Genomic and epigenomic landscapes to predict response to hypomethylating agent therapy in patients with acute myeloid leukemia

TAEKGYU LEE

Department of Medical Science

The Graduate School, Yonsei University

Genomic and epigenomic landscapes to predict response to hypomethylating agent therapy in patients with acute myeloid leukemia

TAEKGYU LEE

Department of Medical Science

The Graduate School, Yonsei University

Genomic and epigenomic landscapes to predict response to hypomethylating agent therapy in patients with acute myeloid leukemia

Directed by Professor SEUNG-TAE LEE

The Master's Thesis
submitted to the Department of Medical Science,
the Graduate School of Yonsei University
in partial fulfillment of the requirements for the degree of
Master of Medical Science

TAEKGYU LEE

December 2022

This certifies that the Master's Thesis
of TAEKGYU LEE is approved.

Thesis Supervisor: SEUNG-TAE LEE

Thesis Committee Member #1: JONG RAK CHOI

Thesis Committee Member #2: JI HYE HA

The Graduate School
Yonsei University

December 2022

ACKNOWLEDGEMENTS

어느덧 2년의 석사과정을 마치고 학위 논문을 제출하게 되었습니다. 지난 2년을 돌이켜보면 많은 경험과 성취를 이뤄냈지만 주변 분들의 도움 없이 결코 혼자서 할 수 없는 일들이었습니다. 그래서 저에게 많은 도움을 주신 모든 분들께 감사의 말씀을 전하고자 합니다.

우선 물심양면으로 지도해 주시고 부족한 점을 가르쳐주신 이승태 교수님께 진심으로 감사드립니다. 교수님의 조언 덕분에 2년의 학위과정을 좋게 마무리 지을 수 있었습니다. 감사합니다. 다음으로는 항상 연구실을 위해 애써주시고 든든한 버팀목이 되어주신 최종락 교수님께도 감사드립니다. 그리고 창의적인 연구를 할 수 있도록 도와주신 하지혜 교수님께도 감사드립니다. 마지막으로 존재만으로도 든든했던 신새암 교수님께도 감사의 인사를 전해드리고 싶습니다.

무엇보다 2년 동안 같이 동고동락했던 우리 연구실 선생님들께도 감사의 인사를 전해드리고 싶습니다. 항상 모르는 게 있어서 물어보면 최선을 다해서 알려주시고 친절하게 답변해 주신 윤우빈 선생님께 진심으로 감사의 말씀 전해드립니다. 다음으로, 바쁘신 와중에도 실험을 도와주신 이현아 선생님께 진심으로 감사드립니다. 선생님 덕분에 좋은 결과로 논문을 작성할 수 있었고 많은 가르침을 얻었습니다. 또한 연구실 살림꾼으로 항상 바쁘지만 연구실 내부의 일들을 도맡아서 챙겨주신 박미리 선생님, 너무 감사드립니다. 마지막으로 지금은 졸업하여 자리에 없지만 처음 입학했을 때부터 학사관련된 일들과 연구실 적응을 도와주었던 심예은 선생님께도 감사의 말씀을 전해드립니다.

이 외에도 주변에 많은 분들의 도움으로 석사학위 과정을 마무리 지을 수 있었습니다. 다시한번 진심으로 감사드리며 지난 2년동안의 좋은 기억과 경험들을 평생 간직하고 감사하며 지내도록 하겠습니다. 모든분들께 감사드립니다.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Genomic and epigenomic landscapes to predict response to hypomethylating agent therapy in patients with acute myeloid leukemia

TAEKGYU LEE

Department of Medical Science
The Graduate School, Yonsei University

(Directed by Professor SEUNG-TAE LEE)

Acute myeloid leukemia (AML) is a cancer of blood and bone marrow. AML is characterized by the rapid growth of abnormal cells and immature blast cells in the bone marrow and peripheral blood. Recently, it has been discovered that aberrant methylation interferes with normal hematopoiesis and is involved in tumor development and progression. As hypomethylating agents (HMAs) can inhibit DNA methylation, they have been used as an alternative treatment in AML. However, only a subset of AML patients respond to HMAs. Furthermore, patients who were sensitive to HMA in primary treatment develop secondary resistance because of repeated chemotherapy. Thus, understanding the drug resistance mechanism is essential for developing therapeutic strategies to overcome development of resistance. However, little is known about the associated predictive biomarkers or molecular markers. To establish novel resistance mechanisms of HMA therapy, this study provides insight into their epigenetic and genomic characteristics. Furthermore, transcriptional analysis reveals the gene expression environment status. A total of 28 acute myeloid patient samples before HMA treatment were obtained to analyze whole genome bisulfite sequencing (WGBS), whole exome sequencing (WES), and total RNA sequencing. Several putative biological pathways related to drug resistance were

identified, including cell differentiation and immune response. The HMA-sensitive group had genes that show low promoter methylation and high expression level associated with myeloid cell differentiation, erythrocyte development, and T cell activation. Furthermore, the responder group indicated relatively hypomethylated transcription factor binding sites around differentially methylated regions compared to the non-responder group. Patients who were resistant to HMA have substantial methylation increases on the hematopoiesis gene promoter and decreased expression. Indeed, inflammatory response was upregulated in the non-responder group. Based on the results, this study suggests that integrated next generation sequencing (NGS) studies such as WGBS, WES, and RNA-sequencing effectively identify biomarkers or biological responses for drug resistance.

Keywords: acute myeloid leukemia, hypomethylating agent, drug resistance, whole genome bisulfite sequencing, whole exome sequencing, total RNA sequencing, biomarker

Genomic and epigenomic landscapes to predict response to hypomethylating agent therapy in patients with acute myeloid leukemia

TAEKGYU LEE

Department of Medical Science
The Graduate School, Yonsei University

(Directed by Professor SEUNG-TAE LEE)

I. INTRODUCTION

Acute myeloid leukemia (AML) is a genetically heterogeneous hematological disease characterized by an increase in the blast cells in the bone marrow and expansion of immature myeloid progenitors. These characteristics result from the accumulation of genetic alterations in hematopoietic progenitor cells that impair normal hematopoiesis, especially self-renewal, differentiation, and proliferation[1-3]. Owing to abnormal hematopoiesis, hematopoietic stem cell and progenitor cell usually become a type of immature cell and cause bone marrow failure.

Recent discoveries revealed that epigenetic changes in genomic regions affect hematopoietic malignancy. It demonstrates a pivotal role in dysregulation of AML progression. The most studied and widely known epigenetic change is DNA methylation. In the human genome, DNA methylation occurs at the fifth carbon of cytosine, which generally comprises CpG dinucleotides[4] (Figure 1A). DNA methylation alters gene expression that is essential for normal physiological and molecular function (Figure 1B). During hematopoietic differentiation and maturation, each stage is controlled by transcription factor networks, and at the epigenetic level, distinct DNA methylation patterns are involved in molecular binding on specific DNA sequences. DNA methylation is also associated with chromatin structures. Once the methylation level of several regions having high

density of CpG dinucleotides is increased, the chromatin structure is altered and becomes compact. When the chromatin structure changes, transcription factors or enzymes for normal differentiation and proliferation cannot bind to specific sequence. In contrast, decreased DNA methylation probably causes upregulation of tumor-promoting gene expression and activates malignant pathways or facilitates transcription factor binding involved in oncogenesis[4-6].

Several studies have found certain genes affecting DNA methylation and binding ability for specific genomic sequences. In humans, DNA methylation is catalyzed by DNA methyltransferases such as *DNMT3A*, *DNMT3B*, and *DNMT1* genes. *DNMT3A* and *DNMT3B* are called de novo DNA methyltransferase that add new methylation patters in unmodified cytosine. These two enzymes are responsible for establishing DNA methylation during cell development and differentiation. Although *DNMT3A* and *DNMT3B* are highly expressed in early embryonic stages, they are decreased in expression at the end of cell differentiation. These genes have similar functions in that they modulate DNA methylation in humans but have distinct functions during embryonic development. *DNMT3A* generally methylates a set of genes and sequences during the late stage of embryonic development. *DNMT3B* methylates a broader region of genomic sequences in early time[7-9]. *DNMT1* is a maintenance methyltransferase that preserves the methylation pattern in the process of DNA replication. DNA is self-replicating, unlike other nucleic acids such as RNA. When self-replication proceeds, it forms a daughter strand. In this process, the *DNMT1* gene is involved in maintaining the methylation pattern of the parent strand in the newly synthesized daughter strand[10]. Unlike genes that form or preserve methylation in DNA, the *TET* gene family is well known for removing DNA methylation. The *TET* gene family is composed of *TET1*, *TET2*, and *TET3* that catalyze the oxidation of 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), and 5-carboxylcytosine (5caC)[11]. These oxidative conversions result in a loss of 5mC, which implies

decreased DNA methylation.

Many studies have been conducted on genes such as *DNMT* and *TET* genes that change DNA methylation over a long period of time. In addition, with the development of sequencing technology, many studies have determined how the patterns of methylation change according to specific mutations of those genes. In AML patients, approximately 25 % of patients have *DNMT3A* mutations, which are involved in de novo cytosine methylation. More than half of these mutations in AML samples are heterozygous missense alterations that cause abnormal catalytic domains in enzymes at residue R882, most commonly resulting in an arginine to histidine change. The R882 mutant protein has been shown to inhibit normal *DNMT3A*. It has also been reported that CGI (CpG island) shows focal DNA hypomethylation by *DNMT3A* R882 mutation compared to wild type[12,13]. *TET2* gene is mutated with high frequency in AML patients[14]. In AML samples, the primary effect of *TET2* mutation is to cause widespread DNA hypermethylation affecting up to 25 % of active enhancer elements. This indicates that *TET2* induces aberrant DNA hypermethylation in the enhancer region involved in myeloid differentiation and proliferation[15,16].

The genetic characteristics of AML patients vary widely, and treatments differ accordingly. AML patients are usually treated using a variety of anticancer drugs to remove leukemic cells from the bone marrow and blood[17]. Radiation therapy is another therapeutic option in case leukemic cells having invaded other organs or nerves[18,19]. Other methods that do not use anticancer drugs include implanting normal human hematopoietic stem cells to restore hematopoietic function. However, these treatments are not always suitable for AML patients. Certain patients are ineligible for high-intensity treatment such as anticancer drugs and hematopoietic stem cell transplantation. Therefore, low-intensity treatment such as HMA is required for elderly AML patients or those who react unfavorably to high-intensity chemotherapy.

Abnormal methylation changes the microenvironment of the cell and leads

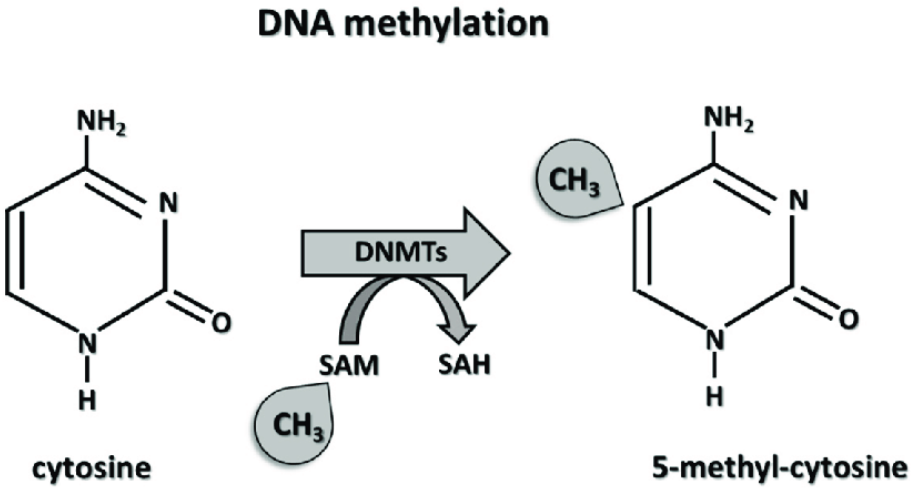
to tumor formation causing repression of tumor suppressor or oncogene activation[20-22]. HMA restores aberrant gene expression by regulating DNA methylation. Earlier, HMA was used in high doses; it turned out to be toxic to patients and failed to have an antitumor effect[20]. Research has shown that using azacitidine and decitabine in low doses has a substantial antitumor effect. Azacitidine was administered for subcutaneous injection at a dose of 75 mg/m^2 for 7 days every 28 days, and intravenous decitabine administration at a dose of 15 mg/m^2 every 8 h for 3 days, repeated every 6 weeks, were shown to have positive effects in patients with myelodysplastic syndromes (MDS). These two regimens were approved by the US Food and Drug Administration (FDA) in 2004 and 2006, respectively, for the treatment of MDS[23-25]. In addition, a 5-day regimen of decitabine given at a dose of 20 mg/m^2 by continuous intravenous infusion over 1 hour daily for 5 days, was approved by FDA in 2010 and has become a clinical standard[26]. Today, azacitidine and decitabine are widely used not only for the treatment of MDS but also older, high-risk AML patients[27,28].

Both azacitidine and decitabine are analogs of the nucleoside cytidine and act on depleting DNA methyltransferase (Figure 2). In brief, HMA metabolic process is comprise of the cellular uptake, intracellular activation, incorporation into nucleic acids, and inhibition of *DNMT* enzymes (Figure 3). The cellular uptake is mediated by nucleoside transporters[29,30]. Intracellular activation through the consecutive phosphorylation result in the active metabolites 5-azacitidine-triphosphate for azacitidine and 5-aza-2′-deoxycytidine-triphosphate (5-aza-dCTP) for decitabine. The first phosphorylation step is progressed by uridine-cytidine kinase (UCK) for azacitidine and deoxycytidine kinase (DCK) for decitabine (Figure 3). Owing to the activity of cytidine deaminase (CDA), which can inactivate cytidine analogs, azacitidine and decitabine are rapidly deactivated in the process of intracellular phosphorylation[31]. After consecutive phosphorylation, 5-aza-CTP and 5-aza-dCTP become incorporated into DNA during replication process. Azacitidine is mainly incorporated into RNA, only 10-

20 % of azacitidine acts on DNA, but decitabine is incorporated entirely into DNA. Incorporated drugs finally induce decreasing DNA methylation[32,33]. The resulting DNA demethylation leads to the reactivation of aberrantly silenced genes involved in multiple different pathways, such as apoptosis, DNA repair, differentiation, and angiogenesis [20,21,23,29] (Figure 3).

However, resistance to HMA is unavoidable despite the responsiveness of HMA initial treatments. In certain patients, even if resistance does not exist at the beginning (baseline), there are patients who gradually show resistance as the treatment cycle progresses. Therefore, it is clinically important to discover biomarkers that can predict the responsiveness of HMA before treatment. However, thus far, there is insufficient relevant research, and the molecular interactions resulting in resistance have not been identified. Therefore, in this study, we conducted whole genome bisulfite sequencing (WGBS) to identify the molecular mechanisms that confirm the resistance before HMA treatment, demonstrating the overall methylation state and differentially methylated regions existing between HMA sensitive (CR) and resistance group (NR). In addition, an analysis of the gene promoter part, which has an important effect on gene expression, was also conducted. Gene expression profiling was performed through RNA-seq analysis. In addition, genes that were differentially expressed in the two groups were identified, and these genes were selected as novel biomarkers that could distinguish the resistance of HMA. Finally, whole exome sequencing (WES) was performed to confirm the mutation landscape of each AML patient. As methylation varies with mutations in residues such as *DNMT*, *IDH*, and *TET* gene, it is crucial to identify genetic characteristics for each patient. In conclusion, based on these three types of next generation sequencing (NGS) techniques, tumor environment and biomarkers that predict the resistance of HMA were identified.

A



B

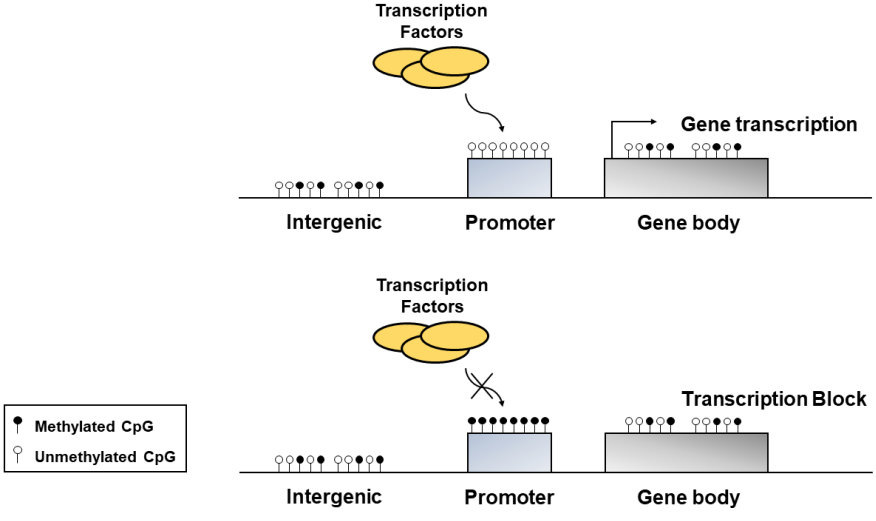


Figure 1. Overview of DNA methylation.

(A) Mechanism of DNA methylation[31]. (B) DNA methylation regulates gene expression.

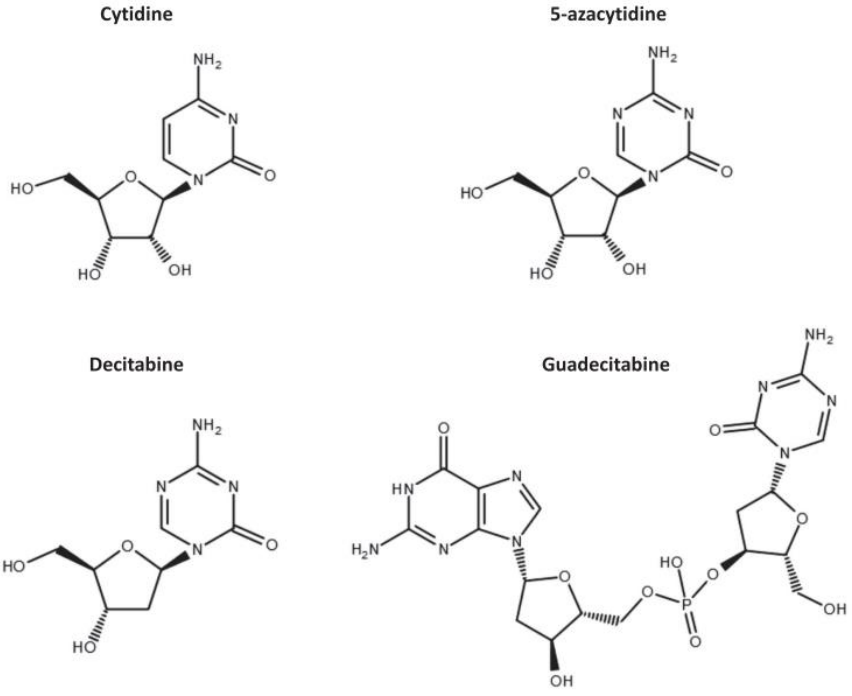


Figure 2. Azanucleoside DNA-hypomethylating agents[23].

