Whole Exome Sequencing Identifies Cohesin Component *STAG1* Mutation in *de novo* Acute Myeloid Leukemia (FAB M2): A Pilot Study with Cytogenetics, Clinical and Prognostic Implications

Kunnathur Murugesan Sakthivel,[a,b] Narayanan Geetha,[c] Thampirajan Vimaladevi Akhila Raj,[a] Ramachandran Krishna Chandran,[a] Kumarapillai Mohanan Nair Jagathnath Krishna,[d] & Hariharan Sreedharan[a,*]

[a]Laboratory of Cytogenetics and Molecular Diagnostics, Division of Cancer Research, Regional Cancer Centre, Medical College Post, Trivandrum, Kerala, India; [b]Department of Biochemistry, PSG College of Arts and Science, Civil Aerodrome Post, Coimbatore, India; [c]Division of Medical Oncology, Regional Cancer Centre, Medical College Post, Trivandrum, Kerala, India; [d]Division of Cancer Epidemiology and Biostatistics, Regional Cancer Centre, Medical College Post, Trivandrum, Kerala, India

*Address all correspondence to: Hariharan Sreedharan, Laboratory of Cytogenetics and Molecular Diagnostics, Division of Cancer Research, Regional Cancer Centre, Medical College Post, Trivandrum, Kerala, India; Tel.: +91-9496750040, E-mail: publicationcytogenetics@gmail.com

ABSTRACT: The clinical implications of cohesin gene complex mutation in acute myeloid leukemia (AML) are not well characterized. In the present study, a cohort of 152 *de novo* unselected adult AML patients underwent conventional and molecular cytogenetic analysis for chromosomal aberrations. Further, we examined the frequency and clinical implications of mutations in cohesin gene complex STAG1, STAG2, RAD21, SMC1, and SMC3 using whole exome sequencing as a pilot study in 10 *de novo* patients with AML-FAB M2. Among the 10 cases, we identified a functionally heterozygous mutation in exon16 of STAG1 in one patient (10%), however no mutation was observed in STAG2, RAD21, SMC1, and SMC3. Sanger sequencing analysis for exon 16 of STAG1 in the remaining 142 AML cases did not reveal any further mutations, which underlined the observation that mutations took place throughout the cohesin gene complex without presence of a mutational hot spot region. The present study identified a positive correlation between serum bilirubin, LDH, and hematological parameters such as Hb, WBC, and platelet count with STAG1 mutation. Our data suggest that the cohesin complex may represent an attractive therapeutic target for future preclinical and clinical studies. However, more studies with a larger number of patients should be performed prospectively to determine the pathogenic involvement of STAG 1 mutation in AML patients.

KEY WORDS: whole exome sequencing, cohesin, STAG1/2, mutations, acute myeloid leukemia, cytogenetics

I. INTRODUCTION

The advent of Next Generation Sequencing (NGS) techniques in cancer biology research has led to new insights into the molecular landscape and heterogeneity of acute myeloid leukemia (AML). Recent advances in high-throughput techniques have significantly enhanced our knowledge in identification of genes being mutated that were not previously implicated in AML patients.[1–3] Similar to other myeloid neoplasms, AML is a biologically complex and clinically heterogeneous disease characterized by clonal expansion of myeloid progenitors (blasts) in the bone marrow/peripheral blood through the acquisition of chromosomal rearrangements and multiple gene mutations. The molecular pathogenesis of acute myeloid leukemia (AML) has not yet been completely defined; its incidence is increasing as the population ages. AML remains a disease with poor outcome with a median overall survival (OS) of less than one year for older patients.[4,5]

Currently, cytogenetic abnormalities, molecular diagnostics, and mutational screening play a major role in prognostic stratification and effective treatment strategies for AML. Clinical guidelines for AML recognize three groups of cytogenetic risk, which include favorable, intermediate, and poor risk group. Although, AML genomes contain hundreds of mutations, small number of mutations were found to be recurrent.[6]

0731-8898/21/$35.00

Accumulating evidence indicates that the mutations in the cohesin multi-protein complex are strongly associated with AML and represent a novel genetic pathway for AML.[7] In humans, the cohesin complex is composed of four primary subunit protein complex: i) SMC1 (structural maintenance of chromosomes protein 1A); ii) SMC3 (structural maintenance of chromosomes protein); iii) RAD21 (double-strand-break repair protein rad21 homolog); and iv) either STAG-1 or STAG-2 (cohesin subunit SA1/2) and four regulatory subunits [WAPL (wings apart-like protein homolog), CDCA5 (soronin), and PDS5A or PDS5B (sister chromatid cohesin protein PDS homolog A and B)], which are responsible for holding sister chromatids together during metaphase and enabling proper segregation of sister chromatids into two daughter cells during cell division.[8] Apart from this, cohesin ring is responsible for many different cellular processes, including facilitation of spindle attachment onto the chromosomes, chromosome condensation, heterochromatin formation, facilitation of DNA repair by recombination, and transcription regulation.[9,10]

Cohesin components containing STAG-1 and STAG-2 have distinct roles in mediating chromatid cohesin. STAG-1 is required for telomere cohesin and STAG-2 is uniquely required for the centromeric cohesin.[11–13] However, the exact functional differences between the two mutually exclusive components STAG-1 and STAG-2 are yet to be documented. Mutations in the cohesin-components have already been described and there has been a link between these mutations and chromosomal instability.[14] It is known that, germline cohesin mutations during human development lead to growth and developmental disorders referred as cohesinopathies represents importance of cohesin genes in the pathogenesis of human disease.[15,16] Several reports have identified that genes encoding cohesin complex are somatically mutated in a wide range of adult and pediatric (human) cancers especially, in myeloid neoplasms, recurrent mutations and deletions have been detected.[17,18] Interestingly, recent studies demonstrates identification of genes in the cohesin complex have been described as novel mutations occurring in ~ 13% of AML cases.[19] However, the impact of cohesin complex mutations and identifying their functional role remains to be a major scientific challenge.

