European Journal of Medical Research

RESEARCH

Open Access

Targeted gene sequencing and transcriptome sequencing reveal characteristics of *NUP98* rearrangement in pediatric acute myeloid leukemia

Jing-Ying Zhang[1,2,3†], Chun-Rong Chen[1,2,3†], Jia-Yue Qin[4†], Di-Ying Shen[1,2,3], Li-Xia Liu[4], Hua Song[1,2,3], Tian Xia[1,2,3], Wei-Qun Xu[1,2,3], Yan Wang[1,2,3], Feng Zhu[1,2,3], Mei-Xin Fang[1,2,3], He-Ping Shen[1,2,3], Chan Liao[1,2,3], Ao Dong[3,5], Shan-Bo Cao[4], Yong-Min Tang[1,2,3] and Xiao-Jun Xu[1,2,3*]

Abstract

Background *NUP98* rearrangements (*NUP98*-r) are rare but overrepresented mutations in pediatric acute myeloid leukemia (AML) patients. *NUP98*-r is often associated with chemotherapy resistance and a particularly poor prognosis. Therefore, characterizing pediatric AML with *NUP98*-r to identify aberrations is critically important.

Methods Here, we retrospectively analyzed the clinicopathological features, genomic and transcriptomic landscapes, treatments, and outcomes of pediatric patients with AML.

Results Nine patients with *NUP98*-r mutations were identified in our cohort of 142 patients. Ten mutated genes were detected in patients with *NUP98*-r. The frequency of *FLT3*-ITD mutations differed significantly between the groups harboring *NUP98*-r and those without *NUP98*-r ($P=0.035$). Unsupervised hierarchical clustering via RNA sequencing data from 21 AML patients revealed that *NUP98*-r samples clustered together, strongly suggesting a distinct subtype. Compared with that in the non-*NUP98*-r fusion and no fusion groups, *CMAHP* expression was significantly upregulated in the *NUP98*-r samples ($P<0.001$ and $P=0.001$, respectively). Multivariate Cox regression analyses demonstrated that patients harboring *NUP98*-r ($P<0.001$) and *WT1* mutations ($P=0.030$) had worse relapse-free survival, and patients harboring *NUP98*-r ($P<0.008$) presented lower overall survival.

Conclusions These investigations contribute to the understanding of the molecular characteristics, risk stratification, and prognostic evaluation of pediatric AML patients.

Keywords *NUP98* rearrangement, Acute myeloid leukemia, Molecular characteristics, Clinical features, Targeted next-generation sequencing

†Jing-Ying Zhang, Chun-Rong Chen and Jia-Yue Qin contributed equally to the work.

*Correspondence:
Xiao-Jun Xu
xuxiaojun@zju.edu.cn
Full list of author information is available at the end of the article

Background

Acute myeloid leukemia (AML) is a heterogeneous hematopoietic malignancy. *NUP98* rearrangements (*NUP98*-r) are fairly common in pediatric AML, occurring in more than 5% of cases [1–3]. This genomic anomaly is, however, rare in adult AML. Abnormal fusion of the *NUP98* gene at chromosome 11p15 may be a pathogenic driver that can be used for disease risk stratification and represents an important therapeutic target. Reverse transcription PCR (RT-PCR) assessment of *NUP98* fusions has significant value for evaluating measurable residual disease (MRD). Early case reports and subsequent cohort studies confirmed that the presence of *NUP98* gene fusion defines a high-risk leukemia subtype with unfavorable outcomes [4–6].

NUP98, part of the nuclear pore complex, plays an important role in mediating the selective transport of RNA molecules between the cytoplasm and nucleus while also participating in transcriptional regulation and mitotic progression [7, 8]. Furthermore, with the advent of genome and transcriptome sequencing, increasing numbers of *NUP98*-r in AML have been identified in pediatric patients, primarily due to balanced translocations and inversions. Since the initial discovery of the *NUP98*::*HOXA9* fusion gene in AML patients bearing t(7;11)(p15;p15), over 40 *NUP98* fusion partners have been identified in patients with hematologic malignancies (including *HOXD13*, *NSD1*, *PHF23*, and *TOP1*) [9].

Recently, our center identified nine cases of *NUP98*-r in a cohort of 142 pediatric AML patients. To improve our understanding of this fusion type, we explored the clinical and biological characteristics of this unique cohort via analysis of next-generation sequencing (NGS) data.

Methods

Study cohort

A retrospective analysis was performed on 9 newly diagnosed AML patients with *NUP98*-r and 133 AML patients without *NUP98*-r. All individuals, aged 0–18 years, were admitted to the Children's Hospital of Zhejiang University School of Medicine between January 1, 2018, and December 31, 2021, without additional exclusion criteria. The induction treatment regimens included standard 3 + 7 chemotherapies, such as daunorubicin and cytarabine (DA regimen), idarubicin and cytarabine (IA regimen), and other treatment regimens, such as homoharringtonine, cytarabine and recombinant human granulocyte colony-stimulating factor (HAG regimen). Consolidation therapy regimen was based on high-dose cytarabine or allogeneic hematopoietic stem cell transplantation (allo-HSCT), and sorafenib was added to chemotherapy for patients with *FLT3*-ITD mutations. The donor sources were haploidentical or umbilical cord blood. The pre-HSCT regimen included modified busulfan, carmustine, cytarabine and cyclophosphamide plus or minus anti-human thymocyte immunoglobulin (BUCY ± ATG regimen).

Among these patients, 100 underwent targeted NGS, 21 underwent RNA sequencing (RNA-Seq), and 94 had more than 1 year of follow-up data. The diagnosis of AML was based on the 2016 World Health Organization (WHO) classification criteria for hematopoietic and lymphoid tissue tumors [10]. All cases were confirmed by comprehensive diagnosis via cytomorphology, immunology, cytogenetics, and molecular biology. All essential and relevant data, including laboratory features at diagnosis, treatment modality, response to therapy, and follow-up data, were collected. Immunophenotypic analysis was conducted via flow cytometry to assess the expression of CD3, CD13, CD14, CD19, CD33, CD34, CD38, CD56, CD64, CD117, HLA-DR, and MPO. Chromosomal karyotyping by G-banding was performed via standard techniques, and karyotypes were described according to the International System for Human Cytogenetic Nomenclature [11].

Studies involving human participants were reviewed and approved by the Institutional Review Board of the Children's Hospital, Zhejiang University School of Medicine. Informed consent was obtained from patients and/or their legal guardians. All methods were performed in accordance with the relevant guidelines and regulations of the Declaration of Helsinki.

NUP98-r testing

NUP98-r testing was performed via RT–PCR. Total mRNA from bone marrow or peripheral blood samples was extracted using TRIzol reagent (TaKaRa, Japan), and the concentration and purity were evaluated using a NanoDrop 2000 (Thermo Fisher, USA). After the RNA was reverse-transcribed into cDNA using the PrimeScript RT Kit (TaKaRa, Japan) according to the manufacturer's instructions, the SYBR Premix Ex Taq TM Kit (TaKaRa, Japan) was used for qRT–PCR, with β-actin used as the endogenous control gene. A real-time quantitative PCR amplification instrument (ABI StepOne Plus) was used to detect the SYBR Green fluorescence signal after each amplification cycle. The primer sequences for the upregulated genes were synthesized by Sangon, China. The relative expression levels of the candidate genes were calculated via the 2-ΔΔCt method. The primer sequences are listed in Table S1.