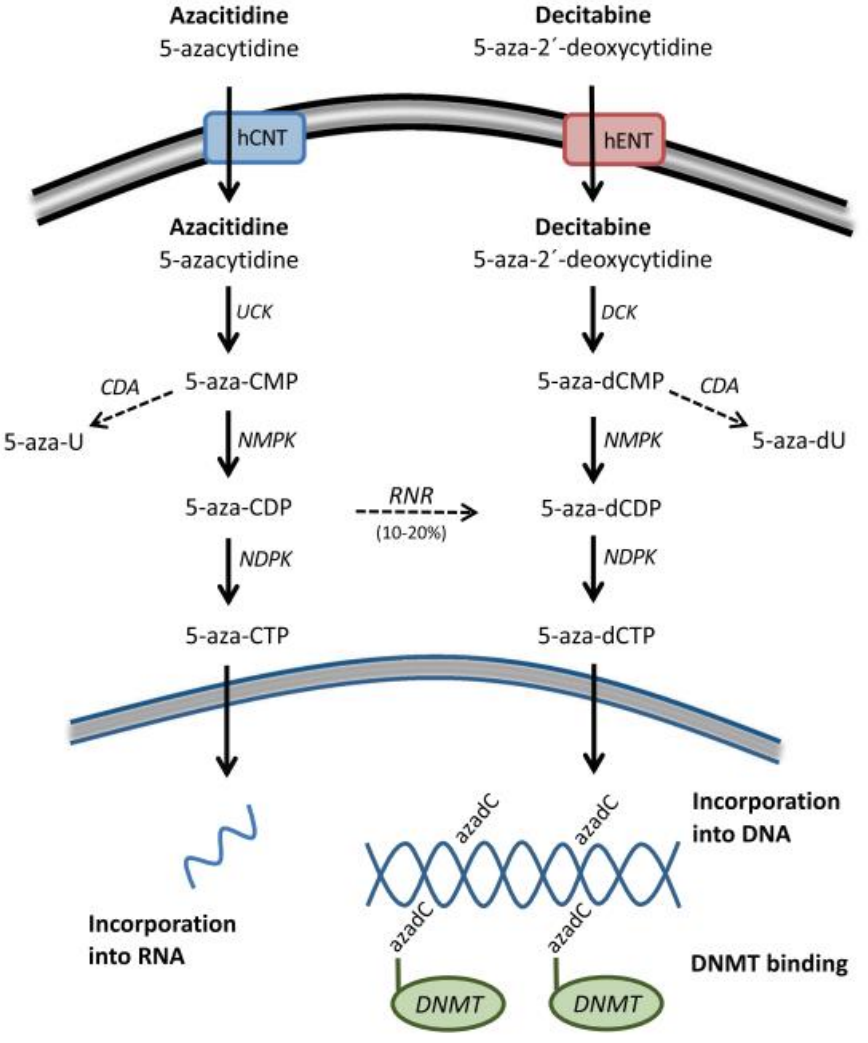


Figure 3. Schematic representation of Azacitidine and Decitabine uptake and metabolism[23].

II. MATERIALS AND METHODS

1. Patient cohort and sample collection

AML samples for epigenetic and genomic studies were collected from the baseline bone marrow (n=27) and peripheral blood (n=1) of AML patients treated using DAC (decitabine). Cohort selection was approved by the institutional review board of the Yonsei University (IRB no.4-2010-0732). Patients were classified after fourth cycle treatment of DAC regimen. A total of 28 AML baseline samples were obtained, and 15 of them show sensitive responses to DAC. This CR includes complete remission and complete remission with incomplete hematologic recovery to DAC, shows over 50 % BM blast reduction, and has 5–25 % BM blast. The other 13 samples were classified as NR, showing under 50 % BM blast reduction or having over 25 % BM blast.

2. DNA and RNA extraction from bone marrow and peripheral blood

Genomic DNA and total RNA were extracted using the AllPrep DNA/RNA kit (QIAGEN, Hilden, Germany) as per the manufacturer's instructions. The quality and concentration of genomic DNA and total RNA were measured using genomic DNA screen Tape and RNA screen Tape (Agilent Technologies, Santa Clara, CA, USA) using the 4150 TapeStation System (Agilent Technologies). To quantify DNA, the Qubit dsDNA HS assay kit (Invitrogen, Carlsbad, CA, USA) with the Qubit 2.0 Fluorometer (Invitrogen) were used.

3. Whole genome bisulfite sequencing, data processing, and analysis

(1) Sequencing

The unmethylated lambda DNA (Promega, WI, USA) (0.5 %) was added to each high molecular weight gDNA sample. The gDNA was bisulfite converted using the EZ DNA Methylation-Gold kit (Zymo

Research, CA, USA) with an elution volume of 15 μL. The bisulfite-treated DNA was assessed by a NanoDrop™ 2000/2000c Spectrophotometer (Thermofisher, MA, USA). The libraries were prepared from the Accel-NGS Methyl-Seq DNA library kit (v3.0) (Swift Biosciences, Ann Arbor, USA) with an input of 600 ng of sheared DNA following the manufacturer's protocol. The quality and quantity of the finished libraries were assessed using a combination of the Agilent D1000 ScreenTape (Agilent Technologies, Inc). 151 base pair (bp) paired-end sequencing was performed on an Illumina NovaSeq6000 sequencer (Illumina, San Diego, CA, USA) with 20 % PhiX.

(2) Data processing

Before mapping to a reference genome, FastQC (v0.11.9) was used to identify overall base and read quality. Adapter sequence was eliminated by Trim galore (v2.8) with default options. For methylation read alignment, reference genome indexing was performed using Bismark (v0.23.1dev) genome preparation command. All bisulfite sequencing data were aligned to the hg19 human reference version using Bismark[34]. After mapping, duplicate reads were removed using deduplicate Bismark option in Bismark software and Bismark bisulfite read mapper and methylation caller calculate methylation values at each CpG site.

(3) Data analysis

CpG methylation values on each strand were used input data for stepwise analysis procedures to divide methylation characteristics based on whether differences exist in CpG methylation between the CR and NR groups. To compare methylation level on specific genomic regions like promoter, exon, intron, and enhancer, principal component analysis (PCA) was conducted. For PCA of WGBS CpG methylation level, we used the

ggbiplot package to perform PCA analysis. PCA analysis indicates the association of principal components with chemoresistance response and DNA methylation. For global DNA mean methylation, genomic regions were divided by 10kb tiles and visualized using circlize R package.

(4) Differentially methylated region identification

The R package 'Methylkit'[35] was used to analyze the methylation level at all CpG and to identify differentially methylated regions (DMRs) with minimum CpG number requirement (CpG ≥ 1). DMRs were defined as those with FDR < 0.05 and absolute mean methylation difference value ≥ 20. Based on the mean methylation difference, hypermethylated and hypomethylated region in CR and NR group were defined. The R package 'annotatr' was used to annotate genes in each DMR. Annotatr is an annotation software using Grange R package. DMRs were annotated and identified where DMRs were located in specific genomic regions. Annotated genes in DMRs that have especially large differences in promoters were used to identify the associated biological pathways. These genes were input into cytoscape software to identify gene ontology enrichment.

(5) Transcription factor analysis

Transcription factors (TFs) that could bind to DMRs region were predicted from the HOMER (v4.11) software. Significant transcription factors were identified by Benjamini–Hochberg < 0.05. These transcription factors binding sites were obtained from JASPAR transcription binding site (TFBS) database in the regulation category of the UCSC table browser. From binding site bed file, each TFBS's average methylation was calculated, the difference between CR and NR group compared, and average methylation levels for visualization were generated.

4. Whole exome sequencing, data processing, and analysis

(1) Sequencing

Genomic DNA quantified using the Qubit BR dsDNA kit (Invitrogen). Approximately 50 ng of genomic DNA, extracted from bone marrow was prepared with the Twist Library Preparation EF Kit (Twist Bioscience, San Francisco, USA). The fragmentation time at 32°C was for 20 minutes followed by an enzyme inactivation at 65°C for 30 minutes. Target enrichment was performed with Twist Exome 2.0 Kit following the manufacturer's instructions (Twist Bioscience). The enriched DNA was sequenced on a NovaSeq6000 sequencer (Illumina), achieving approximately xx million reads per sample. Sequencing with a 151 bp, dual-indexed, paired-end sequencing configuration was performed.

(2) Data processing

Through our custom analysis pipeline, quality control and sequence analysis were conducted. The hg19 built as the reference sequence was applied for mapping and variant calling while using the Burrows–Wheeler alignment (BWA) tool (v 0.7.12). HaplotypeCaller and MuTect2 in the GATK package (v3.8-0) and VarScan2 (v2.4.0) were used to identify single-nucleotide variations (SNVs) and insertion and deletions (indels). Online databases including the Human Gene Mutation Database (HGMD), Online Mendelian Inheritance in Man (OMIM), Clinvar, dbSNP, 1000 Genomes, Exome Aggregation Consortium (ExAC), Exome Sequencing Project (ESP), and Korean Reference Genome Database (KRGDB) were used for analyses and variant annotation.

(3) Data analysis

Variants was classified using a scoring algorithm implemented based on the standards and guidelines established by the American College of

Medical Genetics (ACMG)[36]. We excluded genetic variants classified as benign or likely benign based on ACMG guidelines in NGS clinical reports. Then, variants were lined in order of higher probability of pathogenicity as per ACMG guidelines.

5. Total RNA sequencing, data processing, and analysis

(1) Sequencing

Total RNA sequencing of AML samples was performed on RNA samples that a had high RNA integrative number (RIN >= 5). Total RNA-seq libraries were prepared using Illumina TruSeq Stranded Total RNA Library Kit (Illumina) as per the manufacturer's instructions. The NovaSeq6000 system (Illumina) was used to perform massively parallel sequencing.

(2) Data processing

All sequencing data quality control was conducted using FastQC (v0.11.9) and RSEQC software. Trim Galore was used to trim adapter sequences. Processed sequence data were aligned using the STAR[37] alignment (2.7.10a) and mapped to the hg19 human reference genome with gencode release 19 (hg19). Used parameter for STAR alignment is "--twopassMode basic" to detect more splices' reads mapping to novel junction. Raw read counts were generated by "--quantMode GeneCounts" option to count number of reads per gene while mapping and normalized using DEseq2.

(3) Data analysis

The expression data of each gene was represented for normalized counts. Normalized counts were used to detect differentially expressed genes (DEGs) using DEseq2 following the recommended procedures. Gene

set enrichment analysis (GSEA) and Metascape were utilized to discover the specific biological pathways and molecular functions using differentially expressed genes.

6. Statistical analysis and visualization

Statistical analyses were conducted using R version 4.1.0. Statistical comparison for methylation difference was proceeded by student t-test and FDR correction. In transcription factor analysis, benjamini-hochberg q value was used to identify significantly enriched TFs on DMR. To distinguish significant mutation frequency in all detected mutations, fisher's exact test was conducted. For all tests, $p < 0.05$ and $q < 0.05$ were considered statistically significant. In visualization, the "complexheatmap" and "ggplot2" package in R were used to analyze and visualize.

III. RESULTS

1. Overview of patients cohort

To identify the incidence of hypomethylating agent resistance in acute myeloid leukemia patients, we performed whole genome bisulfite sequencing, whole exome sequencing, and total RNA sequencing. Each DNA or RNA was extracted from acute myeloid leukemia patients prior to hypomethylating agent therapy. This study considered 28 patient samples. In whole exome sequencing, all 28 samples were used for sequencing. However, some samples were excluded owing to improper quality in whole genome bisulfite sequencing (N=24) and total RNA sequencing (N=13). We classified these samples based on the response to hypomethylating agent therapy. The response evaluation time for the drug is the fourth treatment after baseline. The classification criteria for CR includes morphologic CR status and CR with incomplete hematologic recovery. Samples of patients with less than 50 % reduction or over 25 % BM blast were classified as NR group. Details of sequencing descriptions and patient's information are given in table 1, table 2 and appendix table 1.

Table 1. Demographic information of samples classified by HMA response

	Responder (CR) (N = 15)	Non-Responder (NR) (N = 13)
Age		
Range (y)	60-80	66-84
Average (y)	72.9	73.7
SD (y)	5.23	5.22
Sex		
Male (N)	7	10
Female (N)	8	3
Prior Treatment		
Decitabine (N)	15	13
BM Blast		
Mean (%)	52.5	56.67
Median (%)	50.3	47.2
WBC Count		
Mean ($\times 10^3/\mu L$)	54.82	46.98
Median ($\times 10^3/\mu L$)	26.77	21.1
HGB		
Mean (g/dL)	8.87	8.13
Median (g/dL)	9	8.1
PLT		
Mean ($\times 10^3/\mu L$)	103.93	81.76
Median ($\times 10^3/\mu L$)	86	61
Karyotype		
Normal (N)	10	4
Complex (N)	5	9

Table 2. Sequencing information of samples classified by HMA response

	Responder (CR) (N = 15)	Non-Responder (NR) (N = 13)
Whole genome bisulfite sequencing (N) (WGBS)	12	12
Whole exome sequencing (N) (WES)	15	13
Total RNA sequencing (N) (Total RNA-Seq)	8	5

2. Sequencing quality

WGBS. Before analyzing WGBS data, a quality check was conducted to identify bisulfute conversion rate, mapped reads, and average depth of each sample. The bisulfite conversion rate was 90 % or more in each sample. The average depth was 10x and the average alignment stat was approximately 80 % (Table 3).

WES. The average depth of WES data was 286x. More than 99 % of sequencing reads mapped to the human reference genome with an average of 76.5 % of reads matched with exome region (Table 4).

Total RNA-seq. Given the inadequate RNA read quality, more than half of samples were excluded. Considering 20–30 million reads per sample is recommended for expression analysis, we adopted samples that had over 30 million reads (Table 5).

Table 3. Sample quality for whole genome bisulfite sequencing

	BS conv rate (%)	Mean methylation	Mean Coverage (×)	Reads mapped	Alignment stats (%)
YM01	99.50	0.816	7.21	692,732,542	80.1
YM02	99.10	0.818	8.20	738,099,150	82.8
YM03	99.30	0.781	14.44	841,863,390	79.7
YM04	99.70	0.790	12.18	665,153,124	70.2
YM05	99.60	0.822	15.94	807,890,996	79.2
YM06	99.30	0.822	6.51	815,828,392	78.9
YM07	99.50	0.802	11.82	834,813,434	78.2
YM08	99.70	0.789	14.42	800,552,794	79.8
YM09	99.20	0.773	17.98	949,383,316	81.5
YM10	99.00	0.781	8.46	816,328,290	82.6
YM11	99.10	0.819	15.72	936,527,536	81.8
YM12	99.30	0.791	14.69	912,984,078	83.8
YM16	99.30	0.803	14.98	855,882,244	81.4
YM17	99.30	0.801	5.75	709,028,070	80.7
YM18	99.00	0.808	10.29	677,994,402	82.0
YM19	99.20	0.829	12.25	669,248,158	82.8
YM20	99.40	0.770	4.59	725,678,530	79.6
YM21	99.30	0.796	6.23	747,363,574	81.7
YM22	99.40	0.831	9.09	644,383,304	78.3
YM23	99.50	0.821	7.94	644,207,370	77.8
YM24	99.30	0.779	7.14	882,473,634	80.5
YM25	99.20	0.800	5.25	844,381,192	82.6
YM26	99.40	0.804	13.83	705,382,264	82.2
YM27	99.00	0.813	4.20	786,379,798	80.6

Table 4. Sample quality for whole exome sequencing

	Mapped bases	Mapped reads	Paired reads	On Target (%)	Average depth (×)
YM01	17,701,694,988	130,592,361	129,981,664	79.7	362
YM02	19,024,767,450	132,154,291	131,730,212	74.8	366
YM03	15,326,401,347	107,736,517	107,321,236	77.6	305
YM04	13,343,517,782	92,772,405	92,425,032	75.9	260
YM05	13,249,871,024	94,109,299	93,641,200	77.1	263
YM06	12,436,009,494	88,698,204	88,337,026	78.0	249
YM07	13,521,340,910	94,861,554	94,501,100	77.2	268
YM08	11,923,312,193	84,631,353	84,068,452	77.1	236
YM09	15,866,196,381	112,371,095	111,818,118	76.4	311
YM10	13,401,943,412	93,416,008	93,078,368	76.2	262
YM11	12,129,300,329	84,652,940	84,313,156	75.4	235
YM12	17,377,087,951	120,898,187	120,450,194	75.1	335
YM13	11,660,422,605	81,520,539	81,252,562	76.3	229
YM14	14,454,407,236	100,591,096	100,216,328	75.2	279
YM15	16,854,260,925	119,699,257	119,280,886	75.7	327
YM16	16,037,315,990	114,027,137	113,607,838	77.1	317
YM17	13,452,200,037	93,187,519	92,830,694	74.3	257
YM18	10,223,924,731	72,337,789	72,082,540	77.6	204
YM19	13,081,394,081	91,408,148	91,062,118	75.4	254
YM20	18,261,226,195	127,660,914	127,239,298	76.4	358
YM21	15,241,555,200	105,887,638	105,480,504	75.7	297
YM22	16,302,699,366	113,796,125	113,452,470	75.8	318
YM23	11,363,931,365	81,097,492	80,763,134	78.7	230
YM24	17,208,954,325	123,316,630	122,760,374	78.4	346
YM25	17,301,448,977	121,610,291	121,197,498	76.2	339
YM26	14,930,368,728	104,835,550	104,423,130	77.2	296
YM27	13,667,884,471	96,157,186	95,747,722	76.5	269
YM28	13,465,512,378	93,874,362	93,580,164	75.4	261

Table 5. Sample quality for total RNA sequencing

	Total reads	Mapped reads	Uniquely mapped read (%)
YM01	40,955,039	33,649,214	82.16
YM02	113,353,004	56,439,235	49.79
YM03	43,343,273	37,938,497	87.53
YM05	84,600,527	65,671,360	77.63
YM06	66,794,545	36,866,195	55.19
YM10	74,864,206	57,169,470	76.36
YM13	35,970,634	25,457,236	70.77
YM15	53,171,721	27,752,358	52.19
YM16	37,327,556	28,477,655	76.29
YM20	70,613,453	47,915,552	67.86
YM23	72,007,530	51,541,046	71.58
YM24	59,305,277	37,434,873	63.12
YM28	133,846,903	72,433,703	54.12

3. Similar global methylation characteristics between responder and non-responder group

To study overall AML epigenetic features classified according to the response of the hypomethylating agents, we performed whole genome bisulfite sequencing with 24 AML patients. Of the 24 AML patients, 12 patients showed sensitivity to DAC whereas the remaining had poor response to the drug.

To inspect differences in DNA global methylation patterns, we created CpG methylation maps that could identify the mean methylation values corresponding to consecutive 10 kilo base (kb) of tiles for each sample (Figures 4 and 5). These maps showed widespread and variable epigenomic CpG levels. In terms of drug response, we observed that the methylation of CR and NR group was variable.

Variation of methylation appears to be influenced by aberrant gene expression, genetic mutations, and demographic factors such as age, ethnicity[46,47] as well as environmental factors such as nutrient supply, stress, and radiation[48]. Among these factors, we tried to determine whether age affects methylation level in this study, and correlation analysis showed that this association is not significant (p=0.9786, cor=0.00579, n=24). Considering methylation patterns of both groups are similar, we generate a density plot of all 1kb blocks group mean methylation (Figure 6). We confirm that CR and NR group mean methylation is almost distributed within ranges (0.7–1).

Given variation and similarity of methylation in AML patients classified by DAC response, we investigated the methylation patterns of overall and specific region in genome to identify properties that could separate the DAC response. In epigenetics, especially DNA methylation, promoters are considered important regulatory factors that have binding sequences for various transcription factors for transcription or other molecules. In the promtor, the CpG island (CGI), which is spreaded overall genome, is highly distributed. Therefore, these domains are closely related to DNA methylation and play a

pivotal role in gene transcription and expression[51,52]. In this study, we defined promoter regions from transcription start site (TSS) to upstream 2000 bp and investigated personal and group mean methylation. Promoters of each group exhibit methylation canyon that show lower methylation than upstream and gene body regions (Figure 7). This demonstrates that promoter region has low methylation values that enable interaction with various molecules. To inspect the difference of percent methylation of promoter, methylation of each AML patient was calculated by individual CpG sites, and no significant differences ($p < 0.05$) were found (Figure 8). In line with such findings, PCA present that promoter methylation did not show specific patterns for each group and shows high variability. In the same context of promoter methylation, CpG islands (CGI) methylation also shows similar patterns as the promoter (Figures 9 and 10). Thus, methylation of AML patients between CR and NR groups on promoter and CpG islands, known as regulatory region on human genome, showed hypervariable status and comparable methylation level.

Recently, it has been reported that the gene body is also involved in gene expression[53]. Furthermore, enhancer region or downstream region are also affected by DNA methylation and regulate gene expression[54]. To characterize such regions excluding promoters or CpG islands, we adopted the same analysis method previously applied and found no evidence of methylation difference of such regions (Appendix Figure 1).

Taken together, acute myelod leukemia whole genome methylation profiles revealed that overall and global methylation were similar whether patients responded to DAC or not and generally represented hypomethylation on promoter and CGI region in both groups.