Though extensive mutational analysis of large cohort of general AML cases were studied globally, we designed this research analysis as a pilot study to carry out the whole exome sequence analysis (WES) of 10 AML-FAB M2 cases from the cohort of 152 AML cases. The basis for inclusion of patients in the study was (i) to better understand about the incidence of cohesin gene complex mutations in a specific AML FAB subtype; (ii) to maintain homogeneity among the AML subgroups which include a uniform treatment protocol. Ten *de novo* patients with AML-FAB M2 subgroup were chosen randomly from the cohort of 152 cases (5 cases with t(8;21) [favorable risk group]; five cases with Normal Karyotype [intermediate risk group]) were selected; (iii) to check the identified STAG1 (exon16) mutation that was recurrent. The mutational regions were assessed in the remaining 142 cases using Sanger sequencing; and (iv) to check the prognostic impact of mutations in the cohesin complex among the 10 AML-M2 cases. Therefore, the outcome of this study will help to elucidate the clinical characteristics of *de novo* patients with AML harboring cohesin complex mutations with respect to cytogenetics, treatment response, and clinical outcome.

II. MATERIALS AND METHODS

A. Patient Cohorts

This study population included 152 patients, of which 78 were males and 74 females (ratio 1.05:1), with median age 48 years ranging from 16 to 78 years. The frequently used French-American-British (FAB) classification of AML splits this leukemia disease into 8 different subtypes (i.e., M0–M7, based on the type of cell from which the leukemia developed and depending on the degree of maturity of the cell). Diagnostic bone marrow (BM) or peripheral blood (PB) samples were analyzed from these adult *de novo* patients with AML French-American British classification M0–M7. These patients were treated at the Medical Oncology Clinic, Regional Cancer Centre, Thiruvananthapuram, Kerala, India. Written informed consent for research purpose was

obtained from all the patients in accordance with the Declaration of Helsinki protocol. The studies were approved by the Human Ethics Committee (HEC No: 16/2016) and Institutional Review Board (IRB No: 04/2016/02) of Regional Cancer Centre before its initiation. Clinical parameters studied include age, gender, blood counts (WBC, bone marrow/peripheral blast percentage, hemoglobin, and platelet count), LDH level, and overall survival (OS) of these patients.

B. Cytogenetic and Molecular Analysis

Pretreatment specimens from all patients were studied by Giemsa banding (G-Banding) analysis (ASI-Band View software, Applied Biosystems, Foster Hill, CA, USA). Chromosomal abnormalities were described according to the International System for Human Cytogenetic Nomenclature. Cytogenetic risk stratification was established according to the refined Medical Research Council (MRC) criteria. Fluorescence *in situ* hybridization (FISH) analysis were also performed to determine the recurrent abnormalities [t(8;21), (15;17) (DCDF probe) and inv (16) (break apart probe)] associated with FAB subtypes M2, M3, and M4, respectively (using Vysis probes; Abbott Molecular Vysis, Des Plaines, IL, USA).

C. Pilot Study

Separately, this study was designed to determine the correlation between Cohesin gene complex mutations and clinical characteristics of patients with *de novo* AML-M2 subtype, and its association with the prognosis of AML-M2. Generally, AML-M2 subtype comprises 10%–15% of all AML cases and 26% of total population (39 out of 152 patients) in this present study. As described earlier, to maintain homogeneity among the AML cases which include uniform treatment protocol, we had randomly chosen 10 *de novo* AML-M2 patients as a pilot study to screen Cohesin gene complex mutations using NGS analysis. By cytogenetic analysis, t(8;21) is often found in this subgroup, hence 5 patients with t(8;21) and negative, respectively, were included in this study. Clinical parameters and overall survival (OS) of these patients were studied.

D. Analysis of Cohesin Mutations Using Whole Exome Sequencing Method

Pretreatment BM specimens were enriched for mononuclear cells by using Ficoll density gradient centrifugation. Genomic DNA (gDNA) was extracted from cryopreserved mononuclear cells by using a Xcelgen blood gDNA mini kit (cat no. XG2311-01) according to the manufacturer's instructions. For whole exome sequencing, 50 ng of DNA was digested using restriction enzymes, followed by targeted probe enrichment. Paired-end indexed libraries were prepared using HaloPlex Exome Target Enrichment System (Agilent Technology, Santa Clara, CA, USA) as per manufacturer's instruction. The 37 Mb of protein coding exonic sequences were enriched using specific probes supplied with the kit. Capture libraries were pooled at equimolar concentrations and loaded onto paired end flow cells at concentrations of 10 pM to generate clusters following Illumina's standard protocol using Hiseq Rapid cluster kit. Then the flow cells were sequenced as 2 X 150 paired end reads on Illumina HiSeq 2500 (Illumina, San Diego, CA, USA) at Xcelris Genomics (Ahmedabad, Gujarat, India; www.xcelrisgenomics.com).

E. Data Analysis and Annotation

After sequencing, the exonic sequence data was subjected to an in-house framework of bioinformatics pipeline (BWA-GATK) for the detection of cohesin mutations (STAG1, STAG2, RAD21, SMC1, and SMC3) in the target regions, which was very well established at the Xcleris Labs Limited (Ahmedabad, India). The data analysis was mainly focused on 5 standard analytical steps: raw data quality assessment, pre-processing, alignment, post-alignment processing, and variant analysis (variant detection, annotation, and prioritization) as described by Bao et al.[20] Thus, raw data obtained from Illumina Hiseq 2500 was first checked for base quality to remove low quality bases and to remove adapter sequences to obtain high quality (HQ) clean data. These filtration processes was carried out using a Trimmomatic tool (v0.30). Further alignment of these HQ data on to

human reference genome (hg19) was implemented using BWA (v0.7.12). After mapping, 3 recommended steps involving duplicate read removal, indel realignment, and base quality score recalibration (BQSR) were carried out using Picard-tools (v1.60) and GATK (v3.5).[21,22] These steps minimize the sequencing artifacts and false positives which could affect downstream variant calling. Depending on the targeted regions enriched with 37 MB library kit, in each sample approximately 30,000 to 50,000 variants (SNPs) were called using Bayesian model in Unified Genotyper tool. Variant calling was restricted to the target regions interval list. Variants with > ~ 90% truth sensitivity and optimal quality scores were further filtered based on various criteria. Whole exome sequencing-based variants passing all the filtration criteria based on QualByDepth (< 2.0), Strand bias (> 60/> 200), Mapping quality (< 40.0) and Read position (< –8.0/< –20.0), respectively, relevant to SNPs were annotated as either PASS or specific filter name. Final filtered variants were annotated against multiple databases (RefGene, ClinVar, 1000G, ESP6500vi2, ExAC, COSMIC, dbNSFP) using ANNOVAR.[23] Mutations or variants that are in intronic (non-coding) regions of the genome were not covered in the 37 MB exome library kit. During annotation, all synonymous SNPs or synonymous variants were not considered for further interpretation and only non-synonymous SNPs/variants which have known to be pathogenic were interpreted in this study. Candidate pathogenic variants identified were validated and are only reported if they were detected by Sanger sequencing in the gDNA sample. The flowchart of whole exome sequencing methodologies followed in the study is represented in Fig. 1.

