IJMCM
Summer 2021, Vol 10, No 3
DOI: 10.22088/IJMCM.BUMS.10.3.227

Short Communication

Identification of Novel Insertions and Deletions in Haematopoietic Stem/Progenitor Cells in de novo Myelodysplastic Syndromes

Manoj Bandara[1], Hemali Goonasekera[2], Vajira Dissanayake[2*]

1. Department of Pre-Clinical Sciences, Faculty of Medicine, General Sir John Kotelawala Defence University, Rathmalana, Sri Lanka.

2. Human Genetics Unit, Faculty of Medicine, University of Colombo, Colombo 8, Sri Lanka.

Submmited 3 August 2021; Accepted 11 December 2021; Published 10 January 2022

Myelodysplastic Syndromes (MDS) are clonal haematological stem cell disorders. The molecular basis of MDS is heterogeneous and the molecular mechanisms underlying biology of this complex disorder are not fully understood. Genetic variations (GVs) occur in about 90% of patients with MDS. It has been shown that in addition to the single nucleotide variations, insertions and deletions (indels) in the key genes that are known to drive MDS, could also play a role in pathogenesis of MDS. However, only a few genetic studies have analyzed indels in MDS. The present study reports indels of bone marrow (BM) derived $CD34^+$ haematopoietic stem/progenitor cells of 20 newly diagnosed *de novo* MDS patients using next generation sequencing.A total of 88 indels (9 insertions and 79 deletions) across 28 genes were observed. The genes that showed more than five indels are *BCOR* (N=6), *RAD21* (N=6), *TP53* (N=8), *ASXL1* (N=9), *TET2* (N=9) and *BCORL1* (N=10). Deletion in the *BCORL1* gene (c.3957_3959delGGA, TGAG>TGAG/T) was the most recurrent deletion and was observed in 4/20 patients. The other recurrent deletions reported were *EZH2* (W15X, N=2) and *RAD21* (G274X, N=3). The recurrent insertions were detected in the *FLT3* (E598DYVDFREYE, N=3) and in the *NPM1* (L287LCX, N=3) genes. The findings of this study may have a diagnostic, prognostic and a therapeutic value for MDS after validation using a larger cohort.

Key words: Insertions and deletions, myelodysplastic syndromes, haematopoietic stem and progenitor cells, next generation sequencing

Myelodysplastic syndromes (MDS) are clonal haematopoietic stem cell disorders characterized by ineffective hematopoiesis, bone marrow (BM) dysplasia, and peripheral cytopenias with a risk of transforming into acute myeloid leukemia (AML) (1). MDS originates from a malignant transformation of a haematopoietic stem cell (HSC), which shows growth advantage over a normal HSC and its clonal expansion (2). The molecular basis of MDS is heterogeneous and the

*Corresponding author: Faculty of Medicine, University of Colombo, Colombo 8, Sri Lanka.
Email: vajira@anat.cmb.ac.lk

This work is published as an open access article distributed under the terms of the Creative Commons Attribution 4.0 License (http://creativecommons.org/licenses/by-nc/4). Non-commercial uses of the work are permitted, provided the original work is properly cited.

[Downloaded from ijmcmed.org on 2026-11-07]

[DOI: 10.22088/IJMCM.BUMS.10.3.227]

molecular mechanisms underlying pathobiology of this complex disorder is still being understood. About 90% of MDS patients show genetic variations (GVs) in the genes known to be associated with the development of this disorder (3). However, most of the genetic studies on MDS report single nucleotide variations (SNVs) (4, 5) and rarely report insertions and deletions (indels) (6-8). The reported indels in the literature have not been characterized in detail. Nonetheless, investigations have revealed frequent indels in AML (9- 11). Identification of indels in genes that are associated with MDS may provide new insights to understand the molecular basis of the disease to facilitate diagnosis, prognostication, and treatment options. This study investigated the indels of BM derived $CD34^+$ haematopoietic stem/progenitor cells (HSPCs) of MDS patients using a targeted next generation sequencing (NGS).

