PIONEER
ACADEMIC PUBLISHING LIMITED

Current Research in Medical Sciences
ISSN 2958-0390
www.pioneerpublisher.com/crms
Volume 4 Number 5 September 2025

Differential Expression Analysis of RNA-Binding Proteins in Chronic Myeloid Leukemia Progression

Mengzhan Liu[1]

[1] Amador Valley High School, Pleasanton, 94566, US
Correspondence: Mengzhan Liu, Amador Valley High School, Pleasanton, 94566, US.

doi:10.56397/CRMS.2025.09.06

Abstract

Background: Chronic myeloid leukemia (CML) progresses from chronic phase (CP) to blast phase (BP) through poorly understood molecular mechanisms. Previous transcriptomic studies identified widespread splicing alterations in CML, yet systematic analysis of RNA-binding protein (RBP) expression patterns across disease stages remains limited. ***Methods:*** We performed computational analysis of RNA-seq data from peripheral blood mononuclear cells of 5 CP patients, 5 BP patients, and 5 healthy controls. Expression data for 540 RBPs were extracted and compared between groups using a significance threshold of $p < 0.05$. Results were visualized using scatter plots, bar graphs, and Venn diagrams. ***Results:*** Analysis identified 107 significantly dysregulated RBPs in CP versus controls and 61 in BP versus controls. Venn diagram analysis revealed only 39 RBPs (30.2%) were commonly dysregulated across both stages, with 68 RBPs unique to CP and 22 unique to BP. Key dysregulated RBPs included spliceosome components SF3B1 and U2AF1 in CP, and metabolism-associated factors HNRNPC and NPM1 in BP. ***Conclusions:*** RBP dysregulation occurs early in CML pathogenesis and undergoes substantial remodeling during disease progression. Stage-specific RBP expression patterns suggest distinct post-transcriptional regulatory mechanisms operate at different disease phases.

Keywords: chronic myeloid leukemia, RNA-binding proteins, differential expression, chronic phase, blast phase, post-transcriptional regulation

1. Introduction

Chronic myeloid leukemia (CML) is a hematologic malignancy characterized by the BCR-ABL fusion oncogene resulting from the Philadelphia chromosome translocation. The disease exhibits triphasic progression from chronic phase (CP) through accelerated phase (AP) to blast phase (BP). While tyrosine kinase inhibitors have dramatically improved outcomes for CP patients, progression to advanced phases remains associated with poor prognosis and therapeutic resistance, highlighting the need to understand the molecular mechanisms driving disease evolution.

Recent transcriptomic profiling identified extensive alternative splicing alterations in CML, with over 6,000 aberrant splicing events detected between CML patients and healthy controls (Wu Q, et al., 2020). Alternative splicing is regulated primarily by RNA-binding proteins (RBPs), which recognize specific sequence motifs in pre-mRNA and direct spliceosome assembly and splice site selection. Beyond splicing regulation, RBPs control multiple

aspects of RNA metabolism including transcript stability, subcellular localization, and translation efficiency.

RBPs represent particularly relevant targets for investigation in CML for two reasons. First, given that splicing is extensively disrupted in this disease (Wu Q, et al., 2020), the regulatory proteins controlling splicing may themselves be dysregulated. Second, individual RBPs typically regulate hundreds of target transcripts, such that alterations in RBP expression could produce widespread downstream effects on gene expression and cellular phenotype.

Previous studies identified differential splicing of spliceosome pathway genes between CML-BP and CML-CP patients (Wu Q, et al., 2020), suggesting potential RBP involvement. However, systematic analysis of RBP expression patterns across CML disease stages has not been performed.

Research Question: Which RBPs exhibit altered expression in CML, and do expression patterns differ between chronic phase and blast phase disease?

Objective: To identify significantly dysregulated RBPs in CP and BP compared to healthy controls using RNA-seq data, and to determine which RBPs show stage-specific versus pan-stage dysregulation.

