Genome-wide analysis of focal DNA hypermethylation in *IDH*-mutant AML samples

Elisabeth R. Wilson[1], Nichole M. Helton[1], Sharon E. Heath[1], Robert S. Fulton[3], Jacqueline E. Payton[2], John S. Welch[1], Matthew J. Walter[1], Peter Westervelt[1], John F. DiPersio[1], Daniel C. Link[1], Christopher A. Miller[1,3], Timothy J. Ley[1], and David H. Spencer[1,2,3]

[1] Department of Medicine, Division of Oncology, Section of Stem Cell Biology,

[2] Department of Pathology and Immunology,

[3] McDonnell Genome Institute,

Washington University, St. Louis, MO, USA

Running title: Methylation phenotype of *IDH*-mutant AML

The authors have no conflicts of interest relating to the work described in this manuscript.

* Corresponding author

David H. Spencer, MD, PhD

Assistant Professor

Division of Oncology

Department of Medicine

Washington University School of Medicine

Campus Box 8007

660 South Euclid Avenue

St. Louis, MO. 63110

314-273-0739

dspencer@wustl.edu

Abstract

Recurrent mutations in *IDH1* or *IDH2* in acute myeloid leukemia (AML) are associated with increased DNA methylation, but the genome-wide patterns of this hypermethylation phenotype have not been comprehensively studied in AML samples. We analyzed whole-genome bisulfite sequencing data from 15 primary AML samples with *IDH1* or *IDH2* mutations, which identified ~4,000 focal regions that were uniquely hypermethylated vs. normal CD34+ cells. These regions had modest, but significant, hypermethylation in AMLs with biallelic *TET2* mutations, and 5-hydroxymethylation levels that were dependent on functional TET2, indicating that hypermethylation in these regions is caused by inhibition of TET-mediated demethylation. Focal hypermethylation in *IDH*^{mut} AMLs occurred in regions with low methylation in normal CD34+ cells, implying that DNA methylation and demethylation are active at these loci. AML samples containing *IDH* and *DNMT3A*^{R882} mutations were significantly less hypermethylated, suggesting that methylation in these regions is mediated by DNMT3A. *IDH*^{mut}-specific hypermethylation was highly enriched for enhancers that form direct interactions with genes involved in normal hematopoiesis and AML, including *MYC* and *ETV6*. These results suggest that focal hypermethylation in *IDH*-mutant AML occurs by altering the balance between DNA methylation and demethylation, and that disruption of these pathways at enhancers may contribute to AML pathogenesis.

45 **Introduction**

46 DNA methylation changes in acute myeloid leukemia (AML) occur because of disruptions in the
47 balance between processes that add or remove 5-methyl groups to cytosines[1,2]. In both normal
48 and malignant hematopoietic cells, *de novo* DNA methylation is catalyzed primarily by the DNA
49 methyltransferase DNMT3A[3,4], which methylates unmethylated DNA substrates. Demethylation
50 occurs both passively after DNA synthesis in the absence of DNMT1-mediated propagation of
51 hemi-methylated DNA, and actively via hydroxylation of 5mC by the TET family of hydroxylases
52 with subsequent removal of modified cytosine residues. Alterations in these opposing forces
53 result in either increased or decreased DNA methylation in AML cells compared to normal
54 hematopoietic cells. These changes include diffuse hypomethylation across large genomic
55 regions, as well as focal hypermethylation in CpG islands (CGIs). We recently showed that CGI
56 hypermethylation in AML is mediated by DNMT3A and is present in nearly all AML subtypes[5]. In
57 addition to these changes, specific DNA methylation patterns correlate with recurrent AML
58 mutations that influence DNA methylation pathways. This includes the *DNMT3A*^{R882} mutation,
59 which impairs DNA methylation activity and results in a focal, canonical hypomethylation
60 phenotype[5].

61

62 Mutations in *IDH1* and *IDH2* are also associated with altered DNA methylation patterns[6,7] that
63 are thought to occur by disrupting active DNA demethylation. *IDH1* and *IDH2* encode metabolic
64 enzymes not normally involved in DNA methylation, but when mutated produce 2-
65 hydroxyglutarate (2HG)[8] that inhibits the TET family of enzymes[9] thereby reducing active
66 demethylation. Analysis of DNA methylation in primary AML samples using array-based
67 technologies and reduced-representation bisulfite sequencing (RBBS) has shown that DNA
68 methylation is increased in specific genomic regions in samples with *IDH* mutations[6,10]. While
69 the direct effects of these changes on gene regulation have been challenging to identify, the
70 contribution of *IDH* mutations to leukemogenesis has been established in mouse models.

71 Expression of either *IDH1*^{R132H} or *IDH2*^{R140Q} blocks normal hematopoietic differentiation,
72 promotes myeloproliferation[11–13], and can result in AML transformation in the presence of
73 cooperating mutations[13,14]. These studies demonstrate the contribution of mutations in *IDH1* and
74 *IDH2* to AML development, which may occur by disrupting the balance between DNA
75 methylation and demethylation.
76
77 Although previous studies using targeted DNA methylation approaches have established
78 general effects of *IDH1* and *IDH2* mutations on DNA[6,7,10,15], a genome-wide analysis of
79 methylation patterns in primary AML samples has not yet been described. It is therefore unclear
80 whether *IDH1* vs. *IDH2* mutations manifest unique, mutation-specific methylation phenotypes,
81 and whether these methylation changes are distinct from DNMT3A-mediated CGI
82 hypermethylation. In addition, although *IDH* mutations are thought to cause hypermethylation
83 via inhibition of TET enzymes, the extent of overlap in methylation phenotypes between AML
84 samples with mutations in *IDH1/IDH2* and *TET2* remains unclear. Here, we performed a
85 genome-wide analysis of DNA methylation in primary AML samples with recurrent mutations in
86 *IDH1* or *IDH2* using whole-genome bisulfite sequencing (WGBS). WGBS data from normal
87 hematopoietic cells and AML samples with *TET2* mutations and other mutational profiles were
88 included to define the methylation phenotypes specific to *IDH* mutations, and to determine
89 whether these patterns are present in AML samples with *TET2* mutations. We integrated these
90 data with epigenetic modifications and three-dimensional (3D) genome architecture from
91 primary AML samples to characterize the functional genomic elements that are affected upon
92 disruption of the balance between DNA methylation and demethylation in AML cells.
93
94 **Materials and Methods**
95 *Patient samples*

96 Primary AML samples and normal hematopoietic cells for epigenetic studies were obtained from
97 presentation AML and normal bone marrow aspirates, following informed consent using protocol
98 (201011766) approved by the Human Research Protection Office at Washington University as
99 described previously[5] (Table S1). All experiments with AML samples used total bone marrow
100 cells after estimating the leukemic purity.
101
102 *Whole genome bisulfite and oxidative bisulfite sequencing and data analysis*
103 Whole-genome bisulfite sequencing data for 38 samples were described previously[5]. Data for
104 13 additional samples were generated using 50ng of DNA with the Swift Accel-NGS Methyl-Seq
105 library preparation kit. Oxidative bisulfite sequencing libraries were prepared following treatment
106 of 200ng of DNA with the TrueMethyl oxBS module (Cambridge Epigenetix) prior to bisulfite
107 conversion and Swift library construction and sequencing on NovaSeq 6000 instruments (Table
108 S1). Data were aligned to the GRCh38 human reference and processed into methylated read
109 counts using biscuit[16] with default parameters. Differentially methylated regions (DMRs) were
110 identified between AML groups using read count data via DSS[17] and subsequent filtering to
111 retain regions with >10 CpGs and a difference in mean methylation of 0.2. Mutation-specific
112 DMRs were identified using hierarchical clustering where group level mean methylation values
113 at all identified AML DMRs where assessed for euclidean distances via 'dist' with default
114 parameters. Clustering topology was analyzed to identify DMRs where a outlier branch
115 represented the specified AML mutational group. Statistical analysis of differential methylation in
116 DMRs was performed on the sum of the methylated/unmethylated counts using a Fisher's exact
117 test. 5hmC values were obtained by subtracting the methylation ratios from OxBS data from
118 WGBS data at all CpGs with coverage of at least 10x.
119
120 *ChIP-seq for histone modifications*