Materials and methods

Patients

Bone marrow samples of newly diagnosed patients (n=20) with *de novo* MDS were collected from four tertiary care hospitals in Sri Lanka (National Hospital of Sri Lanka, Colombo South Teaching Hospital, Colombo North Teaching Hospital and National Cancer Institute, Maharagama) after obtaining the written informed consent. The study was conducted according to the principles of Declaration of Helsinki (2008). Ethical approvals for the study were obtained from the Ethics Review Committee of the Faculty of Medicine, University of Colombo, Sri Lanka and respective hospitals. Patients with secondary MDS were excluded. Patient's clinical findings and findings of the investigations were recorded. All patients were subtyped according to the WHO classification (12).

Isolation and culture expansion of $CD34^+$ HSPCs

Mononuclear cells were purified from the BM samples using Histopaque (Sigma-Aldrich, Gillingham, United Kingdom) by density gradient centrifugation and were labeled with CD34 MicroBeads (EasySep CD34 selection kit, Stem Cell Technologies, Canada). $CD34^+$ cells were isolated using magnetic cell separation. Flow cytometry (FACScan; Becton Dickinson, Heidelberg, Germany) was used to evaluate the purity of $CD34^+$ cell population. Purified $CD34^+$ cells were cultured at a concentration of 1.0×10^4 cells/ml in Stemline® Haematopoietic Stem Cell Expansion Medium (SIGMA) supplemented with stem cell factor (50 ng/ml), human Flt-3 ligand (20 ng/ml), thrombopoietin (20 ng/ml), interleukin 6 (50 ng/ml) and antibiotics (100 µg/ml). Cultures were maintained in 25 cm^2 tissue culture flasks at 37 °C and 5% CO_2, and cells were expanded for 10 days.

Next generation sequencing

The TruSight Myeloid Sequencing Panel [TMSP (Illumina, San Diego, USA)], which examines mutational hotspots in 54 frequently mutated genes in haematological malignancies, was used to detect variations in the genome of HSPCs. This panel targets 54 genes [15 full genes (exons only) and exonic hotspots of additional 39 genes] (13). DNA sequencing and data analysis were carried out as described previously (14). Briefly, DNA was extracted from HSPCs using QIAamp® DNA Mini kit (Quaigen, USA) and was quantified by "Qubit" fluorometer. DNA libraries were prepared according to the manufacturer's instructions. Following hybridization of oligos to the target regions, an extension and ligation reaction was performed to combine the oligo pairs across the regions of interest. DNA templates were PCR amplified using index primers. Amplified samples were assessed by running an aliquot of amplified DNA (5 µL) on a 4% agarose gel (amplicon size ≈ 250bp). Finally, libraries were purified using AMPure magnetic beads. The purified libraries were then normalized, quantified and pooled. Pooled libraries with unique barcodes

[Downloaded from ijmcmed.org on 2026-11-07]

[DOI: 10.22088/IJMCM.BUMS.10.3.227]

were diluted with chilled HT1 buffer to a final concentration of 12pM and were loaded on a MiSeq Reagent Kit v3 to run on a MiSeq benchtop sequencer (Illumina MiSeq System). Paired end sequencing (2 X 151bp) was conducted according to the default parameters of the Illumina MiSeq System.

Variant calling and data analysis

Genetic variants were identified by using two independent bioinformatic pipelines: in-built (Illumina) and in-house. In the in-built method, alignment and variant calling were done using the TruSeq Amplicon App®(Base space®, Illumina) in somatic mode. Somatic variants were also called by an in-house bioinformatic pipeline, MuTect (version 1.1.7) (15). First, the raw sequencing data obtained from the FASTQ were aligned to GrCh37 human genome assembly using Burrows-Wheeler Aligner [BWA (BWA-0.7.12]-mem algorithm. Then the Genome Analysis Tool Kit (GATK-v.3.6) was used to recalibrate the aligned reads. The Binary Alignment Map (BAM) files of MDS patients were co-realigned in pairs with that of a normal healthy Sri Lankan to increase the specificity of MuTect in retaining causal variants. To eliminate population unique variants, a "panel of normals" (PON) was also created by running MuTect on a set of normals. Input arguments (PON and the control BAM files) were run in MuTect on patient samples. The variants common to both pipelines were considered as "True' variants. Illumina annotation and filtering tool, Variant Studio 3.0® was used to annotate the generated variant caller files. The parameters used in the analysis; filter-pass, quality>30, read depth>250X, minor allele frequency <0.01, alternate allele frequencies >5%, read depth >500X and the alternate allelic depth >20X (16). Catalogue of Somatic Variants in Cancer (COSMIC; http:// cancer. sanger. ac.uk/cosmic) database was used for cross referencing the variants. The analysis was done by two individuals independently. Sanger sequencing was conducted to validate the methodology as described previously (14). In order to predict the potential functional significance, the indels identified by both pipelines were further analyzed with Mutation Taster (www. Mutatio-ntaster.org/).

