). High-risk disease is usually diagnosed in children older than 18 months and, despite multimodal treatment, half of these patient's relapses. *MYCN* genomic amplification is reported in 40% of high-risk neuroblastomas. It is the strongest predictor of poor prognosis and tumor progression although other chromosomal alterations are reported as poor prognostic features, namely the deletion of 1q (30%) and of the 11q (45%) and unbalanced gain of the 17q (60%) (18). Children older than 6 years present unique structural variants with 19p loss and 1q gain among those more recurrent (19). Recent whole genome sequencing analysis in a large cohort of neuroblastoma patients have identified a paucity of recurrent alterations with point mutations in *ALK* (8–10%) and in *ATRX*, being the most frequent (20, 21). The limited burden of mitochondrial DNA mutations with potential pathogenic impact have also been assessed (22). Nevertheless, a recent work has highlighted the involvement of noncoding somatic variants, located in regulatory DNA regions, in neuroblastoma development (23). Pathogenic germline variants can also contribute to neuroblastoma onset (24). Indeed, both common (25, 26) and rare germline (20, 21) variants have been found to associate with neuroblastoma development and progression.

Neuroblastoma shed high amounts of cfDNA in the blood flow depending on the tumoral burden (mean: 1.034 ng/ml of plasma; range: 13.53–26) at diagnosis and which increase with disease progression and decrease after therapy and surgical resection (27, 28). The documented high portion of ctDNA in the cfDNA fraction (mean: 60%, range: 3–99%) in high-risk diseases and metastatic cases further confirms an important shredding of ctDNA into the bloodstream (29). The majority of ctDNA studies in neuroblastoma are based on digital droplet PCR or targeted sequencing that require prior characterization of the biomarkers such as *MYCN* amplification and activating *ALK* mutations (30). The diagnostic utility of high throughput ctDNA sequencing in neuroblastoma has been published in four studies. In all cases genomic alterations detected from cfDNA were not detected from tissue biopsy. WES analysis was successfully applied to cfDNA samples at diagnosis and highlighted that ctDNA profiling performed better than primary tissue profiling in capturing tumor heterogeneity and low frequency variants. Particularly, WES analysis of primary neuroblastoma biopsies and cfDNA identified an overlap of only 41% for Single Nucleotide Variants (SNVs), whereas an overlap of 93% was observed for Copy Number Alterations (CNAs) (29). Longitudinal follow-up studies have been performed in a limited number of patients. Accumulating aberrations were found during the evolution of therapy resistant clones and some of these were potentially targetable (28, 31). Furthermore, shallow whole-genome sequencing of ctDNA to assess copy number profiles has been proposed as a valid and noninvasive genomic test alternative to the analysis of often scarce or small biopsies (32).

In this study, we report the development of a targeted Next Generation Sequencing (NGS) gene panel for ctDNA sequencing that is tailored to the genetic landscape of neuroblastoma. We believe that the ctDNA detection and sequencing, through an *ad hoc* designed gene panel, could complement or replace tissue biopsies necessary for diagnosis of neuroblastoma. The application of the suggested strategy for the diagnosis of neuroblastoma, when optimized, may improve disease stratification at diagnosis and provide useful information for therapy decisions making.

METHODS

Patients

A total of 20 patients, with a diagnosis of neuroblastoma, were recruited at the IRCCS Istituto Giannina Gaslini and Ospedale Pediatrico Santobono-Pausilipon. Upon initial diagnosis, bone marrow biopsies and/or aspirates were obtained for microscopic examination and identification of neuroblastoma cells. Genetic abnormalities (amplification of the *MYCN* gene and deletion of the short arm of chromosome 1 [1p36]) were detected by fluorescence *in situ* hybridization. As described below, only 11 ctDNA samples qualified for deep sequencing. Of these, nine were from Stage 4 and two from Stage 2 patients. Informed consent was obtained through Research Ethics Committee of University of Naples Federico II. All experiments were performed following relevant guidelines and regulations.

Samples Collection

Venous blood samples were collected into ethylenediaminetetraacetic acid-coated tubes and centrifuged at 1,600×*g* for 10 min. Supernatants were transferred to fresh tubes and centrifuged at 16,000×*g* for 10 min. Plasma was removed and stored at − 80 °C until DNA extraction.

Development of Cancer Gene Sequencing Panel

The genes included in the NGS panel were selected according to these following criteria: first, we selected genes with high relevance of mutations in neuroblastoma, as reported in the Catalogue of Somatic Mutations in Cancer (COSMIC) database and PubMed; then we select genes mutated in more than two neuroblastoma samples. We included all the coding exons extended of 10 bp of flanking introns of 68 neuroblastoma-associated genes. The final length of our target was 0.5 Mb.

NGS Library Preparation and Sequencing for Genomic Germline DNA

Genomic germline DNA was extracted from peripheral blood leukocytes using the QIAamp DNA Blood Midi Kit (Qiagen; cat #51183). A total amount of 1.0 μg genomic DNA per sample was used as input material for the DNA library preparation. Sequencing libraries were generated using Agilent SureSelect Human All Exon kit (Agilent Technologies, CA, USA) following manufacturer's recommendations and index codes were added to each sample. Briefly, fragmentation was carried out by hydrodynamic shearing system (Covaris, Massachusetts, USA) to generate 180–280 bp fragments. Remaining overhangs were converted into blunt ends *via* exonuclease/polymerase activities and enzymes were removed. After adenylation of 3' ends of DNA fragments, adapter oligonucleotides were ligated. DNA fragments with ligated adapter molecules on both ends were selectively enriched in a PCR reaction. After PCR reaction, library hybridize with Liquid phase with biotin labeled probe, after which streptomycin-coated magnetic beads are used to capture the exons of genes. Captured libraries were enriched in a PCR reaction to add index tags to prepare for hybridization. Products were purified using AMPure XP system (Beckman Coulter, Beverly, USA) and quantified using the Agilent high sensitivity DNA assay on the Agilent Bioanalyzer 2100 system.

NGS Library Preparation and Sequencing for cfDNA

cfDNA isolation from 20 liquid biopsy samples was done according to the protocol (QIAamp Circulating Nucleic Acid Kit). The cfDNA quantity was assessed with dsDNA HS assay kit using Qubit Fluorometer 4.0 (Thermo Fisher Scientific). cfDNA quality was assessed with the Agilent High Sensitivity D1000 ScreenTape System (Agilent Technologies). After ctDNA quality control procedures, we excluded 9 samples due to low quality or low quantity. Therefore, eleven qualified samples could be used for sequencing. Next generation sequencing experiments on liquid biopsy samples were performed by Genomix4life S.r.l. (Baronissi, Salerno, Italy). Illumina TruSeq DNA indexed libraries were prepared from 1.7 to 10 ng of cfDNA using xGen Custom target capture Library Prep (IDT). This library preparation method incorporates unique molecular identifiers (UMIs). Libraries were quantified used Qubit Fluorometer 4.0 (Thermo Fisher Scientific) and pooled to an equimolar amount of each index-tagged sample to a final concentration of 2 nM. Pooled samples were subject to cluster generation and sequenced on NextSeq platform (Illumina) in a 2 × 150 paired-end format. The raw sequence files generated (fastq files) underwent quality control analysis with FastQC.

Bioinformatic Pipeline

Germline DNA was sequenced by WES. Illumina paired-end reads were mapped versus the reference genome (GRCh37/hg19) using the BWA (Burrows-Wheeler Aligner; Version: 0.7.12) (33) algorithm. PCR duplicated reads were removed with Samtools (Version: 1.9) (34).

cfDNA was sequenced as described above. Sequencing reads in the FASTQ files were processed by partially modifying the IDT analysis guidelines (eu.idtdna.com) as described below.

First, we trimmed sequencing adapters. Then, we moved molecular tags (UMIs) from the read sequence to the read name (see below) with the Fastp tool (Version: 0.20.0) (35). Then, we performed read mapping by using BWA. Then, we used fgbio (Version: 0.8.1; Fulcrum genomics, http://fulcrumgenomics.github.io/fgbio/) to manipulate BAM files and extract consensus reads that were mapped again to the reference genome. In brief, we first sorted and fixed mate information. Then, we identified which reads came from the

same source molecule with fgbio's GroupReadsByUmi tool. GroupReadsByUmi was ran with the adjacency strategy (it implements the directed adjacency graph method introduced by (36) to account for sequencing errors when searching for matching UMIs. Furthermore, we required a minimum mapping quality of 30. Here, reads were aggressively filtered, only high quality read pairs were taken forward, to prevent (i) the grouping of reads that were really from different source molecules; and (ii) the building of two groups from reads that were really from the same molecule. With reads grouped by UMIs, we combined each cluster of reads to generate consensus reads using the CallMolecularConsensusReads tool of fgbio. This step generated unmapped consensus reads in BAM format that were further filtered with FilterConsensusReads tool. Here, we required that a consensus read was supported by at least two reads. Finally, we used Samtools to obtain the consensus reads in FASTQ format and run a second step of alignment with BWA. Duplicate reads marking, here was not be performed given that each read represents a unique source molecule. After the two alignment steps, we collected descriptive statistics as well as coverage metrics of targeted regions, on BAM files, by using SAMtools and Bedtools (Version: 2.29.2), respectively.

Somatic variants were called, within our targeted exons, with the VarDict software (37). VarDict is a variant calling program for SNVs, Multi Nucleotide Variants (MNVs), small Insertions and Deletions (INDELs) and complex variants. We set a minimum allele frequency of 0.01 for somatic variant calling starting from the germline BAM and consensus BAM. We, then annotated VCF files with ANNOVAR (Version: 2019-10-24).

Following the VarDict developer suggestions, we filtered variant files as follows. In brief, we kept only variants marked as "Strong Somatic" or "Likely Somatic" with a calling P value ≤0.05 (Fisher's exact test based on read counts for variant and reference alleles in the normal and tumor samples). We also required a mean base quality of the variant in the tumor BAM > = 40. Then, we filtered out synonymous SNVs and variants with allele frequencies above 1% in European populations of 1,000 Genomes, ExAC and GnomAD. Finally, the lists of variants were filtered from those falling in ENCODE excludable regions of low mappability.

Somatic CNAs were called, within our targeted exons, with the R-Biocondutor package "CNVPanelizer" (Version: 1.14.0) installed on R (Version: 3.5.0). The tool allows reliable CNA detection in targeted sequencing applications. Its approach uses a non-parametric bootstrap subsampling of the available reference samples to estimate the distribution of read counts from targeted sequencing. We used this tool with 10,000 replicates for the bootstrap and a significance level of 0.05.

All the somatic variants that passed the filtering steps, were visually inspected by using the Integrated Genome Viewer (IGV) software.

RESULTS

Patients Characteristics

Nine patients with Stage 4 and two with Stage 2 neuroblastoma were recruited as described in *Methods* (**Supplementary Table 1**). The average age at diagnosis was 36.45 months (range: 7–96). At diagnosis, three patients had *MYCN* amplification and one patient out of four tested, showed 1p36 deletion.

Sequencing and Mapping Yield

We designed a targeted sequencing panel to cover neuroblastoma cancer driver genes and included genes with proven clinical value according to COSMIC database. The full list of neuroblastoma driver genes selected for sequencing is reported in **Supplementary Table 2**.

We have developed an experimental and bioinformatic pipeline to enable next generation sequencing and detection of somatic mutations in cfDNA. Our aim was to develop a method combining accurate sequencing technologies with rare allele amplification strategies, which could potentially be used for personalized medicine at the point of care.

The sequencing of DNA from leukocytes returned, on average 61,876 millions, of high-quality reads per sample that were mapped versus the hg19 human reference genome with a mapping rate of about 95.5%.

cfDNA sequencing yielded, on average, 45,915,594 raw reads per sample with a mean read length of 140 bp (**Supplementary Figure 1A**). Sequence duplication level was of 35.1%, on average (**Supplementary Figure 1B**). The Q20 and Q30 metrics (the percent of bases with phred-scaled quality scores greater than 20 or 30) were 89.2 and 82.1%, respectively (**Supplementary Figure 1C**). With the first mapping step, the mean depth of coverage of our target regions was about 16,752×.

After the calling of consensus reads, by exploiting the presence of molecular tags (UMIs), we obtained an average of 1,078,823 high-quality reads coming from distinct cfDNA molecules (**Supplementary Figure 1D**). Sequence duplication level was about the 0.15%, on average (**Supplementary Figure 1E**). The Q30 was 99% (**Supplementary Figure 1F**).

By aligning consensus reads, we obtained a mean mapping rate of 99.9%. The mean read depth of the targeted bases was 462.6× (**Supplementary Figure 2A**). Overall, the 81.8 and the 62% of the target bases was covered by at least 50 and 100 reads, respectively (**Supplementary Figure 2B**). The fraction of target bases covered by consensus reads was, on average, 99.3%. Indeed, only 10 exons out of 1,702 (with mean length of 112 bp range: 9–532) were not covered by consensus reads. Globally, we obtained an adequate coverage of the target genes (**Supplementary Figure 2B–C**) to perform a reliable variant calling.

Landscape of Somatic Variants Identified in ctDNA

On average, we called about 1,268 raw variants per sample. Of these, roughly the 19.3% were somatic mutations (**Figure 1A**). The mean read depth and base quality of raw somatic variant calls were 181.6 and 44.8, respectively (**Figure 1B**).

The bioinformatic analytic process was established to filter out all germline events, based on the comparison of cfDNA with the leukocyte's DNA, and retaining only somatic alterations for further analysis. With the strong filtering described in *Methods*, we obtained a total of 101 somatic mutations, reported in **Supplementary Table 3**, (mean per sample: 9.2; range: 1–41)

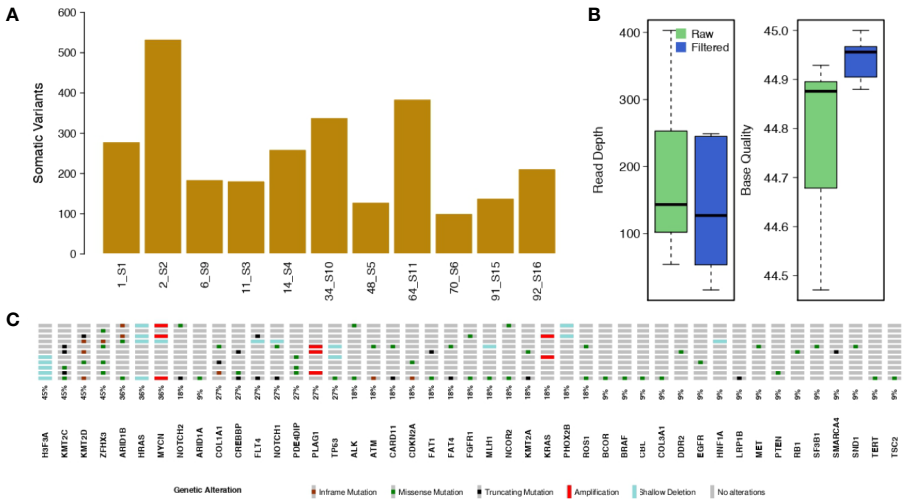


FIGURE 1 | Somatic single nucleotide variants and copy number alterations. **(A)** The bar plot shows the number of raw somatic variant calls. **(B)** Box plots showing the mean Read Depth and the mean Base Quality for raw variant calls and for the filtered somatic variants. **(C)** Oncoprint reporting the somatically altered genes along with their mutation frequency.

with allele frequencies ranging from 1.3 to 100% (mean: 12.4%; median: 6.9%). The filtered variants showed a mean read depth and base quality of 285.4 and 44.8, respectively (**Figure 1B**). Among the filtered somatic mutations, 74 were missense (73.3%), 17 were truncating (16.8%; including nine frameshifts and eight stop codon gains) and 10 (9.0%) non-frameshift INDELs (**Figure 1C**).

In order to identify variants with high pathogenic effects, we focused on those having a CADD (38) score above 20. We obtained a total of 51 (50.5%) filtered somatic mutations, reported in **Supplementary Table 3**, (mean per sample: 5.6, range: 1–24). Nine out of 11 cases (81.8%) harboured at least one pathogenic mutation. Among the filtered somatic mutations, 44 were missense (86.3%) and seven were stop codon gains (13.7%). Of these highly pathogenic mutations, some recurrently involved *KMT2C* (five cases), *NOTCH1/2* (four cases), *CREBBP* (three cases), *ARID1A/B* (three cases), *ALK* (two cases), *FGFR1* (two cases), *FAT4* (two cases) and *CARD11* (two cases), as depicted in **Figure 2A** and **Supplementary Figure 3**. For these mutations the median variant allele frequency was 6.2% (mean: 10.3%; range:

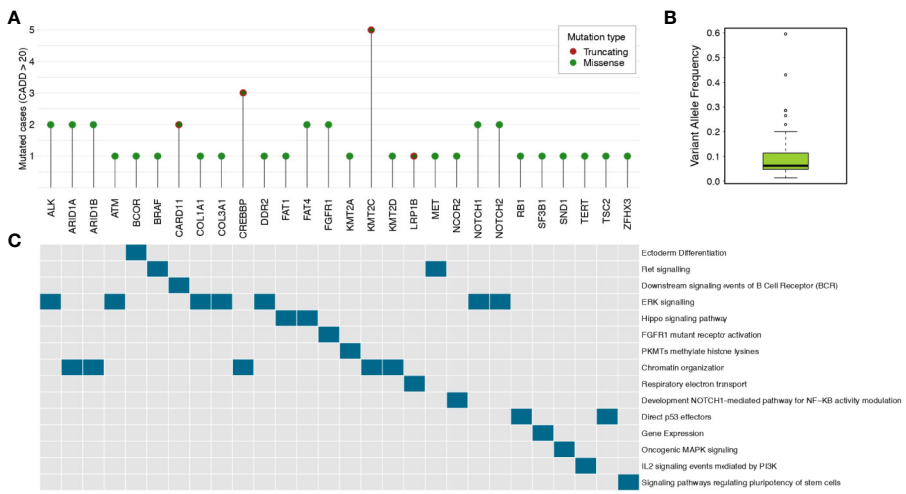


FIGURE 2 | Highly pathogenic somatic mutations. **(A)** The lollipop plot shows the number of cases for genes having somatic mutations with pathogenicity scores higher than 20 as determined by CADD tool. **(B)** The box plot reports the median variant allele frequency of the selected somatic mutations. **(C)** The data matrix shows the pathways in which the mutated genes are involved (pathcards.genecards.org).

1.3–59.6%) (**Figure 2B**). ERK signalling and Chromatin organization were the mainly altered pathways (**Figure 2C**).

As shown in **Supplementary Table 3**, eight out of 11 patients harboured somatic variants in genes that have been previously observed in solid cancers in at least one case (as reported in the COSMIC database). Interestingly, we found the hotspot mutation F1174L in *ALK*, frequently observed in neuroblastoma tumors.

Landscape of Somatic Copy Number Alterations Identified in ctDNA

We predicted CNAs by comparing ctDNA with its control DNA. Significant CNAs are reported in **Figure 1C** and **Supplementary Table 4**. Individual level Copy Number profiles for all the genes in our panel are in **Supplementary Figures 4–13**. We knew, by FISH assay, that three patients had the amplification of *MYCN* locus. For all of these samples we were able to confirm their *MYCN* amplification status. Furthermore, we found the heterozygous deletion of *H3F3A* in five samples. *H3F3A* mutants are drivers of tumorigenesis in pediatric cancers (39), and *H3F3A* is also a susceptibility gene in pheochromocytomas and paragangliomas (40). We also found deletions of *HRAS* in four samples and amplification of *KRAS* in two samples. These data should be treated with caution as *HRAS* is located at 11q5.5 that is a region of allelic imbalance and it can be difficult to distinguish between deletion and amplification. *KRAS* amplification has been reported in several solid cancers and is associated to lack of sensitivity to MAPK inhibitors (41). In two out of eleven samples we detected deletions of *PHOX2B* and *TP53* and amplifications of *PLAG1*. *PHOX2B* germline mutations predispose to neuroblastoma (42) and somatic variants can also occur in sporadic neuroblastoma. Although somatic mutations of *TP53* or other pathway members are rare in primary neuroblastomas obtained at diagnosis (20), *TP53* inactivation was observed in 50% of relapsed neuroblastoma (43). *PLAG1* amplifications or rearrangements have been identified in several tumors (44, 45).

DISCUSSION

Cancer is the leading cause of disease-related death in children. Two thirds of all survivors have late effects related to inadequate therapeutic choices or adverse health-related outcomes. Therefore, precision medicine in pediatric oncology is required to further improve outcomes and decrease toxicity. As the field of precision oncology grows, the application of ctDNA sequencing could complement or substitute the analysis of tissue biopsies and provide the means to explore the heterogeneity and the evolution of tumors. The first randomized controlled trials for pediatric cancers are currently ongoing. They will elucidate the impact of LB and ctDNA analysis on clinical decision making (NGSkids: NCT02546453 and MICCHADO: NCT03496402).

Most retrospective studies, conducted on small sample cohorts, have illustrated the great potential of LB approach in neuroblastoma. In this study, despite the limited number of samples, we highlight the utility of LB approach to facilitate the detection of druggable genes/mutations in ctDNA of neuroblastoma stage 4 patients. This could accelerate the therapeutic choice or may provide clinicians useful disease biomarkers at diagnosis.

We identified nine out of eleven (81.8%) patients who carried at least one pathogenic variation. The high prevalence of pathogenic variants in our cohort demonstrated the utility of a targeted high-throughput sequencing analysis, which excels in terms of the breadth of disease coverage when compared to single-gene tests. Furthermore, the detection of tumor-specific mutations in diagnostic ctDNA samples confirmed the potentiality of our panel to identify druggable mutations. This non-invasive method may address patients towards a personalized therapy. Particularly, we identified ALK F1174L hot spot mutation that was not previously evaluated in tissue biopsies at diagnosis. The presence of *ALK* point mutations, occurring in 8–10% of sporadic neuroblastoma, serve as biomarker of therapeutic sensitivity to small-molecule kinase inhibitors that are currently undergoing clinical assessment in phase I and II trials.

Furthermore, detecting mutations in key cancer genes, that are not actionable per se, could also benefit patient's management by improving diagnosis, prognosis and treatment strategies setting (see (10, 12, 15, 45, 46)). In this context, we observed mutations in *KMT2C* (5/11 cases), in *CREBBP* (3/11 cases) and in ARID1A/B (3/11 cases). All of these genes take part in chromatin remodelling complexes and are frequently mutated in cancer. *KMT2C* and *KMT2D* histone lysine methyltransferases are among the most frequently mutated genes across a variety of cancer types (46). It has been recently reported that *KMT2C* is involved in DNA repair and genomic instability. This, opens up the possibility that *KMT2C*-associated cancers may be targeted by PARP1/2 inhibitors (47). Furthermore, loss of *KMT2D* results in hyperactivation of RAS/MAPK pathways suggesting that these tumors may be treated with MAPK pathway inhibitors (48). *CREBBP* mutations often result in loss of tumor-suppressive functions and are difficult to target therapeutically. In these patients, the use of HDAC inhibitors may serve as an additional therapeutic option (49). Moreover, tumors with *ARID1A* mutations are sensitive to the treatment with EZH2 methyltransferase inhibitors (50). Interestingly, we observed *FGFR1* mutations in two out of eleven. *FGFR1* gene was found mutated in neuroblastoma both at diagnosis and at relapse (21, 51, 52). Targeting of FGFR1 signalling is currently used in adult cancers and may represent an interesting application not yet explored in neuroblastoma.

In our study, we observed a wide range of Minor Allele Frequencies (MAFs) in the lists of somatic mutations for a given sample. This can be mainly explained by clonal mutational events (29). In the setting of treatment strategies, also low MAF somatic mutations should be accounted to avoid the spread of chemo-resistant clones. Indeed, given its potential, we may suggest that our targeted sequencing panel for ctDNA analysis could be also used to monitor treatment responsiveness.

MYCN amplification and 1p36 deletion are genetic alterations routinely detected in diagnostic tissues and serve as important

prognostic biomarkers associated with poor prognosis. We confirmed *MYCN* amplification all of the patients. We could not be able to detect 1p36 deletion in one patient because the panel is not comprehensive of genes located in this region. Given the targeted nature of our sequencing panel, we recognize that we had limited capability in detecting large chromosomal aberrations. However, our choice allowed us to detect exon-level CNAs and drastically reduced both, sequencing costs and analysis time.

Many variables do affect ctDNA concentration into the bloodstream. These include tumor size and stage, metastasis, inflammation and therapy status, among the others. In children, additional variables such as age and tumor type should be considered. Despite neuroblastoma shows higher levels of ctDNA than other pediatric cancers (16), we are aware that highly standardized protocols for LB will become necessary to ensure reliable and reproducible results in routine clinical care.

In our study, we explored the usefulness of sequencing ctDNA by the means of a targeted panel of genes involved in neuroblastoma. By allowing the detection of somatic mutations (including low MAF and druggable mutations) and somatic CNAs, this approach could improve the clinical management of pediatric cancers in a cost-effective manner. Moreover, it represents a valuable alternative tool to avoid invasive tissue biopsies.

One limitation of this study is the relatively low number of analyzed tumors even if our set of samples (n = 11) derived from a careful selection among 20 tumors. Additional ctDNA sequencing studies using larger samples sizes are needed to validate the mutation frequencies we observed.

Our aim for the next future is to implement our custom NGS panel for ctDNA analysis to improve patient stratification at diagnosis. This may also lead to better targeted and timely therapies and reduce ineffective/inappropriate choices. Furthermore, the ctDNA analysis could be useful to monitor patients for early signs of relapse or for early diagnosis in families at high risk of developing neuroblastoma. Of course, well-designed and large-scale validation studies integrated in multicenter trials will be necessary to further delineate the clinical utility and validity of the proposed gene panel for ctDNA analysis.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. This data can be found here: https://www.ncbi.nlm.nih.gov/sra/?term=PRJNA672255.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by University of Naples Federico II. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

FC designed research, performed sample preparation, interpreted the data and drafted the manuscript. VAL performed bioinformatic analysis, interpreted the data and drafted the manuscript. SV contributed with patient samples and clinical information. MC and AI formulated the strategy and supervised the research. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by grants from Associazione Italiana per la Ricerca sul Cancro (Grant no. 19255 to MC and Grant no 20757 to AI); Fondazione Italiana per la Lotta al Neuroblastoma (to MC); Associazione Oncologia Pediatrica e Neuroblastoma (to MC) and Fondazione Umberto Veronesi (to FC); Regione Campania "SATIN" grant 2018-2020 (to MC).

ACKNOWLEDGMENTS

The authors want to thank the IRCCS Istituto Giannina Gaslini, Via G. Gaslini, 5, Genova, Italy; and the Ospedale Pediatrico Santobono-Pausilipon, Via Posillipo, 226, Naples, Italy; for providing us with specimens.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2020.596191/full#supplementary-material

SUPPLEMENTARY FIGURE 1 | Quality reports of sequencing and consensus reads. **(A)** | Total number of sequencing reads. **(B)** | Level (%) of duplicated sequencing reads as result of PCR cycles. **(C)** | Percentage of bases (within sequencing reads) with quality scores above 30. **(D)** | Total number of consensus reads. **(E)** | Level (%) of duplicated consensus reads. **(F)** | Percentage of bases (within consensus reads) with quality scores above 30.

SUPPLEMENTARY FIGURE 2 | Results of consensus reads mapping. **(A)** | Bar plot reporting the Mean Read Depth of target regions across samples. Grey, dashed line shows the overall Mean Read Depth. **(B)** | Percentage of regions covered at specific read depths. **(C)** | The box plots report the Mean Read Depth of target genes. *MYCN*, in panel **(E)**, is reported separately given the different scale of read depths.

SUPPLEMENTARY FIGURE 3 | Recurrently mutated genes. The lollipop plots show the protein-level localization of amino-acid changes in genes mutated in at least two samples.

SUPPLEMENTARY FIGURE 4-13 | Per sample Copy Number profiles. The plots show the Log_2Ratios of the comparisons of the ctDNA (as consensus reads) with the corresponding germline DNA at gene level. Sample names are reported in the top left corner of each figure. Significant CNAs are marked by a red dot.

REFERENCES

1. Crowley E, Di Nicolantonio F, Loupakis F, Bardelli A. Liquid biopsy: monitoring cancer-genetics in the blood. *Nat Rev Clin Oncol* (2013) 10 (8):472–84.
2. Butler TM, Spellman PT, Gray J. Circulating-tumor DNA as an early detection and diagnostic tool. *Curr Opin Genet Dev* (2017) 42:14–21.
3. Esposito A, Criscitiello C, Locatelli M, Milano M, Curigliano G. Liquid biopsies for solid tumors: Understanding tumor heterogeneity and real time monitoring of early resistance to targeted therapies. *Pharmacol Ther* (2016) 157:120–4.
4. Parikh AR, Leshchiner I, Elagina L, Goyal L, Levovitz C, Siravegna G, et al. Liquid versus tissue biopsy for detecting acquired resistance and tumor heterogeneity in gastrointestinal cancers. *Nat Med* (2019) 25(9):1415–21.
5. Garcia-Foncillas J, Alba E, Aranda E, Diaz-Rubio E, Lopez-Lopez R, Tabernero J, et al. Incorporating BEAMing technology as a liquid biopsy into clinical practice for the management of colorectal cancer patients: an expert taskforce review. *Ann Oncol* (2017) 28(12):2943–9.
6. Strickler JH, Loree JM, Ahronian LG, Parikh AR, Niedzwiecki D, Pereira AAL, et al. Genomic Landscape of Cell-Free DNA in Patients with Colorectal Cancer. *Cancer Discov* (2018) 8(2):164–73.
7. Zill OA, Banks KC, Fairclough SR, Mortimer SA, Vowles JV, Mokhtari R, et al. The Landscape of Actionable Genomic Alterations in Cell-Free Circulating Tumor DNA from 21,807 Advanced Cancer Patients. *Clin Cancer Res* (2018) 24(15):3528–38.
8. Lebofsky R, Decraene C, Bernard V, Kamal M, Blin A, Leroy Q, et al. Circulating tumor DNA as a non-invasive substitute to metastasis biopsy for tumor genotyping and personalized medicine in a prospective trial across all tumor types. *Mol Oncol* (2015) 9(4):783–90.
9. Tie J, Wang Y, Tomasetti C, Li L, Springer S, Kinde I, et al. Circulating tumor DNA analysis detects minimal residual disease and predicts recurrence in patients with stage II colon cancer. *Sci Trans Med* (2016) 8(346):346ra92.
10. Lanman RB, Mortimer SA, Zill OA, Sebisanovic D, Lopez R, Blau S, et al. Analytical and Clinical Validation of a Digital Sequencing Panel for Quantitative, Highly Accurate Evaluation of Cell-Free Circulating Tumor DNA. *PLoS One* (2015) 10(10):e0140712.
11. Chiang AC, Fernandes AW, Pavilack M, Wu JW, Laliberte F, Duh MS, et al. EGFR mutation testing and treatment decisions in patients progressing on first- or second-generation epidermal growth factor receptor tyrosine kinase inhibitors. *BMC Cancer* (2020) 20(1):356.
12. Said R, Guibert N, Oxnard GR, Tsimberidou AM. Circulating tumor DNA analysis in the era of precision oncology. *Oncotarget* (2020) 11(2):188–211.
13. Grobner SN, Worst BC, Weischenfeldt J, Buchhalter I, Kleinheinz K, Rudneva VA, et al. The landscape of genomic alterations across childhood cancers. *Nature* (2018) 555(7696):321–7.
14. Ma X, Liu Y, Liu Y, Alexandrov LB, Edmonson MN, Gawad C, et al. Pan-cancer genome and transcriptome analyses of 1,699 paediatric leukaemias and solid tumours. *Nature* (2018) 555(7696):371–6.
15. Van Paemel R, Vlug R, De Preter K, Van Roy N, Speleman F, Willems L, et al. The pitfalls and promise of liquid biopsies for diagnosing and treating solid tumors in children: a review. *Eur J Pediatr* (2020) 179(2):191–202.
16. Abbou SD, Shulman DS, DuBois SG, Crompton BD. Assessment of circulating tumor DNA in pediatric solid tumors: The promise of liquid biopsies. *Pediatr Blood Cancer* (2019) 66(5):e27595.
17. Cohn SL, Pearson AD, London WB, Monclair T, Ambros PF, Brodeur GM, et al. The International Neuroblastoma Risk Group (INRG) classification system: an INRG Task Force report. *J Clin Oncol* (2009) 27(2):289–97.
18. Aygun N. Biological and Genetic Features of Neuroblastoma and Their Clinical Importance. *Curr Pediatr Rev* (2018) 14(2):73–90.
19. Lasorsa VA, Cimmino F, Ognibene M, Mazzocco K, Erminio G, Morini M, et al. 19p loss is significantly enriched in older age neuroblastoma patients and correlates with poor prognosis. *NPJ Genom Med* (2020) 5:18.
20. Pugh TJ, Morozova O, Attiyeh EF, Asgharzadeh S, Wei JS, Auclair D, et al. The genetic landscape of high-risk neuroblastoma. *Nat Genet* (2013) 45(3):279–84.
21. Lasorsa VA, Formicola D, Pignataro P, Cimmino F, Calabrese FM, Mora J, et al. Exome and deep sequencing of clinically aggressive neuroblastoma reveal somatic mutations that affect key pathways involved in cancer progression. *Oncotarget* (2016) 7(16):21840–52.
22. Calabrese FM, Clima R, Pignataro P, Lasorsa VA, Hogarty MD, Castellano A, et al. A comprehensive characterization of rare mitochondrial DNA variants in neuroblastoma. *Oncotarget* (2016) 7(31):49246–58.
23. Capasso M, Lasorsa VA, Cimmino F, Avitabile M, Cantalupo S, Montella A, et al. Transcription Factors Involved in Tumorigenesis Are Over-Represented in Mutated Active DNA-Binding Sites in Neuroblastoma. *Cancer Res* (2020) 80(3):382–93.
24. Tonini GP, Capasso M. Genetic predisposition and chromosome instability in neuroblastoma. *Cancer Metastasis Rev* (2020) 39(1):275–85.
25. Cimmino F, Avitabile M, Diskin SJ, Vaksman Z, Pignataro P, Formicola D, et al. Fine mapping of 2q35 high-risk neuroblastoma locus reveals independent functional risk variants and suggests full-length BARD1 as tumor-suppressor. *Int J Cancer* (2018) 143(11):2828–37.
26. McDaniel LD, Conkrite KL, Chang X, Capasso M, Vaksman Z, Oldridge DA, et al. Common variants upstream of MLF1 at 3q25 and within CPZ at 4p16 associated with neuroblastoma. *PLoS Genet* (2017) 13(5):e1006787.
27. Chicard M, Boyault S, Colmet Daage L, Richer W, Gentien D, Pierron G, et al. Genomic Copy Number Profiling Using Circulating Free Tumor DNA Highlights Heterogeneity in Neuroblastoma. *Clin Cancer Res* (2016) 22 (22):5564–73.
28. Bettegowda C, Sausen M, Leary RJ, Kinde I, Wang Y, Agrawal N, et al. Detection of circulating tumor DNA in early- and late-stage human malignancies. *Sci Trans Med* (2014) 6(224):224ra24.
29. Chicard M, Colmet-Daage L, Clement N, Danzon A, Bohec M, Bernard V, et al. Whole-Exome Sequencing of Cell-Free DNA Reveals Temporo-spatial Heterogeneity and Identifies Treatment-Resistant Clones in Neuroblastoma. *Clin Cancer Res* (2018) 24(4):939–49.
30. Lodrini M, Sprussel A, Astrahantseff K, Tiburtius D, Konschak R, Lode HN, et al. Using droplet digital PCR to analyze MYCN and ALK copy number in plasma from patients with neuroblastoma. *Oncotarget* (2017) 8(49):85234–51.
31. Siravegna G, Mussolin B, Buscarino M, Corti G, Cassingena A, Crisafulli G, et al. Clonal evolution and resistance to EGFR blockade in the blood of colorectal cancer patients. *Nat Med* (2015) 21(7):795–801.
32. Van Roy N, Van Der Linden M, Menten B, Dheedene A, Vandeputte C, Van Dorpe J, et al. Shallow Whole Genome Sequencing on Circulating Cell-Free DNA Allows Reliable Noninvasive Copy-Number Profiling in Neuroblastoma Patients. *Clin Cancer Res* (2017) 23(20):6305–14.
33. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* (2010) 26(5):589–95.
34. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* (2009) 25 (16):2078–9.
35. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* (2018) 34(17):i884–i90.
36. Smith T, Heger A, Sudbery I. UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. *Genome Res* (2017) 27(3):491–9.
37. Lai Z, Markovets A, Ahdesmaki M, Chapman B, Hofmann O, McEwen R, et al. VarDict: a novel and versatile variant caller for next-generation sequencing in cancer research. *Nucleic Acids Res* (2016) 44(11):e108.
38. Rentzsch P, Witten D, Cooper GM, Shendure J, Kircher M. CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res* (2019) 47(D1):D886–D94.
39. Lowe BR, Maxham LA, Hamey JJ, Wilkins MR, Partridge JF. Histone H3 Mutations: An Updated View of Their Role in Chromatin Deregulation and Cancer. *Cancers (Basel)* (2019) 11(5):660.
40. Toledo RA, Qin Y, Cheng ZM, Gao Q, Iwata S, Silva GM, et al. Recurrent Mutations of Chromatin-Remodeling Genes and Kinase Receptors in Pheochromocytomas and Paragangliomas. *Clin Cancer Res* (2016) 22 (9):2301–10.
41. Mai TT, Lito P. A treatment strategy for KRAS-driven tumors. *Nat Med* (2018) 24(7):902–4.
42. Bourdeaut F, Trochet D, Janoueix-Lerosey I, Ribeiro A, Deville A, Coz C, et al. Germline mutations of the paired-like homeobox 2B (PHOX2B) gene in neuroblastoma. *Cancer Lett* (2005) 228(1-2):51–8.
43. Tweddle DA, Pearson AD, Haber M, Norris MD, Xue C, Flemming C, et al. The p53 pathway and its inactivation in neuroblastoma. *Cancer Lett* (2003) 197(1-2):93–8.

44. Van Dyck F, Declercq J, Braem CV, Van de Ven WJ. PLAG1, the prototype of the PLAG gene family: versatility in tumour development (review). *Int J Oncol* (2007) 30(4):765–74.
45. Jo VY, Fletcher CD. Myoepithelial neoplasms of soft tissue: an updated review of the clinicopathologic, immunophenotypic, and genetic features. *Head Neck Pathol* (2015) 9(1):32–8.
46. Fagan RJ, Dingwall AK. COMPASS Ascending: Emerging clues regarding the roles of MLL3/KMT2C and MLL2/KMT2D proteins in cancer. *Cancer Lett* (2019) 458:56–65.
47. Rampias T, Karagiannis D, Avgeris M, Polyzos A, Kokkalis A, Kanaki Z, et al. The lysine-specific methyltransferase KMT2C/MLL3 regulates DNA repair components in cancer. *EMBO Rep* (2019) 20(3):e46821.
48. Tsai IC, McKnight K, McKinstry SU, Maynard AT, Tan PL, Golzio C, et al. Small molecule inhibition of RAS/MAPK signaling ameliorates developmental pathologies of Kabuki Syndrome. *Sci Rep* (2018) 8(1):10779.
49. Hellwig M, Merk DJ, Lutz B, Schuller U. Preferential sensitivity to HDAC inhibitors in tumors with CREBBP mutation. *Cancer Gene Ther* (2020) 27 (5):294–300.
50. Bitler BG, Aird KM, Garipov A, Li H, Amatangelo M, Kossenkov AV, et al. Synthetic lethality by targeting EZH2 methyltransferase activity in ARID1A-mutated cancers. *Nat Med* (2015) 21(3):231–8.
51. Eleveld TF, Oldridge DA, Bernard V, Koster J, Colmet Daage L, Diskin SJ, et al. Relapsed neuroblastomas show frequent RAS-MAPK pathway mutations. *Nat Genet* (2015) 47(8):864–71.
52. Padovan-Merhar OM, Raman P, Ostrovnaya I, Kalletla K, Rubnitz KR, Sanford EM, et al. Enrichment of Targetable Mutations in the Relapsed Neuroblastoma Genome. *PLoS Genet* (2016) 12(12):e1006501.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer AC declared a past co-authorship with several of the authors SV, VAL, MC to the handling editor.

frontiers in Oncology

ORIGINAL RESEARCH
published: 26 February 2021
doi: 10.3389/fonc.2021.608433

A Survival-Related Competitive Endogenous RNA Network of Prognostic lncRNAs, miRNAs, and mRNAs in Wilms Tumor

*HengChen Liu†, MingZhao Zhang†, ManYu Shi, TingTing Zhang, ZeNan Zhang, QingBo Cui, ShuLong Yang and ZhaoZhu Li**

Department of Pediatric Surgery, The Second Hospital Affiliated to Harbin Medical University, Harbin, China

Wilms tumor (WT) commonly occurs in infants and children. We evaluated clinical factors and the expression of multiple RNAs in WT samples in the TARGET database. Eight long non-coding RNAs (lncRNAs; AC079310.1, MYCNOS, LINC00271, AL445228.3, Z84485.1, AC091180.5, AP002518.2, and AC007879.3), two microRNAs (miRNAs; hsa-mir-152 andhsa-mir-181a), and nine messenger RNAs (mRNAs; TCTEX1D4, RNF133, VRK1, CCNE1, HEY1, C10orf71, SPRY1, SPAG11A, and MAGEB18) were screened from differentially expressed RNAs and used to construct predictive survival models. These models showed good prognostic ability and were highly correlated with tumor stage and histological classification. Additionally, survival-related ceRNA network was constructed using 35 RNAs (15 lncRNAs, eight miRNAs, and 12 mRNAs). KEGG pathway analysis suggested the "Wnt signaling pathway" and "Cellular senescence" as the main pathways. In conclusion, we established a multinomial predictive survival model and a survival-related ceRNA network, which provide new potential biomarkers that may improve the prognosis and treatment of WT patients.

Keywords: Wilms tumor, target, lncRNA, miRNA, mRNA, competing endogenous RNA/ceRNA

OPEN ACCESS

Edited by:
Jinhong Zhu,
Harbin Medical University Cancer Hospital, China

Reviewed by:
Joseph Louis Lasky,
Cure 4 The Kids, United States
Jhon A. Guerra,
HIMA San Pablo Oncologic, United States

***Correspondence:**
ZhaoZhu Li
zhaozhu247@163.com

†*These authors share first authorship*

Specialty section:
This article was submitted to Pediatric Oncology, a section of the journal Frontiers in Oncology

Received: *20 September 2020*
Accepted: *25 January 2021*
Published: *26 February 2021*

Citation:
Liu H, Zhang M, Shi M, Zhang T, Zhang Z, Cui Q, Yang S and Li Z (2021) A Survival-Related Competitive Endogenous RNA Network of Prognostic lncRNAs, miRNAs, and mRNAs in Wilms Tumor. Front. Oncol. 11:608433. doi: 10.3389/fonc.2021.608433

INTRODUCTION

Wilms tumor (WT), a renal malignancy originating from the metanephric blastema, is widespread in infants and children (1). WT accounts for 7% of all pediatric malignancies and occurs in one out of every 10,000 children (2). Fortunately, with the development of treatments, the survival rate of children with WT has increased by nearly 60% (3). The International Society of Pediatric Oncology stated that the combination of nephrectomy and chemotherapy significantly improved overall survival (OS) by more than 90% (4). However, 25% of children still have a poor prognosis based on the tumor stage (5). Therefore, clarifying the cellular process involved in WT development and providing prognostic biomarkers are important steps to improving the survival of patients.

In recent years, several studies have suggested non-coding RNAs as key molecules involved in tumorigenesis and tumor progression (6). microRNAs (miRNAs) are non-coding RNAs composed of 18–25 nucleotides that can negatively regulate gene expression (7). By contrast, long non-coding RNAs (lncRNAs) are more than 200 nucleotides in length and they adjust the biological behavior of tumors

through competing endogenous RNAs (ceRNAs) (8). It is suggested that there is a complex regulatory network among lncRNAs, miRNAs, and messenger RNAs (mRNAs). lncRNAs competitively inhibit the function of miRNAs by acting as a sponge, thus indirectly disrupting mRNA expression and ultimately affecting gene expression (9). In general, the instability of a ceRNA network may induce tumorigenesis (10, 11). Several studies have suggested that a ceRNA network can be used as a prognostic biomarker for WT, but did not actually describe the effect of a ceRNA network in WT (12, 13). Therefore, the establishment of a ceRNA network related to the survival of WT patients is of great significance for judging the prognosis of patients. By understanding the role of various RNAs in tumorigenesis, we can find potential targets to improve the prognosis of WT.

In this study, we explored the ability of multiple RNAs to prognosticate WT as a whole and established RNA models that can be used to predict survival. In addition, we established a survival-related ceRNA network and tried to understand its molecular mechanism through functional enrichment. Finally, we provided new potential biological biomarkers to improve the prognosis and treatment of WT.

MATERIALS AND METHODS

Data Collection and Processing

Clinical and RNA sequencing data of all patients were obtained from the Therapeutically Applicable Research to Generate Effective Treatments (TARGET) database, which can be downloaded from The Cancer Genome Atlas (TCGA) portal (https://portal.gdc.cancer.gov/; Data Release 25.0; release time: July 22, 2020). This study met the requirements for using TCGA database and did not require the approval of an ethics committee. We selected only the sequencing data from primary solid tumors for analysis. The mRNA and lncRNA sequencing data included 125 primary WT samples and six normal samples. The miRNA sequencing data included 127 primary WT samples and six normal samples. The clinical data of the 128 patients included in this study are shown in **Table 1**.

TABLE 1 | Corresponding Clinical Features of 128 Patients With Wilms Tumor.

Items	Patients, N = 128	
	N	%
Age		
<5	81	63.28125
≥5	47	36.71875
Gender		
Male	54	42.1875
Female	74	57.8125
Race		
White	95	74.21875
Non-White	33	25.78125
Tumor stage		
Stage I/II	66	51.5625
Stage III/IV	62	48.4375
Histologic classification		
FHWT	84	65.625
DAWT	44	34.375
Survival status		
Alive	76	59.375
Dead	52	40.625

Identification of Differentially Expressed RNAs

The edgeR package in the R 4.0.2 software was used to analyze differentially expressed lncRNAs (DElncRNAs), differentially expressed miRNAs (DEmiRNAs), and differentially expressed mRNAs (DEmRNAs) between the WT and normal samples. The cutoff value for differentially expressed RNAs (DERNAs) was | log2 fold-change (FC)| >1 and false discovery rate (FDR) <0.05. Visualization of DERNAs was performed using ggplot2 package in R software.

Survival Analysis

dentification of survival-associated RNAs was performed *via* univariate Cox regression analyses in the R software. Significantly correlated survival-associated RNAs ($P < 0.003$) were chosen for multivariate Cox regression analysis and to establish predictive survival models. Prognosis index (PI) = $(\text{expression}_{\text{gene1}} \times \beta_{\text{gene1}}) + (\text{expression}_{\text{gene2}} \times \beta_{\text{gene2}}) + \ldots + (\text{expression}_{\text{genen}} \times \beta_{\text{genen}})$. Patients were divided into two groups based on the median PI. Survival prognosis of the two groups was compared by Kaplan–Meier analysis. The receiver operating characteristic (ROC) curve for evaluating the predictive ability of the model was depicted through R software.

Protein–Protein Interaction Network Construction

The Search Tool for the Retrieval of Interacting Genes (STRING; http://string-db.org) database was used to obtain PPI data of the significant survival-associated mRNAs ($P < 0.003$). Establishment of the PPI network was performed using the Cytoscape software.

Construction of the ceRNA Network

Survival-associated RNAs were used to construct ceRNA networks. First, the potential miRNAs interacting with lncRNAs were screened according to the miRcode (http://www.mircode.org/) database. Targeted mRNAs were identified using the miRTarBase (http://mirtarbase.cuhk.edu.cn/), miRDB (http://www.mirdb.org/), and TargetScan (http://www.targetscan.org/) databases. Finally, establishment of the lncRNA-miRNA-mRNA interaction ceRNA network was performed using the Cytoscape software.

Functional Enrichment Analysis

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses of mRNAs in the ceRNA network were conducted using KOBAS 3.0 (http://kobas.cbi.pku.edu.cn/kobas3/). Visualization of the enrichment analyses was conducted using the R software.

Statistical Analysis

The correlation of RNAs with clinical characteristics was analyzed by rank sum test. Differences between survival curves were

analyzed by log-rank test. R 4.0.2, Cytoscape v3.7.2, and GraphPad Prism 8 were used for plotting. SPSS 24 was used for statistical analysis. $P < 0.05$ was considered statistically significant.

RESULTS

Identification of DElncRNAs, DEmiRNAs, and DEmRNAs

We downloaded RNA sequencing data of 128 WT patients from the database and screened multiple DERNAs separately. A total of 10,585 DERNAs were screened, including 3,219 DElncRNAs (1,664 upregulated and 1,555 downregulated), 236 DEmRNAs (153 upregulated and 83 downregulated), and 7,130 DEmRNAs (3,762 upregulated and 3,368 downregulated). Finally, we visualized multiple DERNAs by heat and volcano maps (**Figure S1**).

Identification of Survival-Associated RNAs in WT

The relationship between DERNAs and survival was assessed by univariate Cox regression analysis. RNAs with $P < 0.05$ were selected as survival-associated RNAs, yielding a total of 696 survival-associated RNAs (199 lncRNAs, 17 miRNAs, and 480 mRNAs). The top 15 survival-associated RNAs are shown in **Figure 1**. The gene networks of strongly correlated survival-associated mRNAs ($P < 0.003$) were constructed by STRING (**Figure 2**). The hub genes, including XAB2, SNRPA, PRPF19, and TP53, are shown in the PPI network.

Establishment of Predictive Survival Models

RNAs with strong correlation were selected by univariate Cox regression analysis ($P < 0.003$), and then multivariate Cox regression analysis was used to analyze survival-associated RNAs with strong correlation. Finally, a total of eight lncRNAs (AC079310.1, MYCNOS, LINC00271, AL445228.3, Z84485.1, AC091180.5, AP002518.2, and AC007879.3), two miRNAs (hsa-mir-152 and hsa-mir-181a), and nine mRNAs (TCTEX1D4, RNF133, VRK1, CCNE1, HEY1, C10orf71, SPRY1, SPAG11A, and MAGEB18) were identified. Subsequently, the predictive survival models were built. PI_{lncRNA} = (0.76777 × AC079310.1 expression) + (0.28668 × MYCNOS expression) + (0.53416 × LINC00271 expression) + (0.64448 × AL445228.3 expression) + (0.45112 × Z84485.1 expression) + (− −0.35751 × AC091180.5 expression) + (0.53728 × AP002518.2 expression) + (0.14176 × AC007879.3 expression). PI_{miRNA} = (−0.3952 × hsa-mir-152 expression) + (−0.2490 × hsa-mir-181a expression). PI_{mRNA} = (0.52018 × TCTEX1D4 expression) + (0.24841 × RNF133 expression) + (0.48661 × VRK1 expression) + (0.40846 × CCNE1 expression) + (−0.43723 × HEY1 expression) + (−0.17615 × C10orf71 expression) + (-0.32992 × SPRY1 expression) + (−0.25891 × SPAG11A expression) + (0.29755 × MAGEB18 expression).

PI values were calculated for each patient and divided into two groups. We found that among the three groups of RNAs, the low-risk group had better survival as determined by Kaplan–Meier analysis (**Figures 3A–C**). The ability of the models to predict 3-year survival was evaluated by drawing the ROC curve. The areas under the curves (AUCs) of the three groups were 0.818, 0.701, and 0.848, respectively (**Figures 3D–F**). These results suggest that the three groups of modules have great potential for predicting the clinical prognosis of WT. **Figure 4** shows the risk scores, survival status, and RNA expression profiles in each group.

The correlation of these RNAs with other clinical characteristics was assessed by rank sum test. Clinical characteristics included age (<5/≥5), gender (male/female), race (white/non-white), tumor stage (I–II/III–IV), and pathological classification (FHWT/DAWT). We found that the RNAs in the models were significantly correlated with tumor stage and histological classification (**Figure 5**). This implies that these RNAs can be used as potential indicators to judge the degree of WT tumor development. Next, we identified the relationship between clinical characteristics and OS. Multivariate Cox regression analysis suggested that tumor stage and risk level directly affected tumor prognosis (**Table 2**).

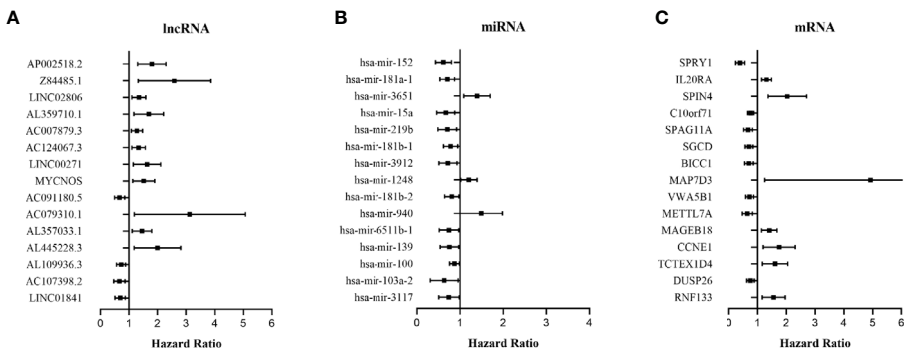


FIGURE 1 | Forest plots of the hazard ratios (HR) of the survival-associated RNAs in WT. **(A)** HR of top 15 survival-associated lncRNAs. **(B)** HR of top 15 survival-associated miRNAs. **(C)** HR of top 15 survival-associated mRNAs. HR <1 indicates the protective RNAs, and HR >1 indicates the risk RNAs.

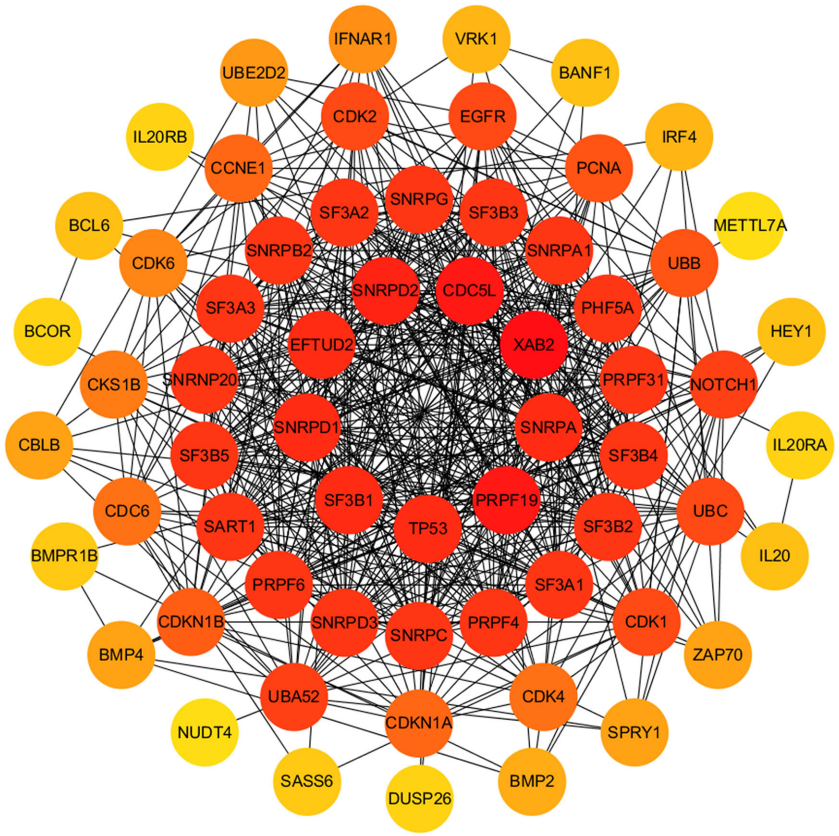


FIGURE 2 | PPI network of significant survival-associated mRNAs by Cytoscape. The brightness of the circle represents the degree of connection. The red circles are hub genes in the networks.

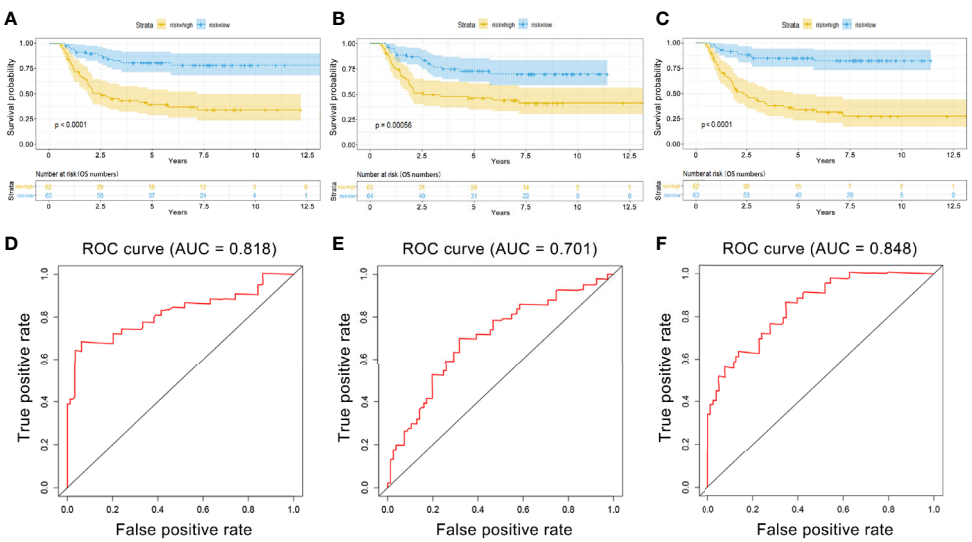


FIGURE 3 | Kaplan–Meier (K–M) and ROC curves for PI in WT patients. **(A)** K–M survival curves between high-risk and low-risk groups based on lncRNA model. **(B)** K–M survival curves between high-risk and low-risk groups based on miRNA model. **(C)** K–M survival curves between high-risk and low-risk groups based on mRNA model. **(D)** Time-dependent ROC curves analysis for survival prediction by PIlncRNA. **(E)** Time-dependent ROC curves analysis for survival prediction by PImiRNA. **(F)** Time-dependent ROC curves analysis for survival prediction by PImRNA.

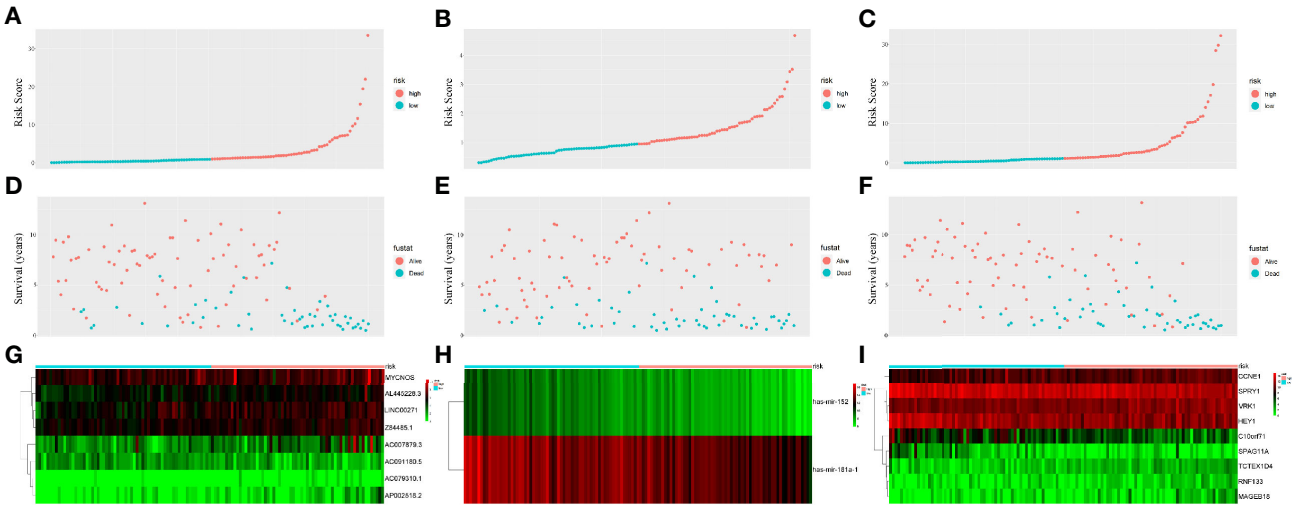


FIGURE 4 | Prognostic classifier analyses in distinguishing patients into low-risk and high-risk groups. **(A)** The lncRNA-based risk score distribution of each patient. **(B)** The miRNA-based risk score distribution of each patient. **(C)** The mRNA-based risk score distribution of each patient. **(D)** The lncRNA-based survival status of each patient in two groups. **(E)** The miRNA-based survival status of each patient in two groups. **(F)** The mRNA-based survival status of each patient in two groups. **(G)** Heatmap of lncRNAs expression between two groups. **(H)** Heatmap of miRNAs expression between two groups. **(I)** Heatmap of mRNAs expression between two groups.

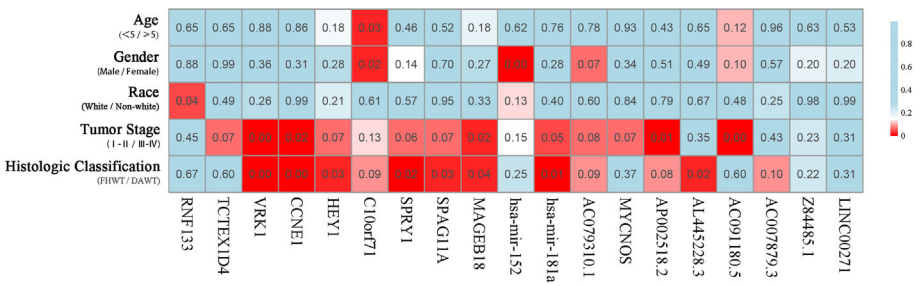


FIGURE 5 | Clinical significance of prognostic RNAs in WT. Correlations between prognostic RNAs and clinical factors. The numbers and brightness of each cell indicate the P-value.

Construction of a Survival-Related ceRNA Network in WT

Based on the survival-associated RNAs, 39 pairs of lncRNA–miRNA and 13 pairs of miRNA–mRNAs were detected from the databases. Subsequently, a ceRNA network containing 15 lncRNAs, eight miRNAs, and 12 mRNAs was constructed (**Figure 6**). Many of these RNAs have been extensively studied as cancer-related molecules, such as miR-181a, CCNE1, and WIF1. Next, we further studied the molecular function of mRNAs in the ceRNA network. A total of 99 functional enrichment terms (68 biological processes, 13 cellular components, and 18 molecular functions) from the GO analysis and 15 KEGG pathways were observed. The biological processes were mainly enriched in "cell surface receptor signaling pathway involved in cell-cell signaling"; the cellular components were mainly enriched in "nucleus"; and the molecular function was mainly enriched in "binding". The KEGG analysis suggested that the "Wnt signaling pathway" and "Cellular senescence" were the main pathways (**Figure 7**).

DISCUSSION

WT is a common pediatric tumor and poor prognosis is the main factor affecting long-term survival of patients (14). However, the exact molecular mechanism of WT is still unclear. The advancement in sequencing technology and the proposed ceRNA hypothesis provide a new perspective for the study of tumorigenesis and tumor progression (9). In the present study, we screened out survival-associated RNAs in the TARGET database and studied molecular events associated with WT. In particular, we proposed a novel survival-related ceRNA network that provides a new dimension for predicting the prognosis of WT patients.

TABLE 2 | Univariate and multivariate Cox regression analyses of overall survival.

Overall survival	Univariate analysis			Multivariate analysis		
	HR	95%CI	P-value	HR	95%CI	P-value
Age (<5/≥5)	0.704	0.386–1.284	0.252			
Gender (Male/Female)	0.591	0.343–1.019	0.058			
Race (White/Non-white)	1.122	0.607–2.071	0.714			
Tumor stage (I-II/III-IV)	0.999	1.792–5.840	0.998			
lncRNA cohort				3.504	1.882–6.524	<0.001
miRNA cohort				3.105	1.717–5.615	<0.001
mRNA cohort				2.454	1.307–4.608	0.005
Histologic classification (FHWT/DAWT)	1.196	0.714–2.209	0.529			
LncRNA signature (Low group/High group)	4.502	2.346–8.637	<0.001	4.822	2.472–9.403	<0.001
miRNA signature (Low group/High group)	2.673	1.495–4.777	0.001	2.498	1.394–4.474	0.002
mRNA signature (Low group/High group)	6.380	3.173–12.830	<0.001	4.883	2.381–10.011	<0.001

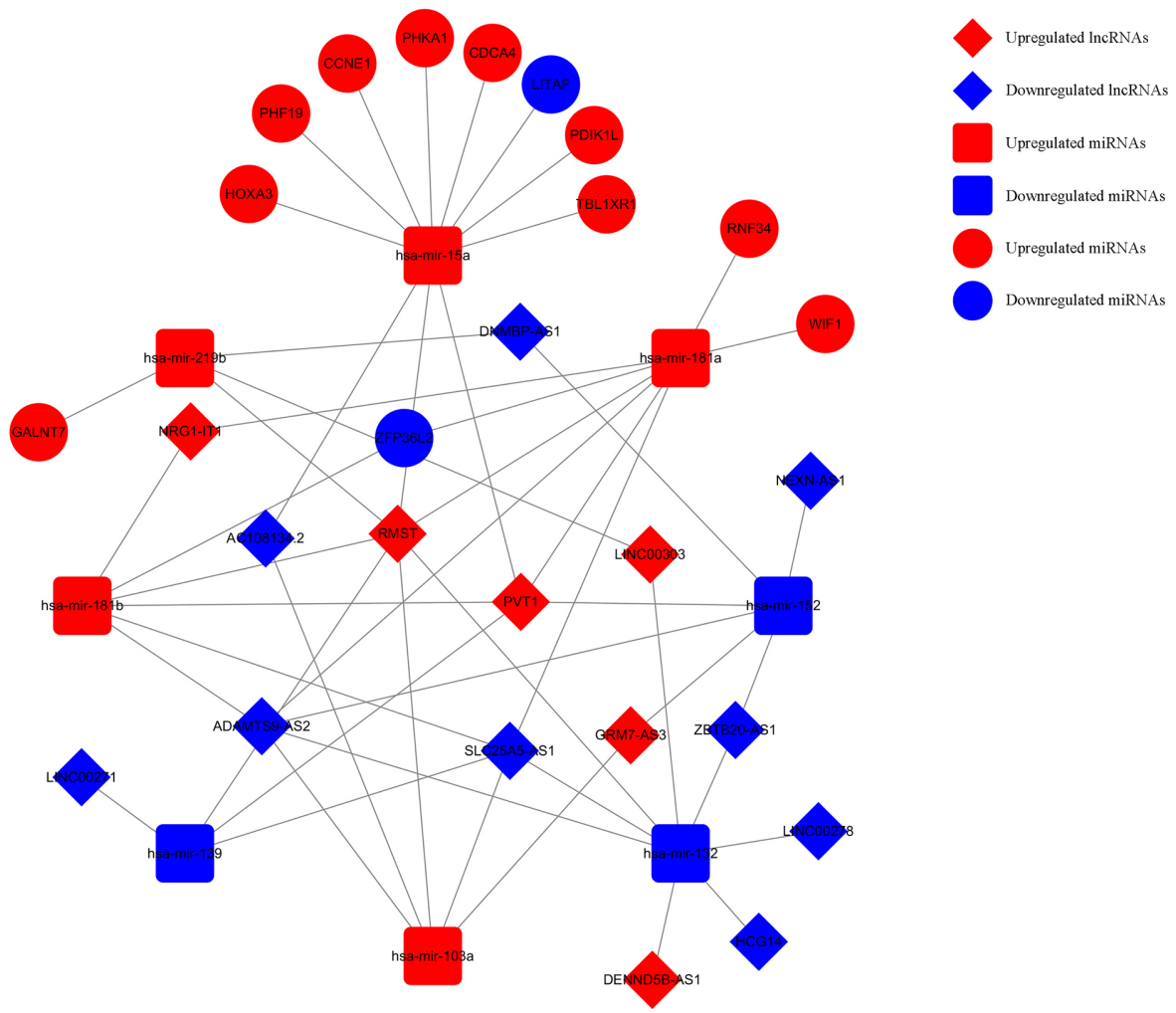


FIGURE 6 | Survival-related ceRNA network in WT. The red diamonds indicate upregulated lncRNAs, and blue diamonds indicate downregulated lncRNAs. The red rectangles indicate upregulated miRNAs, and blue rectangles indicate downregulated miRNAs. The red circles diamonds indicate upregulated mRNAs, and blue circles diamonds indicate downregulated mRNAs.

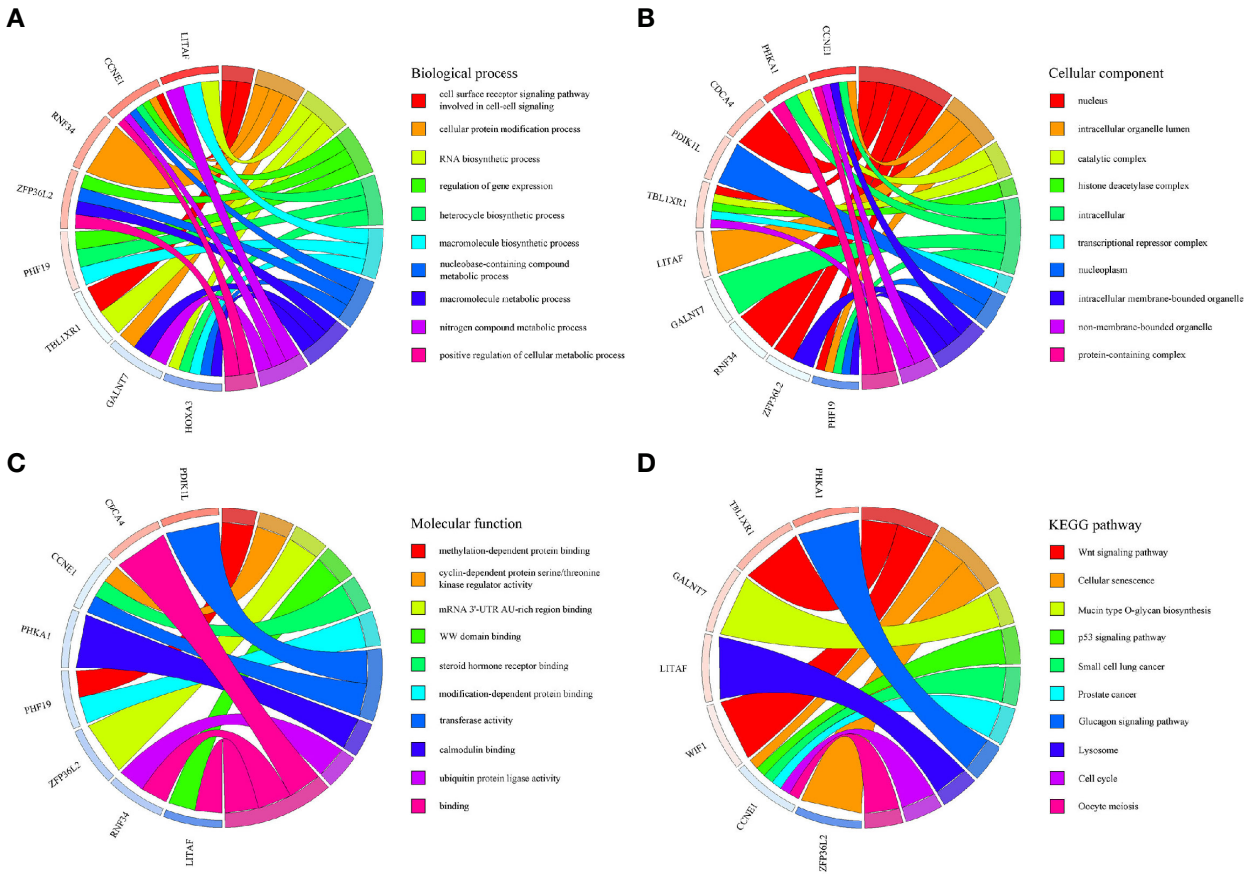


FIGURE 7 | GO and KEGG pathway enrichment analyses of genes included in the ceRNA network. **(A)** Biological process. **(B)** Cellular component. **(C)** Molecular function. **(D)** KEGG pathways.

Among all the DERNAs, we screened a total of 696 survival-related RNAs, including 199 lncRNAs, 14 miRNAs, and 480 mRNAs. mRNAs are key to the realization of molecular functions. PPI network analysis was carried out to identify the hub genes. We found that XAB2, SNRPA, PRPF19, and TP53 had a higher degree of connection among all genes. It is well-known that p53 is a tumor suppressor gene, and some studies have confirmed that TP53 gene mutation worsens the prognosis of WT (15). The other hub genes also affect the progress of other tumors (16, 17), suggesting that these hub genes may also affect the development of WT.

Most recent studies on WT have focused on the search for a single RNA that may influence prognosis. Zong et al. (18) found that miR-30d expression in WT tissues was significantly lower than that in normal tissues. Further animal experiments showed that miR-30d mimic significantly inhibited tumor growth. In addition, overexpression of SOX4 could reverse the effect of mir-30d on WT cells. Bao et al. (19) found that the expression of mirna-203a was low in WT tumor tissues, and that its expression level was directly related to the prognosis of WT. Knockout of miR-203a significantly enhanced the invasiveness of WT cells. Luciferase analysis confirmed that miR-203a targeted JAG1 to exert biological functions. Wang et al. (20) found that overexpression of miR-613 inhibited the G0/G1 phase transition of WT cells and hindered the expansion ability of WT cells. There is also a subset of studies evaluating the relationship between single RNA and OS. Tang et al. (12) predicted 32 DERNAs to be possibly associated with OS in WT patients through Kaplan-Meier analysis, and Zhang et al. (13) predicted 61 DERNAs to be possibly associated with OS by Kaplan-Meier analysis.

However, the development of WT is influenced by multifactorial causes, indicating that we should analyze prognostic markers at a broader level. We established a multinomial predictive survival model based on survival-associated RNAs by multivariate Cox analysis, which could suggest prognostic survival of WT patients. An AUC value >0.70 denotes excellent model performance. The AUC values of the three models were all greater than 0.7 in our study (0.818, 0.701, and 0.848). Gong et al. (21) established a 5-miRNA model for the prognosis of WT patients. The AUC value of their model was 0.767, which was higher than that of the miRNA model in this experiment, but smaller than those of the other two models in this experiment. In addition, compared with other experiments, this

experiment only selected the sequencing data of primary tumor tissues rather than the sequencing data of all tumors including metastatic tumors. Further, we divided the patients into high-risk and low-risk groups according to PI value, and Kaplan–Meier analysis showed that there was a significant difference in survival between the two groups. Visualizing the relationship among survival time, PI score, and grouping intuitively showed the poor survival of high-risk patients. These results indicate that these three RNA models have good specificity and sensitivity in predicting the survival of WT patients. Notably, we found that these predictive RNAs were significantly associated with tumor stage and histological classification, which means that PI may have great value in the prognosis of WT patients. Multivariate Cox regression analysis also suggested that risk level can predict the prognosis of WT patients.

According to the establishment of the predictive survival models, we found some biomarkers that may have important clinical significance. lncRNAs participate in multiple cellular activities and have been shown to regulate tumorigenesis and metastasis (22, 23). miRNAs, as a bridge between lncRNAs and mRNAs, participate in tumor development. MYCNOS promotes the invasion and metastasis of neuroblastoma and rhabdomyosarcoma by regulating MYCN protein (24, 25) and may participate in the development of WT (26). miR-152 targeting DNMT1 inhibits the development of endometrial cancer (27), glioblastoma (28), and lymphomas (29). miR-181a mediates the Wnt/β-catenin pathway to accelerate the progression of colorectal cancer (30), oral squamous cell carcinoma (31), and acute lymphoblastic leukemia (32). The expression of VRK1 directly affects the proliferation of breast cancer and liver cancer (33). CircAGFG1 promotes triple-negative breast cancer progression through the circAGFG1/miR-195/CCNE1 pathway (34). Precancerous lesions in squamous cell carcinoma depend on upregulation of the NOTCH4-HEY1 pathway (35). The HGF-mediated c-Met/FRA1/HEY1 cascade may be the key to inducing the transition from cirrhosis to hepatocellular carcinoma (36). miRNA-21 promotes proliferation of human glioma cells through the PI3K/AKT/SPRY1 pathway (37). The expression of MAGEB18 affects cell proliferation and apoptosis in melanoma (38). However, the effects of the aforementioned RNAs on WT progression have not been reported at present. In addition, it remains unknown whether AC079310.1, LINC00271, AL445228.3, Z84485.1, AC091180.5, AP002518.2, AC007879.3, TCTEX1D4, RNF133, CCNE1, C10orf71, and SPAG11A can influence the development of tumors. These RNAs provide a potential direction for future investigations on WT.

The ceRNA hypothesis explains the complicated RNA molecular regulation mechanisms *via* construction of a lncRNA–miRNA–mRNA network (9). Most studies on the molecular regulation mechanism of WT constructed a network based on DERNAs (12, 13). We proposed a novel survival-related ceRNA network that provided a new dimension for predicting the prognosis of patients with WT. Many cancer-related molecules are included in this network, such as miR-181a, CCNE1, and WIF1. Recent studies have determined that some RNAs in the ceRNA network affect the prognosis of WT patients. For example, miR-15a targeted by SNHG6 (39), cyclin CCNE1 regulated by tumor suppressor gene WWOX (40), and Wnt signaling pathway inhibited by WIF1 (41) all affect tumor prognosis. In view of the fact that mRNA is the executor of ceRNA network function, GO and KEGG analyses were conducted to gain insight into molecular mechanisms. The GO analysis suggested that the cell surface receptor signaling pathway is the main mechanism by which the ceRNA network affects the prognosis of WT patients. KEGG analysis suggested that the "Wnt signaling pathway" and "Cellular senescence" were the main enrichment pathways. Recently, the Wnt/β-catenin signaling pathway was demonstrated to regulate the occurrence and development of WT, and Wnt-targeted agents may have great potential in the treatment of WT (42, 43).

Our study had several limitations. First, due to the scarcity of patients, we only obtained RNA sequencing data from the TARGET database; thus, we could not perform multicenter validation. Second, all the data in this study are from the same pathological tissue and lacked certain accuracy for highly heterogeneous WT. In the future, we will use single-cell sequencing to detect the expression RNAs in multiple WT samples from the same patient to improve the accuracy of the results. Finally, some new RNAs that may be involved in WT progression need further study.

To conclude, we established a multinomial predictive survival model, which has a good application prospect in future clinical practice. It is expected to improve the long-term prognosis of WT patients by screening high-risk patients through sequencing results and strengthening the personalized treatment. Meanwhile, we describe a survival-related ceRNA network, which provides some new potential prognostic indicators and therapeutic targets for improving the prognosis of WT patients.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

HL and ZL designed the study. HL, MZ, and MS performed the data collection. HL, MZ, and TZ analyzed the data. HL, ZZ, and SY wrote the manuscript. HL and QC revised the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the National Natural Science Foundation of China (81871837, 81572117) and the Specialized Research Fund for Doctoral Program of Higher Education of China (20132307110007).

ACKNOWLEDGMENTS

The author would like to thank the staff in scientific research center of the Second Affiliated Hospital of Harbin Medical University.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.608433/full#supplementary-material

REFERENCES

1. Al-Hussain T, Ali A, Akhtar M. Wilms tumor: an update. *Adv Anat Pathol* (2014) 21:166–73. doi: 10.1097/PAP.0000000000000017
2. Treger TD, Chowdhury T, Pritchard-Jones K, Behjati S. The genetic changes of Wilms tumour. *Nat Rev Nephrol* (2019) 15:240–51. doi: 10.1038/s41581-019-0112-0
3. Davidoff AM. Wilms' tumor. *Curr Opin Pediatr* (2009) 21:357–64. doi: 10.1097/MOP.0b013e32832b323a
4. Anvar Z, Acurzio B, Roma J, Cerrato F, Verde G. Origins of DNA methylation defects in Wilms tumors. *Cancer Lett* (2019) 457:119–28. doi: 10.1016/j.canlet.2019.05.013
5. Lin A, Zhou M, Hua RX, Zhang J, Zhou H, Li S, et al. METTL3 polymorphisms and Wilms tumor susceptibility in Chinese children: A five-center case-control study. *J Gene Med* (2020) 22:e3255. doi: 10.1002/jgm.3255
6. Evans JR, Feng FY, Chinnaiyan AM. The bright side of dark matter: lncRNAs in cancer. *J Clin Invest* (2016) 126:2775–82. doi: 10.1172/JCI84421
7. Rupaimoole R, Slack FJ. MicroRNA therapeutics: towards a new era for the management of cancer and other diseases. *Nat Rev Drug Discovery* (2017) 16:203–22. doi: 10.1038/nrd.2016.246
8. Qian X, Zhao J, Yeung PY, Zhang QC, Kwok CK. Revealing lncRNA Structures and Interactions by Sequencing-Based Approaches. *Trends Biochem Sci* (2019) 44:33–52. doi: 10.1016/j.tibs.2018.09.012
9. Salmena L, Poliseno L, Tay Y, Kats L, Pandolfi PP. A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? *Cell* (2011) 146:353–8. doi: 10.1016/j.cell.2011.07.014
10. Bai Y, Long J, Liu Z, Lin J, Huang H, Wang D, et al. Comprehensive analysis of a ceRNA network reveals potential prognostic cytoplasmic lncRNAs involved in HCC progression. *J Cell Physiol* (2019) 234:18837–48. doi: 10.1002/jcp.28522
11. Abdollahzadeh R, Daraei A, Mansoori Y, Sepahvand M, Amoli MM, Tavakkoly-Bazzaz J. Competing endogenous RNA (ceRNA) cross talk and language in ceRNA regulatory networks: A new look at hallmarks of breast cancer. *J Cell Physiol* (2019) 234:10080–100. doi: 10.1002/jcp.27941
12. Tang F, Lu Z, Wang J, Li Z, Wu W, Duan H, et al. Competitive endogenous RNA (ceRNA) regulation network of lncRNAs, miRNAs, and mRNAs in Wilms tumour. *BMC Med Genomics* (2019) 12:194. doi: 10.1186/s12920-019-0644-y
13. Zhang F, Zeng L, Cai Q, Xu Z, Liu R, Zhong H, et al. Comprehensive Analysis of a Long Noncoding RNA-Associated Competing Endogenous RNA Network in Wilms Tumor. *Cancer Control* (2020) 27:1–13. doi: 10.1177/1073274820936991. 1073274820936991.
14. Routh JC, Grundy PE, Anderson JR, Retik AB, Kurek KC. B7-h1 as a biomarker for therapy failure in patients with favorable histology Wilms tumor. *J Urol* (2013) 189:1487–92. doi: 10.1016/j.juro.2012.11.012
15. Lahoti C, Thorner P, Malkin D, Yeger H. Immunohistochemical detection of p53 in Wilms' tumors correlates with unfavorable outcome. *Am J Pathol* (1996) 148:1577–89.
16. Pei N, Cao L, Liu Y, Wu J, Song Q, Zhang Z, et al. XAB2 tagSNPs contribute to non-small cell lung cancer susceptibility in Chinese population. *BMC Cancer* (2015) 15:560. doi: 10.1186/s12885-015-1567-4
17. Dou N, Yang D, Yu S, Wu B, Gao Y, Li Y. SNRPA enhances tumour cell growth in gastric cancer through modulating NGF expression. *Cell Prolif* (2018) 51:e12484. doi: 10.1111/cpr.12484
18. Zong S, Zhao J, Liu L. miR-30d Induced Apoptosis by Targeting Sox4 to Inhibit the Proliferation, Invasion and Migration of Nephroblastoma. *Onco Targets Ther* (2020) 13:7177–88. doi: 10.2147/OTT.S251714
19. Bao JW, Li WJ, Guo JH, Lv ZB, Whang Z. MiRNA-203a-5p alleviates the malignant progression of Wilms' tumor via targeting JAG1. *Eur Rev Med Pharmacol Sci* (2020) 24:5329–35. doi: 10.26355/eurrev_202005_21315
20. Wang HF, Zhang YY, Zhuang HW, Xu M. MicroRNA-613 attenuates the proliferation, migration and invasion of Wilms' tumor via targeting FRS2. *Eur Rev Med Pharmacol Sci* (2017) 21:3360–9.
21. Gong Y, Zou B, Chen J, Ding L, Li P, Chen J, et al. Potential Five-MicroRNA Signature Model for the Prediction of Prognosis in Patients with Wilms Tumor. *Med Sci Monit* (2019) 25:5435–44. doi: 10.12659/MSM.916230
22. Arun G, Diermeier SD, Spector DL. Therapeutic Targeting of Long Non-Coding RNAs in Cancer. *Trends Mol Med* (2018) 24:257–77. doi: 10.1016/j.molmed.2018.01.001
23. Wei JW, Huang K, Yang C, Kang CS. Non-coding RNAs as regulators in epigenetics (Review). *Oncol Rep* (2017) 37:3–9. doi: 10.3892/or.2016.5236
24. Zhao X, Li D, Pu J, Mei H, Yang D, Xiang X, et al. CTCF cooperates with noncoding RNA MYCNOS to promote neuroblastoma progression through facilitating MYCN expression. *Oncogene* (2016) 35:3565–76. doi: 10.1038/onc.2015.422
25. O'Brien EM, Selfe JL, Martins AS, Walters ZS, Shipley JM. The long non-coding RNA MYCNOS-01 regulates MYCN protein levels and affects growth of MYCN-amplified rhabdomyosarcoma and neuroblastoma cells. *BMC Cancer* (2018) 18:217. doi: 10.1186/s12885-018-4129-8
26. Micale MA, Embrey B4, Macknis JK, Harper CE, Aughton DJ. Constitutional 560.49 kb chromosome 2p24.3 duplication including the MYCN gene identified by SNP chromosome microarray analysis in a child with multiple congenital anomalies and bilateral Wilms tumor. *Eur J Med Genet* (2016) 59:618–23. doi: 10.1016/j.ejmg.2016.10.010
27. Tsuruta T, Kozaki K, Uesugi A, Furuta M, Hirasawa A, Imoto I, et al. miR-152 is a tumor suppressor microRNA that is silenced by DNA hypermethylation in endometrial cancer. *Cancer Res* (2011) 71:6450–62. doi: 10.1158/0008-5472.CAN-11-0364
28. Zhang P, Sun H, Yang B, Luo W, Liu Z, Wang J, et al. miR-152 regulated glioma cell proliferation and apoptosis via Runx2 mediated by DNMT1 [published correction appears in Biomed Pharmacother. 2020 Jun;126:110090]. *BioMed Pharmacother* (2017) 92:690–5. doi: 10.1016/j.biopha.2017.05.096
29. Wang QM, Lian GY, Song Y, Peng ZD, Xu SH, Gong Y. Downregulation of miR-152 contributes to DNMT1-mediated silencing of SOCS3/SHP-1 in non-Hodgkin lymphoma. *Cancer Gene Ther* (2019) 26:195–207. doi: 10.1038/s41417-018-0057-7
30. Han P, Li JW, Zhang BM, Lv JC, Li YM, Gu XY, et al. The lncRNA CRNDE promotes colorectal cancer cell proliferation and chemoresistance via miR-181a-5p-mediated regulation of Wnt/β-catenin signaling. *Mol Cancer* (2017) 16:9. doi: 10.1186/s12943-017-0583-1
31. Li GH, Ma ZH, Wang X. Long non-coding RNA CCAT1 is a prognostic biomarker for the progression of oral squamous cell carcinoma via miR-181a-mediated Wnt/β-catenin signaling pathway. *Cell Cycle* (2019) 18:2902–13. doi: 10.1080/15384101.2019.1662257
32. Lyu X, Li J, Yun X, Huang R, Deng X, Wang Y, et al. miR-181a-5p, an inducer of Wnt-signaling, facilitates cell proliferation in acute lymphoblastic leukemia. *Oncol Rep* (2017) 37:1469–76. doi: 10.3892/or.2017.5425
33. Lee N, Kim DK, Han SH, Ryu HG, Park SJ, Kim KT, et al. Comparative Interactomes of VRK1 and VRK3 with Their Distinct Roles in the Cell Cycle of Liver Cancer. *Mol Cells* (2017) 40:621–31. doi: 10.14348/molcells.2017.0108
34. Yang R, Xing L, Zheng X, Sun Y, Wang X, Chen J. The circRNA circAGFG1 acts as a sponge of miR-195-5p to promote triple-negative breast cancer progression through regulating CCNE1 expression. *Mol Cancer* (2019) 18:4. doi: 10.1186/s12943-018-0933-7
35. Fukusumi T, Guo TW, Sakai A, Ando M, Ren S, Haft S, et al. The NOTCH4-HEY1 Pathway Induces Epithelial-Mesenchymal Transition in Head and Neck Squamous Cell Carcinoma. *Clin Cancer Res* (2018) 24:619–33. doi: 10.1158/1078-0432.CCR-17-1366

36. Lau EY, Lo J, Cheng BY, Ma MK, Lee JM, Ng JK, et al. Cancer-Associated Fibroblasts Regulate Tumor-Initiating Cell Plasticity in Hepatocellular Carcinoma through c-Met/FRA1/HEY1 Signaling. *Cell Rep* (2016) 15:1175–89. doi: 10.1016/j.celrep.2016.04.019
37. Chai C, Song LJ, Han SY, Li XQ, Li M. MicroRNA-21 promotes glioma cell proliferation and inhibits senescence and apoptosis by targeting SPRY1 via the PTEN/PI3K/AKT signaling pathway. *CNS Neurosci Ther* (2018) 24:369–80. doi: 10.1111/cns.12785
38. Lin Y, Wen T, Meng X, Wu Z, Zhao L, Wang P, et al. The mouse Mageb18 gene encodes a ubiquitously expressed type I MAGE protein and regulates cell proliferation and apoptosis in melanoma B16-F0 cells. *Biochem J* (2012) 443:779–88. doi: 10.1042/BJ20112054
39. Su L, Wu A, Zhang W, Kong X. Silencing long non-coding RNA SNHG6 restrains proliferation, migration and invasion of Wilms' tumour cell lines by regulating miR-15a. *Artif Cells Nanomed Biotechnol* (2019) 47:2670–7. doi: 10.1080/21691401.2019.1633338
40. Płuciennik E, Nowakowska M, Wujcicka WI, Sitkiewicz A, Kazanowska B, Zielińska E, et al. Genetic alterations of WWOX in Wilms' tumor are involved in its carcinogenesis. *Oncol Rep* (2012) 28:1417–22. doi: 10.3892/or.2012.1940
41. Fukuzawa R, Anaka MR, Heathcott RW, McNoe LA, Morison IM, Perlman EJ, et al. Wilms tumour histology is determined by distinct types of precursor lesions and not epigenetic changes. *J Pathol* (2008) 215:377–87. doi: 10.1002/path.2366
42. Zhu KR, Sun QF, Zhang YQ. Long non-coding RNA LINP1 induces tumorigenesis of Wilms' tumor by affecting Wnt/β-catenin signaling pathway. *Eur Rev Med Pharmacol Sci* (2019) 23:5691–8. doi: 10.26355/eurrev_201907_18306
43. Perotti D, Hohenstein P, Bongarzone I, Maschietto M, Weeks M, Radice P, et al. Is Wilms tumor a candidate neoplasia for treatment with WNT/β-catenin pathway modulators?–A report from the renal tumors biology-driven drug development workshop. *Mol Cancer Ther* (2013) 12:2619–27. doi: 10.1158/1535-7163.MCT-13-0335

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling editor declared a shared affiliation, though no other collaboration, with the authors.

frontiers in Oncology

ORIGINAL RESEARCH
published: 18 March 2021
doi: 10.3389/fonc.2021.611580

Single-Cell RNA-seq Reveals Characteristics of Malignant Cells and Immune Microenvironment in Subcutaneous Panniculitis-Like T-Cell Lymphoma

Zifeng Li[1†], Hongsheng Wang[1†], Rui Dong[2], Jie Man[1], Li Sun[3], Xiaowen Qian[1], Xiaohua Zhu[1], Ping Cao[1], Yi Yu[1], Jun Le[1], Yang Fu[1], Ping Wang[1], Wenjin Jiang[1], Chen Shen[1], Yangyang Ma[4], Lian Chen[4], Yaochen Xu[5], Jiantao Shi[5], Hui Zhang[6], Maoxiang Qian[7] and Xiaowen Zhai[1*]*

[1] Department of Hematology and Oncology, Children's Hospital of Fudan University, National Children's Medical Center, Shanghai, China, [2] Department of Pediatric Surgery, Children's Hospital of Fudan University, National Children's Medical Center, Shanghai, China, [3] Department of Rheumatism and Immunology, Children's Hospital of Fudan University, National Children's Medical Center, Shanghai, China, [4] Department of Pathology, Children's Hospital of Fudan University, National Children's Medical Center, Shanghai, China, [5] Shanghai Institute of Biochemistry and Cell Biology, Center for Excellence in Molecular Cell Science, Chinese Academy of Sciences, Shanghai, China, [6] Department of Hematology/Oncology, Guangzhou Women and Children's Medical Center, Guangzhou, China, [7] Institute of Pediatrics, Children's Hospital of Fudan University, National Children's Medical Center, and the Shanghai Key Laboratory of Medical Epigenetics, Institutes of Biomedical Sciences, Fudan University, Shanghai, China

Background: Subcutaneous panniculitis-like T-cell lymphoma (SPTCL) is a malignant primary T-cell lymphoma that is challenging to distinguish from autoimmune disorders and reactive panniculitides. Delay in diagnosis and a high misdiagnosis rate affect the prognosis and survival of patients. The difficulty of diagnosis is mainly due to an incomplete understanding of disease pathogenesis.

Methods: We performed single-cell RNA sequencing of matched subcutaneous lesion tissue, peripheral blood, and bone marrow from a patient with SPTCL, as well as peripheral blood, bone marrow, lymph node, and lung tissue samples from healthy donors as normal controls. We conducted cell clustering, gene expression program identification, gene differential expression analysis, and cell-cell interaction analysis to investigate the ecosystem of SPTCL.

Results: Based on gene expression profiles in a single-cell resolution, we identified and characterized the malignant cells and immune subsets from a patient with SPTCL. Our analysis showed that SPTCL malignant cells expressed a distinct gene signature, including chemokines families, cytotoxic proteins, T cell immune checkpoint molecules, and the immunoglobulin family. By comparing with normal T cells, we identified potential novel markers for SPTCL (e.g., *CYTOR, CXCL13, VCAM1, and TIMD4*) specifically differentially expressed in the malignant cells. We also found that macrophages and fibroblasts dominated the cell-cell communication landscape with the SPTCL malignant cells.

OPEN ACCESS

Edited by:
Yizhuo Zhang,
Sun Yat-sen University Cancer Center (SYSUCC), China

Reviewed by:
David Michonneau,
Assistance Publique Hopitaux De Paris, France
Adam Lamble,
Seattle Children's Hospital, United States

***Correspondence:**
Xiaowen Zhai
zhaixiaowendy@163.com
Maoxiang Qian
mxqian@fudan.edu.cn

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Pediatric Oncology, a section of the journal Frontiers in Oncology

Received: 29 September 2020
Accepted: 01 March 2021
Published: 18 March 2021

Citation:
Li Z, Wang H, Dong R, Man J, Sun L, Qian X, Zhu X, Cao P, Yu Y, Le J, Fu Y, Wang P, Jiang W, Shen C, Ma Y, Chen L, Xu Y, Shi J, Zhang H, Qian M and Zhai X (2021) Single-Cell RNA-seq Reveals Characteristics of Malignant Cells and Immune Microenvironment in Subcutaneous Panniculitis-Like T-Cell Lymphoma. Front. Oncol. 11:611580. doi: 10.3389/fonc.2021.611580

Conclusions: This work offers insight into the heterogeneity of subcutaneous panniculitis-like T-cell lymphoma, providing a better understanding of the transcription characteristics and immune microenvironment of this rare tumor.

Keywords: single-cell RNA-seq (scRNA-seq), T cell malignancies, pediatric oncology, molecular diagnoses, subcutaneous panniculitis- like T-cell lymphoma

INTRODUCTION

Subcutaneous panniculitis-like T-cell lymphoma (SPTCL) is a rare primary cutaneous lymphoma of mature cytotoxic T cells arising primarily in the skin without the evidence of extracutaneous involvement. According to the 2016 World Health Organization (WHO) and 2018 World Health Organization-European Organization for Research and Treatment of Cancer (WHO-EORTC) classification, SPTCL is defined as subcutaneous lymphomas with an α/β T cell phenotype and neoplastic T cells expressing CD3, CD8, and cytotoxic proteins (GZMB, TIA-1, perforin) (1, 2). Both children (3) and adults can be affected, with a median age at diagnosis of 36 years and female gender bias (4). In a cohort of pediatric patients (3), the median age at diagnosis was 8 years (5 months to 21 years) with a male to female ratio of 1:1.7. The disease response to therapy is usually favorable, with a 5-year survival of more than 80% (5).

However, the clinical manifestations and pathological features of SPTCL are similar to those of benign panniculitis, lupus erythematosus profundus (LEP), and various autoimmune disorders, thus SPTCL is frequently misdiagnosed at the early stage (6). The long diagnosis period and high misdiagnosis rate may affect the prognosis and survival of patients. Although recent studies have provided insight into pathways that may be important to the pathogenesis of this disease (5, 7–11), additional investigations are required to better understand the profile and ecosystem of SPTCL.

Here, we conducted single-cell RNA sequencing (scRNA-seq) to decipher SPTCL at an unprecedented transcriptomic resolution for matched subcutaneous lesion tissue, peripheral blood, and bone marrow from a patient with SPTCL, as well as peripheral blood, bone marrow, lymph node, and lung tissue samples from healthy donors as normal controls. Using this dataset, we investigated the ecosystem of SPTCL and identified novel markers of SPTCL that may advance the detection and diagnosis of this disease.

METHODS

Patient

A male patient diagnosed with SPTCL was recruited from the Children's Hospital of Fudan University in the Department of Hematology and Oncology. At the time of sample collection, the patient was 22 months old with SPTCL.

This study was approved by the Medical Ethics Committee of the Children's Hospital of Fudan University institutional review board and conducted under the Declaration of Helsinki principles (approval reference: No (2020). 307). Informed written consent was obtained from the parents before inclusion in the study.

Healthy Donors

Healthy donors' datasets were downloaded from the Gene Expression Omnibus (GEO, accession number: GSE126030) (12). The samples were obtained from deceased, brain-dead donors at the time of organ acquisition for clinical transplantation. Donors were free of chronic disease, cancer, and chronic infections such as Hepatitis B, C, and HIV. The mononuclear cells were isolated from human lungs (LG), lymph nodes (LN), bone marrow (BM), and blood, and the untouched CD3+ T cells were enriched from single-cell suspensions of all tissues and blood using magnetic negative selection (MojoSort Human CD3+ T cell Isolation Kit; BioLegend) (12).

Single-Cell RNA Sequencing

Experimental procedures followed established techniques using the Chromium Single Cell 3' Library V3 kit (10x Genomics). Briefly, mononuclear cells from enzymatically digested subcutaneous lesion biopsies and bone marrow, as well as peripheral blood by density gradient centrifugation using Lymphocyte Separation Medium, were loaded into the Chromium instrument (10X Genomics), and the resulting barcoded cDNAs were used to construct libraries. RNA-seq was performed on each sample (approximately 200 million reads/sample). Raw sequence data were converted into FASTQs using the Illumina bcl2fastq software. FASTQ files were aligned to the human genome (GRCh38) using the *CellRanger* v3.0.1 (10x Genomics) pipeline according to the manufacturer's instructions.

Single-Cell Data Processing and Analysis

Initial data processing of scRNA-seq for peripheral blood (n = 6,463), bone marrow (n = 11,027), and subcutaneous lesion tissue (n = 19,247) from the patient were performed using Python 3.6 and the Single Cell Analysis in Python (*Scanpy*) (v1.4.6) (13) unless otherwise stated. Healthy donors' scRNA-seq data were also processed in the same way. Individual cells were filtered based on the total number of genes expressed and the percentage of mitochondrial reads. The cells were included with genes greater than 200 but less than 6,000, and the percentage of mitochondrial reads less than 10%. Genes detected in fewer than three cells were filtered out. Read counts of qualified cells were normalized using the deconvolution method implemented in the R package *Scran* (v3.11) (14) and in-transformed.

Single-Sample Analysis

For visualization, a UMAP was calculated by computing the single-cell neighborhood graph (kNN-graph) on the specific principal components using 15 neighbors. The number of principal components utilized in the neighborhood graph was based on the standard deviations of the top 30 principal components. The Leiden graph-clustering method was used to cluster the neighborhood graph of cells.

Cell types were manually assigned to the clusters from the Leiden graph-clustering by comparing the mean expression of known markers across cells in a cluster. Markers used to type cells included *CD19, MS4A1, CD79A* (B cells), *CD2, CD3, CD4, CD8* (T cells), *CCR7, IL7R, LEF1, SELL* (naive T cells), *CD44, CXCR3* (memory T cells), *IL2RA, FOXP3, IKZF2* (Tregs), *CXCR5, BCL6, KLRB1, CCR4, TBX21, GATA3* (Th cells), *NCAM1, NKG7* (NK cells), *CD14, FCGR3A, ITGAM, CD68, ITGB2, ADGRE1, LYZ* (macrophages), *IRF8, CLEC4C* (dendritic cells), *DPP4, TAGLN, COL1A1, PDGFRA* (fibroblasts), and *CD34* (progenitor).

A consensus non-negative matrix factorization (cNMF) algorithm (15) was employed to identified gene expression programs (GEPs) following the protocol on Github https://github.com/dylkot/cNMF. The GEPs obtained were subjected to Gene Ontology (GO) and KEGG analysis using the R package *clusterProfiler* (v3.11) (16).

Integration Sample Analysis

We combined the data generated from isolated cells with CD3 and CD8 positive from peripheral blood (n= 1,812), bone marrow (n=1,143), and subcutaneous lesion tissue (n=5,956) of the patient, and healthy donors (n=13,494) to conduct integration analysis. The Scanorama algorithm (17) was applied to correct the combined dataset for technical batch effects. All reduced dimensions were the same as that in the single-sample analysis. Partition-based graph abstraction (PAGA) was calculated by *Scanpy*.

The top 100 correlated genes were defined as a GEP, and their average relative expression was calculated as a GEP cell score (18). The reference set was randomly sampled from the gene pool for each binned expression value. The number of reference genes to be sampled from each bin was 100.

The Wilcoxon rank-sum test was used to estimate and identify differentially expressed genes. The novel markers utilized a default threshold of 2 for average fold change and a filter for the minimum delta percent of cells ([X (percentage of cluster1) – X (percentage of cluster2)]/X (percentage of cluster1) * 100) greater than 90%.

InferCNV Analysis

Raw gene expression data were extracted from the Scanpy object as recommended in the "Using 10x data" section (inferCNV of the Trinity CTAT Project, https://github.com/broadinstitute/inferCNV). Normal reference cells were identified from annotated Leiden clusters as naïve T cells. Tumor cells were identified as malignant-like cells in Leiden clusters. The inferCNV analysis was performed following the tutorial (https://github.com/broadinstitute/inferCNV/wiki) with parameters including default settings.

Cell-Cell Ligand-Receptor Interactions

Cell-cell ligand-receptor interactions were inferred using the *CellPhoneDB* (v2.0.0) method in Python (19). The lower cutoff for the expression proportion of any ligand or receptor in a given cell type was set to 10%, and the number of permutations was set to 1000.

Whole-Exome Sequencing and Analysis

DNA was extracted from paraffin-embedded (FFPE) SPTCL tissue for whole-exome sequencing (WES). The Agilent SureSelect Human All Exon V6 kit was used for exome capture and library preparation. Paired-end sequencing (2 x 150 bp read length) was performed using the Illumina NovaSeq platform. Reads were mapped to the human genome (GRCh37) reference sequence by the Burrows-Wheeler aligner (bwa mem) algorithm (version 0.7.17) (20). The data processing, including indel realignment, marking duplicates, and recalibrating base quality scores, were performed according to the GATK best practices using GATK (version 3.7) (21) and Picard tools (version 2.18.25, http://broadinstitute.github.io/picard). Variants in the *HAVCR2* gene were manually checked using the Integrative Genomics Viewer (IGV) with the bam file (22).

H&E and Immunohistochemistry Staining

The formalin-fixed and paraffin-embedded tissue was cut into 4-μm thick sections and affixed onto the slides. The slides were subjected to H&E staining and immunohistochemistry. After being deparaffinized and rehydrated, the antigens were retrieved in boiled Tris-EDTA (pH 9.0) buffer for 15 min, cooled off for 1 h in the fume hood, and then blocked according to the protocol of the DAB polymer detection kit (Gene Tech, Shanghai, China) for 10 min. The slides were incubated with primary antibody in 1% bovine serum albumin (BSA)/tris-base solution buffer at 4°C overnight. The next day, the slides were incubated with the secondary antibody and developed with DAB reagent according to the protocol of the DAB polymer detection kit (Gene Tech). Finally, the slides were counterstained with hematoxylin. Anti-CD3 antibody (Catalog Number : AR0042, Talent Biomedical, 1:500), anti-CD4 antibody (Catalog Number : AR0273, Talent Biomedical, 1:500), anti-CD8 antibody (Catalog Number : AM0063, Talent Biomedical, 1:500), anti-TIA-1 antibody (Catalog Number : AM0226, Talent Biomedical, 1:500), anti-Granzyme B antibody (Catalog Number : AM0308, Talent Biomedical, 1:500), anti-Perforin antibody (Catalog Number : AM0311, Talent Biomedical, 1:500), anti-Ki67 antibody (Catalog Number : AR0248, Talent Biomedical, 1:500), anti-CXCL13 (Catalog Number:10927-1-AP, Proteintech, 1:500), anti-TIMD4 (Catalog Number:12008-1-AP, Proteintech, 1:500), and anti-VCAM1 (Catalog Number:11444-1-AP, Proteintech, 1:400) were used.

RESULTS

Clinical Characteristics of the Studied Patient With SPTCL

The clinical manifestations of the studied patient with SPTCL are summarized in **Table 1**. The initial disease onset of this male patient was at six months old, diagnosed with a small hard nodule (diameter: 1 cm) in the left clavicle and enlarged lymph nodes in the groin. When he was 12 months old after a measles vaccination, the initial nodule was significantly enlarged (diameter: 6 cm) with enlarged lymph nodes in the head of the pancreas and did not decrease significantly after antibiotic treatment, puncture, and drainage. At the age of 18 months, the disease progressed with multiple lesions at the root of the patient's right thigh (diameter: 5 cm) accompanied by fever and then at the left hip (diameter: 3 cm) after anti-inflammatory treatment for controlling body temperature. Four months later (22 months old), the patient progressed with a new single lesion at the right shoulder (diameter: 2 cm) with no fever but enlarged lymph nodes in the neck, underarms, mediastinum, and groin. Subcutaneous lesions were more common in the extremities and partly in the trunk. The lesions varied from 1 cm to 6 cm in diameter, with redness and swelling. No ulcerated plaque was observed. Multiple lymphadenopathies were proven by computerized tomography (CT) scans without hepatomegaly. The patient did not receive any chemotherapy but was followed up according to his parents' decision.