121 ChIP-seq was performed using ChIPmentation[18] with the following antibodies: H3K27me3
122 (9733S), and H3K27ac (8173S) from Cell Signaling Technology, and H3K4me1 (ab1012) from
123 Abcam. Sequencing was performed on a NovaSeq 6000 (Illumina, San Diego, CA) to obtain
124 ~50 million 2x150 bp reads. Data were analyzed via adapter trimming with trimgalore and
125 alignment to GRCh38 using bwa mem[19]. Normalized coverage for visualization and analysis
126 used the deeptools "bamCoverage" tool[20], and peaks were called with MACS2[21]. Statistical
127 comparisons with DESeq2[22] used raw fragment counts at peak summits, and visualizations
128 were prepared with Gviz[23]. Superenhancer analysis was conducted using ROSE software[24,25]
129 with default parameters.
130
131 *RNA-seq analysis*
132 RNA-seq data from AML samples were obtained from the AML TCGA study[15]. TPM values were
133 obtained using kallisto[26] and gene counts were generated using the tximport Bioconductor
134 package[27] in R with the tx2gene option set to accomplish gene-level summarization. Previously
135 published RNA-seq data for normal CD34+ cells generated using the same procedures that
136 were used for the AML samples [28,29] were obtained as raw sequencing reads from the short-
137 read archive (GSE48846) and processed as described above.
138
139 *Hi-C data analysis*
140 Hi-C data were obtained from previous studies of 3D genome interactions in primary AML
141 samples[30] and normal hematopoietic stem/progenitors[31]. All libraries were generated using MboI
142 digestion prior to proximity ligation and data were analyzed using the juicer pipeline[32]. Loops
143 were identified with HICCUPs and were analyzed using bedtools[33] to identify overlap with genes
144 and putative enhancers. Visualizations used the GenomicInteractions and Gviz R packages[23].
145
146 **Results**

147 *Primary AML samples with IDH1 or IDH2 mutations are focally hypermethylated at regions with*
148 *low methylation in normal hematopoietic cells*
149 We performed WBGS using 15 primary bone marrow aspirate samples from AML patients with
150 canonical *IDH* mutations, including seven with *IDH1*$^{R132C/G}$, seven with *IDH2*^{R140Q}, and one with
151 an *IDH2*^{R172K} allele (referred to hereafter as *IDH*^{mut}). These data were analyzed with WGBS data
152 from 36 other primary AML samples representing eight mutational categories, including five with
153 biallelic loss-of-function mutations in *TET2*, and primary CD34+ cells from six healthy adult bone
154 marrow donors[5]. All AML samples were previously sequenced using whole genome and/or
155 whole exome sequencing[15,34], which confirmed that the mutations affecting DNA methylation
156 were present in the dominant leukemic clone at presentation (Figure 1A). Importantly, the *IDH*^{mut}
157 AML samples lacked mutations in *DNMT3A* and *TET2* to minimize effects of other mutations on
158 DNA methylation patterns. WGBS produced a mean of 12x coverage (range: 4-16) for at least
159 28 million CpGs in the human reference sequence across all samples. *IDH*^{mut} AML samples had
160 genome-wide methylation levels similar to all other samples (Figure 1B), but were more
161 methylated at CGIs compared with normal CD34+ cells and AMLs with *DNMT3A*^{R882} mutations,
162 which have diminished CGI hypermethylation (*IDH*^{mut} vs. CD34+ cells P<0.0001; *IDH*^{mut} vs.
163 *DNMT3A*^{R882} P<0.0001; Figure 1C). CGI methylation levels were similar between *IDH*^{mut} AML
164 samples and AMLs with wild type *IDH1* and *IDH2* (mean CGI methylation of *IDH*^{mut} cases =
165 0.35, mean of all other groups = 0.33; P= 0.1; Figure 1C), indicating that *IDH* mutations do not
166 result in an exaggerated CGI hypermethylation phenotype.
167
168 Differential methylation analysis was performed using established methods[17] for eight mutation-
169 defined AML groups vs. normal CD34+ cells. This identified a mean of 4,689 (1,114-8,386)
170 Differentially Methylated Regions (DMRs) across the eight categories (minimum methylation
171 difference of 0.2, FDR<0.05); of these, between 6% and 97% were hypermethylated in the AML
172 samples vs. CD34+ cells (see Figure 1D). AMLs with *DNMT3A*^{R882} had the most

173 hypomethylated DMRs (97% of 5,747 DMRs), whereas *IDH*^{mut} AML samples had the most
174 hypermethylated DMRs, with 99% of the 6,960 identified regions having at least 0.2 higher
175 mean methylation than CD34+ cells. *TET2*-mutant samples were also hypermethylated, but at
176 fewer loci, with 74% of the 2,991 identified DMRs having higher methylation. The fewest DMRs
177 were identified in samples with *IDH1* or *IDH2* mutations and *DNMT3A*^{R882}, which is consistent
178 with previous studies suggesting that AML samples with both mutations have abrogated
179 methylation phenotypes[10]. Analysis of DMRs identified in *IDH*^{mut} samples demonstrated that
180 29% were associated with promoters and 43% occurred in CGIs, which was similar to the
181 frequencies observed for commonly hypermethylated DMRs in other AML subtypes (Figure
182 S1A). Interestingly, although *IDH* mutations are thought to lead to hypermethylation by inhibiting
183 active demethylation, most *IDH*^{mut}-specific DMRs had low methylation in normal hematopoietic
184 cells. For example, 72% of the *IDH*^{mut} DMRs had a mean methylation <0.3 in both CD34+ cells
185 (Figure S1B) and more mature myeloid cell populations (Figure S1C), suggesting that DNA
186 methylation pathways must be active in these regions to achieve the level of hypermethylated
187 observed in *IDH*^{mut} AML samples.
188
189 *IDH*^{mut}*-specific methylation changes are distinct from AML-associated CGI hypermethylation*
190 *and are influenced by IDH mutation type*
191 Because AML patients with *IDH1* vs. *IDH2* mutations have distinct clinical and molecular
192 phenotypes[35–39] we next sought to identify methylation changes uniquely associated with *IDH1*
193 and/or *IDH2* mutations and to determine whether these patterns are distinct from AML-
194 associated CGI hypermethylation (Figure 2A). Thus, we performed hierarchical clustering of
195 mean methylation values for either *IDH1*^{mut} or *IDH2*^{mut} and *IDH*^{wt} AML samples at DMRs
196 identified above, and used the clustering topology to identify regions where the single outlier
197 branch represented the *IDH* mutant AML samples (Figure 2B, see Methods). Samples with
198 *TET2* and/or *DNMT3A*^{R882} mutations were excluded, since they may share methylation changes