F. Statistical Methods

Data was analyzed using Statistical Package for the Social Sciences (SPSS) Software version 21. Quantitative variables were summarized using mean (range). For continuous variables we used Student's

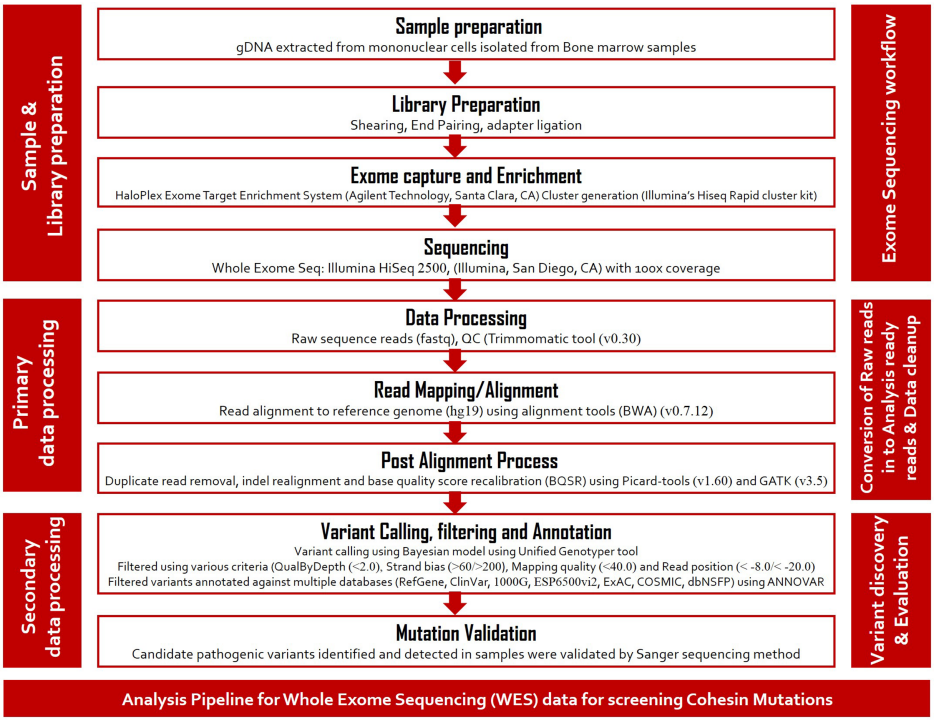

FIG. 1: Analysis pipeline for whole Exome sequencing (WES) data for screening cohesin mutations

t-test and Mann-Whitney U test. Qualitative data were summarized with frequencies and percentages. The Kaplan-Meier survival and log-rank test were used estimate the overall survival (OS). OS up to one year was calculated from the time of diagnosis (i.e., date of first referral) to date of death (uncensored) or last follow-up (censored). *P*-value of < 0.05 was considered statistically significant. The Cox proportional Hazard model was performed to assess the contribution of each independent factor to the probability of survival. Fisher's exact test and Chi-Square test were used to compare the pretreatment characteristics between AML-FAB M2 with t(8;21) and AML-FAB M2 with Normal Karyotype (NK).

III. RESULTS

A. Clinical Characteristics of AML Patients

The clinical and hematologic parameters such as age, gender WBC count, platelet, Hb, BM blast, PB blast, and LDH levels were analyzed between patients with AML M2 and all other subtypes of AML. However, the current study failed to find any significant relationship between these parameters with any of the AML subtypes. Similarly, it was also unable to portray any significant relationship between AML M2 cases with t(8;21) and AML M2 cases with normal karyotypes in any of the above clinical and hematologic parameters as details shown in Tables 1 and 2.

B. Cytogenetic Abnormalities of AML Patients

In our cohort of 152 patients, karyotype was carried out in 131 patients (86.18%). Due to unavailability of good quality metaphase, karyotype analysis failed in 21 cases. Among the 131 patients, 88 patients were identified as cytogenetically normal (CN). In cytogenetically abnormal group, the most frequent abnormality was t(8;21)(q22;q22) identified in 13 cases (8.55%) followed by t(15;17)(q22;q21) in 8 cases (5.26%), inv(16) in 3 cases (1.97%). Apart from these normal and recurrent chromosomal aberrations, 11 patients (7.24%) with numerical abnormalities (+4, +8 in 2 cases, +16, +21, hyperdiploids in 3 cases, and polyploids in 3 cases), 4 patients (2.63%) with structural abnormalities [del(4q), del(7q), dup(11q) and dup(21q)] were identified. Among the 13 patients with t(8;21)(q22;q22), 5 patients were identified to have loss of sex chromosome (-X in 2 case and -Y in 3 case) (Fig. 2A and B).

C. Rare Chromosomal Abnormalities Identified in AML-M2 Subtype

Rare chromosomal aberrations could be detected in a substantial proportion of AML, which are mainly associated with unfavorable prognosis. Four patients (2.63%) with rare abnormalities were identified (i) 46,XY,i(17)(q10), (ii) 46,XY,t(7;11)(p15;q23), (iii) 45,XY,der(21)t(8;21)(q22;q22),–22, (iv) 46,XX, der(13)t(8;21;13),der(21)t(13;21). Another significant finding of this study included most of the numerical and additional abnormalities identified in AML-M2 sub type.

D. Cohesin Gene STAG 1 Mutation in AML-M2 FAB Subtype

In our cohort, a total of 10 *de novo* patients with AML-M2 subgroup from the cohort of 152 cases (5 cases with t(8;21) [favorable risk group] and 5 cases with Normal Karyotype [intermediate risk group]) were selected and screened for cohesin gene complex mutation. No mutation was identified in STAG2, RAD21, SMC1, and SMC3, except STAG1. Only one case harboring STAG1 mutation was identified from the selected AML-M2 FAB 10 cases. STAG1 gene is autosomal located in chromosome 3 region 3q22.3. The patient had a functionally heterozygous mutation for STAG1 (missense; [nucleotide change: c.1577T>C, amino acid change:p.I526T] shown in Fig. 3A–E). The same mutation was recently reported by Thol et al.[19] in an AML patient. In order to check that the mutation was recurrent, the STAG1 mutation (exon16) region was screened using Sanger sequencing in the remaining 142 cases. However, no STAG1 mutation was detected in remaining 142 cases (0%).