RNA-Seq preprocessing and mapping

Total RNA was extracted using the EZ-press RNA Purification Kit according to the manufacturer's protocol. RNA purity and quantification were evaluated using a

NanoDrop 2000 spectrophotometer (Thermo Scientific, USA). RNA integrity was assessed via an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). Samples with an RNA integrity number (RIN) > 7 were subjected to subsequent analysis. Libraries were constructed using the TruSeg Stranded mRNA LT Sample Prep Kit (Illumina, San Diego, CA, USA) according to the manufacturer's instructions. These libraries were sequenced on the Illumina NovaSeq sequencing platform, and 150 bp paired-end reads were generated. HISAT2 (v2.0.5) was used to align the raw sequences to the reference genome Hg38 downloaded from the UCSC Genome Browser (http://hgdownload.soe.ucsc.edu/). Preprocessing steps were carried out according to the GATK best practices pipeline [12]. Arriba v2.1.0 was used for fusion gene calling [13]. To ensure the functionality and reliability of the identified fusions. Specifically, we removed (1) fusion breakpoints with fewer than two split reads or three spanning reads, (2) fusions reported in the healthy population, (3) fusions appearing on the blacklist, and (4) fusions associated with uncharacterized or mitochondrial genes. Out-of-frame fusions were also removed because their protein products were likely to be lost or rapidly degraded. A transcript-level read-count matrix was generated via HTSeq-count (v0.5.4. p3) [14] on the basis of the GENCODE annotation database [15]. BAM files were generated using HISAT2. Differentially expressed genes were identified via DESeq2 (v1.18.1) on the basis of the read count matrix. Fragments per kilobase per million (FPKM) values were then calculated and log-transformed to evaluate gene expression levels.

To compare the functional characteristics of the 3 subgroups within the samples (i.e., *NUP98*-r, non-*NUP98*-r fusion and no fusion AML), gene set enrichment analysis (GSEA) was performed via GSEA software (v3.0, http://software.broadinstitute.org/gsea) with Kyoto Encyclopedia of Genes and Genomes (KEGG) gene sets [16]. The ward.D algorithm in R's cluster method was used for unsupervised clustering, where the different numbers (1–15%) of top-ranked highly variable genes across all samples were selected to evaluate the stability. The Complex Heatmap R package was used to visualize all the samples [17].

Targeted next-generation sequencing

Genomic DNA was extracted from the bone marrow at diagnosis. A 34-gene targeted panel, with mutated genes frequently involved in hematological malignancies, was used for sequencing via the Illumina NovaSeq platform (Table S2). To filter raw variant results, the following criteria were used: average sequencing depth on target ≥ 1000×, base quality ≥ 30, mapping quality ≥ 30, and variant allele frequency (VAF) ≥ 1% for single nucleotide variation (SNV) and insertion or deletion (InDel). Burrows–Wheeler alignment (BWA, version 0.7.12) was used to align trimmed reads. The Mark Duplicates tool from Picard was used to identify PCR duplicates. IndelRealigner and BaseRecalibrator from the Genome Analysis Toolkit (GATK, version 3.8) were used for realignment and recalibration of the BWA data, respectively. Variant calling was performed via Mutect2. ANNOVAR software was used to annotate all the variants, including 1000G projects, COSMIC, SIFT, and PolyPhen.

Statistical analysis

Statistical analyses were performed via SPSS (version 22.0) or R (version 3.5.2) software. For categorical variables, comparisons were evaluated using the chi-square test or Fisher's exact test, and for continuous variables, the Mann–Whitney U test was used. Survival analysis was performed using the Kaplan–Meier method with the log-rank test. Variables with $P < 0.2$ in the univariate analysis were examined via multivariate analysis using a Cox proportional hazards model to identify the statistically significant parameters for relapse-free survival (RFS) and overall survival (OS). RFS was calculated from complete remission (CR) to relapse, death, or the last follow-up. OS was calculated from diagnosis to death or the last follow-up. All tests were two-sided, and statistical significance was set at $P < 0.05$.

Results

Patient characteristics

A total of 142 pediatric patients with AML were included in the study. The median WBC count, hemoglobin concentration, and platelet count were 15.9×10^9/L, 89.0 g/L and 62.0×10^9/L, respectively. The main morphological types were M5 (65 patients, 45.8%) and M2 (48 patients, 33.8%). Among the 140 patients with karyotype results, 20 (14.3%) had complex karyotype abnormalities, and 35 (25.0%) had normal chromosomal karyotypes. Among the 142 patients with fusion results, the main AML fusion gene types detected were *NUP98*-r (9 patients, 6.3%), *RUNX1*::*RUNX1T1* (31 patients, 21.8%), *CBFB*::*MYH11* (8 patients, 5.6%), and *KMT2A* rearrangement (*KMT2A*-r, 30 patients, 21.1%). Rare fusions were detected in 7 patients (4.9%). The initial clinical features of the cohort are shown in Table S3.

In 100 pediatric AML patients whose NGS data were available, the overall mutation prevalence rate was 87.0%. A total of 30 mutated genes were detected, and the most common mutated gene was *NRAS* (25.0%), followed by *FLT3*-ITD (15.0%) and *KRAS* (15%) (Fig. 1A). Gene function classification revealed that the greatest proportion of mutations were related to signaling pathways

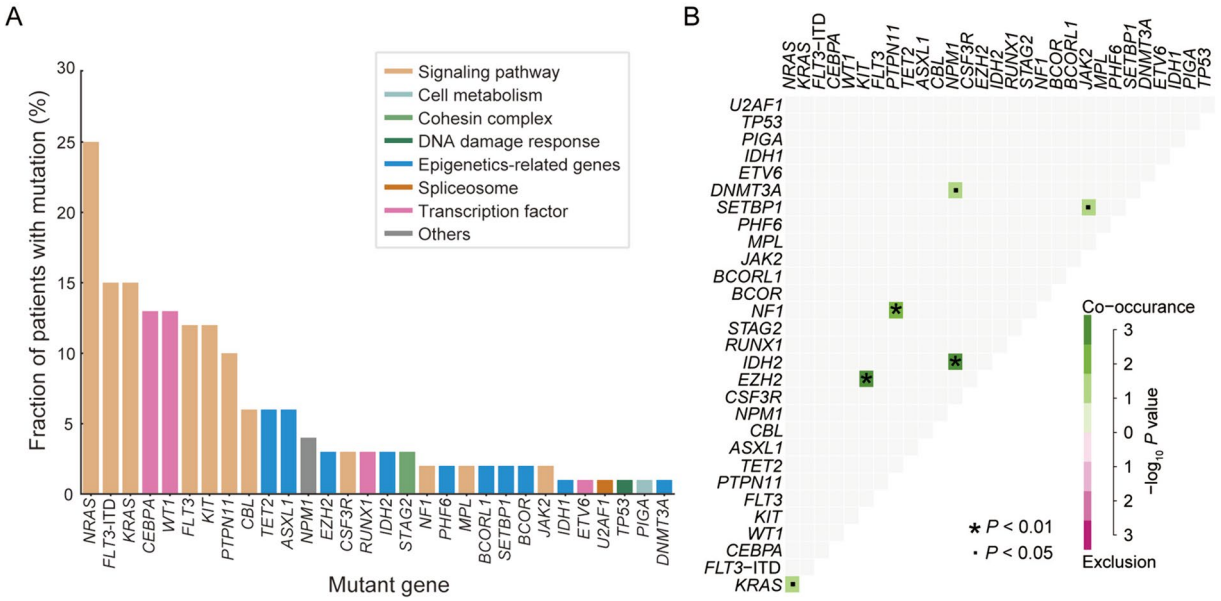

Fig. 1 Functional classification and correlation of detected mutations in 100 patients with pediatric AML. **A** Histogram showing the frequencies of the mutated genes. The different colors represent different genetic pathways. **B** Pairwise associations among the genetic mutations. The co-occurrence of each association is color coded, and the significance level is indicated by an asterisk or dot in each field. *AML* acute myeloid leukemia

(60.5%), followed by transcription factor-related genes (17.4%) (Fig. S1). Significant associations were discovered between mutated *NPM1* and mutations in *DNMT3A* and *IDH2*, and between *KIT* and *EZH2* mutations (Fig. 1B). In addition, a total of 221 variants were detected, of which 130 nonsynonymous SNV sites were found (58.5%), followed by 30 nonframeshift insertions, 28 frameshift insertions, 15 stop-gains, 7 frameshift deletions, 6 nonframeshift deletions, and 5 splice-site variants (Table S4).