199 with *IDH*^{mut} samples, or lack CGI hypermethylation[5,6], respectively. This approach identified
200 3,928 *IDH1*^{mut}-specific and 1,821 *IDH2*^{mut}-specific DMRs, of which 90% and 79% were
201 hypermethylated with respect to normal CD34+ cells, respectively (see Figure 2C-E). Consistent
202 with the analysis above, most DMRs displayed low methylation in normal cells, with 55% of
203 *IDH1*^{mut}-specific and 71% of *IDH2*^{mut}-specific loci having a methylation level <0.3 in CD34+ cells
204 (Figure 2C-E) and mature myeloid cells (Figure S1D). There was extensive overlap between the
205 *IDH* mutation-specific DMRs (94%, 5,403 of 5,749 total DMRs), and AML samples with either
206 mutation were hypermethylated at both DMR sets (Figure 2E). However, hierarchical clustering
207 demonstrated considerable variability in methylation levels between the *IDH1*^{mut} and *IDH2*^{mut}
208 samples (Figure 2F). This was most striking for the *IDH2*^{mut} samples, which included three
209 AMLs with lower methylation at the *IDH2*^{mut}-specific DMRs (Figure 2F-G). *IDH2*^{mut} AML samples
210 were less methylated at the combined set of *IDH1*^{mut}-specific and *IDH2*^{mut}-specific DMR loci
211 (*IDH2*^{mut}= 0.54 vs. *IDH1*^{mut}=0.70; p-value= 0.04), although they remained hypermethylated
212 relative to CD34+ cells (Figure 2G). This difference was not related to mutant *IDH* allele
213 abundance (all samples had VAFs >30%, Table S1), and did not correlate with other recurrent
214 mutations, including *NPM1c* (4 in *IDH1*^{mut} and 3 in *IDH2*^{mut} samples, Figure 2F).
215
216 To determine whether *IDH*^{mut}-specific hypermethylation is similar to the canonical CGI
217 hypermethylation seen in other AML samples, we analyzed the CpG content and genomic
218 features of the *IDH*^{mut}-specific DMRs. Interestingly, the CpG density and overlap with genomic
219 annotations was markedly different for both *IDH1*^{mut}-specific and *IDH2*^{mut}-specific DMRs
220 compared to regions with AML-associated CGI hypermethylation. For example, the combined
221 set of *IDH*^{mut}-specific DMRs displayed significantly lower CpG density (mean CpG density of
222 0.89 vs. 1.26; P<0.0001) and less overlap with annotated CGIs (23% vs. 54%) compared to
223 4,573 commonly hypermethylated regions identified in at least 2 other AML mutation categories
224 (Figure 2H-I). Promoter regions were also underrepresented among the *IDH*^{mut}-specific DMRs

225 (21% vs 31%; see Figure 2H), consistent with depletion of promoter-associated CGIs among
226 *IDH*^{mut}-specific DMRs, further suggesting that *IDH*-associated hypermethylation is distinct from
227 AML-associated CGI hypermethylation.
228
229 *Hypermethylation in TET2*^{mut} *AMLs overlaps with IDH*^{mut}*-specific hypermethylation, but does not*
230 *phenocopy the extent of methylation changes*
231 We next determined whether AML samples with biallelic loss-of-function mutations in *TET2*
232 shared similar genome-wide patterns of hypermethylation with *IDH*^{mut} AMLs. Initial comparison
233 of the *TET2*^{mut} AMLs vs. normal CD34+ cells yielded fewer DMRs and a lower proportion of
234 hypermethylated regions compared to *IDH*^{mut} samples (2,512 vs. 6,523 DMRs, and 78% vs 99%
235 hypermethylated regions, respectively), consistent with previous reports that inactivation of
236 *TET2* has a less profound effect on DNA methylation[6,10]. Identification of *TET2*^{mut}-specific DMRs
237 using the clustering approach described above (with *IDH*^{mut} and *DNMT3A*^{R882} AMLs excluded
238 from the analysis) produced only 51 *TET2*^{mut}-specific DMRs, confirming that these AMLs lack a
239 strong hypermethylation phenotype (Figure 3A). Although many DMRs were hypermethylated
240 relative to CD34+ cells (31 of 51), the fraction was significantly less than in either *IDH1*^{mut} or
241 *IDH2*^{mut} AMLs (60% vs 91% and 73%, respectively). *TET2*^{mut}-specific DMRs were also not
242 enriched for either CGIs or promoters, compared to the set of commonly hypermethylated
243 DMRs (17% of *TET2* DMRs vs. 54% generic DMRs; 11% of *TET2* DMRs vs. 31% generic
244 DMRs; see Figure S2A-B), suggesting these regions are unlikely to reflect CGI
245 hypermethylation.
246
247 To investigate the interaction between *IDH* mutations and *TET2*-mediated demethylation, we
248 compared *TET2*^{mut}-specific and *IDH*^{mut}-specific DMRs, and performed oxidative bisulfite
249 sequencing[40] to measure 5-hydroxymethylation (5hmC) in samples with and without biallelic
250 inactivating *TET2* mutations. This analysis showed that 82% (42 of 51) of the *TET2*^{mut}-specific

251 DMRs overlapped an *IDH*^{mut}-specific hypermethylated region (Figure 3B). Methylation in
252 *TET2*^{mut} AMLs at the 4008 *IDH*^{mut}-specific DMRs was also significantly increased compared to
253 CD34+ cells (mean methylation of 0.38 vs. 0.31; 61% of DMRs with increased methylation via
254 Fisher's exact test with $q<0.05$; Figure 3C-D). We next analyzed levels of 5hmC at *IDH*^{mut} DMRs
255 using paired oxidative and standard whole-genome bisulfite sequencing (oxWGBS and WGBS)
256 of primary AML samples with (N=3) and without (N=1) *TET2* mutations (Figure S2C).
257 Subtraction of oxWGBS from WGBS data for CpGs with >10x coverage demonstrated low
258 levels of 5hmC across the genomes of these samples (mean: 0.51-0.71% in *TET2*^{mut}, 1.35% in
259 *TET2*^{wt}; Figure S2D), and identifiable peaks at selected loci (Figure S2E). *IDH*^{mut} DMRs had
260 higher levels of 5hmC than in either constitutively methylated heterochromatin or regions with
261 CGI hypermethylation, and 5hmC levels were lower in all three of these regions in *TET2*^{mut} AML
262 samples (Figure 3E, Figure S2F), indicating that 5hmC in these regions is dependent on TET2
263 activity.
264
265 *DNA hypermethylation in IDH*^{mut} *AML cells requires DNMT3A*
266 To assess whether *de novo* DNA methylation by DNMT3A contributes to *IDH*^{mut}-associated
267 hypermethylation, we analyzed methylation levels at *IDH*^{mut}-specific DMRs in seven AML
268 samples with co-occurring *IDH1* (N=5) or *IDH2* (N=2) and *DNMT3A*^{R882} mutations, which have a
269 dominant negative phenotype and a more severe hypomethylation than other *DNMT3A*
270 mutations[4]. Interestingly, although *DNMT3A*^{R882}/*IDH*^{mut} AMLs were still hypermethylated at
271 *IDH*^{mut}-specific DMRs, the degree of hypermethylation was diminished, with 51% of these
272 regions having significantly lower DNA methylation levels than samples with *IDH* mutations
273 alone (Figures 4A-C, S3A). Similar findings were observed in 7 additional *DNMT3A*^{R882}/*IDH*^{mut}
274 AML samples using methylation array data from the TCGA AML study[15] (Figure S3B). To further
275 characterize the extent of this interaction at regions known to be methylated by DNMT3A, we
276 analyzed DNA methylation levels in *DNMT3A*^{R882}/*IDH*^{mut} AML samples at hypomethylated DMRs