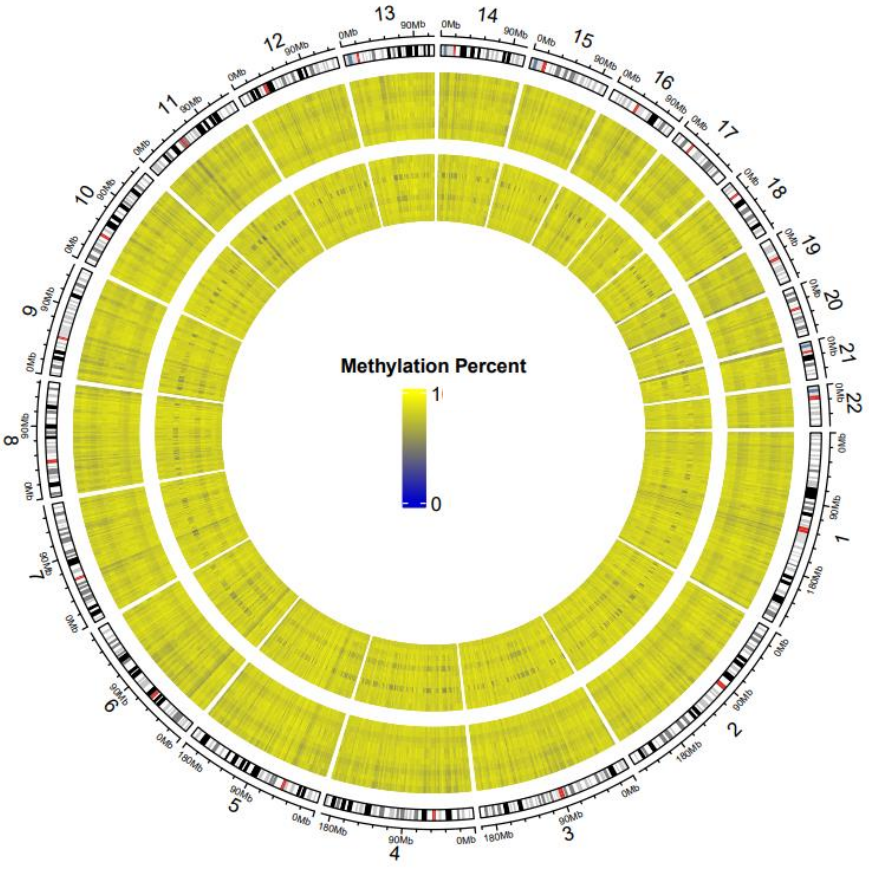


Figure 4. Methylation circlize plot.

The circlize plot shows global methylation levels for each chromosome except for sex chromosomes. The innermost circle indicates methylation of decitabine resistance patients. The middle circle represents methylation of responder group.

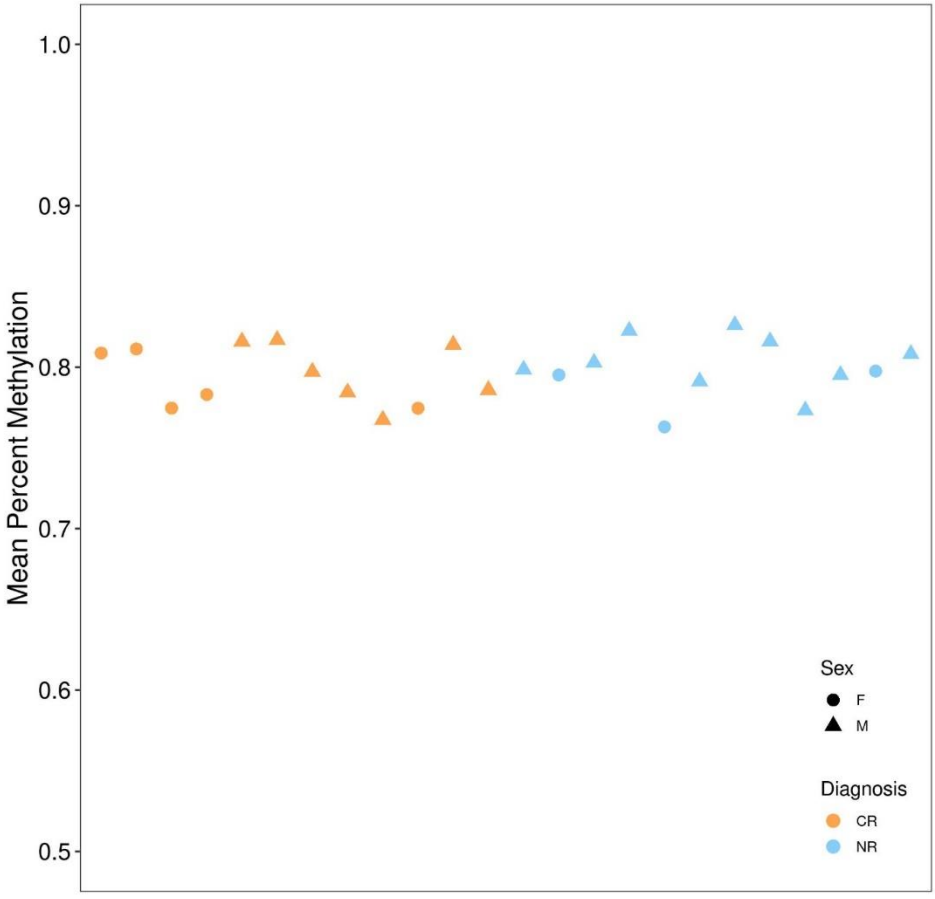


Figure 5. Average methylation plot.

All CpG methylation data was calculated to obtain the mean methylation level. This plot shows mean methylation of each sample as well as the methylation according to gender. CR, Responder, NR, Non-responder, F, Female, M, Male.

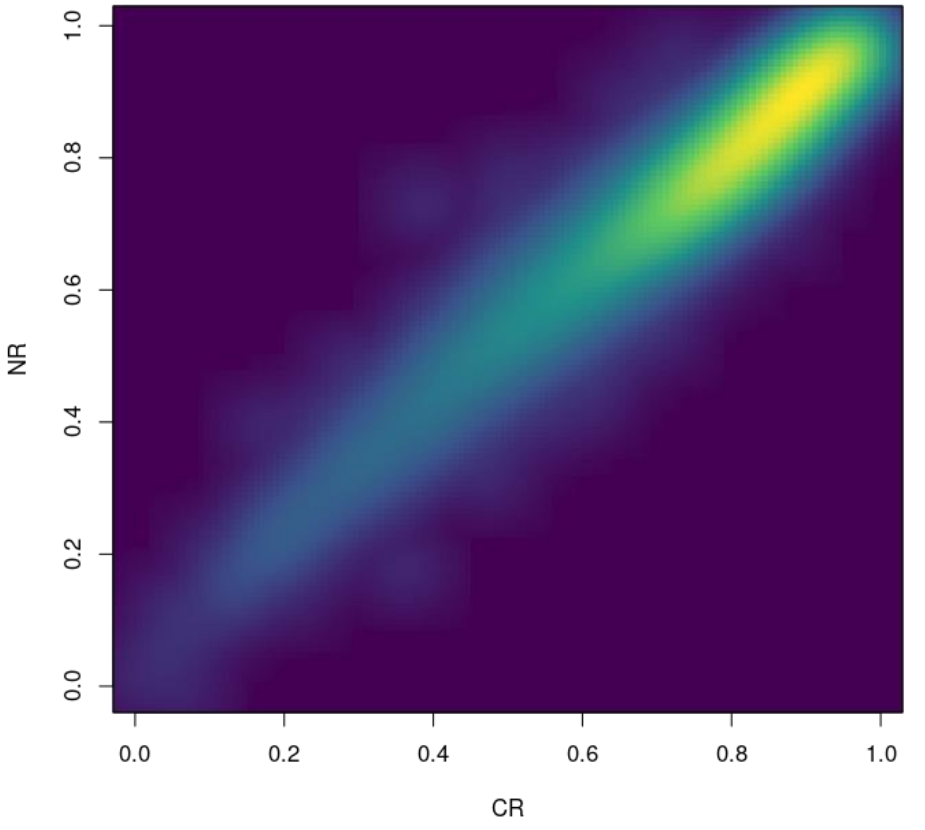


Figure 6. Methlylation scatter plot

The scatter plot shows the distribution of methylation levels divided into 1kb window. In both CR and NR groups, the distribution of methtylation level is more prevalent between range 0.7–1. CR, Responder, NR, Non-responder.

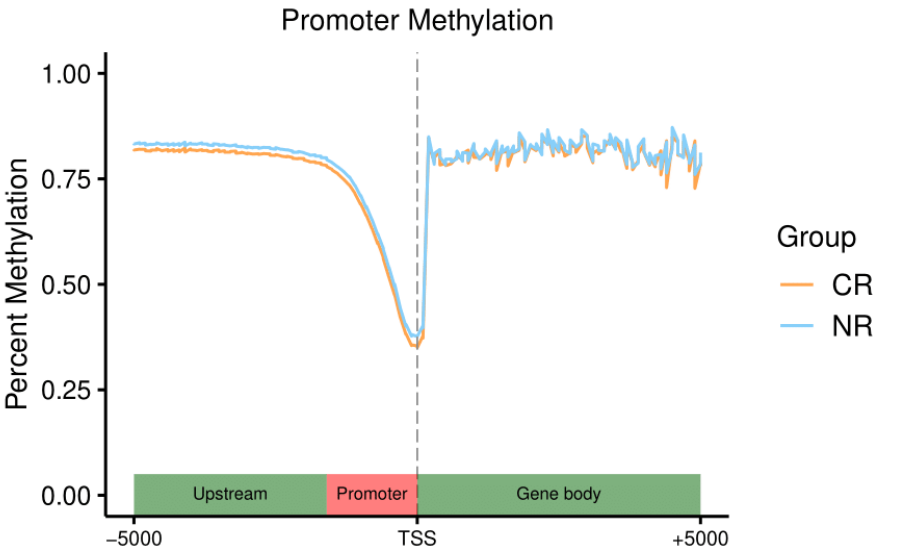


Figure 7. Average promoter methylation

The promoter regions were defined 2000 bp upstream from the TSS. Each line represents the average methylation level of each group. In both CR and NR groups, a canyon was formed in the promotor region. The difference in methylation between the two groups was not significant ($p > 0.05$). TSS, Transcription start site, CR, Responder, NR, Non-responder.

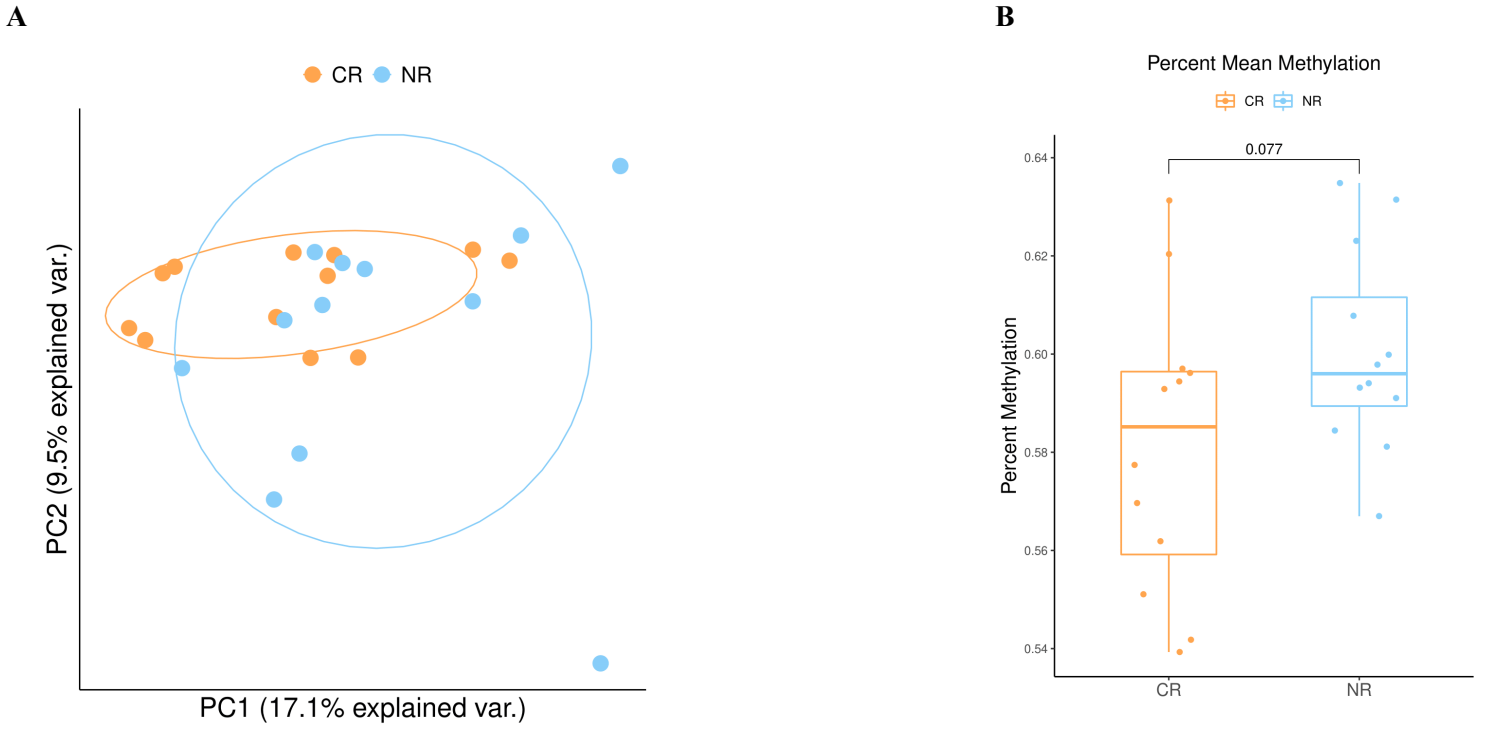


Figure 8. Promoter methylation characteristics.

This plot indicates promoter methylation of CR and NR group. (A) PCA of promoter methylation (B) Boxplot shows individual promoter methylation. CR, Responder, NR, Non-responder.

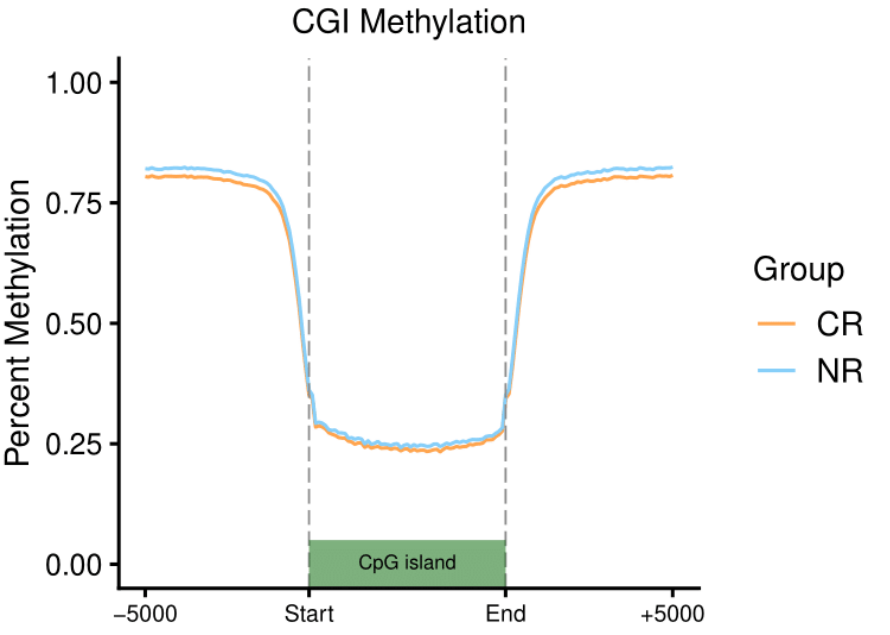


Figure 9. Average CpG island methylation.

Methylation of all CpG islands. The plot shows that methylation of CpG islands decreased compared to up- or downstream regions. The difference in methylation between the two groups was not significant ($p > 0.05$). CR, Responder, NR, Non-responder.

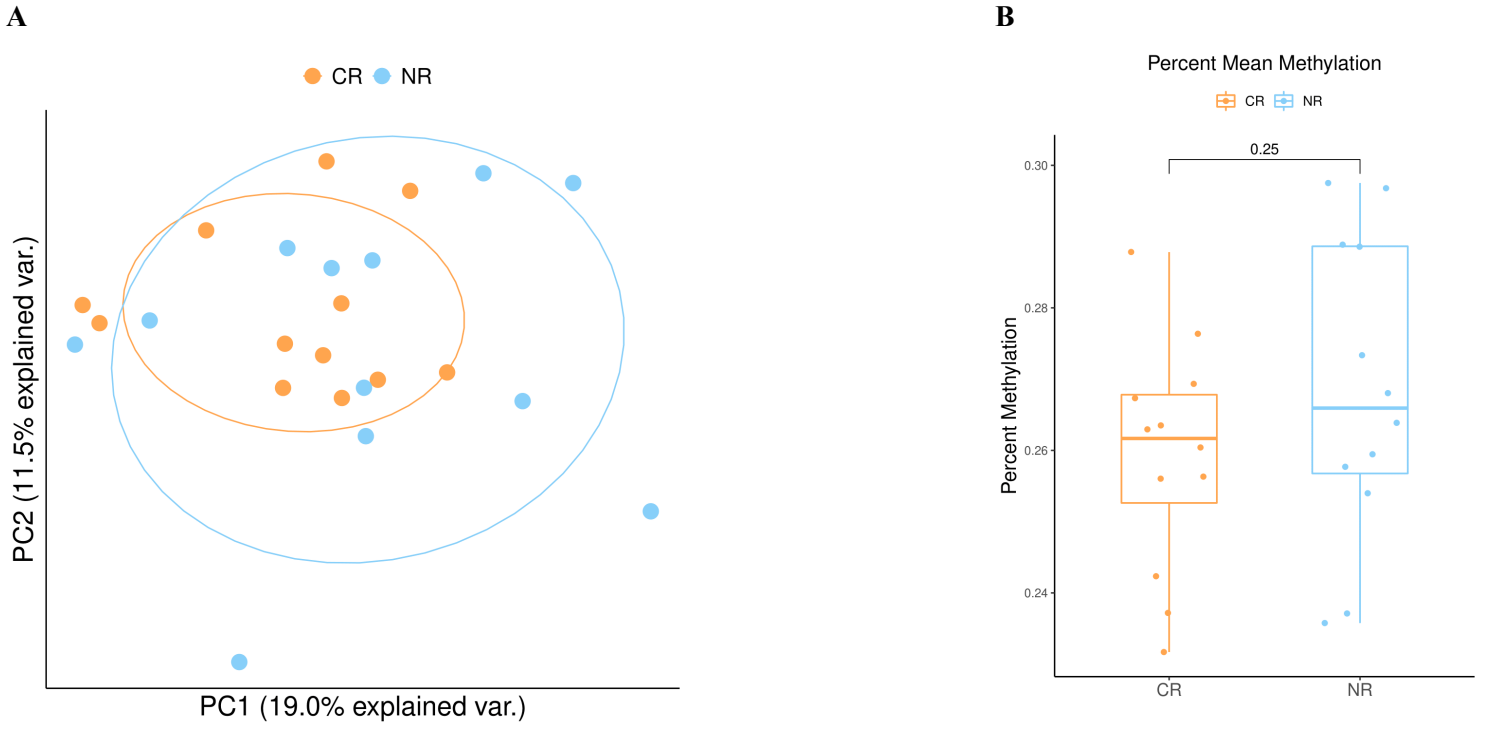


Figure 10. CpG island methylation characteristics.

This plot indicates CpG island methylation of CR and NR group. (A) PCA of CpG island methylation (B) Boxplot shows individual CpG island methylation. CR, Responder, NR, Non-responder.

4. Hypermethylated genes are associated with normal hematopoiesis and hematopoietic cell differentiation

Previous studies demonstrated that overall DNA methylation of AML patients are variable and have similar global mean methylation per group. Therefore, drug response cannot be distinguished only in overall methylation. To identify DNA methylation difference at baseline associated with drug response, we analyzed differentially methylated regions (DMRs).

To assess DMRs between CR and NR group, we used established methods (Methylkit). This analysis identified a total of 3,598 DMRs with a 20 % mean methylation difference ($q < 0.05$) (Figure 11). Particularly, 2,118 DMRs were hypermethylated in the NR group. These DMRs were almost enriched in intergenic region and introns. In contrast, 1,480 DMRs were hypomethylated in NR. DMRs were also enriched in CpG sparse region (Figure 12). The difference of methylation in the hypermethylated or hypomethylated DMRs region indicates a substantial increase or decrease (Figure 13).