TABLE 1: Comparison of pretreatment characteristics between patients with AML-M2 sub group and all other subgroups

Clinical parameters	*De novo* AML-FAB M2 Subtype (n = 39)	*De novo* AML (subtypes except FAB M2) (n = 113)	*P*-value
Age, Years			0.674
Mean	44.38	45.72	
Range	16–68	16–78	
Gender			0.460
Male	22 (56%)	56 (50%)	
Female	17 (44%)	57 (50%)	
FAB subtype			
M0	0	9 (8%)	
M1	0	18 (16%)	
M2	39 (100%)	0	
M3	0	12 (11%)	
M4	0	30 (27%)	
M5	0	34 (30%)	
M6	0	1 (1%)	
M7	0	0	
Missing data	0	9 (8%)	
Cytogenetic risk group			0.0001
Favorable	14 (36%)	11 (10%)	
Intermediate	22 (55%)	81 (72%)	
Unfavorable	3 (7%)	1 (1%)	
Missing data	1 (2%)	20 (17%)	
Peripheral blood blasts			0.21
Mean	45.4	52.3	
Missing data	2 (5%)	10 (9%)	
Bone marrow blasts			0.007
Mean	51.28	62.37	
Missing data	1 (1%)	9 (8%)	
WBC count			0.815
Mean ($\times 10^9$/L)	37505	41187	
Range ($\times 10^9$/L)	1000–256,100	300–378,700	
Missing data	0 (0%)	0 (0%)	
Hemoglobin			0.709
Mean (g/L)	8.03	8.18	
Range (g/L)	3.1–12.8	4.1–13.4	
Missing data	0 (0%)	1 (1%)	
Platelet count			0.044
Mean ($\times 10^9$/L)	53282	66701	

TABLE 1: (*continued*)

Range ($\times 10^9$/L)	6000–301,000	5200–289,000	
Missing data	0 (0%)	1 (1%)	
LDH			0.182
Mean (IU/L)	1473.05	1264.09	
Range (IU/L)	263–4054	314–5025	
Missing data	1 (1%)	3 (3%)	

TABLE 2: Comparison of pretreatment characteristics between AML-FAB M2 with t(8;21) and AML-FAB M2 with normal karyotype (NK)

Clinical Parameters	*De novo* AML-M2 with t(8;21) (n = 14)	*De novo* AML-M2 with NK (n = 25)	*P*-value
Age, Years			0.022
Median	36.71	48.68	
Range	16–64	22–68	
Gender			0.458
Male	9	13	
Female	5	12	
Peripheral blood blasts			0.682
Mean	43.43	46.52	
Missing data	0 (0%)	2 (8%)	
Bone marrow blasts			0.434
Mean	55.71	50.83	
Missing data	0 (0%)	1 (4%)	
WBC count			0.219
Mean ($\times 10^9$/L)	13250	16000	
Range ($\times 10^9$/L)	3300–37,400	1000–256,100	
Missing data	0 (0%)	0 (0%)	
Hemoglobin			0.181
Mean (g/L)	7.4	8.3	
Range (g/L)	3.1–10.6	4.8–12.8	
Missing data	0 (0%)	0 (0%)	
Platelet count			0.235
Mean ($\times 10^9$/L)	30,642	65,960	
Range ($\times 10^9$/L)	9000–131,000	6000–301,000	
Missing data	0 (0%)	0 (0%)	
LDH			0.619
Mean (IU/L)	1552.57	1428.52	
Range (IU/L)	473–3860	263–4054	
Missing data	0 (0%)	1 (4%)	

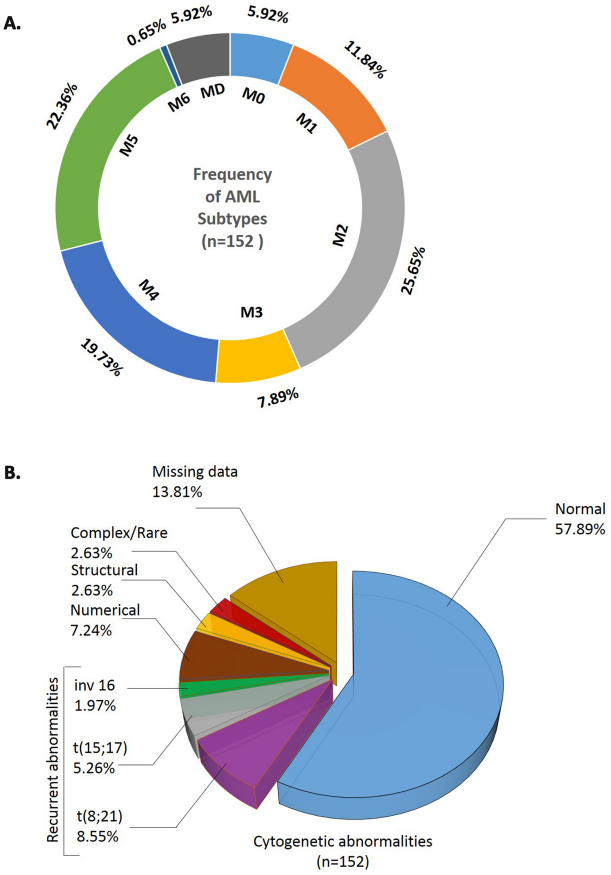

FIG. 2: (A) Frequency of AML subtypes; (B) Distribution of cytogenetically and molecularly defined chromosomal aberrations in a cohort of 152 AML patients

E. Association of Identified STAG1 Gene Mutation with Clinical Characteristics

The patient with STAG1 possessed an abnormal karyotype with t(8;21)(q22;q22) and was associated with adverse treatment response. The detailed case report as follows, a 25-year-old male patient was evaluated at a local hospital and referred to Outpatient Clinic of Medical Oncology, Regional Cancer Centre, for generalized tiredness for 1 month and multiple ecchymoti patches for 1 week (February, 2017). Patient had no family history of cancer. Patient showed normal chest X-ray and mild splenomegaly. Hematology parameters, immunophenotyping, quantitative analysis of BCR/ABL and cytogenetics of the patient were determined at the time of diagnosis. Morphology and immunophenotyping of peripheral blood cells and bone marrow aspiration were examined. The abnormal cell population were positive for CD13, CD117, MPO, CD11c, CD19dim, CD34, and HLA-DR positive (human leukocyte antigen DR isotype) and negative for CD33. Peripheral blood and bone marrow aspiration revealed blast cells of 79% and 61%, respectively. Less than 3% blast cells were peroxidase positive. The biochemical parameters such as urea, uric acid, creatinine, and liver enzymes were normal. The elevated level of LDH indicate the form of tissue damage. Bilirubin was slightly raised and serum LDH was highly increased (1985 µ/L). Hematology parameters at the time of diagnosis were: total WBC count was 20,300 cells/mm (19% neutrophils, 22% lymphocytes, and 59% of other abnormal cells), hemoglobin level of 10.5 gm% and platelet count of 26,000 cells/cmm. Conventional and molecular cytogenetic profiling identified the presence of t(8;21) and the RT- PCR report was negative for *BCR- ABL1*. The diagnosis was confirmed as AML-M2.