277 in AMLs with the *DNMT3A*^{R882} allele[5]. Surprisingly, these regions remained nearly fully
278 methylated in the *DNMT3A*^{R882}/*IDH*^{mut} double mutant samples, with 93% of the regions having
279 significantly higher methylation than AMLs with *DNMT3A*^{R882} alone (Figures 4D-F, Figure S3C).
280 Similar findings were observed in the array-based methylation data from the TCGA AML study[15]
281 (Figure S3D), further implying that DNMT3A-mediated methylation and TET-mediated
282 demethylation occur at the same places in the genome.
283
284 *IDH*^{mut}*-specific hypermethylated DMRs are enriched for enhancers*
285 We next asked whether these *IDH*^{mut}-specific DMRs were associated with certain chromatin
286 states. Annotation of these DMRs with chromatin states in CD34+ cells[41] demonstrated that
287 44% occurred the enhancer chromatin state, which was a 2-fold enrichment over regions
288 commonly hypermethylated (Figure 5A). Similar analysis performed on DMRs identified in other
289 AML subtypes showed a different distribution of chromatin states (Figure S4A-C). We further
290 defined this association using ChIP-seq peaks for enhancer-associated H3K27ac and H3K4me1
291 modifications from 16 primary AML samples, including 14 *IDH*^{wt} and 2 *IDH*^{mut}. Analysis defined
292 44,762 and 6,917 consensus peaks for H3K27ac and H3K4me1, respectively, which along with
293 H3K27me3 ChIP-seq data from 9 AML patients were used to identify active, poised, and weak
294 enhancer loci. Analysis of the *IDH*^{mut} DMRs showed that 47% overlapped an active enhancer,
295 compared to 3% and 1% that overlapped poised and weak regions, respectively (Figures 5B-D).
296 In comparison, commonly hypermethylated regions showed less overlap with active enhancers
297 (13% of DMRs) and greater intersection with repressive H3K27me3 marks (Figure 5D). We
298 analyzed *IDH*^{mut}-specific DMRs for transcription factor (TF) binding motifs, which identified
299 binding sites for hematopoietic-associated TFs, including *SPI1*, *RUNX1*, and *MYC* (Figure 5E),
300 further supporting the occurrence of *IDH*^{mut}-specific hypermethylation at regions with potential
301 regulatory activity. However, quantitative analysis of H3K27ac signal over these regions in
302 samples with and without *IDH* mutations did not identify appreciable differences in wild type vs.

303 mutant samples (p-value=0.24, Figure 5F), suggesting that hypermethylation does not modify
304 H3K27ac levels within these regions.
305
306 *IDH*^{mut}*-specific DMRs occur in enhancers that form direct interactions with highly expressed*
307 *genes in AML cells*
308 We next asked whether enhancers with *IDH*^{mut} DMRs could be involved in controlling the
309 expression of genes relevant for AML pathogenesis. To directly link these enhancers to their
310 target genes, we analyzed three-dimensional (3D) genome interactions generated using *in situ*
311 HiC from both normal CD34+ cells[31] and 3 primary AML samples[30] (all wild-type for *DNMT3A*,
312 *IDH1*, *IDH2*, and *TET2*). This analysis demonstrated that 25% (1047/4008) of all *IDH*^{mut}-specific
313 DMRs and 32% (322/1021) of the DMRs in putative enhancers overlapped the 'loop anchor' of a
314 genome interaction (Figure 6A, Figure S5A). *IDH*^{mut} DMRs in these loop anchors were highly
315 enriched in 'superenhancers', with between 37 and 39% of superenhancers defined in 3 primary
316 *IDH*^{mut} AML samples containing at least one *IDH*^{mut}-specific DMR (Figure 6B-C, Figure S5B-C).
317 We next analyzed gene expression in 750 genes with promoters that formed 3D interactions
318 with *IDH*^{mut}-specific DMRs using RNA-seq data from 179 AML samples from the TCGA AML
319 study. This showed that the genes linked to *IDH*^{mut}-specific DMRs were highly expressed, with
320 68% of these genes ranked in the top 25th percentile of gene expression (Figure 6D, Figure
321 S5D). Further analysis of 3D genome interactions containing *IDH*^{mut}-specific DMRs identified
322 known and novel enhancers of genes important in hematopoiesis and AML, including an
323 enhancer of *MYC*[42–44] (Figure 6E), and previously unreported putative enhancers that form
324 interactions with *ETV6* (Figure 6F), *DOT1L,* and *SRSF3* (Figure S5E-F). Although we did not
325 observe significant changes in expression of these genes between *IDH*^{mut} and *IDH*^{wt} AMLs, their
326 high expression in AML samples and CD34+ cells was consistent with the enrichment of *IDH*^{mut}-
327 specific DMRs in enhancers of active genes (Figure 6D-F).
328

329 **Discussion**

330 Recurrent gain-of-function *IDH1* and *IDH2* mutations are known to increase DNA methylation,
331 though the extent and functional consequences of these changes have not been clearly defined.
332 We used whole-genome bisulfite sequencing of primary AML samples to demonstrate that the
333 effects of *IDH* mutations on methylation do not manifest as diffuse changes across the genome,
334 but instead occur in thousands of focal regions that are uniquely hypermethylated compared to
335 both normal CD34+ cells and other AML samples. These regions had lower CpG density and
336 fewer CGIs than loci that are commonly hypermethylated in AML, suggesting that mechanisms
337 of IDH^{mut}-associated methylation changes are distinct from 'traditional' CGI hypermethylation
338 associated with AML. The *IDH2*-mutant AML samples in our dataset had less pronounced
339 hypermethylation than those with *IDH1* mutations, but both were hypermethylated at a highly
340 overlapping set of loci. AMLs with biallelic inactivating mutations in *TET2* had a far less dramatic
341 methylation phenotype, though many IDH^{mut}-specific DMRs were significantly hypermethylated
342 in $TET2^{mut}$ AML samples. Further, oxidative bisulfite sequencing demonstrated increased levels
343 of 5hmC in these regions, which was absent in the $TET2^{mut}$ AML samples, providing additional
344 evidence in primary AML samples that both *IDH* and *TET2* mutations cause increased DNA
345 methylation by impairing the TET-mediated DNA demethylation pathway. Regions with IDH^{mut}-
346 specific hypermethylation were enriched for active enhancers, many of which formed direct
347 interactions with genes that are highly expressed in AML cells, including regulatory sequences
348 that interact with the promoters of *MYC* and *ETV6*. Although the increased methylation at these
349 loci was not associated with repressed chromatin or lower gene expression in IDH^{mut} AML
350 samples, this finding demonstrates that IDH^{mut}-associated hypermethylation affects the
351 regulatory sequences of genes that may contribute to AML pathogenesis.

352

353 This study adds new context to the dynamics of *de novo* DNA methylation and active
354 demethylation pathways in normal hematopoietic cells and in AML. The fact that IDH^{mut}-

355 associated hypermethylation occurs at regions with low levels of DNA methylation in normal
356 CD34+ cells means that *de novo* DNA methylation and TET-mediated demethylation can both
357 be active in these regions, despite their low steady-state methylation levels. This is supported
358 by the observation that AML samples with co-occurring *IDH* and *DNMT3A*^{R882} mutations show
359 significantly attenuated hypermethylation, and that there are appreciable levels of 5hmC in
360 *IDH*^{mut}-specific DMRs, which is produced from 5mC as a substrate. Remodeling of DNA
361 methylation by these processes in specific regions has been reported previously in studies of
362 embryonic stem cells, which have shown that methylation and active demethylation are in
363 equilibrium at many loci[1,2], and may be maintained by the occupancy of methylation and
364 demethylation complexes[45]. Our analysis suggests this equilibrium exists in normal CD34+ cells
365 and is disrupted in the presence of mutant *IDH* alleles, leaving *de novo* DNA methylation
366 unopposed. The focal nature of *IDH*^{mut}-associated hypermethylation implies that activity (or
367 occupancy) of DNMT3A and TET enzymes may be 'concentrated' in specific genomic regions.
368 The genomic or epigenetic features directing this activity remain unclear[46], but the enrichment of
369 *IDH*^{mut} DMRs in active enhancers suggests that components of active chromatin may recruit
370 methylation and demethylation machinery. The convergence of these processes at enhancers
371 could provide clues as to why mutations with opposite effects on DNA methylation both
372 contribute to AML development via dysregulation of common target genes.
373
374 Our analysis of 3D genome interactions involving *IDH*^{mut}-specific DMRs found that these
375 sequences directly interact with genes that are highly expressed in hematopoiesis and AML
376 (e.g., *MYC* and *ETV6*). Contrary to the canonical relationship between DNA methylation and
377 activity, hypermethylation in the *IDH*^{mut} AML samples does not appear to repress either the
378 enhancer elements or expression of their target genes. Other regulatory factors may therefore
379 be dominant to DNA methylation at these loci, and result in persistently high gene expression. It
380 could also be that regions of active chromatin, such as enhancers (and superenhancers), have

381 high rates of methylation turnover and are therefore susceptible to perturbations in methylation
382 and demethylation[1,2]. Focal hypermethylation may occur in DNA elements bound by factors that
383 contribute to 'fine-tuning' of these enhancers in specific cellular or developmental contexts, but
384 that do not drive activity in AML cells. Additional studies will be necessary to understand
385 whether hypermethylation of these enhancer elements is a consequence of decreased
386 occupancy of these modulating factors[47], or whether it directly prevents proper regulation in
387 ways that contribute to AML development.
388

Acknowledgements

This work was supported by the National Cancer Institute (K08CA190815) and The Cancer Research Foundation Young Investigator Award to D.H.S. Support for procurement, annotation, and sequencing of human samples was provided by the Genomics of AML Program Project (P01CA101937, to Dr. Ley) and the Specialized Program of Research Excellence in AML (P50CA171963, to Dr. Link) from the NCI. Sequencing was performed by the McDonnell Genome Institute at Washington University in St. Louis.

396 **References**

397 1. Charlton, J. *et al.* TETs compete with DNMT3 activity in pluripotent cells at thousands of
398 methylated somatic enhancers. *Nat Genet* **52**, 819–827 (2020).

399 2. Ginno, P. A. *et al.* A genome-scale map of DNA methylation turnover identifies site-specific
400 dependencies of DNMT and TET activity. *Nat Commun* **11**, 2680 (2020).

401 3. Challen, G. A. *et al.* Dnmt3a is essential for hematopoietic stem cell differentiation. *Nat Genet*
402 **44**, 23 (2012).

403 4. Russler-Germain, D. A. *et al.* The R882H DNMT3A Mutation Associated with AML
404 Dominantly Inhibits Wild-Type DNMT3A by Blocking Its Ability to Form Active Tetramers.
405 *Cancer Cell* **25**, 442–454 (2014).

406 5. Spencer, D. H. *et al.* CpG Island Hypermethylation Mediated by DNMT3A Is a Consequence
407 of AML Progression. *Cell* **168**, 801-816.e13 (2017).

408 6. Figueroa, M. E. *et al.* Leukemic IDH1 and IDH2 Mutations Result in a Hypermethylation
409 Phenotype, Disrupt TET2 Function, and Impair Hematopoietic Differentiation. *Cancer Cell* **18**,
410 553–567 (2010).

411 7. Akalin, A. *et al.* Base-Pair Resolution DNA Methylation Sequencing Reveals Profoundly
412 Divergent Epigenetic Landscapes in Acute Myeloid Leukemia. *Plos Genet* **8**, e1002781 (2012).

413 8. Losman, J.-A. *et al.* (*R*)-2-Hydroxyglutarate Is Sufficient to Promote Leukemogenesis and Its
414 Effects Are Reversible. *Science* **339**, 1621–1625 (2013).

415 9. Xu, W. *et al.* Oncometabolite 2-Hydroxyglutarate Is a Competitive Inhibitor of α-Ketoglutarate-
416 Dependent Dioxygenases. *Cancer Cell* **19**, 17–30 (2011).

417 10. Glass, J. L. *et al.* Epigenetic Identity in AML Depends on Disruption of Nonpromoter
418 Regulatory Elements and Is Affected by Antagonistic Effects of Mutations in Epigenetic
419 Modifiers. *Cancer Discov* **7**, 868–883 (2017).

420 11. Moran-Crusio, K. *et al.* Tet2 Loss Leads to Increased Hematopoietic Stem Cell Self-
421 Renewal and Myeloid Transformation. *Cancer Cell* **20**, 11–24 (2011).

422 12. Sasaki, M. *et al.* IDH1(R132H) mutation increases murine haematopoietic progenitors and
423 alters epigenetics. *Nature* **488**, 656 (2012).

424 13. Kats, L. M. *et al.* Proto-Oncogenic Role of Mutant IDH2 in Leukemia Initiation and
425 Maintenance. *Cell Stem Cell* **14**, 329–341 (2014).

426 14. Yoshimi, A. *et al.* Coordinated alterations in RNA splicing and epigenetic regulation drive
427 leukaemogenesis. *Nature* **574**, 273–277 (2019).

428 15. Ley, T. J. *et al.* Genomic and Epigenomic Landscapes of Adult De Novo Acute Myeloid
429 Leukemia. *New Engl J Medicine* **368**, 2059–2074 (2013).

430 16. GitHub - huishenlab/biscuit: BISulfite-seq CUI Toolkit. https://github.com/huishenlab/biscuit
431 (n.d.).

432 17. Feng, H., Conneely, K. N. & Wu, H. A Bayesian hierarchical model to detect differentially
433 methylated loci from single nucleotide resolution sequencing data. *Nucleic Acids Res* **42**, e69–
434 e69 (2014).

435 18. Schmidl, C., Rendeiro, A. F., Sheffield, N. C. & Bock, C. ChIPmentation: fast, robust, low-
436 input ChIP-seq for histones and transcription factors. *Nat Methods* **12**, 963–965 (2015).

437 19. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM.
438 (2013).

439 20. Ramírez, F. *et al.* deepTools2: a next generation web server for deep-sequencing data
440 analysis. *Nucleic Acids Res* **44**, W160–W165 (2016).

441 21. Zhang, Y. *et al.* Model-based Analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137 (2008).

442 22. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for
443 RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).

444 23. Hahne, F. & Ivanek, R. Methods in Molecular Biology. *Methods Mol Biology Clifton N J*
445 **1418**, 335–351 (2016).

446 24. Lovén, J. *et al.* Selective Inhibition of Tumor Oncogenes by Disruption of Super-Enhancers.
447 *Cell* **153**, 320–334 (2013).

448 25. Whyte, W. A. *et al.* Master Transcription Factors and Mediator Establish Super-Enhancers
449 at Key Cell Identity Genes. *Cell* **153**, 307–319 (2013).

450 26. Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq
451 quantification. *Nat Biotechnol* **34**, 525–527 (2016).

452 27. Soneson, C., Love, M. I. & Robinson, M. D. Differential analyses for RNA-seq: transcript-
453 level estimates improve gene-level inferences. *F1000research* **4**, 1521 (2016).

454 28. Pabst, C. *et al.* GPR56 identifies primary human acute myeloid leukemia cells with high
455 repopulating potential in vivo. *Blood* **127**, 2018–2027 (2016).

456 29. MacRae, T. *et al.* RNA-Seq Reveals Spliceosome and Proteasome Genes as Most
457 Consistent Transcripts in Human Cancer Cells. *Plos One* **8**, e72884 (2013).

458 30. Ghasemi, R., Struthers, H., Wilson, E. R. & Spencer, D. H. Contribution of CTCF binding to
459 transcriptional activity at the HOXA locus in NPM1-mutant AML cells. *Leukemia* 1–13 (2020)
460 doi:10.1038/s41375-020-0856-3.

461 31. Zhang, X. *et al.* Large DNA Methylation Nadirs Anchor Chromatin Loops Maintaining
462 Hematopoietic Stem Cell Identity. *Mol Cell* **78**, 506-521.e6 (2020).

463 32. Durand, N. C. *et al.* Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C
464 Experiments. *Cell Syst* **3**, 95–98 (2016).

465 33. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic
466 features. *Bioinformatics* **26**, 841–842 (2010).

467 34. Klco, J. M. *et al.* Association Between Mutation Clearance After Induction Therapy and
468 Outcomes in Acute Myeloid Leukemia. *Jama* **314**, 811–822 (2015).

469 35. Paschka, P. *et al.* IDH1 and IDH2 Mutations Are Frequent Genetic Alterations in Acute
470 Myeloid Leukemia and Confer Adverse Prognosis in Cytogenetically Normal Acute Myeloid
471 Leukemia With NPM1 Mutation Without FLT3 Internal Tandem Duplication. *J Clin Oncol* **28**,
472 3636–3643 (2010).

473 36. Marcucci, G. *et al.* IDH1 and IDH2 Gene Mutations Identify Novel Molecular Subsets Within
474 De Novo Cytogenetically Normal Acute Myeloid Leukemia: A Cancer and Leukemia Group B
475 Study. *J Clin Oncol* **28**, 2348–2355 (2010).

476 37. Tefferi, A. *et al.* IDH1 and IDH2 mutation studies in 1473 patients with chronic-, fibrotic- or
477 blast-phase essential thrombocythemia, polycythemia vera or myelofibrosis. *Leukemia* **24**,
478 1302–1309 (2010).

479 38. Patel, K. P. *et al.* Acute Myeloid Leukemia With IDH1 or IDH2 MutationFrequency and
480 Clinicopathologic Features. *Am J Clin Pathol* **135**, 35–45 (2011).

481 39. Kattih, B. *et al.* IDH1/2 mutations in acute myeloid leukemia patients and risk of coronary
482 artery disease and cardiac dysfunction—a retrospective propensity score analysis. *Leukemia* 1–
483 16 (2020) doi:10.1038/s41375-020-01043-x.

484 40. Kernaleguen, M. *et al.* Epigenome Editing, Methods and Protocols. *Methods Mol Biology*
485 **1767**, 311–349 (2018).

486 41. Ernst, J. & Kellis, M. Chromatin-state discovery and genome annotation with ChromHMM.
487 *Nat Protoc* **12**, 2478–2492 (2017).

488 42. Shi, J. *et al.* Role of SWI/SNF in acute leukemia maintenance and enhancer-mediated Myc
489 regulation. *Gene Dev* **27**, 2648–2662 (2013).

490 43. Pulikkan, J. A. *et al.* CBFβ-SMMHC Inhibition Triggers Apoptosis by Disrupting MYC
491 Chromatin Dynamics in Acute Myeloid Leukemia. *Cell* **174**, 172-186.e21 (2018).

492 44. Bahr, C. *et al.* A Myc enhancer cluster regulates normal and leukaemic haematopoietic stem
493 cell hierarchies. *Nature* **553**, 515–520 (2018).

494 45. Dixon, G. *et al.* QSER1 protects DNA methylation valleys from de novo methylation. *Science*
495 **372**, eabd0875 (2021).

496 46. Weinberg, D. N. *et al.* The histone mark H3K36me2 recruits DNMT3A and shapes the
497 intergenic DNA methylation landscape. *Nature* **573**, 281–286 (2019).

498 47. Thurman, R. E. *et al.* The accessible chromatin landscape of the human genome. *Nature*
499 **489**, 75 (2012).

500

501

502

503 **Figure 1.** Genome-wide DNA methylation patterns in seven AML mutation groups and normal
504 CD34+ cells. A) Summary of mutation status in 51 primary AML patients with WGBS data. B)
505 Average methylation levels across ~28 million CpGs in CD34+ cells (N=6) and AML subtypes
506 (*IDH1*^{mut} or *IDH2*^{mut}, n=15; *TET2*^{mut}, n=5; *DNMT3A*^{R882}, n=6; *DNMT3A*^{R882}/*IDH*^{mut}, n=7; normal
507 karyotype with *NPM1c* and wild-type *IDH1*, *IDH2*, *TET2*, and *DNMT3A,* n=4; Normal karyotype
508 with wild-type *NPM1*, *IDH1*, *IDH2*, *TET2*, and *DNMT3A*, n=4; *CBFB-MYH11*, n=3; *KMT2A-ELL*,
509 n=3; *RUNX1-RUNX1T1*, n=3). C) Mean CpG island methylation levels in CD34+ cells and AML
510 subtypes. D) Number of differentially methylated regions (DMRs) identified for each AML
511 subtype compared with normal CD34+ cells. Blue and red bars represent hypomethylated and
512 hypermethylated DMRs with respect to normal CD34+ cells, respectively. Mean number of
513 CpGs per DMR (top panel) and DMR length (bottom panel) are shown for each AML subtype.
514 E) Mean CpG density across DMRs identified in each AML subtype.
515
516

517 **Figure 2.** Characterization of *IDH*^{mut}-specific DMRs. A) Distribution of shared hypermethylated
518 DMRs across AML subtypes. B) Example of the hierarchical clustering approach used to identify
519 *IDH*^{mut}-specific DMRs. Aggregate DMR methylation for individual AML subtypes are shown in
520 bottom panel with *IDH* samples indicated in red and CD34+ cells indicated in blue. C) Mean
521 DMR methylation across *IDH*^{mut}-specific DMRs in *IDH*^{mut} samples versus mean methylation of
522 all other AMLs. D) Aggregate methylation levels across all *IDH1/2*^{mut}-specific DMRs in *IDH*^{mut}
523 and *IDH*^{wt} AML subtypes. E) Locus heatmap showing mean methylation values by group for all
524 *IDH*^{mut}-specific DMRs, where each column is centered over the DMR with the window extending
525 5kb up- or down-stream the DMR center. F) Mean DMR methylation across all *IDH1/2*^{mut}-
526 specific DMRs (rows) in 15 individual *IDH*^{mut} cases (columns). G) Aggregate DMR methylation
527 across 3,928 *IDH1*^{mut} specific DMRs and 1,821 *IDH2*^{mut} specific DMRs respectively. H) Fraction
528 of functional genomic elements overlapping generically hypermethylated DMRs, *IDH1*^{mut} specific
529 DMRs, and *IDH2*^{mut} specific DMRs. I) Distribution of CpG densities across generically
530 hypermethylated regions in primary AML and *IDH1*^{mut}- and *IDH2*^{mut}-specific DMRs.
531
532

533 **Figure 3.** *TET2*^{mut} AMLs have modest hypermethylation that overlaps *IDH*^{mut}-specific DMRs. A)
534 Mean DMR methylation across 2512 in *TET2*^{mut} DMRs called vs. CD34+ cells (black points) and
535 52 in *TET2*^{mut}-specific DMRs (red points) in *TET2* mutant samples versus mean methylation in
536 all other AMLs. B) Intersection of in *TET2*^{mut}-specific and *IDH*^{mut}-specific DMRs. C) Aggregate
537 methylation over *IDH*^{mut}-specific DMRs in *IDH*^{mut} and in *TET2*^{mut} AML and CD34+ cells. D)
538 Locus heatmap of mean methylation values for all *IDH*^{mut}-specific DMRs (rows), where each
539 column is centered over the DMR with the window extending 5kb up- and down-stream the
540 DMR center point. E) Mean 5hmC (WGBS minus oxWGBS) levels in in *TET2*^{mut} and in *TET2*^{wt}
541 AML samples at 4650 ChromHMM heterochromatic regions, 4586 generically hypermethylated
542 regions, and 4008 *IDH*^{mut}-specific hypermethylated DMRs.
543
544

545 **Figure 4.** *DNMT3A*^{R882}/*IDH*^{mut} double mutant AMLs display an attenuated focal
546 hypermethylation phenotype. A) Locus heatmap of mean methylation at *IDH*^{mut} DMRs (rows) in
547 *IDH1* or *IDH2* mutant, *DNMT3A*^{R882}/*IDH*^{mut} double mutant, and *DNMT3A*^{R882} AMLs, and CD34+
548 cells. B) Distribution of *IDH*^{mut}-specific DMR methylation levels by AML subtype. C) Example
549 *IDH*^{mut}-specific DMR locus within the *ETV6* gene demonstrating an intermediate methylation
550 phenotype of double mutant samples with respect to *IDH*^{mut} and *DNMT3A*^{R882} mutant AMLs. D)
551 Methylation locus heatmap of average subtype methylation across *DNMT3A*^{R882} DMRs called
552 vs. CD34+ cells in *IDH*^{mut}, *DNMT3A*^{R882}/*IDH*^{mut} double mutant, and *DNMT3A*^{R882} AMLs, and
553 CD34+ cells. E) Distribution of *DNMT3A*^{R882} DMR methylation levels by AML subtype. F)
554 Example *DNMT3A*^{R882} DMR locus within the *MLLT1* gene, demonstrating the hypomethylation
555 phenotype of *DNMT3A*^{R882} mutant samples with respect to *IDH*^{mut} and *DNMT3A*^{R882}/*IDH*^{mut}
556 double mutant AML samples.
557
558

559 **Figure 5.** *IDH*^{mut}-specific DMRs are enriched for putative enhancers. A) Distribution of
560 ChromHMM chromatin states from CD34+ cells represented in *IDH*^{mut}-specific DMRs.
561 Enrichment of chromatin states within *IDH*^{mut}-specific DMRs is shown with respect to the
562 frequency of states overlapping regions of common CpG island hypermethylation. B) Enhancer-
563 based annotation of common hypermethylated regions, *IDH1*^{mut}, and *IDH2*^{mut} DMRs, where
564 DMRs intersecting an H3K27ac peak alone or in combination with H3K4me1 constitute active
565 enhancers, an H3K27ac peak in combination with H3K27me3 constitutes poised enhancers,
566 and H3K4me1 alone constitutes a weak enhancer profile. C) Examples of intragenic and genic
567 enhancer regions exhibiting *IDH1*^{mut}, *IDH2*^{mut}, or *IDH1/2*^{mut} hypermethylation compared with
568 CD34+ cells and other AML subtypes. D) Heatmap of enhancer histone modifications and
569 heterochromatin modifications over *IDH*^{mut}-specific DMRs (left) and generic hypermethylation
570 (right) in CD34+ cells (N=4 H3K27ac, N=7 H3K3me1, and N=7 H3K27me3), *IDH*^{mut} AML (n=3),
571 and *IDH*^{wt} AML samples (N=9 H3K27ac, N=10 H3K3me1, and N=24 H3K27me3). E)
572 Differential active enhancer signal (H3K27ac) for all AML-associated putative enhancers (black
573 points) compared to putative enhancers intersecting an *IDH1/2*^{mut}-specific DMR (red points). F)
574 HOMER motif enrichment analysis of *IDH1/2*^{mut}-specific DMRs with respect to a background set
575 of generically hypermethylated regions. G) Enrichment analysis of TF binding events for 445
576 TFs within *IDH1/2*^{mut}-specific DMRs.
577
578

579 **Figure 6**. *IDH*^{mut}-specific DMRs are enriched in superenhancers and interact with highly
580 expressed genes in AML. A) Schematic of DMR and enhancer-associated DMRs (eDMR) and
581 their interaction with target genes based on intersection HiC-defined genome loops. B) Rank
582 ordered enhancer regions based on H3K27ac signal in a representative *IDH*^{mut} AML sample,
583 annotated by presence of overlapping *IDH*^{mut}-specific DMRs (absence of DMRs indicated by
584 green points, greater than one DMR indicated by orange points) and computationally-defined
585 'superenhancer' (above red line). Enhancers of specific hematopoietic genes are designated. C)
586 Distribution of number of *IDH*^{mut}-specific DMRs overlapping a set of AML consensus
587 superenhancers (SEs) from H3K27ac data from 4 primary samples (N=779). D) Distribution of
588 normalized gene expression values for all expressed genes (orange histogram) and a set of 750
589 eDMR target genes (blue histogram) in *IDH*^{mut} AML samples. E) Example *IDH*^{mut} -eDMR locus
590 displaying interactions with the *MYC* promoter. A zoomed in view of the locus demonstrates
591 focal enhancer hypermethylation in *IDH1*^{mut} (purple) and *IDH2*^{mut} (green) samples compared
592 with CD34+ cells (blue). Normalized *MYC* expression is shown for 17 CD34+ cord blood cell
593 samples, 6 and 14 *IDH1*^{mut} and *IDH2*^{mut} samples, and 91 *IDH*^{wt} samples. F) Example *IDH*^{mut}-
594 DMR locus in a candidate enhancer that displays robust interactions with the *ETV6* promoter. A
595 zoomed in locus view demonstrates focal enhancer hypermethylation in *IDH1*^{mut} (purple) and
596 *IDH2*^{mut} (green) samples compared with CD34+ cells (blue). Normalized *ETV6* expression is
597 shown for CD34+ cells, *IDH1*^{mut} and *IDH2*^{mut} samples, and *IDH*^{wt} samples (see E for sample
598 sizes).

Figure 1.