Histopathological, immunophenotypical, and molecular features of the patient samples are also summarized in **Table 1**. All skin biopsy specimens demonstrated a dense lymphoid infiltrate located in the subcutaneous tissue, with the overlying epidermis and dermis involved. Atypical lymphocytes were pleomorphic small to medium-sized to diffusely large T cells with irregular hyperchromatic nuclei and were admixed with small lymphocytes and histiocytes, which were found in both biopsy specimens at 18 and 22 months old. Areas of karyokinesis and karyorrhexis were seen. These atypical lymphocytes showed a $CD3^+$, $CD4^-$, $CD8^+$, Granzyme B^+, $Perforin^+$, and $TIA1^+$ phenotype (**Figure 1**) with a high proliferation rate. Epstein-Barr virus (EBV) detection by EBV-encoded RNA (EBER) in situ hybridization was negative. Clonal rearrangement of the TCR beta gene was found in the biopsy at 22 months old. In all episodes, bone marrow examination showed no evidence of lymphoma.

The SPTCL-Specific Ecosystem at Single-Cell Resolution

We used scRNA-seq to profile gene expression in cells obtained from the enzymatically digested subcutaneous lesion tissue of the biopsy before any treatment at 22 months old. Transcriptomic data were obtained from a total of 17,598 cells, with a median of 1,672 genes detected per cell. Cells were grouped according to their expression profiles by principal component analysis (PCA) and Uniform Manifold Approximation and Projection (UMAP) dimensional reduction. Unsupervised graph-based Leiden clustering by *Scanpy* identified 17 clusters of cells that were

TABLE 1 | The clinical characteristics of multiple episodes in the patient with SPTCL analyzed in this study.

The onset of symptoms (months)	Subcutaneous lesions		Other lesions	Systemic symptoms	HLH	EBV infection	Histopathological characteristics			Immunophenotype			TCR gene rearrangement
	Location	Diameter					Infiltration range	Infiltration cell	Types	IHC	Level of Ki-67		
6	Single lesion: left clavicle	1cm	Enlarged lymph nodes of the groin	None	No	No	NA	NA	NA	NA	NA	NA	NA
12	Single lesion: left clavicle	6cm	Enlarged lymph nodes in the head of the pancreas	None	No	No	Dermis and subcutaneous tissue	Lymphocytes and histocytes	Inflammation	CD1α (-), CD34 (+), CD45 (+), CD68 (+)	5%	NA
18	Multiple lesions: right thigh and left hip	3-5cm	Enlarged lymph nodes of the neck, underarms, mediastinum, and groin	Fever	No	No	Epidermis, dermis and subcutaneous fatty tissue	Heterotypic lymphocytes, histocytes and karyokinesis	Panniculitis-like	CD20 (-), CD3 (+), CD5 (+), CD7 (+), CD4 (-), CD8 (+), TIA (+), GB (+/-), CD56 (-), EBER (-)	40-50%	NA
22	Single lesion: right shoulder	2cm	Enlarged lymph nodes of the neck, underarms, mediastinum, and groin	None	No	No	Dermis and subcutaneous fatty tissue	Heterotypic lymphocytes, histocytes, karyokinesis and karyorrhexis	Panniculitis-like	CD20 (-), CD3 (+), CD5 (+), CD7 (+), CD4 (+/-), CD8 (+), TIA (+/-), GB (+/-), Perforin (+/-), CD56 (-), EBER (-)	20%	Clonal TCR-Beta gene rearrangements

EBV, Epstein-Barr virus; HLH, hemophagocytic lymphohistiocytosis; SPTCL, subcutaneous panniculitis-like T-cell lymphoma.

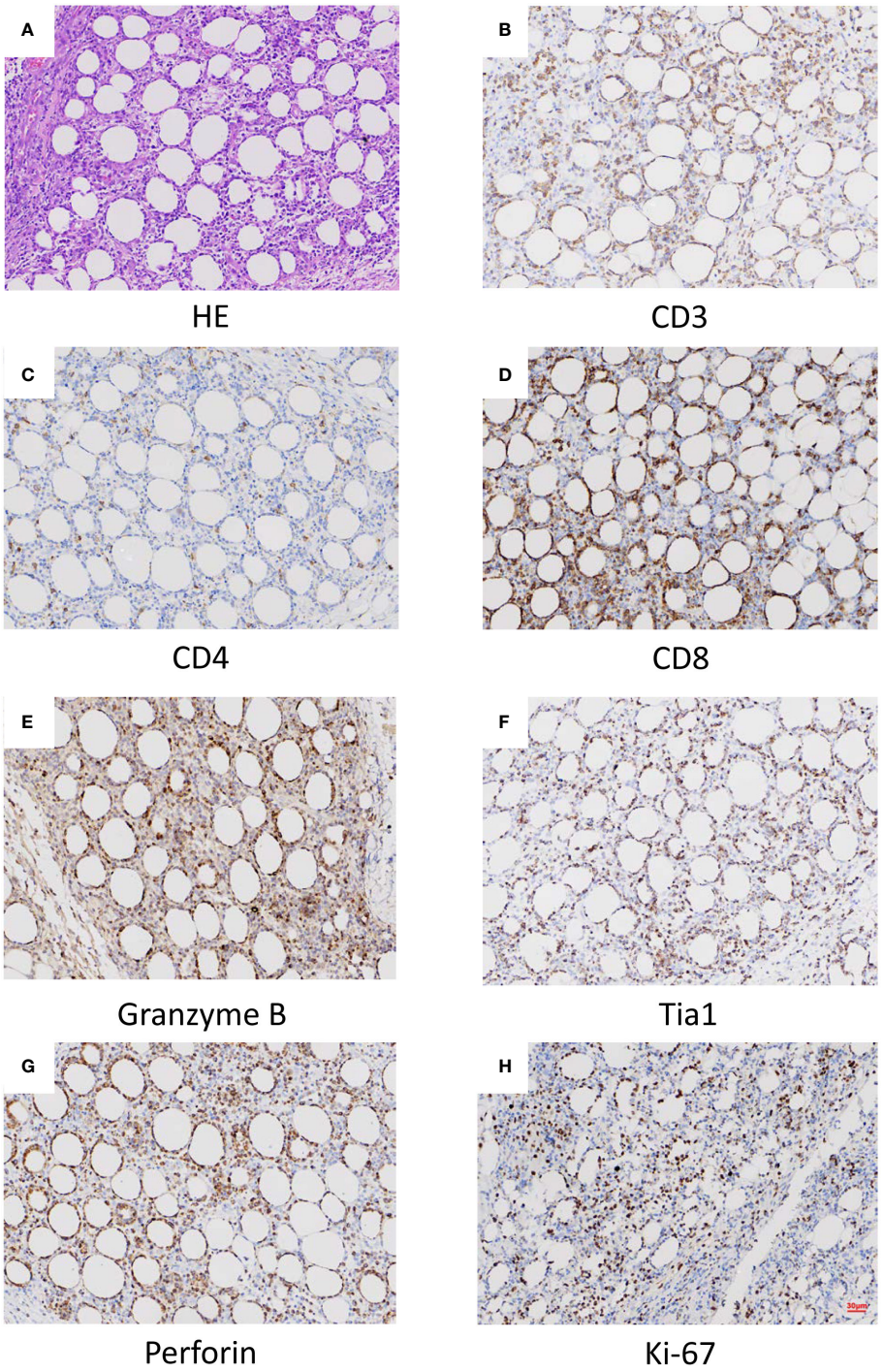


FIGURE 1 | Histopathological **(A)** and histochemical **(B–H)** results of the lesion. **(A)** Sections at low power stained with hematoxylin and eosin showing a heavy lymphocytic infiltrate predominantly in the subcutis (x40). **(B)** CD3 positive (x40). **(C)** CD4 in approximately 5% of cells (x40). **(D)** CD8 positive (x40). **(E)** Granzyme B positive (x40). **(F)** Tia1 positive (x40). **(G)** Perforin positive (x40). **(H)** Ki-67 positive (x40).

annotated and assigned with a cell type based on the expression of genes described in known canonical markers and published transcriptome data (see Methods for details) (**Figures 2A, B** and **Supplementary Figure S1**). These included one naïve T cell cluster, one Treg cell cluster, two $CD8^+$ T cell clusters, two NK cell clusters, one naïve B cell cluster, six macrophage clusters, one dendritic cell cluster, two fibroblast clusters, and one progenitor cell cluster. Macrophages were the most abundant immune cells

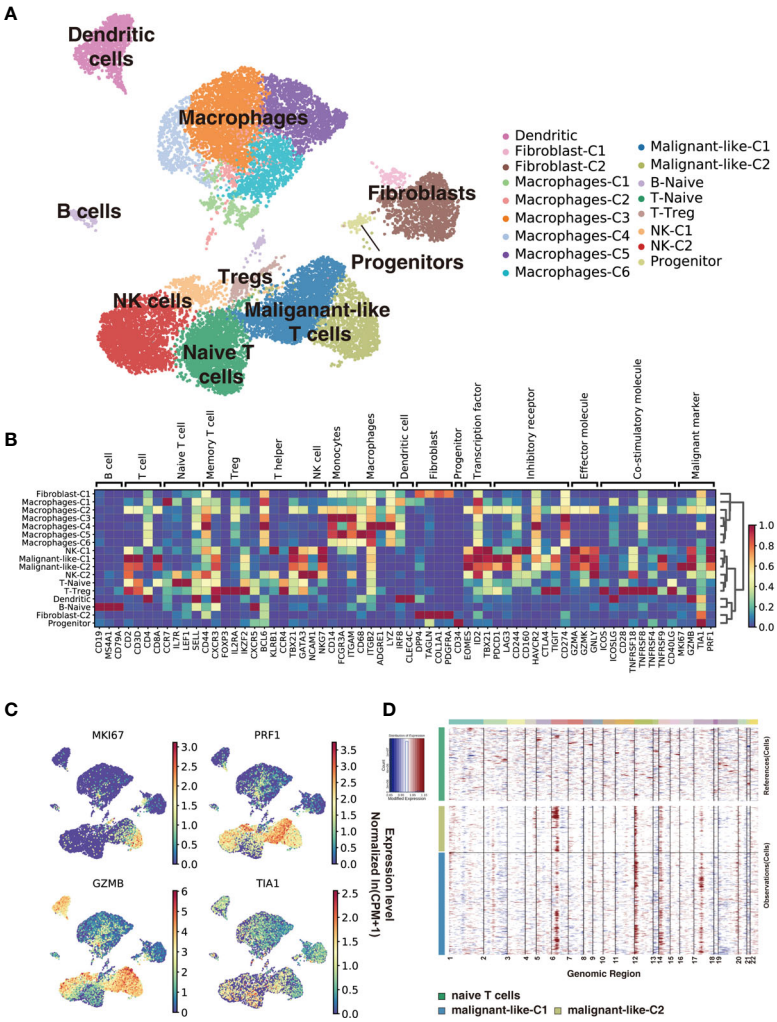


FIGURE 2 | Subcutaneous panniculitis-like T-cell lymphoma ecosystem at single-cell resolution. Cells from the patient's subcutaneous lesion tissue were clustered using the Leiden community detection algorithm to identify groups of cells with similar expression patterns. **(A)** Single-cell expression of the subcutaneous lesions' cells in UMAP space (first two dimensions). Cells are color-coded according to the clusters generated by the Leiden algorithm. **(B)** Heatmap summarizes the mean expression (normalized and log-transformed) of selected canonical markers in each cluster. The gene expression value has been scaled for visualization. The covariate bar on the top side indicates the component associated with each gene, and red boxes highlight the prominent expression of genes for the known subtypes. **(C)** UMAP plots of malignant markers (*MKI67, PRF1, TIA1, GZMB*) expression in subcutaneous lesions' cells. **(D)** Chromosomal landscape of inferred large-scale copy number variations (CNVs) distinguishes malignant from non-malignant cells. Amplifications (red) or deletions (blue) were inferred by averaging expression over 100-gene stretches on the respective chromosomes.

in our study, with a low proportion of B cells. Malignant-like T cells were identified based on conventional SPTCL markers (i.e., *MKI67*, *PRF1, TIA1*, and *GZMB*; **Figure 2C**), which were highly expressed in these two $CD8^+$ T cell clusters. However, we cannot rule out the possibility that there were a few normal $CD8^+$ T cells in these two clusters since some markers such as *GZMB* and *PRF1* were also expressed to a certain extent in normal $CD8^+$ T cells. To validate the identification of malignant-like T cells, we further distinguished malignant from non-malignant T cells by inferring large-scale chromosomal copy-number variations (CNVs) based on transcriptomes (**Figure 2D**). As expected, almost all the identified malignant-like cells (>99%) showed clear evidence of a gain of 6p, 12p, and 14p compared with normal reference cells, supporting that most of them were real malignant cells.

Next, we use cNMF (15) to infer potential GEPs underlying the expression profiles and which cells expressed the GEPs. We identified 23 distinct programs in this dataset, which were further divided into identity programs (n=19) and activity programs (n=4) based on the criterion that the former represents a unique cell type while the latter can occur in multiple diverse cell types (**Supplementary Table S1**;

Supplementary Figure S2A). Most cells had only one GEP, which represents their identity program. In addition to the 17 primary cell-type clusters initially generated by Scanpy and refined by the identity GEPs, we also identified epithelial cells, endothelial cells, and mast cells in this SPTCL-specific ecosystem (**Supplementary Figure S2A**). We noticed that the identified malignant T cells expressed one identity GEP that was significantly enriched for genes involved in cell killing and T cells activation (**Supplementary Figures S2A, B**), including the chemokines family *(i.e., CCL5, CCR5, CXCR3, CXCR6)*, cytotoxic proteins *(i.e., NKG7, GZMA, GZMB, GZMH, GZMK, GNLY, PRF1)*, and immune checkpoint genes *(i.e., LAG3, CD27, TIGIT, HAVCR2, PDCD1, CTLA4)* (**Supplementary Figure S2C**; **Supplementary Table S1**). Some malignant T cells also expressed an activity GEP named Proliferation, which was strongly enriched for genes associated with cell cycle (e.g., Mitotic Nuclear Division; **Supplementary Figures S2A, B**). Moreover, parts of malignant T cells expressed an activity GEP named Act.T, which was also expressed in naïve T cells and NK cells (**Supplementary Figure S2A**).

Comparison of Malignant and Normal T Cells by Expression Profiling

To investigate the difference between malignant and normal T cells, we paired isolated T cells from the subcutaneous lesion tissue of the patient with normal T cells from donors' peripheral blood, bone marrow, lung tissues, and lymph nodes. We applied Scanorama to correct the potential batch effects between two datasets and merged the neighbor sets *via* the UMAP algorithm as a combined dataset. Based on the cell-type clusters in SPTCL, we found that the naïve T cells from the patient overlapped with normal T cells from donors, while malignant T cells were obviously separated from them (**Supplementary Figure S3**). Using the Leiden clustering algorithm, we identified nine UMAP clusters presenting the normal versus malignant classification clearly (**Figure 3A**), actively supporting the separation within the UMAP. The graph-like maps of cells generated by the partition-based graph abstraction (PAGA) also confirmed these two distinct populations without secure connections (**Figure 3B**).

As reported by Gayden et al. (8), germline *HAVCR2* mutations altering TIM-3 were significantly overrepresented in SPTCL patients, especially with hemophagocytic lymphohistiocytosis (HLH). They also observed elevated serum levels of IFN-γ-induced CXCL10, inflammasome-activated interleukin-18 (IL-18), and soluble CD25 in a *HAVCR2* mutant SPTCL patient at the time of active disease, and increased amounts of tumor necrosis factor-α (TNF- α) and IL-2 produced *in vitro* by T lymphoblasts from *HAVCR2* mutant patients with SPTCL. Thus, we checked the genotype of *HAVCR2* by examining the whole-exome sequencing (WES) data of the patient's SPTCL tissue and did not observe any coding mutation in the *HAVCR2* gene (**Supplementary Figure S4**). We also checked the expression of genes (*HAVCR2, TNF, IL2, CXCL10, IL18*, and *IL2RA/CD25*) in our scRNA-seq data for malignant and normal T cells and found regular expression of *HAVCR2* and *CXCL10* and low expression of *TNF*, *IL2, IL18*, and *IL2RA(CD25)* in the SPTCL malignant cells (**Supplementary Figure S5**) compared with normal T cells, suggesting the difference between *HAVCR2*-wild-type and mutant SPTCL patients.

It has previously been proposed that regulatory T lymphocytes (Treg) could play an essential role in SPTCL pathogenesis, especially in the skin (8, 23). In particular, Gayden et al. identified a drastic decrease in $FOXP3^+CD4^+$ T cells in TIM-3 mutants compared with TIM-3 wild-type SPTCL (8). For comparison, we isolated $FOXP3^+CD4^+$ T cells in our SPTCL scRNA-seq data and found that the proportion (21.97%) of $FOXP3^+CD4^+$ T cells in $CD4^+$ T cells in our patient was similar to that in TIM-3 wild-type SPTCLs and higher than that in TIM-3 mutants in the reported cohort, consistent with their finding (**Supplementary Figure S6**).

Next, we used the conventional SPTCL markers to examine the separation of normal and malignant cells above. As expected, we found that the classical SPTCL marker *MKI67* was very specifically observed in malignant cells but mostly not seen in normal cells, while the markers *GZMB* and *PRF1* were not only expressed in the tumor T cells but also in part of normal $CD8^+$ T cells (**Supplementary Figure S7**). Then we scored each cell by their gene expression correlation to malignant GEPs, including the previously identified malignant identity GEP (Int.SPTCL) and activity GEPs (Proliferation and Act.T). There were significant differences in malignant and normal T cells scored with all these GEPs ($P< 0.001$) (**Figure 3C**).

Meanwhile, we also paired the isolated T cells from the subcutaneous lesion tissue with T cells from the peripheral blood and bone marrow of the patient using the same process. Interestingly, there were a small amount of $CD8^+$ T cells from peripheral blood and bone marrow in proximity to malignant T cells, and PAGA analysis also showed connections between them (**Supplementary Figures S8A, B**). Furthermore, we also found that the malignant T cells and the proximate T cells from matched peripheral blood and bone marrow were scored significantly higher than others with GEPs named Int.SPTCL and Proliferation (**Supplementary Figure S9**), suggesting that malignant-like or pre-malignant cells may exist in the circulation of the patient resulting in malignant recurrence.

To identify potential novel markers and/or therapeutic targets of SPTCL, we performed differential gene analysis by comparing the malignant cells to normal T cells. In total, we identified 45 significantly overexpressed genes in the malignant cells as potential markers for SPTCL ($P_{adj} < 0.05$ and average fold change >2). As expected, the top upregulated genes in the malignant cells were *GNLY* and the granzyme subfamily (e.g., *GZMA, GZMK*) (**Supplementary Figure S10**). We further identified potential novel markers for SPTCL including *CYTOR, CXCL13, VCAM1*, and *TIMD4*, which were specifically differentially expressed in the SPTCL cells as defined by average fold change > 2 and delta percentage > 90% in malignant T cells versus normal T cells (**Figure 3D**). Moreover, we also examined previously reported SPTCL-

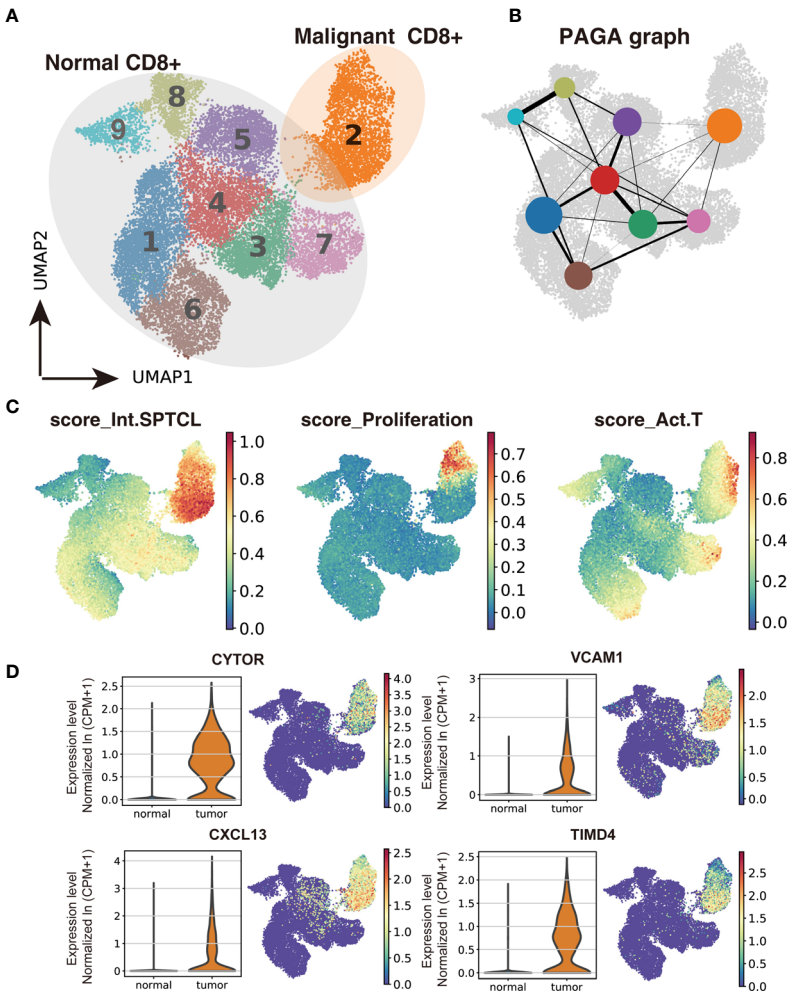


FIGURE 3 | Transcriptomic comparison of malignant versus normal CD8$^+$ T cells. **(A)** UMAP projection of cells from healthy donors and the patient's subcutaneous lesion tissue with normal CD8$^+$ T cells outlined in grey and malignant CD8$^+$ T cells in orange. **(B)** Results of partition-based graph abstraction (*PAGA*). Each node represents a cluster, and edges show the connectivity between clusters. The size of nodes indicates the number of cells in each cluster, and the edge thickness shows the connection strength. **(C)** Results of GEP-program cell scoring in UMAP space (first two dimensions). **(D)** Potential novel markers of SPTCL cells with a Δ percentage of cells expressed greater than 90% and P_{adj} < 1e-100. CPM, counts per million.

related genes (10) and found a group of genes significantly differentially expressed in the SPTCL cells (i.e., *APOBEC3G, CCL4, CCL5, CXCL10, CXCR3, FASLG, GBP5, IFNG, IKZF3, KLRD1, PRF1, and TNFRSF9* (P_{adj} < 1 x 10^{-10}; **Supplementary Figure S11**). The complete results for differential expression analysis are included in **Supplementary Table S2**.

Next, we focused on three of these potential novel markers, *CXCL13*, *VCAM1, and TIMD4*, which are protein-coding genes and presented no or shallow expression in normal lymphocytes. Their expression was examined by immunohistochemistry in the patient's subcutaneous lesion and additional samples from patients with panniculitis (PA) (**Supplementary Figure S12**). Results showed that PA lesions were negative or weakly positive for the expression of these markers, while SPTCL lesions exhibited high numbers of positive cells for all three markers.

Single-Cell Expression Patterns of Novel SPTCL-Specific Immune Subsets

To further characterize immune cells in the tumor environment of SPTCL, we annotated and dissected macrophages and fibroblasts based on the expression of genes described in known canonical markers (**Figure 4A**). We found that most of the macrophages were of the M1-type (classically activated macrophage) and M2-type (alternatively activated macrophage) with similar proportions (48.9% vs. 41.1%). The other two clusters of macrophages were not in polarized activation states; thus, they may be in M0 resting states. Intriguingly, we also identified a group of cancer-associated fibroblasts (CAFs), a type of perpetually activated fibroblasts, based on the "CAF markers", suggesting that these cells could emerge as players in immune regulation.

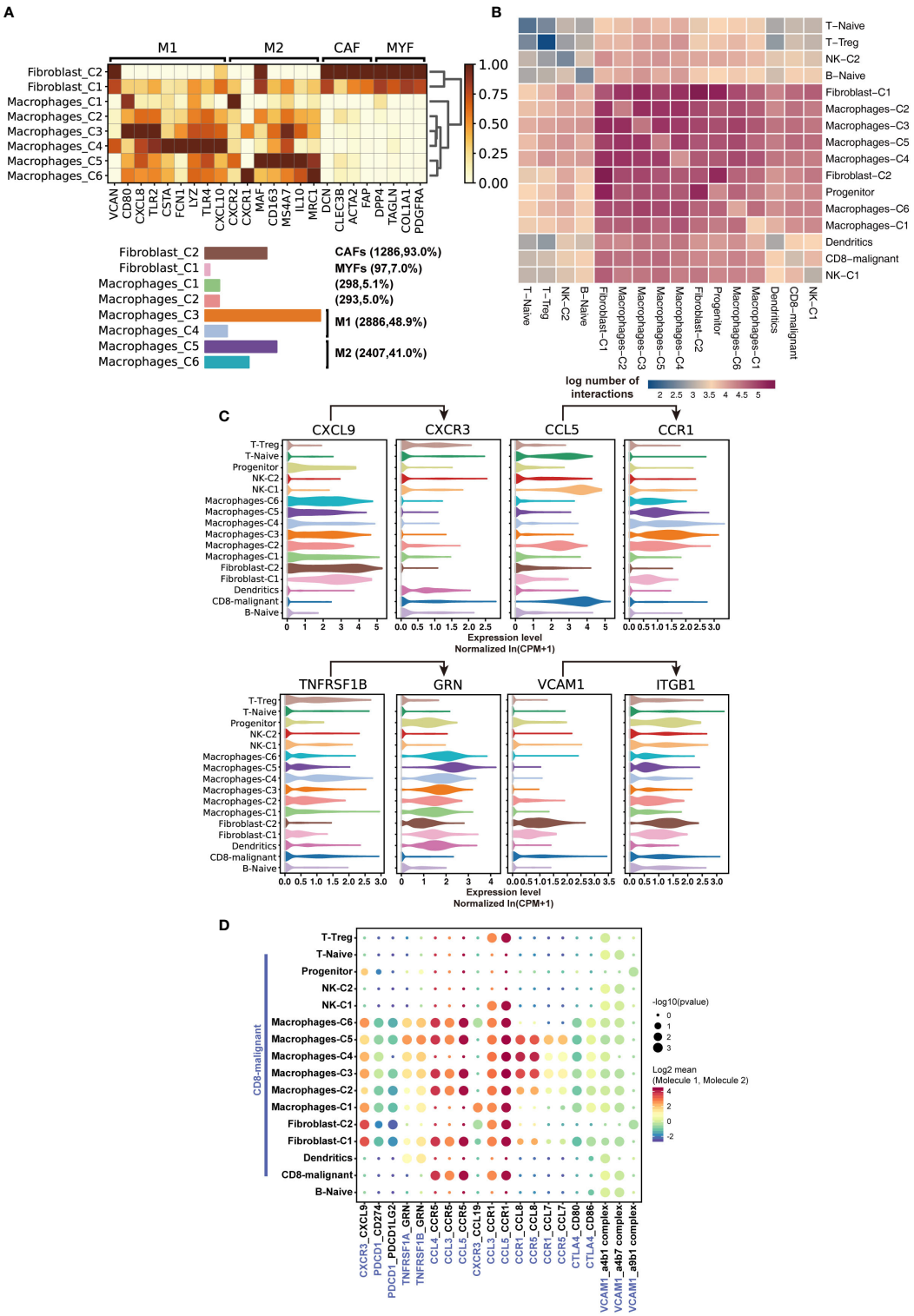


FIGURE 4 | Characteristics of SPTCL-specific immune subsets. **(A)** Heatmap summarizing mean expression (normalized and log-transformed) of M1, M2, CAF, and MYF markers in each cluster (above). Bar plot showing the cell fraction of subsets of macrophages and fibroblasts (below). M1, classically activated macrophage; M2, alternatively activated macrophage; CAF, cancer-associated fibroblasts; MYF, myofibroblasts. **(B)** Heatmap depicting the log number of all possible interactions between the clusters analyzed. **(C)** Violin plots showing expression of ligands *CXCR3, CCL5, TNFRSF1B*, and *VCAM1* and cognate receptors *CXCL9, CCR1, GRN*, and *IGTB1* on respective stromal populations. **(D)** Dot plot depicting selected tumor-immune interactions enriched in the microenvironments.

Next, we sought to elucidate the interactions of the malignant T cells with the immune populations by examining the cross-talk between them. We systematically predicted cell-cell communication networks based on *CellPhoneDB* (19), a manually curated repository of ligands, receptors, and their interactions integrated with a statistical framework to infer cell-cell communication networks from single-cell transcriptomic data. We found that the interactions of malignant T cells occurred more frequently with macrophages, fibroblasts, and dendritic cells, compared with naïve T cells, Treg, NK cells, and B cells (**Figures 4B, C**). Notably, macrophages and fibroblasts dominated the cell-cell communication landscape in this microenvironment, suggesting they might play the primary role in tumor-immune interactions of SPTCL. There was no significant difference between M1 and M2 macrophages in tumor-immune interactions. We identified multiple tumor-immune interactions, for example between *CXCR3, CCL5, TNFRSF1B*, and *VCAM1*-expressing malignant T cells and macrophages/fibroblasts positive for *CXCL9, CCR1, GRN*, and *IGTB1*, respectively (**Figure 4C**). Interestingly, we found that the recruited macrophages might promote the inflammatory activity of malignant T cells *via* suppressing the *PDCD1* and *CTLA4* axis (**Figure 4D**), because the *PDCD1*(PD-1)-*CD274*(PD-L1) and *CTL4-CD80/86* interactions can inhibit activation, expansion, and acquisition of effector functions of $CD8^+$ T cells (24).

DISCUSSION

SPTCL is a rare disease facing significant diagnostic challenges. The clinical manifestations of SPTCL are complex with only a few consistent characteristics. Subcutaneous tissue infiltration and/or infiltration by $CD3^+CD8^+$ cells expressing cytotoxic proteins (GZMB, TIA-1, perforin) is the typical pathological change of SPTCL (25). However, this change can also occur in benign panniculitis and lupus erythematosus profundus caused by autoimmune attacks (26, 27). The patient in this study experienced multiple subcutaneous mass in 16 months, and the results of the biopsy have shown that the mass evolved from benign to malignant. Because no standardized therapeutic approach has been established for SPTCL, the patient received two surgeries to remove the tumor without chemotherapy or radiotherapy. Interestingly, after the final operation, the patient has been followed up for more than one year and has not suffered a relapse. Thus, if there are proper approaches to effectively diagnose the disease and specific markers to distinguish malignant cells of SPTCL, timely surgical resection could be an effective therapy.

Here, we used scRNA-seq profiling of the malignant and normal cells from the SPTCL patient and normal cells from healthy donors to characterize the molecular events of SPTCL. To our knowledge, this is the first study exploring gene expression signatures, summarizing the tumor microenvironment of SPTCL in single-cell resolution. We identified a unique GEP that was expressed significantly higher in SPTCL cells than in normal T cells, which could be a characteristic of SPTCL. We found four genes (i.e., *CYTOR, CXCL13, VCAM1*, and *TIMD4*) explicitly expressed in malignant T cells, which may be potential novel markers for SPTCL. We also investigated interplays between different stromal populations and malignant T cells and found the leading role of macrophages and fibroblasts (especially CAFs) in the SPTCL microenvironment, suggesting their contribution to malignant T cell dysfunction. More specifically, the recruited macrophages might suppress the *PDCD1* and *CTLA4* axis to enhance the inflammatory activity of malignant cells, consistent with the clinical manifestation of SPTCL.

CYTOR (or Linc00152) is a long non-coding RNA that is overexpressed in multiple cancer cells, and it can promote cell proliferation and epithelial-mesenchymal transition (28). Given its crucial role in the pathogenesis of cancers, *CYTOR* (average fold change = 3.40, Δ percentage =95%) may play a role in SPTCL development.

CXCL13, initially identified as a B-cell chemoattractant, exerts essential functions in lymphoid neogenesis and has been widely implicated in the pathogenesis of several autoimmune diseases and inflammatory conditions, as well as in lymphoproliferative disorders (29). This chemokine has been proposed as a marker for certain lymphomas, such as angioimmunoblastic T-cell lymphoma (AITL), an aggressive nodal T-cell lymphoma derived from T_{FH} cells (2, 30). The SPTCL malignant cells highly expressed *CXCL13*, suggesting that its role in SPTCL is intriguing and worth exploring.

Vascular adhesion molecule-1 (*VCAM1*), a member of the immunoglobulin family of cell-cell adhesion receptors, is expressed aberrantly in some tumor cells, such as renal, breast, or gastric carcinomas (31–33). Clustering of VCAM-1 on the cell surface, acting through Ezrin, triggers Akt activation and protects cancer cells from proapoptotic cytokines such as the TNF-related apoptosis-inducing ligand (TRAIL) (32, 34). VCAM-1 can tether macrophages to cancer cells *via* counter-receptor α4β1-integrins, and we found that macrophages and fibroblasts in the SPTCL microenvironment highly expressed *ITGA4* and *ITGB1*, which constitute α4β1-integrins. The interaction between malignant T cells and immune cells may possess similar effects like *VCAM1*-mediated mechanisms in breast cancer cells (32, 34).

TIMD4, a member of the TIM family of immunoregulatory proteins, is overexpressed in multiple tumor tissues, which has been proven to promote tumor cell growth and proliferation both *in vitro* and *in vivo* in lung cancer (35). As reported in recent studies, *TIMD4* is expressed in professional antigen-presenting cells (APCs), pro-B cells (36), and NKT cells (37) but not in normal $CD8^+$ T cells. Its role of aberrant expression in SPTCL cells needs to be explored further in the future.

Single-cell methods allow researchers to characterize the tumor transcriptome and microenvironment in an unprecedented resolution. Our study offered a new insight into the heterogeneity of subcutaneous panniculitis-like T-cell lymphoma, providing a better understanding of the transcription characteristics and immune microenvironment of this rare tumor. This new level of data provided an opportunity for clinically meaningful advances in SPTCL.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: the Genome Sequence Archive (GSA) for Human in National Genomics Data Center (NGDC) under accession number HRA000370.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Children's Hospital of Fudan University Research Ethics Board. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

XWZ and MQ are the principal investigators of this study and take responsibility for the integrity of the data and the accuracy of the data analysis. ZL analyzed the single-cell RNA sequencing data. XWZ, RD, MQ, and ZL wrote the manuscript. HW, JM, LS, XQ, XHZ, PC, HM, YY, YF, PW, WJ, CS, YM, CL, XY, SJ, HZ, and RD provided advice on the study design and participated in data collection. XWZ, RD, MQ, and ZL interpreted the data and the research findings. All authors contributed to the article and approved the submitted version.

FUNDING

The work was supported by the Health and Family Planning Commission of Shanghai Municipality (201740011), the National Natural Science Foundation of China (81973997), and the Cyrus Tang Foundation. MQ is supported by the Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning.

ACKNOWLEDGMENTS

The authors thank the patients in this study. We also thank the Biobank staff at Children's Hospital of Fudan University for the clinical sample collection.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.611580/full#supplementary-material

REFERENCES

1. Willemze R, Cerroni L, Kempf W, Berti E, Facchetti F, Swerdlow SH, et al. The 2018 update of the WHO-EORTC classification for primary cutaneous lymphomas. *Blood* (2019) 133(16):1703–14. doi: 10.1182/blood-2018-11-881268
2. Swerdlow SH, Campo E, Pileri SA, Harris NL, Stein H, Siebert R, et al. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood* (2016) 127(20):2375–90. doi: 10.1182/blood-2016-01-643569
3. Huppmann AR, Xi L, Raffeld M, Pittaluga S, Jaffe ES. Subcutaneous panniculitis-like T-cell lymphoma in the pediatric age group: a lymphoma of low malignant potential. *Pediatr Blood Cancer* (2013) 60(7):1165–70. doi: 10.1002/pbc.24462
4. Willemze R. Cutaneous lymphomas with a panniculitic presentation. *Semin Diagn Pathol* (2017) 34(1):36–43. doi: 10.1053/j.semdp.2016.11.009
5. Hahtola S, Burghart E, Jeskanen L, Karenko L, Abdel-Rahman WM, Polzer B, et al. Clinicopathological Characterization and Genomic Aberrations in Subcutaneous Panniculitis-Like T-Cell Lymphoma. *J Invest Dermatol* (2008) 128(9):2304–9. doi: 10.1038/jid.2008.6
6. Pincus LB, LeBoit PE, McCalmont TH, Ricci R, Buzio C, Fox LP, et al. Subcutaneous panniculitis-like T-cell lymphoma with overlapping clinicopathologic features of lupus erythematosus: coexistence of 2 entities? *Am J Dermatopathol* (2009) 31(6):520–6. doi: 10.1097/DAD.0b013e3181a84f32
7. Polprasert C, Takeuchi Y, Kakiuchi N, Yoshida K, Assanasen T, Sitthi W, et al. Frequent germline mutations of HAVCR2 in sporadic subcutaneous panniculitis-like T-cell lymphoma. *Blood Adv* (2019) 3(4):588–95. doi: 10.1182/bloodadvances.2018028340
8. Gayden T, Sepulveda FE, Khuong-Quang D-A, Pratt J, Valera ET, Garrigue A, et al. Germline HAVCR2 mutations altering TIM-3 characterize subcutaneous panniculitis-like T cell lymphomas with hemophagocytic lymphohistiocytic syndrome. *Nat Genet* (2018) 50(12):1650–7. doi: 10.1038/s41588-018-0251-4
9. Fernandez-Pol S, Costa HA, Steiner DF, Ma L, Merker JD, Kim YH, et al. High-throughput Sequencing of Subcutaneous Panniculitis-like T-Cell Lymphoma Reveals Candidate Pathogenic Mutations. *Appl Immunohistochem Mol Morphol* (2019) 27(10):740–8. doi: 10.1097/PAI.0000000000000683
10. Maliniemi P, Hahtola S, Ovaska K, Jeskanen L, Vakeva L, Jantti K, et al. Molecular characterization of subcutaneous panniculitis-like T-cell lymphoma reveals upregulation of immunosuppression- and autoimmunity-associated genes. *Orphanet J Rare Dis* (2014) 9:160. doi: 10.1186/s13023-014-0160-2
11. Sen F, Rassidakis GZ, Jones D, Medeiros LJ. Apoptosis and proliferation in subcutaneous panniculitis-like T-cell lymphoma. *Mod Pathol* (2002) 15(6):625–31. doi: 10.1038/modpathol.3880577
12. Szabo PA, Levitin HM, Miron M, Snyder ME, Senda T, Yuan J, et al. Single-cell transcriptomics of human T cells reveals tissue and activation signatures in health and disease. *Nat Commun* (2019) 10(1):4706. doi: 10.1038/s41467-019-12464-3
13. Wolf FA, Angerer P, Theis FJ. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol* (2018) 19(1):15. doi: 10.1186/s13059-017-1382-0
14. Lun AT, Bach K, Marioni JC. Pooling across cells to normalize single-cell RNA sequencing data with many zero counts. *Genome Biol* (2016) 17:75. doi: 10.1186/s13059-016-0947-7
15. Kotliar D, Veres A, Nagy MA, Tabrizi S, Hodis E, Melton DA, et al. Identifying gene expression programs of cell-type identity and cellular activity with single-cell RNA-Seq. *eLife* (2019) 8:e43803. doi: 10.7554/eLife.43803
16. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *OMICS: A J Integr Biol* (2012) 16(5):284–7. doi: 10.1089/omi.2011.0118
17. Hie B, Bryson B, Berger B. Efficient integration of heterogeneous single-cell transcriptomes using Scanorama. *Nat Biotechnol* (2019) 37(6):685–91. doi: 10.1038/s41587-019-0113-3

18. Tirosh I, Izar B, Prakadan SM, Wadsworth MH, Treacy D, Trombetta JJ, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* (2016) 352(6282):189–96. doi: 10.1126/science.aad0501
19. Efremova M, Vento-Tormo M, Teichmann SA, Vento-Tormo R. CellPhoneDB: inferring cell–cell communication from combined expression of multi-subunit ligand–receptor complexes. *Nat Protoc* (2020) 15(4):1484–506. doi: 10.1038/s41596-020-0292-x
20. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* (2009) 25(14):1754–60. doi: 10.1093/bioinformatics/btp324
21. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* (2010) 20(9):1297–303. doi: 10.1101/gr.107524.110
22. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. *Nat Biotechnol* (2011) 29(1):24–6. doi: 10.1038/nbt.1754
23. Dixon KO, Das M, Kuchroo VK. Human disease mutations highlight the inhibitory function of TIM-3. *Nat Genet* (2018) 50(12):1640–1. doi: 10.1038/s41588-018-0289-3
24. Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. *Nat Rev Cancer* (2012) 12(4):252–64. doi: 10.1038/nrc3239
25. Willemze R, Jansen PM, Cerroni L, Berti E, Santucci M, Assaf C, et al. Subcutaneous panniculitis-like T-cell lymphoma: definition, classification, and prognostic factors: an EORTC Cutaneous Lymphoma Group Study of 83 cases. *Blood* (2008) 111(2):838–45. doi: 10.1182/blood-2007-04-087288
26. LeBlanc RE, Tavallaee M, Kim YH, Kim J. Useful Parameters for Distinguishing Subcutaneous Panniculitis-like T-Cell Lymphoma From Lupus Erythematosus Panniculitis. *Am J Surg Pathol* (2016) 40(6):745–54. doi: 10.1097/PAS.0000000000000596
27. Bosisio F, Boi S, Caputo V, Chiarelli C, Oliver F, Ricci R, et al. Lobular panniculitic infiltrates with overlapping histopathologic features of lupus panniculitis (lupus profundus) and subcutaneous T-cell lymphoma: a conceptual and practical dilemma. *Am J Surg Pathol* (2015) 39(2):206–11. doi: 10.1097/PAS.0000000000000307
28. Tang Y, He Y, Zhang P, Wang J, Fan C, Yang L, et al. LncRNAs regulate the cytoskeleton and related Rho/ROCK signaling in cancer metastasis. *Mol Cancer* (2018) 17(1):77. doi: 10.1186/s12943-018-0825-x
29. Kazanietz MG, Durando M, Cooke M. CXCL13 and Its Receptor CXCR5 in Cancer: Inflammation, Immune Response, and Beyond. *Front Endocrinol* (2019) 10:471. doi: 10.3389/fendo.2019.00471
30. Dupuis J, Boye K, Martin N, Copie-Bergman C, Plonquet A, Fabiani B, et al. Expression of CXCL13 by neoplastic cells in angioimmunoblastic T-cell lymphoma (AITL): a new diagnostic marker providing evidence that AITL derives from follicular helper T cells. *Am J Surg Pathol* (2006) 30(4):490–4. doi: 10.1097/00000478-200604000-00009
31. Kuai WX, Wang Q, Yang XZ, Zhao Y, Yu R, Tang XJ. Interleukin-8 associates with adhesion, migration, invasion and chemosensitivity of human gastric cancer cells. *World J Gastroenterol* (2012) 18(9):979–85. doi: 10.3748/wjg.v18.i9.979
32. Chen Q, Zhang XH, Massagué J. Macrophage binding to receptor VCAM-1 transmits survival signals in breast cancer cells that invade the lungs. *Cancer Cell* (2011) 20(4):538–49. doi: 10.1016/j.ccr.2011.08.025
33. Lin KY, Lu D, Hung CF, Peng S, Huang L, Jie C, et al. Ectopic expression of vascular cell adhesion molecule-1 as a new mechanism for tumor immune evasion. *Cancer Res* (2007) 67(4):1832–41. doi: 10.1158/0008-5472.Can-06-3014
34. Lu X, Mu E, Wei Y, Riethdorf S, Yang Q, Yuan M, et al. VCAM-1 promotes osteolytic expansion of indolent bone micrometastasis of breast cancer by engaging α4β1-positive osteoclast progenitors. *Cancer Cell* (2011) 20(6):701–14. doi: 10.1016/j.ccr.2011.11.002
35. Zhang Q, Wang H, Wu X, Liu B, Liu W, Wang R, et al. TIM-4 promotes the growth of non-small-cell lung cancer in a RGD motif-dependent manner. *Br J Cancer* (2015) 113(10):1484–92. doi: 10.1038/bjc.2015.323
36. Toda S, Hanayama R, Nagata S. Two-step engulfment of apoptotic cells. *Mol Cell Biol* (2012) 32(1):118–25. doi: 10.1128/MCB.05993-11
37. Kim HS, Kim HS, Lee CW, Chung DH. T cell Ig domain and mucin domain 1 engagement on invariant NKT cells in the presence of TCR stimulation enhances IL-4 production but inhibits IFN-gamma production. *J Immunol* (2010) 184(8):4095–106. doi: 10.4049/jimmunol.0901991

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

ORIGINAL RESEARCH
published: 30 March 2021
doi: 10.3389/fonc.2021.659709

Clinical and Molecular Differentiation Between Malignant Rhabdoid Tumor of the Kidney and Normal Tissue: A Two-Case Report

Chenghao Zhanghuang[1†], Shuo Chen[2†], Li Li[3], Zhen Yang[4], Yucheng Xie[5], Jiwei Li[5], Haoyu Tang[1], Xiaoli He[3], Liuyi Dong[2] and Bing Yan[1*]*

[1] Department of Urology, Kunming Children's Hospital, Kunming, China, [2] Department of Pharmacology, Anhui Medical University, Hefei, China, [3] Yunnan Key Laboratory of Children's Major Disease Research Kunming Children's Hospital, Kunming, China, [4] Department of Oncology, Kunming Children's Hospital, Kunming, China, [5] Department of Pathology, Kunming Children's Hospital, Kunming, China

OPEN ACCESS

Edited by:
Jing He,
Guangzhou Medical University, China

Reviewed by:
Yi Ji,
Sichuan University, China
Liu Xing,
Children's Hospital of Chongqing Medical University, China

****Correspondence:***
Bing Yan
zhanghuangchenghao@etyy.cn
Liuyi Dong
dongly@ahmu.edu.cn

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Pediatric Oncology, a section of the journal Frontiers in Oncology

Received: *28 January 2021*
Accepted: *11 March 2021*
Published: *30 March 2021*

Citation:
Zhanghuang C, Chen S, Li L, Yang Z, Xie Y, Li J, Tang H, He X, Dong L and Yan B (2021) Clinical and Molecular Differentiation Between Malignant Rhabdoid Tumor of the Kidney and Normal Tissue: A Two-Case Report. Front. Oncol. 11:659709. doi: 10.3389/fonc.2021.659709

Background: Malignant rhabdoid tumor of the kidney (MRTK) is a rare type of tumor that lacks typical clinical manifestations. Herein, we presented clinical data of 2 children with MRTK. In addition, we used a high-throughput RNA-sequencing (RNA-seq), GO analysis, and KEGG signaling pathway analysis to examine gene expression differences at the transcripts level between 2 patients with MRTK and 3 patients with non-tumor diseases without other symptoms.

Case report: Preoperative B-scan ultrasonography and computed tomography (CT) examination in 2 cases suggested nephroblastoma. Both patients were treated with radical nephrectomy. After the operation, MRTK was confirmed by pathological examination. Child 1 and Child 2 then received 7 courses and 12 courses of regular chemotherapy, respectively. Child 1 was followed up for 2 years, and Child 2 for 3.1 years without showing symptoms. RNA-seq results showed 2203 differential genes (DEGs) in the kidney tissue of children with MRTK compared to normal tissue ($p < 0.01$). GO analysis suggested that most DEGs participate in protein binding. KEGG results showed that the DEGs were mainly involved in the PI3K-Akt signaling pathway and microRNA-related proteins.

Conclusion: The PI3K-Akt signaling pathway and microRNA-related proteins as targets have extremely high potential value for the diagnosis and treatment of MRTK.

Keywords: malignant rhabdoid tumor of kidney, RNA-sequencing, KEGG, GO, child preschool, BioSystems

INTRODUCTION

Malignant rhabdoid tumor of the kidney (MRTK), a rare type of malignant rhabdoid tumor (MRT), is a highly aggressive tumor that occurs in infants and young children. The tumor has a poor prognosis, and the incidence rate in men is slightly higher than in women (1.5:1) (1). The overall 3-year survival rate in patients with MRTK ranges from 12% to 38.4%; it is the highest in children aged 24 months or older and the lowest in those between 0 and 5 months of age (1, 2). In China, a recent study reported

that within a 5-year follow-up, out of 35 cases, one survived (3). In addition, in children younger than 6 months, this condition is often accompanied by distant metastases, including metastases to the brain (4). Surgery plus postoperative radiotherapy and chemotherapy is the main approach to treat patients with MRTK. Yet, so far, there is still no uniform standard for the treatment of MRTK.

The disease was first reported in 1978 by Beckwith et al. (5) and Haas et al. (6). MRTK has the same characteristics as all MRTs. Patients with MRTK have an abnormal expression of the tumor suppressor gene SMARCB1/INI-1 (7–9). Moreover, the detection of SMARCB1 gene mutation has been associated with a worse prognosis (1). At the same time, cytogenetic studies have also found that most MRTK patients have chromosome 22 monomer deletion or 22q111.2 hSNF/INI1 gene mutation inactivation. This, in turn, leads to loss of INI-1 protein expression in the nucleus (10), which is of great significance as it can be differentiated from renal rhabdomyosarcoma.

MRTK lacks typical clinical manifestations. Most children develop abdominal masses, hematuria, or abdominal pain as the first symptoms. The tumor is often misdiagnosed as a Wilms tumor. B-scan ultrasonography can be used as a preliminary screening test, but it lacks specificity. Recent studies have found that CT and magnetic resonance imaging (MRI) can be useful for detecting early small tumors with subrenal hematoma/effusion (4, 11). Agrons et al. found that two-thirds of MRTK patients have dark crescent-shaped areas; the MRTK is unilateral and single, while Wilms Cell tumors can be bilateral or multiple (12). Preoperative imaging examination can be performed to distinguish Wilms tumor from renal blastocytoma. The final diagnosis needs to be confirmed by SMARCB1 gene detection or immunohistochemical INI-1 negative.

There are various signs, which indicate that the occurrence and development of MRTK are closely related to epigenetics (13). In this study, we used RNA-seq technology to investigate the transcription level of children with MRTK. The current report on MRTK mainly focused on the study of related proteins of traditional hSNF/INI1 and other genes, ignoring other information at the whole RNA level. The aim of this study was to screen new possible targets related to the occurrence and development of MRTK through global detection of RNA level, which can provide new ideas and theoretical basis for the prevention, diagnosis, and treatment of MRTK.

MATERIALS AND METHODS

Intraoperative Pathological Sample Acquisition

Intraoperative pathological tissues from 2 MRTK patients and 3 patients with non-tumor diseases without other symptoms were collected and analyzed after their family signed the informed consent. After identification by the pathology department, the normal tissue area in the intraoperative pathological tissue of children with non-tumor diseases without other symptoms was taken as a control. The central area of the tumor taken by MRTK children was used as the pathological sample of MRTK. After collection, save the sample in the RNA storage solution.

High-Throughput RNA-Sequencing (RNA-Seq) Analysis

The total RNA was extracted using TRIzol reagent. RNA-Seq was conducted according to a previously reported method (14). Briefly, after quality control and purification, the total RNA from each sample was used to enrich mRNA. The mRNA was fragmented and reverse-transcribed to cDNA. The cDNA was then purified, linked with the adaptor, size selected, and PCR amplified to obtain the cDNA library. After the quality check, libraries were sequenced through the Illumina platform.

Clean reads were obtained by using the preprocessing tool Trimmomatic (v0.36) to remove adapters and low-quality reads from the raw sequencing reads. Clean reads were mapped to Ensembl GRCh38 reference genome using the alignment program TOPHAT2. The mapped data were assembled into transcripts and quantified to obtain a matrix of expression values in FPKMusing StringTie v1.3.1c. Differentially expressed genes (DEGs) were screened out according to the logarithmic value of gene expression fold change (FC) between the PBS-treated cells and the OxLDL-treated cells. If logFC>1 or <1, and there was a significant statistical difference, it was judged as aDEGs. The distribution of the volcanic map of the DEGs was obtained by R software, and functional annotations were performed using Ensembl GRCh38 annotation files. Sequencing data have been uploaded from the GO database (https://www.ncbi.nlm.nih.gov/GO/query/acc.cgi?acc=GSE167547).

Gene Ontology (GO) Construction and Kyoto Encyclopedia of Genes and Genomes (KEGG) Pathway Analysis

To explore the potential interactions between all genes at the protein level, online retrieval tool GO (https://GO.org/) was used. The KEGG pathway analysis was performed using online biotool KEGG Mapper 83.0 (http://www.kegg.jp/kegg/mapper.html).

Statistical Analysis

The experimental data were analyzed using the R toolkit. A P-value <0.05 was considered to be statistically significant.

RESULTS

1. Case Description

Review of the Course of 2 Renal Malignant Rhabdoid Tumors

The study was approved by the Ethics Committee of Kunming Children's Hospital.

Case 1

Child 1 was a 1.8-year-old girl who was admitted to the hospital due to painless naked hematuria that lasted 5 days. The family of the child reported the presence of dark brown urine that lasted

for 5 days. The child had no frequent urination and no abdominal pain. On November 28, 2018, a urine routine examination at the first people's Hospital of Xundian City indicated increased red blood cells in the urine, after which the patient was transferred to our hospital. B-scan ultrasonography suggested the left lower kidney solid mass sonogram (considering nephroblastoma). Since the onset of hematuria (November 29, 2018), the patient had a normal stool, a slight increase in urine, and no significant body weight change.

Blood and fecal routine, biochemistry testing, coagulation function, and tumor markers were negative, including human chorionic gonadotropin (HCG), alpha-fetoprotein (AFP), and carcinoembryonic antigen (CEA). Urine red blood cells were 40355.9/UL, the cortisol (for measuring adrenal function) was 1262 nmol/L, and the adrenocorticotropic hormone was 305.6 pg/ml. Enhanced CT scan of the head, chest, and abdomen (**Figure 1A**) suggested a round-soft tissue density in the middle and lower part of the left kidney, CT about 40 HU, size of approx. 4.4 cm×3.9 cm×4.8 cm with a clear border. The enhancement scan showed mild to moderate uneven enhancement. There was no obvious abnormality in the head and lungs. The bilateral hip joint and sacrococcygeal X line were both normal.

Radical resection of the left renal tumor was performed under general anesthesia on December 7, 2018. The size of the renal region was about 7.5 cm×5.0 cm×4.5 cm cystic mass. The boundary between tumor and residual renal tissue was clear. There were adhesions between the tumor and the ipsilateral adrenal gland, peritoneum and inferior vena cava, and a varicose artery's nutritional tumor tissues in the renal hilum. The operation was successful, the operation time was 130 min, and the intraoperative bleeding was 60 ml without blood transfusion.

Under histological examination (**Figure 2**), the tumor cells showed patchy diffuse distribution. The tumor cells were round in shape, short fusiform, with a big nucleus and bubbles. Eosinophilic vitreous inclusion bodies were found in some cytoplasm. Nuclear fission and multiple necroses were also observed. No tumor components were found in 4 lymph nodes of the ureteric stump, perirenal fat, perirenal, and

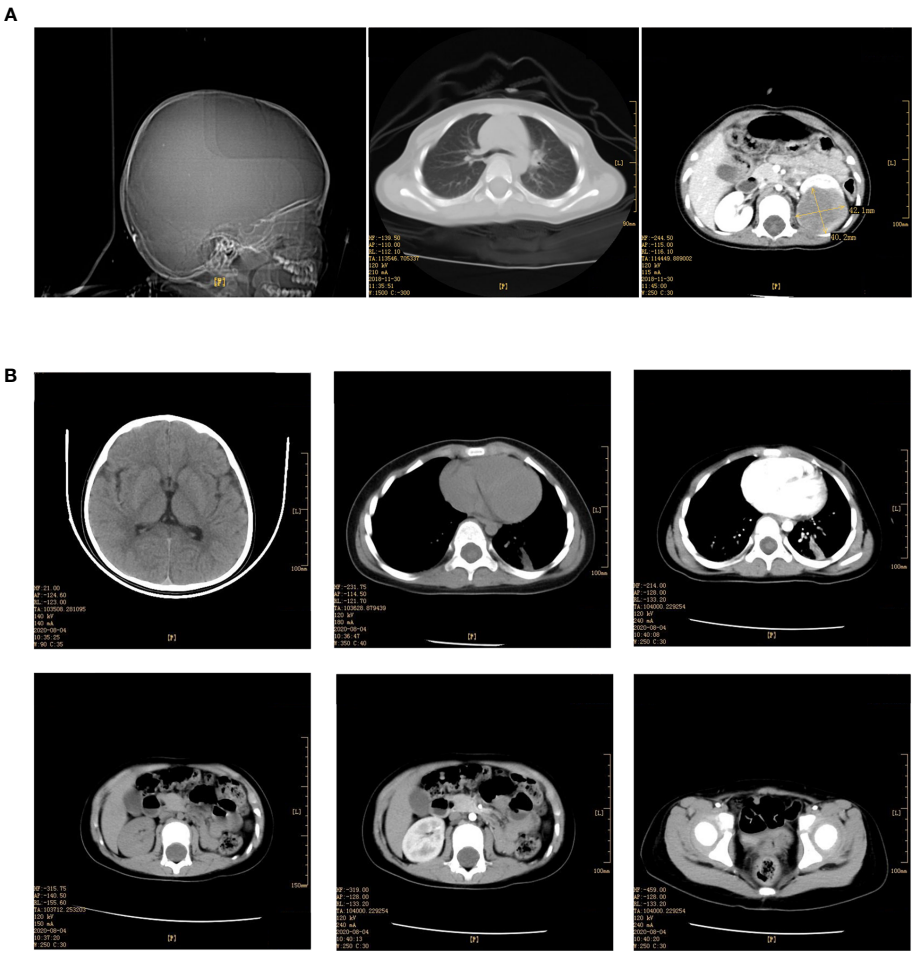


FIGURE 1 | **(A)** Enhancement of skull, chest, whole abdomen CT plain scan before the operation (December 03, 2018). **(B)** Enhancement of skull, chest, whole abdomen CT plain scan after the operation (August 04, 2020).

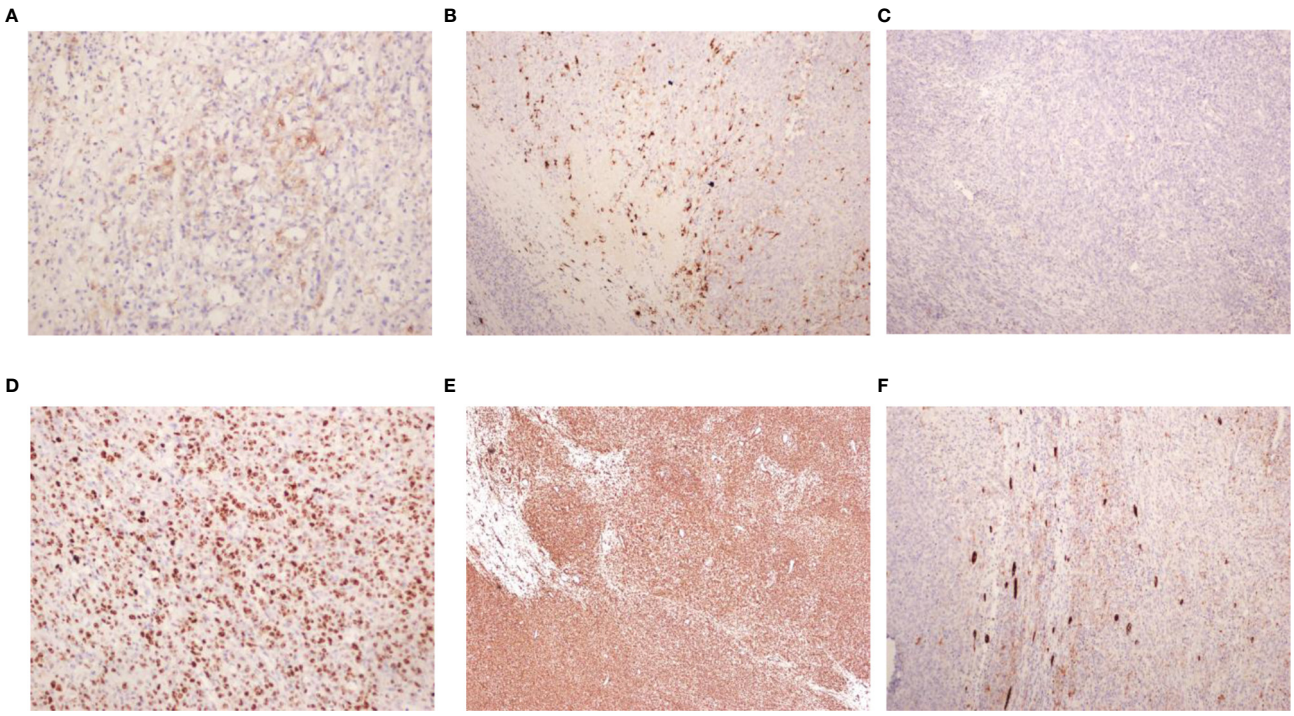


FIGURE 2 | Pathological findings. **(A)** EMA (x200); **(B)** S-100 (x100); **(C)** INI-1 (x100); **(D)** Ki-67 (x200); **(E)** Vim (x40); **(F)** CK (x100).

retroperitoneal. Immunohistochemistry results were the following: INI-1(-), CyclinD1(+), CK(+), EMA(+), S-100(+), Ki-67 hot spot 80%(+), Vimentin(+), Desmin(-), CD10(-), Myogenin(-), MyoD1(-), TFE3(-). The pathological results were: (left renal tumor) combined with immunohistochemical results. Thus, MRTK was confirmed. No tumor components were found in ureteral stumps, perirenal fat, and lymph nodes. According to the National Wilms Tumor Study of the American Association for the Study of Nephroblastoma NWTS Standards, it was considered as clinical stage II. The child recovered and was discharged 7 days after the operation. She received further treatment in Shanghai Fudan Children's Hospital: carboplatin 16.7".mg/(kg·d) on Day1 and Day2, etoposide [vp163.3 mg/(kg·d) on Day1-3, and cyclophosphamide [CTX14.7 mg/(kg·d) on Day1-5" regimen chemotherapy once a day. She received 7 rounds of chemotherapy in total. The treatment was completed on June 27, 2019. The most recent follow-up (August 4, 2020) suggested no metastasis (**Figure 1B**). At present, 2 years and 1 month after surgery, the patient continues to be asymptomatic.

Case 2

Child 2 was a 2.1 years old girl who was admitted to the hospital on November 29, 2017, due to painless naked hematuria that lasted for more than 10 days. B-scan ultrasonography suggested left renal mass; thus, nephroblastoma was suspected. The child's mental health, diet, sleep, and body weight were normal. No relevant physical examination and family history were reported. CT showed a large irregular mixed density (approx. 7.0 cm×8.6 cm×8.0 cm) in the left abdomen, which was further suggestive of nephroblastoma (**Figure 3A**).

The patient underwent radical resection of the left kidney on December 6, 2017. The left renal area was enlarged with solid mass (size:7.0 cm×8.6 cm×8.0 cm). The exposed part was smooth. The envelope was purplish-red and grayish-white; varicose veins were seen in the capsule; there was no obvious normal kidney tissue. Part of the macula adhered to the renal hilum and peritoneum, and varicose vein nutritional tumor tissue was visible. Along with the perirenal serosal space, the renal tumors and the upper ureter were isolated layer by layer and released the tumor body along this gap. The release included a total, complete resection of tumor tissue and ureteral stump, as well as perirenal fascia resection, abdominal para-aortic lymph node dissection. The operation was successful; intraoperative bleeding was 30 ml, and blood transfusion of 1U red blood cells was required. Operation time was 145 min.

Postoperative pathology indicated the following: the tumor cells were epithelioid, with some dense areas. The cytoplasm was vacuolated, nesting, and lamellar; dendritic blood vessels could be seen. No tumor invasion was found in the lymph nodes. Immunohistochemistry indicated FLI-1(+), CD99(+), Vimentin(+), CyclinD1(+), WT-1(+), EMA(+), Bcl2(-), Ki-67(+) 60%, CD56(-), CK(-), CD34(-), Desmin(-), and SMA(-), which suggested that renal clear cell sarcoma may not rule out renal cell carcinoma. The pathological sections were then sent to Beijing Children's Hospital for re-examination, which indicated the following: INI-1(-),

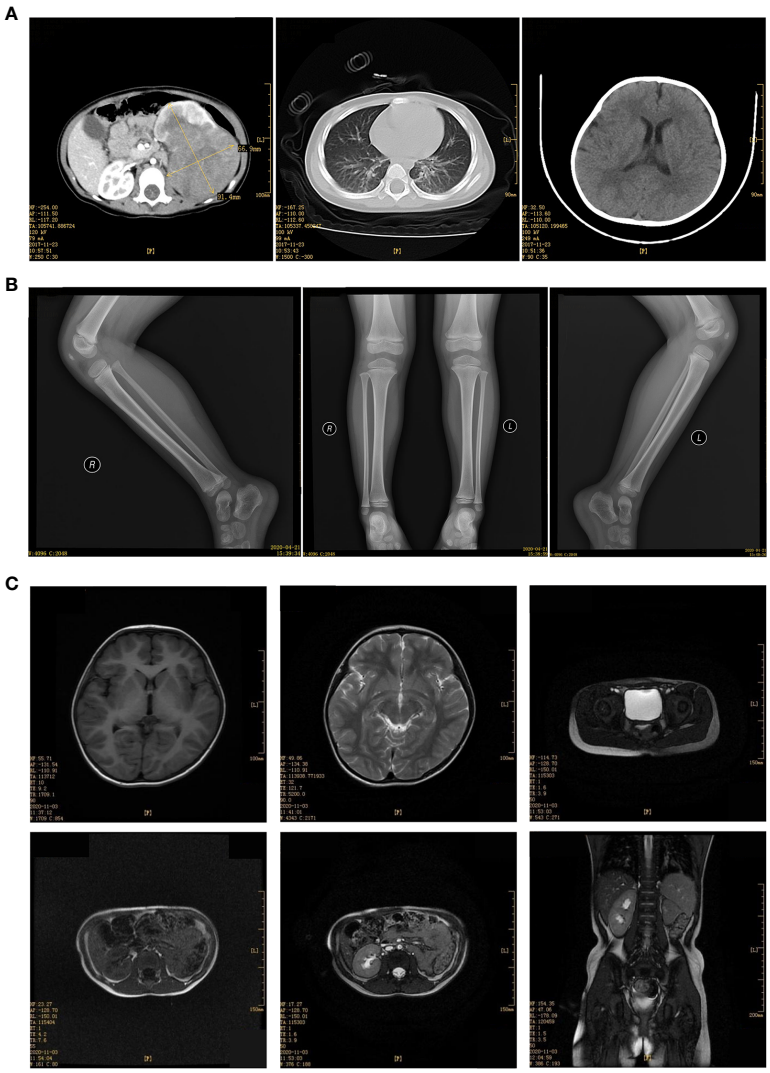


FIGURE 3 | **(A)** Enhancement of CT plain scan of chest and abdomen before the operation (December 29, 2017). **(B)** Bilateral hip joint X tablets after the operation (April 21,2020). **(C)** MRI scan of the head, chest, abdomen, and pelvis after the operation (December 07, 2020).

CyclinD1 (+), CK (+), EMA(+), and S-100 (+). The final diagnosis was MRTK.

The patient received the following treatment: Programme I "vincristine [VCR0.5mg/(kg·d) on day 1] + cyclophosphamide [CTX100 mg/(kg·d) on day 1-2] + actinomycin D [Act-D150μg/(kg· d) on day 1-5] + epirubicin [E-ADM 10 mg/(kg·d) on day 3-4]". Programme II "VCR0.5mg/(kg·d) on day 1] + carboplatin [Carbopla-tin16.7 mg/(kg·d) on day 2-3] + etoposide [vp16 30 mg/(kg·d) on day 1-5]"Alternate use of the two schemes, once a day. Chemotherapy began on December 20, 2017. On December 11, 2018, the head, chest, and abdomen were examined by MRI; the hip X line did not appear abnormal on April 21, 2018 (**Figure 3B**). During the most recent follow-up on December 7, 2020, MRI suggested no metastases of the head, chest, and total abdominal and pelvic (**Figure 3C**). Currently, 3 years after surgery, the patient continues to be asymptomatic.

2. Effect of the Gene Expression at the Transcript Level in Patients With MRTK

Differential Gene (DEGs) Screening and Analysis

The PCA analysis results (**Figure 4A**) showed that the overall expression of the transcriptome of MRTK pathological tissues was different compared with normal tissues. A total of 2203 DEGs were found (p-value <0.01); 1080 were up-regulated genes, and 1123 were down-regulated genes (**Figures 4B, D**). Heat map cluster analysis showed a big difference in gene expression between the normal tissues and MRTK pathological tissues. The INI1 protein corresponding gene SMARCB1 showed a significant downward trend, which is consistent with the pathology report.

GO Function Analysis

GO function analysis (**Figures 4E, F**) indicated that 5665 genes were involved in the biological process, 812 genes in cellular

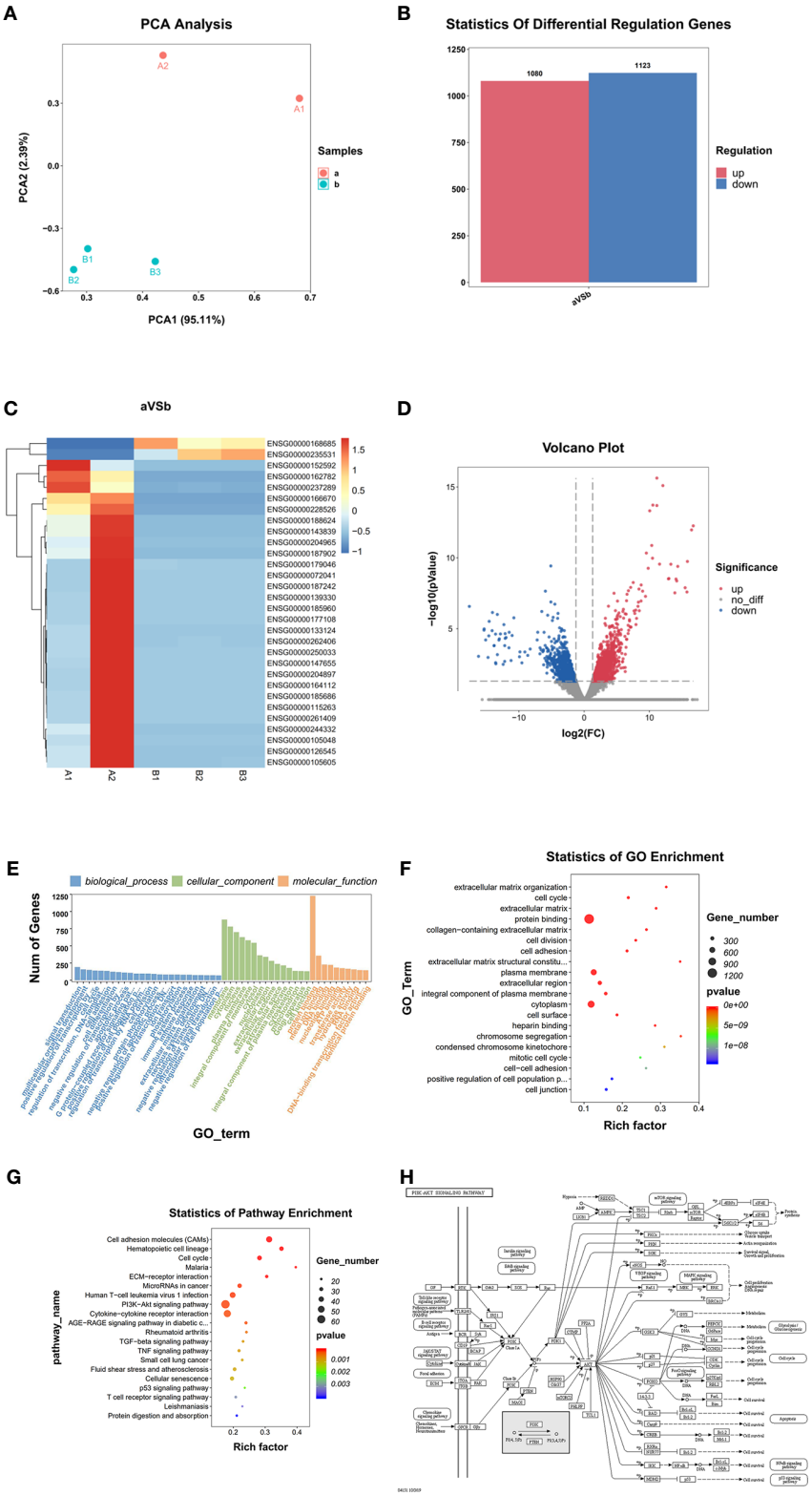


FIGURE 4 | Differential gene screening and analysis. **(A)** PCA Analysis; **(B)** Differential Regulation Genes. a: Tumor tissue, B: Normal tissues; **(C)** Heat map. A1, A2: Tumor tissue (Child 1 and Child 2), B1, B2, B3: Normal tissues; **(D)** Volcano Plot. Blur: Down regulation DEGs. Red: Up regulation DEGs; **(E, F)** GO Analysis, **(G)** Statistics of Pathway Enrichment; **(H)** PI3K-AKT signaling pathway (From: http://www.kegg.jp/kegg/mapper.html).

components, and 1565 genes in molecular functions. The genes involved in the protein binding function represented the highest group (1227 genes).

KEGG Signal Pathway Analysis

KEGG signaling pathway analysis results showed that DEGs involved in the PI3K-Akt signaling pathway (**Figure 4G**) were the most abundant, with 62 (**Figure 4H**), indicating that the PI3K-Akt signaling pathway has an important role in the occurrence and development of MRTK and microRNAs are also listed in the same position in cancer. The top 20 suggest that MicroRNA also has an irreplaceable role in the development of MRTK.

DISCUSSION

We believe that the above two points can be used to identify Wilms tumor during non-invasive preoperative examination. In patient 1, the tumor was 4.0 cm×3.5 cm in size. CT showed that the mass was homogeneous, which was consistent with the appearance of early small tumor foci. Also, no cyst-solid alternate dark areas characteristic of Wilms tumor were seen, which are the two main features found in postoperative pathological reports of children with renal clear cell sarcoma. The diagnosis was confirmed after consultation with the pathology department of Beijing Children's Hospital, thus suggesting that it is difficult to distinguish the disease only from the CCSK in pathology. Experienced pathologists need to be consulted to achieve an accurate diagnosis purpose of symptomatic treatment. In addition, the most common cause of death due to MRTK is tumor metastasis. In patients with MRTK, the most common metastatic sites are the lungs and brain, followed by abdominal organs such as the liver, sacrum, and hip joints (1). The two patients' preoperative CT showed no tumors in the lungs, brain, and liver. Moreover, X-rays of the hip and sacrococcyx showed no abnormalities. Considering that there was no metastasis, the patients were treated with radical nephrectomy.

Complete tumor resection is considered as the prerequisite and basis for long-term survival after surgery (15). It has been reported that the prognosis of children undergoing surgery after chemotherapy is worse than that of primary surgery (16). Thus, surgery should be performed as soon as possible after the discovery of MRTK. As long as the clinical metastasis is not confirmed, one or more operations should be performed as far as possible to remove all tumor tissues. If the tumor cannot be completely removed, high-dose chemotherapy and local gamma knife treatment can also improve the prognosis (17). Chemotherapy is also an important part of MRTK treatment. At present, the alternative chemotherapy regimen, including vincristine, adriamycin, cyclophosphamide, etoposide, and ifosfamide have been universally recognized by experts at home and abroad and have been suggested as a treatment for MRTK (18–20).

In this study, the two patients were 1.8 years old and 2.1 years old. No metastasis was found in the preoperative imaging examination. Both patients were treated by complete tumor resection directly. It is worth noting that patient 2 did not complete the entire chemotherapy. Radiotherapy was not added during the process, and the patient continues to be asymptomatic.

Standard chemotherapy is effective for children after MRTK. The aforementioned alternative chemotherapy regimens and doses can provide references for postoperative treatment of MRTK. Yet, it remains unclear whether there are individual differences in doses and whether chemotherapy supplemented with radiotherapy is effective for the prognosis of children with MRTK.

Compared with the normal tissue parts of children's kidneys removed from other non-tumor diseases, RNA-seq analysis showed that children with MRTK have 2203 differential genes (DEG); 1080 DEGs were up-regulated, and 1123 DEGs were down-regulated. Moreover, the GO analysis results showed that most DEGs were involved in protein binding (up to 1227 genes) (**Figures 4B, D**). This result suggests that many proteins involved in protein binding are included in the development of MRTK. Many DEGs were also involved in membrane composition, thus suggesting that cell membrane constituent proteins may be used as new therapeutic targets when treating patients with MRTK. These data suggested that the MRTK pathogenic genes and cancer-causing genes were down-regulated, which further proved the correctness of the previous diagnosis of the patient's condition, and also suggested that our subsequent research and analysis are true and reliable (**Table 1**).

Through the analysis of the KEGG signaling pathway, we found that up to 62 genes were involved in the PI3K-AKT signaling pathway (**Figure 4G**), which suggested that the PI3K-AKT signaling pathway has a vital role in the occurrence and development of MRTK (**Figure 4H**). More importantly, this is the first time that the signal pathway has been involved in MRTK. Various growth factors/ligands specific to fibroblast growth factors of different receptor tyrosine kinases (RTKs) can activate PI3-kinase (21). Growth factors mediate PI3K activation by stimulating RTK. Akt is the main mediator of the PI3K pathway. It interacts with PDK1 (phosphoinositide-dependent kinase 1) to cause Akt phosphorylation at breast cancer to play a role (22). A number of studies have shown that the PI3K/Akt pathway in human cancers, especially the expression of the two important genes PIK3CA and PTEN, usually significantly changes in human cancers, which have been found in more than 70% of tumor types (23, 24).

PI3K-Akt has a certain role in cell proliferation, apoptosis, migration, and other physiological activities. Its activity level is one of the key factors affecting the physiological activities of cancer cells. At present, more than 20 PI3K and Akt inhibitors have entered the stage of clinical trials (25). Further in-depth studies of the PI3K-Akt signaling pathway mechanism are necessary to gain a more comprehensive understanding of the formation and evolution of tumors to treat cancer diseases.

Among the other top 20 high-abundance signal pathways shown by KEGG, most of them are very common tumor signal pathways, such as cell cycle, calcium signal pathway, etc. (26–28).

TABLE 1 | MRTK pathogenic genes and cancer-causing genes.

gene_name	gene_id	transcript_id	Description	trans_type	FPKM.A1	FPKM.A2	FPKM.B1	FPKM.B2	FPKM.B3	fc	regulation	chr	start	end	strand
MRTK pathogenic genes															
SMARCB1	ENSG00000099956	ENST00000407422; ENST00000644036; ENST00000344921; ENST00000646421; ENST00000491967; ENST00000647057; ENST00000644619; ENST00000263121; ENST00000407082; ENST00000417137; ENST00000646723; ENST00000643421; ENST00000646911; ENST00000644462; ENST00000634926; ENST00000635578; ENST00000642727; ENST00000644467; ENST00000642275; ENST00000477836; ENST00000645799	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily b, member 1 [Source: HGNC Symbol;Acc: HGNC:11103]	protein_coding; retained_intron; nonsense_mediated_decay	1.225805 11878971	0.7662761 15400908	2.1732073 3276853	2.288547 27842663	1.8908610 2263238	0.470376 61705428	down	chr22	23786931	23838008	+
Cancer-causing genes(Cancer core genes (list only the top five)															
TACSTD2	ENSG00000184292	ENST00000371225	tumor associated calcium signal transducer 2 [Source: HGNC Symbol;Acc: HGNC:11530]	protein_coding	266.599854	0.189811	1.935009	10.726301	0.628289	30.112608 9282303	up	chr1	58575423	58577773	–
PTTG1	ENSG00000164611	ENST00000524244; ENST00000352433; ENST00000517480; ENST00000523659; ENST00000520452; ENST00000393964; ENST00000519287	pituitary tumor-transforming 1 [Source: HGNC Symbol;Acc: HGNC:9690]	retained_intron; protein_coding; processed_transcript	3.877970 38936709	21.484409 5782278	1.584697 19594937	1.8656379 7670886	0.8709661 15443038	8.803730 03755733	up	chr5	160421855	160428739	+
TPD52	ENSG00000076554	ENST00000379096; ENST00000448733; ENST00000518937; ENST00000517427; ENST00000520527; ENST00000379097; ENST00000517462; ENST00000519303; ENST00000523395; ENST00000523193; ENST00000523319; ENST00000521354; ENST00000520877; ENST00000521618; ENST00000518517; ENST00000524194; ENST00000521241; ENST00000517445; ENST00000523753; ENST00000522364; ENST00000519250; ENST00000518500; ENST00000520795; ENST00000520741; ENST00000521561; ENST00000523783; ENST00000523564;	tumor protein D52 [Source: HGNC Symbol;Acc: HGNC:12005]	protein_coding; nonsense_mediated_decay; processed_transcript; retained_intron	3.782222 04660242	3.626550 71721361	0.8107420 42210565	0.8657080 26595096	0.6478829 30828352	4.781225 04282902	up	chr8	80034745	80231172	–

(Continued)

TABLE 1 | Continued

gene_name	gene_id	transcript_id	Description	trans_type	FPKM.A1	FPKM.A2	FPKM.B1	FPKM.B2	FPKM.B3	fc	regulation	chr	start	end	strand
		ENST00000520035; ENST00000602950													
TUSC3	ENSG00000104723	ENST00000503191; ENST00000382020; ENST00000515859; ENST00000506802; ENST00000510836; ENST00000509380; ENST00000503731; ENST00000507400; ENST00000509177; ENST00000511783; ENST00000508446; ENST00000507316; ENST00000511342	tumor suppressor candidate 3 [Source: HGNC Symbol;Acc: HGNC:30242]	processed_transcript; protein_coding; nonsense_mediated_decay; retained_intron	15.89841 17990366	18.136619 1400771	6.344768 36088632	6.806862 92697495	5.821504 9495183	2.690780 57364551	up	chr8	15417215	15766649	+
TP53	ENSG00000141510	ENST00000413465; ENST00000635293; ENST00000359597; ENST00000504290; ENST00000504937; ENST00000510385; ENST00000610623; ENST00000618944; ENST00000619186; ENST00000610292; ENST00000269305; ENST00000620739; ENST00000617185; ENST00000420246; ENST00000455263; ENST00000610538; ENST00000622645; ENST00000445888; ENST00000619485; ENST00000576024; ENST00000615910; ENST00000509690; ENST00000514944; ENST00000574684; ENST00000505014; ENST00000508793; ENST00000604348; ENST00000503591; ENST00000571370	tumor protein p53 [Source: HGNC Symbol;Acc: HGNC:11998]	protein_coding; nonsense_mediated_decay; processed_transcript; retained_intron	9.997909 31997187	11.517389 2960619	3.878623 23136428	4.477840 09388186	4.3725784 0963432	2.535379 22148652	up	chr17	7661779	7687538	–

However, so far, no study has examined the changes in microRNA in MRTK cells. The KEGG results in this study suggest that many microRNA-related genes (up to 33) have changed in MRTK. Although most of them have been reported in other diseases, these microRNAs have never been deeply investigated in MRTK. Interestingly, these genes have been proven to be extremely safe and effective new targets with considerable potential in other tumors (29–31).

Although this study provides many new diagnostic and therapeutic targets with greater feasibility, further in-depth experimental research is still needed. We plan to further explore the new targets' feasibility and safety in subsequent clinical treatment and basic research.

Because children's MRTK accumulates throughout the whole kidney, it is difficult to obtain satisfactory paracancerous tissues. Therefore, it is possible to select the normal tissue part of the pathological kidney tissue in non-tumor patients with no other symptoms to replace the adjacent or normal kidney tissue only through pathology identification.

CONCLUSION

Our data suggest that the PI3K-Akt signaling pathway and microRNA-related proteins as targets have extremely high potential value for the diagnosis and treatment of MRTK.

LIMITATIONS AND STRENGTHS

This study has a few limitations. The sample size is too small to identify key genes in this signaling pathway associated with MRTK disease.

The research was done through high-throughput RNA sequence analysis, and the results and information were complete with high-quality. With the improvement of high-throughput sequencing, its higher recognition accuracy and sensitivity greatly reduced the error caused by too few experimental samples. Therefore, the evaluation of signal pathways in this study is representative.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: GEO Submission (GSE167547, link: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE167547).

ETHICS STATEMENT

The study was approved by the Ethics Committee of Kunming Children's Hospital. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin. Written informed consent was obtained from the minor(s)' legal guardian/next of kin for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

CZ and SC: study design, data collection and analysis, statistical analysis, and manuscript drafting. LL: study design, and data collection and analysis. JL: study design and data collection. HT, XH, and YX: data collection. ZY: manuscript revision. BY: study design and critical revision of the manuscript. LD: study design and manuscript revision. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the Yunnan Education Department of Science Research Fund (No.:2020 J0228), Kunming City Health Science and Technology Talent "1000" training Project (No.:2020- SW (Reserve)-112), Kunming Health and Health Commission Health Research Project (No.:2020-0201-001), Kunming Xishan District Science and Technology Project (No.:2020-Xike word 23), Kunming Medical Joint Project of Yunnan Science and Technology Department (No.:202001 AY070001-271).

ACKNOWLEDGMENTS

We would like to thank LL (Key Laboratory of Pediatric Major Diseases, Kunming Children's Hospital) and SC (Department of Pharmacology, Anhui Medical University) for optimization of experimental design and result analysis.

REFERENCES

1. Wang HM, Cheng HY. Current Situation and Prospect of Treatment of Malignant Rhabdomyoid Tumor in Children. *Chin J Pediatr Blood Cancer* (2018) 23(1):2–6. doi: CNKI:SUN:ZGXY.0.2018-01-002
2. Cai MX, Tang JY. Advances in the study of malignant rhabdomyoid tumors in children. *Int J Blood Transfusion Hematol* (2016) 39(02):169–73. doi: CNKI:SUN:GWSX.0.2016-02-018
3. Cheng HY, Yang S, Cai SY, Fu LB, Qin H, Ma XL, et al. Clinical and prognostic features of renal malignant rhabdomyosarcoma in 35 children. *Chin J Pediatr Blood Cancer* (2018) 23(03):118–23. doi: CNKI:SUN:ZGXY.0.2018-03-002
4. Furtwängler R, Nourkami-Tutibi N, Leuschner I, Niggli F, Kager L, Frühwald MC, et al. Malignant Rhabdoid Tumor of the Kidney (MRTK) – Data of 52 Patients Treated According to Protocols of the GPOH (German Society of Paediatric Oncology and Haematology). *Cancer Genet* (2014) 207(9):454–4. doi: 10.1016/j.cancergen.2014.09.032
5. Beckwith JB, Palmer NF. Histopathology and prognosis of Wilms tumors: results from the First National Wilms' Tumor Study. *Cancer* (1978) 41(5):1937–48. doi: 10.1002/1097-0142(197805)41:5<1937::aid-cncr2820410538>3.0.co;2-u

6. Haas JE, Palmer NF, Weinberg AG, Beckwith JB. Ultrastructure of malignant rhabdoid tumor of the kidney. A distinctive renal tumor of children. *Hum Pathol* (1981) 12(7):646–57. doi: 10.1016/s0046-8177(81)80050-0
7. Wang X, Lee RS, Alver BH, Haswell JR, Wang S, Mieczkowski J, et al. SMARCB1-mediated SWI/SNF complex function is essential for enhancer regulation. *Nat Genet* (2017) 49(2):289–95. doi: 10.1038/ng.3746
8. Wang W, Xue Y, Zhou S, Kuo A, Cairns BR, Crabtree GR. Diversity and specialization of mammalian SWI/SNF complexes. *Genes Dev* (1996) 10 (17):2117–30. doi: 10.1101/gad.10.17.2117
9. Hulsebos TJ, Plomp AS, Wolterman RA, Robanus-Maandag EC, Baas F, Wesseling P. Germline mutation of INI1/SMARCB1 in familial schwannomatosis. *Am J Hum Genet* (2007) 80(4):805–10. doi: 10.1086/513207
10. Uno K, Takita J, Yokomori K, Tanaka Y, Ohta S, Shimada H, et al. Aberrations of the hSNF5/INI1 gene are restricted to malignant rhabdoid tumors or atypical teratoid/rhabdoid tumors in pediatric solid tumors. *Genes Chromosomes Cancer* (2002) 34(1):33–41. doi: 10.1002/gcc.10052
11. Ahmed HU, Arya M, Levitt G, Duffy PG, Mushtaq I. Sebire NJ. Part I: Primary malignant non-Wilms' renal tumours in children. *Lancet Oncol* (2007) 8 (8):730–7. doi: 10.1016/s1470-2045(07)70241-3
12. Agrons GA, Kingsman KD, Wagner BJ, Sotelo-Avila C. Rhabdoid tumor of the kidney in children: a comparative study of 21 cases. *AJR Am J Roentgenol* (1997) 168(2):447–51. doi: 10.2214/ajr.168.2.9016225
13. Tang J, Xie Y, Xu X, Yin Y, Jiang R, Deng L, et al. Bidirectional transcription of Linc00441 and RB1 via H3K27 modification-dependent way promotes hepatocellular carcinoma. *Cell Death Dis* (2017) 8(3):e2675. doi: 10.1038/cddis.2017.81
14. Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. *Bioinf (Oxford England)* (2009) 25(9):1105–11. doi: 10.1093/bioinformatics/btp120
15. Martini N, Rusch VW, Bains MS, Kris MG, Downey RJ, Flehinger BJ, et al. Factors influencing ten-year survival in resected stages I to IIIa non-small cell lung cancer. *J Thoracic Cardiovasc Surg* (1999) 117(1):32–6; discussion 7-8. doi: 10.1016/s0022-5223(99)70467-8
16. D'Angio GJ, Breslow N, Beckwith JB, Evans A, Baum H, deLorimier A, et al. Treatment of Wilms' tumor. Results of the Third National Wilms' Tumor Study. *Cancer* (1989) 64(2):349–60. doi: 10.1002/1097-0142(19890715)64:2<349::aid-cncr2820640202>3.0.co;2-q
17. Wang JH, Cai JB, Li MJ, Shu Q. Interpretation of International and Domestic Diagnosis and Treatment of Renal Cytoma in Children. *J Clin Pediatr Surg* (2020) 19(09):765–74. doi: CNKI:SUN:LCXR.0.2020-09-003
18. Hu HM, Zhang WL, Huang DS, Zhang Y, Han T, Zhi T, et al. Clinical analysis and literature review of 3 cases of malignant rhabdomyosarcoma in children. *Chin J Pediatr Blood Cancer* (2013) 18(06):274–6. doi: CNKI:SUN:ZGXY.0.2013-06-015
19. Yamamoto M, Suzuki N, Hatakeyama N, Mizue N, Hori T, Kuroiwa Y, et al. Treatment of stage IV malignant rhabdoid tumor of the kidney (MRTK) with ICE and VDCy: a case report. *J Pediatr Hematol/Oncol* (2006) 28(5):286–9. doi: 10.1097/01.mph.0000212901.84146.5a
20. Green DM, Breslow NE, Beckwith JB, Ritchey ML, Shamberger RC, Haase GM, et al. Treatment with nephrectomy only for small, stage I/favorable histology Wilms' tumor: a report from the National Wilms' Tumor Study Group. *J Clin Oncol Off J Am Soc Clin Oncol* (2001) 19(17):3719–24. doi: 10.1200/jco.2001.19.17.3719
21. Lemmon MA, Schlessinger J. Cell signaling by receptor tyrosine kinases. *Cell* (2010) 141(7):1117–34. doi: 10.1016/j.cell.2010.06.011
22. Fedele CG, Ooms LM, Ho M, Vieusseux J, O'Toole SA, Millar EK, et al. Inositol polyphosphate 4-phosphatase II regulates PI3K/Akt signaling and is lost in human basal-like breast cancers. *Proc Natl Acad Sci USA* (2010) 107 (51):22231–6. doi: 10.1073/pnas.1015245107
23. Gojo I, Perl A, Luger S, Baer MR, Norsworthy KJ, Bauer KS, et al. Phase I study of UCN-01 and perifosine in patients with relapsed and refractory acute leukemias and high-risk myelodysplastic syndrome. *Invest New Drugs* (2013) 31(5):1217–27. doi: 10.1007/s10637-013-9937-8
24. Pons-Tostivint E, Thibault B, Guillermet-Guibert J. Targeting PI3K Signaling in Combination Cancer Therapy. *Trends Cancer* (2017) 3(6):454–69. doi: 10.1016/j.trecan.2017.04.002
25. McCubrey JA, Steelman LS, Abrams SL, Lee JT, Chang F, Bertrand FE, et al. Roles of the RAF/MEK/ERK and PI3K/PTEN/AKT pathways in malignant transformation and drug resistance. *Adv Enzyme Regul* (2006) 46:249–79. doi: 10.1016/j.advenzreg.2006.01.004
26. Nagata T, Takahashi Y, Ishii Y, Asai S, Sugahara-Kobayashi M, Nishida Y, et al. Molecular genetic alterations and gene expression profile of a malignant rhabdoid tumor of the kidney. *Cancer Genet Cytogenet* (2005) 163(2):130–7. doi: 10.1016/j.cancergencyto.2005.05.009
27. Parkash J, Asotra K. Calcium wave signaling in cancer cells. *Life Sci* (2010) 87 (19-22):587–95. doi: 10.1016/j.lfs.2010.09.013
28. Monteith GR, Davis FM, Roberts-Thomson SJ. Calcium channels and pumps in cancer: changes and consequences. *J Biol Chem* (2012) 287(38):31666–73. doi: 10.1074/jbc.R112.343061
29. Gandellini P, Profumo V, Folini M, Zaffaroni N. MicroRNAs as new therapeutic targets and tools in cancer. *Expert Opin Ther Targets* (2011) 15 (3):265–79. doi: 10.1517/14728222.2011.550878
30. Nana-Sinkam SP, Croce NSM. MicroRNAs as therapeutic targets in cancer. *Trans Res* (2011) 157(4):216–25. doi: 10.1016/j.trsl.2011.01.013
31. Soriano A, Jubierre L, Almazán-Moga A, Molist C, Roma J, de Toledo JS, et al. microRNAs as pharmacological targets in cancer. *Pharmacol Res* (2013) 75:3–14. doi: 10.1016/j.phrs.2013.03.006

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

REVIEW
published: 01 April 2021
doi: 10.3389/fonc.2021.620831

In Vivo and *Ex Vivo* Pediatric Brain Tumor Models: An Overview

Zhiqin Li and Sigrid A. Langhans *

Nemours Biomedical Research, Alfred I. duPont Hospital for Children, Wilmington, DE, United States

After leukemia, tumors of the brain and spine are the second most common form of cancer in children. Despite advances in treatment, brain tumors remain a leading cause of death in pediatric cancer patients and survivors often suffer from life-long consequences of side effects of therapy. The 5-year survival rates, however, vary widely by tumor type, ranging from over 90% in more benign tumors to as low as 20% in the most aggressive forms such as glioblastoma. Even within historically defined tumor types such as medulloblastoma, molecular analysis identified biologically heterogeneous subgroups each with different genetic alterations, age of onset and prognosis. Besides molecularly driven patient stratification to tailor disease risk to therapy intensity, such a diversity demonstrates the need for more precise and disease-relevant pediatric brain cancer models for research and drug development. Here we give an overview of currently available *in vitro* and *in vivo* pediatric brain tumor models and discuss the opportunities that new technologies such as 3D cultures and organoids that can bridge limitations posed by the simplicity of monolayer cultures and the complexity of *in vivo* models, bring to accommodate better precision in drug development for pediatric brain tumors.

Keywords: medulloblastoma, glioma, pediatrics, preclinical models, *in vivo* models, *in vitro* models, cancer

OPEN ACCESS

Edited by:
Jing He,
Guangzhou Medical University, China

Reviewed by:
Moise Danielpour,
Cedars Sinai Medical Center, United States
Xinhui Du,
Zhengzhou University, China

***Correspondence:**
Sigrid A. Langhans
sigrid.langhans@nemours.org

Specialty section:
This article was submitted to Pediatric Oncology, a section of the journal Frontiers in Oncology

Received: *23 October 2020*
Accepted: *15 March 2021*
Published: *01 April 2021*

Citation:
Li Z and Langhans SA (2021) In Vivo and Ex Vivo Pediatric Brain Tumor Models: An Overview. Front. Oncol. 11:620831. doi: 10.3389/fonc.2021.620831

INTRODUCTION

Brain tumors are the most common solid tumors and the leading cause of cancer-related death in children. The incidence and mortality rate of primary brain and other central nervous system tumors have not changed significantly in recent years, with an average incidence rate of 5.65 per 100,000 population and an average mortality rate of 0.72 per 100,000 population for the 0 to 14 years age group from 2011 to 2015 in the United States (1). In the past, the diagnosis and classification of brain tumors had largely relied on histological characteristics derived from hematoxylin and eosin-staining, and immunohistochemical detection of lineage-associated proteins. However, more and more evidence shows that histologically similar brain tumors sometimes have distinct molecular features; they respond differently to the treatment and have various prognosis as well. In addition, some histologically ambiguous tumors may largely rely on their molecular characterization for their diagnosis and treatment plan. In 2016, the World Health Organization (WHO) updated classification of central nervous system tumors by incorporating molecular features into traditional histological characteristics for more accurate diagnosis, prognosis predictions, and treatments (2–6). With the overall success rate of new anticancer drugs remaining low (7, 8), we will need to switch from "one size fits all" treatments to more specific individualized strategies, to increase treatment efficacy, to reduce complications due to treatment, and to improve the translation rate of anti-cancer drugs. Brain

tumor models that can mimic tumor initiation and progression, and predict a tumor's response to treatments *in vivo* are fundamental to achieve this goal. In this review, we provide an overview of the most common pediatric brain tumors and currently available well-annotated *in vitro* and *in vivo* tumor models. We also discuss the advantages and limitations of each model, which need to be considered when choosing an appropriate tumor model that best suits the experimental purpose.

Common Pediatric Brain Tumors and Molecular Subgrouping

Gliomas

Glioma is the most common pediatric primary brain tumor, representing approximately 47% of brain tumor cases in the age group of 0-19 years. Glioma can originate from all glia cell types and 75% of these glial tumors are astrocytoma (1). Glioma are highly heterogeneous tumors, ranging from low-grade glioma (LGG) to high-grade glioma (HGG) depending on the tumor malignant status.

LGG is the most common glioma, which is typically nonmalignant and slow growing. Histologically, LGGs include pilocytic astrocytoma (PA), pilomyxoid astrocytoma (PMA), oligoastrocytoma, subependymal giant cell astrocytoma (SEGA), pleomorphic xanthoastrocytomas (PXA), oligodendroglioma, gangliogliomа, dysembryoplastic neuroepithelial tumors, etc., among which pilocytic astrocytoma is the most common form. The aberrant Ras-mitogen activated protein kinase (MAPK) signaling pathway is mainly reported in LGG. The mutations usually occur at BRAF in this pathway, including the KIAA1549-BRAF fusion and BRAF V600E mutant, which lead to constitutive activation of the MAPK pathway. Furthermore, kRAS, FGFR1, MYB/MYBL1, NTRK2, NF1, TSC1/2 and other genetic alteration have also been identified in pediatric LGG. Unlike adult LGG, IDH mutations are almost absent in children (3–6, 9–11). In some cases, the molecular alteration is associated with a specific tumor type. For example, KIAA1549- BRAF fusion is mostly found in pilocytic astrocytoma (PA), while BRAF V600E is frequently detected in pleomorphic xanthoastrocytoma (PMA) and gangliogliomas (12).

HGG is relatively uncommon in pediatric glioma, accounting for around 20% of cases. However, HGGs are diffusely infiltrating malignant tumors and they are usually aggressive with an overall very poor prognosis; some patients succumb to the tumor within one year after diagnosis. Based on distinct histological and radiological features, HGG is subclassified into anaplastic astrocytoma, diffuse intrinsic pontine glioma (DIPG) and glioblastoma multiforme (GBM). Mutations in histone genes were first discovered in pediatric HGGs, and now serve as a hallmark of this glioma type. Histone mutations often vary according to HGG locations. In tumors arising from the midline and pons, K27M mutations in H3F3A (encoding histone H3.3) or HIST1H3B/C (encoding histone H3.1) are very common, which lead to a global decrease of H3 K27 trimethylation by inhibiting polycomb repressive complex 2 (PRC2) activity through sequestration of its catalytic subunit EZH2; while G34R (or rarely G34V) mutations in H3F3A (encoding histone H3.3) are mostly reported in hemispheric HGGs. In addition, the RTK/RAS/PI3K pathway (e.g., PDGFRA, PIK3CA, PIK3R1, or PTEN) and the p53/Rb pathway (e.g., TP53, CDKN2A, CDK4/6, CCND1-3) are also dysregulated in pediatric HGG (3, 4, 9–11, 13–15). Recent studies discovered some overlap in molecular profiling between LGG and HGG. BRAF V600E and FGFR1 mutations are found both in LGG and HGG (9, 10), which suggests that LGG and HGG might share a similar biological mechanism of tumor pathogenesis.

Ependymal Tumors

Ependymomas represent 5.5% of all pediatric primary brain tumor cases in the age group of 0 to 14 (1). Ependymomas are thought to originate from radial glia cells of the ependymal lining of the ventricles and the central canal. Histologically, ependymomas are classified into 4 groups: subependymoma, myxopapillary ependymoma, classic ependymoma, and anaplastic ependymoma, of which classic and anaplastic ependymoma are the most common subtypes in children. Classic ependymoma is further subclassified into 3 subtypes: papillary, clear cell, and tanycytic ependymoma based on their histological features (4, 5).

The molecular characteristics of ependymoma is usually associated with its location. Over 90% of pediatric ependymomas arise in the infratentorial and supratentorial regions. The infratentorial posterior fossa (PF) ependymomas are generally subclassified into Group A (PF-EPN-A) and Group B (PF-EPN-B) based on their DNA methylation profiling. PF-EPN-A tumors are hypermethylated, and mostly found in infants and young children, who have a poorer outcome compared to those with PF-EPN-B tumors, which are typically seen in adolescents and adults. Supratentorial (ST) ependymomas in children have two major subgroups: RELA fusion-positive (ST-EPN-RELA) ependymoma and YAP1 fusion-positive (ST-EPN-YAP1) ependymoma. ST-EPN-RELA usually harbors the fusion protein of C11orf95 and RELA, which constitutively activates the NF-κB pathway by enriching a RELA-encoded transcription factor p65. In ST-EPN-YAP1 ependymoma, transcriptional coactivator YAP1 fuses with other genes such as MAMLD1 and FAM118B and can upregulate Notch signaling. Compared to ST-EPN-YAP1, ST-EPN-RELA is more frequently observed in children and has worse prognosis. The major treatment plan for ependymoma is surgical resection plus adjuvant radiological therapies. Benefits of chemotherapy have not been reported, yet (3–6, 16–18).

Medulloblastoma

Medulloblastoma is the most common pediatric embryonal tumor originating from precursor cells in the cerebellum or dorsal brainstem. Like other embryonal tumors, medulloblastoma is highly proliferative and predisposed to metastasis. Histologically, medulloblastoma is classified into four different types: classic, desmoplastic/nodular, extensive nodularity and large cell/anaplastic. Medulloblastoma is one of the most heterogeneous brain tumors and currently has the best characterized molecular features. There are four distinct subgroups: wingless/integrated (WNT), sonic hedgehog (SHH), Group 3, and Group 4, which have different genetic alterations, phenotypes and prognostics.

The WNT subgroup accounts for around 10% of medulloblastoma; it mostly occurs in older children. The most common mutation in this subgroup is in the *CTNNB1* gene, which encodes β-catenin, a major player in cell cycle control and embryogenesis. The overexpression of nuclear β-catenin is often used as a diagnostic indicator in this subgroup. Monosomy chromosome 6 is another hallmark of the WNT subgroup, occurring in ~80-85% of patients, usually in conjunction with *CTNNB1* mutations. *DDX3X, SMARCA4*, and *TP53* mutations also have been reported in WNT-activated medulloblastomas (3–6, 19).

The SHH subgroup represents approximately 30% of medulloblastoma. SHH-activated medulloblastoma is highly heterogeneous and many key molecules in the SHH signaling pathway such as SUFU, smoothened (SMO), PTCH1, GLI1 and GLI2 have been dysregulated in this subgroup. Besides those, other genetic aberrations, like MYCN amplification or TP53 mutation, are also involved in SHH-activated medulloblastoma formation (19, 20). The outcome varies in SHH-activated medulloblastoma and although metastasis is not common, if a patient has a metastatic tumor, the outcome is usually worse. Moreover, patients with TP53 mutations or MYCN amplification usually have a poorer prognosis. The ongoing therapies using small molecule inhibitors target almost all affected molecules in the SHH pathway (3–6, 13, 19).

Group 3 composes around 25% of medulloblastoma. It is the most aggressive form and metastasis is very common in this subgroup. Unlike WNT- and SHH-activated medulloblastoma, the Group 3 tumors are less defined; some studies showed MYC amplification leading to tumor formation in this subgroup. Other possible pathways, such as TNFβ, have been found in around 20% of Group 3 medulloblastoma. The prognosis is overall poor for this subgroup, especially for patients with MYC amplification (4–6).

Group 4 is the most prevalent subgroup, comprising approximately 35% of medulloblastoma. Like Group 3, it has not been biologically characterized. The loss of chromosome 8, 11 and 17p or gain of chromosome 7 and 17q have been identified in this subgroup. In addition, amplification of CDK6, MYCN and SNCAP1 as well as aberrant ERBB4-SRC signaling and nuclear factor kappa B (NF-κB) have also been observed in Group 4 medulloblastoma (3–6, 19, 21, 22).

In Vivo Brain Tumor Models

Most animals rarely develop spontaneous brain tumors (23) and they can be used to generate experimental models for brain tumor studies (24). A good animal model should have high incidence rate, can recapitulate original tumor's histopathological and molecular features, and can manifest the human response to drug treatment. Numerous animal brain tumor models have been developed so far. These models can be used to investigate biological mechanisms of brain tumors and their microenvironment and for preclinical testing of novel, promising therapeutic regimens. To date, most animal models are generated with rodents, and in this review we will focus on rat and mouse models of brain tumors and discuss zebrafish brain tumor models. There are three major methods to generate animal models in brain tumor research: carcinogen induced animal models, xenograft animal models and genetically engineered animal models (23–26).

