2019

Hypermethylation in the promoter of PRDM16 is associated with decreased survival in AML patients

CONNOR SMITH, BS

School of Medicine

Oregon Health & Science University

Certificate of Approval

This is to certify that the Master's Thesis of

Connor J. Smith

"Hypermethylation in the promoter of PRDM16 is associated with decreased survival in AML patients"

Has been approved

_____

Ted G. Laderas, Ph.D.

_____

Lucia Carbone, Ph.D.

_____

Michael A. Mooney, Ph.D.

Abstract

Acute Myeloid Leukemia (AML) is the most common acute adult leukemia, accounting for approximately 1% of all new cancers in 2018. DNA methylation is an epigenetic modification that has been shown to play a strong role in the development and progression of AML. Disruption of DNA methylation in patients is often due to somatic mutations of epigenetic modifier genes (e.g. (*DNMT3A, TET2, IDH1/2, WT1, EZH2, ASXL1*). Typically, studies of DNA methylation in AML and other cancers focus on regions of differential methylation (DMRs) between cancer patients and their healthy counterparts. Using the data from The Cancer Genome Atlas study on 194 AML patients, we investigated differential methylation between patients with and without mutations of the epigenetic modification genes commonly mutated in AML. We identified 20,567 differentially methylated probes within 2,893 DMRs and analyzed the affected regions for involvement in AML. In particular, we identified a region within the promoter of PRDM16 which, when hypermethylated, is associated with decreased survival ($p = 0.017$). In conclusion, we find DMRs between AML patients with and without somatic mutations of epigenetic modification genes that affect biologically relevant pathway and show association with the development and progression of AML.

Introduction

Acute Myeloid Leukemia (AML) is a cancer of the blood and bone marrow characterized by abnormal growth and differentiation of hematopoietic stem cells, the precursor to normal blood cells. As of 2016, AML was the most common form of acute leukemia in adults. [1] With an incidence rate of 4.8 (per 100,000) in 2014, AML continues to affect many patients every year. [2]

Cytogenetics have been established and verified as a strong prognostic predictor of AML [3, 4], though only 40 – 50% of AML patients have an abnormal karyotype [1], leaving open a need for investigation to other potential factors. Aside from karyotypic level abnormalities, known genomic mutations have been identified in the pathogenesis of AML, though there are fewer than in most solid tumors. [1] The Cancer Genome Atlas (TCGA) conducted a study of AML and found that in the 200 patients of their cohort, there was an average of only 13 genetic mutations in each patient. Of those, there were an average of 5 genes that would be recurrently mutated in AML. [5] Mutations in several genes have been determined as prognostically relevant in AML, including several genes directly and indirectly involved in DNA methylation (*DNMT3A, TET2, IDH1/2, WT1, EZH2, ASXL1*). [6, 7] DNA methylation, primarily *de novo* methylation, is believed to be a primary epigenetic mechanism involved in gene regulation during cell differentiation [8] and development. [9] In general, DNA methylation events in the promoter of a gene has an inverse effect on expression [10, 11] and, in some genes, has even been shown to mimic the effects of genomic mutations. [12] Since DNA methylation is a normal epigenetic process, [9] the abundance of mutations in the genes involved in establishing this mark is indicative of a possible predisposition to AML due to disrupted epigenetic machinery. [13]

Events of disrupted DNA methylation have been observed in leukemia [14] and have been implicated as a strong driver in the progression from intermediate myelodysplastic syndromes to AML. [15] With DNA methylation being a reversible process, targeted therapies have been developed to attempt to remove irregularities in the methylome of patients. [16] Identifying regions affected by

differential methylation is essential in order to drive such targeted therapies. Moreover, documenting which modifications are more likely to occur in patients carrying mutations in epigenetic modification genes may aid in the early detection and tailoring of such targeted therapies. Previous studies have found a number of mutation events in *DNMT3A* that lead to disruption of tetramer formation and processive function [17] and a global hypomethylation phenotype. [14] Additionally, mutations in histone modification genes such as *EXH2* and *ASXL1* have been shown to lead to changes in methylation [18] due to DNMT3A binding to them during histone modification. [19] Lastly, studies have found that *IDH1/2* and *TET2*, enzymes involved in the demethylation of CpG sites, disrupt the methylome with a general hypermethylation phenotype when mutated, [20] as well as when mutations to *WT1* cause loss of *TET2* function. [21]

Given the high frequency of mutations in key genes involved in establishing DNA methylation, we hypothesized that patients with mutations in epigenetic modifier genes (*DNMT3A, IDH1/2, TET2, WT1, EZH2, and ASXL1*) would display different DNA methylation patterns when compared to patients without these mutations. Moreover, we hypothesized that differential methylation would occur in regions biologically and functionally relevant to AML and therefore have a role in disease progression and/or survival. To test these hypotheses we leveraged data from The Cancer Genome Atlas (TCGA) study of AML [5] where methylation was measured using the Illumina Infinium® HumanMethylation450 BeadChip array.

Results

Characterization of samples

Methylation data for 194 samples were obtained from the Genomic Data Commons (GDC) Legacy Data Portal, for which 183 samples were retained for analysis after quality control. Mutation data was acquired from cBioPortal [22, 23]. Mutation data for each patient was screened for somatic

mutations in genes in the following mechanisms: DNA methylation (*DNMT3A*), DNA demethylation (*TET2, IDH1, IDH2, WT1*), or histone modification (*EZH2, ASXL1*). Due to small sample size within each group the samples were not divided in separate groups based on which gene group was mutated. . Of all AML samples, 81 were determined to have some form of somatic mutation in at least one of these genes and were labeled as "mutated"; the remaining 102 samples were labeled as "non-mutated". Patient characteristics are summarized between the two groups in Table 1. Cytogenetic risk is the only variable which is significantly imbalanced between the "mutated" and "non-mutated" groups (Pearson Chi-Square test, $p < 0.0001$).

Differential methylation analysis

Probe-wise differential methylation analysis was conducted between the "mutated" and "non-mutated" samples. The raw unprocessed data utilized in this study [5] was collected by TCGA using the Illumina Infinium HD HumanMethylation450K Array . This data was normalized and then scored as a β-value at each site. This value, ranging from 0 to 1, represents the percentage of methylation at that site for a given sample. The distribution of this value at each site for each group was compared to determine whether one group, the "mutated" patients, were differentially methylated compared to the "non-mutated" patients.