A 7+3 chemotherapy regimen was planned for the induction cycle. Chemotherapy started immediately after diagnosis and, due to febrile neutropenia, chemo was skipped on day 7. After 7+3 induction, consolidation chemotherapy with high dose Ara-C (HiDAC) was given. As the bone marrow biopsy shows leukemic cells, re-induction chemotherapy was recommended for the patient. The patient attained hematological remission and tolerated well with the post-chemo complications like febrile neutropenia, perianal pain, and thrombophlebitis. After 1 year of diagnosis, the patient relapsed with the disease.

The data for overall survival of patients with AML FAB-M2 subtype and patients with AML FAB subtypes except M2 and also Kaplan-Meier plots showing the prognostic impact of t(8;21) in the overall survival of AML-FAB M2 subtype with t(8;21) and AML-FAB M2 subtype with normal karyotype (Fig. 4 and Table 3).

IV. DISCUSSION

Mutations in the cohesin complex have been well characterized in solid tumors such as colorectal

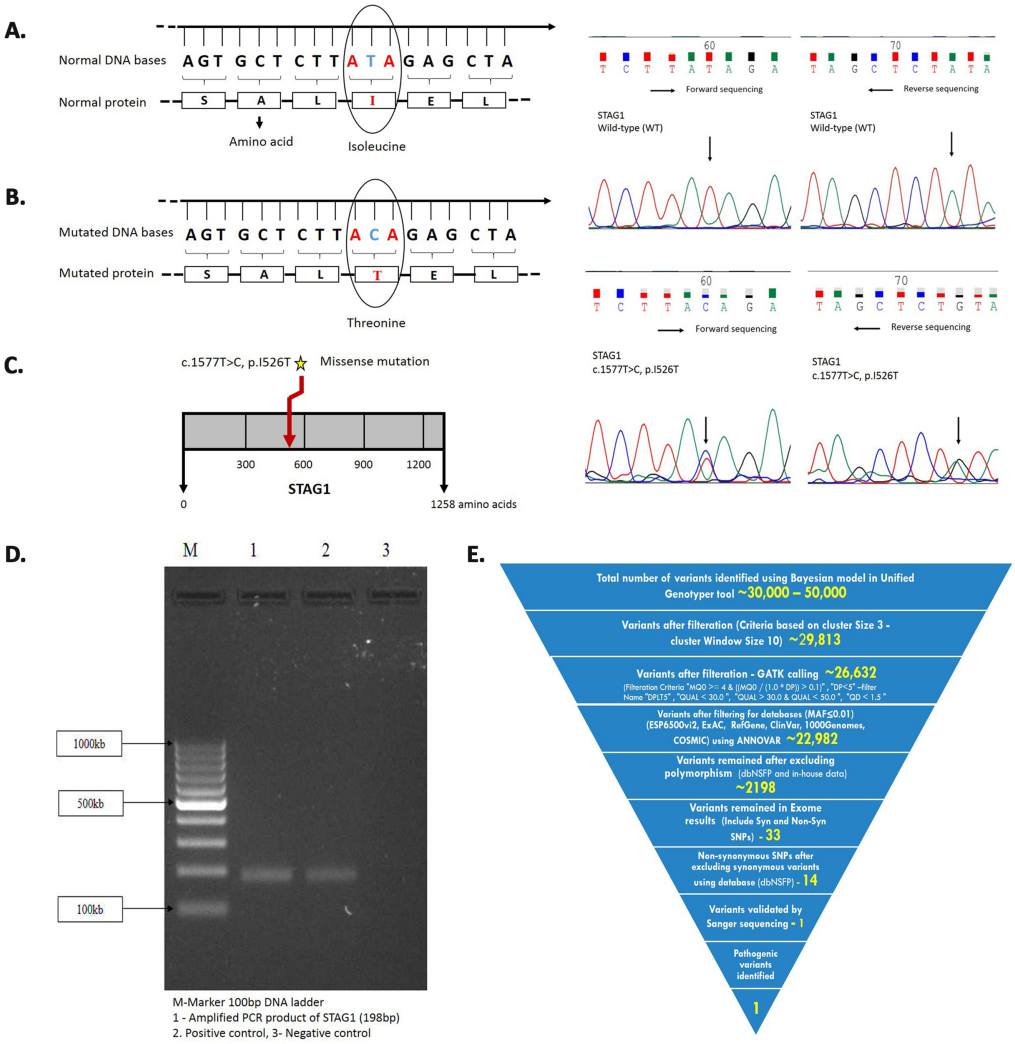

FIG. 3: (A) Schematic representation of normal and mutant nucleotide and amino acid sequences of identified STAG1 missense mutation; (B) Location and type of identified mutation in Exon 16 (nucleotide change: c.1577T>C, Amino acid change:p.I526T) of STAG1 gene of the cohesin complex. Sequence numbering is in accordance to the DNA coding sequence of Ensembl transcript ENST00000383202.6 and the protein sequence of ENSEMBL protein ENSP00000372689.2 for STAG1; (C) Sanger validation of forward and reverse sequencing chromatographs; (D) Amplified PCR product of STAG1 gene (198 bp) along with positive and negative control; (E) Filtering strategy for identification of variant found in STAG1. MQ and MQ0 Root Mean Square Mapping Quality and Mapping Quality Zero total count; DP Coverage (reads that passed quality metrics); QD, Variant Quality /depth of non-ref samples; FS Test (Fisher) - Phred score *P*-value for strand bias.

carcinoma and elucidated a link between these mutations and chromosomal instability.[24] However, only a few studies have been reported regarding the clinical implications of cohesin complex mutation in myeloid neoplasms such as AML. Hence, it is very necessary to investigate the frequency, incidence, and prognostic impact of cohesin complex mutation in various subtypes of AML and whether do they have any association with AML cytogenetics. All these parameters will help to identify a novel prognostic biomarker in AML and thereby illuminate the path of effective management of AML patients.