Carcinogen-Induced Brain Tumor Models

Rats are widely used when generating a carcinogen-induced brain tumor model, since the tumor induction in rat strains is much more effective than in mice (23). The most common carcinogens used to generate animal brain tumors are chemical carcinogens and viruses. N-nitrosourea and its derivatives have been reported to induce most common gliomas in rats, including astrocytoma, oligodendroglioma, and ependymal tumors. The embryos are much more susceptible to the chemical carcinogens, and transplacental injection is often used to administer the chemical compounds to pregnant animals (24). Injection of ethylnitrosourea to pregnant rats at gestational day 20 induced brain tumors in all of 25 pups born (27). Chemical carcinogens can also be applied to rodents through oral, intravenous or local exposure after they are born, but repeated administration might be necessary to increase induction efficacy, especially when working with older animals (23–26). The cell lines established from these chemical induced glioma models include C6, 9L, T9, F98, RG2, BT4C and CNS-1 and have been widely used in brain tumor studies (28–31). In addition to chemical carcinogens, oncogenic viruses may also be used to induce brain tumors. Both RNA viruses, such as Rous sarcoma virus-1 (RSV-1) and DNA viruses, such as adenovirus can induce brain tumors. Intracerebral injection of RSV caused malignant brain tumors in newborn pups (32). Different injection sites caused distinct tumor types (33). It has also been reported that injection of human adenovirus 12 virus (AD12) into mouse brain induced medulloblastoma or glioblastoma (34). These chemical carcinogens and oncogenic viruses are prevalent in the human environment; thus, this model can imitate natural tumorigenesis especially when animals are exposed in early development. The induced tumors can be continuously passed in animals and retain relatively stable biological characteristics. However, carcinogen induced tumor models lack consistency in tumor types, locations and biological characteristics. Moreover, the induced brain tumors are histologically and biologically different from human tumors.

Xenograft Models of Human Brain Tumors

Xenograft models are usually made by transplanting established cancer cell lines or brain tumor tissues derived from patients (patient-derived xenograft, PDX) or animal models into host animals. The established cancer cell lines grow very fast *in vitro* with well-defined biological characteristics, which makes them applicable to generate xenograft models. The cell lines generated from carcinogen-induced rodent tumors or from transgenic mice can be cultured and transplanted into syngeneic hosts with competent immune system (35). However, the cancer cell lines are a homogenous population lacking tumor heterogeneity and the induced tumors are rarely infiltrative. Moreover, cancer cell lines will gradually lose the original tumor phenotypes and genetic features during *in vitro* culture. Patient tissues can also be dissociated and cultured in neurobasal serum-free medium. This selects highly tumorigenic subpopulations with stem cell-like characteristics that can be grown as neurospheres before

implantation into host mice (23, 24, 35, 36). In addition, the tumor tissues derived from patients can be directly transplanted into recipient animals without *in vitro* culture. Engraftments grown in these animals include tumor tissues as well as their surrounding stroma in early passage. They retain histological and molecular characteristic of original tumors, interaction between tumor and host, and a tumor's responses to drug treatment. With this they are a more representative and reliable *in vivo* brain tumor model than those generated from cultured cells (23, 25).

In most cases to generate PDX, host animals are immunodeficient mice. The early xenografts were transplanted into nude mice, which are the first generation of immunodeficient mice. Nude mice not only lack body fur but also have no thymus. Thus, these mice have a defective adaptive immune response as they do not have T lymphocytes. Nevertheless, they still have functional B and NK cells, and an intact innate immune response causes a low engraftment rate in these mice. Later, the severe combined immunodeficient (SCID) mice that lack both functional T and B lymphocytes were generated. The engraftment efficacy has improved on SCID mice, but these mice still have remnant NK cells, hindering the engraftment rate. To eliminate the effect of NK cells, SCID mice were crossbred with Beige mice to establish SCID/Beige mice that have severely impaired NK cells and macrophages, and no mature T and B lymphocytes. SCID/Beige mice display a better engraftment rate, leading to more feasible PDX models. Since then, more immunodeficient mice strains have been established to improve engraftment and increase the success rate of PDX, such as non-obese diabetic (NOD)/SCID mice and its derivative mice (NOG, NSG and NOJ), and BALB/c background immunocompromised mice (BRG and BRJ) (37, 38). Different immunocompromised mouse strains have various sensitivity to chemotherapy or radiation, which needs to be considered when choosing an appropriate animal model. For example, BALB/c mice are very sensitive to radiation and SCID mice are sensitive to γ-irradiation and thus are not useful for radiotherapy related studies (39). Immunodeficient mice can also be modified by receiving human bone marrow to reconstitute a human immune response. These humanized mice provide an opportunity to even more closely recapitulate human brain tumors, to study the effect of the immune system on brain tumor pathogenesis, and to evaluate immunotherapies (24, 38, 39).

Xenografts can be administrated in two different ways: heterotopic xenograft and orthotopic xenograft. Heterotopic xenografts, which most typically are achieved through subcutaneous injection, are a popular method in cancer research. They are simple and convenient to observe and monitor tumors and to evaluate drug efficacies by measuring the tumor volume. However, the microenvironment of tumors induced in this way is different from the original tumor and it cannot faithfully recapitulate the original tumor initiation and progression. In addition, there is no blood brain barrier around these subcutaneous tumors, so this model cannot accurately reflect the anti-cancer drug efficacies. Orthotopic xenografts usually apply tumor cells/or tissues to the location where the original tumor is found in patients. Orthotopic xenografts can better mimic the original tumor pathogenesis, retain histological and molecular characteristics of original tumors as well as tumor host interactions (23, 25, 38, 39). However, even orthotopic xenografts may not completely maintain the histological characteristics of human tumors. Some intracranial glioblastoma xenograft models lack necrotic features and fail to show endothelial proliferation (40). To date, most available pediatric brain tumor PDX models represent glioblastoma, diffuse midline glioma, ependymoma, and medulloblastoma. The establishment of PDX models for less aggressive brain tumors, such as pilocytic astrocytoma, has been less successful due to a very low tumor engraftment rate (39).

The development of pediatric brain tumor PDX models emerged over thirty years ago (24, 38). In recent years, with the raised interest in some pediatric brain tumor types and increased availability of tumor tissues, more and more PDX models have been generated. In 2018, a brain tumor biology study sponsored by the Children's Oncology Group led to generation of 30 orthotopic pediatric brain tumor PDX models, including medulloblastoma, high grade glioma and ependymoma. These PDX models are valuable tools to investigate subtype specific pediatric brain tumors, since they preserve the original tumors' histological and molecular features and remain relatively stable when being passaged in mice (41). The scientists from St. Jude Children's Research Hospital also successfully generated 37 novel orthotopic PDX models derived from pediatric brain tumor patients including 22 medulloblastomas and 5 ependymomas, which also maintain original tumors' histological features and are genetically faithful to corresponding patient tumors (42). The Mayo Clinic Brain Tumor Patient-Derived Xenograft (PDX) National Resource has also established a repository of glioblastoma PDX models with highly characterized molecular subtype and phenotype. Another study collected DIPG samples from patient autopsies and biopsies at 8 different international institutions and generated 22 *in vivo* xenograft models, covering the main molecular subtypes including H3.3 K27M and H3.1 K27M mutations (43, 44).

With the advance in gene editing technology, neural stem cells (NSC) can be genetically engineered to acquire tumorigenic capability and used to generate xenograft models (35). Transplantation of NSCs that overexpressed *myc* alone, or with oncogene *gfi1* or *gfi1b* into the cerebella of immunocompromised mice induced Group 3 medulloblastoma (45–48). Funato et al. successfully transformed neural progenitor cells derived human embryonic stem cells with a constitutively active form of the PDGFRA, a small hairpin RNA (shRNA) against p53 and H3.3K27M to model pediatric DIPG with H3.3K27M mutation (49). In addition, NSCs co-expressing PDGFRB and H3.3K27M were injected into the pons of SCID mice and induced tumors similar to human H3K27M DIPGs (50). The first mouse model of ependymoma was generated by implanting embryonic cerebral *Ink4a/Arf*−/− NSCs overexpressing Ephb2 into the cerebrum of immunocompromised mice (51). More recently, induced pluripotent stem (iPS) cells are also being used to generate brain tumor xenograft models. iPS-derived neural stem cells generated from Gorlin syndrome patients, who are carrying a germline mutation in PTCH1 and are predisposed to medulloblastoma, were transplanted into mouse cerebellum. These

cells formed tumors that mimic SHH-driven medulloblastoma (52, 53).

Xenograft models are a valuable tool in cancer research and drug screening. The National Cancer Institute recently decided to use PDX models to replace a panel of 60 human cancer cell lines (NCI-60) as a model for drug screening (54). However, there are some limitations of xenograft models. First, the generation of some xenograft models is challenging, but the success rate is increased with more aggressive and highly malignant tumors. Second, xenografts usually require many cells at a time, which is not naturally occurring in patients. Third, the transplantation procedure can disrupt the blood brain barrier, which is a key factor when evaluating drug efficacy. Fourth, the host animal for xenografts are usually immunodeficient mice, which cannot be used to discover the contribution of immune system in tumor initiation and development. Fifth, the engrafted human tumor stroma structure will be lost over time, replaced by the host mice's own microenvironment. Sixth, genetic and phenotypic drifts gradually occur as the xenograft tumors are propagated through mice. Last, maintenance of PDX models is costly and labor intensive (22–25, 35, 38, 39, 55).

Mouse Models With Genetic Engineering of Brain Tumors

In recent years, with rapid advances in gene editing techniques, genetically engineered mouse models (GEMMs) have gained popularity in brain tumor research. Unlike PDXs, GEMMs can recapitulate tumor initiation and development in animals with native immune system and intact blood brain barrier and undisrupted microenvironment. This makes GEMMs more attractive as models for tumor mechanism and drug discovery studies (23, 24, 26, 56). Moreover, other genetically engineered animal models, such as mice expressing enhanced green fluorescent protein (EGFP) or humanized mice carrying human functional biological system are valuable tools in brain tumor research, especially in studies about tumor host interactions and human-specific pathogenesis and therapies (57, 58).

GEMMs for cancer research can be generated by introduction of oncogenes or disruption of tumor suppressor genes in embryonic stem (ES) cells or zygotes (25) and include both transgenic mice and knockout mice. In addition to oncogenes and tumor suppressor genes, key molecules in tumor signaling pathways can also be utilized to develop GEMMs (**Table 1**). The conventional knockout models alter target gene expression in all tissues throughout the whole mouse. These constitutive changes often lead to more severe phenotypes with contribution by the brain tumor itself as well as other conditions. This makes data analysis and interpretation more difficult and less accurate. In some cases, this global knockout can even be lethal in animals. To overcome this issue, conditional or inducible conditional knockout has been developed, in which the target gene can be edited in a tissue specific and/or time-dependent way. The most common tool to make conditional knockout mice is the Cre-loxP system. Cre is a recombinase and its expression can be driven under the control of a tissue-specific promoter. When Cre is induced, it can recognize the loxP sites and catalyze the recombination, so the target gene flanked with two loxP sites in the same orientation will be excised. To achieve precise temporal specificity in the Cre-loxP system, Cre can be fused with a hormone responsive element, and induced by the exogenous inducers tamoxifen or tetracycline (35, 106).

The generation of germline GEMMs usually needs an extensive breeding scheme, which is time-consuming and expensive. Thus, virus mediated gene transfer is introduced to deliver Cre recombinase to somatic cells to establish non-germline GEMMs, which retains the ability of spatial and temporal gene regulation, at the same time also reduces the cost and time by bypassing complicated breeding (25). Replication-competent avian sarcoma-leukosis virus long terminal repeat with splice acceptor/tumor virus A (RCAS/TVA) is a commonly used system. RCAS is a retrovirus that enters specific cells *via* binding to its specific cell surface receptor TVA. TVA is only expressed in avian cells, but mammalian cells can gain the expression through genetic engineering (107). RACS/TVA based GEMMs have some advantages over Cre-loxP based models. The virus transduction rate is quite low, so only a small fraction of cells can acquire the expression of target genes. This makes the model close to natural tumorigenesis, since studies have shown that only small amounts of cancer stem cells are key players in tumor initiation (35). Moreover, genetically-engineered mammalian cells can get multiple RACS infection simultaneously or sequentially, which makes this model suitable to study the effect of multiple genes on tumorigenesis (107).

Recently, short palindromic clustered regularly interspaced repeats/CRISPR associated protein 9 (CRISPR/Cas9) technology has become a powerful tool to generate GEMMs. The CRISPR/Cas9 system is a groundbreaking gene editing technique; it consists of two necessary components: single strand guide RNA, which can recognize the target genomic DNA sequence, and endonuclease cas9, which can break the double-stranded DNA at the target sequence site. Then random or targeted gene editing can be achieved by DNA repair through error-prone non-homologous end joining (NHEJ) and high-fidelity homology directed repair (HDR) pathways. CRISPR/Cas9 can efficiently introduce gene modification on virtually any genetic background, both in germline and somatic cells. It turns conventionally tedious and expensive genetic engineering into a simple, fast and affordable procedure and dramatically broadens the application of GEMMs in tumor research (35, 108).

In 2019, MADR (mosaic analysis by dual recombinase-mediated cassette exchange) was introduced as a simpler, higher-throughput method to generate stable, defined copy number somatic transgenic animals. MADR was designed to overcome limitations of some of the previously described methods to generate mouse models (98). This includes the limited payloads and possible immune reactions when using viruses, the unpredictable genomic integration patterns, epigenetic transgenic silencing, transgene copy number variability, and overexpression artifacts such as cytotoxicity and transcriptional squelching when using viruses or

TABLE 1 | The common GEMMs of pediatric brain tumors.

Tumor type	Molecular subtypes	Mouse model name	References
Medulloblastoma	WNT	*Blbp-Cre*$^{+/-}$: Ctnnb1$^{+/lox(ex3)}$; Tp53$^{flx/flx}$	(59)
		Blbp-Cre$^{+/-}$; *Ctnnb1*$^{+/lox(ex3)}$; *Tp53*$^{+/flx}$; *Pik3ca*$^{loxE545K/loxE545K}$	(60)
	SHH	Ptch1$^{+/-}$	(61)
		Ptch1$^{+/-}$; Math1-Cre	(62)
		Ptch1$^{+/-}$; hGFAP-Cre	(62)
		Ptch1$^{+/-}$; Math1-CreER	(63)
		Ptc1$^{+/-}$; p53$^{-/-}$	(64)
		Ptc1$^{+/-}$; Ink4c$^{-/-}$	(65)
		Ptc1$^{+/-}$; Kip1$^{-/-}$	(66)
		Ptch1$^{+/-}$; Hic1$^{+/-}$	(67)
		Ptch1$^{+/-}$; Ptch2$^{-/-}$	(68)
		NeuroD2-SmoA1 (W539L)	(69)
		Smo/smo (homozygous smoA1)	(70)
		NeuroD2-SmoA2 (S537N)	(71)
		CAGGS-CreER; R26-SmoM2	(72)
		p53$^{-/-}$; Sufu$^{+/-}$	(73)
		Trp53$^{-/-}$; PTEN$^{-/-}$	(74)
		Trp53$^{-/-}$; Parp$^{-/-}$	(75)
		Nestin-tv-α mice infected with RCAS-Shh + N-Myc	(76)
		Nestin-tv-α mice infected with RCAS-Shh + N-Myc (T50A)	(76)
		Nestin-tv-α mice infected with RCAS-Shh + Bcl2	(77)
	Group3	Gtl1-tTA : TRE-MYCN/luciferase (GTML)	(78)
		GTML; Trp53$^{-/-}$	(79)
		Mll4$^{-/-}$; Nestin-Cre	(80)
		Nestin-tv-α; Trp53$^{-/-}$ mice infected with RCAS-Myc	(81)
		Nestin-tv-α mice infected with RCAS-Myc + Bcl2	(81)
		Co-electroporation of Myc and trp53DN into embryonic cerebellar progenitor cells	(82)
	Group4	Co-electroporation of SRC-CA and DNp53 into E13.5 developing cerebella	(83)
Gliomas		Nf1$^{+/-}$; p53$^{+/-}$	(84)
		p53–/–;NF1flox/flox;hGFAP-cre+	(85)
		cisp53+/–;NF1+/flox;hGFAP-cre+	(85)
		cisp53+/–;NF1+/flox; Pten$^{f/+}$*;hGFAP-cre+*	(86)
		TgGFAPT121	(87)
		GFAP-V^{12}Ha-ras	(88)
		GFAP-V^{12}Ha-ras;GFAP-EGFRvIII	(89)
		GFAP-V^{12}Ha-ras*;Ptenf/f*; hGFAP-Cre	(90)
		S100β-*v-erbB*	(91)
		S100β-*v-erbB*; Ink4a/Arf$^{-/-}$	(91)
		S100β-*v-erbB; p53* +/-	(91)
		Nestin-tv-α mice infected with RCAS K-ras and Akt	(92)
		Nestin-tv-α; *PtenloxP/loxP* mice infected with R*CAS KRAS and* RCAS cre	(93)
		Nestin-tv-α mice infected with RCAS–PDGF-B	(94)
		GFAP-tv-α mice infected with RCAS–PDGF-B	(94)
		Nestin-tv- α mice infected with RCAS-PDGFB	(95)
		Nestin-tv-a; Ink4a-arf$^{-/-}$ mice infected with RCAS-PDGFB	(95)
		Nestin-tv-a; p53$^{fl/fl}$ mice infected with RCAS-PDGF-B and RCAS-Cre	(96)
		GFAP tv-a; p53$^{fl/fl}$ mice infected with RCAS-PDGFB + RCAS Cre	(97)
		Nestin tv-a; p53$^{fl/fl}$ mice infected with RCAS-PDGFB+RCAS-Cre	(97)
		Pdgfra (D842V); Trp53 (R270H); H3f3a (G34R) (MDRA mice)	(98)
		Pdgfra (D842V); Trp53 (R270H); H3f3a (K27M) (MDRA mice)	(98)
		Erbb2-V664E; *PiggyBac* transposon	(99)
		Hras-G12V; *PiggyBac* transposon	(99)
		Kras-G12V; *PiggyBac* transposon	(99)
		Pdgfra-D842V; *PiggyBac* transposon	(99)
		PBCAG-Ngn2/PBCAG-HRasV12/Akt; in utero *Piggy Bac* transposon in different cell lineages *	(100)
		PBCAG-NeuroD1/PBCAG-HRasV12/Akt; in utero *Piggy Bac* transposon in different cell lineages *	(100)
		PTEN; NF1; P53 PiggyBac transposon-CRISPR/Cas9*	(101)
Ependymoma	ST-EPN-RELA	Nestin-tv- α mice infected with RCAS-*RELA*^{FUS1}	(102)
	ST-EPN-RELA	GFAP-tv- α mice infected with RCAS-*RELA*^{FUS1}	(102)
	ST-EPN-RELA	BLBP-tv- α mice infected with RCAS-*RELA*^{FUS1}	(102)
	ST-EPN-YAP1	YAP1-MAMLD1	(103)
	ST-EPN-YAP1	LATS1$^{f/f}$ LATs2$^{f/f}$: NEX$^{Cre/+}$	(104, 105)
	ST-EPN-YAP1	nlsYAP5SA/+: NEX$^{Cre/+}$	(104, 105)

**denotes a rat model.*

transposons, or the variability and unintended off-target genomic alterations of CRISPR/Cas9 systems (98).

To date, the majority of pediatric brain tumor GEMMs are SHH-activated medulloblastoma (**Table 1**), which are generated by modifying SHH signaling genes, such as *PTCH*, *SMO*, or *SUFU* (22, 24). The first mouse model of medulloblastoma was established by disrupting *ptch1* (61). Thereafter, the combination of a *ptch* mutation with inactivation of tumor suppressors, including TP53 or cyclin D-dependent kinase inhibitor $p18^{Ink4c}$, was used to generate different SHH driven medulloblastoma models with shorter latency and higher penetrance (22, 55, 109). SHH medulloblastoma models were also developed by overexpressing Shh, alone or in combination with *mycn* or *bcl2* with the RACS-TVA system (76, 77). Most WNT activated medulloblastoma models were generated by targeting the gene *ctnnb1* in progenitor cells of the dorsal brainstem. However, *ctnnb1* aberration alone was not sufficient to form medulloblastoma and the combination with TP53 mutation was needed to drive tumor initiation (59). Co-occurrence of *pik3ca* mutation significantly accelerated the formation of WNT medulloblastoma and dramatically increased tumor penetrance in mice (60). Most GEMMs of Group 3 medulloblastoma were developed by targeting *myc*. Newborn mice with *myc* overexpression in the cerebellum through the RCAS/TVA system induced Group 3 medulloblastoma, but tumor formation required *tp53* loss or *bcl-2* overexpression (81). Conditional enforced co-expression of *myc* and a dominant-negative form of *Trp53* (*Trp53DN)* in embryonic cerebellar progenitor cells by in utero electroporation also induced Group 3 medulloblastoma in mice (82). The first mouse model of Group 4 medulloblastoma was recently developed by overexpression of an activated SRC combined with p53 inactivation in the developing cerebellum (83). It is worthy of note that in some animal models induced tumors simultaneously possess multiple molecular characteristics. For example, the GTML (Glt1-tTA/TRE-MYCN-Luc) model in which MYCN aberration is driven by the glutamate transporter 1 (Glt1) promoter expressed in hindbrain progenitors develops tumors that closely resemble Group 3, but also shows the features of WNT, SHH, and Group 4 medulloblastoma (24).

Various medulloblastoma mouse models have been used to dissect mechanisms, including those influenced by the tumor microenvironment, that regulate the progression from precancerous lesions to medulloblastoma tumors (110). In general, tumors not only consist of the heterogenous tumor cell population but also of the extracellular matrix (ECM) surrounding the cells, resident and infiltrating cells such as tumor-associated fibroblasts, endothelial cells, pericytes, adipocytes, and immune cells including lymphocytes and macrophages as well as soluble factors, including cytokines, chemokines, growth factors, matrix remodeling enzymes and inflammatory enzymes (111, 112). The tumor microenvironment is known to contribute to tumor progression, metastasis formation and therapeutic response (113–122). In brain tumors, macrophages are the most abundant type of immune cells and are particularly high in Shh-driven medulloblastoma. In humans, decreased macrophage numbers are correlated with significant poorer outcome and indeed, a recent study in *NeuroD2:SmoA1* mice and derivative mouse lines was able to demonstrate that tumor-associated macrophages have properties that kill tumor cells (123). The *NeuroD2:SmoA1* medulloblastoma model was also used to show that blocking TGF-β signaling promoted memory T cell development thereby conferring antitumor immunity (124) and in *Atoh1-Cre;Ptch1$^{fl/fl}$* mice, tumor astrocyte-derived Shh induced the proliferation of medulloblastoma tumor cells (125). Cancer stem cells reside in specialized, anatomically distinct niches within the tumor microenvironment (126) and medulloblastoma stem cells ($Nestin^+$, $Prominin^+$) are closely associated with capillaries in the perivascular niche. Using mice infected with RCAS-Shh RCAS-SHH in combination with RCAS-N-myc-T50A or RCAS-AKT-Myr Δ11–60 of Ntv-a wild-type p53 and Ntv-a p53-null background, Hambardzumyan et al. showed that similar to human medulloblastomas, nestin-expressing perivascular stem cells survive radiation, activate PI3K/Akt signaling, undergo PTEN/p53-dependent cell cycle arrest and shortly thereafter re-enter the cell cycle (127). Medulloblastoma mouse models have also been used to elucidate pathways involved in tumor angiogenesis, in medulloblastoma metastasis, and in cell senescence and reprogramming (110).

Most mouse models of glioma are generated by altering key signaling pathways disrupted in human gliomas, including Ras, EGFR, Akt, Rb, Pten, Nf1 and platelet-derived growth factor (PDGF) (**Table 1**). GFAP-V12Ha-ras mice were generated by overexpressing oncogenic V12Ha-ras in astrocytes, and 95% of these mice died from low- and high-grade astrocytoma within 2-6 months (88). Further expressing a mutant EGFRVIII or inactivating PTEN in GFAP-V12Ha-ras mice demonstrated earlier tumor onset, higher tumor grade and a dramatic reduction in survival (89, 90). Introduction of activated Ras (KRas) into neural progenitors with the RCAS/TVA system, combined with activated Akt or PTEN loss induced high-grade gliomas in mice that resembled human GBMs (92, 93). Further deleting ink4a/arf increased tumor incidence and grades in these mice (128). Silencing of Bcl6 in neuronal precursor cells suppressed, but did not abolish, the formation of tumors in a somatic KrasG12V-driven glioma mouse model (99, 129). Transgenic S100β-v-erbB mice in which a transforming allele of EGFR, v-erbB, is expressed under the control of murine S100β promotor developed low-grade oligodendroglioma, and further deleting *ink4a/arf* or *p53* increased tumor grade and penetrance (91). An Nf1+/-; p53+/- mouse model shows a range of astrocytoma stages, from low-grade astrocytoma to glioblastoma multiforme (84). Conditionally deleting NF1 in glial progenitors and astrocytes of p53 null mice dramatically increased the penetrance of induced astrocytoma and the incidence of non-CNS neoplasms (85). Further loss *of Pten* in glial progenitors and astrocytes of this mouse model significantly accelerated tumor growth and animal mortality (86). TgGFAPT121 mice generated by a truncated SV40 T antigen (T121) to inactivate the Rb pathway in astrocytes develop high grade astrocytoma and die perinatally (87). PDGF B-chain

(PDGF-B) is another common target used to generate glioma models. PDGF-B delivered to nestin-positive neural progenitors or GFAP positive astrocytes induced low grade glioma in mice. Loss of *Ink4a–Arf* dramatically shortened tumor latency and enhanced malignancy of gliomas. p53 loss can also enhance PDGF-B driven glioma in mouse models (94–96). A more recent DIPG model was developed by overexpressing PDGF-B and H3.3K27M together with p53 loss in nestin-positive neural progenitors. The induced tumors in the brainstem of these mice demonstrated DIPG-like features that recapitulate the histopathological and molecular characteristic of human DIPG (96, 130). The MADR method was used to generate pediatric glioma mice modeling simultaneous H3f3a, Pdgfra, and Trp53 mutations with two missense mutation variants G34R or K27M that recapitulated human tumor heterogeneity and developmental hierarchy (98). Other recent pediatric brain tumor models have also been successful in capturing tumor heterogeneity and spatiotemporal characteristics of pediatric gliomas. *PiggyBac* transposon systems not only can circumvent the loss or inactivation of episomal plasmids delivered to glial cells *via in utero* electroporation but also allow for expression of multiple oncogenes in selected cell populations at different times in brain development (131). Using in utero electroporation of *piggyBac* transposons, Chen and colleagues generated rat tumor models by directing HRasV12 and AKT to different cell populations. Using the same transgene under the control of different promoters resulted in tumors ranging from glioblastoma multiforme to anaplastic oligoastrocytomas and atypical teratoid/rhabdoid-like tumors that could be distinguished at the cellular and the molecular level (100, 131). Moreover, targeting different genes, PTEN or NF1, in the same lineage resulted in distinct neuropathologies and when PTEN, NF1 and P53 were targeted simultaneously caused the formation of GBM (101).

The generation of ependymoma GEMMS began recently. $RELA^{FUS}1$ fusion gene expressed in nestin, GFAP, or BLBP positive cells in the mouse brain induced tumors which recapitulate the histology and transcriptome panel of human ST-EPN-RELA ependymomas (102). The YAP1-MAMLD1 fusion gene delivered to mice by in utero electroporation drove tumor formation and tumors share histological and molecular characteristics of human ST-EPN-YAP1 (103). Recently, Eder and colleagues reported that ectopic expression of active nuclear YAP1 (nlsYAP5SA) or conditional deletion of YAP1's negative regulators LATS1 and LATS2 kinases in neural progenitor cells in ventricular zone also induced tumors which display molecular and ultrastructural characteristics of human ependymoma (104, 105).

Zebrafish Brain Tumor Models

Zebrafish are an alternative model to study human cancer as they can develop tumors that are histologically and genetically similar to those in humans (132). Zebrafish models are also amenable to high-throughput screening for drug discovery as well as transplantation of primary patient tumors. This makes zebrafish a cost- and time-effective alternative to other *in vivo* tumor models such as rodents. In recent years, several pediatric brain tumor models have been developed in zebrafish and have been used to identify molecular mechanisms driving tumor formation. This includes the analysis of unique and shared molecular pathways driving pediatric HGG within and outside the brainstem (133) and to identify three molecular subgroups of DIPG (134). Ependymoma, glioma and choroid plexus carcinoma cells from mouse models of pediatric brain tumors were conditioned to grow at 34°C and used for orthotopic xenografts in zebrafish. These cells not only readily formed tumors but also spinal metastasis. The tumors retained the histological characteristics of the corresponding mouse tumor and formed tumor vasculature by recruiting fish endothelial cells (135). Lin et al. used zebrafish to experimentally validate subgroup-specific enhancers in medulloblastoma (13), Modzelewska et al. used tumors grown in zebrafish to demonstrate that MEK inhibitors can reverse the growth of embryonal brain tumors derived from oligoneural precursor cells (14) and Idilli et al. used them to study telomere maintenance mechanisms in pediatric brain tumors (136). With protocols for developing zebrafish tumor models evolving, long-term orthotopic transplantation of tumor cells is now possible (137). This allows for the long-term *in vivo* studies of tumor cell behaviors including tumor invasion and dissemination as well as testing for more durable response of tumors to novel anticancer therapeutics and the development of cancer drug resistance. Overall, zebrafish may provide an opportunity to develop pediatric brain tumor models in a timely and affordable manner for preclinical drug discovery in a model system with intact blood-brain barrier.

In Vitro Brain Tumor Models

Cancer Cell Lines

Cancer cell lines play an important role in brain tumor research. The cells lines can be established directly from patients' samples or from animal models. These cells often retain original tumor features, are easy to grow and propagate, and can be stored for a long time. They are well-suited models to explore a tumors' molecular features *in vitro* and predict the tumors' response to therapeutic regimens. Cancer cell lines are particularly useful in high-throughput drug screening to identify and evaluate potential targets for chemotherapies. Most established pediatric brain tumor cell lines are medulloblastoma cell lines. Less than half of these cell lines have been molecularly defined, among which the majority represent the SHH or Group 3 subtypes; only a few are for WNT or Group 4 tumors (**Table 2**). In addition, around half SHH medulloblastoma cell lines have mutations in *TP53*, and almost all Group 3 cell lines bear *MYC* amplification, while only a small part of SHH and Group 3 medulloblastoma patients typically have these mutations (24, 55, 145, 150). This discrepancy might be because the medulloblastomas with TP53 mutation and MYC amplification are more aggressive with poorer prognosis, and more aggressive cells are easier to grow *in vitro*. Similarly, although gliomas are the most common brain tumors in children, most gliomas are low grade gliomas with less malignancy and more favorable prognosis. Moreover, some high grade gliomas, such as DIPG, have a limited tissue availability due to tumor locations and established glioma cell lines are fewer

than medulloblastoma (150). In recent years, however, with the refinement of surgical skills and advances in DIPG biology, biopsy becomes more feasible in DIPG patients and new patient derived DIPG cell lines are becoming available (**Table 2**).

Cell culture *in vitro* has intrinsic drawbacks. Most cell lines are maintained as a monolayer culture in serum containing media, and a genetic and phenotypic drift from original tumors will gradually occur with passage (156). The cell lines are homogenous populations that cannot fully recapitulate the heterogeneity of tumors, and are not suitable to study tumor host interactions during tumor development. In monolayer culture, all cells receive the same level of nutrition and oxygen, which is different from tumor growth *in vivo*. Moreover, tumor cell lines are typically grown on borosilicate glass or clear plastics *in vitro*, which are much more rigid compared to extracellular matrix on which cells are naturally grown *in vivo* (157).

The advent of neurosphere cultures addressed some limitations of traditional cell cultures. Neurospheres are typically cultured in serum-free medium and can maintain tumor heterogeneity and preserve the phenotype and genotype of primary tumors (158). Some pediatric brain tumors, including DIPG, have been successfully cultured in neurospheres and used to generate xenografts that recapitulate the histological features and infiltrative growth of original patient tumors (24, 130, 152, 159). Neurosphere formation also can independently predict clinical outcome in malignant glioma (160). Therefore, neurospheres are a more representative and reliable cell model compared to traditional cell lines (36). However, neurosphere culture has some limitations, too. For example, the lack of a tumor microenvironment highly enriches glioma stem cell (GSC)-like cells, which only represent a relatively small subpopulation in native tumors (158, 161).

Three-Dimensional Culture

3D culture such as spheroids and scaffold-based cultures are other techniques that have been developed to overcome the limitations with traditional monolayer culture. In scaffold-based 3D cultures, extracellular matrix can be synthesized to simulate a tumor's natural microenvironment and a gradient of oxygen and nutrient level can be constituted to mimic a tumor's hypoxic core *in vivo*. Moreover, gene expression panels in 3D culture more closely resemble human tumors *in vivo* (157, 162, 163). To date, there are two major types of 3D culture: anchorage-dependent and anchorage-independent 3D models.

Anchorage-Independent 3D Models

Anchorage-independent 3D models are achieved mainly by self-assembly of cells grown in special tissue culture plates, such as hanging drop microplates and low attachment plates; they do not need any scaffold to facilitate the culture. The hanging drop culture is a well-known 3D culture technology. Typically, there is a micro-hole at the top of wells, which allows the medium to pass through and form a small droplet. Since there is no surface available for the cells in the droplet to attach, these cells tend to form spheroids. Spheroids can also be generated when they are grown in ultra-low attachment plates. Ependymoma cell lines cultured in ultra-low attachment plates better recapitulated the histological and transcriptional features of the primary tumors when compared to a monolayer (164). In addition, magnetic levitation is a newly developed method for spheroid formation. In this method, cells coated with magnetic nanoparticles are cultured in a magnetic field and the cells are floated toward the air/liquid interface within a low adhesion plate to form spheroids (162, 163, 165). Larger tumor spheroids formed by anchorage-independent models may consist of a peripheral layer with proliferating cells, an intermediate layer with quiescent cells and an inner necrotic core, which may closer reproduce human tumor architecture *in vivo* (165, 166). In recent years, with the refinement of technologies on spheroid culture, anchorage-independent 3D models have become common methods for cancer drug discovery, even applicable for high-throughput drug screening. Spheroids can be established with a few different types of cells and are especially suitable for studies about cell-cell interactions during brain tumor development (167). However, there is typically no extracellular matrix (ECM) in spheroids, and thus they are unsuitable for studying cell-host interactions which is a key game player in tumor pathologies.

Anchorage-Dependent 3D Models

The cells inside the body are usually surrounded by ECM, a network of extracellular molecules, which not only provides the structural scaffold for the surrounding cells, but also plays an important role in cell proliferation, differentiation, migration, survival and adhesion (168). The composition of ECM is highly heterogeneous and tissue-specific. The brain ECM ingredients include proteoglycans, hyaluronic acids, tenascins, collagen, fibronectin, vitronectin and laminin (162, 169). In anchorage-dependent 3D models, cells are encapsulated into scaffold materials, which can mimic the composition and key physical properties of ECM. Hydrogels are the most commonly used scaffolds for anchorage-dependent 3D models. Hydrogels are water-swollen networks of polymers and can mimic salient components of ECM. The highly hydrated and porous nature of hydrogel make them ideal to encapsulate cells and render ECM-like functions, such as supporting cell survival, growth, differentiation and modulating the response to chemotherapy, immunotherapy and radiation therapy (161).

Hydrogels may come from natural sources or can be synthetic. The widely used natural hydrogels for neural cell culture are collagen I and Matrigel. Matrigel is extracted from the Engelbreth-Holm-Swarm (EHS) mouse sarcoma, a tumor rich in ECM components, such as laminin, collagen, heparan sulfate proteoglycans, entactin/nidogen, and several growth factors. Matrigel is minimally processed and it can better mimic *in vivo* ECM (170). However, two major components of Matrigel are laminin and collagen, which are in low concentration in the brain ECM (158). Thus, collagen and Matrigel are not ideal choices as *in vivo*-like 3D scaffolds for brain tumor cells. In addition, collagen and Matrigel are derived from natural sources, they are heterogenous and not well defined, and exhibit considerable batch-to-batch variability. Moreover, collagen and Matrigel are available in liquid form and require handling at cold temperatures (below 10°C) to avoid premature

TABLE 2 | Established pediatric brain tumor cell lines with defined molecular characteristics.

Tumor Type	Molecular subtype	Cell line name	Mutations	Sources	References
Medulloblastoma	WNT	MED5R	β-catenin		(138)
	SHH	DAOY	CDKN2A NF1 TP53	ATCC	(139)
		ONS76		JCRB	(139)
		UW228	TP53		(139, 140)
		UW426			(139)
	Group 3	D341 Med	myc amplification	ATCC	(141, 142)
		D384 MED	myc amplification		(143)
		D425 MED	myc amplification p53	Millipore Sigma	(144)
		D458	Myc amplification		(145)
		D283 Med	myc amplification	ATCC	(144, 146)
		MED8A	myc amplification		(147)
		HD-MB03	myc amplification	DSMZ	(148)
		MB002	myc amplification		(149)
		Med-114FHTC	myc amplification	BTRL	https://www.btrl.org/product/med-114fhtc/
		Med-411FHTC	myc amplification Isochromosome 17	BTRL	https://www.btrl.org/product/med-411fhtc/
		Med-2112FHTC	myc Isochromosome17	BTRL	https://www.btrl.org/product/med-2112fhtc/
	Group 4	CHLA-01-MED	Myc amplification	ATCC	(150)
		CHLA-01R-MED	Myc amplification	ATCC	(150)
High grade glioma	MYCN	PBT-04FHTC	mycn, id2, nras	BTRL	(41)
		PBT-05FHTC	mycn, id2, egfr amplification	BTRL	(41)
	pedRTK1	GBM-511FHTC	cdkn2	BTRL	(41)
	pedRTK2	GBM-110FHTC	cdkn2, braf	BTRL	(41)
	Myc	CHLA-200	myc	COGcell.org	(151)
DIPG	H3.3 K27M	SF7761	Histone	Millipore Sigma	(152, 153)
		SF8628	Histone	Millipore Sigma	(153)
		PED8	Histone		(130)
		PED17	Histone		(130)
		PED36	histone		(130)
		HSJD-DIPG-007	histone		(50)
		HSJD-DIPG012	histone		(50)
		HSJD-DIPG017	histone		(50)
		SU-DIPG-VI	histone		(154)
		SU-DIPG-XIII	histone		(154)
		VUMC-DIPG-A	histone		(154)
		JHH-DIPG-1	histone		(154)
	H3.1 K27M	VUMC-DIPG-B	histone		(154)
		SU-DIPG-IV	histone		(154)
		HSJD-DIPG018	histone		(50)
GBM	H3.3 G34R	GBM002	histone		(50)
	H3K27M	GBM003	histone		(50)
Ependymoma	PF-EPN-A	EPD-210FHTC	1q gain	BTRL	(41)
		EPN-811	1q gain		(155)
		EPN-928	1q gain		(155)

gelation. The need for handling these hydrogels at low temperatures makes them poorly suited for common liquid handling equipment used for high-throughput screens in drug discovery (158, 162). Some of these limitations might be overcome by synthetic hydrogels. Synthetic hydrogels are derived from polymeric materials, such as polyethylene glycol (PEG), polylactic acid (PA) and polyglycolic acid (PGA). These hydrogels can simulate the composition and function of ECM; they often have engineered tunable properties to achieve desired stiffness and porosity, to enhance cell proliferation and differentiation by encapsulating bioactive molecules, such as growth factors or hormones. However, these polymers are biological inert, so they must be modified by addition of cell adhesion ligands or mixing with other natural ECM components to acquire the properties of cell adhesion (158, 162, 163). To date, PEG is a widely used synthetic hydrogel in neural cell 3D culture. A PEG-based hydrogel has been successfully used to grow GBM cell lines. In this system, the PEG-based hydrogel was modified with CRGDS and a MMP-cleavable peptide to facilitate cell proliferation, migration; hyaluronic acid (HA) was also added to mimic brain extracellular matrix (171). A synthetic MAX8 β-hairpin hydrogel was successfully used to culture pediatric

medulloblastoma cell lines in a high-throughput screening setting (172, 173). Hydrogel-based models are not suitable for long-term culture since they degrade fast. Solid porous scaffolds can be adapted to bypass this issue. Solid porous scaffolds are prepared from natural or synthetic polymers with mechanical stability and pore interconnectivity. The cells can be directly added to these solid porous scaffolds and maintain their 3D properties with continuous supply of nutrients. A recently developed tunable 3D brain tissue model integrated the porous scaffold with hydrogels. In this model, the donut shaped silk fibroin protein scaffold was infused with ECM hydrogels and brain tumor cells can grow into spheroids within the stiff silk scaffold, or migrate toward the central hydrogel. Thus, the outer-ring scaffold can be used to anchor neuronal cells, and the central soft hydrogel allows axonal penetration and connectivity (174, 175). A pediatric anaplastic ependymoma has been successfully cultured by this model (176). In addition, culturing glioblastoma tumor-initiating cells (TICs) in microscale alginate hydrogel tubes (AlgTubes) has been reported. This culture system allows for long-term and scalable production of glioblastoma cells for drug discovery (177). Self-assembling peptide (SAP) hydrogels are an evolving field for neural cell culture. These synthetic peptides can self-assemble under physiological conditions and support neural cell attachment, differentiation and synapse formation. SAP hydrogels are highly versatile, their material properties can be modulated by substituting amino acids, extending or shortening the peptide sequence, or by the addition of functional epitopes. A widely used peptide hydrogel is RADA16. However, peptide-based hydrogels may have poor mechanical properties, and some exhibit impaired cell viability caused by low pH, making it difficult to culture sensitive brain tumor cells (163).

Brain Organoids

Organoids are an emerging technology to study pediatric brain tumors. Organoids are typically generated with embryonic stem cells (ESC) or induced pluripotent stem cells (iPSC) and have the potential to grow into a 3D architecture in a way similar to *in vivo* tissue development by virtue of their capacity to self-renew and differentiate. Early organoid models were typically heterogeneous and lacked reproducibility since it was difficult to control the differentiation pattern of stem cells. However, with technical advances on directed differentiation, stems cells can now be differentiated into virtually any specific lineages. This has significantly moved forward the application of organoid models including in biomarker and drug discovery (163). Significant effort is being made in developing neural-based spheroids with cerebral organoids being one of the early ones (178). Cerebral organoids can be used as platforms for human brain tumor cells, or tumors can be initiated in cerebral organoids by introducing oncogenes and/or disrupting tumor repressor genes using gene editing technologies. Human cerebellar organoids derived from iPS cells electroporated with Otx2/c-MYC induced Group 3 medulloblastoma (179). Injecting cancer stem cells derived from GBM patients into cerebral organoids or genetic engineering of cerebral organoids by introducing $HRas^{G12V}$ and disrupting *p53* initiated tumorigenesis that closely recapitulated patient GBMs (180, 181). Organoids can also be established from patient brain tumors. Hubert et al. generated GBM organoids directly from patient samples that present hypoxic gradients and regional tumor heterogeneity (182). Organoids can theoretically resemble any *in vivo* brain niche with preserved cell distribution, can retain genetic and phenotypic stabilities, and are capable of long term culture; this makes them a valuable model to discover tumor initiation and progression, and a more accurate tool to predict the responses to tumor treatments. However, organoid cultures typically lack blood vessels and immune cells, which makes them unsuitable for testing tumor treatments targeting angiogenesis, or studying the contribution of immune system on tumorigenesis and relevant therapies. In addition, although organoids have proper cell composition and functions, they typically lack correct anatomical organization (162, 183). Another drawback of organoids and also found in spheroid models is that they often have a necrotic core, which sets a limit on the culture size and longevity. To overcome this limitation, microfluidic devices can be incorporated into 3D models. Microfluidic devices are designed for cell cultures under perfusion and allow for steady supplies of oxygen and nutrients while at the same time removing waste (163).

CONCLUSIONS

A precise *in vivo* pediatric brain tumor model is the one, which can faithfully recapitulate tumor's histopathological and molecular features; exhibit tumor's spatiotemporal characterization; demonstrate a tumor's microenvironment; predict patients' response to treatments; show high rate of incidence and short latency; and is reproducible, timesaving and cost-effective (184). Such accuracy in tumor models can best be achieved when genetic insults match the cell of origin and are introduced at developmental stages that are critical to tumor development. For effective *in vitro* drug discovery of novel cancer therapeutics, *in vitro* brain tumor models should not only recapitulate tumor biology but culture methods should also be suitable for high-throughput screening (HTS). New technologies and with it the possibilities of more complex screening platforms may be integrated to optimize the model systems for pediatric brain tumors. For example, the recently developed brain cancer-on-a-chip models incorporate multiple tissue types in 3D cultures into microphysiological system (MPS) and provide precise control of a cellular microenvironment and real-time monitoring on cell behavior and response. Nevertheless, while brain cancer-on-a-chip models can better mimic the physiological function of brain, challenges remain. Brain tumors demonstrate profound inter- and intra-tumoral heterogeneity and cellular plasticity to adapt their phenotypes to the surrounding. With more accurate *in vitro* and *in vivo* tumor models, however, it is possible to improve the current low approval rate of anticancer drugs, to offer more treatment options for pediatric brain tumor patients.

Although pediatric brain tumor models have been expanded immensely in the past decades, there is no single model that

meets all criteria and thus, experimental design and purpose will need to guide the choice of the brain tumor model (22, 24). The rapid advancement of genomic characterization of pediatric brain tumors and with it new genomic signatures of tumor subgroups add to the complexity of developing precise pediatric brain tumor models. Moreover, in recent years the genome landscape of pediatric brain tumors, both somatic and epigenetic, has been complemented by the analysis of tumor transcriptomes. Despite the plethora of data generated through such approaches, the finding that impaired differentiation of specific neural progenitors is a common mechanism underlying pediatric cancers (185) provides hope that a rational approach towards developing *in vitro* and *in vivo* pediatric brain tumor models can achieve a manageable library of research platforms for the development of impactful therapeutic interventions for pediatric brain cancers.

AUTHOR CONTRIBUTIONS

ZL and SL perceived and wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Nemours Foundation and the DoBelieve Foundation.

ACKNOWLEDGMENTS

We thank Karen Sperle for suggestions and critical reading of the manuscript.

REFERENCES

1. Ostrom QT, Gittleman H, Truitt G, Boscia A, Kruchko C, Barnholtz-Sloan JS. CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2011-2015. *Neuro Oncol* (2018) 20(suppl_4):iv1–iv86. doi: 10.1093/neuonc/noy131
2. Louis DN, Perry A, Reifenberger G, von Deimling A, Figarella-Branger D, Cavenee WK, et al. The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. *Acta Neuropathol* (2016) 131(6):803–20. doi: 10.1007/s00401-016-1545-1
3. Wells EM, Packer RJ. Pediatric brain tumors. *Continuum (Minneap Minn) 21(2 Neuro-oncol)* (2015) 21(2 Neuro-oncology):373–96. doi: 10.1212/01.CON.0000464176.96311.d1
4. Dang M, Phillips PC. Pediatric Brain Tumors. *Continuum (Minneap Minn)* (2017) 23(6, Neuro-oncology):1727–57. doi: 10.1212/CON.0000000000000545
5. Udaka YT, Packer RJ. Pediatric Brain Tumors. *Neurol Clin* (2018) 36 (3):533–56. doi: 10.1016/j.ncl.2018.04.009
6. Pollack IF, Agnihotri S, Broniscer A. Childhood brain tumors: current management, biological insights, and future directions. *J Neurosurg Pediatr* (2019) 23(3):261–73. doi: 10.3171/2018.10.PEDS183772018.10.PEDS18377
7. Wong CH, Siah KW, Lo AW. Corrigendum: Estimation of clinical trial success rates and related parameters. *Biostatistics* (2019) 20(2):366. doi: 10.1093/biostatistics/kxy072
8. Wong CH, Siah KW, Lo AW. Estimation of clinical trial success rates and related parameters. *Biostatistics* (2019) 20(2):273–86. doi: 10.1093/biostatistics/kxx069
9. Filbin MG, Sturm D. Gliomas in Children. *Semin Neurol* (2018) 38(1):121–30. doi: 10.1055/s-0038-1635106
10. Sturm D, Pfister SM, Jones DTW. Pediatric Gliomas: Current Concepts on Diagnosis, Biology, and Clinical Management. *J Clin Oncol* (2017) 35 (21):2370–7. doi: 10.1200/JCO.2017.73.0242
11. Ferris SP, Hofmann JW, Solomon DA, Perry A. Characterization of gliomas: from morphology to molecules. *Virchows Arch* (2017) 471(2):257–69. doi: 10.1007/s00428-017-2181-4
12. Penman CL, Faulkner C, Lowis SP, Kurian KM. Current Understanding of BRAF Alterations in Diagnosis, Prognosis, and Therapeutic Targeting in Pediatric Low-Grade Gliomas. *Front Oncol* (2015) 5:54. doi: 10.3389/fonc.2015.00054
13. Lin CY, Erkek S, Tong Y, Yin L, Federation AJ, Zapatka M, et al. Active medulloblastoma enhancers reveal subgroup-specific cellular origins. *Nature* (2016) 530(7588):57–62. doi: 10.1038/nature16546
14. Modzelewska K, Boer EF, Mosbruger TL, Picard D, Anderson D, Miles RR, et al. MEK Inhibitors Reverse Growth of Embryonal Brain Tumors Derived from Oligoneural Precursor Cells. *Cell Rep* (2016) 17(5):1255–64. doi: 10.1016/j.celrep.2016.09.081
15. Diaz AK, Baker SJ. The genetic signatures of pediatric high-grade glioma: no longer a one-act play. *Semin Radiat Oncol* (2014) 24(4):240–7. doi: 10.1016/j.semradonc.2014.06.003
16. Pajtler KW, Witt H, Sill M, Jones DT, Hovestadt V, Kratochwil F, et al. Molecular Classification of Ependymal Tumors across All CNS Compartments, Histopathological Grades, and Age Groups. *Cancer Cell* (2015) 27(5):728–43. doi: 10.1016/j.ccell.2015.04.002
17. Khatua S, Ramaswamy V, Bouffet E. Current therapy and the evolving molecular landscape of paediatric ependymoma. *Eur J Cancer* (2017) 70:34–41. doi: 10.1016/j.ejca.2016.10.013
18. Vitanza NA, Partap S. Pediatric Ependymoma. *J Child Neurol* (2016) 31 (12):1354–66. doi: 10.1177/0883073815610428
19. Northcott PA, Robinson GW, Kratz CP, Mabbott DJ, Pomeroy SL, Clifford SC, et al. Medulloblastoma. *Nat Rev Dis Primers* (2019) 5(1):11. doi: 10.1038/s41572-019-0063-610.1038/s41572-019-0063-6
20. Huang SY, Yang JY. Targeting the Hedgehog Pathway in Pediatric Medulloblastoma. *Cancers (Basel)* (2015) 7(4):2110–23. doi: 10.3390/cancers7040880
21. Cavalli FMG, Remke M, Rampasek L, Peacock J, Shih DJH, Luu B, et al. Intertumoral Heterogeneity within Medulloblastoma Subgroups. *Cancer Cell* (2017) 31(6):737–54 e6. doi: 10.1016/j.ccell.2017.05.005
22. Neumann JE, Swartling FJ, Schuller U. Medulloblastoma: experimental models and reality. *Acta Neuropathol* (2017) 134(5):679–89. doi: 10.1007/s00401-017-1753-3
23. Huszthy PC, Daphu I, Niclou SP, Stieber D, Nigro JM, Sakariassen PO, et al. In vivo models of primary brain tumors: pitfalls and perspectives. *Neuro Oncol* (2012) 14(8):979–93. doi: 10.1093/neuonc/nos135
24. Dobson THW, Gopalakrishnan V. Preclinical Models of Pediatric Brain Tumors-Forging Ahead. *Bioengineer (Basel)* (2018) 5(4):81. doi: 10.3390/bioengineering5040081
25. Day CP, Merlino G, Van Dyke T. Preclinical mouse cancer models: a maze of opportunities and challenges. *Cell* (2015) 163(1):39–53. doi: 10.1016/j.cell.2015.08.068
26. Simeonova I, Huillard E. In vivo models of brain tumors: roles of genetically engineered mouse models in understanding tumor biology and use in preclinical studies. *Cell Mol Life Sci* (2014) 71(20):4007–26. doi: 10.1007/s00018-014-1675-3
27. Koestner A, Swenberg JA, Wechsler W. Transplacental production with ethylnitrosourea of neoplasms of the nervous system in Sprague-Dawley rats. *Am J Pathol* (1971) 63(1):37–56.
28. Barth RF, Kaur B. Rat brain tumor models in experimental neuro-oncology: the C6, 9L, T9, RG2, F98, BT4C, RT-2 and CNS-1 gliomas. *J Neurooncol* (2009) 94(3):299–312. doi: 10.1007/s11060-009-9875-7
29. Kruse CA, Molleston MC, Parks EP, Schiltz PM, Kleinschmidt-DeMasters BK, Hickey WF. A rat glioma model, CNS-1, with invasive characteristics similar to those of human gliomas: a comparison to 9L gliosarcoma. *J Neurooncol* (1994) 22(3):191–200. doi: 10.1007/BF01052919

30. Ko L, Koestner A, Wechsler W. Morphological characterization of nitrosourea-induced glioma cell lines and clones. *Acta Neuropathol* (1980) 51(1):23–31. doi: 10.1007/BF00688846
31. Benda P, Lightbody J, Sato G, Levine L, Sweet W. Differentiated rat glial cell strain in tissue culture. *Science* (1968) 161(3839):370–1. doi: 10.1126/science.161.3839.370
32. Cuatico W, Cho JR, Spiegelman S. Molecular evidence for a viral etiology of human CNS tumors. *Acta Neurochir (Wien)* (1976) 35(1-3):149–60. doi: 10.1007/BF01405943
33. Rabotti GF, Raine WA. Brain Tumours Induced in Hamsters Inoculated Intracerebrally at Birth with Rous Sarcoma Virus. *Nature* (1964) 204:898–9. doi: 10.1038/204898a0
34. Ogawa K, Hamaya K, Fujii Y, Matsuura K, Endo T. Tumor induction by adenovirus type 12 and its target cells in the central nervous system. *Gan* (1969) 60(4):383–92.
35. Robertson FL, Marques-Torrejon MA, Morrison GM, Pollard SM. Experimental models and tools to tackle glioblastoma. *Dis Model Mech* (2019) 12(9):dmm040386. doi: 10.1242/dmm.040386
36. Patrizii M, Bartucci M, Pine SR, Sabaawy HE. Utility of Glioblastoma Patient-Derived Orthotopic Xenografts in Drug Discovery and Personalized Therapy. *Front Oncol* (2018) 8:23. doi: 10.3389/fonc.2018.00023
37. Okada S, Vaeteewoottacharn K, Kariya R. Application of Highly Immunocompromised Mice for the Establishment of Patient-Derived Xenograft (PDX) Models. *Cells* (2019) 8(8):889. doi: 10.3390/cells8080889
38. Zarzosa P, Navarro N, Giralt I, Molist C, Almazan-Moga A, Vidal I, et al. Patient-derived xenografts for childhood solid tumors: a valuable tool to test new drugs and personalize treatments. *Clin Transl Oncol* (2017) 19(1):44–50. doi: 10.1007/s12094-016-1557-2
39. Hermans E, Hulleman E. Patient-Derived Orthotopic Xenograft Models of Pediatric Brain Tumors: In a Mature Phase or Still in Its Infancy? *Front Oncol* (2019) 9:1418. doi: 10.3389/fonc.2019.01418
40. Giannini C, Sarkaria JN, Saito A, Uhm JH, Galanis E, Carlson BL, et al. and PDGFRA gene amplifications retained in an invasive intracranial xenograft model of glioblastoma multiforme. *Neuro Oncol* (2005) 7(2):164–76. doi: 10.1215/S1152851704000821
41. Brabetz S, Leary SES, Grobner SN, Nakamoto MW, Seker-Cin H, Girard EJ, et al. A biobank of patient-derived pediatric brain tumor models. *Nat Med* (2018) 24(11):1752–61. doi: 10.1038/s41591-018-0207-3
42. Smith KS, Xu K, Mercer KS, Boop F, Klimo P, DeCupyere M, et al. Patient-derived orthotopic xenografts of pediatric brain tumors: a St. Jude resource. *Acta Neuropathol* (2020) 140(2):209–25. doi: 10.1007/s00401-020-02171-5
43. Tsoli M, Shen H, Mayoh C, Franshaw L, Ehteda A, Upton D, et al. Correction to: International experience in the development of patient-derived xenograft models of diffuse intrinsic pontine glioma. *J Neurooncol* (2019) 141(2):265. doi: 10.1007/s11060-018-03060-4
44. Tsoli M, Shen H, Mayoh C, Franshaw L, Ehteda A, Upton D, et al. International experience in the development of patient-derived xenograft models of diffuse intrinsic pontine glioma. *J Neurooncol* (2019) 141(2):253–63. doi: 10.1007/s11060-018-03038-2
45. Vo BT, Kwon JA, Li C, Finkelstein D, Xu B, Orr BA, et al. Mouse medulloblastoma driven by CRISPR activation of cellular Myc. *Sci Rep* (2018) 8(1):8733. doi: 10.1038/s41598-018-24956-1
46. Pei Y, Moore CE, Wang J, Tewari AK, Eroshkin A, Cho YJ, et al. An animal model of MYC-driven medulloblastoma. *Cancer Cell* (2012) 21(2):155–67. doi: 10.1016/j.ccr.2011.12.021
47. Northcott PA, Lee C, Zichner T, Stutz AM, Erkek S, Kawauchi D, et al. Enhancer hijacking activates GFI1 family oncogenes in medulloblastoma. *Nature* (2014) 511(7510):428–34. doi: 10.1038/nature13379
48. Kawauchi D, Robinson G, Uziel T, Gibson P, Rehg J, Gao C, et al. A mouse model of the most aggressive subgroup of human medulloblastoma. *Cancer Cell* (2012) 21(2):168–80. doi: 10.1016/j.ccr.2011.12.023
49. Funato K, Major T, Lewis PW, Allis CD, Tabar V. Use of human embryonic stem cells to model pediatric gliomas with H3.3K27M histone mutation. *Science* (2014) 346(6216):1529–33. doi: 10.1126/science.1253799
50. Mohammad F, Weissmann S, Leblanc B, Pandey DP, Hojfeldt JW, Comet I, et al. EZH2 is a potential therapeutic target for H3K27M-mutant pediatric gliomas. *Nat Med* (2017) 23(4):483–92. doi: 10.1038/nm.4293
51. Johnson RA, Wright KD, Poppleton H, Mohankumar KM, Finkelstein D, Pounds SB, et al. Cross-species genomics matches driver mutations and cell compartments to model ependymoma. *Nature* (2010) 466(7306):632–6. doi: 10.1038/nature09173
52. Huang M, Tailor J, Zhen Q, Gillmor AH, Miller ML, Weishaupt H, et al. Engineering Genetic Predisposition in Human Neuroepithelial Stem Cells Recapitulates Medulloblastoma Tumorigenesis. *Cell Stem Cell* (2019) 25(3):433–46 e7. doi: 10.1016/j.stem.2019.05.013
53. Susanto E, Marin Navarro A, Zhou L, Sundstrom A, van Bree N, Stantic M, et al. Modeling SHH-driven medulloblastoma with patient iPS cell-derived neural stem cells. *Proc Natl Acad Sci USA* (2020) 117(33):20127–38. doi: 10.1073/pnas.1920521117
54. Ledford H. US cancer institute to overhaul tumour cell lines. *Nature* (2016) 530(7591):391. doi: 10.1038/nature.2016.19364
55. Roussel MF, Stripay JL. Modeling pediatric medulloblastoma. *Brain Pathol* (2020) 30(3):703–12. doi: 10.1111/bpa.12803
56. Huse JT, Holland EC. Genetically engineered mouse models of brain cancer and the promise of preclinical testing. *Brain Pathol* (2009) 19(1):132–43. doi: 10.1111/j.1750-3639.2008.00234.x
57. Niclou SP, Danzeisen C, Eikesdal HP, Wiig H, Brons NH, Poli AM, et al. A novel eGFP-expressing immunodeficient mouse model to study tumor-host interactions. *FASEB J* (2008) 22(9):3120–8. doi: 10.1096/fj.08-109611
58. Scheer N, Snaith M, Wolf CR, Seibler J. Generation and utility of genetically humanized mouse models. *Drug Discovery Today* (2013) 18(23-24):1200–11. doi: 10.1016/j.drudis.2013.07.007
59. Gibson P, Tong Y, Robinson G, Thompson MC, Currle DS, Eden C, et al. Subtypes of medulloblastoma have distinct developmental origins. *Nature* (2010) 468(7327):1095–9. doi: 10.1038/nature09587
60. Robinson G, Parker M, Kranenburg TA, Lu C, Chen X, Ding L, et al. Novel mutations target distinct subgroups of medulloblastoma. *Nature* (2012) 488(7409):43–8. doi: 10.1038/nature11213nature11213
61. Goodrich LV, Milenkovic L, Higgins KM, Scott MP. Altered neural cell fates and medulloblastoma in mouse patched mutants. *Science* (1997) 277(5329):1109–13. doi: 10.1126/science.277.5329.1109
62. Yang ZJ, Ellis T, Markant SL, Read TA, Kessler JD, Bourboulas M, et al. Medulloblastoma can be initiated by deletion of Patched in lineage-restricted progenitors or stem cells. *Cancer Cell* (2008) 14(2):135–45. doi: 10.1016/j.ccr.2008.07.003
63. Li P, Du F, Yuelling LW, Lin T, Muradimova RE, Tricarico R, et al. A population of Nestin-expressing progenitors in the cerebellum exhibits increased tumorigenicity. *Nat Neurosci* (2013) 16(12):1737–44. doi: 10.1038/nn.3553
64. Wetmore C, Eberhart DE, Curran T. Loss of p53 but not ARF accelerates medulloblastoma in mice heterozygous for patched. *Cancer Res* (2001) 61(2):513–6.
65. Uziel T, Zindy F, Xie S, Lee Y, Forget A, Magdaleno S, et al. The tumor suppressors Ink4c and p53 collaborate independently with Patched to suppress medulloblastoma formation. *Genes Dev* (2005) 19(22):2656–67. doi: 10.1101/gad.1368605
66. Ayrault O, Zindy F, Rehg J, Sherr CJ, Roussel MF. Two tumor suppressors, p27Kip1 and patched-1, collaborate to prevent medulloblastoma. *Mol Cancer Res* (2009) 7(1):33–40. doi: 10.1158/1541-7786.MCR-08-0369
67. Briggs KJ, Corcoran-Schwartz IM, Zhang W, Harcke T, Devereux WL, Baylin SB, et al. Cooperation between the Hic1 and Ptch1 tumor suppressors in medulloblastoma. *Genes Dev* (2008) 22(6):770–85. doi: 10.1101/gad.1640908
68. Lee Y, Miller HL, Russell HR, Boyd K, Curran T, McKinnon PJ. Patched2 modulates tumorigenesis in patched1 heterozygous mice. *Cancer Res* (2006) 66(14):6964–71. doi: 10.1158/0008-5472.CAN-06-0505
69. Hallahan AR, Pritchard JI, Hansen S, Benson M, Stoeck J, Hatton BA, et al. The SmoA1 mouse model reveals that notch signaling is critical for the growth and survival of sonic hedgehog-induced medulloblastomas. *Cancer Res* (2004) 64(21):7794–800. doi: 10.1158/0008-5472.CAN-04-1813
70. Hatton BA, Villavicencio EH, Tsuchiya KD, Pritchard JI, Ditzler S, Pullar B, et al. The Smo/Smo model: hedgehog-induced medulloblastoma with 90%

incidence and leptomeningeal spread. *Cancer Res* (2008) 68(6):1768–76. doi: 10.1158/0008-5472.CAN-07-5092

71. Dey J, Ditzler S, Knoblaugh SE, Hatton BA, Schelter JM, Cleary MA, et al. A distinct Smoothened mutation causes severe cerebellar developmental defects and medulloblastoma in a novel transgenic mouse model. *Mol Cell Biol* (2012) 32(20):4104–15. doi: 10.1128/MCB.00862-12
72. Mao J, Ligon KL, Rakhlin EY, Thayer SP, Bronson RT, Rowitch D, et al. A novel somatic mouse model to survey tumorigenic potential applied to the Hedgehog pathway. *Cancer Res* (2006) 66(20):10171–8. doi: 10.1158/0008-5472.CAN-06-0657
73. Lee Y, Kawagoe R, Sasai K, Li Y, Russell HR, Curran T, et al. Loss of suppressor-of-fused function promotes tumorigenesis. *Oncogene* (2007) 26 (44):6442–7. doi: 10.1038/sj.onc.1210467
74. Zhu G, Rankin SL, Larson JD, Zhu X, Chow LM, Qu C, et al. PTEN Signaling in the Postnatal Perivascular Progenitor Niche Drives Medulloblastoma Formation. *Cancer Res* (2017) 77(1):123–33. doi: 10.1158/0008-5472.CAN-16-1991
75. Tong WM, Ohgaki H, Huang H, Granier C, Kleihues P, Wang ZQ. Null mutation of DNA strand break-binding molecule poly(ADP-ribose) polymerase causes medulloblastomas in p53(-/-) mice. *Am J Pathol* (2003) 162(1):343–52. doi: 10.1016/S0002-9440(10)63825-4
76. Browd SR, Kenney AM, Gottfried ON, Yoon JW, Walterhouse D, Pedone CA, et al. N-myc can substitute for insulin-like growth factor signaling in a mouse model of sonic hedgehog-induced medulloblastoma. *Cancer Res* (2006) 66(5):2666–72. doi: 10.1158/0008-5472.CAN-05-2198
77. McCall TD, Pedone CA, Fults DW. Apoptosis suppression by somatic cell transfer of Bcl-2 promotes Sonic hedgehog-dependent medulloblastoma formation in mice. *Cancer Res* (2007) 67(11):5179–85. doi: 10.1158/0008-5472.CAN-06-4177
78. Swartling FJ, Grimmer MR, Hackett CS, Northcott PA, Fan QW, Goldenberg DD, et al. Pleiotropic role for MYCN in medulloblastoma. *Genes Dev* (2010) 24(10):1059–72. doi: 10.1101/gad.1907510
79. Hill RM, Kuijper S, Lindsey JC, Petrie K, Schwalbe EC, Barker K, et al. and P53 defects emerge at medulloblastoma relapse and define rapidly progressive, therapeutically targetable disease. *Cancer Cell* (2015) 27 (1):72–84. doi: 10.1016/j.ccell.2014.11.002
80. Dhar SS, Zhao D, Lin T, Gu B, Pal K, Wu SJ, et al. MLL4 Is Required to Maintain Broad H3K4me3 Peaks and Super-Enhancers at Tumor Suppressor Genes. *Mol Cell* (2018) 70(5):825–41 e6. doi: 10.1016/j.molcel.2018.04.028
81. Jenkins NC, Rao G, Eberhart CG, Pedone CA, Dubuc AM, Fults DW. Somatic cell transfer of c-Myc and Bcl-2 induces large-cell anaplastic medulloblastomas in mice. *J Neurooncol* (2016) 126(3):415–24. doi: 10.1007/s11060-015-1985-9
82. Kawauchi D, Ogg RJ, Liu L, Shih DJH, Finkelstein D, Murphy BL, et al. Novel MYC-driven medulloblastoma models from multiple embryonic cerebellar cells. *Oncogene* (2017) 36(37):5231–42. doi: 10.1038/onc.2017.110
83. Forget A, Martignetti L, Puget S, Calzone L, Brabetz S, Picard D, et al. Aberrant ERBB4-SRC Signaling as a Hallmark of Group 4 Medulloblastoma Revealed by Integrative Phosphoproteomic Profiling. *Cancer Cell* (2018) 34 (3):379–95 e7. doi: 10.1016/j.ccell.2018.08.002
84. Reilly KM, Loisel DA, Bronson RT, McLaughlin ME, Jacks T. Nf1;Trp53 mutant mice develop glioblastoma with evidence of strain-specific effects. *Nat Genet* (2000) 26(1):109–13. doi: 10.1038/79075
85. Zhu Y, Guignard F, Zhao D, Liu L, Burns DK, Mason RP, et al. Early inactivation of p53 tumor suppressor gene cooperating with NF1 loss induces malignant astrocytoma. *Cancer Cell* (2005) 8(2):119–30. doi: 10.1016/j.ccr.2005.07.004
86. Kwon CH, Zhao D, Chen J, Alcantara S, Li Y, Burns DK, et al. Pten haploinsufficiency accelerates formation of high-grade astrocytomas. *Cancer Res* (2008) 68(9):3286–94. doi: 10.1158/0008-5472.CAN-07-6867
87. Xiao A, Wu H, Pandolfi PP, Louis DN, Van Dyke T. Astrocyte inactivation of the pRb pathway predisposes mice to malignant astrocytoma development that is accelerated by PTEN mutation. *Cancer Cell* (2002) 1 (2):157–68. doi: 10.1016/s1535-6108(02)00029-6
88. Ding H, Roncari L, Shannon P, Wu X, Lau N, Karaskova J, et al. Astrocyte-specific expression of activated p21-ras results in malignant astrocytoma formation in a transgenic mouse model of human gliomas. *Cancer Res* (2001) 61(9):3826–36.
89. Ding H, Shannon P, Lau N, Wu X, Roncari L, Baldwin RL, et al. Oligodendrogliomas result from the expression of an activated mutant epidermal growth factor receptor in a RAS transgenic mouse astrocytoma model. *Cancer Res* (2003) 63(5):1106–13.
90. Wei Q, Clarke L, Scheidenhelm DK, Qian B, Tong A, Sabha N, et al. High-grade glioma formation results from postnatal pten loss or mutant epidermal growth factor receptor expression in a transgenic mouse glioma model. *Cancer Res* (2006) 66(15):7429–37. doi: 10.1158/0008-5472.CAN-06-0712
91. Weiss WA, Burns MJ, Hackett C, Aldape K, Hill JR, Kuriyama H, et al. Genetic determinants of malignancy in a mouse model for oligodendroglioma. *Cancer Res* (2003) 63(7):1589–95.
92. Holland EC, Celestino J, Dai C, Schaefer L, Sawaya RE, Fuller GN. Combined activation of Ras and Akt in neural progenitors induces glioblastoma formation in mice. *Nat Genet* (2000) 25(1):55–7. doi: 10.1038/75596
93. Hu X, Pandolfi PP, Li Y, Koutcher JA, Rosenblum M, Holland EC. mTOR promotes survival and astrocytic characteristics induced by Pten/AKT signaling in glioblastoma. *Neoplasia* (2005) 7(4):356–68. doi: 10.1593/neo.04595
94. Dai C, Celestino JC, Okada Y, Louis DN, Fuller GN, Holland EC. PDGF autocrine stimulation dedifferentiates cultured astrocytes and induces oligodendrogliomas and oligoastrocytomas from neural progenitors and astrocytes *in vivo*. *Genes Dev* (2001) 15(15):1913–25. doi: 10.1101/gad.903001
95. Becher OJ, Hambardzumyan D, Walker TR, Helmy K, Nazarian J, Albrecht S, et al. Preclinical evaluation of radiation and perifosine in a genetically and histologically accurate model of brainstem glioma. *Cancer Res* (2010) 70 (6):2548–57. doi: 10.1158/0008-5472.CAN-09-2503
96. Barton KL, Misuraca K, Cordero F, Dobrikova E, Min HD, Gromeier M, et al. PD-0332991, a CDK4/6 inhibitor, significantly prolongs survival in a genetically engineered mouse model of brainstem glioma. *PloS One* (2013) 8 (10):e77639. doi: 10.1371/journal.pone.0077639
97. Halvorson KG, Barton KL, Schroeder K, Misuraca KL, Hoeman C, Chung A, et al. A high-throughput *in vitro* drug screen in a genetically engineered mouse model of diffuse intrinsic pontine glioma identifies BMS-754807 as a promising therapeutic agent. *PloS One* (2015) 10(3):e0118926. doi: 10.1371/journal.pone.0118926
98. Kim GB, Rincon Fernandez Pacheco D, Saxon D, Yang A, Sabet S, Dutra-Clarke M, et al. Rapid Generation of Somatic Mouse Mosaics with Locus-Specific, Stably Integrated Transgenic Elements. *Cell* (2019) 179(1):251–67 e24. doi: 10.1016/j.cell.2019.08.013
99. Breunig JJ, Levy R, Antonuk CD, Molina J, Dutra-Clarke M, Park H, et al. Ets Factors Regulate Neural Stem Cell Depletion and Gliogenesis in Ras Pathway Glioma. *Cell Rep* (2015) 12(2):258–71. doi: 10.1016/j.celrep.2015.06.012
100. Chen F, Becker AJ, LoTurco JJ. Contribution of tumor heterogeneity in a new animal model of CNS tumors. *Mol Cancer Res* (2014) 12(5):742–53. doi: 10.1158/1541-7786.MCR-13-0531
101. Chen F, Rosiene J, Che A, Becker A, LoTurco J. Tracking and transforming neocortical progenitors by CRISPR/Cas9 gene targeting and piggyBac transposase lineage labeling. *Development* (2015) 142(20):3601–11. doi: 10.1242/dev.118836
102. Ozawa T, Arora S, Szulzewsky F, Juric-Sekhar G, Miyajima Y, Bolouri H, et al. A De Novo Mouse Model of C11orf95-RELA Fusion-Driven Ependymoma Identifies Driver Functions in Addition to NF-kappaB. *Cell Rep* (2018) 23(13):3787–97. doi: 10.1016/j.celrep.2018.04.099
103. Pajtler KW, Wei Y, Okonechnikov K, Silva PBG, Vouri M, Zhang L, et al. YAP1 subgroup supratentorial ependymoma requires TEAD and nuclear factor I-mediated transcriptional programmes for tumorigenesis. *Nat Commun* (2019) 10(1):3914. doi: 10.1038/s41467-019-11884-5
104. Eder N, Roncaroli F, Domart MC, Horswell S, Andreiuolo F, Flynn HR, et al. Author Correction: YAP1/TAZ drives ependymoma-like tumour formation in mice. *Nat Commun* (2020) 11(1):4934. doi: 10.1038/s41467-020-18851-5
105. Eder N, Roncaroli F, Domart MC, Horswell S, Andreiuolo F, Flynn HR, et al. YAP1/TAZ drives ependymoma-like tumour formation in mice. *Nat Commun* (2020) 11(1):2380. doi: 10.1038/s41467-020-16167-y

106. Kim H, Kim M, Im SK, Fang S. Mouse Cre-LoxP system: general principles to determine tissue-specific roles of target genes. *Lab Anim Res* (2018) 34 (4):147–59. doi: 10.5625/lar.2018.34.4.147
107. von Werder A, Seidler B, Schmid RM, Schneider G, Saur D. Production of avian retroviruses and tissue-specific somatic retroviral gene transfer *in vivo* using the RCAS/TVA system. *Nat Protoc* (2012) 7(6):1167–83. doi: 10.1038/nprot.2012.060
108. Weber J, Rad R. Engineering CRISPR mouse models of cancer. *Curr Opin Genet Dev* (2019) 54:88–96. doi: 10.1016/j.gde.2019.04.001
109. Wu X, Northcott PA, Croul S, Taylor MD. Mouse models of medulloblastoma. *Chin J Cancer* (2011) 30(7):442–9. doi: 10.5732/cjc.011.10040
110. Tamayo-Orrego L, Charron F. Recent advances in SHH medulloblastoma progression: tumor suppressor mechanisms and the tumor microenvironment. *F1000Res* (2019) 8(F1000 Faculty Rev):1823. doi: 10.12688/f1000research.20013.1
111. Balkwill FR, Capasso M, Hagemann T. The tumor microenvironment at a glance. *J Cell Sci* (2012) 125(23):5591–6. doi: 10.1242/jcs.116392
112. Egeblad M, Nakasone ES, Werb Z. Tumors as organs: complex tissues that interface with the entire organism. *Dev Cell* (2010) 18(6):884–901. doi: 10.1016/j.devcel.2010.05.012S1534-5807(10)00248-0
113. Dickreuter E, Cordes N. The cancer cell adhesion resistome: mechanisms, targeting and translational approaches. *Biol Chem* (2017) 398(7):721–35. doi: 10.1515/hsz-2016-0326
114. Holle AW, Young JL, Spatz JP. In vitro cancer cell-ECM interactions inform *in vivo* cancer treatment. *Adv Drug Delivery Rev* (2016) 97:270–9. doi: 10.1016/j.addr.2015.10.007S0169-409X(15)00232-X
115. Henke E, Nandigama R, Ergun S. Extracellular Matrix in the Tumor Microenvironment and Its Impact on Cancer Therapy. *Front Mol Biosci* (2019) 6:160. doi: 10.3389/fmolb.2019.00160
116. Holohan C, Van Schaeybroeck S, Longley DB, Johnston PG. Cancer drug resistance: an evolving paradigm. *Nat Rev Cancer* (2013) 13(10):714–26. doi: 10.1038/nrc3599nrc3599
117. Paolillo M, Schinelli S. Extracellular Matrix Alterations in Metastatic Processes. *Int J Mol Sci* (2019) 20(19):4947. doi: 10.3390/ijms20194947
118. Junttila MR, de Sauvage FJ. Influence of tumour micro-environment heterogeneity on therapeutic response. *Nature* (2013) 501(7467):346–54. doi: 10.1038/nature12626nature12626
119. McMillin DW, Negri JM, Mitsiades CS. The role of tumour-stromal interactions in modifying drug response: challenges and opportunities. *Nat Rev Drug Discovery* (2013) 12(3):217–28. doi: 10.1038/nrd3870nrd3870
120. Li F, Simon MC. Cancer Cells Don't Live Alone: Metabolic Communication within Tumor Microenvironments. *Dev Cell* (2020) 54(2):183–95. doi: 10.1016/j.devcel.2020.06.018
121. Turley SJ, Cremasco V, Astarita JL. Immunological hallmarks of stromal cells in the tumour microenvironment. *Nat Rev Immunol* (2015) 15(11):669–82. doi: 10.1038/nri3902nri3902
122. Hanahan D, Coussens LM. Accessories to the crime: functions of cells recruited to the tumor microenvironment. *Cancer Cell* (2012) 21(3):309–22. doi: 10.1016/j.ccr.2012.02.022
123. Maximov V, Chen Z, Wei Y, Robinson MH, Herting CJ, Shanmugam NS, et al. Tumour-associated macrophages exhibit anti-tumoural properties in Sonic Hedgehog medulloblastoma. *Nat Commun* (2019) 10(1):2410. doi: 10.1038/s41467-019-10458-9
124. Gate D, Danielpour M, Rodriguez JJr., Kim G-B, Levy R, Bannykh S, et al. T-cell TGF-β signaling abrogation restricts medulloblastoma progression. *Proc Natl Acad Sci USA* (2014) 111(33):E3458–E66. doi: 10.1073/pnas.1412489111
125. Liu Y, Yuelling LW, Wang Y, Du F, Gordon RE, O'Brien JA, et al. Astrocytes Promote Medulloblastoma Progression through Hedgehog Secretion. *Cancer Res* (2017) 77(23):6692–703. doi: 10.1158/0008-5472.Can-17-1463
126. Plaks V, Kong N, Werb Z. The cancer stem cell niche: how essential is the niche in regulating stemness of tumor cells? *Cell Stem Cell* (2015) 16(3):225–38. doi: 10.1016/j.stem.2015.02.015
127. Hambardzumyan D, Becher OJ, Rosenblum MK, Pandolfi PP, Manova-Todorova K, Holland EC. PI3K pathway regulates survival of cancer stem cells residing in the perivascular niche following radiation in medulloblastoma. *Vivo Genes Dev* (2008) 22(4):436–48. doi: 10.1101/gad.1627008
128. Uhrbom L, Dai C, Celestino JC, Rosenblum MK, Fuller GN, Holland EC. Ink4a-Arf loss cooperates with KRas activation in astrocytes and neural progenitors to generate glioblastomas of various morphologies depending on activated Akt. *Cancer Res* (2002) 62(19):5551–8.
129. Xu L, Chen Y, Dutra-Clarke M, Mayakonda A, Hazawa M, Savinoff SE, et al. BCL6 promotes glioma and serves as a therapeutic target. *Proc Natl Acad Sci USA* (2017) 114(15):3981–6. doi: 10.1073/pnas.1609758114
130. Welby JP, Kaptzan T, Wohl A, Peterson TE, Raghunathan A, Brown DA, et al. Current Murine Models and New Developments in H3K27M Diffuse Midline Gliomas. *Front Oncol* (2019) 9:92. doi: 10.3389/fonc.2019.00092
131. Chen F, Becker A, LoTurco J. Overview of Transgenic Glioblastoma and Oligoastrocytoma CNS Models and Their Utility in Drug Discovery. *Curr Protoc Pharmacol* (2016) 72:14 37 1–14 37 12. doi: 10.1002/0471141755.ph1437s72
132. Casey MJ, Stewart RA. Pediatric Cancer Models in Zebrafish. *Trends Cancer* (2020) 6(5):407–18. doi: 10.1016/j.trecan.2020.02.006
133. Wu G, Diaz AK, Paugh BS, Rankin SL, Ju B, Li Y, et al. The genomic landscape of diffuse intrinsic pontine glioma and pediatric non-brainstem high-grade glioma. *Nat Genet* (2014) 46(5):444–50. doi: 10.1038/ng.2938
134. Buczkowicz P, Hoeman C, Rakopoulos P, Pajovic S, Letourneau L, Dzamba M, et al. Genomic analysis of diffuse intrinsic pontine gliomas identifies three molecular subgroups and recurrent activating ACVR1 mutations. *Nat Genet* (2014) 46(5):451–6. doi: 10.1038/ng.2936
135. Eden CJ, Ju B, Murugesan M, Phoenix TN, Nimmervoll B, Tong Y, et al. Orthotopic models of pediatric brain tumors in zebrafish. *Oncogene* (2015) 34(13):1736–42. doi: 10.1038/onc.2014.107
136. Idilli AI, Pagani F, Kerschbamer E, Berardinelli F, Bernabe M, Cayuela ML, et al. Changes in the Expression of Pre-Replicative Complex Genes in hTERT and ALT Pediatric Brain Tumors. *Cancers (Basel)* (2020) 12(4):1028. doi: 10.3390/cancers12041028
137. Casey MJ, Modzelewska K, Anderson D, Goodman J, Boer EF, Jimenez L, et al. Transplantation of Zebrafish Pediatric Brain Tumors into Immune-competent Hosts for Long-term Study of Tumor Cell Behavior and Drug Response. *J Vis Exp* (2017) 123:e55712. doi: 10.3791/55712
138. Othman RT, Kimishi I, Bradshaw TD, Storer LC, Korshunov A, Pfister SM, et al. Overcoming multiple drug resistance mechanisms in medulloblastoma. *Acta Neuropathol Commun* (2014) 2:57. doi: 10.1186/2051-5960-2-57
139. Triscott J, Lee C, Foster C, Manoranjan B, Pambid MR, Berns R, et al. Personalizing the treatment of pediatric medulloblastoma: Polo-like kinase 1 as a molecular target in high-risk children. *Cancer Res* (2013) 73(22):6734–44. doi: 10.1158/0008-5472.CAN-12-43310008-5472.CAN-12-4331
140. Keles GE, Berger MS, Srinivasan J, Kolstoe DD, Bobola MS, Silber JR. Establishment and characterization of four human medulloblastoma-derived cell lines. *Oncol Res* (1995) 7(10-11):493–503.
141. Snuderl M, Batista A, Kirkpatrick ND, Ruiz de Almodovar C, Riedemann L, Walsh EC, et al. Targeting placental growth factor/neuropilin 1 pathway inhibits growth and spread of medulloblastoma. *Cell* (2013) 152(5):1065–76. doi: 10.1016/j.cell.2013.01.036
142. Friedman HS, Burger PC, Bigner SH, Trojanowski JQ, Brodeur GM, He XM, et al. Phenotypic and genotypic analysis of a human medulloblastoma cell line and transplantable xenograft (D341 Med) demonstrating amplification of c-myc. *Am J Pathol* (1988) 130(3):472–84.
143. Friedman GK, Moore BP, Nan L, Kelly VM, Etminan T, Langford CP, et al. Pediatric medulloblastoma xenografts including molecular subgroup 3 and CD133+ and CD15+ cells are sensitive to killing by oncolytic herpes simplex viruses. *Neuro Oncol* (2016) 18(2):227–35. doi: 10.1093/neuonc/nov123
144. Thompson EM, Keir ST, Venkatraman T, Lascola C, Yeom KW, Nixon AB, et al. The role of angiogenesis in Group 3 medulloblastoma pathogenesis and survival. *Neuro Oncol* (2017) 19(9):1217–27. doi: 10.1093/neuonc/nox033
145. Ivanov DP, Coyle B, Walker DA, Grabowska AM. In vitro models of medulloblastoma: Choosing the right tool for the job. *J Biotechnol* (2016) 236:10–25. doi: 10.1016/j.jbiotec.2016.07.028S0168-1656(16)31438-9
146. Friedman HS, Burger PC, Bigner SH, Trojanowski JQ, Wikstrand CJ, Halperin EC, et al. Establishment and characterization of the human medulloblastoma cell line and transplantable xenograft D283 Med. *J Neuropathol Exp Neurol* (1985) 44 (6):592–605. doi: 10.1097/00005072-198511000-00005
147. Northcott PA, Shih DJ, Peacock J, Garzia L, Morrissy AS, Zichner T, et al. Subgroup-specific structural variation across 1,000 medulloblastoma genomes. *Nature* (2012) 488(7409):49–56. doi: 10.1038/nature11327

148. Milde T, Lodrini M, Savelyeva L, Korshunov A, Kool M, Brueckner LM, et al. HD-MB03 is a novel Group 3 medulloblastoma model demonstrating sensitivity to histone deacetylase inhibitor treatment. *J Neurooncol* (2012) 110(3):335–48. doi: 10.1007/s11060-012-0978-1
149. Bandopadhayay P, Bergthold G, Nguyen B, Schubert S, Gholamin S, Tang Y, et al. BET bromodomain inhibition of MYC-amplified medulloblastoma. *Clin Cancer Res* (2014) 20(4):912–25. doi: 10.1158/1078-0432.CCR-13-2281
150. Xu J, Margol A, Asgharzadeh S, Erdreich-Epstein A. Pediatric brain tumor cell lines. *J Cell Biochem* (2015) 116(2):218–24. doi: 10.1002/jcb.24976
151. Xu J, Erdreich-Epstein A, Gonzalez-Gomez I, Melendez EY, Smbatyan G, Moats RA, et al. Novel cell lines established from pediatric brain tumors. *J Neurooncol* (2012) 107(2):269–80. doi: 10.1007/s11060-011-0756-5
152. Hashizume R, Smirnov I, Liu S, Phillips JJ, Hyer J, McKnight TR, et al. Characterization of a diffuse intrinsic pontine glioma cell line: implications for future investigations and treatment. *J Neurooncol* (2012) 110(3):305–13. doi: 10.1007/s11060-012-0973-6
153. Chan KM, Fang D, Gan H, Hashizume R, Yu C, Schroeder M, et al. The histone H3.3K27M mutation in pediatric glioma reprograms H3K27 methylation and gene expression. *Genes Dev* (2013) 27(9):985–90. doi: 10.1101/gad.217778.113
154. Grasso CS, Tang Y, Truffaux N, Berlow NE, Liu L, Debily MA, et al. Functionally defined therapeutic targets in diffuse intrinsic pontine glioma. *Nat Med* (2015) 21(6):555–9. doi: 10.1158/1535-7163.TARG-15-LB-B06
155. Donson AM, Amani V, Warner EA, Griesinger AM, Witt DA, Levy JMM, et al. Identification of FDA-Approved Oncology Drugs with Selective Potency in High-Risk Childhood Ependymoma. *Mol Cancer Ther* (2018) 17(9):1984–94. doi: 10.1158/1535-7163.MCT-17-1185
156. Lee J, Kotliarova S, Kotliarov Y, Li A, Su Q, Donin NM, et al. Tumor stem cells derived from glioblastomas cultured in bFGF and EGF more closely mirror the phenotype and genotype of primary tumors than do serum-cultured cell lines. *Cancer Cell* (2006) 9(5):391–403. doi: 10.1016/j.ccr.2006.03.030
157. Kapalczynska M, Kolenda T, Przybyla W, Zajaczkowska M, Teresiak A, Filas V, et al. 2D and 3D cell cultures - a comparison of different types of cancer cell cultures. *Arch Med Sci* (2018) 14(4):910–9. doi: 10.5114/aoms.2016.63743
158. Caragher S, Chalmers AJ, Gomez-Roman N. Glioblastoma's Next Top Model: Novel Culture Systems for Brain Cancer Radiotherapy Research. *Cancers (Basel)* (2019) 11(1):44. doi: 10.3390/cancers11010044
159. Monje M, Mitra SS, Freret ME, Raveh TB, Kim J, Masek M, et al. Hedgehog-responsive candidate cell of origin for diffuse intrinsic pontine glioma. *Proc Natl Acad Sci USA* (2011) 108(11):4453–8. doi: 10.1073/pnas.1101657108
160. Laks DR, Masterman-Smith M, Visnyei K, Angenieux B, Orozco NM, Foran I, et al. Neurosphere formation is an independent predictor of clinical outcome in malignant glioma. *Stem Cells* (2009) 27(4):980–7. doi: 10.1002/stem.15
161. Xiao W, Sohrabi A, Seidlits SK. Integrating the glioblastoma microenvironment into engineered experimental models. *Future Sci OA* (2017) 3(3):FSO189. doi: 10.4155/fsoa-2016-0094
162. Langhans SA. Three-Dimensional in Vitro Cell Culture Models in Drug Discovery and Drug Repositioning. *Front Pharmacol* (2018) 9:6. doi: 10.3389/fphar.2018.00006
163. Lovett ML, Nieland TJF, Dingle YTL, Kaplan DL. Innovations in 3D Tissue Models of Human Brain Physiology and Diseases. *Adv Funct Mater* (2020) n/a(n/a):1909146. doi: 10.1002/adfm.201909146
164. Amani V, Donson AM, Lummus SC, Prince EW, Griesinger AM, Witt DA, et al. Characterization of 2 Novel Ependymoma Cell Lines With Chromosome 1q Gain Derived From Posterior Fossa Tumors of Childhood. *J Neuropathol Exp Neurol* (2017) 76(7):595–604. doi: 10.1093/jnen/nlx040
165. Nath S, Devi GR. Three-dimensional culture systems in cancer research: Focus on tumor spheroid model. *Pharmacol Ther* (2016) 163:94–108. doi: 10.1016/j.pharmthera.2016.03.013S0163-7258(16)30021-3
166. Sant S, Johnston PA. The production of 3D tumor spheroids for cancer drug discovery. *Drug Discovery Today Technol* (2017) 23:27–36. doi: 10.1016/j.ddtec.2017.03.002
167. Costa EC, Moreira AF, de Melo-Diogo D, Gaspar VM, Carvalho MP, Correia IJ. 3D tumor spheroids: an overview on the tools and techniques used for their analysis. *Biotechnol Adv* (2016) 34(8):1427–41. doi: 10.1016/j.biotechadv.2016.11.002
168. Frantz C, Stewart KM, Weaver VM. The extracellular matrix at a glance. *J Cell Sci* (2010) 123(Pt 24):4195–200. doi: 10.1242/jcs.023820
169. Mouw JK, Ou G, Weaver VM. Extracellular matrix assembly: a multiscale deconstruction. *Nat Rev Mol Cell Biol* (2014) 15(12):771–85. doi: 10.1038/nrm3902nrm3902
170. Hughes CS, Postovit LM, Lajoie GA. Matrigel: a complex protein mixture required for optimal growth of cell culture. *Proteomics* (2010) 10(9):1886–90. doi: 10.1002/pmic.200900758
171. Wang C, Tong X, Yang F. Bioengineered 3D brain tumor model to elucidate the effects of matrix stiffness on glioblastoma cell behavior using PEG-based hydrogels. *Mol Pharm* (2014) 11(7):2115–25. doi: 10.1021/mp5000828
172. Worthington P, Drake KM, Li Z, Napper AD, Pochan DJ, Langhans SA. Beta-hairpin hydrogels as scaffolds for high-throughput drug discovery in three-dimensional cell culture. *Anal Biochem* (2017) 535:25–34. doi: 10.1016/j.ab.2017.07.024
173. Worthington P, Drake KM, Li Z, Napper AD, Pochan DJ, Langhans SA. Implementation of a High-Throughput Pilot Screen in Peptide Hydrogel-Based Three-Dimensional Cell Cultures. *SLAS Discovery* (2019) 24(7):714–23. doi: 10.1177/2472555219844570
174. Tang-Schomer MD, White JD, Tien LW, Schmitt LI, Valentin TM, Graziano DJ, et al. Bioengineered functional brain-like cortical tissue. *Proc Natl Acad Sci USA* (2014) 111(38):13811–6. doi: 10.1073/pnas.1324214111
175. Chwalek K, Sood D, Cantley WL, White JD, Tang-Schomer M, Kaplan DL. Engineered 3D Silk-collagen-based Model of Polarized Neural Tissue. *J Vis Exp* (2015) 105):e52970. doi: 10.3791/52970
176. Sood D, Tang-Schomer M, Pouli D, Mizzoni C, Raia N, Tai A, et al. 3D extracellular matrix microenvironment in bioengineered tissue models of primary pediatric and adult brain tumors. *Nat Commun* (2019) 10(1):4529. doi: 10.1038/s41467-019-12420-1
177. Li Q, Lin H, Rauch J, Deleyrolle LP, Reynolds BA, Viljoen HJ, et al. Scalable Culturing of Primary Human Glioblastoma Tumor-Initiating Cells with a Cell-Friendly Culture System. *Sci Rep* (2018) 8(1):3531. doi: 10.1038/s41598-018-21927-4
178. Lancaster MA, Knoblich JA. Generation of cerebral organoids from human pluripotent stem cells. *Nat Protoc* (2014) 9(10):2329–40. doi: 10.1038/nprot.2014.158
179. Ballabio C, Anderle M, Gianesello M, Lago C, Miele E, Cardano M, et al. Modeling medulloblastoma *in vivo* and with human cerebellar organoids. *Nat Commun* (2020) 11(1):583. doi: 10.1038/s41467-019-13989-3
180. Ogawa J, Pao GM, Shokhirev MN, Verma IM. Glioblastoma Model Using Human Cerebral Organoids. *Cell Rep* (2018) 23(4):1220–9. doi: 10.1016/j.celrep.2018.03.105
181. Linkous A, Balamatsias D, Snuderl M, Edwards L, Miyaguchi K, Milner T, et al. Modeling Patient-Derived Glioblastoma with Cerebral Organoids. *Cell Rep* (2019) 26(12):3203–11 e5. doi: 10.1016/j.celrep.2019.02.063
182. Hubert CG, Rivera M, Spangler LC, Wu Q, Mack SC, Prager BC, et al. A Three-Dimensional Organoid Culture System Derived from Human Glioblastomas Recapitulates the Hypoxic Gradients and Cancer Stem Cell Heterogeneity of Tumors Found In Vivo. *Cancer Res* (2016) 76(8):2465–77. doi: 10.1158/0008-5472.CAN-15-2402
183. Drost J, Clevers H. Organoids in cancer research. *Nat Rev Cancer* (2018) 18(7):407–18. doi: 10.1038/s41568-018-0007-6
184. Perrin S. Preclinical research: Make mouse studies work. *Nature* (2014) 507(7493):423–5. doi: 10.1038/507423a
185. Jessa S, Blanchet-Cohen A, Krug B, Vladoiu M, Coutelier M, Faury D, et al. Stalled developmental programs at the root of pediatric brain tumors. *Nat Genet* (2019) 51(12):1702–13. doi: 10.1038/s41588-019-0531-7

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

ORIGINAL RESEARCH
published: 01 April 2021
doi: 10.3389/fonc.2021.591507

Identification of Ten Core Hub Genes as Potential Biomarkers and Treatment Target for Hepatoblastoma

Rui Sun[1,2], *Simin Li*[3], *Ke Zhao*[4], *Mei Diao*[1] *and Long Li*[1]*

[1] Department of Pediatric Surgery, Capital Institute of Pediatrics, Beijing, China, [2] Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China, [3] Stomatological Hospital, Southern Medical University, Guangzhou, China, [4] Department of Ophthalmology, Ningbo Hangzhou Bay Hospital, Ningbo, China

OPEN ACCESS

Edited by:
Jing He,
Guangzhou Medical University, China

Reviewed by:
Xiaoxi Meng,
University of Minnesota Twin Cities, United States
Xiao Han,
Fuzhou University, China
Wenlong Ren,
Nantong University, China

***Correspondence:**
Long Li
lilong_pediatric@163.com

Specialty section:
This article was submitted to Pediatric Oncology, a section of the journal Frontiers in Oncology

Received: *04 August 2020*
Accepted: *25 February 2021*
Published: *01 April 2021*

Citation:
Sun R, Li S, Zhao K, Diao M and Li L (2021) Identification of Ten Core Hub Genes as Potential Biomarkers and Treatment Target for Hepatoblastoma. Front. Oncol. 11:591507. doi: 10.3389/fonc.2021.591507

Background: This study aimed to systematically investigate gene signatures for hepatoblastoma (HB) and identify potential biomarkers for its diagnosis and treatment.

Materials and Methods: GSE131329 and GSE81928 were obtained from the Gene Expression Omnibus (GEO) database. Differentially expressed genes (DEGs) between hepatoblastoma and normal samples were identified using the Limma package in R. Then, the similarity of network traits between two sets of genes was analyzed by weighted gene correlation network analysis (WGCNA). Cytoscape was used to visualize and select hub genes. PPI network of hub genes was construed by Cytoscape. GO enrichment and KEGG pathway analyses of hub genes were carried out using ClueGO. The random forest classifier was constructed based on the hub genes using the GSE131329 dataset as the training set, and its reliability was validated using the GSE81928 dataset. The resulting core hub genes were combined with the InnateDB database to identify the innate core genes.

Results: A total of 4244 DEGs in HB were identified. WGCNA identified four modules that were significantly correlated with the disease status. A total of 114 hub genes were obtained within the top 20 genes of each node rank. 6982 relation pairs and 3700 nodes were contained in the PPI network of 114 hub genes. GO enrichment and KEGG pathway analyses of hub genes were focused on MAPK, cell cycle, p53, and other crucial pathways involved in HB. A random forest classifier was constructed using the 114 hub genes as feature genes, resulting in a 95.5% true positive rate when classifying HB and normal samples. A total of 35 core hub genes were obtained through the mean decrease in accuracy and mean decrease Gini of the random forest model. The classification efficiency of the random forest model was 81.4%. Finally, *CDK1*, *TOP2A*, *ADRA1A*,

FANCI, *XRCC1*, *TPX2*, *CCNB2*, *CDK4*, *GLYATL1*, and *CFHR3* were identified by cross-comparison with the InnateDB database.

Conclusion: Our study established a random forest classifier that identified 10 core genes in HB. These findings may be beneficial for the diagnosis, prediction, and targeted therapy of HB.

Keywords: hepatoblastoma, gene expression omnibus, random forest classifier, nomogram, diagnosis

INTRODUCTION

Hepatoblastoma (HB) is the most common pediatric liver tumor, affecting mainly children under 4 years of age (1). Although its incidence has increased markedly over the last few decades, HB is a rare pediatric malignancy with an annual incidence of 1.5 cases per million (2). Complete surgical resection and chemotherapy have contributed to improving the survival rate of up to 80% in all diagnosed patients (3). However, the prognosis for patients with clinically advanced HB remains relatively low. Furthermore, surviving patients can suffer severe and lifelong side effects due to chemotherapy and immunosuppression (4). Lacking of an effective means of early diagnosis is the main reason contributed to the relative worse prognosis for patients with HB. At present, clinicians rely primarily on clinical symptoms, imaging, and alpha-fetoprotein levels to diagnose the disease. Among these methods, no novel biomarker had been showed except the conventional AFP levels. However, the sensitivity and specificity were not satisfied due to the various sources of AFP from different patients.In the previous study of Liu et al, it claimed there were 5 patients with a normal AFP level were diagnosed as HB (5). Consequently, novel biomarkers must be identified to develop efficient diagnostic methods and therapeutic strategies for patients affected with HB.

Recent studies have demonstrated that some RNAs are aberrantly expressed in HB thanks to the advancements in gene chips and high-throughput sequencing. A recent study reported by Liu et al. revealed that the increase of N6-Methyladenosine modification is an oncogenic mechanism in HB (6). Multiple studies have also shown that different genes, including genes encoding for long non-coding RNAs, are involved in the proliferation, apoptosis, and glutaminolysis of HB, such as zinc finger antisense 1 (7), 3-hydroxy-3-methylglutaryl-CoA synthase 1 (8), and TUG1 (9). Since the analysis pipeline, experimental methods, and sample size of each research are different, the conclusions have been controversial. Thus, a further bioinformatics exploration of data published in public databases could consolidate data and reveal novel additional genes associated with HB.

In this study, we investigated two HB datasets obtained from the Gene Expression Omnibus (GEO) database to identify reliable differentially expressed genes (DEGs) in HB. Through deep and comprehensive bioinformatics analysis, we identified hub genes, which we used to construct a diagnosis classification for HB. Moreover, we identified the core genes using our classification and cross-comparing it with the congenital immune-related genes present in the InnateDB database. The identification of a list of core genes may provide new diagnostic, prognostic, and potential therapeutic biomarkers for HB.

MATERIALS AND METHODS

Acquisition of Microarray Profiles

The flow chart for the study was showed in **Figure 1**. Microarrays that met the following criteria were collected: (1) studies including at least 20 samples and (2) examination expression of both cancerous tissue and adjacent noncancerous tissue from HB patients. Microarrays without useful data for analysis were excluded. Finally, 2 independent microarrays data, GSE131329 and GSE81928 databases, were obtained from the GEO database (http://www.ncbi.nlm.nih.gov/geo). The characteristics of the 2 datasets were presented in **Table 1**. Probes were converted into the corresponding gene symbols according to the annotation information in the dataset.

Since GSE131329 is chip data and GSE81928 is sequencing data, we used different procedures to deal with 2 datasets. For the GSE131329 dataset, platform annotation files were used to match probes to the gene symbol. If multiple probes matched a single gene, the median ranking value was used as the expression value. Then, the disease and normal gene expression spectrum of GSE131329 was constructed. For the GSE81928 dataset, we excluded from the analysis genes whose expression value was 0 in 80% of the samples. We analyzed a total of 17920 genes from the two data sets, which were then used for the subsequent analyses.

Identification of Differentially Expressed Genes (DEGs)

The Limma package in R was used to identify DEGs between HB and non-tumor samples. The cutoff value was set to $|Log_2FC$ (fold-change)$| > 0.58$ in both datasets to obtain more DEGS for further analysis in accordance with protocols of previous studies (11, 12). Because the experimental assay and platform of the 2 datasets were different, the P value was < 0.05 for GSE131329, and was < 0.01 for GSE81928 to obtain more significant DEGs, which was used the previous researches as reference (13, 14).

Weighted Gene Correlation Network Analysis (WGCNA)

WGCNA is a systematic biological method used to describe gene association patterns among different samples (15). It can be used

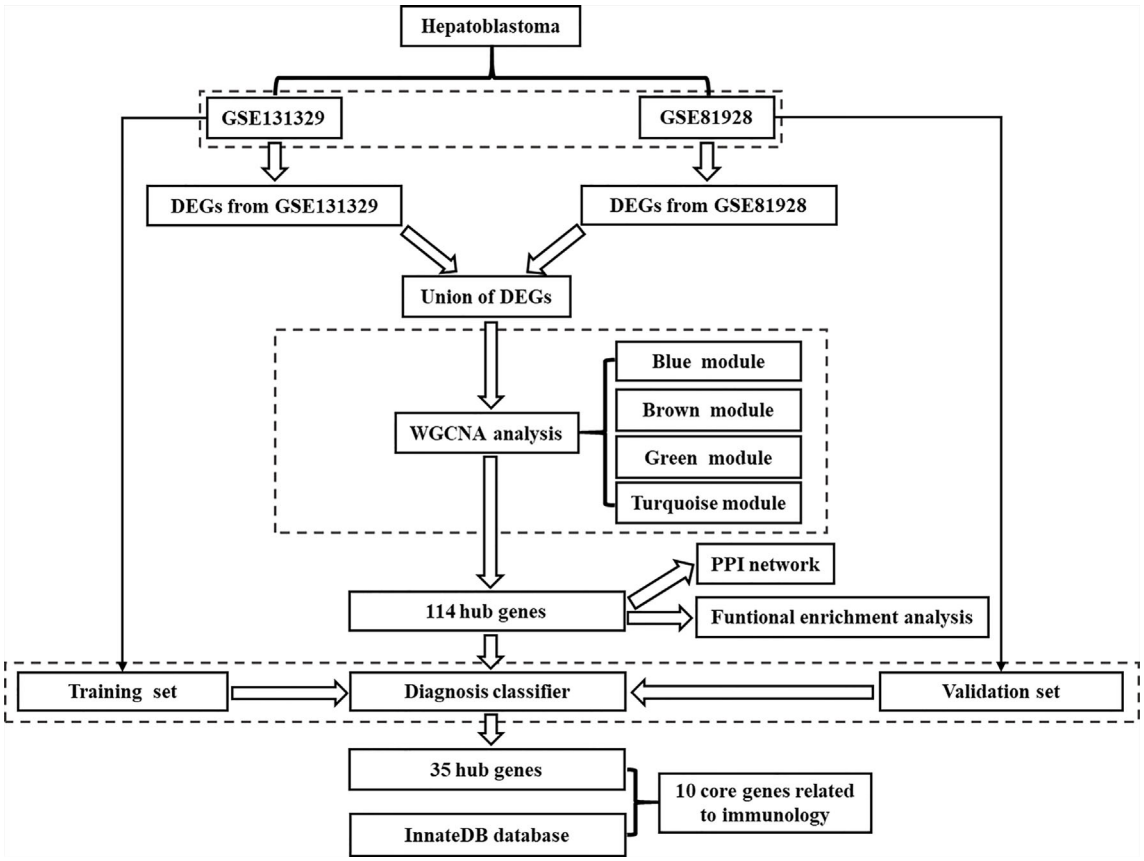


FIGURE 1 | Flowchart showing the protocol of the study. DEGs, differentially expressed genes; WGCNA, Weighted gene correlation network analysis; PPI, protein-protein interaction.

TABLE 1 | The characteristics of the 2 datasets in the study.

Datesets	Country	Researchers/ References	Experiment type	Tumor site	Sample size (normal/tumor)	Platform
GSE131329	Japan	Contributed by Hiyama E, et al.	Expression profiling by array	hepatoblastoma	67 (14/53)	GPL6244 [HuGene-1_0-st] Affymetrix Human Gene 1.0 ST Array
GSE81928	USA	(10)	Expression profiling by high throughput sequencing	hepatoblastoma	26 (3/23)	GPL16791 Illumina HiSeq 2500

to identify highly collaborative gene sets and to identify candidate biomarker genes or therapeutic targets based on gene set interconnection and the correlation between gene sets and phenotypes. Using the GSE131329 dataset as reference, the potential DEGs expression profile of HB was constructed. Then, we identified the related modules of HB, and analyzed the relationship between those modules and either HB or normal samples, using the WGCNA package in R. The identified network of HB modules was visualized using Cytoscape v. 3.8.0 (https://cytoscape.org/) to identify the hub genes in each module.