Results

Patient characteristics

There were 7 males and 13 females in the study population with mean age of 64.5 years (range 31-75 years). Patient's subtypes, their blood and BM characteristics and karyotypes have been previously published (17).

Distribution of indels within the genes

The patient cohort carried 88 indels across 28 genes. These variants included 79 deletions and 9 insertions. HSPCs harbored indels in genes associated with DNA methylation (*DNMT3A* and *TET2*), chromatin remodeling (*ASXL1*, *ATRX*, *EZH2*, *KDM6A*, *PHF6*, *FLT3*, and *IKZF1*), transcription regulation (*ETV6*, *BCOR*, *BCORL1*, *RUNX1*, and *CUX1)*, cohesion complex (*SMC1A*, *SMC3*, *STAG2*, and *RAD21*), spliceosome machinery (*U2AF1* and *ZRSR2*), signal transduction (*NOTCH1*, *KRAS*, and *KIT*) and other cellular pathways (*FBXW7*, *PTEN*, and *NPM1*). The genes that showed more than 5 indels included *BCOR* (N=6), *RAD21* (N=6), *TP53* (N=8), *ASXL1* (N=9), *TET2* (N=9), and *BCORL1* (N=10) (Figure 1). Out of these 88 indels, 15 have been previously reported by other research groups (Table 1). The indels per patient in refractory anaemia with excess blasts (RAEB) (N=11) was higher than that of refractory cytopenia with unilineage dysplasia (RCUD) (N=4) and refractory cytopenia with multilineage dysplasia (RCMD) (N=2).

Recurrent indels

Three deletions and two insertions were identified as recurrent indels in our patient cohort (Table 2). The recurrent deletions were W15X in the *EZH2* gene (NM_004456.4: c.43delT),

[Downloaded from ijmcmed.org on 2026-11-07]

[DOI: 10.22088/IJMCM.BUMS.10.3.227]

G274X in the *RAD21* gene (NM_006265.2:c.822delG), and TGAG>TGAG/T in the *BCORL1* gene (NM_ 021946.4: c.3957_ 3959delGGA). The recurrent insertions were identified in the *FLT3:* E598DYVDFREYE (NM_ 004119.2:c. 1770_ 1793dup CTACGTTGATTTC AGAGAATATGA) and in the *NPM1*: L287LCX (NM_002520. 6:c.860_ 863dupTCTG) genes. Number of patients with recurrent insertions and deletions were: *EZH2* (N=2), *FLT3*, *NPM1* and *RAD21* (N=3), and *BCORL1* (N=4).

Analysis of downstream effects of indels

The affected domains of the proteins and the downstream effects are shown in Table 3.

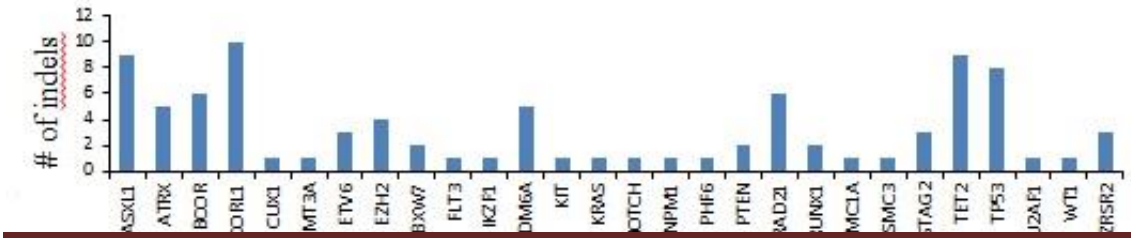

Fig.1. Genes presenting indels in Myelodysplastic syndromes patients.

Table 1. Previously reported indels

Gene	Variant	Type	COSMIC ID	Primary site
ASXL1	G>G/GC	insertion	110708	haematopoietic and lymphoid tissue
ASXL1	AC>AC/A	deletion	307360	haematopoietic and lymphoid tissue
BCOR	GT>GT/G	deletion	5453546	large intestine
DNMT3A	TC>TC/T	deletion	1583095	haematopoietic and lymphoid tissue
FBXW7	CA>CA/C	deletion	5007903	large intestine
FLT3	T>T/TTCATATTCTCT GAAATCAACGTAG	insertion	1317912	haematopoietic and lymphoid tissue
FLT3	T>T/TTCATATTCTCT GAAATCAACGTAG	insertion	1317912	haematopoietic and lymphoid tissue
FLT3	T>T/TTCATATTCTCT GAAATCAACGTAG	insertion	1317912	haematopoietic and lymphoid tissue
KDM6A	TA>TA/T	deletion	255005	urinary tract
NPM1	C>C/CTCTG	insertion	1319222	haematopoietic and lymphoid tissue
NPM1	C>C/CTCTG	insertion	1319222	haematopoietic and lymphoid tissue
NPM1	C>C/CTCTG	insertion	1319222	haematopoietic and lymphoid tissue
PTEN	AC>AC/A	deletion	125653	endometrium
PTEN	TC>TC/T	deletion	5176	Endometrium
RUNX1	GC>GC/G	deletion	444420	breast
TET2	TA>TA/T	deletion	211709	haematopoietic and lymphoid tissue
TET2	TG>TG/T	deletion	120173	haematopoietic and lymphoid tissue
TET2	CT>CT/C	deletion	4170105	haematopoietic and lymphoid tissue
TET2	TC>TC/T	deletion	87187	haematopoietic and lymphoid tissue

Table 2. Recurrent indels

Gene	Variant	Amino acid change	Type	No of patients
EZH2	CA>CA/C	W15X	deletion	2
FLT3	T>T/TTCATATTCTCTGAAATCAACGTAG	E598DYVDFREYE	insertion	3
NPM1	C>C/CTCTG	L287LCX	insertion	3
RAD21	GC>GC/G	G274X	deletion	3
BCORL1	TGAG>TGAG/T	E1316	deletion	4

[Downloaded from ijmcmed.org on 2026-11-07]

[DOI: 10.22088/IJMCM.BUMS.10.3.227]

Table 3. Downstream effects of variants

Gene	Amino acid change	Type	Affected region/domain & possible effect
EZH2	W15X	deletion	Interaction with DNMT1, DNMT3A and DNMT3B is lost
FLT3	E598DYVDFREYE	insertion	The region important for normal regulation of the kinase activity is lost.
NPM1	L287LCX	insertion	The region required for nucleolar localization is lost
RAD21	G274X	deletion	D->A: Abolishes cleavage by caspase-3 is lost
BCORL1	E1316	deletion	Nuclear localization signal might get lost

Discussion

The current study analyzed indels of $CD34^+$ HSPCs derived from a cohort of patients with newly diagnosed *de novo* MDS. Deletion in the *BCORL1* at position 1316 was detected as the most recurrent indel (n=4). This was identified as an inframe deletion within the nuclear localization signal motif of BCORL1. BCORL1 is a transcriptional corepressor that binds to class II histone deacetylases (HDAC4, HDAC5, HDAC7), and interacts with the CTBP1 corepressor to regulate the repression of E-cadherin (18). It has been suggested that the loss of function of *BCORL1* may play a role in the progression of MDS to AML (19). The GVs in *BCORL1* have been reported in 5% of the MDS patients (20). Interestingly, in our study cohort this deletion was observed in 20% of patients, thus this could possibly be a molecular marker for the diagnosis of MDS.

Recurrent deletions were also observed in the *EZH2* and *RAD21* genes. EZH2 is an essential component of the polycomb repressive complex 2 (PRC2), which is involved in gene silencing through trimethylation of H3K27 (21). Dysregulation of *EZH2* has been shown to play an oncogenic role in various cancers (22). GVs in *EZH2* gene are associated with poor prognosis in MDS (23). RAD21 protein is a structural component of the cohesin complex and SNVs in *RAD21* have been reported in haematopoietic neoplasms (24). GVs in *RAD21* gene are associated with proliferation and differentiation of blood cells (25). Therefore, the detected deletions in *RAD21* gene may contribute to the pathogenies of MDS.

Recurrent insertions were observed in the *FLT3* and *NPM1* genes. FLT3 is a class III receptor tyrosine kinase consisting of a juxtamembrane domain (JMD), two tyrosine kinase domains (TKD1 and TKD2), and five extracellular immunoglobulin-like domains (26). GVs of *FLT3* gene have been observed in the internal tandem duplication (ITD) in exon 14 causing the duplication and tandem insertion of the juxtamembrane (JM) domain (27). FLT3 plays a major role in proliferation, differentiation and survival of haematopoietic cells (28). Studies have shown that GVs in internal tandem duplication (ITD) of the *FLT3* gene are associated with poor prognosis (29). *NPM1* encodes for a phospho-protein that belongs to the nucleophosmin/ nucleoplasmin family of proteins, which shuttles between the nucleus and cytoplasm (30). *NPM1* knocked down mice have shown dysplasia in megakaryocyte and erythrocyte lineages in the BM (31). GVs in *NPM1* with a prognostic significance have been previously reported in AML (32). Thus, the insertions in the *FLT3* and *NPM1* genes identified in this study may have a clinical importance in MDS pathogenesis, and possess a potential to be used as markers in the diagnosis of MDS.

Our study identified some recurrent indels in MDS which could possibly have a diagnostic, prognostic, and therapeutic importance in MDS.

[Downloaded from ijmcmed.org on 2026-11-07]

[DOI: 10.22088/IJMCM.BUMS.10.3.227]

However, further investigations with larger cohorts are needed to explore the potential of these indels to be used as biomarkers in MDS.

Conflict of Interest

The authors declare that they have no conflict of interest

Acknowledgement

This research was funded by the University Grants Commission and National Science Foundation of Sri Lanka (UGC/VC/DRIC/08 and RPHS/ 2016/C04) We would like to thank the staff of the Human Genetics Unit, Faculty of Medicine, University of Colombo, Sri Lanka and the consultant haematologists; Dr. Nishadya Ranasinghe, Dr. Yasintha J de Costa, Dr. Lallindra V. Gooneratne, Dr. Baddika R. Jayaratne and Dr. Sashikala Suresh of the haematology units of National Hospital of Sri Lanka, Colombo 10, Colombo South Teaching Hospital, Kalubowila, Colombo North Teaching Hospital, Ragama and the National Cancer Institute, Maharagama, Sri Lanka.

References

1. Pang WW, Pluvinage JV, Price EA, et al. Hematopoietic stem cell and progenitor cell mechanisms in myelodysplastic syndromes. Proc Natl Acad Sci U S A 2013;110:3011-6.

2. Shastri A, Will B, Steidl U, et al. Stem and progenitor cell alterations in myelodysplastic syndromes. Blood 2017;129: 1586-94.

3. Cazzola M, Della Porta MG, Malcovati L. The genetic basis of myelodysplasia and its clinical relevance. Blood 2013;122:4021-34.

4. Ogawa S. Genetics of MDS. Blood 2019;133:1049-59.

5. Kennedy JA, Ebert BL. Clinical Implications of Genetic Mutations in Myelodysplastic Syndrome. J Clin Oncol 2017;35:968-74.

6. Xu F, Wu LY, He Q, et al. Exploration of the role of gene mutations in myelodysplastic syndromes through a sequencing design involving a small number of target genes. Sci Rep 2017;7:43113.

7. Haferlach T, Nagata Y, Grossmann V, et al. Landscape of genetic lesions in 944 patients with myelodysplastic syndromes. Leukemia 2014;28:241-7.

8. Yoshida K, Sanada M, Shiraishi Y, et al. Frequent pathway mutations of splicing machinery in myelodysplasia. Nature 2011;478:64-9.

9. Deeb KK, Smonskey MT, DeFedericis H, et al. Deletion and deletion/insertion mutations in the juxtamembrane domain of the FLT3 gene in adult acute myeloid leukemia. Leuk Res Rep 2014;3:86-9.

10. Kiyoi H, Ohno R, Ueda R, et al. Mechanism of constitutive activation of FLT3 with internal tandem duplication in the juxtamembrane domain. Oncogene 2002;21:2555-63.

11. Young DJ, Nguyen B, Zhu R, et al. Deletions in FLT-3 juxtamembrane domain define a new class of pathogenic mutations: case report and systematic analysis. Blood Adv 2021;5:2285-93.

12. Swerdlow SH, Campo E, Harris NL, et al. WHO classification of tumours of haematopoietic and lymphoid tissues. 4th ed: International agency for research on cancer Lyon, France; 2008.