Based on the fact that the promoter or 5'UTR closely located to the TSS is involved in gene expression, we annotated each DMR to define the gene symbol and where they are located on genome. Then, we used these genes, which are located in the regulatory region, to identify which biomolecular mechanisms and biological pathways affect the DAC response. Gene ontology analysis of the genes linked to differentially methylated CpG rich loci between CR and NR revealed enrichment terms such as granulocyte differentiation, hemoglobin complex and erythrocyte development, and leukocyte differentiation (Figure 14) (Appendix Figure 2). Data from gene enrichment analysis revealed that methylation of genes that are important to normal hematopoiesis are hypomethylated in CR, and the gene set that determines this ontology was primarily comprised of *GATA1, GATA2, BCL11A, TAL1, RUNX1,* and *NFE2* genes (Figure 15). The results of enrichment analysis for

hypomethylated genes in NR were not detected because the number of genes were not sufficient.

Taken together, a number of hypermethylated regions were slightly higher in the group that shows resistance response to DAC. The DMR methylation status of each individual differs enough to distinguish the group. Regions that have highly different frequencies in DNA methylation were intergenic, exon, and intron regions, not promoters, which are regulatory regions involved in gene expression. A small subset of genes that were differentially methylated in regulatory regions, especially hypomethylated genes in CR group, suggests that biological pathways such as hemoglobin complex, granulocyte differentiation, and erythrocyte development were significantly enriched in the DAC-sensitive CR group.

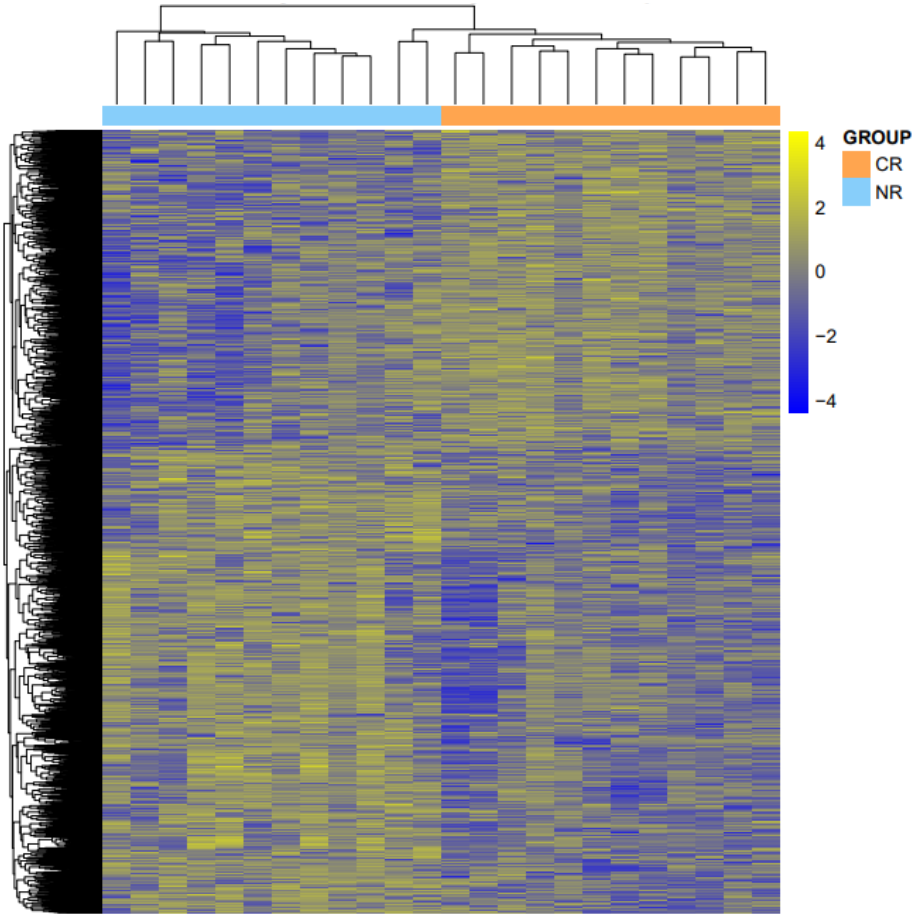


Figure 11. Heatmap of differentially methylated regions.

Total 3,598 differentially methylated region of CR and NR group. Heatmap indicates normalized (Z score) methylation values. CR, Responder, NR, Non-responder.

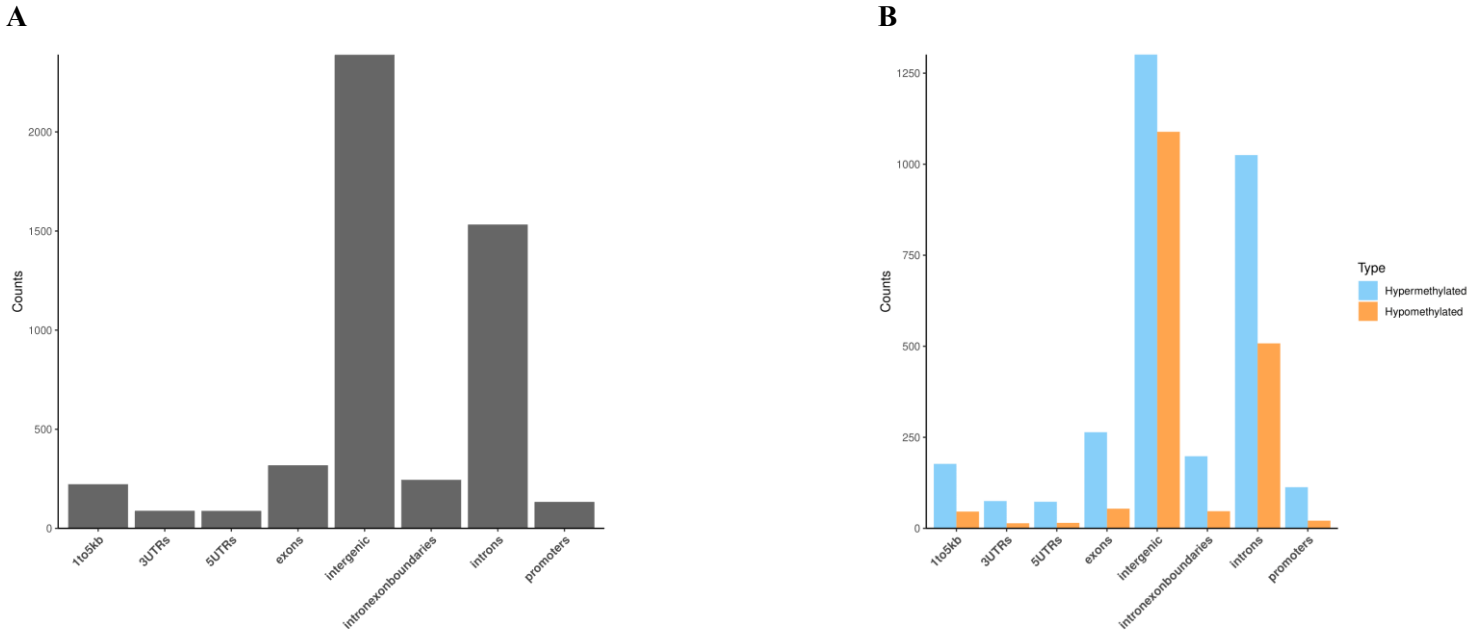


Figure 12. Annotated differentially methylated regions.

This result indicates enriched regions for differentially methylated regions (A) total enriched regions on genomic regions (B) Bar plot represents the counts for the hypomethylated region and the hypomethylated region in NR group. Most DMRs are enriched in the intergenic or intron region. DMR, Differentially methylated region. NR, Non-responder.

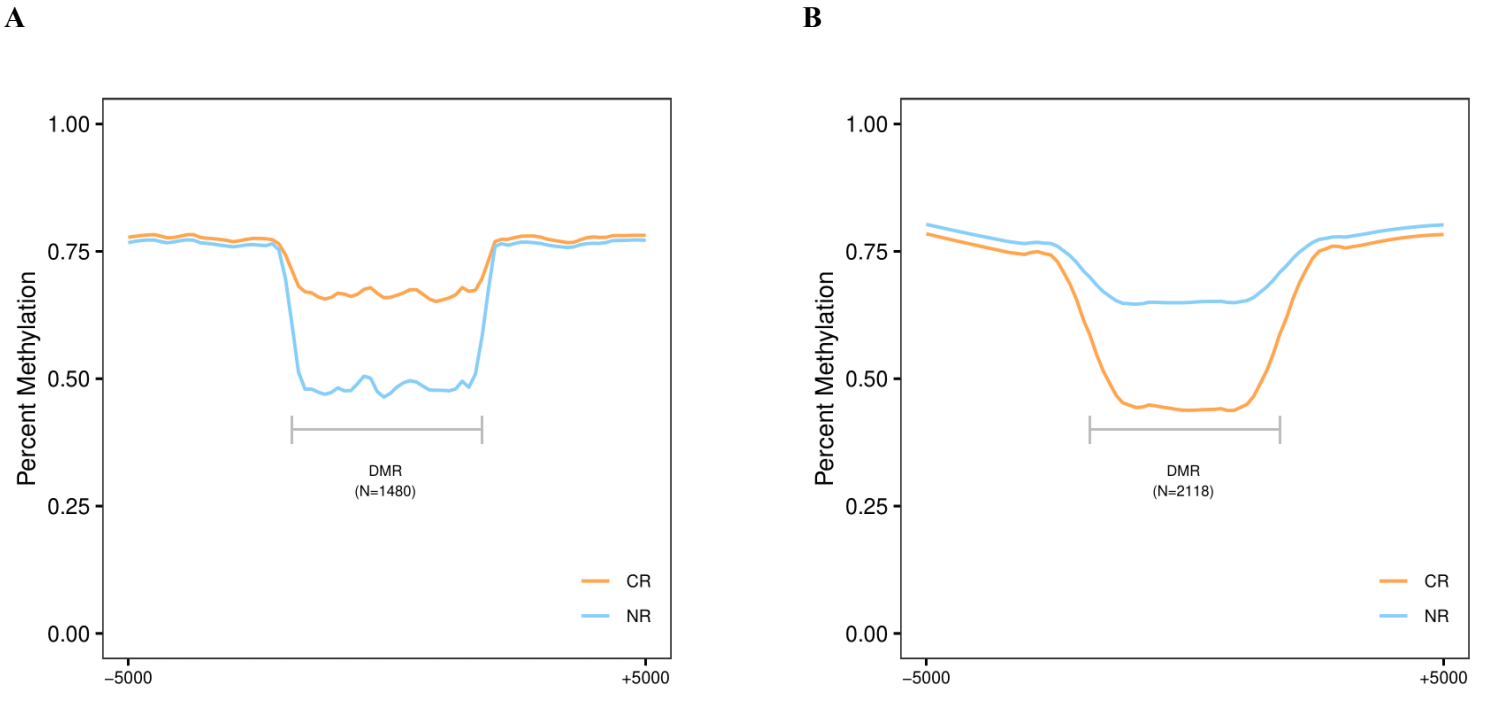


Figure 13. Methylation of differentially methylated regions.

Average methylation of DMRs shows substantial decrease or increase (A) Mean methylation values of 2,118 hypermethylated DMRs and (B) Mean methylation values of 1,480 hypomethylated DMRs for CR and NR. DMR, Differentially methylated region, CR, Responder, NR, Non-responder.

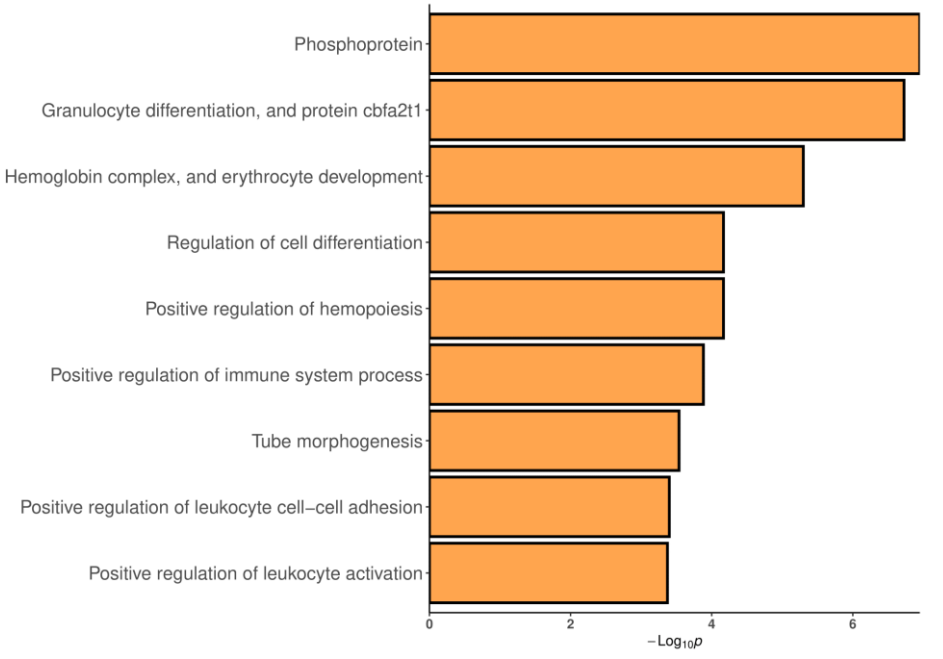
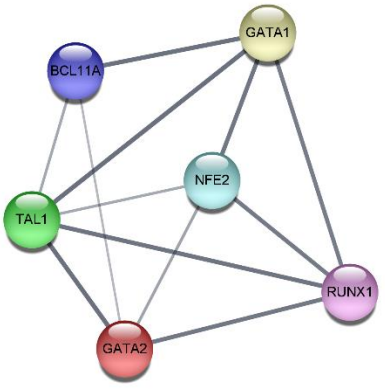


Figure 14. Gene ontology plot consist of promoter hypomethylated genes in responder group.

A



B

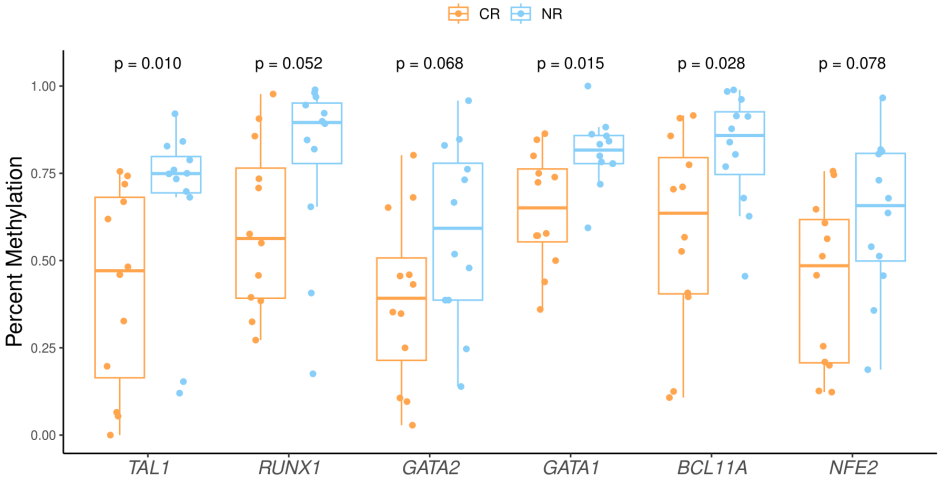


Figure 15. The gene set associated with hematopoiesis.

Promoter hypomethylated genes related to hematopoiesis (A) The core gene set associated with normal hematopoiesis (B) The promoter methylation of hemoglobin complex, erythrocyte development and granulocyte differentiation associated genes in CR and NR group. CR, Responder, NR, Non-responder.

5. Enrichment of specific transcription factors related to cell differentiation and hematopoietic development

After analyzing differentially methylated regions, we defined that they correspond to the specific transcription binding site (TFBS). Transcription factor (TF) is a protein that modulates the transcription from DNA to messenger RNA. Epigenetic modifications such as DNA methylation have the potential to modulate TFs binding ability. To define the binding sites in DMRs, we used HOMER software, which is a binding sites prediction tool for methyl-seq analysis. A total of 57 transcription factors were enriched around DMRs while 45 of 57 TFs were significantly enriched ($q < 0.05$) (Table 6). While analyzing TF binding sites in DMRs, we further investigated the methylation changes of TF binding sites between CR and NR group using the WGBS single CpG data. To obtain binding sites of each TF, we gained data from JASPAR TFs binding sites database. The average methylation difference was calculated by confirming the binding sites for significantly enriched TFs. The difference in average methylation of 30 of 45 TFs was significant ($q < 0.05$). This result shows that for all TFs, binding sites methylation were observed to be higher in the NR than CR group (Figure 16). In other words, the elevated methylation level of these TFs binding sites in the NR group suggest that the binding potential of TFs could be inhibited, and the resulting biological pathway also be impaired.

Our TFBS methylation study found that the transcription factors that bind to differentially methylated region and average methylation at each binding sites were far higher in NR group. We next performed gene ontology analysis on these transcription factors. Gene ontology analysis shows that these TFs are significantly associated with cell fate commitment. This includes pathways such as leukocyte differentiation, hematopoietic or lymphoid organ development, regulation of hemopoiesis, and myeloid leukocyte differentiation. Subsequently, embryonic organ development and embryonic morphogenesis

were also enriched (Appendix Figure 4). Therefore, probabilities for inhibited normal hematopoiesis or embryonic development exist in the NR group, and these results have the same context as the DMR analysis.

Consequently, methylation difference in several transcription binding sites involved in normal hematopoiesis was confirmed. Although methylation of transcription binding sites was observed in DMR of both groups, mean methylation values were significantly different.

Table 6. Top 21 known motifs significantly enriched in DMRs and associated with hematopoiesis

Motif	Transcription factor	q-value (Benjamini)	% of Targets Sequences with Motif
	*STAT3*	0.0002	28.77 %
	*STAT1*	0.0004	19.71 %
	*GATA2*	0.0013	38.74 %
	*GATA4*	0.0023	55.98 %
	*RUNX2*	0.0038	36.99 %
	*NF1*	0.0038	11.34 %
	*EBF1*	0.0109	21.57 %
	*STAT5*	0.0124	20.07 %
	*BCL6*	0.0126	54.70 %
	*RUNX1*	0.0126	42.16 %
	*STAT6*	0.0126	29.77 %
	*GATA1*	0.0126	33.52 %
	*ETV2*	0.0152	33.07 %
	*ETS1*	0.0162	33.55 %
	*ELF4*	0.0210	33.41 %
	*GATA6*	0.0372	52.75 %
	*HOXD10*	0.0424	64.56 %
	*JUNB*	0.0443	23.79 %
	*IRF4*	0.0455	24.99 %
	*HNF6B*	0.0474	58.37 %
	*BATF*	0.0476	29.35 %

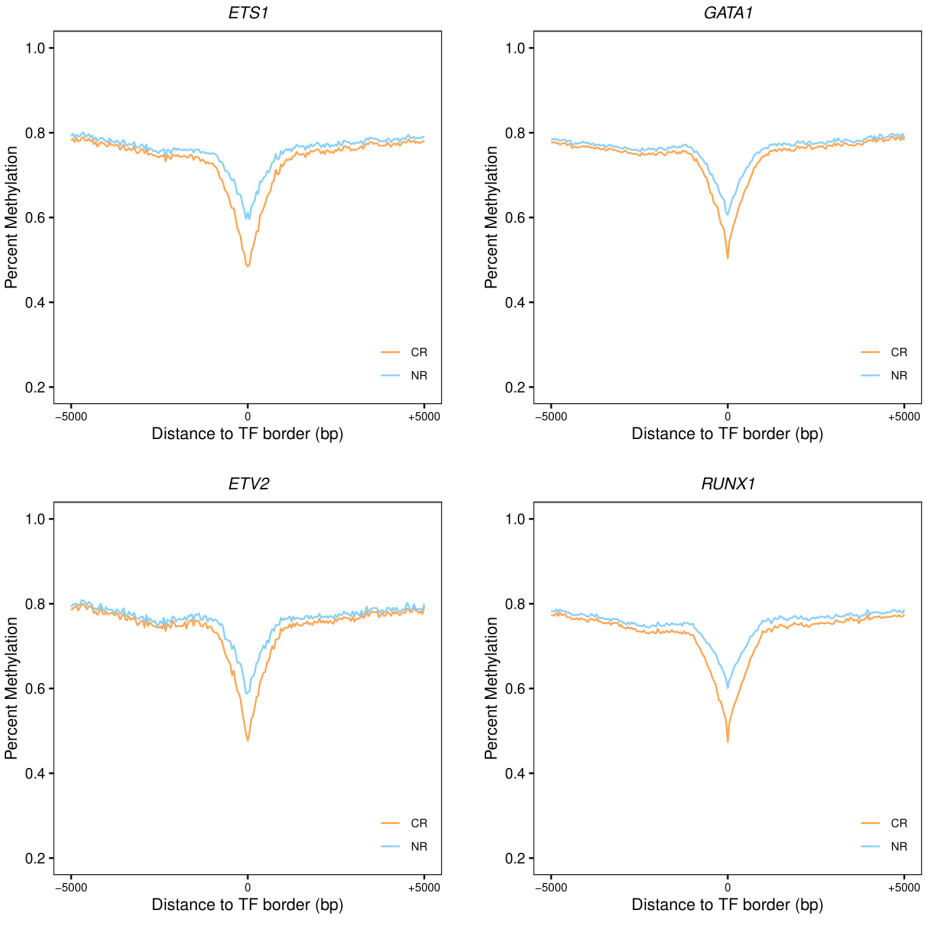


Figure 16. Enrichment of specific transcription factors binding sites on differentially methylated regions.