Protein-Protein Interaction (PPI) Network Construction of Hub Genes

In order to analyze the role of 114 hub genes in the global human biological network, we constructed a PPI network of modular genes. We downloaded and integrated human interaction protein data from the following database: HPRD release9 (http://www.hprd.org/), IntAct (http://www.ebi.ac.uk/intact/), MINT (http://mint.bio.uniroma2.it/mint/Welcome.do), BioGRID Release 3.4.132 (http://thebiogrid.org/), DIP (http://dip.doe-mbi.ucla.edu/dip/Main.cgi), String (https://string-db.org). We extracted 114 protein interaction pairs of hub genes

from the integrated human interaction protein pairs. Even if there was only one protein interacting with one of the 114 module genes, it would be extracted. The PPI network of these 114 modules was visualized by Cytoscape. In the network, 114 hub genes were marked with the color of their modules. Network analyzer, a Cytoscape tool, was used to calculate network topology properties.

Bioinformatic Analysis of Hub Genes

Gene ontology (GO) analysis was used to identify potential biological processes, cellular components, and molecular functions associated with DEGs. The Kyoto Encyclopedia of Genes and Genomes (KEGG) is a collection of databases for the systematic analysis of gene functions that link genomic information with higher-order functional information (16). GO enrichment and KEGG pathway analysis of the top 20 DEGs in HB were revealed using the ClueGO software. ClueGO software is a Cytoscape App that extracts representative functional biological information from a large list of genes or proteins (17). $P < 0.05$ was regarded as the cut-off criterion with statistic difference.

Construction and Validation of the HB Classifier

Random forest is a classification method that uses multiple trees to train and predict samples and is characterized by high accuracy (18). Therefore, we constructed a random forest model for HB, using GSE131329 as the training set, the top 20 genes in the module as the classification feature, and disease and normal samples as the variables. Then, we validated the model using the GSE81928 dataset as an independent validator. The model feature files of training set (**Supplementary Table 1**) and verification set (**Supplementary Table 2**) were shown in the **Supplementary**.

Cross-Comparison of Biological Markers of HB in InnateDB

InnateDB (http://www.innatedb.com) is a publicly available database of genes, proteins, and experimentally verified interacttions and signaling pathways involved in innate immunity (19). We intersected hub genes related to immunity in HB as revealed by our bioinformatics analysis with genes present in the InnateDB database.

RESULTS

Identification of DEGs

In total, 4244 DEGs (2839 in GSE131329 and 1863 in GSE81928) were identified between tumor and normal tissues (**Supplementary Table 3**), of which in GSE131329, 1368 were downregulated and 1471 were upregulated (**Supplementary Table 4**), while in GSE81928, 28 were downregulated and 1835 were upregulated (**Supplementary Table 5**). There were 453 overlapping DEGs of 2 datasets. The Venn diagrams (Available online: http://bioinformatics.psb.ugent.be/webtools/Venn/) was showed in **Figure 2**. The heat map and Volcano plot of the two datasets were showed in **Figures 3–D**.

WGCNA

Using the GSE131329 dataset and the WGCNA package in R to analyze the co-expression with default parameters, we constructed the expression spectrum of the 4244 DEGs. We obtained six different modules (indicated in blue, brown, green, turquoise and yellow) (**Figure 4**). The blue, brown, green, and turquoise modules were significantly correlated with HB and normal samples (**Figure 4**). The blue and brown modules were negatively correlated with HB disease, whereas the green and turquoise modules were positively correlated with HB disease. The modules contained 408 genes (blue), 188 genes (brown), 123 genes (green), and 666 genes (turquoise). Sample clustering is shown in **Figure 4**. The red component represents HB samples, while green represents the non-tumor samples.

Modules Network Construction and Hub Genes Identification

The network of HB-related modules (blue, brown, green, and turquoise modules) is shown in **Figures 5A–D**. Then, we analyzed the network using Cytoscape, selecting the top 20 genes of each module as the HB hub genes (genes with the same degree were taken out at the same time). Degree refers to the number of connections between one point and other points in the network. We identified a total of 114 hub genes (**Supplementary Table 6**). The larger the point is, the greater the degree of the representative node.

PPI Network Construction of Hub Genes

We constructed the PPI network of 114 hub genes. Finally, the network was consisted of 6982 relation pairs and 3700 nodes (**Figure 6**). For the topological properties of nodes, we arranged them in descending order according to the interaction degrees, and selected the top 20 genes to display, including RPS2, PPP2R1A, CDK1, FBL, PLK1, TRIM28, CDK4, PRMT1, SF3A2, ITCH, ANLN, USP15, CCNB1, EHMT2, CCNA2, USP9X, HCFC1, KIF11, TOP2A, as shown in **Table 2**. These genes play an important role in the global biological network.

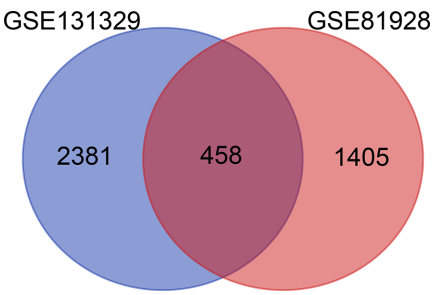


FIGURE 2 | Total of 4,244 DEGs identified from 2 datasets (2839 in GSE131329 and 1863 in GSE81928). DEGs, differentially expressed genes.

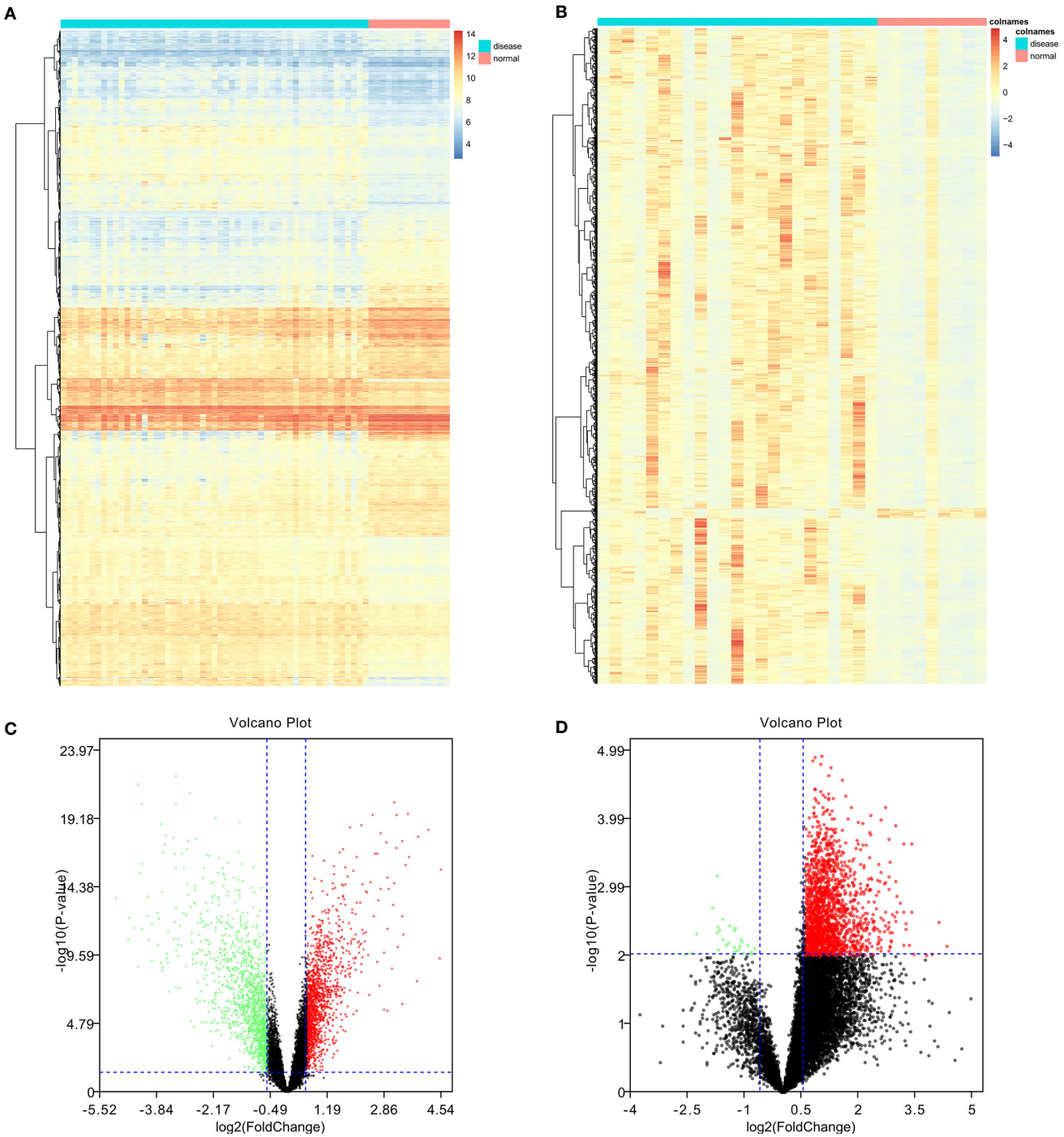


FIGURE 3 | Heat map of differentially expressed genes (DEGs) and volcano plot of genes in the two datasets. Upregulated DEGs are shown in red; downregulated DEGs are shown in green; non-DEGs are shown in black. **(A, B)** GSE131329 dataset; **(C, D)** GSE81928 dataset.

Bioinformatic Analysis of Hub Genes

We performed the GO enrichment and KEGG pathway analysis of the top 20 genes using ClueGO. Of the 114 hub genes identified, 21 were from the blue module, 24 from the brown module, 46 from the green module, and 23 from the turquoise module. GO function enrichment results are shown in **Figure 7**. Hub genes were enriched in multiple biological functions, including regulation of DNA demethylation, nuclear chromosome isolation, protein targeting to the peroxisomes, negative regulation of stress-activated MAPK cascade, signal

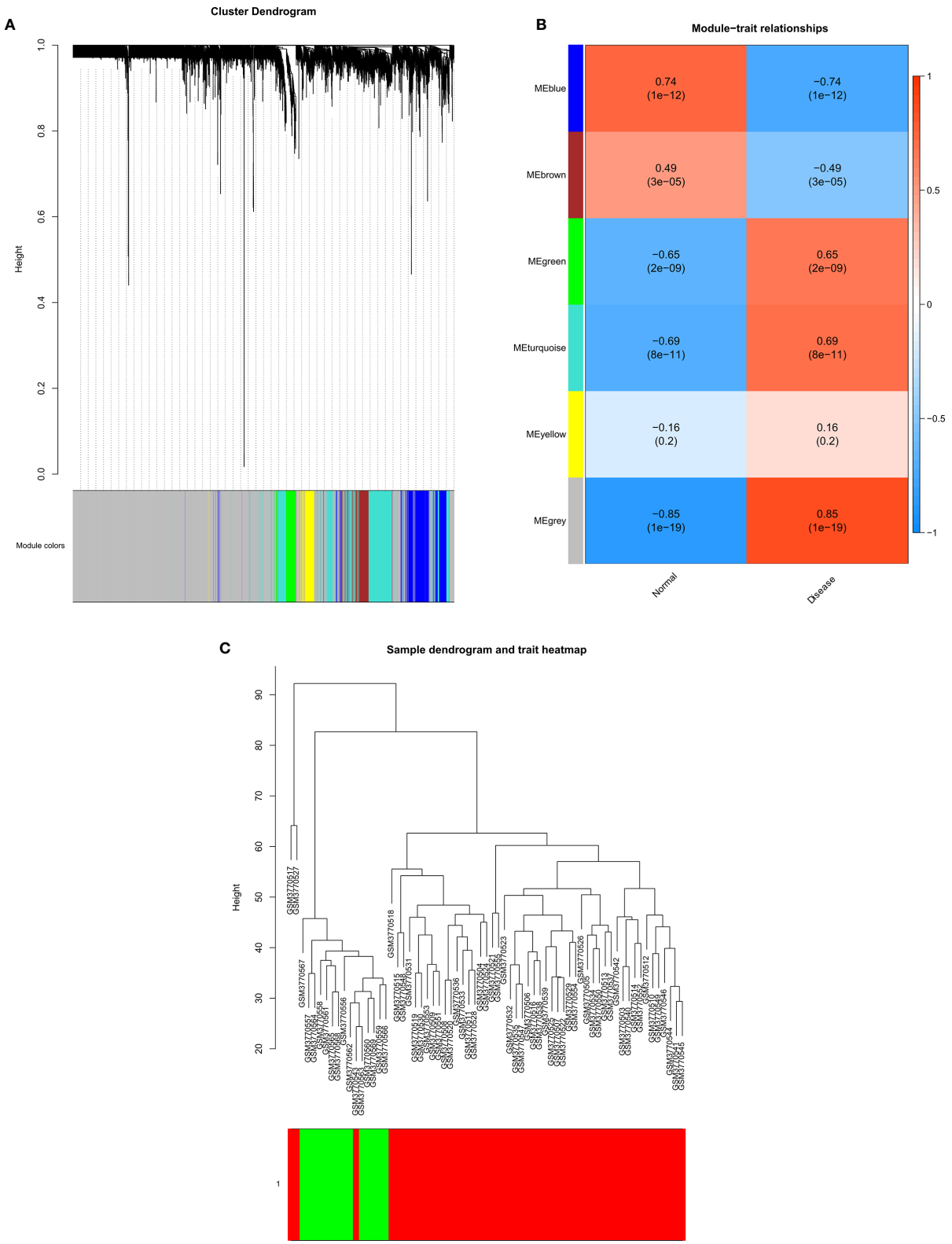


FIGURE 4 | Gene modules identified by WGCNA. **(A)** Cluster dendrogram of the coexpression network modules; **(B)** Gene relation between hepatoblastoma and normal samples; **(C)** Cluster tree of hepatoblastoma and normal samples.

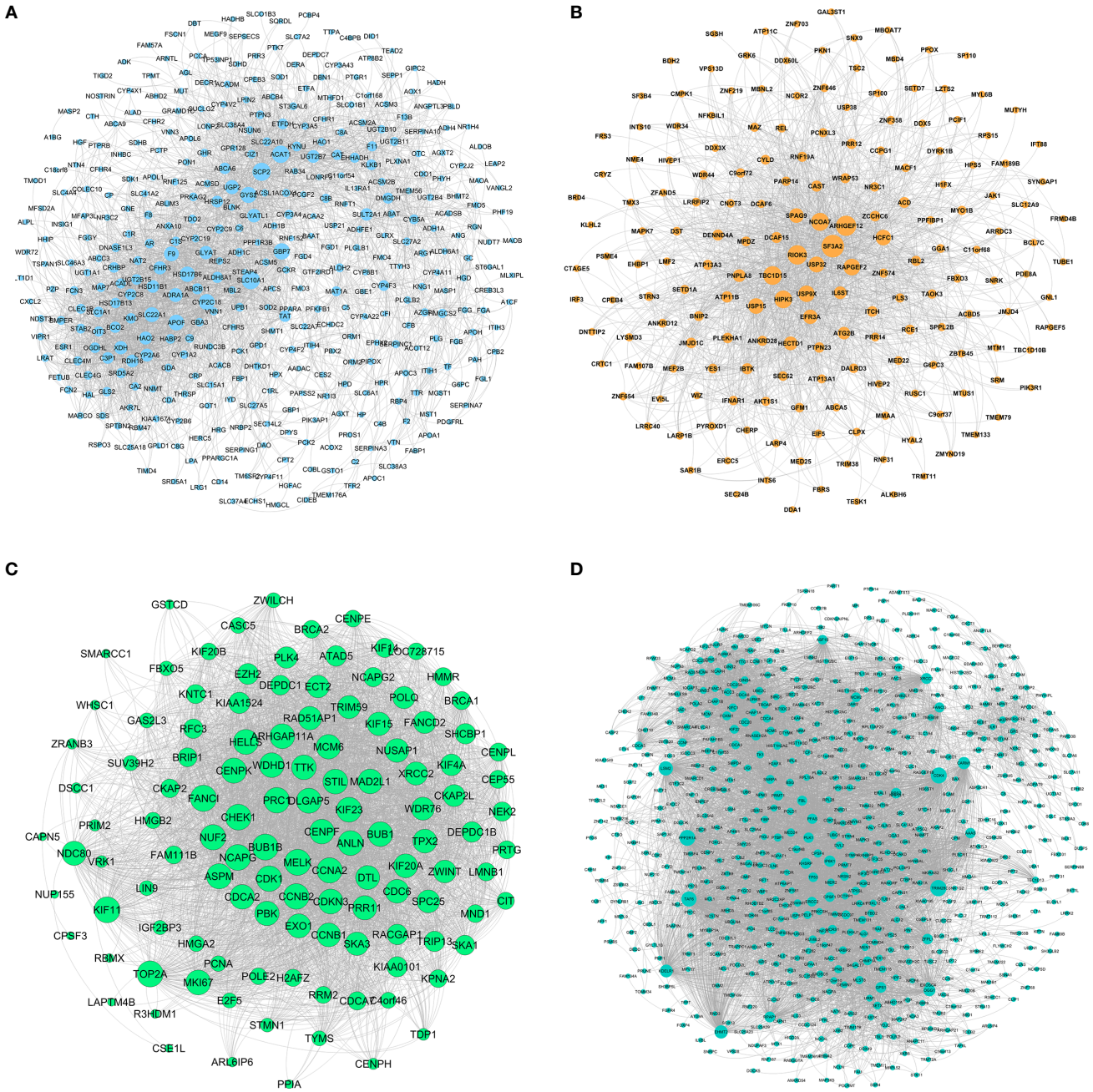


FIGURE 5 | Gene symbols and gene interaction in the four modules, as determined by ClueGO. **(A)** Blue module; **(B)** Brown module; **(C)** Green module; **(D)** Turquoise module.

transduction by p53 class mediator resulting in cell cycle arrest, toroid dehydrogenase activity with the CH-OH group acting as donors and NAD or NADP as acceptors et al. (**Supplementary Table 7**). The KEGG pathway analysis results showed that these hub genes also participated in the P53 signaling pathway, cell aging, cell cycle, meiotic maturation process of oocytes, progesterone-mediated oocyte maturation, steroid biosynthesis, retinol metabolism, chemical carcinogenesis, and other biological pathways (**Figure 7**) (**Supplementary Table 8**).

Construction and Validation of the HB Classification Method

The random forest method can calculate the importance of a single feature and screen the feature against the selected dataset.

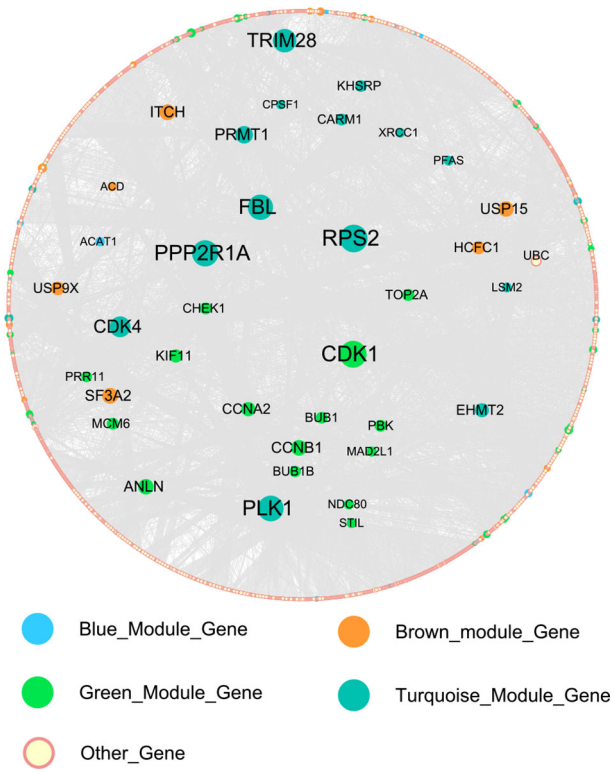


FIGURE 6 | The protein-protein interaction (PPI) network of the 114 hub genes.

TABLE 2 | Network topological characteristic of top 20 nodes in PPI network.

Gene	label	Degree	Average ShortestPath Length	BetweennessCentrality	ClosenessCentrality	ClusteringCoefficient	Stress	TopologicalCoefficient
RPS2	Turquoise	304	2.796578	0.088088	0.35758	0.002084	35651276	0.011482
PPP2R1A	Turquoise	285	2.733026	0.10859	0.365895	0.001112	32731778	0.008927
CDK1	Green	277	2.606464	0.086532	0.383662	0.007953	23033244	0.008038
FBL	Turquoise	272	2.764259	0.083119	0.361761	0.003337	26783624	0.009266
PLK1	Turquoise	271	2.749593	0.087714	0.36369	0.002952	24705604	0.009257
TRIM28	Turquoise	251	2.847094	0.07047	0.351235	0.000829	35607200	0.02181
CDK4	Turquoise	215	2.763987	0.064648	0.361796	0.002478	17600494	0.009869
PRMT1	Turquoise	174	2.833786	0.051374	0.352885	0.003654	15009342	0.013589
SF3A2	Brown	156	2.933188	0.048426	0.340926	0	11812770	0.028122
ITCH	Brown	155	2.893808	0.049782	0.345565	0.000922	10898838	0.016011
ANLN	Green	149	2.937534	0.055011	0.340422	0.000998	7721342	0.020761
USP15	Brown	146	2.891092	0.04818	0.34589	0.001606	9324218	0.015811
CCNB1	Green	144	2.747691	0.02827	0.363942	0.019814	5878754	0.013833
EHMT2	Turquoise	132	2.939164	0.038808	0.340233	0	8115194	0.02823
CCNA2	Green	126	2.771863	0.026821	0.360768	0.01981	4666808	0.015196
USP9X	Brown	125	2.898968	0.034266	0.34495	0.001419	9220858	0.020271
HCFC1	Brown	120	2.885117	0.032388	0.346606	0.003641	7730212	0.018792
KIF11	Green	118	2.847366	0.039087	0.351202	0.001883	5902744	0.013225
TOP2A	Green	106	2.892178	0.020279	0.34576	0.004672	7719634	0.024587

Therefore, we used the 114 hub genes as the feature, HB and normal as the variables, and the GSE131329 dataset as the training set to construct the model. The receiving operator curve (ROC) of the GES131329 training set is shown in **Figure 8**. The area under the curve was 0.955. The mean decrease accuracy (MDA) of the random forest model was positively correlated with the predictive variable, and the mean decrease Gini (MDG) is positively correlated with the most important variable (20). Therefore, 30 hub genes were established using MDA and MDG (**Figure 8**). Furthermore, a

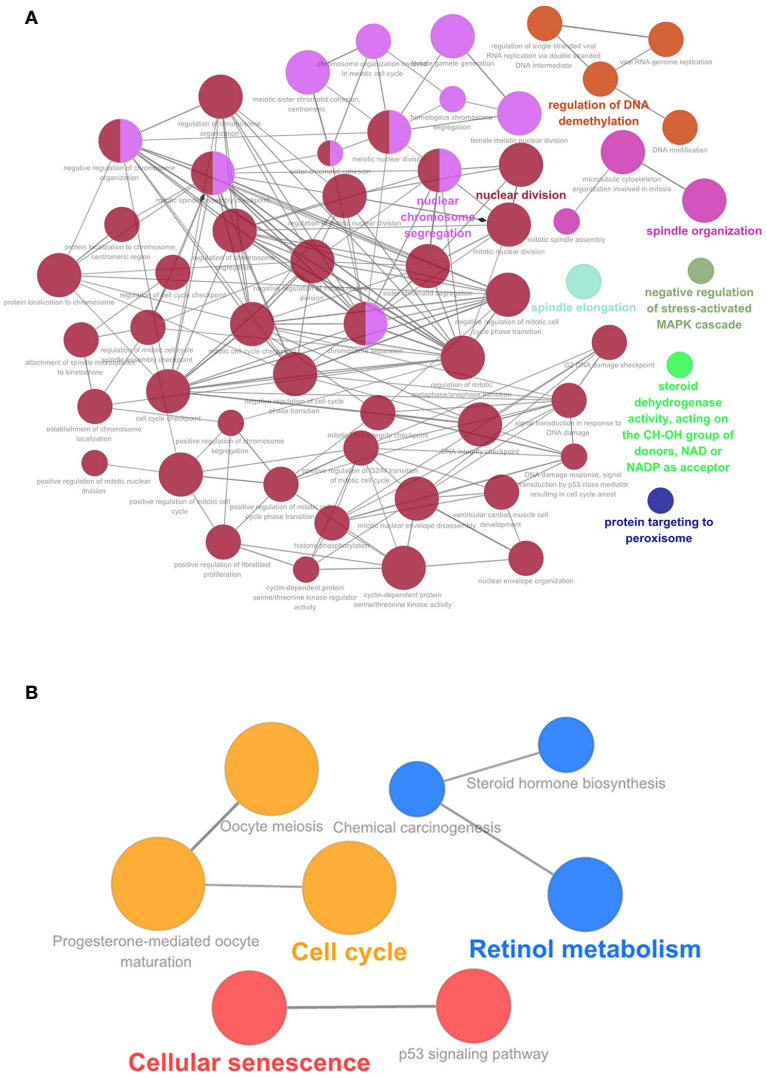


FIGURE 7 | **(A)** GO analysis of the 114 hub genes. **(B)** KEGG pathway of the 114 hub genes, GO.

total of 35 core genes of HB were obtained by cross-comparison with InnateDB database (**Supplementary Table 9**). The random forest model was then validated using the independent GSE81928 dataset, which was also contained the 114 hub genes. The area under the ROC curve was 0.814 (**Figure 8**).

Validation of HB Biological Markers Through InnateDB Cross-Comparison

We selected the immune-related hub gene, containing 35 genes, as the HB innate immune core genes and compared it with the immune-related genes present in the InnateDB database. We obtained 10 core genes: *CDK1*, *TOP2A*, *ADRA1A*, *FANCI*, *XRCC1*, *TPX2*, *CCNB2*, *CDK4*, *GLYATL1*, and *CFHR3*. For nine genes at least one molecular interaction was annotated in the InnateDB database, except *GLYATL1*. These interactions were mainly protein-protein and DNA-protein, as shown in **Supplementary Figures 1–9**.

DISCUSSION

In this study, we established for the first time a diagnosis classifier model based on the random tree method for HB using 114 hub genes. We also validated the classification efficiency of this model using an independent dataset. Consequently, this model may contribute to improving the diagnosis of HB. We also performed GO and KEGG analyses, revealing that the identified hub genes were mainly involved in the p53 pathway and cell cycle. We also identified 10 core genes by cross-referencing our analysis with the InnateDB database. Among the 10 core genes, the molecular interactions for 9 genes were annotated, which may provide new therapeutic targets.

Among the 10 identified core genes, *CDK1* and *CDK4* were previously reported to be associated with HB (21, 22). *CDK1* and *CDK4* both belong to the family of cyclin-dependent kinases (CDKs). CDK complexes are critical regulatory enzymes that

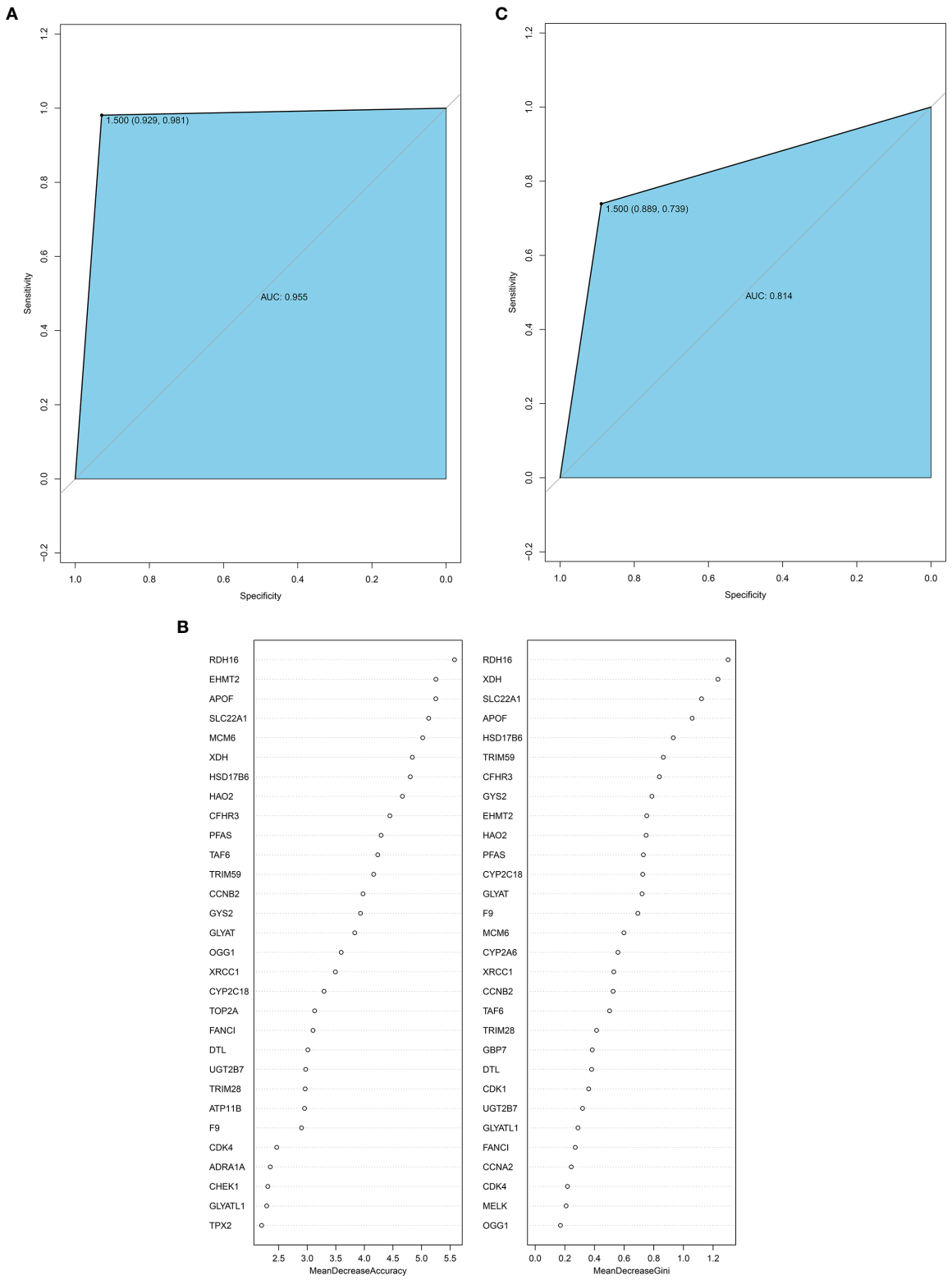


FIGURE 8 | **(A)** ROC curve for the GSE131329 dataset; **(B)** 30 hub genes from the random forest classifier extracted through MDA and MDG; **(C)** ROC curve for the GSE81928 dataset. AUC, area under the curve; ROC, receiver operating characteristic; MDA, mean decrease accuracy; MDG, mean decrease Gini.

drive the transition of different phases of the cell cycle and ensure successful cell division through their activity (23). Almost all malignant cells exhibit some features that derange the normal controls over the cell cycle (24). Therefore, various drugs targeting different CDKs have been developed and have been applied in the clinic over the past decades. *CDK1* can bind to different cyclins and regulate all the steps required for cell division (25). For this reason, *CDK1* is essential for mammalian cell proliferation (21) and is the only CDK that can initiate mitosis (26). *CDK1* is a key determinant of mitotic progression and thus it is also a pivotal tumorigenic event. It has been reported that treatment with a *CDK1* inhibitor could decrease tumor growth of HB and prolong the survival rate in an HB murine model (21). Therefore, *CDK1* is considered an ideal target for HB treatment. *CDK4* can mediate the transition from the G0 or G1 phase into the S phase of the cell cycle (27). The activity of *CDK4* is primarily controlled by its association with D-type cyclins, with cyclin D1 being the best characterized. Kim et al. revealed that *CDK4* and cyclin D1 were significantly overexpressed in HB tissues compared with normal tissues (22). They also suggested that *CDK4* may be correlated with tumorigenesis, tumor recurrence, and metastasis of HB. Although there is still no available *CDK4* inhibitor for HB, multiple selective *CDK4* inhibitors targeting other types of cancer have been used in the clinic. The progression-free survival rate of patients with estrogen receptor-positive breast cancer can improve when *CDK4/6* inhibitors are added to antiestrogen therapy (28). Therefore, the role of *CDK4* in HB progression and treatment requires further studies.

The role of the other 8 core genes in HB has never been reported before. Among them, 6 genes have been reported to be associated with hepatocellular carcinoma (HCC). TOP2A was one of the top 20 genes with the highest degree of interaction in the PPI network complex. *TOP2A* encodes a DNA topoisomerase that controls and alters the topologic states of intertwined DNA during anaphase. Therefore, *TOP2A* is involved in chromosome condensation and chromatid separation (29). Overexpression of *TOP2A* is correlated with a more aggressive tumor phenotype, microvascular invasion, and early age onset of HCC (30). Moreover, *TOP2A* has also been a valuable prognostic marker for tumor advancements, recurrences, and predictors of poor survival in a variety of cancers, such as breast, ovarian, colon, and small cell lung cancer (29). *ADRA1A* encodes the alpha-1 adrenergic receptor subtype with catecholamines ligands (31), which is located on chromosome 8p (32). *ADRA1A* can stimulate the sympathetic nervous system to compete with some functions (33). It was reported by Chen et al. that the mean methylation level of the *ADRA1A* promoter region was significantly increased in HCC tissues compared with the normal tissues (32). They also demonstrated that the mean methylation levels of the *ADRA1A* gene in HCC samples were not only associated with clinical characteristics but could also discriminate between HCC tissues and adjacent normal tissues, thus being suitable as a diagnostic marker. *XRCC1* is a DNA repair gene that plays a crucial role in maintaining genomic integrity and stability and in the pathogenesis and carcinogenesis of various type of cancer (34). *XRCC1* is significantly correlated with the number of tumors, tumor size, and location, and is also an independent risk factor for the poor prognosis of HCC (34, 35). *TPX2*, a nuclear proliferation microtubule-associated protein, is essential for spindle formation and stabilizes spindle microtubules (36). The overexpression of *TPX2* induces abnormal centrosome amplification and aneuploidy formation, leading to malignant transformation of cells (37). Multiple studies have shown that the expression levels of *TPX2* were significantly upregulated in HCC tissues compared with the adjacent normal tissues (36–38). They also confirmed that *TPX2* may improve the viability of HCC cells and inhibit cell apoptosis. However, knockdown of *TPX2* expression or TPX2 inhibition could reduce the migration and invasion ability of HCC cells. *CCNB2* was one of the top 20 genes with the highest degree of interaction in the PPI network complex. *CCNB2* belongs to the B-type cyclin family and regulates the activity of CDKs by binding to them during the cell cycle (23). The overexpression of *CCNB2* was positively correlated with tumor number, tumor size, tumor thrombus, and metastasis of HCC, which may contribute to the poor prognosis of HCC patients (39–41). However, *CCNB2* knockdown could slow cell growth and promote apoptosis of HCC cells, indicating that *CCNB2* may be a novel treatment marker (41). *CFHR3*, a member of the human factor H protein family, is a negative complement activation regulator, which is an essential component of the innate immune system (42). The expression level of *CFHR3* in HCC tissues was lower than that in normal tissues (43). In addition, the expression level of the *CFHR3* gene was the highest in the liver than in other organs (44). *CFHR3* is correlated to the HCC stage. In addition, the overall survival of patients affected with HCC was significantly better when *CFHR3* was highly expressed than when its expression was low (43, 44). Therefore, *CFHR3* may be a novel prognostic biomarker for HCC. Although these 6 genes were never reported in the context of HB, our bioinformatics analysis suggests that they deserve further attention as potential targets in HB.

FANCI and *GLYATL1* have never been reported in either HB or HCC. However, their abnormal expression has been found in other tumor types. *FANCI* has a key role in the Fanconi anemia DNA repair pathway, where it forms a heterodimer with *FANCD2* and recruits DNA repair proteins to promote the interstrand cross-link DNA damage repair (45). Moreover, *FANCI* may promote cellular metabolism when it is not needed for DNA repair, according to a recent study (46). *FANCI* mRNA and protein were both found to be overexpressed in lung adenocarcinoma tumor tissues compared with adjacent normal tissues (47). It was demonstrated that the expression level of *FANCI* was positively associated with lymphatic metastasis and distant metastasis of lung adenocarcinoma tumor, whereas knockdown of *FANCI* decreased lung adenocarcinoma tumor cell proliferation and invasion *in vitro*. *FANCI* has also been reported to regulate breast cancer survival (48). These findings suggest that *FANCI* has a novel oncogenic role and may be useful as a prognostic biomarker and/or therapeutic target for different tumors.

GLYATL1 belongs to the glycine-N-acyltransferase gene family and is normally expressed in the liver and kidney (49). *GLYATL1* encodes an enzyme with phenylacetyl-CoA glutamine N-acyltransferase activity, which regulates mitochondrial ATP production, glycine availability, CoASH availability, and the detoxification of various organic acids (50). In a previous study, the expression of *GLYATL1* was higher in localized prostate cancers than in benign prostatic tissue and metastatic prostate cancer (49, 51). This study also demonstrated that *GLYATL1* may be associated with the grade of prostate cancer since the expression of *GLYATL1* was significantly high in low-grade tumors. Therefore, *GLYATL1* could be a potential early-stage biomarker. In addition, *GLYATL* was also found to be overexpressed in ER-negative compared to ER-positive breast cancer (52).

We also conducted GO enrichment and KEGG pathway analysis to identify pathways correlated with the hub genes. KEGG pathway analysis revealed that the largest number of genes were enriched in the cell cycle, including 13 hub genes. Most of them, including *CDK1* (21), *CDK4* (27), *BUB1* (53), *BUB1B* (54), *CCNA2* (55), *CCNB1* (56), *CCNB2* (39), *CDC6* (57), *MAD2L1* (58), *MCM6* (59), and *PLK1* (60) have been already reported to be associated with cell cycle-related proliferation and tumor differentiation. GO analysis further showed that hub genes are involved in different cell cycle-related processes, including mitotic nuclear division, cell division, chromosome separation, sister chromatid cohesion, microtubule cytoskeleton organization involved in mitosis, and DNA integrity checkpoint. Furthermore, GO enrichment and KEGG pathway analysis also demonstrated that the hub genes were associated with the p53 signaling pathway, a tumor suppression pathway through a variety of responses, including cell-cycle arrest, apoptosis, senescence, and DNA repair (61, 62), suggesting that the p53 signaling pathway is also involved in the cell cycle. It was reported that p53 gene mutations may contribute to the development of sporadic HB (63). Moreover, hepatic p53 expression could cause lysis of implanted hepatoblastoma cells in a chimeric mouse (64). Although p53 may play a crucial role in HB development, the specific mechanism needs further studies. Taken together, based on the GO and KEGG analyses, we suggest that targeting the cell cycle could be a potential strategy for HB therapy. Compared with the traditional clinical manifestations, imaging, AFP and other diagnostic methods, our study considered the underlying genetic dysregulations. Genes are more objective and stable; thus, they may not be beneficial for early diagnosis.

CONCLUSION

In the present study, we established a 114 genes random forest classifier for HB and identified 10 core genes. These 10 core genes are closely related to the progression and prognosis of cancers and thus are also potential therapeutic targets. Our classifier model and the identified core genes may give novel insight into the diagnosis and development of therapeutic options for HB.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

RS performed the data analyses and wrote the manuscript. SL, KZ, and MD contributed significantly to data analyses and manuscript revision. LL conceived and designed the study. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by grants from the Special and Key Projects in the Pediatrics of Beijing Hospitals Authority and Pediatric Collaborative Development (No. XTZD20180302) and Fundamental Research Funds for the Central Universities (No. 3332019166).

ACKNOWLEDGMENTS

We thank those authors who provided us with the full text and relevant data from their studies.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.591507/full#supplementary-material

REFERENCES

1. Kremer N, Walther AE, Tiao GM. Management of hepatoblastoma: an update. *Curr Opin Pediatr* (2014) 26(3):362–9. doi: 10.1097/mop.0000000000000081
2. Meyers RL, Maibach R, Hiyama E, Häberle B, Krailo M, Rangaswami A, et al. Risk-stratified staging in paediatric hepatoblastoma: a unified analysis from the Children's Hepatic tumors International Collaboration. *Lancet Oncol* (2017) 18(1):122–31. doi: 10.1016/s1470-2045(16)30598-8
3. Hooks KB, Audoux J, Fazli H, Lesjean S, Ernault T, Dugot-Senant N, et al. New insights into diagnosis and therapeutic options for proliferative hepatoblastoma. *Hepatology* (2018) 68(1):89–102. doi: 10.1002/hep.29672
4. Carrillo-Reixach J, Torrens L, Simon-Coma M, Royo L, Domingo-Sabat M, Abril-Fornaguera J, et al. Epigenetic footprint enables molecular risk stratification of hepatoblastoma with clinical implications. *J Hepatol* (2020) 73(2):328–37. doi: 10.1016/j.jhep.2020.03.025
5. Liu W, Chen S, Liu B. Diagnostic and prognostic values of serum exosomal microRNA-21 in children with hepatoblastoma: A Chinese population-based

study. *Pediatr Surg Int* (2016) 32(11):1059–65. doi: 10.1007/s00383-016-3960-8
6. Liu L, Wang J, Sun G, Wu Q, Ma J, Zhang X, et al. m(6)A mRNA methylation regulates CTNNB1 to promote the proliferation of hepatoblastoma. *Mol Cancer* (2019) 18(1):188. doi: 10.1186/s12943-019-1119-7
7. Cui X, Wang Z, Liu L, Liu X, Zhang D, Li J, et al. The Long Non-coding RNA ZFAS1 Sponges miR-193a-3p to Modulate Hepatoblastoma Growth by Targeting RALY via HGF/c-Met Pathway. *Front Cell Dev Biol* (2019) 7:271. doi: 10.3389/fcell.2019.00271
8. Zhen N, Gu S, Ma J, Zhu J, Yin M, Xu M, et al. CircHMGCS1 Promotes Hepatoblastoma Cell Proliferation by Regulating the IGF Signaling Pathway and Glutaminolysis. *Theranostics* (2019) 9(3):900–19. doi: 10.7150/thno.29515
9. Dong R, Liu GB, Liu BH, Chen G, Li K, Zheng S, et al. Targeting long non-coding RNA-TUG1 inhibits tumor growth and angiogenesis in hepatoblastoma. *Cell Death Dis* (2016) 7(6):e2278. doi: 10.1038/cddis.2016.143
10. Valanejad L, Cast A, Wright M, Bissig KD, Karns R, Weirauch MT, et al. PARP1 activation increases expression of modified tumor suppressors and pathways underlying development of aggressive hepatoblastoma. *Commun Biol* (2018) 1:67. doi: 10.1038/s42003-018-0077-8
11. McCarthy DJ, Smyth GK. Testing significance relative to a fold-change threshold is a TREAT. *Bioinformatics* (2009) 25(6):765–71. doi: 10.1093/bioinformatics/btp053
12. Ji Z, He L, Rotem A, Janzer A, Cheng CS, Regev A, et al. Genome-scale identification of transcription factors that mediate an inflammatory network during breast cellular transformation. *Nat Commun* (2018) 9(1):2068. doi: 10.1038/s41467-018-04406-2
13. Liang Y, Zhang C, Dai DQ. Identification of differentially expressed genes regulated by methylation in colon cancer based on bioinformatics analysis. *World J Gastroenterol* (2019) 25(26):3392–407. doi: 10.3748/wjg.v25.i26.3392
14. Zhou Y, Yang L, Zhang X, Chen R, Chen X, Tang W, et al. Identification of Potential Biomarkers in Glioblastoma through Bioinformatic Analysis and Evaluating Their Prognostic Value. *BioMed Res Int* (2019) 2019:6581576. doi: 10.1155/2019/6581576
15. Langfelder P, Horvath S. WGCNA: An R package for weighted correlation network analysis. *BMC Bioinf* (2008) 9:559. doi: 10.1186/1471-2105-9-559
16. Kanehisa M, Goto S, Kawashima S, Nakaya A. The KEGG databases at GenomeNet. *Nucleic Acids Res* (2002) 30(1):42–6. doi: 10.1093/nar/30.1.42
17. Mlecnik B, Galon J, Bindea G. Comprehensive functional analysis of large lists of genes and proteins. *J Proteomics* (2018) 171:2–10. doi: 10.1016/j.jprot.2017.03.016
18. Speiser JL, Durkalski VL, Lee WM. Random forest classification of etiologies for an orphan disease. *Stat Med* (2015) 34(5):887–99. doi: 10.1002/sim.6351
19. Lynn DJ, Chan C, Naseer M, Yau M, Lo R, Sribnaia A, et al. Curating the innate immunity interactome. *BMC Syst Biol* (2010) 4:117. doi: 10.1186/1752-0509-4-117
20. Wang H, Yang F, Luo Z. An experimental study of the intrinsic stability of random forest variable importance measures. *BMC Bioinf* (2016) 17:60. doi: 10.1186/s12859-016-0900-5
21. Goga A, Yang D, Tward AD, Morgan DO, Bishop JM. Inhibition of CDK1 as a potential therapy for tumors over-expressing MYC. *Nat Med* (2007) 13(7):820–7. doi: 10.1038/nm1606
22. Kim H, Ham EK, Kim YI, Chi JG, Lee HS, Park SH, et al. Overexpression of cyclin D1 and cdk4 in tumorigenesis of sporadic hepatoblastomas. *Cancer Lett* (1998) 131(2):177–83. doi: 10.1016/s0304-3835(98)00151-7
23. Asghar U, Witkiewicz AK, Turner NC, Knudsen ES. The history and future of targeting cyclin-dependent kinases in cancer therapy. *Nat Rev Drug Discovery* (2015) 14(2):130–46. doi: 10.1038/nrd4504
24. Malumbres M, Barbacid M. To cycle or not to cycle: a critical decision in cancer. *Nat Rev Cancer* (2001) 1(3):222–31. doi: 10.1038/35106065
25. Xie B, Wang S, Jiang N, Li JJ. Cyclin B1/CDK1-regulated mitochondrial bioenergetics in cell cycle progression and tumor resistance. *Cancer Lett* (2019) 443:56–66. doi: 10.1016/j.canlet.2018.11.019
26. Santamaría D, Barrière C, Cerqueira A, Hunt S, Tardy C, Newton K, et al. Cdk1 is sufficient to drive the mammalian cell cycle. *Nature* (2007) 448(7155):811–5. doi: 10.1038/nature06046
27. O'Leary B, Finn RS, Turner NC. Treating cancer with selective CDK4/6 inhibitors. *Nat Rev Clin Oncol* (2016) 13(7):417–30. doi: 10.1038/nrclinonc.2016.26
28. Goel S, DeCristo MJ, McAllister SS, Zhao JJ. CDK4/6 Inhibition in Cancer: Beyond Cell Cycle Arrest. *Trends Cell Biol* (2018) 28(11):911–25. doi: 10.1016/j.tcb.2018.07.002
29. Jain M, Zhang L, He M, Zhang YQ, Shen M, Kebebew E. TOP2A is overexpressed and is a therapeutic target for adrenocortical carcinoma. *Endocr Relat Cancer* (2013) 20(3):361–70. doi: 10.1530/erc-12-0403
30. Wong N, Yeo W, Wong WL, Wong NL, Chan KY, Mo FK, et al. TOP2A overexpression in hepatocellular carcinoma correlates with early age onset, shorter patients survival and chemoresistance. *Int J Cancer* (2009) 124(3):644–52. doi: 10.1002/ijc.23968
31. Chen G, Fan X, Li Y, He L, Wang S, Dai Y, et al. Promoter aberrant methylation status of ADRA1A is associated with hepatocellular carcinoma. *Epigenetics* (2020) 15(6-7):684–701. doi: 10.1080/15592294.2019.1709267
32. Loo SK, Fisher SE, Francks C, Ogdie MN, MacPhie IL, Yang M, et al. Genome-wide scan of reading ability in affected sibling pairs with attention-deficit/hyperactivity disorder: unique and shared genetic effects. *Mol Psychiatry* (2004) 9(5):485–93. doi: 10.1038/sj.mp.4001450
33. Peng L, Peng W, Hu P, Zhang HF. Clinical significance of expression levels of serum ADRA1A in hysterocarcinoma patients. *Oncol Lett* (2018) 15(6):9162–6. doi: 10.3892/ol.2018.8465
34. Guan Q, Chen Z, Chen Q, Zhi X. XRCC1 and XPD polymorphisms and their relation to the clinical course in hepatocarcinoma patients. *Oncol Lett* (2017) 14(3):2783–8. doi: 10.3892/ol.2017.6522
35. Xiong Y, Zhang Q, Ye J, Pan S, Ge L. Associations between three XRCC1 polymorphisms and hepatocellular carcinoma risk: A meta-analysis of case-control studies. *PloS One* (2018) 13(11):e0206853. doi: 10.1371/journal.pone.0206853
36. Huang DH, Jian J, Li S, Zhang Y, Liu LZ. TPX2 silencing exerts anti–tumor effects on hepatocellular carcinoma by regulating the PI3K/AKT signaling pathway. *Int J Mol Med* (2019) 44(6):2113–22. doi: 10.3892/ijmm.2019.4371
37. Liang B, Jia C, Huang Y, He H, Li J, Liao H, et al. TPX2 Level Correlates with Hepatocellular Carcinoma Cell Proliferation, Apoptosis, and EMT. *Dig Dis Sci* (2015) 60(8):2360–72. doi: 10.1007/s10620-015-3730-9
38. Liu Q, Tu K, Zhang H, Zheng X, Yao Y, Liu Q. TPX2 as a novel prognostic biomarker for hepatocellular carcinoma. *Hepatol Res* (2015) 45(8):906–18. doi: 10.1111/hepr.12428
39. Gao CL, Wang GW, Yang GQ, Yang H, Zhuang L. Karyopherin subunit-α 2 expression accelerates cell cycle progression by upregulating CCNB2 and CDK1 in hepatocellular carcinoma. *Oncol Lett* (2018) 15(3):2815–20. doi: 10.3892/ol.2017.7691
40. Gao X, Wang X, Zhang S. Bioinformatics identification of crucial genes and pathways associated with hepatocellular carcinoma. *Biosci Rep* (2018) 38(6). doi: 10.1042/bsr20181441
41. Li R, Jiang X, Zhang Y, Wang S, Chen X, Yu X, et al. Cyclin B2 Overexpression in Human Hepatocellular Carcinoma is Associated with Poor Prognosis. *Arch Med Res* (2019) 50(1):10–7. doi: 10.1016/j.arcmed.2019.03.003
42. Pouw RB, Gómez Delgado I, López Lera A, Rodríguez de Córdoba S, Wouters D, Kuijpers TW, et al. High Complement Factor H-Related (FHR)-3 Levels Are Associated With the Atypical Hemolytic-Uremic Syndrome-Risk Allele CFHR3*B. *Front Immunol* (2018) 9:848. doi: 10.3389/fimmu.2018.00848
43. Liu J, Li W, Zhao H. CFHR3 is a potential novel biomarker for hepatocellular carcinoma. *J Cell Biochem* (2020) 121(4):2970–80. doi: 10.1002/jcb.29551
44. Liu H, Zhang L, Wang P. Complement factor H–related 3 overexpression affects hepatocellular carcinoma proliferation and apoptosis. *Mol Med Rep* (2019) 20(3):2694–702. doi: 10.3892/mmr.2019.10514
45. Nalepa G, Clapp DW. Fanconi anaemia and cancer: an intricate relationship. *Nat Rev Cancer* (2018) 18(3):168–85. doi: 10.1038/nrc.2017.116
46. Sondalle SB, Longerich S, Ogawa LM, Sung P, Baserga SJ. Fanconi anemia protein FANCI functions in ribosome biogenesis. *Proc Natl Acad Sci USA* (2019) 116(7):2561–70. doi: 10.1073/pnas.1811557116
47. Zheng P, Li L. FANCI Cooperates with IMPDH2 to Promote Lung Adenocarcinoma Tumor Growth via a MEK/ERK/MMPs Pathway. *Onco Targets Ther* (2020) 13:451–63. doi: 10.2147/ott.S230333
48. Hu WF, Krieger KL, Lagundžin D, Li X, Cheung RS, Taniguchi T, et al. CTDP1 regulates breast cancer survival and DNA repair through BRCT-specific interactions with FANCI. *Cell Death Discovery* (2019) 5:105. doi: 10.1038/s41420-019-0185-3
49. Eich ML, Chandrashekar DS, Rodriguez Pen AM, Robinson AD, Siddiqui J, Daignault-Newton S, et al. Characterization of glycine-N-acyltransferase like

1 (GLYATL1) in prostate cancer. *Prostate* (2019) 79(14):1629–39. doi: 10.1002/pros.23887
50. van der Sluis R, Badenhorst CP, Erasmus E, van Dyk E, van der Westhuizen FH, van Dijk AA. Conservation of the coding regions of the glycine N-acyltransferase gene further suggests that glycine conjugation is an essential detoxification pathway. *Gene* (2015) 571(1):126–34. doi: 10.1016/j.gene.2015.06.081
51. Nalla AK, Williams TF, Collins CP, Rae DT, Trobridge GD. Lentiviral vector-mediated insertional mutagenesis screen identifies genes that influence androgen independent prostate cancer progression and predict clinical outcome. *Mol Carcinog* (2016) 55(11):1761–71. doi: 10.1002/mc.22425
52. Wang J, Shidfar A, Ivancic D, Ranjan M, Liu L, Choi MR, et al. Overexpression of lipid metabolism genes and PBX1 in the contralateral breasts of women with estrogen receptor-negative breast cancer. *Int J Cancer* (2017) 140 (11):2484–97. doi: 10.1002/ijc.30680
53. Mur P, De Voer RM, Olivera-Salguero R, Rodríguez-Perales S, Pons T, Setién F, et al. Germline mutations in the spindle assembly checkpoint genes BUB1 and BUB3 are infrequent in familial colorectal cancer and polyposis. *Mol Cancer* (2018) 17(1):23. doi: 10.1186/s12943-018-0762-8
54. Ding Y, Hubert CG, Herman J, Corrin P, Toledo CM, Skutt-Kakaria K, et al. Cancer-Specific requirement for BUB1B/BUBR1 in human brain tumor isolates and genetically transformed cells. *Cancer Discovery* (2013) 3(2):198–211. doi: 10.1158/2159-8290.Cd-12-0353
55. Li J, Ying Y, Xie H, Jin K, Yan H, Wang S, et al. Dual regulatory role of CCNA2 in modulating CDK6 and MET-mediated cell-cycle pathway and EMT progression is blocked by miR-381-3p in bladder cancer. *FASEB J* (2019) 33(1):1374–88. doi: 10.1096/fj.201800667R
56. Zhang H, Zhang X, Li X, Meng WB, Bai ZT, Rui SZ, et al. Effect of CCNB1 silencing on cell cycle, senescence, and apoptosis through the p53 signaling pathway in pancreatic cancer. *J Cell Physiol* (2018) 234(1):619–31. doi: 10.1002/jcp.26816
57. Jiang W, Yu Y, Liu J, Zhao Q, Wang J, Zhang J, et al. Downregulation of Cdc6 inhibits tumorigenesis of osteosarcoma in vivo and in vitro. *BioMed Pharmacother* (2019) 115:108949. doi: 10.1016/j.biopha.2019.108949
58. Wang Y, Wang F, He J, Du J, Zhang H, Shi H, et al. miR-30a-3p Targets MAD2L1 and Regulates Proliferation of Gastric Cancer Cells. *Onco Targets Ther* (2019) 12:11313–24. doi: 10.2147/ott.S222854
59. Liu Z, Li J, Chen J, Shan Q, Dai H, Xie H, et al. MCM family in HCC: MCM6 indicates adverse tumor features and poor outcomes and promotes S/G2 cell cycle progression. *BMC Cancer* (2018) 18(1):200. doi: 10.1186/s12885-018-4056-8
60. Merrick DT, Edwards MG, Franklin WA, Sugita M, Keith RL, Miller YE, et al. Altered Cell-Cycle Control, Inflammation, and Adhesion in High-Risk Persistent Bronchial Dysplasia. *Cancer Res* (2018) 78(17):4971–83. doi: 10.1158/0008-5472.Can-17-3822
61. Duffy MJ, Synnott NC, Crown J. Mutant p53 as a target for cancer treatment. *Eur J Cancer* (2017) 83:258–65. doi: 10.1016/j.ejca.2017.06.023
62. Wang X, Simpson ER, Brown KA. p53: Protection against Tumor Growth beyond Effects on Cell Cycle and Apoptosis. *Cancer Res* (2015) 75(23):5001–7. doi: 10.1158/0008-5472.Can-15-0563
63. Curia MC, Zuckermann M, De Lellis L, Catalano T, Lattanzio R, Aceto G, et al. Sporadic childhood hepatoblastomas show activation of beta-catenin, mismatch repair defects and p53 mutations. *Mod Pathol* (2008) 21(1):7–14. doi: 10.1038/modpathol.3800977
64. Teoh N, Pyakurel P, Dan YY, Swisshelm K, Hou J, Mitchell C, et al. Induction of p53 renders ATM-deficient mice refractory to hepatocarcinogenesis. *Gastroenterology* (2010) 138(3):1155–65.e1-2. doi: 10.1053/j.gastro.2009.11.008

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

ORIGINAL RESEARCH
published: 07 May 2021
doi: 10.3389/fonc.2021.631682

Immune Microenvironment in Langerhans Cell Histiocytosis: Potential Prognostic Indicators

Chuchu Feng[1], Yang Li[1], Huang Ke[1]*, Xiaomin Peng[1], Haixia Guo[2], Liping Zhan[1], Xilin Xiong[1], Wenjun Weng[1], Jiaqiang Li[1] and Jianpei Fang[1]*

[1] Department of Pediatric Hematology/Oncology, Sun Yat-Sen Memorial Hospital, Sun Yat-Sen University, Guangzhou, China, [2] Department of Pediatrics, Nanfang Hospital, Southern Medical University, Guangzhou, China

OPEN ACCESS

Edited by:
Yizhuo Zhang,
Sun Yat-sen University Cancer Center (SYSUCC), China

Reviewed by:
Matteo Chinello,
Integrated University Hospital Verona, Italy
Sebastian Dorin Asaftei,
Ospedale Città della Salute e della Scienza, Italy

***Correspondence:**
Yang Li
liyang5@mail.sysu.edu.cn
Huang Ke
hke@mail.sysu.cn

Specialty section:
This article was submitted to Pediatric Oncology, a section of the journal Frontiers in Oncology

Received: *20 November 2020*
Accepted: *13 April 2021*
Published: *07 May 2021*

Citation:
Feng C, Li Y, Ke H, Peng X, Guo H, Zhan L, Xiong X, Weng W, Li J and Fang J (2021) Immune Microenvironment in Langerhans Cell Histiocytosis: Potential Prognostic Indicators. Front. Oncol. 11:631682. doi: 10.3389/fonc.2021.631682

In this study, the immune microenvironment in Langerhans cell histiocytosis (LCH) was characterized to determine if immune indices are predictive of severity. Serum samples from 54 treatment-naïve patients were analyzed quantitatively for inflammatory cytokines and immunoglobulins before and after the induction of chemotherapy. The initial serum sIL-2R, TNF-α, and IL-10 of untreated LCH patients with risk organ involvement (RO+) were significantly higher than those with single-system (SS) involvement. LCH patients with hematologic involvement exhibited a significantly higher sIL-2R, TNF-α, IL-10, and IL-1β expression, as compared to the group without involvement. sIL-2R, TNF-α, and IL-10 were increased in patients with liver or spleen involvement. Th cells have decreased in the liver+ and spleen+ group, and Ts cells were significantly decreased in non-response group after induction chemotherapy. The serum level of immune indices represents, to some extent, the severity of the disease. Pertinent laboratory inspections can be used to improve risk stratification and guide immunotherapy.

Keywords: Langerhans cell histiocytosis, tumor immunology, children, cytokines, inflammation

INTRODUCTION

Langerhans cell histiocytosis (LCH) is a histiocytic disorder arising from the mononuclear phagocyte system, which results in the abnormal accumulation and proliferation of LCH cells. LCH is more common in children than in adults, with the clinical manifestations varying from isolated osseous, mucocutaneous, and pulmonary involvement to multi-system (MS) involvement, such as lymph node, bone marrow, liver, spleen, gastrointestinal tract, thymus, endocrine gland, and central nervous system involvement, causing hyperplasia, fibrosis, necrosis and other pathological changes, eventually leading to organ dysfunction (1). In 2010, Badalian-very found 57% of BRAF V600E gene positive-mutation in 61 LCH patients (2). The disease is an inflammatory myeloid neoplasia with the characteristics of both an abnormal reactive process and a neoplastic process. Subsequent studies showed that the mutation rate of BRAF gene in LCH cases was 45% to 65%, suggesting that the BRAF V600E gene was closely related to LCH incidence (3–11). De Graaf et al. found a variety of cytokines expressed in LCH lesions, such as IL-1, TGF-α, TGF-β, TNF-α, and TFN-γ (11).

Kannourakis et al. extracted and cultured monocytes from eosinophilic granulomatous tissues in patients with LCH, and found that such monocytes could highly express IL-1, TNF-α, GM-CSF, IL-8, and LIF-21 (12). In addition to the "cytokine storm" of local lesions, the apparent level of some inflammatory cytokines in the serum of LCH patients increased, suggesting that cytokines may be associated with the pathogenesis of LCH. At present, the etiology and pathogenesis of LCH remains speculative, bringing about uneven curative effect and lacking effective prognosis indicator. In the present study, the immune function of LCH children admitted to our hospital in the past 7 years was reviewed before and after induction treatment. Clarifying the immune status of LCH children will help provide insights into LCH prognosis and, ultimately, optimize and personalize therapy.

METHODS

Patients

This study was performed at the Sun Yat-sen Memorial Hospital (Guangzhou, China) between March 2013 and September 2020. A total of 54 children (37 males and 17 females; median age, 3.6 years; age range, 2.0 months to 12.0 years) were enrolled in this study. All patients fulfilled accepted diagnostic criteria established by the Histiocyte Society in 2009. Evaluation after the induction chemotherapy of JLSG-96/02 or Chinese Children's Histiocytic Group (CCHG)-LCH-2019 regimens was performed based on the following criteria (13): i) non active disease (NAD), no evidence of disease, Resolution of all signs or symptoms; ii) active disease-better (AD-B), regression of signs or symptoms, no new lesions; iii) active disease-mixed (AD-I), new lesions in one site, regression in another site; iv) active disease-stable (AD-S), persistence of signs or symptoms, no new lesions; vi) worse, progression of signs or symptoms and/or appearance of new lesions. In isolated bone disease progression is defined as appearance of new bone lesions or lesions in other organs. Risk organs include the hematologic system, the spleen and the liver. All methods were carried out in accordance with relevant guidelines and regulations. All peripheral blood samples were obtained with written informed consent from the legal guardian of patients. The study was approved by the ethics committee at Sun Yat-sen Memorial Hospital.

Cytokine, Immunoglobulin, and Lymphocyte Subset Measurement

Prior to treatment, and 2 weeks after the initial induction of chemotherapy, the serum levels of the cytokines IL-8, IL-6, IL-10, IL-1β, and sIL-2R were measured using the IMMULITE-1000 Immunoassay System (Siemens Healthineers, Erlangen, Germany) while the immunoglobulins IgA, IgG, IgM and IgE were detected by the BN II system (Siemens Healthineers). Lymphocyte subsets in the peripheral blood, including T, B, natural killer, cytokine-induced killer, T helper (Th), suppressor T (Ts), Th/Ts, and regulatory T cells (Treg) were tested using a FACSCanto™ II flow cytometer (Beckman Coulter, Inc., Brea, CA, USA), using a BD Multitest™ 6-color TBNK kit and DIVA software.

Statistical Analysis

Data were statistically analyzed using SPSS software (version 21.0). The continuous variables were expressed as the mean ± standard deviation and the count variables were expressed in frequency/rate. Original data of inflammatory factors IL-10 and IL-1β were expressed as <5 pg/ml, which were transformed into qualitative variables in the correlation analysis. The non-paired t and Kruskal-Wallis H tests were used to assess the statistical significance between the groups. To compare qualitative data we used the Fisher's exact test. $P<0.05$ was considered to indicate a statistically significant difference.

RESULTS

Patient Characteristics (**Table 1**).

Analysis of Immune Indices and Clinical Types

A total of 17 patients had multifocal single system disease (SS), 15 patients had multiple system disease without risk organ involvement (RO-), and the remaining 22 patients had risk organ involvement (RO+). Among the inflammatory cytokines, as shown in **Figure 1**, significantly higher serum levels of sIL-2R (2244.2 ± 2790.9 *vs* 595.4 ± 366.7 μ/ml), TNF-α (31.0 ± 24.0 *vs* 11.8 ± 4.6 pg/ml), and IL-10 were observed in the RO+ group, as compared to the SS group. The percentages of T cells in peripheral blood were obviously lower in the RO+ group than RO-, while B cells obviously increased (**Figure 2**). No statistical significance was identified in immunoglobulin (**Table 2**).

Next, patients were divided into the groups depending on the type of organ involvement. As shown in **Figure 3**, patients with hematologic involvement exhibited a significantly higher sIL-2R, TNF-α, IL-10, and IL-1β expression, as compared to the group without involvement. sIL-2R, TNF-α, and IL-10 were increased

TABLE 1 | Clinical characteristics of 54 children with LCH.

	No. of patients
Gender	Gender
Male	37
Female	17
Age at diagnosis (yrs)	
Median	3.6
Disease classification	
Multifocal single system (SS)	17
Multiple system without risk organ involvement (MS-RO-)	15
Multiple system with risk organ involvement (MS-RO+)	22
Risk organ involvement	
Liver	18
Spleen	6
Hematologic	3
Response at week 6	
NAD	9
AD-B	33
AD-I	1
AD-S	8
Worse	3

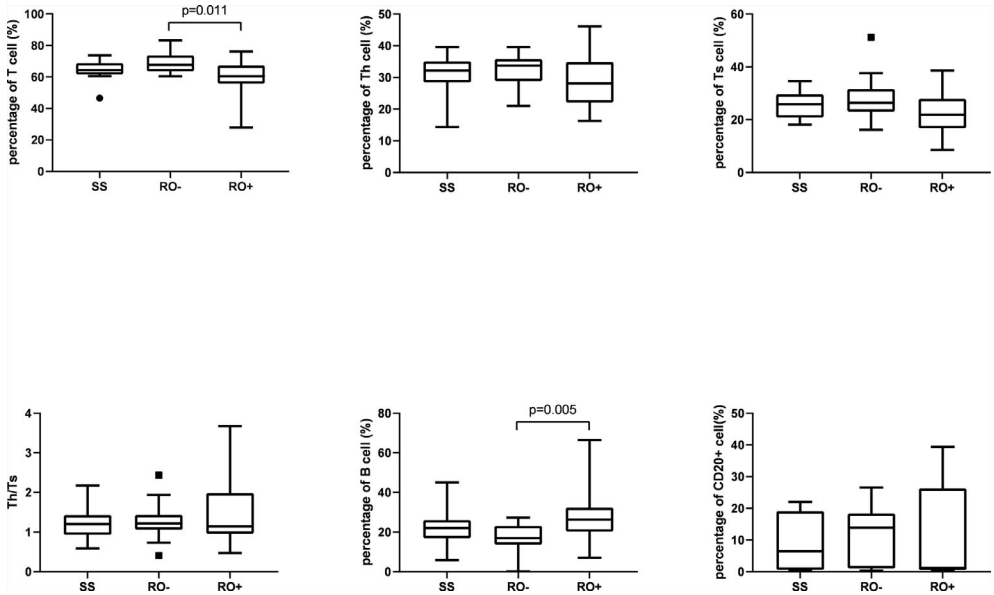


FIGURE 1 | The serum level of inflammatory cytokines in disease classification.

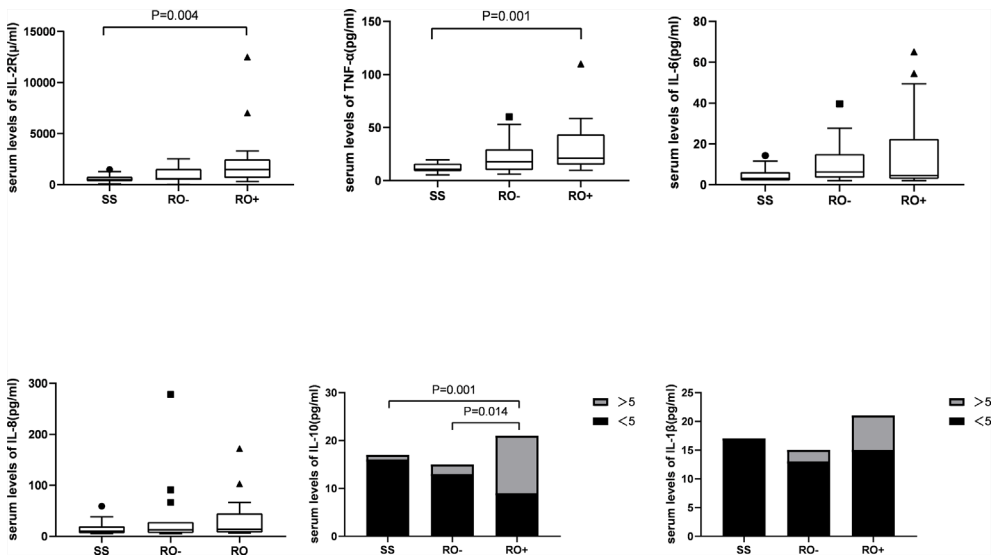


FIGURE 2 | The lymphocyte subset in disease classification.

in patients with liver or spleen involvement. The percentage of T and Th cells were significantly lower in the spleen+ group, but B cells subset increased by contrast. Similarly, the percentage of Th cells and Th/Ts ratio has decreased in the liver+ group (**Table 3**).

Analysis of Immune Indices Before and After Induction Chemotherapy

In order to clarify whether the change in immune indices can be used as a marker of efficacy, serum levels of cytokines, immunoglobulins and lymphocyte subsets were measured prior to treatment. All cases underwent initial chemotherapy of the JLSG-96/02 or CCHG-LCH-2019 regimen, based on disease status. Following the induction of chemotherapy, Ts cells were significantly decreased in non-response group. However, no significant changes were observed in cytokine and immunoglobulin between the response (NAD+AD-B+AD-I) and non-response (AD-S+Worse) groups (**Table 4**). Similarly, no significant changes were observed when we further divided the data into SS, MS, RO- and RO+ sets.

TABLE 2 | Immunoglobulin in different disease classification.

Immune indices	Disease classification			
	SS	MS-RO-	MS-RO+	P-value
IgA(g/l)	2.0 ± 1.0	1.2 ± 0.7	1.4 ± 1.0	0.052
IgG(g/l)	12.5 ± 2.9	10.3 ± 3.6	10.6 ± 3.5	0.178
IgM(g/l)	1.5 ± 0.5	1.3 ± 0.5	1.4 ± 0.7	0.729
IgE(IU/ml)	543.4 ± 1275.1	166.3 ± 284.0	152.6 ± 226.8	0.466

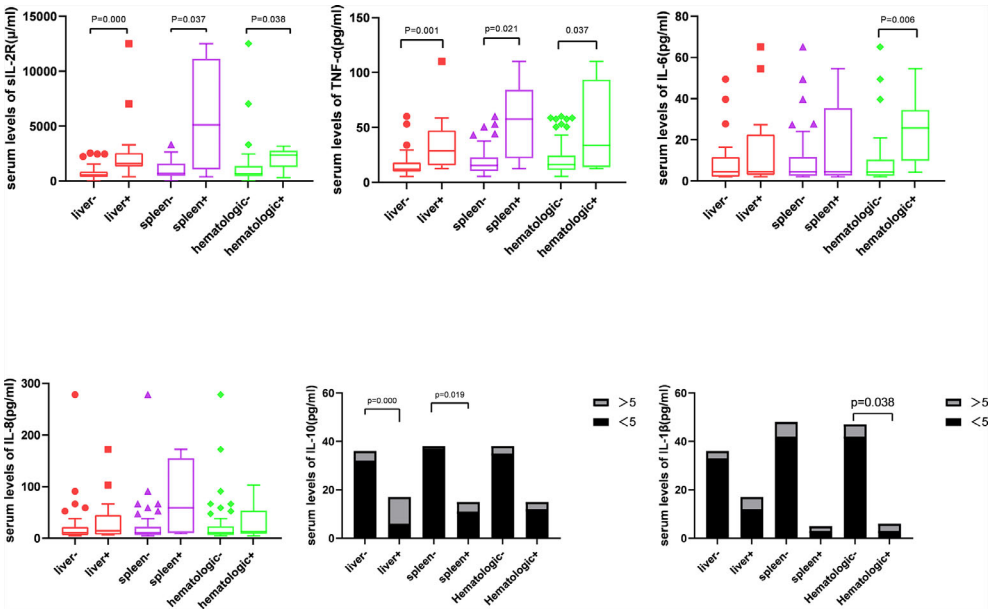


FIGURE 3 | Analysis of inflammatory cytokines and affected organs.

TABLE 3 | Lymphocyte subset in different affected organs.

	Affected organs								
	Liver+	Liver-	P-value	Spleen+	Spleen-	P-value	hematologic+	hematologic-	P-value
T cells (%)	57.3 ± 10.6	66.9 ± 6.9	0.001*	56.7 ± 7.7	64.4 ± 9.3	0.058	60.2 ± 14.2	64.3 ± 8.6	0.259
B cells (%)	28.8 ± 12.8	20.0 ± 8.9	0.018*	27.7 ± 10.0	22.6 ± 11.2	0.341	25.4 ± 15.8	22.7 ± 10.4	0.389
Th cells (%)	27.0 ± 7.6	32.5 ± 6.2	0.005*	22.9 ± 3.0	31.5 ± 6.9	0.010*	27.7 ± 8.3	31.2 ± 6.9	0.395
Ts cells (%)	23.0 ± 6.9	26.1 ± 7.7	0.241	27.6 ± 5.0	24.9 ± 7.7	0.333	27.6 ± 12.2	24.8 ± 6.8	0.927
Th/Ts	1.3 ± 0.7	1.4 ± 0.7	0.579	0.9 ± 0.2	1.4 ± 0.7	0.024*	1.2 ± 0.6	1.4 ± 0.7	0.446
CD20+ cells (%)	10.5 ± 10.0	10.6 ± 14.8	0.943	13.8 ± 15.7	10.1 ± 11.3	0.454	19.5 ± 16.9	9.3 ± 10.6	0.165

**Statistical difference between the affected organs groups (P < 0.05).*
Th, T helper; Ts, suppressor T.

DISCUSSION

LCH is a heterogeneous disease and it can affect from single, localized lesions to multiple systems/organs, including risk organs. The prognosis of different subtypes of LCH is highly variable. Partial cases with a solitary bone lesion can be cured by curettage, and children with skin-isolated LCH usually require no specific therapy, as spontaneous healing may occur. MS-LCH with risk organ involvement, such as liver and hematopoietic system, has a poor response to therapy, resulting in a particularly dismal prognosis (14). To date, the pathogenesis of LCH remains unclear. In recent years, it was found that the MAPK pathway is involved in the pathogenesis of LCH, with the mutation of more than half of BRAF, 20% of MAP2K1 and rare ARAF, MAP3K1 (2, 15). However, studies have shown that cytokines are essential for local infiltration and metastasis of LCH cells (16, 17). De Graaf et al. found a variety of cytokines in LCH lesions, such as IL-1, TGF-α, TGF-β, GM-CSF, TNF-α, and TFN-γ (11). Kannourakis

TABLE 4 | Comparison of immune indices before and after induction chemotherapy.

Immune index rangeability (After-before)	Response	Non-response	P-value
ΔsIL-2R (μ/ml)	50.9 ± 609.6	13.3 ± 879.0	0.870
ΔIL-6 (pg/ml)	−2.4 ± 18.4	4.3 ± 20.2	0.299
ΔTNF-α (pg/ml)	18.6 ± 66.2	2.9 ± 12.8	0.438
ΔIL-8 (pg/ml)	−0.2 ± 101.0	−31.8 ± 134.4	0.389
ΔIgA (g/l)	0.2 ± 0.7	0.2 ± 0.3	0.963
ΔIgG (g/l)	1.8 ± 4.3	1.3 ± 2.4	0.677
ΔIgM (g/l)	0.3 ± 0.8	0.5 ± 0.3	0.471
ΔIgE (IU/ml)	−13.0 ± 473.2	77.3 ± 158.3	0.539
ΔT cell (%)	−3.4 ± 23.3	−20.4 ± 17.1	0.036*
ΔB cell (%)	2.1 ± 17.5	9.0 ± 11.7	0.150
ΔNK cell (%)	−0.9 ± 8.2	0.9 ± 8.2	0.531
ΔTh cell (%)	−1.7 ± 13.4	−5.9 ± 17.6	0.384
ΔTs cell (%)	−0.2 ± 12.8	−9.6 ± 12.3	0.035*
ΔTh/Ts	0.5 ± 3.3	−0.03 ± 1.0	0.347
ΔCD20+ cell (%)	2.7 ± 16.9	13.2 ± 10.5	0.153

Th, T helper; Ts, suppressor T.

**Statistical difference between the Affected organs groups (P < 0.05).*

et al. extracted and cultured monocytes from eosinophilic granulomatous tissues in patients with LCH, and found that such monocytes could produce a large number of IL-1, TNF-α, GM-CSF, IL-8, and LIF-21 (12). Pathologically, one of the cardinal manifestations of LCH is the accumulation of pathologic LCH cells in target tissues, surrounding by the varying degrees of lymphocyte infiltration, such as T cells, macrophages, eosinophils, and B cells, as well as multinucleated giant cells. Egeler used immunohistochemical techniques to detect LCH and T cells, macrophages and eosinophils in 14 children with LCH; it was found that cytokines mainly originated from LCH and T cells (18). On the other hand, LCH is seen as a result of a misguided differentiation of myeloid dendritic cell (DC) precursors originating from multiple hematopoietic stem cells, whose differentiation, maturation and migration are regulated by diverse cytokines (19). For example, Cumberbatch found that DCs were significantly concentrated in the lymph nodes of mice following a subcutaneous injection of TNF-α. However, this phenomenon did not occur after injecting the same dose of TNF-α directly into the lymph nodes, suggesting that TNF-α may contribute to the migration of DC/LC lineage cells (20).

In 1994, Kannourakis reported elevated peripheral blood levels of GM-CSF and IL-3 in MS-LCH patients, recognizing that cytokines in LCH lesions were likely to be released to the circulation (12). Our results demonstrated that the serum levels of sIL-2R, TNF-α, and IL-10 in the MS-LCH patients with RO+ were significantly higher than SS-LCH patients. In particular, sIL-2R, TNF-α, and IL-10 were noticeably increased in patients with liver spleen and hematologic involvement. This indicated that serum levels of sIL-2R, TNF-α, and IL-10 may reflect the severity of the disease in LCH to a certain extent. Morimoto's study, which found the serum levels of 9 humoral factors, including IL-2R, IL-8, IL-18, and M-CSF, substantially higher in patients with MS-LCH than in those with SS-LCH (21). sIL-2R consists of three chains; α (also termed IL-2Rα, CD25, or Tac antigen), β (also termed IL-2Rβ or CD122), and γ (also termed IL-2Rγ or CD132). The principal functions of IL-2Rα are to bind with IL-2 and promote optimal IL-2 signaling through its association with the IL-2Rβ and IL-2Rγ chains, while inhibiting the clonal proliferation of activated T cells. IL-2, IL-2Rβ, and IL-2Rγ are rapidly degraded, but IL-2Rα is recycled to the cell surface. Thus, the available concentration of the soluble form of IL-2R (sIL2Rα) determines the tempo, magnitude and extent of T cell immune responses. TNF-α is mainly produced by activated macrophages with a wide range of biological functions, including the induction of inflammation, anti-tumor effect, activation of T cells, and mediated immune response. IL-2R and TNF-α have been reported to play an important role in inducing the generation and maturation of LCs *in vitro* (22). It has also been found that serum sIL-2R and TNF-α are significantly elevated in LCH patients (23, 24). IL-10 could bind with IL-10 receptors on tumor cells to activate STAT3, which thus promotes the proliferation of tumor cells *via* the activation of cell cycle-related proteins (25, 26). It has been hypothesized that IL-10 may play a role in the assessment of LCH, since a study reported an increased expression of IL-10 in LCH lesions (27).

Cytokine elevation indicates a disturbance in cellular immunity. The number of T cells in an active state in LCH lesions was second only to the number of LCH cells (27). Treg cells were also found to be increased in the peripheral blood of LCH patients (28, 29). Our findings have shown that the percentage of Th cells and Th/Ts ratio in the peripheral blood of LCH patients with liver or spleen involvement was lower than those without involvement. After that, the level of Ts cell dropped during the induction treatment in non-response group. Suppressor T cell (Ts), also call regulatory cells (Tregs), were the second most common type of infiltrating immune cell in LCH tissue (30). Ts could inhibit immune responses against LCH cells, which lead to increasing survival of LCH cells, granuloma maintenance, and dissemination (28). It seems to reflect that the abnormal regulation of T lymphocyte may affect the disease progression.

The present study failed to reflect the effects of the changes in immune indicators on the assessment of efficacy. Some clinical observation showed that the ratio of serum TNF-α (23), IL-2R,

RANKL, OPG, and SRANKL/OPG (24) significantly decreased following chemotherapy. Patients with IL-18 serum levels of >500 pg/ml were insensitive to JLSG treatment (21). To date, there have been a number of studies with similar results, at home and abroad, but it is necessary to explore and confirm the underlying mechanisms by a large-scale, multicenter trial. Therefore, changes in serum cytokines may be used as a marker of the curative effect of clinical treatment, but whether it is a sensitive and specific marker requires further research.

Children with LCH often present with MS damage at onset and frequently involved risk organs. These cases are characterized by a long disease course, low cure rate and easy recurrence. Increased understanding of the pathogenesis and pathological changes of different clinical types of LCH will help optimize and personalize therapy, which can, in turn, improve the curative effect. Initial reports have indicated that immunocyte and cytokine immunoregulatory disorders might be linked to the occurrence, and development of LCH. Pertinent laboratory inspections can be used as prognostic indices for children with LCH, to improve risk-stratification and guide immunotherapy. Besides, the short assessment time, the change of the clinical classification and complications, such as infections during treatment may also affect the results. It may be possible to collect immune indicators of patients during the 12 weeks of induction and maintenance treatment in a follow-up study.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary mterial. Further inquiries can be directed to the corresponding authors.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Sun Yat-Sen Memorial Hospital. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

YL: conceptualization, writing—review and editing and supervision. HK: conceptualization, methodology, and supervision. CF: formal analysis, writing—original draft, and visualization. XP: data curation. HG: validation. LZ: data curation. XX: formal analysis. WW: supervision. JL: data curation. JF: supervision. All authors contributed to the article and approved the submitted version.

REFERENCES

1. Morimoto A, Oh Y, Shioda Y, Kudo K, Imamura T. Recent Advances in Langerhans Cell Histiocytosis. *Pediatr Int* (2014) 56:451–61. doi: 10.1111/ped.12380
2. Badalian-Very G, Vergilio JA, Degar BA, MacConaill LE, Brandner B, Calicchio ML, et al. Recurrent BRAF Mutations in Langerhans Cell Histiocytosis. *Blood* (2010) 116:1919–23. doi: 10.1182/blood-2010-04-279083
3. Haroche J, Charlotte F, Arnaud L, von Deimling A, Helias-Rodzewicz Z, Hervier B, et al. High Prevalence of BRAF V600E Mutations in Erdheim-Chester Disease But Not in Other non-Langerhans Cell Histiocytoses. *Blood* (2012) 120:2700–3. doi: 10.1182/blood-2012-05-430140
4. Sahm F, Capper D, Preusser M, Meyer J, Stenzinger A, Lasitschka F, et al. BRAFV600E Mutant Protein is Expressed in Cells of Variable Maturation in Langerhans Cell Histiocytosis. *Blood* (2012) 120:e28–34. doi: 10.1182/blood-2012-06-429597
5. Satoh T, Smith A, Sarde A, Lu HC, Mian S, Trouillet C, et al. B-RAF Mutant Alleles Associated With Langerhans Cell Histiocytosis, a Granulomatous Pediatric Disease. *PLoS One* (2012) 7:e33891. doi: 10.1371/journal.pone.0033891
6. Bubolz AM, Weissinger SE, Stenzinger A, Arndt A, Steinestel K, Bruderlein S, et al. Potential Clinical Implications of BRAF Mutations in Histiocytic Proliferations. *Oncotarget* (2014) 5:4060–70. doi: 10.18632/oncotarget.2061
7. Chilosi M, Facchetti F, Calio A, Zamo A, Brunelli M, Martignoni G, et al. Oncogene-Induced Senescence Distinguishes Indolent From Aggressive Forms of Pulmonary and non-Pulmonary Langerhans Cell Histiocytosis. *Leuk Lymphoma* (2014) 55:2620–6. doi: 10.3109/10428194.2014.887713
8. Mehes G, Irsai G, Bedekovics J, Beke L, Fazakas F, Rozsa T, et al. Activating BRAF V600E Mutation in Aggressive Pediatric Langerhans Cell Histiocytosis: Demonstration by Allele-Specific PCR/direct Sequencing and Immunohistochemistry. *Am J Surg Pathol* (2014) 38:1644–8. doi: 10.1097/PAS.0000000000000304
9. Roden AC, Hu X, Kip S, Parrilla Castellar ER, Rumilla KM, Vrana JA, et al. Braf V600E Expression in Langerhans Cell Histiocytosis: Clinical and Immunohistochemical Study on 25 Pulmonary and 54 Extrapulmonary Cases. *Am J Surg Pathol* (2014) 38:548–51. doi: 10.1097/PAS.0000000000000129
10. Berres ML, Lim KP, Peters T, Price J, Takizawa H, Salmon H, et al. Braf-V600E Expression in Precursor Versus Differentiated Dendritic Cells Defines Clinically Distinct LCH Risk Groups. *J Exp Med* (2015) 212:281. doi: 10.1084/jem.2013097701202015c
11. de Graaf JH, Tamminga RY, Dam-Meiring A, Kamps WA, Timens W. The Presence of Cytokines in Langerhans' Cell Histiocytosis. *J Pathol* (1996) 180:400–6. doi: 10.1002/(SICI)1096-9896(199612)180:4<400::AID-PATH701>3.0.CO;2-W
12. Kannourakis G, Abbas A. The Role of Cytokines in the Pathogenesis of Langerhans Cell Histiocytosis. *Br J Cancer Suppl* (1994) 23:S37–40.
13. Morimoto A, Shioda Y, Imamura T, Kudo K, Kawaguchi H, Sakashita K, et al. Intensified and Prolonged Therapy Comprising Cytarabine, Vincristine and Prednisolone Improves Outcome in Patients With Multisystem Langerhans Cell Histiocytosis: Results of the Japan Langerhans Cell Histiocytosis Study Group-02 Protocol Study. *Int J Hematol* (2016) 104:99–109. doi: 10.1007/s12185-016-1993-3
14. Arico M. Langerhans Cell Histiocytosis in Children: From the Bench to Bedside for an Updated Therapy. *Br J Haematol* (2016) 173:663–70. doi: 10.1111/bjh.13955
15. Kolenova A, Schwentner R, Jug G, Simonitsch-Klupp I, Kornauth C, Plank L, et al. Targeted Inhibition of the MAPK Pathway: Emerging Salvage Option for Progressive Life-Threatening Multisystem LCH. *Blood Adv* (2017) 1:352–6. doi: 10.1182/bloodadvances.2016003533
16. Garabedian L, Struyf S, Opdenakker G, Sozzani S, Van Damme J, Laureys G. Langerhans Cell Histiocytosis: A Cytokine/Chemokine-Mediated Disorder? *Eur Cytokine Netw* (2011) 22:148–53. doi: 10.1684/ecn.2011.0290
17. Berres ML, Allen CE, Merad M. Pathological Consequence of Misguided Dendritic Cell Differentiation in Histiocytic Diseases. *Adv Immunol* (2013) 120:127–61. doi: 10.1016/B978-0-12-417028-5.00005-3
18. Egeler RM, Favara BE, van Meurs M, Laman JD, Claassen E. Differential In Situ Cytokine Profiles of Langerhans-like Cells and T Cells in Langerhans Cell Histiocytosis: Abundant Expression of Cytokines Relevant to Disease and Treatment. *Blood* (1999) 94:4195–201.
19. Rizzo FM, Cives M, Simone V, Silvestris F. New Insights Into the Molecular Pathogenesis of Langerhans Cell Histiocytosis. *Oncologist* (2014) 19:151–63. doi: 10.1634/theoncologist.2013-0341

20. Cumberbatch M, Kimber I. Dermal Tumour Necrosis Factor-Alpha Induces Dendritic Cell Migration to Draining Lymph Nodes, and Possibly Provides One Stimulus for Langerhans' Cell Migration. *Immunology* (1992) 75:257–63.
21. Morimoto A, Oh Y, Nakamura S, Shioda Y, Hayase T, Imamura T, et al. Inflammatory Serum Cytokines and Chemokines Increase Associated With the Disease Extent in Pediatric Langerhans Cell Histiocytosis. *Cytokine* (2017) 97:73–9. doi: 10.1016/j.cyto.2017.05.026
22. Zou GM, Tam YK. Cytokines in the Generation and Maturation of Dendritic Cells: Recent Advances. *Eur Cytokine Netw* (2002) 13:186–99.
23. Rosso DA, Ripoli MF, Roy A, Diez RA, Zelazko ME, Braier JL. Serum Levels of Interleukin-1 Receptor Antagonist and Tumor Necrosis Factor-Alpha are Elevated in Children With Langerhans Cell Histiocytosis. *J Pediatr Hematol Oncol* (2003) 25:480–3. doi: 10.1097/00043426-200306000-00010
24. Ishii R, Morimoto A, Ikushima S, Sugimoto T, Asami K, Bessho F, et al. High Serum Values of Soluble CD154, IL-2 Receptor, RANKL and Osteoprotegerin in Langerhans Cell Histiocytosis. *Pediatr Blood Cancer* (2006) 47:194–9. doi: 10.1002/pbc.20595
25. Li L, Shaw PE. Autocrine-Mediated Activation of STAT3 Correlates With Cell Proliferation in Breast Carcinoma Lines. *J Biol Chem* (2002) 277:17397–405. doi: 10.1074/jbc.M109962200
26. Chen CL, Cen L, Kohout J, Hutzen B, Chan C, Hsieh FC, et al. Signal Transducer and Activator of Transcription 3 Activation is Associated With Bladder Cancer Cell Growth and Survival. *Mol Cancer* (2008) 7:78. doi: 10.1186/1476-4598-7-78
27. Quispel WT, Stegehuis-Kamp JA, Santos SJ, Egeler RM, van Halteren AG. Activated Conventional T-Cells Are Present in Langerhans Cell Histiocytosis Lesions Despite the Presence of Immune Suppressive Cytokines. *J Interferon Cytokine Res* (2015) 35:831–9. doi: 10.1089/jir.2014.0190
28. Senechal B, Elain G, Jeziorski E, Grondin V, Patey-Mariaud de Serre N, Jaubert F, et al. Expansion of Regulatory T Cells in Patients With Langerhans Cell Histiocytosis. *PLoS Med* (2007) 4:e253. doi: 10.1371/journal.pmed.0040253
29. Tong C, Jia X, Jia Y, He Y. Langerhans Cell Histiocytosis in Chinese Adults: Absence of BRAF Mutations and Increased FOXP3(+) Regulatory T Cells. *Int J Clin Exp Pathol* (2014) 7:3166–73.
30. Paredes SEY, Almeida LY, Trevisan GL, Polanco XBJ, Silveira HA, Vilela Silva E, et al. Immunohistochemical Characterization of Immune Cell Infiltration in Paediatric and Adult Langerhans Cell Histiocytosis. *Scand J Immunol* (2020) 92:e12950. doi: 10.1111/sji.12950

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling editor declared a shared affiliation, though no other collaboration, with several of the authors CF, YL, HK, XP, HG, LZ, XX, WW, JL, and JF.

frontiers in Oncology

ORIGINAL RESEARCH
published: 14 May 2021
doi: 10.3389/fonc.2021.633106

The Application of and Factors Influencing, the NB5 Assay in Neuroblastomas

Zuopeng Wang[1], Chengyun Wang[2], Yibing Xu[3], Jun Le[4], Yuan Jiang[5], Wei Yao[1], Hongsheng Wang[4] and Kai Li[1]**

[1] Department of Pediatric Surgery, Children's Hospital of Fudan University, Shanghai, China, [2] Department of Pediatric Surgery, Zaozhuang Maternal and Child Health Care Hospital, Shandong, China, [3] Institute of Translational Medicine, Zhejiang University School of Medicine, Hangzhou, China, [4] Department of Hematology, Children's Hospital of Fudan University, Shanghai, China, [5] Department of Clinical Epidemiology, Children's Hospital of Fudan University, Shanghai, China

OPEN ACCESS

Edited by:
Jing He,
Guangzhou Medical University, China

Reviewed by:
Yi Ji,
Sichuan University, China
Yuanmei Liu,
Affiliated Hospital of Zunyi Medical College, China
Zhibao LV,
Shanghai Children's Hospital, China

***Correspondence:**
Kai Li
likai2727@163.com
Hongsheng Wang
honswang@hotmail.com

Specialty section:
This article was submitted to Pediatric Oncology, a section of the journal Frontiers in Oncology

Received: 24 November 2020
Accepted: 29 March 2021
Published: 14 May 2021

Citation:
Wang Z, Wang C, Xu Y, Le J, Jiang Y, Yao W, Wang H and Li K (2021) The Application of and Factors Influencing, the NB5 Assay in Neuroblastomas. Front. Oncol. 11:633106. doi: 10.3389/fonc.2021.633106

Purpose: The NB5 assay was performed in bone marrow (BM) and peripheral blood (PB) to detect neuroblastomas (NBs) with micrometastases. The sensitivity and factors influencing the NB5 assay were preliminarily evaluated.

Methods: The NB5 assay uses RT-PCR to detect the co-expression of five mRNAs from the neuroblastoma-associated genes, CHGA, DCX, DDC, PHOX2B, and TH. We enrolled 180 cases of neuroblastoma and 65 cases of non-neuroblastoma. Bone marrow and peripheral blood were collected from every patient. The gold standard for the diagnosis of NB was pathological evaluation of solid tumor specimens or bone marrow biopsies (BMBs) from hematological tumors. STATA version 15 and SPSS version 17 software were used for analysis.

Results: We found that 17 patients were BMB (+), and they were diagnosed as the International Neuroblastoma Staging System (INSS) stage IV and the high-risk group. All 17 patients were BM (+), while 15 patients were PB (+) (15/17, 88.2%). Among the 163 children who were BMB (–), 56 were BM (+), 40 were PB (+), and 36 were BM (+) and PB (+). The sensitivity of the NB5 assay in BM (40.5%) and PB (30.5%) was significantly higher than the sensitivity of BMB (9.4%, $P = 0.000$). In the non-NB group, four cases were BM (+) and one case was PB (+). The specificity of the NB5 assay in BM and PB was 93.8% and 98.5%, respectively. The sensitivity of the NB5 assay in both BM and PB in INSS stage IV patients was significantly higher than that in INSS stage I–II patients ($P < 0.05$). The sensitivity of the NB5 assay in both BM and PB in the high-risk group was significantly higher than that in the middle-low-risk groups ($P = 0.0001$). Logistic regression analyses indicated that liver metastases and bone metastases were the primary factors influencing the sensitivity of the NB5 assay in BM and PB ($P < 0.05$).

Conclusions: The NB5 assay had significantly higher sensitivity than the pathological analysis of BMB in detecting NB with micrometastases. The NB5 assay had higher

sensitivity in INSS stage IV or the high-risk group. Liver metastases and bone metastases were the primary factors that affected the sensitivity of the NB5 assay.

Keywords: neuroblastoma, NB5 assay, sensitivity, micrometastases, bone marrow

INTRODUCTION

Neuroblastoma (NB) is the most common extracranial solid tumor in children and accounts for 15% of all pediatric malignancy deaths (1). Although NB has a heterogeneous clinical course and may regress spontaneously, most patients with NB experience early onset and progress rapidly. Approximately 45% of patients have distant metastatic lesions when diagnosed during infancy (2). NB with micrometastases are known as the minimal residual disease (MRD) and contribute to relapse, but are difficult to detect (3). The persistence of MRD is also predictive of worse patient survival and poorer outcomes (4). Bone marrow biopsies (BMBs) are routinely used for the diagnosis of bone marrow (BM) metastasis by cytological and histological examinations, which exhibit an analytical sensitivity of less than 1×10^{-3}. Thus, such sensitivities could severely underestimate the prevalence of bone marrow involvement (1, 2, 5, 6). Immunocytological investigations or flow cytometry had an analytical sensitivity of up to $(1 \times 10^{-4}) - (1 \times 10^{-5})$ for MRD in NB (5). With well-defined reverse transcriptase polymerase chain reaction (RT-PCR) markers that could provide a sensitivity of MRD of 1×10^{-6}, one NB cell among 10^6 normal cells could be detected for the early diagnosis of NB.

Several MRD diagnostic methods based on the expression of multiple gene markers have been reported (7–11). As a novel diagnostic method for NB, the mRNAs of five neuroblastoma markers (NB5 assay), including chromogranin A (CHGA), doublecortin (DCX), dopadecarboxylase (DDC), paired-like homeobox 2b (PHOX2B), and tyrosine hydroxylase (TH) were detected in the BM and peripheral blood (PB) from patients using RT-PCR (8). The NB5 assay was developed at the Children's Hospital of Los Angeles and is very sensitive in detecting MRD in BM and PB. Shanghai Bokang Biotechnology Co. LTD performed the detection services for MRD in NB using the NB5 assay.

The current consensus is that MRD remains in a dormant state, until "awakened" to progress towards overt metastases (12). MRD existing in the blood and bone marrow could contain tumor-initiating cells that generate tumors through abnormal proliferation and differentiation (13). A good MRD detection marker should be exclusively expressed in NB cells and not in non-NB cells; however, no clinical studies evaluated the sensitivity and effectiveness of these methods with NB markers. In this study, the NB5 assay was performed to assist in the clinical diagnosis of NB, and its sensitivity and specificity were preliminarily evaluated.

METHODS

Patients

Patients were regarded as having micrometastases when the expression of three or more genes was detected. Between August 2015 and May 2019, we enrolled 180 NB cases and 65 non-NB cases from the Children's Hospital of Fudan University. This study received approval from the local Research Ethics Committee of our hospital and was performed in accordance with the approved guidelines. Written informed consent was obtained from the guardians of each patient. The gold standard for the diagnosis of NB was based on pathological evaluations of solid tumor specimens or the BMBs of hematologic tumors.

Sample Processing and NB5 Assay

BM (2 ml) and PB (2 ml) were collected from each patient with a solid tumor or hematologic tumor, respectively. Peripheral blood mononuclear cells (PBMCs) were isolated from heparinized blood and bone marrow by density separation with Ficoll–Hypaque (14). Total RNA was prepared using the TRIzol® reagent (Invitrogen) and processed with the RNeasy® Mini Kit (QIAGEN). The RNA Integrity Number (RIN) was obtained using the Agilent Bioanalyzer, and only specimens with RIN >5.5 were tested. Reverse transcription of 2,500 ng of total RNA in 20 µl was carried out with M-MLV Reverse Transcriptase (Life technologies). The NB5 assay quantified the expression of the neuroblastoma-associated genes, CHGA, DCX, DDC, PHOX2B, and TH, as well as the housekeeping gene, beta-2-microglobulin (B2M). Predesigned and preoptimized probes and primer sets (**Supplementary Table S1**) were included with 2,500 ng of cDNA and amplified using standard cycling conditions using the 7900HT Fast Real-Time PCR System (Applied Biosystems). The ΔCt was chosen instead of Ct alone as ΔCt accounts for the RNA quality of the samples obtained. Note that ΔCt values may be equivalent when one specimen has no detectable NB mRNA, and other specimen is mildly positive. The detection service of NB5 assay was supplied by Shanghai Bokang Biotechnology Co. LTD.

Statistical Methods and Software

The sensitivity and specificity of the NB5 assay were determined using STATA version 15 software. The chi-square test and logistic regression analyses were carried out using SPSS version 17 software. P-values <0.05 were considered statistically significant.

RESULTS

Patient Demographics

The NB group included 101 males and 79 females aged from 1 month to 10 years. Patient characteristics, International Neuroblastoma Staging System (INSS) stage, risk group, and clinical data are listed in **Table 1**. The non-NB group contained 65 diverse hematological tumors and solid tumors, including nephroblastoma (n = 12), pancreatoblastoma (n = 4), retinoblastoma (n = 4), pheochromocytoma (n = 1), adrenocortical carcinoma (n = 1), rhabdomyosarcoma (n = 5), teratoma (n = 5), lymphoma (n = 10), lymphocytic leukemia (n = 16), hepatoblastoma (n = 5), primitive neuroectodermal tumor

TABLE 1 | Clinical characteristics of NB patients.

Total cases	N	(%)
Age (mo)		
<12	43	23.89%
≥12	137	76.11%
Sex		
Female	79	43.89%
Male	101	56.11%
Primary site		
Abdomen	162	90.00%
Thorax and other	18	10.00%
Tumor size		
>10 cm	65	36.11%
≤10 cm	115	63.89%
Tumor stage		
I	37	20.56%
II	17	9.44%
III	21	11.67%
IV	89	49.44%
Ivs	16	8.89%
Risk group		
Low risk group	97	53.89%
medium risk group	42	23.33%
high-risk group	41	22.78%
MYCN gene		
Amplification	27	15.00%
Nonamplification	153	85.00%
NSE (ng/mL)		
<370	131	72.78%
≥370	49	27.22%
Metastatic site		
Bone	35	19.44%
Bone marrow	17	9.44%
Lymph node	97	53.89%
Liver	23	12.78%

(n = 1), and an endodermal sinus tumor (n = 1). The clinical features of metastasis in the non-NB group are listed in supplementary **Table S2**.

The Sensitivity and Specificity of the NB5 Assay

We identified NB cells in 17 patients based on the evaluation of BMBs. All such individuals were clinically diagnosed as the INSS stage 4, high-risk group. All 17 patients were BM (+), while 15 patients were PB (+) (15/17, 88.2%). Among the 163 children who were BMB (–), 56 were BM (+), 40 were PB (+), and 36 were BM (+) and PB (+). The sensitivity of the NB5 assay in BM (40.6%) was significantly higher than that of BMBs (9.4%, P = 0.000). Similarly, the sensitivity of the NB5 assay in PB (30.5%) was significantly higher than that of BMBs (9.4%, P = 0.000). However, there was no significant difference between the sensitivity of the NB5 assay in BM and PB (P = 0.103).

In the non-NB group, four cases were BM (+), including those with retinoblastoma (n = 2), nephroblastoma (n = 1), and teratoma (n = 1). One case was PB (+), which was an endodermal sinus tumor (n = 1).

The sensitivity and specificity of the NB5 assay in BM were 40.5% (95%CI, 33.3–48.1%) and 93.8% (95%CI, 85–98.3%), respectively. The likelihood ratio (–), the likelihood ratio (+), the negative predictive value, and the positive predictive value are shown in **Table 2**.

TABLE 2 | Sensitivity and specificity of the NB5 assay in BM.

			[95% Confidence Interval]
Sensitivity	Pr(+|A)	40.60%	33.30–48.10%
Specificity	Pr(–|N)	93.80%	85–98.30%
ROC area	(Sens. + Spec.)/2	0.672	0.626–0.718
Likelihood ratio (+)	Pr(+|A)/Pr(+|N)	6.59	2.51–17.3
Likelihood ratio (-)	Pr(-|A)/Pr(–|N)	0.633	0.553–0.726
Odds ratio	LR(+)/LR(–)	10.4	3.77–28.6
Positive predictive value	Pr(A|+)	94.80%	87.20–98.60%
Negative predictive value	Pr(N|–)	36.30%	29–44.10%

The sensitivity and specificity of the NB5 assay in BM were 30.6% (95%CI, 23.9–37.8%) and 98.5% (95%CI, 92–100%), respectively. The likelihood ratio (–), the likelihood ratio (+), the negative predictive value, and the positive predictive value are shown in **Table 3**. The receiver operating characteristic (ROC) curves of the sensitivity of the NB5 assay and BMB indicated that the NB5 assay in PB and BM samples exhibited significantly higher sensitivity than BMB (**Figure 1**).

Analyses of the Sensitivity of the NB5 Assay

The sensitivity of the NB5 assay in BM and PB was compared to BMBs from different INSS stages (**Table 4**). The results revealed significant differences between the NB5 assay and BMBs in INSS stage III and IV samples. The sensitivity of the NB5 assay in BM and PB from INSS stage IV samples was significantly higher than that in INSS stage I–II samples (P < 0.05). The sensitivity of the NB5 assay in BM and PB in samples from the high-risk group was significantly higher than that in samples from the middle-low-risk groups (P = 0.000). The sensitivity of the NB5 assay between samples from different INSS stages are shown in **Table 5**.

The Factors Influencing the Analysis of the NB5 Assay

We carried out logistic regression analyses of liver metastases, bone metastases, lymph node metastases, tumor size >10 cm, neuron-specific enolase (NSE) ≥370 ng/ml, and MYCN with the NB5 assay. The results revealed that liver metastases (P = 0.028) and bone metastases (P = 0.002) affected the sensitivity of the NB5 assay in BM. Three factors had significant differences in the NB5 assay in PB, including liver metastases (P = 0.0001), bone

TABLE 3 | Sensitivity and specificity of the NB5 assay in PB.

			[95% Confidence Interval]
Sensitivity	Pr(+|A)	30.60%	23.90–37.80%
Specificity	Pr(–|N)	98.50%	92–100.00%
ROC area	(Sens. + Spec.)/2	0.645	0.608–0.682
Likelihood ratio (+)	Pr(+|A)/Pr(+|N)	19.9	2.81–141
Likelihood ratio (-)	Pr(–|A)/Pr(–|N)	0.705	0.637–0.781
Odds ratio	LR(+)/LR(–)	28.2	4.8–0
Positive predictive value	Pr(A|+)	98.20%	90.40–100.00%
Negative predictive value	Pr(N|–)	33.90%	27–41.10%

Pr, proportion; Sens, sensitivity; Spec, specificity; LR, likelihood ratio; +, positive; –, negative; A, Abnormal; N, Normal.

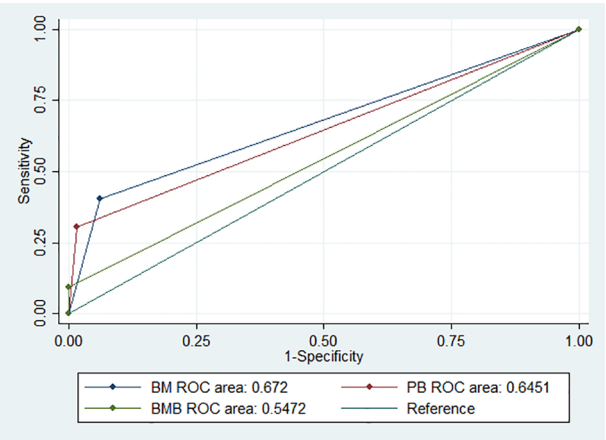


FIGURE 1 | ROC curves indicating the sensitivity of the NB5 assay in BM [area, 0.672 (95% CI 0.625–0.718)]; PB [area, 0.645 (95% CI 0.608–0.682)]; and BMB [area, 0.547 (95% CI 0.525–0.568)].

