TET2 Mutations Affect Non-CpG Island DNA Methylation at Enhancers and Transcription Factor–Binding Sites in Chronic Myelomonocytic Leukemia

Jumpei Yamazaki1,2, Jaroslav Jelinek1,2, Yue Lu3, Matteo Cesaroni1, Jozef Madzo1,4, Frank Neumann2, Rong He2, Rodolphe Taby2, Aparna Vasanthakumar4, Trisha Macrae4, Kelly R. Ostler4, Hagop M. Kantarjian2, Shoudan Liang5, Marcos R. Estecio2,3, Lucy A. Godley4, and Jean-Pierre J. Issa1,2

Abstract

TET2 enzymatically converts 5-methylcytosine to 5-hydroxymethylcytosine as well as other covalently modified cytosines and its mutations are common in myeloid leukemia. However, the exact mechanism and the extent to which TET2 mutations affect DNA methylation remain in question. Here, we report on DNA methylomes in TET2 wild-type (TET2-WT) and mutant (TET2-MT) cases of chronic myelomonocytic leukemia (CMML). We analyzed 85,134 CpG sites [28,114 TET2-WT and 57,020 in non-CpG islands (NCGI)]. TET2 mutations do not explain genome-wide differences in DNA methylation in CMML, and we found few and inconsistent differences at CGIs between TET2-WT and TET2-MT cases. In contrast, we identified 409 (0.71%) TET2-specific differentially methylated CpGs (tet2-DMCs) in NCGIs, 86% of which were hypermethylated in TET2-MT cases, suggesting a strikingly different biology of the effects of TET2 mutations at CGIs and NCGIs. DNA methylation of tet2-DMCs at promoters and nonpromoters repressed gene expression. Tet2-DMCs showed significant enrichment at hematopoietic-specific enhancers marked by H3K4me1 and at binding sites for the transcription factor p300. Tet2-DMCs showed significantly lower 5-hydroxymethylcytosine in TET2-MT cases. We conclude that leukemia-associated TET2 mutations affect DNA methylation at NCGI regions containing hematopoietic-specific enhancers and transcription factor–binding sites. Cancer Res; 75(14): 2833–43. ©2015 AACR.

Introduction

TET2 [ten-eleven translocation (TET) oncogene family member 2] is a tumor suppressor gene on chromosome 4q24 (1). TET2 mutations were first described in myeloproliferative neoplasms (MPN; ref. 1), and were later also described in systemic mastocytosis (2), chronic myelodysplastic syndrome (MDS; ref. 3), myelodysplastic syndrome (MDS)/myeloproliferative neoplasms (MPN; ref. 1), and were later also described in systemic mastocytosis (2), chronic myelomonocytic leukemia (CMML; ref. 3), myelodysplastic syndrome (MDS; ref. 4), MDS/MPN (5), and acute myelogenous leukemia (AML; ref. 6). The incidence of TET2 gene alterations ranges from 10% to 50% in myeloid malignancies, with the highest frequency of mutations found in CMML, where TET2 mutations were noted in 35% to 50% of cases (5–7). As first reported for TET1 (8), TET2 converts 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC; ref. 9) as well as other covalently modified cytosines (10, 11) in embryonic stem cells, and thus mutations of TET2 were theorized to contribute to leukemogenesis by altering the epigenetic regulation of transcription through DNA methylation. In fact, among the three members of the TET gene family (TET1, TET2, and TET3), TET2 is the sole gene found to be frequently mutated in myeloid malignancies (6) and to disrupt hematopoietic differentiation (12, 13). Furthermore, in murine models, Tet2 deficiency impairs hematopoietic differentiation with the expansion of myeloid precursors (14, 15). However, the exact mechanism and the extent to which TET2 mutations affect DNA methylation remain in question. There are conflicting reports (12, 13, 16, 17) on the effect of TET2 mutations on DNA methylation. It has been reported that overall loss of 5mC content (hypermethylation; ref. 12) was a remarkable characteristic of CMML patients with TET2 mutations. Two groups studied TET2-mutant AMLs and CMML and identified a promotor hypermethylation phenotype (13, 17). Another group reported predominantly hypomethylation in TET2-mutant CMML, and we previously reported no effect of TET2 mutations on DNA methylation in CpG islands (CGI; ref. 16).
genome-wide studies reporting conflicting results used microarray analysis with limited validation. Here, we used the quantitative, sequence-based digital restriction enzyme analysis of methylation (DREAM; ref. 18) method to study this issue and found that hypermethylated sites in TET2-MT are mostly in non-CpG islands (NCGI) and are enriched at hematopoietic-specific enhancers marked by H3K4me1, and at binding sites for the transcription factor p300.

Patients and Methods

Patients

We analyzed whole bone marrow or peripheral blood samples (bone marrow was not available in one TET2-WT case) before treatment from 40 patients with CMML referred to The University of Texas MD Anderson Cancer Center (Houston, TX) or The University of Chicago (Chicago, IL), or enrolled in a multi-institution phase III trial comparing decitabine with supportive care (19). The Institutional Review Board at The University of Texas MD Anderson Cancer Center and The University of Chicago approved each institution’s respective protocols, and all patients gave informed consent for the collection of residual tissues as per institutional guidelines and in accordance with the Declaration of Helsinki.

Mutation analysis

For TET2 gene analysis, PCR and direct sequencing of exons 3–11 were performed starting from 20 ng of genomic DNA, as previously described (1). PCR amplicons were sequenced by Beckman Coulter Genomics (Beckman Coulter Genomics). All TET2 mutations were scored on both strands. Sequence traces were analyzed with SeqMan Pro (DNASTAR, Inc.) and reviewed visually. TET2 anomalies were numbered according to the European Molecular Biology Laboratory nucleotide sequence reference Fm3992369. Previously annotated SNPs in the HapMap database (20) were discarded. SIFT software (21) was used to determine the probability that a particular amino acid substitution is tolerated. We used pyrosequencing to analyze mutations of the R132 residue in IDH1, and residues R140 and R173 in IDH2, which have been reported in MDS (22) and glioblastoma (23). Mutations encoding amino acid R882 residue in the R132 residue in IDH1, and residues R140 and R173 in IDH2, which have been reported in MDS (22) and glioblastoma (23). Mutations encoding amino acid R882 residue in IDH1, and residues R140 and R173 in IDH2, which have been reported in MDS (22) and glioblastoma (23). Mutations encoding amino acid R882 residue in IDH1, and residues R140 and R173 in IDH2, which have been reported in MDS (22) and glioblastoma (23).

Digital restriction enzyme analysis of methylation

Genome-wide DNA methylation analysis using next-generation sequencing (18, 26) was performed for 20 samples for which a sufficient amount of DNA was available. Briefly, genomic DNA (5 μg) was digested with 5 μL of FastDigest SmaI endonuclease (Fermentas) for 3 hours at 37°C. Subsequently, 50 U (5 μL) of Xmal endonuclease (NEB) was added, and digestion continued for an additional 16 hours. The digested DNA was purified using a QIAquick PCR Purification Kit (Qiagen). The 3’ recessed ends of the DNA created by Xmal digestion were filled in with 3’-da tails added by Klenow DNA polymerase lacking 3’-to-5’ exonuclease activity (New England Biolabs) and a dCTP, dGTP, and dATP mix (0.4 mmol/L of each). Illumina paired-end sequencing adaptors were ligated using Rapid T4 DNA ligase (Enzymatics). The ligation mix was size selected by electrophoresis in 2% agarose. A slice corresponding to a 250- to 500-bp window, according to the DNA ladder, was cut out, and DNA was extracted from the agarose. Eluted DNA was amplified with Illumina paired-end PCR primers using iProof High-Fidelity DNA Polymerase (Bio-Rad) and 18 cycles of amplification. The resulting sequencing library was cleaned with AMPure magnetic beads (Beckman Coulter Genomics) and sequenced on an Illumina Genome Analyzer II or HiSeq 2000 (Illumina). Sequencing reads were mapped to Sma1 sites in the human genome (hg18), and signatures corresponding to methylated and unmethylated CpGs were enumerated for each Sma1 site. Methylation frequencies for individual Sma1 sites were then calculated. The methylation ratio is the ratio of the number of tags starting with CCGGG divided by the total number of tags mapped to a given Sma1 site. We used at least 10 sequencing reads to analyze methylation levels at individual Sma1 sites. On the basis of technical replicates (data not shown), we could distinguish differences in methylation of >10% with an FDR of 2.4%. We used the UCSC definition of CpG islands: GC content of 50% or greater, length >200 bp, ratio greater than 0.6 of observed number of CG dinucleotides to the expected number on the basis of the number of Gs and Cs in the segment (27). Sites at promoter regions are defined as being located within −1 kb to +500 b from transcription start sites of RefSeq genes.

Identification of Tet2-DMCs

P values for the methylation difference between TET2-MT and WT for each CpG site were calculated using a combinatorial approach. We created 1,000 pseudo-datasets by sampling patients between the two groups of TET2-mutant and TET2 wild-type patients without replacement. For each CpG site, we created the sampling distribution of the average difference between methylation levels in group one and group two. The proportion of shuffled datasets where the difference was greater than or equal to the real difference was defined as the P value. We identified 506 CpG sites with a P value less than or equal to 0.01. After we set the cutoff for the minimum difference of 5% to exclude low-level changes of uncertain significance, the number of CpG sites differentially methylated in TET2-mutant and TET2 wild-type CMML patients dropped from 506 to 472. We applied the same approach to a control dataset generated from 20 samples of total white blood cells (WBC) from age-matched healthy subjects to evaluate the probability of obtaining 472 differentially methylated CpG sites by creating random combinations of two groups containing 8 and 12 samples. After reshuffling all possible 125,970 combinations and applying the filter for a minimum 5% difference between the two groups, the probability of obtaining 472 differentially methylated CpG sites by random reshuffling data from healthy WBCs was 0.0074 (935 combinations of 125,970 total).

Quantitative DNA methylation analyses by bisulfite pyrosequencing

We used bisulfite pyrosequencing to quantitatively assess DNA methylation (28) for differentially methylated genes from the DREAM analysis as well as for ATM and SP140 from a previous report (16). We analyzed the same sites of these genes that were analyzed by DREAM. The number of patients with successful results (mostly >90% success rate) varied slightly for each gene. Primer sequences are listed in Supplementary Table S1.
Quantitative real-time PCR
RNA was isolated using TRIzol (Invitrogen). cDNA was synthesized using High Capacity cDNA Reverse Transcription Kits (Applied Biosystems). qPCR was performed on a StepOne Real-time PCR System (Applied Biosystems) using SYBR Green gene expression assays (Bio-Rad). Gene expression data were normalized to GAPDH. Primer sequences are listed in Supplementary Table S1.

Gene Set Enrichment Analysis
For Gene Set Enrichment Analysis, gene sets were downloaded from the Broad Institute’s MSigDB website (29). Gene set permutations were used to determine the statistical enrichment of the gene sets using the difference in gene expression set permutations were used to determine the statistical enrich-

Analysis for enrichment of transcription factor–binding sites and enhancer sites
The locations of transcription factor–binding sites and enhancer sites were downloaded from ENCODE project data (31). We annotated individual Smal sites in the DREAM analysis depending upon whether each site was in peaks for transcription factors and enhancers (marked by H3K4me1). We calculated fold enrichment scores of tet2-DMCs over all sites analyzed for each transcription factor–binding site and enhancer site. We also mapped individual Smal sites in the DREAM analysis to regulatory regions by publicly available Ensembl Regulator Build data (32).

5hmC pull-down assay followed by qPCR
The 5hmC affinity purification was performed as previously described (33). Briefly, sonicated genomic DNA (15–30 μg) was labeled with chemically modified uridine diphosphoglu-
cose glucose (UDP-6-N3-Glu). The click chemistry reaction was performed by addition of 150 μmol/L Biotin-S-S-DBCO (dibenzylcylooctyne). After pull-down with streptavidin mag-
cetic beads (Dynabeads, MyOne Streptavidin C1, Invitrogen), DNA was released with 50 mmol/L DTT and purified by MinElute Reaction Cleanup Kit (Qiagen). DNA concentration after affinity enrichment was measured by the Quant-IT Pico-
Green dsDNA quantitation assay (Invitrogen). The enrichment for target loci was assessed by qPCR using the Power SYBR Green assay (Applied Biosystems) and the 7500 Applied Bio-
systems PCR machine. Primer sequences are listed in Supple-
mentary Table S1.

Statistical analysis
Statistical analyses were performed using PRISM (GraphPad Software, Inc.). We used the Mann–Whitney test to compare continuous variables of DNA methylation levels between TET2-
MT and TET2-WT cases. All P values were two tailed. Unsupervised hierarchical analyses were performed by ArrayTrack (http://ekdb.fda.gov/webstart/arraytrack/) with standard crite-
ria. Principal component analysis was performed in R using the princomp function in the stats package. Fisher exact test was used to calculate the enrichment of tet2-DMCs over all sites analyzed for enhancers and transcription factors within known regulatory regions of CD14+ monocytes and the GM12878 lymphoblastoid cell line.

Results
TET2 mutation status in CMML
We first analyzed the mutational status of the TET2 coding sequence (exons 3–11) in samples from 20 patients with CMML, according to WHO criteria. TET2 missense or nonsense mutations were detected in 8 out of 20 patients (40%) studied. Five patients had a single heterozygous mutation, two had a biallelic or homozygous mutation, and one had two mutations. Altogether, nine mutations were identified, including two missense, four nonsense, and three frameshift mutations. Detailed mutation information is shown in Table 1. Using SIFT software (21), both of the identified missense mutations were predicted to affect protein function. Furthermore, we identified an IDH2 R140Q mutation in 1 out of 20 CMML patients (5%), and a mutation at the R882 residue in DNMT3A was found in the same patient (TET2-WT).

Genome-wide DNA methylation analysis
We used DREAM (18) using next-generation sequencing, which allowed us to identify differentially methylated sites in the human genome for TET2-MT and TET2-WT cases at high resolution, independently of bisulfite treatment. From all the samples used for DREAM, 5 to 94 million unique usable reads (quality filtered and aligned to the human genome) were successfully generated for DNA methylation analyses (Supplementary Table S2). Supplementary Figure S1 shows representative DREAM data for DNA methylation from two TET2-MT and two TET2-WT cases compared with normal peripheral blood. Compared with normal blood, hypermethylation in CGIs and hypomethylation in NCGIs were found in a considerable number of sites in all patients regardless of TET2 mutation status (Supplementary Fig. S1). This was also the case when averages of DNA methylation levels in each population were analyzed. Direct comparison of averages of DNA methylation in TET2-MT versus TET2-WT cases revealed a slight increase of CGI hypermethylation in TET2-WT cases and NCGI hypermethylation in TET2-MT cases (Fig. 1). There were no differences in the methylation status of 7 classes of repeat sequences examined (SINE, LINE, LTR, etc.; Supplementary Fig. S2).

Next, we analyzed 38,282 CpG sites (all those with >9 reads in all 20 patients) alongside those in five normal blood samples using unsupervised hierarchical clustering analysis. Using either
average DNA methylation in each population for each site analyzed which had >9 reads in more than 10 out of 20 patients. In this setting, 85,134 CpG sites were analyzed (28,114 sites in CGIs and 57,020 sites in NCGIs). Volcano plots of these sites revealed that TET2-MT cases have more NCGI methylated sites, whereas minor differences were seen in CGIs (Fig. 3A–C). Using permutation analysis to control for overfitting, we found 472 CpG sites (0.55%) that were differentially methylated in TET2-MT and TET2-WT cases (TET2-specific differentially methylated CpGs; tet2-DMCs; see Materials and Methods). We found more methylated sites in TET2-MT (375 sites) than in TET2-WT cases (97 sites), supporting previous findings of an increased amount of 5mC in TET2-MT compared with TET2-WT cases (16). Interestingly, we found a strikingly different biology at CGI and NCGI sites. Of these tet2-DMCs, 13% are in CGIs (63 sites) and 87% are in NCGIs (409 sites). Thus, 0.22% of CGI sites and 0.71% of NCGI sites were affected by TET2 mutations (P < 0.0001 for the difference between CGI and NCGI). Furthermore, among the 63 CGI sites, 62% were more methylated in TET2-WT (39 sites). In contrast, among the 409 NCGI sites, 86% (351 sites) were more methylated in TET2-MT (P < 0.0001). The differences were confirmed in additional cases by bisulfite-pyrosequencing for several genes (Fig. 3D and E). Next, we performed supervised hierarchical clustering analysis for only tet2-DMCs and found clearer clusters than in the analysis of all the sites (Supplementary Fig. S4), suggesting that TET2 mutations mostly affect tet2-DMCs.

To further gain insight into the characteristics of tet2-DMCs, we compared their DNA methylation levels to that seen in normal blood. This analysis gave us an idea of the role of the temporal and spatial difference of tet2-DMCs in leukemogenesis in CMML. We again found a striking difference between CGIs and NCGIs. Tet2-DMCs in CGI sites were often unmethylated in normal blood and hypermethylated in TET2-WT cases but not in TET2-MT cases (Fig. 4A). On the other hand, there is a considerable number of tet2-DMCs in NCGIs whose DNA methylation levels in normal blood range from intermediate to high (Fig. 4B). For these sites, TET2-WT cases showed hypomethylation, whereas TET2-MT cases showed identical to higher methylation levels when compared with normal blood. To clarify these findings, we plotted the difference in DNA methylation between TET2-MT and TET2-WT cases for tet2-DMCs against DNA methylation levels in normal peripheral blood. Overall, tet2-DMCs in CGIs are mainly the sites whose methylation levels are higher in TET2-WT cases than in TET2-MT cases and that are unmethylated in normal blood (Fig. 4C). In contrast, tet2-DMCs in NCGIs are the sites whose methylation levels are higher in TET2-MT than in TET2-WT cases, and are immediately to fully methylated in normal blood (Fig. 4D). In other words, TET2-WT CMML is characterized by loss of methylation at these normally methylated sites, whereas TET2-MT CMML shows preserved or enhanced methylation at these sites.

**Functional relevance of TET2 mutations to leukemogenesis in CMML**

To address the possible mechanism of TET2 mutations in leukemogenesis in CMML, we sought a functional relevance for methylated sites in TET2-MT cases by correlating them with gene expression levels in TET2-MT, TET2-WT, and normal peripheral blood. We looked at expression profiles for genes with tet2-DMCs.
at their promoters (−1,000 bp to +500 bp from transcription start site), because DNA methylation at promoters is well known to be correlated with gene expression. We measured gene expression levels for GGA2, AIM2, and SP140, which have NCGI promoters hypermethylated in TET2-MT cases. Within the complement of NCGI sites, TET2 mutations affected promoters and nonpromoter sites equally. Indeed, the CpG sites previously validated as potential TET2 targets were AIM2 and SP140, both in NCGI promoters (16). As expected, we found a strong correlation between DNA methylation and gene repression (Fig. 5A).

We also selected three genes with tet2-DMCs in NCGIs more methylated in TET2-MT at nonpromoter regions and found that gene expression changes also correlated with DNA methylation at these sites (Fig. 5B). Interestingly, TET2-WT cases showed expression levels similar to TET2-MT cases in at least two out of three genes analyzed, but had lower DNA methylation levels compared with TET2-MT cases.

Finally, to understand the mechanisms by which these genes are regulated by tet2-DMCs in NCGI, we focused on enhancers and transcription factors, which are known to control gene expression in cis and trans from distal regions such as gene-body and outside the genes (34). We found that 25% (104 sites out of...
409) of tet2-DMCs in NCGI were shown to be at enhancer sites marked by H3K4me1 in a lymphoblastoid cell line, which was significantly enriched compared with all NCGI sites analyzed (16%, \( P < 0.001 \); Fig. 5C and Supplementary Fig. S5). We also noticed that these sites are localized at several transcription factor–binding sites as well. Among all analyzed, binding sites for p300 in the lymphoblastoid cell line were colocalized with 2.0% of these sites, which was a 4-fold enrichment compared with all NCGI sites analyzed (0.5%, \( P = 0.001 \); Fig. 5D and Supplementary Fig. S5). We also observed that tet2-DMCs in NCGIs were significantly enriched in enhancer regions and in regions flanking promoters but depleted within active promoters in CD14\(^+\) monocytes and the GM12878 lymphoblastoid cell line (supplementary Fig. S6). No enrichment was observed in CTCF-binding sites. Altogether, our data suggest that methylation at tet2-DMCs in NCGIs is linked with dysregulation of gene expression through altering hematopoietic specific transcription factor–binding sites and enhancers.

**Figure 3.** Difference in DNA methylation levels of TET2-MT and TET2-WT cases. Volcano plots with the difference in DNA methylation between averages of TET2-MT versus TET2-WT on the x-axis, and the unadjusted \( P \) value for each site on the y-axis for sites in CGIs + NCGIs (A), sites in CGIs (B), and sites in NCGIs (C). Also shown is validation of DNA methylation levels by bisulphite-pyrosequencing for two tet2-DMCs in CGIs (D) and three tet2-DMCs in NCGIs (E) for TET2-MT and TET2-WT cases. DNA methylation levels of normal blood samples are shown in the light gray rectangle.

Decreased 5hmC at tet2-DMCs in TET2-MT cases

Because neither DREAM nor bisulphite-pyrosequencing can discriminate 5hmC from 5mC, we developed an assay for assessing enrichment of 5hmC at tet2-DMCs by 5hmC labeling followed by an affinity enrichment method. We quantified 5hmC amounts at 6 tet2-DMCs in 13 CML patients (4 TET2-MT and 9 TET2-WT cases) and found that 3 out 6 tet2-DMCs analyzed showed
significantly lower 5hmC enrichment in TET2-MT cases (Fig. 6A) while the remaining 3 tet2-DMCs also showed a trend toward lower 5hmC enrichment in TET2-MT cases. We calculated a z-score of 5hmC enrichment for each tet2-DMC and found that average of z-scores for 6 tet2-DMCs in TET2-MT cases were significantly lower than TET2-WT cases (median z-scores/0.73 vs. 0.40, P = 0.03; Fig. 6B). Hierarchical clustering analysis of 5hmC enrichment clearly separated subsets of TET2-MT cases and TET2-WT cases (Fig. 6C).

**Discussion**

Genome-wide screening for DNA methylation by DREAM for eight TET2-MT and twelve TET2-WT CMML cases revealed that the general tumor phenotype in DNA methylation, hypermethylation in CGIs and hypomethylation in NCGIs, is found in patients with both TET2-MT and TET2-WT. Unsupervised hierarchical clustering analysis revealed that TET2-MT and TET2-WT cases are not clearly separated, suggesting that the effects of TET2 mutations are relatively minor in the human CMML methylome. We moved on to further analyses to clarify the characteristics of the difference in DNA methylation between TET2-MT and TET2-WT cases. We found that 0.55% of all sites analyzed were differentially methylated with a high proportion of hypermethylation in TET2-MT, supporting previous findings of an increased amount of 5mC in mutants compared with wild-types (16).

Strikingly, tet2-DMCs are primarily at NCGI sites that had intermediate to high DNA methylation levels in normal blood. For these sites, TET2-WT cases showed hypomethylation, whereas TET2-MT cases showed identical to higher methylation levels compared with normal blood. This suggests that TET2 mutations block hypomethylation in NCGIs during tumorigenesis. Interestingly, although TET1 has a CXXC domain responsible for binding to unmethylated sites, and is enriched at CGIs (35, 36), TET2 lacks the CXXC domain (37).
This could explain why TET2 mutations primarily affect NCGI sites that are methylated in normal blood.

To gain insight into the functional relevance of TET2 mutations in leukemogenesis, we investigated the effects of tet2-DMC methylation on gene expression. Although we found relatively small numbers of genes with tet2-DMCs at promoter regions, we found good correlations between DNA methylation and gene expression. Importantly, these genes have promoters in NCGIs. We also found a substantial number of NCGI sites hypermethylated at nonpromoter regions in TET2-MT cases and found good correlations between DNA methylation and gene expression that are validated for three genes with tet2-DMCs in NCGIs at nonpromoter regions.

Interestingly, TET2-WT cases showed expression levels similar to at least two out of three genes analyzed, but had lower DNA methylation levels compared with TET2-MT cases, implying that there might be mechanisms different from DNA methylation that result in gene expression deregulation in CMMI.

This finding prompted us to analyze the link to transcription factor-binding sites (and enhancers), which are known to control gene expression in cis and trans from distal regions in order to achieve precise differentiation (34). We found that tet2-DMCs in NCGI case are significantly enriched in enhancer sites (marked by H3K4me1) with tet2-DMCs in NCGIs overall sites analyzed in NCGIs in several cell lines. Red bars, significant enrichment. D, enrichment for transcription factor-binding sites in GM12878, a lymphoblastoid cell line, with tet2-DMCs in NCGIs over all sites analyzed in NCGIs. Red bars, significant enrichment.

Figure 5.
Deregulation of gene expression for genes with tet2-DMCs and their enrichment at enhancer and transcription factor-binding sites. Shown are correlations between DNA methylation and gene expression for genes with tet2-DMCs at promoters (A) and nonpromoters (B). Red, TET2-MT; blue, TET2-WT; green, normal blood. The sample from the patient with IDH2/DNMT3A mutations is shown in purple. Gene expression levels are calculated as 40ΔC T to GAPDH. Linear regression curve is indicated with a black line. C, enrichment for enhancer sites (marked by H3K4me1) with tet2-DMCs in NCGIs overall sites analyzed in NCGIs in several cell lines. Red bars, significant enrichment. D, enrichment for transcription factor-binding sites in GM12878, a lymphoblastoid cell line, with tet2-DMCs in NCGIs over all sites analyzed in NCGIs. Red bars, significant enrichment.
important to prove that not only 5mC but also 5hmC is variable at tet2-DMCs. Therefore, we developed a 5hmC pull-down assay to address this question. We found that 5hmC amounts at tet2-DMCs are significantly lower in TET2-MT than TET2-WT cases, suggesting that TET2 mutations lead to enzymatically deficient TET2 function at tet2-DMCs. Other covalently-modified cytosines such as 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) are also possibly variable between TET2-MT and TET2-WT, and this has to be verified with more sensitive and specific assays to detect these further oxidized derivatives that are far lower in quantity.

The effect of TET2 mutations on DNA methylation has been controversial. There are three reports in CMML including our previous report, two of which supported hypermethylation phenotypes in TET2-MT CMML (16, 17), whereas one report showed a remarkable hypomethylation phenotype in TET2-MT CMML (12). The dominant hypermethylation phenotypes were also supported by three other reports in AML (13), diffuse large B-cell lymphoma (44), and normal elderly individuals (45). Among these, one only report utilized genome-wide profiling of DNA methylation like our method and showed that TET2 mutations were primarily associated with hypermethylation within CGI and CpG-rich promoters. Although our data where we found that tet2-DMCs was associated with NCGI sites are inconsistent with this result, it is possible that the effects of TET2 mutations could vary depending on differences in disease-origins such as myeloid and lymphoid cells. Similar studies with genome-wide profiling need to be performed to clarify this question.

Although the fact that we have used normal peripheral blood rather than sorted cells as a control is a drawback of our study, we were mostly interested in a case–case comparison of TET2-WT to TET2-MT CMML, which is unaffected by normal peripheral blood data. Moreover, we have previously shown that there are very few differences in DNA methylation between bone marrow and blood in MDS and CMML and relatively little variation in methylation in different subsets in blood from patients with AML (sorted CD34⁺ and CD3⁻/CD19⁻ cells; ref. 46).

In conclusion, our data suggest that TET2 mutations have a minor effect (<1%) on DNA methylation throughout the human genome, and preferentially result in hypermethylation at selected NCGI sites that are enriched at transcription factor-binding sites and enhancers. Our data are consistent with a model whereby transcription factors such as p300 recruit TET2 as part of their mechanism of gene regulation, and provide an explanation for the differentiation block seen in TET2-mutant hematopoietic cells.

Disclosure of Potential Conflicts of Interest
No potential conflicts of interest were disclosed.

Authors’ Contributions
Conception and design: J. Yamazaki, H.M. Kantarjian, L.A. Godley, J.-P.J. Issa
Development of methodology: J. Yamazaki, J. Jelinek, M. Cesaroni, L.A. Godley
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): J. Yamazaki, J. Jelinek, J. Madzo, F. Neumann, R. Taby, A. Vasanthakumar, T. Macrae, K.R. Ostler, L.A. Godley
Writing, review, and/or revision of the manuscript: J. Yamazaki, J. Jelinek, M. Cesaroni, F. Neumann, H.M. Kantarjian, L.A. Godley, J.-P.J. Issa
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): J. Yamazaki, F. Neumann, R. He, M.R. Estecio, J.-P.J. Issa
Study supervision: J. Yamazaki, J.-P.J. Issa

Acknowledgments
The authors thank the patients who have contributed to their understanding of these disorders and thank Jenna Al-Malawi for excellent editing of the article.
Grant Support

This work was supported by the NIH grants CA100632, CA121104, and CA146969 (J.-P.J. Issa) and CA129831 and CA129831-08S1 (J.A. Godley), and supported by a Stand Up to Cancer Dream Team Translational Research Grant, Grant Number SU2C-AACR-DT1019. Stand Up To Cancer is a program of the Entertainment Industry Foundation Funded through the American Association for Cancer Research. J.-P.J. Issa is an American Cancer Society Clinical Research Professor supported by a generous gift from the F. M. Kirby Foundation.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received March 16, 2014; revised February 23, 2015; accepted March 7, 2015; published OnlineFirst May 13, 2015.

References


TET2 Mutations Affect Non-CpG Island DNA Methylation at Enhancers and Transcription Factor–Binding Sites in Chronic Myelomonocytic Leukemia

Jumpei Yamazaki, Jaroslav Jelinek, Yue Lu, et al.

Cancer Res 2015;75:2833-2843. Published OnlineFirst May 13, 2015.