Significantly enriched transcription factors and its binding sites methylation is almost hypomethylated in CR in average than NR. CR, Responder, NR, Non-responder.

6. Mutational landscape of the samples and methylation associated with mutations

Somatic mutations in leukemogenesis and epigenomic modifying genes are prevalent in AML. Many recent studies have reported that *TET2* and *DNMT3A* are associated with drug response to HMA in AML and related hematological disorders. To determine whether specific genetic alterations are associated with DAC sensitivity or resistance in acute myeloid leukemia, we conducted whole exome sequencing of AML patients before DAC treatment. Mutational analysis of 28 AML baselines were identified to reveal their molecular profile. Genetic alterations of *TET2*, *DNMT3A*, *IDH2* and *TP53,* known as modifying DNA methylation and most frequently mutated genes in AML, were identified in more than one patient. Of these genes, *TET2* (18 %) was most frequently mutated in our cohort, followed by *DNMT3A* (14 %), *IDH2* (7 %), and *TP53* (7 %) (Figure 17). Mutations only found in each group are shown in Appendix Figure 5.

***TET2* mutations.** Only 18 % (5/28) of cases had somatic *TET2* mutations regardless of their response to DAC. Two patients indicated missense and one had missense and duplicate mutation in CR (S509*, L1398P, Q831*, Q876*). In detected mutations in *TET2* gene, Q876* results in premature truncation of the *TET2* protein. In DAC resistance group, one patient had *TET2* (H721Qfs*2 and Q910*) mutation. The other had V1417N mutation.

***DNMT3A* mutations.** *DNMT3A* mutations were observed in NR group. Only one mutation appeared in CR group (L815P). Three *DNMT3A* mutations (R882H, Q249*, R736H) were detected in NR group. Each mutation occurred in the catalytic domain of *DNMT3A* and truncate normal *DNMT3A* function, resulting in hypomethylation.

***IDH2* mutations.** One *IDH2* missense mutation (R172K) occurred in each CR and NR group. This mutation confers a gain of function mutation as

indicated by the increased conversion of alpha-ketoglutarate to the onco-metabolite (2HG).

***TP53* mutations.** *TP53* mutations only appeared in CR group. All detected mutations were misssense mutations. One is an R282W mutation that lies within the DNA binding domain of the *TP53* protein. The other C124* mutation disrupts the DNA binding domain and causes loss of the oligomerization domain.

The mutation frequencies of the two groups were compared, revealing that no somatic mutation was significantly correlated with response to DAC in our AML cohort ($p > 0.05$). Finally, we classified the observed mutations based on their normal function and identified which pathway had a large distribution of gene mutation. However, there was also no significant difference in the frequency of mutations occurring in CR and NR ($p > 0.05$) (Figure 18). Furthermore, it is challenging to find the aberrant DNA methylation that occurs related to the mutation (Appendix Figure 3).

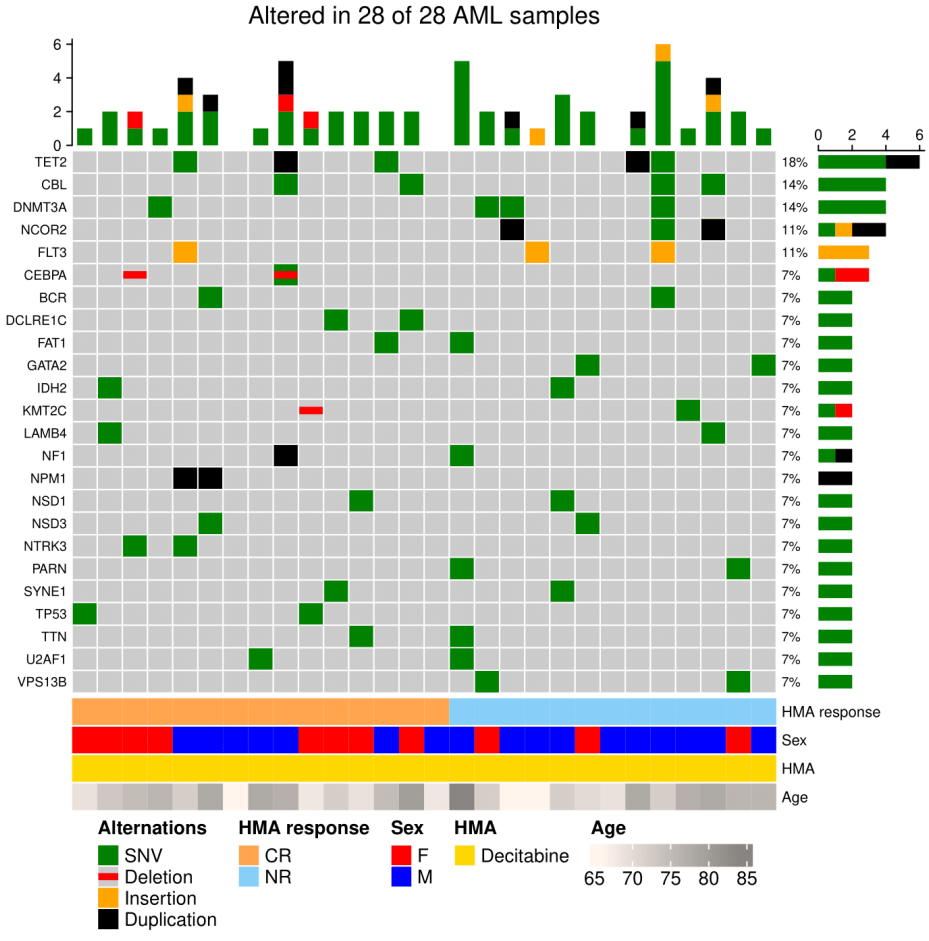


Figure 17. Mutational profiles of 28 AML patients.

Oncoprint showing mutated genes that exist at least two samples. The central plot shows the types of mutations in each sample. The upper bar graph indicates the number of mutations for each patient. The right side bar plot show overall frequencies for all samples in each gene. The bar plot showed in bottom part of the plot shows clinical features of each patient. CR, Reponder, NR, Non-respoonder, F, Gemale, M, Male, Del, Deletopn, Ins, Insertion, Dup, Duplication, HMA, Hypomethylating agent.

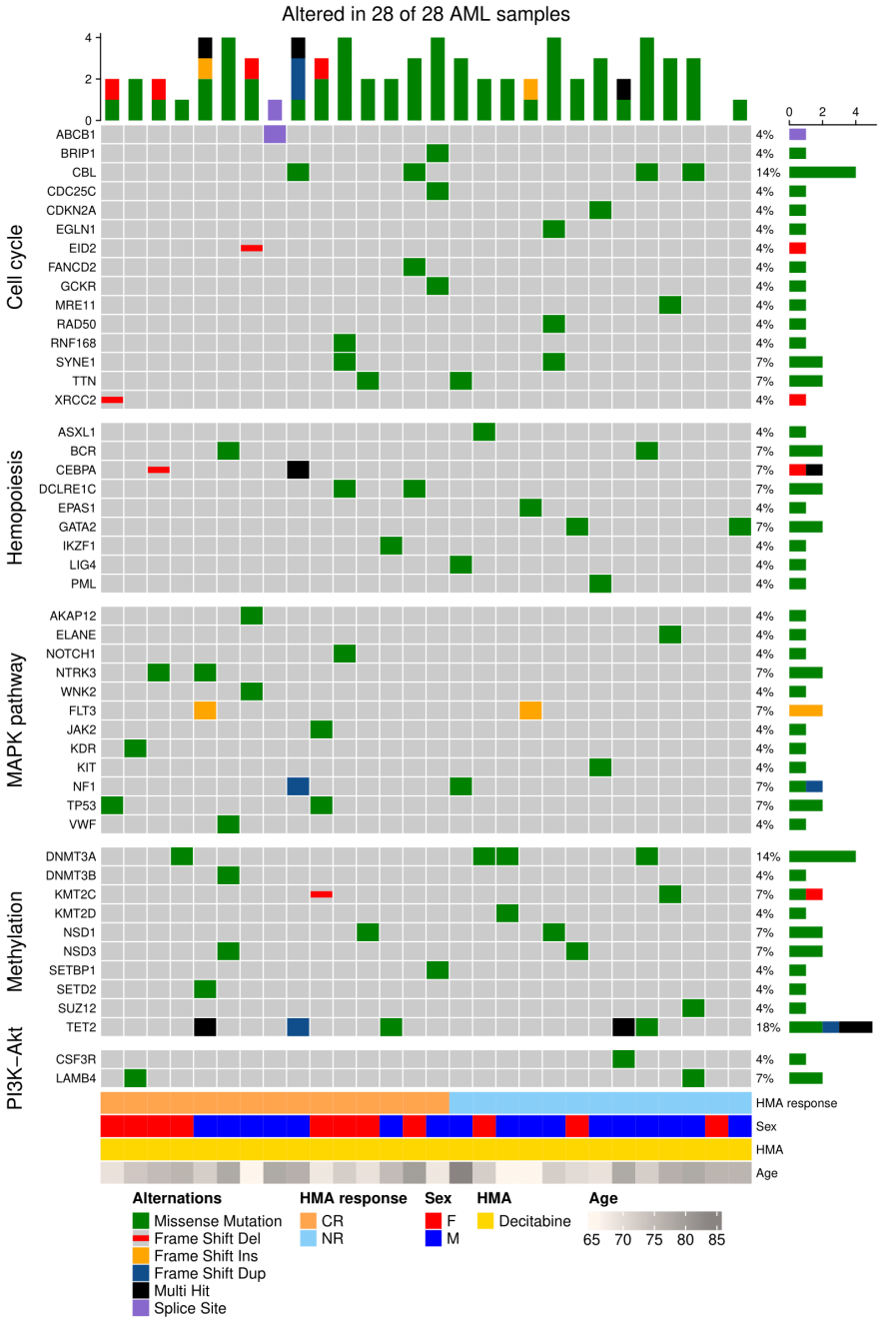


Figure 18. Functional mutational profiles of 28 AML patients.

7. Transcriptional landscape between DAC sensitive and resistance group

As shown in previous analyses, we confirmed that overall methylation was similar in a group of patients with diverse responses to DAC. However, DMR analysis revealed that the genes of regulatory regions related to hematopoiesis and erythrocyte differentiation were hypomethylated in CR. Based on these facts, an expression analysis was conducted to prove whether methylation analysis results and gene expression analysis show consistent results. First, by identifying differentially expressed genes (DEGs), the expression difference in the two groups was defined. A total of 450 DEGs were confirmed, of which 316 were upregulated in CR and 134 were upregulated in NR group (Figure 19). In addition, individual expression of the genes was confirmed in patients and sample clustering through DEGs, which identified that CR and NR groups were actually distinguished by specific gene expression (Figure 20). Then, ontology analysis was performed using Metascape to determine which biological pathways are related to each group of DEGs. Cell adhesion and hematopoietic cell differentiation had a high significant score (Figure 21). We further subdivided the pathway using gene set enrichemt analysis (GSEA). First, we used gene sets that were associated with hematopoiesis and erythrocyte differentiation observed in the methylation results to explain the association with promoter methylation changes and transcriptional characteristics. This analysis revealed that genes associated with erythrocyte differentiation were highly expressed and hypomethylated in the regulatory region of CR group. Through these results, the consistency of methylation and expression analysis was confirmed and gene expression increased when regions such as promoters, which play an important role in gene expression, became hypomethylated. In addition to these pathways, cell adhesion pathway and T cell receptor complex show high enrichment score (Figure 22). Mainly, a number of genes involved in these biological pathways and molecular functions correspond to the immune response required for T cell

activation. CR group had upregulated molecules constituting the T cell receptor complex, specifically, *TRAC, CD8A, CD8B,* and *CTLA* genes (Figure 23). This result implies that the DAC resistance group had downregulated gene expression related to T cell activation. Conversely, when gene ontology analysis was conducted on genes that were upregulated in NR, it was confirmed that the inflammatory pathway had a much higher score (Figure 24). GSEA analysis showed that the inflammatory reaction largely consisted of toll-like receptor cascade molecules (Figure 25). Unlike the CR group, the expression of the toll-like receptor related gene was increased, but the TLR3 and TLR4 genes did not differ between the two groups (Figure 26). The overall enrichment score for toll-like cascade was confirmed for each individual, and it shows that patients in the NR group had an upregulated TLR pathway (Figure 27).

To summarize the results, each group had unique gene expression characteristics before the DAC treatment. In CR group, the receptor gene expression required for T cell and immune response increased significantly. In addition, unlike NR, the expression of genes involved in the differentiation of hematopoietic stem cells, especially erythrocyte differentiation and development, had increased more than that in the NR group. In contrast, in the NR group, not only were the immune response genes by T cell receptors downregulated, but the expression of certain genes related to hematopoietic action had also been reduced, suggesting decreased functions with normal stem cell or progenitor cell differentiation.

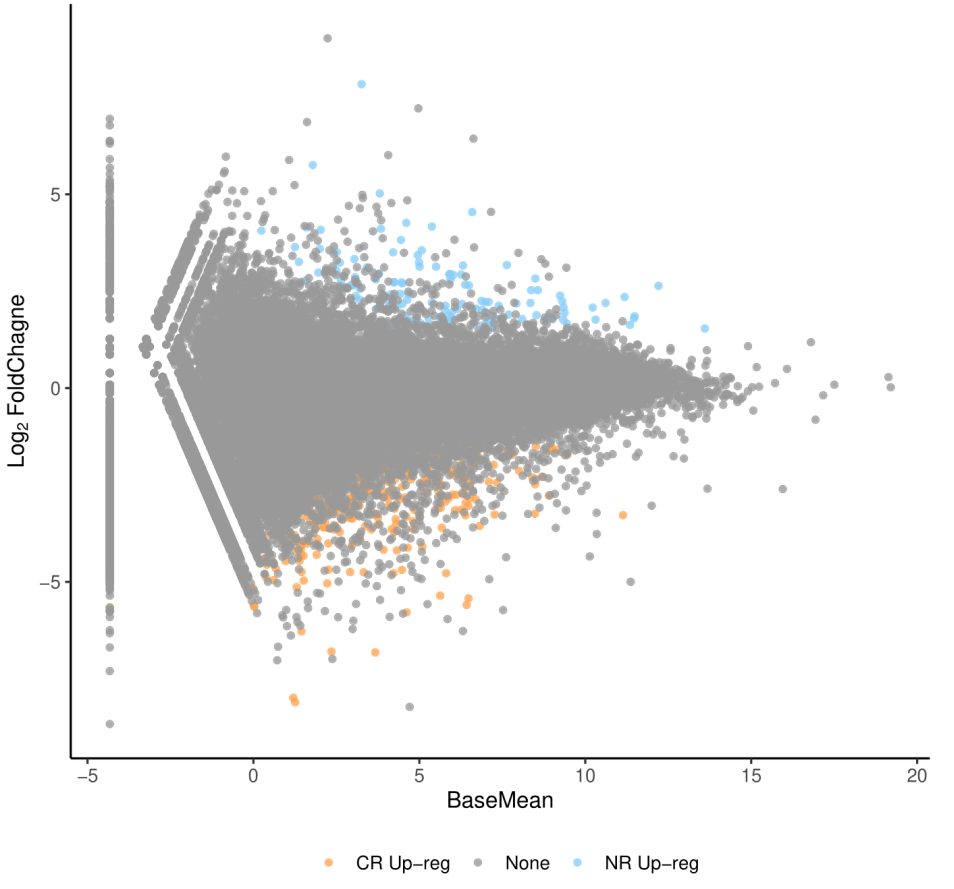


Figure 19. MA plot of differentially expressed genes.

MA plot for DESeq2. Points that are significantly different with a *p*-value less than 0.05 are in blue or orange. Orange dot indicates upregulated genes in CR and blue dot indicates NR upregulated genes. CR, Responder, NR, Non-responder.

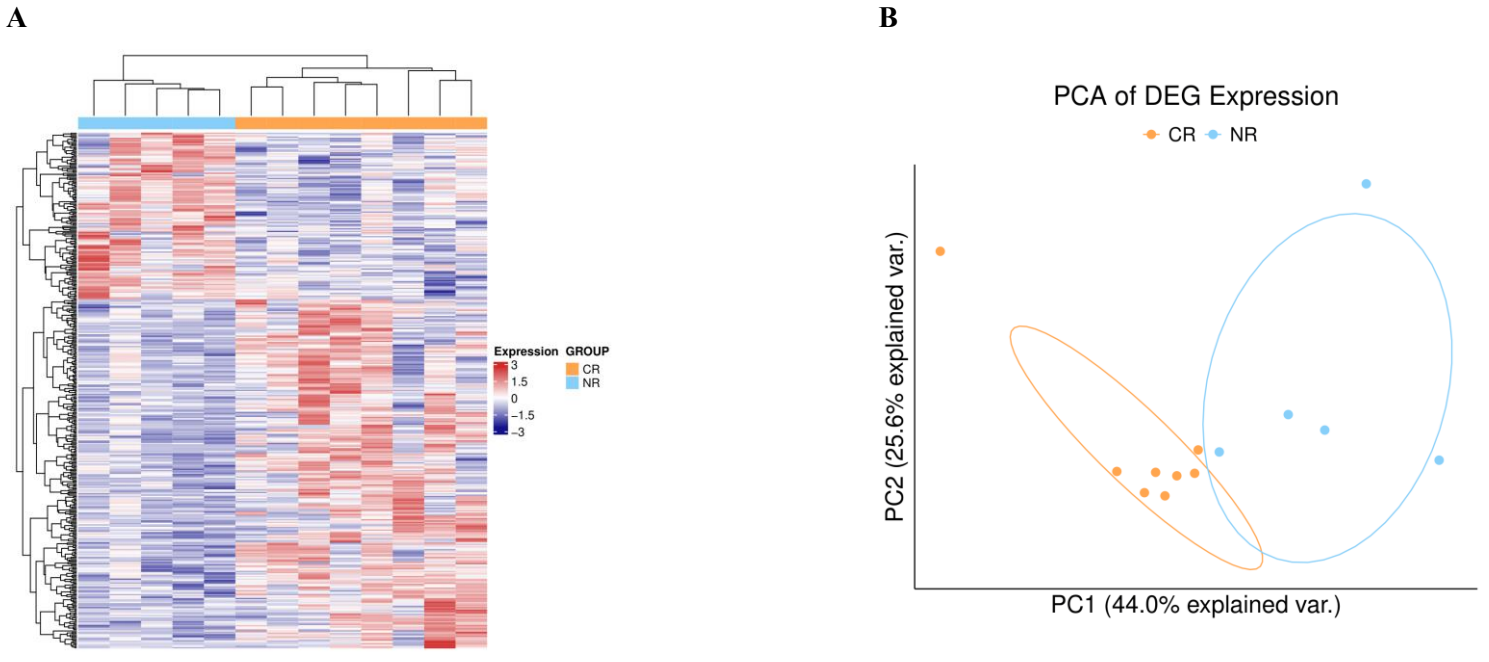


Figure 20. Features of differentially expressed genes.

Differentially expressed genes were observed between CR and NR group. (A) Normalized gene expression heatmap (Z score) of total 450 differentially expressed genes (B) PCA plot of differentially expressed genes. CR, Responder, NR, Non-responder.