2. Methods

2.1 Data Source and Patient Samples

RNA-seq data were obtained from Gene Expression Omnibus accession GSE100026 (Wu Q, et al., 2020). The dataset comprises gene expression profiles from peripheral blood mononuclear cells (PBMCs) of 5 chronic phase CML patients, 5 blast phase CML patients, and 5 healthy controls. Samples were collected with informed consent under approval from the medical ethics committee of the Second Affiliated Hospital of Nanchang University, China.

2.2 Original Study: RNA-seq Data Generation and Processing

The following methods were performed by the original investigators (Wu Q, et al., 2020) and are summarized here for context:

Library preparation: Total RNA was extracted from PBMCs using TRIzol Reagent. For each sample, 1 μg total RNA underwent poly(A) selection, fragmentation at 95°C, end repair, adaptor ligation, reverse transcription, and PCR amplification. Library products of 200-500 bp were size-selected for sequencing.

Sequencing and alignment: Libraries were sequenced on Illumina NextSeq 500 (150 bp paired-end reads). Raw reads were processed using FASTX-Toolkit (v0.0.13) for adapter trimming and quality filtering. Quality metrics were assessed using FastQC. Clean reads were aligned to the human reference genome (GRCh37/hg19) using TopHat2, and uniquely mapped reads were quantified as reads per kilobase per million mapped reads (RPKM).

Differential expression analysis: The original study used edgeR to identify differentially expressed genes with criteria of $|\log2 \text{ fold change}| > 1$ and false discovery rate (FDR) < 0.05.

2.3 Current Analysis: RBP Expression Profiling

RBP selection: A curated list of 540 known RNA-binding proteins was compiled from established RBP databases and literature. For detailed functional discussion, four RBPs were selected based on biological relevance to hematologic malignancies rather than magnitude of expression change: SF3B1 and U2AF1 (CP-dysregulated spliceosome components) and HNRNPC and NPM1 (BP-dysregulated factors with established roles in leukemia). This approach enables contextualization of findings within existing leukemia literature.

Data extraction and processing: Custom Python scripts were developed to:

1). Extract RPKM expression values for all 540 RBPs from processed RNA-seq data

2) Calculate mean expression for each RBP within each sample group (Control, CP, BP)

3) Perform statistical comparisons between CP versus Control and BP versus Control

Statistical analysis: Differential expression was assessed using Student's t-test with significance threshold of $p < 0.05$. Multiple testing correction was not applied in this exploratory analysis.

Visualization: Results were visualized using:

- Scatter plots (Figures 1A, 1B) displaying all 540 RBPs with significantly dysregulated RBPs highlighted

- Bar graphs (Figures 2A, 2B) quantifying numbers of significantly dysregulated RBPs

- Venn diagram (Figure 3) illustrating overlap of dysregulated RBPs between disease stages

3. Results

3.1 Global RBP Expression Patterns Across CML Disease Stages

Analysis of 540 RBPs revealed substantial dysregulation in both disease stages relative to controls. In chronic phase, 107 RBPs (19.8%) showed significant differential expression ($p < 0.05$), while blast phase exhibited 61 significantly dysregulated RBPs (11.3%). The greater number of altered RBPs in CP compared to BP, despite BP representing more advanced disease, suggests extensive post-transcriptional reprogramming occurs early in CML pathogenesis.

3.2 Visualization of Differential RBP Expression

Scatter plot analysis (Figures 1A and 1B): Scatter plots comparing average expression between disease groups and controls revealed bidirectional dysregulation, with RBPs showing both increased and decreased expression relative to controls. This pattern indicates that CML progression involves both gain and loss of specific RBP functions. The visual distribution of significantly dysregulated RBPs (highlighted in red) demonstrates greater density in the CP comparison (Figure 1A) compared to BP (Figure 1B), consistent with quantitative findings.

