blood

American Society of Hematology
2021 L Street NW, Suite 900,
Washington, DC 20036
Phone: 202-776-0544 | Fax 202-776-0545
editorial@hematology.org

Fusion Oncoproteins and Cooperating Mutations Define Disease Phenotypes in NUP98-Rearranged Leukemia

Tracking no: BLD-2025-028993R1

Masayuki Umeda (St. Jude Children's Research Hospital, United States) Ryan Hiltenbrand (St. Jude Children's Research Hospital, United States) Nicole Michmerhuizen (St Jude Children's Research Hospital, United States) Juan Barajas (St. Jude Children's Research Hospital, United States) Melvin Thomas (St. Jude Children's Research Hospital, United States) Bright Arthur (St Jude children's hospital, United States) Michael Walsh (St. Jude Children's Research Hospital, United States) Guangchun Song (St Jude Children's Research Hospital, United States) Jing Ma (St Jude Children's Research Hospital, United States) Tamara Westover (St Jude Children's Research Hospital, United States) Amit Kumar (St. Jude Children's Research Hospital, United States) Petri Pölönen (St Jude Children's Research Hospital, United States) Cristina Mecucci (University of Perugia, Italy) Danika Di Giacomo (University of Perugia, Italy) Franco Locatelli (Ospedale Bambino Gesù, Italy) Riccardo Masetti (IRCCS Azienda Ospedaliero Universitaria di Bologna, Italy) Salvatore Nicola Bertuccio (IRCCS Azienda Ospedaliero-Universitaria di Bologna, Italy) Martina Pigazzi (Pediatric Hematology, Oncology and Stem Cell Transplant Division, University of Padova, Italy) Shondra Pruett-Miller (St. Jude Children's Research Hospital, United States) Stanley Pounds (St. Jude Children's Research Hospital, United States) Jeffrey Rubnitz (St. Jude Children's Research Hospital, United States) Hiroto Inaba (St. Jude Children's Research Hospital, United States) Kyriakos Papadopoulos (XenoSTART/ The START Center for Cancer Research, United States) Michael Wick (The START Center for Cancer Research, United States) Ilaria Iacobucci (St. Jude Children's Research Hospital, United States) Charles Mullighan (St Jude Children's Research Hospital, United States) Jeffery Klco (St Jude Children's Research Hospital, United States)

Abstract:
Leukemias with NUP98 rearrangements exhibit heterogeneous phenotypes such as acute myeloid leukemia (AML), T-acute lymphoblastic leukemia (T-ALL), or myelodysplastic syndrome/neoplasms (MDS) associated with fusion partners, whereas the mechanism responsible for this heterogeneity is poorly understood. Through genome-wide mutational and transcriptional analyses of 177 NUP98-rearranged leukemias, we show that cooperating alterations are associated with differentiation status even among leukemias sharing the same NUP98 fusions, such as NUP98::KDM5A acute megakaryocytic leukemia (AMKL) with RB1 loss or T-ALL with NOTCH1 mutations. CUT&RUN profiling of in vitro cord blood CD34+ cell (cbCD34) models of major NUP98 fusions revealed that NUP98 fusion oncoproteins directly regulate differentiation-related genes contributing to the disease phenotypes, represented by NUP98::KDM5A binding to MEIS2 or GFI1B for megakaryocyte differentiation. In patient samples, NUP98-fusion oncoprotein binding patterns are heterogeneous, potentially shaped by somatic mutations and differentiation status. Using cbCD34 models and CRISPR/Cas9 gene editing, we show that RB1 loss cooperates with NUP98::KDM5A by blocking terminal differentiation toward platelets and expanding megakaryocyte-like cells, whereas WT1 frameshift mutations skew differentiation toward dormant lymphoid-myeloid primed progenitor cells and cycling granulocyte-monocyte progenitor cells, providing evidence for NUP98-rearranged leukemia phenotypes affected by cooperating alterations. NUP98::KDM5A cbCD34 models with RB1 or WT1 alterations have different sensitivities to menin inhibition, suggesting that cellular differentiation provides stage-specific menin dependencies and resistance mechanisms that can be leveraged for future treatment strategies for NUP98-rearranged leukemia.

Conflict of interest: COI declared - see note

COI notes: Kyriakos P. Papadopoulos and Michael J. Wick are employees of XenoSTART/ The START Center for Cancer Research, which provided ST1653 PDX cells for this manuscript.

Preprint server: Yes; medRxiv https://doi.org/10.1101/2025.01.21.25320683

Author contributions and disclosures: J.M.K and C.G.M conceptualized the study. J.M.K, M.P.W, G.S, J.M, T.W, P.P, I.I, and C.G.M collected data. J.M.K, M.U M.P.W, G.S, J.M, and T.W curated data. M.U and J.M conducted formal analyses. M.U, R.H, N.L.M, J.B, T.M. M.E.T, B.A, and A.K performed experiments. M.U and J.M performed statistical tests. MU made figures. P.P, C.M, D.D.G, F.L, R.M, S.N.B, M.P, S.M.P-M, S.P, J.R, H.I, K.P.P, M.J.W, I.I, and C.G.M provided resources. M.U and J.M.K wrote the original draft, and all the co-authors reviewed it and were involved in the edition.

Non-author contributions and disclosures: No;

Agreement to Share Publication-Related Data and Data Sharing Statement: Data availability Genomic analyses in this study were based on the GENCODE GRCh38/hg38, and gnomAD (v2.1.1, RRID: SCR_014964) was used for classification for germline and somatic mutations. The sequencing data from patient samples newly generated in this study (RNAseq: n=10, WGS: n=2, WES: n=2, scRNAseq: n=9, CUT&RUN: n=64) and undeposited data from previous studies (RNAseq: n=3) have been deposited in the European Genome-Phenome Archive (EGA, RRID:SCR_004944), which is hosted by the European Bioinformatics Institute (EBI), under accession EGAS00001005760. For the remaining RNAseq data for 175 samples, 56 samples are sequenced in St. Jude and available either on EGA or St. Jude Cloud(1-5) or on the original publication.(6) For 9 published cases(7), we downloaded the BAM files from EGA (EGAS00001004701). Of the remaining WGS data for 73 samples, 34 are sequenced at St. Jude samples, data for all 34 samples from the original publications(1, 2, 4, 5) are available on either EGA or St. Jude Cloud, and for other 7 published cases(7), we downloaded the BAM files from EGA (EGAS00001004701). For the remaining WES data for 46 samples, 21 are sequenced at St. Jude and all data from the original publications(1-5, 8) are available either on St. Jude Cloud or EGA, and for the other 9 published cases(7), we downloaded the BAM files from EGA (EGAS00001004701). The data generated by the TARGET-AML initiative(9, 10) (RNA:n=18, WGS:n=18, WES:n=2) is available under accession phs000218 (TARGET-AML) and phs000465 (TARGET sub-study, data is available as a part of phs000218), managed by the NCI. Data from the AAML1031 study led by COG is also available at the GDC data portal under phs000218 (TARGET-AML). Public data from the COG AALL0434 trial (RNA:n=14, WGS:n=14, WES:n=14) is available on dbGaP under phs002276.v2.p1 (phs000218 and phs000464 for T-ALL TARGET samples) and the Kids First data portal (https://portal.kidsfirstdrc.org/dashboard). Information about TARGET can be found at http://ocg.cancer.gov/programs/target. Background RNAseq data of other T-ALL(11) was also obtained from SYNAPSE under project ID of syn54032669, and data of other AML subtypes were from publication(5). Output files from Cell Ranger for scRNAseq data of 9 samples from 8 NUP98r patients and 4 samples from cbCD34 models in this study were deposited in GEO under accession number GSE287716. Bulk RNAseq data (log2CPM count data) from cbCD34 models was deposited in GEO under accession number GSE287297. Processed CUT&RUN data (bigwig files and binding peaks) from primary patient samples and CD34 models were deposited in GEO under accession number GSE287298. Public patient scRNAseq data from the COG43 (GSM7494309, 7494310) is available at GEO under accession number GSE235063, and the patient data was obtained from supplemental tables from the publication. Other data generated in this study are available in Supplemental tables or Source data files or upon request to the corresponding author. We did not use any custom code in this study. 1.Schwartz JR, Ma J, Lamprecht T, et al. The genomic landscape of pediatric myelodysplastic syndromes. Nat Commun. 2017;8(1):1557. 2.Umeda M, Ma J, Huang BJ, et al. Integrated Genomic Analysis Identifies UBTF Tandem Duplications as a Recurrent Lesion in Pediatric Acute Myeloid Leukemia. Blood Cancer Discov. 2022;3(3):194-207. 3.de Rooij JD, Branstetter C, Ma J, et al. Pediatric non-Down syndrome acute megakaryoblastic leukemia is characterized by distinct genomic subsets with varying outcomes. Nat Genet. 2017;49(3):451-456. 4.Schwartz JR, Ma J, Kamens J, et al. The acquisition of molecular drivers in pediatric therapy-related myeloid neoplasms. Nat Commun. 2021;12(1):985. 5.Umeda M, Ma J, Westover T, et al. A new genomic framework to categorize pediatric acute myeloid leukemia. Nat Genet. 2024. 6.Buelow DR, Pounds SB, Wang YD, et al. Uncovering the Genomic Landscape in Newly Diagnosed and Relapsed Pediatric Cytogenetically Normal FLT3-ITD AML. Clin Transl Sci. 2019;12(6):641-647. 7.Fornerod M, Ma J, Noort S, et al. Integrative Genomic Analysis of Pediatric Myeloid-Related Acute Leukemias Identifies Novel Subtypes and Prognostic Indicators. Blood Cancer Discov. 2021;2(6):586-599. 8.Iacobucci I, Wen J, Meggendorfer M, et al. Genomic subtyping and therapeutic targeting of acute erythroleukemia. Nat Genet. 2019;51(4):694-704. 9.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):103-112. 10.McNeer NA, Philip J, Geiger H, et al. Genetic mechanisms of primary chemotherapy resistance in pediatric acute myeloid leukemia. Leukemia. 2019;33(8):1934-1943. 11.Polonen P, Di Giacomo D, Seffernick AE, et al. The genomic basis of childhood T-lineage acute lymphoblastic leukaemia. Nature. 2024;632(8027):1082-1091.

Clinical trial registration information (if any):

Fusion Oncoproteins and Cooperating Mutations Define Disease Phenotypes in *NUP98*-Rearranged Leukemia

Masayuki Umeda[1], Ryan Hiltenbrand[1], Nicole L. Michmerhuizen[1], Juan M. Barajas[1], Melvin E. Thomas[1], Bright Arthur[1], Michael P Walsh[1], Guangchun Song[1], Jing Ma[1], Tamara Westover[1], Amit Kumar[2], Petri Pölönen[1], Cristina Mecucci[3], Danika Di Giacomo[1,4], Franco Locatelli[5,6], Riccardo Masetti[7,8], Salvatore N. Bertuccio[9], Martina Pigazzi[10], Shondra M. Pruett-Miller[11,12], Stanley Pounds[13], Jeffrey Rubnitz[14], Hiroto Inaba[14], Kyriakos P. Papadopoulos[15], Michael J. Wick[15], Ilaria Iacobucci[1], Charles G. Mullighan[1,2], Jeffery M. Klco[1,2,*]

1 Department of Pathology, St. Jude Children's Research Hospital, Memphis, TN, US
2 Center of Excellence for Leukemia Studies (CELS), St. Jude Children's Research Hospital, Memphis, TN, US
3 Department of Medicine and Surgery, Hematology and Bone Marrow Transplantation Unit, University of Perugia, Perugia, Italy
4 Department of Chemistry, Biology and Biotechnology, University of Perugia, Perugia, Italy
5 Department of Onco-Haematology and Cell and Gene Therapy, Bambino Gesù Children's Hospital, IRCCS, Rome, Italy
6 Department of Life Sciences and Public Health, Catholic University of the Sacred Heart, Rome, Italy
7 Pediatric Oncology and Hematology, IRCCS Azienda Ospedaliero-Universitaria di Bologna, Bologna, Italy
8 Department of Medical and Surgical Sciences (DIMEC), University of Bologna, Bologna, Italy
9 IRCCS Azienda Ospedaliero-Universitaria di Bologna, Bologna, Italy
10 Department of Women's and Children's Health, Onco-hematology lab and clinic, University of Padova, Padova, Italy
11 Center for Advanced Genome Engineering, St. Jude Children's Research Hospital, Memphis, TN, US
12 Department of Cell and Molecular Biology, St. Jude Children's Research Hospital, Memphis, TN, US
13 Department of Biostatistics, St. Jude Children's Research Hospital, Memphis, TN, US
14 Department of Oncology, St. Jude Children's Research Hospital, Memphis, TN, US
15 XenoSTART/ The START Center for Cancer Research, San Antonio, TX, US

*Corresponding Author
Jeffery M. Klco, MD, PhD
262 Danny Thomas Place, MS 342, D4047B, Memphis, TN 38105
Jeffery.Klco@STJUDE.ORG
Phone: 901-595-6807
Fax: 901-595-5947

Running title : Molecular basis of *NUP98*-rearranged leukemia phenotypes

Scientific category : Oncology, Genomics

Text word counts : 3997/4000

Abstract word counts : 223/250

Figures : 7/7

References : 71/100

ABSTRACT

Leukemias with *NUP98* rearrangements exhibit heterogeneous phenotypes such as acute myeloid leukemia (AML), T-acute lymphoblastic leukemia (T-ALL), or myelodysplastic syndrome/neoplasms (MDS) associated with fusion partners, whereas the mechanism responsible for this heterogeneity is poorly understood. Through genome-wide mutational and transcriptional analyses of 177 *NUP98*-rearranged leukemias, we show that cooperating alterations are associated with differentiation status even among leukemias sharing the same *NUP98* fusions, such as *NUP98::KDM5A* acute megakaryocytic leukemia (AMKL) with *RB1* loss or T-ALL with *NOTCH1* mutations. CUT&RUN profiling of *in vitro* cord blood CD34+ cell (cbCD34) models of major NUP98 fusions revealed that NUP98 fusion oncoproteins directly regulate differentiation-related genes contributing to the disease phenotypes, represented by NUP98::KDM5A binding to *MEIS2* or *GFI1B* for megakaryocyte differentiation. In patient samples, NUP98-fusion oncoprotein binding patterns are heterogeneous, potentially shaped by somatic mutations and differentiation status. Using cbCD34 models and CRISPR/Cas9 gene editing, we show that *RB1* loss cooperates with NUP98::KDM5A by blocking terminal differentiation toward platelets and expanding megakaryocyte-like cells, whereas *WT1* frameshift mutations skew differentiation toward dormant lymphoid-myeloid primed progenitor cells and cycling granulocyte-monocyte progenitor cells, providing evidence for *NUP98*-rearranged leukemia phenotypes affected by cooperating alterations. NUP98::KDM5A cbCD34 models with *RB1* or *WT1* alterations have different sensitivities to menin inhibition, suggesting that cellular differentiation provides stage-specific menin dependencies and resistance mechanisms that can be leveraged for future treatment strategies for *NUP98*-rearranged leukemia.