TABLE 4 | Sensitivity of the NB5 Assay between BM, PB, and BMBs from different INSS stages.

	N	BM	BMB	p value	
				BM *vs* BMB	PB *vs* BMB
I	37	4	0	0.071	0.135
II	17	2	0	0.271	0.514
III	21	7	0	**0.014**	**0.024**
IV	89	55	17	**0.0001**	**0.004**
IVs	16	5	0	**0.047**	0.082
Total	180	73	17	**0.0001**	**0.0001**

Bold value means p < 0.05 which showed significant difference.

metastases (P = 0.018), and NSE (P = 0.035). No other factors were significantly different (P > 0.05) (**Table 6**).

DISCUSSION

Previously, CHGA, DCX, DDC, PHOX2B, and TH were suggested as indicators of micrometastases in NB by single mRNA or combinations (9, 11). Several studies suggested that the analysis of multiple genes was more informative than the analysis of a single gene in detecting NB cells (15–18). In fact, the combination of all five genes is more sensitive in detecting NB micrometastases (8).

TABLE 5 | Sensitivity of the NB5 assay in BM and PB from different INSS stages.

	BM				PB			
	II	III	IV	IVs	II	III	IV	IVs
I	0.626	0.087	**0.00001**	0.135	0.635	0.087	**0.01**	0.159
II		0.197	**0.012**	0.248		0.137	**0.015**	0.205
III			0.131	0.597			0.236	0.574
IV				0.149				0.219

Bold value means p < 0.05 which showed significant difference.

TABLE 6 | Factors influencing the analysis of the BN5 assay in BM and PB.

P value	Liver	Lymph node	Bone	Tumor size>10 cm	NSE>370	MYCN
BM	**0.028**	0.332	**0.002**	0.391	0.261	0.069
PB	**0.018**	0.445	**0.0001**	0.153	**0.035**	0.284

Bold value means p < 0.05 which showed significant difference.

PHOX2B encodes a homeodomain transcription factor involved in the differentiation and development of several major noradrenergic neural cells (19). PHOX2B is highly expressed in NB and has been reported as a specific marker for MRD in NB (17, 20). DCX is specifically expressed in migrating neurons of the central and peripheral nervous systems, and regulates the microtubule cytoskeleton by signaling pathway (21). CHGA is a neuroendocrine marker that participates in coding for neurosecretory granules that promote the differentiation of NB cells (22). TH encodes the first enzyme involved in the catecholamine synthesis pathway, which serves a functional role in the detection of NB with micrometastases, as catecholamines are mainly produced by NB cells (3). Similar to TH, DDC is a key enzyme involved in catecholamine synthesis and has been claimed to be a sensitive marker for NB (23).

The present study focused on assessing the sensitivity and specificity of the NB5 assay in clinical applications (8) and explored the factors influencing the assay. The sensitivity and specificity of the assay in BM were 40.5% and 93.8%, while they were 30.6% and 98.5%, respectively, for PB samples. We found that the sensitivity of the NB5 assay in BM and PB was significantly higher than that of BMB; however, there was no significant difference between the assay sensitivity in BM and PB (P = 0.103). The sensitivity of the NB5 assay in BM and PB samples from INSS stage IV samples was significantly higher than that in INSS stages I–II samples, while the sensitivity in BM and PB in the high-risk group was significantly higher than that in the middle-low-risk groups. Logistic regression analyses indicated that liver metastases and bone metastases were the primary factors influencing the sensitivity of the NB5 assay in BM and PB. Liver and bone metastases are transferred through the blood, and thus, there is a significant correlation with the detection efficiency of the NB5 assay. Lymph nodes are transferred through the lymphatic system, and therefore, have little impact on the NB5 assay in BM and PB. The NB5 assay results were not associated with tumor size, but may be associated with tumor biology. NSE >370 ng/ml was the factor influencing the NB5 assay in PB (P = 0.035), while MYCN was a possible influencing factor in BM (P = 0.069).

Except as diagnostic methods, Virginie et al. suggested that the mRNA levels of PHOX2B, TH, and DCX in BM could be used as predictors of event-free survival (EFS) and overall survival (OS) and also to monitor the statuses of patients throughout their course of treatment for NB (16). Janine et al. found that the expression of these mRNAs varied greatly during the treatment of NB or during relapse, which rendered them excellent markers (24). However, Alexander et al. found that flow cytometric analysis for NB cells in BM was much stronger than mRNA detection in prognostic impact (25). Thus, based on this study, the prediction of the NB5 assay in the relapse and prognosis of NB warranted further evaluation, especially compared with flow cytometric analyses.

Illegitimate expressions were investigated in the non-NB group, including in retinoblastoma (n = 2), nephroblastoma (n = 1), and teratoma samples (n = 1) in BM (+) cases, and in an endodermal sinus tumor sample (n = 1) in a PB (+) case. However, most of such samples were weakly positive in the NB5 assay. Among the five genes evaluated for mRNA expression, PHOX2B was a more specific biomarker for NB than TH, DDC, CHGA, or DCX. The mRNA expression of TH, DDC and DCX was observed in hematopoietic cells and PB from healthy donors, while PHOX2B expression was limited to only NB samples (7, 9, 26). Despite the small number of false positives and false negatives in the NB5 assay, we will now conduct a larger multicenter study, combined with clinical characteristics and multiple other tests for further verification of NB. To maximize the sensitivity and specificity of the NB5 assay, we will optimize the assay to improve its accuracy of diagnosis.

CONCLUSION

In summary, the NB5 assay had significantly higher sensitivity in detecting NB with micrometastases in BM and PB than BMB. The NB5 assay had higher sensitivity in patients with INSS stage IV or in the high-risk group. Liver metastases and bone metastases were the factors that affected the sensitivity of the NB5 assay in BM and PB samples. In the future, we will analyze the relationship between the NB5 assay and prognosis and explore the relationship between tumor relapse and PHOX2B, TH, DDC, CHGA, and DCX expression.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the local Research Ethics Committee of Children's Hospital of Fudan University. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

KL, HW and ZW conceived and designed the study. CW, ZW, JL and YX collected the clinical data, performed data analysis, and wrote the paper. YJ and WY offered the assist in data collection. KL, ZW, HW and YX reviewed and edited the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This study was sponsored by the Personnel Training Program for Distinguished Medical Young Scholar (2017, to KL), the New Hundred People Plan of Shanghai Health and Family Planning Commission (2017 BR052, to KL), the Shanghai Sailing Program (17YF1401400, to ZW), the Youth Fund of Shanghai Health and Family Planning Commission Clinical Research Special Plan (20184Y0212, to ZW), the Clinical Research Plan of SHDC (no. SHDC2020CR2009A) and the Shanghai Municipal Key Clinical Specialty (no. shslczdzk05703).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.633106/full#supplementary-material

REFERENCES

1. Maris JM, Hogarty MD, Bagatell R, Cohn SL. Neuroblastoma. *LANCET* (2007) 369:2106–20. doi: 10.1016/S0140-6736(07)60983-0
2. Monclair T, Brodeur GM, Ambros PF, Brisse HJ, Cecchetto G, Holmes K, et al. The International Neuroblastoma Risk Group (INRG) staging system: an INRG Task Force report. *J Clin Oncol* (2009) 27:298–303. doi: 10.1200/JCO.2008.16.6876
3. Burchill SA, Lewis IJ, Abrams KR, Riley R, Imeson J, Pearson AD, et al. Circulating neuroblastoma cells detected by reverse transcriptase polymerase chain reaction for tyrosine hydroxylase mRNA are an independent poor prognostic indicator in stage 4 neuroblastoma in children over 1 year. *J Clin Oncol* (2001) 19:1795–801. doi: 10.1200/JCO.2001.19.6.1795
4. Cai J, Pan C, Tang Y, Chen J, Ye Q, Zhou M, et al. Minimal Residual Disease is a Prognostic Marker for Neuroblastoma With Bone Marrow Infiltration. *Am J Clin Oncol* (2012) 35:275–8. doi: 10.1097/COC.0b013e318210f51b
5. Druy AE, Shorikov EV, Tsaur GA, Popov AM, Zaychikov AN, Tuponogov SN, et al. Prospective investigation of applicability and the prognostic significance of bone marrow involvement in patients with neuroblastoma detected by quantitative reverse transcription PCR. *Pediatr Blood Cancer* (2018) 65:e27354. doi: 10.1002/pbc.27354
6. Burchill SA, Beiske K, Shimada H, Ambros PF, Seeger R, Tytgat GAM, et al. Recommendations for the standardization of bone marrow disease assessment and reporting in children with neuroblastoma on behalf of the International Neuroblastoma Response Criteria Bone Marrow Working Group. *Cancer* (2017) 123:1095–105. doi: 10.1002/cncr.30380
7. Hartomo TB, Kozaki A, Hasegawa D, Van Huyen PT, Yamamoto N, Saitoh A, et al. Minimal residual disease monitoring in neuroblastoma patients based on the expression of a set of real-time RT-PCR markers in tumor-initiating cells. *Oncol Rep* (2013) 29:1629–36. doi: 10.3892/or.2013.2286
8. Marachelian A, Villablanca JG, Liu CW, Liu B, Goodarzian F, Lai HA, 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) 23:5374–83. doi: 10.1158/1078-0432.CCR-16-2647
9. Yáñez Y, Hervás D, Grau E, Oltra S, Pérez G, Palanca S, et al. TH and DCX mRNAs in peripheral blood and bone marrow predict outcome in metastatic neuroblastoma patients. *J Cancer Res Clin* (2016) 142:573–80. doi: 10.1007/s00432-015-2054-7
10. Greze V, Brugnon F, Chambon F, Halle P, Canis M, Amiot C, et al. Highly sensitive assessment of neuroblastoma minimal residual disease in ovarian tissue using RT-qPCR-A strategy for improving the safety of fertility restoration. *Pediatr Blood Cancer* (2017) 64:1–6. doi: 10.1002/pbc.26287
11. Drui AE, Tsaur GA, Popov AM, Tuponogov SN, Shorikov EV, Tsvirenko SV, et al. [The TH, ELAVL4 and GD2 gene expression as diagnostic markers of

bone marrow lesions in patients with neuroblastoma]. *Vopr Onkol* (2012) 58:514–20. doi: 10.1007%2F978-94-007-6217-6
12. Maman S, Witz IP. "The metastatic microenvironment". In: MR Shurin, V Umansky, A Malyguine, editors. *The Tumor Immunoenvironment*. N.Y: Springer (2013). p. 15–38.
13. Visvader JE, Lindeman GJ. Cancer stem cells in solid tumours: accumulating evidence and unresolved questions. *Nat Rev Cancer* (2008) 8:755–68. doi: 10.1038/nrc2499
14. Seeger RC, Reynolds CP, Gallego R, Stram DO, Gerbing RB, Matthay KK. Quantitative tumor cell content of bone marrow and blood as a predictor of outcome in stage IV neuroblastoma: a Children's Cancer Group Study. *J Clin Oncol* (2000) 18:4067–76. doi: 10.1200/JCO.2000.18.24.4067
15. Träger C, Vernby Å, Kullman A, Øra I, Kogner P, Kågedal B. mRNAs of tyrosine hydroxylase and dopa decarboxylase but not of GD2 synthase are specific for neuroblastoma minimal disease and predicts outcome for children with high-risk disease when measured at diagnosis. *Int J Cancer* (2008) 123:2849–55. doi: 10.1002/ijc.23846
16. Viprey VF, Gregory WM, Corrias MV, Tchirkov A, Swerts K, Vicha A, et al. Neuroblastoma mRNAs predict outcome in children with stage 4 neuroblastoma: a European HR-NBL1/SIOPEN study. *J Clin Oncol* (2014) 32:1074–83. doi: 10.1200/JCO.2013.53.3604
17. Cheung IY, Feng Y, Gerald W, Cheung NKV. Exploiting Gene Expression Profiling to Identify Novel Minimal Residual Disease Markers of Neuroblastoma. *Clin Cancer Res* (2008) 14:7020–7. doi: 10.1158/1078-0432.CCR-08-0541
18. Stutterheim J, Gerritsen A, Zappeij-Kannegieter L, Yalcin B, Dee R, van Noesel MM, et al. Detecting minimal residual disease in neuroblastoma: the superiority of a panel of real-time quantitative PCR markers. *Clin Chem* (2009) 55:1316–26. doi: 10.1373/clinchem.2008.117945
19. Bachetti T, Di Zanni E, Ravazzolo R, Ceccherini I. miR-204 mediates post-transcriptional down-regulation of PHOX2B gene expression in neuroblastoma cells. *Biochim Biophys Acta* (2015) 1849:1057–65. doi: 10.1016/j.bbagrm.2015.06.008
20. Stutterheim J, Gerritsen A, Zappeij-Kannegieter L, Kleijn I, Dee R, Hooft L, et al. PHOX2B Is a Novel and Specific Marker for Minimal Residual Disease Testing in Neuroblastoma. *J Clin Oncol* (2008) 26:5443–9. doi: 10.1200/jco.2007.13.6531
21. Oltra S, Martinez F, Orellana C, Grau E, Fernandez JM, Canete A, et al. The doublecortin gene, a new molecular marker to detect minimal residual disease in neuroblastoma. *Diagn Mol Pathol* (2005) 14:53–7. doi: 10.1097/01.pas.0000149876.32376.c0
22. Gaetano C, Manni I, Bossi G, Piaggio G, Soddu S, Farina A, et al. Retinoic acid and cAMP differentially regulate human chromogranin A promoter activity during differentiation of neuroblastoma cells. *Eur J Cancer* (1995) 31A:447–52. doi: 10.1016/0959-8049(95)00038-k
23. Bozzi F, Luksch R, Collini P, Gambirasio F, Barzanò E, Polastri D, et al. Molecular detection of dopamine decarboxylase expression by means of reverse transcriptase and polymerase chain reaction in bone marrow and peripheral blood: utility as a tumor marker for neuroblastoma. *Diagn Mol Pathol* (2004) 13:135–43. doi: 10.1097/01.pdm.0000128699.14504.06
24. Stutterheim J, Zappeij-Kannegieter L, Ora I, van Sluis PG, Bras J, den Ouden E, et al. Stability of PCR targets for monitoring minimal residual disease in neuroblastoma. *J Mol Diagn* (2012) 14:168–75. doi: 10.1016/j.jmoldx.2011.12.002
25. Popov A, Druy A, Shorikov E, Verzhbitskaya T, Solodovnikov A, Saveliev L, et al. Prognostic value of initial bone marrow disease detection by multiparameter flow cytometry in children with neuroblastoma. *J Cancer Res Clin Oncol* (2019) 145(2):535–42. doi: 10.1007/s00432-018-02831-w
26. Corrias MV, Haupt R, Carlini B, Cappelli E, Giardino S, Tripodi G, et al. Multiple target molecular monitoring of bone marrow and peripheral blood samples from patients with localized neuroblastoma and healthy donors. *Pediatr Blood Cancer* (2012) 58:43–9. doi: 10.1002/pbc.22960

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

ORIGINAL RESEARCH
published: 10 June 2021
doi: 10.3389/fonc.2021.637482

Classifying Medulloblastoma Subgroups Based on Small, Clinically Achievable Gene Sets

Sivan Gershanov[1], Shreyas Madiwale[2,3†], Galina Feinberg-Gorenshtein[2†], Igor Vainer[1], Tamar Nehushtan[1], Shalom Michowiz[3,4], Nitza Goldenberg-Cohen[5,6,7], Yehudit Birger[2], Helen Toledano[3,8] and Mali Salmon-Divon[1,9]*

[1] Department of Molecular Biology, Ariel University, Ariel, Israel, [2] Hemato-Oncology Laboratory, Division of Pediatric Hematology Oncology, Schneider Children's Medical Center of Israel, Petach Tikva, Israel, [3] Sackler Faculty of Medicine, Tel-Aviv University, Tel-Aviv, Israel, [4] Department of Pediatric Neurosurgery, Schneider Children's Medical Center of Israel, Petach-Tikva, Israel, [5] Department of Ophthalmology, Bnai Zion Medical Center, Haifa, Israel, [6] The Krieger Eye Research Laboratory, Felsenstein Medical Research Center, Rabin Medical Center, Petach-Tikva, Israel, [7] The Ruth and Bruce Rappaport Faculty of Medicine, Technion, Haifa, Israel, [8] Department of Pediatric Oncology, Schneider Children's Medical Center of Israel, Petach-Tikva, Israel, [9] Adelson School of Medicine, Ariel University, Ariel, Israel

OPEN ACCESS

Edited by:
Jing He,
Guangzhou Medical University, China

Reviewed by:
Qiangqiang Xia,
University of Alabama at Birmingham, United States
Yiyang Wu,
Vanderbilt University Medical Center, United States

***Correspondence:**
Mali Salmon-Divon
malisa@ariel.ac.il

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Pediatric Oncology, a section of the journal Frontiers in Oncology

Received: *03 December 2020*
Accepted: *10 May 2021*
Published: *10 June 2021*

Citation:
Gershanov S, Madiwale S, Feinberg-Gorenshtein G, Vainer I, Nehushtan T, Michowiz S, Goldenberg-Cohen N, Birger Y, Toledano H and Salmon-Divon M (2021) Classifying Medulloblastoma Subgroups Based on Small, Clinically Achievable Gene Sets. Front. Oncol. 11:637482. doi: 10.3389/fonc.2021.637482

As treatment protocols for medulloblastoma (MB) are becoming subgroup-specific, means for reliably distinguishing between its subgroups are a timely need. Currently available methods include immunohistochemical stains, which are subjective and often inconclusive, and molecular techniques—e.g., NanoString, microarrays, or DNA methylation assays—which are time-consuming, expensive and not widely available. Quantitative PCR (qPCR) provides a good alternative for these methods, but the current NanoString panel which includes 22 genes is impractical for qPCR. Here, we applied machine-learning–based classifiers to extract reliable, concise gene sets for distinguishing between the four MB subgroups, and we compared the accuracy of these gene sets to that of the known NanoString 22-gene set. We validated our results using an independent microarray-based dataset of 92 samples of all four subgroups. In addition, we performed a qPCR validation on a cohort of 18 patients diagnosed with SHH, Group 3 and Group 4 MB. We found that the 22-gene set can be reduced to only six genes (*IMPG2*, *NPR3*, *KHDRBS2*, *RBM24*, *WIF1*, and *EMX2*) without compromising accuracy. The identified gene set is sufficiently small to make a qPCR-based MB subgroup classification easily accessible to clinicians, even in developing, poorly equipped countries.

Keywords: medulloblastoma, subgroup classification, biomarkers, machine learning, gene expression

INTRODUCTION

Medulloblastoma (MB)—the most common malignant brain tumor in children—demonstrates extremely high biological and clinical heterogeneity (1). Accordingly, it is divided into four subgroups, each representing distinct clinical, biological, and genetic profiles and involves a distinct activation pathway (2–7): WNT (or Group 1) involves Wingless pathway signaling (3); SHH (or Group 2) involves sonic hedgehog pathway signaling (4); Group C (or Group 3) involves photoreceptor

and GABAergic pathway signaling; and Group D (or Group 4) involves neuronal and glutamatergic signaling (6). Importantly, although the histological presentation of the different subgroups is often similar, their response to treatment and the clinical outcomes are subgroup-specific (8); therefore, the World Health Organization has recently recommended that molecular markers be integrated as part of MB tumor diagnostic criteria (9). This recommendation is currently limited to distinguishing between the WNT and SHH subgroups, but means for distinguishing between Group 3 and Group 4 are already clinically required.

Transcriptomic analyses have shown promising potential for distinguishing between the four MB subgroups. Most notably, Northcott et al. (10) employed the NanoString technology that is based on a direct molecular barcoding of target molecules, followed by digital detection of their expression, to identify a set of 22 genes that can distinguish between the four MB subgroups (11); this set is currently used in many clinical laboratories worldwide. However, NanoString has two important limitations vis-à-vis its clinical use for MB subgroup classification: first, it is expensive and currently unavailable in most medical institutes, especially in developing countries; and second, it is not sufficiently reliable and shows relatively high rates of MB misdiagnosis and subgroup misclassification, especially between groups C and D (12). DNA methylation is more reliable in MB subgroup classification (13), but it is even more costly than NanoString and is unavailable in most medical institutes. Thus, there is a need to develop a reliable—yet simple and cost-effective—means of MB subgroup classification, which could be utilized through readily available technologies, such as qPCR. Indeed, Kunder et al. (14) used a quantitative PCR (qPCR) analysis, based on 21 biomarkers (including 12 protein-coding genes and nine microRNA expression profiles), but this number of genes is still high, hence impractical for qPCR test in the clinic.

To meet this need, this study aimed to identify sets of genes that comprise the minimal number of genes required for reliably differentiating between all four MB subgroups. To achieve this goal, we fed published data from microarray studies of MB, which comprehensively characterized the expression pattern of thousands of genes simultaneously, as input for machine-learning-based classifiers for cancer classification (15–17). Such classifiers were previously applied to discriminate anaplastic from non-anaplastic MB image regions (18) and to predict subtypes of the four MB subgroups (19), but, to the best of our knowledge, they have not been used to extract sets of potential biomarkers from microarray data. Indeed, this approach has enabled us to identify both protein-coding genes and non-coding RNAs as potential biomarkers for MB subgroup classification. These biomarkers could reliably be used in MB-related diagnosis, prognosis, and clinical decision-making, and they could later be used to identify potential drug targets.

METHODS

Public Datasets

To identify minimal gene sets for MB subgroup classification, we used the dataset GSE85217 (19) to train and test the algorithms, and the datasets GSE37418 (20) and GSE41842 (21) for validation. All datasets are publicly available, quality-controlled, mRNA expression matrixes that were generated using Affymetrix microarrays. The datasets were downloaded from the gene expression omnibus (GEO) (22) database, which contains data on subjects diagnosed with any of the four MB subgroups. Specifically, the GSE85217 dataset comprises 763 samples (70 WNT samples, 223 SHH samples, 144 Group 3 samples, and 326 Group 4 samples), which were molecularly classified by inferring the expression levels of 22 MB signature genes, using the NanoString technology. The GSE37418 dataset comprises 73 samples (14 WNT samples, 13 SHH samples, 18 Group 3 samples, and 47 Group 4 samples), which were segregated into four MB subgroups using mRNA expression profiling and immunohistochemistry. The GSE41842 dataset comprises 19 samples (six WNT samples, three SHH samples, two Group 3 samples, and eight Group 4 samples), which were molecularly classified using unsupervised hierarchical clustering with the 1000 most differentially expressed genes. All samples included in these datasets were collected from fresh frozen tissue samples. Demographic and clinical data available for the above datasets is provided in **Supplementary File 1** - **Public Datasets**.

Public Dataset Normalization

For the datasets GSE85217 and GSE41842, we downloaded the robust multi-array average normalized matrixes. For the GSE37418 dataset, we normalized the gene expression data by using the MAS 5.0 algorithm; therefore, we downloaded the raw CEL files and performed a robust multi-array average normalization by using the affy R package (23).

Microarray Gene Annotation

To identify and match gene symbols to the probe ID of molecules in the two Affymetrix microarray datasets mentioned above, we used the biomaRt R package (24).

Machine Learning Algorithms for Classification

We used the Waikato environment for knowledge analysis (WEKA) workbench software (25)—a Java-based machine learning algorithm collection—for all classification analyses. We initially employed four well-known algorithms: C4.5 Decision Tree (DT) (algorithm J48) (26); Decision Rules (RIPPER Rule Induction algorithm JRip) (27); Random Forest (28); and Support Vector Machines (SVM) using Sequential Minimal Optimization (SMO) (29–31). We chose the default parameters for all algorithms and used a 10-fold cross-validation to prevent overfitting. A detailed description of the methodology is provided in the **Supplementary Information** section.

In addition to the four well-established algorithms mentioned above, we designed and developed a novel algorithm that we termed SVM Attribute Ranking and Combinations (SARC). The main steps of the algorithm included: 1. building six pairwise models for the four MB subgroups, using the SVM classification model with a linear kernel; 2. for each binary classifier, ranking the attributes according to their squared weight; 3. for each subgroup, performing an aggregation of

attribute ranks by summarizing each attribute rank to produce final ranks, leading to a list of top attributes; 4. using a combination of 0–12 top attributes (**Supplementary Table S1**) of each subgroup as the de-facto feature-selection method for the final classifier; and 5. producing an SVM classifier based on the 134 combinations, eventually choosing the smallest, best-performing combinations for each accuracy level. When using the NanoString 22-gene set to build the classification model, we used combinations of all 22 attributes. A more detailed description is provided in the Supplementary Information section (**Supplementary Tables S2–S5** and **Supplementary Figure S1**). We used the top nine reduced gene sets output by the SARC classifier (**Supplementary Table S3**) as input for the independent public dataset validation.

Visualization

We generated clustering plots by using t-SNE, a non-linear dimensionality-reduction algorithm, with the Rtsne (32, 33) R package, version 0.15. Each plot was made with 1,000 iterations and the perplexity set to 30.

Patient Cohort and Tumor Collection for Validation

An independent cohort of pediatric and young adult patients diagnosed with MB was collected at the Pediatric Hematology & Oncology Department at the Schneider Children's Medical Center, Israel, and from the Pathology Department at the Rabin Medical Center, Israel. Since 2013, the standard of care has been to assign MB subgroup by using the NanoString nCounter Technology (NanoString Technologies, Seattle, WA), as described previously (10). We selected only the patients with MBs whose tumor subgroup had been classified by NanoString for clinical purposes and who had remaining RNA for real time PCR validation. Group-A MB (WNT) samples were not available to us, hence only SHH, Group 3 and Group 4 were included in qPCR analysis. The RNA was obtained from primary tumors for the initial clinical standard of care test at the time of diagnosis before any treatment; we did not extract any new RNA for this study. Altogether, the cohort used for validation comprised 18 children and young adults (8 males, 10 females; mean age at diagnosis: 6.53 ± 4.5 years), who were classified by NanoString as either SHH, Group 3, Group 4, or non-WNT/SHH (i.e., either Group 3 or Group 4) MBs (n = 5, 3, 8, and 2 respectively; **Supplementary Table S6**). Of the 18 patients, 11 were diagnosed with a localized disease and six were diagnosed with a metastatic disease (four M1 and two M2); data were unavailable for one patient (SHH4). All patients were treated with chemotherapy, eight patients underwent autologous bone marrow transplantation, and 14 patients received radiation therapy in addition to chemotherapy. Four patients did not receive radiation therapy due to their young age (<3 y). Disease recurrence was recorded in three patients. Four patients died altogether, including one who died from disease progression and three who died from other causes: patient SHH5 died as a result of secondary AML, patient C2 died of secondary diffuse intrinsic pontine glioma (DIPG) despite not receiving radiation, and patient D7 died from post-operative complications prior to therapy. All tissue samples, were from freshly frozen (FF) tissues. The study design adhered to the tenets of the Declaration of Helsinki and was approved by the local IRB and the National Review Board of the Israel Ministry of Health.

Reverse-Transcription (RT) and qPCR

The cDNA synthesis was performed using the cDNA Reverse Transcription Kit (ABI High Capacity cDNA reverse-transcription kit, Cat No. 4368813) and was followed by a quantitative expression analysis using the SYBR Green qPCR Kit (PowerUP SYBR green master mix ABI, Cat No. A25776) according to the manufacturer's instructions. The expression levels of each gene were normalized to those of *GAPDH*. Data and melting curves were analyzed by using the QuantStudio3 real-time instrument (Applied Biosystems, Waltham, Massachusetts(and associated software. Primer sequences are provided in **Supplementary Table S7**.

qPCR Expression Level Analysis

The expression level of each protein-coding gene was normalized to that of GAPDH, as determined by the delta cycle threshold (dCt) method. Since we did not have a control (non-MB) cerebellum sample, we used dCt for unsupervised hierarchical clustering, generated using the pvclust (34) R package, version 2.0-0. Euclidean was used as the distance measure and ward.D2 was used as the linkage method. For each cluster in the dendrogram, p-values were calculated by multiscale bootstrap resampling (nboot = 1000).

RESULTS

Applying Machine-Learning Algorithms for MB Subgroup Classification

To detect the minimal set of genes that accurately distinguishes between MB subgroups, we employed four well-known machine-learning algorithms, including Decision Tree, Decision Rules, Random Forest, and Support Vector Machines (SVM-SMO). The different algorithms were run in two modes. In the first, all 21,641 attributes (defined as Probe ID, **Supplementary File 2**) were used as input to the algorithm; in the second, the algorithms were fed with the known NanoString 22-gene set. The attributes selected by each algorithm for classification in either mode, as well as the classification accuracy, are indicated in **Table 1**. All four algorithms were highly accurate, as compared with the known 22-gene set of the NanoString panel. The Decision Tree and Decision Rules models resulted in a reduced gene sets (9 and 10 genes, respectively) with a similar or a slightly higher accuracy than that of the 22-gene signature set, while Random Forest and SVM-SMO used all input attributes and demonstrated the highest accuracy (**Table 1**).

The SVM Attribute Ranking and Combinations (SARC) Classifier Displays the Highest Accuracy

Despite the high accuracy of the Random Forest and SVM-SMO algorithms, they did not enable us to derive a gene-set output

TABLE 1 | The accuracies of the sets of attributes selected for classification by each algorithm, based on the GSE85217 dataset (n = 763 MB samples).

Algorithm	Input[1]	Accuracy (%)	Attributes required for classification (output)[2]	Number of attributes required for classification
Decision tree[3]	All attributes	95.5	*OTX2, TMEM51, AIF1L, RASSF4, DYNC1I1, TRAK2, RPL3, C1orf112, RABGAP1*	9
	22 genes	94.5	*ATOH1, WIF1, RBM24, PDLIM3, NRL, TNC, GABRA5, KHDRBS2, SFRP1, IMPG2*	10
Decision rules[3]	All attributes	94.2	*PDLIM4, NPR3, PDE10A, PDK2, RALGPS2, SHD, BSG, ARNTL2, USP2, FBXL21*	10
	22 genes	94	*GAD1, PDLIM3, WIF1, EYA1, NPR3, EYS, RBM24, GABRA5, EOMES, EMX2, KCNA1, ATOH1, IMPG2*	13
Random forest	All attributes	97.8	*All attributes*	21,641
	22 genes	97.1	*All attributes*	22
SVM-SMO	All attributes	98.4	*All attributes*	21,641
	22 genes	97.8	*All attributes*	22

[1]*Attribute sets that were used as inputs for the algorithm.*
[2]*Attributes chosen by each algorithm for classification.*
[3]*Detailed results obtained from these algorithms can be found in* ***Supplementary Figure S1*** *and* ***Supplementary Table S5****.*

because they are non-interpretative regarding the attributes being used during the classification process. Therefore, we developed a novel algorithm—the SVM Attribute Ranking and Combinations (SARC)—in an attempt to obtain accuracy levels that are comparable to or higher than those of the Random Forest and SVM-SMO algorithms, while allowing a tailored feature selection.

When we introduced all genes as input, the SARC classifier provided a list of gene sets (between four and 32 biomarkers in each set; **Figure 1A** and **Supplementary Table S2**), of which the top 14 sets demonstrated accuracy levels between 92.4% and 98.56%. In most sets, the lowest number of genes necessary for classification was in the WNT and SHH group, while the highest number necessary was in Group 4. When we introduced the NanoString 22-gene set as input, the SARC classifier provided nine gene sets (**Figure 1B** and **Supplementary Table S3**) that comprised between three and 15 biomarkers and demonstrated an accuracy between 92.01% for the smallest set (three genes) and 98.3% for the largest set (15 genes). **Table 2** indicates the gene sets that demonstrated the highest accuracy levels; these include a set of 32 genes obtained when all genes were introduced to the SARC classifier as input, and a set of 15 genes obtained when the NanoString 22-gene set was introduced as input. Indeed, the SARC algorithm demonstrated the highest accuracy of all five tested algorithms.

The SARC Classifier Reduces the Number of Biomarkers Required for Accurate Classification to Only Six Genes – Validation in an Independent Dataset

The best-performing sets used by the SARC algorithm for classification comprised either 32 or 15 attributes. This number of biomarkers is too large to be practically used for qPCR in the clinic. The performances of the various reduced sets of genes (**Supplementary Tables S2**, **S3**) suggested that the number of biomarkers can be reduced to only six genes (*IMPG2*, *NPR3*, *KHDRBS2*, *RBM24*, *WIF1*, and *EMX2*) without compromising accuracy (**Supplementary Table S3** and **Supplementary Figure S2**). To validate this assumption, we tested the classification accuracy of these nine reduced sets (listed in **Supplementary Table S3**) in two independent public datasets, GSE37418 (20) and GSE41842 (21), which, together, contain 92 samples (73 and 19 samples, respectively) of all four MB subgroups. The classification accuracy of the six-gene set was 93.48%, which is higher than the accuracy observed when all 22 NanoString genes were introduced to the model (**Table 3**) accuracy, sensitivity, and specificity is specified in **Supplementary File 3 – Confusion Matrix**.

Next, we created t-SNE plots (**Figure 2**) to visualize the performance of the full NanoString and the reduced gene sets on the validation dataset (n = 92 samples). Both gene sets performed well in separating the MB groups, with a slightly

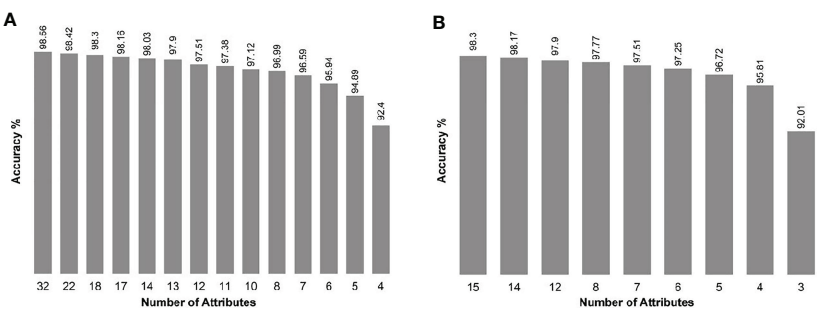


FIGURE 1 | Accuracy of the smallest best-performing gene sets output by the SARC classifier, applied on the GSE85217 dataset (n = 763 samples), **(A)** when introducing all 21,641 attributes as input, and **(B)** when introducing the Nanostring 22-gene set as input.

TABLE 2 | The accuracies of the top set of attributes selected for classification by the SARC algorithm for each input, based on the GSE85217 dataset (n = 763 MB samples).

Input[1]	Accuracy (%)	Attributes required for classification (output)[2]	Number of attributes required for classification
All attributes	98.6	*AL513318.2, NPR3, LMX1A, BARHL1, SIX6, GRM8, NID2, CA4, ZIC2, RBM24, ZIC5, DDX31, SNCAIP, NEUROG1, ATOH1, KCNA5, PEX5L, GLRA1, NDP, ZFHX4, RPGRIP1, PAX3, WIF1, TMEM51, ADGRL3, DLX3, TMEM51-AS1, TMEM132C, PGM5, PDE11A, NKD1, FZD10*	32
22 genes	98.3	*KHDRBS2, RBM24, EMX2, PDLIM3, NPR3, UNC5D, IMPG2, TNC, GABRA5, GAD1, OAS1, ATOH1, EYA1, EOMES, SFRP1*	15

[1]*Attribute that were used as inputs for the algorithm.*
[2]*Attributes chosen by the algorithm for classification.*

better separation observed by the 12-gene set, whose performance was similar to that of the full 22-gene set. Not surprisingly, the WNT and SHH groups are presented as clearly separated clusters, while the separation between Group 3 and Group 4 is less pronounced.

Classifying MB Subgroups in an Independent Clinical Cohort Based on the SARC Reduced Gene Set, Using qPCR

As a proof-of-concept that the suggested gene sets can be used to classify MB subgroups in patients by using gene expression levels generated by qPCR, we validated our results on an independent cohort of 18 patients, whose MB subgroup was previously classified by NanoString. The cohort included five patients with SHH MB, three patients with Group 3 MB, eight patients with Group 4 MB, and two patients who were classified as non-WNT/SHH MB, i.e., with either Group 3 or Group 4 MB (**Figure 3** and **Supplementary Table S6**). At the time of completion of this study, we did not have samples from patients with a WNT MB; hence, this subgroup was not included in the validation.

The unsupervised hierarchical clustering was performed using the expression levels (namely, dCt) of the reduced six-gene set (*IMPG2*, *NPR3*, *KHDRBS2*, *RBM24*, *WIF1*, and *EMX2*).

The reduced gene set performed well in classifying the patients to their diagnosed MB subgroups (**Figure 4A** and **Supplementary Figure S3A**). Adding the two patients whose subgroup was undefined resulted in the clustering of patient GrpC.D14 with patients from Group 3, and of patient GrpC.D15 with patients from Group 4 (**Figure 4B** and **Supplementary Figure S3B**). Hence, our data demonstrate the potential of using this small set of genes for an easy and accessible qPCR-based MB subgroup classification.

TABLE 3 | Classification accuracy of the reduced genes sets (12 genes or fewer), as compared with the full, 22-gene NanoString set, used on the independent validation datasets GSE37418 and GSE41842 (n = 92 MB samples altogether).

Number of attributes	Accuracy (%)	Input set for validation[1]
22	91.30	*EYS, TNC, IMPG2, OAS1, EYA1, SFRP1, KCNA1, RBM24, KHDRBS2, NPR3, GAD1, NRL, PDLIM3, DKK2, WIF1, UNC5D, EOMES, HHIP, EMX2, ATOH1, MAB21L2, GABRA5*
12	96.74	*IMPG2, NPR3, EMX2, RBM24, SFRP1, NRL, TNC, PDLIM3, KHDRBS2, UNC5D, ATOH1, WIF1*
8	90.22	*IMPG2, KHDRBS2, RBM24, EMX2, PDLIM3, NPR3, UNC5D, WIF1*
7	93.48	*IMPG2, KHDRBS2, RBM24, EMX2, PDLIM3, NPR3, WIF1*
6	93.48	*IMPG2, NPR3, KHDRBS2, RBM24, WIF1, EMX2*
5	82.61	*IMPG2, NPR3, KHDRBS2, RBM24, WIF1*
4	81.52	*IMPG2, KHDRBS2, RBM24, WIF1*

[1]*Attribute sets that were used as input for the validation based on the SARC classifier output, chosen from the GSE85217 dataset (**Supplementary Table S3**).*

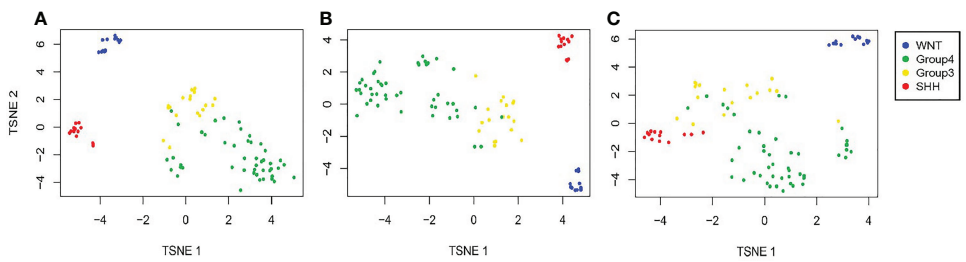


FIGURE 2 | Validation of the predicted classification set outputs created by the SARC classifier. Expression t-SNE of the independent datasets GSE37418 and GSE41842 (n = 92) based on **(A)** a 22-gene NanoString panel set, **(B)** 12 genes out of the 22 Nanostring panel, and **(C)** six genes out of the 22 Nanostring panel.

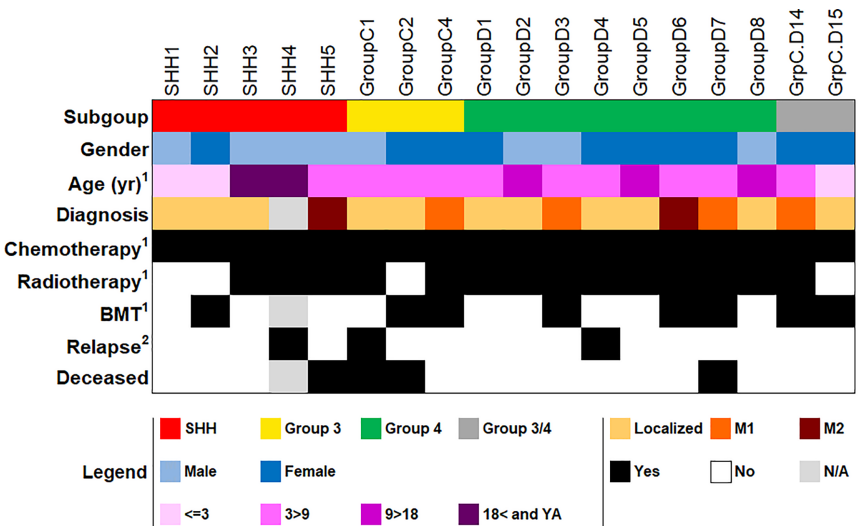


FIGURE 3 | Demographic and clinical data of the patient cohort used for qPCR validation (n = 18). BMT, bone marrow transplantation; YA, young adult; N/A, not available. [1]At first diagnosis. [2]As of the completion of this study. More detailed information in **Supplementary Table S6**.

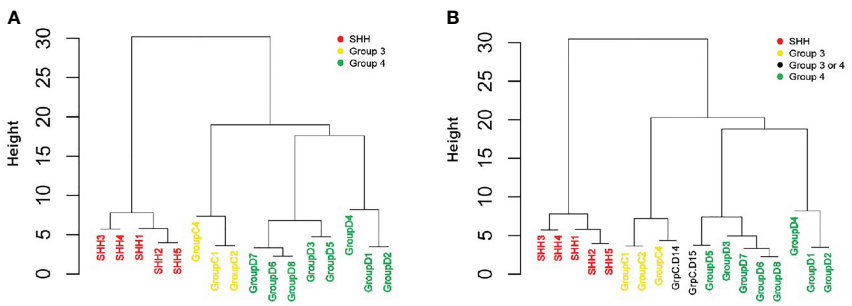


FIGURE 4 | qPCR-based classification of an independent cohort, using reduced six-gene setout of the 22-gene NanoString set (*IMPG2*, *NPR3*, *KHDRBS2*, *RBM24*, *WIF1*, and *EMX2*). An unsupervised hierarchical clustering of gene expression levels was generated by using qPCR (dCt) values. **(A)** A cohort of 16 patients who were classified by NanoString as having either SHH, Group 3, or Group 4 MBs (n = 5, 3, and 8, respectively; see **Figure 3** and **Supplementary Table S6**). **(B)** The same cohort, but with the addition of two patients who were classified as having a non-WNT/SHH MB. The Height (y axis) is a measure of closeness of either individual data points or clusters.

DISCUSSION

Using feature selection and machine learning classification, we were able to identify potential gene sets with fewer attributes and a more accurate subgroup classification of MB tumors, as compared with the NanoString 22-gene set currently used in several clinical institutions. More specifically, our SARC algorithm was able to reduce the 22-gene set to only six genes that reliably differentiated between the four MB subgroups. The reduced gene set includes *WIF1* and *EMX2* which are known activated Wingless pathway signaling in WNT subgroup. Widely accepted biomarkers *IMPG2*, and *NPR3* identity Group 3, as well as *KHDRBS2*, and *RBM24* recognise Group 4 MB tumors (10). Notably, none of these genes are classical biomarkers of SHH subgroup, and probably the combination of these genes' expression contributes to accurate SHH group classification.

All genes in this set are known and have commercially available primers, which should enable most clinical laboratories to accurately classify MB subgroups at a reasonable price and within a reasonable timeframe, to the benefit of both patients and clinicians alike.

The tumor subgroups in the GSE85217 dataset that we used to construct the model were originally determined according to the expression levels of the 22 genes by the NanoString technology. Hence, it was not surprising that the accuracy levels of all tested algorithms were very high when they were based on this 22-gene set. However, in the independent validation datasets, the subgroups were classified by using a different approach: in the GSE41842, the subgroups were classified according to unsupervised hierarchical clustering using the 1000 most differentially expressed genes, while in GSE37418, the subgroups were classified using the mRNA expression of 2,750 probes with the highest median absolute

difference (MAD) score and with immunohistochemistry to provide an additional assessment for WNT and SHH subgroups (20, 21). Therefore, the high accuracy obtained at the validation step demonstrates the promising potential of using fewer biomarkers, such as 12 or six genes having higher accuracy (96.74% and 93.48% respectively) than the 22-gene set (91.3%). This potential was further demonstrated by the qPCR-based classification that we obtained by using the reduced six-gene set in the cohort of 22 pediatric patients. We included in this qPCR validation two patients whose subgroup was defined as "non-SHH/WNT", one clustered with patients from Group 3, and one clustered with patients from Group 4. Methylation may help to determine the subgroup of these patients, to check if the reduced gene set model classified them correctly. Unfortunately, methylation was unavailable at the Schneider Children's Medical Center as it is in most clinical centers. Future studies on larger cohorts are required to test the effectiveness of the reduced six-gene set in decreasing MB misclassification, in general, and in accurately distinguishing between Group 3 and Group 4 MBs, in particular.

Our study has several limitations; first, due to a lack of WNT samples, we were unable to add this subgroup to the qPCR validation step. Nevertheless WNT subgroup is easily identifiable by other currently available methods, e.g. using a combination of immunohistochemistry for nucleopositive beta-catenin, and FISH for monosomy of chromosome 6 (35). Future studies should use qPCR to test the reduced gene set of all MB subgroups. Second, since both our modeling and validation steps were performed on primary tumors, we cannot comment on the performance of the reduce set on metastasis, relapse, or progression disease samples. Third, our models do not distinguish between the different subtypes of each subgroup; instead, the algorithm was trained to classify the different subgroups regardless of their molecular states, especially since the current clinical recommendations focus only on the main subgroups and do not consider the different subtypes. Future studies should take intertumoral heterogeneity within MB subgroups into consideration. Finally, the current study focused on the minimal set of genes required for MB subgroup classification, but implementation in a clinical setting requires that the suggested gene set is adapted to an individual patient setting. Such a setting should include a cut-off of the detection of expression level for each gene, a definition of the reference that should be used, a statement of the type of normalization that should be employed, etc.

CONCLUSIONS

Since personalized treatment in oncology assumes that each tumor harbors a unique variation of the human genome and should be treated accordingly, it is crucial to correctly classify the molecular subgroup of the tumor. Indeed, as treatment (e.g., radiation and chemotherapy) protocols are becoming subgroup-specific and usually commence within 28 days of operation, our machine-learning approach, which yielded concise and reliable gene sets, provides a significant clinical advantage over available MB subgroup classification methods.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The study design adhered to the tenets of the Declaration of Helsinki and was approved by the local IRB and the National Review Board of the Israel Ministry of Health. Written informed consent from the participants' legal guardian/next of kin was not required to participate in this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

Conceptualization: SG and MS-D. Formal analysis: SG, IV, and TN. Investigation: SG, HT, and MS-D. Methodology: SG, IV, and MS-D. Resources: GF-G, SMi, NG-C, and HT. Supervision: YB, HT, and MS-D. Validation: SMa, GF-G, and YB. Visualization: SG. Writing – original draft: SG, HT, and MS-D. Writing – review and editing: SG, SM, GF-G, IV, TN, NG-C, YB, HT, and MS-D. All authors contributed to the article and approved the submitted version.

FUNDING

This study was funded by the Levi Eshkol Fund, Ministry of Science, Technology & Space, Israel, grant number 3-12624, which provided SG's scholarship.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.637482/full#supplementary-material

REFERENCES

1. Louis DN, Ohgaki H, Wiestler OD, Cavenee WK, Burger PC, Jouvet A, et al. The 2007 WHO Classification of Tumours of the Central Nervous System. *Acta Neuropathol* (2007) 114:97–109. doi: 10.1007/s00401-007-0243-4
2. Gupta T, Shirsat N, Jalali R. Molecular Subgrouping of Medulloblastoma: Impact Upon Research and Clinical Practice. *Curr Pediatr Rev* (2015) 11:106–19. doi: 10.2174/1573396311666150702104030
3. Clifford SC, Lusher ME, Lindsey JC, Langdon JA, Gilbertson RJ, Straughton D, et al. Wnt/Wingless Pathway Activation and Chromosome 6 Loss Characterize a Distinct Molecular Sub-Group of Medulloblastomas Associated With a Favorable Prognosis. *Cell Cycle* (2006) 5:2666–70. doi: 10.4161/cc.5.22.3446
4. Kool M, Korshunov A, Remke M, Jones DTW, Schlanstein M, Northcott PA, et al. Molecular Subgroups of Medulloblastoma: An International Meta-Analysis of Transcriptome, Genetic Aberrations, and Clinical Data of

WNT, SHH, Group 3, and Group 4 Medulloblastomas. *Acta Neuropathol* (2012) 123:473–84. doi: 10.1007/s00401-012-0958-8
5. Sexton-Oates A, MacGregor D, Dodgshun A, Saffery R. The Potential for Epigenetic Analysis of Paediatric CNS Tumours to Improve Diagnosis, Treatment and Prognosis. *Ann Oncol* (2015) 26:1314–24. doi: 10.1093/annonc/mdv024
6. Taylor MD, Northcott PA, Korshunov A, Remke M, Cho YJ, Clifford SC, et al. Molecular Subgroups of Medulloblastoma: The Current Consensus. *Acta Neuropathol* (2012) 123:465–72. doi: 10.1007/s00401-011-0922-z
7. Northcott PA, Korshunov A, Witt H, Hielscher T, Eberhart CG, Mack S, et al. Medulloblastoma Comprises Four Distinct Molecular Variants. *J Clin Oncol* (2011) 29:1408–14. doi: 10.1200/JCO.2009.27.4324
8. Ramaswamy V, Remke M, Bouffet E, Bailey S, Clifford SC, Doz F, et al. Risk Stratification of Childhood Medulloblastoma in the Molecular Era: The Current Consensus. *Acta Neuropathol* (2016) 131:821–31. doi: 10.1007/s00401-016-1569-6
9. Louis DN, Perry A, Reifenberger G, Von Deimling A, Figarella-Branger D, Cavenee WK, et al. *The 2016 World Health Organization Classification of Tumors of the Central Nervous System: A Summary*. Berlin Heidelberg: Springer (2016) p. 803–20. doi: 10.1007/s00401-016-1545-1
10. Northcott PA, Shih DJH, Remke M, Cho YJ, Kool M, Hawkins C, et al. Rapid, Reliable, and Reproducible Molecular Sub-Grouping of Clinical Medulloblastoma Samples. *Acta Neuropathol* (2012) 123:615–26. doi: 10.1007/s00401-011-0899-7
11. Leal LF, Evangelista AF, de Paula FE, Caravina Almeida G, Carloni AC, Saggioro F, et al. Reproducibility of the NanoString 22-Gene Molecular Subgroup Assay for Improved Prognostic Prediction of Medulloblastoma. *Neuropathology* (2018) 38(5):475–83. doi: 10.1111/neup.12508
12. Korshunov A, Chavez L, Northcott PA, Sharma T, Ryzhova M, Jones DTW, et al. DNA-Methylation Profiling Discloses Significant Advantages Over NanoString Method for Molecular Classification of Medulloblastoma. *Acta Neuropathol* (2017) 134(6):965–7. doi: 10.1007/s00401-017-1776-9
13. Gomez S, Garrido-Garcia A, Garcia-Gerique L, Lemos I, Suñol M, de Torres C, et al. A Novel Method for Rapid Molecular Subgrouping of Medulloblastoma. *Clin Cancer Res* (2018) 24:1355–63. doi: 10.1158/1078-0432.CCR-17-2243
14. Kunder R, Jalali R, Sridhar E, Moiyadi A, Goel N, Goel A, et al. Real-Time PCR Assay Based on the Differential Expression of MicroRNAs and Protein-Coding Genes for Molecular Classification of Formalinfixed Paraffin Embedded Medulloblastomas. *Neuro Oncol* (2013) 15:1644–51. doi: 10.1093/neuonc/not123
15. Nguyen DV, Rocke DM. Tumor Classification by Partial Least Squares Using Microarray Gene Expression Data. *Bioinformatics* (2002) 18(1):39–50. doi: 10.1093/bioinformatics/18.1.39
16. Statnikov A, Aliferis CF, Tsamardinos I, Hardin D, Levy S. A Comprehensive Evaluation of Multicategory Classification Methods for Microarray Gene Expression Cancer Diagnosis. *Bioinformatics* (2005) 21(5):631–43. doi: 10.1093/bioinformatics/bti033
17. Statnikov A, Wang L, Aliferis CF. A Comprehensive Comparison of Random Forests and Support Vector Machines for Microarray-Based Cancer Classification. *BMC Bioinformatics* (2008) 22(9):319. doi: 10.1186/1471-2105-9-319
18. Cruz-Roa A, Arévalo J, Judkins A, Madabhushi A, González F. *A Method for Medulloblastoma Tumor Differentiation Based on Convolutional Neural Networks and Transfer Learning*. E Romero, N Lepore, JD García-Arteaga, J Brieva, editors. International Society for Optics and Photonics. (2015). p. 968103. doi: 10.1117/12.2208825
19. Cavalli FMG, Remke M, Rampasek L, Peacock J, Shih DJH, Luu B, et al. Intertumoral Heterogeneity Within Medulloblastoma Subgroups. *Cancer Cell* (2017) 31:737–54.e6. doi: 10.1016/j.ccell.2017.05.005
20. Robinson G, Parker M, Kranenburg TA, Lu C, Chen X, Ding L, et al. Novel Mutations Target Distinct Subgroups of Medulloblastoma. *Nature* (2012) 488:43–8. doi: 10.1038/nature11213
21. Gokhale A, Kunder R, Goel A, Sarin R, Moiyadi A, Shenoy A, et al. Distinctive MicroRNA Signature of Medulloblastomas Associated With the WNT Signaling Pathway. *J Cancer Res Ther* (2010) 6:521–9. doi: 10.4103/0973-1482.77072
22. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI Gene Expression and Hybridization Array Data Repository. *Nucleic Acids Res* (2002) 30:207–10. doi: 10.1093/nar/30.1.207
23. Gautier L, Cope L, Bolstad BM, Irizarry RA. Affy - Analysis of Affymetrix GeneChip Data at the Probe Level. *Bioinformatics* (2004) 20(3):307–15. doi: 10.1093/bioinformatics/btg405
24. Durinck S, Spellman PT, Birney E, Huber W. Mapping Identifiers for the Integration of Genomic Datasets With the R/ Bioconductor Package Biomart. *Nat Protoc* (2009) 4:1184–91. doi: 10.1038/nprot.2009.97
25. Frank E, Hall MA, Witten IH. The WEKA Workbench. In: *Morgan Kaufmann, 4th ed.* (2016). p. 553–71. doi: 10.1016/B978-0-12-804291-5.00024-6
26. Salzberg SL. C4.5: Programs for Machine Learning by J. Ross Quinlan. Morgan Kaufmann Publishers, Inc., 1993. *Mach Learn* (1994) 16:235–40. doi: 10.1007/BF00993309
27. Cohen WW. Fast Effective Rule Induction. *Proc Twelfth Int Conf Mach Learn* (1995) 95:115–23. doi: 10.1.1.50.8204
28. Breiman L. Random Forests. *Mach Learn* (2001) 45:5–32. doi: 10.1023/A:1010933404324
29. Platt JC. Fast Training of Support Vector Machines Using Sequential Minimal Optimization. In: *Advances in Kernel Methods - Support Vector Learning. 2008 3rd International Conference on Intelligent System and Knowledge Engineering*. (2008). p. 185–208. doi: 10.1109/ISKE.2008.4731075
30. Keerthi SS, Shevade SK, Bhattacharyya C, Murthy KRK. Improvements to Platt's SMO Algorithm for SVM Classifier Design. *Neural Comput* (2001) 13:637–49. doi: 10.1162/089976601300014493
31. Hastie T, Tibshirani R. Classification by Pairwise Coupling. *Ann Stat* (1998) 26:451–71. doi: 10.1214/aos/1028144844
32. Van Der Maaten L, Hinton G. Visualizing Data Using t-SNE. *J Mach Learn Res* (2008) 9(86):2579–605.
33. Van Der Maaten L. Accelerating t-SNE Using Tree-Based Algorithms. *J Mach Learn Res* (2015) 151:3221–45.
34. Suzuki R, Shimodaira H. Pvclust: An R Package for Assessing the Uncertainty in Hierarchical Clustering. *Bioinformatics* (2006) 22:1540–2. doi: 10.1093/bioinformatics/btl117
35. Pickles JC, Hawkins C, Pietsch T, Jacques TS. CNS Embryonal Tumours: WHO 2016 and Beyond. *Neuropathol Appl Neurobiol* (2018) 44:151–62. doi: 10.1111/nan.12443

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

ORIGINAL RESEARCH
published: 24 June 2021
doi: 10.3389/fonc.2021.688203

Myeloid Sarcoma Type of Acute Promyelocytic Leukemia With a Cryptic Insertion of RARA Into FIP1L1: The Clinical Utility of NGS and Bioinformatic Analyses

Yongren Wang[1,2†], *Yaoyao Rui*[1,2†], *Ying Shen*[3], *Jian Li*[1,2], *Poning Liu*[1,2], *Qin Lu*[1,2] *and Yongjun Fang*[1,2*]

[1] Department of Hematology and Oncology, Children's Hospital of Nanjing Medical University, Nanjing, China, [2] Key Laboratory of Hematology, Nanjing Medical University, Nanjing, China, [3] Department of Gastroenterology, Children's Hospital of Nanjing Medical University, Nanjing, China

OPEN ACCESS

Edited by:
Hua Tan,
University of Texas Health Science Center at Houston, United States

Reviewed by:
Sebastian Dorin Asaftei,
Ospedale Città della Salute e della Scienza, Italy
E. Anders Kolb,
Alfred I. duPont Hospital for Children, United States

***Correspondence:**
Yongjun Fang
fyj322@189.com

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Pediatric Oncology, a section of the journal Frontiers in Oncology

Received: 30 March 2021
Accepted: 03 June 2021
Published: 24 June 2021

Citation:
Wang Y, Rui Y, Shen Y, Li J, Liu P, Lu Q and Fang Y (2021) Myeloid Sarcoma Type of Acute Promyelocytic Leukemia With a Cryptic Insertion of RARA Into FIP1L1: The Clinical Utility of NGS and Bioinformatic Analyses. Front. Oncol. 11:688203. doi: 10.3389/fonc.2021.688203

Background: Acute promyelocytic leukemia (APL) is characterized by the presence of coagulopathy at onset and translocation t (15; 17) (q22; 21), meanwhile, other translocation variants of APL have also been reported. The *FIP1L1–RARA* fusion gene has recently been reported as a novel RARA-associated fusion gene.

Objectives: We report a case of *de novo* myeloid sarcoma (MS) type of APL with *FIP1L1–RARA* found by next-generation sequencing (NGS) that was not detected by conventional analyze analysis for RARA translocations.

Methods: We performed typical morphological, magnetic resonance imaging (MRI), conventional tests for PML–RARA dual-fusion translocation probe, high-through sequencing and NGS. Meanwhile, bioinformatics analyses were done by using public repositories, including ONCOMINE, COSMIC, and GeneMANIA analysis.

Results: A 28-month-old girl with a complex karyotype that includes 46,XX,t(4;17)(q12; q22)[9]/46,idem,del(16)(q22)[3]/45,idem,-x,-4,-9,-15,del(16)(q22),+marl,+mar2,+mar3 [7]/46,xx[3], c.38G>A (p.Gly13Asp) in the *KRAS* gene, and a cryptic insertion of *RARA* gene into the *FIP1L1* gene was diagnosed with APL complicated by the *de novo* MS.

Conclusion: We report a *FIP1L1–RARA* fusion in a child with APL who presented with an extramedullary tumor in the skull without the classic karyotype using NGS, whom we treated with good results. NGS analysis should be considered for APL variant cases. Further experimental studies to the association between the mutation in *KRAS* gene and *FIP1L1–RARA* fusion on the clinical phenotype and progression of APL are needed to identify more effective therapeutic targets for APL.

Keywords: acute promyelocytic leukemia, myeloid sarcoma, *FIP1L1–RARA* fusion, next-generation sequencing, *KRAS* gene, bioinformatic analyses

INTRODUCTION

Acute promyelocytic leukemia (APL) is a subtype of acute myeloid leukemia (AML) that is cytogenetically characterized by the t (15;17) (q24;q21) translocation and gene fusion between the promyelocytic leukemia (*PML*) and the retinoic acid receptor alpha (*RARA*). Additionally, fewer than 5% of APL patients have other fusion gene between *NPM1*, *NUMA1*, *STAT5B*, *BCOR*, *FIP1L1*, *IRF2BP2*, *FNDC3B*, *PRKAR1A*, *OBCF2A*, *GTF2I* and *RARA*, respectively (1). In 2007, the *FIP1L1/RARA* fusion gene was first reported as a novel *RARA*-associated fusion gene in a patient with juvenile myelomonocytic leukemia (JMML) (2). In addition, it was found in patients with APL in 2008 (3) and 2011 (4). Their fusion gene reported was generated rearrange exon 3 of RARA with exon 15 or 13 of FIP1L1, respectively.

Myeloid sarcoma (MS) in APL is a rare condition and is mainly associated with cases of relapse. Herein, we report a primary MS typing case with *FIP1L1/RARA* that was not detected by conventional tests for RARA-associated translocation.

MATERIALS AND METHODS

Patients and DNA Samples

DNA samples extracted from bone marrow of the patient with APL were sent to the Nanjing Key Laboratory of Pediatrics at the Children's Hospital of Nanjing Medical University in Nanjing, China, for genetic analysis. These works were approved by the ethics committee of Children's Hospital of Nanjing Medical University. The patients/participants provided their written informed consent to participate in this study.

Morphology, Flow Cytometry, and FISH Studies

Cytogenetic analysis at diagnosis was carried out according to a standard procedure on a BM sample processed after short-term culture (24 h). G-banded chromosome was determined according to International System for Human Cytogenetic Nomenclature. FISH analysis was performed using commercially available PML/RARα dual color dual fusion DNA probe (Abbott Molecular, Des Plaines, IL).

Molecular Analysis

KRAS mutation, *FIP1L1/RARA* and the *RARA/FIP1L1* fusion transcripts monitoring were performed by NGS. The size and quality of the pre- and post-captured libraries were evaluated using a 2100 BioAnalyzer (Agilent). The HiSeq X PE 2*150 method was used for sequencing, and each sample was sequenced with 15G PF data.

Abbreviations: AML, acute myeloid leukemia; APL, acute promyelocytic leukemia; MS, Myeloid sarcoma; JMML, juvenile myelomonocytic leukemia; WBC, white blood cell; ATO, arsenictrioxide; ATRA, all-trans retinoic acid; CR, complete remission; FISH, fluorescence *in situ* hybridization; NGS, next-generation sequencing, MRI, magnetic resonance imaging, PML, promyelocytic leukemia protein; *RARA*, retinoic acid receptor alpha; COSMIC, Catalog of Somatic Mutations in Cancer; CT, computed tomography; SNP, single nucleotide polymorphisms.

Public Database

The summary of the distribution of different types of mutations for *KRAS* mutations in hematopoietic neoplasm was performed using the Catalog of Somatic Mutations in Cancer (COSMIC, release v92, 27th August 2020) database (http://www.sanger.ac.uk/cosmic/) and visualized in the bar chart. All data were extracted on Dec. 2, 2020. ONCOMINE (https://www.ONCOMINE.org) was used to analyze the messenger RNA levels of *KRAS*, *RARA*, and *FIP1L1* in different cancer tissues compared with that in normal control with the significance was generated using Students t-test. Statistically significant values and fold change: P value <1E−4 and fold change >2. The comparison heatmap and boxplot of *KRAS*, *RARA*, and *FIP1L1* expression value across different cancers was generated.

Statistical Analysis

GraphPad Prism software (version 6, GraphPad Software Inc., La Jolla, CA, USA) and R statistical software (RStudio) were used for statistical analyses. Student's t-test (two-tailed) was used to compare the means between two groups, mRNA expression data are presented as fold change, and a statistically significant difference was considered at the level of $P < 0.05$.

Case Report

The 28-month-old girl was presented with pain in her right upper limb. Initial laboratory evaluation presented a high white blood cell (WBC) level of 20×10^9 with 65.2% of abnormal promyelocyte cells, hemoglobin of 9.5g/dl, and a platelet level of 101×10^9/L. Her coagulation function and thromboela-stogram were normal, while lactate dehydrogenase was 740 U/L. The BM aspirate showed 74.5% promyelocytes that had numerous hypergranularity with Auer bundle; the rate of positive peroxidase staining was 100%, suggestive of APL (**Figure 1A**). The immunophenotype was positive for CD4, CD2, MPO, CD33, CD13, CD11b, CD64, CD15, CD71, CD9 and CD65, and negative for CD34, HLA-DR and TDT. Fluorescence *in situ* hybridization (FISH) analysis with the PML–RARA dual-fusion translocation probe identified no dual fusion signal but the presence of increased signals of RARA gene (73%; **Figure 1B**). Besides, cytogenetics revealed a complex karyotype in 19 metaphase cells with the following formula: 46,XX,t(4;17)(q12;q22)[9]/46,idem,del(16)(q22)[3]/45,idem,-x,-4,-9,-15,del(16)(q22),+marl,+mar2,+mar3[7]/46,xx[3] according to ISCN2016 (**Figure 1C**). Mutational analysis of myelogenous leukemia related genes using a high-through sequencing technique showed a KRAS mutation [c.38G >A; p.Gly13Asp with 11.9% variant allele frequency (VAF)] and 28 single nucleotide polymorphisms (SNP) which considered to be no pathogenic significant related to myeloid hematological tumors. To further characterize, we arranged a NGS strategy for hematopoietic malignancies. Sequencing of this sample confirmed *FIP1L1/RARA* and the *RARA/FIP1L1* fusion transcripts. Besides, the fusion gene *FIP1L1–RARA* was generated between exon 12 of *FIP1L1* and exon 3 of *RARA*, while the fusion gene *RARA–FIP1L1* was generated between intron 2 of RARA and exon 14 of *FIP1L1* (**Figure 1D**). This patient was diagnosed with microgranular variant (M3v) of

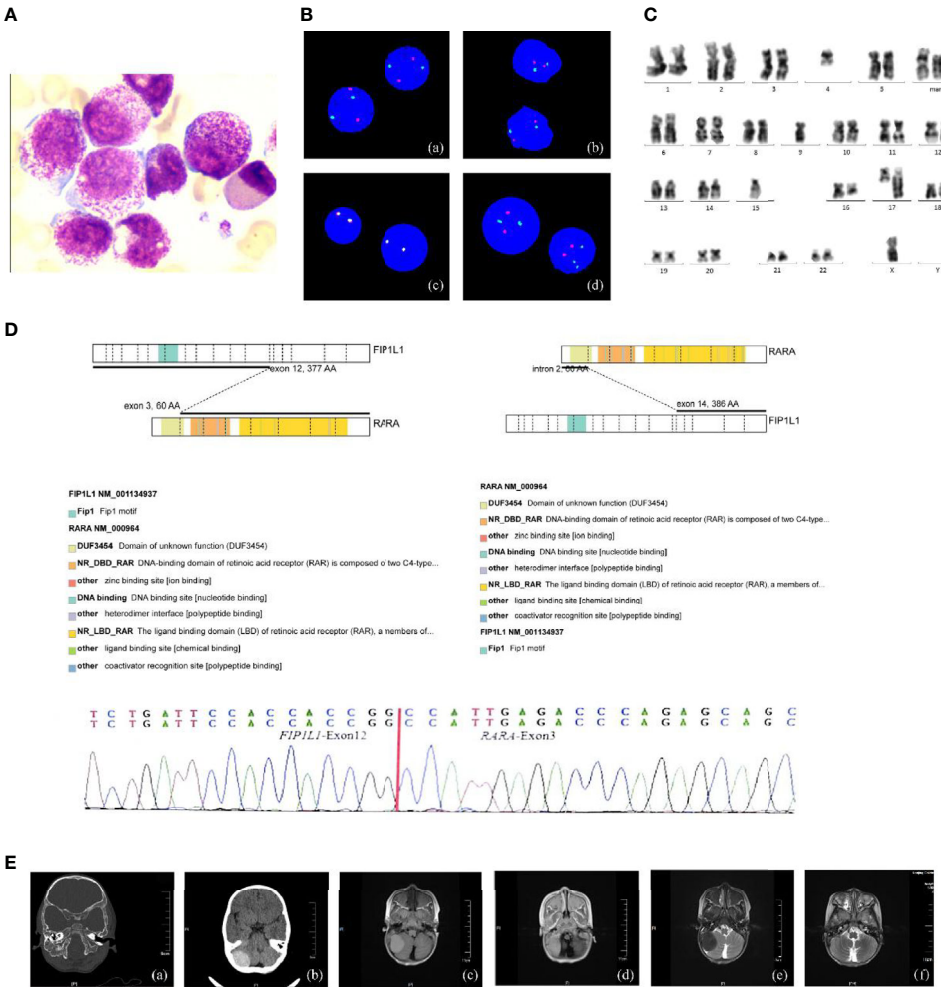


FIGURE 1 | The clinical information of the APL patient. **(A)** Bone marrow morphology at initial diagnosis. ×400; **(B)** FISH using the *PML/RARA* dual-color, dual-fusion translocation probe indicated the absence of the normal *PML/RARA*: (a) nuc ish(D5S23/D5S721,CSF1R)×2[400], (b) nuc ish (D7Z1/D7S486)×2[400], (c) nuc ish (MLL×2) [400], (d) nuc ish (PML×2, RARA×3)[292/400]; **(C)** G-banded karyotype of the patient; **(D)** Sequencing analysis of *FIP1L1–RARA* fusion transcripts atonset. Diagrammatic representation and sequencing information of *FIP1L1–RARA* fusion transcripts. *FIP1L1* transcripts consisting of exon 12 joined to RARA intron 3 (variant form). **(E)** Intracranial tumor: (a, b) CT detects mass with some skull changes in the right posterior fossa, (c–e) MR image: infiltrative lesion involving the right posterior fossa, bilateral mandible, antrum maxillae, skull base, and partial vertebra, (f) complete resolution of the intracranial mass after the treatment.

APL according to bone marrow morphology, immunophenotype, cytogenetics, and the transcriptome sequencing.

What's more, an urgent cerebral computed tomography (CT) scan revealed the presence of high-density shadows in the right posterior fossa with partial skull changes. Then, head and spine MRI showed intracranial mas formation in the right posterior fossa which is considered as MS with extensive infiltrative lesions involving the bilateral mandible, the antrum maxillae, the skull base, and partial vertebra (**Figure 1E**).

After a treatment combining all-trans retinoic acid (ATRA, 20 mg/d, divided twice a day) with DA regimen (daunorubicin [20 mg/d for 5 days] and cytarabine [40 mg/day for 8 days]) as induction chemotherapy, the patient had pain in her eyes on the 14th day of treatment, which was considered to be differentiation syndrome. A second BM smear showed 1.5% blasts, and a second head MR showed that the chloroma had disappeared, indicating that the patient had achieved initial complete remission by day 30. However, FIP1L1/RARA was still positive after the second course. Since November 11, 2020, the patient has received three cycles of ATRA and idarubicin [10 mg/d for 3 days] in the following consolidation treatment. She received continual therapy with ATRA and chemotherapy as previously described and kept leukemia-free after a 5-month follow-up.

Bioinformatics Analysis of Genetic Mutations in the *KRAS*

First, we searched for the mutants of *KRAS* using the COSMIC database (http://cancer.sanger.ac.uk/cosmic): there were mutations in 11 of 1,713 (0.6%) patients in hematopoietic neoplasm, and c.35G >A is the most frequent mutation (**Figure 2A**).

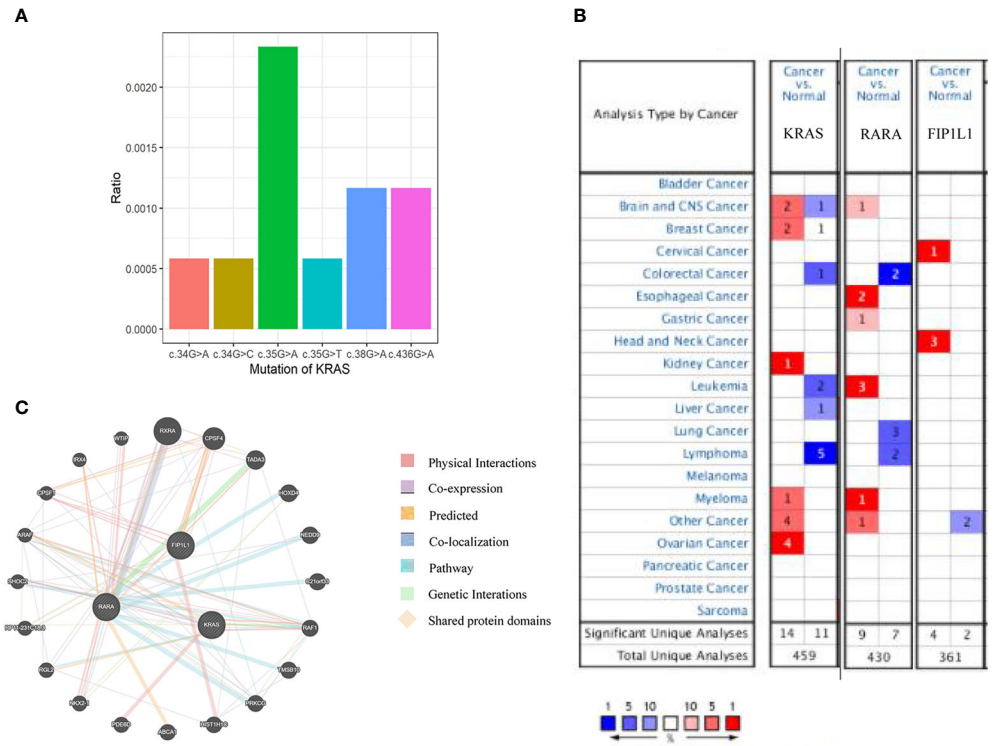


FIGURE 2 | Bioinformatic analyses of *KRAS*, *RARA*, and *FIP1L1* genes. **(A)** An overview of the mutations in *KRAS* in the samples from patients with hematopoietic malignancies, according to the COSMIC database. **(B)** The number of datasets that had mRNA overexpression (red) or down-regulated expression (blue) of the *KRAS*, *RARA*, and *FIP1L1* gene in ONCOMINE. The threshold was designed with the following parameters: p-value of 1E–4 and fold change of 2. **(C)** Network of *KRAS*, *RARA*, and *FIP1L1* and their 20 related genes was analyzed by GeneMANIA.

We compared the mRNA levels of *KRAS*, *RARA*, and *FIP1L1* among samples of different cancers by using ONCOMINE databases (https://www.ONCOMINE.org). For *RARA*, more datasets showed an increased trend in leukemia patients (**Figure 2B**). For detailed analysis in ONCOMINE datasets, we further searched their expression in APL patients. GeneMANIA was used to analyze associations in terms of co-expression, physical interactions, shared protein domains, predicted, pathway and co-localization among *KRAS*, *RARA*, and *FIP1L1* (**Figure 2C**).

DISCUSSION

APL is a subtype of AML and is commonly characterized by the expression of the oncogenic PML–RARα fusion protein (more than 95% of APL cases). The incidence of APL in infants is very rare (median age of pediatric APL is 9–12 years) (5). The typical feature of APL is a life-threatening coagulopathy which can lead the patients to death (6). Early deaths (occur in the first 30 days after diagnosis) were associated with a high WBC count (7). The classical diagnosis of APL is represented by the morphology and the flow-cytometry analysis on bone marrow aspirate and confirmed by FISH probes for cytogenetic translocation t (15; 17) or other molecular biology techniques such as reverse transcription PCR for PML–RARA fusion transcript.

Although PML–RARA was failed to be identified by FISH in our case, her abnormal promyelocytes in morphology and immunophenotype were fully consistent with APL, suggesting that her disease may be caused by other X-RARA fusions. In this situation, we use NGS as a tool to discover her fusion gene called *FIP1L1–RARA*. To date, several other partner genes including *ZBTB16*, *BCOR*, *NPM1*, *NABP1*, *Stat5B*, *PRKAR1A*, etc. have been reported (8). Only three cases with cryptic *FIP1L1–RARA* have been reported, one was a 20-month-old boy with JMML (2), one was a 90 year-old woman diagnosed with APL (3), and the last one was a 77-year old female APL patient (4). Both of the two APL patients were old with no DIC. Their fusion gene was respectively generated between exon 15 or 13 of *FIP1L1* and exon 3 of *RARA*, while the reason for two different phenotypes of leukemia caused by *FIP1L1–RARA* is still unknown. Meanwhile, Kondo (3) put forward the reciprocal *RARA–FIP1L1* has none functional role in leukemogenesis. Genetic involvement in APL has clinical value for the choice of the therapy and evaluating prognosis (9). Thus, it is important to combine NGS with karyotype analysis, FISH, and RT-PCR for accurate diagnosis, especially when RARA rearrangements are failed to be identified by conventional methods.

The mutational spectrum of APL differs from other AML subtypes. The molecular feature of somatic mutations in newly

diagnosed and relapsed APL is defined by frequent alterations of *FLT3*, *NRAS*, *KRAS*, and *ARID1A/B* genes, and the lack of mutations in non-M3 AML genes (e.g. *DNMT3A*, *NPM1*, *IDH1/2*, and *ASXL1*) (10). *PML–RARA* acts as the main driver mutation in each APL exome. TCGA consortium showed that the expected recurrent somatic mutations in patient with APL were almost twice lower than those in patients with other AML subtypes. Additionally, the number of somatic mutations in leukemia patients was three to seven times lower than in patients with solid tumors (11). Previous studies have reported that the numerous genes involved in APL had similar cell functions, but lacked recurrence and consistency. The impairment of these genes on APL is weaker than the interaction of mutated genes with different functionally related categories (12). Next, it was found that RARA gene and KRAS were co-expressed through PPI network analysis by GeneMANIA. However, the impairment of c.38G >A (p.Gly13Asp) in the *KRAS* gene and the co-expressed *FIP1L1–RARA* fusion APL is not completely clear, further studies have to be performed in the future.

MS is a relatively rare disease lacks specificity in morphology. Our patient with lesions in the head lack of tissue biopsy is one of the important limits. Considering her clinical history, the mass by MRI and CT imaging is very likely extending from the right mandibular region to the skull and the posterior cranial fossa. It needs to be clinically differentiated from hamartomas, abscesses, meningioma, primitive neuroectodermal tumors, Langerhans cell histiocytosis, and so on. Although the bony destruction persisted in the first repeat MRI, there was complete resolution of the Intracranial mass. The initial MRI findings (combined with APL and significant radiological improvement after treatment) confirmed the diagnosis of MS. however, as the tissue biopsy was not performed, residual disease could not be completely ruled out. In addition, the clinical features of MS in our case occur in the absence of coagulation abnormalities that may differ from extramedullary diseases in advanced APL.

MS in APL occurs rarely but frequently at the time of relapse, that involves the skin, lymph nodes, and central nervous system. It has been suggested that MS may be associated with a direct effect of ATRA on adhesion molecules, a consequence of the prolonged survival, a high WBC ($>10 \times 10^9$/L) at presentation, and the presence of bcr3 *PML/RARA* fusion transcripts (13). Typically, MS may develop *de novo* or concurrently in AML with 2.5–9.11% occurrence (14). Furthermore, MS in AML was enriched with mutations of the RTK-RAS pathway genes (*KRAS*, *NRAS*, *BRAF*, *PTPN11*, and *CBL*) (15), while *KRAS* and *NRAS* mutations occupy 70% (16). The pathogenesis of extramedullary AML tumor is related to the abnormal cellular adhesion molecules and RAS-MAPK/ERK signaling (17). Presently, there are relatively few data for the prevalence of different mutations and mechanisms of extramedullary APL. What's more, there might be some connection(s) that APL patient with combined *FIP1L1–RARA* and KRAS mutations had a predilection to develop MS.

To the best of our knowledge, there is no consensus on MS treatment (18). Generally, AML-type therapy is effective for *de novo* MS (19). As our patient has APL combined with intracranial MS, we considered the combination of ATRA with conventional chemotherapy (14). In our case, despite the short follow-up, the patient got CR after 1 month course of treatment without recurrence, which emphasized the efficacy of the combination ATRA–chemotherapy therapy on MS. Therapeutic function of ATRA in APL includes activating the gene transcription in myeloid lineage differentiation and degrading the PML–RARα oncoprotein, while arsenictrioxide (ATO) degrades all PML containing molecules and promotes apoptosis. Interestingly, the reported variety of X-*RARA* fusions including *PLZF*, *NuMA*, *NPM*, *STAT5b*, *FIP1L1*, *PRKAR1A*, *ZBTB16*, *OBFC2A*, *TBLR1*, *GTF2I*, *IRF2BP2*, and *FNDC3B* may cause the resistance to ATO treatment due to the lack of ATO binding sites (20).

In conclusion, patients with APL presenting with the fusion gene *FIP1L1/RARA* are rare but have been reported in the literature (4). We found the first case of APL with *FIP1L1/RARA* by using NGS and concurrent MS, while there is no clear treatment guideline so far. The presented patient had achieved complete remission following systemic chemotherapy. Molecular analysis of APL variants is insufficient only through routine analysis, because variant APL has many partners with RARA. We should consider NGS analysis as a conventional method for patients with variant APL. Lastly, further studies are needed to address the cooperation with *FIP1L1–RARA* and *KRAS* in the MS formation of APL.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/supplementary material.

ETHICS STATEMENT

These works were approved by the ethics committee of Children's Hospital of Nanjing Medical University. Written informed consent was obtained from the individual(s), and minor(s)' legal guardian/next of kin, for the publication of any potentially identifiable images or data included in this article. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

Patient Management and Data curation: YW, YR, YS, JL, PL, QL and YF. Data analysis YS and JL. Project administration: YW and YF. Writing—original draft: YW and YR. Writing—review and editing: YF. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by the National Natural Science Foundation of China (81602913, 81670155, 81903383).

REFERENCES

1. Osumi T, Watanabe A, Okamura K, Nakabayashi K, Yoshida M, Tsujimoto SI, et al. Acute Promyelocytic Leukemia With a Cryptic Insertion of RARA Into TBL1XR1. *Genes Chromosomes Cancer* (2019) 58(11):820–3. doi: 10.1002/gcc.22791
2. Buijs A, Bruin M. Fusion of FIP1L1 and RARA as a Result of a Novel T(4;17)(q12;q21) in a Case of Juvenile Myelomonocytic Leukemia. *Leukemia* (2007) 21(5):1104–8. doi: 10.1038/sj.leu.2404596
3. Kondo T, Mori A, Darmanin S, Hashino S, Tanaka J, Asaka M. The Seventh Pathogenic Fusion Gene FIP1L1-RARA Was Isolated From a T(4;17)-Positive Acute Promyelocytic Leukemia. *Haematologica* (2008) 93(9):1414–6. doi: 10.3324/haematol.12854
4. Menezes J, Acquadro F, Perez-Pons de la Villa C, Garcia-Sanchez F, Alvarez S, Cigudosa JC. FIP1L1/RARA With Breakpoint at FIP1L1 Intron 13: A Variant Translocation in Acute Promyelocytic Leukemia. *Haematologica* (2011) 96 (10):1565–6. doi: 10.3324/haematol.2011.047134
5. Testi AM, D'Angio M, Locatelli F, Pession A, Lo Coco F. Acute Promyelocytic Leukemia (Apl): Comparison Between Children and Adults. *Mediterr J Hematol Infect Dis* (2014) 6(1):e2014032. doi: 10.4084/MJHID.2014.032
6. Sanz MA, Montesinos P. Advances in the Management of Coagulopathy in Acute Promyelocytic Leukemia. *Thromb Res* (2020) 191:S63–7. doi: 10.1016/s0049-3848(20)30399-6
7. Sanz MA, Lo Coco F, Martín G, Avvisati G, Rayón C, Barbui T, et al. Definition of Relapse Risk and Role of Nonanthracycline Drugs for Consolidation in Patients With Acute Promyelocytic Leukemia_ A Joint Study of the PETHEMA and GIMEMA Cooperative Groups. *Blood* (2000) 96:1247–53. doi: 10.1182/blood.V96.4.1247
8. Braekeleer ED, Douet-Guilbert N, Braekeleer MD. RARA Fusion Genes in Acute Promyelocytic Leukemia: A Review. *Expert Rev Hematol* (2014) 7 (3):347–57. doi: 10.1586/17474086.2014.903794
9. Zhao J, Liang JW, Xue HL, Shen SH, Chen J, Tang YJ, et al. The Genetics and Clinical Characteristics of Children Morphologically Diagnosed as Acute Promyelocytic Leukemia. *Leukemia* (2019) 33(6):1387–99. doi: 10.1038/s41375-018-0338-z(PMID:30575821
10. Madan V, Shyamsunder P, Han L, Mayakonda A, Nagata Y, Sundaresan J, et al. Comprehensive Mutational Analysis of Primary and Relapse Acute Promyelocytic Leukemia. *Leukemia* (2016) 30(8):1672–81. doi: 10.1038/leu.2016.69
11. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Kinzler KW. Cancer Genome Landscapes. *Science* (2013) 339(6127):1546–58. doi: 10.1126/science.1235122
12. Ibanez M, Carbonell-Caballero J, Garcia-Alonso L, Such E, Jimenez-Almazan J, Vidal E, et al. The Mutational Landscape of Acute Promyelocytic Leukemia Reveals an Interacting Network of Co-Occurrences and Recurrent Mutations. *PloS One* (2016) 11(2):e0148346. doi: 10.1371/journal.pone.0148346
13. de Botton S, Sanz MA, Chevret S, Dombret H, Martin G, Thomas X, et al. Extramedullary Relapse in Acute Promyelocytic Leukemia Treated With All-Trans Retinoic Acid and Chemotherapy. *Leukemia* (2006) 20(1):35–41. doi: 10.1038/sj.leu.2404006
14. Magdy M, Abdel Karim N, Eldessouki I, Gaber O, Rahouma M, Ghareeb M. Myeloid Sarcoma. *Oncol Res Treat* (2019) 42(4):224–9. doi: 10.1159/000497210
15. Choi M, Jeon YK, Sun C-H, Yun H-S, Hong J, Shin D-Y, et al. Rtk-RAS Pathway Mutation Is Enriched in Myeloid Sarcoma. *Blood Cancer J* (2018) 8 (5):43. doi: 10.1038/s41408-018-0083-6
16. Shallis RM, Gale RP, Lazarus HM, Roberts KB, Xu ML, Seropian SE, et al. Myeloid Sarcoma, Chloroma, or Extramedullary Acute Myeloid Leukemia Tumor: A Tale of Misnomers, Controversy and the Unresolved. *Blood Rev* (2020) 42:100773. doi: 10.1016/j.blre.2020.100773
17. Fu JF, Yen TH, Chen Y, Huang YJ, Hsu CL, Liang DC, et al. Involvement of Gpr125 in the Myeloid Sarcoma Formation Induced by Cooperating MLL/AF10(OM-LZ) and Oncogenic KRAS in a Mouse Bone Marrow Transplantation Model. *Int J Cancer* (2013) 133(8):1792–802. doi: 10.1002/ijc.28195
18. Pileri SA, Ascani S, Cox MC, Campidelli C, Bacci F, Piccioli M, et al. Myeloid Sarcoma: Clinico-Pathologic, Phenotypic and Cytogenetic Analysis of 92 Adult Patients. *Leukemia* (2007) 21(2):340–50. doi: 10.1038/sj.leu.2404491
19. Yamashita T, Nishijima A, Noguchi Y, Narukawa K, Oshikawa G, Takano H. Acute Promyelocytic Leukemia Presenting as Recurrent Spinal Myeloid Sarcomas 3 Years Before Developing Leukemia: A Case Report With Review of Literature. *Clin Case Rep* (2019) 7(2):316–21. doi: 10.1002/ccr3.1991
20. Noguera NI, Catalano G, Banella C, Divona M, Faraoni I, Ottone T, et al. Acute Promyelocytic Leukemia: Update on the Mechanisms of Leukemogenesis, Resistance and on Innovative Treatment Strategies. *Cancers* (2019) 11(10):1591. doi: 10.3390/cancers11101591

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

ORIGINAL RESEARCH
published: 14 July 2021
doi: 10.3389/fonc.2021.653863

A Deep-Learning Model With the Attention Mechanism Could Rigorously Predict Survivals in Neuroblastoma

Chenzhao Feng[1‡], Tianyu Xiang[2,3‡], Zixuan Yi[4‡], Xinyao Meng[1], Xufeng Chu[5], Guiyang Huang[5], Xiang Zhao[1], Feng Chen[6], Bo Xiong[5*] and Jiexiong Feng[1*†]*

[1] Department of Pediatric Surgery, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, [2] Department of Control Science and Engineering, College of Electronics and Information Engineering, Tongji University, Shanghai, China, [3] State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, China, [4] School of Mathematics and Statistics, College of Arts and Sciences, Wuhan University, Wuhan, China, [5] Department of Forensic Medicine, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, [6] Department of Pediatric Surgery, Fujian Medical University Union Hospital, Fuzhou, China

Background: Neuroblastoma is one of the most devastating forms of childhood cancer. Despite large amounts of attempts in precise survival prediction in neuroblastoma, the prediction efficacy remains to be improved.

Methods: Here, we applied a deep-learning (DL) model with the attention mechanism to predict survivals in neuroblastoma. We utilized 2 groups of features separated from 172 genes, to train 2 deep neural networks and combined them by the attention mechanism.

Results: This classifier could accurately predict survivals, with areas under the curve of receiver operating characteristic (ROC) curves and time-dependent ROC reaching 0.968 and 0.974 in the training set respectively. The accuracy of the model was further confirmed in a validation cohort. Importantly, the two feature groups were mapped to two groups of patients, which were prognostic in Kaplan-Meier curves. Biological analyses showed that they exhibited diverse molecular backgrounds which could be linked to the prognosis of the patients.

Conclusions: In this study, we applied artificial intelligence methods to improve the accuracy of neuroblastoma survival prediction based on gene expression and provide explanations for better understanding of the molecular mechanisms underlying neuroblastoma.

Keywords: neuroblastoma, survival, deep-learning (DL), individual therapy, transcriptome

OPEN ACCESS

Edited by:
Jing He,
Guangzhou Medical University, China

Reviewed by:
Xinhui Du,
Henan Provincial Cancer Hospital, China
Jinfa Tou,
Zhejiang University, China

***Correspondence:**
Jiexiong Feng
fengjiexiong@126.com
Bo Xiong
bxiong@hust.edu.cn
Feng Chen
cfeng3000@163.com

†ORCID:
Jiexiong Feng
orcid.org/0000-0001-9596-2330

‡These authors have contributed equally to this work

Specialty section:
This article was submitted to Pediatric Oncology, a section of the journal Frontiers in Oncology

Received: *15 January 2021*
Accepted: *24 June 2021*
Published: *14 July 2021*

Citation:
Feng C, Xiang T, Yi Z, Meng X, Chu X, Huang G, Zhao X, Chen F, Xiong B and Feng J (2021) A Deep-Learning Model With the Attention Mechanism Could Rigorously Predict Survivals in Neuroblastoma. Front. Oncol. 11:653863. doi: 10.3389/fonc.2021.653863

INTRODUCTION

Neuroblastoma, arising from the developing sympathetic nervous system, is the most common form of malignancy in children (1). Although diverse treatments have been developed for different stages of neuroblastoma, survival rates only improved in low and intermediate risk patients (2, 3). Whole genome sequencing and RNA sequencing (RNA-seq) delineated the genomic and transcriptomic traits of neuroblastoma, in which *MYCN* amplification, *ALK* mutations, *PHOX2B* mutations, *TERT* rearrangements, abnormally expressed microRNAs (miRNA) such as Mir17-92a, etc., occur mostly (4–6). Utilizing these data, a number of previous studies attempted to quantitatively predict outcomes for neuroblastoma patients. For instance, chromosomal gain or loss status were used to construct a cox regression model in one attempt, whereas most studies implemented gene expression data into multivariable score models (7, 8).

In recent years, machine-learning (ML) has been widely applied in medical sciences, especially in radiography, healthcare monitoring and genomics (9–12). ML was adopted to predict the outcomes and survival time by different approaches, such as Artificial Neural Network, Supported Vector Machine, Decision Tree and so on in many types of cancer (13–15), a vast majority of which outperformed the traditional cox regression models.

Deep-learning (DL) is a subdiscipline of ML that allows computers to transform raw data through multiple levels of representations. DL-based image detection has been widely studied in the diagnosis of diabetes and cancers (16, 17). In genomics, a multilayer perceptron could predict survival in an unsupervised or supervised way and was extended in lung cancer and hepatocellular carcinoma (18–20). DL-models were also utilized to predict stages and clinical outcomes in neuroblastoma (21, 22). However, reports examining the accuracy of DL-model with survival time are lacking.

Here, we developed a DL-based model to predict outcomes using gene expression matrices. First, 172 features were selected by the chi-square test between gene expression levels and patient survival outcomes in the training cohort. K-means clustering method was used to divide these gene features into two groups. A two-layer neural network decoder was then used to predict survival probabilities and status. F-score, accuracy, sensitivity and specificity were calculated to demonstrate that our model could precisely classify patients. To examine the robustness of our approach, we applied the same procedure in the validation cohort. Indeed, the area under the curve (AUC) of our model was 0.974 in the 5-year-survival receiver operating characteristic (ROC) curve, outperforming existing prognostic models. Furthermore, we partitioned the patients into two subgroups according to their feature expression levels. These two subgroups diverged in survival by log-rank test in Kaplan-Meier (KM) curve with $p < 0.001$. Gene Ontology (GO) enrichment analysis showed that the gene feature group 1 was enriched in the JAK-STAT pathway, while genes involved in bone morphogenesis were enriched in group 2. Therefore, this DL-based approach could rigorously predict neuroblastoma survivals and shine lights over the molecular mechanisms underlying neuroblastoma.

MATERIALS AND METHODS

Data Acquisition

A total of 721 microarray samples, including two datasets named GSE49710 and E-MTAB-8248 (for short, GSE49710 and EMTAB), both detected on Agilent-020382 Human Custom Microarray 44k, were retrieved from NCBI Gene Expression Omnibus and ArrayExpress. The GSE49710 cohort, part of the SEQC project, portrayed Chinese neuroblastoma atlas, while EMTAB displayed German characteristics. Gene expression matrices accompanying clinical information were downloaded directly for the following analyses. F-score, accuracy, sensitivity and specificity were calculated on the whole GSE49710 cohort. Besides, GSE49711, the RNA-seq result of the same samples from GSE49710, was also fetched for lncRNA-related analysis.

Data Preprocessing

To reduce the biases between the two datasets, we normalized the expression levels by equation (1) since the data should be better limited in 0 to 1 in the neural network.

$$f_i\prime = \frac{f_i - \min(f_i)}{\max(f_i) - \min(f_i)} \quad (1)$$

f_i here indicates the expression of each RNA. $f_i\prime$ is designated for the transformed level.

Feature Selection

After data normalizing, significant gene expression features were selected by chi-square test which is implemented by 'chi2' function in the python package *sklearn* (https://scikit-learn.org/). Genes whose FDR in the chi-square test was less than 0.05 were filtered. Following this principle, only 172 differentially expressed genes were chosen.

Another common feature selection approach, the Principal Component Analysis (PCA), was used to transfer gene expression matrix into principal components. The cox proportional regression was used to filter the components. We then compared the results of the PCA method with the chi-square method.

Feature Classification

After feature selection, a classifier was built to classify genes into different subgroups. Genes with similar biological functions were clustered into the same group. The K-means model in *sklearn* divided the selected genes into two clusters.

Model Construction

Then a supervised classification model based on deep neural networks was built. Both of two feature groups were used as inputs of our classification. The output of this neural network was a patient's probability which ranged between 0 and 1. 0 would indicate that the patient is likely to be alive and 1 would indicate that the patient would probably be dead.

The structure of this classifier can be seen in **Figure 1**. It consisted of two parts, the encoder and the decoder. For the encoder part, we encoded two different groups of features into

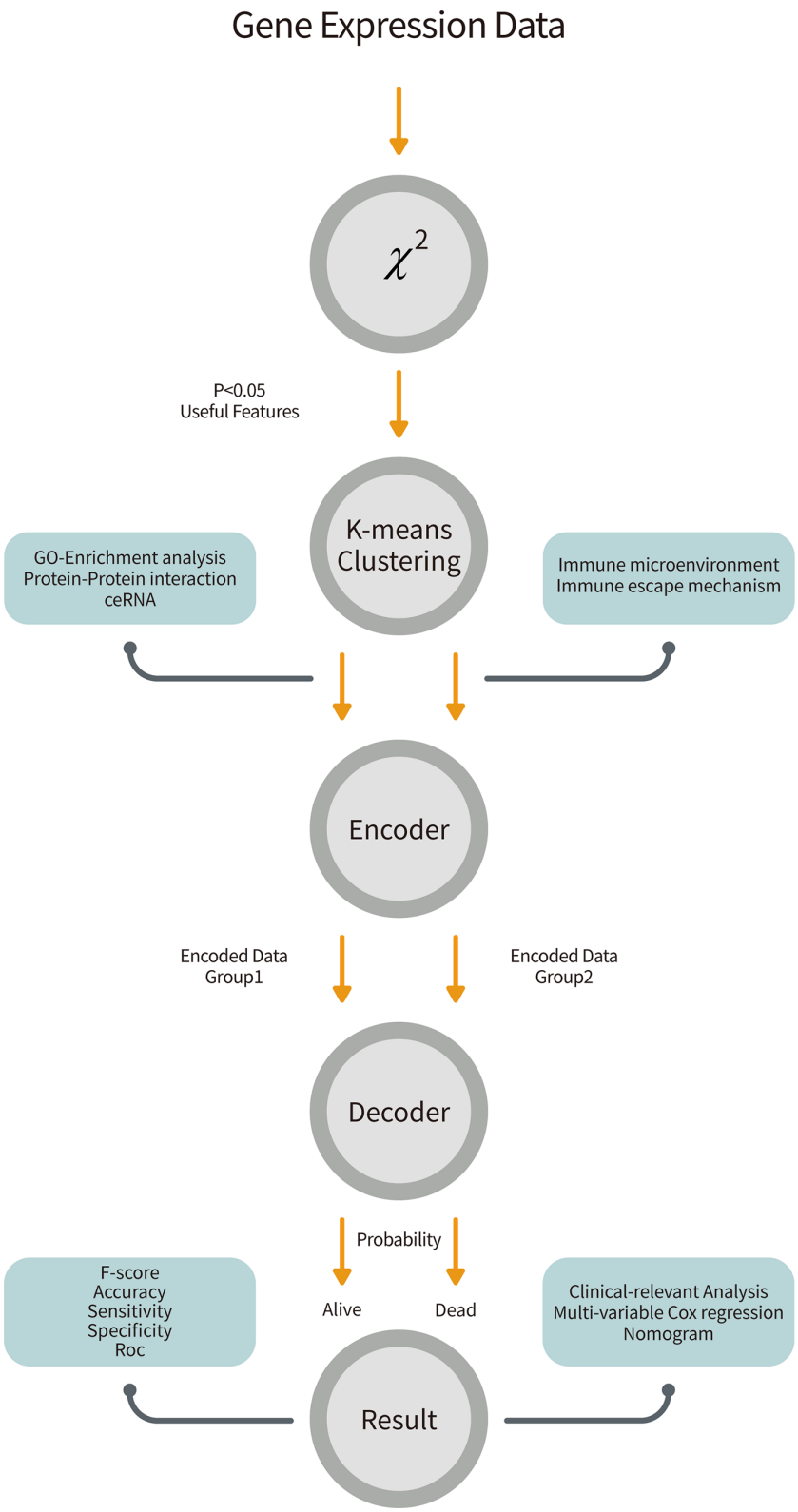


FIGURE 1 | The overall workflow of our pipeline. Gene expression data from GSE49710 was retrieved and performed a chi-square test to filter 172 features. The K-means clustering method partitioned patients and genes into two groups. We trained two neural networks for two groups of features and combined them by the attention mechanism to predict survivals. Further, we analyzed biological effects between two groups and did clinical-relevant analysis.

two 10 dimensional features by two different two-layer networks.

$$g_1' = f(w_{12}f(w_{11}f_{g1} + b_{11}) + b_{12}) \quad (2)$$

$$g_2' = f(w_{22}f(w_{21}f_{g2} + b_{21}) + b_{22}) \quad (3)$$

$g_1'g_2'$ are encoded features. w_{ij} are weights of the networks. f_{g1} and f_{g2} are transformed expressions using formula (1) and b_{ij} are biases of the networks. The function f indicates the activation function which is a nonlinear part of the encoder. Here, we used the ReLU as this nonlinear function:

$$f(x) = \begin{cases} x\ x > 0 \\ 0\ x \le 0 \end{cases} \quad (4)$$

To combine these two different encoded features together, we applied the attention mechanism to this model (23). With the difference of simply concatenating different features, this attention mechanism can learn the relationship between them.

$$G = \left[sig\left(g_2'\right) * g_1', sig\left(g_1'\right) * g_2'\right] \quad (5)$$

This step is illustrated in equation 5. G refers to combined features. sig is a nonlinear function.

$$sig(x) = \frac{e^{-x}}{1 + e^{-x}} \quad (6)$$

After seizing combined features, we input them into the decoder part. The decoder part is also a two-layer neural network.

$$y' = sig(w_{32}f(w_{31}G + b_{31}) + b_{32}) \quad (7)$$

y' is the output of this classifier.

To train this network, we defined the loss function as equation 8.

$$L = \begin{cases} -\alpha(1 - y')^{\gamma} log y'\ y = 1 \\ -(1 - \alpha)y'^{\ \gamma} \log(1 - y')\ y = 0 \end{cases} \quad (8)$$

α and γ are parameters. In this system we set α to 0.2 and γ to 2. y is the true label of each patient. The patients of each group were uneven, so we used Equation 8, which is called the focal loss and was designed to solve this problem instead of cross entropy loss function (24). To avoid over-fitting, we drop 20% neurons in each layer by the dropout method. GSE49710 was chosen to train the network and EMTAB was to validate our model. For GSE49710, 70% of samples were used to train and the rest were to test.

All of the algorithms mentioned in this subsection were realized by *tensorflow* 2.2 (https://tensorflow.google.cn/). To optimize this neural network, we applied Adam optimizer and set the learning rate to 0.01 (25). The weights of networks were initialized by *glorot* uniform distribution (26).

Model Appraisal

To further evaluate our model, we calculated the accuracy, sensitivity, specificity as well as F-score of our model in two cohorts (27).

$$Accuracy = \frac{TP + TN}{Total}$$

$$Sensitivity = \frac{TP}{TP + FN}$$

$$Specificity = \frac{TN}{TN + FP}$$

$$F1 = 2 * \frac{Precision * Recall}{Precision + Recall}$$

ROC curves and AUCs were estimated by *pROC* package in R to assess the performance of the classifier (28). The time-dependent ROC (tROC) curve and its AUC were estimated by *survivalROC* R package to introduce survival time into our classifier (29). The curves were plotted by *ggplot2* R package (30). Besides, traditional cox regression models, devised by Zhong et al. (31) and De Preter et al. (32), were compared with ours.

Patients Clustering

In order to correspond patients to those two groups of gene features, we performed K-means clustering on patients. *ConsensusClusterPlus* was used to determine the best k with parameters (cluster algorithmml: km, distance: Euclidean, replicate time: 1000) (33). The CDF plot and the consensus matrix instructed us to cluster patients into 2 groups. A heatmap showing expression levels of features across samples were made by *complexHeatmap* R package (34). Survival rates between these two subgroups were measured by the log-rank test in KM curves. R packages *survival* and *survminer* were used to fit KM equation and plot the curves (35, 36).

Clinical-Relevance Analyses

In order to test whether our model was independent of other clinical factors and beneficial for clinicians to identify patients' conditions, we first applied univariable cox regression on the age, *MYCN* status, gender, tumor stage, INSS-Risk and our probability score. A multivariable cox regression determined whether a covariate involved was decisive. Forest plots were plotted by *forestplot* package (https://CRAN.R-project.org/package=forestplot). Decision curve analysis was done by ggDCA (https://cran.r-project.org/web/packages/ggDCA/index.html). The construction and plot of the nomogram which can help clinicians to predict survival were done by *rms* (https://CRAN.R-project.org/package=rms) and *regplot* (https://CRAN.R-project.org/package=regplot) R package. Finally, an alluvial diagram was used to visualize the characteristics and disease progressions of each patient. This was achieved by the *ggalluvial* R package (37).

Biological Function Prediction

All biological analyses were done on GSE49710. We executed GO enrichment analysis on the two groups of features by *clusterprofiler* R package respectively (38). Gene Set Enrichment Analysis (GSEA) was done by *GSEA* software (Broad Institute, Inc., version 4.0.3)

with gene set 'c5.all.v7.1.symbols.gmt' and default parameters. *String* (https://string-db.org/) was used to identify protein-protein interactions (PPI) between the 172 features and *Cytoscape* software was used to visualize the interaction networks. *CytoHubba*, a module inside *Cytoscape*, was carried out to identify the hub genes with 12 algorithms (39, 40). For lncRNAs searching, we extracted 250 lncRNA expressions by sorting lncRNA names in *gencode.v34.long_noncoding_RNAs.gtf*, a collection of known lncRNAs downloaded from GENCODE (https://www.gencodegenes.org/). Only lncRNAs that owned a standard error > 0.2 could be enrolled in the correlation test between the 18 mRNAs. The cut-off values were: $p < 0.05$ and |coefficient|>0.5. *TarBase v.8* and *LncBase Predicted v.2* were used to query for mRNA-miRNA and lncRNA-miRNA pairs respectively (41, 42). TarBase v.8 gave 78 *CCNB1*-binding miRNAs with filters (Species: Homo Sapiens, Method Type: High-throughput, Regulation Type: DOWN, Validation Type: Direct). LncBase Predicted v.2 provided predicted lncRNA-miRNA pairs with cut-off 0.7. Finally, a competing endogenous RNA (ceRNA) network was constructed using *Cytoscape*.

Immune Microenvironment Estimation

Inferred abundances of immune cells and normal tissue cells were calculated by single-sample gene set enrichment analysis (ssGSEA) using *GSVA* R package (43). Gene sets, also known as the markers of each cell, were collected by Charoentong et al. (44). Univariable cox proportional regression tests were exerted on all cells to reveal prognostic immune cells.

Statistical Analysis

For categorical and continuous data with normal distribution, we applied chi-square tests and student t tests to distinguish the differences between groups. When continuous data was not normal distributed, Wilcoxon sum rank tests and ANOVA were utilized. The Pearson correlation test was used to find linear connections between two groups of observations. A p-value<0.05 was considered statistically significant except for emphasis. To account for multiple-testing, the p-values were adjusted using the Benjamini-Hochberg FDR correction. All statistical analyses were two-tailed and done by Python (Python Software Foundation, version 3.8.2) and R (R Foundation, version 3.7.0).