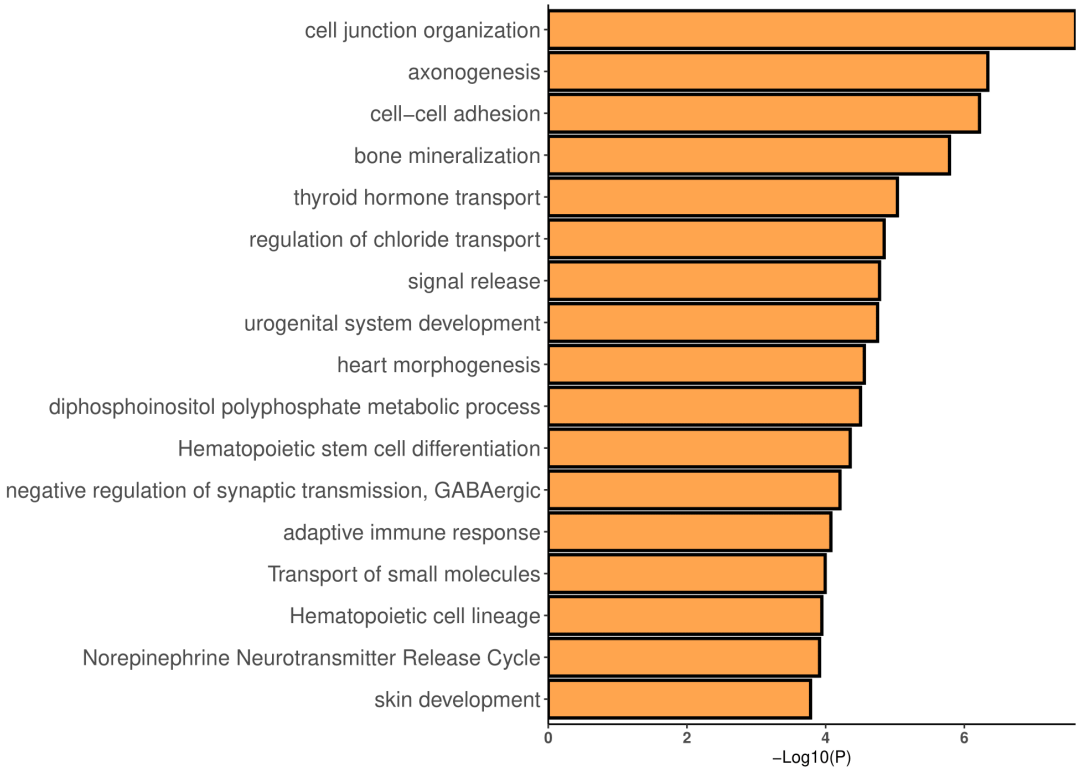


Figure 21. Gene ontology plot consist of responder upregualted genes.

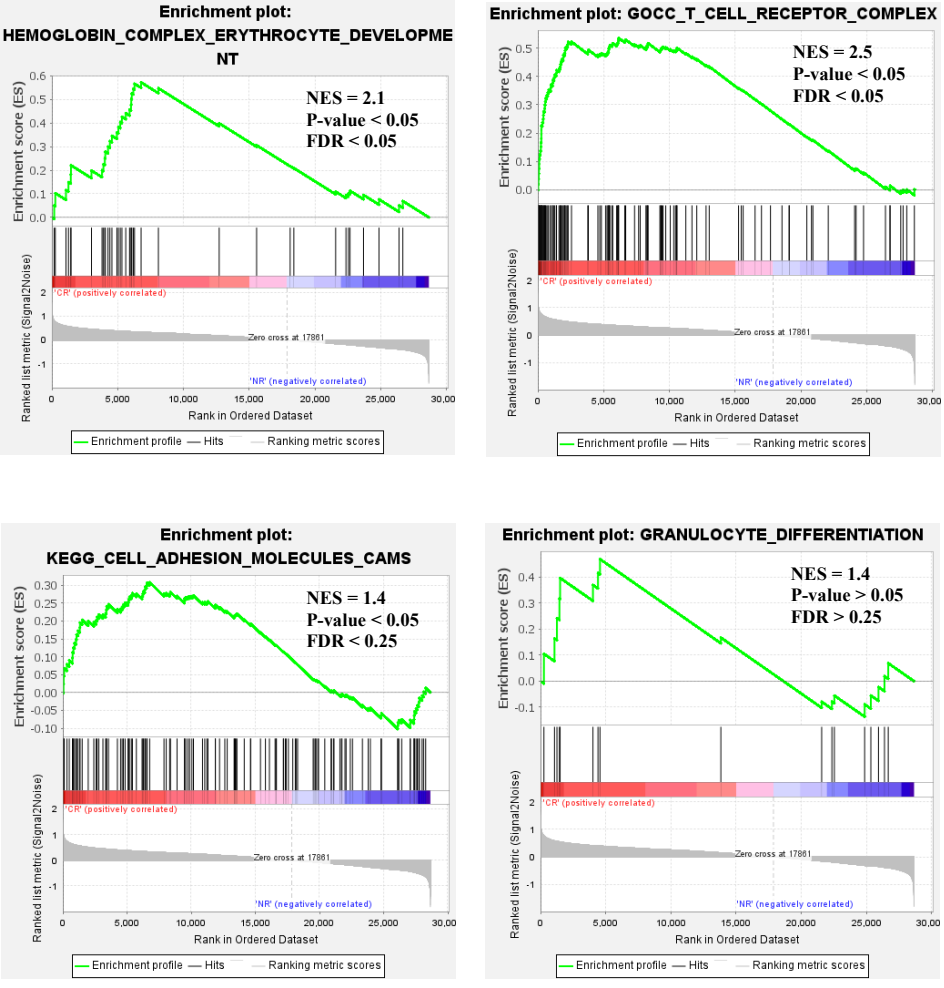


Figure 22. Gene set enrichment analysis of responder group.

Pathways with hematopoiesis and T cell receptor were recorded high NES in CR. In case of hemoglobin, erythrocyte development and granulocyte differentiation are already enriched in methylation study. These results indicate consistent findings between methylation and expression. Although granulocyte differentiation pathway was not significant, but it has borderline significant *p* value. NES, Normalized enrichment score, CR, Responder.

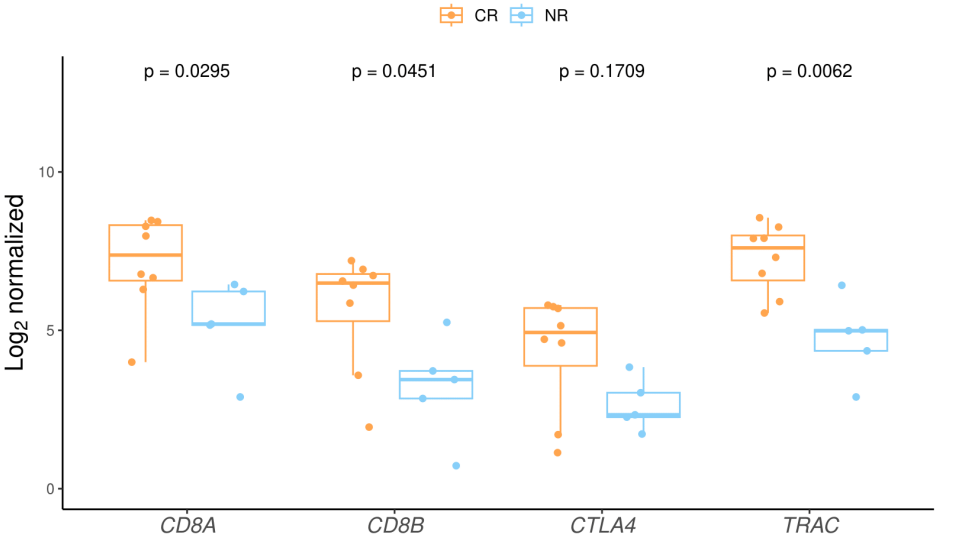


Figure 23. T cell receptor complex gene expression.

Expression of each gene is significantly different between CR and NR group. CR, Responder, NR, Non-responder.

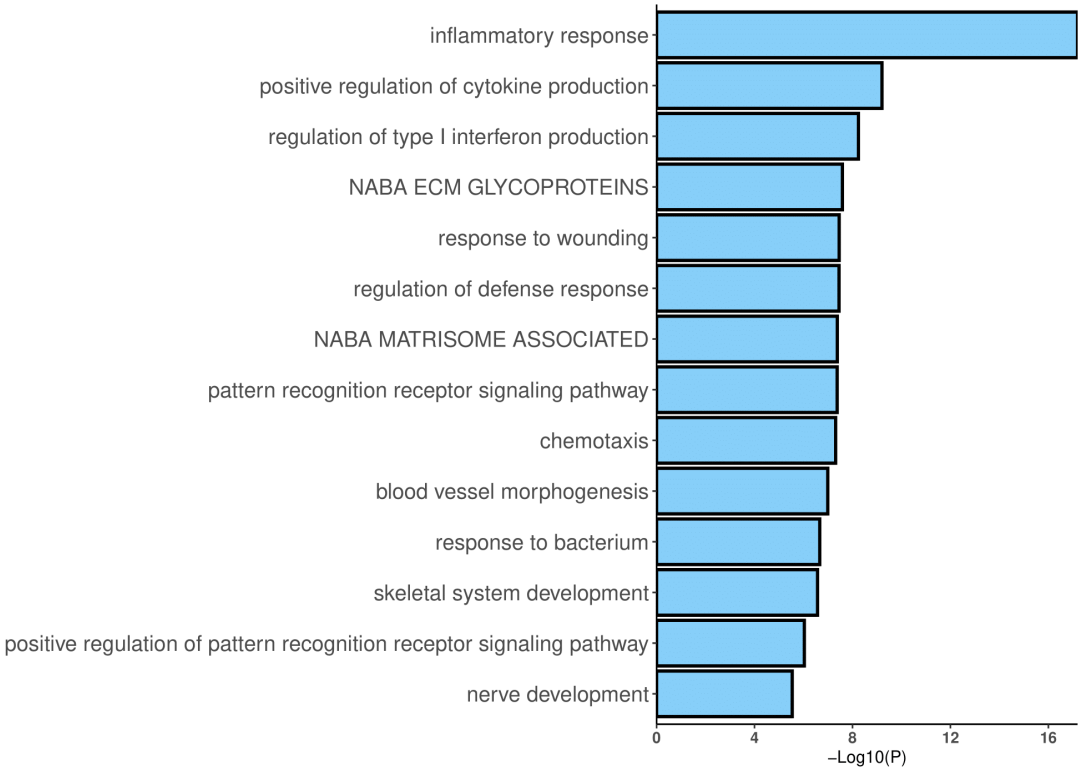


Figure 24. Gene ontology plot consist of non-responder upregualted gene.

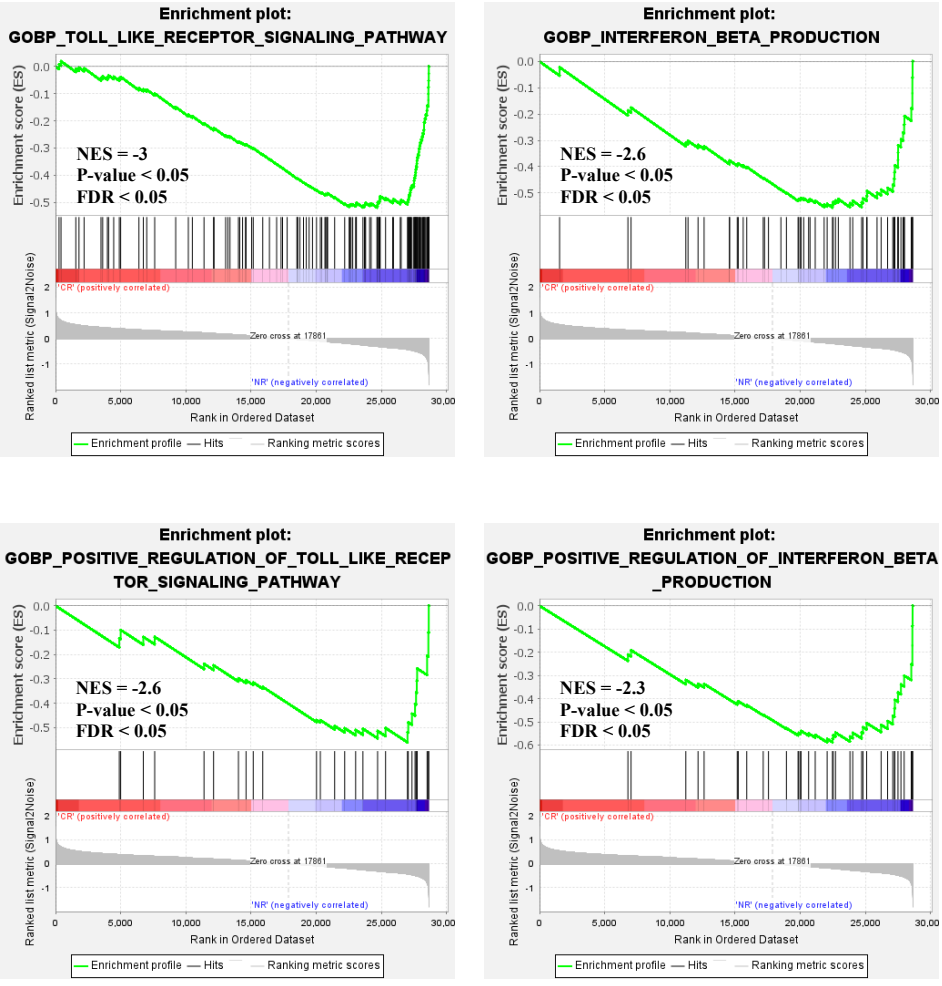


Figure 25. Gene set enrichment analysis of non-responder group.

Almost enriched pathway in non-responder (NR) group is inflammatory response such as toll-like receptor or interferon production. All these pathways have significant statistic value. NES, Normalized enrichment score, NR, Non-responder.

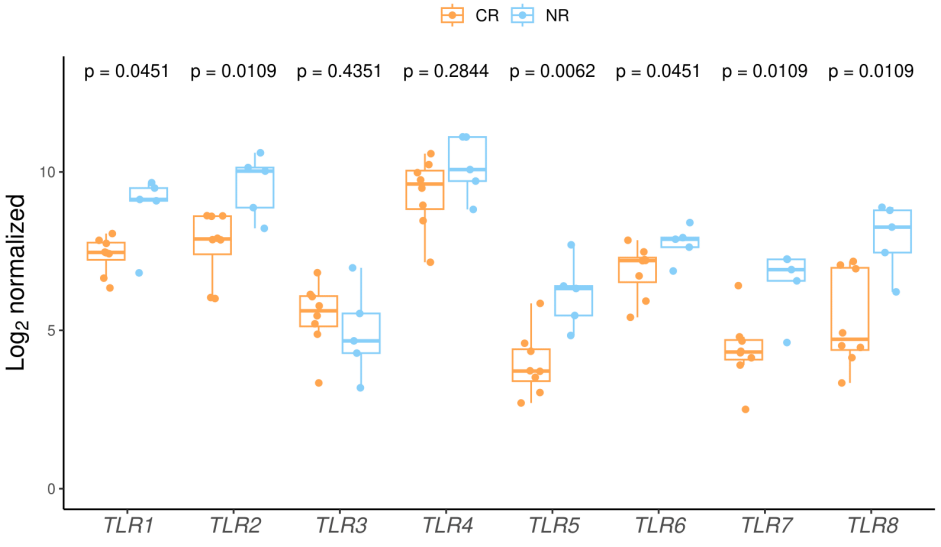


Figure 26. Toll-like receptor gene expression between responder and non-responder group.

Except for *TLR3 and TLR4*, gene expression of toll-like receptors is higher in NR group and significantly different. CR, responder, NR, non-responder.

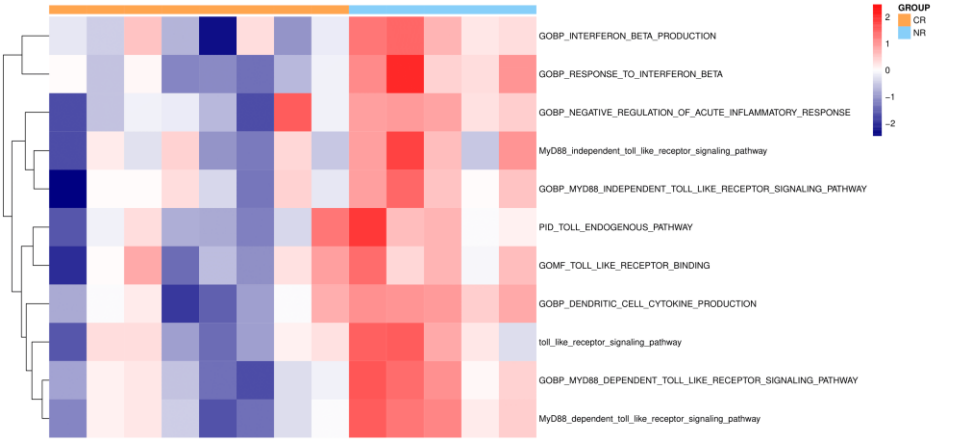


Figure 27. Enrichment score heatmap of toll-like receptor pathway.

Single sample gene set enrichment analysis indicates that toll-like receptor cascade pathway is almost highly enriched in NR. CR shows decreased expression with toll-like receptor pathway. CR, Responder, NR, Non-responder.

IV. DISCUSSION

Hypomethylating agents, which inhibit DNA methyltransferase, can be appropriate for treating acute myeloid leukemia in patients who are not eligible for intensive therapy. Two kinds of hypomethylating agents, azacitidine and decitabine, are FDA-approved for use in myelodysplastic syndrome, acute myeloid leukemia, and several other cancers. These drugs alter the epigenetic characteristics of DNA and gene expression by modulating DNA methylation changes. However, the prognosis of HMA treatment failure is extremely poor, with median survival for these patients barely reaching 6 months and approximately 50 % of patients never even achieving a response in the first treatment[56-58]. This low rate of therapeutic response is a substantial problem, and little is known about what causes drug resistance or developmental resistance. For this reason, understanding drug metabolism and resistance mechanisms is important to predict the response to a hypomethylating agent. Therefore, we set up to investigate the epigenetic, genetic, and transcriptional landscapes associated with response to DAC in a cohort of acute myeloid leukemia patients in order to explain the molecular character and mechanism of primary resistance of this drug. To better understand the DAC response, we used next-generation sequencing technology to study the presence of DNA methylation, gene mutation, and gene expression difference at the time of diagnosis so as to distinguish DAC-sensitive and DAC-resistant patients.

In our methtylation analysis, data suggests that one of the most characteristic cancer-associated methylation changes, promoter and CpG island hypomethylation, appears to be similar between two groups that show sensitivity or resistance to the drugs. However, we revealed that the promoter region of *GATA1, GATA2, NFE2, BCL11A, TAL1,* and *RUNX1* is hypermethylated in AML patients with poor response to DAC, implying that the expression of these genes involved in normal hematopoiesis and cell differentiation could be limited owing to the high methylation. This finding supports the results of others suggesting the

importance of genes associated normal hematopoiesis in contributing towards better prognosis of AML/MDS patients. Reviews describe the importance of *GATA1* and *GATA2* genes in regulating hematopoietic differentiation, and these genes are controlled in a cell-type-specific manner[59-61]. *NFE2* gene is essential for regulating erythrocyte and megakaryocytic maturation and differentiation. This gene is also involved in transcription of multiple enzymes in the heme biosynthesis process in *in silico* annlysis[62]. However, overexpression of *NFE2* perturbs erythrocyte maturation[63]. In case of *BCL11A*, it is expressed in most hematopoietic cells and common lymphoid progenitors (CLPs). In a mouse model, depletion of *BCL11A* gene causes apoptosis of CLPs and modifies the lymphoid developmental possibility associated with immune response[64]. Genes such as *TAL1* and *RUNX1* participate in multipotency, maintain normal myeloid differentiation, or determine the differentiation process[65,66].