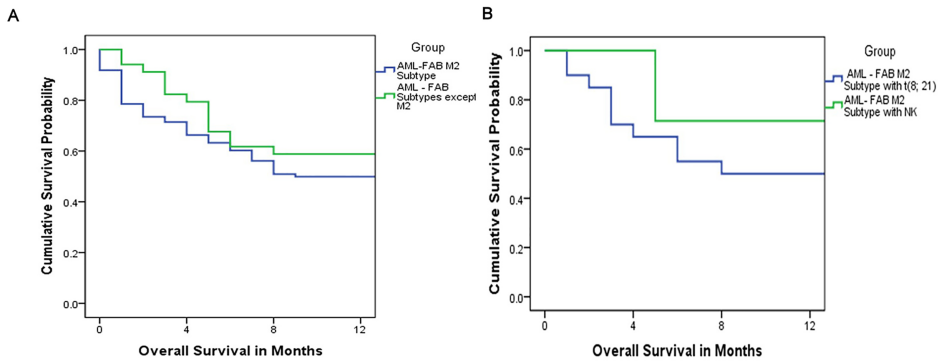

FIG. 4: (A) Kaplan-Meier plots comparing the overall survival of patients with AML FAB M2 subtype and patients with AML FAB subtypes except M2; (B) Kaplan-Meier plots showing the prognostic impact of t(8;21) in the overall survival of AML-FAB M2 subtype with t(8;21) and AML-FAB M2 subtype with normal karyotype

TABLE 3: Univariate analysis for overall survival (12 months) in AML patients according to FAB subtypes

Samples (n = 152)	12 months survival probability		*P*-value	HR	95% CI
	Percentage (%)	SE			
AML-FAB subtypes except M2 (n = 113)	49.9	5.1	0.298	0.738 (*P*-value: 0.316)	0.407–1.337
AML-FAB M2 subtype (n = 39)	58.8	8.4			
AML-FAB M2 t(8;21) (n = 14)	50	11.2	0.160	0.455 (*P*-value: 0.184)	0.142–1.454
AML-FAB M2 with NK (n = 25)	71.4	12.1			

With this notion, in the present study, we analyzed both cytogenetic profile and cohesin complex mutational signatures of 152 newly diagnosed unselected adult AML patients and investigated their correlation with clinical parameters. Furthermore, to the best of our knowledge, this is the first report portraying the clinical and prognostic impact of cohesin complex mutation in AML patients from India.

In our cohort, successful cytogenetic analyses were carried out in 131 patients (86.18%). In the remaining 21 cases analysis was not possible due to non-availability of good quality metaphases. Among the 131 patients, 88 patients revealed normal karyotype pattern, while 43 patients revealed both structural and numerical chromosomal aberrations. Most of the previous cytogenetic studies from different geographical areas reported higher frequency (> 50%) of chromosomal abnormalities in AML patients. But in our study, by performing combined conventional and molecular cytogenetic techniques, we identified 32.82% chromosomal anomalies. The lower frequency of cytogenetic abnormalities in the present study was in agreement with studies reported from Germany and Malaysia.[25,26]

The most frequent cytogenetic abnormality identified in our study was t(8;21)(q22;q22) and it was observed in 13 cases (8.55%). Karyotype with t(15;17)(q22;q21) were observed in 8 cases (5.26%) and inv(16) in 3 cases (1.97%). Concerning recurrent balanced chromosomal translocations and inversions, the frequency of t(8;21) (RUNX/RUNX1T1) in our study was comparable with previous Asian reports (8.3% and 7.5%).[27,28] But, the frequency of t(15;17)(PML/RARA) was slightly lower than previous western and Asian

studies (10.8%).[28,29] Majority of the studies report a frequency of 2–8% inv(16) in AML patients.[30,31] 1.97% of inv(16) in our AML cohort was in agreement with those reports. In addition, 11 patients (7.24%) unveiled karyotype with numerical aberrations. The main numerical aberrations observed in our study were +4, +8 (2 cases) +16, +21, hyperdiploids (3 cases) and polyploids (3 cases). Furthermore, 4 patients (2.63%) displayed structural chromosomal abnormalities like del(4q), del(7q), dup(21q) and dup(11q). Among the 13 patients with t(8;21)(q22;q22), 5 patients showed a loss of sex chromosome (-X in 2 case and -Y in 3 case) (Fig. 2A and B).

Chromosomal translocations involving the alterations of *MLL* gene located on the 11q23 are common in adult AML patients. Here, by using G-banding and FISH analysis, we identified a balanced reciprocal rare t(7;11)(p15;q23) in a 49-year-old AML-M2 subtype male patient. There are no substantial previous reports on t(7;11)(p15;q23) in AML, however, only few cases were reported in AML-M2. In the current scenario, more than 60 different MLL fusion partner genes have been characterized in different chromosomes at the molecular level.[32] In the present study, further studies are needed for identifying the MLL fusion partner located on 7p15.

Next, we analyzed cohesin complex mutation signatures in our study samples by NGS and Sanger sequencing methods. We selected 10 *de novo* AML-M2 patients at random as a pilot study to monitor cohesin gene complex mutations using NGS analysis. Among this 10 *de novo* AML-M2 cases, 5 cases were positive for t(8;21) and the remaining 5 patients possessed normal karyotypes. Our whole exome scale sequencing study in 10 AML-M2 FAB subtype identified STAG1 gene mutation in only one case (10%), however, no mutation was observed in STAG2, RAD21, SMC1, and SMC3 which was in agreement with a previous report.[19] As per their report, STAG1, STAG2, and SMC3 were the most frequently mutated genes in the cohesin complex, whereas mutations in SMC1A were very rare. However, in contrast to our finding, recent report by Cancer Genome Atlas Research Network[33] showed that the mutation frequency of STAG2, SMC1A, SMC3, and RAD21 was slightly higher in their study population; their study failed to identify any such mutation in STAG1.

Similar to previous reports,[19,33,34] our study also demonstrated that mutations in the cohesin complex genes occurred in 10% of patients with myeloid malignancy. Among the myeloid neoplasms, cohesin mutations were more predominant in AML, especially in 20% of secondary AML[33] Most of the identified cohesin mutations were either nonsense or frame shift types, pointing to the fact that these mutations lead to decreased function of cohesin complex genes. Similarly, the identified STAG1 mutation in the current study was frame shift type, reflecting that it may lead to decreased or altered function of cohesin complex in that patient.