Comparison of clinical characteristics between patients with *NUP98*-r and those without *NUP98*-r

The median age of the AML patients with *NUP98*-r was 4.4 years, with a male-to-female ratio of 2:1, whereas the median age and male-to-female ratio of patients without *NUP98*-r were 5.9 years and 1.89:1, respectively. The WBC count, hemoglobin level, and platelet count were comparable between the groups (Table 1).

On the basis of flow cytometry analysis of 136 patients, leukemia blasts with *NUP98*-r were usually positive for CD33 (100%), CD13 (66.7%), CD117 (77.8%), CD45 (100%), CD34 (66.7%), and MPO (44.4%) but were only occasionally positive for CD56 (11.1%) and CD11b (22.2%). There were no statistically significant differences in the expression of CD surface markers between AML patients with *NUP98*-r and those without *NUP98*-r (Table 1, Fig. S2).

In *NUP98*-r AML patients, leukemic blasts of M2 (4 patients; 44.4%) and M5 (3 patients; 33.3%) morphology accounted for approximately 77.8%, which was similar to patients without *NUP98*-r (106 patients, 79.7%). However, Patient 9, with *NUP98*-r, had a cytomorphology similar to that of M3 variant leukemia (high granular feature bundles were randomly distributed in the cytoplasm), a morphology not otherwise observed in patients without *NUP98*-r (Fig. S3). We found no statistically significant difference in the incidence of complex karyotypes between patients with *NUP98*-r and those without *NUP98*-r (3/9, 33.3% vs. 17/131, 12.8%, $P=0.232$; Table 1).

Comparison of the molecular characteristics between patients with *NUP98*-r and without *NUP98*-r

Of the nine *NUP98* rearranged cases, three had *NUP98*::*KDM5A*, five had *NUP98*::*NSD1*, and one had *NUP98*::*RARG*. The detailed fusion sites in the six patients with available RNA-Seq data are shown in Fig. S4. The *NUP98* fusion sites included exons 21, 22, and 23 on chromosome 11. As shown in Fig. 2A, the frequencies of concurrent mutations were compared between the patients with *NUP98*-r and those without *NUP98*-r. A total of 10 mutated genes were detected in *NUP98*-rearranged AML patients. The frequency of *FLT3*-ITD mutations differed significantly between the two groups (44.4% vs. 12.1%, $P=0.035$). The total

Table 1 Comparison of patient characteristics between AML with *NUP98*-r and those without *NUP98*-r

Characteristics	Patients with *NUP98*-r (n = 9)	Patients without *NUP98*-r (n = 133)	*P* value
Male, n (%)	6 (66.7%)	87 (65.4%)	0.775
Age, M (range) years	4.4 (1.5–14.2)	5.9 (0.3–18.0)	0.864
WBC, M (range) × 10^9/L	34.4 (3.3–375.2)	15.8 (1.1–557.2)	0.273
Hb, M (range) × g/L	94.0 (36.0–107.0)	88.5 (22.0–135.0)	0.826
PLT, M (range) × 10^9/L	90.0 (10.0–595.0)	62.0 (2.0–571.0)	0.607
BM blast (%)	75.0 (38.0–92.0)	69.5 (9.0–97.0)	0.717
Immunotyping			
CD33+	9/9 (100%)	117/127 (92.1%)	1.000
CD13+	6/9 (66.7%)	86/127 (67.7%)	0.761
MPO+	4/9 (44.4%)	57/127 (44.9%)	0.748
CD117+	7/9 (77.8%)	106/127 (83.5%)	0.984
HLA-DR+	6/9 (66.7%)	103/126 (81.7%)	0.502
CD34+	6/9 (66.7%)	69/127 (54.3%)	0.710
CD11b+	2/9 (22.2%)	41/127 (32.3%)	0.798
CD45+	9/9 (100%)	125/127 (98.4%)	1.000
CD15+	4/9 (44.4%)	74/127 (58.3%)	0.644
CD4+/CD7+	4/9 (44.4%)	67/127 (52.8%)	0.891
CD56+	1/9 (11.1%)	46/126 (36.5%)	0.237
Morphology			
M0	0 (0%)	5 (3.8%)	1.000
M1	0 (0%)	2 (1.5%)	1.000
M2	4 (44.4%)	44 (33.1%)	0.739
M4	1 (11.1%)	17 (12.8%)	0.710
M5	3 (33.3%)	62 (46.6%)	0.668
M6	0 (0%)	1 (0.8%)	1.000
M7	0 (0%)	2 (1.5%)	1.000
APL-like	1 (11.1%)	0 (0.0%)	0.063
Cytogenetics			
Normal karyotype	5/9 (55.6%)	30/131 (22.6%)	0.073
Complex karyotype	3/9 (33.3%)	17/131 (12.8%)	0.232

AML acute myeloid leukemia, *NUP98-r NUP98* rearrangement, *WBC* white blood cell, *HB* haemoglobin, *PLT* platelet, *BM* bone marrow

number of mutations per patient ranged from 0 to 5 (Fig. 2B). The median VAF of *FLT3*-ITD was significantly different between the patients with *NUP98*-r and those without *NUP98*-r (48.5% vs. 34.7%, $P = 0.037$; Fig. 2C).

Unsupervised hierarchical clustering of gene expression profiles was performed to compare *NUP98*-r samples (n = 6) with non-*NUP98*-r fusion samples (n = 12), no fusion samples (n = 3), and a normal control sample (n = 1). The fusion type details of *NUP98*-r and non-*NUP98*-r fusions for the RNA-Seq data are shown in Table S5. All *NUP98*-r samples clustered together, strongly suggesting that it was a subgroup of AML (Fig. 3A). Figure 3B shows the differentially expressed genes between the different groups (*NUP98*-r vs. non-*NUP98*-r fusion, *NUP98*-r vs. no fusion, and *NUP98*-r vs. normal control). Compared with that in the non-*NUP98*-r fusion and no fusion groups, *cytidine monophospho-N-acetylneuraminic acid hydroxylase pseudogene* (*CMAHP*) expression was significantly upregulated in the *NUP98*-r group ($P < 0.001$ and $P = 0.001$, respectively, Fig. 3B), indicating that *NUP98*-r promotes *CMAHP* expression. Owing to the small sample size and limited power of the RNA-Seq data generated, we explored the Therapeutically Applicable Research to Generate Effective Treatments (TARGET) database to verify these results, and we found that *CMAHP* was highly expressed in the *NUP98*-r group ($P < 0.001$, Fig. S5). Moreover, we discovered that the expression of differentially expressed *HOX* cluster genes was upregulated in the *NUP98*-r group compared with the non-*NUP98*-r group ($P = 0.015$, Fig. 3C, D), strongly indicating that *HOX* genes play important roles in *NUP98*-r-mediated leukemogenesis. All *HOX* gene expression details are shown in Fig. S6. Gene set enrichment analysis (GSEA) based on KEGG demonstrated that in the *NUP98*-r group, genes were significantly upregulated in the calcium signaling pathway and significantly downregulated in the B-cell receptor signaling pathway (both $P < 0.050$, Fig. S7).