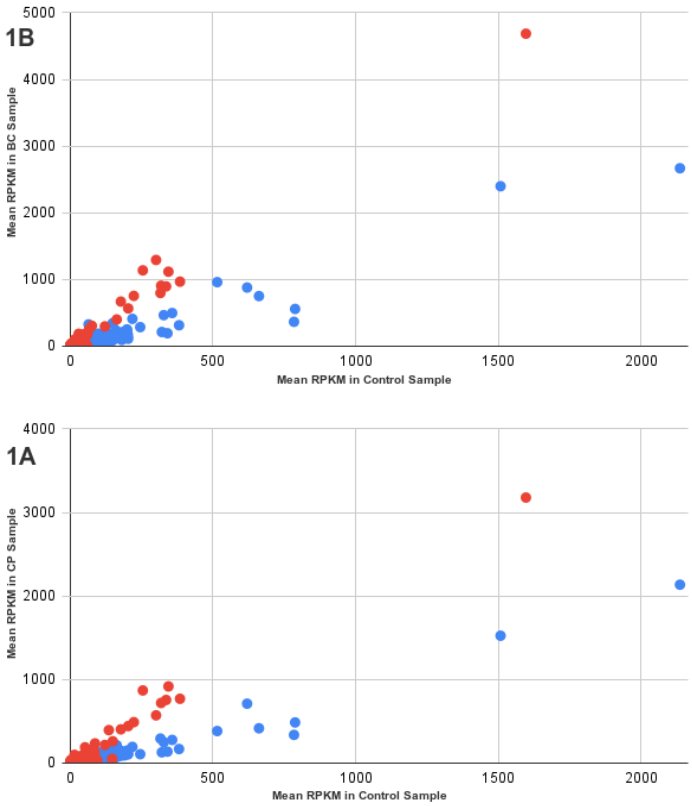

Figure 1. Scatter plot visualization of RBP expression in CML disease stages compared to healthy controls

(A) Chronic phase versus control comparison. Each point represents one RBP (n=540 total). Red points indicate significantly dysregulated RBPs ($p<0.05$); black points indicate non-significant RBPs.

(B) Blast phase versus control comparison. Layout and color scheme identical to panel A.

Quantification of dysregulated RBPs (Figures 2A and 2B): Bar graphs confirm that 107 of 540 RBPs (19.8%) are significantly altered in CP versus controls (Figure 2A), while 61 of 540 RBPs (11.3%) are significantly altered in BP versus controls (Figure 2B). These proportions indicate that approximately one-fifth of RBPs are dysregulated in CP, with proportionally fewer affected in BP.