KEY POINTS:

Genome-wide profiling reveals somatic alterations in *NUP98*-rearranged leukemia associated with fusion partners and phenotypes.

The dependency of *NUP98*-rearranged leukemia on menin can be affected by differentiation state and cooperating mutations.

DATA SHARING STATEMENT:

The sequencing data from patient samples newly generated in this study (RNAseq: n=10, WGS: n=2, WES: n=2, scRNAseq: n=9, CUT&RUN: n=64) and undeposited data from previous studies (RNAseq: n=3) have been deposited in the European Genome-Phenome Archive (EGA), which is hosted by the European Bioinformatics Institute (EBI), under accession EGAS00001005760.

Output files from Cell Ranger for scRNAseq data of 9 samples from 8 *NUP98*r patients and 4 samples from cbCD34 models in this study were deposited in GEO under accession number GSE287716. Bulk RNAseq data (log2CPM count data) from cbCD34 models were deposited in GEO under accession number GSE287297. Processed CUT&RUN data (bigwig files and binding peaks) from primary patient samples and CD34 models were deposited in GEO under accession number GSE287298.

Patient samples were de-identified and replaced with unique research IDs, which are maintained using an honest broker, in accordance with institutional protocols and ethical guidelines. This unique research ID cannot be linked to patient identification outside of the research team.

Other data generated in this study are available in Supplemental tables or upon request to the corresponding author. We did not use any custom code in this study.

INTRODUCTION

The prognosis of children with acute myeloid leukemia (AML) remains unsatisfactory with high relapse rates[1-3]. This is partly attributed to subtypes associated with poor treatment response, such as *KMT2A* rearrangements (*KMT2A*r)[4,5], *UBTF* tandem duplications (*UBTF*-TD)[6], *CBFA2T3*::*GLIS2*[7,8], and *NUP98* rearrangements (*NUP98*r).[5,9,10] Understanding the genetic background and disease mechanisms of these subtypes is needed to develop efficient treatment strategies based on biology and molecular mechanisms.

*NUP98*r with various fusion partners are known to be associated with specific hematologic malignancies.[11] *NUP98*r accounts for roughly 7% of pediatric AML, with *NUP98::NSD1* predominantly found in adolescents with myelomonocytic AML, whereas *NUP98::KDM5A* is often found in acute megakaryocytic leukemia (AMKL) in young children.[5,8,9] Other *NUP98*r fusion partners have also been reported in T-cell acute lymphoblastic leukemia (T-ALL)[12], myelodysplastic neoplasms (MDS)[13], or mixed phenotype acute leukemia (MPAL)[14]. Recent genomic studies focusing on *NUP98*r AML confirmed known associations of *NUP98::NSD1* with *FLT3* internal tandem duplications (ITD) and *WT1* mutations or *NUP98::KDM5A* with chromosome 13 loss involving *RB1*.[8,10,15] However, genome-wide studies on various *NUP98*r leukemias are needed to better understand the association of genomic features with disease phenotypes.

Mechanistically, recent studies revealed that NUP98 fusion oncoproteins (FOs) bind to gene loci active in *NUP98*r AML, such as *MEIS1*, *HOXA/B* clusters, and *CDK6.*[16-18] This binding is partly driven by intrinsically disordered regions of the N-terminal NUP98, which mediate transcriptional condensate formation[19,20], and DNA/chromatin binding domains of the C-terminal partners. This knowledge has mostly been derived from mouse or non-hematopoietic human models due to a lack of human *NUP98*r leukemia models. An understanding of molecular mechanisms, including NUP98-FO-binding profiles, in human hematopoietic cells is necessary for developing mechanism-based therapeutics for *NUP98*r leukemias.

To reveal phenotype-specific mutational profiles, we obtained and reanalyzed 175 publicly available AML and T-ALL datasets along with genome-wide characterization of 10 new pediatric and young adult *NUP98*r cases using RNA sequencing (RNAseq) and whole genome/exome sequencing (WGS/WES). These data were corroborated by single cell RNAseq (scRNAseq) and epigenetic profiling by CUT&RUN[21] of samples with major *NUP98* fusions, revealing NUP98-FO binding profiles associated with fusion partners and disease phenotypes. Functional characterization of recurrent gene alterations using cord blood CD34 models showed the NUP98 fusion- and phenotype-dependent impact on cell growth and transcriptional profiles associated with drug responses, which can be a basis for overcoming these refractory hematopoietic neoplasms.

METHODS

Experimental details are found in Supplementary Methods.

Patient cohort: Tumor samples from patients with acute leukemia were obtained with written informed consent using a protocol approved by the institutional review board (IRB) in each institute. Samples for RNAseq (n=10), WGS (n=2), and WES (n=2) were newly sequenced in this study. The rest of the data were obtained from previous publications (Tables S1-4)[3,6,8,10,13,22-30] or public databases (see details in Supplementary Methods), establishing a cohort with 185 samples from 177 patients, including 8 patients with data at multiple time points (Table S1). RNAseq data were available for all samples (n=185) and DNA data for 91 samples (WGS:43, WES:16, both:32, Tables S2-9). Fusion genes were identified by CICERO[31] and manual inspection of sequencing data, confirming *NUP98* fusion transcripts in all cases.

Cord blood CD34+ modeling: Cord blood CD34+ cells (cbCD34) were purchased from Lonza (2C-101) or the Carolinas Cord Blood Bank/Duke University. Cells were cultured in StemSpan SFEM II media (STEMCELL Technologies) supplemented with penicillin–streptomycin, l-glutamine, and recombinant human SCF, FLT-3 ligand,

TPO, and IL6 (all 50 ng/mL, PeproTech), SR-1 (1 μmol/L), and UM171 (35 nmol/L, for Fig.4) or UM729 (10 nmol/L, for Fig.6-7, all STEMCELL Technologies). Cells were transduced with lentivirus expressing NUP98::KDM5A, NUP98::NSD1, NUP98::HOXA9, or empty vector controls, and transduced mCherry+ cells were enriched by flow sorting on day 7 of transduction. For the introduction of somatic alterations, cbCD34 models were transduced with lentivirus expressing Cas9 to establish cell lines with constitutive Cas9 expression. Transduction of gRNA-expressing vectors (for *WT1*, *RB1*, and the AAVS locus) was performed around day 40-60 from the transduction of NUP98-FOs.

RESULTS

Transcriptional heterogeneity of pediatric NUP98r leukemia

In our pediatric and young adult *NUP98*r leukemia cohort (median age: 8.6, range: 0.4-28), AML was the dominant disease (n=162), followed by T-ALL (n=18), therapy-related myeloid neoplasms (t-MN, n=4), and MDS (n=1), reflecting heterogeneous *NUP98*r hematologic neoplasms (Fig.1A). Among fusion partners, *NSD1* was the most frequent (n=92, 51.7%), followed by *KDM5A* (n=45, 25.3%, Fig.1B, Fig.S1B, Table S8). *RAP1GDS1* (n=11, 6.2%) was the third most common partner, with nine *NUP98::RAP1GDS1* found in T-ALL cases. *NUP98::KDM5A* is enriched in infants, whereas *NUP98::NSD1* and other partners were broadly found in adolescents and young adults (Fig.1C), consistent with its known occurrence in adult AML.[9,15] Most of the other partners fall into the functional categories of epigenetic regulators (n=12, 6.7%) or homeobox proteins (n=9, 5.1%). Partners with epigenetic regulatory functions, including NSD1 and KDM5A, frequently have plant homeodomain (PHD) fingers in the retained C-terminus, suggesting structural redundancy (Fig.S1B-C). Functions of the remaining 6 partners were broad, ranging from RNA-binding (DDX1/10), maintenance of DNA structure (HMGB3 and CEP295), or without known function (LNP1 and CCDC28A), but all had domains with multiple α-helices with linker sequences resembling homeobox domains or PHD.

Appreciating the known heterogeneity of the disease phenotypes of *NUP98*r leukemia, we performed transcriptional analyses with other subtypes of AML[5] (n=816) and T-ALL[30] (n=1,320) as a background cohort (Fig.S2A, Table S10). UMAP (Uniform Manifold Approximation and Projection) and unsupervised clustering revealed 10 clusters driven by driver alterations and differentiation status. *NUP98*r malignancies were distributed to 6 of these clusters (Fig.1D, Fig.S2B-D). Most *NUP98::NSD1* samples were clustered with AML with *NPM1* mutations or *UBTF*-TD showing high *HOXA/B* gene expression (the *HOXB* cluster, n=114). By contrast, samples with *NUP98::KDM5A* were predominantly clustered with other AEL/AMKL subtypes (the AEL/AMKL cluster, n=37), although rare cases clustered with subtypes with *KMT2A*-rearranged and *KAT6A*-rearranged AML with *HOXA* expression (the *HOXA* cluster, n=10) or with mature ALL (the *TLX* cluster, n=3). *NUP98*::*RAP1GDS1* samples were mainly clustered with other early T-cell precursor-like ALL (the ETP-like cluster, n=16), whereas they were also found in the *HOXB*, *TLX*, or immature AML clusters (n=5). Although *HOXA/B* gene expression is recognized as a hallmark of *NUP98*r leukemia, we observed that *NUP98::NSD1* samples lack posterior *HOXB* (*HOXB6-9*) expression and immature/T-ALL cases lack *HOXB* or both *HOXA/B* expression (Fig.S2E). These data show that *NUP98*r leukemias are heterogeneous, and even the same fusion partners can contribute to variable phenotypes.

Mutational background of NUP98r leukemia associated with disease phenotypes

To assess genetic factors contributing to the transcriptional heterogeneity of *NUP98*r, we investigated the genomic landscape using DNA data and an RNAseq-based pipeline[5,6] (Fig.1A), including single nucleotide variants (SNV), insertions and deletions (indel), ITD, copy number variations (CNV), and fusion genes other than *NUP98* rearrangements (Tables S5-9). Among 53 recurrently altered genes, the most frequent alteration was *FLT3*-ITD (n=70, 37.8%), followed by *WT1* (n=44, 23.8%) and *NRAS* mutations (n=30, 16.2%) (Fig.2A-C). Chromosome 13 deletions involving *RB1*[8,10] were found predominantly in *NUP98::KDM5A* cases (n=23, 14.1%), with deleted regions extending in both 3' and 5' directions of *RB1*, often affecting the tumor suppressor *BRCA2* (Fig.S3A). We also

identified *RB1* loss in cases with only RNAseq by assessing for allelic imbalance of single nucleotide polymorphisms (SNPs) within and around the *RB1* locus; however, *RB1* loss could still be underestimated as AMKL cases without detectable CNV also showed low *RB1* RNA expression (Fig.S3B-C). *RB1* loss was associated with high expression of stem cell markers (*GATA2, CD96,* and *BMI1*) among *NUP98::KDM5A* AMKL samples, and gene set enrichment analysis (GSEA) showed enrichment of genes involving AMKL phenotypes (Fig.S3D-F). GSEA also showed changes in genes related to cell growth and mitochondrial function, consistent with the known function of RB1/E2F[32,33] (Fig.S3F). Additionally, genome-wide profiling revealed *JAK2* (n=6, 3.2%), *GATA1* (n=3, 1.6%), and *MPL* (n=3) mutations associated with the AEL/AMKL cluster (Fig.2C), all of which were frequently found in non-Down syndrome AMKL.[8]

FLT3-ITD and *WT1* were highly co-occurring and enriched in *NUP98::NSD1* AMLs (Fig.2A-C). Variant allele frequencies (VAFs) of *WT1* and *FLT3*-ITD were generally high both at diagnosis and relapse (Fig.S4A-C). *FLT3*-ITD and *RAS*-related mutations were less likely to co-occur, especially among *NUP98*::*NSD1* samples, and VAFs of *RAS* mutations were variable (Fig.2B, Fig.S4A-B), indicating *RAS* mutations existing in independent clones or subclonal to *FLT3*-ITD[34]. Differentially expressed gene (DEG) analysis showed high expression of immunity-related genes, such as *MRC1* (CD206) or *IL2RA* (CD25), in *FLT3*-ITD+ *NUP98*::*NSD1* AML, consistent with previous findings in *FLT3*-ITD+ AML, where these markers are associated with inferior outcomes.[35,36] GSEA confirmed known enrichment of inflammatory-related genes[37] (Fig.S4D-E). Samples with *WT1* mutations had significantly high expression of genes such as *DNTT* (TdT: Terminal deoxynucleotidyl transferase), a known marker for lymphoid-myeloid primed progenitors (LMPP)[38,39], and gene sets related to mitochondria and translation.

Clinically, all 12 *NUP98::RAP1GDS1* leukemias were either T-ALL or immature AML (FAB M0-1). Those with *NOTCH1* mutations (n=7, 58.3%) were predominantly found in the *TLX* and ETP-like clusters (6/7, 85.7%) and associated with T cell-related expression profiles, whereas *NRAS* mutations were mainly found in the ETP and immature AML clusters (Fig.S4F-H). Mutational profiling of samples in T-ALL/immature AML clusters showed *NOTCH1* mutations enriched in the ALL clusters, including those in *NUP98::KDM5A* ALL, whereas *IL7R* mutations were enriched in immature AML clusters (Fig.2C). Thus, among immature *NUP98*r leukemias, *NOTCH1* mutations likely contribute to T-ALL phenotypes[40]. These data collectively show the association among fusion partners, cooperating mutations, and disease phenotypes of *NUP98*r malignancies.

Varieties of cellular hierarchies in NUP98r leukemia associated with fusions and cooperating alterations

Recent research showed associations between differentiation status with outcomes or drug responses in AML.[41-43] To investigate the differentiation status of *NUP98*r leukemia, we performed scRNAseq on 9 samples with representative fusions (*NUP98*::*NSD1*, ::*KDM5A*, and ::*RAP1GDS1*) and obtained publicly available data from one case with diagnosis-relapse paired (D-R) samples of *NUP98::NSD1* AML (PAXLWH)[43] (Fig.3A, Fig.S5A, Table S11). UMAP showed distinct clusters for each sample, indicating unique expression profiles of each tumor, whereas T-cell or B-cell clusters consisted of cells from multiple patients, reflecting normal lymphocytes[43,44] (Fig.3B-C, Fig.S5B). To infer the differentiation stages corresponding to normal hematopoiesis, we projected scRNAseq data onto reference normal bone marrow and thymocyte scRNAseq data[44], which revealed unique distributions of tumor cells in each sample (Fig.3D-F, Fig.S5C). Three *NUP98::NSD1* samples (SJAML015363_D5, SJAML016582_D1, and SJAML061252_R1) showed various cellular hierarchies ranging from LMPP-like to monocyte-like cells (Fig.3E-F). SJAML001441 D-R pair (*FLT3*-ITD+/*WT1*+) had enriched monocyte and granulocyte-macrophage progenitor (GMP)-like cells at diagnosis, with increasing LMPP-like cells at relapse with an additional *WT1* mutation. PAXLWH D-R pair (*FLT3*-ITD+) showed homogenous LMPP-like cells at diagnosis but increased GMP or monocytic populations at relapse with an acquired *RUNX1* mutation, suggesting different patterns of relapse of these *NUP98::NSD1* cases. Leukemic cells from *NUP98::NSD1* patients showed aberrantly high *HOXA/B* gene expression along with low expression of marker genes (*e.g.*, *S100A12* in monocytes and *ELANE* in GMP) compared with normal counterparts (Fig.S5D), confirming aberrant expression profiles. Deconvolution of bulk RNA samples using CIBERSORT[45] revealed heterogeneous hierarchies among the entire cohort and variable patterns among both D-R pairs and somatic mutations (Fig.3G, Fig.S5E). Relapse samples commonly showed low monocyte and

high LMPP signatures, indicating expansion of stem-like populations at relapse in a subset of *NUP98::NSD1* AML, as reported in other subtypes[43].

A *NUP98::KDM5A* AMKL case with *RB1* loss contained a cluster corresponding to megakaryocyte/erythroid progenitors (MEP) in the normal reference (Fig.3B,E-F), showing aberrant co-expression of *HOXB* genes and immune response-related *IFI27* at the single cell level (Fig.S6A-B). *NOTCH1*$^{+}$ *NUP98::KDM5A* ALL cells were enriched at Pro- and Pre-T cell stages with aberrant expression of hemoglobin or *HOXB* genes (Fig.S6A).

NOTCH1^{+} *NUP98::RAP1GDS1* cells were distributed broadly at the LMPP-ETP-Pro-T stages, whereas *NUP98::RAP1GDS1* AML cells were enriched in the LMPP cluster (Fig.3E-F, Fig.S6C). Deconvoluted bulk *NUP98::RAP1GDS1* samples in immature clusters showed LMPP-CLP (common lymphoid progenitor) signatures, while those in the ALL clusters showed high Pro-Pre-T cell signatures, suggesting that the differentiation block of *NUP98::RAP1GDS1* occurs at various stages during T-cell development, likely affected by cooperating mutations[40] (Fig.S6C-E). Two *NUP98::KDM5A* ALL cases showed higher Pre-T signals than *NUP98::RAP1GDS1* cases (Fig.S6F), indicating that NUP98::KDM5A could contribute to more mature ALL than NUP98::RAP1GDS1, both cooperating with *NOTCH1* mutations.

Cord blood CD34 models recapitulate the phenotypes of NUP98r leukemia

With limited human *NUP98*r leukemia models that recapitulate the disease phenotypes to study the underlying mechanisms[40,46-49], we tested the transforming potential of various NUP98 FOs by transducing cord blood CD34+ cells (cbCD34) with lentiviral particles encoding major NUP98-FOs (pediatrics-NUP98::NSD1 and NUP98:KDM5A, adult-NUP98::HOXA9). Expression of these NUP98-FOs increased clonogenic potential in methylcellulose and enhanced cell growth in liquid culture compared with an empty vector control (Fig.4A-C), despite a gradual decrease in CD34 expression. Flow cytometry analyses showed low CD11b expression in all conditions compared with the control, whereas only NUP98::KDM5A demonstrated consistent CD41a expression, indicating preferential differentiation toward megakaryocytes (Fig.4D, Fig.S7A-B). RNAseq from liquid cultures at multiple time points revealed dynamic changes of transcriptional status unique to each NUP98-FO (Fig.4E-G, Fig.S7C-E). Consistent with patient samples, NUP98-FO-transduced conditions showed high *TAL1* and *MYCN* expression in addition to upregulation of *HOXA*/*B* genes, whereas control conditions showed reduced *TAL1* and increased *ITGAM* (CD11b) expression, suggesting myeloid differentiation (Fig.4F). We identified 320 common genes with high expression in conditions with NUP98-FOs, which enriched homeobox genes (Fig.4G). Notably, we also identified uniquely high genes with each FO related to specific functions, such as *GFI1B or GATA1* related to erythromegakaryocytic differentiation in cells with NUP98::KDM5A. Comparison with the patient cohort showed clustering of cbCD34 models with patient samples with similar fusions/phenotypes (Fig.S7E). Two cell lines with *NUP98::KDM5A*, CHRF-288-11[46] (megakaryocytic) and ST1653[50] (monocytic patient-derived xenograft: PDX), also showed unique patterns of expression (*e.g., ITGA2B* and *CD34* in CHRF-288-11, and *ITGAM* in ST1653), confirming the heterogeneity of leukemia with *NUP98::KDM5A* (Fig.4E-F, Fig.S7E).

Differential gene regulation by NUP98 fusion oncoproteins

To further study how NUP98-FOs drive leukemogenesis, we profiled the genomic binding of N-terminal HA-tagged NUP98-FOs (Fig.S8A), using anti-HA antibody in cbCD34 models along with histone modifications (H3K4me3 and H3K27ac) and menin using CUT&RUN[21] (Fig.5A-B, Fig.S8B-D). FO-binding peaks were enriched in the promoter and intronic regions of protein-coding genes with high reproducibility among biological replicates (Fig.S8C). All three FOs showed binding to the *HOXA-B* clusters as previously reported using different models[16-19], whereas the binding patterns along the anterior-posterior axes were unique to each fusion, corresponding to RNA expression, active histone modifications, and menin binding (Fig.5A, Fig.S8D). Cells with empty vector controls showed only weak H3K4me3 and H3K27ac signals on the *HOXA/B* clusters, consistent with the observed low RNA expression (Fig.5A) . Comparisons of protein-coding genes with FO-binding peaks showed 97 common target genes for all

three NUP98-FOs, including genes involved in leukemogenesis and hematopoiesis (e.g., *RUNX1*, *CDK6*), whereas each FO also showed specific targets and binding motifs (Fig.5B, Sig.S8B). Notably, NUP98::KDM5A uniquely showed peaks at ETS motifs and on *GFI1B* and *MEIS2,* which are megakaryocyte/erythroid differentiation regulators[51-54], indicating that direct regulation of differentiation-related genes could drive phenotypes of *NUP98::KDM5A* AMKL.

We next sought to profile endogenous NUP98-FO binding without epitope tags in patient samples or cell lines. We tested N-terminus NUP98 (N-NUP98) antibodies in cbCD34, which showed similar binding patterns albeit with weaker signals yielding fewer peaks than HA-antibody, possibly underestimating weak FO-binding (Fig.S9A). CHRF-288-11 and ST1653 showed NUP98 signals at genes highly expressed in these cell lines, such as *HOXA/B* genes and *MEIS1*, with menin co-localization (Fig.S9B). Contrarily, weak or low signals were observed in an empty vector control from cdCD34 models or *RUNX1*::*RUNX1T1* AML (lacking *HOX* gene expression) or *UBTF*-TD AML (resembling *NUP98::NSD1* AML expression[6]) (Fig.5A, Fig.S9B). We then applied CUT&RUN using antibodies for N-NUP98 and histone modifications (H3K4me3, H3K27ac, H3K27me3) for 14 *NUP98*r patient samples covering *NUP98::NSD1*, *NUP98::KDM5A*, and *NUP98::RAP1GDS1* (Fig.5C, Fig.S9C, Table S12). N-NUP98 antibodies revealed target genes for each FO, including *HOXA/B* genes with heterogeneous patterns even within the same NUP98-FO groups (Fig.5D-E). Also, N-NUP98 antibodies showed NUP98::KDM5A-specific targets (1,518 genes), whereas 21 genes, including *MEIS2* or *MECOM,* were *NUP98::KDM5A* AMKL-specific. Genes with H3K4me3 peaks showed a broad overlap among samples, whereas 61 genes, including *IGF2BP3* associated with megakaryocyte development[55], were unique to *NUP98::KDM5A* AMKL (Fig.S9D). For H3K27ac and H3K27me3 signals with broad peaks that could span multiple genes, we applied the PBS[56] (probability of being signal) approach to quantify the genome-wide signals. This analysis revealed that genome-wide H3K27ac status correlated with disease phenotypes, whereas H3K27me3 status was associated with fusion partners (Fig.5F). Comparisons of *NUP98::KDM5A* with the other fusions (*NUP98*::*NSD1* and *NUP98*::*RAP1GDS1*) showed limited regions with differential H3K27ac signals (Fig.5G). In contrast, a comparison between *NUP98*::*KDM5A* AML and AMKL showed enrichment of *MECOM*, *IGF2BP1*, or *GATA2* regions in *NUP98::KDM5A* AMKL, indicating phenotype-specific epigenetic status (Fig.5G-H, Fig.S9E). These data suggest that both fusion partners and cellular differentiation status can affect FO-binding patterns to establish the heterogeneous epigenetic and expression profiles in *NUP98*r leukemia.

Functional characterization of recurrent somatic alterations in NUP98r leukemia

Given the differential NUP98-FO binding and mutation patterns related to disease phenotypes, we investigated the impact of co-occurring alterations on *NUP98*r cbCD34 models. We first established cbCD34 models with constitutive Cas9 expression, followed by lentiviral transduction of gRNA designed to mimic *WT1* frameshift mutations or *RB1* loss (Fig.6A, Fig.S10A-B). In NUP98::KDM5A cbCD34 models, *WT1* gRNAs enhanced cell growth compared with AAVS-targeting controls, whereas out-of-frame indels in *RB1* were enriched at later time points, indicating growth advantages of cells with *RB1* loss (Fig.6B-C). Morphologically, AAVS-controls showed cells with various degrees of myeloid and erythrocyte/megakaryocyte differentiation, whereas *RB1*-gRNAs yielded larger cells with features of maturing megakaryocytes, and *WT1*-gRNAs enriched homogenously small blast-like cells (Fig.6B). Flow cytometric analysis showed that *RB1*-gRNAs enriched CD34+CD41a+ cells, whereas *WT1*-gRNAs depleted the CD41a+ cells with increased CD34+CD41a- cells (Fig.6D, Fig.S10C). *NUP98::NSD1* cbCD34 models showed growth advantages with *WT1*-gRNAs, whereas *RB1*-gRNAs did not show either growth advantage or consistent enrichment of out-of-frame indels, suggesting that *RB1* loss specifically synergizes with NUP98::KDM5A (Fig.S10D-E). RNAseq revealed global transcriptional changes in *RB1*- or *WT1*-gRNA conditions (Fig.6E). *RB1*-gRNAs led to upregulation of *CD34* and *CDKN2A*, which is downstream of RB1/E2F and forms a feedback loop[57,58], and changes in immunity genes compared with AAVS-controls (Fig.6F, Fig.S11A). *WT1*-gRNAs also increased *CD34* expression, along with increased MHC-class II and decreased megakaryocyte/platelet-related gene expression (Fig.6G).

ScRNAseq from these conditions collectively showed differentiation trajectories resembling normal hematopoiesis (Fig.6H-I, Fig.S11B). A control condition showed broad distributions with enrichment at the platelet-like populations

adjacent to the MEP-like or megakaryocyte (MK)-like clusters. By contrast, *RB1*-gRNAs led to enrichment at the MEP or MK-like clusters and lacked platelet-like populations, and *WT1*-gRNA resulted in a bias toward the GMP- and LMPP-like populations rarely found in control and *RB1*-gRNA conditions. Pseudotime analyses showed that *RB1* was upregulated through platelet differentiation and highest in the platelet-like cluster, whereas *CDKN2A* expression was high in the MEP or MK clusters but low or undetectable in the platelet cluster (Fig.6J-K, Fig.S11C). We also observed negative correlations with *RB1* and *CDKN2A* among *NUP98::KDM5A* patient samples (Fig.S11D). The platelet-like cluster had higher expression of translation-related genes than the MK-like cluster (Fig.6L), recapitulating changes in translation in normal platelet production from MK or proplatelets.[59] These data suggest that *RB1* loss inhibits terminal platelet differentiation in the NUP98::KDM5A model, leading to the accumulation of cells at the MK-like or MEP-like stages, resembling the megakaryocytic differentiation block seen in E2F transgenic mice[60] (Fig.6M). The LMPP population enriched in the *WT1*-gRNA condition was characterized by high MHC-class II gene and low cell-division gene expression, and cell-cycle inference indicated enrichment of cells at G1-stage in the LMPP cluster and S/G2M in the GMP cluster (Fig.S11E-G), suggesting a unique hierarchy of proliferating GMP-like cells and dormant LMPP-like cells with *WT1*-gRNA-induced mutations.

Cooperating alterations and differentiation status affect sensitivity to menin inhibition

*NUP98*r leukemia has been reported to be dependent on KMT2A/menin complex[18], and menin inhibitors are currently being tested in clinical trials enrolling patients with AML, including those with *NUP98*r[61]. Mouse models showed that menin inhibition led to myeloid differentiation in *NPM1*-mutant and *NUP98*r leukemia models[18,62-64], although whether menin is required for megakaryocytic lineages or AMKL is currently unclear. We tested a menin inhibitor (revumenib) in NUP98::KDM5A CD34 models with cooperating alterations to assess the sensitivity and impact on cellular populations (Fig.7A). Two-week treatment with revumenib (1μM) decreased cell growth in AAVS conditions (14.1±4.6%; mean ± standard error) compared with DMSO, whereas *RB1*- or *WT1*-gRNA conditions were less sensitive (*RB1*-gRNA: 64.3±10.3%, *WT1*-gRNA, 56.8±4.0%, Fig.7B). Flow cytometric analysis showed significant increases in the CD11b+ differentiated myeloid population in AAVS-control and *WT1*-gRNA conditions, whereas *RB1*-gRNA conditions increased the CD34+CD41a+ population (Fig.7C), suggesting adaptive changes of cellular hierarchies. RNAseq on day 7 of treatment showed increased expression of genes previously reported to be downregulated in *NUP98*r models in all conditions, whereas *CDKN2* family[65], inflammation, and erythroid/megakaryocyte-related genes were enriched in AAVS conditions compared with *RB1*-gRNA conditions, suggesting differentiation induced by menin inhibition (Fig.7D). Menin inhibition in *WT1*-gRNA showed immunity-related changes as in AAVS controls, along with activation of monocyte/lymphocyte genes (e.g., *CD48*, *MAFB,* Fig.S11H), suggesting partial differentiation from dominant GMP or LMPP clusters. We also tested revumenib on ST1653 (monocytic) and CHRF-288-11 (megakaryocytic), which showed that ST1653 was sensitive to menin inhibition but CHRF-288-11 was not, further confirming the association of differentiation status with menin-dependency (Fig.7E). These collective data indicate that cellular differentiation and cooperating mutations can both influence menin dependency of *NUP98*r leukemias (Fig.7F).

DISCUSSION

Biology-informed treatment of subtypes of pediatric leukemia is paving the way for improved clinical outcomes[3,66]; however, the biology driving the heterogeneous phenotypes of *NUP98*r leukemia has not yet been fully characterized. Using genome-wide mutational and transcriptional analyses of a large collective cohort of 185 *NUP98*r leukemia, we expand on the known association of *NUP98*r with *WT1*, *FLT3*-ITD, and *RB1* loss to highlight *JAK2*, *GATA1*, and *MPL* alterations associated with AMKL and *NOTCH1* alterations with T-ALL phenotypes. Profiling with scRNAseq and deconvolution of bulk RNAseq data additionally shows heterogeneous cellular hierarchies even with the same fusion partners, suggesting that both fusion partners and cooperating mutations can contribute to the expression profiles and cellular hierarchies.[41,43]

Despite broadly shared binding to the *HOXA/B* clusters, NUP98 FO-binding patterns are also unique to each fusion partner, whereas profiling of patient samples and cell lines showed that the binding patterns are not static and could be dynamic according to the differentiation status, represented by AMKL-specific NUP98::KDM5A-binding to *MEIS2* or *MECOM.* These data highlight the importance of using human hematopoietic cells to profile NUP98 FO-binding and epigenetic status to determine their contribution to leukemogenesis. We also found that *RB1* deletion and *WT1* frameshift mutations could render cells with growth advantages and alter the cellular hierarchies of *NUP98*r *in vitro* models in a fusion partner-specific manner. Previously, transplant models of cbCD34 NUP98::KDM5A cells showed heterogeneous disease phenotypes[49], and whether these cooperating mutations can affect phenotypes *in vivo* is to be investigated. Although acquired mutations in the *MEN1* gene have been clinically observed as mechanisms of resistance to menin inhibitors in other AML subtypes[67,68], our *in vitro* models also suggest that differentiation-specific menin dependency can be a different resistance mechanism for *NUP98*r leukemia. It will be critical to investigate clinical outcomes of *NUP98*r AML treated with menin inhibitors according to fusion partners, cooperating mutations, cellular differentiation, and expression profiles to identify patients who will not benefit from menin inhibition and treat them with alternative approaches.

Mechanistically, RB1 is a repressor of E2F1 proteins, which in turn activate the expression of cell cycle-related genes, including *CDKN2* families.[57,58] E2F1 has been shown to inhibit normal platelet differentiation, leading to the accumulation of abnormal megakaryocytes in E2F1 transgenic mice[60], whereas *Cdkn2a*-deficient mice showed enhanced platelet production.[69] NUP98-FO binding data from cbCD34 models suggests that preferential differentiation toward the megakaryocytic lineage is NUP98::KDM5A-intrinsic, whereas *RB1* loss could lead to the accumulation of megakaryocytic cells and ultimately development of AMKL via high E2F activity and *CDKN2* family expression. Another consideration is the cell-of-origin of fetal hematopoiesis, which has been shown to affect the AMKL phenotype of *CBFA2T3*(*ETO*)-*GLIS2* leukemia.[70] Similarly, myelogenic phenotypes of *NUP98::NSD1*, which are common in adolescents and adults, may be supported preferentially in the adult hematopoietic niche.[71] Future studies using mice at the developmental stages or transplantation models could dissect the contribution of cell-of-origin (fetal or adult definitive hematopoiesis) or microenvironments (fetal liver or bone marrow) to the *NUP98*r leukemogenesis and provide mechanistic insights needed for better treatment for these refractory diseases.

ACKNOWLEDGMENTS

We thank all the patients and their families at St. Jude Children's Research Hospital (SJCRH) for their contribution to the biological specimens used in this study. We also thank the Biorepository, the Flow Cytometry and Cell Sorting Core, the Center for Applied Bioinformatics (CAB), the Center for Advanced Genome Engineering (CAGE), and the Hartwell Center for Biotechnology at SJCRH for their essential services, which are supported by the NIH (P30 CA021765, Cancer Center Support Grant). This work was funded by the American Lebanese and Syrian Associated Charities of St. Jude Children's Research Hospital and grants from the NIH (U54 CA243124, Fusion Oncoproteins in Childhood Cancers (FusOnC2) Consortium granted for C.G. Mullighan and J.M. Klco, R35CA197695 for C.G. Mullighan, R01 CA276079 and CA285272 to J.M. Klco, and F32 CA261011 and K99 CA283256 to N.L. Michmerhuizen). The content, however, does not necessarily represent the official views of the NIH and is solely the responsibility of the authors. J.M. Klco holds a Career Award for Medical Scientists from the Burroughs Wellcome Fund and is a previous recipient of the V Foundation Scholar Award (Pediatric). D.D. Giacomo received financial support under the National Recovery and Resilience Plan (NRRP), Mission 4, Component 2, Investment 1.1, Call for tender No. 1409 published on 14.9.2022 by the Italian Ministry of University and Research (MUR), funded by the European Union-NextGenerationEU, Project Title "Experimental modeling to identify molecular mechanisms and therapeutic vulnerabilities in NUP98-rearranged leukemia", CUP:J53D23017510001- Grant Assignment Decree No.1384 adopted on 01/09/2023 by the Italian Ministry of Ministry of University and Research (MUR). C. Mecucci was supported by AIRC 5×1000 called "Metastatic disease: the key unmet need in oncology" to MYNERVA project, #21267 (MYeloid NEoplasms Research Venture AIRC). R. Masetti was supported by AIRC Associazione Italiana per la Ricerca sul Cancro (AIRC IG 26039). J.M. Barajas is a Fellow of The Jane Coffin Childs Fund for Medical Research. F. Locatelli is a recipient of "European Union - Next Generation EU, Mission 4, Component 2, CUP B93D21010860004, National Center for Gene Therapy and Drugs Based on RNA Technology", "Hub Life Science – Terapia Avanzata (LSH-TA) PNC-E3-2022-23683269 – CUP E83C22006230001 – from the Italian Ministry of Health - Piano Nazionale Complementare Ecosistema Innovativo della Salute - cod PNC-E.3.", and Ministero dell'Università e della Ricerca (Grant PRIN 2020 and Grant PRIN 2022).

AUTHORSHIP CONTRIBUTION

J.M.K and C.G.M conceptualized the study. J.M.K, M.P.W, G.S, J.M, T.W, P.P, I.I, and C.G.M collected data. J.M.K, M.U M.P.W, G.S, J.M, and T.W curated data. M.U and J.M conducted formal analyses. M.U, R.H, N.L.M, J.B, T.M. M.E.T, B.A, and A.K performed experiments. M.U and J.M performed statistical tests. MU made figures. P.P, C.M, D.D.G, F.L, R.M, S.N.B, M.P, S.M.P-M, S.P, J.R, H.I, K.P.P, M.J.W, I.I, and C.G.M provided resources. M.U and J.M.K wrote the original draft, and all the co-authors reviewed it and were involved in the edition.

DISCLOSURE OF CONFLICTS OF INTEREST

The authors do not have any conflicts of interest regarding the content of this manuscript.

REFERENCES

1. Rubnitz, J.E. & Kaspers, G.J.L. How I treat pediatric acute myeloid leukemia. *Blood* **138**, 1009-1018 (2021).
2. Rubnitz, J.E. *et al.* Clofarabine Can Replace Anthracyclines and Etoposide in Remission Induction Therapy for Childhood Acute Myeloid Leukemia: The AML08 Multicenter, Randomized Phase III Trial. *J Clin Oncol* **37**, 2072-2081 (2019).
3. Pollard, J.A. *et al.* Sorafenib in Combination With Standard Chemotherapy for Children With High Allelic Ratio FLT3/ITD+ Acute Myeloid Leukemia: A Report From the Children's Oncology Group Protocol AAML1031. *J Clin Oncol* **40**, 2023-2035 (2022).
4. Balgobind, B.V. *et al.* Novel prognostic subgroups in childhood 11q23/MLL-rearranged acute myeloid leukemia: results of an international retrospective study. *Blood* **114**, 2489-96 (2009).
5. Umeda, M. *et al.* A new genomic framework to categorize pediatric acute myeloid leukemia. *Nat Genet* (2024).
6. Umeda, M. *et al.* Integrated Genomic Analysis Identifies UBTF Tandem Duplications as a Recurrent Lesion in Pediatric Acute Myeloid Leukemia. *Blood Cancer Discov* **3**, 194-207 (2022).
7. Gruber, T.A. *et al.* An Inv(16)(p13.3q24.3)-encoded CBFA2T3-GLIS2 fusion protein defines an aggressive subtype of pediatric acute megakaryoblastic leukemia. *Cancer Cell* **22**, 683-97 (2012).
8. de Rooij, J.D. *et al.* Pediatric non-Down syndrome acute megakaryoblastic leukemia is characterized by distinct genomic subsets with varying outcomes. *Nat Genet* **49**, 451-456 (2017).
9. Hollink, I.H. *et al.* NUP98/NSD1 characterizes a novel poor prognostic group in acute myeloid leukemia with a distinct HOX gene expression pattern. *Blood* **118**, 3645-56 (2011).
10. Bertrums, E.J.M. *et al.* Comprehensive molecular and clinical characterization of NUP98 fusions in pediatric acute myeloid leukemia. *Haematologica* (2023).
11. Michmerhuizen, N.L., Klco, J.M. & Mullighan, C.G. Mechanistic insights and potential therapeutic approaches for NUP98-rearranged hematologic malignancies. *Blood* **136**, 2275-2289 (2020).
12. Polonen, P., Mullighan, C.G. & Teachey, D.T. Classification and risk stratification in T-lineage acute lymphoblastic leukemia. *Blood* (2024).
13. Schwartz, J.R. *et al.* The genomic landscape of pediatric myelodysplastic syndromes. *Nat Commun* **8**, 1557 (2017).
14. Takahashi, K. *et al.* Integrative genomic analysis of adult mixed phenotype acute leukemia delineates lineage associated molecular subtypes. *Nat Commun* **9**, 2670 (2018).
15. Tian, J. *et al.* The landscape of NUP98 rearrangements clinical characteristics and treatment response from 1491 acute leukemia patients. *Blood Cancer J* **14**, 81 (2024).
16. Xu, H. *et al.* NUP98 Fusion Proteins Interact with the NSL and MLL1 Complexes to Drive Leukemogenesis. *Cancer Cell* **30**, 863-878 (2016).
17. Oka, M. *et al.* Chromatin-prebound Crm1 recruits Nup98-HoxA9 fusion to induce aberrant expression of Hox cluster genes. *Elife* **5**, e09540 (2016).
18. Heikamp, E.B. *et al.* The menin-MLL1 interaction is a molecular dependency in NUP98-rearranged AML. *Blood* **139**, 894-906 (2022).
19. Ahn, J.H. *et al.* Phase separation drives aberrant chromatin looping and cancer development. *Nature* **595**, 591-595 (2021).
20. Chandra, B. *et al.* Phase Separation Mediates NUP98 Fusion Oncoprotein Leukemic Transformation. *Cancer Discov* **12**, 1152-1169 (2022).
21. Skene, P.J. & Henikoff, S. An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *Elife* **6**(2017).

22. Togni, M. *et al.* Identification of the NUP98-PHF23 fusion gene in pediatric cytogenetically normal acute myeloid leukemia by whole-transcriptome sequencing. *J Hematol Oncol* **8**, 69 (2015).
23. Bolouri, H. *et al.* The molecular landscape of pediatric acute myeloid leukemia reveals recurrent structural alterations and age-specific mutational interactions. *Nat Med* **24**, 103-112 (2018).
24. Iacobucci, I. *et al.* Genomic subtyping and therapeutic targeting of acute erythroleukemia. *Nat Genet* **51**, 694-704 (2019).
25. McNeer, N.A. *et al.* Genetic mechanisms of primary chemotherapy resistance in pediatric acute myeloid leukemia. *Leukemia* **33**, 1934-1943 (2019).
26. Buelow, D.R. *et al.* Uncovering the Genomic Landscape in Newly Diagnosed and Relapsed Pediatric Cytogenetically Normal FLT3-ITD AML. *Clin Transl Sci* **12**, 641-647 (2019).
27. Fornerod, M. *et al.* Integrative Genomic Analysis of Pediatric Myeloid-Related Acute Leukemias Identifies Novel Subtypes and Prognostic Indicators. *Blood Cancer Discov* **2**, 586-599 (2021).
28. Schwartz, J.R. *et al.* The acquisition of molecular drivers in pediatric therapy-related myeloid neoplasms. *Nat Commun* **12**, 985 (2021).
29. Umeda, M. *et al.* A new genomic framework to categorize pediatric acute myeloid leukemia. *Nat Genet* **56**, 281-293 (2024).
30. Polonen, P. *et al.* The genomic basis of childhood T-lineage acute lymphoblastic leukaemia. *Nature* **632**, 1082-1091 (2024).
31. Tian, L. *et al.* CICERO: a versatile method for detecting complex and diverse driver fusions using cancer RNA sequencing data. *Genome Biol* **21**, 126 (2020).
32. Chellappan, S.P., Hiebert, S., Mudryj, M., Horowitz, J.M. & Nevins, J.R. The E2F transcription factor is a cellular target for the RB protein. *Cell* **65**, 1053-61 (1991).
33. Benevolenskaya, E.V. & Frolov, M.V. Emerging links between E2F control and mitochondrial function. *Cancer Res* **75**, 619-23 (2015).
34. McMahon, C.M. *et al.* Clonal Selection with RAS Pathway Activation Mediates Secondary Clinical Resistance to Selective FLT3 Inhibition in Acute Myeloid Leukemia. *Cancer Discov* **9**, 1050-1063 (2019).
35. Gonen, M. *et al.* CD25 expression status improves prognostic risk classification in AML independent of established biomarkers: ECOG phase 3 trial, E1900. *Blood* **120**, 2297-306 (2012).
36. Xu, Z.J. *et al.* The M2 macrophage marker CD206: a novel prognostic indicator for acute myeloid leukemia. *Oncoimmunology* **9**, 1683347 (2020).
37. Lasry, A. *et al.* An inflammatory state remodels the immune microenvironment and improves risk stratification in acute myeloid leukemia. *Nat Cancer* **4**, 27-42 (2023).
38. de Pooter, R.F. *et al.* Cutting Edge: Lymphomyeloid-Primed Progenitor Cell Fates Are Controlled by the Transcription Factor Tal1. *J Immunol* **202**, 2837-2842 (2019).
39. Kim, Y. *et al.* Terminal deoxynucleotidyl transferase and CD84 identify human multi-potent lymphoid progenitors. *Nat Commun* **15**, 5910 (2024).
40. Crescenzi, B. *et al.* NUP98/11p15 translocations affect CD34+ cells in myeloid and T lymphoid leukemias. *Leuk Res* **39**, 769-72 (2015).
41. Zeng, A.G.X. *et al.* A cellular hierarchy framework for understanding heterogeneity and predicting drug response in acute myeloid leukemia. *Nat Med* **28**, 1212-1223 (2022).
42. Bottomly, D. *et al.* Integrative analysis of drug response and clinical outcome in acute myeloid leukemia. *Cancer Cell* **40**, 850-864 e9 (2022).
43. Lambo, S. *et al.* A longitudinal single-cell atlas of treatment response in pediatric AML. *Cancer Cell* **41**, 2117-2135 e12 (2023).
44. Xu, J. *et al.* A multiomic atlas identifies a treatment-resistant, bone marrow progenitor-like cell population in T cell acute lymphoblastic leukemia. *Nat Cancer* (2024).
45. Newman, A.M. *et al.* Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat Biotechnol* **37**, 773-782 (2019).
46. Fugman, D.A., Witte, D.P., Jones, C.L., Aronow, B.J. & Lieberman, M.A. In vitro establishment and characterization of a human megakaryoblastic cell line. *Blood* **75**, 1252-61 (1990).
47. Takeda, A., Goolsby, C. & Yaseen, N.R. NUP98-HOXA9 induces long-term proliferation and blocks differentiation of primary human CD34+ hematopoietic cells. *Cancer Res* **66**, 6628-37 (2006).
48. Chung, K.Y. *et al.* Enforced expression of NUP98-HOXA9 in human CD34(+) cells enhances stem cell proliferation. *Cancer Res* **66**, 11781-91 (2006).
49. Cardin, S. *et al.* Human models of NUP98-KDM5A megakaryocytic leukemia in mice contribute to uncovering new biomarkers and therapeutic vulnerabilities. *Blood Adv* **3**, 3307-3321 (2019).
50. Diaz, A., 3rd *et al.* Abstract C099: Establishment and characterization of an NUP98-KDM5A-driven AML XPDX model. *Molecular Cancer Therapeutics* **22**, C099-C099 (2023).

51. Saleque, S., Cameron, S. & Orkin, S.H. The zinc-finger proto-oncogene Gfi-1b is essential for development of the erythroid and megakaryocytic lineages. *Genes Dev* **16**, 301-6 (2002).
52. Cai, M. *et al.* Dual actions of Meis1 inhibit erythroid progenitor development and sustain general hematopoietic cell proliferation. *Blood* **120**, 335-46 (2012).
53. Pang, L. *et al.* Maturation stage-specific regulation of megakaryopoiesis by pointed-domain Ets proteins. *Blood* **108**, 2198-206 (2006).
54. Levay, K. & Slepak, V.Z. Tescalcin is an essential factor in megakaryocytic differentiation associated with Ets family gene expression. *J Clin Invest* **117**, 2672-83 (2007).
55. Elagib, K.E. *et al.* Neonatal expression of RNA-binding protein IGF2BP3 regulates the human fetal-adult megakaryocyte transition. *J Clin Invest* **127**, 2365-2377 (2017).
56. Hecht, V. *et al.* Analyzing histone ChIP-seq data with a bin-based probability of being signal. *PLoS Comput Biol* **19**, e1011568 (2023).
57. Foulkes, W.D., Flanders, T.Y., Pollock, P.M. & Hayward, N.K. The CDKN2A (p16) gene and human cancer. *Mol Med* **3**, 5-20 (1997).
58. Stott, F.J. *et al.* The alternative product from the human CDKN2A locus, p14(ARF), participates in a regulatory feedback loop with p53 and MDM2. *EMBO J* **17**, 5001-14 (1998).
59. Mills, E.W., Green, R. & Ingolia, N.T. Slowed decay of mRNAs enhances platelet specific translation. *Blood* **129**, e38-e48 (2017).
60. Guy, C.T., Zhou, W., Kaufman, S. & Robinson, M.O. E2F-1 blocks terminal differentiation and causes proliferation in transgenic megakaryocytes. *Mol Cell Biol* **16**, 685-93 (1996).
61. Swaminathan, M., Bourgeois, W., Armstrong, S.A. & Wang, E.S. Menin Inhibitors in Acute Myeloid Leukemia-What Does the Future Hold? *Cancer J* **28**, 62-66 (2022).
62. Uckelmann, H.J. *et al.* Therapeutic targeting of preleukemia cells in a mouse model of NPM1 mutant acute myeloid leukemia. *Science* **367**, 586-590 (2020).
63. Uckelmann, H.J. *et al.* Mutant NPM1 Directly Regulates Oncogenic Transcription in Acute Myeloid Leukemia. *Cancer Discov* **13**, 746-765 (2023).
64. Wang, X.Q.D. *et al.* Mutant NPM1 Hijacks Transcriptional Hubs to Maintain Pathogenic Gene Programs in Acute Myeloid Leukemia. *Cancer Discov* **13**, 724-745 (2023).
65. Soto-Feliciano, Y.M. *et al.* A Molecular Switch between Mammalian MLL Complexes Dictates Response to Menin-MLL Inhibition. *Cancer Discov* **13**, 146-169 (2023).
66. Issa, G.C. *et al.* The menin inhibitor revumenib in KMT2A-rearranged or NPM1-mutant leukaemia. *Nature* **615**, 920-924 (2023).
67. Perner, F. *et al.* MEN1 mutations mediate clinical resistance to menin inhibition. *Nature* **615**, 913-919 (2023).
68. Tiong, I.S., Ritchie, D.S. & Blombery, P. Response and Resistance to Menin Inhibitor in UBTF-Tandem Duplication AML. *N Engl J Med* **390**, 2323-2325 (2024).
69. Wang, W. *et al.* Enhanced Megakaryopoiesis and Platelet Activity in Hypercholesterolemic, B6-Ldlr-/-, Cdkn2a-Deficient Mice. *Circ Cardiovasc Genet* **9**, 213-22 (2016).
70. Lopez, C.K. *et al.* Ontogenic Changes in Hematopoietic Hierarchy Determine Pediatric Specificity and Disease Phenotype in Fusion Oncogene-Driven Myeloid Leukemia. *Cancer Discov* **9**, 1736-1753 (2019).
71. Li, Y. *et al.* FLT3ITD drives context-specific changes in cell identity and variable interferon dependence during AML initiation. *Blood* **141**, 1442-1456 (2023).

FIGURE LEGENDS

Figure 1. ***Heterogeneity of pediatric NUP98r leukemia***
A. Details of *NUP98*-rearranged (*NUP98*r) leukemia samples (n=185, **left**) and analytical pipelines (**right**). **B.** Numbers and functional annotations of fusion partners in the study cohort. Colors indicate protein functional groups. **C**. Age distribution related to fusion partners and disease types. Colors indicate disease types. **D**. UMAP (Uniform Manifold Approximation and Projection) plots of the transcriptional cohort (n=2,321) colored according to leukemia subtypes (**left**) and *NUP98* fusion partners (**mid**) and enrichment of fusion partners in transcriptional clusters (**right**). The shapes of dots indicate disease types (circles-acute myeloid leukemia: AML, triangles-acute lymphoblastic leukemia: ALL), and colors in the heatmap indicate enrichment of fusions in each cluster, asterisks indicate statistically significant adjusted *P*-values from two-sided Fisher's exact tests and the Benjamini-Hochberg adjustment (*<0.05, **<0.01, ***<0.001, black: enriched, blue: exclusive). In **A** and **C**, lines of the box plots represent the 25% quantile, median, and 75% quantile, and the upper whisker represents the higher value of maxima or 1.5 x interquartile range (IQR), and the lower whisker represents the lower value of minima or 1.5 x IQR. Abbreviations. tMN: therapy-related myeloid neoplasm, MDS/MPN: myelodysplastic syndrome/ myeloproliferative neoplasms, AEL/AMKL: acute erythroid leukemia/ acute megakaryocytic leukemia, ETP: Early T-cell precursor ALL. Abbreviations in leukemia subtypes are found in Table S10.

Figure 2. ***Mutational background of NUP98r leukemia associated with disease phenotypes***
A. Genetic profiles of *NUP98*r samples in the cohort. Colors indicate patient annotations (**top**) and types of gene alterations (**bottom**). **B**. Co-occurrence and mutual exclusivity among recurrent alterations (n≥3). **C.** Enrichment of somatic alterations in transcriptional clusters (**left**) and fusion partners (**right**). In **B** and **C**, colors indicate adjusted *P*-values by two-sided Fisher's exact tests and the Benjamini-Hochberg adjustment (red: co-occurring, blue: mutually exclusive), and asterisks indicate statistically significant values (adjusted *P*-values *<0.05, **<0.01, ***<0.001). Annotations of genes in mutational heatmaps depend on known general functions.

Figure 3. ***Varieties of cellular hierarchies in NUP98r leukemia***
A. Strategies for single cell RNAseq (scRNAseq) and deconvolution of bulk RNAseq data. **B**. UMAP plots of patient samples colored by sample source (**left**) and transcriptional clusters (**right**). **C**. Enrichment of cells in each cluster indicated by colors and sizes (**left**) and marker gene expression indicated by colors (averaged expression) and size (ratio of expressing cells: count>0). **D**. UMAP plot of reference bone marrow and thymocyte scRNAseq data, colored according to cell labels from the original reference data44. **E**. Distribution of patient sample scRNAseq on the reference data inferred by the MapQuery function in the Seurat package. Cells in normal hematopoietic cell clusters were excluded. Cells are colored according to the cell density on UMAP. Cooperating mutations found in bulk samples are also shown. **F**. Enrichment of cells with each cell label inferred by Seurat, indicated by colors and sizes. **G**. Cellular component of bulk RNA samples (n=185) inferred by CIBERSORT using a signature matrix derived from reference scRNAseq data. Bars are colored by cell populations in each sample. Abbreviations. HSPC: hematopoietic stem and progenitor cell, LMPP: Lympho-myeloid primed progenitor, GMP: Granulocyte-monocyte progenitor, cDC: classic dendritic cell, pDC: plasmacytoid dendritic cell, CLP: common lymphoid progenitor, DP: CD4-CD8 double-positive T-cell, NK: natural killer T-cell, MEP: Megakaryocyte–erythroid progenitor

Figure 4. ***Cord blood CD34 models recapitulate the phenotypes of NUP98r leukemia***
A. Experimental schema using cord blood CD34+ cell models (cbCD34). **B**. Colony-forming unit assays of cbCD34 models with empty control vectors or *NUP98::NSD1*, *NUP98::KDM5A*, or *NUP98::HOXA9*-expressing vectors. **C**. Cell growth assays of cbCD34 models in liquid culture. **D**. Flow cytometric analysis of cbCD34 models in liquid culture (**top**: CD34+, **mid**: CD11b+, **bottom**: CD41a+ population ratio: % in mCherry+ live cells). **E**. Principal component analysis (PCA) of RNAseq data from liquid culture. Colors indicate NUP98 fusions, and shapes indicate days after transduction. **F**. Heatmap showing expression of representative genes related to stemness or differentiation of hematopoietic cells. The colors of cells indicate expression levels normalized among samples, and genes are annotated on the left. **G**. Comparison of differentially expressed genes (DEGs) in each cbCD34 model compared with empty vector controls at day 42. Venn diagram showing overlaps of highly expressed genes in each model (**mid**) and GO term analyses of shared or specific DEGs are shown (**left**, **right**). Data was obtained from three biological replicates (different lots of cord blood). In **B**-**D**, statistical tests were performed by a generalized

linear mixed effect model with Poisson (**B**) and Gaussian (**C**-**D**) distributions, followed by comparison with empty vector control and the Benjamini-Hochberg adjustment, asterisks indicating adjusted *P*-values *<0.05. Error bars indicate mean ± s.e.m. Abbreviations. FO: fusion oncoprotein, PDX: patient-derived xenograft, FDR: false-discovery rate.

Figure 5. ***Differential gene regulation by NUP98 fusion oncoproteins***

A. IGV tracks of the *HOXA-B* clusters from CUT&RUN using HA, H3K4me3, and H3K27ac antibodies in HA-tagged *NUP98*r cbCD34 models (**top**: Empty vector control-gray, HA-NUP98::KDM5A-red, HA-NUP98::NSD1-blue, HA-NUP98::HOXA9-black) and heatmap showing expression levels of *HOXA-B* genes (**bottom**). N-terminal NUP98 antibody was applied for the empty vector control. **B**. IGV tracks of differentiation-related gene loci (**top**: *RUNX1*, *GFI1B*, and the *HOXC* cluster and Venn diagram showing overlap of protein-coding genes with annotated peaks (**bottom**: FDR<0.00001). **C**. CUT&RUN strategy from primary patient samples or *NUP98::KDM5A* cell lines (CHRF-288-11 and ST1653). **D**. Counts of peaks from the N-terminus NUP98 antibody in primary samples (**top**) and overlaps of target genes among non-NUP98::KDM5A and NUP98::KDM5A. NUP98::KDM5A AMKL-specific 21 target genes are highlighted. Colors indicate peak annotations. **E**. IGV tracks of the *HOXA-B* cluster from CUT&RUN using N-terminal NUP98, H3K4me3, and H3K27ac antibodies in primary leukemia samples and *NUP98::KDM5A* cell lines. **F**. PCA of genome-wide PBS (probability being signals) scores of H3K27ac (**left**) and H3K27me3 (**right**) from primary samples. Colors indicate expression clusters, and shapes indicate fusion partners. **G**. Differential signal analysis using H3K27ac PBS scores between *NUP98::KDM5A* and other (*NUP98::NSD1* and *NUP98::RAP1GDS1*) samples (**left**) and *NUP98::KDM5A* AMKL and non-AMKL (**right**) calculated by limma, followed by the Benjamini-Hochberg adjustment. Only regions with significant enrichment (adjusted *P*-values < 0.05) are shown. **H**. IGV tracks of the *MECOM* and *MEIS2* gene loci from CUT&RUN using N-terminal NUP98, H3K4me3, and H3K27ac antibodies in primary leukemia samples and *NUP98::KDM5A* cell lines.

Figure 6. ***Functional characterization of recurrent somatic alterations in NUP98r leukemia***

A. Experimental schema of induction of cooperating alterations (*RB1*, *WT1*) in cbCD34/Cas9 models. **B**. Cell growth assays of cbCD34/Cas9 NUP98::KDM5A with gRNAs targeting the AAVS, *RB1*, or *WT1* loci (**left**), and cytospin of cells on day 35 (**right**). **C**. Induction rates of indel (insertions and deletions) at days 4 and 39 in each condition. Bars represent fractions of indel rates in all target sequence reads, and dots represent the out-of-frame indel ratio among total indels. **D**. Flow gating (**left**), CD34+ CD41a+ positivity (**mid**), and CD34+ CD41a- (**right**) among mCherry+ GFP+ mAmetrine+ live cells. **E**. PCA of RNAseq of gRNA-transduced NUP98::KDM5A cbCD34 models at day 35. **F**. DEG analysis between AAVS controls and *RB1*-gRNA conditions (**left**) and gene ontology (GO) term analysis of DEGs (**right**). Colors indicate DEGs and GO terms (red: high in *RB1*-gRNA conditions, blue: low in *RB1*-gRNA conditions)**. G**. DEG analysis between AAVS controls and *WT1*-gRNA conditions as shown in **F**. **H**. UMAP plots of scRNAseq data from gRNA-transduced NUP98::KDM5A cbCD34 models at day 35, showing marker gene expression (**left**), annotated clusters (**mid**), and cell distributions among conditions (**right**). Colors in plots indicate relative expression levels, clusters, and cell density, respectively. **I**. Enrichment of cells with each cluster indicated by colors and sizes. **J**. Pseudotime along myeloid (HSC→ GMP→ monocytes) and platelet (HSC→ MEP→ MK→ platelet) trajectories. Colors represent pseudotime scores of each single cell inferred by Slingshot. **K**. *RB1* (**top**) and *CDKN2A* (**bottom**) expression along the pseudotime axis in each condition, with red curves showing average expressions. **L**. DEG analysis between the platelet-like and MK-like clusters in the AAVS-control condition (**left**) and GO term analysis (**right**) of genes high in the platelet-like cluster (red) and the MK-like cluster (blue). **M.** Schematics created with BioRender.com illustrating platelet differentiation in normal hematopoiesis and NUP98::KDM5A models. Assay data were obtained in technical triplicates from an established NUP98::KDM5A/Cas9 line and independent experiments. One data point in **C** was not obtained due to technical errors. RNAseq data were obtained from six independent experiments. In **B-D**, statistical tests were performed by linear mixed effect model (**B**) or two-sided Student's t-test by comparing day 4 and day 39 (**C**) or gRNA conditions and AAVS controls (**D**), and limma (**F**, **G**), followed by the Benjamini-Hochberg adjustment when applicable. DEGs in scRNAseq (**L**) were identified using the FindMarker function in the Seurat package with default settings, which calculates adjusted *P*-values with limma implementation of the Wilcoxon rank-sum test followed by Bonferroni correction. Asterisks indicating *P*-values or adjusted *P*-values <0.05. Error bars indicate mean ± s.e.m.

Figure 7. ***Cooperating alterations and differentiation status affect sensitivity to menin inhibition***

A. Experimental schema showing revumenib treatment. **B**. Relative cell growth of cbCD34 NUP98::KDM5A models with gRNA treated with revumenib (0.1-1μM) compared with DMSO controls. **C**. Flow gating (**top**), CD34+CD41a+ population (**left**), CD34-CD41a+ population (**mid**), and CD34-CD11B+ populations (**right**) among mCherry+ GFP+

mAmetrine+ live population at day 15. **D.** GSEA analyses between DMSO and revumenib-treated conditions, colors showing NES (**top-left**), comparison of expression changes between AAVS and *RB1*-gRNA1 conditions (**right**), and GO term analysis of changes enriched (difference of fold changes >1) in AAVS conditions (**bottom-left**). **E.** Relative cell growth of unedited cdCD34 NUP98::KDM5A (control, black), CHRF-288-11 (red), and ST1653 PDX (blue) treated with revumenib (0.1-1μM) compared with DMSO controls. **F.** Schematics illustrating a cellular hierarchy of NUP98::KDM5A models created with BioRender.com. Data was obtained from three technical replicates using gRNA-transduced cells in Fig.6. One data point at day 15 was excluded for technical errors. Statistical tests were performed by a generalized linear mixed effect model with Gaussian distribution, followed by the Benjamini-Hochberg adjustment (**B, E**) or Student's t-test by comparing gRNA conditions with AAVS controls (**C**). Asterisks indicating *P*-values or adjusted *P*-values *<0.05. Error bars indicate mean ± s.e.m.

Figure 1

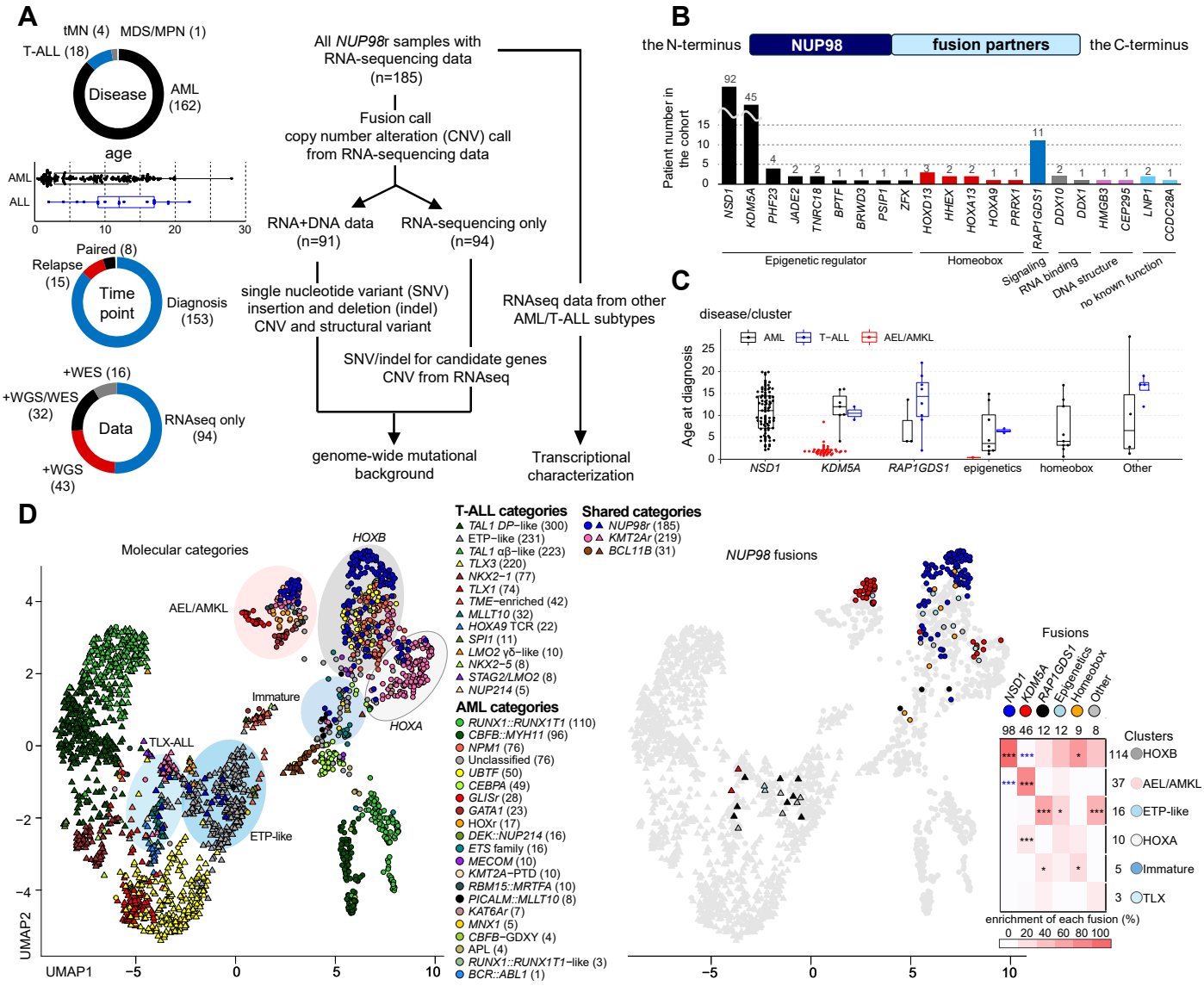


Fig.2

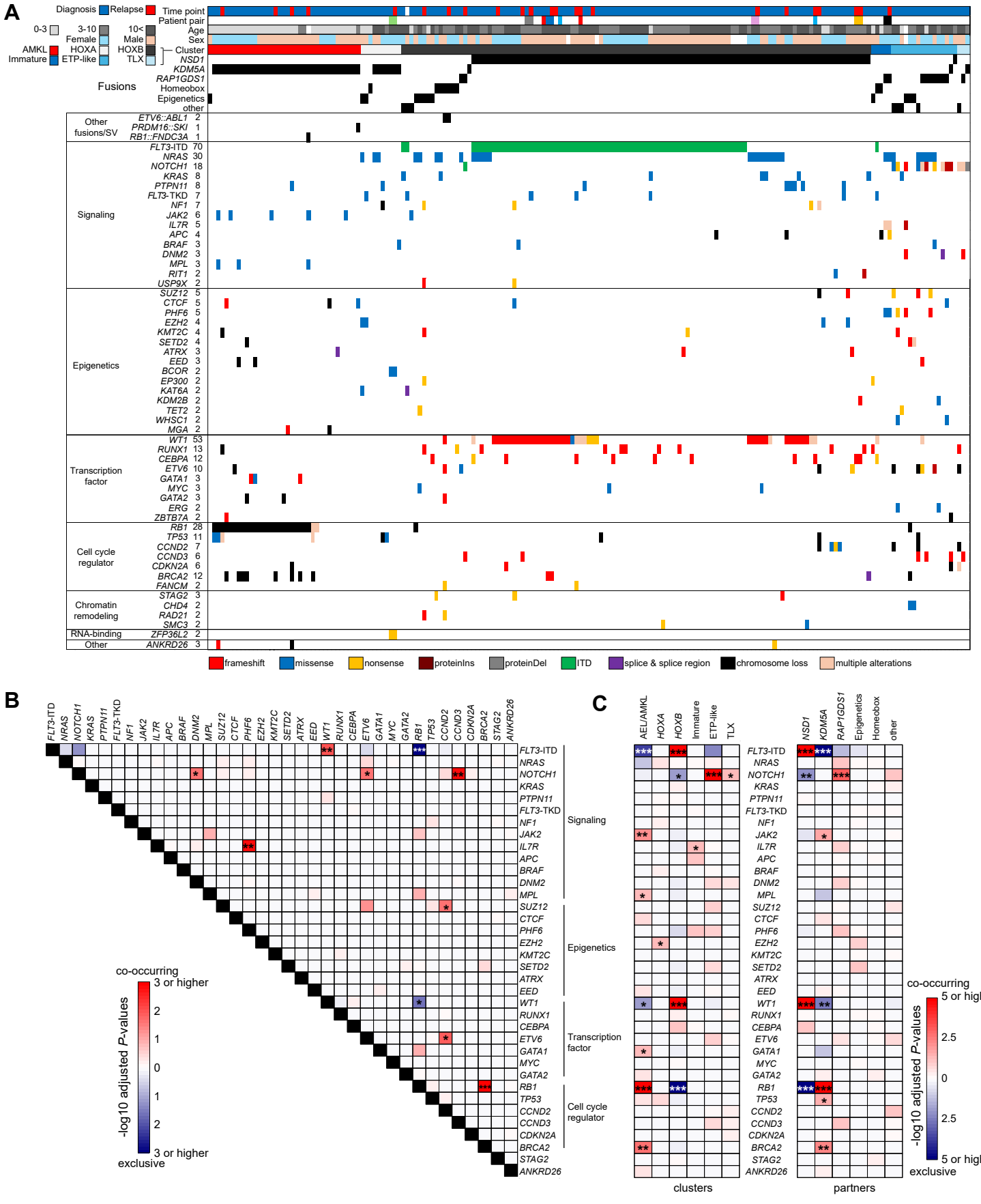


Fig.3

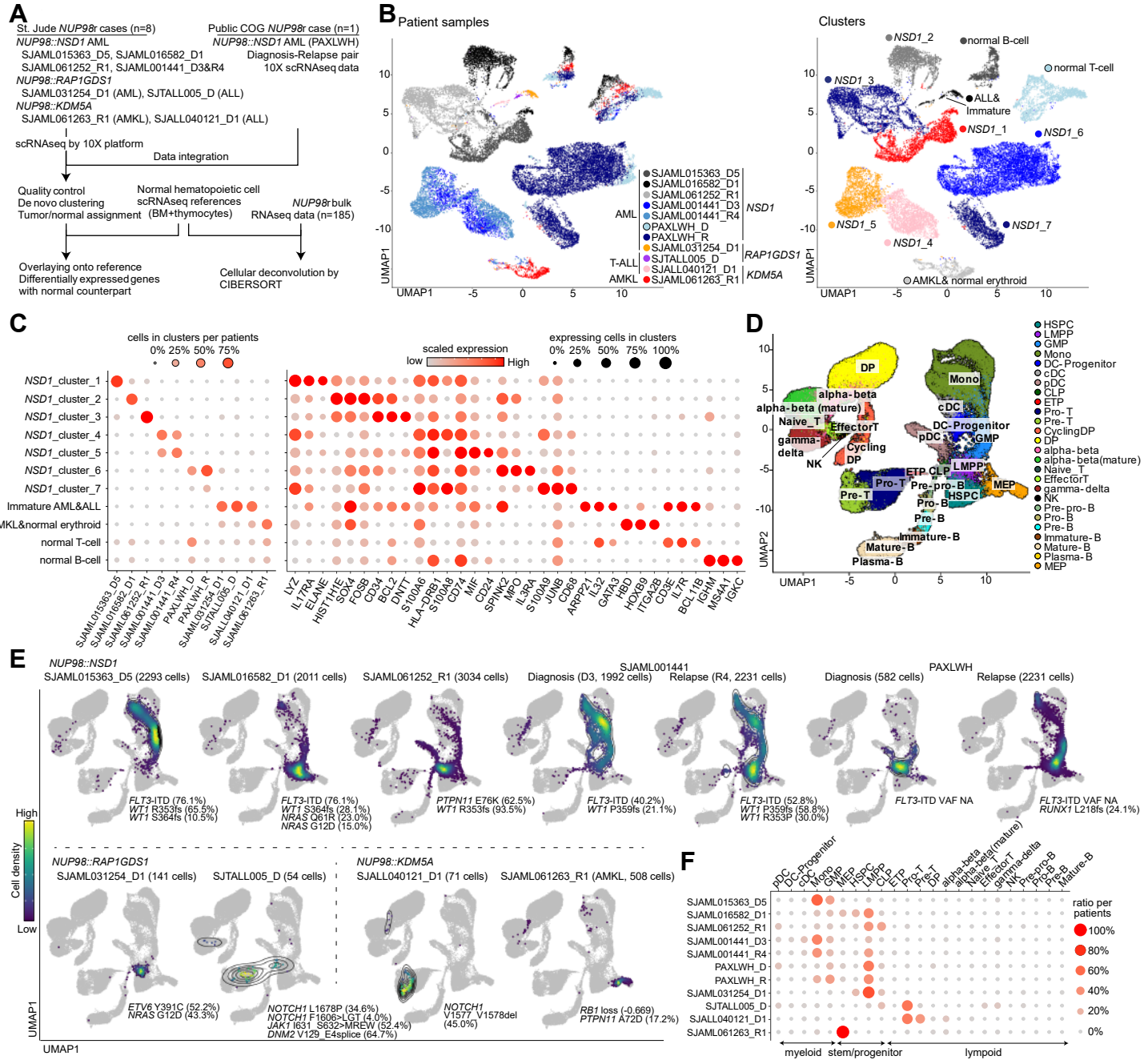


Fig.4

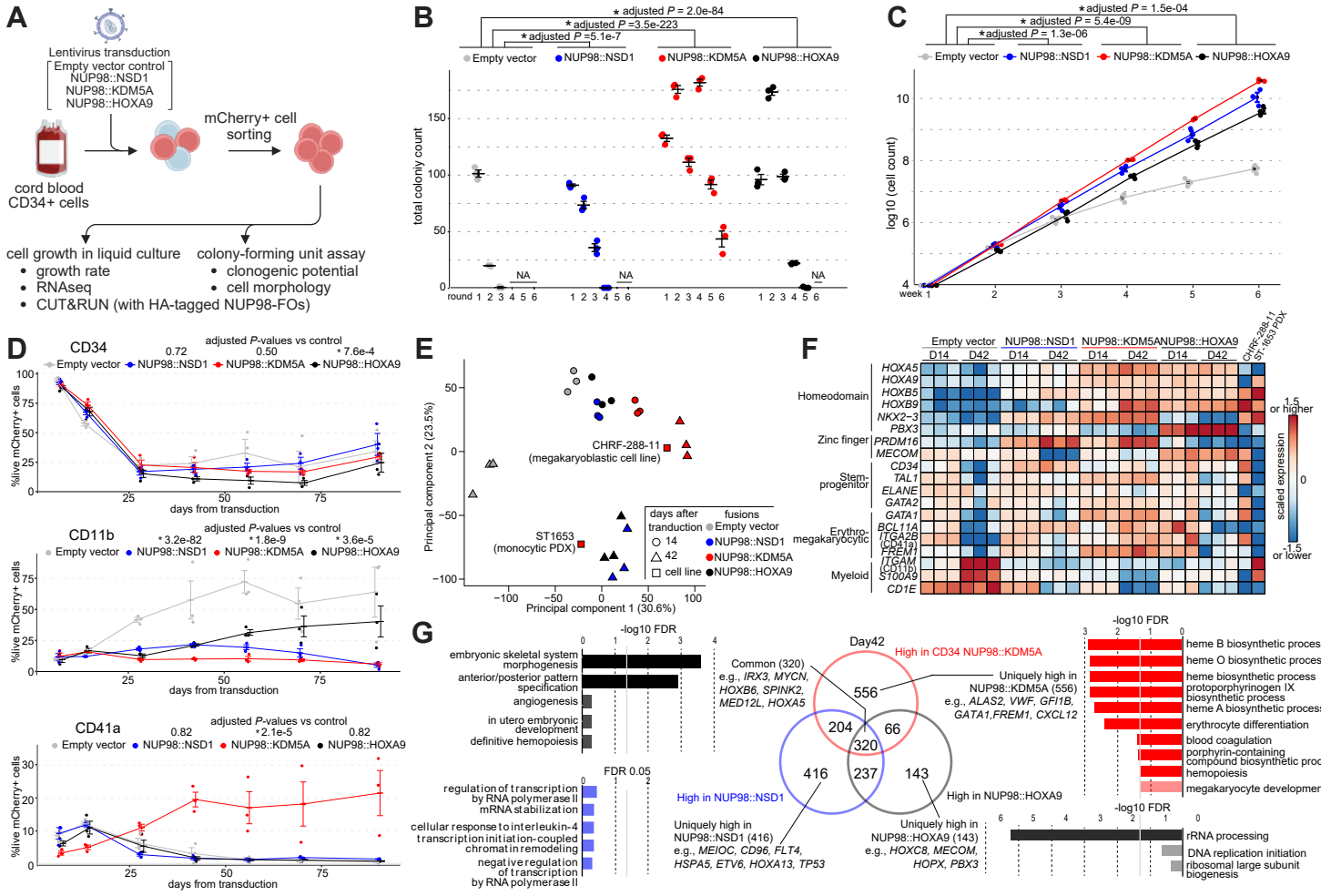


Fig.5

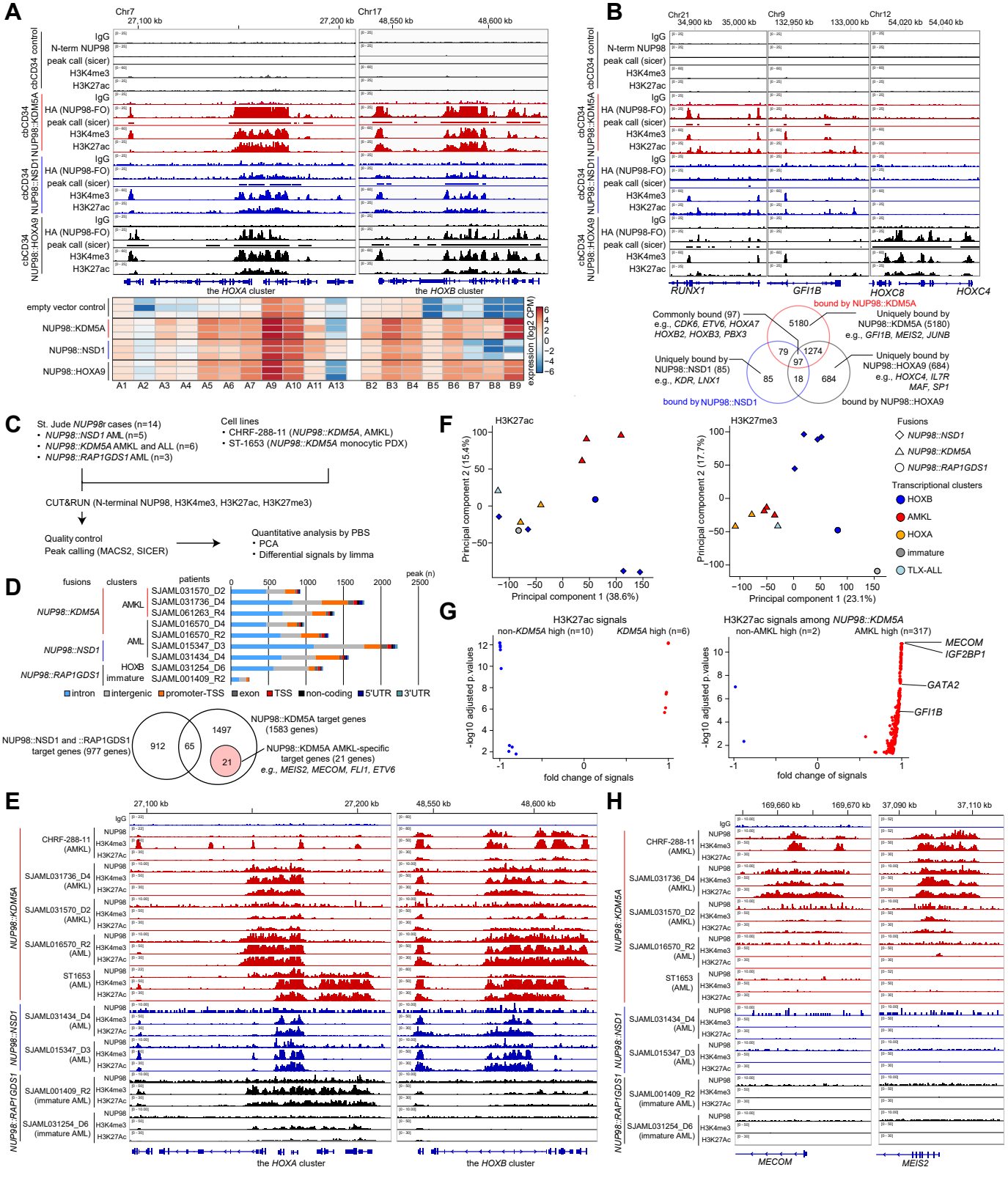


Fig.6

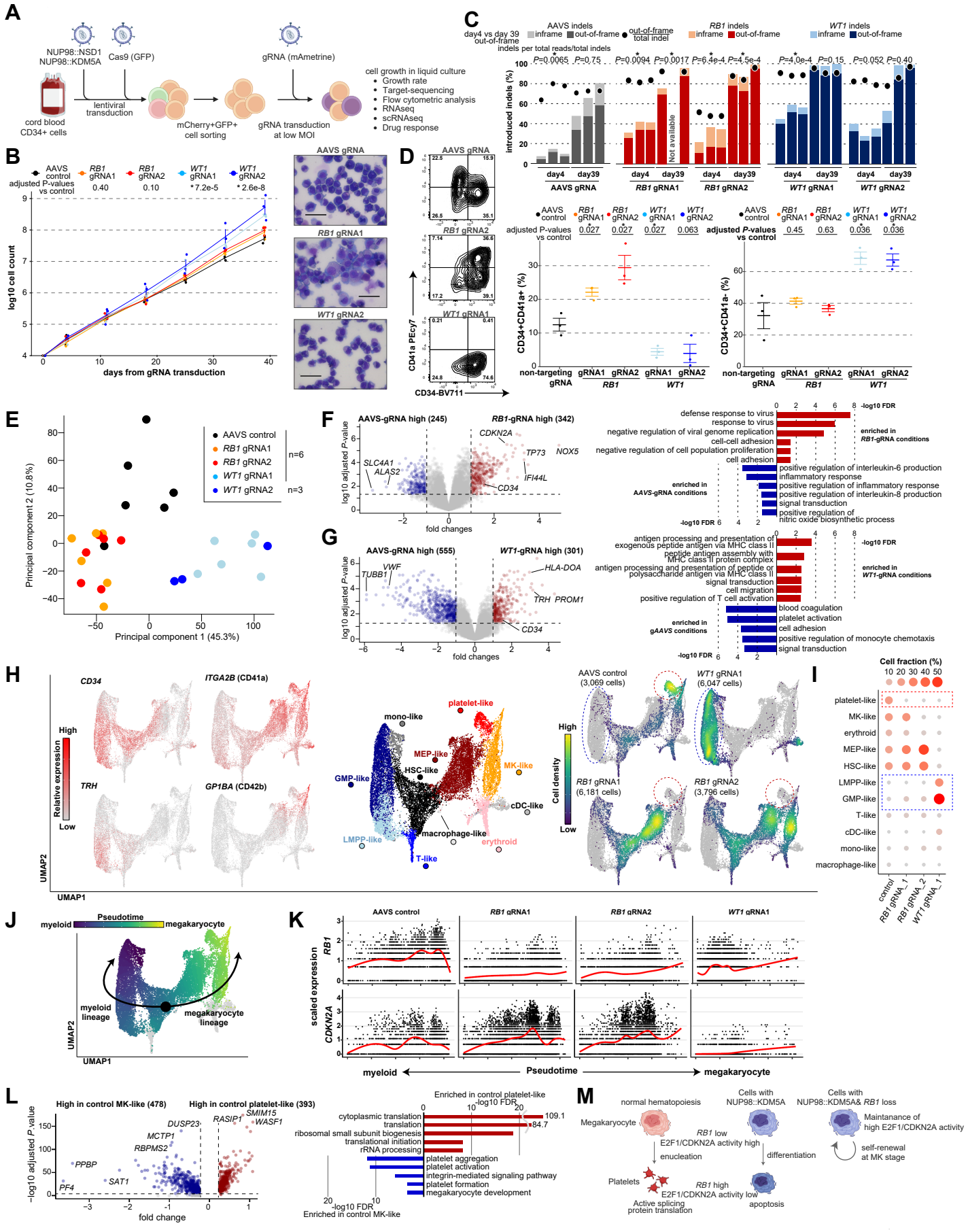
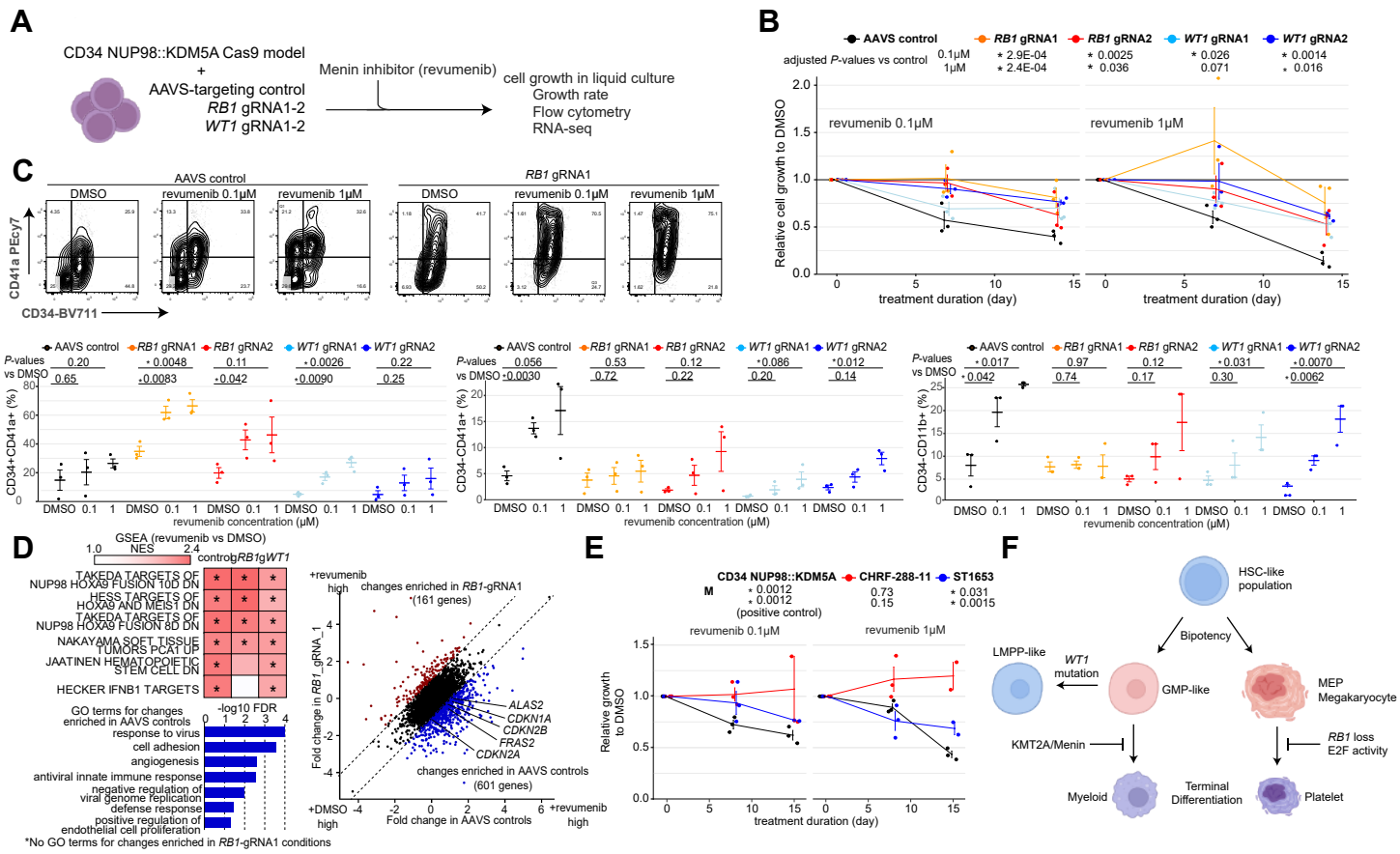


Fig.7



Contribution of Fusion Oncoproteins and Cooperating Mutations to Disease Phenotypes in *NUP98*-Rearranged Leukemia

Context of Research

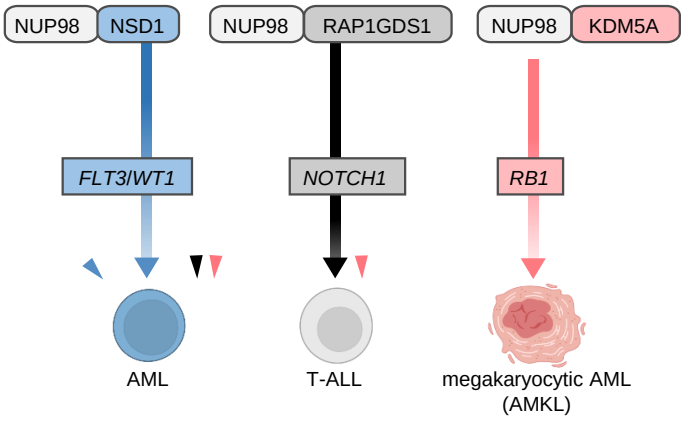
Various *NUP98* fusion partners show heterogeneous leukemia subtypes such as acute myeloid leukemia (AML) and T-cell acute lymphoblastic leukemia (T-ALL) with poor outcomes

Aim of This Study

To study the genomic background of pediatric leukemia with *NUP98* rearrangements (*NUP98*r) and its contribution to the disease phenotypes and therapeutic vulnerabilities