RESULTS

Neuroblastoma Genomic Atlas Was Depicted by 172 Features

The overall workflow is shown in **Figure 1**. After implementing the chi-square test into each feature and the survival in GSE49710, 172 features were selected (**Figure 2A**). The consensus clustering method determined the best k value as 2 to partition the features based on the expression matrix (**Supplementary Figure 1**). After that, K-means method was used to cluster the features and patients into 2 groups. Fifty genes were the markers of subgroup 1 of 336 patients (for short, S1), whereas the other 122 genes were the markers of subgroup 2 of 162 patients(S2). Basic characteristics were distributed diversely between the two subgroups except for gender (**Supplementary Table 1**). It is noteworthy that no *MYCN* amplification was detected in S1 and 92 were detected in S2 in the GSE49710 cohort while only 1 such case was detected in S1 and 45 in S2 in the EMTAB cohort, suggesting that these 172 features and the corresponding subgroups were *MYCN*-relevant (chi-square test $p < 0.001$, **Supplementary Table 1**). Also, these subgroups exhibited significant differences in overall survival and event-free survival (both log-rank test $p < 0.001$, **Figures 2B, C** and **Supplementary Figure 2**).

To understand the potential biological functions of the genes in each group, we performed GO-enrichment analyses. Notably, many features in group 1 (F1) are related to the JAK-STAT signaling while features in group 2 (F2) aggregated in the cell migration, bone morphogenesis and ubiquitin-protein transferase activities (**Figures 2D, E**). The JAK-STAT pathway promotes tumor cell proliferation, invasion and immunosuppression through a membrane-nucleus cascade (45). A Previous study has shown that the JAK1/2 inhibitor, AZD1480, could abate neuroblastoma tumor cells growth and extend survivals, suggesting that S1 patients not only maintained better survivals with neuroblastoma, but also might potentially respond to drugs such as AZD1480 to recover (46). Next, the GSEA analysis revealed that S1 showed a higher level of metabolism compared to S2 and S2 developed an intensive immune response (**Supplementary Figure 3**). This might be attributed to the mild symptoms in S1 where patients kept a normal or slightly elevated metabolism. However, accompanying the progression of tumors in S2, the patients started a fierce immune reaction and finally exhausted. These findings suggest that the subgrouping method could help to understand the molecular pathology underlying the differences in prognoses of neuroblastoma patients.

The Neural Network Model Manifested Great Performance in Classifying Neuroblastoma

An encoder-decoder model was then trained on the GSE49710 dataset to predict survivals (**Figure 1**). Since F1 and F2 contributed unequally to the body responses and outcomes, two neural networks were created for them separately in the encoder. In this encoder, a widely used activation function, the ReLU function in the hidden layer; and a binary classification function, the sigmoid function (or say logistic function), in the output layer were employed. The attention mechanism, inspired by human physiology that people would only concentrate on tasks at hand to improve the efficacy of the encoder-decoder framework with rich information, was used to combine the two encoder parts into the decoder (47). The sigmoid function was also used in the final layer, which outputted survival probabilities. If the probability is less than 0.5, we predicted this patient as alive and vice versa (**Supplementary Table 2**).

To assess the prediction quality of the overall survival status, we calculated the accuracy, sensitivity, specificity and F-score of our model, which reflected the proportion of correct predictions

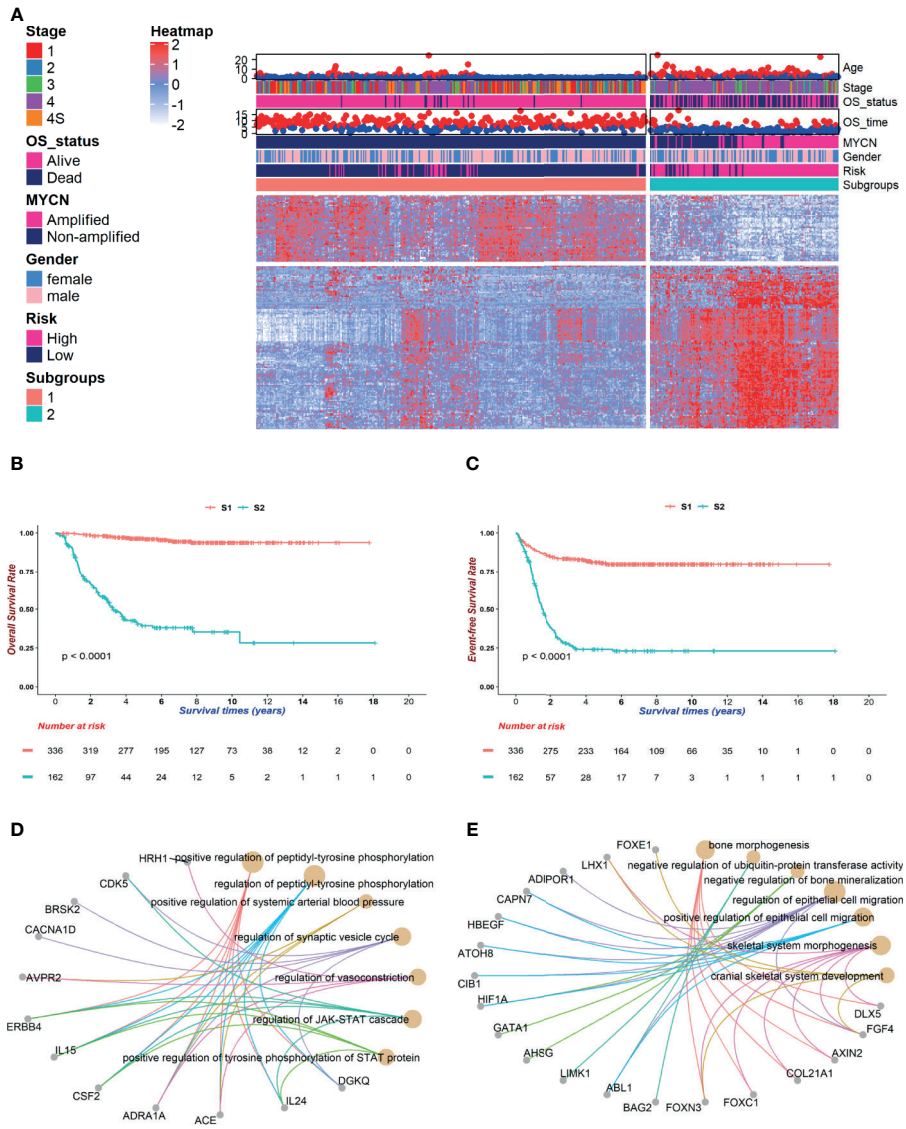


FIGURE 2 | The genomic atlas of neuroblastoma was characterized by 172 genes. After implementing the chi-square test between gene expressions and survival status, a total of 172 genes were selected for following investigations. Patients were clustered into two groups, named subgroup 1 (S1) and subgroup 2 (S2), which owned 50 and 122 markers respectively. **(A)** The heatmap portraited the neuroblastoma genomic landscape in GSE49710. Gene expressions were normalized among samples. The higher expressions reached red while lower reached white. Corresponding clinical information, including age, stage, vital status, survival time, *MYCN* amplification, gender, INSS risk and subgroups, was attached on the top of the heatmap. **(B, C)** Kaplan-Meier (KM) curves showed distinct survivals between S1 (coralline line) and S2 (atroceruleous line) in overall survival (OS) **(B)** and event-free survival (EFS) **(C)** (log rank test $p < 0.001$ for each). **(D, E)** GO-enrichment analysis for feature 1 (F1) and feature 2 (F2) showed that S1 was up-regulated in the JAK-STAT pathway **(D)** and S2 was up-regulated in bone morphogenesis **(E)**.

in all samples, true positives in all positives, true negatives in all negatives and the harmonic mean of precision and recall. In the training set, the accuracy (0.918), sensitivity (0.913) and specificity (0.944) were all greater than 0.9, suggesting that it could efficiently forecast whether a patient would be alive or dead using 172 features. Moreover, a descent efficacy was achieved in the test set (accuracy: 0.852, sensitivity: 0.911, specificity: 0.605), however, F1 score was slightly higher (GSE49710: 0.881, EMTAB: 0.886), indicating that the model was suitable for cohorts of various genetic background. The ROC and tROC curves were then generated which further demonstrate the eminence of the neural network (**Figures 3A–C** and **Table 1**). The AUCs of the training set and the test set achieved 0.968 and 0.891 respectively (**Figure 3A**), alluding the robustness of our neural network prediction model. Adding survival time into ROC curves, we found that the 5-year-survival AUCs could be boosted to 0.974 in GSE49710 (**Figure 3B**) and 0.896 in EMTAB (**Figure 3C**), which validated that our neural network could classify patients with high precision. Twenty times of 10-fold cross validation showed the stability and robustness of our neural

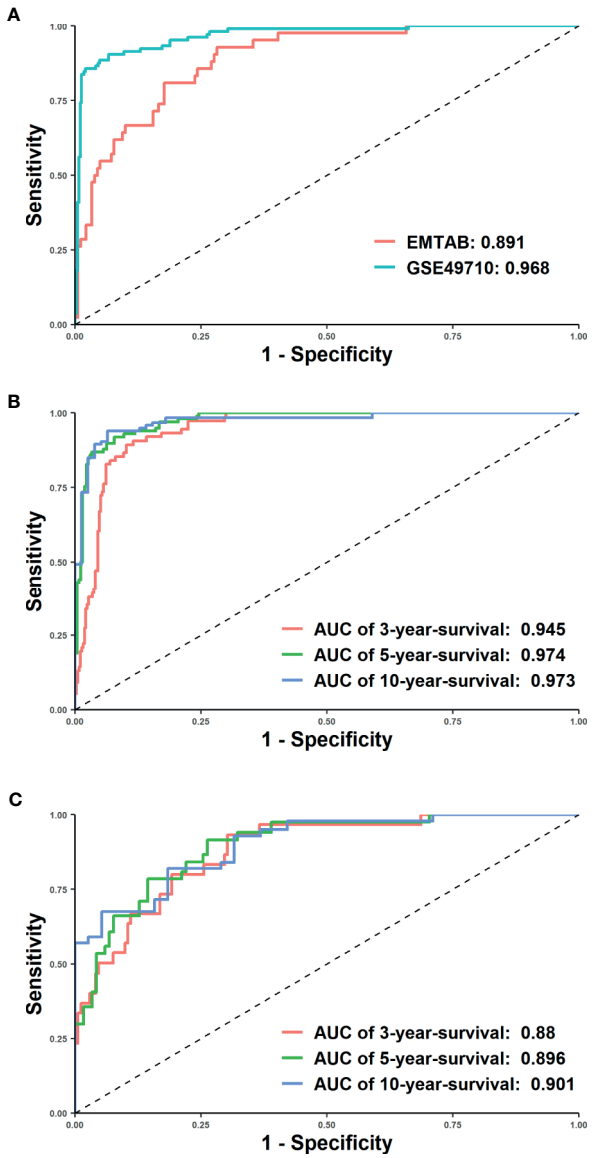


FIGURE 3 | Receiver operating characteristic (ROC) curves demonstrated the superiority of our model. X-axis represented false positive rate (1-specificity) and y axis represented true positive rate. **(A)** The ROC curves of the EMTAB (coralline line) and GSE46960 (atroceruleous line) cohort. **(B, C)** 3-year, 5-year and 10-year time-dependent ROC curves of the GSE46960 **(B)** and EMTAB **(C)** cohort.

network architecture, suggesting that our attempts of the attention mechanism would be extended into more datasets.

Performance Comparison With Alternative Methods

Next, we compared the performance of alternative methods which varied in either feature selection or model construction with our model. To demonstrate that our feature selection was more closely related to prognoses, a broadly used dimension reduction and feature selection method, PCA, was utilized to select features on GSE49710. Kaiser-Harris Criterion suggests that those principal components whose eigenvalue were more than 1 would be retained. In our study, all components had an eigenvalue greater than 1. Variances explained in each component were similar (**Supplementary Figure 4**). The top 200 principal components were chosen with cumulative variance percent at 89.532% for further analyses. Since the survival data has not been utilized, a univariable cox regression model was implemented to principal components, resulting in 17 components being selected with $p < 0.05$. The Consensus cluster determined the best k=value as 3 using K-means clustering (**Supplementary Figure 5**). However, the 3 subgroups were not significantly different in OS (**Figure 4A**, log-rank test $p = 0.088$) but in EFS (**Figure 4B**, log-rank test $p = 0.029$), which indicated that our chi-square-based feature selection could highlight the hub genes in neuroblastoma and partitioned patients into high and low risk groups.

We then compared the performance of our model with several existing models. In order to reduce the biases between the cohorts used in this study and in the literatures, the expression data was normalized before calculating risk scores. We selected two previously published models as well as the 'gold standard', *MYCN* status, and then performed survival prediction in the GSE49710 and EMTAB cohorts (**Supplementary Table 2**) (31, 32). Our DL-model generated the highest AUC in 5-year tROC curves in both GSE49710 and EMTAB cohorts (**Figures 4C, D**), indicating that our model outperformed the existing model in survival prediction.

DL-Model Probabilities Were of Clinical Significance

To test whether the DL-based prediction model is widely useful for patients of various background conditions, we implemented the univariable cox regressions among the age, *MYCN* status, gender, diagnostic stage, risk and our output probability. All variables were converted to binaries in this test (**Supplementary Figure 6**). Only gender failed in this test as $p > 0.05$, which was discarded in the multivariable regression. In the multivariable cox model, the probability risk was still significant (**Figure 5A**), indicating that our DL-based model had a broad prognostic ability regardless of clinical covariates.

We further used the decision curve analysis (DCA) to evaluate the net benefit of different models (48). We constructed 3 models: only DL-probabilities, only clinical covariates in multivariable cox analysis as well as a combined model. The combined model achieved the highest net benefit no matter how risk threshold was set (**Supplementary Figure 7**). The data implied that combining DL-model and clinical information could be profitable for clinicians to diagnose and to predict survivals. Therefore, we build up a nomogram which could help clinicians to predict the potential outcomes of the patients beforehand the medical treatments (**Figure 5B** and **Supplementary Table 3**). A *C*-index 0.889 of the nomogram along with the calibration curve predicted 5-year-survival, demonstrated that this scoring system would be handy and practical in the first-line diagnosis (**Figure 5C**).

TABLE 1 | AUCs of ROCs, 3-year-ROC, 5-year-ROC and 10-year-ROC with 20 times 10-fold cross validation.

Dataset	AUC	AUC of 3-year-ROC	AUC of 5-year-ROC	AUC of 10-year-ROC
Train	0.996 ± 0.009	0.963 ± 0.009	0.991 ± 0.006	0.993 ± 0.007
Test	0.878 ± 0.024	0.879 ± 0.028	0.907 ± 0.021	0.904 ± 0.030
Validation	0.862 ± 0.088	0.867 ± 0.076	0.895 ± 0.066	0.910 ± 0.095
EMTAB	0.865 ± 0.017	0.838 ± 0.022	0.863 ± 0.021	0.853 ± 0.032

The GSE49710 cohort was split at 7:3 into train and test cohort to perform 20 times 10-fold cross validations. In each calculation, 10% of the train cohort was randomly chosen into the validation group. The best model with the most AUC was further validated in EMTAB cohort. The probabilities of patients from the output layer were used in time-dependent ROC analyses. If the probabilities were less than 0.5, we predicted corresponding patients would be alive and if were greater than 0.5, they would be dead. These binary predictions would be compared with true labels in ROC analyses. All AUCs were expressed as mean ± standard deviation.

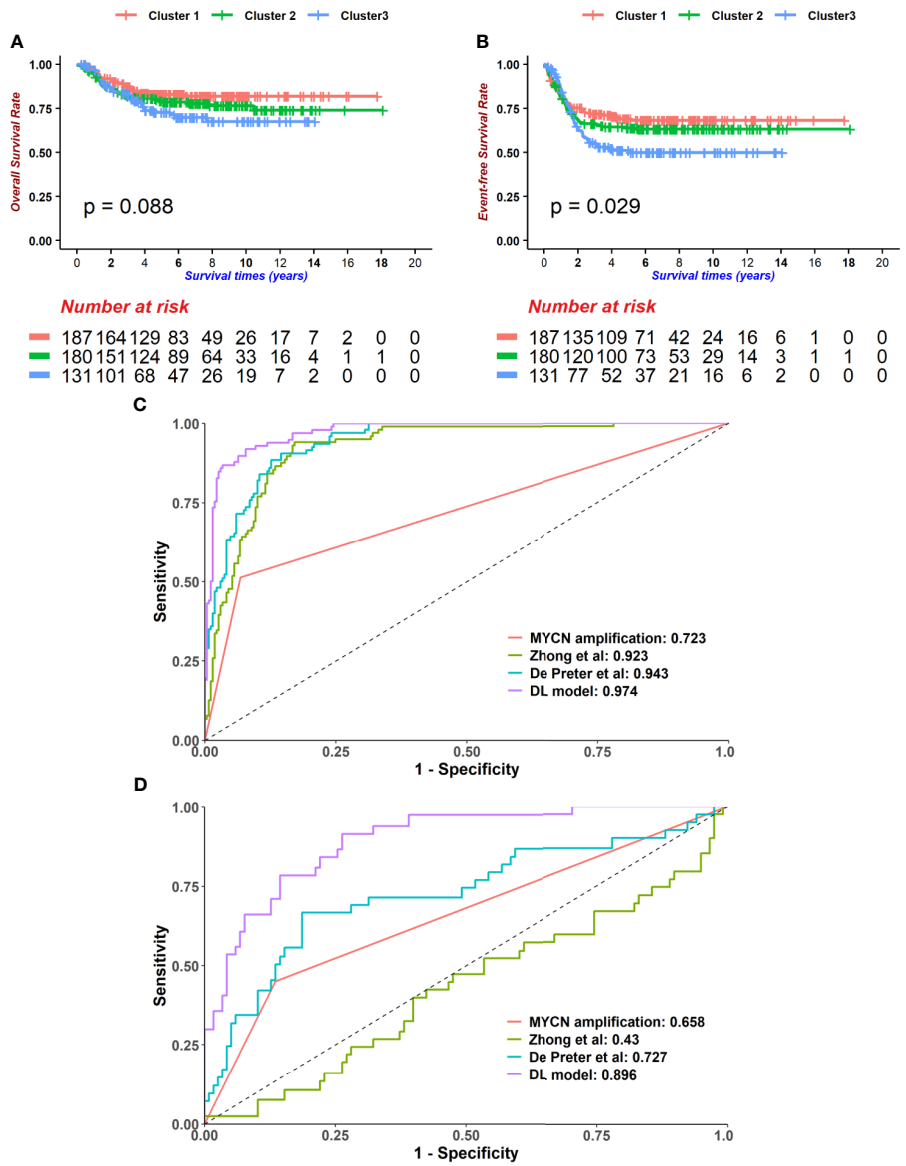


FIGURE 4 | The Deep-learning-model (DL-model) outperformed alternatives in two aspects. **(A, B)** We employed the Principal component analysis (PCA) method to cut features down to 200. Using 200 PCA dimensions, we distributed patients into 3 groups, which was determined by the consensus clustering method. These groups did not show prognostic value in overall survival (OS) (**A**, log-rank test $p = 0.088$, n = 498) but event-free survival (EFS) (**B**, log-rank test $p = 0.029$, n = 498). **(C, D)** Compare with other models (*MYCN* status, 4-gene-signature by Zhong et al. and 42-gene-signature by De Preter et al.), our DL-model received the highest AUCs of 5-year-survival ROC curves in GSE49710 **(C)** and EMTAB **(D)** cohorts.

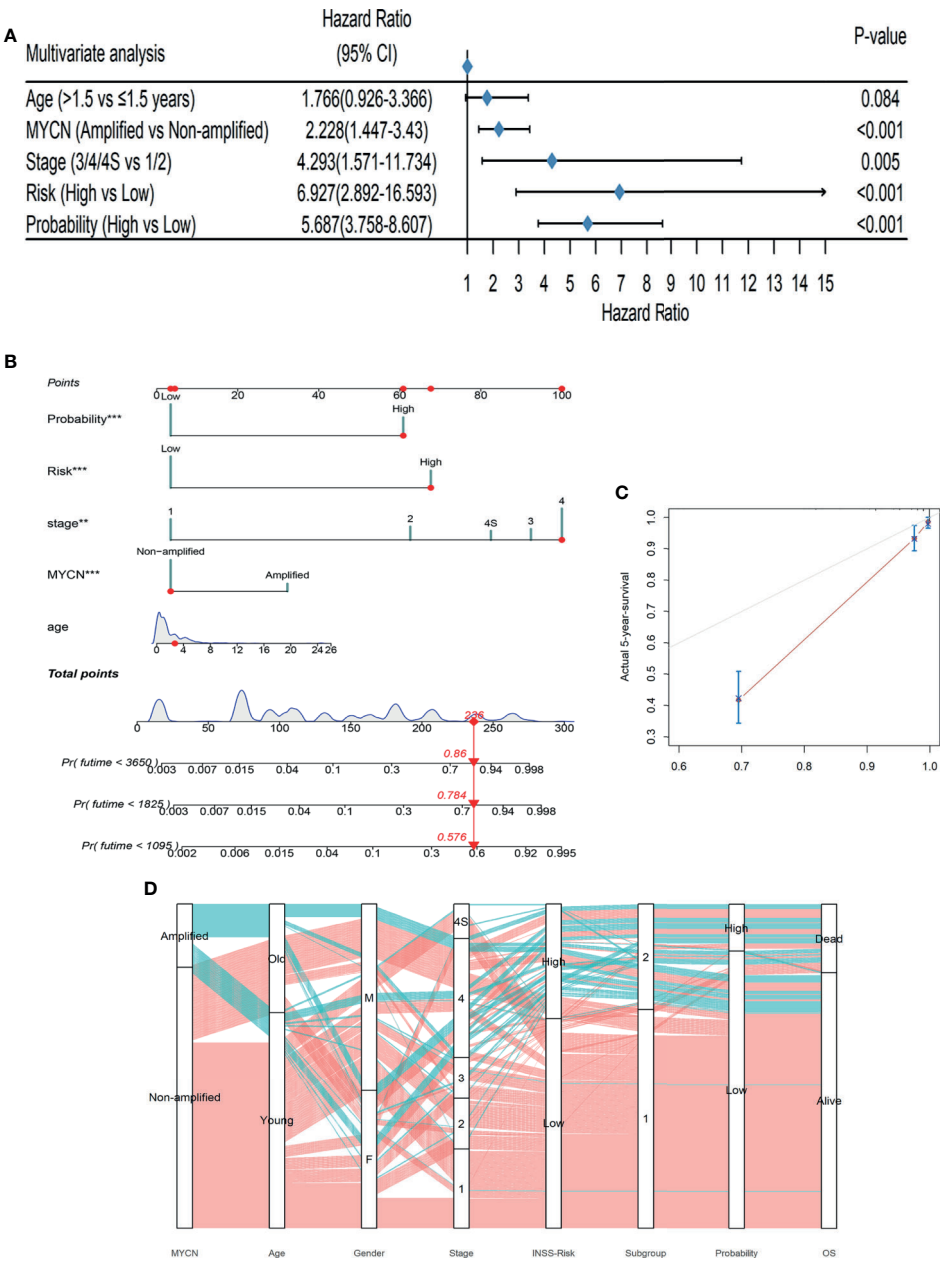


FIGURE 5 | The DL-model was independent of clinical covariates and could aid to diagnose. **(A)** The DL output probability was significant in the multivariable cox regression with clinical covariates ($p < 0.001$). **(B)** A nomogram could be beneficial for survival time prediction. **(C)** The calibration curve of the nomogram with a *C*-index: 0.889. **(D)** The alluvial diagram visualized the general conditions of patients. Coralline lines represented patients without *MYCN* amplified and atroceruleous lines represented those with *MYCN* amplified.

The alluvial diagram summarized the samples in our study (**Figure 5D**). 69.07% (67/97) of *MYCN*-amplified patients would be at stage 4 and this tendency was notable (chi-square test: $p < 0.001$). Only 21 patients diagnosed INSS low risk were classified into S2 and 35 with high risk into S1, showing that our prognostic subgroups were highly clinical-relevant (chi-square test $p < 0.001$). Subgroups and probabilities were also correlated (chi-square test $p < 0.001$). In the alluvial diagram, we observed that if patients had *MYCN* amplified, whether they were old or young, male or female, most of them would be at stage 4, INSS high risk, Subgroup 2. Whereas *MYCN* was not amplified, an antithetical conclusion would be drawn.

A *CCNB1*-Associated ceRNA Network Is Related to the Survivals of Neuroblastoma Patients

In order to isolate the hub genes of these 172 features, we first retrieved their PPI on the *String* website (**Figure 6A**). We then input them into *Cytoscape* software, and used the *cytoHubba* module to uncover hub genes by using 12 different algorithms. We summed up the top 5 genes in each algorithm and finally selected 18 genes as pivotal molecules in the network.

Next, we aimed to identify lncRNAs that could participate in the regulation of hub genes. GSE49711 is the RNA-seq of the same sample with GSE49710 and was used to uncover potential lncRNA-mRNA pairs. Subsequently, 5 mRNAs (*FBXO17*, *GNG11*, *CCNB1*, *KLF2* and *CD9*) were highly connected with 17 lncRNAs (**Figure 6B** and **Supplementary Table 4**). We noticed that both *CCNB1* and *CD9* could interact with 4 lncRNAs (*MYCNOS*, *TERC*, *SNHG1*, *MIR17HG*), along with positive coefficients between *CCNB1* and lncRNAs and opposite trends on *CD9*. *CCNB1*, an oncogene that controls the cell cycle at G2/M, had been found to be overexpressed in hepatocellular carcinoma and pancreatic cancer (49, 50). *CD9* is a tetraspanin involved in cell adhesion, metastasis and

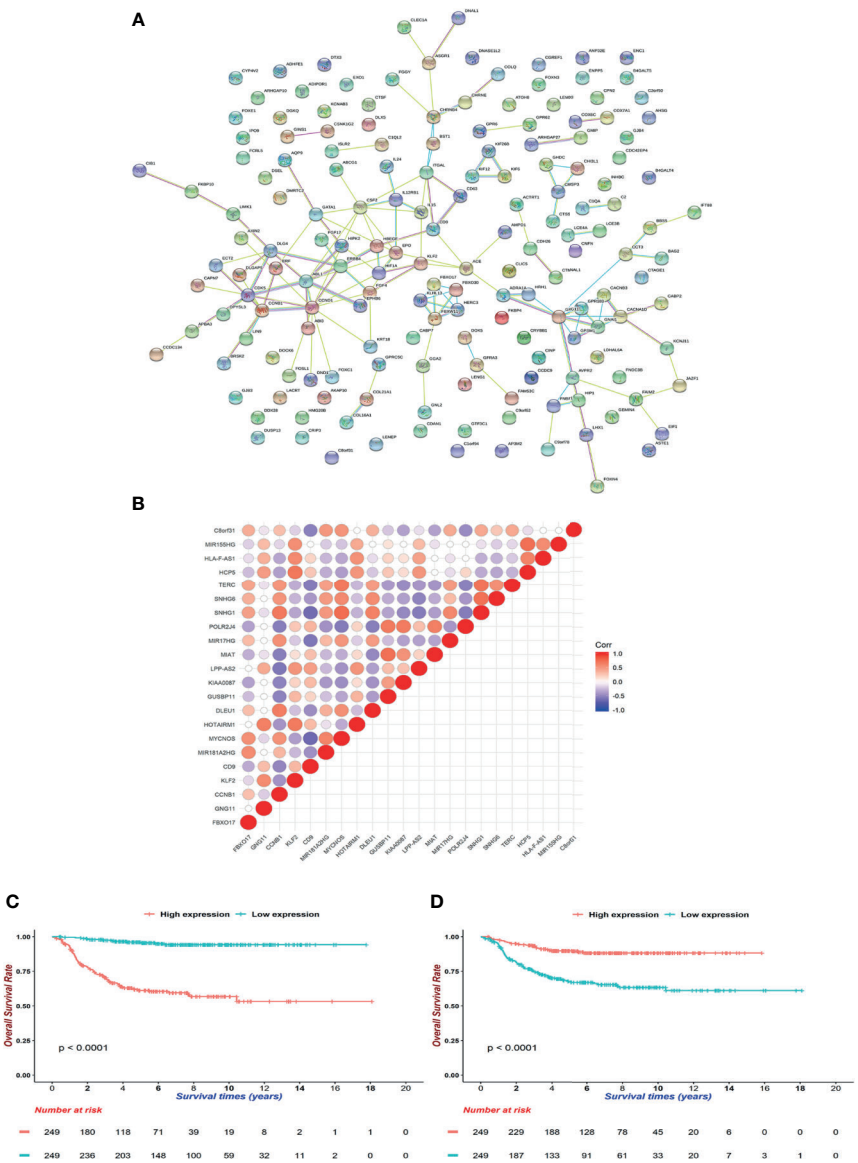


FIGURE 6 | Underlying hub genes and associated interactions in neuroblastoma. **(A)** The protein-protein-interaction (PPI) network was constructed by *String* website. **(B)** Mutual correlations among 18 hub genes filtered by *cytoHubba* module in *Cytoscape* software. The dots are colored red when Pearson correlation coefficients approach 1 and dots are colored blue when coefficients reach -1. **(C, D)** KM curves for CCNB1 and CD9 **(D)**. Expressions were cut by median levels. (both log-rank test $p < 0.0001$, n = 498).

inflammation in cancer (51, 52). *CCNB1* curtailed and *CD9* increased survivals in GSE49710 and *R2* (https://hgserver1.amc.nl/cgi-bin/r2/main.cgi), which were consistent with previous reports (both $p < 0.001$, **Figures 6C, D**) (49–52).

The ceRNA theory proposed that lncRNAs and mRNAs competed to interact with shared miRNAs, up-regulating downstream RNAs by impairing miRNA activities (53). We created a *CCNB1*-associated ceRNA network as described in Methods. (**Supplementary Figure 8**, **Supplementary Table 5**). Of note, the mir-302 family (hsa-mir-302-a, -b, -c and -d) which was highly expressed in embryonic stem cells, was associated with *CCNB1* and *MIR17HG*. This indicated that mir-302 might reduce the proliferation of neuroblastoma as it did in other cancers (54, 55).

Inhibitory Cells and Cytokines Increased in S2

Since distinct immune response patterns were observed between the two subgroups in GO-enrichment analysis and GSEA (**Supplementary Figure 3**), we further analyzed the immune microenvironment in neuroblastoma. The ESTIMATE algorithm was used to infer the purity of the microenvironment by scoring immune and stromal cells (56). Two groups did differ in stromal scores but not in immune scores, suggesting that S1 might preserve more normal stromal cells (**Supplementary Figure 9**). We used the ssGSEA algorithm to convert gene expression data into relative cell proportions (**Supplementary Figure 10A**). The numbers of the T regulatory cells (Tregs), Natural Killer (NK) cells, Monocytes, MDSCs, Eosinophils and central memory CD4 T cells were up-regulated in S2 compared to S1, while the Memory B cells, Macrophages, Gamma Delta T cells and central memory CD8 T cells exhibited opposite trends. Despite a rich amount of innate cytotoxic NK cells in S2, inhibitory immune cells like Tregs and MDSCs might contribute to deficient cytolytic activities. Since *GZMA*PRF1* could represent tumor microenvironment cytolytic activities (57), these data indicated that S1 may possess superior cytolytic activities which might eliminate tumor cells conspicuously (**Supplementary Figure 11A**). In addition, only activated CD8+ T cells were connected with survival events in two subgroups at $p < 0.05$, however, they anticipated contradicting outcomes (**Supplementary Figure 10B**, **Supplementary Table 6**). This implied that CD8+ T cells might play dual roles in neuroblastoma patients, i.e., CD8+ T cells functioned as a normal beneficial factor in malignant tumors in S2, however, impeding patients of S1 from recovering.

Then, we examined the intrinsic immune escape mechanism in neuroblastoma. Down-regulations of interferon signals and droppings of two γ-IFN receptors were observed in S2 (**Supplementary Figures 3**, **11B**), whereas IL-2 was increased in S1 which might stimulate T cell differentiation. A loss of HLA-class I/II can aid tumor cells to escape from immune monitoring. HLA-A and HLA-C were lower in S2, making the tumors prone to survive (**Supplementary Figure 11C**) (58, 59). The expression levels of *PDCD1*, *PDL1* and *CTLA4*, which are critical immune checkpoint genes, were also affected in S2 (60). Overall, these data suggest that disturbance of the immune system may be underneath the poor outcomes of the patients in S2.

DISCUSSION

One of the cruxes for neuroblastoma treatments is the heterogeneity. *MYCN* amplification and INSS risk classification have improved the efficacy to herald survivals, which many studies have unraveled genetic polymorphisms among. However, the current staging and grading systems are mainly based on clinical phenotypes, while it is steadily accepted that patients should be categorized by genetic associations.

Machine-learning and deep-learning methods have been used in medicine for many years. Generally, a deep-learning model receives multi-omics data and predicts outcomes by one or more neural networks. Chaudhary et al. used RNA-seq, miRNA-seq and DNA methylation data to train an autoencoder and partition patients into two prognostic groups (20). Chabon et al. sequenced SNV and CNV data of cell-free DNA in patients with lung cancers and controls. They established a 'Lung-CLip' machine-learning model to score each patient and determined whether a patient got lung cancer by the relative score (19). In this study, we used a DL-based classifier to significantly improve the prediction of neuroblastoma outcomes. We fed 172 genes expression data to the neural network and enrolled the attention mechanism into the survival classifier. The output probability could tell whether a patient could be dead or alive. Moreover, the 172 features selected for survival prediction could help characterize the genetic heterogeneities among the neuroblastoma patients.

A special attention mechanism was employed to combine two different parts of RNAs together (23). The attention mechanism is firstly presented by Vaswani et al. and widely used in computer vision and natural language processing (61), which is helpful to find interactions among different features, such as importance, relationship and so on. The attention mechanism can help the network learn how these two different groups of genes interact with each other. Information learned by the network can help it achieve a better performance. Indeed, our model outperformed traditional cox models, gaining a 5-year-survival AUC 0.974 and 0.896 in GSE49710 and EMTAB cohorts respectively. Besides, the PCA method failed to partition patients into appropriate prognostic groups, suggesting the superiority of our methods. Finally, we ran the gamut from all the samples in two cohorts, showing the robustness of our DL-based model.

For a long period, lncRNAs have been thought fruitless until recent advances that they might participate in chromosome stabilization, transcriptional initiation, localization, etc., thus broadening the cancer epigenetic network and making it possible for new drugs (62–64). Here, we identified four critical lncRNAs: *MYCNOS*, *TERC*, *SNHG1* and *MIR17HG*. *MYCNOS*, the antisense of *MYCN*, functions as the regulator of upstream *MYCN* promotor to enhance *MYCN* expressions. *TERC*, the telomerase RNA component, part of the telomerase, could proliferate prostate cancer cells (65, 66). *SNHG1* up-regulates in colorectal, liver, prostate and gastric cancers, which is the

biomarker for decreased survivals (67). Also it contributes to the neuroinflammation in Parkinson's disease (68). *MIR17HG* promotes colorectal as well as gastric cancer progression and up-regulates *PD-L1* expression, which could be inhibited by γ-IFN (69, 70). Investigations about those lncRNAs indicated that they could be engaged in the oncogenesis of neuroblastoma.

Our DL-based approach evinced a pathbreaking conjecture for survivals of neuroblastoma patients, still, there are some caveats should be aware of. First, neural networks are thought to be uninterpreted for now. We tried to exploit an attention mechanism to decipher underlying juxtapositions of genes involved in neuroblastoma, however, we could not declare how these neural networks work explicitly. Second, we only applied our model into two datasets that provided high-quality sequencing results as well as unequivocal labels and clinical annotations for each patient. We expected to test the reliability in more large cohorts. Last but not the least, we exerted neural networks on 172 features, which would be an obstacle for massive use in clinical examination due to its costs.

In summary, a DL-based model was constructed using 172 gene expressions to forecast survival status of neuroblastoma. Patients were split into two groups, which presented distinct microenvironments and clinical denouements. Our work paved the way for applications of artificial intelligence in medicine, not only in survival prediction, but also biological interpretations and associated accurate medicine.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE49710 https://wwwdev.ebi.ac.uk/arrayexpress/experiments/E-MTAB-8248/.

ETHICS STATEMENT

This study did not need approvals by the Ethics Committee of Tongji Hospital, Wuhan, China according to the regulations. This study fully complied guidelines of GEO and ArrayExpress.

AUTHOR CONTRIBUTIONS

CF, TX, and ZY contributed to devising the overall pipeline. TX trained neural networks and CF conducted the following analysis. XM, XC, GH, and XZ contributed to providing biological and technical instructions. FC, BX and JF contributed to concrete guides and funding in this project. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Key Research and Development Program of China [2016YFE0203900].

ACKNOWLEDGMENTS

We thank all colleagues and mentors for their contributions and instructions in this project.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.653863/full#supplementary-material

Supplementary Figure 1 | Consensus clustering results for GSE49710. **(A)** Cumulative density functions (CDF) for k=2 to 8. **(B)** Relative changes in CDF curves. **(C–F)** Consensus matrices for k=2 to 5.

Supplementary Figure 2 | KM curves of the EMTAB cohort for OS **(A)** and EFS **(B)**. S1 (coralline line) and S2 (atroceruleous line) were determined by the same procedure as GSE49710 (both log-rank test $p < 0.0001$).

Supplementary Figure 3 | GSEA plots showed that S1 exhibited higher biochemical activities **(A–E)** while S2 owned immune responses **(F–I)**.

Supplementary Figure 4 | Explained variances of each top 10 PCA dimensions.

Supplementary Figure 5 | Consensus clustering results for GSE49710 using PCA dimensions. **(A)** Cumulative density functions (CDF) for k=2 to 8. **(B)** Relative changes in CDF curves. **(C–F)** Consensus matrices for k=2 to 5.

Supplementary Figure 6 | The univariable cox regression result of age, *MYCN* status, gender, stage, INSS-risk and DL-probability.

Supplementary Figure 7 | Decision curve analysis for 3 models: DL-model (red), clinical covariates (palm green) and combined model (green).

Supplementary Figure 8 | The ceRNA network associated with *CCNB1*, *CD9*, *MYCNOS*, *TERC*, *SNHG1* and *MIR17HG*.

Supplementary Figure 9 | ESTIMATE scores for samples in GSE49710. **(A)** The total ESTIMATE scores. **(B)** Immune socres. **(C)** Stromal scores.

Supplementary Figure 10 | The immune microenvironment in neuroblastoma. **(A)** Relative proportions of cell types in S1 and S2. Wilcoxon rank sum tests were used to detect differences between two subgroups (n=498). ns, not significant; *:0.05, **:0.01, ***:0.001, ****:0.0001 **(B)** Cox regressions for individual cell types in S1, S2 and the whole cohort. Dots are colored red when hazard ratios are higher than 1 and are colored blue when hazard ratios are less than 1. Also, a larger circle means a lower *p*-value.

Supplementary Figure 11 | Immune microenvironment molecules. Differences between groups were examined by Wilcoxon rank sum tests. ns: not significant, *:0.05, **:0.01, ***:0.001, ****:0.0001 **(A)** S1 owned higher cytolytic activities, which were calculated by *GZMA*PRF1* (Wilcoxon rank sum tests: $p < 0.001$). **(B)** Cytokines in S1 and S2. **(C)** HLA molecules in S1 and S2. **(D)** Immune checkpoint molecules in S1 and S2.

REFERENCES

1. Matthay KK, Maris JM, Schleiermacher G, Nakagawara A, Mackall CL, Diller L, et al. Neuroblastoma. *Nat Rev Dis Primers* (2016) 2:16078. doi: 10.1038/nrdp.2016.78
2. Maris JM, Hogarty MD, Bagatell R, Cohn SL. Neuroblastoma. *Lancet* (2007) 369(9579):2106–20. doi: 10.1016/S0140-6736(07)60983-0
3. Pinto NR, Applebaum MA, Volchenboum SL, Matthay KK, London WB, Ambros PF, et al. Advances in Risk Classification and Treatment Strategies for Neuroblastoma. *J Clin Oncol* (2015) 33(27):3008–17. doi: 10.1200/JCO.2014.59.4648
4. Cheung NK, Dyer MA. Neuroblastoma: Developmental Biology, Cancer Genomics and Immunotherapy. *Nat Rev Cancer* (2013) 13(6):397–411. doi: 10.1038/nrc3526
5. Jones DTW, Banito A, Grunewald TGP, Haber M, Jager N, Kool M, et al. Molecular Characteristics and Therapeutic Vulnerabilities Across Paediatric Solid Tumours. *Nat Rev Cancer* (2019) 19(8):420–38. doi: 10.1038/s41568-019-0169-x
6. Louis CU, Shohet JM. Neuroblastoma: Molecular Pathogenesis and Therapy. *Annu Rev Med* (2015) 66:49–63. doi: 10.1146/annurev-med-011514-023121
7. Vandesompele J, Baudis M, De Preter K, Van Roy N, Ambros P, Bown N, et al. Unequivocal Delineation of Clinicogenetic Subgroups and Development of a New Model for Improved Outcome Prediction in Neuroblastoma. *J Clin Oncol* (2005) 23(10):2280–99. doi: 10.1200/JCO.2005.06.104
8. Vermeulen J, De Preter K, Naranjo A, Vercruysse L, Van Roy N, Hellemans J, et al. Predicting Outcomes for Children With Neuroblastoma Using a Multigene-Expression Signature: A Retrospective SIOPEN/COG/GPOH Study. *Lancet Oncol* (2009) 10(7):663–71. doi: 10.1016/S1470-2045(09)70154-8
9. Camacho DM, Collins KM, Powers RK, Costello JC, Collins JJ. Next-Generation Machine Learning for Biological Networks. *Cell* (2018) 173(7):1581–92. doi: 10.1016/j.cell.2018.05.015
10. Niazi MKK, Parwani AV, Gurcan MN. Digital Pathology and Artificial Intelligence. *Lancet Oncol* (2019) 20(5):e253–61. doi: 10.1016/S1470-2045(19)30154-8
11. Ngiam KY, Khor IW. Big Data and Machine Learning Algorithms for Health-Care Delivery. *Lancet Oncol* (2019) 20(5):e262–73. doi: 10.1016/S1470-2045(19)30149-4
12. Libbrecht MW, Noble WS. Machine Learning Applications in Genetics and Genomics. *Nat Rev Genet* (2015) 16(6):321–32. doi: 10.1038/nrg3920
13. Chen YC, Ke WC, Chiu HW. Risk Classification of Cancer Survival Using ANN With Gene Expression Data From Multiple Laboratories. *Comput Biol Med* (2014) 48:1–7. doi: 10.1016/j.compbiomed.2014.02.006
14. Xu X, Zhang Y, Zou L, Wang M, Li A. eds. A Gene Signature for Breast Cancer Prognosis Using Support Vector Machine, in: *2012 5th International Conference on BioMedical Engineering and Informatics*. IEEE 2012 16-18 Oct (2012).
15. Delen D, Walker G, Kadam A. Predicting Breast Cancer Survivability: A Comparison of Three Data Mining Methods. *Artif Intell Med* (2005) 34(2):113–27. doi: 10.1016/j.artmed.2004.07.002
16. Ting DSW, Cheung CY-L, Lim G, Tan GSW, Quang ND, Gan A, et al. Development and Validation of a Deep Learning System for Diabetic Retinopathy and Related Eye Diseases Using Retinal Images From Multiethnic Populations With Diabetes. *JAMA* (2017) 318(22):2211–23. doi: 10.1001/jama.2017.18152
17. Bi WL, Hosny A, Schabath MB, Giger ML, Birkbak NJ, Mehrtash A, et al. Artificial Intelligence in Cancer Imaging: Clinical Challenges and Applications. *CA Cancer J Clin* (2019) 69(2):127–57. doi: 10.3322/caac.21552
18. Eraslan G, Avsec Z, Gagneur J, Theis FJ. Deep Learning: New Computational Modelling Techniques for Genomics. *Nat Rev Genet* (2019) 20(7):389–403. doi: 10.1038/s41576-019-0122-6
19. Chabon JJ, Hamilton EG, Kurtz DM, Esfahani MS, Moding EJ, Stehr H, et al. Integrating Genomic Features for Non-Invasive Early Lung Cancer Detection. *Nature* (2020) 580(7802):245–51. doi: 10.1038/s41586-020-2140-0
20. Chaudhary K, Poirion OB, Lu L, Garmire LX. Deep Learning-Based Multi-Omics Integration Robustly Predicts Survival in Liver Cancer. *Clin Cancer Res* (2018) 24(6):1248–59. doi: 10.1158/1078-0432.CCR-17-0853
21. Tranchevent L-C, Azuaje F, Rajapakse JC. A Deep Neural Network Approach to Predicting Clinical Outcomes of Neuroblastoma Patients. *BMC Med Genomics* (2019) 12(Suppl 8):178–. doi: 10.1186/s12920-019-0628-y
22. Park A, Nam S. Deep Learning for Stage Prediction in Neuroblastoma Using Gene Expression Data. *Genomics Inform* (2019) 17(3):e30–e. doi: 10.5808/GI.2019.17.3.e30
23. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al eds. Attention Is All You Need. In: *Advances in Neural Information Processing Systems*. NIPS
24. Lin T-Y, Goyal P, Girshick R, He K, Dollár P. eds. Focal Loss for Dense Object Detection. In: *Proceedings of the IEEE International Conference on Computer Vision*. ICCV.
25. Kingma DP, Ba J. Adamml: A Method for Stochastic Optimization. (2014). arXiv preprint arXiv:14126980.
26. Glorot X, Bengio Y. eds. Understanding the Difficulty of Training Deep Feedforward Neural Networks. In: *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*. PMLR.
27. Eusebi P. Diagnostic Accuracy Measures. *Cerebrovasc Dis* (2013) 36(4):267–72. doi: 10.1159/000353863
28. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, et al. pROC: An Open-Source Package for R and S+ to Analyze and Compare ROC Curves. *BMC Bioinf* (2011) 12:77. doi: 10.1186/1471-2105-12-77
29. Heagerty PJ, Lumley T, Pepe MS. Time-Dependent ROC Curves for Censored Survival Data and a Diagnostic Marker. *Biometrics* (2000) 56(2):337–44. doi: 10.1111/j.0006-341X.2000.00337.x
30. Wickham H. *Ggplot2: Elegant Graphics for Data Analysis*. German: Springer (2016).
31. Zhong X, Liu Y, Liu H, Zhang Y, Wang L, Zhang H. Identification of Potential Prognostic Genes for Neuroblastoma. *Front Genet* (2018) 9:589. doi: 10.3389/fgene.2018.00589
32. De Preter K, Vermeulen J, Brors B, Delattre O, Eggert A, Fischer M, et al. Accurate Outcome Prediction in Neuroblastoma Across Independent Data Sets Using a Multigene Signature. *Clin Cancer Res* (2010) 16(5):1532–41. doi: 10.1158/1078-0432.CCR-09-2607
33. Wilkerson MD, Hayes DN. ConsensusClusterPlus: A Class Discovery Tool With Confidence Assessments and Item Tracking. *Bioinformatics* (2010) 26(12):1572–3. doi: 10.1093/bioinformatics/btq170
34. Gu Z, Eils R, Schlesner M. Complex Heatmaps Reveal Patterns and Correlations in Multidimensional Genomic Data. *Bioinformatics* (2016) 32(18):2847–9. doi: 10.1093/bioinformatics/btw313
35. Therneau TM, Grambsch PM. The Cox Model. In: *Modeling Survival Data: Extending the Cox Model*, vol. p. German: Springer (2000). p. 39–77.
36. Kassambara A, Kosinski M, Biecek P, Fabian S. Survminer: Drawing Survival Curves Using'ggplot2'. *R Package version 03.* (2017) 1.
37. Brunson J. Ggalluvial: Layered Grammar for Alluvial Plots. *J Open Source Software* (2020) 5(49). doi: 10.21105/joss.02017
38. Yu G, Wang LG, Han Y, He QY. Clusterprofiler: An R Package for Comparing Biological Themes Among Gene Clusters. *OMICS: A J Integr Biol* (2012) 16(5):284–7. doi: 10.1089/omi.2011.0118
39. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res* (2003) 13(11):2498–504. doi: 10.1101/gr.1239303
40. Chin C-H, Chen S-H, Wu H-H, Ho C-W, Ko M-T, Lin C-Y. Cytohubba: Identifying Hub Objects and Sub-Networks From Complex Interactome. *BMC Syst Biol* (2014) 8 Suppl 4(Suppl 4):S11–S. doi: 10.1186/1752-0509-8-S4-S11
41. Karagkouni D, Paraskevopoulou MD, Chatzopoulos S, Vlachos IS, Tastsoglou S, Kanellos I, et al. DIANA-TarBase V8: A Decade-Long Collection of Experimentally Supported miRNA–Gene Interactions. *Nucleic Acids Res* (2017) 46(D1):D239–45. doi: 10.1093/nar/gkx1141
42. Paraskevopoulou MD, Vlachos IS, Karagkouni D, Georgakilas G, Kanellos I, Vergoulis T, et al. DIANA-LncBase V2: Indexing microRNA Targets on Non-Coding Transcripts. *Nucleic Acids Res* (2016) 44(D1):D231–8. doi: 10.1093/nar/gkv1270
43. Hänzelmann S, Castelo R, Guinney J. GSVA: Gene Set Variation Analysis for Microarray and RNA-Seq Data. *BMC Bioinf* (2013) 14(1):7. doi: 10.1186/1471-2105-14-7

44. Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, et al. Pan-Cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade. *Cell Rep* (2017) 18(1):248–62. doi: 10.1016/j.celrep.2016.12.019
45. Yu H, Lee H, Herrmann A, Buettner R, Jove R. Revisiting STAT3 Signalling in Cancer: New and Unexpected Biological Functions. *Nat Rev Cancer* (2014) 14 (11):736–46. doi: 10.1038/nrc3818
46. Yan S, Li Z, Thiele CJ. Inhibition of STAT3 With Orally Active JAK Inhibitor, AZD1480, Decreases Tumor Growth in Neuroblastoma and Pediatric Sarcomas *In Vitro* and *In Vivo*. *Oncotarget* (2013) 4(3):433–45. doi: 10.18632/oncotarget.930
47. Young T, Hazarika D, Poria S, Cambria E. Recent Trends in Deep Learning Based Natural Language Processing [Review Article]. *IEEE Comput Intell Magazine* (2018) 13(3):55–75. doi: 10.1109/MCI.2018.2840738
48. Vickers AJ, Elkin EB. Decision Curve Analysis: A Novel Method for Evaluating Prediction Models. *Med Decis Making* (2006) 26(6):565–74. doi: 10.1177/0272989X06295361
49. Zhang H, Zhang X, Li X, Meng WB, Bai ZT, Rui SZ, et al. Effect of CCNB1 Silencing on Cell Cycle, Senescence, and Apoptosis Through the P53 Signaling Pathway in Pancreatic Cancer. *J Cell Physiol* (2018) 234(1):619–31. doi: 10.1002/jcp.26816
50. Chai N, Xie HH, Yin JP, Sa KD, Guo Y, Wang M, et al. FOXM1 Promotes Proliferation in Human Hepatocellular Carcinoma Cells by Transcriptional Activation of CCNB1. *Biochem Biophys Res Commun* (2018) 500(4):924–9. doi: 10.1016/j.bbrc.2018.04.201
51. Murayama Y, Oritani K, Tsutsui S. Novel CD9-Targeted Therapies in Gastric Cancer. *World J Gastroenterol* (2015) 21(11):3206–13. doi: 10.3748/wjg.v21.i11.3206
52. Brosseau C, Colas L, Magnan A, Brouard S. CD9 Tetraspanin: A New Pathway for the Regulation of Inflammation? *Front Immunol* (2018) 9:2316. doi: 10.3389/fimmu.2018.02316
53. Thomson DW, Dinger ME. Endogenous microRNA Sponges: Evidence and Controversy. *Nat Rev Genet* (2016) 17(5):272–83. doi: 10.1038/nrg.2016.20
54. Maadi H, Moshtaghian A, Taha MF, Mowla SJ, Kazeroonian A, Haass NK, et al. Multimodal Tumor Suppression by miR-302 Cluster in Melanoma and Colon Cancer. *Int J Biochem Cell Biol* (2016) 81:121–32. doi: 10.1016/j.biocel.2016.11.004
55. Ahmadalizadeh Khanehsar M, Hoseinbeyki M, Fakhr Taha M, Javeri A. Repression of TGF-β Signaling in Breast Cancer Cells by miR-302/367 Cluster. *Cell J* (2020) 21(4):444–50. doi: 10.22074/cellj.2020.6193
56. Yoshihara K, Shahmoradgoli M, Martinez E, Vegesna R, Kim H, Torres-Garcia W, et al. Inferring Tumour Purity and Stromal and Immune Cell Admixture From Expression Data. *Nat Commun* (2013) 4:2612. doi: 10.1038/ncomms3612
57. Rooney MS, Shukla SA, Wu CJ, Getz G, Hacohen N. Molecular and Genetic Properties of Tumors Associated With Local Immune Cytolytic Activity. *Cell* (2015) 160(1-2):48–61. doi: 10.1016/j.cell.2014.12.033
58. Garrido F. HLA Class-I Expression and Cancer Immunotherapy. *Adv Exp Med Biol* (2019) 1151:79–90. doi: 10.1007/978-3-030-17864-2_3
59. Raffaghello L, Prigione I, Airoldi I, Camoriano M, Morandi F, Bocca P, et al. Mechanisms of Immune Evasion of Human Neuroblastoma. *Cancer Lett* (2005) 228(1–2):155–61. doi: 10.1016/j.canlet.2004.11.064
60. Nallasamy P, Chava S, Verma SS, Mishra S, Gorantla S, Coulter DW, et al. PD-L1, Inflammation, non-Coding RNAs, and Neuroblastoma: Immuno-Oncology Perspective. *Semin Cancer Biol* (2018) 52(Pt 2):53–65. doi: 10.1016/j.semcancer.2017.11.009
61. Hu D. ed. An Introductory Survey on Attention Mechanisms in NLP Problems. In: *Proceedings of SAI Intelligent Systems Conference*. German: Springer.
62. Quinn JJ, Chang HY. Unique Features of Long non-Coding RNA Biogenesis and Function. *Nat Rev Genet* (2016) 17(1):47–62. doi: 10.1038/nrg.2015.10
63. Schmitt AM, Chang HY. Long Noncoding RNAs in Cancer Pathways. *Cancer Cell* (2016) 29(4):452–63. doi: 10.1016/j.ccell.2016.03.010
64. Matsui M, Corey DR. Non-Coding RNAs as Drug Targets. *Nat Rev Drug Discovery* (2017) 16(3):167–79. doi: 10.1038/nrd.2016.117
65. Laudadio I, Orso F, Azzalin G, Calabro C, Berardinelli F, Coluzzi E, et al. AGO2 Promotes Telomerase Activity and Interaction Between the Telomerase Components TERT and TERC. *EMBO Rep* (2019) 20(2). doi: 10.15252/embr.201845969
66. Baena-Del Valle JA, Zheng Q, Esopi DM, Rubenstein M, Hubbard GK, Moncaliano MC, et al. MYC Drives Overexpression of Telomerase RNA (hTR/TERC) in Prostate Cancer. *J Pathol* (2018) 244(1):11–24. doi: 10.1002/path.4980
67. Thin KZ, Tu JC, Raveendran S. Long non-Coding SNHG1 in Cancer. *Clin Chim Acta* (2019) 494:38–47. doi: 10.1016/j.cca.2019.03.002
68. Cao B, Wang T, Qu Q, Kang T, Yang Q. Long Noncoding RNA SNHG1 Promotes Neuroinflammation in Parkinson's Disease via Regulating miR-7/NLRP3 Pathway. *Neuroscience* (2018) 388:118–27. doi: 10.1016/j.neuroscience.2018.07.019
69. Xu J, Meng Q, Li X, Yang H, Xu J, Gao N, et al. Long Noncoding RNA MIR17HG Promotes Colorectal Cancer Progression via miR-17-5p. *Cancer Res* (2019) 79(19):4882–95. doi: 10.1158/0008-5472.CAN-18-3880
70. Yuan J, Tan L, Yin Z, Zhu W, Tao K, Wang G, et al. MIR17HG-miR-18a/19a Axis, Regulated by Interferon Regulatory Factor-1, Promotes Gastric Cancer Metastasis via Wnt/beta-Catenin Signalling. *Cell Death Dis* (2019) 10(6):454. doi: 10.1038/s41419-019-1685-z

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

REVIEW
published: 21 July 2021
doi: 10.3389/fonc.2021.690641

The Genetic Changes of Hepatoblastoma

Huitong Chen[1], *Qian Guan*[1,2], *Huiqin Guo*[1,2], *Lei Miao*[1] *and Zhenjian Zhuo*[1*]

[1] *Department of Pediatric Surgery, Guangzhou Institute of Pediatrics, Guangdong Provincial Key Laboratory of Research in Structural Birth Defect Disease, Guangzhou Women and Children's Medical Center, Guangzhou Medical University, Guangzhou, China,* [2] *School of Medicine, South China University of Technology, Guangzhou, China*

Hepatoblastoma is the most common malignant liver cancer in childhood. The etiology of hepatoblastoma remains obscure. Hepatoblastoma is closely related to genetic syndromes, hinting that hepatoblastoma is a genetic predisposition disease. However, no precise exposures or genetic events are reported to hepatoblastoma occurrence. During the past decade, significant advances have been made in the understanding of etiology leading to hepatoblastoma, and several important genetic events that appear to be important for the development and progression of this tumor have been identified. Advances in our understanding of the genetic changes that underlie hepatoblastoma may translate into better patient outcomes. Single nucleotide polymorphisms (SNPs) have been generally applied in the research of etiology's exploration, disease treatment, and prognosis assessment. Here, we reviewed and discussed the molecular epidemiology, especially SNPs progresses in hepatoblastoma, to provide references for future studies and promote the study of hepatoblastoma's etiology.

Keywords: hepatoblastoma, etiology, genetics, single nucleotide polymorphism, epidemiology

OPEN ACCESS

Edited by:
Hua Tan,
University of Texas Health Science Center at Houston, United States

Reviewed by:
Moira Garraus,
Hospital Sant Joan de Déu Barcelona, Spain
Zhimin Liu,
Janssen Pharmaceuticals, Inc., United States

***Correspondence:**
Zhenjian Zhuo
zhenjianzhuo@163.com

Specialty section:
This article was submitted to Pediatric Oncology, a section of the journal Frontiers in Oncology

Received: *03 April 2021*
Accepted: *05 July 2021*
Published: *21 July 2021*

Citation:
Chen H, Guan Q, Guo H, Miao L and Zhuo Z (2021) The Genetic Changes of Hepatoblastoma. Front. Oncol. 11:690641. doi: 10.3389/fonc.2021.690641

INTRODUCTION

Hepatoblastoma arising from the hepatocyte precursor is the most common malignant liver tumor among children (1). The typical clinical symptoms of hepatoblastoma are alpha-fetoprotein (AFP) rising and abdominal mass (2). Due to the rarity of hepatoblastoma, diagnosis and treatment are facing challenges. With medical-technical development such as adjuvant chemotherapy and hepatectomy in decades, the 5-year survival rate is greater than 70% nowadays (3, 4). Despite the improved survival rate, numerous survivors suffer treatment-related side effects, such as hearing loss or cardiomyopathy (5). In addition, the prognosis of advanced stage hepatoblastoma patients with unresectable tumors remains poor (6).

Unlike the hepatocellular carcinoma which has clear pathogenesis (HCC), the etiology of hepatoblastoma has no connection with hepatitis B virus or cirrhosis (7). The first study of genetic-molecular changes in hepatoblastoma was conducted in the late 1980s (8). However, the etiology of hepatoblastoma remains unclear by far. In this review, we aimed at giving a brief overview of the molecular epidemiology for hepatoblastoma, focusing on the SNPs that influence hepatoblastoma risk. We further discussed the clinical challenges for elucidating the etiology of hepatoblastoma and provided theoretical basis for future prevention, diagnosis, and therapeutic approaches for hepatoblastoma.

EPIDEMIOLOGY

Because of the rarity of hepatoblastoma, the hepatoblastoma's epidemiology has not been investigated comprehensively. Majority of hepatoblastoma are sporadic and are commonly found in children in their first 5 years, with a predominance in boys (9). The incidence of hepatoblastoma remains at a lower level worldwide comparing with other solid tumors in children, including neuroblastoma and Wilms tumor (10–12). Employing the Surveillance, Epidemiology and End Results (SEER) database, the incidence of hepatoblastoma in the United States was 1.5–1.9 per million with an upward tendency (5, 13). It is similar to the incidence rate of roughly 1.1 per million in China and 1.7 per million in the Nordic countries (14, 15). There is little difference of hepatoblastoma's incidence in diverse countries, but these data were not collected from the same period. The difference in ethnicity is also correlated to the incidence rate. It was reported that the incidence rate in blacks was relative lower (16). A population-based analysis conducted in the United States revealed that the higher the maternal education, the lower the incidence rate is of hepatoblastoma (17).

GENETIC SYNDROMES

Although most hepatoblastomas are sporadic (18), hepatoblastoma was reported to be closely related to genetic syndromes, including familial adenomatous polyposis (FAP), Beckwith-Wiedemann syndrome (BWS), and trisomy 18 (19–21).

The association between FAP and hepatoblastoma was originally perceived in 1983 (22). Mutation of the *APC* was detected in hepatoblastoma patients with a family history of FAP (23). *APC* mutation was also found in other cancers, including gastric and colorectal cancer (24, 25). However, in subsequent research, Harvey et al. demonstrated that no *APC* mutation was found in sporadic hepatoblastoma (26). BWS, an overgrowth syndrome associated with alteration of genomic imprinting on chromosome 11p15.5, is characterized by macroglossia, high birth weight, overgrowth of abdominal organs, and neonatal hypoglycemia (27, 28). Comparing with children without BWS, the relative rate of hepatoblastoma among children with BWS was 2,280 (95% CI: 928–11,656) (29). Trisomy 18, regarded as a fatal disease, affects approximately 1 per 6,000 newborns (30). The correlation between trisomy 18 and hepatoblastoma has been reported in several previous studies (21, 31–33). Tomlinson et al. raised an interesting point that the hepatoblastoma cases with trisomy 18 almost were females, and this situation was contrary to hepatoblastoma with higher prevalence in males (34). In addition to the syndromes mentioned above, some genetic syndromes have been reported, including Prader-Willi syndrome, and Simpson-Golabi-Behmel syndrome (35–37) (**Table 1**). However, these cases are too rare to consider as a close risk factor for hepatoblastoma.

TABLE 1 | Genetic syndromes of hepatoblastoma.

Syndrome	Chromosome	Correlated gene	Relative risk	Reference
FAP	5q21	*APC*	750–7,500	(38)
BWS	11p15.5	IGF2-H19	2280	(29)
Trisomy 18	18	/	Unknown	(21)
SGBS	Xq26	*GPC3*	Unknown	(37)
PWS	46XY del (15) (q11, q13)	/	Unknown	(36)

FAP, familial adenomatous polyposis; BWS, Beckwith-Wiedemann syndrome; SGBS, Simpson-Golabi-Behmel syndrome; PWS, Prader-Willi syndrome.

RISK FACTORS

Although the etiology of hepatoblastoma remains unclear, some risk factors have been identified (**Table 2**). The United Kingdom Childhood Cancer Study (UKCCS) reported that parental smoking is a risk factor for hepatoblastoma (OR = 4.74, 95% CI:1.68–13.35), although only 28 hepatoblastoma patients were recruited in this study (39). The result of subsequent studies from the United States and China also supported this conclusion (40, 41). In 2009, parental smoking has been declared as the significant high-risk factor for hepatoblastoma by the International Agency for Research on Cancer (47).

As mentioned above, overgrowth syndrome is firmly correlated with hepatoblastoma; not only that, premature birth and very low birth weight (VLBW, <1,500 g) is closely associated with hepatoblastoma. The data analyzation, which was from California's population-based cancer registry, indicated that hepatoblastoma risk is remarkably increase in VLBW children (OR=50.57, 95% CI: 6.59–387.97) (42). Although the ORs were diverse in different regions, whole ORs were greater than one and proved VLBW was an obvious risk factor for hepatoblastoma (15, 40, 43).

Janitz et al. had confirmed that maternal and paternal occupational exposures to paints were etiologically relevant to hepatoblastoma (44). Other studies indicated that parental occupational exposures to wood dust, metal fumes, and petroleum products also could be the risk factors (45, 46). On account of the rarity of hepatoblastoma, the investigations of risk factors are relatively limited.

SINGLE NUCLEOTIDE POLYMORPHISMS WITH HEPATOBLASTOMA

In the late 1990s, research showed that the existence of mutation of *CTNNB1* (β-catenin gene) might lead to β-catenin

TABLE 2 | Risk factors of hepatoblastoma.

Risk factors	Location/ category	OR	95% CI	Reference
Parental smoking	UK	4.74	1.68–13.35	(39)
	China	2.9	1.1–4.2	(40)
	US	2.69	1.18–6.13	(41)
VLBW	US	50.57	6.59–387.97	(42)
	China	26.0	14.0–65.7	(40)
	Japan	15.6	7.6–31.1	(43)
	Nordic countries	9.5	2.3–38.2	(15)
Parental occupational exposure	Paints	1.71	1.04–2.81	(44)
	Wood dust	2.41	0.99–5.88	(45)
	Metal fumes	8.0	1.5–148.4	(46)
	Petroleum	2.3	1.2–4.6	(46)

accumulation, resulting in the development of hepatoblastoma (48). Some pathways related to hepatoblastoma molecular mechanisms were detected, including the well-studied Wnt/β-catenin and MYC pathways (49). However, research on the molecular basis in hepatoblastoma was limited. It is urgent to identify early diagnostic molecular-genetic markers for timely and valid therapeutic choices.

Single nucleotide polymorphisms (SNPs), comparing with rare gene mutation, were identified to abundantly exist in human genome by Human Genome Project (50). The susceptibilities and pathogenesis of disease and genetic heterogeneity are tightly correlated with SNPs (51–54). When present in non-coding regions, SNPs are regarded as critical genetic markers and can regulate protein expression (55). SNPs have been generally applied in the research of etiology's exploration, treatment of disease, and prognosis assessment. Comparing to other solid tumors in children, studies of the association between hepatoblastoma and gene polymorphism are relatively few. Here we summarized the significant hepatoblastoma susceptibility SNPs in **Table 3**.

Cancer-Related Genes

Pakakasama et al. addressed the association between myeloperoxidase (MPO) promotor gene polymorphism located on chromosome 17q23 and hepatoblastoma in 2003 (56). They demonstrated that MPO-463 G>A was associated with the reduced susceptibility of hepatoblastoma. Their study represents the first case-control study regarding genetic polymorphism and hepatoblastoma risk, although the cases were only less than 100. The significant roles of MPO-463 G>A polymorphism were also reported in other cancers, including cervical, lung, breast, and bladder cancer (63–66). MPO is an oxidative enzyme located in neutrophils and monocytes. It can catalyze an oxidation reaction to generate hypochlorous acid (HOCl), which is involved in DNA damage and inhibition of DNA repair (67). Carrying G/A or A/A genotype affects the expression of MPO and reduces the generation of oxygen radicals to decrease the risk of cancer.

In the following year, Pakakasama et al. conducted another study to elaborate that *CCND1* gene rs9344 G>A polymorphism affecting gene splicing was associated with the age of onset of hepatoblastoma (57). *CCND1* was identified as the core gene in the β-catenin/LEF pathway, which is relevant to hepatoblastoma's development (68, 69). The same as the study mentioned above, the cases of these two studies were less than 100, which limited the reliability of the statistical result in the subgroups. In order to affirm these conclusions, study subjects are supposed to enlarge in a future study.

RAS Gene

More than a decade after that, the progression of studies about SNPs and hepatoblastoma was stagnant. In 2019, the third study, a relatively large-scale case-control study that recruited 213 cases in Chinese children, was conducted by our research group (70). As a famous oncogenic role, the *RAS* gene (*KRAS*, *NRAS*, and *HRAS*) is commonly mutated in human cancers (71–73). However, in this study, we regrettably identified that one *NRAS* polymorphism and three *KRAS* polymorphisms do not correlate with hepatoblastoma susceptibility.

Long Non-coding RNAs (lncRNAs)

LncRNAs, with over 200 nucleotides in length and involved in diverse gene regulation, account for a large number of ncRNAs (74). Various studies have verified that lncRNAs play a vital role in transcription processes, regulation of cellular contexts, assembly of protein, tumor suppressor dysregulation, and other crucial biological function (75–80). *LINC00673*, located on chromosome 17q24.3, has been reported as an oncogene in diverse cancers (81–83). Childs et al. performed a genome-wide association study (GWAS) to confirm that *LINC00673* rs11655237 polymorphism is associated with pancreatic cancer susceptibility (84). Considering the involvement of *LINC00673* in the occurrence and development of diverse cancers while there were no previous studies linking *LINC00673* to hepatoblastoma, our research group conducted a case-control study selecting this polymorphism and confirmed that the *LINC00673* rs11655237 C>T polymorphism may be correlated with hepatoblastoma susceptibility. In the stratified analysis, significant result was also found in the subgroup of clinical stages III+IV (58). The patients carrying this SNP seemed to tend to suffer severe hepatoblastoma. The conjecture based on the statistical result needs further validation.

TABLE 3 | Summary of hepatoblastoma susceptibility SNPs.

Chromosome	Variant	Candidate gene	Alternate allele	Effect	Reference
17q23	G-463-A	MPO	G>A	Protective factor	(56)
11q13	rs9344	CCND1	G>A	Risk factor	(57)
17q24.3	rs11655237	LINC00673	C>T	Risk factor	(58)
12q15	rs968697	HMGA2	T>C	Risk factor	(52)
11p15.5	rs2839698	H19	G>A	Risk factor	(59)
11p15.5	rs3024270	H19	C>G	Risk factor	(59)
11p15.5	rs217727	H19	G>A	Protective factor	(59)
20q13	rs6090311	YTHDF1	A>G	Protective factor	(60)
6q21	rs9404590	LIN28B	T>G	Risk factor	(61)
6q21	rs314276	LIN28B	C>A	Risk factor	(61)
6q25	rs7766006	WTAP	G>T	Protective factor	(62)
3p25.3	rs23795	hOGG1	A>G	Risk factor	

LncRNA *H19* gene, a maternally imprinted gene, is located on chromosome 11p5.5 and highly expressed during the stage of embryonic development (85, 86). H19 plays a vital role in tumorigenesis and the development of malignant tumors *via* regulation of transcription (87). Tan et al. identified that rs2839698 G>A and rs3024270 C>G, which decreased long non-coding RNA MRPL23 antisense RNA 1 (MRPL23-AS1) expression, were significantly correlated with increased hepatoblastoma risk. In contrast, rs217727 G>A increased MRPL23-AS1 expression to reduced hepatoblastoma risk in the Han population (59). Carrying GGG and AGG haplotypes (order rs2839698, rs3024270, rs217727), children have a tendency to suffer hepatoblastoma. These three polymorphisms were reported to affect the folding structures of *H19* mRNA (88). These results revealed that even though SNPs are in the same gene, their effects of hepatoblastoma may be different. These SNPs are expected to be the biomarkers of early diagnosis of hepatoblastoma. However, the functions of *H19* polymorphism in hepatoblastoma still need to be further validated.

LIN28

LIN28A and *LIN28B* are two paralogs of *LIN28*, located in chromosome 1p36.11 and 6q21, respectively (89). They can bind to the target RNAs, involved separately or jointly in human development and metabolism, to affect cancer occurrence *via* inhibition of let-7 miRNA (90). Yang et al. enrolled 275 hepatoblastoma cases and 1,018 healthy controls to prove *LIN28B* SNPs (rs94904590 T>G and rs314276 C>A) could increase the risk of hepatoblastoma (61). The *LIN28A* SNP (rs3811464 G>A) in hepatoblastoma affected hepatoblastoma in a low-penetrating manner because the significant result was only found in the stratified analysis (91). Four *LIN28A* SNPs (rs3811464 G>A, rs3811463 T>C, rs34787247 G>A, and rs11247957 G>A) were analyzed in hepatoblastoma, neuroblastoma, and Wilms tumor. Nevertheless, the association between the same SNPs with different malignant tumors is diverse (91–93). Interestingly, these findings suggested that the effects of *LIN28A* polymorphisms were specific in a specific cancer.

HMGA2

HMGA2, a member of the high mobility group (HMG) proteins family, carries a typical functional sequence motif named AT-hooks (94). HMGA2 regulates gene transcription in a modification of chromatin construction way (95). Moreover, it is mainly expressed in embryonic stem cells during embryogenesis rather than in adult tissue cells (96). Many studies identified that aberrant HMGA2 expression is associated with diverse cancer (94). Li et al. detected that *HMGA2* rs968697 T>C polymorphism was related to hepatoblastoma susceptibility in Chinese children (52).

Base Excision Repair Pathway Genes

DNA damage is common in humans, and DNA repair systems maintain the stability and integrity of DNA. If the damaged DNA is not repaired, genomic instability may eventually evolve into tumorigenesis (97). The base excision repair (BER) pathway is a critical part of DNA repair systems (98). Zhuo et al. conducted a case-control study exploring the relationship between six BER pathway genes (*PARP1*, *hOGG1*, *FEN1*, *APEX1*, *LIG3*, and *XRCC1*) and hepatoblastoma. *hOGG1* gene rs293795 A>G was significantly correlated with hepatoblastoma risk (99).

Other Genes

As a tumor suppressor gene studied widely, *TP53* is located on human chromosome 17p13.1 and plays a vital role in apoptosis and tumorigenesis (100). Among genes, *TP53* has the highest pertinence with human tumors. Aberrant expression and dysfunction of *TP53* have been detected in various human tumor cases. The *TP53* rs1042522 C>G polymorphism leads to an amino acid alteration (Arg to Pro) and therefore influences the susceptibility of various malignant tumors (101–103). Our research group conducted two studies to explore the association between *TP53* rs1042522 C>G polymorphism and hepatoblastoma. The significant result could not be found in the first study, which enrolled 213 hepatoblastoma cases and 958 cancer-free controls (104). After enlarging the study subjects to 313 cases and 1,446 controls and adding the analysis of rs4938723 T>C of *miR-34b/c*, Liu et al. did not observe any significant result either (105). *CMYC* is also a critical oncogene and reported that the expression of c-Myc increased in hepatoblastoma tissue (106). However, Yang et al. conducted a study in Chinese children and the result showed that CMYC rs4645943 and rs2070583 polymorphisms were not correlated with hepatoblastoma risk (107).

N6-Methyladenosine (m^6A) Modification Genes

Mainly occurring on the N6-position of adenosine, m^6A as an invertible epigenetic modification is prevalent in various eucaryotes (108). m^6A prefers appearing in 3'untranslated regions (3'UTRs), around termination codons, and within long internal exons (109). Although m^6A modification does not disturb base paring or coding, it was reported that it is involved in various RNA metabolism, including RNA expression, alternative splicing, and export, and therefore plays a critical role in tumor occurrence and development (110). Cui et al. demonstrated that the majority of m^6A-related genes were overexpressed in hepatoblastoma tissues (111). However, at present, the research about m^6A-related gene polymorphism in hepatoblastoma is few.

m^6A proteins can be divided into three categories, namely, "writers", "erasers", and "readers", which have the function of adding, removing, and recognizing, respectively (112). Methyltransferase-like 3 (METTL3) is a catalytic enzyme that, combined with methyltransferase-like 14 (METTL14), becomes a heterocomplex; and Wilms' tumor 1–associated protein (WTAP) as an assistant protein interacts with this heterocomplex (113). *METTL3* was localized at nuclear speckles and was involved in mRNA splicing to regulate mRNA metabolism (110). Liu et al. demonstrated that high expression of *METTL3* is the main factor of the aberration of m^6A and thereby promotes hepatoblastoma growth *via* the Wnt/β-catenin pathway (114). Downregulation of *METTL14* was demonstrated to correlate with the prognosis of hepatocellular carcinoma (47). In contrast, the expression of *WTAP* was significantly upregulated in hepatocellular carcinoma (115). m^6A demethylases, known as "erasers", include obesity-associated protein (FTO) and alkylation repair homolog protein 5 (ALKBH5).

FTO and ALKBH5 are α-ketoglutarate (α-KG) and Fe (II)-dependent demethylases that remove the RNA m^{6}A modification (116). There are many kinds of "readers" enzymes including YTH domain-containing family (YTHDF1-3 and YTHDC1-2), HNRNPA2BFHU1, Mrb1, ELAVL1, IGF2BPs, and eIFs. "Writers", as binding proteins with the function of specific recognition, impact the gene expression after RNA transcription (117).

To our knowledge, only our research group investigated the relationship between m^{6}A modification core gene polymorphisms and hepatoblastoma risk. Several research results have been published. Enrolled 313 cases and 1,446 controls, a Chinese seven-center case-control study was conducted. The *WTAP* rs7766006 G>T was significantly correlated with reduced hepatoblastoma risk. Preliminary annotation revealed that *WTAP* mRNA levels were upregulated in the liver in those who carried the rs7766006 T genotype (62). The *YTHDF1* rs6090311 G allele was identified as a protective factor of hepatoblastoma, and expression quantitative trait loci (eQTL) analyses showed that rs6090311 A>G might affect the mRNA level (60). Similar studies detected the association between *YTHDC1* and *ALKBH5* polymorphism and hepatoblastoma. However, significant relationships between *YTHDC1* rs2293596 T>C and *ALKBH5* rs8400 G>A polymorphism and hepatoblastoma risk were observed in the subgroup of clinical stage III+IV in stratification analysis (118, 119), requiring further study to identify whether these SNPs correlate to the prognosis of hepatoblastoma. Results indicated that the m^{6}A gene SNPs might affect the m^{6}A modification and thereby influence the hepatoblastoma growth (**Figure 1**). Our research group has also conducted the study of *METTL3* and *METTL14* polymorphisms and hepatoblastoma risk, and these results await to be published.

Considering the importance of m^{6}A modification in malignant tumors, the studies of m^{6}A modification core gene polymorphisms and hepatoblastoma risk are still inadequate. Besides, the number of SNPs included in the research is not enough, and more SNPs are urging to be enrolled to analyze.

DISCUSSION AND FUTURE DIRECTIONS

The treatment approaches of hepatoblastoma mainly include chemotherapy, surgical resection, and liver transplantation. In order to avoid overtreatments and improve the efficiency of treatment, individualized approaches are needed to provide to patients with different conditions. According to the recommendation of International Childhood Liver Tumor Strategy Group (SIOPEL), patients with high risk are suggested to be given dose-dense cisplatin weekly (120). Up to now, only alpha-fetoprotein (AFP) is used as a biomarker in clinics (121). If more biomarkers are applied in clinical practice for early diagnosis, the survival rate could be vastly improved. We reviewed the development of molecular epidemiology of hepatoblastoma. Advances in genotyping technologies facilitate the measuring of polymorphisms in a mass of samples. The sample sizes are finite in previous studies. However, the sample sizes have been distinctly enlarged with the launching of multicenter studies.

There some limitations of the research that actually need to be addressed. In recent years, the study subjects mentioned above are all from the Han population. Therefore, the results could not be generalized to other ethnicities. The precise functional role of SNPs in hepatoblastoma still awaits to be explored. Besides, there are plenty of cancer-related gene polymorphisms waiting to be detected.

In the past decades, genome-wide association study (GWAS), which is a method of conducting high-throughput sequencing technologies to measure plenty of polymorphisms, has discovered abundant significant cancer-related loci to improve

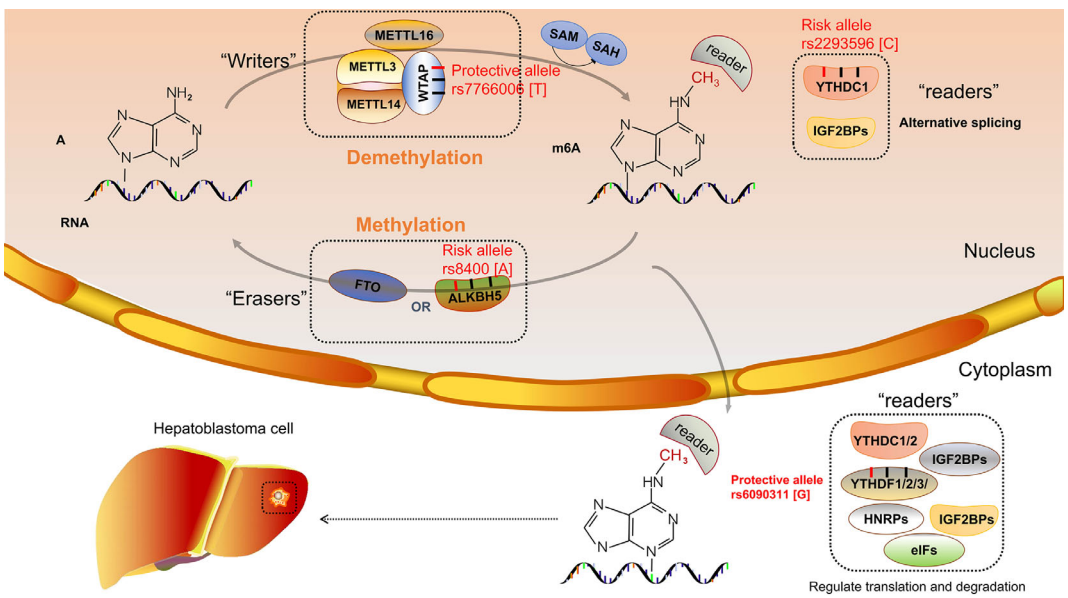


FIGURE 1 | N6-methyladenosine (m^{6}A) modification core gene polymorphisms and hepatoblastoma risk.

the methods of genetic research (122, 123). This technology was applied in other childhood solid tumors such as neuroblastoma and has determined multiple disease-related loci (124, 125). GWAS needs a larger sample to apply in multistages and multicenters to ensure the reliability of the results. Based on the result that abundant SNPs were detected by GWAS, it is challenging to seek out the actual cancer-related loci and explain their biological function in hepatoblastoma. Identifying SNPs with important functionalities could be applied in prenatal screening to diminish birth defects. There is no doubt that the application of GWAS in hepatoblastoma is necessary. It will contribute to figuring out the biomarkers of hepatoblastoma and guiding people to understand hepatoblastoma's etiology to improve the prevention and treatment of hepatoblastoma.

AUTHOR CONTRIBUTIONS

LM and ZZ: conceptualization and supervision. HC and QG: resources. HC: writing—original draft preparation. QG, HG, and LM: writing—review and editing. ZZ: project administration. LM and ZZ: funding acquisition. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by grants from the National Natural Science Foundation of China (nos. 82002636, 82002635) and China Postdoctoral Science Foundation (nos. 2020T130132, 2020M682668).

REFERENCES

1. Sia D, Villanueva A, Friedman SL, Llovet JM. Liver Cancer Cell of Origin, Molecular Class, and Effects On Patient Prognosis. *Gastroenterology* (2017) 152(4):745–61. doi: 10.1053/j.gastro.2016.11.048
2. Kalish JM, Doros L, Helman LJ, Hennekam RC, Kuiper RP, Maas SM, et al. Surveillance Recommendations for Children With Overgrowth Syndromes and Predisposition to Wilms Tumors and Hepatoblastoma. *Clin Cancer Res* (2017) 23(13):e115–22. doi: 10.1158/1078-0432.CCR-17-0710
3. Musick SR, Smith M, Rouster AS, Babiker HM. Hepatoblastoma. In: *StatPearls*. Treasure Island (FL: StatPearls Publishing Copyright © 2021, StatPearls Publishing LLC (2021).
4. Feng J, Polychronidis G, Heger U, Frongia G, Mehrabi A, Hoffmann K. Incidence Trends and Survival Prediction of Hepatoblastoma in Children: A Population-Based Study. *Cancer Commun (London England)* (2019) 39 (1):62. doi: 10.1186/s40880-019-0411-7
5. Sivaprakasam P, Gupta AA, Greenberg ML, Capra M, Nathan PC. Survival and Long-Term Outcomes in Children With Hepatoblastoma Treated With Continuous Infusion of Cisplatin and Doxorubicin. *J Pediatr Hematol/Oncol* (2011) 33(6):e226–30. doi: 10.1097/MPH.0b013e31821f0eaf
6. Lake CM, Tiao GM, Bondoc AJ. Surgical Management of Locally-Advanced and Metastatic Hepatoblastoma. *Semin Pediatr Surg* (2019) 28(6):150856. doi: 10.1016/j.sempedsurg.2019.150856
7. Wiwanitkit V. Hepatitis Virus B is Not a Risk Factor in Hepatoblastoma Patients. *Asian Pacific J Cancer Prev APJCP* (2005) 6(2):213–4.
8. Ding SF, Michail NE, Habib NA. Genetic Changes in Hepatoblastoma. *J Hepatol* (1994) 20(5):672–5. doi: 10.1016/s0168-8278(05)80359-9
9. Spector LG, Birch J. The Epidemiology of Hepatoblastoma. *Pediatr Blood Cancer* (2012) 59(5):776–9. doi: 10.1002/pbc.24215
10. Hafberg E, Borinstein SC, Alexopoulos SP. Contemporary Management of Hepatoblastoma. *Curr Opin Organ Transplant* (2019) 24:113–7. doi: 10.1097/MOT.0000000000000618
11. Maris JM. Recent Advances in Neuroblastoma. *New Engl J Med* (2010) 362 (23):2202–11. doi: 10.1056/NEJMra0804577
12. Mahamdallie S, Yost S, Poyastro-Pearson E, Holt E, Zachariou A, Seal S, et al. Identification of New Wilms Tumour Predisposition Genes: An Exome Sequencing Study. *Lancet Child Adolesc Health* (2019) 3(5):322–31. doi: 10.1016/s2352-4642(19)30018-5
13. Linabery AM, Ross JA. Trends in Childhood Cancer Incidence in the U.S. (1992-2004). *Cancer* (2008) 112(2):416–32. doi: 10.1002/cncr.23169
14. Bao PP, Li K, Wu CX, Huang ZZ, Wang CF, Xiang YM, et al. [Recent Incidences and Trends of Childhood Malignant Solid Tumors in Shanghai, 2002-2010]. *Zhonghua er ke za zhi = Chin J Pediatr* (2013) 51(4):288–94.
15. de Fine Licht S, Schmidt LS, Rod NH, Schmiegelow K, Lähteenmäki PM, Kogner P, et al. Hepatoblastoma in the Nordic Countries. *Int J Cancer* (2012) 131(4):E555–61. doi: 10.1002/ijc.27351
16. Friedrich P, Itriago E, Rodriguez-Galindo C, Ribeiro K. Racial and Ethnic Disparities in the Incidence of Pediatric Extracranial Embryonal Tumors. *J Natl Cancer Inst* (2017) 109(10). doi: 10.1093/jnci/djx050
17. Kehm RD, Spector LG, Poynter JN, Vock DM, Osypuk TL. Socioeconomic Status and Childhood Cancer Incidence: A Population-Based Multilevel Analysis. *Am J Epidemiol* (2018) 187(5):982–91. doi: 10.1093/aje/kwx322
18. Calvisi DF, Solinas A. Hepatoblastoma: Current Knowledge and Promises From Preclinical Studies. *Trans Gastroenterol Hepatol* (2020) 5:42. doi: 10.21037/tgh.2019.12.03
19. Trobaugh-Lotrario AD, López-Terrada D, Li P, Feusner JH. Hepatoblastoma in Patients With Molecularly Proven Familial Adenomatous Polyposis: Clinical Characteristics and Rationale for Surveillance Screening. *Pediatr Blood Cancer* (2018) 65(8):e27103. doi: 10.1002/pbc.27103
20. Hamada Y, Takada K, Fukunaga S, Hioki K. Hepatoblastoma Associated With Beckwith-Wiedemann Syndrome and Hemihypertrophy. *Pediatr Surg Int* (2003) 19(1-2):112–4. doi: 10.1007/s00383-002-0734-2
21. Pereira EM, Marion R, Ramesh KH, Kim JS, Ewart M, Ricafort R. Hepatoblastoma in a Mosaic Trisomy 18 Patient. *J Pediatr Hematol/Oncol* (2012) 34(4):e145–8. doi: 10.1097/MPH.0b013e3182459ee8
22. Kingston JE, Herbert A, Draper GJ, Mann JR. Association Between Hepatoblastoma and Polyposis Coli. *Arch Dis Childhood* (1983) 58 (12):959–62. doi: 10.1136/adc.58.12.959
23. Giardiello FM, Petersen GM, Brensinger JD, Luce MC, Cayouette MC, Bacon J, et al. Hepatoblastoma and APC Gene Mutation in Familial Adenomatous Polyposis. *Gut* (1996) 39(6):867–9. doi: 10.1136/gut.39.6.867
24. Fodde R. The APC Gene in Colorectal Cancer. *Eur J Cancer* (2002) 38 (7):867–71. doi: 10.1016/s0959-8049(02)00040-0
25. Fang DC, Luo YH, Yang SM, Li XA, Ling XL, Fang L. Mutation Analysis of APC Gene in Gastric Cancer With Microsatellite Instability. *World J Gastroenterol* (2002) 8(5):787–91. doi: 10.3748/wjg.v8.i5.787
26. Harvey J, Clark S, Hyer W, Hadzic N, Tomlinson I, Hinds R. Germline APC Mutations are Not Commonly Seen in Children With Sporadic Hepatoblastoma. *J Pediatr Gastroenterol Nutr* (2008) 47(5):675–7. doi: 10.1097/MPG.0b013e318174e808
27. Kim SY, Jung SH, Kim MS, Han MR, Park HC, Jung ES, et al. Genomic Profiles of a Hepatoblastoma From a Patient With Beckwith-Wiedemann Syndrome With Uniparental Disomy on Chromosome 11p15 and Germline Mutation of APC and PALB2. *Oncotarget* (2017) 8(54):91950–7. doi: 10.18632/oncotarget.20515
28. Weksberg R, Shuman C, Beckwith JB, Weksberg R. Beckwith-Wiedemann Syndrome. *Eur J Hum Genet* (2010) 18(1):1–14. doi: 10.1038/ejhg.2009.106
29. DeBaun MR, Tucker MA. Risk of Cancer During the First Four Years of Life in Children From The Beckwith-Wiedemann Syndrome Registry. *J Pediatr* (1998) 132(3 Pt 1):398–400. doi: 10.1016/s0022-3476(98)70008-3
30. Wang TJ, Li YY, Wu WJ, Lin CK, Wang CK, Wang CY, et al. Dandy-Walker Syndrome With Duplex Kidney Abnormalities in Trisomy 18 - A Rare Case Report. *Taiwanese J Obstetrics Gynecol* (2017) 56(5):697–9. doi: 10.1016/j.tjog.2017.08.022
31. Ahmad N, Wheeler K, Stewart H, Campbell C. Hepatoblastoma in a Mosaic Trisomy 18 Child With Hemihypertrophy. *BMJ Case Rep* (2016) 2016. doi: 10.1136/bcr-2015-211380

32. Inoue A, Suzuki R, Urabe K, Kawamura Y, Masuda M, Kishi K, et al. Therapeutic Experience With Hepatoblastoma Associated With Trisomy 18. *Pediatr Blood Cancer* (2018) 65(8):e27093. doi: 10.1002/pbc.27093
33. Lucas DJ, Rubinstein J, Gosain A, Tiao G, Head T, Pratap JN, et al. Surgical and Anesthetic Management for Hepatectomy in Two Pediatric Patients With Trisomy 18, Pulmonary Hypertension, and Hepatoblastoma. *Pediatr Blood Cancer* (2019) 66(6):e27678. doi: 10.1002/pbc.27678
34. Tomlinson GE, Kappler R. Genetics and Epigenetics of Hepatoblastoma. *Pediatr Blood Cancer* (2012) 59(5):785–92. doi: 10.1002/pbc.24213
35. Ozawa MG, Cooney T, Rangaswami A, Hazard FK. Synchronous Hepatoblastoma, Neuroblastoma, and Cutaneous Capillary Hemangiomas: A Case Report. *Pediatr Dev Pathol* (2016) 19(1):74–9. doi: 10.2350/14-11-1573-cr.1
36. Takayasu H, Motoi T, Kanamori Y, Kitano Y, Nakanishi H, Tange T, et al. Two Case Reports of Childhood Liver Cell Adenomas Harboring Beta-Catenin Abnormalities. *Hum Pathol* (2002) 33(8):852–5. doi: 10.1053/hupa.2002.125771
37. Mateos ME, Beyer K, López-Laso E, Siles JL, Pérez-Navero JL, Peña MJ, et al. Simpson-Golabi-Behmel Syndrome Type 1 and Hepatoblastoma in a Patient With a Novel Exon 2-4 Duplication of the GPC3 Gene. *Am J Med Genet Part A* (2013) 161a(5):1091–5. doi: 10.1002/ajmg.a.35738
38. Aretz S, Koch A, Uhlhaas S, Friedl W, Propping P, von Schweinitz D, et al. Should Children at Risk for Familial Adenomatous Polyposis be Screened for Hepatoblastoma and Children With Apparently Sporadic Hepatoblastoma be Screened for APC Germline Mutations? *Pediatr Blood Cancer* (2006) 47 (6):811–8. doi: 10.1002/pbc.20698
39. Pang D, McNally R, Birch JM. Parental Smoking and Childhood Cancer: Results From the United Kingdom Childhood Cancer Study. *Br J Cancer* (2003) 88(3):373–81. doi: 10.1038/sj.bjc.6600774
40. Pu CL, Guo CB, Jin XQ, Deng C, Zhang MM, Li YC, et al. [Retrospective Analysis of Maternal and Infant Birth Features of Hepatoblastoma Patients]. *Zhonghua gan zang bing za zhi = Zhonghua Ganzangbing Zazhi = Chin J Hepatol* (2009) 17(6):459–61.
41. Sorahan T, Lancashire RJ. Parental Cigarette Smoking and Childhood Risks of Hepatoblastoma: OSCC Data. *Br J Cancer* (2004) 90(5):1016–8. doi: 10.1038/sj.bjc.6601651
42. Reynolds P, Urayama KY, Von Behren J, Feusner J. Birth Characteristics and Hepatoblastoma Risk in Young Children. *Cancer* (2004) 100(5):1070–6. doi: 10.1002/cncr.20061
43. Tanimura M, Matsui I, Abe J, Ikeda H, Kobayashi N, Ohira M, et al. Increased Risk of Hepatoblastoma Among Immature Children With a Lower Birth Weight. *Cancer Res* (1998) 58(14):3032–5.
44. Janitz AE, Ramachandran G, Tomlinson GE, Krailo M, Richardson M, Spector L. Maternal and Paternal Occupational Exposures and Hepatoblastoma: Results From the HOPE Study Through the Children's Oncology Group. *J Exposure Sci Environ Epidemiol* (2017) 27(4):359–64. doi: 10.1038/jes.2017.1
45. Volk J, Heck JE, Schmiegelow K, Hansen J. Parental Occupational Organic Dust Exposure and Selected Childhood Cancers in Denmark 1968-2016. *Cancer Epidemiol* (2020) 65:101667. doi: 10.1016/j.canep.2020.101667
46. Buckley JD, Sather H, Ruccione K, Rogers PC, Haas JE, Henderson BE, et al. A Case-Control Study of Risk Factors for Hepatoblastoma. A Report From the Childrens Cancer Study Group. *Cancer* (1989) 64(5):1169–76. doi: 10.1002/1097-0142(19890901)64:5<1169::aid-cncr2820640534>3.0.co;2-i
47. Ma JZ, Yang F, Zhou CC, Liu F, Yuan JH, Wang F, et al. METTL14 Suppresses the Metastatic Potential of Hepatocellular Carcinoma by Modulating N(6) -Methyladenosine-Dependent Primary MicroRNA Processing. *Hepatology* (2017) 65(2):529–43. doi: 10.1002/hep.28885
48. Bläker H, Hofmann WJ, Rieker RJ, Penzel R, Graf M, Otto HF. Beta-Catenin Accumulation and Mutation of the CTNNB1 Gene in Hepatoblastoma. *Genes Chromosomes Cancer* (1999) 25(4):399–402. doi: 10.1002/(SICI)1098-2264(199908)25:4<399::AID-GCC14>3.0.CO;2-X
49. Shen G, Shen H, Zhang J, Yan Q, Liu H. DNA Methylation in Hepatoblastoma-A Literature Review. *Ital J Pediatr* (2020) 46(1):113. doi: 10.1186/s13052-020-00877-6
50. Shah T, Joshi K. Analysis of FOXO3a Gene Polymorphism Associated With Asthma. *Methods Mol Biol (Clifton NJ)* (2019) 1890:259–66. doi: 10.1007/978-1-4939-8900-3_22
51. Liu X, Han Z, Yang C. Associations of microRNA Single Nucleotide Polymorphisms and Disease Risk and Pathophysiology. *Clin Genet* (2017) 92(3):235–42. doi: 10.1111/cge.12950
52. Yang S, Zheng Y, Zhou L, Jin J, Deng Y, Yao J, et al. miR-499 Rs3746444 and miR-196a-2 Rs11614913 Are Associated With the Risk of Glioma, But Not the Prognosis. *Mol Ther Nucleic Acids* (2020) 22:340–51. doi: 10.1016/j.omtn.2020.08.038
53. Zhou L, Dong S, Deng Y, Yang P, Zheng Y, Yao L, et al. GOLGA7 Rs11337, a Polymorphism at the MicroRNA Binding Site, Is Associated With Glioma Prognosis. *Mol Ther Nucleic Acids* (2019) 18:56–65. doi: 10.1016/j.omtn.2019.08.006
54. Yang Y, Shu X, Shu XO, Bolla MK, Kweon SS, Cai Q, et al. Re-Evaluating Genetic Variants Identified in Candidate Gene Studies of Breast Cancer Risk Using Data From Nearly 280,000 Women of Asian and European Ancestry. *EBioMedicine* (2019) 48:203–11. doi: 10.1016/j.ebiom.2019.09.006
55. Kim S, Misra A. SNP Genotyping: Technologies and Biomedical Applications. *Annu Rev Biomed Eng* (2007) 9:289–320. doi: 10.1146/annurev.bioeng.9.060906.152037
56. Pakakasama S, Chen TT, Frawley W, Muller C, Douglass EC, Tomlinson GE. Myeloperoxidase Promotor Polymorphism and Risk of Hepatoblastoma. *Int J Cancer* (2003) 106(2):205–7. doi: 10.1002/ijc.11191
57. Pakakasama S, Chen TT, Frawley W, Muller CY, Douglass EC, Lee R, et al. CCND1 Polymorphism and Age of Onset of Hepatoblastoma. *Oncogene* (2004) 23:4789–92. doi: 10.1038/sj.onc.1207499
58. Yang T, Li J, Wen Y, Tan T, Yang J, Pan J, et al. LINC00673 Rs11655237 C>T Polymorphism Impacts Hepatoblastoma Susceptibility in Chinese Children. *Front Genet* (2019) 10:506. doi: 10.3389/fgene.2019.00506
59. Tan T, Li J, Wen Y, Zou Y, Yang J, Pan J, et al. Association Between lncRNA-H19 Polymorphisms and Hepatoblastoma Risk in an Ethic Chinese Population. *J Cell Mol Med* (2021) 25(2):742–50. doi: 10.1111/jcmm.16124
60. Luo Z, Li G, Wang M, Zhu J, Yang Z, Li Y, et al. YTHDF1 Rs6090311 a>G Polymorphism Reduces Hepatoblastoma Risk: Evidence From a Seven-Center Case-Control Study. *J Cancer* (2020) 11(17):5129–34. doi: 10.7150/jca.46120
61. Yang Z, Deng Y, Zhang K, Bai Y, Zhu J, Zhang J, et al. LIN28B Gene Polymorphisms Modify Hepatoblastoma Susceptibility in Chinese Children. *J Cancer* (2020) 11(12):3512–8. doi: 10.7150/jca.42798
62. Zhuo ZJ, Hua RX, Chen Z, Zhu J, Wang M, Yang Z, et al. WTAP Gene Variants Confer Hepatoblastoma Susceptibility: A Seven-Center Case-Control Study. *Mol Ther Oncol* (2020) 18:118–25. doi: 10.1016/j.omto.2020.06.007
63. Hung RJ, Boffetta P, Brennan P, Malaveille C, Gelatti U, Placidi D, et al. Genetic Polymorphisms of MPO, COMT, MnSOD, NQO1, Interactions With Environmental Exposures and Bladder Cancer Risk. *Carcinogenesis* (2004) 25(6):973–8. doi: 10.1093/carcin/bgh080
64. Lin SC, Chou YC, Wu MH, Wu CC, Lin WY, Yu CP, et al. Genetic Variants of Myeloperoxidase and Catechol-O-Methyltransferase and Breast Cancer Risk. *Eur J Cancer Prev* (2005) 14(3):257–61. doi: 10.1097/00008469-200506000-00010
65. Shi X, Li B, Yuan Y, Chen L, Zhang Y, Yang M, et al. The Possible Association Between the Presence of an MPO -463 G > A (Rs2333227) Polymorphism and Cervical Cancer Risk. *Pathol Res Pract* (2018) 214 (8):1142–8. doi: 10.1016/j.prp.2018.05.018
66. Kantarci OH, Lesnick TG, Yang P, Meyer RL, Hebrink DD, McMurray CT, et al. Myeloperoxidase -463 (G–>A) Polymorphism Associated With Lower Risk of Lung Cancer. *Mayo Clin Proc* (2002) 77(1):17–22. doi: 10.4065/77.1.17
67. Ohnishi S, Murata M, Kawanishi S. DNA Damage Induced by Hypochlorite and Hypobromite With Reference to Inflammation-Associated Carcinogenesis. *Cancer Lett* (2002) 178(1):37–42. doi: 10.1016/s0304-3835(01)00812-6
68. Govatati S, Singamsetty GK, Nallabelli N, Malempati S, Rao PS, Madamchetty VK, et al. Contribution of Cyclin D1 (CCND1) and E-Cadherin (CDH1) Alterations to Colorectal Cancer Susceptibility: A Case-Control Study. *Tumour Biol* (2014) 35(12):12059–67. doi: 10.1007/s13277-014-2505-9
69. Takayasu H, Horie H, Hiyama E, Matsunaga T, Hayashi Y, Watanabe Y, et al. Frequent Deletions and Mutations of the Beta-Catenin Gene are

Associated With Overexpression of Cyclin D1 and Fibronectin and Poorly Differentiated Histology in Childhood Hepatoblastoma. *Clin Cancer Res* (2001) 7(4):901–8.
70. Yang T, Wen Y, Li J, Tan T, Yang J, Pan J, et al. NRAS and KRAS Polymorphisms are Not Associated With Hepatoblastoma Susceptibility in Chinese Children. *Exp Hematol Oncol* (2019) 8:11. doi: 10.1186/s40164-019-0135-z
71. Chen S, Li F, Xu D, Hou K, Fang W, Li Y. The Function of RAS Mutation in Cancer and Advances in its Drug Research. *Curr Pharm Design* (2019) 25 (10):1105–14. doi: 10.2174/1381612825666190506122228
72. Meng L, Liu S, Liu F, Sang M, Ju Y, Fan X, et al. ZEB1-Mediated Transcriptional Upregulation of Circwwc3 Promotes Breast Cancer Progression Through Activating Ras Signaling Pathway. *Mol Ther Nucleic Acids* (2020) 22:124–37. doi: 10.1016/j.omtn.2020.08.015
73. Wang H, Lv Q, Xu Y, Cai Z, Zheng J, Cheng X, et al. An Integrative Pharmacogenomics Analysis Identifies Therapeutic Targets in KRAS-Mutant Lung Cancer. *EBioMedicine* (2019) 49:106–17. doi: 10.1016/j.ebiom.2019.10.012
74. Esteller M. Non-Coding RNAs in Human Disease. *Nat Rev Genet* (2011) 12 (12):861–74. doi: 10.1038/nrg3074
75. Gong J, Liu W, Zhang J, Miao X, Guo AY. lncRNASNP: A Database of SNPs in lncRNAs and Their Potential Functions in Human and Mouse. *Nucleic Acids Res* (2015) 43(Database issue):D181–6. doi: 10.1093/nar/gku1000
76. Tang JY, Lee JC, Chang YT, Hou MF, Huang HW, Liaw CC, et al. Long Noncoding RNAs-Related Diseases, Cancers, and Drugs. *Scientific WorldJournal* (2013) 2013:943539. doi: 10.1155/2013/943539
77. Tan YT, Lin JF, Li T, Li JJ, Xu RH, Ju HQ. LncRNA-Mediated Posttranslational Modifications and Reprogramming of Energy Metabolism in Cancer. *Cancer Commun (London England)* (2021) 41 (2):109–20. doi: 10.1002/cac2.12108
78. Yang Y, Deng X, Li Q, Wang F, Miao L, Jiang Q. Emerging Roles of Long Noncoding RNAs in Cholangiocarcinoma: Advances and Challenges. *Cancer Commun (London England)* (2020) 40(12):655–80. doi: 10.1002/cac2.12109
79. Song H, Liu Y, Li X, Chen S, Xie R, Chen D, et al. Long Noncoding RNA CASC11 Promotes Hepatocarcinogenesis and HCC Progression Through EIF4A3-Mediated E2F1 Activation. *Clin Transl Med* (2020) 10(7):e220. doi: 10.1002/ctm2.220
80. Hua RX, Liu J, Fu W, Zhu J, Zhang J, Cheng J, et al. ALKBH5 Gene Polymorphisms and Wilms Tumor Risk in Chinese Children: A Five-Center Case-Control Study. *J Clin Lab Anal* (2020) 34(6):e23251. doi: 10.1002/jcla.23251
81. Qiao K, Ning S, Wan L, Wu H, Wang Q, Zhang X, et al. LINC00673 is Activated by YY1 and Promotes the Proliferation of Breast Cancer Cells *via* the miR-515-5p/MARK4/Hippo Signaling Pathway. *J Exp Clin Cancer Res CR* (2019) 38(1):418. doi: 10.1186/s13046-019-1421-7
82. Zheng J, Huang X, Tan W, Yu D, Du Z, Chang J, et al. Pancreatic Cancer Risk Variant in LINC00673 Creates a miR-1231 Binding Site and Interferes With PTPN11 Degradation. *Nat Genet* (2016) 48(7):747–57. doi: 10.1038/ng.3568
83. Lu W, Zhang H, Niu Y, Wu Y, Sun W, Li H, et al. Long non-Coding RNA Linc00673 Regulated non-Small Cell Lung Cancer Proliferation, Migration, Invasion and Epithelial Mesenchymal Transition by Sponging miR-150-5p. *Mol Cancer* (2017) 16(1):118. doi: 10.1186/s12943-017-0685-9
84. Childs EJ, Mocci E, Campa D, Bracci PM, Gallinger S, Goggins M, et al. Common Variation at 2p13.3, 3q29, 7p13 and 17q25.1 Associated With Susceptibility to Pancreatic Cancer. *Nat Genet* (2015) 47(8):911–6. doi: 10.1038/ng.3341
85. Verhaegh GW, Verkleij L, Vermeulen SH, den Heijer M, Witjes JA, Kiemeney LA. Polymorphisms in the H19 Gene and the Risk of Bladder Cancer. *Eur Urol* (2008) 54(5):1118–26. doi: 10.1016/j.eururo.2008.01.060
86. Tanos V, Prus D, Ayesh S, Weinstein D, Tykocinski ML, De-Groot N, et al. Expression of the Imprinted H19 Oncofetal RNA in Epithelial Ovarian Cancer. *Eur J Obstetrics Gynecol Reprod Biol* (1999) 85(1):7–11. doi: 10.1016/s0301-2115(98)00275-9
87. Lecerf C, Le Bourhis X, Adriaenssens E. The Long non-Coding RNA H19: An Active Player With Multiple Facets to Sustain the Hallmarks of Cancer. *Cell Mol Life Sci CMLS* (2019) 76(23):4673–87. doi: 10.1007/s00018-019-03240-z
88. Li S, Hua Y, Jin J, Wang H, Du M, Zhu L, et al. Association of Genetic Variants in lncRNA H19 With Risk of Colorectal Cancer in a Chinese Population. *Oncotarget* (2016) 7(18):25470–7. doi: 10.18632/oncotarget.8330
89. Wang H, Zhao Q, Deng K, Guo X, Xia J. Lin28: An Emerging Important Oncogene Connecting Several Aspects of Cancer. *Tumour Biol* (2016) 37 (3):2841–8. doi: 10.1007/s13277-015-4759-2
90. Balzeau J, Menezes MR, Cao S, Hagan JP. The LIN28/let-7 Pathway in Cancer. *Front Genet* (2017) 8:31. doi: 10.3389/fgene.2017.00031
91. Yang Z, Deng Y, Zhang K, Bai Y, Zhu J, Zhang J, et al. LIN28A Polymorphisms and Hepatoblastoma Susceptibility in Chinese Children. *J Cancer* (2021) 12(5):1373–8. doi: 10.7150/jca.52621
92. Zhuo Z, Fu W, Liu J, Cheng J, Zhou H, Zhang J, et al. LIN28A Gene Polymorphisms Confer Wilms Tumour Susceptibility: A Four-Centre Case-Control Study. *J Cell Mol Med* (2019) 23(10):7105–10. doi: 10.1111/jcmm.14561
93. Hua RX, Zhuo Z, Ge L, Zhu J, Yuan L, Chen C, et al. LIN28A Gene Polymorphisms Modify Neuroblastoma Susceptibility: A Four-Centre Case-Control Study. *J Cell Mol Med* (2020) 24(1):1059–66. doi: 10.1111/jcmm.14827
94. Zhang BH, Yan LN, Yang JY. Pending Role of METTL14 in Liver Cancer. *Hepatobil Surg Nutr* (2019) 8(6):669–70. doi: 10.21037/hbsn.2019.10.16
95. Fedele M, Battista S, Kenyon L, Baldassarre G, Fidanza V, Klein-Szanto AJ, et al. Overexpression of the HMGA2 Gene in Transgenic Mice Leads to the Onset of Pituitary Adenomas. *Oncogene* (2002) 21(20):3190–8. doi: 10.1038/sj.onc.1205428
96. Zhou X, Benson KF, Ashar HR, Chada K. Mutation Responsible for the Mouse Pygmy Phenotype in the Developmentally Regulated Factor HMGI-C. *Nature* (1995) 376(6543):771–4. doi: 10.1038/376771a0
97. Kumar N, Raja S, Van Houten B. The Involvement of Nucleotide Excision Repair Proteins in the Removal of Oxidative DNA Damage. *Nucleic Acids Res* (2020) 48(20):11227–43. doi: 10.1093/nar/gkaa777
98. Liu Y, Wilson SH. DNA Base Excision Repair: A Mechanism of Trinucleotide Repeat Expansion. *Trends Biochem Sci* (2012) 37(4):162–72. doi: 10.1016/j.tibs.2011.12.002
99. Zhuo Z, Lin A, Zhang J, Chen H, Li Y, Yang Z, et al. Genetic Variations in Base Excision Repair Pathway Genes and Risk of Hepatoblastoma: A Seven-Center Case-Control Study. *Am J Cancer Res* (2021) 11(3):849–57.
100. Xu H, El-Gewely MR. P53-Responsive Genes and the Potential for Cancer Diagnostics and Therapeutics Development. *Biotechnol Annu Rev* (2001) 7:131–64. doi: 10.1016/s1387-2656(01)07035-1
101. Icen-Taskin I, Irtegun-Kandemir S, Munzuroglu O. TP53 Rs1042522 Polymorphism and Early-Onset Breast Cancer. *J Res Med Sci* (2020) 25:25. doi: 10.4103/jrms.JRMS_506_19
102. Li Z, Weng H, Su R, Weng X, Zuo Z, Li C, et al. FTO Plays an Oncogenic Role in Acute Myeloid Leukemia as a N(6)-Methyladenosine RNA Demethylase. *Cancer Cell* (2017) 31:127–41. doi: 10.1016/j.ccell.2016.11.017
103. Kamiza AB, Kamiza S, Singini MG, Mathew CG. Association of TP53 Rs1042522 With Cervical Cancer in the Sub-Saharan African Population: A Meta-Analysis. *Trop Med Int Health TM IH* (2020) 25(6):666–72. doi: 10.1111/tmi.13397
104. Yang T, Wen Y, Li J, Yang J, Tan T, Pan J, et al. Association of the TP53 Rs1042522 C>G Polymorphism and Hepatoblastoma Risk in Chinese Children. *J Cancer* (2019) 10(15):3444–9. doi: 10.7150/jca.33063
105. Liu P, Zhuo ZJ, Zhu J, Yang Z, Xin Y, Li S, et al. Association of TP53 Rs1042522 C>G and miR-34b/C Rs4938723 T>C Polymorphisms With Hepatoblastoma Susceptibility: A Seven-Center Case-Control Study. *J Gene Med* (2020) 22(7):e3182. doi: 10.1002/jgm.3182
106. Ranganathan S, Tan X, Monga SP. Beta-Catenin and Met Deregulation in Childhood Hepatoblastomas. *Pediatr Dev Pathol* (2005) 8(4):435–47. doi: 10.1007/s10024-005-0028-5
107. Yang T, Wen Y, Li J, Tan T, Yang J, Pan J, et al. Association of CMYC Polymorphisms With Hepatoblastoma Risk. *Transl Cancer Res* (2020) 9 (2):849–55. doi: 10.21037/tcr.2019.12.19

108. Yue Y, Liu J, He C. RNA N6-Methyladenosine Methylation in Post-Transcriptional Gene Expression Regulation. *Genes Dev* (2015) 29 (13):1343–55. doi: 10.1101/gad.262766.115
109. Ke S, Alemu EA, Mertens C, Gantman EC, Fak JJ, Mele A, et al. A Majority of M6a Residues are in the Last Exons, Allowing the Potential for 3' UTR Regulation. *Genes Dev* (2015) 29(19):2037–53. doi: 10.1101/gad.269415.115
110. Yang Y, Hsu PJ, Chen YS, Yang YG. Dynamic Transcriptomic M(6)A Decoration: Writers, Erasers, Readers and Functions in RNA Metabolism. *Cell Res* (2018) 28(6):616–24. doi: 10.1038/s41422-018-0040-8
111. Cui X, Wang Z, Li J, Zhu J, Ren Z, Zhang D, et al. Cross Talk Between RNA N6-Methyladenosine Methyltransferase-Like 3 and miR-186 Regulates Hepatoblastoma Progression Through Wnt/β-Catenin Signalling Pathway. *Cell Proliferation* (2020) 53(3):e12768. doi: 10.1111/cpr.12768
112. He PC, He C. M(6) A RNA Methylation: From Mechanisms to Therapeutic Potential. *EMBO J* (2021) 40(3):e105977. doi: 10.15252/embj.2020105977
113. Zaccara S, Ries RJ, Jaffrey SR. Reading, Writing and Erasing mRNA Methylation. *Nat Rev Mol Cell Biol* (2019) 20(10):608–24. doi: 10.1038/s41580-019-0168-5
114. Liu L, Wang J, Sun G, Wu Q, Ma J, Zhang X, et al. M(6)A mRNA Methylation Regulates CTNNB1 to Promote the Proliferation of Hepatoblastoma. *Mol Cancer* (2019) 18(1):188. doi: 10.1186/s12943-019-1119-7
115. Chen Y, Peng C, Chen J, Chen D, Yang B, He B, et al. WTAP Facilitates Progression of Hepatocellular Carcinoma *via* M6a-HuR-Dependent Epigenetic Silencing of ETS1. *Mol Cancer* (2019) 18(1):127. doi: 10.1186/s12943-019-1053-8
116. Gerken T, Girard CA, Tung YC, Webby CJ, Saudek V, Hewitson KS, et al. The Obesity-Associated FTO Gene Encodes a 2-Oxoglutarate-Dependent Nucleic Acid Demethylase. *Sci (New York NY)* (2007) 318(5855):1469–72. doi: 10.1126/science.1151710
117. Zhao Y, Shi Y, Shen H, Xie W. M(6)A-Binding Proteins: The Emerging Crucial Performers in Epigenetics. *J Hematol Oncol* (2020) 13(1):35. doi: 10.1186/s13045-020-00872-8
118. Chen H, Li Y, Li L, Zhu J, Yang Z, Zhang J, et al. YTHDC1 Gene Polymorphisms and Hepatoblastoma Susceptibility in Chinese Children: A Seven-Center Case-Control Study. *J Gene Med* (2020) 22(11):e3249. doi: 10.1002/jgm.3249
119. Ren H, Zhuo ZJ, Duan F, Li Y, Yang Z, Zhang J, et al. ALKBH5 Gene Polymorphisms and Hepatoblastoma Susceptibility in Chinese Children. *J Oncol* (2021) 2021:6658480. doi: 10.1155/2021/6658480
120. Zsiros J, Brugieres L, Brock P, Roebuck D, Maibach R, Zimmermann A, et al. Dose-Dense Cisplatin-Based Chemotherapy and Surgery for Children With High-Risk Hepatoblastoma (SIOPEL-4): A Prospective, Single-Arm, Feasibility Study. *Lancet Oncol* (2013) 14(9):834–42. doi: 10.1016/s1470-2045(13)70272-9
121. Maibach R, Roebuck D, Brugieres L, Capra M, Brock P, Dall'Igna P, et al. Prognostic Stratification for Children With Hepatoblastoma: The SIOPEL Experience. *Eur J Cancer (Oxford Engl 1990)* (2012) 48(10):1543–9. doi: 10.1016/j.ejca.2011.12.011
122. Kinnersley B, Houlston RS, Bondy ML. Genome-Wide Association Studies in Glioma. *Cancer Epidemiol Biomarkers Prev* (2018) 27(4):418–28. doi: 10.1158/1055-9965.Epi-17-1080
123. Benafif S, Kote-Jarai Z. Eeles RA. A Review of Prostate Cancer Genome-Wide Association Studies (GWAS). *Cancer Epidemiol Biomarkers Prev* (2018) 27(8):845–57. doi: 10.1158/1055-9965.Epi-16-1046
124. He J, Zou Y, Wang T, Zhang R, Yang T, Zhu J, et al. Genetic Variations of GWAS-Identified Genes and Neuroblastoma Susceptibility: A Replication Study in Southern Chinese Children. *Trans Oncol* (2017) 10(6):936–41. doi: 10.1016/j.tranon.2017.09.008
125. Nguyen le B, Diskin SJ, Capasso M, Wang K, Diamond MA, Glessner J, et al. Phenotype Restricted Genome-Wide Association Study Using a Gene-Centric Approach Identifies Three Low-Risk Neuroblastoma Susceptibility Loci. *PloS Genet* (2011) 7(3):e1002026. doi: 10.1371/journal.pgen.1002026

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

ORIGINAL RESEARCH
published: 29 July 2021
doi: 10.3389/fonc.2021.666470

Targeted Next-Generation Sequencing of Circulating Tumor DNA, Bone Marrow, and Peripheral Blood Mononuclear Cells in Pediatric AML

Min Ruan[1†], Lipeng Liu[1†], Benquan Qi[1], Xiaoyan Chen[1], Lixian Chang[1], Aoli Zhang[1], Fang Liu[1], Shuchun Wang[1], Xiaoming Liu[1], Xiaojuan Chen[1], Li Zhang[1], Ye Guo[1], Yao Zou[1], Yingchi Zhang[1], Yumei Chen[1], LiXia Liu[3], Shanbo Cao[2], Feng Lou[2], Chengcheng Wang[3] and Xiaofan Zhu[1]*

[1] Division of Pediatric Blood Diseases Center, State Key Laboratory of Experimental Hematology, National Clinical Research Center for Blood Diseases, Institute of Hematology & Blood Diseases Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Tianjin, China, [2] Executive President Office, Acornmed Biotechnology Co., Ltd., Beijing, China, [3] Medical Department, Acornmed Biotechnology Co., Ltd., Beijing, China

Background: The aim of the study was to validate the diagnostic role of circulating tumor DNA (ctDNA) in genetics aberration on the basis of next-generation sequencing (NGS) in pediatric acute myeloid leukemia (AML).