The identified STAG1 mutation by whole exome scale sequencing analysis was heterozygous, which was in agreement to the recent report by Thol et al.[19] in AML patient and thereby it necessitates the urgency to check whether the mutated region could act as hotspot or not. Sanger sequencing analysis in the remaining 142 cases for STAG1 mutation in exon16. However, Sanger sequencing did not identify any such mutation in the above region. This finding once again underlined the conclusion made by Thol et al.[19] that mutations took place throughout the cohesin gene complex were without the presence of any mutational hot spot region.

In line with other studies, cohesin gene mutations did not alter the global chromosomal integrity in AML, since most of the mutated AML patients harbored a normal karyotype pattern. Instead, they were able to control patient's genome by modifying the regulation at transcriptional level. It have been reported that cohesin gene extended their tumor suppressor activity through direct or indirect binding with CCCTC-binding factor, a sequence-specific transcription factor which is found to be interacted with NPM1.[35–37] Furthermore, it was also proven that cohesin played an important role in the arrangement of sister chromatids during the metaphase stage of cell cycle, further illustrating that cohesin defects appearing during leukemogenic consequences did not convey disrupted mitosis and global chromosomal instability.[38] Contradictory to the above findings, in our study, conventional and molecular cytogenetic profiling identified the presence of t(8;21) in patient

with STAG1 mutation. Although its biochemical parameters were normal, serum bilirubin and LDH level was highly elevated. In addition, its hematology parameters such as Hb, WBC, and platelet count were also noted to be abnormal. In the present study, we identified a positive correlation between serum bilirubin, LDH, and hematological parameters with STAG1 mutation. However, in order to identify the significance of this correlation it is necessary to conduct further additional studies with a greater number of patients.

Treatment response analysis on STAG1 mutated patient showed that the patient achieved hematological remission by 7+3 induction and consolidation chemotherapy with high dose Ara-C (HiDAC). However, after 1 year the patient relapsed with the disease, highlighting that the prognostic impact of STAG1 mutation was poor in our study. These data implement the prognostic effect of STAG1 mutation in only 1 case, and therefore further more studies are needed to make a final conclusion about the prognostic impact cohesin mutation. A previous study[19] strongly recommended that there exists a strong association between NPM1 mutations and genes in the cohesin complex. Therefore, the adverse impact on prognosis of cohesin mutation might have been attenuated by the favorable prognostic impact of NPM1 mutations. Finally, contradictory to our report, their data pinpointed the fact that cohesin mutations did not exert much impact on patient prognosis.

V. CONLCUSION

In summary, our study shows that cohesin complex mutations are not recurrent molecular genetic event in AML. Sanger sequencing analysis for *STAG1* mutation in exon 16 region points the fact that it could not be considered as hot spot region for AML mutagenesis. Molecular studies to rule out their functional and pathophysiological role are currently underway, but our genetic data suggest that cohesin mutations can contribute to hematopoietic transformation through altered chromosomal instability. It was also highlighted that cohesin complex mutation had a poor prognostic impact in cytogenetically abnormal AML patients. In the present study, the recent advancement in exome and whole genome sequencing provided some relevant information regarding cohesin complex mutations like frequency of incidence, prognostic significance, and correlation with cytogenetics as well as clinical parameters in AML. Clinical and preclinical research in the future will be necessary to understand the involvement of other functional proteins with the mutated cohesin complex proteins, which will hopefully shed further light on the pathogenic mechanism of cohesin complex mutations in AML incidence and progression. Optimization of AML treatment is a continuous process. Knowledge is still accumulating, and the therapeutic scenario is still evolving. In this context, the identification of cohesin as a key protein complex in leukemogenesis could certainly contribute to the development of new drugs that specifically target mutant stem cells in AML.

ACKNOWLEDGMENTS

The authors received financial support by Science and Engineering Research Board (SERB), Department of Science and Technology (DST), Government of India, New Delhi (Ref No. PDF/2016/000631) in the form of National Post-Doctoral Fellowship (N-PDF) to Kunnathur Murugesan Sakthivel. We also thank the patients who participated in the study and staff members of the various clinical departments of Regional Cancer Centre, Trivandrum. Authors thank Sequencing team members, Xcelris Genomics, Ahmedabad, Gujarat-380015 India for their support for sequencing and analysis.

REFERENCES

1. Ley TJ, Mardis ER, Ding L, Fulton B, McLellan MD, Chen K, Dooling D, Dunford-Shore BH, McGrath S, Hickenbotham M, Cook L. DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome. Nature. 2008;456(7218):66–72.
2. Michael RS, Campbell PJ, Futreal PA. The cancer genome. Nature. 2009;458(7239):719–24.
3. Shiba N, Yoshida K, Shiraishi Y, Okuno Y, Yamato G, Hara Y, Nagata Y, Chiba K, Tanaka H, Terui K, Kato M. Whole-exome sequencing reveals the spectrum of gene mutations and the clonal evolution patterns in paediatric acute myeloid leukaemia. Br J Haematol. 2016;175(3):476–89.