Between the two groups, 20,567 probes (of 427,462) remaining after preprocessing were identified as significantly different using a 5% False Discovery Rate cut-off (Figure 1) and the *limma* package from Bioconductor. [24] Probes where considered hypermethylated if the distribution of β values was significantly higher in the "mutated" group as compared to the "non-mutated" group and conversely for hypomethylated. Of these, 12,428 probes were identified as hypermethylated and 8,139 as hypomethylated. The distributions of these differentially methylated probes (DMPs) in relation to CpG islands were similar in both groups, with roughly 28.28% of probes in CpG Islands, 32.65% in open sea, 31.26% in CpG shores, and 7.8% in CpG shelves (Supplementary Figure S1).

Identified Differentially Methylated Regions and their involvement in cancer

Differentially Methylated Regions (DMRs) were identified (Materials and Methods), lifted over to the human assembly (GRCh38), then annotated using an in-house pipeline which matches regions to known coding regions, then promoters, and finally to the closest transcription start site (TSS) when found in intergenic regions. In total we identified 2894 DMRs, one of which was lost when we LiftOver from hg19 to hg38. DMRs range in size from 2 CpGs to 62 CpGs, with a median of 5 CpGs and a mean of roughly 7 CpGs. Distributions of DMR size and number of CpGs within DMRs by each chromosome can be seen in (Supplementary Figure S2).

In order to further functionally annotate DMRs, we employed the Genomic Regions Enrichment of Annotations Tool (GREAT) [25] that associates each gene with a 'regulatory domain' defined as 5 kbp upstream and1 kb downstream from the TSS and an extension within 1 Mb up to the regulatory domain of the nearest upstream or downstream genes. In the GREAT tool, biological processes related to cell fate commitment and hemopoiesis were found to be significantly enriched with DMRs at a false discovery rate of 5%, though their ranks were relatively low (93, 149, and 160 of 365, Table S1, Supplemental Data). GOrilla [26, 27] was used to look for enrichment of gene ontology terms. Several biological processes involved in DNA repair and histone methylation were identified as enriched with unadjusted p-values less than 0.001, though none of these were considered statistically significant after adjusting for multiple testing (Table S2, Supplemental Data).

Further investigation of the DMRs found were done by looking for previous associations in AML and other forms of cancer. This was done using the COSMIC Cancer Gene Census, a highly curated collection of genes and their involvement in different cancers, as a filter for genes with a strong, demonstrated involvement in AML [28]. Using this method, 12 genes associated with AML were found with DMRs in the gene body and/or promoter in 34 unique DMRs. A table of these genes can be found in

Table 2. The two genes with the highest number of DMRs, PRDM16 and DNMT3A, were chosen for further analysis.

PRDM16 was found to have 18 DMRs in our dataset with one located in the promoter, the highest among any genes related to AML or other cancers. Previous studies have found PRDM16 expression to have prognostic impact in adult [29] and pediatric [30] AML. Additionally, a mouse study found that epigenetic silencing of PRDM16 was associated with a fast progression from mixed lineage leukemia (MLL) to AML. [31] DNMT3A was found to have four DMRs in our dataset with one being in the promoter. While genomic mutations in DNMT3A are found frequently in patients with AML, [14] another study found that epimutations, which are regions of differential methylation within the promoter of DNMT3A, were found to have similar effects to global methylation as patients with DNMT3A mutations. [12]

Survival Analysis

Overall survival was available for all 183 patients via the clinical data acquired from cBioPortal. Univariate survival analysis was performed using the Kaplan-Meier estimator. [32] Epigenetic modifier mutation status, as labeled during this study, was first tested and found to not be significantly associated with overall survival ($p = 0.16$) (Supplementary Figure S3). In order to conduct survival analyses based on the methylation of the promoter for PRDM16 and DNMT3A, samples would need to be labeled as hypermethylated, hypomethylated, or intermediate methylation. To do this, mean methylation β values across the probe sites in the DMR were calculated for each patient. Once this was calculated, distributions of the mean methylation values were plotted, allowing threshold decisions to be made with the help of a biostatistician.

The mean methylation values within the promoter of PRDM16 appeared to take a bimodal distribution with the median close to 0.5 (Figure 2A). A β value of 0.5 was used to separate hypermethylated ($\beta>0.5$) and hypomethylated ($\beta<0.5$). Univariate survival analysis via the Kaplan-Meier

estimator found that hypermethylation of the promoter of PRDM16 was significantly associated ($p = 0.017$) with decreased survival (Figure 2B).

The mean methylation values within the promoter of DNMT3A were found to be closer to a normal distribution. After consultation with a biostatistician, the decision was made to label all samples in the lower quartile of mean β values as hypomethylated, all samples in the upper quartile as hypermethylated, and all remaining samples as intermediate methylation (Figure 3A). Univariate survival analysis via the Kaplan-Meier estimator found that aberrant methylation of the promoter of DNMT3A was not significantly associated ($p = 0.23$) with survival (Figure 3B).

Due to the strong prognostic impact of Cytogenetic Risk in AML [3, 5] and the disproportionate distribution of cytogenetic risk in the two patient groups (Table 1), it is possible that the results from the univariate model are detecting other factors. Multivariate survival analysis was conducted using a Cox Proportional Hazards model [33] to combine the age, cytogenetic risk, and PRDM16 hypermethylation covariates as a single model (Figure 4). 3 samples were removed as no data was available for cytogenetic risk group. Univariate analysis of cytogenetic risk of the patient cohort showed that poor and intermediate cytogenetic risk status were found to be significantly associated ($p < 0.001$ and $p = 0.049$, respectively) with increased risk ($HR > 1$), along with Age as a linear variable ($p < 0.001$). In this multivariate model, PRDM16 was suggestive, but not significantly associated with increased risk ($p = 0.059$).

Discussion

Differential Methylation between mutated and non-mutated patients

Typically, differential methylation is investigated between healthy control patients and patients with cancer. While this method may be able to illuminate regions of differential methylation due to the broad cancer phenotype, it is possible that smaller changes between patients will be left undiscovered.