Methods: Bone marrow (BM) and peripheral blood (PB) were collected from 20 AML children at the time of initial diagnosis, and a ctDNA sample was isolated from PB. Detection of mutation was performed on ctDNA, BM, and peripheral blood mononuclear cell (PBMC) by NGS based on a 185-gene panel.

Results: Among 185 genes sequenced by the NGS platform, a total of 82 abnormal genes were identified in 20 patients. Among them, 61 genes (74.39%) were detected in ctDNA, PBMC, and BM samples, while 11 (13.41%) genes were found only in ctDNA and 4 (4.88%) were detected only in the BM sample, and 2 (2.44%) were detected only in PBMC. A total of 239 mutations were detected in three samples, while 209 in ctDNA, 180 in bone marrow, and 184 in PBMC. One hundred sixty-four mutations in ctDNA were shared by matched BM samples, and the median variant allelic frequency (VAF) of these mutations was 41.34% (range, 0.55% to 99.96%) and 44.36% (range, 0.56% to 99.98%) in bone marrow and ctDNA. It was found that 65.79% (75/114) of mutations with clinical significance were detected in three samples, with 9 mutations detected both in ctDNA and BM, and 2 mutations detected both in PBMC and BM. The consistency of mutations with clinical significance between ctDNA and BM was 77.06% (84/109). Among the 84

OPEN ACCESS

Edited by:
Hua Tan,
University of Texas Health Science Center at Houston, United States

Reviewed by:
Xinhui Du,
Henan Provincial Cancer Hospital, China
Joseph Louis Lasky,
Cure 4 The Kids, United States

***Correspondence:**
Xiaofan Zhu
xfzhu@ihcams.ac.cn

[†]*These authors have contributed equally to this work and share first authorship*

Specialty section:
This article was submitted to Pediatric Oncology, a section of the journal Frontiers in Oncology

Received: *10 February 2021*
Accepted: *09 July 2021*
Published: *29 July 2021*

Citation:
Ruan M, Liu L, Qi B, Chen X, Chang L, Zhang A, Liu F, Wang S, Liu X, Chen X, Zhang L, Guo Y, Zou Y, Zhang Y, Chen Y, Liu L, Cao S, Lou F, Wang C and Zhu X (2021) Targeted Next-Generation Sequencing of Circulating Tumor DNA, Bone Marrow, and Peripheral Blood Mononuclear Cells in Pediatric AML. Front. Oncol. 11:666470. doi: 10.3389/fonc.2021.666470

mutations with clinical significance detected in both sources, the concordance of VAF assessment by both methods was high (R^2 = 0.895).

Conclusion: This study demonstrates that ctDNA was a reliable sample in pediatric AML and can be used for mutation detection. Consistency analysis showed that ctDNA can mirror the genomic information from BM. In addition, a subset of mutations was exclusively detected in ctDNA. These data support the fact that monitoring ctDNA with next-generation sequencing-based assays can provide more information about gene mutations to guide precision treatment in pediatric AML.

Keywords: acute myeloid leukemia, targeted next-generation sequencing, circulating tumor DNA, mutation (genetics), pediatric

INTRODUCTION

Acute myeloid leukemia (AML), which accounts for 25% of childhood acute leukemia, is a rapidly progressing hematopoietic malignancy characterized by the differentiation block and aberrant proliferation of leukemic blasts (1). In pediatric AML patients, the achieved 5-year overall survival (OS) is 60–70%, while the event-free survival (EFS) is 50% (2). With the development of molecular biology technology, the molecular landscape of pediatric AML becomes clearer (3, 4). Mutations in FLT3, TP53, NPM1, CEBPA, RUNX1, and ASXL1, which are common in AML children, have received more and more attention. And the clinical significance of new mutations, such as STAG2, RAD21, SRSF2, and U2AF1, have been gradually clarified. These diverse genomic molecular markers reflect the heterogeneity of AML, and accurate molecular profiling in AML is important for risk stratification and selection of targeted therapies (5).

Circulating tumor DNA (ctDNA), which is contained in circulating-free cell DNA (cfDNA) and released by necrosis or apoptosis tumor cell, allows for noninvasive peripheral blood sampling of cancer-associated mutations (6–9). When compared with other samples, ctDNA is more like a genomic library of different tumor cells and can mirror the heterogeneity of AML; moreover, ctDNA has a relatively short half-life, which may better reflect the latest status of the disease (10–12).

Nowadays, noninvasive detection of mutations by ctDNA was widely used in various solid tumors, but its role in hematological malignancies is still not clear. The current "gold standard" for molecular testing in pediatric AML is from bone marrow (BM) aspirate DNA. However, BM aspiration is an invasive procedure, which severely limited its application in clinical research. To date, there are very limited studies on the potential role of ctDNA, as a relatively non-invasive source, in monitoring leukemia-associated mutations and providing prognostic information in patients with hematologic malignancies (13, 14). Furthermore, it is still unknown whether ctDNA can fully supplant BM assessment for molecular profiling in pediatric AML.

Therefore, we aim to validate the diagnostic role of ctDNA in molecular profiles in pediatric AML patients, when compared with hybrid capture-targeted next-generation sequencing of BM, peripheral blood mononuclear cell (PBMC).

MATERIALS AND METHODS

Patients and Patient Specimens

For this prospective analysis, the source population included 20 children (age < 18 years) with AML at the Division of Pediatric Blood Diseases Center in Institute of Hematology and Blood Diseases Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College. BM and PBMC were collected for diagnostic purposes from all enrolled patients (excluding Down syndrome or acute promyelocytic leukemia, secondary AML), while ctDNA was isolated from the PB samples. The data collected included information regarding age, sex, peripheral blood white blood cell counts (WBC), blast percentages in BM and PBMC, chromosome karyotypes, and gene mutation signatures.

The study design and methods complied with the Declaration of Helsinki and were approved by the Ethics Committee and Institutional Review Board of Institute of Hematology and Blood Diseases Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College. Informed consent was obtained from all subjects. The raw sequence data reported in this paper have been deposited in the Genome Sequence Archive (Genomics, Proteomics & Bioinformatics 2017) in National Genomics Data Center (Nucleic Acids Res 2021), China National Center for Bioinformation/Beijing Institute of Genomics, Chinese Academy of Sciences, under accession number HRA000912, which are publicly accessible at https://ngdc.cncb.ac.cn/gsa-human.

Next-Generation Sequencing and Mutation Analysis

cfDNAs were extracted by a customized QIAamp Circulating Nucleic Acid kit (Qiagen GmbH) from 20 patient's PB samples at diagnosis, while DNA were extracted by a customized Genomic DNA kit (Qiagen GmbH) from the patient's BM and PBMC samples. Gene library amplification was based on a KAPA Hyper

Prep Kit. The gene panel from Acornmed Biotechnology was used to capture the target region. Detailed sequencing information is provided in **Table S1**.

Multiplexed libraries were sequenced with Illumina Novaseq and then analyzed for data including Sequencing mapping, coverage and quality assessment, Insertion/Deletion detection, annotation for sequence mutations: Average raw sequencing depth on target per sample ≥10000x(ctDNA)≥1000x (DNA), Allele mutation frequency ≥0.5% for single Nucleotide variation and insertion or deletion, respectively. All reads were filtered by high Mapping quality (≥30) and Base quality (≥30). The mutant reads were supported by positive and negative strands. Reads were aligned to the human genome using the Burrows-Wheeler Alignment tool (BWA, version 0.7.12). PCR duplicates were marked by the MarkDuplicates tool in Picard. IndelRealigner and BaseRecalibrator on Genome Analysis Toolkit (GATK; version 3.8) were used for the realignment and recalibration of the BWA alignment results, respectively. Mutect2 was used for identifying SNV and INDEL. We obtained candidate variations through background database filtering of normal samples. Pindel was used for detecting FLT3-ITD. FLT3-ITD quantitative analysis was performed by in-house tools based on machine learning development. All the variants were annotated by the ANNOVAR software using some resources, including 1000G projects, COSMIC, SIFT, and Polyphen. Our gene panel was mainly from NCCN guidelines, EMSO guidelines, authoritative databases, and literature reports of hematologic tumors.

Statistical Analysis

Patient characteristics were summarized using median (range) for continuous variables and frequencies (percentages) for categorical variables. The Fisher exact test was used to test the association between two categorical variables. Concordance of BM and ctDNA and PBMC results were assessed using Pearson correlation analysis. *P* values<0.05 were considered significant. All statistical tests were performed using SPSS 24.0 (IBM Corporation).

RESULTS

Baseline Characteristics

Twenty patients with newly diagnosed AML were evaluated. The baseline characteristics of the study cohort are shown in **Table 1**. Four patients (20%) had absolute PB blast count<1×10^9/L, and one patient had no peripheral blood circulating blasts. ctDNA, PBMC, and BM targeted sequencing were performed in all 20 patients at diagnosis simultaneously.

Detection of Molecular Profiles by Three Methods

The molecular profiles of all patients were detected by target-Next-generation Sequencing (t-NGS) (Acornmed Biotechnology Co., Ltd.), which covers the most frequent mutations in 137 genes in AML patients, *via* ctDNA, BM, and PBMC samples (**Table S1**).

TABLE 1 | Baseline characteristics of the study population.

Characteristic	Median [range] or n (%)
Sex	
Male	7 (35)
Female	13 (65)
Age, years	11 (1-15)
White blood cells, $\times10^9$/L	25.82 (1.54-182.5)
Peripheral blood blasts, $\times10^9$/L	12.41 (0-124.1)
Peripheral blood blasts, %	36 (0-98.0)
Bone marrow blasts, %	66.5 (21.0-90.0)
Cytogenetics	
t (8;21)	6 (30)
inv (16)	1 (5)
t (6;11)	1 (5)
+8	2 (10)
-7	1 (5)
Complex karyotype	1 (5)
Normal karyotype	8 (40)
FAB type	
M1	2 (10)
M2	8 (40)
M4	6 (30)
M5	3 (15)
M7	1 (5)

The sequencing depths of the three samples were all greater than 2000 X, namely 3460X (1837X-4270X) in ctDNA, 2530X (1633X-2862X) in BM, and 2324X (1208X-3720X) in PBMC.

A total of 82 abnormal genes were identified in 20 included patients. Among them, 61 genes (74.39%) were detected in ctDNA, PBMC, and BM samples, while 11 (13.41%) genes were found only in ctDNA, 4 (4.88%) were detected only in BM sample, and 2 (2.44%) were detected only in PBMC. There were 18 genes with mutation frequency ≥10% in this study, and 38.89% (7/18) of them were identified by both methods. Eleven genes, namely NRAS (8, 10 and 11), KIT (9, 9 and 10), KRAS (6, 5 and 6), ASXL2 (4, 4 and 4), CEBPA (4, 4 and 4), CSF3R (3, 3 and 3), GATA2 (2, 2 and 3), FLT3-ITD (3, 2 and 2), FBXW7 (2, 2 and 2), EP300 (2, 2 and 2), and TET2 (2, 2 and 2), were with mutation frequency ≥10% of BM, PBMC, and ctDNA sequencing (**Figure 1**).

What's more, a total of 239 mutation forms in 80 abnormal genes were detected in three samples; among all mutations, there were 180 in BM, 203 in ctDNA, and 184 in PBMC. Variant allelic frequencies (VAFs) of 180 mutations in BM were from 0.52% to 99.96% (median 37.97%), including 131 single nucleotide variations (SNVs) and 49 indels. For each patient, the median number of mutation is 9 (3-14); specifically, the number of SNVs was 6.5 (3 to 12) and that of indels was 2.0 (0-12). When compared with the BM sample, more mutations were found in ctDNA with VAFs from 0.50% to 99.98% (median 34.49%), including 150 SNVs and 53 indels. For each patient, the median number of the mutation was 10 (3 to 20), and 7.0 (3 to 20) SNVs and 3.0 (0 to 11) indels were identified. A total of 184 mutations were detected in PBMC with VAFs from 0.51% to 100.00% (median 34.72%), including 140 SNVs and 44 indels. For each patient, the number of mutations ranged from 3 to 20 (median: 8.5), and the median number of SNVs was 6.5 (3 to 12) and that of indels was 2.0 (0 to 8).

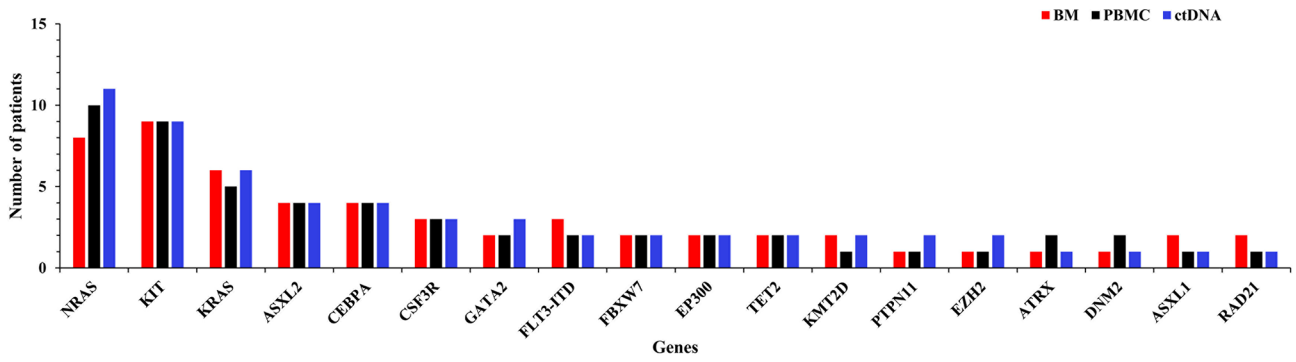


FIGURE 1 | Frequency of AML-related mutated genes as detected by targeted sequencing of bone marrow (BM), peripheral blood mononuclear cell (PBMC), and circulating tumor DNA (ctDNA) samples.

Concordance of Mutation Detection in ctDNA and BM

A total of 219 mutations were identified in BM and ctDNA, while 164 mutations (74.88%) were detected both in ctDNA and BM (**Figure 2**), including 121 SNVs (75.63%) and 43 indels (72.88%). Five patients (UPN3, UPN7, UPN14, UPN15, and UPN16) were with the same mutation sites according to BM and ctDNA sequencing (**Figure S1**). The median absolute blast count in PB of these patients was 19.33 $\times 10^9$/L (0.89 to 43.80 $\times 10^9$/L), which was higher than other patients. There was only one patient (UPN6) with more mutation sites detected in BM than ctDNA (**Figure S1**), and the absolute blast count in PB was 1.23 $\times 10^9$/L. There were 11 patients with more mutation sites detected in ctDNA than BM (**Figure S1**); the median absolute blast count of PB in these patients was 11.51 $\times 10^9$/L, which was similar to others.

Although the number of mutations was similar in three patients (UPN9, UPN10, and UPN13) on the basis of BM and ctDNA, all sites of mutation were completely different (**Figure S1**). For each patient, 8 (2 to 14) mutations were detected in both samples, with high concordance of the number of mutation assessment by both methods ($R^2 = 0.816$, $P < 0.0001$; **Figure 3**).

The median VAF of the 164 individual mutations detected by both assays was 41.34% (range, 0.55% to 99.96%) and 44.36% (range, 0.56% to 99.98%), and the concordance was high in all mutations ($R^2 = 0.945$; $P<0.0001$, **Figure 4**), both in SNVs ($R^2 = 0.948$; $P<0.0001$, **Figure S2A**) and indels ($R^2 = 0.934$; $P<0.0001$, **Figure S2B**). The median VAFs of 16 mutations only detected by BM was 1.22% (0.52% to 14.63%), while it was 0.93% (0.50% to 21.14%) in 39 mutations tested by ctDNA only, and the VAFs were <1% in most of these mutations (**Figure S3**). In view of this, small subclonal populations with lower VAFs<1% were more likely to be missed. A total of 37 mutated genes with clinical significance were detected in all patients, involving 109 mutation sites (93 in BM and 100 in ctDNA). A total of 84 (77.06%)

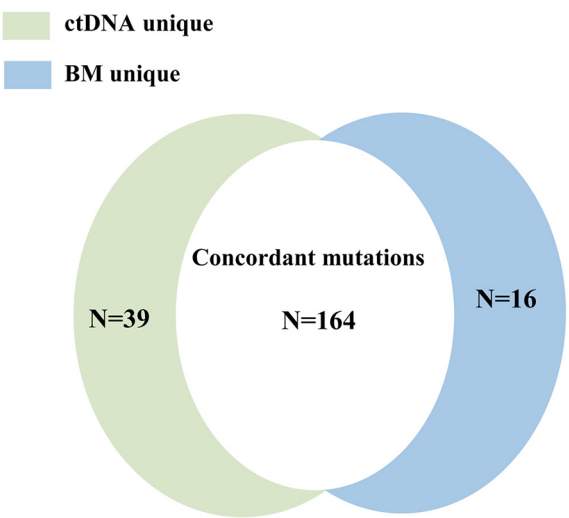


FIGURE 2 | The concordance of the number of mutation detected in BM and ctDNA.

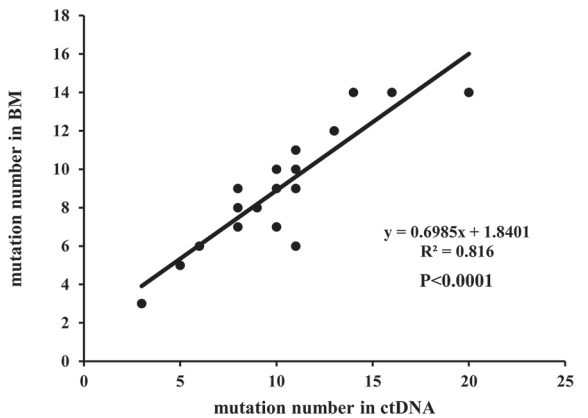


FIGURE 3 | The correlation of the number of mutations in a separate patient detected in BM and ctDNA.

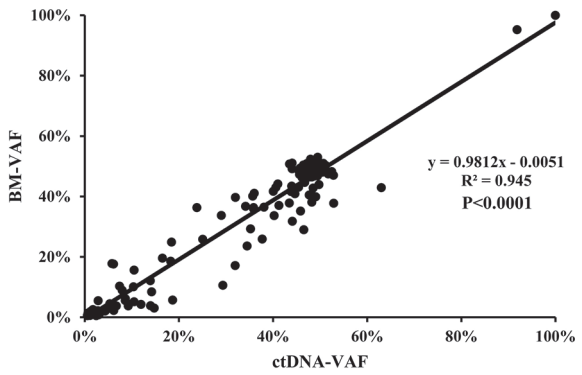


FIGURE 4 | The correlation of the variant allelic frequencies (VAFs) in the same mutation site detected in BM and ctDNA among the 164 individual mutations detected by both assays.

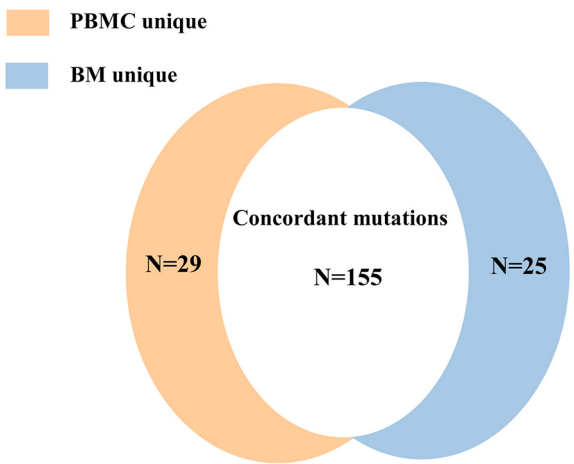


FIGURE 6 | The concordance of the number of mutation detected in BM and PBMC.

mutations were detected in both samples with high concordance of VAF assessment (R^2 = 0.895, P<0.0001; **Figure 5**).

Concordance of Mutation Detection in BM and PBMC

Mutations were detected in PBMC samples of all patients at diagnosis in comparison to BM. A total of 155 mutations (74.16%) were found both in PBMC and BM (**Figure 6**), with high concordance of VAF assessment (R^2 = 0.953, P<0.0001; **Figure 7**).

Analysis of Mutations With Clinical Significance Detected in BM, ctDNA, and PBMC

It was found that 65.79% (75/114) of mutations with clinical significance were detected in three samples, with 9 mutations detected both in ctDNA and BM, and 2 mutations detected both in PBMC and BM (**Figure 8**). The same mutated genes with clinical significance were detected in three samples in five patients (UPN4, UPN7, UPN12, UPN16, UPN20) with high concordance of VAF assessment (**Figure S4**). In addition to three sample co-detecting mutations, the remaining mutations were mostly detected in BM or ctDNA (**Figure S5**). These results suggested that PBMC cannot accurately reflect the mutations of bone marrow.

Comparison of Bone Marrow and ctDNA in Minimal Residual Disease Assessment in AML Patients

To assess whether the dynamic change of ctDNA could reflect the status of MRD in AML patients, we investigated the concordance between the BM sample and ctDNA statuses of 5 out of 20 ctDNA-positive children at diagnosis (**Figure 9**). Of these five patients, one patient experienced relapse on the basis of the BM sample while four patient were relapse-free during the following surveillance. Notably, the results of serial plasma samples showed that four patients under the condition of

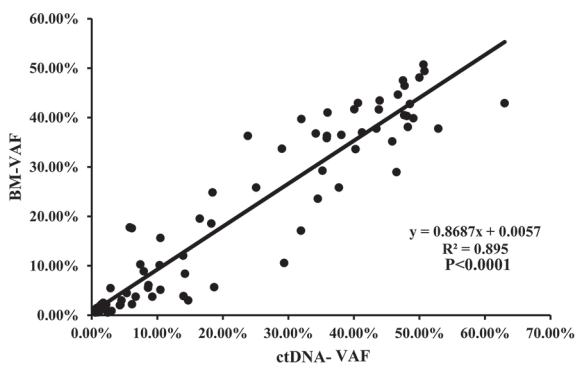


FIGURE 5 | The correlation of the VAFs in the same mutation site detected in BM and ctDNA among the 84 mutations with clinical significance detected by both assays.

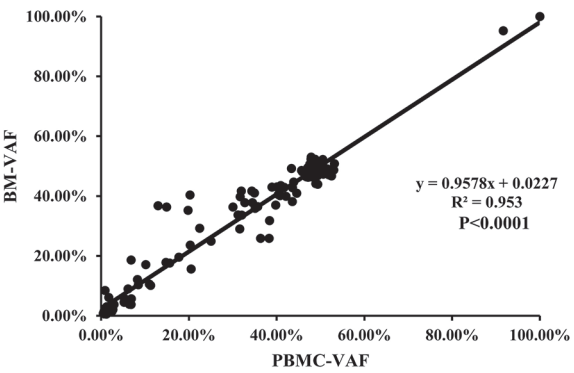


FIGURE 7 | The correlation of the VAFs in the same mutation site detected in BM and PBMC among the 155 individual mutations detected by both assays.

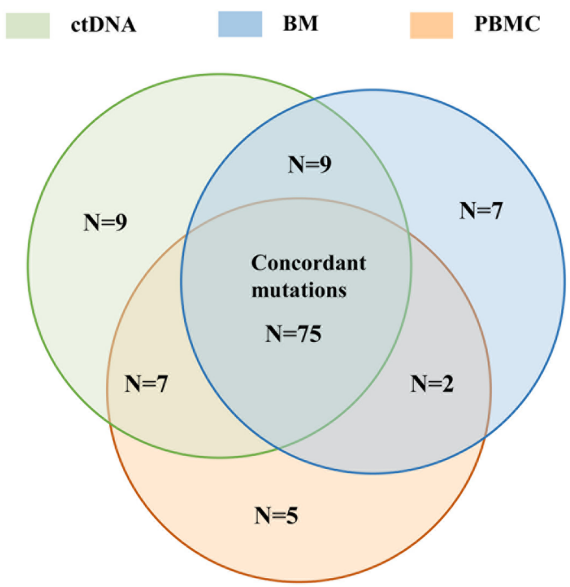


FIGURE 8 | The concordance of the number of mutation detected in BM, ctDNA, and PBMC with clinical significance detected by all assays.

relapse-free were with complete ctDNA clearance after chemotherapy and remained negative at the last follow-up. As for one patient who experienced relapse, ctDNA positivity regained after a temporary ctDNA clearance by chemotherapy, and recurrence of these cytogenetic abnormalities in ctDNA was observed nearly 2 months earlier than BM relapse. These results showed that ctDNA was basically consistent with the results from BM samples, and the shifting level of ctDNA may be a useful tool for MRD monitoring in children with AML.

DISCUSSION

In this study, we aim to evaluate the potential value of MRD based on positive ctDNA status in patients with AML, and the result reported that surveillance of matched serum ctDNA in residual driver mutation persistence may be regarded as an independent sample of MRD testing, which was comparable and with high concordance with sequencing of BM samples for the diagnosis of gene alterations in the AML children.

AML is a highly heterogeneous disease, and its diagnosis and treatment require a comprehensive analysis of morphology, immunology, genetics, and molecular biology. NGS, as a new molecular biological technology, has the advantages of high throughput, high sensitivity, and low cost and is an important means to explore the molecular pathogenesis of blood tumors and guide clinical diagnosis and treatment. Previously, detection of gene mutation by BM was the standard method to identify DNA aberration in AML patients. However, acquisition of the BM sample is traumatic, and it is usually difficult to collect specimens in succession for the close monitoring of MRD, which greatly limits its application in clinical practices. Moreover, the sensitivity of MRD monitoring from PB was much lower than that from BM (15, 16). This is true even under the circumstances of highly sensitive real-time PCR-based methods targeting leukemia-related gene alterations.

In 1948, Mandel and Metais firstly advanced the presence of cfDNA in human blood (17). Subsequently, Koffler et al. (18) found the higher concentration of cfDNA in the circulation of patients with cancer when compared with healthy people, indicating that the presence of the cancer patients may be simply screened through a test of PB. In 1994, cfDNA was regarded as an independent sample in distinguishing RAS mutations in patients with hematological oncology (19, 20).

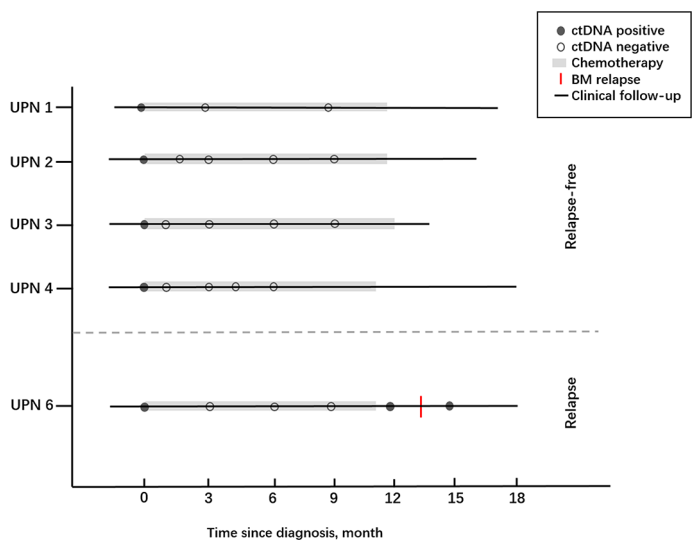


FIGURE 9 | The clinical courses together with ctDNA statuses of five AML children who received chemotherapy.

The recent introduction of NGS-based molecular approaches has further refined such MRD measurements with regard to broader applicability. ctDNA was a kind of noninvasive method that showed its great potential in identifying the gene mutation, and specifically for patients for whom no conventional genetic marker for MRD testing was available or conventional MRD approaches such as flow cytometry or cytogenetics were negative in AML children in recent years (21–24).

To date, data on the utility of ctDNA from PB in AML children are relatively sparse; moreover, the results regarding the diagnostic value in this population were still unspecified. In this study, it is the first time that ctDNA was used for the detection of genetics aberration in pediatric AML, and consistent results were found in this sample when compared with BM and PBMC samples on the distribution of targeted sites. Moreover, the absolute blast count in PB did not affect the result of ctDNA in identifying gene mutations. It was found that ctDNA has good consistency with BM in the analysis of mutation frequency, and ctDNA may identify some potential mutations that cannot be detected by NGS in the BM and PBMC sample.

MRD monitoring has been used as a vital tool for early prediction of the efficacy of chemotherapy in AML children. For MRD evaluation, the sample of choice is BM, although peripheral blood is easy to obtain and lacks immature normal populations of cells that may interfere with the analysis. ctDNA has the potential to capture intratumor heterogeneity that may be missed by BM analysis. In addition, ctDNA has an advantage of faster turn-around time as well as an acceptable running cost for serial monitoring of MRD. The current practice for the assessment of MRD for response assessments relies on BM sampling, whereas dynamic ctDNA monitoring may be adequate for reflecting the remission status in some AML cases. In view of this, these findings potentially introduce the utility of this noninvasive means at the time of diagnosis.

In summary, our results confirm that ctDNA may be used as a complementary method in reflecting the mutation spectrum and MRD monitoring of AML children, which may be particularly relevant in the context of subclonal mutations with lower VAF. However, these findings warranted a larger, prospective study to investigate the prognostic stratification and MRD monitoring in pediatric AML.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://ngdc.cncb.ac.cn/gsa-human/browse/HRA000912.

ETHICS STATEMENT

The study design and methods complied with the Declaration of Helsinki and was approved by the Ethics Committee and Institutional Review Board of Institute of Hematology and Blood Diseases Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College. Informed consent was obtained from all subjects.

AUTHOR CONTRIBUTIONS

XZ and MR contributed to the idea, conception, and study design. MR and LpL wrote the manuscript and contributed to data analysis. BQ, FaL, SW, XL, YG, YZ, and YC collected, organized, and provided AML patient data. XyC, LC, LxL, CW, FeL, and SC conducted experiments, analyzed the data, and contributed to generating figures. AZ, XjC, and YcZ helped in writing and revising the manuscript, and generating the figures. LZ revised the manuscript and prepared it for submission.

FUNDING

This work was supported by the CAMS Innovation Fund for Medical Sciences (CIFMS) (2020-I2M-C&T-B-087) and the National Natural Science Foundation of China (81500156, 81170470).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.666470/full#supplementary-material

REFERENCES

1. Steliarova-Foucher E, Colombet M, Ries LAG, Moreno F, Dolya A, Bray F, et al. IICC-3 Contributors. International Incidence of Childhood Cancer, 2001-10: A Population-Based Registry Study. *Lancet Oncol* (2017) 18(6):719–31. doi: 10.1016/S1470-2045(17)30186-9
2. Gamis AS, Alonzo TA, Perentesis JP, Meshinchi S. COG Acute Myeloid Leukemia Committee. Children's Oncology Group's 2013 Blueprint for Research: Acute Myeloid Leukemia. *Pediatr Blood Cancer* (2013) 60 (6):964–71. doi: 10.1002/pbc.24432
3. Grimwade D, Ivey A, Huntly BJP. Molecular Landscape of Acute Myeloid Leukemia in Younger Adults and its Clinical Relevance. *Blood* (2016) 127 (1):29–41. doi: 10.1182/blood-2015-07-604496
4. Pui CH, Carroll WL, Meshinchi S, Arceci RJ. Biology, Risk Stratification, and Therapy of Pediatric Acute Leukemias: An Update. *J Clin Oncol Off J Am Soc Clin Oncol* (2011) 29(5):551. doi: 10.1200/JCO.2010.30.7405
5. Bolouri H, Farrar JE, Triche T Jr, Ries RE, Lim EL, Alonzo TA, et al. The Molecular Landscapeof Pediatric Acute Myeloid Leukemia Reveals Recurrent Structural Alterations and Age-Specific Mutational Interactions. *Nat Med* (2018) 24(1):103–12. doi: 10.1038/nm.4439
6. Jahr S, Hentze H, Englisch S, Hardt D, Fackelmayer FO, Hesch RD, et al. DNA Fragments in the Blood Plasma of Cancer Patients: Quantitations and Evidence for Their Origin From Apoptotic and Necrotic Cells. *Cancer Res* (2001) 61(4):1659–65.
7. Jiang P, Chan CW, Chan KC, Cheng SH, Wong J, Wong VW, et al. Lengthening and Shortening of Plasma DNA in Hepatocellular Carcinoma Patients. *Proc Natl Acad Sci USA* (2015) 112(11):E1317–25. doi: 10.1073/pnas.1500076112
8. Snyder MW, Kircher M, Hill AJ, Daza RM, Shendure J. Cell-Free DNA Comprises an In Vivo Nucleosome Footprint That Informs Its Tissues-of-Origin. *Cell* (2016) 164(1-2):57–68. doi: 10.1016/j.cell.2015.11.050
9. Stroun M, Lyautey J, Lederrey C, Olson-Sand A, Anker P. About the Possible Origin and Mechanism of Circulating DNA Apoptosis and Active DNA

Release. *Clin Chim Acta* (2001) 313(1-2):139–42. doi: 10.1016/S0009-8981(01)00665-9
10. Yu SC, Lee SW, Jiang P, Leung TY, Chan KC, Chiu RW, et al. High-Resolution Profiling of Fetal DNA Clearance From Maternal Plasma by Massively Parallelsequencing. *Clin Chem* (2013) 59(8):1228–37. doi: 10.1373/clinchem.2013.203679
11. Minchin RF, Carpenter D, Orr RJ. Polyinosinic Acid and Polycationic Liposomes Attenuate the Hepatic Clearance of Circulating Plasmid DNA. *J Pharmacol Exp Ther* (2001) 296(3):1006–12.
12. Lo YM, Zhang J, Leung TN, Lau TK, Chang AM, Hjelm NM. Rapid Clearance of Fetal DNA From Maternal Plasma. *Am J Hum Genet* (1999) 64(1):218–24. doi: 10.1086/302205
13. Yeh P, Dickinson M, Ftouni S, Hunter T, Sinha D, Wong SQ, et al. Molecular Disease Monitoring Using Circulating Tumor DNA in Myelodysplastic Syndromes. *Blood* (2017) 129(12):1685–90. doi: 10.1182/blood-2016-09-740308
14. Yeh P, Hunter T, Sinha D, Ftouni S, Wallach E, Jiang D, et al. Circulating Tumour DNA Reflects Treatment Response and Clonal Evolution in Chronic Lymphocytic Leukaemia. *Nat Commun* (2017) 8:14756. doi: 10.1038/ncomms14756
15. Shayegi N, Kramer M, Bornhäuser M, Schaich M, Schetelig J, Platzbecker U, et al. The Level of Residual Disease Based on Mutant NPM1 Is an Independent Prognostic Factor for Relapse and Survival in AML. *Blood* (2013) 122(1):83–92. doi: 10.1182/blood-2012-10-461749
16. Nakamura S, Yokoyama K, Shimizu E, Yusa N, Kondoh K, Ogawa M, et al. Prognostic Impact of Circulating Tumor DNA Status Post-Allogeneic Hematopoietic Stem Cell Transplantation in AML and MDS. *Blood* (2019) 133(25):2682–95. doi: 10.1182/blood-2018-10-880690
17. Mandel P, Métais P. Les Acides Nucléiques Duplasma Sanguin Chez L'homme. *C R Acad Sci* (1948) 142(3-4):241–3.
18. Koffler D, Agnello V, Winchester R, Kunkel HG. The Occurrence of Single-Stranded DNA in the Serum of Patients With Systemiclupus Erythematosus and Other Diseases. *J Clin Invest* (1973) 52(1):198–204. doi: 10.1172/JCI107165
19. Sorenson GD, Pribish DM, Valone FH, Memoli VA, Bzik DJ, Yao SL. Soluble Normal and Mutated DNA Sequences From Single-Copy Genes in Human Blood. *Cancer Epidemiol Biomarkers Prev* (1994) 3(1):67–71.
20. Vasioukhin V, Anker P, Maurice P, Lyautey J, Lederrey C, Stroun M. Point Mutations of the N-Ras Gene in the Blood Plasma DNA of Patients With Myelodysplastic Syndrome or Acute Myelogenous Leukaemia. *Br J Haematol* (1994) 86(4):774–9. doi: 10.1111/j.1365-2141.1994.tb04828.x
21. Jongen-Lavrencic M, Grob T, Hanekamp D, Kavelaars FG, Al Hinai A, Zeilemaker A, et al. Molecular Minimal Residual Disease in Acute Myeloid Leukemia. *N Engl J Med* (2018) 378(13):1189–99. doi: 10.1056/NEJMoa1716863
22. Gaksch L, Kashofer K, Heitzer E, Quehenberger F, Daga S, Hofer S, et al. Residual Disease Detection Using Targeted Parallel Sequencing Predicts Relapse in Cytogenetically Normal Acute Myeloid Leukemia. *Am J Hematol* (2018) 93(1):23–30. doi: 10.1002/ajh.24922
23. Ivey A, Hills RK, Simpson MA, Jovanovic JV, Gilkes A, Grech A, et al. Assessment of Minimal Residual Disease in Standard-Risk AML. *N Engl J Med* (2016) 374(5):422–33. doi: 10.1056/NEJMoa1507471
24. Morita K, Kantarjian HM, Wang F, Yan Y, Bueso-Ramos C, Sasaki K, et al. Clearance of Somatic Mutations at Remission and the Risk of Relapse in Acute Myeloid Leukemia. *J Clin Oncol* (2018) 36(18):1788–97. doi: 10.1200/JCO.2017.77.6757

Conflict of Interest: Author LxL, SC, FeL, and CW were employed by the company Acornmed Biotechnology Co., Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

ORIGINAL RESEARCH
published: 06 August 2021
doi: 10.3389/fonc.2021.628531

Genotypic Characteristics of Hepatoblastoma as Detected by Next Generation Sequencing and Their Correlation With Clinical Efficacy

Huimin Hu[†], *Weiling Zhang*[†], *Tian Zhi*, *Jing Li*, *Yuan Wen*, *Fan Li*, *Yanyan Mei* and *Dongsheng Huang*[*]

Department of Pediatrics, Beijing Tongren Hospital, Capital Medical University, Beijing, China

OPEN ACCESS

Edited by:
Jing He,
Guangzhou Medical University, China

Reviewed by:
Junhui Hu,
University of California, Los Angeles, United States
Shen Hu,
UCLA Jonsson Comprehensive Cancer Center, United States

***Correspondence:**
Dongsheng Huang
hdsmed@163.com

[†]*These authors have contributed equally to this work*

Specialty section:
This article was submitted to Pediatric Oncology, a section of the journal Frontiers in Oncology

Received: *12 November 2020*
Accepted: *20 July 2021*
Published: *06 August 2021*

Citation:
Hu H, Zhang W, Zhi T, Li J, Wen Y, Li F, Mei Y and Huang D (2021) Genotypic Characteristics of Hepatoblastoma as Detected by Next Generation Sequencing and Their Correlation With Clinical Efficacy. Front. Oncol. 11:628531. doi: 10.3389/fonc.2021.628531

Background: Hepatoblastoma (HB) is the most common malignant embryonic liver tumor type in children under 3 years of age. In the present study, the next generation sequencing (NGS) method was used to detect the genotype characteristics of HB and summarize the correlation between the common mutation genotypes noted in this disease and the clinical treatment and prognosis. The results may aid clinical prognosis and the successful application of targeted drugs.

Methods: Initially, DNA was extracted from tumor tissue specimens and peripheral blood derived from 19 pediatric patients with HB. Subsequently, DNA panel and NGS methods were used to detect tumor diagnosis and the expression levels of treatment-associated genes, followed by the summary of genotype characteristics. In addition, in order to further assess the application of immunotherapy in HB, immunohistochemical detection of programmed cell death 1 ligand 1 (PDL1) was performed in combination with tumor mutation burden (TMB) and DNA mismatch repair status analysis. Furthermore, the clinical treatment effect and prognosis of the pediatric patients were statistically analyzed according to the characteristics of the genotype. Overall prognosis and prognostic analyses in different groups were performed by Kaplan-Meier and log-rank tests, respectively. Finally, expression validation and diagnostic analysis of commonly reported genes were performed in the GSE75271 dataset, which was obtained from the Gene Expression Omnibus (GEO) database.

Results: In the present study, certain mutated genes, including nuclear factor erythroid 2-related factor 2 (NFE2L2), catenin β1 (CTNNB1), MYCN, tumor protein p53, axis inhibition protein 1 (AXIN1) and adenomatous polyposis coli (APC) were associated with the pathogenesis of HB. During TMB and DNA mismatch repair status analyses, pediatric patients had a low TMB. All of them did not present with microsatellite instability. The immunohistochemical results indicated lower expression levels of PDL1 in HB. The complete remission (CR) rate of pediatric patients in the gene abnormality group was

lower than that of the non-reported disease-associated gene abnormality group. The 2-year overall survival rate and disease-free survival rate of 19 pediatric patients with HB were 72.1% and 42.4%, respectively. Receiver operating characteristic (ROC) analysis demonstrated that CTNNB1, NFE2L2, AXIN1, APC, MYCN and insulin growth factor 2 (IGF2) may be potential biomarkers that could be used for the diagnosis of HB.

Conclusion: The genotype changes in HB were more common and the CR rate of the pediatric patients with an altered genotype was lower than that of pediatric patients without an altered genotype. In addition, pediatric patients with HB exhibited lower TMB compared with adult patients. Moreover, the data indicated that *CTNNB1*, *NFE2L2*, *AXIN1*, *APC*, *MYCN* and *IGF2* may be potential biomarkers that can be used for the diagnosis of HB.

Keywords: hepatoblastoma, CTNNB1, next generation sequencing, tumor mutation burden, PDL1

INTRODUCTION

Hepatoblastoma (HB) is the most common malignant embryonic tumor of the liver in children (most common under 3 years of age), accounting for approximately 79% of all pediatric malignant liver tumors, with an average annual incidence of 1.5 per million population (1–3). Abdominal masses are major clinical manifestations of HB (4). Clinically, the treatment of HB mainly includes surgery and chemotherapy (2, 5). The cause of HB remains unclear. However, previous studies have shown that the incidence of HB in premature infants with very low birth weight is high (6). In addition, children with very low birth weight exhibit a higher risk of HB than those with normal weight (7). The incidence of HB in pediatric patients with Beckwith-Wiedemann syndrome has recently increased 10-fold (from 1,000 to 10,000) (8, 9).

HB is a disease mainly caused by the activation of the WNT pathway, which involves the activating mutation/deletion of exon 3 of the *catenin β1* (*CTNNB1*) gene (10). Certain rare gene mutations lead to the activation of the WNT pathway, such as those occurring in the genes *axis inhibition protein 1* (*AXIN1*), *axis inhibition protein 2* (*AXIN2*) and *adenomatous polyposis coli* (*APC*) (can only be observed in cases associated with familial adenomatous polyposis) (11). Previous studies reported that the mutation frequency of *CTNNB1* and *nuclear factor, erythroid 2 like 2* (*NFE2L2*) was 80% and 13%, respectively (12). An additional study demonstrated that the mutation frequency of *NFE2L2* was approximately 10% (10). To date, a high number of studies have been performed on the genotype of patients with HB. However, a lower number of reports have been conducted on the correlation between genotype and HB prognosis. In the present study, statistical methods were used to analyze the genotype characteristics of 19 pediatric patients with HB. Subsequently, the clinical data, clinical efficacy and prognosis of these patients were analyzed. Finally, the correlations between genotype and clinical phenotype and between genotype and clinical efficacy of HB were summarized. The present study may provide a basis for the application of targeted drugs for the treatment of HB.

MATERIALS AND METHODS

Patients

A total of 19 Han nationality pediatric patients with HB were selected who were hospitalized in our hospital between November 1, 2018 and March 31, 2020. These patients included 17 patients with recurrence or metastasis and 2 patients with unsatisfactory decrease in the levels of alpha-fetoprotein (AFP) prior to surgery. The tumor and plasma samples were collected for genetic testing in Rendong Medical Laboratory.

The parents of the pediatric patients with HB signed the informed consent form for their participation in the study protocol, which included examination and treatment. The present study was approved by the Medical Ethics Committee of the Beijing Tongren Hospital, Capital Medical University (approval no. TRECKY2019-033).

Sample Source and Processing

The plasma samples were centrifuged at 1,600 x g and 16,000 x g for 10 min. A Blood Genomic DNA Mini kit (CW Biotech) was used to extract genomic DNA (gDNA) from white blood cells. The gDNA was used as a reference genome. A QIAamp DNA FFPE Tissue kit (Qiagen, Inc.) was used to extract tumor DNA from 5-10 formalin fixed paraffin-embedded (FFPE) sections (5-mm thick). Qubit detection was performed on the extracted DNA. In case the total amount of DNA was >3 μg and the A260/280 ratio was within the range of 1.8-2.0, the quality of DNA was determined to meet the requirements for subsequent experiments.

Genetic Testing

The NimbleGen SeqCap EZ capture panel was used to capture the coding regions of the genes. In each sample, 200-500 ng FFPE DNA or 500 ng gDNA were used for library preparation and quantification using KAPA Hyper Prep protocols. Briefly, the DNA samples were fragmented by nebulization and the fragmented DNA was repaired. An 'A' was ligated to the 3′ end. Subsequently, Illumina adapters were ligated to the fragments and the samples were size selected aiming for a

350-400 base pair product. The size selected product was PCR amplified and the concentration of the final product was determined using the QubitDsDNHS Assay kit. Nimblegen/IDT was used for 16-h hybridization capture of 4-6 libraries at 47/65°C. Washing, recovery and amplification were carried out in sequence according to the standard procedures of the NimbleGen SeqCap EZ and IDT panels. The AMPure XP (Beckman Coulter, Inc.) and Qubit™ dsDNA HS Assay kits (Thermo Fisher Scientific, Inc.) were used for library purification and quantification, respectively. LuminaNextseq500 (pe75) sequencer was used for library sequencing.

Data Processing and Bioinformatics Analysis

The original data (13), bioinformatics and mutation (14–16) analyses were performed according to the methods reported in the literature. Initially, the processed data were compared with the reference genome to delete duplicate readings. Subsequently, the data sequence that aligned to a single mode of the genome was also aligned to the exome region. Finally, associated mutations were annotated and analyzed. The depth distribution and coverage uniformity of single bases was assessed in the target region. The sources of the mutation databases included the following: Catalogue Of Somatic Mutations In Cancer, Oncology Knowledge Base, MD Anderson, China Kadoorie Biobank, 1000 Genomes Project, Single Nucleotide Polymorphism Database, ClinVar, NHLBI GO Exome Sequencing Project, Exome Aggregation Consortium, Ensembl, Human Gene Mutation Database and University of California Santa Cruz.

Diagnosis, Staging and Grouping of HB

The stage of HB was based on the pretreatment extent of disease (PRETEXT) stage and Children's Oncology Group (COG) Evans stage system. The risk group was based on the risk factors that affected the prognosis (12). The detailed inclusion criteria for the patients with HB were as follows (12): Pediatric patients younger than 14 years and a clear diagnosis of HB. In addition, the pediatric patients who met one of the following conditions were considered as high-risk cases: (1) Concentration of serum AFP <100 ng/ml; (2) patients were PRETEXT IV stage prior to operation; (3) postoperative patients with COG IV stage; (4) patients who had invasion of portal vein (P+), inferior vena cava or hepatic vein (V+). The patients who did not receive regular treatment and follow-up were excluded.

Clinical Treatment

According to the Chinese Guidelines for the Diagnosis and Treatment of Childhood HB, the main therapeutic drugs used were platinum and anthracycline. However, in order to reduce the cardiotoxicity of anthracyclines, different chemotherapeutic regimens were used for pediatric patients with different stage HB in China. The pediatric patients in the low-risk group were mainly treated with the "cisplatin + 5-fluorouracil + vincristine" combination, whereas the pediatric patients in the medium-risk group were mainly treated with the "cisplatin + 5-fluroracil + vincristine + doxorubicin" combination. In addition, the pediatric patients in high-risk group were mainly treated with "cisplatin + adriamycin", "carboplatin + adriamycin" and "ifosfamide + carboplatin + etoposide". VIT is a second-line solution. If the treatment regimen recommended by the guidelines was ineffective, "vincristine + lrinotecan + temozolomide" or "vincristine + lrinotecan + cisplatin + cyclophosphamide" and other regimens were applied at a later stage and the relevant results were further explained.

Analysis of DNA Mismatch Repair Status and Tumor Mutation Burden (TMB)

Deficient mismatch repair results in a strong mutator phenotype known as microsatellite instability (MSI) (17). In the present study, NGS method was used to evaluate the length distribution of 309 microsatellite loci to determine the MSI status (18). More than 20% of unstable microsatellite sites are unstable. MSI was considered when the unstable loci accounted for more than 20% of total loci. TMB was defined as the number of somatic, coding, base substitution, and indel mutations per megabase of genome examined.

Efficacy Assessment and Prognosis

According to the Chinese Guidelines for the Diagnosis and Treatment of Childhood HB, the efficacy assessment and prognostic criteria of HB were as follows: Complete tumor disappearance upon physical examination and computed tomography or magnetic resonance imaging and normal concentration of AFP for >4 weeks. These criteria were considered necessary for complete remission (CR). Tumor shrinkage ≥50% in the absence of any evidence of new lesions or disease progression was regarded as the partial response (PR). Tumor shrinkage <50% in the absence of any evidence of new lesions or tumor growth was defined as stable disease. Tumor enlargement ≥25%, development of a new tumor or increased AFP levels were considered disease progression (PD). Biopsy confirmation, clear imaging evidence and 3 increase in the serum AFP levels within 4 weeks was considered disease recurrence.

In the present study, the deadline for the follow-up period was May 31, 2020 or the time of death due to cancer progression. Survival analysis was performed. CR and PR represented effective treatment and PD and time of death represented ineffective treatment.

Diagnostic Analysis and Expression Verification in the GSE75271 Dataset

Expression validation and diagnostic analyses of commonly reported genes were performed in the GSE75271 dataset (involving 50 HB samples and five normal controls), which was obtained from the Gene Expression Omnibus (GEO) database (19). Receiver operating characteristic (ROC) analysis was also performed using pROC package in R language. The sensitivity and specificity at the cut-offs was calculated as determined in a previous report (20). The diagnostic ability was evaluated by the area under curve (AUC) values in the ROC curve. In the expression verification, the Wilcoxon signed-rank test was used to analyze the statistical difference between the normal control and the HB groups.

Statistical Analysis

In the present study, all statistical analyses were performed using SPSS21 (IBM Corp.). Normally distributed and skewed distribution data were presented as the mean ± standard deviation and median (or quartile), respectively. The Fisher exact probability test analysis of the chi-squared test was used for the analysis of the comparison of count data in different groups. The overall prognosis analysis was analyzed by the Kaplan-Meier method. The prognostic analysis in the different groups was analyzed using the log-rank test. Expression verification was performed using the Wilcoxon signed-rank test in order to analyze significant differences. $P<0.05$ was considered to indicate a statistically significant difference.

RESULTS

Clinical Features of 19 Pediatric Patients With HB

The "pwr" package in R was used for power analysis, which was analyzed by the chi-squared power calculation. The effect value (w)=0.5060481, sample size (n)=19, degree of freedom=1 and significance threshold (sig. level)=0.05 were initially input followed by the output power level (power)=0.5971157. The closer the power to 1, the more accurate fit of the sample size on the experiment. In the present study, the power was estimated to 0.5971157, indicating that the sample size was medium but not optimal. In total, 9 males and 10 females were enrolled as pediatric patients with HB. The age of onset range and the median age of onset were 11-120 months and 33 months, respectively. No family history of familial adenomatous polyposis and Beckwith-Wiedemann syndrome was noted in all cases. The specific clinical characteristics of 19 pediatric patients with HB are shown in **Table 1**.

TABLE 1 | Clinical features of 19 pediatric patients with HB.

Feature	Number of cases	Proportion (%)
Initial symptom		
Abdominal mass and liver space are occupied	9	47.37%
Abdominal pain and bloating	9	47.37%
Fever, fatigue and wasting	1	5.26%
AFP value at onset (ng/mL)		
726-2170000, median: 31877		
<100	0	0
≥100	19	100%
PRETEXT staging		
PRETEXT II	5	26.32%
PRETEXT III	11	57.89%
PRETEXT IV	3	15.79%
Intrahepatic lesion at onset		
Single lesion	13	68.42%
Many lesions	6	31.58%
COG staging		
COG II	4	21.05%
COG III	6	31.58%
COG IV	9	47.37%
Pulmonary metastatic disease	7	36.84%
Single lung metastasis	2	10.53%
Double lung metastasis	5	26.32%
Other distant metastases		
Osseous metastasis	1	5.26%
Brain metastases	1	5.26%
Other high risk factors affecting prognosis		
V+	2	10.53%
P+	2	10.53%
E+	2	10.53%
H+	1	5.26%
N+	2	10.53%
Pathological type		
Epithelium type	12	63.16%
Fetal type	2	10.53%
Embryonal	1	5.26%
Fetal and embryonic mixed type	7	36.84%
Giant beam type	1	5.26%
Embryo and giant beam mixed type	1	5.26%
Epithelium and lobus intermedius mixed type	7	36.84%

HB, hepatoblastoma; AFP, alpha-fetoprotein; PRETEXT, pretreatment extent of disease; COG, children's oncology group; V+, invade inferior vena cava or hepatic vein; P+, invasion of the portal vein; E+, extra-hepatic and intra-abdominal disease; H+, rupture of tumor or intraperitoneal hemorrhage; N+, lymph node invasion.

Genotypic Characteristics of Next-Generation Sequencing

In the present study, 642 genes with coding regions were captured. The genes with mutation abundance ≥5% are shown in **Table 2**. In total, 7 out of 19 pediatric (36.84%) patients with HB exhibited *CTNNB1* gene alterations, among which, 5 patients had single nucleotide variants and two gene indel deletion alterations. In the present study, no *CTNNB1* gene deletions were detected, which may be associated with the small sample size of the population or ethnic differences. There was 1 patient who exhibited both single nucleotide variation of *tumor protein p53* (*TP53*, also known as *p53*) gene and point mutation of the *CTNNB1* gene. The *MYCN proto-oncogene* and the *basic helix-loop-helix transcription factor* gene were amplified in 1 patient. Moreover, 2 patients exhibited single nucleotide variation of the *NFE2L2* gene. Indel deletion alterations of the *AXIN1* gene were noted in 1 patient with chronic hepatitis B. It is interesting to note that 1 patient presented with single nucleotide variations of *APC*, *insulin like growth factor 2* (*IGF2*), *dicer 1, ribonuclease III* (*DICER1*), *notch receptor 1* (*NOTCH1*) and *phosphatidylinositol-4-phosphate 3-kinase catalytic subunit type 2 beta* (*PIK3C2B*) genes. In addition, 1 patient exhibited a single nucleotide variant of the *B cell lymphoma 6* (*BCL6*) gene.

The tumor mutation burden (TMB) was assessed in 17 pediatric patients with HB. The range and median of TMB was 0-22 mut/Mb and 2.25 mut/Mb, respectively. The TMB in only 1 patient reached 22 mut/Mb. Moreover, the patient exhibited a partially missing base of the *AXIN1* gene. The DNA mismatch repair status analysis indicated that all patients with HB did not exhibit microsatellite instability (**Table 2**).

Immunohistochemical Analysis

Due to the infrequent use of immunotherapy for pediatric solid tumors, only 8 patients were tested for the expression of programmed cell death 1 ligand 1 (PDL1). In order to further

investigate the application of immunotherapy in HB, immunohistochemical detection of PDL1 was performed in 8 pediatric patients with HB. The percentage of tumor cells was measured that indicated partial or complete membrane staining at any intensity. The results further indicated that the percentage of all of PDL1 staining in the tumor cells was ≤2% (**Table 2**), which demonstrated the lower positive expression of PDL1 in HB. Immunohistochemical analysis of PDL1 is shown in **Figure 1**. Immunohistochemical detection of β-catenin was performed in 19 pediatric patients with HB (**Table 2**). The degree of cell staining accounted for >1/4 of the total cells, which were considered positive. Among them, only case 15 was negative and exhibited a splice point deletion mutation of the *CTNNB1* gene (c.53_241+117del306). In addition, case 7 was focally positive and exhibited an amplification mutation of the *MYCN* gene. The remaining immunohistochemical results of β-catenin included positive to strong positive cases.

Genotype and Efficacy Analysis

A total of 17 out of 19 pediatric patients with HB relapsed and were accompanied by distant metastasis. CR was achieved in 2 patients without recurrence or metastasis (**Table 2**). These patients were divided into the gene abnormality group and the non-reported disease-associated gene abnormality group based the presence of the commonly reported genes was associated with the pathogenesis of childhood HB (*CTNNB1*, *NFE2L2*, *AXIN1*, *TP53*, *APC*, *IGF2*). A total of 12 patients were identified in the gene abnormality group, among whom 11 patients, exhibited recurrence and metastasis. In total, 7 patients (6 patients exhibited recurrence and metastasis) were present in the non-reported disease-related gene abnormality group. Non-significant differences were noted with regards to recurrence and metastasis between the two groups (P=0.614). The prognosis of the two groups was compared. Among 12 patients in the gene abnormality group, 4 cases did not survive, 1 developed PD, 3 exhibited PR and 4 CR. Moreover, 7 and 5 cases were noted with effective and ineffective treatment, respectively. Among the 7 patients in the non-reported disease-associated gene abnormality group, 1 case of PR and 6 cases of CR were present, whereas all the cases received effective treatment. There were 4 cases of CR (33.33%, 4/12) in the gene abnormality group. There were 6 cases of CR (85.71%, 6/7) in the gene abnormality group. There was a statistical difference between the two groups in view of CR rate (P=0.027).

Overall Prognostic Analysis

The follow-up time range of 19 pediatric patients with HB was 9-50 months (4 cases died), with a median follow-up time of 21 months. The 2-year overall survival rate of 19 pediatric patients with HB was 72.1% and their survival time was 37.7 ± 5.99 months. The survival curve of these pediatric patients with HB is shown in **Figure 2A**. Among the 19 patients, 10 cases exhibited a CR. The 2-year disease-free survival rate was 42.4%. The disease-free survival time was 29.71 ± 3.92 months. The disease-free survival curve of 19 pediatric patients with HB is shown in **Figure 2B**.

Genotype and Prognosis Analysis

The prognostic analysis between the gene abnormality group and the non-reported disease-associated gene abnormality group demonstrated that the 2-year overall survival rates of the two groups were 64.3% and 100%, respectively (**Figure 3A**). The difference was not statistically significant (chi-squared=0.536, P=0.464). According to the TMB, the patients were divided into TMB ≥5 mut/Mb group and TMB <5 mut/Mb group. The 2-year overall survival rates of the two groups were 100% and 57.1%, respectively (**Figure 3B**). No significant differences were noted between the two groups (chi-squared=2.616, P=0.106).

Diagnostic Analysis and Expression Verification

The GSE75271 dataset was used to investigate commonly reported genes. ROC analysis and expression verification were performed. *CTNNB1*, *NFE2L2*, *AXIN1*, *APC*, *MYCN*, *IGF2*, *TP53* and *BCL6* were selected for diagnostic analysis (**Figure 4**). ROC analysis demonstrated that with the exception of *TP53* and *BCL6*, the AUC of other genes was >0.7. The data indicated that the *CTNNB1*, *NFE2L2*, *AXIN1*, *APC*, *MYCN* and *IGF2* genes may be potential biomarkers for the diagnosis of HB. Simultaneously, expression verification of these genes was performed. With the exception of *TP53* and *BCL6*, the expression levels of other genes were significantly different between case and control groups (P<0.05) (**Supplementary Figure 1**). This suggested that *CTNNB1*, *NFE2L2*, *AXIN1*, *APC*, *MYCN* and *IGF2* may play an important regulatory role in the pathological mechanism of HB. In addition, the expression trends of specific genes, such as *MYCN* and *BCL6*, were inconsistent with those reported in the literature, which may be associated with the small number of normal control samples in the GSE75271 dataset. Unfortunately, the prognostic data for HB were not available in public databases. Therefore, prognostic analysis of these genes could not be performed.

DISCUSSION

The two main primary liver malignancies derived from hepatocytes in children are HB and hepatocellular carcinoma (HCC), among which HB is more common than HCC (21). HB is the most common hepatic tumor noted in children, which accounts for 43% of liver tumors or two-thirds of liver malignancies in children (22). Several unique genetic features have been found in HB and HCC, including germline mutations of certain genes, hallmark cytogenetic changes and repeated mutations in the main somatic cells described in these tumors (21). In addition, the present study performed a comprehensive genetic analysis of HB and HCC using NGS, which not only confirmed the previously discovered frequent mutations of *CTNNB1* and *p53* genes in HB and HCC, but also identified new genetic changes in tumor-associated genes, including *AXIN*, *APC*, *cyclin-dependent kinase inhibitor 2a* (*CDKN2A*), *inverted formin 2* (*INF2*) and *AT-rich interactive domain 2* (*ARID2*) (21).

Previous studies have shown that changes in the expression levels of the p53 gene are associated with poor prognosis of HCC

TABLE 2 | Specific genotypes of 19 clinical pediatric patients with HB. HB, hepatoblastoma.

Medical record No	Genetic changes associated with HB paroxysm	Abundance / Amplification multiple	Unknown genetic change	Abundance / Amplification multiple	MSI	TMB	PDL1	Pathological type	β-catenin	Recurrence or metastasis	Survival time	Prognosis
1	**NFE2L2** c.230A>G p.D77G	48.30%	**KMT2D** c.5467G>T p.G1823X **NUP210L** c.4475G>A p.G1492E **ARID1A** c. 2714C>G p.A905G **TRIM23** c.868G>T p.A290S **PTPRB** c.5432A>G p.K1811R **AURKA** amplification	11.70% 20.00% 18.50% 16.80% 11.90% 1.51	MSS	3.9	<1%	Epithelium and lobus intermedius mixed type	Positive	Metastasis	9	PR
2					MSS	0	<1%	Epithelium and lobus intermedius mixed type	Positive	Recurrence	17	CR
3			**ARID1A** c.4764-4770delT p.A1589fs	37.20%	MSS	1.69	1.00%	Fetal and embryonic mixed type	Positive	Recurrence	29	CR
4	**NFE2L2** c.92G>C p.G31A	22.00%	**EP300** c.3244_3245delinsTT p.Q1082L **PARP1** c.2680G>T p.G894N	48.70% 29.00%	MSS	1.69	<1%	Fetal type	Positive	Recurrence	21	CR
5	**CTNNB1** c.101G>T p.G34V	11.60%			MSS	3.18		Epithelium and lobus intermedius mixed type	Positive	Recurrence	23	PR
6					MSS	0		Fetal and embryonic mixed type	Positive	No	12	CR
7	**MYCN** amplification	1.5	**TCF7L2** c.1258C>T p.R420W **MED12** c.5382G>T p.Q1794H	26.00% 24.70%	MSS	2.25	1-2%	Epithelium and lobus intermedius mixed type	Focal positive	Recurrence	22	Dead
8			**USP9X** c.1916C>T p.P639L **NOTCH3** c.3991C>G p.P1331A	37.90% 10.20%	MSS	3.93	2.00%	Embryo and giant trabecular mixed type	Positive	Recurrence	15	PR
9	**CTNNB1** c.1003A>C p.K335Q	45.50%			MSS	1.69		Fetal and embryonic mixed type	Positive	Recurrence	30	Dead
10			**BCL6** c.959A>G p.N320S **HIST1H3F** c.148C>G p.R50G	12.70% 11.90%	MSS	2.25		Fetal type	Positive	Metastasis	21	CR
11	**CTNNB1** c.94G>A p.D32N **TP53** c.743G>A p.R248Q	23.00% 47.30%	**INPP4A** c.959A>G p.N320S **RFWD2** c.173C>G p.S58W	14.80% 13.90%	MSS			Fetal and embryonic mixed type	Positive	Recurrence or metastasis	17	PD
12	**CTNNB1** c.121A>G p.T41A	27.50%			MSS			Fetal and embryonic mixed type	Positive	No	20	CR
13	**AXIN1** c.516_537del CATGAAGCAGCTGATCGATCCT p.I172Mfs*63	13.11%	**EPHA7** c.1909C>T p.R637C **NRAS** c.182A>G p.Q61R	7.16% 6.26%	MSS	22		giant beam type	Positive	Recurrence after PD	50	Dead
14	**APC** c.5465T>A p.V1822D	11.20%	**PIK3C2B** c.533C>T p.P178L **IGF2**	36.50% 31.61%	MSS	5	0	Fetal type	Positive	Recurrence or metastasis	40	PR

(Continued)

TABLE 2 | Continued

Medical record No	Genetic changes associated with HB paroxysm	Abundance / Amplification multiple	Unknown genetic change	Abundance / Amplification multiple	MSI	TMB	PDL1	Pathological type	β-catenin	Recurrence or metastasis	Survival time	Prognosis
			c.517_538dup ACACCTGGAAGCAGTCCACC p.Q180Rfs*78									
			NORCH1 c.1820G>T p.C607F									
			DICER1 c.4082A>G p.K1361R	28.96%								
			c.4064A>G p.N1355S	25.40% 25.06%								
15	**CTNNB1** c.53_241+117del306	47.50%			MSS	0		Epithelium and lobus intermedius mixed type	negative	PD	25	Dead
16	**CTNNB1** c.66_95del p.His24_Ser33del	21.20%	**NUP93** c.180-2A>T	41.40%	MSS	1.34	1-2%	Epithelium and lobus intermedius mixed type	Positive	Recurrence or metastasis	24	CR
			PREX2 c.2741C>G p.A914G	38.30%								
			PAK7 c.1573T>C p.F525L	28.60%								
17			**ATRX** c.2701C>G p.I901V	8.14%	MSS	9		Epithelium and lobus intermedius mixed type	Positive	Recurrence	9	CR
			PTCH1 c.2479A>G p.S827G	5.60%								
			MPL c.173C>T p.A58VG	6.40%								
18	**CTNNB1** c.94G>A p.D32N	14.14%			MSS	4.8		Fetal and embryonic mixed type	Positive	Recurrence	15	CR
19					MSS	1.11		Fetal and embryonic mixed type	Positive	Recurrence	23	CR

Only genetic changes with mutation abundance ≥5% are listed; MSI, microsatellite instability; MSS, microsatellite stability; TMB, tumor mutation burden; CR, complete remission; PR, partial response; PD, disease progression.

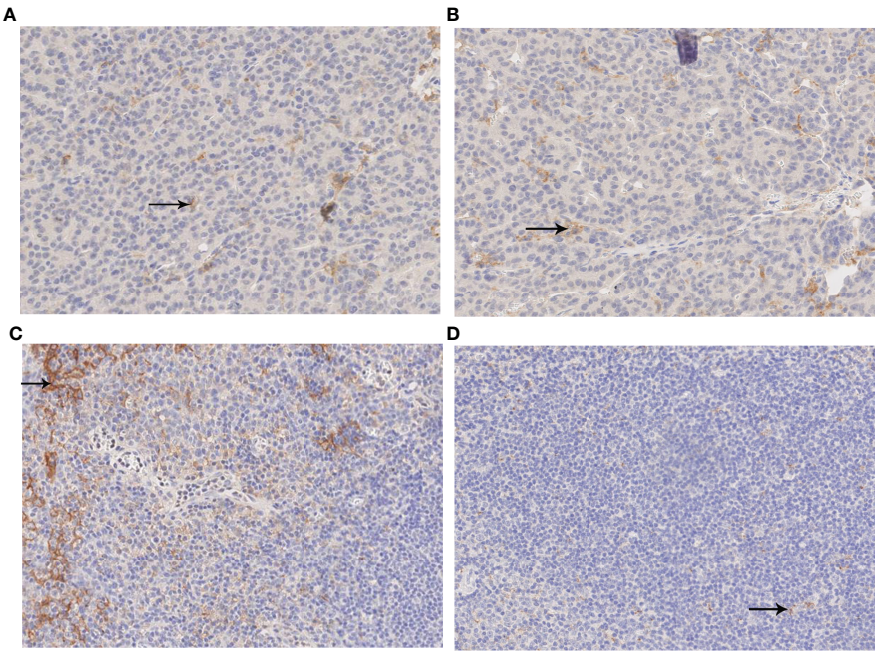


FIGURE 1 | PDL1 immunohistochemical analysis. **(A, B)** Immunohistochemical analysis of PDL1. The percentage of positive-stained cells was 1% (X200); **(C)** Positive control image of immunohistochemical analysis of PDL1 (X200). **(D)** Negative control imaging of immunohistochemical analysis of PDL1 (X200). PDL1, programmed cell death ligand 1.

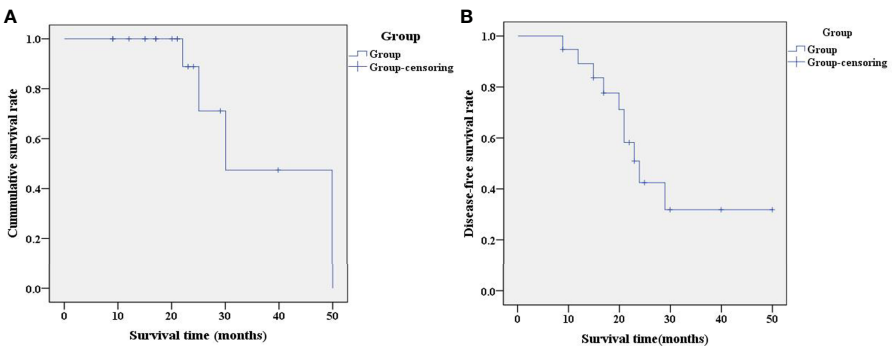


FIGURE 2 | Analysis of overall prognosis by survival curve. **(A)** Survival curve of 19 pediatric patients with HB; **(B)** Disease-free survival curve of 19 pediatric patients with HB. Group-censoring represents the cases that did not survive. HB, hepatoblastoma.

(23, 24). The results of all-exon detection noted in 15 cases with HB indicated that the *CTNNB1* and *NFE2L2* mutations were present in 12 and 2 cases, respectively (12). In addition, it was found that the repeated mutations of *CTNNB1* and the activation of the NFE2L2/kelch-like ECH-associated protein 1 pathway played an important role in the occurrence of HB, which confirmed the stability loss of the genome and the deletions of the telomerase reverse transcriptase promoter as prominent characteristics of aggressive HB with HCC features (12). A previous report examined 27 patients with HB and demonstrated that *CTNNB1* (point mutations and deletion mutations) and *AXIN1* (point mutations) gene mutations accounted for 70.4% and 7.4% of the cases, respectively. This confirmed that the WNT signaling pathway played an important role in the pathogenesis of HB (25). In addition, hepatitis B virus (HBV)-associated HCC is more likely to indicate high frequency of *p53* or *AXIN1* mutations that cause chromosomal instability, while the most common *CTNNB1* gene mutation in HCC is not associated with chronic HBV infection (26, 27). It has been reported that *APC* gene mutations play an important role in sporadic HB and WNT pathway activation (11, 28). In the present study, 7 patients (36.84%) with *CTNNB1* gene alteration

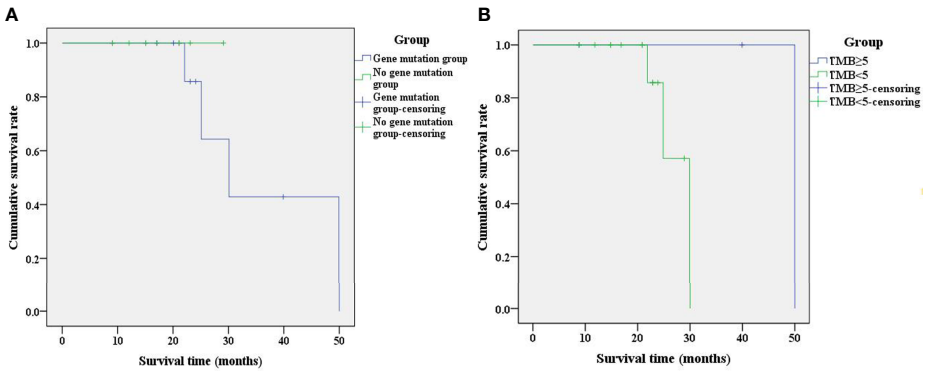


FIGURE 3 | Analysis of prognostic survival curves corresponding to different patient groups. **(A)** Survival curves between the gene abnormality group and the non-reported disease-related gene abnormality group; **(B)** Survival curves of different groups of TMB. The group-censoring represents the cases that did not survive. TMB, tumor mutation burden.

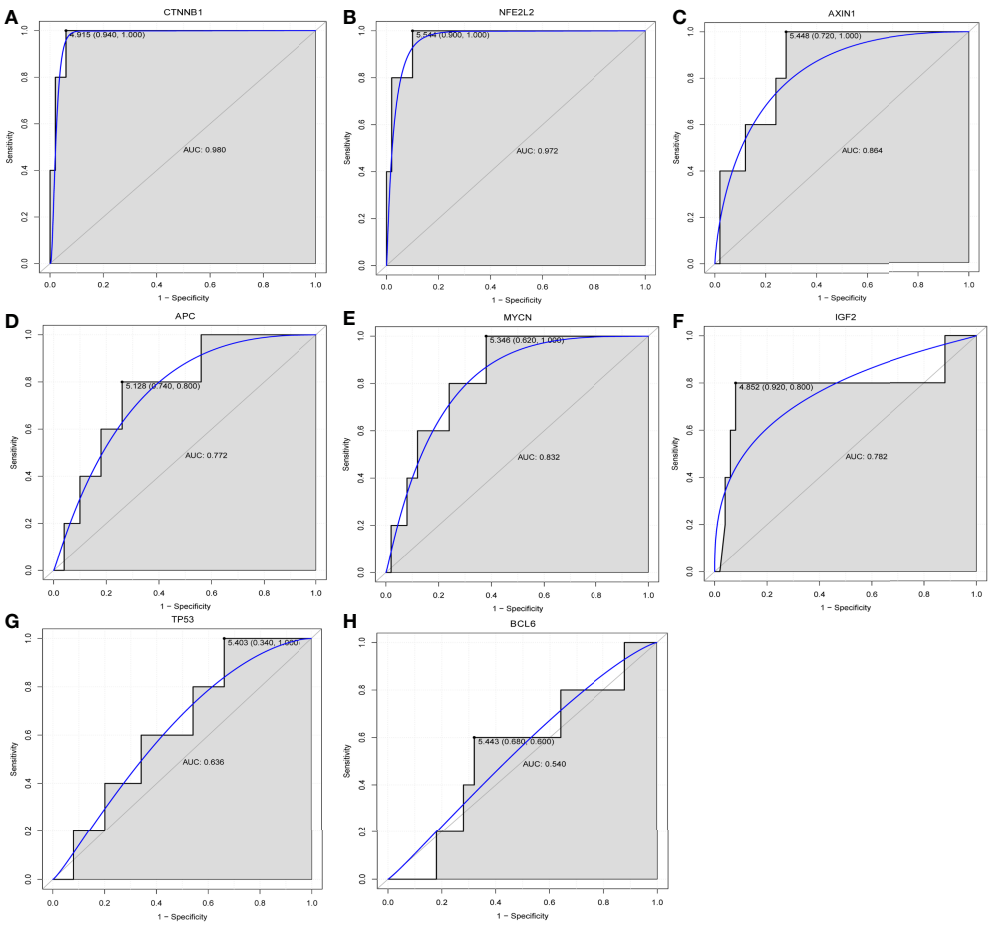


FIGURE 4 | Diagnostic analysis of **(A)** CTNNB1, **(B)** NFE2L2, **(C)** AXIN1, **(D)** APC, **(E)** MYCN, **(F)** IGF2, **(G)** TP53 and **(H)** BCL6. CTNNB1, catenin β1; NFE2L2, factor erythroid 2-related factor 2; AXIN1, axis inhibition protein 1; APC, adenomatous polyposis coli; IGF2, insulin growth factor 2; BCL6, B cell lymphoma 6; AUC, area under curve.

were noted, which suggested a lower percentage than that reported in the literature (10, 21, 25). Single nucleotide variation and gene insertion or deletion of *CTNNB1* was found in 5 and 2 patients, respectively. In addition, 1 patient with *TP53* gene alteration presented with *CTNNB1* gene point mutations. Concomitantly, both *APC* and *AXIN1* exhibited one genetic alteration. Although the proportion of patients with various genetic changes was low, the positive β-catenin expression was present in 18 cases. This indicated that the activation of the WNT pathway played an important role in the occurrence of HB. It is interesting to note that the *CTNNB1* genotype (c.53_241+117del306) of case 15, the *CTNNB1* genotype (c.66_95del p.His24_Ser33del) of case 16 and the *AXIN1* genotype (c.516_537del CATGAAGCAGCTGATCGATCCT p.I172Mfs*63) of case 13 have not been previously reported in the relevant literature. This provides an additional direction for subsequent research.

The *BCL6* gene is closely associated with the pathogenesis of B-cell lymphoma (29). The expression or activity of *BCL6* is decreased in glioblastoma tumor specimens and cell lines, whereas induction of cell apoptosis is increased and proliferation is reduced (30). In addition, *BCL6* can also increase the sensitivity of glioma to targeted therapy (30). The *MYCN* gene is a member of the MYC family of proto-oncogenes, which participates in the development of human and animal cancers by regulating cell proliferation and cell death (31, 32). In neuroblastoma, *MYCN* expansion was associated with poor prognosis and treatment failure (31–33). Previous studies have not examined *MYCN* gene amplification in HB (34, 35). However, previous studies in recent years demonstrated that *MYCN* expression was significantly increased in HB, whereas *MYCN* knockdown inhibited the proliferation of HB cells (36, 37). This indicated that *BCL6* and *MYCN* may have an important regulatory role in the development of HB.

The prognosis of HB is associated with several factors, such as initial AFP level, age, co-morbidity and pathological subtype (38). At present, the research on the genotypes of HB-associated disease is mostly associated with the research of targeted therapy (39, 40). In the present study, the CR rate of the gene abnormality group was lower than that of the non-reported disease-associated gene abnormality group. No significant differences were noted (P=0.013). Notably, the patients with *TP53* gene mutation presented with *MYCN* gene amplification and *AXIN1* gene deletion and demonstrated a poor prognosis. In addition, the 2-year overall survival rates of patients in the gene abnormality group and the non-reported disease-associated gene abnormality group were 64.3% and 100%, respectively. Although no significant differences were noted between the two groups (chi-squared=0.536, P=0.464), the survival rate of the patients in the gene abnormality group was lower than that of the non-reported disease-associated gene abnormality group. This further suggested that changes in the genotype of different genes can be used to predict the prognosis of patients with HB.

In the present study, TMB was assessed in 17 patients. The range and median of TMB were 0-22 mut/Mb and 2.25 mut/Mb, respectively. TMB was present in only 1 patient with HB, who reached 22 mut/Mb. The microsatellite state analysis indicated that all of the 19 patients did not exhibit microsatellite instability. This result is consistent with a previous study, which demonstrated that recurrent and metastatic HB exhibited a lower TMB (5). The interaction of PDL1 with its receptor programmed cell death 1 (PD1) inhibited T cell activity (41). Various types of cancer express high levels of the PDL1 protein. The PDL1/PD1 signaling pathway is activated to evade T-cell immunity (42). Inhibition of the PDL1/PD1 pathway can enhance T cell response and mediate antitumor activity (43). The expression of the PDL1 protein can be used as a biomarker for predicting which patients are more likely to respond to immunotherapy (44). Immunohistochemical detection of PDL1 in 8 patients with HB demonstrated low percentage of PDL1 positive expression (≤2%), which was consistent with previously reported results (45). In addition, the expression levels of PD-L1 in common solid tumors of pediatric patients are generally deficient (46). Although immunohistochemical analysis of PD-L1 can be used to predict the treatment response of patients receiving anti-PD1 or anti-PDL1 therapy, certain patients who present with negative PD-L1 expression may also benefit from immunotherapy (47). In the present study, immunohistochemical detection of PDL1 expression in patients with HB indicated a low percentage of PDL1, which may be caused by the small sample size. Additional verification in larger sample-size studies is required. In addition, a limited number of studies have been performed on the immunohistochemical detection of PDL1 expression in patients with HB. Whether immunotherapy is effective in the treatment of HB remains to be further confirmed.

The present study contains certain limitations. Firstly, the sample size was small, which led to a certain degree of error. Additional patients are required to expand the sample size. Secondly, the follow-up time of each group was inconsistent and the follow-up time period had to be adjusted for further confirmation. Thirdly, the molecular mechanisms of the identified mutant genes in HB remain unclear and should be further studied.

In the present study, statistical methods were used to analyze the genotype characteristics, clinical data, clinical efficacy and prognosis of 19 pediatric patients with HB. The results of the present study demonstrated that different genotypes may play an important regulatory role in the physiology and pathology of HB, which is helpful for the assessment of the clinical prognosis and the application of targeted drugs and immunotherapy. In addition, the data indicated that CTNNB1, NFE2L2, AXIN1, APC, MYCN and IGF2 may be potential biomarkers that can be used for the diagnosis of HB.

DATA AVAILABILITY STATEMENT

The data presented in the study are deposited in the SRA repository, accession number PRJNA751247.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Medical Ethics Committee of the Beijing Tongren

Hospital, Capital Medical University (TRECKY2019-033). Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

Conception and design, HH, WZ and DH. Administrative support, DH. Provision of materials and samples, WZ and DH. Data collection and collation, HH, TZ and JL. Data analysis and interpretation, YW, FL and YM. All authors contributed to the article and approved the submitted version.

FUNDING

This present study was funded by the "Dengfeng" Talent Cultivation Plan of Beijing Municipal Hospital Administration (Project Number: DFL20180201), the Pediatric Medical Coordinated Development Center of Beijing Municipal Administration of Hospitals (No: XTZD20180205) and the Beijing Yizhuang Economic and Technological Development Zone Leading Program [No: (2017)-8].

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.628531/full#supplementary-material

Supplementary Figure 1 | Electronic verifications of **(A)** CTNNB1, **(B)** NFE2L2, **(C)** AXIN1, **(D)** APC, **(E)** MYCN, **(F)** IGF2, **(G)** TP53 and **(H)** BCL6. CTNNB1, catenin β1; NFE2L2, factor erythroid 2-related factor 2; AXIN1, axis inhibition protein 1; APC, adenomatous polyposis coli; IGF2, insulin growth factor 2; control, normal control group; case, hepatoblastoma group.

REFERENCES

1. Meyers RL, Maibach R, Hiyama E, Häberle B, Krailo M, Rangaswami A, et al. Risk-Stratified Staging in Paediatric Hepatoblastoma: A Unified Analysis From the Children's Hepatic Tumors International Collaboration. *Lancet Oncol* (2017) 18:122–31. doi: 10.1016/S1470-2045(16)30598-8
2. Marin JJG, Cives-Losada C, Asensio M, Lozano E, Briz O, Macias RIR. Mechanisms of Anticancer Drug Resistance in Hepatoblastoma. *Cancers (Basel)* (2019) 11:407. doi: 10.3390/cancers11030407
3. Shen G, Shen H, Zhang J, Yan Q, Liu H. DNA Methylation in Hepatoblastoma-A Literature Review. *Ital J Pediatr* (2020) 46:113. doi: 10.1186/s13052-020-00877-6
4. Thyagarajan MS, Sharif K. Space Occupying Lesions in the Liver. *Indian J Pediatr* (2016) 83:1291–302. doi: 10.1007/s12098-016-2240-x
5. ee H, El Jabbour T, Ainechi S, Gay LM, Elvin JA, Vergilio JA, et al. General Paucity of Genomic Alteration and Low Tumor Mutation Burden in Refractory and Metastatic Hepatoblastoma: Comprehensive Genomic Profiling Study. *Hum Pathol* (2017) 70:84–91. doi: 10.1016/j.humpath.2017.10.007
6. McLaughlin CC, Baptiste MS, Schymura MJ, Nasca PC, Zdeb MS. Maternal and Infant Birth Characteristics and Hepatoblastoma. *Am J Epidemiol* (2006) 163:818–28. doi: 10.1093/aje/kwj104
7. Tanimura M, Matsui I, Abe J, Ikeda H, Kobayashi N, Ohira M, et al. Increased Risk of Hepatoblastoma Among Immature Children With a Lower Birth Weight. *Cancer Res* (1998) 58:3032–5. doi: http://cancerres.aacrjournals.org/content/58/14/3032
8. Weksberg R, Shuman C, Smith AC. Beckwith-Wiedemann Syndrome. *Am J Med Genet C Semin Med Genet* (2005) 137c:12–23. doi: 10.1002/ajmg.c.30058
9. Trobaugh-Lotrario AD, Venkatramani R, Feusner JH. Hepatoblastoma in Children With Beckwith-Wiedemann Syndrome: Does it Warrant Different Treatment? *J Pediatr Hematol Oncol* (2014) 36:369–73. doi: 10.1097/MPH.0000000000000129
10. Aguiar TFM, Carneiro TN. The Genetic and Epigenetic Landscapes of Hepatoblastomas. *Appl Cancer Res* (2017) 37:20. doi: 10.1186/s41241-017-0021-0
11. Hiyama E, Kurihara S, Onitake Y, Morihara N, Ikeda K, Hiyama K. Abstract 5188: Integrated Exome Analysis in Childhood Hepatoblastoma: Biological Approach for Next Clinical Trial Designs. *Cancer Res* (2014) 74:5188–8. doi: 10.1158/1538-7445.AM2014-5188
12. Eichenmüller M, Trippel F, Kreuder M, Beck A, Schwarzmayr T, Häberle B, et al. The Genomic Landscape of Hepatoblastoma and Their Progenies With HCC-Like Features. *J Hepatol* (2014) 61:1312–20. doi: 10.1016/j.jhep.2014.08.009
13. Bolger AM, Lohse M, Usadel B. Trimmomatic: A Flexible Trimmer for Illumina Sequence Data. *Bioinformatics* (2014) 30:2114–20. doi: 10.1093/bioinformatics/btu170
14. Li H, Durbin R. Fast and Accurate Short Read Alignment With Burrows-Wheeler Transform. *Bioinformatics* (2009) 25:1754–60. doi: 10.1093/bioinformatics/btp324
15. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: A MapReduce Framework for Analyzing Next-Generation DNA Sequencing Data. *Genome Res* (2010) 20:1297–303. doi: 10.1101/gr.107524.110
16. Wang K, Li M, Hakonarson H. ANNOVAR: Functional Annotation of Genetic Variants From High-Throughput Sequencing Data. *Nucleic Acids Res* (2010) 38:e164. doi: 10.1093/nar/gkq603
17. De' Angelis GL, Bottarelli L, Azzoni C, De' Angelis N, Leandro G, Di Mario F, et al. Microsatellite Instability in Colorectal Cancer. *Acta BioMed* (2018) 89:97–101. doi: 10.23750/abm.v89i9-S.7960
18. Salipante SJ, Scroggins SM, Hampel HL, Turner EH, Pritchard CC. Microsatellite Instability Detection by Next Generation Sequencing. *Clin Chem* (2014) 60:1192–9. doi: 10.1373/clinchem.2014.223677
19. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI Gene Expression and Hybridization Array Data Repository. *Nucleic Acids Res* (2002) 30:207–10. doi: 10.1093/nar/30.1.207
20. Šimundić AM. Measures of Diagnostic Accuracy: Basic Definitions. *Ejifcc* (2009) 19:203–11.
21. Luo M, Conrad M, Lin HC. Genetics of Pediatric Hepatoblastoma and Hepatocellular Carcinoma and Their Clinical Application. *Am J Digest Dis* (2014) 1(2):97–111.
22. Meyers RL. Tumors of the Liver in Children. *Surg Oncol* (2007) 16:195–203. doi: 10.1016/j.suronc.2007.07.002
23. Zhan P, Ji YN, Yu LK. TP53 Mutation is Associated With a Poor Outcome for Patients With Hepatocellular Carcinoma: Evidence From a Meta-Analysis. *Hepatobiliary Surg Nutr* (2013) 2:260–5. doi: 10.3978/j.issn.2304-3881.2013.07.06
24. Liu J, Ma Q, Zhang M, Wang X, Zhang D, Li W, et al. Alterations of TP53 are Associated With a Poor Outcome for Patients With Hepatocellular Carcinoma: Evidence From a Systematic Review and Meta-Analysis. *Eur J Cancer* (2012) 48:2328–38. doi: 10.1016/j.ejca.2012.03.001
25. Taniguchi K, Roberts LR, Aderca IN, Dong X, Qian C, Murphy LM, et al. Mutational Spectrum of Beta-Catenin, AXIN1, and AXIN2 in Hepatocellular Carcinomas and Hepatoblastomas. *Oncogene* (2002) 21:4863–71. doi: 10.1038/sj.onc.1205591
26. Huang H, Fujii H, Sankila A, Mahler-Araujo BM, Matsuda M, Cathomas G, et al. Beta-Catenin Mutations Are Frequent in Human Hepatocellular Carcinomas Associated With Hepatitis C Virus Infection. *Am J Pathol* (1999) 155:1795–801. doi: 10.1016/S0002-9440(10)65496-X
27. Laurent-Puig P, Legoix P, Bluteau O, Belghiti J, Franco D, Binot F, et al. Genetic Alterations Associated With Hepatocellular Carcinomas Define Distinct Pathways of Hepatocarcinogenesis. *Gastroenterology* (2001) 120:1763–73. doi: 10.1053/gast.2001.24798

28. Yang A, Sisson R, Gupta A, Tiao G, Geller JI. Germline APC Mutations in Hepatoblastoma. *Pediatr Blood Cancer* (2018) 65:e26892. doi: 10.1002/pbc.26892
29. Yang H, Green MR. Epigenetic Programing of B-Cell Lymphoma by BCL6 and Its Genetic Deregulation. *Front Cell Dev Biol* (2019) 7:272. doi: 10.3389/fcell.2019.00272
30. Fabre MS, Stanton NM, Slatter TL, Lee S, Senanayake D, Gordon RMA. The Oncogene BCL6 Is Up-Regulated in Glioblastoma in Response to DNA Damage. *Drives Survival After Ther* (2020) 15:e0231470. doi: 10.1371/journal.pone.0231470
31. Schwab M. MYCN in Neuronal Tumours. *Cancer Lett* (2004) 204:179–87. doi: 10.1016/S0304-3835(03)00454-3
32. Ruiz-Pérez MV, Henley AB, Arsenian-Henriksson M. The MYCN Protein in Health and Disease. *Genes (Basel)* (2017) 8:113. doi: 10.3390/genes8040113
33. Mertens F, Mandahl N, Mitelman F, Heim S. Cytogenetic Analysis in the Examination of Solid Tumors in Children. *Pediatr Hematol Oncol* (1994) 11:361–77. doi: 10.3109/08880019409140536
34. Tsuda H, Shimosato Y, Upton MP, Yokota J, Terada M, Ohira M, et al. Retrospective Study on Amplification of N-Myc and C-Myc Genes in Pediatric Solid Tumors and Its Association With Prognosis and Tumor Differentiation. *Lab Invest* (1988) 59:321–7. doi: 10.1084/jem.168.3.1205
35. Mares J, Polanská V, Görgens H, Sedlácek Z, Maríková T, Bocek P, et al. Oncogene Amplification and Expression in Pediatric Solid Tumors. *Neoplasma* (1998) 45:123–7. doi: 10.1021/cr9500747
36. Shin E, Lee KB, Park SY, Kim SH, Ryu HS, Park YN, et al. Gene Expression Profiling of Human Hepatoblastoma Using Archived Formalin-Fixed and Paraffin-Embedded Tissues. *Virchows Arch* (2011) 458:453–65. doi: 10.1007/s00428-011-1043-8
37. Eberherr C, Beck A, Vokuhl C, Becker K, Häberle B, Von Schweinitz D, et al. Targeting Excessive MYCN Expression Using MLN8237 and JQ1 Impairs the Growth of Hepatoblastoma Cells. *Int J Oncol* (2019) 54:1853–63. doi: 10.3892/ijo.2019.4741
38. Czauderna P, Haeberle B, Hiyama E, Rangaswami A, Krailo M, Maibach R, et al. The Children's Hepatic Tumors International Collaboration (CHIC): Novel Global Rare Tumor Database Yields New Prognostic Factors in Hepatoblastoma and Becomes a Research Model. *Eur J Cancer* (2016) 52:92–101. doi: 10.1016/j.ejca.2015.09.023
39. Rokita JL, Rathi KS, Cardenas MF, Upton KA, Jayaseelan J, Cross KL, et al. Genomic Profiling of Childhood Tumor Patient-Derived Xenograft Models to Enable Rational Clinical Trial Design. *Cell Rep* (2019) 29:1675–1689.e1679. doi: 10.1016/j.celrep.2019.09.071
40. Khater F, Vairy S, Langlois S, Dumoucel S, Sontag T, St-Onge P, et al. Molecular Profiling of Hard-To-Treat Childhood and Adolescent Cancers. *JAMA Netw Open* (2019) 2:e192906. doi: 10.1001/jamanetworkopen.2019.2906
41. Daassi D, Mahoney KM. The Importance of Exosomal PDL1 In Tumour Immune Evasion. *Nat Rev Immunol* (2020) 20:209–15. doi: 10.1038/s41577-019-0264-y
42. Cha JH, Chan LC, Li CW, Hsu JL, Hung MC. Mechanisms Controlling PD-L1 Expression in Cancer. *Mol Cell* (2019) 76:359–70. doi: 10.1016/j.molcel.2019.09.030
43. Dermani FK, Samadi P, Rahmani G, Kohlan AK, Najafi R. PD-1/PD-L1 Immune Checkpoint: Potential Target for Cancer Therapy. *J Cell Physiol* (2019) 234:1313–25. doi: 10.1002/jcp.27172
44. Yu H, Boyle TA, Zhou C, Rimm DL, Hirsch FR. PD-L1 Expression in Lung Cancer. *J Thorac Oncol* (2016) 11:964–75. doi: 10.1016/j.jtho.2016.04.014
45. Aoki T, Hino M, Koh K, Kyushiki M, Kishimoto H, Arakawa Y, et al. Low Frequency of Programmed Death Ligand 1 Expression in Pediatric Cancers. *Pediatr Blood Cancer* (2016) 63:1461–4. doi: 10.1002/pbc.26018
46. Davis KL, Fox E, Merchant MS, Reid JM, Kudgus RA, Liu X, et al. Nivolumab in Children and Young Adults With Relapsed or Refractory Solid Tumours or Lymphoma (ADVL1412): A Multicentre, Open-Label, Single-Arm, Phase 1-2 Trial. *Lancet Oncol* (2020) 21:541–50. doi: 10.1016/S1470-2045(20)30023-1
47. Ancevski Hunter K, Socinski MA, Villaruz LC. PD-L1 Testing in Guiding Patient Selection for PD-1/PD-L1 Inhibitor Therapy in Lung Cancer. *Mol Diagn Ther* (2018) 22:1–10. doi: 10.1007/s40291-017-0308-6

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Pediatrics

ORIGINAL RESEARCH
published: 26 August 2021
doi: 10.3389/fped.2021.663054

Identification of CDC20 as a Novel Biomarker in Diagnosis and Treatment of Wilms Tumor

Qinlin Shi[1,2†], *Bo Tang*[1,2†], *Yanping Li*[1,2], *Yonglin Li*[1,2], *Tao Lin*[1,2], *Dawei He*[1,2*] and *Guanghui Wei*[1,2*]

[1] Ministry of Education Key Laboratory of Child Development and Disorders, Chongqing Key Laboratory of Pediatrics, Chongqing Key Laboratory of Children Urogenital Development and Tissue Engineering, China International Science and Technology Cooperation Base of Child Development and Critical Disorders, Pediatric Research Institute, Children's Hospital of Chongqing Medical University, Chongqing, China, [2] Department of Pediatric Urology Surgery, Children's Hospital of Chongqing Medical University, Chongqing, China

Objective: Wilms tumor (WT) is a common malignant solid tumor in children. Many tumor biomarkers have been reported; however, there are poorly targetable molecular mechanisms which have been defined in WT. This study aimed to identify the oncogene in WT and explore the potential mechanisms.