Findings

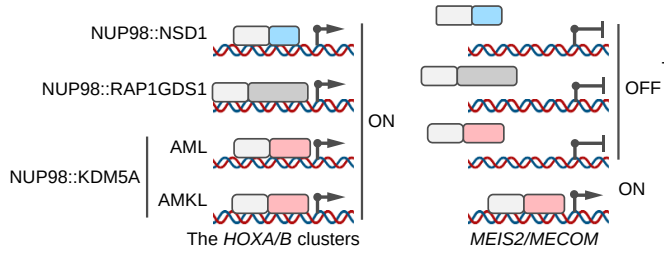
NUP98 fusion partners and cooperating alterations are associated with disease phenotypes



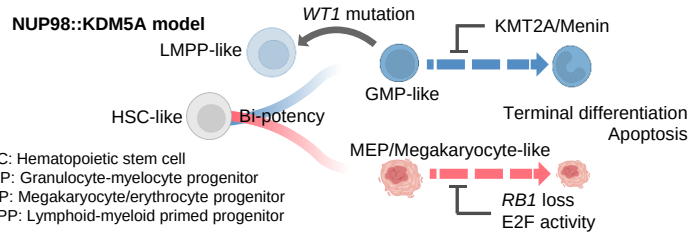
The same fusion partners can contribute to various phenotypes with unique cooperating mutations

Created by BioRender.com

Fusion and differentiation-associated gene regulation



Cooperating alterations act in a lineage-specific manner



Conclusions: 1) *NUP98* fusion partners and cooperating mutations are associated with specific disease phenotypes with unique cellular hierarchies. 2) NUP98 fusion oncoproteins bind to the genome, differentially regulating differentiation-related genes, whereas cooperating mutations also affect differentiation status in a lineage-specific manner.

*Umeda et al. DOI: 10.xxxx/**blood**.2024xxxxxx*