Treatment and outcomes

Among the nine *NUP98* rearranged patients evaluable for treatment response evaluation, five patients achieved complete remission (CR) after 2–3 cycles of induction chemotherapy (anthracycline combined with cytarabine), and four patients experienced CR after one cycle of a homoharringtonine-based regimen. With a median follow-up time of 22.3 months (range 0.2–60.6 months), survival analysis revealed that *NUP98*-r was associated with both lower RFS ($P < 0.001$, Fig. 4A) and lower OS ($P < 0.001$, Fig. 4B).

We analyzed several prognostic variables via a Cox regression model to investigate the prognostic effects of molecular genetics and clinical variables. In the univariate analysis of RFS, a WBC count ≥ 100 × 10^9/L, a MRD after one cycle of induction ≥ 5%, a MRD ≥ 1 × 10^{-3} after two cycles of induction, complex karyotype, *KMT2A* aberration (including *KMT2A*-r and *KMT2A*-PTD), *RUNX1*::*RUNX1T1* fusion, and mutations in *KIT*, *FLT3*-ITD, *FLT3* and *NRAS* were not significant predictors of RFS, whereas *NUP98*-r ($P < 0.001$), *RUNX1*::*RUNX1T1* fusion ($P = 0.022$) and *WT1* mutation ($P = 0.030$) were associated with RFS (Table 2, Fig. 4A). Multivariate analyses revealed that patients with *NUP98*-r ($P < 0.001$) and *WT1* mutations ($P = 0.030$) had significantly lower RFS rates, with hazard ratios (HRs) of 8.72 and 3.15,

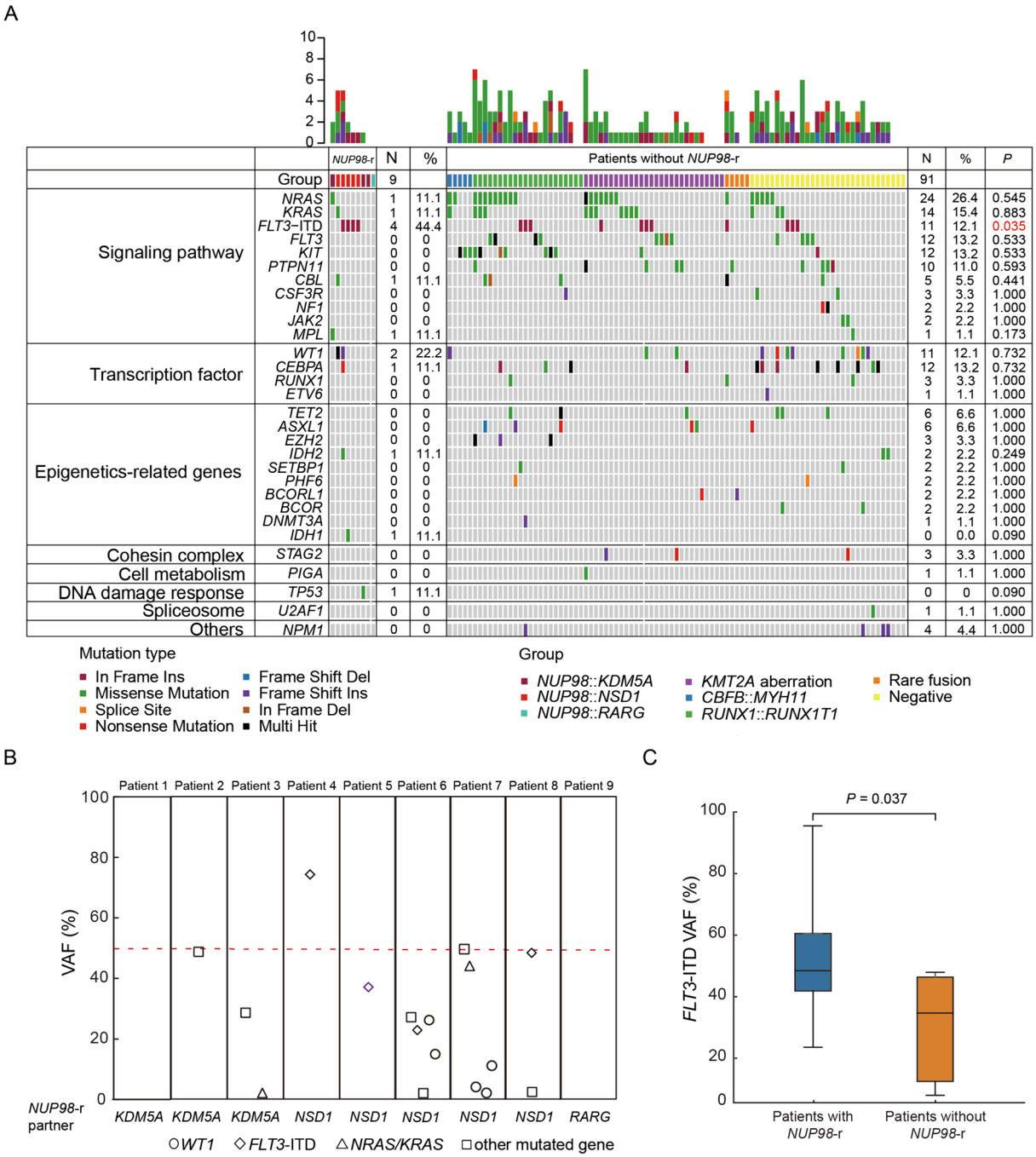

Fig. 2 Distribution of genetic mutations in 100 patients with pediatric AML. **A** Heatmap showing the mutation details for each of the 100 pediatric AML patients. Each column represents one patient, and each row corresponds to a mutation in the defined genes. The colors represent mutated genes for missense variants (green), nonsense variants (bright red), frameshift indel variants (blue), in-frame indel variants (brown), frameshift ins (purple), and in-frame ins (dark red). The top bar indicates the number of mutations, and the right bar indicates the frequency of the different mutated genes. **B** Mutated genes in each patient with *NUP98*-r. **C** Comparison of VAFs in mutated *FLT3*-ITD between patients with *NUP98*-r and without *NUP98*-r. *AML* acute myeloid leukemia, *VAF* variant allele frequency, *NUP98-r NUP98* rearrangement

respectively (Table 2). Univariate analysis of OS data revealed that a MRD after one cycle of induction ≥ 5% ($P=0.013$) and *NUP98*-r ($P<0.001$) were both adverse prognostic factors, whereas *RUNX1*:*RUNX1T1* fusion ($P=0.014$) was a favorable prognostic factor (Table 3, Fig. 4B). Multivariate analyses revealed that *NUP98*-r

was a unique poor prognostic factor for OS (HR = 4.24, $P = 0.008$; Table 3).

Discussion

NUP98 fusions are most commonly observed in pediatric AML. Since its first description in 1996 with the initial identification of the *NUP98*::*HOXA9* fusion gene, over 40 *NUP98* fusion partners have been identified in patients with hematological malignancies [18, 19]. In this study, we identified nine patients carrying *NUP98*-r with three different fusion gene partners. *NSD1* (5q35) was the most frequent partner detected in five patients. *NSD1* encodes a lysine methyltransferase responsible for the methylation of proteins such as histones H3 and H4 [20]. Previous reports have suggested that its frequency is greater in AML patients with a normal karyotype. As demonstrated in our study, all five cases of AML with *NUP98*::*NSD1* had normal karyotypes (5/9, 55.6%), which is consistent with other reports [2, 5]. The second partner identified in 3 of 9 patients was the lysine demethylase *KDM5A* (12p13) [21]. *NUP98*::*KDM5A* was detected in 9% of acute megakaryoblastic leukemia (M7) patients according to reports by de Rooij et al. [22]; however, another study by Struski et al. [2, 23] found no *NUP98* rearrangement when an AML M7 patient series was analyzed. As observed in our study, *NUP98*::*KDM5A* was found mainly in M2 and M5 rather than in AML M7. We discovered that the *NUP98*::*KDM5A* fusion rarely cooccurred with the *WT1* mutation in our cohort. In addition, two patients with *NUP98*::*KDM5A* and chr13 alterations have recently been identified, which is consistent with the report that chr13 abnormalities are predominantly present in the *NUP98*::*KDM5A* subgroup [24]. Owing to the limited number of patients in our series, the prognostic value of chr13 abnormalities in *NUP98*::*KDM5A* cases was not observed, whereas a recent report revealed that patients harboring chr13 alterations within the *NUP98*::*KDM5A* group were associated with increased OS and event-free survival (EFS) compared with patients without chr13 alterations [24]. The last *NUP98* fusion gene partner identified in one patient was *RARG* (12q13). On the basis of our previous findings, *RARG* is a different form of *RARa*, and AML with *NUP98*::*RARG* has strikingly similar features to acute promyelocytic leukemia (APL). However, these patients do not respond to standard treatment with all-trans retinoic acid (ATRA) and arsenic trioxide (ATO) and have poor outcomes [23]. Wei et al. discovered that an alkaloid-based regimen was beneficial for a patient with APL-like AML harboring an *NUP98*::*RARG* fusion [25].

Larger cohorts of patients have confirmed that the presence of *NUP98* gene fusion defines a high-risk leukemia subset. For example, a report by Struski et al. demonstrated that patients with *NUP98*-r were refractory to standard 3 + 7 induction chemotherapy and had a low CR rate (67%) [2]. In this study, we observed a similar phenomenon in patients with *NUP98* gene fusions, with higher rates of one course of induction failure (55.6%). Surprisingly, we found that AML patients with *NUP98*-r were sensitive to homoharringtonine and that 4 patients achieved CR after receiving one course of HAG therapy. The HAG regimen has achieved good effects in induction therapy, and several patients' tolerances were changed to obtain the opportunity for allo-HSCT. However, the effectiveness of this treatment regimen needs to be validated in more patients and explored both in vitro and in vivo.

Furthermore, we combined specific mutations, including those in *FLT3*-ITD, *WT1* and *CEBPA*, and RNA-Seq analysis of *NUP98* rearranged pediatric AML patients to investigate the associated oncogenic pathways. *FLT3*-ITD and *WT1* mutations were frequently associated with *NUP98* fusion genes. *FLT3*-ITD was associated with *NUP98*::*NSD1* in 67 to 91% of reported cases [26, 27], similar to our frequency of 80%. Among pediatric AML patients with *NUP98*::*NSD1*, the 3-year OS was significantly lower than that in those without the *FLT3*-ITD mutation [28]. Our dataset frequency of 22.2% for *WT1* mutations was lower than reported frequencies ranging from 33 to 50% [27]. In our series, 4/9 patients had ≥ 2 mutations, and all 4 patients had died by the end of follow-up. Interestingly, we found a single-allele variant in *CEBPA* (Patient 6), as well as variants in epigenetic modifier genes such as *IDH1*, *IDH2*, *ASXL1*, *DNMT3A*

(See figure on next page.)

Fig. 3 Distribution of RNA clustering and expression in 21 pediatric AML patients. **A** Unsupervised hierarchical clustering identified a unique gene expression pattern of *NUP98*-r in our cohort. AML patients with non-*NUP98*-r fusion and no fusion were included for comparison. The columns indicate AML patients, and the rows represent the gene expression levels for each patient. Genes showing over- or underexpression in the heatmap are shown in red or blue, respectively. **B** DEG comparisons in AML patients with *NUP98*-r, non-*NUP98*-r fusion, no fusion and a normal control in our cohort. Volcano plots showing DEGs between different groups (*NUP98*-r vs. non-*NUP98*-r-fused AML, *NUP98*-r vs. nonfused AML, and *NUP98*-r vs. the normal control). The most significant DEGs are labeled, and *CMAHP* is highlighted. **C** Heatmap showing differentially expressed *HOX* cluster gene expression in different subgroups. **D** Expression comparison of differentially expressed *HOX* genes between patients in the *NUP98*-r group and those in the non-*NUP98*-r group. *NUP98-r NUP98* rearrangement, *AML* acute myeloid leukemia, *NA* not applicable, *DEG* differentially expressed gene

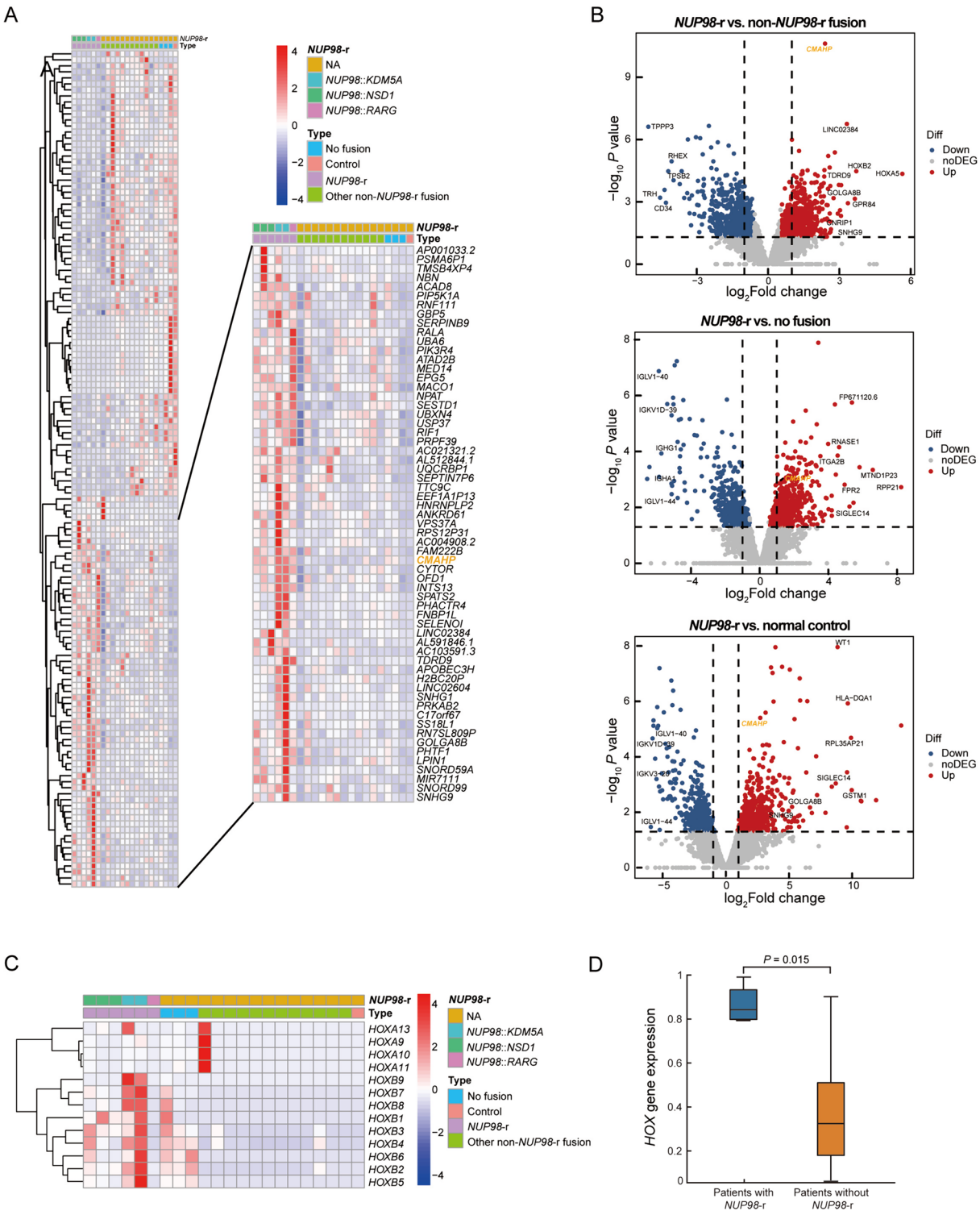

Fig. 3 (See legend on previous page.)

and *TET2*. Epigenetic regulation genes play important roles in the pathogenesis of *NUP98*-r AML. The *NUP98* fusion protein has a SET domain methyltransferase site, which participates in cell proliferation and differentiation by regulating the methylation of domain sites and histone acetylation, thereby promoting AML occurrence [29].

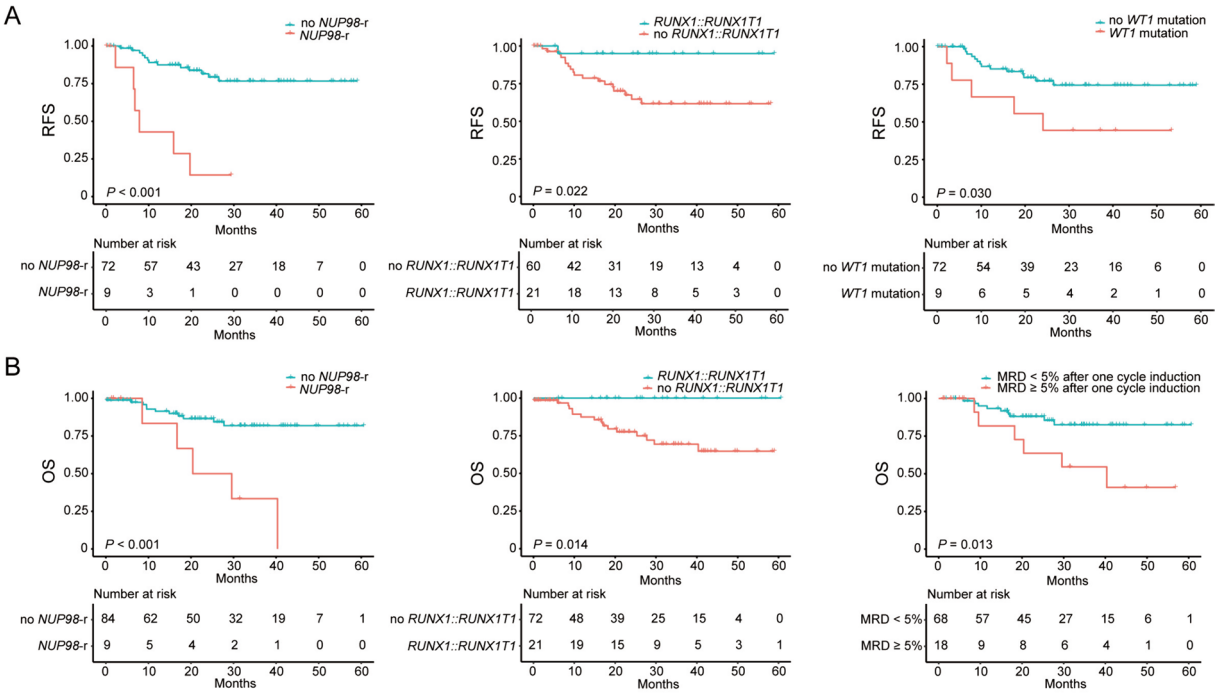

Fig. 4 Survival analysis based on Kaplan–Meier plots and two-sided log-rank tests. **A** RFS was stratified by *NUP98*-r, *RUNX1*::*RUNX1T1* fusion, and *WT1* mutation. **B** OS was stratified by *NUP98*-r, *RUNX1*::*RUNX1T1* fusion, and a MRD after one cycle induction ≥ 5%. *RFS* relapse-free survival, *NUP98-r NUP98* rearrangement, *OS* overall survival, *MRD* measurable residual disease

Table 2 Univariate and multivariate Cox regression analysis of RFS in pediatric AML

Variable	Univariate analysis		Multivariate analysis	
	HR	*P* value	HR	*P* value
WBC ≥ 100×10^9/L	0.783	0.743		
MRD after one cycle induction ≥ 5%	2.451	0.080		
MRD after two cycle induction ≥ 1×10^{-3}	1.155	0.752		
Complex karyotype	0.828	0.800		
NUP98-r	7.934	< 0.001	8.720	< 0.001
KMT2A aberration	0.993	0.989		
RUNX1::*RUNX1T1* fusion	0.135	0.022		
WT1 mutation	2.930	0.030	3.152	0.030
KIT mutation	0.290	0.199		
FLT3-ITD mutation	2.394	0.109		
FLT3 mutation	0.283	0.189		
NRAS mutation	0.363	0.157		

RFS relapse-free survival, *AML* acute myeloid leukemia, *HR* hazard ratio, *WBC* white blood cell, *MRD* minimal residual disease, *NUP98-r NUP98* rearrangement

Table 3 Univariate and multivariate Cox regression analysis of OS in pediatric AML

Variable	Univariate analysis		Multivariate analysis	
	HR	*P* value	HR	*P* value
WBC ≥ 100×10^9/L	1.569	0.477		
MRD after one cycle induction ≥ 5%	3.428	0.013		
MRD after two cycle induction ≥ 1×10^{-3}	0.629	0.373		
Complex karyotype	1.341	0.645		
NUP98-r	5.529	< 0.001	4.241	0.008
KMT2A aberration	1.801	0.268		
RUNX1::*RUNX1T1* fusion	0.000	0.014		
WT1 mutation	2.260	0.146		
FLT3-ITD mutation	2.638	0.080		
FLT3 mutation	0.000	0.095		

OS overall survival, *AML* acute myeloid leukemia, *HR* hazard ratio, *WBC* white blood cell, *MRD* minimal residual disease, *NUP98-r NUP98* rearrangement

Transcriptome profiling further defined the *NUP98*-r functional classifications. Through RNA-Seq analysis, we also detected significantly high expression of *CMAHP* in *NUP98*-r AML. The expression of *CMAHP* clearly segregated *NUP98*-r subsets and seemed to be a common feature of *NUP98* translocations. *CMAHP* is considered a pseudogene, and it is inactivated by a 92 bp exon deletion due to Alu-mediated replacement of the genome [30, 31]. A recent report demonstrated that *CMAHP* is involved

in gastric cancer progression through the modulation of metastasis and angiogenesis [30]. In addition, *CMAHP* is more highly expressed in primitive $CD133^+$ hematopoietic stem cells than in $CD133^-$ cells, making *CMAHP* a novel stem cell marker [32]. Tuborgh et al. [33] reported that the apparent fusion gene *CMAHP-KMT2A* was identified in an infant myeloid leukemia patient, implying that it may contribute to the malignant transformation of hematopoietic stem cells. *CMAHP* could be a promising target for antitumor therapy in *NUP98*-r AML. Previous studies have shown that *NUP98* translocations are associated with the overexpression of the *PRDM16*, *MECOM* and *HOXA*/*B* genes [34, 35]. *PRDM16* and *MECOM* encode H3K9 methyltransferases that are important for the maintenance of heterochromatin integrity and are selectively expressed in hematopoietic stem cells and linked to oncogenic transformation. The *HOX* cluster was shown to be in a locked, transcriptionally active position due to the H3K4me3-binding PHD domain when it was fused to *NUP98* [35]. Translationally, the upregulation of the above genes indicates a potential therapeutic role of menin inhibitors in *NUP98*-r AML, as has been recently shown in mice [36]. Via gene set enrichment analysis, we also discovered that genes in the *NUP98*-r group were significantly upregulated in the calcium signaling pathway, which plays an important role in AML cell proliferation and differentiation and in the quiescence of hematopoietic stem cells [37]. Based on multivariate analyses, we discovered that patients with *WT1* mutations had significantly lower RFS than those without *WT1* mutations ($P=0.030$). As the fusion types of *NUP98*-r were heterogeneous, the outcome was susceptible to change and ultimately dependent on a larger sample size to verify its validity.

Our study has several limitations, including the small sample size and heterogeneity of the molecular subtypes. *NUP98*-r fusions, classified by the WHO as a distinct entity within the AML subcategory, show important biological diversity and may originate from differences in the defining gene rearrangement event and/or a different spectrum of cooperative mutations. In the future, comprehensive whole-genome sequencing and single-cell transcriptomics are needed to clarify the molecular and genetic landscapes of pediatric AML.

Conclusion

In summary, our research elucidated the clinical and molecular characteristics of *NUP98*-r in pediatric AML and revealed that *NUP98*-rearranged AML constitutes a homogeneous group associated with a poor prognosis. These investigations contribute to the understanding of the molecular characteristics, risk stratification, and prognostic evaluation of pediatric AML patients.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40001-024-02042-9.

Supplementary Material 1: Figure S1. Pie chart showing the fractions of mutations according to the different pathways.

Supplementary Material 2: Figure S2. Heatmap of immunotyping markers. Immunotyping markers included CD33, CD13, MPO, CD117, HLA-DR, CD34, CD11b, CD45, CD15, CD4/CD7, and CD56. *NUP98*-r, *NUP98* rearrangement.

Supplementary Material 3: Figure S3. Typical representative bone marrow morphology. The bone marrow morphology included the M2 type, M3-like type, M4 type, and M5 type.

Supplementary Material 4: Figure S4. The genetic landscape of *NUP98*-r. Schematic diagrams showing the structural composition of fusion proteins among samples with *NUP98*-r. Functional domains of proteins are separately colored according to the keys. For each protein, the longest isoform with the support of RNA-Seq read alignment was chosen. *NUP98*-r, *NUP98* rearrangement; *RNA-Seq*, RNA sequencing.

Supplementary Material 5: Figure S5. Comparisons of *CMAHP* gene expression levels in different subgroups based on an external TARGET database are shown in a box diagram. The X-axis indicates the different subgroups, whereas the Y-axis represents the gene expression level. *NUP98*-r, *NUP98* rearrangement; *TARGET*, Therapeutically Applicable Research to Generate Effective Treatments.

Supplementary Material 6: Figure S6. Heatmap showing the expression of all *HOX* cluster genes discovered in the different subgroups.

Supplementary Material 7: Figure S7. KEGG enrichment analysis of upregulated and downregulated pathways. *KEGG*, Kyoto Encyclopedia of Genes and Genomes.

Supplementary Material 8. Table S1. *NUP98*-r testing primers. *NUP98-r*, *NUP98* rearrangement.

Supplementary Material 9. Table S2. The 34-gene targeted sequencing panel.

Supplementary Material 10. Table S3. Patient characteristics in pediatric AML. *AML*, acute myeloid leukemia; *WBC*, white blood cell; *Hb*, hemoglobin; *PLT*, platelet; *BM*, bone marrow; *NUP98-r*, *NUP98* rearrangement; *KMT2A-r*, *KMT2A* rearrangement.

Supplementary Material 11. Table S4. Mutation sites detected in pediatric AML. *AML*, acute myeloid leukemia; *SNV*, single nucleotide variation.

Supplementary Material 12. Table S5. The details of fusion genes for RNA-Seq. *RNA-Seq*, RNA sequencing; *NUP98-r*, *NUP98* rearrangement.

Acknowledgements

We thank all the participants who kindly agreed to provide the data for this study. We also thank Springer Nature Author Services for editing the English text of the draft of this manuscript.

Author contributions

XX designed this study. JZ, CC, DS, LL, HS, TX, WX, YW, FZ, MF, HS, CL, AD, and YT collected data. JQ, LL, and SC performed NGS. JZ, JQ, and LL performed statistical analyses. JZ, CC and JQ drafted the manuscript. JZ, JQ and XX reviewed and revised the manuscript. All the authors have read and agreed to the published version of the manuscript.

Funding

This work was supported by the Pediatric Leukemia Diagnosis and Therapeutic Technology Research Center of Zhejiang Province (No. JBZX-201904), the Zhejiang University Special Funds Project of Fundamental Research Funds (No. 226-2022-00021) and the Natural Science Foundation of Zhejiang Province, China (LY19H190005).

Availability of data and materials
The datasets generated and/or analyzed during the current study are available in the CNGB Sequence Archive (CNSA) of the China National GeneBank DataBase (CNGBdb), CNP0004231. The remaining datasets analyzed in the current study are publicly available from the TARGET database (https://ocg.cancer.gov/programs/target), phs000218, and the data used for this analysis are available at https://portal.gdc.cancer.gov/projects.

Declarations

Ethics approval and consent to participate
This study was conducted in accordance with the principles of the Declaration of Helsinki. This study was approved by the Ethics Committee of the Children's Hospital, Zhejiang University School of Medicine. Informed consent was obtained from all the patients and/or their legal guardians.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Division/Center of Hematology-Oncology, Children's Hospital, Zhejiang University School of Medicine, Hangzhou 310005, China. [2]The Pediatric Leukemia Diagnostic and Therapeutic Technology Research Center of Zhejiang Province, Hangzhou 310005, China. [3]National Clinical Research Center for Child Health, Hangzhou 310005, China. [4]Department of Medical Affairs, Acornmed Biotechnology Co., Ltd., Beijing 100176, China. [5]Department of Clinical Laboratory, Children's Hospital, Zhejiang University School of Medicine, Hangzhou 310005, China.

Received: 13 March 2024 Accepted: 23 August 2024
Published online: 02 September 2024

References

1. Bolouri H, Farrar JE, Triche T Jr, et al. The molecular landscape of pediatric acute myeloid leukemia reveals recurrent structural alterations and age-specific mutational interactions. Nat Med. 2018;24(1):103–12.
2. Struski S, Lagarde S, Bories P, et al. NUP98 is rearranged in 3.8% of pediatric AML forming a clinical and molecular homogenous group with a poor prognosis. Leukemia. 2017;31(3):565–72.
3. Michmerhuizen NL, Klco JM, Mullighan CG. Mechanistic insights and potential therapeutic approaches for NUP98-rearranged hematologic malignancies. Blood. 2020;136(20):2275–89.
4. Iacobucci I, Wen J, Meggendorfer M, et al. Genomic subtyping and therapeutic targeting of acute erythroleukemia. Nat Genet. 2019;51(4):694–704.
5. Hollink IH, van den Heuvel-Eibrink MM, Arentsen-Peters ST, et al. NUP98/NSD1 characterizes a novel poor prognostic group in acute myeloid leukemia with a distinct HOX gene expression pattern. Blood. 2011;118(13):3645–56.
6. Noort S, Wander P, Alonzo TA, et al. The clinical and biological characteristics of NUP98-KDM5A in pediatric acute myeloid leukemia. Haematologica. 2021;106(2):630–4.
7. Miyamura T, Moritake H, Nakayama H, et al. Clinical and biological features of paediatric acute myeloid leukaemia (AML) with primary induction failure in the Japanese paediatric leukaemia/lymphoma study group AML-05 study. Br J Haematol. 2019;185(2):284–8.
8. Laurell E, Beck K, Krupina K, et al. Phosphorylation of Nup98 by multiple kinases is crucial for NPC disassembly during mitotic entry. Cell. 2011;144(4):539–50.
9. Xu H, Valerio DG, Eisold ME, et al. NUP98 fusion proteins interact with the NSL and MLL1 complexes to drive leukemogenesis. Cancer Cell. 2016;30(6):863–78.
10. Arber DA, Orazi A, Hasserjian R, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. Blood. 2016;127(20):2391–2405.
11. ISCN 2020. An international system for human cytogenomic nomenclature (2020). S. Karger AG; 2020.
12. DePristo MA, Banks E, Poplin R, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet. 2011;43(5):491–8.
13. Uhrig S, Ellermann J, Walther T, et al. Accurate and efficient detection of gene fusions from RNA sequencing data. Genome Res. 2021;31(3):448–60.
14. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics. 2015;31(2):166–9.
15. Frankish A, Diekhans M, Jungreis I, et al. Gencode 2021. Nucleic Acids Res. 2021;49(D1):D916–23.
16. Kanehisa M, Furumichi M, Sato Y, Ishiguro-Watanabe M, Tanabe M. KEGG: integrating viruses and cellular organisms. Nucleic Acids Res. 2021;49(D1):D545–51.
17. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics. 2016;32(18):2847–9.
18. Nakamura T, Largaespada DA, Lee MP, et al. Fusion of the nucleoporin gene NUP98 to HOXA9 by the chromosome translocation t(7;11)(p15;p15) in human myeloid leukaemia. Nat Genet. 1996;12(2):154–8.
19. Borrow J, Shearman AM, Stanton VP Jr, et al. The t(7;11)(p15;p15) translocation in acute myeloid leukaemia fuses the genes for nucleoporin NUP98 and class I homeoprotein HOXA9. Nat Genet. 1996;12(2):159–67.
20. Qiao Q, Li Y, Chen Z, Wang M, Reinberg D, Xu RM. The structure of NSD1 reveals an autoregulatory mechanism underlying histone H3K36 methylation. J Biol Chem. 2011;286(10):8361–8.
21. van Zutven LJ, Onen E, Velthuizen SC, et al. Identification of NUP98 abnormalities in acute leukemia: JARID1A (12p13) as a new partner gene. Genes Chromosomes Cancer. 2006;45(5):437–46.
22. de Rooij JD, Masetti R, van den Heuvel-Eibrink MM, et al. Recurrent abnormalities can be used for risk group stratification in pediatric AMKL: a retrospective intergroup study. Blood. 2016;127(26):3424–30.
23. Zhang J, Shen H, Song H, et al. A novel NUP98/RARG gene fusion in pediatric acute myeloid leukemia resembling acute promyelocytic leukemia. J Pediatr Hematol Oncol. 2022;44(3):e665–71.
24. Bertrums EJM, Smith JL, Harmon L, et al. Comprehensive molecular and clinical characterization of NUP98 fusions in pediatric acute myeloid leukemia. Haematologica. 2023;108(8):2044–58.
25. Wei W, Liu Q, Song F, et al. Alkaloid-based regimen is beneficial for acute myeloid leukemia resembling acute promyelocytic leukemia with NUP98/RARG fusion and RUNX1 mutation: a case report. Medicine. 2020;99(40): e22488.
26. Shiba N, Ichikawa H, Taki T, et al. NUP98-NSD1 gene fusion and its related gene expression signature are strongly associated with a poor prognosis in pediatric acute myeloid leukemia. Genes Chromosomes Cancer. 2013;52(7):683–93.
27. Fasan A, Haferlach C, Alpermann T, Kern W, Haferlach T, Schnittger S. A rare but specific subset of adult AML patients can be defined by the cytogenetically cryptic NUP98-NSD1 fusion gene. Leukemia. 2013;27(1):245–8.
28. Ostronoff F, Othus M, Gerbing RB, et al. NUP98/NSD1 and FLT3/ITD coexpression is more prevalent in younger AML patients and leads to induction failure: a COG and SWOG report. Blood. 2014;124(15):2400–7.
29. Wang GG, Cai L, Pasillas MP, Kamps MP. NUP98-NSD1 links H3K36 methylation to Hox-A gene activation and leukaemogenesis. Nat Cell Biol. 2007;9(7):804–12.
30. Huang HW, Chen CY, Huang YH, et al. CMAHP promotes metastasis by reducing ubiquitination of Snail and inducing angiogenesis via GM-CSF overexpression in gastric cancer. Oncogene. 2022;41(2):159–72.
31. Hu X, Yang L, Mo YY. Role of pseudogenes in tumorigenesis. Cancers. 2018;10(8):256.
32. Jaatinen T, Hemmoranta H, Hautaniemi S, et al. Global gene expression profile of human cord blood-derived CD133+ cells. Stem Cells. 2006;24(3):631–41.
33. Tuborgh A, Meyer C, Marschalek R, Preiss B, Hasle H, Kjeldsen E. Complex three-way translocation involving MLL, ELL, RREB1, and CMAHP genes in an infant with acute myeloid leukemia and t(6;19;11)(p22.2;p13.1;q23.3). Cytogenet Genome Res. 2013;141(1):7–15.
34. Robertson G, Schein J, Chiu R, et al. De novo assembly and analysis of RNA-seq data. Nat Methods. 2010;7(11):909–12.

35. Shiba N, Ohki K, Kobayashi T, et al. High PRDM16 expression identifies a prognostic subgroup of pediatric acute myeloid leukaemia correlated to FLT3-ITD, KMT2A-PTD, and NUP98-NSD1: the results of the Japanese paediatric leukaemia/lymphoma study group AML-05 trial. Br J Haematol. 2016;172(4):581–91.
36. Heikamp EB, Henrich JA, Perner F, et al. The menin-MLL1 interaction is a molecular dependency in NUP98-rearranged AML. Blood. 2022;139(6):894–906.
37. Lewuillon C, Laguillaumie MO, Quesnel B, Idziorek T, Touil Y, Lemonnier L. Put in a "Ca(2+)ll" to acute myeloid leukemia. Cells. 2022;11(3):543.

Publisher's Note