This study took advantage of the dataset available from TCGA to determine if mutations in epigenetic modification genes significantly affect the methylation of these patients and potentially affect survival. Ideally, group labeling for this sort of study would consider the impact of the different types of mutations present. While mutations to *DNMT3A* lead to a generally hypomethylated methylome [34], mutations in *IDH1* and *IDH2* result in a generally hypermethylated phenotype. [20] This potentially antagonistic effect may lead to a cancellation of signal due to grouping samples that are significantly different from each other. In the case of this study, sub-grouping by the potential impact of mutations was not possible due to the low sample count.

The differential methylation analysis revealed many differentially methylated probes between the two groups, which then identified differentially methylated regions between the two groups. Using the pathway analysis tools GREAT and GOrilla, several biological processes in DNA repair, cell differentiation, and epigenetic modification were identified as enriched in the dataset. The genes of these pathways and processes are good candidates for further investigation. The use of the COSMIC Census allowed for an efficient filtering of a large list of DMRs. While this ensured high confidence of the region's involvement in AML, it causes the findings to be limited to previously discovered associations, potentially masking new discoveries. Future study of this dataset should endeavor to investigate regions of unknown significance to further advance understanding of the effects.

Hypermethylation of PRDM16 and Survival

In the univariate survival analysis, hypermethylation of the promoter of PRDM16 was found to be significantly associated with decreased survival. While this effect was not considered significant in the multivariate model ($p=0.059$), it was suggestive of an association with decreased survival. If hypermethylation of the promoter typically results in decreased expression [10, 11], then it is possible that the reduced survival is due to reduced expression of PRDM16; however, we were unable to test this hypothesis in this study. A previous mouse study found that the epigenetic silencing of PRDM16, which

acts as a histone methyltransferase, led to a decreased latency in development of AML from MLL. [31] This could have a similar effect in humans, leading to accelerated development of AML and potentially decreased survival. Conversely, another study into the prognostic impact of PRDM16 expression in AML patients found that increased expression of PRDM16 was associated with worse survival. [29] These competing results illustrate a potentially interesting effect of PRDM16 or a complex interplay between under and overexpression of PRDM16. The integration of expression data would need to be performed and analyzed in order to properly investigate this effect with the data from this study.

In summary, methylation values differ between AML patients with and without mutations in genes associated with methylation and chromatin modification. This study has discovered differentially methylated regions within genes pertinent to AML. Further research should look to continue investigating the effects of mutations to these genes and whether the changes to the methylome are a primary driving force in the development of AML.

Materials and Methods

Data Collection and Assembly

Raw methylation array data (TCGA Level 1) for 194 patients from the 2013 TCGA study of AML [5] was collected from the Genomic Data Commons (GDC) Data Portal using the GDC Legacy Archive. The methylation data was created using the Illumina Infinium HumanMethylation450K Array. Clinical and somatic mutation data for 200 patients were collected from cBioPortal [22, 23]. Patients were then screened for any somatic mutations in genes related to DNA methylation (DNMT3A), DNA demethylation (TET2, IDH1, IDH2, WT1), or histone modification (EZH2, ASXL1). If patients had a mutation in any of these genes, they were labeled as “mutated”. No distinction was made for type of mutation or existence of complex karyotypes. The remaining patients were labeled as “non-mutated”. The mutation status was then combined with the clinical data to be assembled into a sample sheet,

which is used as the phenotypic data for methylation array analysis. Of the 200 patients, 194 were retained for methylation analysis. All data was read into R using the *minfi* [35] package.

Quality Control

The methylation array includes 850 control probes used for quality control. Quality control of control probe intensities was conducted using Illumina's BeadArray Controls Reporter in conjunction with their Methylation Module v1.9 for Genome Studio 2011.1. The default thresholds recommended by Illumina were considered when flagging samples. Additionally, the *minfi* [35] and *wateRmelon* [36] packages from Bioconductor were used to assess and visualize the quality of the control probes for samples. Four samples were flagged for removal during this phase.

Probes with a low level of confidence, indicated by detection p-value and bead count thresholds, were removed using methods described by *wateRmelon*. [36] This resulted in the removal of 4760 probes.

Samples were assessed for their distribution of overall intensities, color bias, and sample relationship using methods described and implemented in the *lumi* [37] package from Bioconductor. One sample was identified as an outlier based on the distribution of intensities. When plotting the relationship of samples using a subset of probes with a coefficient of variance above 0.1, four samples separated completely from the rest of the samples. These samples were considered as outliers and flagged as removal. Finally, two additional samples were removed due to lack of documentation and metadata from the original study. Overall, 183 samples were retained.

Normalization

Eight methods of normalization were evaluated on the data, including Stratified Quantile Normalization [35], Subset-quantile Within Array Normalization (SWAN) [38], Functional Normalization [39], Normal-exponential out-of-band (NOOB) [40], DASEN and two variants (DANES and NANES) [36], and Beta-Mixture Quantile (BMIQ) [41]. Normalization methods were compared using

Color Bias and Beta by Probe Type density plots adapted from those in the *lumi* [37] package. SQN and BMIQ were not assessed for color bias correction via these plots as the raw intensities are unavailable after normalization. Normalization metrics from the *wateRmelon* [36] package were used to assess data quality after normalization, though only the normalization methods implemented in *wateRmelon* contained the needed data to test the SNP probe metrics. The DASEN method was chosen as the method to move forward with based on its performance on the data. After normalization, 81 "mutated" samples and 102 "non-mutated" samples remained.

Differential Methylation Analysis

Preparation of the normalized data and identification of Differentially Methylated Probes (DMPs) and Differentially Methylated Regions (DMRs) was conducted using methods detailed by the third version of *A cross-package Bioconductor workflow for analysing methylation array data.* [42] Probes were removed if they were on a sex chromosome (n = 11132), probes with known SNPs (n = 15879), and probes known to be cross-reactive (n = 26279) [43]. DMPs were identified through pairwise comparisons between "mutated" and "non-mutated" patients. A false discovery rate cut-off of 5% was used. After comparing the mutated samples to the non-mutated, 12428 probes were identified as hypermethylated, 8139 as hypomethylated, and 406895 as insignificant.

To investigate the regional effects of differential methylation, the final list of individual DMPs was used as input to determine differentially methylated regions (DMRs). 2894 DMRs were obtained using *DMRcate* [44], an annotation-agnostic method which only considers the location of DMPs within a 1000 base-pair window. The identified DMRs were then lifted over to the human genome build 38 using the UCSC Genome Browser *liftOver* tool. [45, 46] DMRs were then annotated using a pipeline developed by the Knight Cardiovascular Institute Epigenetics Consortium. The pipeline uses the current Ensembl annotation [47] to annotate each DMR for gene-body overlap (exon or intron), overlap with promoter regions, and finally the nearest transcription start site for intergenic DMRs.

Analysis of DMRs

The identified DMRs were then explored for pathway relationships using two complimentary tools, GOrilla [26, 27] and GREAT [25]. The input for GOrilla was a gene name list extracted from the resulting file from the annotation pipeline, using only records where the gene body was overlapped, retaining duplicate gene names. The input for GREAT was the unannotated DMR coordinates. The GREAT tool uses its own algorithm to annotate for cis-regulatory regions. Resulting pathways from each tool were assessed for potential involvement in AML and cancer.

To narrow the list of candidate genes, the annotation of the DMRs was filtered using data from the COSMIC Census [28]. This highly curated list of cancer genes was used to identify DMRs which overlapped genes previously shown to be pertinent in AML. DMRs located within PRDM16 (n = 18) and DNMT3A (n = 4) were chosen for further focus. Of these, one DMR for each gene was located in the promoter of its respective gene.

Survival Analysis

Univariate survival analysis was conducted using the Kaplan-Meier estimator [32] based on mutation status, methylation in the promoter of PRDM16, and methylation in the promoter of DNMT3A. Methylation of the promoters was defined by taking the average methylation of all probes in the DMR, then analyzing the distribution of these averages for all patients. For PRDM16, a bimodal distribution was seen and a β value of 0.5 was used to distinguish between hypermethylated ($\beta > 0.5$) and hypomethylated ($\beta < 0.5$) samples (Figure 2B). For DNMT3A, a near normal distribution was observed and methylation status were determined using quartiles to separate into hypermethylated ($\beta \geq$ top 25%), intermediate methylation (top 25% < B < top 75%), and hypermethylation ($\beta \leq$ top 75%) (Figure 3B).

Multivariate survival analysis was conducted using the Cox Proportional Hazards Model [33] on PRDM16 hypermethylation after it was found to be significant in the univariate model. This was done to

account for the strong prognostic power of cytogenetic risk [3, 5] and determine whether the addition of PRDM16 promoter methylation status further informs survival analysis.

Tables

Characteristic	"Mutated" (n = 81)	"Non-mutated" (n = 102)
Sex – no. (%)		
Male	37 (45.7%)	59 (57.8%)
Female	44 (54.3%)	43 (42.2%)
Age – years		
Mean	59.5	53.2
Median (Interquantile Range)	61 (52 - 68)	54.5 (42.3 - 65.8)
Range	21 - 88	22 - 83
Cytogenetic Risk – no. (%)		
Good	1 (1.2%)	33 (32.4%)
Intermediate	65 (80.2%)	40 (39.2%)
Poor	13 (16.0%)	28 (27.5%)
No Data	2 (2.5%)	1 (1.0%)
Survival – no. (%)		
Living	23 (28.4%)	39 (38.2%)
Deceased	58 (71.6%)	63 (61.8%)
Survival Time (Deceased) – months		
Mean	12.2	12.7
Median (Interquantile Range)	8.8 (4.2 - 17.2)	7.7 (2.2 - 17.1)
Range	0.1 - 46.8	0 - 55.4

Table 1 Patient characteristics within labeled groups. Patients were labeled as mutated if they had any somatic mutation in DNA methylation, DNA demethylation, or histone modification genes.

Gene Name	*Number of DMRs*	*DMRs Located In:*		
		Promoter	Exon	Intron
CBFA2T3	2	1	0	1
CDX2	1	1	1	0
CREBBP	1	1	1	1
CUX1	1	0	0	1
DNMT3A	4	1	0	4
HOXA13	1	1	1	1
HOXC13	1	1	1	1
KRAS	1	1	0	0
MNX1	1	1	0	1
PRDM16	18	2	3	17
PRRX1	1	1	0	0
RUNX1T1	2	2	1	2

Table 2 DMRs overlapping the gene body or promoter region of genes implicated in AML were identified for further analysis. Filtering was done using the COSMIC Census. DMRs may be overlap multiple elements (promoter, exons, and/or introns), so the values in the "DMRs Located In" columns will not add to those in the "Number of DMRs".

Figures

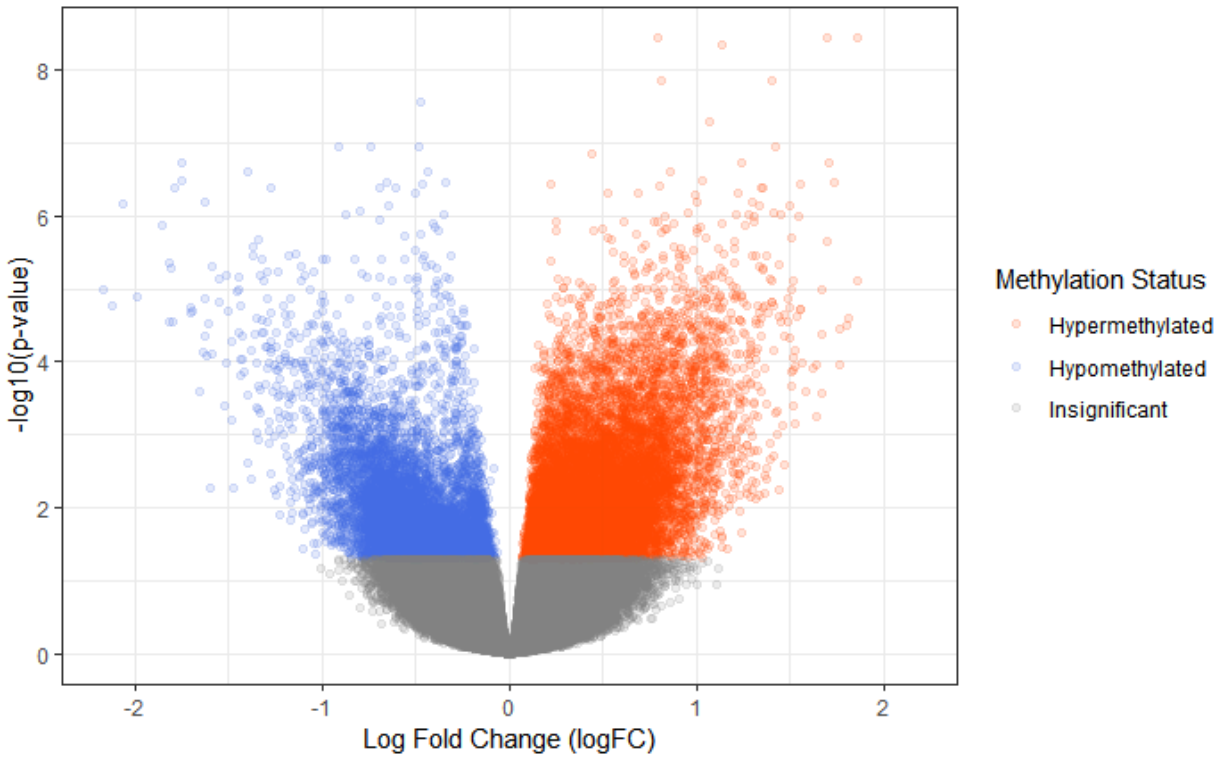

Figure 1 Volcano plot of differentially methylated probes. Probes with an adjusted p-value less than 0.05 were considered significant. Probes were labeled hypermethylated and hypomethylated based on their log fold change value being positive or negative, respectively.

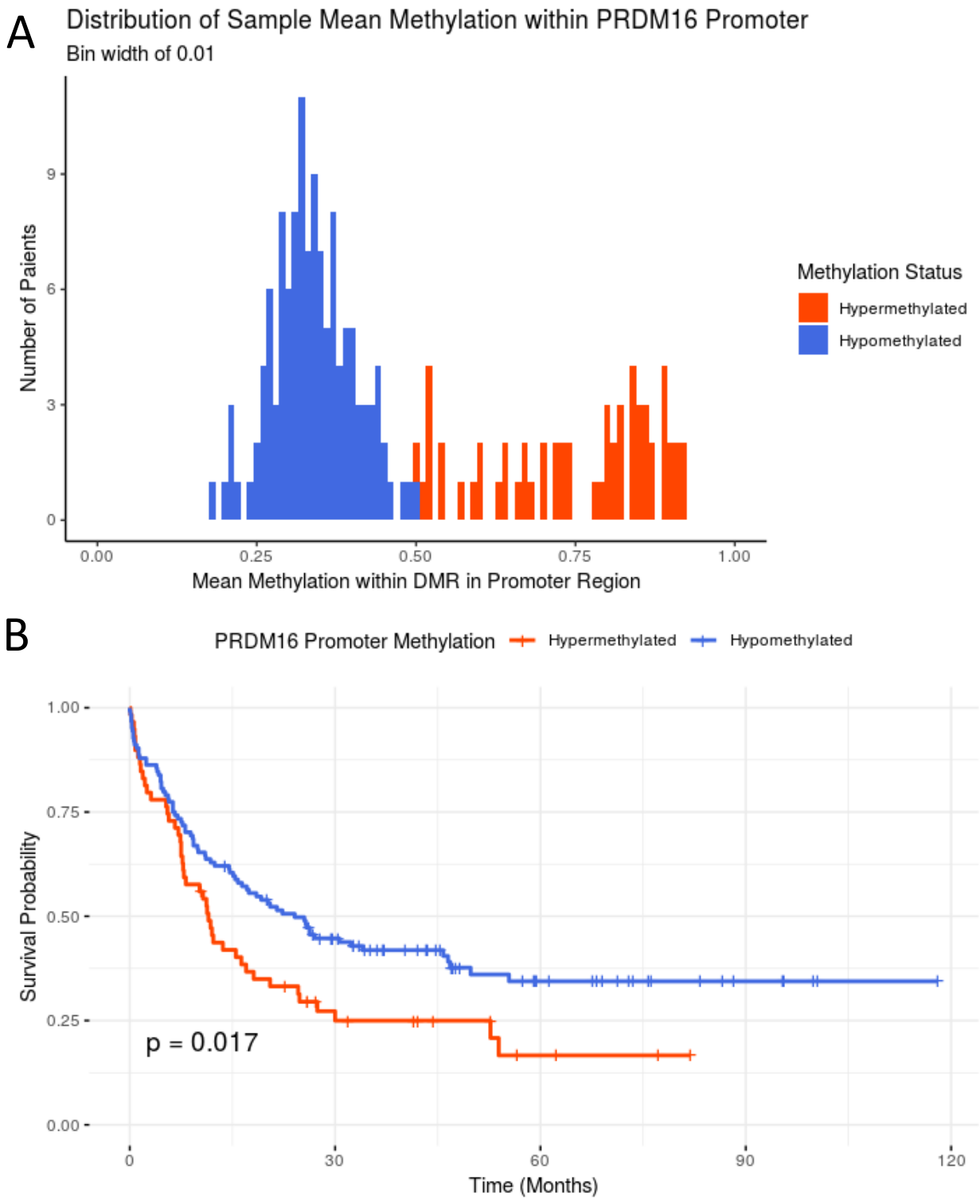

Figure 2 (A) Mean methylation values across the DMR in the promoter of PRDM16 were calculated by patient. These values were then plotted by frequency in a histogram with a bin width of 0.01. \β value of 0.5 was established as the threshold for determining methylation status. (B)Kaplan-Meier survival analysis of patients and their methylation status in the DMR located within the promoter of PRDM16. There is a significant association between hypermethylated patients (those with average methylation level of more than 0.5) in this region with a shortened overall survival time.

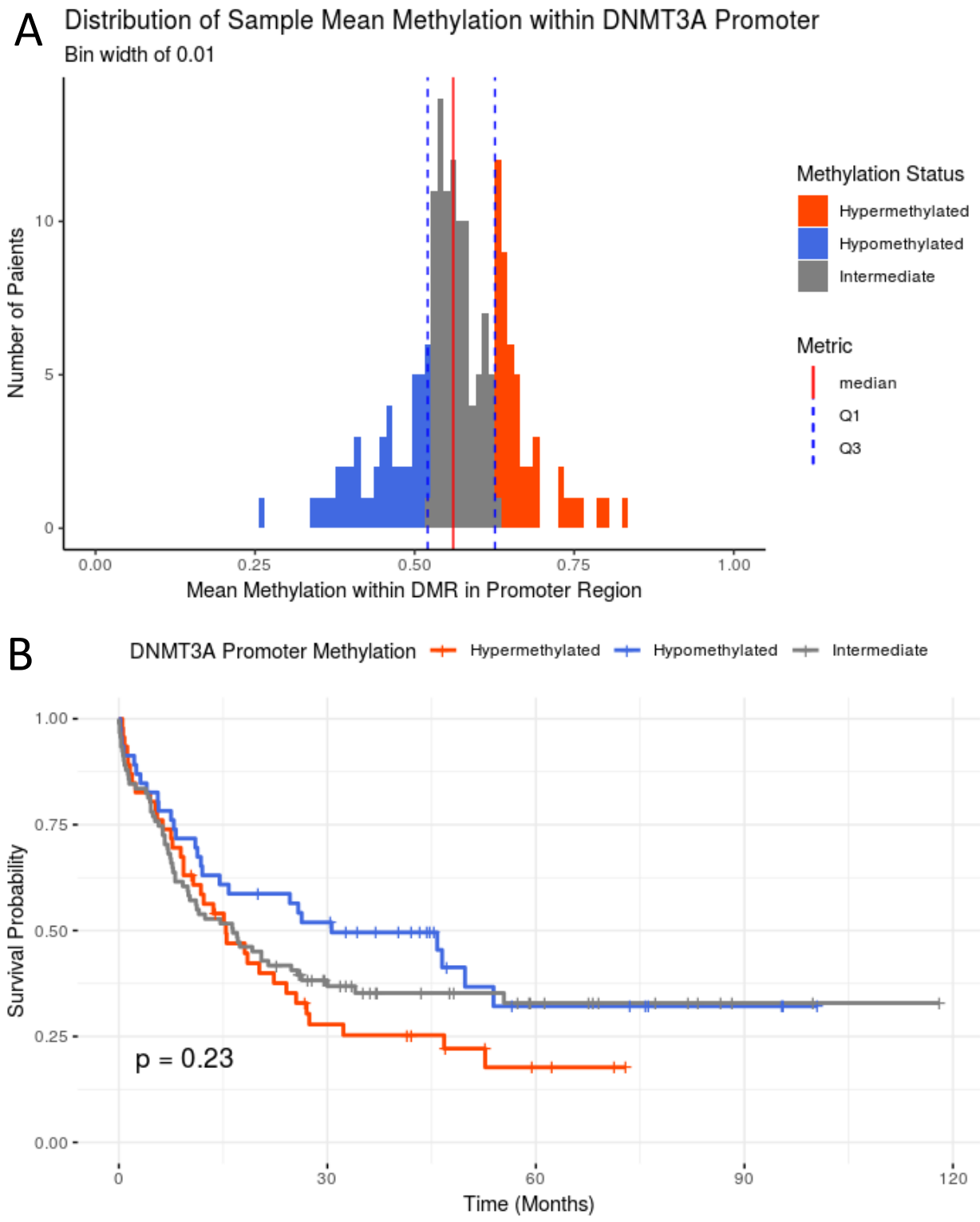

Figure 3 (A) Mean methylation values across the DMR in the promoter of DNMT3A were calculated by sample. These values were then plotted by frequency in a histogram with a bin width of 0.01. All samples in the lower quartile were labeled as hypomethylated, samples in the upper quartile as hypermethylated, and all remaining samples as intermediate. (B) Kaplan-Meier survival analysis of patients and their methylation status in the DMR located within the promoter of DNMT3A. No significant association was found (p= 0.23).

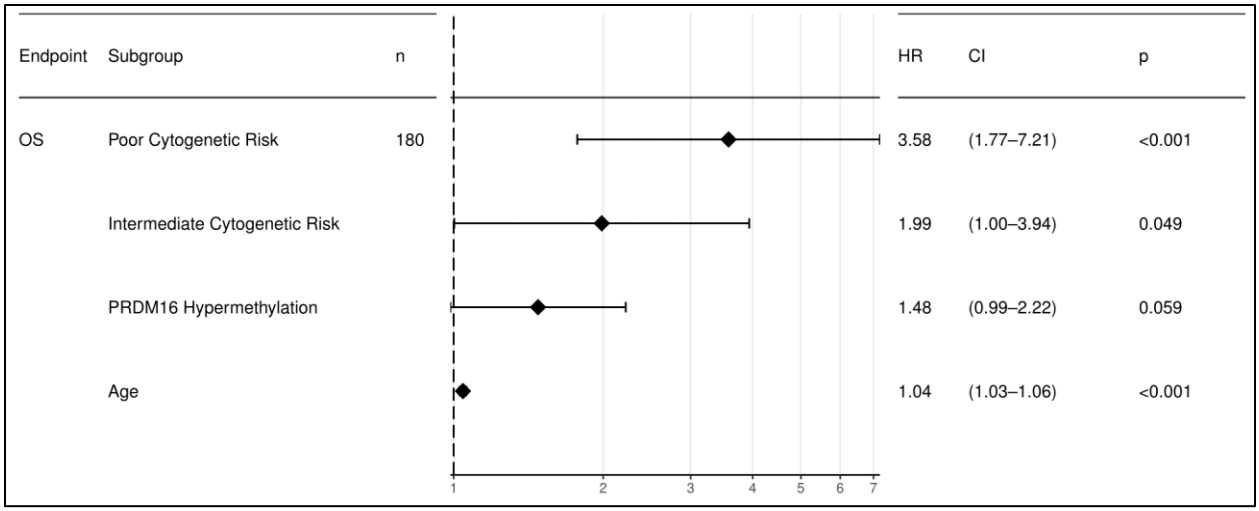

Figure 4 Multivariate survival analysis of age, cytogenetic risk, and PRDM16 hypermethylation by way of the Cox Proportional Hazards model.

References

1. Khwaja A, Bjorkholm M, Gale RE, et al. Acute myeloid leukaemia. Nature Reviews Disease Primers. 2016 2016-03-10;2:16010. doi: 10.1038/nrdp.2016.10. En.
2. Cancer Facts & Figures 2018. Cancer Facts & Figures. Atlanta, GA: American Cancer Society; 2018.
3. Grimwade D, Walker H, Oliver F, et al. The Importance of Diagnostic Cytogenetics on Outcome in AML: Analysis of 1,612 Patients Entered Into the MRC AML 10 Trial. Blood. 1998 1998/10/01;92(7):2322-2333. en.
4. Döhner H, Estey E, Grimwade D, et al. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. Blood. 2017 2017-1-26;129(4):424-447. doi: 10.1182/blood-2016-08-733196.
5. Network TCGAR. Genomic and Epigenomic Landscapes of Adult De Novo Acute Myeloid Leukemia. New England Journal of Medicine. 2013 May 30, 2013;368(22):2059-2074. doi: 10.1056/NEJMoa1301689.
6. Patel JP, Gonen M, Figueroa ME, et al. Prognostic relevance of integrated genetic profiling in acute myeloid leukemia. N Engl J Med. 2012 Mar 22;366(12):1079-89. doi: 10.1056/NEJMoa1112304. PubMed PMID: 22417203; PubMed Central PMCID: PMCPMC3545649.
7. Shen Y, Zhu Y-M, Fan X, et al. Gene mutation patterns and their prognostic impact in a cohort of 1185 patients with acute myeloid leukemia. Blood. 2011 2011/11/17;118(20):5593-5603. doi: 10.1182/blood-2011-03-343988. en.
8. Jasielec J, Saloura V, Godley LA. The mechanistic role of DNA methylation in myeloid leukemogenesis. Leukemia. 2014 Sep 2014;28(9):1765-1773. doi: 10.1038/leu.2014.163. eng.
9. Smith ZD, Meissner A. DNA methylation: roles in mammalian development. Nature Reviews Genetics. 2013 2013-03;14(3):204-220. doi: 10.1038/nrg3354. en.
10. Zahava S, Itamar S. DNA methylation and gene expression. Wiley Interdisciplinary Reviews: Systems Biology and Medicine. 2010 2010-05-01;2(3):362-371. doi: 10.1002/wsbm.64. en.
11. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nature reviews Genetics. 2012 May 29;13(7):484-92. doi: 10.1038/nrg3230. PubMed PMID: 22641018.
12. Jost E, Lin Q, Weidner CI, et al. Epimutations mimic genomic mutations of DNMT3A in acute myeloid leukemia. Leukemia. 2014 2014-06;28(6):1227-1234. doi: 10.1038/leu.2013.362.
13. Qu Y, Lennartsson A, Gaidzik VI, et al. Differential methylation in CN-AML preferentially targets non-CGI regions and is dictated by DNMT3A mutational status and associated with predominant hypomethylation of HOX genes. Epigenetics. 2014 August 6, 2014;9(8):1108-1119. doi: 10.4161/epi.29315.
14. Guillamot M, Cimmino L, Aifantis I. The Impact of DNA Methylation in Hematopoietic Malignancies. Trends in Cancer. 2016 2016-02-01;2(2):70-83. doi: 10.1016/j.trecan.2015.12.006. English.
15. Jiang Y, Dunbar A, Gondek LP, et al. Aberrant DNA methylation is a dominant mechanism in MDS progression to AML. Blood. 2009 2009/02/05;113(6):1315-1325. doi: 10.1182/blood-2008-06-163246. en.
16. Wouters BJ, Delwel R. Epigenetics and approaches to targeted epigenetic therapy in acute myeloid leukemia. Blood. 2016 Jan 7;127(1):42-52. doi: 10.1182/blood-2015-07-604512. PubMed PMID: 26660432.
17. Holz-Schietinger C, Matje DM, Reich NO. Mutations in DNA Methyltransferase (DNMT3A) Observed in Acute Myeloid Leukemia Patients Disrupt Processive Methylation. Journal of

Biological Chemistry. 2012 09/07/2012;287(37):30941-30951. doi: 10.1074/jbc.M112.366625. en.
18. Chaudry SF, Chevassut TJT. Epigenetic Guardian: A Review of the DNA Methyltransferase DNMT3A in Acute Myeloid Leukaemia and Clonal Haematopoiesis. BioMed Research International. 2017 2017;2017:5473197. doi: 10.1155/2017/5473197. eng.
19. Conway O'Brien E, Prideaux S, Chevassut T. The Epigenetic Landscape of Acute Myeloid Leukemia [Research article]. 2014 [updated 2014].
20. Figueroa ME, Abdel-Wahab O, Lu C, et al. Leukemic IDH1 and IDH2 Mutations Result in a Hypermethylation Phenotype, Disrupt TET2 Function, and Impair Hematopoietic Differentiation. Cancer Cell. 2010 December 14, 2010;18(6):553-567. doi: 10.1016/j.ccr.2010.11.015.
21. Rampal R, Alkalin A, Madzo J, et al. DNA hydroxymethylation profiling reveals that WT1 mutations result in loss of TET2 function in acute myeloid leukemia. Cell Rep. 2014 Dec 11;9(5):1841-1855. doi: 10.1016/j.celrep.2014.11.004. PubMed PMID: 25482556; PubMed Central PMCID: PMCPMC4267494.
22. Gao J, Aksoy BA, Dogrusoz U, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal. 2013 Apr 2;6(269):pl1. doi: 10.1126/scisignal.2004088. PubMed PMID: 23550210; PubMed Central PMCID: PMCPMC4160307.
23. Cerami E, Gao J, Dogrusoz U, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. 2012 May;2(5):401-4. doi: 10.1158/2159-8290.CD-12-0095. PubMed PMID: 22588877; PubMed Central PMCID: PMCPMC3956037.
24. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015 Apr 20;43(7):e47. doi: 10.1093/nar/gkv007. PubMed PMID: 25605792; PubMed Central PMCID: PMCPMC4402510.
25. McLean CY, Bristor D, Hiller M, et al. GREAT improves functional interpretation of cis-regulatory regions. Nature Biotechnology. 2010 05/02/online;28:495. doi: 10.1038/nbt.1630 https://www.nature.com/articles/nbt.1630#supplementary-information.
26. Eden E, Lipson D, Yogev S, et al. Discovering motifs in ranked lists of DNA sequences. PLoS Comput Biol. 2007 Mar 23;3(3):e39. doi: 10.1371/journal.pcbi.0030039. PubMed PMID: 17381235; PubMed Central PMCID: PMCPMC1829477.
27. Eden E, Navon R, Steinfeld I, et al. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics. 2009 Feb 3;10:48. doi: 10.1186/1471-2105-10-48. PubMed PMID: 19192299; PubMed Central PMCID: PMCPMC2644678.
28. Sondka Z, Bamford S, Cole CG, et al. The COSMIC Cancer Gene Census: describing genetic dysfunction across all human cancers. Nat Rev Cancer. 2018 Nov;18(11):696-705. doi: 10.1038/s41568-018-0060-1. PubMed PMID: 30293088.
29. Yamato G, Yamaguchi H, Handa H, et al. Clinical features and prognostic impact of PRDM16 expression in adult acute myeloid leukemia. Genes Chromosomes Cancer. 2017 Nov;56(11):800-809. doi: 10.1002/gcc.22483. PubMed PMID: 28710806.
30. Shiba N, Ohki K, Kobayashi T, et al. High PRDM16 expression identifies a prognostic subgroup of pediatric acute myeloid leukaemia correlated to FLT3-ITD, KMT2A-PTD, and NUP98-NSD1: the results of the Japanese Paediatric Leukaemia/Lymphoma Study Group AML-05 trial. Br J Haematol. 2016 Feb;172(4):581-91. doi: 10.1111/bjh.13869. PubMed PMID: 26684393.
31. Zhou B, Wang J, Lee SY, et al. PRDM16 Suppresses MLL1r Leukemia via Intrinsic Histone Methyltransferase Activity. Mol Cell. 2016 Apr 21;62(2):222-236. doi: 10.1016/j.molcel.2016.03.010. PubMed PMID: 27151440; PubMed Central PMCID: PMCPMC5061593.

32. Meier P. Nonparametric Estimation from Incomplete Observations AU - Kaplan, E. L. Journal of the American Statistical Association. 1958 1958/06/01;53(282):457-481. doi: 10.1080/01621459.1958.10501452.
33. Cox DR. Regression Models and Life-Tables. Journal of the Royal Statistical Society Series B (Methodological). 1972;34(2):187-220.
34. Qu Y. Aberrant DNA methylation in acute myeloid leukemia. Inst för medicin, Huddinge / Dept of Medicine, Huddinge; 2017 2017-05-18. eng.
35. Aryee MJ, Jaffe AE, Corrada-Bravo H, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics. 2014 May 15;30(10):1363-9. doi: 10.1093/bioinformatics/btu049. PubMed PMID: 24478339; PubMed Central PMCID: PMCPMC4016708.
36. Pidsley R, CC YW, Volta M, et al. A data-driven approach to preprocessing Illumina 450K methylation array data. BMC Genomics. 2013 May 1;14:293. doi: 10.1186/1471-2164-14-293. PubMed PMID: 23631413; PubMed Central PMCID: PMCPMC3769145.
37. Du P, Kibbe WA, Lin SM. lumi: a pipeline for processing Illumina microarray. Bioinformatics. 2008 Jul 1;24(13):1547-8. doi: 10.1093/bioinformatics/btn224. PubMed PMID: 18467348.
38. Maksimovic J, Gordon L, Oshlack A. SWAN: Subset-quantile within array normalization for illumina infinium HumanMethylation450 BeadChips. Genome Biol. 2012 Jun 15;13(6):R44. doi: 10.1186/gb-2012-13-6-r44. PubMed PMID: 22703947; PubMed Central PMCID: PMCPMC3446316.
39. Fortin J-P, Labbe A, Lemire M, et al. Functional normalization of 450k methylation array data improves replication in large cancer studies. Genome Biology. 2014 2014;15(11). doi: 10.1186/s13059-014-0503-2.
40. Triche TJ, Jr., Weisenberger DJ, Van Den Berg D, et al. Low-level processing of Illumina Infinium DNA Methylation BeadArrays. Nucleic Acids Res. 2013 Apr;41(7):e90. doi: 10.1093/nar/gkt090. PubMed PMID: 23476028; PubMed Central PMCID: PMCPMC3627582.
41. Marabita F, Almgren M, Lindholm ME, et al. An evaluation of analysis pipelines for DNA methylation profiling using the Illumina HumanMethylation450 BeadChip platform. Epigenetics. 2013 Mar;8(3):333-46. doi: 10.4161/epi.24008. PubMed PMID: 23422812; PubMed Central PMCID: PMCPMC3669124.
42. Maksimovic J, Phipson B, Oshlack A. A cross-package Bioconductor workflow for analysing methylation array data. F1000Research. 2017 2017-4-5;5:1281. doi: 10.12688/f1000research.8839.3. en.
43. Chen YA, Lemire M, Choufani S, et al. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. Epigenetics. 2013 Feb;8(2):203-9. doi: 10.4161/epi.23470. PubMed PMID: 23314698; PubMed Central PMCID: PMCPMC3592906.
44. Peters TJ, Buckley MJ, Statham AL, et al. De novo identification of differentially methylated regions in the human genome. Epigenetics Chromatin. 2015;8:6. doi: 10.1186/1756-8935-8-6. PubMed PMID: 25972926; PubMed Central PMCID: PMCPMC4429355.
45. Kent WJ, Sugnet CW, Furey TS, et al. The human genome browser at UCSC. Genome Res. 2002 Jun;12(6):996-1006. doi: 10.1101/gr.229102. PubMed PMID: 12045153; PubMed Central PMCID: PMCPMC186604.
46. Haeussler M, Zweig AS, Tyner C, et al. The UCSC Genome Browser database: 2019 update. Nucleic Acids Res. 2019 Jan 8;47(D1):D853-D858. doi: 10.1093/nar/gky1095. PubMed PMID: 30407534; PubMed Central PMCID: PMCPMC6323953.
47. Zerbino DR, Achuthan P, Akanni W, et al. Ensembl 2018. Nucleic Acids Res. 2018 Jan 4;46(D1):D754-D761. doi: 10.1093/nar/gkx1098. PubMed PMID: 29155950; PubMed Central PMCID: PMCPMC5753206.