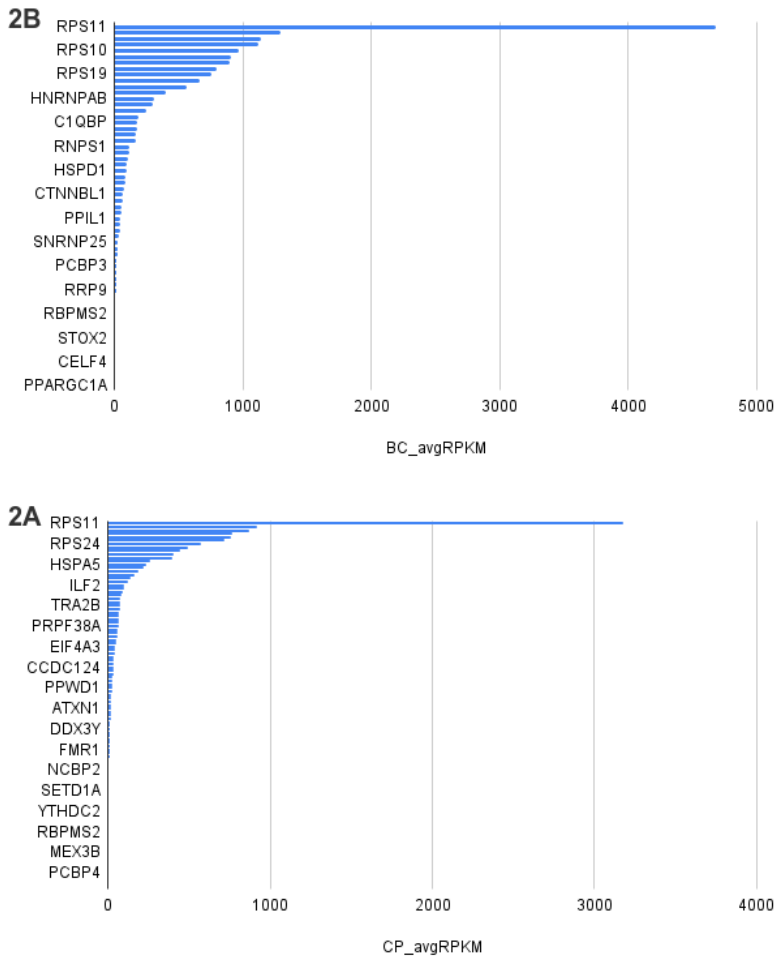

Figure 2. Quantification of significantly dysregulated RBPs by disease stage

(A) Bar graph showing number of significantly dysregulated RBPs in chronic phase versus control (107 of 540 RBPs, $p<0.05$).

(B) Bar graph showing number of significantly dysregulated RBPs in blast phase versus control (61 of 540 RBPs, $p<0.05$).

3.3 Stage-Specific and Shared RBP Dysregulation

Venn diagram analysis (Figure 3): Comparison of dysregulated RBPs between CP and BP revealed limited overlap between disease stages:

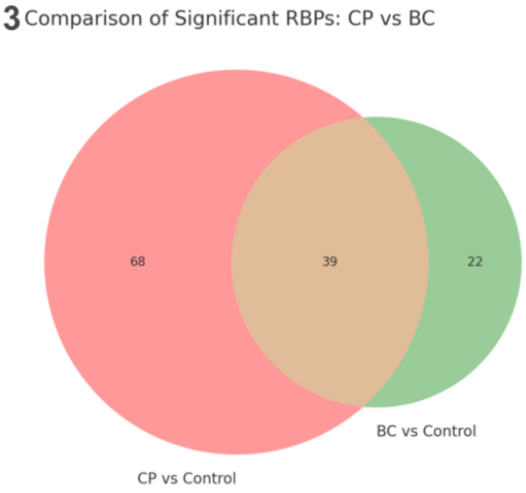

Figure 3. Venn diagram analysis of RBP dysregulation overlap between disease stages

Venn diagram comparing significantly dysregulated RBPs ($p<0.05$) in CP versus control (left circle) and BP versus control (right circle). Left-only region: 68 RBPs unique to chronic phase. Overlap region: 39 RBPs dysregulated in both stages. Right-only region: 22 RBPs unique to blast phase. Total: 129 distinct dysregulated RBPs across both comparisons.

- **39 RBPs (30.2% of total significant RBPs)** were dysregulated in both CP and BP, representing a core set of pan-stage alterations. This group includes multiple ribosomal proteins (RPS2, RPS3, RPS5, RPS10, RPS11, RPS19, RPS24, RPLP0), spliceosome components (SNRPB, SNRPC, SNRPD2, SNRPF, SNRPG, SNRNP25), and other factors such as ALYREF, MAGOH, and HNRNPAB.

- **68 RBPs (52.7% of total significant RBPs)** were uniquely dysregulated in CP, including core spliceosome factors SF3B1 and U2AF1, as well as MBNL2, MBNL3, and numerous splicing regulators. This CP-specific signature suggests extensive restructuring of splicing machinery occurs during chronic phase.

- **22 RBPs (17.1% of total significant RBPs)** were uniquely dysregulated in BP, including HNRNPC, NPM1, MSI2, and DKC1. These BP-specific alterations represent late-emerging changes potentially associated with blast transformation.

The limited overlap (30.2%) indicates substantial remodeling of the RBP landscape during CML progression, with different post-transcriptional regulatory programs predominating at different disease stages.

3.4 Functionally Relevant RBPs in Chronic Phase

SF3B1 (Splicing Factor 3b Subunit 1): SF3B1 is a core component of the U2 small nuclear ribonucleoprotein complex essential for spliceosome catalytic activity. SF3B1 mutations occur in 20-30% of myelodysplastic syndrome (MDS) cases and define a distinct disease subtype characterized by ring sideroblasts (Malcovati L, Stevenson K, Papaemmanuil E, et al., 2020; Papaemmanuil E, Cazzola M, Boultwood J, et al., 2011). These mutations alter branch point recognition, leading to aberrant 3′ splice site selection (Pellagatti A & Boultwood J., 2021). The observed alteration in SF3B1 expression (rather than mutation) in CP CML suggests that quantitative changes in spliceosome component abundance may contribute to splicing dysregulation independent of mutational mechanisms.

U2AF1 (U2 Small Nuclear RNA Auxiliary Factor 1): U2AF1 mediates 3′ splice site recognition through sequence-specific binding to the AG dinucleotide at exon-intron boundaries. Recurrent U2AF1 mutations occur in approximately 11% of MDS patients and alter splice site recognition patterns (Ilagan JO, Ramakrishnan A, Hayes B, et al., 2015; Shirai CL, White BS, Tripathi M, et al., 2015). Mouse models demonstrate that mutant U2AF1 impairs hematopoiesis and promotes leukemogenesis (Fei DL, Zhen T, Durham B, et al., 2018). Wild-type U2AF1 is required for hematopoietic stem cell survival and function (Yoshida K & Ogawa S., 2021). Dysregulation of U2AF1 expression in CP may therefore affect splice site selection and hematopoietic differentiation through mechanisms paralleling those of U2AF1 mutations in MDS.

3.5 Functionally Relevant RBPs in Blast Phase

HNRNPC (Heterogeneous Nuclear Ribonucleoprotein C): HNRNPC belongs to the heterogeneous nuclear ribonucleoprotein family that complexes with nascent pre-mRNA to regulate processing. Recent studies identified HNRNPC as a driver of metabolic reprogramming in drug-resistant acute myeloid leukemia (AML) through the HNRNPC/CELF2 pathway (Chen Y, Zhang L, Wang Q, et al., 2025). Additionally, HNRNPC functions as an m6A reader coordinating oncogenic transcription and metabolism in T-cell acute lymphoblastic leukemia (Ntziachristos P, Glytsou C, Kloetgen A, et al., 2025). The dysregulation of HNRNPC in BP suggests potential involvement in metabolic adaptation and therapeutic resistance during blast transformation.

NPM1 (Nucleophosmin 1): NPM1 is a nucleolar phosphoprotein involved in ribosome biogenesis, centrosome duplication, and regulation of the ARF-p53 tumor suppressor pathway. NPM1 represents the most frequently mutated gene in AML, with mutations detected in approximately 30% of adult cases (Falini B, Brunetti L, Sportoletti P & Martelli MP., 2020; Falini B, Mecucci C, Tiacci E, et al., 2005). NPM1 mutations cause aberrant cytoplasmic localization and are associated with distinct clinical features (Heath EM, Chan SM, Minden MD, et al., 2017; Falini B, Mecucci C, Tiacci E, et al., 2005). While BP samples likely harbor altered

NPM1 expression rather than mutations, quantitative changes in NPM1 levels may nonetheless affect ribosome production, genomic stability, and cell cycle regulation. The identification of NPM1 dysregulation in BP supports molecular convergence between blast crisis CML and de novo AML.

4. Discussion

4.1 RBP Dysregulation as an Early Event in CML Pathogenesis

This analysis demonstrates that RBP dysregulation is extensive in CML, with approximately 20% of examined RBPs showing significant expression changes in chronic phase. This finding indicates that post-transcriptional regulatory alterations are not secondary consequences of advanced disease but rather represent early events in CML development.

The extent of RBP dysregulation is consistent with prior observations of widespread splicing alterations in CML (Wu Q, et al., 2020). RBPs function as master regulators of splicing, and alterations in RBP expression would be expected to produce downstream effects on splicing patterns. The identification of core spliceosome components (SF3B1, U2AF1) among dysregulated RBPs provides a mechanistic basis for the observed splicing abnormalities.

4.2 Stage-Specific RBP Expression Patterns

The observation that CP exhibits more dysregulated RBPs (107) than BP (61) initially appears counterintuitive given that BP represents advanced disease. Several factors may contribute to this pattern.

First, transformation from normal hematopoiesis to chronic phase CML may require extensive reprogramming of RNA processing machinery. The large number of dysregulated RBPs in CP, including numerous spliceosome components, supports this interpretation. Establishment of the leukemic phenotype appears to involve fundamental restructuring of post-transcriptional regulatory networks.

Second, while fewer RBPs are dysregulated in BP overall, the specific RBPs altered in this stage may be particularly significant for blast transformation. The 22 BP-specific RBPs include factors with established roles in leukemia metabolism (HNRNPC) and proliferation (NPM1). These targeted alterations may be sufficient to drive blast crisis without requiring the extensive RBP remodeling observed in CP.

Third, the limited overlap between CP and BP dysregulated RBPs (30.2%) indicates stage-specific regulatory programs. CP dysregulation centers on core splicing machinery, while BP alterations involve metabolism and proliferation-associated factors. This shift suggests that different aspects of post-transcriptional control become rate-limiting at different disease stages.

4.3 Connections to Established Hematologic Malignancy Biology

The dysregulated RBPs identified here show notable parallels to genetic alterations in related myeloid malignancies. SF3B1 and U2AF1 mutations are recurrent drivers in MDS (Malcovati L, Stevenson K, Papaemmanuil E, et al., 2020; Papaemmanuil E, Cazzola M, Boultwood J, et al., 2011; Ilagan JO, Ramakrishnan A, Hayes B, et al., 2015; Shirai CL, White BS, Tripathi M, et al., 2015), raising the possibility that altered expression of wild-type proteins produces similar functional consequences. This concept—that quantitative changes in splicing factor abundance can phenocopy mutational effects—warrants further investigation.

HNRNPC dysregulation in BP parallels its role in AML metabolic reprogramming (Chen Y, Zhang L, Wang Q, et al., 2025; Ntziachristos P, Glytsou C, Kloetgen A, et al., 2025). Blast transformation in CML shares phenotypic features with de novo AML, including rapid proliferation and treatment resistance. The identification of HNRNPC as a BP-specific alteration suggests that common metabolic adaptations may underlie the aggressive behavior of both diseases.

NPM1 dysregulation in BP further supports molecular convergence between blast crisis and AML (Falini B, Brunetti L, Sportoletti P & Martelli MP., 2020; Heath EM, Chan SM, Minden MD, et al., 2017; Falini B, Mecucci C, Tiacci E, et al., 2005). While NPM1 alterations in BP likely involve expression changes rather than the cytoplasmic-localizing mutations characteristic of AML, both mechanisms may disrupt nucleolar function and genomic stability.

4.4 Therapeutic Implications

The identification of stage-specific RBP dysregulation suggests potential therapeutic approaches. For chronic phase, targeting core spliceosome components such as SF3B1 may prevent or delay disease progression. Small

molecule splicing modulators have shown activity in other hematologic malignancies (Malcovati L, Stevenson K, Papaemmanuil E, et al., 2020; Pellagatti A & Boultwood J., 2021) and represent a potential therapeutic avenue for CML.

For blast phase, targeting metabolism-associated RBPs such as HNRNPC may overcome therapeutic resistance. HNRNPC drives glycolytic reprogramming in drug-resistant AML (Chen Y, Zhang L, Wang Q, et al., 2025), suggesting that HNRNPC inhibition could restore treatment sensitivity in BP CML.

The stage-specific nature of RBP dysregulation implies that optimal therapeutic strategies may differ between disease phases, with splicing-targeted approaches potentially more relevant for CP and metabolism-targeted approaches for BP.

4.5 Study Limitations and Future Directions

This analysis has several limitations that should be addressed in future work. First, the sample size (n=5 per group) limits statistical power and increases the risk of false positive findings. Future analyses should incorporate additional datasets from public repositories to increase sample size and improve robustness.

Second, the significance threshold ($p < 0.05$ without multiple testing correction) is not stringent given the number of comparisons performed. Application of false discovery rate correction (FDR < 0.05) would provide more conservative and reliable identification of dysregulated RBPs.