We also found that transcription factors that bind to differentially methylated regions had a significant methylation difference in average methylation. Generally, the average methylation level of significantly enriched transcription factors is higher in non-responder group than responder group. Almost-enriched transcription factors are linked with hematopoietic cell differentiation and development process. Early studies reported that elevated methylation acts as a transcriptional repressor[67]. These promote better understanding for our results that hypomethylated transcription factor binding sites were prevalent in the CR group. Taken together, these findings demonstrate that patients who are sensitive to DAC were observed to have a low methylation value of transcription factor binding sites and genes that regulate normal hematopoiesis and hematopoietic cell differentiation. In other words, the characteristics of the DAC-sensitive CR group had reduced methylation of genes that plays a pivotal role in the formation of hematopoietic lineage process of becoming mature cells from human hematopoietic stem cell (HSC) through progenitor cells.

While previous reports on MDS and hematologic malignancies have pinpointed certain mutations, *TET2* and *DNMT3A,* specifically, were involved in HMA response. Although our study identified mutated genes that are widely observed in AML patients and affect drug response, we could not find any correlation between mutational profile and response to DAC. In addition, methylation levels among patients with the same genetic mutation varied. In previous studies, AML patients who have the *DNMT3A* R882 mutation, which is known as inducing focal hypomethylation in AML[48], have a similar methylation level compared to other AML patients. Therefore, genetic changes involved in methylation cannot fully explain the correlation between the properties of methylation difference and the drug resistance. Similar study also suggest that mutation was not significantly correlated to DAC response in CMML patients[68]. However, a large cohort study with acute myeloid leukemia and myelodysplastic syndromes identified that response rates to DAC were higher in patients with *TP53* mutations than among patients with wild-type *TP53*[69]. Likewise, although only a small subset of patients who were sensitive to DAC treatment had *TP53* mutations, a similar tendency was observed in our study.

In addition to epigenetic and genetic analysis, our study also demonstrated the baseline gene expression level correlated with DAC response. Analysis of response-associated gene signature revealed high enrichment for gene sets involved in erythrocyte differentiation and T cell receptor complex among the genes upregulated in DAC-sensitive patients. This finding is in line with the methylation analysis that hypomethylated genes are involved in hematopoiesis and T cell activation pathway in CR. In fact, in a study investigating the anti-tumor effect of HMA, the activation of $CD8^+$ T cell increased both *in vivo* and *ex vivo* after HMA use[70]. By contrast, among upregulated genes in DAC resistance group, we found that the toll-like receptor pathway and TLR genes were overexpressed. The toll-like receptor pathway is an innate immune response in humans to bacteria or various antigens. Recent studies demonstrated that primary

induction of toll-like receptor cascade stimulates hematopoietic stem cell proliferation and differentiation, but the long term stimulus causes a decrease in normal hematopoietic cell functions and cell differentiation[71-73]. Therefore, in our study, it is possible that the NR group exhibits repressed HSC differentiation through increased inflammatory response via the toll-like receptor cascade pathway, and CR patients had a relatively high level of expression associated with hematopoiesis because of the absence of the TLR pathway.

Overall, this study was able to identify genetic and epigenetic biomarkers in patients who were sensitive to DAC response and those who were not sensitive to DAC. Substantially, the distinct response originated from differences in methylation and expression of genes related to HSC differentiation. While these findings provide some environmental insights into the methylation and gene expression differences associated with DAC resistance, downstream cellular and molecular consequences are still unclear. Our studies revealed not only promoter regions that affect gene expression or modulate transcriptional potential but also regions distributed throughout the genome. However, many genes that show substantial methylation difference were not expressed regardless of their methylation status. Therefore, our results failed to examine exact trends of gene expression in DMRs linked with regulatory regions or gene bodies. However, these results suggest that transcription is not regulated only by methylation of regulatory elements; many highly sophisticated molecular interations are involved with transcription. Indeed, we observed that immunse response and HSC differentiation are the main factors that affect DAC resistance, but it is difficult to determine exactly how HSC differentiation affects the DAC response.

Our integrative NGS studies promote comprehensive understanding of DAC resistance in AML patients. Although this result may not be associated with other hematohematopoietic malignancy or other cancer with DAC treatment, large-scale computational study and experiments designed to test this mechanistic

pathway may reveal new insights into the exact role of HSC differentiation and immune response in the resistance to hypomethylating agents.

V. CONCLUSION

This study provides insights into the baseline epigenomic and genomic characteristics of acute myeloid leukemia patients treated with hypomethylating agents. Methylation analysis implies that region- or gene-specific methylation provides distinctive methylation properties for drug resistance. Furthermore, differentially methylated regions and transcription factors promote better understanding of HMA resistance mechanisms. WES analysis identified the mutational landscape to identify genetic profiles related to HMA responses, but this analysis suggests limited information for drug resistance. Total RNA sequencing has the same results in a methylation study with respect to biologocal pathways such as hematopoiesis and immune response. Furthermore, inflammatory response was determined to be a novel therapeutic marker. While computational results should be validated in subsequent functional studies, this integrated understanding shed light on the utility of detecting predictive biomarkers and estimating suitability for HMA treatment.

REFERENCES

1. Dohner H, Weisdorf DJ, Bloomfield CD. acute myeloid leukemia. N Engl J Med. 2015;373(12):1136–52.
2. Dohner K, Dohner H. Molecular characterization of acute myeloid leukemia. Haematologica. 2008;93(7):976–82.
3. Yang X, Wong MPM, Ng RK. Aberrant DNA methylation in acute myeloid leukemia and its clinical implications. Int J Mol Sci. 2019;20(18).
4. Cedar H, Bergman Y. Epigenetics of haematopoietic cell development. Nat Rev Immunol. 2011;11(7):478–88.
5. Ji H, Ehrlich LI, Seita J, Murakami P, Doi A, Lindau P, et al. Comprehensive methylome map of lineage commitment from haematopoietic progenitors. Nature. 2010;467(7313):338–42.
6. Guillamot M, Cimmino L, Aifantis I. The Impact of DNA Methylation in hematopoietic malignancies. Trends Cancer. 2016;2(2):70–83.
7. Li E, Zhang Y. DNA methylation in mammals. Cold Spring Harb Perspect Biol. 2014;6(5):a019133.
8. Smith ZD, Meissner A. DNA methylation: roles in mammalian development. Nat Rev Genet. 2013;14(3):204–20.
9. Zhang W, Xu J. DNA methyltransferases and their roles in tumorigenesis. Biomark Res. 2017;5:1.
10. Sen GL, Reuter JA, Webster DE, Zhu L, Khavari PA. DNMT1 maintains progenitor function in self-renewing somatic tissue. Nature. 2010;463(7280):563–7.
11. Rasmussen KD, Helin K. Role of TET enzymes in DNA methylation, development, and cancer. Genes Dev. 2016;30(7):733–50.
12. Spencer DH, Russler-Germain DA, Ketkar S, Helton NM, Lamprecht TL, Fulton RS, et al. CpG Island Hypermethylation Mediated by DNMT3A is a consequence of AML progression. Cell.

2017;168(5):801–16 e13.

13. Russler-Germain DA, Spencer DH, Young MA, Lamprecht TL, Miller CA, Fulton R, et al. The R882H DNMT3A mutation associated with AML dominantly inhibits wild-type DNMT3A by blocking its ability to form active tetramers. Cancer Cell. 2014;25(4):442–54.
14. Weissmann S, Alpermann T, Grossmann V, Kowarsch A, Nadarajah N, Eder C, et al. Landscape of TET2 mutations in acute myeloid leukemia. Leukemia. 2012;26(5):934–42.
15. Tulstrup M, Soerensen M, Hansen JW, Gillberg L, Needhamsen M, Kaastrup K, et al. TET2 mutations are associated with hypermethylation at key regulatory enhancers in normal and malignant hematopoiesis. Nat Commun. 2021;12(1):6061.
16. Rasmussen KD, Jia G, Johansen JV, Pedersen MT, Rapin N, Bagger FO, et al. Loss of TET2 in hematopoietic cells leads to DNA hypermethylation of active enhancers and induction of leukemogenesis. Genes Dev. 2015;29(9):910–22.
17. Carter JL, Hege K, Yang J, Kalpage HA, Su Y, Edwards H, et al. Targeting multiple signaling pathways: the new approach to acute myeloid leukemia therapy. Signal Transduct Target Ther. 2020;5(1):288.
18. Nardi V, Winkfield KM, Ok CY, Niemierko A, Kluk MJ, Attar EC, et al. Acute myeloid leukemia and myelodysplastic syndromes after radiation therapy are similar to de novo disease and differ from other therapy-related myeloid neoplasms. J Clin Oncol. 2012;30(19):2340–7.
19. Patel N, BJ RI, Patel S, Watts JM, Benveniste R, Abramowitz M, et al. Emergent Radiotherapy for Leukemia-Induced Cranial Neuropathies Refractory to Intrathecal Therapy. Cureus. 2021;13(5):e15212.
20. Cruijsen M, Lubbert M, Wijermans P, Huls G. Clinical Results of Hypomethylating Agents in AML Treatment. J Clin Med. 2014;4(1):1–17.

21. Gardin C, Dombret H. Hypomethylating agents as a therapy for AML. Curr Hematol Malig Rep. 2017;12(1):1–10.
22. Al-Ali HK, Jaekel N, Niederwieser D. The role of hypomethylating agents in the treatment of elderly patients with AML. J Geriatr Oncol. 2014;5(1):89–105.
23. Stomper J, Rotondo JC, Greve G, Lubbert M. Hypomethylating agents (HMA) for the treatment of acute myeloid leukemia and myelodysplastic syndromes: mechanisms of resistance and novel HMA-based therapies. Leukemia. 2021;35(7):1873–89.
24. Lubbert M, Suciu S, Baila L, Ruter BH, Platzbecker U, Giagounidis A, et al. Low-dose decitabine versus best supportive care in elderly patients with intermediate- or high-risk myelodysplastic syndrome (MDS) ineligible for intensive chemotherapy: final results of the randomized phase III study of the European Organisation for Research and Treatment of Cancer Leukemia Group and the German MDS Study Group. J Clin Oncol. 2011;29(15):1987–96.
25. Kantarjian H, Issa JP, Rosenfeld CS, Bennett JM, Albitar M, DiPersio J, et al. Decitabine improves patient outcomes in myelodysplastic syndromes: results of a phase III randomized study. Cancer. 2006;106(8):1794–803.
26. Steensma DP, Baer MR, Slack JL, Buckstein R, Godley LA, Garcia-Manero G, et al. Multicenter study of decitabine administered daily for 5 days every 4 weeks to adults with myelodysplastic syndromes: the alternative dosing for outpatient treatment (ADOPT) trial. J Clin Oncol. 2009;27(23):3842–8.
27. Kantarjian HM, Thomas XG, Dmoszynska A, Wierzbowska A, Mazur G, Mayer J, et al. Multicenter, randomized, open-label, phase III trial of decitabine versus patient choice, with physician advice, of either supportive care or low-dose cytarabine for the treatment of older

patients with newly diagnosed acute myeloid leukemia. J Clin Oncol. 2012;30(21):2670–7.

28. Dombret H, Seymour JF, Butrym A, Wierzbowska A, Selleslag D, Jang JH, et al. International phase 3 study of Azacitidine vs conventional care regimens in older patients with newly diagnosed AML with >30 % blasts. Blood. 2015;126(3):291–9.
29. Diesch J, Zwick A, Garz A-K, Palau A, Buschbeck M, Götze KS. A clinical-molecular update on azanucleoside-based therapy for the treatment of hematologic cancers. Clinical Epigenetics. 2016;8(1).
30. Rius M, Stresemann C, Keller D, Brom M, Schirrmacher E, Keppler D, et al. Human concentrative nucleoside transporter 1-mediated uptake of 5-Azacitidine enhances DNA demethylation. Mol Cancer Ther. 2009;8(1):225–31.
31. Mahfouz RZ, Jankowska A, Ebrahem Q, Gu X, Visconte V, Tabarroki A, et al. Increased CDA expression/activity in males contributes to decreased cytidine analog half-life and likely contributes to worse outcomes with 5-Azacitidine or decitabine therapy. Clin Cancer Res. 2013;19(4):938–48.
32. Marcucci G, Silverman L, Eller M, Lintz L, Beach CL. Bioavailability of Azacitidine subcutaneous versus intravenous in patients with the myelodysplastic syndromes. J Clin Pharmacol. 2005;45(5):597–602.
33. Cashen AF, Shah AK, Todt L, Fisher N, DiPersio J. Pharmacokinetics of decitabine administered as a 3-h infusion to patients with acute myeloid leukemia (AML) or myelodysplastic syndrome (MDS). Cancer Chemother Pharmacol. 2008;61(5):759–66.
34. Krueger F, Andrews SR. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. Bioinformatics. 2011;27(11):1571-72.
35. Akalin A, Kormaksson M, Li S, et al. methylKit: a comprehensive R

package for the analysis of genome-wide DNA methylation profiles. Genome Biol. 2012;13(10):R87.

36. Wei Y, Dimicoli S, Bueso-Ramos C, Chen R, Yang H, Neuberg D, et al. Toll-like receptor alterations in myelodysplastic syndrome. Leukemia. 2013;27(9):1832-40.
37. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29(1):15-21.
38. Rogstad DK, Herring JL, Theruvathu JA, Burdzy A, Perry CC, Neidigh JW, et al. Chemical decomposition of 5-aza-2'-deoxycytidine (Decitabine): kinetic analyses and identification of products by NMR, HPLC, and mass spectrometry. Chem Res Toxicol. 2009;22(6):1194–204.
39. Patel K, Dickson J, Din S, Macleod K, Jodrell D, Ramsahoye B. Targeting of 5-aza-2'-deoxycytidine residues by chromatin-associated DNMT1 induces proteasomal degradation of the free enzyme. Nucleic Acids Res. 2010;38(13):4313–24.
40. Ghoshal K, Datta J, Majumder S, Bai S, Kutay H, Motiwala T, et al. 5-Aza-deoxycytidine induces selective degradation of DNA methyltransferase 1 by a proteasomal pathway that requires the KEN box, bromo-adjacent homology domain, and nuclear localization signal. Mol Cell Biol. 2005;25(11):4727–41.
41. Issa JP, Kantarjian HM. Targeting DNA methylation. Clin Cancer Res. 2009;15(12):3938–46.
42. Kordella C, Lamprianidou E, Kotsianidis I. Mechanisms of Action of Hypomethylating Agents: Endogenous Retroelements at the Epicenter. Front Oncol. 2021;11:650473.
43. Karimi Kelaye S, Najafi F, Kazemi B, Foruzandeh Z, Seif F, Solali S, et al. The contributing factors of resistance or sensitivity to epigenetic

drugs in the treatment of AML. Clin Transl Oncol. 2022;24(7):1250–61.

44. Saliba AN, John AJ, Kaufmann SH. Resistance to venetoclax and hypomethylating agents in acute myeloid leukemia. Cancer Drug Resist. 2021;4:125–42.
45. Simonicova K, Janotka L, Kavcova H, Sulova Z, Breier A, Messingerova L. Different mechanisms of drug resistance to hypomethylating agents in the treatment of myelodysplastic syndromes and acute myeloid leukemia. Drug Resist Updat. 2022;61:100805.
46. Nazha A, Sekeres MA, Bejar R, Rauh MJ, Othus M, Komrokji RS, et al. Genomic biomarkers to predict resistance to hypomethylating agents in patients with myelodysplastic syndromes using artificial intelligence. JCO Precis Oncol. 2019;3.
47. Ciechomska M, Roszkowski L, Maslinski W. DNA methylation as a Future Therapeutic and Diagnostic Target in Rheumatoid Arthritis. Cells. 2019;8(9).
48. Xiao FH, Kong QP, Perry B, He YH. Progress on the role of DNA methylation in aging and longevity. Brief Funct Genomics. 2016;15(6):454–9.
49. Jones MJ, Goodman SJ, Kobor MS. DNA methylation and healthy human aging. Aging Cell. 2015;14(6):924–32.
50. Lam LL, Emberly E, Fraser HB, Neumann SM, Chen E, Miller GE, et al. Factors underlying variable DNA methylation in a human community cohort. Proc Natl Acad Sci U S A. 2012;109 Suppl 2:17253–60.
51. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat Rev Genet. 2012;13(7):484–92.
52. Moarii M, Boeva V, Vert JP, Reyal F. Changes in correlation between promoter methylation and gene expression in cancer. BMC Genomics. 2015;16:873.
53. Yang X, Han H, De Carvalho DD, Lay FD, Jones PA, Liang G. Gene

body methylation can alter gene expression and is a therapeutic target in cancer. Cancer Cell. 2014;26(4):577-90.

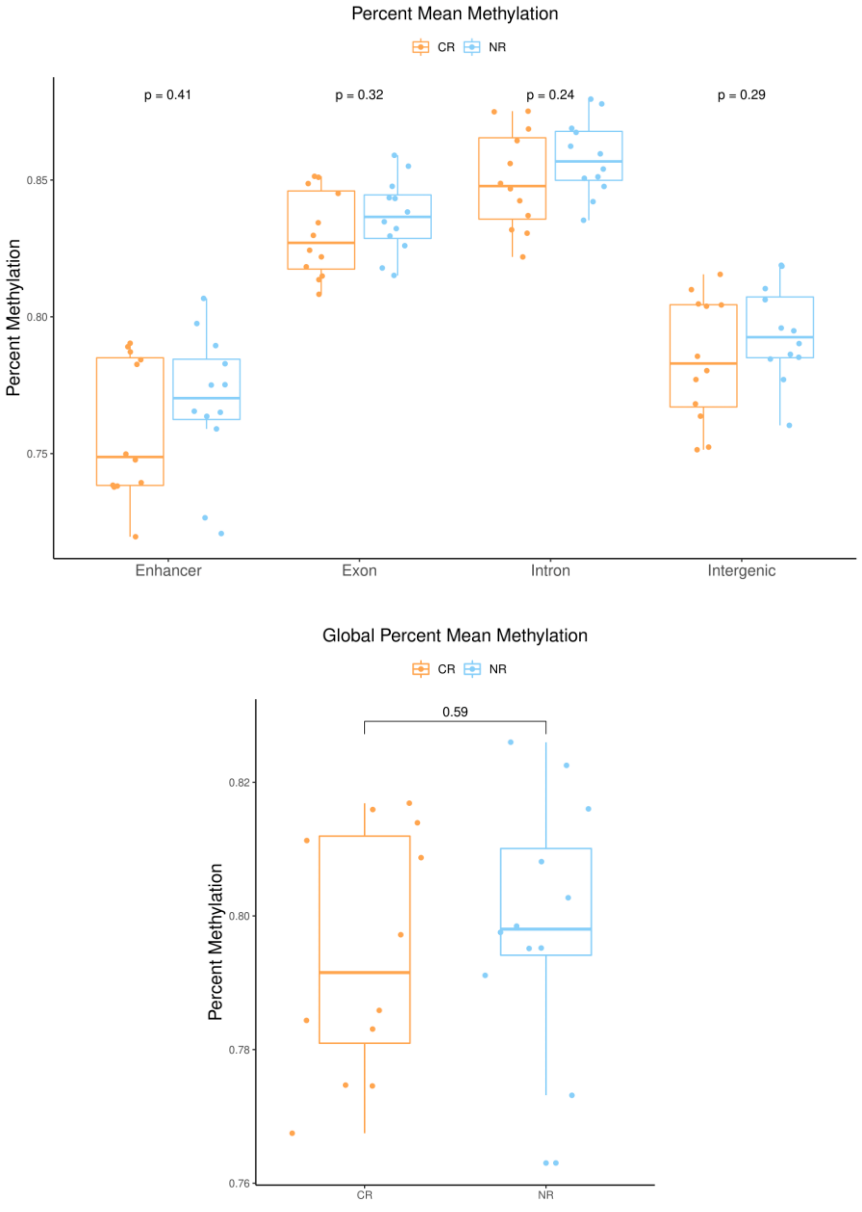
54. DNMT3A R882H Mutants Dominantly Inhibit Active DNMT3A Tetramers in AML. Cancer Discovery. 2014;4(5):508-.
55. Aran, D. and A. Hellman, DNA methylation of transcriptional enhancers and cancer predisposition. Cell, 2013. 154(1): p. 11–3
56. Prébet T, Gore SD, Esterni B, Gardin C, Itzykson R, Thepot S, et al. Outcome of high-risk myelodysplastic syndrome after azacitidine treatment failure. J Clin Oncol. 2011;29(24):3322-7.
57. Yoo KH, Cho J, Han B, Kim SH, Shin D-Y, Hong J, et al. Outcomes of decitabine treatment for newly diagnosed acute myeloid leukemia in older adults. PLOS ONE. 2020;15(8):e0235503.
58. Griffiths, E. A., & Gore, S. D. (2008). DNA methyltransferase and histone deacetylase inhibitors in the treatment of myelodysplastic syndromes. Seminars in hematology, 45(1), 23–30.
59. Gothwal M, Wehrle J, Aumann K, Zimmermann V, Gründer A, Pahl HL. A novel role for nuclear factor-erythroid 2 in erythroid maturation by modulation of mitochondrial autophagy. Haematologica. 2016;101(9):1054-64.
60. Moriguchi T, Suzuki M, Engel JD, Yamamoto M. GATA1 and GATA2 Function in Hematopoietic Differentiation. In: Kondo M, editor. Hematopoietic Stem Cell Biology. Totowa, NJ: Humana Press; 2010. p. 117-42.
61. Ohneda K, Yamamoto M. Roles of hematopoietic transcription factors GATA-1 and GATA-2 in the development of red blood cell lineage. Acta Haematol. 2002;108(4):237-45.
62. Rheinemann L, Seeger TS, Wehrle J, Pahl HL. NFE2 regulates transcription of multiple enzymes in the heme biosynthesis pathway. Haematologica. 2014;99(10):e208-10.