13. Au CH, Wa A, Ho DN, et al. Clinical evaluation of panel testing by next-generation sequencing (NGS) for gene mutations in myeloid neoplasms. Diagn Pathol 2016;11:11.

14. Bandara W, Rathnayake A, Neththikumara NF, et al. Comparative Analysis of the Genetic Variants in Haematopoietic Stem/Progenitor and Mesenchymal Stem Cell Compartments in de novo Myelodysplastic Syndromes. Blood Cells Mol Dis 2021;88:102535.

15. Wang Q, Jia P, Li F, et al. Detecting somatic point mutations in cancer genome sequencing data: a comparison of mutation callers. Genome Med 2013;5:91.

16. Kanagal-Shamanna R, Singh RR, Routbort MJ, et al. Principles of analytical validation of next-generation sequencing based mutational analysis for hematologic neoplasms in a CLIA-certified laboratory. Expert Rev Mol Diagn 2016;16:461-72.

17. Bandara MS, Goonasekera HWW, Dissanayake VHW. The utility of hematopoietic stem cell karyotyping in the diagnosis of de novo myelodysplastic syndromes. J Hematopathol 2016;9:121-8.

[Downloaded from ijmcmed.org on 2026-11-07]

[DOI: 10.22088/IJMCM.BUMS.10.3.227]

18. Pagan JK, Arnold J, Hanchard KJ, et al. A novel corepressor, BCoR-L1, represses transcription through an interaction with CtBP. J Biol Chem 2007;282:15248-57.

19. Plevin MJ, Mills MM, Ikura M. The LxxLL motif: a multifunctional binding sequence in transcriptional regulation. Trends Biochem Sci 2005;30:66-9.

20. Damm F, Chesnais V, Nagata Y, et al. BCOR and BCORL1 mutations in myelodysplastic syndromes and related disorders. Blood 2013;122:3169-77.

21. Rinke J, Chase A, Cross NCP, et al. EZH2 in Myeloid Malignancies. Cells 2020;9.

22. Margueron R, Reinberg D. The Polycomb complex PRC2 and its mark in life. Nature 2011;469:343-9.

23. Nakamura S, Takayama N, Hirata S, et al. Expandable megakaryocyte cell lines enable clinically applicable generation of platelets from human induced pluripotent stem cells. Cell Stem Cell 2014;14:535-48.

24. Cheng H, Zhang N, Pati D. Cohesin subunit RAD21: From biology to disease. Gene 2020;758:144966.

25. Tsai CH, Hou HA, Tang JL, et al. Prognostic impacts and dynamic changes of cohesin complex gene mutations in de novo acute myeloid leukemia. Blood Cancer J 2017;7:663.

26. Boddu P, Kantarjian H, Borthakur G, et al. Co-occurrence of FLT3-TKD and NPM1 mutations defines a highly favorable prognostic AML group. Blood Adv 2017;1:1546-50.

27. Stirewalt DL, Kopecky KJ, Meshinchi S, et al. Size of FLT3 internal tandem duplication has prognostic significance in patients with acute myeloid leukemia. Blood 2006;107:3724-6.

28. McKenna HJ, Stocking KL, Miller RE, et al. Mice lacking flt3 ligand have deficient hematopoiesis affecting hematopoietic progenitor cells, dendritic cells, and natural killer cells. Blood 2000;95:3489-97.

29. Dohner K, Thiede C, Jahn N, et al. Impact of NPM1/FLT3-ITD genotypes defined by the 2017 European LeukemiaNet in patients with acute myeloid leukemia. Blood 2020;135:371-80.

30. Yip SP, Siu PM, Leung PHM, et al. The multifunctional nucleolar protein nucleophosmin/NPM/B23 and the nucleoplasmin family of proteins. The nucleolus: Springer; 2011. p. 213-52.

31. Loberg MA, Bell RK, Goodwin LO, et al. Sequentially inducible mouse models reveal that Npm1 mutation causes malignant transformation of Dnmt3a-mutant clonal hematopoiesis. Leukemia 2019;33:1635-49.

32. Liu Y, He P, Liu F, et al. Prognostic significance of NPM1 mutations in acute myeloid leukemia: A meta-analysis. Mol Clin Oncol 2014;2:275-81.

[Downloaded from ijmcmed.org on 2026-11-07]

[DOI: 10.22088/IJMCM.BUMS.10.3.227]