Methods: Differentially expressed genes (DEGs) in three independent RNA-seq datasets were downloaded from The Cancer Genome Atlas data portal and the Gene Expression Omnibus database (GSE66405 and GSE73209). The common DEGs were then subjected to Gene Ontology enrichment analysis, protein–protein interaction (PPI) network analysis, and gene set enrichment analysis. The protein expression levels of the hub gene were analyzed by immunohistochemical analysis and Western blotting in a 60 WT sample. The univariate Kaplan–Meier analysis for overall survival was performed, and the log-rank test was utilized. A small interfering RNA targeting cell division cycle 20 (CDC20) was transfected into G401 and SK-NEP-1 cell lines. The Cell Counting Kit-8 assay and wound healing assay were used to observe the changes in cell proliferation and migration after transfection. Flow cytometry was used to detect the effect on the cell cycle. Western blot was conducted to study the changes of related functional proteins.

Results: We commonly identified 44 upregulation and 272 downregulation differentially expressed genes in three independent RNA-seq datasets. Gene and pathway enrichment analyses of the regulatory networks involving hub genes suggested that cell cycle changes are crucial in WT. The top 15 highly connected genes were found by PPI network analysis. Furthermore, we demonstrated that one candidate biomarker, CDC20, for the diagnosis of WT was detected, and its high expression predicted poor prognosis of WT patients. Moreover, the area under the curve value obtained by receiver operating characteristic curve analysis from paired WT samples was 0.9181. Finally, we found that the suppression of CDC20 inhibited proliferation and migration and resulted in G2/M phase arrest in WT cells. The mechanism may be involved in increasing the protein level of securin, cyclin B1, and cyclin A

OPEN ACCESS

Edited by:
Jing He,
Guangzhou Medical University, China

Reviewed by:
Fang Chen,
Shanghai Children's Hospital, China
Daniel Green,
Kite Pharma, United States

***Correspondence:**
Dawei He
400116@hospital.cqmu.edu.cn
Guanghui Wei
u806806@cqmu.edu.cn

[†] *These authors have contributed equally to this work*

Specialty section:
This article was submitted to Pediatric Oncology, a section of the journal Frontiers in Pediatrics

Received: *02 February 2021*
Accepted: *24 June 2021*
Published: *26 August 2021*

Citation:
Shi Q, Tang B, Li Y, Li Y, Lin T, He D and Wei G (2021) Identification of CDC20 as a Novel Biomarker in Diagnosis and Treatment of Wilms Tumor. Front. Pediatr. 9:663054. doi: 10.3389/fped.2021.663054

Conclusion: Our results suggest that CDC20 could serve as a candidate diagnostic and prognostic biomarker for WT, and suppression of CDC20 may be a potential approach for the prevention and treatment of WT.

Keywords: Wilms tumor, biomarkers, cell division cycle 20, diagnosis, cell proliferation

INTRODUCTION

Wilms tumor (WT) is a common pediatric solid retroperitoneal tumor. The incidence of WT was ∼6 per 100,000 to 7 per 100,000 for children younger than 15 years (1, 2). Thanks to the continuous efforts by the Children's Oncology Group and the National Wilms Tumor Society (NWTS), the overall survival rate of WT has improved from 30 to 90% in the last 30 years (3). However, some cases still result in poor outcomes, which is associated with metastasis, recurrence, anaplastic WT, and chemoradiotherapy resistance (4). Moreover, chronic health conditions secondary to treatment impact nearly one quarter of survivors of WT and include renal failure, infertility, cardiac toxicity, restrictive pulmonary disease, and the development of subsequent malignancies (5, 6). Hence, finding a novel strategy for the diagnosis and treatment of WT has become a hotspot in recent years. Most research on WT biomarkers has focused on the genetic components of WT development including WT1, WTX, MYCN, CTNNB1, SIX1/SIX2, TP53, loss of heterozygosity 11p15, 16q, and 1p and 1q gain of function (7–9). A recent whole-exome study has identified that DROSHA and DICER1 mutations impair expression of tumor-suppressing miRNAs (10). Unfortunately, the frequency of alterations in genes is similarly uncommon, and there is no clear gene for clinical application (11).

The Gene Expression Omnibus (GEO) is an international public repository that archives and freely distributes microarray, next-generation sequencing, and The Cancer Genome Atlas (TCGA) is a large-scale cancer genome project that provides researchers with multidimensional maps of the key genomic changes (12, 13). Both GEO and TCGA have significantly increased our understanding of cancer. Therefore, in this study, we first identified the common differentially expressed genes (DEGs) from multiple microarrays and TCGA WT RNA-sequence dataset. The upregulation DEGs were then subjected to Gene Ontology (GO) enrichment analysis, protein–protein interaction (PPI) network analysis, and gene set enrichment analysis (GSEA). According the bioinformatics results, one candidate biomarker, cell cycle 20 (CDC20) (cell division cycle 20 homolog, also called Fizzy), was performed to detect the expression level in 60 paired WT samples. Receiver operating characteristic (ROC) analysis and Kaplan–Meier (KM) analysis were performed to identify diagnostic and prognosis markers for WT. In addition, we predicted and verified the effect of knockdown of CDC20 on WT cell lines. CDC20 small interfering RNA (siRNA) can knock down CDC20 expression at protein levels and thereby lead to cell cycle arrest in the G2/M phase in WT cells. Taken together, the present findings provide more valuable strategies for the diagnosis and treatment of patients with WT.

MATERIALS AND METHODS

Study Population

RNA-sequence data for WT patients were downloaded from the TCGA data portal (https://tcga-data.nci.nih.gov/tcga/) and the GEO database (GSE66405 and GSE73209, http://www.ncbi.nlm.nih.gov/geo), which contains 184 WT tissues and 12 adjacent non-tumor tissues. The TCGA Target-WT sample clinic data were downloaded using package 'TCGAbiolinks' in R.

DEG Analysis

GEO database (GSE66405 and GSE73209) and TCGA database analyses of DEGs between WT and their non-tumor counterparts were performed using package "DESeq2" in R. The DEGs were screened using $p < 0.05$ and $|\log FC| > 1.5$ as the thresholds. Next, heatmaps and volcano plots based on the upregulated and downregulated genes in each dataset were plotted using the "pheatmap" and "ggplots" package of R software. Then, the downregulated and upregulated genes on the three databases were intersected using the "gridBase" and "VennDiagram" package of R software.

GO Enrichment Analysis

GO enrichment analysis was performed using Database for Annotation, Visualization, and Integrated Discovery (DAVID; http://david.abcc.ncifcrf.gov/). The DAVID tool was used for obtaining the enriched GO terms of differentially expressed mRNA genes based on the hypergeometric distribution to compute values, which was described in a previous study (14). The enriched biological processes (BPs), cellular component (CC), and molecular function (MF) were obtained to analyze the common DEGs at the functional level. $p < 0.05$ was set as the threshold value.

PPI Network Construction and Pathway Analysis

STRING (Search Tool for the Retrieval of Interacting Genes/Proteins, http://string-db.org/) is a biological database and Web resource of known and predicted PPIs. Based on the STRING database, PPIs of DEGs were selected with a score (median confidence) of >0.7, and the PPI network was then visualized by Cytoscape (http://www.cytoscape.org/). The hub protein was selected based on its association with other proteins. The DEGs with more association with other DEGs indicate important roles in the PPI network. In addition, the CDC20

single GSEA was performed using the "clusterProfiler," "ggplots" R package of R software.

Patient Tissue

We obtained WT tissues and adjacent kidney tissues from 60 patients who underwent surgery for WT at the Department of Urology Surgery of the Children's Hospital of Chongqing Medical University from January 2015 to January 2020. All specimens were histopathologically identified as WT, and all WT tissues were classified according to the American National Wilms Tumor Study 5 (NWTS-5) typing and TNM staging system by pathologists at the Children's Hospital of Chongqing Medical University who were blinded to the results. After the specimens were extracted, they were placed immediately in liquid nitrogen and for further examination by immunohistochemical (IHC) analysis and Western blotting (WB) experiments.

Cell Lines and Cell Culture

The human Wilms cell lines (G401 and SK-NEP-1) were purchased from the American Type Culture Collection. Both SK-NEP-1 and G401 cells were maintained in McCoy 5A medium (Sigma-Aldrich, Shanghai, China) and supplemented with 15% fetal bovine serum (FBS) and 1% penicillin/streptomycin (Gibco, NY, USA); the cells were cultured at 37°C in a humidified atmosphere with 5% CO_2.

siRNA Transfection

Three segments of CDC20 siRNA and a negative control (NC) were synthesized and purified by Guangzhou RuiBo Company (Guangzhou, China). Target sequences for siRNAs were ACCAACCCAUCACCUCAGU tt ACUGAGGUGAUGGG UUGGU tt (CDC20 si1), GGAGCUCAUCUCAGGC-CAU tt AUGGCCUGAGAUGAGCUCC tt (CDC20 si2), and CAAGA AGGAA-CAUCAGAAA tt UUUCUGAUGUUCCUUCUUG tt (CDC20 si). The G401 and SK-NEP-1 cells were plated onto 6- or 12-well plates and transiently transfected using Lipofectamine™ RNAiMAX (Invitrogen, USA) according to the manufacturer's protocol.

Cell Proliferation and Migration

Cell Counting Kit-8 (CCK-8) assays (Dojindo, Japan) were performed to determine cell proliferation. Approximately 1×10^4 G401 or SK-NEP-1 cells were seeded into 96-well plates and transfected with si-CDC20-1, si-CDC20-2, si-CDC20-3, or NC oligonucleotides. At the indicated time points (hours 0, 24, 48, and 72), the culture medium was removed, and 100 μL of CCK-8 medium was added to each well. The cells were incubated for an additional 4 h, and the optical density was measured at an absorbance wavelength of 450 nm on a microplate reader (Bio-Rad, USA).

Wound healing assays were used to evaluate cell migration. Briefly, G401 cells were seeded in 6-well plates and incubated for 24 h, followed by transfection with an si-CDC20-1, si-CDC20-2, si-CDC20-3, or NC oligonucleotides. Then, scratching was performed with 10-μL pipette tips when the cell confluence reached 100%. Next, the cells were washed several times with phosphate-buffered saline (PBS) to remove the floating cells, and the medium was replaced with fresh cell culture medium without FBS. Images were taken of non-overlapping fields in each well at 0, 24, and 48 h after the scratching step using ImageJ software (http://imagej.en.softonic.com).

Cell Cycle Analysis

The transfected cells were detached by EDTA-free trypsin (Gibco, NY, USA), washed with precooled PBS, and fixed in 75% ethanol at 4°C overnight. The cells were resuspended in 0.2 mL of PI/RNase Staining Buffer (BD Biosciences, Shanghai, China) and incubated in the dark for 30 min. The cells were analyzed using a flow cytometer (BD Biosciences).

Immunohistochemistry

Imunohistochemistry studies were performed on formalin-fixed, paraffin-embedded WT and adjacent tissue sections obtained from untreated patients with WT according to standard procedures. Briefly, 4-μm-thick paraffin sections were deparaffinized and rehydrated, and antigen retrieval was performed. Then, the sections were incubated with 3% H_2O_2 and 0.5% bull serum albumin (BSA). The primary antibodies used were a CDC20 rabbit antibody (1:200, Absin, Shanghai, China). Histochemistry score [H score $= \sum$ (PI $\times$ I) = (percentage of cells of weak intensity $\times$ 1) + (percentage of cells of moderate intensity $\times$ 2) + (percentage of cells of strong intensity $\times$ 3)] (15) was obtained with Quant Center Analysis tool.

Western Blot

Total protein was extracted from tissues and transfected cells using radioimmunoprecipitation assay lysis buffer (Beyotime, China) supplemented with phenylmethanesulfonyl fluoride, and the concentrations were determined by bicinchoninic acid assay. Following protein extraction, sodium dodecyl sulfate–polyacrylamide gel electrophoresis was performed. Then, the electrophoretic bands were transferred to polyvinylidene fluoride membranes (Millipore, USA). Next, the membranes were incubated in 5% BSA (ZSGB-BIO, Beijing, China)–Tris-buffered saline with Tween 20 for 1 h. We used a CDC20 rabbit antibody (1:1,000, Absin, China), securin (1:5,000, Abcam, Shanghai, China), cyclin B1 (1:3,000, Abcam, USA), and cyclin A (1:2,000, Abcam, USA) and GAPDH mouse antibody (1:800, ZSGB-BIO, China) as primary antibodies. After incubating the membranes with primary antibodies and the corresponding secondary antibodies, we detected positive bands with a chemiluminescent reaction. Image collection and densitometry analysis were executed with Quantity One (Bio-Rad, Shanghai, China).

Statistical Analysis

The KM analysis for overall survival proceeded based on the gene's expression level, the cutoff level of which was set at the median value with the aid of GraphPad Prism 7 software and the log-rank test was utilized. One-way analysis of variance and two-tailed Student t-tests were used for expression data comparisons by using GraphPad Prism 7 software. Each experiment was repeated three times or more, and all data were presented as mean ± standard deviation (SD). Statistical significance was described

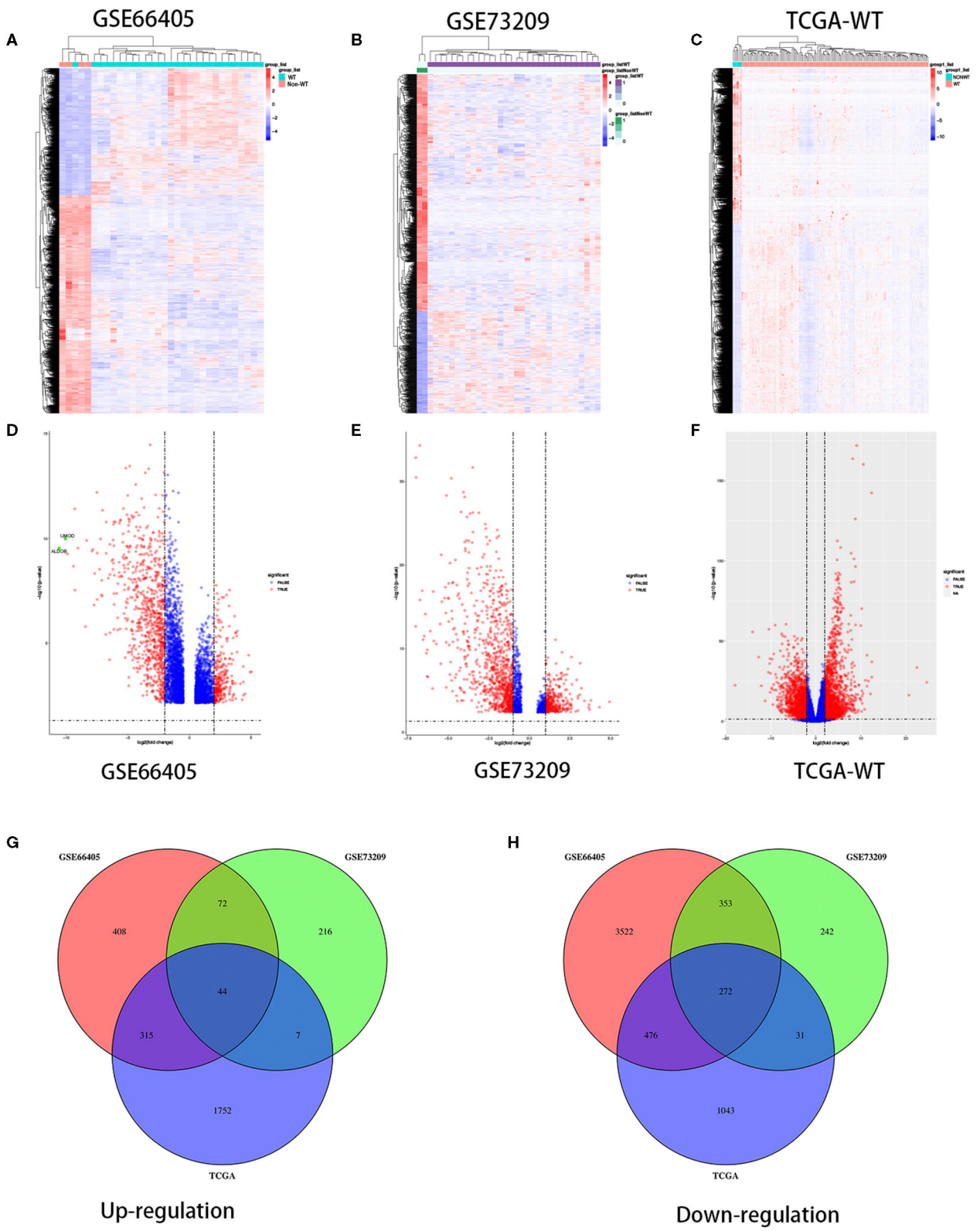


FIGURE 1 | Differentially expressed genes in three independent datasets. Heatmaps of the DEGs in the WT gene expression datasets GSE66405, GSE73209, and TCGA-WT, respectively **(A–C)**. Volcano plots of genes that are significantly different between WT tissues and normal controls in datasets GSE66405, GSE73209, and TCGA-WT, respectively **(D–F)**. X axis indicates the fold change (log-scaled), whereas the Y axis shows the *p*-values (log-scaled). Each symbol represents a different gene, and the red color of the symbols categorizes the upregulated/downregulated genes falling under different criteria (*p* value and fold-change threshold). $p < 0.05$ is considered as statistically significant, whereas fold change = 1.5 is set as the threshold **(D–F)**. The common differentially expressed genes among GSE66405, GSE73209, and TCGA **(G,H)**.

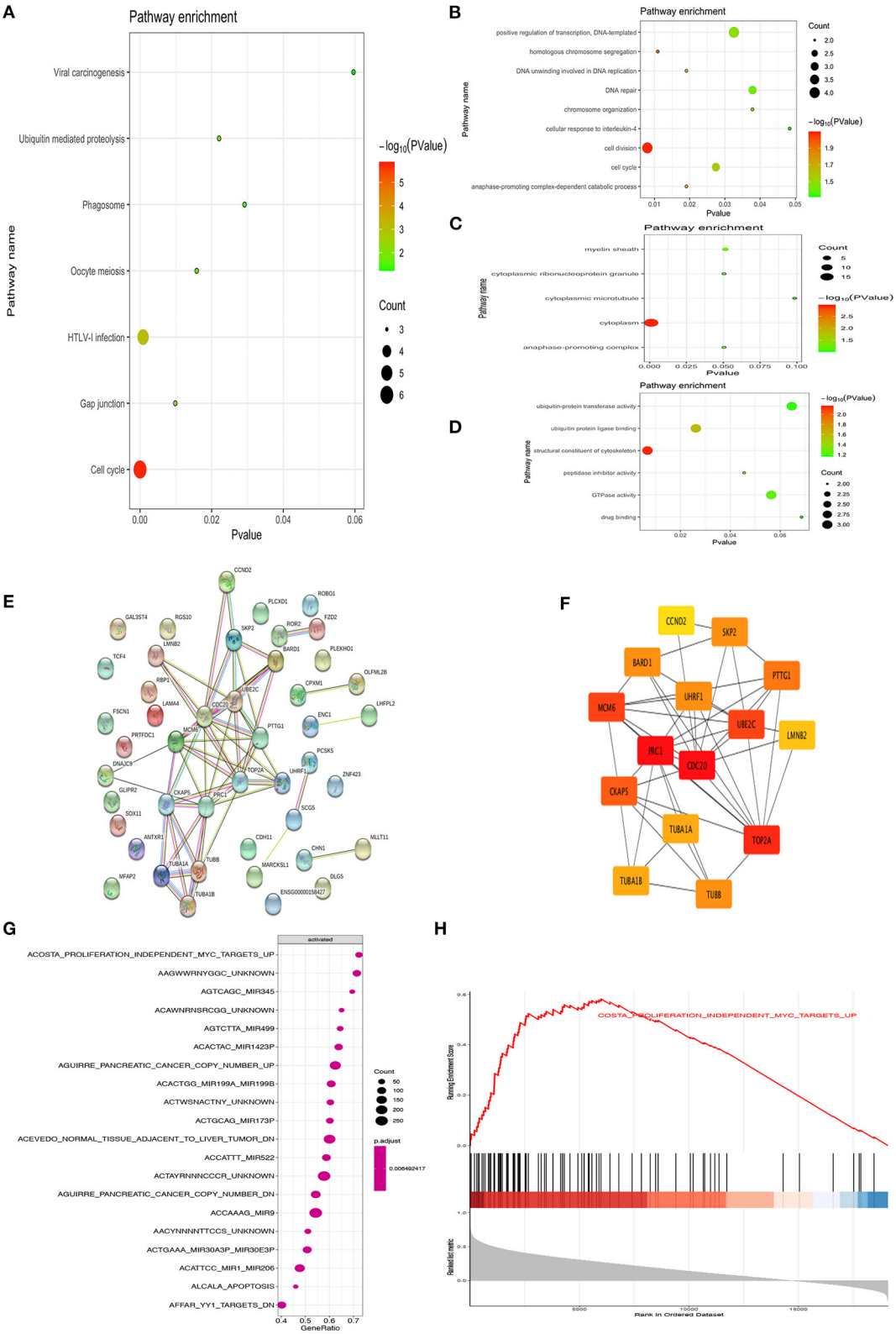


FIGURE 2 | Bioinformatics analysis of different genes. Gene ontology analyses of the common up-regulation DEGs according to biological process, cellular component and molecular function **(A–D)**. PPI network of the common DEGs identified from GSE66405, GSE73029, and TCGA was constructed **(E)**. The sub-networks were identified by Cytoscape MCODE plugin **(F)**. Gene set enrichment analysis of CDC20 related genes from TCGA datasets **(G,H)**.

as follows: #$p > 0.05$, not significant; *$p \leq 0.05$; **$p \leq 0.01$; ***$p \leq 0.001$; ****$p \leq 0.0001$.

RESULTS

The DEGs Among GSE66405, GSE73209, and TCGA

To determine the different mRNA expression profiles in WT, our study performed three mRNA microarray analyses of 184 WT tissues and 12 non-tumor adjacent tissues (**Figures 1A–C**). As the volcano plots illustrated, gene expression profiles from GSE66405 identified 5,462 DEGs with 839 genes upregulated and 4,623 genes downregulated in WT samples compared with the non-tumor adjacent tissues (**Figure 1D**). From GSE73209 data, we recognized 1,237 DEGs, of which 339 genes were upregulated and 898 genes were downregulated in WT (**Figure 1E**). We identified 3,940 differentially expressed mRNAs, including 2,118 upregulated mRNAs and 1,822 upregulated mRNAs from TCGA database (**Figure 1F**). We identified 44 commonly upregulated genes and 272 downregulated genes in the above datasets via Venn diagram (**Figures 1G,H**).

GO and Pathway Enrichment Analysis

DAVID was used to analyze the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway and GO analysis of 44 common

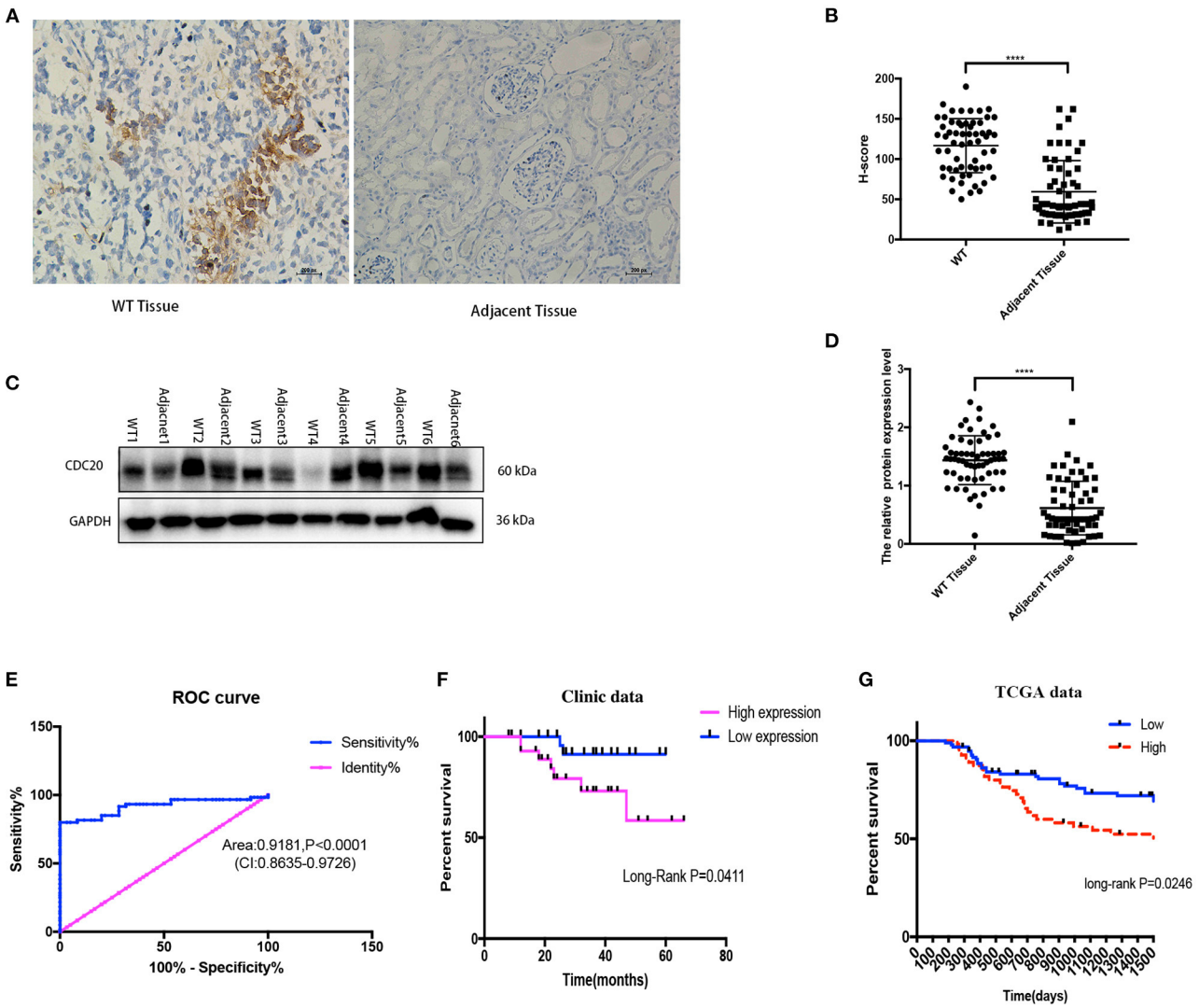


FIGURE 3 | CDC20 expression is upregulated in clinical WT samples. Representative images of IHC staining for CDC20 in adjacent tissues and different histologic grades of WT tissues **(A)**. The protein levels of CDC20 in six pairs of WT tissues and adjacent non-tumor tissues measured by Western blot **(C)**. Quantification of CDC20 IHC staining and Western blot in paired WT and adjacent tissues, respectively **(B,D)**. Receiver operating characteristic curve from IHC staining shows CDC20 is a marker to distinguish WT tissues from adjacent tissues **(E)**. Kaplan–Meier analysis of overall survival was performed to indicate that higher expression of CDC20 was correlated with poor survival of WT patients **(F,G)**. *p*-values were obtained from the log-rank test. ****$p \leq 0.0001$ were obtained by Student *t*-test. All data are represented by mean ± SD.

upregulation genes. The KEGG disease enrichment analysis demonstrated that targets were associated with the cell cycle, HTLV-I infection, oocyte meiosis, phagosome, gap junction, ubiquitin-mediated proteolysis, and viral carcinogenesis (**Figure 2A**). The GO analysis showed that, for BPs, genes significantly enriched in cell cycle, cell division, positive regulation of transcription, DNA template, homologous chromosome segregation, DNA unwinding involved in DNA replication, DNA repair, anaphase-promoting complex-dependent catabolic process, chromosome organization, and cellular response to interleukin 4 (**Figure 2B**). For MF, genes were primarily enriched in drug-binding ubiquitin–protein transferase activity, ubiquitin protein ligase binding, structural constituent of cytoskeleton, peptidase inhibitor activity, GTPase activity, and drug binding (**Figure 2C**). For CC, genes were particularly enriched in the cytoplasm, anaphase-promoting complex, cytoplasmic ribonucleoprotein granule, myelin sheath, and cytoplasmic microtubule (**Figure 2D**).

Key Candidate Genes Identification With DEG PPI Network

The PPI network of DEGs was constructed by using the STRING online database and Cytoscape (**Figure 2E**). MCODE plugin was used for module analysis of the PPI network, and the most significant modules were chosen for further pathway analyses based on the degree of importance. Then, the central node genes (more than 10 connections/interactions) were identified, and the top 15 highly connected genes were TOP2A, PTTG1, SKP2, TUBB, TUBA1A, UHRF1, TUBA1B, UBE2C, CDC20, CCND2, BARD1, MCM6, CKAP5, LMNB2, and PRC1 (**Figure 2F**). The genes in the module were mainly associated with increased cell cycle, cell division, cell cycle process, regulation of cell cycle, mitotic cell cycle process, and G2/M transition of mitotic cell cycle. As previously reported, CDC20 is an oncogene that plays a crucial role in cell cycle, cell division, and cell process (16, 17). Hence, we further investigated the role of CDC20 in WT. Furthermore, we applied single GSEA on the TCGA dataset and found that CDC20 was mainly regulated by MYC, Mir-345, Mir-449, Mir-1423P, Mir-199A/B, Mir-522, Mir-9, and Mir-206 (**Figure 2G**). Moreover, MYC is significantly correlated with CDC20 (**Figures 2G,H**).

Expression of CDC20 Was Higher in WT Tissues Compared With Adjacent Normal Tissues

In order to verify the results of the above bioinformatics analysis, WB (quantitative) and imunohistochemistry (semiquantitative and localization) methods were used for examination expression of CDC20 in WT clinic samples. The protein expression of CDC20 was detected by WB and CDC20 staining, and the results were similar to the IHC result in WT tissues. In IHC staining, H score revealed that CDC20 was significantly highly expressed in WT tissues compared with paired adjacent normal kidney tissues ($p < 0.0001$, **Figures 3A,B**). WB results, which were similar to IHC results, revealed high expression in WT tissues compared with paired adjacent normal kidney tissues ($p < 0.0001$, **Figures 3C,D**). In order to sequence the CDC20 diagnostic sensitivity, the area under the curve value obtained by ROC curve analysis from paired WT samples was 0.9181, which held statistical significance to support the diagnostic value of CDC20 for WT (**Figure 3E**). Furthermore, to detect the relationship between high expression of CDC20 and clinical prognosis, we used the KM survival analysis and log-rank test. Interestingly, we found that the high expression of CDC20 (median value) had a markedly lower overall survival rate. The results were similar in our clinical samples and in the TCGA database (log-rank $p < 0.05$, **Figures 3F,G**). Altogether, these data implied the potential oncogenic role of CDC20 in WT, and high expression of CDC20 may influence the survival rate of WT patients.

CDC20 Promotes WT Cell Proliferation and Migration and Controls Cell Cycle Progression *in vitro*

As in the results mentioned previously, CDC20 may be involved in the tumorigenesis of WT. However, the potential mechanism is unknown. To explore whether CDC20 can be used as a new strategy for the treatment of WT, three silenced RNA segments were used for CDC20 in G401 and SK-NEP-1 WT cell lines. We performed CCK-8 assays to examine the proliferation effect of si-CDC20 WT cells. As determined by the CCK-8 assay, si-CDC20-1 and si-CDC20-3 significantly slowed cell proliferation in a time-dependent manner in G401and SK-NEP-1 cells compared with the cells transfected with NC siRNA and siCDC20-2 ($p < 0.001$, **Figures 4A,B**). These results indicate that si-CDC20 could decrease WT cell proliferation. In addition, we used wound healing assays to examine the migration ability after downregulation of CDC20. Compared with the si-NC, the si-CDC20-1 and CDC20-3 could significantly impair the migration of G401 cells lines in 24 and 48 h ($p < 0.05$ and $p < 0.001$, respectively; **Figures 4C,D**). Next, the cell cycle distribution was altered by the si-CDC20-1 in SK-NEP-1 and G401 cell lines. Compared with the si-NC, the proportion of G0/G1 phase cells was significantly decreased in SK-NEP-1 and G401 ($p < 0.05$, **Figures 4E,F**). On the contrary, the proportion of G2/M phase cells was reduced by si-CDC20 in SK-NEP-1 and G401 cell lines ($p < 0.01$, **Figures 4E,F**). Based on these data, we hypothesized that downregulation of CDC20 may inhibit proliferation and migration by inducing cell cycle arrest in G2/M phase.

Cell Cycle–Related Proteins Levels Were Suppressed by Inhibition of CDC20 in WT Cells

It has been previously reported that CDC20 plays an important role during the metaphase-to-anaphase transition by targeting critical cell cycle regulators including securin and cyclin B1 and cyclin A for ubiquitination-mediated destruction (18–20). In addition, in human malignant tumors, inhibition of CDC20 in growing cells leads to G2 arrest with a consequent decrease of cyclin B1, securin, and cyclin A (21). In the study, the WB results showed that protein expression level of CDC20

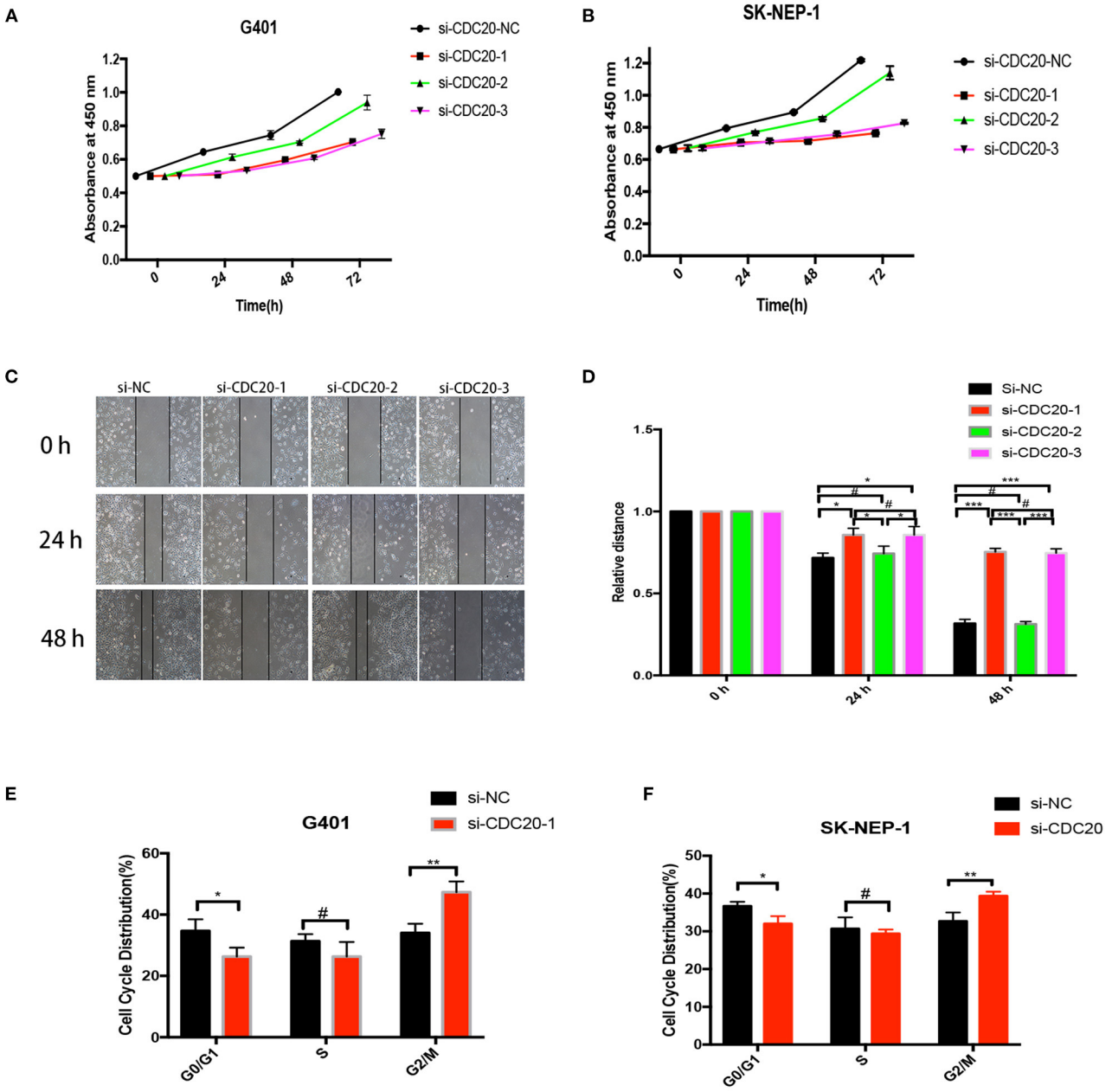


FIGURE 4 | CDC20 controls cell proliferation, migration, and cell cycle *in vitro*. CDC20 siRNA suppression of proliferation of G401 and SK-NEP-1, respectively, *in vitro* **(A,B)**. Wound healing assays were performed to determine the migration rate of G401 cells at 24 and 48 h after transfection of siRNA **(C,D)**. The G0/G1, S, and G2/M phase proportions of G401 and SK-NEP-1 cells transfected with siCDC20 or NC **(E,F)**. Results are shown as the mean ± SD. For comparisons, the Student *t*-test was performed. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, $\#p > 0.05$.

could be significantly inhibited by si-CDC20 in SK-NEP-1 and G401 cell lines (**Figures 5A,B**). Meanwhile, compared with the si-NC group, the expression levels of securin and cyclin B1 and cyclin A were markedly decreased in the si-CDC20 group (**Figures 5A,B**), supporting the results of the cell cycle analysis. Taken together, the aforementioned findings suggest that silence of CDC20 arrests the cell in G2/M phase of WT cell.

DISCUSSION

The carcinogenesis of WT involves many factors that lead the cells to undergo uncontrolled proliferation (22). However, the underlying molecular mechanisms remain unclear. A recent study showed eight genes (EGF, CDK1, ENDRA, NGFR, OIP5, NUF2, and CDCA8) are predicted to be involved in carcinogenesis pathways (23). But, the study involved only TCGA

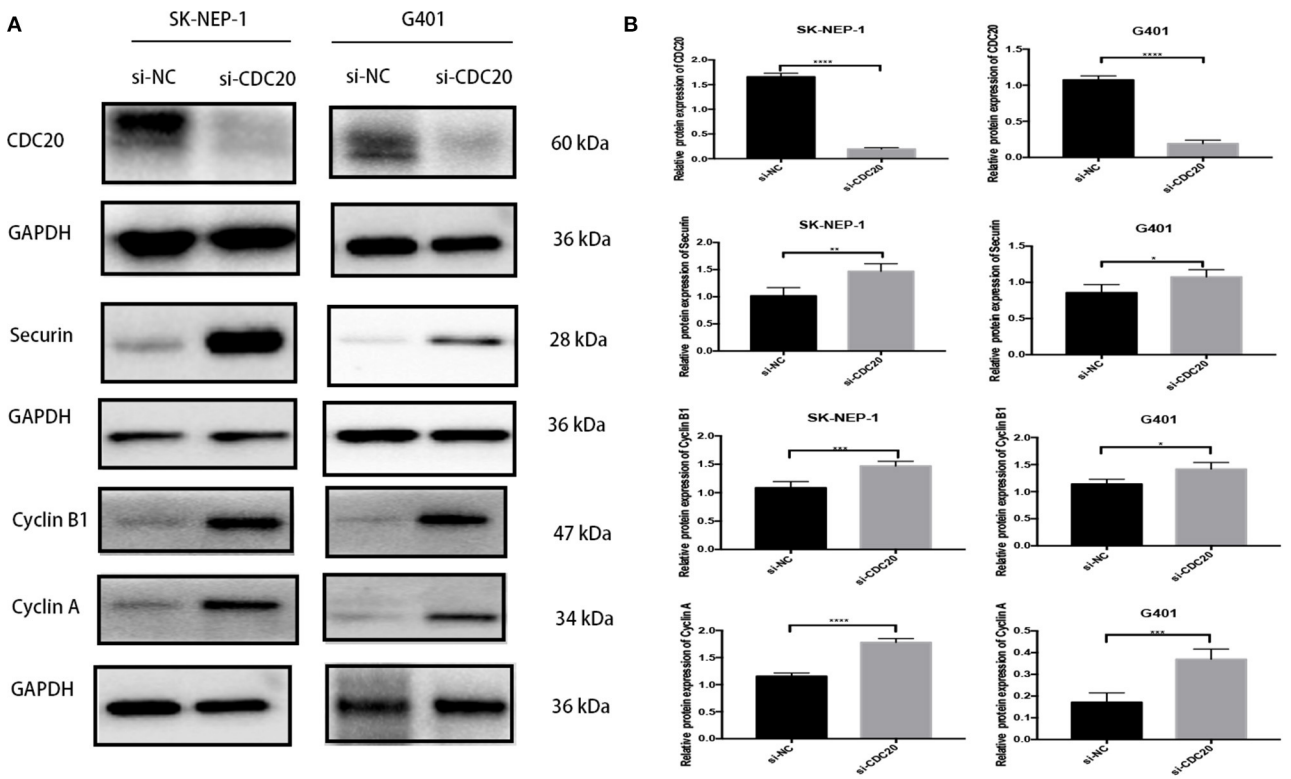


FIGURE 5 | Western blot analyses of proteins that regulate G2-to-M transition in WT cells lines inhibiting CDC20. Suppression of CDC20 protein expression by siRNA in SK-NEP-1 and G401 cells. Western blot analysis of CDC20, securin, cyclin B1, and cyclin A protein levels of the siCDC20 group compared with the NC group in SK-NEP-1 and G401 cells **(A)**. Quantification of CDC20, securin, cyclin B1, and cyclin A Western bolt in SK-NEP-1 and G401 cell lines, respectively **(B)**. Analysis was performed by Western blotting 72 h after the siRNA transfection. Results are shown as the mean ± SD. For comparison, the Student *t*-test was performed. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, $^{****}p < 0.0001$.

dataset which did not represent a generalization. Moreover, the study did not exclude 6 metastatic specimens according to TCGA nomenclature principles. More importantly, TARGET dataset included only unfavorable histology WT cases that relapsed and anaplastic WT cases, indicating that this dataset is not a representative random sampling of WT but rather a highly selected set. In this study, we have further identified common significant DEGs from three independent studies. The PPI network of DEGs revealed the top 15 highly connected genes, and CDC20 plays a crucial role in WT as the node connecting core. Functional analyses demonstrated that these DEGs are mainly associated with the cell division and cell cycle process. Meanwhile, many studies showed that CDC20 plays an oncogenic role in human tumorigenesis. Overexpression of CDC20 was observed in a variety of human tumors including pancreatic cancer, breast cancer, prostate cancer, lung cancer, colorectal cancer, hepatocellular carcinoma, glioblastoma, gastric cancer, and other types of human cancer (24–27). Therefore, CDC20 is usually identified as an oncogene (16). A recent study identified that nine key genes including CDC20 were potential diagnosis genes in clear cell renal cell carcinoma (28). Meanwhile, the study by Gayyed et al. showed that high expression of CDC20 was associated with high tumor grade in RCC (29). However, there are no further studies on the relationship of high expression of CDC20 between WT and RCC. Although both WT and RCC occur in the kidney, the difference is that WT originates in embryonic cells, and more than 95% of WTs occur in children. Moreover, an early study showed that CDC20 expression in RCC may be involved in cytochrome P450 1B1 (CYP1B1) (30).

In this study, we assessed the expression level of CDC20 in 60 paired WT tissues and corresponding non-tumor samples. The results indicated that the protein level of CDC20 in Wilms tumor tissues was much higher than that in matched nontumor tissues. Immunohistochemistry was used to investigate the subcellular location of CDC20 and its relationship with clinical pathological parameters of WT patients. By ROC analysis, we found that the high expression of CDC20 may provide diagnostic value in paired WT samples. In addition, by KM analysis and log-rank test, we found that higher CDC20 protein expression level was associated with poor survival rate. To investigate the potential biological function and molecular mechanism of CDC20 in WT, we designed a double-stranded, siRNA targeting CDC20 to interfere with its expression level in the WT cell lines.

By cellular proliferation assay, migrate assay, and fluorescence-activated cell sorting test, we found that cells transfected with siCDC20 oligonucleotides showed decreased

growth speed, a reduced rate of migration, and an increased proportion of cells in the G2/M stage. The specific knockdown of CDC20 by siRNA showed a suppressed effect against WT cell proliferation and migration *in vitro*, which indicated that the overexpression of CDC20 might be expected to accelerate cell proliferation and promote tumor initiation and progression of WT. The WT cells with suppressed CDC20 expression were induced to accumulate in the G2/M phase, which may be responsible for the inhibition of cell growth. Taken together, the overexpression of CDC20 might be expected to lead to accelerated proliferation of cells, and the specific knockdown of CDC20 by siRNA did, in fact, show an inhibitory effect against cell growth *in vitro*.

The accurate transition from the S phase to the G2/M phase is crucial for the control of eukaryotic cell proliferation (31). It was previously reported that in metaphase to anaphase, APC/C-Cdc20 mediates the ubiquitination of securin and cyclin B1, allowing the activation of separase and the onset of anaphase and mitotic exit (18). CDC20 plays an indispensable role during the metaphase-to-anaphase transition by targeting critical cell cycle regulators including securin and cyclin B1 for ubiquitination-mediated destruction (19, 32, 33). Cyclin A was essential for the control of the cell cycle at the G1/S and the G2/M transitions (20, 34). In mitosis, it may contribute to the control of cyclin B1 stability (35). In this study, we found that the securin, cyclin B1, and securin protein levels were regulated by high expression of CDC20.

In conclusion, our study demonstrated the high expression of CDC20 involvement in tumorigenesis in WT. Functional experiments verified that suppression of CDC20 could inhibit WT cell proliferation, migration, and arrested cell cycle in G2/M phase. However, more underlying molecular mechanisms upstream of CDC20 still need further research. What is more, our study has limitations on the WT cell model such that G401 and SK-NEP-1 cells were formerly classified as WT cell lines, but they have since had more correct classifications (36). Overall, this finding provides a new focus that CDC20 may be a clinically relevant indicator and a promising therapeutic target of WT.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Children Hospital of Chongqing Medical University. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

QS, BT, DH, and GW conceived of and designed the study. BT, QS, YaL, and YoL made contributions in acquisition, analysis of the data, and drafted the manuscript. BT, QS, and YoL completed the experiments. DH, TL, and GW helped draft the manuscript and assisted in the analysis of the data. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by The National clinic key specialty project [2013(544)], China. Special Key Project of Chongqing Technology Innovation and Application Development (Approval Number: cstc2019jscx-tjsbX003) and Chongqing Special Science of Technology Innovation Project for Social Undertakings and People's Livelihood Security (Approval Number: cstc2017shmsA130103), China.

REFERENCES

1. Hung IJ, Chang WH, Yang CP, Jaing TH, Liang DC, Lin KH, et al. Taiwan Pediatric Oncology Group. Epidemiology, clinical features and treatment outcome of Wilms' tumor in Taiwan: a report from Taiwan Pediatric Oncology Group. *J Formos Med Assoc*. (2004) 103:104–11. doi: 10.1109/91.983277
2. Termuhlen AM, Tersak JM, Liu Q, Yasui Y, Stovall M, Weathers R, et al. Twenty-five year follow-up of childhood wilms tumor: a report from the childhood cancer survivor study. *Cancer*. (2011) 57:1210–6. doi: 10.1002/pbc.23090
3. Ko EY, Ritchey ML. Current management of Wilms' tumor in children. *J Pediatr Urol*. (2009) 5:56–65. doi: 10.1016/j.jpurol.2008.08.007
4. Ehrlich PF, Anderson JR, Ritchey ML, Dome JS, Green DM, Grundy PE, et al. Clinicopathologic findings predictive of relapse in children with stage III favorable-histology Wilms tumor. *J Clin Oncol*. (2013) 31:1196–201. doi: 10.1200/JCO.2011.41.1165
5. Wong KF, Reulen RC, Winter DL, Guha J, Fidler MM, Kelly J, et al. Risk of adverse health and social outcomes up to 50 years after wilms tumor: the British childhood cancer survivor study. *J Clin Oncol*. (2016) 2016:JCO.2015.2064.4344. doi: 10.1200/JCO.2015.64.4344
6. Wright KD, Green DM, Daw NC. Late effects of treatment for wilms tumor. *Pediatr Hematol Oncol*. (2009) 26:407–13. doi: 10.1080/08880010903019344
7. Cone EB, Dalton SS, Van Noord M, Tracy ET, Rice HE, Routh JC. Biomarkers for wilms tumor: a systematic review. *J Urol*. (2016) 2016:1530–5. doi: 10.1016/j.juro.2016.05.100
8. Bown N, Cotterill SJ, Roberts P, Griffiths M, Larkins S, Hibbert S, et al. Cytogenetic abnormalities and clinical outcome in Wilms tumor: A study by the U.K. cancer cytogenetics group and the U.K. Children's Cancer Study Group. *Pediatr Oncol*. (2002) 38:11–21. doi: 10.1002/mpo.1258
9. Changkai D, Rong D, Xuliang L, Liu F. Genetic variation frequencies in Wilms' tumor: A meta-analysis and systematic review. *Cancer Sci*. (2016) 107:690–9. doi: 10.1111/cas.12910
10. Rakheja D, Chen KS, Liu Y, Shukla AA, Schmid V, Chang TC, et al. Somatic mutations in DROSHA and DICER1 impair microRNA biogenesis through distinct mechanisms in Wilms tumours. *Nat Commun*. (2014) 2:4802. doi: 10.1158/1538-7445.AM2014-LB-204
11. Ruteshouser EC, Robinson SM, Huff V. Wilms tumor genetics: mutations in WT1, WTX, and CTNNB1 account for only about one-third of tumors. *Genes Chromoso Cancer*. (2008) 47:461–70. doi: 10.1002/gcc.20553
12. Clough E, Barrett T. The gene expression omnibus database. *Methods Mol Biol*. (2016) 1418:93–110. doi: 10.1007/978-1-4939-3578-9_5

13. Lee JS. Exploring cancer genomic data from the cancer genome atlas project. *BMB Rep.* (2016) 49:607–11. doi: 10.5483/BMBRep.2016.49.11.145
14. Su P, Wen S, Zhang Y, Li Y, Xu Y, Zhu Y, et al. Identification of the key genes and pathways in esophageal carcinoma. *Gastroenterol Res Pract.* (2016) 2016:2968106. doi: 10.1155/2016/2968106
15. Azim HA Jr, Peccatori FA, Brohée S, Branstetter D, Loi S, Viale G, et al. RANK-ligand (RANKL) expression in young breast cancer patients and during pregnancy. *Breast Cancer Res.* (2015) 17:24. doi: 10.1186/s13058-015-0538-7
16. Penas C, Ramachandran V, Ayad NG. The APC/C ubiquitin ligase: from cell biology to tumorigenesis. *Front Oncol.* (2012) 1:60. doi: 10.3389/fonc.2011.00060
17. Kim S, Yu H. Mutual regulation between the spindle checkpoint and APC/C. *Semin Cell Dev Biol.* (2011) 22:551–8. doi: 10.1016/j.semcdb.2011.03.008
18. Li M, York JP, Zhang P. Loss of Cdc20 causes a securin-dependent metaphase arrest in two-cell mouse embryos. *Mol Cell Biol.* (2007) 27:3481–8. doi: 10.1128/MCB.02088-06
19. Clute P, Pines J. Temporal and spatial control of cyclin B1 destruction in metaphase. *Nat Cell Biol.* (1999) 1:82–7. doi: 10.1038/10049
20. Van Zon W, Wolthuis RM. Cyclin A and Nek2A: APC/C-Cdc20 substrates invisible to the mitotic spindle checkpoint. *Biochem Soc Trans.* (2010) 38:72–7. doi: 10.1042/BST0380072
21. Kidokoro T, Tanikawa C, Furukawa Y, Katagiri T, Nakamura Y, Matsuda K. CDC20, a potential cancer therapeutic target, is negatively regulated by p53. *Oncogene.* (2008) 27:1562–71. doi: 10.1038/sj.onc.1210799
22. Treger TD, Chowdhury T, Pritchard-Jones K, Behjati S. The genetic changes of Wilms tumour. *Nat Rev Nephrol.* (2019) 15:240–51 doi: 10.1038/s41581-019-0112-0
23. Wang X, Song P, Huang C, Yuan N, Zhao X, Xu C. Weighted gene co expression network analysis for identifying hub genes in association with prognosis in Wilms tumor. *Mol Med Rep.* (2019) 19:2041–50. doi: 10.3892/mmr.2014.2921
24. Li J, Gao JZ, Du JL, Huang ZX, Wei LX. Increased CDC20 expression is associated with development and progression of hepatocellular carcinoma. *Int J Oncol.* (2014) 45:1547–55. doi: 10.3892/ijo.2014.2559
25. Karra H, Repo H, Ahonen I, Löyttyniemi E, Pitkänen R, Lintunen M, et al. Cdc20 and securin overexpression predict short-term breast cancer survival. *Br J Cancer.* (2014) 110:2905–13. doi: 10.1038/bjc.2014.252
26. Kato T, Daigo Y, Aragaki M, Ishikawa K, Sato M, Kaji M. Overexpression of CDC20 predicts poor prognosis in primary non-small cell lung cancer patients. *J Surg Oncol.* (2012) 106:423–30. doi: 10.1002/jso.23109
27. Wang L, Zhang J, Wan L, Zhou X, Wang Z, Wei W. Targeting Cdc20 as a novel cancer therapeutic strategy. *Pharmacol Ther.* (2015) 151:141–51. doi: 10.1016/j.pharmthera.2015.04.002
28. Luo Y, Shen D, Chen L, Wang G, Liu X, Qian K, et al. Identification of 9 key genes and small molecule drugs in clear cell renal cell carcinoma. *Aging (Albany NY).* (2019) 11:6029–52. doi: 10.18632/aging.102161
29. Gayyed MF, El-Maqsoud NM, Tawfiek ER, El Gelany SA, Rahman MF. A comprehensive analysis of CDC20 overexpression in common malignant tumors from multiple organs: its correlation with tumor grade and stage. *Tumour Biol.* (2016) 37:749–62. doi: 10.1007/s13277-015-3808-1
30. Mitsui Y, Chang I, Fukuhara S, Hiraki M, Arichi N, Yasumoto H, et al. CYP1B1 promotes tumorigenesis via altered expression of CDC20 and DAPK1 genes in renal cell carcinoma. *BMC Cancer.* (2015) 15:942. doi: 10.1186/s12885-015-1951-0
31. Garrett MD. Cell cycle control and cancer. *Current.* (2001) 81:515–22. doi: 10.1126/science.7997877
32. Michaelis C, Ciosk R, Nasmyth K. Cohesins chromosomal proteins that prevent premature separation of sister chromatids. *Cell.* (1997) 91:35–45. doi: 10.1016/S0092-8674(01)80007-6
33. Nasmyth K. Disseminating the genome: joining, resolving, and separating sister chromatids during mitosis and meiosis. *Annu Rev Genet.* (2001) 35:673–745. doi: 10.1146/annurev.genet.35.102401.091334
34. Ohtoshi A, Maeda T, Higashi H, Ashizawa S, Hatakeyama M. Human p55(CDC)/Cdc20 associates with cyclin A and is phosphorylated by the cyclin A-Cdk2 complex. *Biochem Biophys Res Commun.* (2000) 268:530–4. doi: 10.1006/bbrc.2000.2167
35. Yam CH, Fung TK, Poon RY. Cyclin A in cell cycle control and cancer. *Cell Mol Life Sci.* (2002) 59:1317–26. doi: 10.1007/s00018-002-8510-y
36. Pritchard-Jones K, Perotti D. WARNING: G-401 and SK-NEP-1 cell lines are not Wilms tumor cell lines. *Pediatr Blood Cancer.* (2019) 66:e27741. doi: 10.1002/pbc.27741

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

ORIGINAL RESEARCH
published: 09 September 2021
doi: 10.3389/fonc.2021.635251

Novel Associations Between *METTL3* Gene Polymorphisms and Pediatric Acute Lymphoblastic Leukemia: A Five-Center Case-Control Study

Xiaoping Liu[1†], Libin Huang[2†], Ke Huang[3], Lihua Yang[4], Xu Yang[1], Ailing Luo[1], Mansi Cai[1], Xuedong Wu[5], Xiaodan Liu[1,6], Yaping Yan[1], Jianyun Wen[5], Yun Cai[7], Ling Xu[1] and Hua Jiang[1*]*

[1] Department of Hematology, Guangzhou Women and Children's Medical Center, Guangzhou Medical University, Guangzhou, China, [2] Pediatrics Department, The First Affiliated Hospital, Sun Yat-Sen University, Guangzhou, China, [3] Department of Pediatrics, Sun Yat-sen Memorial Hospital, Sun Yat-sen University, Guangzhou, China, [4] Pediatric Center of Zhujiang Hospital, Southern Medical University, Guangzhou, China, [5] Department of Pediatrics, Nanfang Hospital, Southern Medical University, Guangzhou, China, [6] Division of Birth Cohort Study, Guangzhou Women and Children's Medical Center, Guangzhou Medical University, Guangzhou, China, [7] Department of Pediatrics, The Third Affiliated Hospital, Sun Yat-sen University, Guangzhou, China

Objective: To reveal the contributing role of METTL3 gene SNPs in pediatric ALL risk.

Patients and Methods: A total of 808 pediatric ALL cases and 1,340 cancer-free controls from five hospitals in South China were recruited. A case-control study by genotyping three SNPs in the *METTL3* gene was conducted. Genomic DNA was abstracted from peripheral blood. Three SNPs (rs1263801 C>G, rs1139130 A>G, and rs1061027 A>C) in the METTL3 gene were chosen to be detected by taqman real-time polymerase chain reaction assay.

Results: That rs1263801 C>G, rs1139130 A>G, and rs1061027 A>C polymorphisms were significantly associated with increased pediatric ALL risk was identified. In stratification analyses, it was discovered that rs1263801 CC, rs1061027 AA, and rs1139130 GG carriers were more likely to develop ALL in subgroups of common B-ALL, MLL gene fusion. Rs1263801 CC and rs10610257 AA carriers were more possible to increase the risk of ALL in subgroups of low hyperdiploid, and all of these three SNPs exhibited a trend toward the risk of ALL. All of these three polymorphisms were associated with the primitive/naïve lymphocytes and MRD in marrow after chemotherapy in ALL children. Rs1263801 CC and rs1139130 AA alleles provided a protective effect on MRD ≥0.01% among CCCG-treated children. As for rs1139130, AA alleles provided a protective effect on MRD in marrow ≥0.01% on 33 days and 12 weeks among CCCG-treated children, but provided a risk effect on MRD in the marrow ≥0.01% among SCCLG-treated children. As for rs1263801 CC and rs1139130 AA, these two alleles provided a protective effect on MRD in the marrow ≥0.01% among CCCG-treated children.

OPEN ACCESS

Edited by:
Hua Tan,
University of Texas Health Science Center at Houston, United States

Reviewed by:
Matteo Chinello,
Integrated University Hospital Verona, Italy
Meenakshi Devidas,
St. Jude Children's Research Hospital, United States

***Correspondence:**
Hua Jiang
jiang_hua18@sina.cn
Ling Xu
luoxul64@126.com

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Pediatric Oncology, a section of the journal Frontiers in Oncology

Received: *30 November 2020*
Accepted: *23 July 2021*
Published: *09 September 2021*

Citation:
Liu X, Huang L, Huang K, Yang L, Yang X, Luo A, Cai M, Wu X, Liu X, Yan Y, Wen J, Cai Y, Xu L and Jiang H (2021) Novel Associations Between METTL3 Gene Polymorphisms and Pediatric Acute Lymphoblastic Leukemia: A Five-Center Case-Control Study. Front. Oncol. 11:635251. doi: 10.3389/fonc.2021.635251

Conclusion: In this study, we revealed that *METTL3* gene polymorphisms were associated with increased pediatric ALL risk and indicated that *METTL3* gene polymorphisms might be a potential biomarker for choosing ALL chemotherapeutics.

Keywords: methyltransferase-like 3, acute lymphoid leukemia, polymorphism, pediatric, susceptibility

INTRODUCTION

Acute lymphoblastic leukemia (ALL) is the most common type of pediatric cancer in the world; in China, it accounts for 70–80% of pediatric leukemia (1). ALL can be classified by immune cell phenotype as B-cell ALL and T-cell ALL. B-cell ALL is the most common ALL; T-cell ALL is typically more aggressive (2). As traditional chemotherapy combined with novel therapies makes great progress, higher survival rates and reduced morbidities have been achieved in ALL. Recently, the 5-year overall survival rate of ALL children younger than 14 years has been achieved >90% (3). However, recurrence occurs in 15–20% of ALL children, and 15% pediatric ALL patients were therapeutic failures, which resulted in early age mortality (4). ALL is characterized by multiple genetic alterations (5).

Heritable variations in genes are risk factors for ALL and play a strong role in the development of pediatric ALL (6). Populations with different races are well distinguished by genetic polymorphisms. Genome-wide association studies (GWAS) have identified a number of loci, and single nucleotide polymorphism (SNP) associations in several genes are associated with the risk of ALL. Genetic alterations in pediatric ALL are found to be very different from those in adult ALL (7).

N^6-methyladenosine (m6A) is the most abundant internal modification of messenger RNAs (mRNAs) in eukaryotic organisms. Methylation at the sixth N atom on adenine base is m6A. M6A regulates mRNA expression posttranscriptionally in a dynamic and reversible manner (8). M6A modification is regulated by several key regulators, including writers [RNA methyltransferase complex methyltransferase-like 3 (METTL3)/ methyltransferase-like 14 (METTL14)/Wilms' tumor 1-associating protein (WTAP)], erasers [demethylases fat mass and obesity-associated protein (FTO) and AlkB homolog 5 (ALKBH5)], and readers (YTHD family proteins) (9). It was reported that dysregulation of m6A is associated with multiple tumors including acute myeloid leukemia (AML) (10). M6A methylation writer METTL3 was discovered playing an oncogenic role in carcinogenesis, such as colorectal carcinoma (11), bladder cancer (12), breast cancer (13), etc. METTL3 mRNA and protein are expressed abundantly in AML cells, and their depletion induces cell differentiation and apoptosis and delays leukemia progression (14). Some genetic variations in m6A-related gene regions may affect m6A methylation, subsequently regulating mRNA expression (15). Studies identified that m6A-associated SNPs were potential functional variants for periodontitis (16) and coronary artery disease (17). Genetic alterations in the m6A demethyltransferase FTO gene were shown to be associated with ALL and AML risk, and there is evidence that indicates dysregulation of m6A methyltransferase METTL3 in AML (18, 19). However the relationship between genetic variations of the METTL3 gene and ALL is still unclear.

In the present study, a total of three SNPs were selected to assess the relationship between *METTL3* polymorphisms and pediatric ALL. The current study was a case-control study that was performed using samples from five hospitals in South China.

MATERIALS AND METHODS

Study Subjects

A Southern Chinese population that included 808 pediatric ALL patients and 1,340 age-matched, gender-matched, and ethnicity-matched healthy controls is summarized in **Table S1**. ALL cases were collected from Guangzhou Women and Children's Medical Center (GWCMC), Guangzhou Medical University (n=582), The First Affiliated Hospital, Sun Yat-sen University (n=74), Sun Yat-sen Memorial Hospital, Sun Yat-sen University (n=26), Nanfang Hospital, Southern Medical University (SMU) (n=100), and Zhujiang Hospital, Southern Medical University (n=26), from January 2017 to May 2019. All children were diagnosed with ALL by at least two hematologists. The control subjects were free from any type of hematological diseases or any other malignancy or autoimmune disorder and were recruited from the same hospital.

The major clinical and biological characteristics of the ALL children, including age, gender, immunophenotype, gene fusion type, risk level, karyotype, clinical manifestations, rate of primitive/naive lymphocytes in the marrow, and minimal residual disease on 19 days, 33 days, and 12 weeks after chemotherapy and chemotherapy regimen were collected. The information is summarized in **Table S1**.

The study was approved by the institutional ethics committee of every participating hospital, and written informed consent was acquired from all participants in accordance with the Declaration of Helsinki.

METTL3 SNPs Selection and Genotyping

The included potentially functional candidate SNPs were selected as follows: located in the 5' untranslated region, 3' untranslated region, 5' flanking region, and exon of the *METTL3* gene. The NCBI dbSNP database (http://www.ncbi.nlm.nih.gov/projects/SNP) and the SNPinfo (https://snpinfo.niehs.nih.gov/snpinfo/snpfunc.html) online software were used to perform the above selection. Three SNPs (rs1263801 C>G,

rs1139130 G>A, and rs1061027 A>C) in the *METTL3* gene were chosen. Genomic DNA was extracted from peripheral blood. The reaction system and condition of the Taqman RT-PCR assay was according to the published reference (20, 21). To ensure the accuracy of these genotyping results, 10% of the samples were randomly selected to be genotyped by a DNA sequencing method. A concordance rate of 100% for the quality control samples was obtained (21).

Statistical Analysis

The goodness-of-fit χ^2 test was performed to assess if the selected *METTL3* SNPs deviated from Hardy–Weinberg equilibrium among controls. The two-sided χ^2 test was used to compare demographic variables and genotype frequencies of the cases and controls. ORs and their corresponding 95% CIs were computed by unconditional logistic regression analyses with or without adjustment for age and gender. The SAS statistical package (version 9.1; SAS Institute, Cary, NC) was used to perform all statistical analyses. All reported p values were two sided, and a p value < 0.05 was considered statistically significant.

RESULTS

Population Characteristics

The demographic and clinical characteristics data of ALL cases and cancer-free controls are summarized in **Table S1**. No significant differences were observed between cases and controls for the Southern Chinese children regarding age (p=0.082) and gender (p=0.059). Among ALL cases, 28.22% (228 cases) were pro B cell ALL, 35.27% (285 cases) were common B cell ALL, 20.67% (167 cases) were pre-B cell ALL, 0.67% (3 cases) were mature B ALL, 8.54% (69 cases) were T cell ALL, and 6.93% (56 cases) were undefined immunophenotype. Regarding the gene fusion type, 3.34% (27 cases) had BCR-ABL gene fusion, 16.83% (136 cases) had TEL-AML, 2.97% (24 cases) had E2A-PBX, 0.99% (8 cases) had SIL-TAL, 1.98% (16 cases) had MLL, 3.09% (25 cases) had other gene fusions, 68.19% (551 cases) were normal, and 21 were undefined. A total of 258 patients (33.73%) were with low risk, 360 cases (47.06%) were with medium risk, 77 cases (10.07%) were with high risk, and 70 cases (9.15%) were undefined. Regarding the karyotype, 64.40% (517 cases) were normal diploid, 5.25% (45 cases) were abnormal diploid, 2.69% (22 cases) were hypodiploid, 3.46% (27 cases) were low hyperdiploid, and 7.81% (61 cases) were high hyperdiploid.

Correlation of *METTL3* Gene Polymorphisms With ALL Risk

The genotype frequencies of *METTL3* associated with ALL risk are shown in **Table 1**. In the single-locus analysis, carriers of rs1263801 (CC *vs*. GG: adjusted OR= 4.18, 95% CI=3.21–5.43, p<0.001) and rs1061027 (CA *vs*. CC: adjusted OR=2.42, 95% CI=2.00–2.94, p<0.001; AA *vs*. CC: adjusted OR=6.21, 95% CI=4.38–8.81, p<0.001) variant alleles showed significant enhanced risk of pediatric ALL. On the contrary, rs1139130 (GA *vs*. GG: adjusted OR=1.41, 95% CI=1.15–1.73, p=0.001; AA *vs*. GG: adjusted OR=1.52, 95% CI=1.81–3.06, p<0.001) variant alleles contribute to decreased risk of pediatric ALL.

TABLE 1 | Logistic regression analysis of associations between *METTL3* polymorphisms and ALL susceptibility.

Genotype	Cases (N = 808)	Controls (N = 1340)	P[a]	Crude OR (95% CI)	P	Adjusted OR (95% CI) b	P[b]
rs1263801 (HWE=0.0971)							
GG	269 (33.50)	600 (44.88)		1.00		1.00	
GC	304 (37.86)	611 (45.70)		1.11 (0.91-1.35)	0.305	1.12 (0.92-1.37)	0.254
CC	230 (28.64)	126 (9.42)		4.07 (3.14-5.28)	0.001	4.18 (3.21-5.43)	0.001
Additive			0.001	1.82 (1.61-2.07)	0.001	1.84 (1.63-2.09)	0.001
Dominant	534 (66.50)	737 (55.12)	0.001	1.62 (1.35-1.94)	0.001	1.64 (1.37-1.97)	0.001
Recessive	573 (71.36)	1211 (90.58)	0.001	3.86 (3.04-4.90)	0.001	3.93 (3.09-5.00)	0.001
rs1139130 (HWE=0.3401)							
GG	220 (28.17)	511 (38.51)		1.00		1.00	
GA	383 (49.04)	638 (48.08)		1.39 (1.14-1.71)	0.002	1.41 (1.15-1.73)	0.001
AA	178 (22.79)	178 (13.41)		2.32 (1.79-3.02)	0.001	2.36 (1.81-3.06)	0.001
Additive			0.001	1.51 (1.32-1.71)	0.001	1.52 (1.33-1.73)	0.001
Dominant	561 (71.83)	816 (61.49)	0.001	1.60 (1.32-1.93)	0.001	1.62 (1.34-1.96)	0.001
Recessive	603 (77.21)	1149 (86.59)	0.001	1.90 (1.51-2.40)	0.001	1.92 (1.52-2.42)	0.001
rs1061027 (HWE=0.6433)							
CC	319 (39.78)	859 (64.73)		1.00		1.00	
CA	364 (45.39)	414 (31.20)		2.37 (1.96-2.87)	0.001	2.42 (2.00-2.94)	0.001
AA	119 (14.84)	54(4.07)		5.93 (4.20-8.39)	0.001	6.21 (4.38-8.81)	0.001
Additive			0.001	2.41 (2.09-2.78)	0.001	2.46 (2.13-2.84)	0.001
Dominant	483 (60.33)	468 (35.27)	0.001	2.78 (2.32-3.33)	0.001	2.85 (2.38-3.42)	0.001
Recessive	683 (85.16)	1273 (95.93)	0.001	4.11 (2.94-5.74)	0.001	4.23 (3.02-5.93)	0.001

[a]χ^2 *test for genotype distributions between ALL cases and cancer-free controls.*
[b]*Adjusted for age and gender.*

TABLE 2 | Stratification analysis of *METTL3* polymorphisms with ALL susceptibility.

Variables	rs1263801 (cases/controls)		Adjusted OR[a]	*P*[a]	rs1139130 (cases/controls)		Adjusted OR[a]	*P*[a]	rs1061027 (cases/controls)		Adjusted OR[a]	*P*[a]
	GG/GC	CC	(95% CI)		GG/GA	AA	(95% CI)		CC/CA	AA	(95% CI)	
Age, month												
<120	513/1095	192/110	3.69 (2.85-4.77)	0.001	534/1036	154/159	1.87 (1.46-2.40)	0.001	610/1149	94/48	3.64 (2.53-5.22)	0.001
≥120	60/116	38/16	4.65 (2.39-9.05)	0.001	69/113	24/19	2.07 (1.06-4.05)	0.034	73/124	25/6	7.10 (2.78-18.1)	0.001
Gender												
Females	227/435	96/50	3.72 (2.55-5.44)	0.001	226/426	85/60	2.62 (1.81-3.78)	0.001	268/462	53/20	4.84 (2.83-8.28)	0.001
Males	346/776	134/76	4.06 (2.97-5.54)	0.001	377/723	93/118	1.53 (1.13-2.06)	0.005	415/811	64/34	3.82 (2.47-5.92)	0.001
Immunophenotyping												
Pro B	198/1211	29/126	1.47 (0.95-2.28)	0.083	181/1149	41/178	1.51 (1.04-2.21)	0.032	214/1273	13/54	1.51 (0.80-2.83)	0.201
Common B	149/1211	133/126	8.59 (6.37-11.6)	0.001	177/1149	98/178	3.58 (2.67-4.81)	0.001	199/1273	82/54	9.72 (6.67-14.2)	0.001
Pre B	136/1211	30/126	2.22 (1.43-3.44)	0.001	140/1149	22/178	1.04 (0.65-1.68)	0.863	158/1273	8/54	1.25 (0.58-2.68)	0.570
Mature B	3/1211	0/126	0.001 (0.00-999)	0.973	3/1149	0/178	0.001 (0.00-999)	0.968	3/1273	0/54	0.001(0.00-999)	0.982
T ALL	49/1211	20/126	3.78 (2.15-6.63)	0.001	58/1149	8/178	0.86 (0.40-1.83)	0.688	61/1273	8/54	2.88 (1.29-6.42)	0.010
Mix	38/1211	18/126	4.41 (2.43-7.99)	0.001	44/1149	9/178	1.31 (0.63-2.73)	0.474	48/1273	8/54	3.82 (1.71-8.53)	0.002
Gene fusion type												
BCR-ABL	15/1211	12/126	6.05 (2.66-13.7)	0.001	20/1149	6/178	1.63 (0.62-4.26)	0.319	20/1273	7/54	6.15 (2.32-16.3)	0.001
TEL-AML	96/1211	39/126	4.08 (2.68-6.21)	0.001	103/1149	30/178	1.89 (1.22-2.93)	0.004	114/1273	21/54	4.61 (2.67-7.94)	0.001
E2A-PBX	21/1211	3/126	1.40 (0.41-4.79)	0.588	21/1149	2/178	0.65 (0.15-2.80)	0.562	23/1273	1/54	1.05 (0.14-7.93)	0.965
SIL-TAL	7/1211	1/126	1.31 (0.16-10.9)	0.797	7/1149	0/178	0.001 (0.00-999)	0.961	7/1273	1/54	3.27 (0.39-27.2)	0.274
MLL	7/1211	9/126	13.1 (4.77-35.9)	0.001	10/1149	6/178	3.95 (1.41-11.0)	0.009	10/1273	6/54	14.8 (5.15-42.6)	0.001
Others	10/1211	15/126	14.2 (6.22-32.3)	0.001	13/1149	11/178	5.53 (2.44-12.6)	0.001	16/1273	9/54	13.2 (5.56-31.3)	0.001
Normal	405/1211	142/126	3.41 (2.61-4.45)	0.001	413/1149	118/178	1.86 (1.43-2.41)	0.001	474/1273	72/54	3.63 (2.50-5.25)	0.001
Karyotype												
Hypo-diploid	17/1211	5/126	2.83 (1.02-7.85)	0.046	15/1149	5/178	1.84 (0.67-5.07)	0.236	20/1273	2/54	2.29 (0.52-10.1)	0.274
Normal diploid	371/1211	144/126	3.78 (2.89-4.95)	0.001	389/1149	113/178	1.88 (1.45-2.45)	0.001	434/1273	80/54	4.44 (3.08-6.39)	0.001
Abnormal diploid	34/1211	11/126	2.99 (1.47-6.06)	0.002	35/1149	7/178	1.28 (0.56-2.93)	0.562	40/1273	5/54	2.87 (1.09-7.58)	0.034
Low hyperdiploid	16/1211	11/126	6.16 (2.78-13.7)	0.001	20/1149	5/178	1.58 (0.58-4.27)	0.372	20/1273	7/54	7.83 (3.17-19.5)	0.001
High hyperdiploid	45/1211	16/126	3.66 (1.99-6.71)	0.001	49/1149	12/178	1.67 (0.87-3.22)	0.126	57/1273	4/54	1.81 (0.63-5.22)	0.271
Primitive/naive lymphocytes in marrow(%, 19d)												
<5	362/1211	134/126	3.66 (2.79-4.82)	0.001	377/1149	104/178	1.80 (1.38-2.36)	0.001	441/1273	55/54	3.07 (2.07-4.56)	0.001
≥5	45/1211	24/126	4.83 (2.86-8.24)	0.001	50/1149	18/178	2.27 (1.29-4.00)	0.004	55/1273	14/54	5.69 (2.96-10.9)	0.001
MRD in marrow(%, 19d)												
<0.01	22/1211	25/126	10.6 (5.81-19.5)	0.001	26/1149	19/178	4.66 (2.52-8.60)	0.001	30/1273	17/54	13.1 (6.80-25.3)	0.001
≥0.01	298/1211	159/126	5.27 (4.03-6.89)	0.001	337/1149	108/178	2.09 (1.60-2.73)	0.001	378/1273	79/54	5.18 (3.58-7.49)	0.001
Primitive/naïve lymphocytes in marrow(%, 33d)												
<5	367/1211	144/126	3.90 (2.98-5.10)	0.001	396/1149	104/178	1.72 (1.32-2.26)	0.001	460/1273	51/54	2.72 (1.82-4.05)	0.001
≥5	27/1211	5/126	1.83 (0.69-4.87)	0.223	23/1149	6/178	1.80 (0.72-4.50)	0.211	30/1273	2/54	1.61 (0.37-6.95)	0.523
MRD in marrow(%, 33d)												
<0.01	242/1211	50/126	2.06 (1.44-2.95)	0.001	231/1149	48/178	1.36 (0.96-1.94)	0.082	262/1273	30/54	2.83 (1.77-4.52)	0.001
≥0.01	149/1211	86/126	5.64 (4.07-7.80)	0.001	175/1149	59/178	2.19 (1.56-3.06)	0.001	209/1273	26/54	3.02 (1.84-4.92)	0.001
Primitive/naïve lymphocytes in marrow(%, 12w)												
<5	288/1211	47/126	1.65 (1.15-2.36)	0.007	269/1149	52/178	1.29 (0.92-1.81)	0.139	312/1273	23/54	1.90 (1.14-3.16)	0.014
≥5	12/1211	1/126	0.88 (0.11-6.92)	0.907	10/1149	3/178	2.14 (0.58-7.96)	0.257	13/1273	0/54	0.001 (0.00-999)	0.980
MRD in marrow(%, 12w)												
<0.01	282/1211	36/126	1.31 (0.88-1.94)	0.187	262/1149	43/178	1.10 (0.77-1.58)	0.600	299/1273	19/54	1.67 (0.97-2.88)	0.066
≥0.01	20/1211	6/126	2.89 (1.14-7.35)	0.026	18/1149	8/178	2.91 (1.24-6.81)	0.014	25/1273	1/54	0.95 (0.13-7.16)	0.962

[a]*Adjusted for age and gender.*

Stratification Analysis of Identified SNPs

We further analyzed whether the selected *METTL3* polymorphisms (rs1263801 C>G, rs1139130 A>G, and rs1061027 A>C) preferentially predispose to any specific subtype of ALL (**Table 2**). A stronger risk effect of rs1263801 was found among children older than 10 years (adjusted OR= 4.65, 95% CI=2.39–9.05, p<0.001), male (adjusted OR= 4.06, 95% CI=2.97–5.54, p=0.001), common B subtype ALL (adjusted OR= 8.59, 95% CI=6.37–11.6, p<0.001), MLL gene fusion type (adjusted OR= 13.1, 95% CI=4.77–35.9, p<0.001), low hyperdiploid (adjusted OR= 6.16, 95% CI=2.78–13.7, p<0.001), primitive/naive lymphocytes in marrow ≥ 5% on 19 days (adjusted OR=4.83, 95% CI=2.86–8.24, p<0.001) after chemotherapy, primitive/naive lymphocytes in marrow < 5% on 33 days (adjusted OR= 3.90, 95% CI=2.98–5.10, p<0.001) and 12 weeks (adjusted OR=1.65, 95% CI=1.15–2.36, p=0.007) after chemotherapy, MRD in marrow < 0.01% on 19 days (adjusted OR=10.6, 95% CI=5.81–19.5, p<0.001), MRD ≥ 0.01% on 33 days (adjusted OR=5.64, 95% CI=4.07–7.80–8.24, p<0.001), and ≥0.01% on 12 weeks (adjusted OR=2.89, 95% CI=1.14–7.35, p=0.026). As for the rs1139130 polymorphism, a more significant risk association was identified with those children age ≥10 years (adjusted OR=2.07, 95% CI= 1.06–4.05, p=0.034), female (adjusted OR=2.62, 95% CI= 18.1–3.78, p<0.001), common B subtype (adjusted OR=3.58, 95% CI= 2.67–4.81, p<0.001), MLL gene fusion type (adjusted OR=3.95, 95% CI=1.41–11.0, p=0.009), normal diploid (adjusted OR=1.88, 95% CI=1.45–2.45, p<0.001), primitive/naive lymphocytes in marrow ≥ 5% on 19 days (adjusted OR= 2.27, 95% CI=1.29–4.00, p<0.001) and <5% on 33 days (adjusted OR=1.72, 95% CI=1.32–2.26, p<0.001) after chemotherapy, MRD in marrow <0.01% on 19 days (adjusted OR= 4.66, 95% CI=2.52–8.60, p<0.001), MRD ≥ 0.01% on 33 days (adjusted OR= 2.19, 95% CI=1.56–3.06, p<0.001), and ≥0.01% on 12 weeks (adjusted OR= 2.91, 95% CI=1.24–6.81, p=0.014). As for the rs1061027 polymorphism, a stronger risk association was revealed with those children age ≥10 years (adjusted OR=7.10, 95% CI= 2.78–18.1, p<0.001), female (adjusted OR=4.84, 95% CI= 2.83–8.28, p<0.001), common B subtype (adjusted OR=9.72, 95% CI= 6.67–14.2, p<0.001), MLL gene fusion type (adjusted OR=14.8, 95% CI= 5.15–42.6, p<0.001), low hyperdiploid (adjusted OR=7.83, 95% CI= 3.17–19.5, p<0.001), primitive/naive lymphocytes in marrow ≥5% on 19 days (adjusted OR= 5.69, 95% CI=2.96–10.9, p<0.001) and <5% on 33 days (adjusted OR=2.72, 95% CI= 1.82–4.05, p<0.001) after chemotherapy, MRD in marrow <0.01% on 19 days (adjusted OR= 13.1, 95% CI=6.80–25.3, p<0.001), and MRD ≥0.01% on 33 days (adjusted OR= 3.02, 95% CI=1.84–4.92, p<0.001). No correlation was found between the rs1061027 polymorphism and MRD on 12 weeks.

Association of *METTL3* Polymorphisms With Chemotherapeutics in Southern Chinese Pediatric ALL Patients

All patients were treated with Chinese Children Cancer Group chemotherapeutics (CCCG) or South China Children Leukemia Group chemotherapeutics (SCCLG). We compared the MRD in

TABLE 3 | The correlation between METTL3 polymorphisms and South China pediatric ALL patients' response to different chemotherapeutics.

SNP	Variables		MRD in marrow (%, 19d)				MRD in marrow (%, 33d)				MRD in marrow (%, 12w)			
			Case (%)		P^a	Adjusted OR[a]	Case (%)		P^a	Adjusted OR[a]	Case (%)		P^a	Adjusted OR[a]
			<0.01	≥0.01		(95% CI)	<0.01	≥0.01		(95% CI)	<0.01	≥0.01		(95% CI)
rs1263801	CCCG-ALL-2015	GG/GC	4 (1.54)	256 (98.46)			194 (59.88)	130 (40.12)			238 (93.70)	16 (6.30)		
		CC	6 (4.69)	122 (95.31)	0.09	3.00 (0.83-10.9)	19 (20.88)	72 (79.12)	0	0.17 (0.10-0.30)	21 (80.77)	5 (19.23)	0.03	0.30 (0.10-0.90)
	SCCLG-ALL-2016	GG/GC	7 (23.33)	23 (76.67)			25 (71.43)	10 (28.57)			26 (89.66)	3 (10.34)		
		CC	9 (34.62)	17 (65.38)	0.31	1.89 (0.55-6.48)	19 (73.08)	7 (26.92)	0.6	1.39 (0.41-4.72)	9 (90.00)	1 (10.00)	0.64	1.85 (0.14-25.4)
rs1139130	CCCG-ALL-2015	GG/GA	4 (1.41)	279 (98.59)			173 (53.07)	153 (46.93)			215 (93.89)	14 (6.11)		
		AA	4 (4.40)	87 (95.60)	0.14	2.88 (0.70-11.9)	27 (36.00)	48 (64.00)	0.01	0.50 (0.29-0.83)	31 (81.58)	7 (33.33)	0.03	0.32 (0.12-0.87)
	SCCLG-ALL-2016	GG/GA	8 (20.00)	32 (80.00)			32 (71.11)	13 (28.89)			28 (89.66)	7 (10.34)		
		AA	8 (50.00)	8 (50.00)	0.02	5.70 (1.37-23.7)	12 (75.00)	4 (25.00)	0.53	1.56 (0.39-6.31)	7 (100.0)	0 (00.00)	0.97	999 (0.00-999)
rs1061027	CCCG-ALL-2015	CC/CA	4 (1.22)	324 (98.78)			205 (52.03)	189 (47.97)			251 (92.28)	21 (7.72)		
		AA	6 (10.00)	54 (90.00)	0	8.63 (2.31-32.3)	8 (38.10)	13 (61.90)	0.19	0.54 (0.22-1.35)	8 (100.0)	0 (0.00)	0.98	999 (0.00-999)
	SCCLG-ALL-2016	CC/CA	9 (25.71)	26 (74.29)			31 (75.61)	10 (24.39)			28 (90.32)	3 (9.68)		
		AA	7 (33.33)	14 (66.67)	0.4	1.74 (0.47-6.37)	13 (65.00)	7 (35.00)	0.61	0.72 (0.21-2.51)	7 (87.50)	1 (12.50)	0.77	1.50 (0.10-21.8)

[a]*Adjusted for age and gender.*

CCCG, Chinese Children Cancer Group; SCCLG, South China Children Leukemia Group.

the marrow of patients with different alleles after being treated with CCCG and SCCLG (**Table 3**). As for rs1263801, CC alleles provided a protective effect on MRD in the marrow ≥0.01% on 33 days (adjusted OR= 0.17, 95% CI= 0.10–0.30, p<0.001) and 12 weeks (adjusted OR= 0.30, 95% CI= 0.10–0.90, p=0.030) among CCCG-treated children. As for rs1139130, AA alleles provided a protective effect on MRD in marrow ≥0.01% on 33 days (adjusted OR= 0.50, 95% CI= 0.29–0.83, p=0.008) and 12 weeks (adjusted OR= 0.32, 95% CI= 0.12–0.87, p=0.030) among CCCG-treated children but provided a risk effect on MRD in marrow ≥0.01% among SCCLG-treated children (adjusted OR=5.70, 95% CI=1.37–23.7, p=0.017). As for rs1061.27, AA alleles provided a risk effect on MRD in the marrow ≥0.01% among CCCG-treated children (adjusted OR=8.63, 95% CI=2.31–32.3, p=0.002). These results indicated that SCCLG chemotherapeutics is more suitable for rs1263801 CC and rs1139130 AA carriers; CCCG chemotherapeutics is more efficient for rs1061027 AA carriers.

DISCUSSION

In the current case-control study with 808 pediatric ALL case and 1,340 healthy controls from Southern Chinese populations, we explored the potential association between *METTL3* gene polymorphisms and pediatric ALL risk. We certificated that three polymorphisms, namely rs1263801 C>G, rs1139130 A>G, and rs1061027 A>C, were associated with an increased susceptibility of pediatric ALL. To our knowledge, this study is the first to identify the association between *METTL3* polymorphisms and pediatric ALL susceptibility.

Epigenetic alterations, including DNA methylation, histone modifications, and noncoding RNAs, have been reported to contribute to ALL progression (22). In recent years, another epigenetic modification, RNA methylation, is considered to play an important role in carcinogenesis (11). M6A is the most common modification of RNA on the posttranscriptional level, mainly in mRNA and long noncoding RNA (lncRNA) (23). The complex composed of METTL3, METTL14, and WTAP induces m6A-methylation of mRNA or lncRNA. METTL3 is the essential component of the complex. Dysregulation of METTL3 was identified to be a key role in the progression of multiple malignant tumors, such as endometrial cancer (24), bladder cancer (25), pancreatic cancer (26), etc. A number of articles infer that METTL3 can promote tumor progression through multiple mechanisms. METTL3 can promote growth, survival, and invasion by interacting with the translation initiation element to enhance mRNA translation in lung adenocarcinoma (27). Lin et al. (28) revealed that deletion of METTL3 could impair the epithelial-mesenchymal transition (EMT). In breast cancer, METTL3 is upregulated by HBXIP and promotes the cancer progression by suppressing let-7g (29). METTL3 promotes self-renewal of glioblastoma stem cells to induce tumorigenesis (30). METTL3 can directly interact with the eukaryotic translation initiation factor e subunit h (eIF3h). The interaction between METTL3 and eIF3h is essential for translation and oncogenic transformation in lung cancer (31). Promoter-bound METTL3 promotes m6A modification within the coding region of mRNA transcript and enhances translation by inhibiting ribosome stalling. METTL3 regulates mRNA expression in this way to facilitate the progression of acute myeloid leukemia (32). However, the function of METTL3 in ALL is still unknown.

Herein, we investigated whether *METTL3* gene polymorphisms could influence the susceptibility of ALL in South China children for the first time. With regard to the remaining three *METTL3* gene polymorphisms (rs1263801 C>G, rs1139130 A>G, and rs1061027 A>C), we identified the association between these three SNPs and pediatric ALL susceptibility. The location and predicted function was analyzed using the online software SNPinfo. The rs1263801 C>G polymorphism was located in intron 1 of the *METTL3* gene and was predicted to be the transcriptional factor binding site. The rs1139130 A>G located in the exon 5 of the *METTL3* gene was predicted to affect splicing and protein function. The rs1061027 A>C polymorphism located in intron 8 was predicted to be associated with miRNA function. In 2018, Bertero et al. reported that the interactome of transcriptional factors SMAD2/3 promoted another transcriptional factor TGFβ to control the METTL3/METTL14/WTAP complex mediated m6A mRNA methylation in human pluripotent stem cells (33). Xia et al. reported that *Zmettl3* mutation disrupts gamete maturation and reduces fertility in zebrafish (34). Other studies identified that METTL3 mRNA could be targeted by miR-600 (35) and miR-33a (36). However, there was no evidence certifying that genetic variations of METTL3 could affect the transcriptional factor or miRNAs binding to METLL3 and the coding of METTL3 mRNA. Our results suggested that the rs1263801 CC phenotype, rs1139130 GG phenotypes, and rs1061027 CA/CC phenotypes are associated with an increased risk of pediatric ALL in South China. Lin et al. reported that the combination of rs1139130, rs1263801, rs1061026, and rs1061027 reduced the risk of Wilms tumor in Chinese children (37). Bian et al. (38) identified that these four polymorphisms were associated with an increased risk of neuroblastoma. It suggested that *METTL3* polymorphisms function diversely in different tumors.

We next examined whether the *METTL3* SNP genotype preferentially predisposes to any pediatric ALL subtype, including immunophenotype, gene fusion type, karyotype, primitive/naïve lymphocytes, and MRD in the marrow after chemotherapy. The *METTL3* rs1263801 CC phenotype and the rs1061027 AA phenotype were considered to increase the risk of ALL in the B-ALL, mature B ALL, and T-ALL subtype. In BCR-ABL, TEL-AML, and MLL gene fusion types, rs1263801 CC phenotype and rs1061027 AA phenotype carriers showed a higher risk for ALL. The rs1139130 GG carriers were revealed to have a higher risk for ALL in B-ALL, mature B ALL subtype, and medium risk level subtype. We failed to identify the association between the FAB subtype and these three *METTL3* polymorphisms.

In stratification analysis, we tried to reveal the relationship between clinical characteristic, response to different

chemotherapeutics, and *METTL3* polymorphisms. The results showed that rs1263801 C>G, rs1139130 A>G, and rs1061027 A>C could remarkably increase the risk of the common B type and MLL fusion type ALL in Southern Chinese children. All these three selected polymorphisms were more strongly associated with the primitive/naïve lymphocytes over 5% and MRD less than 0.01% on the 19th day, and also with the primitive/naïve lymphocytes less than 5% and MRD more than 0.01% on the 33rd day after chemotherapy. After chemotherapy treatment of 12 weeks, rs1263801 C>G and rs1061027 A>C were identified to increase susceptibility to primitive/naïve lymphocytes less than 5%; rs1263801 C>G and rs1139130 A>G may increase susceptibility to MRD more than 0.01% in ALL patients. And we also identified that SCCLG chemotherapeutics was more suitable for rs1263801 CC and rs1139130 AA carriers; CCCG chemotherapeutics was more efficient for rs1061027 AA carriers.

Several limitations should be noted in the current study. First, the sample size was not large enough. Second, this was a retrospective study; information bias and selection bias were inevitable. We have reduced these biases by frequency-matching of cases and controls by age and gender, and recruiting subjects from six hospitals in South China. Third, our study focused on the analysis of genetic factors in pediatric ALL risk. However, other important information such as environment and dietary intake was not available for analysis. Finally, the association between *METTL3* gene polymorphisms and prognosis of pediatric ALL was not analyzed in the current study.

In summary, our results suggest that polymorphisms rs1263801 C>G, rs1139130 A>G, and rs1061027 A>C in the *METTL3* gene were significantly associated with increased pediatric ALL risk, and SCCLG chemotherapeutics is more suitable for rs1263801 CC and rs1139130 AA carriers; CCCG chemotherapeutics is more efficient for rs1061027 AA carriers in the Southern Chinese ALL children. Further studies are necessary to elucidate the biological function of *METTL3* gene risk SNPs in the etiology of pediatric ALL.

CONCLUSION

METTL3 gene polymorphisms were associated with increased pediatric ALL risk. These three polymorphisms (rs1263801 C>G, rs1139130 A>G, and rs1061027 A>C) were likely to contribute to the sensitivity of different chemotherapies in pediatric ALL. The results indicated that METTL3 gene polymorphisms might be a potential biomarker for ALL susceptibility and when choosing chemotherapeutics.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

ETHICS STATEMENT

The study was approved by the institutional ethics committee of Guangzhou Women and Children's Medical Center, Guangzhou Medical University; The First Affiliated Hospital, Sun Yat-sen University; Sun Yat-sen Memorial Hospital, Sun Yat-sen University; Nanfang Hospital, Southern Medical University; and Zhujiang Hospital, Southern Medical University. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

XpL and LH are equal to this work in writing the manuscript. KH, LY, XW, JW, and YC collected ALL blood samples. XY analyzed the data. AL, MC, XdL, and YY performed qPCR. LX and HJ supplied the idea and funding. All authors contributed to the article and approved the submitted version.

FUNDING

National Natural Science Foundation of China (81672496 and 81870115), Natural Science Foundation of Guangdong Province (2020A1515010188), and Guangzhou Municipal Science and Technology Project (201804010042).

ACKNOWLEDGMENTS

We thank the Clinical Biological Resource Bank of Guangzhou Women and Children's Medical Center for providing part of the clinical samples.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.635251/full#supplementary-material

REFERENCES

1. Wu C, Li W. Genomics and Pharmacogenomics of Pediatric Acute Lymphoblastic Leukemia. *Crit Rev Oncol Hematol* (2018) 126:100–11. doi: 10.1016/j.critrevonc.2018.04.002
2. Schiffman JD. Applying Molecular Epidemiology in Pediatric Leukemia. *J Investig Med* (2016) 64(2):355–60. doi: 10.1097/JIM.0000000000000204
3. Jabbour E, Pui CH, Kantarjian H. Progress and Innovations in the Management of Adult Acute Lymphoblastic Leukemia. *JAMA Oncol* (2018) 4(10):1413–20. doi: 10.1001/jamaoncol.2018.1915
4. Eryilmaz E, Canpolat C. Novel Agents for the Treatment of Childhood Leukemia: An Update. *Onco Targets Ther* (2017) 10:3299–306. doi: 10.2147/OTT.S126368
5. Horowitz NA, Akasha D, Rowe JM. Advances in the Genetics of Acute Lymphoblastic Leukemia in Adults and the Potential Clinical

Implications. *Expert Rev Hematol* (2018) 11(10):781–91. doi: 10.1080/17474086.2018.1509702
6. Enciso-Mora V, Hosking FJ, Sheridan E, Kinsey SE, Lightfoot T, Roman E, et al. Common Genetic Variation Contributes Significantly to the Risk of Childhood B-Cell Precursor Acute Lymphoblastic Leukemia. *Leukemia* (2012) 26(10):2212–5. doi: 10.1038/leu.2012.89
7. de Smith AJ, Walsh KM, Francis SS, Zhang C, Hansen HM, Smirnov I, et al. BMI1 Enhancer Polymorphism Underlies Chromosome 10p12.31 Association With Childhood Acute Lymphoblastic Leukemia. *Int J Cancer* (2018) 143(11):2647–58. doi: 10.1002/ijc.31622
8. Deng X, Su R, Weng H, Huang H, Li Z, Chen J. RNA N(6)-Methyladenosine Modification in Cancers: Current Status and Perspectives. *Cell Res* (2018) 28(5):507–17. doi: 10.1038/s41422-018-0034-6
9. Weng H, Huang H, Chen J. RNA N (6)-Methyladenosine Modification in Normal and Malignant Hematopoiesis. *Adv Exp Med Biol* (2019) 1143:75–93. doi: 10.1007/978-981-13-7342-8_4
10. Huang Y, Su R, Sheng Y, Dong L, Dong Z, Xu H, et al. Small-Molecule Targeting of Oncogenic FTO Demethylase in Acute Myeloid Leukemia. *Cancer Cell* (2019) 35(4):677–91.e10. doi: 10.1016/j.ccell.2019.03.006
11. Li T, Hu PS, Zuo Z, Lin JF, Li X, Wu QN, et al. METTL3 Facilitates Tumor Progression *via* an M(6)A-IGF2BP2-Dependent Mechanism in Colorectal Carcinoma. *Mol Cancer* (2019) 18(1):112. doi: 10.1186/s12943-019-1038-7
12. Han J, Wang JZ, Yang X, Yu H, Zhou R, Lu HC, et al. METTL3 Promote Tumor Proliferation of Bladder Cancer by Accelerating Pri-Mir221/222 Maturation in M6a-Dependent Manner. *Mol Cancer* (2019) 18(1):110. doi: 10.1186/s12943-019-1036-9
13. Wu L, Wu D, Ning J, Liu W, Zhang D. Changes of N6-Methyladenosine Modulators Promote Breast Cancer Progression. *BMC Cancer* (2019) 19(1):326. doi: 10.1186/s12885-019-5538-z
14. Vu LP, Pickering BF, Cheng Y, Zaccara S, Nguyen D, Minuesa G, et al. The N (6)-Methyladenosine (M(6)A)-Forming Enzyme METTL3 Controls Myeloid Differentiation of Normal Hematopoietic and Leukemia Cells. *Nat Med* (2017) 23(11):1369–76. doi: 10.1038/nm.4416
15. Zheng Y, Nie P, Peng D, He Z, Liu M, Xie Y, et al. M6avar: A Database of Functional Variants Involved in M6a Modification. *Nucleic Acids Res* (2018) 46(D1):D139–45. doi: 10.1093/nar/gkx895
16. Lin W, Xu H, Wu Y, Wang J, Yuan Q. *In Silico* Genome-Wide Identification of M6a-Associated SNPs as Potential Functional Variants for Periodontitis. *J Cell Physiol* (2020) 235(2):900–8. doi: 10.1002/jcp.29005
17. Mo XB, Lei SF, Zhang YH, Zhang H. Detection of M(6)A-Associated SNPs as Potential Functional Variants for Coronary Artery Disease. *Epigenomics* (2018) 10(10):1279–87. doi: 10.2217/epi-2018-0007
18. Szymon S, Bik-Multanowski M, Balwierz W, Pietrzyk JJ, Surmiak M, Strojny W, et al. Homozygosity for the Rs9939609t Allele of the FTO Gene may Have Protective Effect on Becoming Overweight in Survivors of Childhood Acute Lymphoblastic Leukaemia. *J Genet* (2011) 90(2):365–8. doi: 10.1007/s12041-011-0089-3
19. Kwiecinska K, Strojny W, Pietrys D, Bik-Multanowski M, Siedlar M, Balwierz W, et al. Late Effects in Survivors of Childhood Acute Lymphoblastic Leukemia in the Context of Selected Gene Polymorphisms. *Ital J Pediatr* (2018) 44(1):92. doi: 10.1186/s13052-018-0526-5
20. He J, Wang F, Zhu J, Zhang R, Yang T, Zou Y, et al. Association of Potentially Functional Variants in the XPG Gene With Neuroblastoma Risk in a Chinese Population. *J Cell Mol Med* (2016) 20(8):1481–90. doi: 10.1111/jcmm.12836
21. He J, Zou Y, Liu X, Zhu J, Zhang J, Zhang R, et al. Association of Common Genetic Variants in Pre-microRNAs and Neuroblastoma Susceptibility: A Two-Center Study in Chinese Children. *Mol Ther Nucleic Acids* (2018) 11:1–8. doi: 10.1016/j.omtn.2018.01.003
22. Cruz-Rodriguez N, Combita AL, Zabaleta J. Epigenetics in Hematological Malignancies. *Methods Mol Biol* (2018) 1856:87–101. doi: 10.1007/978-1-4939-8751-1_5
23. Alarcon CR, Lee H, Goodarzi H, Halberg N, Tavazoie SF. N6-Methyladenosine Marks Primary microRNAs for Processing. *Nature* (2015) 519(7544):482–5. doi: 10.1038/nature14281
24. Liu J, Eckert MA, Harada BT, Liu SM, Lu Z, Yu K, et al. M(6)A mRNA Methylation Regulates AKT Activity to Promote the Proliferation and Tumorigenicity of Endometrial Cancer. *Nat Cell Biol* (2018) 20(9):1074–83. doi: 10.1038/s41556-018-0174-4
25. Jin H, Ying X, Que B, Wang X, Chao Y, Zhang H, et al. N(6)-Methyladenosine Modification of ITGA6 mRNA Promotes the Development and Progression of Bladder Cancer. *EBioMedicine* (2019) 47:195–207. doi: 10.1016/j.ebiom.2019.07.068
26. Zhang J, Bai R, Li M, Ye H, Wu C, Wang C, et al. Excessive miR-25-3p Maturation *via* N(6)-Methyladenosine Stimulated by Cigarette Smoke Promotes Pancreatic Cancer Progression. *Nat Commun* (2019) 10(1):1858. doi: 10.1038/s41467-019-09712-x
27. Lin S, Choe J, Du P, Triboulet R, Gregory RI. The M(6)A Methyltransferase METTL3 Promotes Translation in Human Cancer Cells. *Mol Cell* (2016) 62(3):335–45. doi: 10.1016/j.molcel.2016.03.021
28. Lin X, Chai G, Wu Y, Li J, Chen F, Liu J, et al. RNA M(6)A Methylation Regulates the Epithelial Mesenchymal Transition of Cancer Cells and Translation of Snail. *Nat Commun* (2019) 10(1):2065. doi: 10.1038/s41467-019-09865-9
29. Cai X, Wang X, Cao C, Gao Y, Zhang S, Yang Z, et al. HBXIP-Elevated Methyltransferase METTL3 Promotes the Progression of Breast Cancer *via* Inhibiting Tumor Suppressor Let-7g. *Cancer Lett* (2018) 415:11–9. doi: 10.1016/j.canlet.2017.11.018
30. Cui Q, Shi H, Ye P, Li L, Qu Q, Sun G, et al. M(6)A RNA Methylation Regulates the Self-Renewal and Tumorigenesis of Glioblastoma Stem Cells. *Cell Rep* (2017) 18(11):2622–34. doi: 10.1016/j.celrep.2017.02.059
31. Choe J, Lin S, Zhang W, Liu Q, Wang L, Ramirez-Moya J, et al. mRNA Circularization by METTL3-Eif3h Enhances Translation and Promotes Oncogenesis. *Nature* (2018) 561(7724):556–60. doi: 10.1038/s41586-018-0538-8
32. Barbieri I, Tzelepis K, Pandolfini L, Shi J, Millan-Zambrano G, Robson SC, et al. Promoter-Bound METTL3 Maintains Myeloid Leukaemia by M(6)A-Dependent Translation Control. *Nature* (2017) 552(7683):126–31. doi: 10.1038/nature24678
33. Bertero A, Brown S, Madrigal P, Osnato A, Ortmann D, Yiangou L, et al. The SMAD2/3 Interactome Reveals That TGFbeta Controls M(6)A mRNA Methylation in Pluripotency. *Nature* (2018) 555(7695):256–9. doi: 10.1038/nature25784
34. Xia H, Zhong C, Wu X, Chen J, Tao B, Xia X, et al. Mettl3 Mutation Disrupts Gamete Maturation and Reduces Fertility in Zebrafish. *Genetics* (2018) 208(2):729–43. doi: 10.1534/genetics.117.300574
35. Wei W, Huo B, Shi X. miR-600 Inhibits Lung Cancer *via* Downregulating the Expression of METTL3. *Cancer Manag Res* (2019) 11:1177–87. doi: 10.2147/CMAR.S181058
36. Du M, Zhang Y, Mao Y, Mou J, Zhao J, Xue Q, et al. MiR-33a Suppresses Proliferation of NSCLC Cells *via* Targeting METTL3 mRNA. *Biochem Biophys Res Commun* (2017) 482(4):582–9. doi: 10.1016/j.bbrc.2016.11.077
37. Lin A, Zhou M, Hua RX, Zhang J, Zhou H, Li S, et al. METTL3 Polymorphisms and Wilms Tumor Susceptibility in Chinese Children: A Five-Center Case-Control Study. *J Gene Med* (2020) 22(11):e3255. doi: 10.1002/jgm.3255
38. Bian J, Zhuo Z, Zhu J, Yang Z, Jiao Z, Li Y, et al. Association Between METTL3 Gene Polymorphisms and Neuroblastoma Susceptibility: A Nine-Centre Case-Control Study. *J Cell Mol Med* (2020) 24(16):9280–6. doi: 10.1111/jcmm.15576

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership