Aberrant DNA Methylation is a Dominant Mechanism in MDS Progression to AML

Ying Jiang¹, Andrew Dunbar¹, Lukasz P. Gondek¹, Sanjay Mohan¹,², Manjot Rataul¹, Christine O’Keefe¹, Mikkael Sekeres², Yogen Saunthararajah*¹,², Jaroslaw P. Maciejewski*¹,²

¹Experimental Hematology and Hematopoiesis Section; ²Department of Hematologic Oncology and Blood Disorders, Taussig Cancer Center, Cleveland Clinic, Cleveland OH 44195

*These authors contributed equally to the work

Running title: Methylation arrays in AML and MDS

Key words: Methylation, FZD9, CpG, MDS, AML

Correspondence:

Yogen Sauntharajah, MD and Jaroslaw P. Maciejewski, M.D., Ph.D
Taussig Cancer Institute/R40
9500 Euclid Avenue
Cleveland OH USA 44195
Phone: 216-445-5962; FAX: 216-636-2498
E-mail: saunthy@ccf.org or macieji@ccf.org
Abstract

Myelodysplastic syndromes (MDS) are clonal hematological disorders that frequently represent an intermediate disease stage before progression to acute myeloid leukemia (AML). As such, study of MDS/AML can provide insight into the mechanisms of neoplastic evolution. In 184 MDS and AML patients, DNA methylation microarray and high-density SNP-A karyotyping were used to assess the relative contributions of aberrant DNA methylation and chromosomal deletions to tumor-suppressor gene (TSG) silencing during disease progression. Aberrant methylation was seen in every sample, on average affecting 91/1505 CpG loci in early MDS and 179/1505 loci after blast transformation (RAEB/AML). In contrast, chromosome aberrations were seen in 79% of early MDS samples, 90% of RAEB/AML samples, and were not as widely distributed over the genome. Analysis of the most frequently aberrantly methylated genes identified FZD9 as a candidate TSG on chromosome 7. In patients with chromosome deletion at the FZD9 locus, aberrant methylation of the remaining allele was associated with the poorest clinical outcome. These results indicate that aberrant methylation can cooperate with chromosome deletions to silence TSG. However, the ubiquity, extent and correlation with disease progression suggest that aberrant DNA methylation is the dominant mechanism for TSG silencing and clonal variation in MDS evolution to AML.
Introduction

Loss of heterozygosity (LOH) is an important event in neoplasia. For example, in the myeloid malignancies myelodysplastic syndrome (MDS) and acute myeloid leukemia (AML), chromosome 7 deletions are amongst the most important predictors of outcome. The contribution of chromosomal deletions to the neoplastic process may involve loss of critical growth regulating genes that would otherwise suppress the malignant phenotype. Such genes have been termed tumor suppressor genes (TSGs).

Methylation of promoter CpG sites is an important factor regulating gene expression including expression of TSGs. In cancer, aberrant DNA methylation in the promoters of TSGs is an alternative to chromosomal deletion for silencing TSGs. The clinical responses of MDS and AML to drugs that reverse aberrant hypermethylation, such as 5-aza-2'-deoxycytidine and 5-azacytidine, suggest that aberrant hypermethylation plays a causative role in disease and is not just a side effect resulting from dysregulated proliferation or DNA damage. Finally, TSGs may be silenced by mutation (reviewed in Payne et al).

Evolutionary principles provide an important framework for understanding the process of neoplastic transformation. In essence, the three forms of genomic instability described above generate clonal variation, with subsequent selection by the micro-environment for the most favorable variants (reviewed in Breivik et al). Since MDS frequently represents an intermediate stage before transformation to more aggressive disease (refractory anemia with excess blasts [RAEB] or AML) characterized by accumulation of myeloblasts, the study of MDS and AML can provide an insight into the relative contributions of aberrant methylation, chromosome instability or mutational instability to the process of neoplastic evolution. With the advent of gene array technologies, it has become practical to karyotype malignant cells at high resolution using high density single nucleotide polymorphism arrays (SNP-A) and more recently, to simultaneously screen a large number of promoters for aberrant methylation. Genome-wide sequencing of individual tumors to assess the mutation spectrum in its entirety is still not widely practical. In any case, it has been estimated that instability at the chromosome level is more prevalent than mutational instability at the nucleotide level.

3
In this study, we utilize gene array technologies to assess the extent of aberrant DNA methylation and chromosome aberrations in bone marrow samples from 184 MDS and AML patients. Findings were correlated with clinical outcomes. In addition to the identification of previously unrecognized TSGs, by analyzing aberrant methylation and chromosome abnormalities at different stages of disease, the relative contribution of these processes to myeloid neoplastic evolution was revealed.
Methods

Patients and cell-lines. Bone marrow aspirates (total N=184) were collected from patients with MDS or AML seen between 2002-2007 and grouped according to the World Health Organization (WHO) classification system \(^\text{15}\) and IPSS \(^\text{16}\) classifications of MDS. Aspirates obtained from healthy individuals patients (N=42) (mean age 43 years, range 15-76) were used as normal bone marrow controls for the methylation array. CD34+ hematopoietic precursors used as controls were isolated from total bone marrow from 9 normal donors.

Six human myeloid leukemia cell-lines were studied, of which K-562 and TF1 are frequently described as erythro-leukemia cell-lines. Of the cell-lines analyzed, KG1 and TF1 are CD34+ cell-lines.

For SNP-A analysis, samples from 116 healthy control individuals were used. Informed consent for sample and data collection was obtained according to the Declaration of Helsinki and conformed to protocols approved by the Cleveland Clinic IRB. In addition, a cohort of 61 CEPH (Utah residents with ancestry from northern and western Europe; CEU) HapMap individuals were used for comparison \(^\text{17}\); however, it should be noted that the criteria used to assign membership in the CEPH population have not been specified, except that all donors were residents of Utah.

Cytogenetic analysis. Cytogenetic analysis was performed on marrow aspirates according to standard methods. Chromosome preparations were G-banded using trypsin and Giemsa (GTG) and karyotypes were described according to ISCN \(^\text{18}\).

DNA extraction. DNA was extracted using the ArchivePure DNA Blood Kit (5Prime, Gaithersburg, MD, USA) from bone marrow aspirates. CD34+ cells were isolated from total bone marrow using immunomagnetic beads (StemCell Technologies Inc, Vancouver, BC, Canada).

Methylation array. Methylation detection was performed using Methylation Cancer Panel I and GoldenGate Assay kit with UDG from Illumina (San Diego, CA) according to the manufacturer’s protocol \(^\text{19,20}\). Briefly, bisulfite conversion of DNA samples was performed using the EZ DNA methylation kit (Zymo Research, Orange, CA). For each CpG site, two pairs of probes were designed which corresponded to
either the methylated or unmethylated state of the CpG site. Through allele-specific extension and ligation, PCR templates are generated and then amplified by PCR using fluorescently labeled common primers. The resulting PCR products were hybridized to a bead array at sites bearing complementary address sequences. These hybridized targets contained a fluorescent label that denotes a methylated or unmethylated state for a given locus. Methylation status of the interrogated CpG site is then calculated as the ratio of fluorescent signal from one allele relative to the sum of both methylated and unmethylated alleles (β-value).

The β-value provides a continuous measure of levels of DNA methylation at a CpG site, ranging from 0 in the case of completely unmethylated sites to 1 in completely methylated sites. Frequently, we provide the β-value of all 1505 CpG sites analyzed in a particular sample as an overall average β-value. The overall average β-value may mask the fact that some CpG sites may demonstrate increased methylation compared to controls, whereas other CpG sites may demonstrate decreased methylation compared to controls. Therefore, methylation status of a CpG site was also dichotomized by comparing β-values between disease samples and controls: for each CpG site, a control methylation level was determined by analysis of 42 bone marrow samples from healthy individuals or normal CD34+ hematopoietic cells. β-values from disease samples were compared to this control value and aberrant hypermethylation was defined as a β-value for a CpG site that was significantly greater (p<0.001) than the average β-value of the same CpG site in the control group. The image extraction and statistical analysis were done with Beadstudio Methylation Module (Illumina). Several different control oligos were used to ensure the highest quality data including allele-specific extension control, bisulfite conversion controls, extension gap controls, gender controls, first hybridization controls, second hybridization controls, negative controls and contamination detection controls.

**Methylation-specific PCR analysis.** The methylated DNA was enriched with MethylCollector Kit (Active Motif, Carlsbad, CA) according to the manufacturer’s protocol. CpG-methylated DNA fragments prepared by Mse I digestion specifically bind a His-tagged recombinant MBD2b protein. These protein-DNA complexes are captured with nickel-coated magnetic beads and subsequent wash steps are
performed with a stringent high salt buffer to remove fragments with little or no methylation. The methylated DNA is then eluted from the beads in the presence of proteinase K. Primers used were as follows: for *DCC*, 5'-TTGCCTGATCCAGCGTTTCCCT-3', 5'-TCCTTGTCGTCGGTTGGGTAAAGT-3'; for *DBC1*, 5'-AGCTCTTGGCTTCAATCTGGAGGT-3', 5'-TGTGTCCGAG-ACCTGAGTGGTTT-3'.

**Expression analysis using semi-quantitative RT-PCR.** *FZD9* expression was analyzed by semi-quantitative RT-PCR. Total RNA was extracted from total bone marrow cells with TRIzol reagent (Invitrogen, Carlsbad, CA) and cDNA was generated by first strand cDNA synthesis using the superscript II RT kit (Invitrogen, Carlsbad, CA). Primers used were as follows: 5'-TGATGAGCTGACTGGGCTTTGCTA-3', 5'-TCCATGTTGAGGCGTTCGTAGACA-3'.

**SNP-A analysis.** The Gene Chip Mapping 250K Assay Kit (Affymetrix, Santa Clara, CA, USA) was used for SNP-A analysis and utilized per the manufacturer’s instructions as previously described. When available, somatic changes not seen on metaphase cytogenetics were confirmed by re-running sorted CD3+ lymphocytes. In addition, for confirmation a number of samples (N=107) were run on ultra-high density Affymetrix 6.0 arrays and analyzed using Genotyping Console v2.0 (Affymetrix). Signal intensity was analyzed and SNP calls determined using Gene Chip Genotyping Analysis Software Version 4.0 (GTYPE). Copy number (CN) and areas of UPD were investigated using a Hidden Markov Model and CN Analyzer for Affymetrix GeneChip Mapping 250K arrays (CNAG v3.0) as previously described. Lesions identified by SNP-A were compared with the Cancer Genome Anatomy Project database (http://cgap.nci.nih.gov) and our own internal control series (N=116) to exclude known copy number variants.

**Statistical analysis.** Statistical significance of the difference between values of methylation level for different samples/groups was assessed using Beadstudio Methylation Module (Illumina). Spearman correlation was determined between methylation levels of replicate samples. Survival analyses were performed using the Kaplan-Meier product-limit method and the Log-rank test.
Results

Principles and validation of global methylation array. An array-based DNA methylation profiling platform was applied to assess the methylation status of 1505 CpG sites selected from the 5'-regulatory regions of 807 genes. The genes on this array were chosen on the basis of their importance to cancer development and progression, including TSG and oncogenes. The β-value provides a continuous measure of levels of DNA methylation at a CpG site, ranging from 0 in the case of completely unmethylated sites to 1 in completely methylated sites. Methylation status of a CpG site was also dichotomized by comparing β-values between disease samples and controls: aberrant hypermethylation was defined as an average methylation level (β-value) for a CpG site that was significantly (p<0.001) greater than the average methylation level of the same CpG site in the control cohort.

Methylation was analyzed in 6 leukemia cell lines, 184 patients with MDS or AML, 42 healthy control total bone marrow samples and 9 healthy control CD34+ hematopoietic precursor samples (Table 1). The MDS patients were divided into 3 groups for analysis: patients with disease that had transformed into a myeloblast-containing aggressive stage: a “RAEB/AML” group containing RAEB1, RAEB2, primary AML and secondary AML, a “low risk MDS” group containing refractory anemia (RA), refractory cytopenia with multilineage dysplasia (RCMD), refractory anemia with ringed sideroblasts (RARS), refractory cytopenia with multilineage dysplasia and ringed sideroblasts (RCMD-RS), 5q- syndrome, myelodysplasia unclassifiable (MDS-U), MDS/myeloproliferative disorders unclassifiable (MPD-U) and refractory anemia with ringed sideroblasts associated with thrombocytosis (RARS-T), and a group with a specific myeloproliferative variant of MDS – chronic myelomonocytic leukemia (CMML). To test the variation of the array, technical replicate samples were tested on 2 separate arrays. The average $R^2$ was 0.989 (Figure 1A), indicating the high reproducibility of the analyses.

To validate the results obtained through the methylation microarray, methylation-enriched PCR was used to detect the methylation level of exemplary genes (DBC1 and DCC). A strong correlation was observed between PCR results and array-based methylation analysis in the corresponding samples (Figure 1B).

To provide a proof of principle, we performed methylation array analysis on 6 leukemia cell lines (K562, TF1, KG1, Kasumi, U937 and HL60). Average (and median) methylation levels in the cell-lines
were higher than levels in the normal CD34+ hematopoietic precursor cell controls. The number of aberrantly methylated loci was significantly higher in all six leukemia cell-lines analyzed (Figure 2A), even when overall average methylation levels were not significantly increased (K562). This is because overall methylation levels can mask the fact that some CpG loci may be significantly hypermethylated while others may be significantly hypomethylated. Each cell line showed a distinct methylation profile, and hierarchical clustering revealed that all malignant cell lines could be easily separated from both normal total bone marrow and CD34+ cells. K562 and TF1, frequently described as erythro-leukemia cell-lines, clustered as a markedly distinct group (data not shown). Of the 1505 CpG sites studied, 58 were aberrantly hypermethylated in all 6 leukemia cell lines (concordantly methylated) compared to controls (Figure 2B), these included genes involved in apoptosis, cycling, differentiation, proliferation, adhesion, chromosome architecture and signal transduction (Figure 2B).

Promoter methylation correlates with disease evolution in patients with MDS/AML. We compared overall average methylation β-values, and the number of aberrantly methylated CpG sites, in patients with low risk MDS versus those with MDS advanced to the blast stage (RAEB) or AML. Average methylation and the number of aberrantly methylated CpG sites were significantly greater in patients with RAEB/AML (average β-value 0.353, average number of aberrantly methylated CpG sites 179/1505 in 138/807 genes) compared to patients with low risk MDS (average β-value 0.331, average number of aberrantly methylated CpG sites 91/1505 in 60/807 genes) or controls (average β-value 0.328, no aberrantly methylated genes) (Figure 3A). The difference between low risk MDS and controls was not statistically significant. CMML patients also displayed higher methylation levels (average β-value 0.336, average number of aberrantly methylated CpG sites 125/1505 in 94/807 genes) than controls; however, the difference was not statistically significant (P=0.076), possibly due to smaller sample size (N=14).

We reasoned that genes aberrantly methylated in a large proportion of patients (concordantly methylated genes) may more likely represent pathologically consequential events. For each patient group, we defined concordantly hypermethylated sites as those that were hypermethylated in at least 50% of the patients within that group. Based on this criterion, patients with RAEB/AML had the most concordantly methylated CpG sites (51 sites) followed by CMML (39 CpG sites) and low risk MDS (7 CpG sites). Among
the concordantly hypermethylated sites in patients with RAEB/AML, 22 were also found to be concordantly hypermethylated in patients with CMML (Figure 3B). In RAEB/AML (advanced myeloid malignancies), the hypermethylated genes include tumor suppressor genes (DCC, HIC1), genes involved in DNA repair (OGG1), cell cycle control (DBC1), development and differentiation (EPHA5, HHIP, NTRK1, MYOD1, TDGF1, FGF2, HOXA5) and apoptosis (ALOX12) (Figure 3B).

Four CpG sites (TSP50, NTRK1, 2 sites of GSTM1) were concordantly hypermethylated in both low risk MDS and RAEB/AML patients. In low risk MDS, concordant methylation was also noted in two CpG sites of STK23, a gene which plays a role in differentiation and development and 1 CpG site of BRCA1, a tumor suppressor gene.

Methylation arrays can be applied to identify differences between morphologically defined groups and also serially in the same patient during disease progression. As an example of serial detection, we analyzed an MDS patient with RA who progressed to RCMD to characterize the evolution of the aberrant methylation profile (Figure 3C). 25 out of 1505 CpG sites became hypermethylated during disease progression. A high proportion of the genes affected were in the group of concordantly hypermethylated genes in RAEB/AML (Figure 3B, D).

Concordantly hypermethylated CpG sites were dispersed over 19 chromosomes (Figure 4B). Focal concentration of hypermethylation was also observed in some chromosomal segments. For example, chromosome region 17p13.1p13.3 contained hypermethylation in a CpG site of HIC1, 2 CpG sites of ALOX12 and a CpG site of TNK1.

Comparing frequency of chromosome aberrations with frequency of aberrant DNA methylation. A combined study using both methylation array and chromosome analysis was performed in 112 patients. We looked for chromosome abnormalities using standard metaphase and 250K SNP array (SNP-A) based karyotyping (SNP-A can detect chromosome aberrations [including somatic uniparental disomy] at a high resolution allowing for identification of cryptic chromosome lesions not noted on standard metaphase karyotyping).

With routine metaphase cytogenetic analysis (standard karyotyping), 59 of 112 patients showed abnormal and 53 of 112 had normal chromosomes. Of the 53 patients with normal chromosomes by
standard karyotyping, 33 demonstrated chromosomal aberrations by SNP-A, i.e., no abnormalities were found by both standard karyotyping and SNP-A in 20 patients. The proportion of patients with chromosomal aberrations detected by either standard karyotyping or by SNP-A increased from low risk MDS (79%) to RAEB/AML (90%), however, this increase was not statistically significant (Figure 4A).

Aberrant methylation (of 182 CpG sites on average) was noted in all 20 patients with normal karyotype. In this group with normal cytogenetics by standard karyotyping and SNP-A, concordantly hypermethylated genes (defined as genes aberrantly methylated in >50% of patients) included *ALOX12*, *FZD9*, *HIC1*, *HPN* and *DBC1*. When individual chromosomes were studied, the proportion of patients with aberrant methylation detected on the specific chromosomes was much higher than the proportion of patients with an identified chromosome lesion (Figure 4C). Therefore, aberrant methylation was seen in all patients and also appeared to be more widely distributed over the genome than chromosome aberrations.

The average number of aberrantly methylated CpG sites was higher in patients with complex cytogenetic abnormalities, but not to a statistically significant extent (Figure 3A). Similarly, the presence or absence of cytogenetic abnormalities did not correlate with the average DNA methylation score as measured by logistic regression.

**Methylation at the FZD9 locus is an independent predictor of prognosis for MDS/AML patients.** The significant increase in aberrant CpG hypermethylation with progression of disease from low risk MDS to RAEB/AML suggests that hypermethylation of specific genes could be predictive of clinical outcomes. We reasoned that the genes most frequently methylated in patients with advanced disease (RAEB/AML) may be most useful in this regard. Therefore, methylation of *ALOX12*, *FZD9*, *GSTM1*, *HS3ST2* and *HIC1* (genes that were concordantly methylated in >70% of patients with RAEB/AML) were correlated with overall survival.

**FZD9** is a receptor for Wnt in the Wnt pathway that is important in carcinogenesis. In these analyses, aberrant methylation of a CpG site -458 base pairs from the transcription start-site of *FZD9* was significantly predictive of decreased survival at 12 months from diagnosis (Figure 5A) (p=0.005). The β-value of methylation at this site was also significantly predictive of survival in proportional hazards regression analysis (p=0.002) (data not shown). This CpG site was aberrantly methylated in 72% of
RAEB/AML cases, 48% of cases with low risk MDS and 2% of controls (Figure 5B). The hypermethylation of this CpG site inversely correlated with \( FZD9 \) transcript expression measured by semi-quantitative RT-PCR (Figure 5C).

Of 3 CpG sites at the \( FZD9 \) promoter locus, only one was concordantly methylated. Furthermore, 50 CpG sites from 23 genes surrounding the methylated \( FZD9 \) site on chromosome 7 were not extensively hypermethylated (Figure 5D). Therefore, hypermethylation of the \( FZD9 \) CpG site does not appear to be a component of a more widespread or propagated hypermethylation.

Aberrant methylation and chromosomal abnormalities may cooperate to completely silence a TSG (functional knock-out). Similarly, duplication of a methylated allele through UPD could potentially lead to complete TSG silencing. Therefore, we assessed the frequency of \( FZD9 \) methylation in patients with deletion or duplication of one allele through chromosomal abnormalities detected on standard metaphase karyotyping or SNP-A. Seventeen patients had chromosome 7 lesions that affected the \( FZD9 \) locus (11 with monosomy 7, 3 with del7/7q lesions and 3 with UPD7q) (Figure 5E illustrates lesions detected by SNP-A). Aberrant methylation was noted in 12/17 (71%) of patients with chromosome abnormalities at the \( FZD9 \) locus compared to 99/157 (63%) of patients without chromosome abnormalities at the \( FZD9 \) locus; this difference was not statistically significant (\( p=0.54 \) Chi-test) (Figure 5D, Table 2). Clinical outcome was poorest in patients with the combination of chromosome deletion of one \( FZD9 \) locus and aberrant methylation of the remaining \( FZD9 \) locus (Figure 5F). In multi-variate analysis, chromosome 7 abnormalities and blasts \( \geq 5\% \) were independent predictors of prognosis, and aberrant \( FZD9 \) methylation approached statistical significance (Table 2). Complex cytogenetic abnormalities were associated with prognosis in univariate analysis but this significance was lost in multi-variate analyses that included \( FZD9 \) methylation, blast \( \geq 5\% \) and chromosome 7 abnormalities (Table 2).

The mechanisms that drive aberrant DNA methylation are not known. One possibility is that aberrant DNA methylation could in some way be triggered by chromosome instability. Although aberrant methylation of \( FZD9 \) was more prevalent in patients with blasts \( \geq 5\% \), it was not associated with complex cytogenetic abnormalities (Table 2).
Discussion

Although aberrant methylation of specific genes has been noted in MDS and AML, identifying $p15$, E-cadherin, ER, MYOD1 and HIC1 as examples of TSG silenced by hypermethylation, most earlier studies were limited to analysis of a few specific loci. Therefore, the true extent of promoter methylation in hematological malignancy remains largely unknown. The data presented here represents the first systematic study of bone marrow from a large cohort of MDS and AML patients with a methylation array. Furthermore, the concurrent application of high density SNP-A allowed a comparison of aberrant methylation with chromosome abnormalities in the same individual cases.

The Golden Gate chip hybridization technology utilized by Illumina methylation arrays constitutes a robust analytic methylation platform and produces reliable and highly reproducible data. Technical replicates in 2 individual assay plates generated comparable results. Fidelity of the array was also confirmed using methylation-specific PCR for randomly selected genes. Finally, methylation status of $FZD9$, a gene selected for more in depth analysis, showed that $FZD9$-associated CpG site methylation correlated with decreased $FZD9$ transcript levels. The high density SNP-A has been validated in our previous work, in which we confirmed the clinical significance and acquired somatic nature of chromosome abnormalities detected by SNP-A in patients with normal cytogenetics by standard karyotyping.

Using methylation arrays, we show abnormal promoter methylation in more than one sixth of the genes analyzed in CD34+ and CD34- AML cell lines as well as primary RAEB or AML cells as compared to control CD34+ cells. The methylated genes have designated roles in multiple cellular functions including DNA repair, cell cycle control, regulation of development, differentiation and apoptosis. This observation agrees with other reports that many fundamental pathways related to tumor pathogenesis are inactivated by methylation. Our systematic analysis of a large cohort of patients with MDS and AML supports that aberrant methylation is a progressive process, noted in low risk MDS and significantly increasing in extent with more advanced disease. Aberrant DNA methylation was seen in every sample and was more common and more extensive than chromosome abnormalities detected by a high density SNP-A. The detected extent of chromosome aberrations is in keeping with studies using other techniques such as fluorescence in-situ hybridization (FISH) for recurrent chromosome abnormalities. Genome-wide DNA
sequencing of a high volume of primary tissue to assess the contribution of mutation to TSG silencing is not generally practical. When the coding sequences of more than 13,000 genes were analyzed in breast and colon cancer cell lines and xenografts, the estimate was for approximately 80 amino-acid altering mutant genes accumulated in each individual tumor. Mutation rates may be higher in individual cancer cases (‘mutator phenotype’ cases), but overall, genomic instability at the chromosome level has been shown to be more prevalent than genomic instability at the nucleotide level.

The extent and ubiquity of aberrant DNA methylation suggests that it could be the dominant force for TSG silencing and clonal variation that is then subject to positive selection in neoplastic evolution. While the wide-spread nature of aberrant hypermethylation suggests that some ‘passenger’ or epiphenomenon effect is plausible, the CpG sites represented in this array are believed to play a role in regulating gene expression, suggesting their methylation may well produce gene-expression changes that provide incremental competitive advantage. The association of methylation of specific genes with clinical outcome lends further support to the hypothesis that aberrant methylation plays a pathogenic role in MDS and its evolution to AML. In an effort to identify the biologically and clinically most relevant sites of promoter hypermethylation, we reasoned that concordantly (frequently) hypermethylated CpG sites/genes might be more likely to represent events that provide clonal advantage. In RAEB/AML, 51 CpG sites were hypermethylated in >50% of patients and 5 genes (ALOX12, GSTM1, HIC1, FZD9, HS3ST2) were hypermethylated in >70% of patients (methylation of HIC1 in MDS and AML has also been reported by others). We then examined if the methylation status of these sites predicted clinical outcomes. Analysis focused on FZD9 (frizzled 9), a receptor for Wnt in the Wnt pathway that is important in carcinogenesis. FZD9 methylation was an independent predictor of decreased survival. As an illustration of cooperation between chromosome aberrations and aberrant DNA methylation to completely silence TSGs, clinical outcome was poorest in patients with chromosome deletion of one allele of FZD9 and aberrant methylation of the remaining allele. This finding also suggests that in cases of chromosomal deletion, DNA methylation analysis of the remaining chromosome might identify a recessive TSG. Although studies to explain the mechanisms by which FZD9 repression impacts MDS and AML outcomes were beyond the scope of this study, the strong clinical correlations of FZD9 methylation suggests that it is at least one of the important TSGs on chromosome 7.
Loss of part of chromosome 7 (rather than the whole chromosome) is also frequently seen in MDS and AML and associated with a poor prognosis. We demonstrate that sometimes, small chromosome aberrations are confined to the 7q11 region that incorporates FZD9. However, there are more frequent deletions in the region of 7q22 and 7q33.\textsuperscript{17,42}

The mechanisms responsible for producing aberrant DNA methylation are unknown. Only one of three CpG sites at the FZD9 locus was a frequent target of aberrant methylation and only the methylated CpG site contained an E2F binding site. The E2F family has an important role in hematopoiesis and E2F binding is prevented by DNA methylation.\textsuperscript{43} However, we do not know if the E2F binding site is relevant to the aberrant methylation of this site. An analysis of the chromosome region around FZD9 demonstrated that methylation of this particular CpG site did not appear to be a part of propagated methylation of a wider chromosome region. Another possibility is that aberrant DNA methylation could in some way be triggered by chromosome instability. However, aberrant DNA methylation and complex cytogenetic abnormalities did not correlate to a statistically significant extent. Therefore, aberrant DNA methylation and cytogenetic abnormalities can cooperate to determine clinical outcome, but they appear to be mechanistically independent processes.

Other genes that were highly concordantly hypermethylated in RAEB/AML, including HIC1, ALOX12 and TNK1, were also noted to be frequently deleted or subject to uniparental disomy on SNP array analysis. TSP50 and NTRK1 were concordantly hypermethylated in all three disease groups analyzed (low risk MDS, RAEB/AML and CMML). Thus, it is possible that these 2 genes may be involved in pathogenetic steps that are common across the spectrum of myeloid malignancy. However, TSP50, which encodes a testis-specific protease, has been noted to be frequently hypomethylated in breast cancer.\textsuperscript{44} Therefore, aberrant DNA methylation may have differences between cancer histologies as well as pan-cancer similarities.\textsuperscript{45} NTRK1 is one of three known neurotrophic tyrosine kinase receptors and plays roles in cell differentiation, nervous system development and tyrosine kinase signaling.\textsuperscript{47}

Previous studies of specific methylated loci in MDS and AML have focused on genes such as p15, ER and E-cadherin.\textsuperscript{11,25,28-31} In our analysis, frequent aberrant methylation was noted at CpG sites at each one of these gene loci, with the highest aberrant methylation rate being 48% for a CpG site in the p15 promoter. However, as discussed earlier, our broader analysis demonstrates that a number of other CpG
sites at other gene loci were even more frequently aberrantly methylated. Even our study, limited to 1505 CpG sites, is only a reconnaissance into the extent and significance of aberrant methylation in leukemogenesis.

In conclusion, in this first comparison of high-throughput microarray promoter CpG methylation and SNP karyotyping analysis in MDS and AML, the ubiquity, extent and progression of aberrant hypermethylation with disease stage suggests that it is the dominant source of TSG silencing and clonal variation in the neoplastic evolution of myeloid disease. This observation provides further incentive to the development and refinement of drug strategies to reverse aberrant DNA methylation. Methylation array analysis may identify candidate TSG, for example FZD9 on chromosome 7, that may account in part for the poor prognosis associated with deletions of certain chromosomes in myeloid malignancies. Accurate clinical insight into malignant disease stage and prognosis may benefit from routine measurement of aberrant methylation at specific loci.

Acknowledgements

This work was supported by NIH R01 HL082983 (JPM), U54 RR019391 (JPM), K24 HL077522 (JPM), a grant from AA & MDS International Foundation and a charitable donation from the Robert Duggan Cancer Research Fund.

Authorship

Contribution: Y.J. performed experiments, helped Y.S analyze data, and prepared the manuscript under supervision by Y.S and J.P.M. J.P.M conceived the idea for this work which was further developed by Y.S. J.P.M obtained the funding and together with Y.S. designed the experiments and supervised the work. In addition, Y.S. analyzed and interpreted the data, and designed and wrote major portions of the manuscript. A.D., L.G, S.M. M.R., COK have performed some of the experiments and have compiled the databases. M.S. contributed to patient recruitment.

The authors declare no conflict of interest.
References


(49) Maurillo L, Buccisano F, Spagnoli A et al. Monitoring of minimal residual disease in adult acute myeloid leukemia using peripheral blood as an alternative source to bone marrow. Haematologica. 2007;92:605-611.
### Table 1. General characteristics of patients (N=226)

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>N</th>
<th>Sex(M/F)</th>
<th>Age(yrs)</th>
<th>Blast(%)</th>
<th>IPSS(mean)</th>
<th>Cytogenetics (normal/abnormal)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients</td>
<td>184</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AML</td>
<td>102</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary AML</td>
<td>56</td>
<td>23/33</td>
<td>54±15</td>
<td>43.0</td>
<td>29/25</td>
<td></td>
</tr>
<tr>
<td>Secondary AML</td>
<td>46</td>
<td>26/20</td>
<td>68±11</td>
<td>43.0</td>
<td>11/29</td>
<td></td>
</tr>
<tr>
<td>MDS advanced</td>
<td>28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAEB1/2</td>
<td>28</td>
<td>18/10</td>
<td>63±16</td>
<td>10.6</td>
<td>11/16</td>
<td></td>
</tr>
<tr>
<td>MDS Low risk</td>
<td>32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RA/RCMD</td>
<td>19</td>
<td>11/8</td>
<td>65±13</td>
<td>1.7</td>
<td>0.63&lt;sup&gt;a&lt;/sup&gt;</td>
<td>9/10</td>
</tr>
<tr>
<td>RARS/RCMD-RS</td>
<td>6</td>
<td>5/1</td>
<td>77±3</td>
<td>1.3</td>
<td>0.21</td>
<td>5/1</td>
</tr>
<tr>
<td>5q syndrome</td>
<td>3</td>
<td>0/3</td>
<td>55±31</td>
<td>0.3</td>
<td>0.17</td>
<td>0/3</td>
</tr>
<tr>
<td>MDS-U</td>
<td>4</td>
<td>2/2</td>
<td>74±8</td>
<td>2.0</td>
<td>0.75&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1/2</td>
</tr>
<tr>
<td>MDS/MPD</td>
<td>22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MDS/MPD-U</td>
<td>3</td>
<td>2/1</td>
<td>72±6</td>
<td>5.0</td>
<td>0.83</td>
<td>1/2</td>
</tr>
<tr>
<td>RARS-T</td>
<td>5</td>
<td>3/2</td>
<td>73±7</td>
<td>1.0</td>
<td>0</td>
<td>5/0</td>
</tr>
<tr>
<td>CMML1/2</td>
<td>14</td>
<td>5/9</td>
<td>67±10</td>
<td>18.3</td>
<td>11/3</td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>42</td>
<td>20/22</td>
<td>43±29</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Contains patients that cannot be analyzed for IPSS

Age data are given as the average ± the standard deviation (SD).

Abbreviations: refractory anemia (RA), refractory cytopenia with multilineage dysplasia (RCMD), refractory anemia with ringed sideroblasts (RARS), refractory cytopenia with multilineage dysplasia and ringed sideroblasts (RCMD-RS), myelodysplasia unclassifiable (MDS-U), myeloproliferative disorders unclassifiable (MPD-U), refractory anemia with ringed sideroblasts associated with marked thrombocytosis (RARS-T), chronic myelomonocytic leukemia (CMML).

In some analyses, primary AML and secondary AML were analyzed together with RAEB1/2. MDS low risk was analyzed together with MDS/MPD-U and RARS-T.
Table 2 FZD9 methylation is an independent predictor of prognosis in MDS/AML

<table>
<thead>
<tr>
<th></th>
<th>FZD9 methylated</th>
<th>FZD9 unmethylated</th>
</tr>
</thead>
<tbody>
<tr>
<td>P=0.005</td>
<td>Blasts ≥5% (n=126)</td>
<td>Blasts &lt;5% (n=58)</td>
</tr>
<tr>
<td></td>
<td>50% (29/58)</td>
<td>50% (29/58)</td>
</tr>
<tr>
<td></td>
<td>71% (90/126)</td>
<td>29% (36/126)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>FZD9 methylated</th>
<th>FZD9 unmethylated</th>
</tr>
</thead>
<tbody>
<tr>
<td>P=0.54</td>
<td>Normal Chr7 (n=157)</td>
<td>Abnormal Chr7 (n=17)</td>
</tr>
<tr>
<td></td>
<td>63% (99/157)</td>
<td>37% (58/157)</td>
</tr>
<tr>
<td></td>
<td>71% (12/17)</td>
<td>29% (5/17)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>FZD9 methylated</th>
<th>FZD9 unmethylated</th>
</tr>
</thead>
<tbody>
<tr>
<td>P=0.84</td>
<td>&lt;3 chr abnorm. (n=131)</td>
<td>≥3 chr abnorm. (n=43)</td>
</tr>
<tr>
<td></td>
<td>65% (28/43)</td>
<td>35% (15/43)</td>
</tr>
<tr>
<td></td>
<td>63% (83/131)</td>
<td>37% (48/131)</td>
</tr>
</tbody>
</table>

Chr 7 abnormalities and ≥5% blast are independent predictors of 12 month survival by multi-variate analysis. FZD9 methylation approaches significance.

<table>
<thead>
<tr>
<th>Cox Proportional Hazards Regression Analysis</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blasts ≥5%</td>
<td>0.019</td>
</tr>
<tr>
<td>Chromosome 7 abnormality</td>
<td>0.0008</td>
</tr>
<tr>
<td>Complex chr abnormality</td>
<td>0.17</td>
</tr>
<tr>
<td>FZD9 methylation</td>
<td>0.064</td>
</tr>
</tbody>
</table>

Chr = chromosome. Blasts % = bone marrow myeloblast %. The variables used in multi-variate analysis were FZD9 methylation status, presence of chromosome 7 abnormalities, presence of complex chr abnormalities (≥3 abnormalities on cytogenetic analysis), and blasts ≥5%.
Figure 1. Validation of the methylation array. (A) Methylation profiles of replicate samples show high reproducibility. Data shown are methylation β-values of 1505 CpG sites in which 1 and 0 represent high and low methylation levels respectively. Each dot represents a CpG site. (B) Methylation array data correlates with methylation enriched PCR results. Methylation enriched PCR was used to measure the methylation status of CpG sites of DBC1 and DCC genes in bone marrow mononuclear cells from 3 healthy control donors and 5 AML patients. Black bars indicate the methylation enriched PCR results represented as fold change compared to the results in the AML patient marked with *. Grey bars indicate the array average methylation values (β-value) for the same CpG sites.
Figure 2. Hypermethylated genes in leukemic cell lines belong to many functional groups. (A) Myeloid leukemia cell lines have increased CpG site methylation levels ($\beta$-values), and a high frequency of aberrantly methylated CpG sites, compared to normal CD34+ hematopoietic cells. The vertical boxes delineate the inter-quartile range in $\beta$-values and the T-bars the 95% range of values. The horizontal line in the boxes indicates the median values and the + sign indicates the mean. The asterix indicates significant $\beta$-value difference with CD34+ cells in a Tukey Studentized Range t-test. The numbers below the graph are the number of aberrantly methylated loci (compared to normal CD34+ control). The text explains why the frequency of aberrant methylation may be under-estimated if only $\beta$-values are analyzed. (B) Heat-map analysis showing the methylation $\beta$-values of 58 CpG loci which were concordantly hypermethylated in all six leukemia cell lines. Red, black and green colors correspond to high, medium and low methylation levels respectively. Genes were grouped according to their functions: group I, apoptosis and cell cycle; group II, cell development; group III, cell proliferation and adhesion; group IV, cell differentiation; group V, chromosome architecture; group VI, signal transduction; group VII, others.
Figure 3. Aberrant promoter methylation correlates with disease evolution. (A) RAEB/AML patients show a higher average and median methylation (β-value) (p=0.001) and a higher number of aberrantly methylated sites (p<0.0001) than those with low risk MDS (aberrant methylation was defined as a methylation β-value for a CpG site that was significantly greater (p<0.001) than the methylation β-value for the corresponding site in the group of normal bone marrow controls). The number of aberrantly methylated CpG sites in patients with complex cytogenetic abnormalities (defined as ≥3 abnormalities by standard metaphase karyotyping) is significantly higher (p=0.05) than in patients without complex cytogenetic abnormalities, however, the average methylation level (β-value) was not significantly different (p=0.23). Each dot represents the average array methylation level in a patient, the horizontal
dashed-line indicates the mean value for the patient group. Numbers below the graph represent the average number of aberrantly methylated CpG sites (of 1505 sites analyzed) in each patient group. ‘Normal’ is normal whole bone marrow. ‘CD34+’ are normal CD34+ selected hematopoietic cells. (B) The number of concordantly hypermethylated sites by patient group, defined as CpG loci that were aberrantly methylated in >50% of RAEB/AML, low risk MDS and CMML patients respectively. Each circle represents a patient group and the overlapping areas represent common aberrantly methylated genes. (C) CpG sites that became hypermethylated during evolution of RA to RCMD. Each dot represents a CpG site. The hypermethylated CpG sites are listed. (D, E) Concordantly hypermethylated loci in low risk MDS and RAEB/AML patients respectively. Each dot represents a CpG site. Genes shown more than once represent those with more than one hypermethylated loci. The percentage refers to the proportion of patients with aberrant methylation at the specified locus.
Figure 4. Aberrant DNA hypermethylation was detected much more frequently than chromosome aberrations. (A). Proportion of patients with chromosome aberrations detected by standard metaphase karyotyping or SNP-A by patient group. (B) Idiogram representing the chromosomal localization of concordantly hypermethylated loci (loci hypermethylated in more than 50% of low risk MDS and RAEB/AML patients). Chromosomes are numbered at the top. Gene names are indicated beside each hypermethylated locus. Genes marked with a box are hypermethylated in both low risk MDS and AML/RAEB. Genes shown more than once represent those with more than one hypermethylated locus. (C) On a chromosome by chromosome basis, the proportion of patients with aberrant methylation
detected on that chromosome was higher than the proportion of patients with a chromosome aberration involving that chromosome.
% of Patients with Aberrant Methylation

- Patients with monosomy 7 or UPD/deletion at FZD9 locus N=17
- Other AML/MDS patients N=157

Survival

- FZD9 unmethylated n=65
- FZD9 methylated n=119
  \( P = 0.005 \) log-rank

Graphs showing survival rates and methylation levels for various conditions.
Figure 5. Aberrant methylation at the FZD9 locus is a predictor of prognosis in MDS/AML. (A) FZD9 hypermethylation is associated with decreased 12 month survival. (B) FZD9 methylation levels in individual patients classified by disease group. Each dot represents the methylation β-value of the FZD9 CpG site in an individual patient. The average methylation β-value of this CpG site for each patient group is indicated by the horizontal line. (C) FZD9 methylation is inversely correlated to FZD9 transcript expression. Grey bars represent the degree of FZD9 methylation in the sample; the black bars represent the FZD9 expression levels determined by semi-quantitative RT-PCR, defined as fold change compared to the control marked with an asterisk. (D) FZD9 CpG methylation is not part of a wider hypermethylation at its chromosome 7 locus. The height of the vertical bars represents the frequency of aberrant hypermethylation of the CpG sites designated along the horizontal line. The space between bars marked with 0 indicates analyzed CpG sites that were not hypermethylated in any patient. (E) Regions of chromosome deletion and uniparental disomy (UPD), detected by SNP-A, that involve the FZD9 locus. FZD9 is designated by the vertical red bar. Red lines depict single SNP signal intensity, while green lines present an average value of SNP signal intensity. The horizontal purple bar represents areas of chromosome deletion, while the horizontal blue bar represents areas of loss of heterozygosity through UPD. (F) Chromosomal deletion or duplication of an FZD9 allele, combined with aberrant methylation of the remaining allele, is associated with a worse prognosis than either abnormality alone.
Aberrant DNA methylation is a dominant mechanism in MDS progression to AML

Ying Jiang, Andrew Dunbar, Lukasz P. Gondek, Sanjay Mohan, Manjot Rataul, Christine O'Keefe, Mikkael Sekeres, Yogen Saunthararajah and Jaroslaw P. Maciejewski