4. Roboz GJ. Current treatment of acute myeloid leukemia. Curr Opin Oncol. 2012; 24(6):711–9.
5. Yan XJ, Xu J, Gu ZH, Pan CM, Lu G, Shen Y, Shi JY, Zhu YM, Tang L, Zhang XW, Liang WX. Exome sequencing identifies somatic mutations of DNA methyltransferase gene DNMT3A in acute monocytic leukemia. Nat Genet. 2011;43(4): 309–15.
6. Welch JS, Ley TJ, Link DC, Miller CA, Larson DE, Koboldt DC, Wartman LD, Lamprecht TL, Liu F, Xia J, Kandoth C. The origin and evolution of mutations in acute myeloid leukemia. Cell. 2012;150(2):264–78.
7. Leeke B, Marsman J, O'Sullivan JM, Horsfield JA. Cohesin mutations in myeloid malignancies: underlying mechanisms. Exper Hematol Oncol. 2014;3(1):13.
8. Nasmyth K, Haering CH. Cohesin: its roles and mechanisms. Annu Rev Genet. 2009;17:43:525–58.
9. Peters JM, Tedeschi A, Schmitz J. The cohesin complex and its roles in chromosome biology. Genes Dev. 2008; 15:22(22):3089–114.
10. Mehta GD, Kumar R, Srivastava S, Ghosh SK. Cohesin: Functions beyond sister chromatid cohesion. FEBS Lett. 2013;2;587(15):2299–312.
11. Van Der Lelij P, Lieb S, Jude J, Wutz G, Santos CP, Falkenberg K, Schlattl A, Ban J, Schwentner R, Hoffmann T, Kovar H. Synthetic lethality between the cohesin subunits STAG1 and STAG2 in diverse cancer contexts. Elife. 2017;10;6:e26980.
12. Losada A. Cohesin in cancer: Chromosome segregation and beyond. Nat Rev Cancer. 2014;14(6):389–93.
13. Murayama Y, Uhlmann F. Biochemical reconstitution of topological DNA binding by the cohesin ring. Nature. 2014;505(7483):367–71.
14. Kon A, Shih LY, Minamino M, Sanada M, Shiraishi Y, Nagata Y, Yoshida K, Okuno Y, Bando M, Nakato R, Ishikawa S. Recurrent mutations in multiple components of the cohesin complex in myeloid neoplasms. Nature Gen. 2013;45(10):1232–7.
15. Ball Jr AR, Chen YY, Yokomori K. Mechanisms of cohesin-mediated gene regulation and lessons learned from cohesinopathies. Biochimica Biophysica Acta (BBA)-Gene Regulatory Mechan. 2014;1839(3):191–202.
16. Zakari M, Yuen K, Gerton JL. Etiology and pathogenesis of the cohesinopathies. Wiley Interdiscip Rev Dev Biol. 2015;4(5):489–504.
17. Hill VK, Kim JS, Waldman T. Cohesin mutations in human cancer. Biochim Biophys Acta Rev Cancer. 2016;1866(1):1.
18. Solomon DA, Kim JS, Waldman T. Cohesin gene mutations in tumorigenesis: From discovery to clinical significance. BMB Rep. 2014;47(6):299.
19. Thol F, Bollin R, Gehlhaar M, Walter C, Dugas M, Suchanek KJ, Kirchner A, Huang L, Chaturvedi A, Wichmann M, Wiehlmann L. Mutations in the cohesin complex in acute myeloid leukemia: Clinical and prognostic implications. Blood. 2014;123(6):914–20.
20. Bao R, Huang L, Andrade J, Tan W, Kibbe WA, Jiang H, Feng G. Review of current methods, applications, and data management for the bioinformatics analysis of whole exome sequencing. Cancer Inform. 2014:CIN-S13779.
21. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, Jordan T, Shakir K, Roazen D, Thibault J, Banks E. From FastQ data to high-confidence variant calls: The genome analysis toolkit best practices pipeline. Curr Protoc Bioinform. 2013;43(1):11–33.
22. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, Del Angel G, Rivas MA, Hanna M, McKenna A. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet. 2011;43(5):491.
23. Wang K, Li M, Hakonarson H. ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010;38(16):e164.
24. Barber TD, McManus K, Yuen KW, Reis M, Parmigiani G, Shen D, Barrett I, Nouhi Y, Spencer F, Markowitz S, Velculescu VE. Chromatid cohesion defects may underlie chromosome instability in human colorectal cancers. Proc Natl Acad Sci U S A. 2008;105(9):3443–8.
25. Bacher U, Kern W, Schnittger S, Hiddemann W, Schoch C, Haferlach T. Further correlations of morphology according to FAB and WHO classification to cytogenetics in de novo acute myeloid leukemia: A study on 2,235 patients. Ann Hematol. 2005;84(12):785–91.
26. Meng CY, Noor PJ, Ismail A, Ahid MF, Zakaria Z. Cytogenetic profile of de novo acute myeloid leukemia patients in Malaysia. Int J Biomed Sci. 2013;9(1):26.
27. Cheng Y, Wang Y, Wang H, Chen Z, Lou J, Xu H, Qian W, Meng H, Lin M, Jin J. Cytogenetic profile of de novo acute myeloid leukemia: A study based on 1432 patients in a single institution of China. Leukemia. 2009;23(10):1801–6.
28. Enjeti AK, Tien SL, Sivaswaren CR. Cytogenetic abnormalities in de novo acute myeloid leukemia in adults: Relation to morphology, age, sex and ethnicity-a single center study from Singapore. Hematol J. 2004;5(5):419–25.
29. Grimwade D, Walker H, Harrison G, Oliver F, Chatters S, Harrison CJ, Wheatley K, Burnett AK, Goldstone AH. The predictive value of hierarchical cytogenetic classification in older adults with acute myeloid leukemia (AML): Analysis of 1065 patients entered into the United Kingdom Medical Research Council AML11 trial. Blood. 2001;98(5):1312–20.
30. Mrózek K, Heinonen K, Bloomfield CD. Clinical importance of cytogenetics in acute myeloid leukaemia. Best Pract Res Clin Hematol. 2001;14(1):19–47.
31. Preiss BS, Kerndrup GB, Schmidt KG, Sørensen AG, Clausen NA, Gadeberg OV, Mourits-Andersen T, Pedersen NT, AML Study Group Region of Southern Denmark. Cytogenetic findings in adult de novo acute myeloid leukaemia. A population-based study of 303/337 patients. Br J Haematol. 2003;123(2):219–34.
32. Coenen EA, Zwaan CM, Meyer C, Marschalek R,

Pieters R, van der Veken LT, Beverloo HB, van den Heuvel-Eibrink MM. KIAA1524: A novel MLL translocation partner in acute myeloid leukemia. Leukemia Res. 2011;35(1):133–5.
33. Cancer Genome Atlas Research Network. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. N Engl J Med. 2013;30;368(22):2059–74.
34. Thota S, Viny AD, Makishima H, Spitzer B, Radivoyevitch T, Przychodzen B, Sekeres MA, Levine RL, Maciejewski JP. Genetic alterations of the cohesin complex genes in myeloid malignancies. Blood. 2014;124(11):1790–8.
35. Lee BK, Iyer VR. Genome-wide studies of CCCTC-binding factor (CTCF) and cohesin provide insight into chromatin structure and regulation. J Biol Chem. 2012; 287(37):30906–13.
36. Zlatanova J, Caiafa P. CTCF and its protein partners: Divide and rule? J Cell Sci. 2009;122(9):1275–84.
37. Schmidt D, Schwalie PC, Ross-Innes CS, Hurtado A, Brown GD, Carroll JS, Flicek P, Odom DT. A CTCF-independent role for cohesin in tissue-specific transcription. Genome Res. 2010;20(5):578–88.
38. Michaelis C, Ciosk R, Nasmyth K. Cohesins: Chromosomal proteins that prevent premature separation of sister chromatids. Cell. 1997;91(1):35–45.