Third, the dataset lacks accelerated phase samples. CML progression is classically triphasic (CP → AP → BP), and the absence of AP data prevents determination of whether RBP changes occur gradually or in discrete transitions. Identification or generation of AP samples would enable more complete characterization of RBP dynamics during disease evolution.

Fourth, this analysis examined only steady-state expression levels. RBP function is also regulated by subcellular localization, post-translational modifications, and protein-protein interactions. Future studies should incorporate these regulatory layers.

Fifth, these findings require experimental validation. Functional studies in CML cell lines or patient samples are needed to establish causal relationships between RBP dysregulation and disease phenotypes.

Specific improvements for future analyses include: (1) aggregating multiple CML RNA-seq datasets to increase statistical power, (2) applying FDR correction for multiple testing, (3) incorporating fold change thresholds in addition to p-value cutoffs, (4) performing pathway enrichment analysis on dysregulated RBP sets, (5) integrating RBP expression data with splicing event data to identify RBP-splicing relationships, and (6) validating findings through RBP knockdown or overexpression experiments in cellular models.

References

Brunetti L, Gundry MC, Goodell MA. (2017). DNMT3A in leukemia. *Cold Spring Harb Perspect Med., 7*(2), a030320.

Chen Y, Zhang L, Wang Q, et al. (2025). The HNRNPC/CELF2 signaling pathway drives glycolytic reprogramming and mitochondrial dysfunction in drug-resistant acute myeloid leukemia. *Cell Biosci., 15*(1), 25.

Falini B, Brunetti L, Sportoletti P, Martelli MP. (2020). NPM1-mutated acute myeloid leukemia: from bench to bedside. *Blood., 136*(15), 1707-1721.

Falini B, Mecucci C, Tiacci E, et al. (2005). Cytoplasmic nucleophosmin in acute myelogenous leukemia with a normal karyotype. *N Engl J Med., 352*(3), 254-266.

Fei DL, Zhen T, Durham B, et al. (2018). Impaired hematopoiesis and leukemia development in mice with a conditional knock-in allele of a mutant splicing factor gene U2af1. *Proc Natl Acad Sci USA., 115*(44), E10437-E10446.

Heath EM, Chan SM, Minden MD, et al. (2017). Biological and clinical consequences of NPM1 mutations in AML. *Leukemia., 31*(4), 798-807.

Ilagan JO, Ramakrishnan A, Hayes B, et al. (2015). U2AF1 mutations alter splice site recognition in hematological malignancies. *Genome Res., 25*(1), 14-26.

Malcovati L, Stevenson K, Papaemmanuil E, et al. (2020). SF3B1-mutant MDS as a distinct disease subtype: a proposal from the International Working Group for the Prognosis of MDS. *Blood., 136*(2), 157-170.

Ntziachristos P, Glytsou C, Kloetgen A, et al.

(2025). HNRNPC and m6A RNA methylation control oncogenic transcription and metabolism in T-cell leukemia. *Blood., 145*(16), 1699-1713.

Papaemmanuil E, Cazzola M, Boultwood J, et al. (2011). Somatic SF3B1 mutation in myelodysplasia with ring sideroblasts. *N Engl J Med., 365*(15), 1384-1395.

Pellagatti A, Boultwood J. (2021). SF3B1 mutant myelodysplastic syndrome: Recent advances. *Adv Biol Regul., 79*, 100776.

Shirai CL, White BS, Tripathi M, et al. (2015). Mutant U2AF1 expression alters hematopoiesis and pre-mRNA splicing in vivo. *Cancer Cell., 27*(5), 631-643.

Wu Q, et al. (2020). Transcriptome profiling reveals the high incidence of hnRNPA1 exon 8 inclusion in chronic myeloid leukemia. *J Adv Res., 24*, 375-387.

Yoshida K, Ogawa S. (2021). U2af1 is required for survival and function of hematopoietic stem/progenitor cells. *Leukemia., 35*(1), 76-86.