63. Suwabe N, Takahashi S, Nakano T, Yamamoto M. GATA-1 regulates growth and differentiation of definitive erythroid lineage cells during in vitro ES cell differentiation. Blood. 1998;92(11):4108-18.
64. Yokomizo T, Hasegawa K, Ishitobi H, Osato M, Ema M, Ito Y, et al. Runx1 is involved in primitive erythropoiesis in the mouse. Blood. 2008;111(8):4075-80.
65. Yu Y, Wang J, Khaled W, Burke S, Li P, Chen X, et al. Bcl11a is essential for lymphoid development and negatively regulates p53. J Exp Med. 2012;209(13):2467-83.
66. Vagapova ER, Spirin PV, Lebedev TD, Prassolov VS. The Role of TAL1 in Hematopoiesis and Leukemogenesis. Acta Naturae. 2018;10(1):15–23.
67. Medvedeva YA, Khamis AM, Kulakovskiy IV, Ba-Alawi W, Bhuyan MSI, Kawaji H, et al. Effects of cytosine methylation on transcription factor binding sites. BMC Genomics. 2014;15(1):119.
68. Meldi K, Qin T, Buchi F, Droin N, Sotzen J, Micol JB, et al. Specific molecular signatures predict decitabine response in chronic myelomonocytic leukemia. J Clin Invest. 2015;125(5):1857-72.
69. Welch JS, Petti AA, Miller CA, Fronick CC, O'Laughlin M, Fulton RS, et al. TP53 and Decitabine in Acute Myeloid Leukemia and Myelodysplastic Syndromes. New England Journal of Medicine. 2016;375(21):2023-36
70. Loo Yau H, Bell E, Ettayebi I, de Almeida FC, Boukhaled GM, Shen SY, et al. DNA hypomethylating agents increase activation and cytolytic activity of CD8(+) T cells. Mol Cell. 2021;81(7):1469-83.e8.
71. Esplin BL, Shimazu T, Welner RS, Garrett KP, Nie L, Zhang Q, et al. Chronic exposure to a TLR ligand injures hematopoietic stem cells. J Immunol. 2011;186(9):5367-75.
72. Takizawa H, Fritsch K, Kovtonyuk LV, Saito Y, Yakkala C, Jacobs K, et

al. Pathogen-Induced TLR4-TRIF Innate Immune Signaling in Hematopoietic Stem Cells Promotes Proliferation but Reduces Competitive Fitness. Cell Stem Cell. 2017;21(2):225-40.e5.

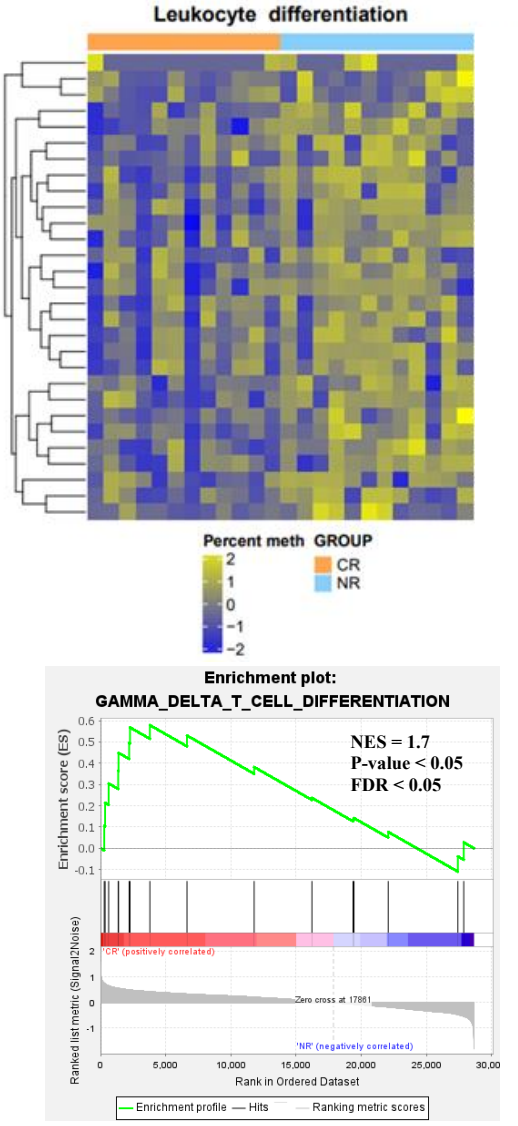
73. Trowbridge JJ, Starczynowski DT. Innate immune pathways and inflammation in hematopoietic aging, clonal hematopoiesis, and MDS. J Exp Med. 2021;218(7).

APPENDICES



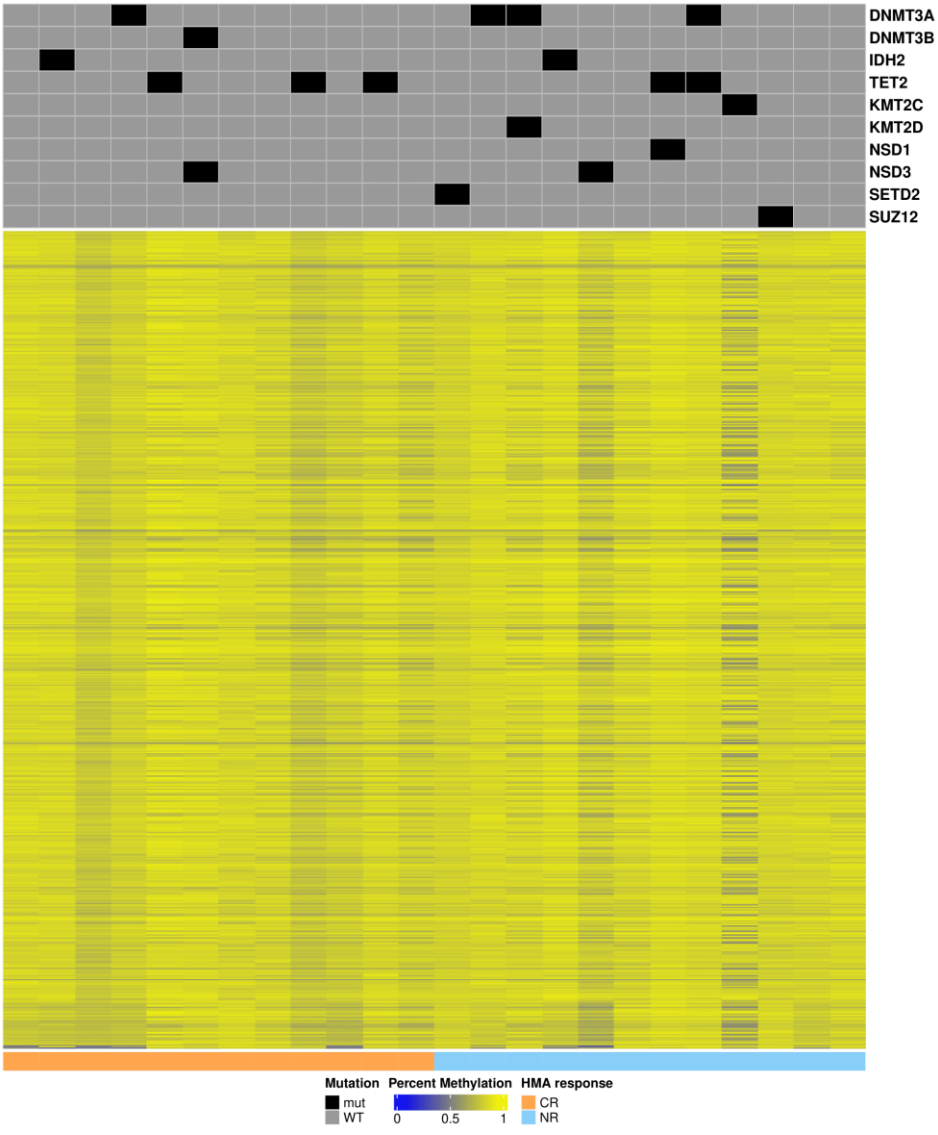
Appendix Figure 1. Percent methylation of global CpG sites and other genomic regions.

Individual methylation of genomic regions (A) (B) Percent mean methylation is similar within CR and NR group. CR, Responder, NR, Non-responder.



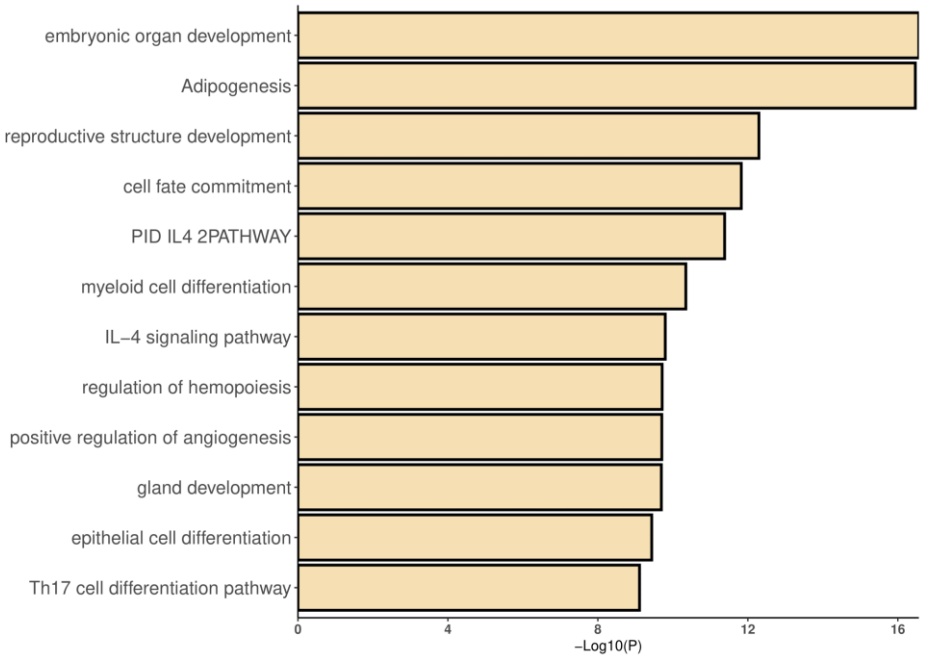
Appendix Figure 2. Leukocyte differentiation gene and subset pathway.

(A) Heatmap shows normalized promoter percent methylation of genes associated with leukocyte differentiation. (B) GSEA result indicates that gamma delta T cell differentiation corresponding to a subset of leukocyte differentiation had significant NES in line with promoter hypomethylation of CR group. GSEA, Gene set enrichment test, NES, Normalized enrichment score, CR, Responder.



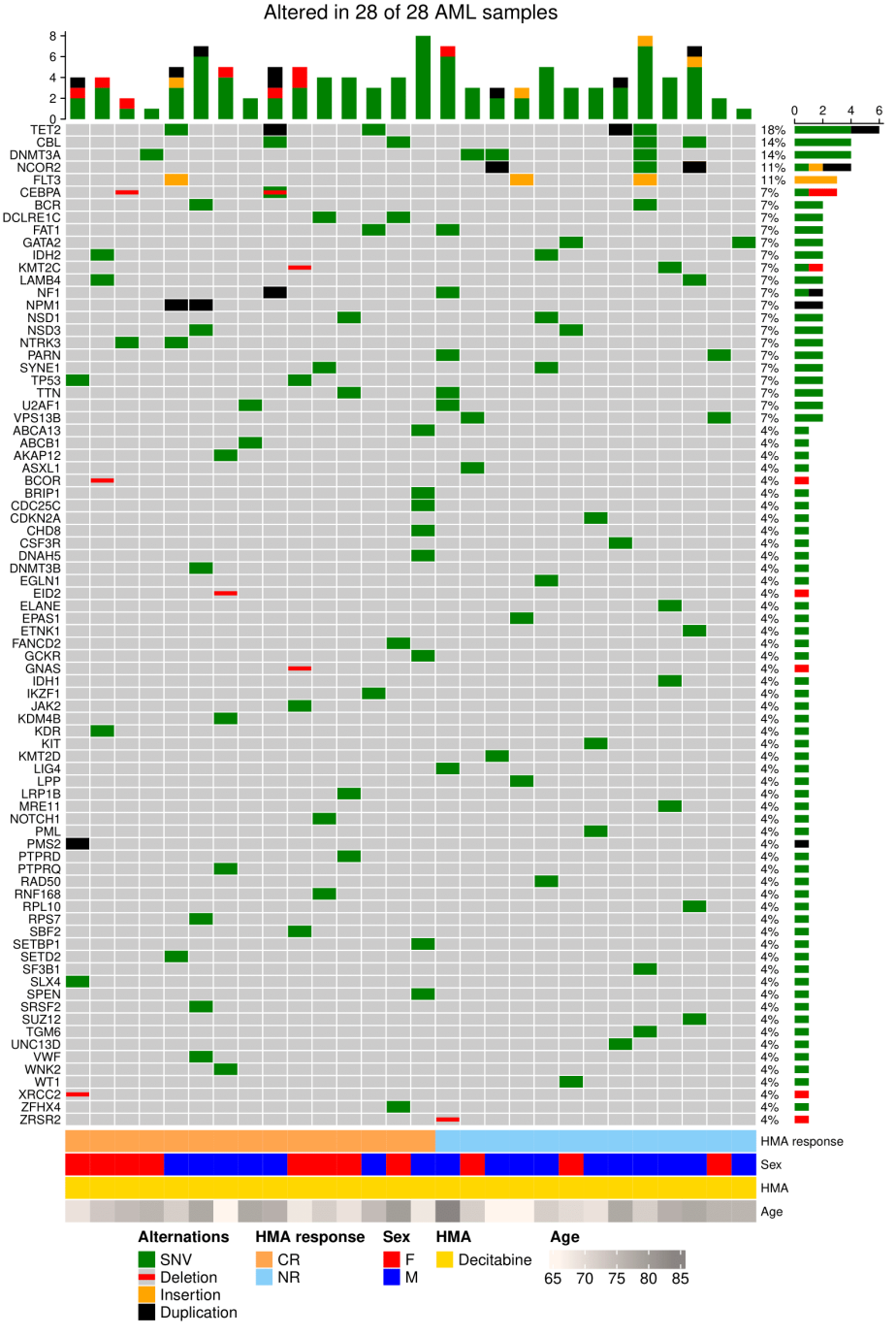
Appendix Figure 3. Methylation heatmap for 1 megabase (MB) window and gene mutations which affect the DNA methylation.

Mutations associated with DNA methylation were observed and heatmap shows DNA methylation for 1 MB tiles. Each methylation was hypervariable.



Appendix Figure 4. Enrichment of transcription factors ontology on differentially methylated regions.

This result indicates biological pathways for enriched TF gene sets. Almost pathways were associated with development and differentiation. TF, Transcription factor



Appendix Figure 5. All mutational profiles of 28 AML patients

Appendix Table 1. Detailed demographic variables of 28 AML patients

	Sex	Age	BM Blast (%)	PB Blast (%)	WBC Count ($\times 10^3/\mu L$)	HGB (g/dL)	PLT ($\times 10^3/\mu L$)	Karyotype
YM01	F	69	51.4	64	11.16	7.4	37	hypodiploid
YM02	F	73	69.4	4	0.68	9.4	140	46, XX[25]
YM03	F	75	40.3	4	1.78	11.0	94	46, XX[20]
YM04	F	76	33.2	23	57.20	9.0	149	46, XX[20]
YM05	M	72	75.0	92	26.92	9.9	32	46, XY[20]
YM06	M	78	70.3	51	13.01	10.4	105	46, XY[22]
YM07	M	60	46.3	83	26.77	7.8	86	46, XY[30]
YM08	M	78	36.6	4	346.7	4.9	153	46, XY[20]
YM09	M	77	74.8	85	101.73	8.0	67	46, XY[20]
YM10	F	68	17.3	6	2.31	8.0	52	46,XX,del(5)(q31),-20,+22[20]
YM11	F	72	77.5	69	103.70	8.9	48	46, XX[20]
YM12	F	69	47.5	0	14.30	6.9	79	46,XX,t(11;19)(q23;p13.1)[6]/46,XX[8]
YM13	M	75	75.1	63	81.10	10.7	80	46,XY,del(7)(q22)[10]
YM14	F	80	50.3	24	27.96	9.3	223	46, XX[22]

YM15	M	68	22.5	21	7.05	11.5	214	46, XY[20]
YM16	M	84	25.5	0	2.73	9.1	39	47, XX, +21[36]/46, XX[7]
YM17	F	72	39.5	78	126.68	8.5	61	47, XY, +6,del(9)(q13q22),del(16)(q22q24)[15]
YM18	M	66	53.1	2	41.05	8.5	54	47, XY, inv(9)(p11q13),+21[20]
YM19	M	66	45.5	35	57.50	9.1	330	47,XX,+11[8]/46,XX[12]
YM20	M	72	47.2	61	6.96	7.7	61	46,XX,t(11;12)(p15;q13)[20]
YM21	F	70	36.5	65	9.14	10.2	96	46,XY,t(8;21)(q22;q22)[15]/46,XY[3]
YM22	M	69	87.1	37	14.06	8.1	31	46,XY[25]
YM23	M	78	36.4	47	19.88	6.8	48	47,XY,+13[18]/46,XY[2]
YM24	M	72	85	69	21.10	7.3	66	46,XY[21]
YM25	M	77	88.3	92	122.50	11.7	40	47,XY,1qh+c,+8[12]/46,XY,1qh+c[8]
YM26	M	78	22.4	0	63.82	7.2	34	46,XX[20]
YM27	F	76	85.1	91	110.59	4.6	61	46,XY,del(7)(q22)[7]/46,XY[10]
YM28	M	76	85.2	71	14.72	7.0	142	46,XY[20]

ABSTRACT (IN KOREAN)

급성 골수성 백혈병 환자의 저메틸화제 치료에 대한 반응을 예측하기 위한 유전학적 및 후성유전학적 특성

<지도교수 이 승 태>

연세대학교 대학원 의과학과

이 택 규

급성골수성 백혈병은 골수성 백혈구가 분화 단계 중 전구 세포 혹은 줄기세포에 변이가 발생하여 과도한 분열과 골수 내부에 축적으로인해 말초혈액을 따라 전신으로 퍼지는 질병이다. 결국 정상적인 골수 기능의 마비로 인해 조혈작용의 저하가 발생한다. 최근 DNA 저메틸화제를 사용하여 생존율 개선 효과를 입증하는 연구를 기반으로 해당 약물을 치료에 이용하고 있다. 하지만 DNA 저메틸화제에 대한 저항성이 초기에 발생하거나, 연속된 치료에 의해 발생한다. 이와 같은 이유로 인해 치료반응을 예측할 수 있는 생물학적인 지표를 정립하는 것이 중요한 문제로 떠오르고 있다. 본 연구에서는 DNA 저메틸화제를 이용해 치료를 진행한 급성골수성 백혈병 환자들에게서 약물 투여 이전 골수에 존재하는 DNA와 RNA를 통해 유전정보를 분석하였다. 구체적으로 후성유전학적 분석, 유전자변이 분석, 유전자발현 분석법을 통해 저항성을 예측할 수 있는 기전과 유전자를 확인하였다. 결과적으로 차세대 염기서열 분석법을 통해 저항성을 갖는 환자들의 특성과 기전을 정립하였다.

핵심되는 말: 급성골수성 백혈병, DNA저메틸화제, 차세대 염기서열 분석법