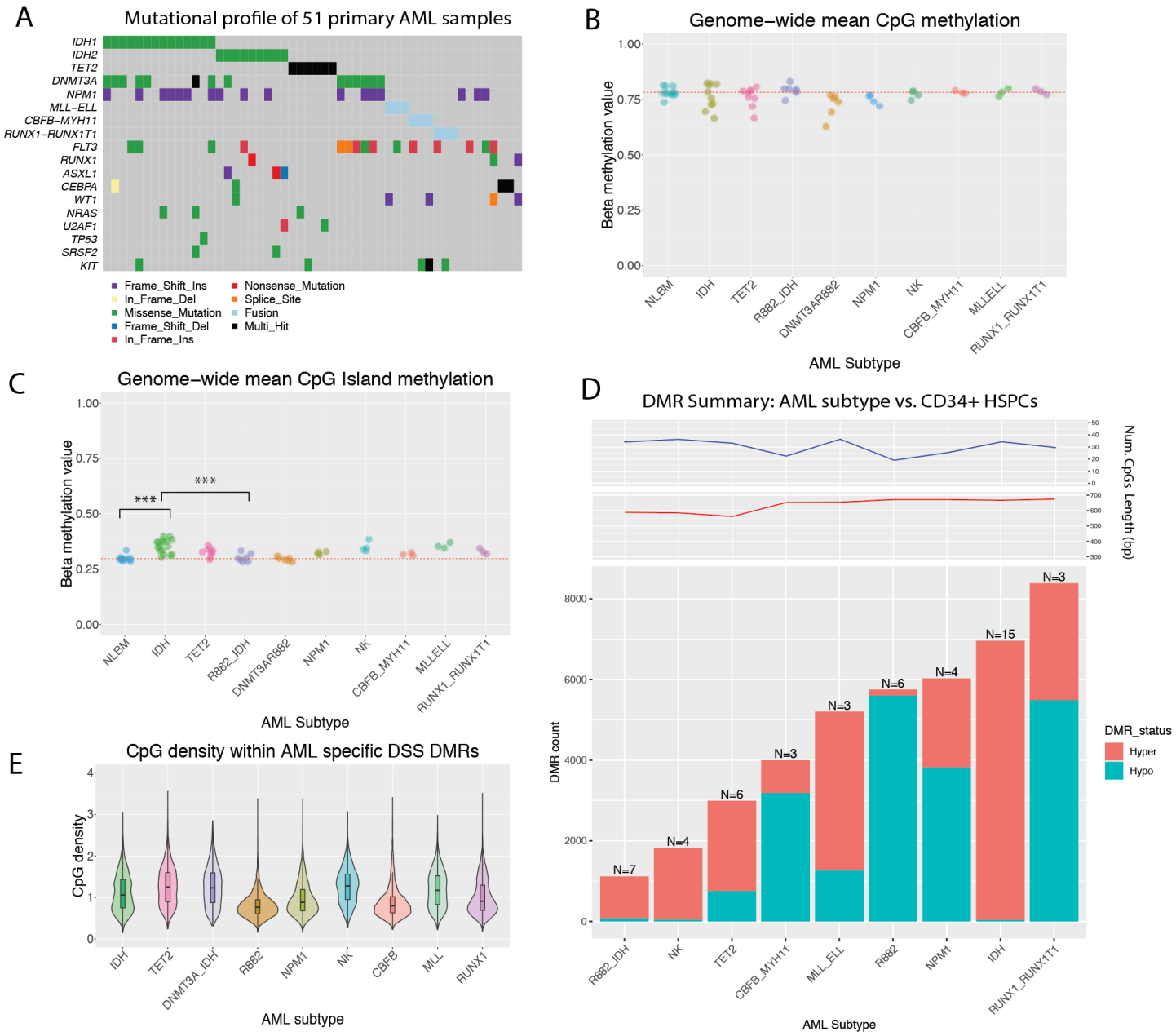

A
Mutational profile of 51 primary AML samples
IDH1
IDH2
TET2
DNMT3A
NPM1
MLL–ELL
CBFB–MYH11
RUNX1–RUNX1T1
FLT3
RUNX1
ASXL1
CEBPA
WT1
NRAS
U2AF1
TP53
SRSF2
KIT
Frame_Shift_Ins
In_Frame_Del
Missense_Mutation
Frame_Shift_Del
In_Frame_Ins
Nonsense_Mutation
Splice_Site
Fusion
Multi_Hit
B
Genome–wide mean CpG methylation
Beta methylation value
AML Subtype
NLBM
IDH
TET2
R882_IDH
DNMT3AR882
NPM1
NK
CBFB_MYH11
MLLELL
RUNX1_RUNX1T1
C
Genome–wide mean CpG Island methylation
Beta methylation value
***
***
AML Subtype
D
DMR Summary: AML subtype vs. CD34+ HSPCs
Num. CpGs Length (bp)
DMR count
N=7
N=4
N=6
N=3
N=3
N=6
N=4
N=15
N=3
DMR_status
Hyper
Hypo
R882_IDH
NK
TET2
CBFB_MYH11
MLL_ELL
R882
NPM1
IDH
RUNX1_RUNX1T1
AML Subtype
E
CpG density within AML specific DSS DMRs
CpG density
IDH
TET2
DNMT3A_IDH
R882
NPM1
NK
CBFB
MLL
RUNX1
AML subtype

Figure 2.

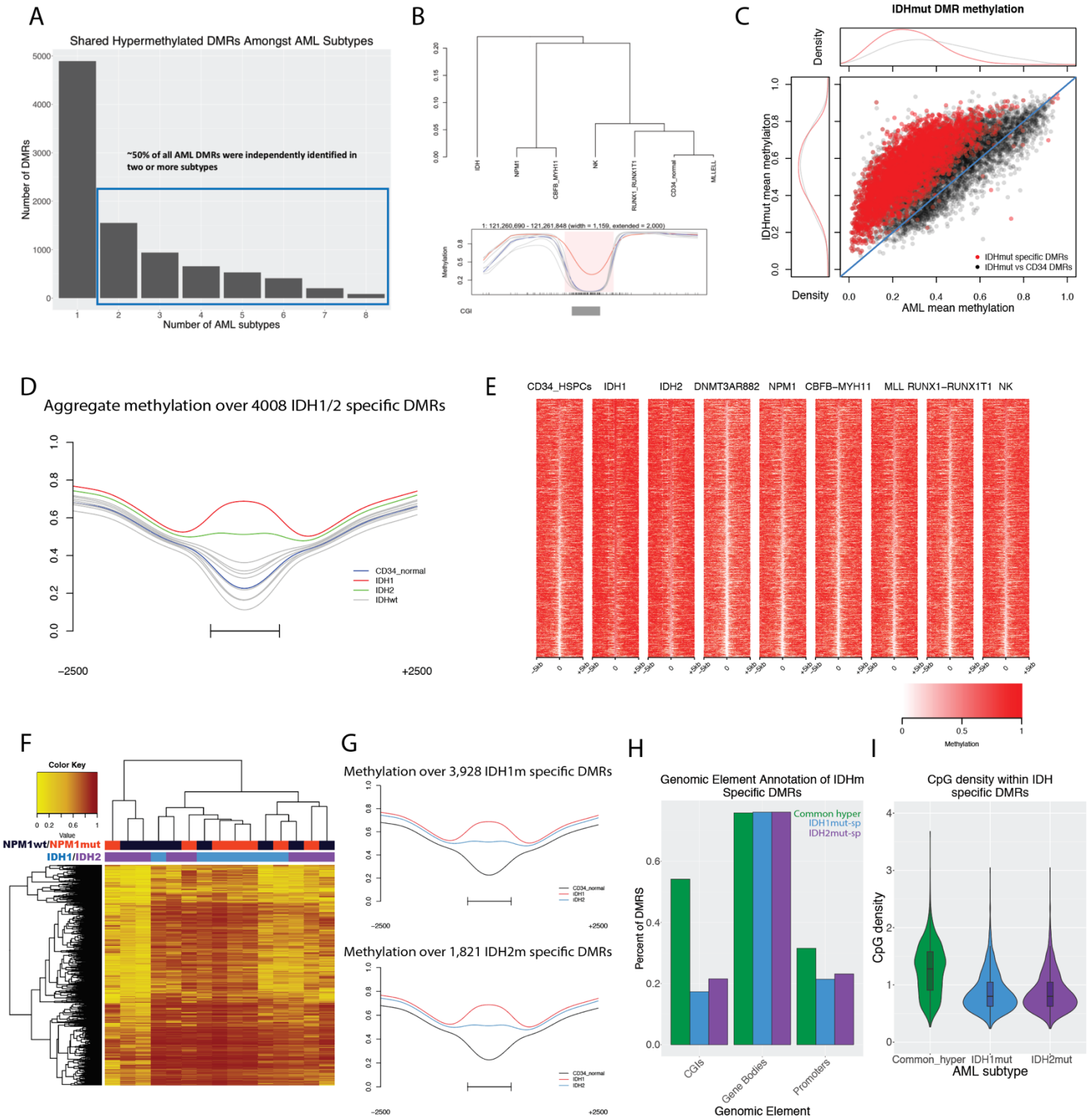

A
Shared Hypermethylated DMRs Amongst AML Subtypes
~50% of all AML DMRs were independently identified in two or more subtypes
Number of DMRs
Number of AML subtypes
B
IDH
NPM1
CBFB_MYH11
NK
RUNX1_RUNX1T1
CD34_normal
MLLELL
1: 121,260,690 - 121,261,848 (width = 1,159, extended = 2,000)
Methylation
CGI
C
IDHmut DMR methylation
Density
IDHmut mean methylaiton
AML mean methylation
IDHmut specific DMRs
IDHmut vs CD34 DMRs
D
Aggregate methylation over 4008 IDH1/2 specific DMRs
CD34_normal
IDH1
IDH2
IDHwt
−2500
+2500
E
CD34_HSPCs
IDH1
IDH2
DNMT3AR882
NPM1
CBFB−MYH11
MLL RUNX1−RUNX1T1
NK
Methylation
F
Color Key
Value
NPM1wt/NPM1mut
IDH1/IDH2
G
Methylation over 3,928 IDH1m specific DMRs
Methylation over 1,821 IDH2m specific DMRs
CD34_normal
IDH1
IDH2
−2500
+2500
H
Genomic Element Annotation of IDHm Specific DMRs
Common hyper
IDH1mut-sp
IDH2mut-sp
Percent of DMRS
CGIs
Gene Bodies
Promoters
Genomic Element
I
CpG density within IDH specific DMRs
CpG density
Common_hyper
IDH1mut
IDH2mut
AML subtype

Figure 3.

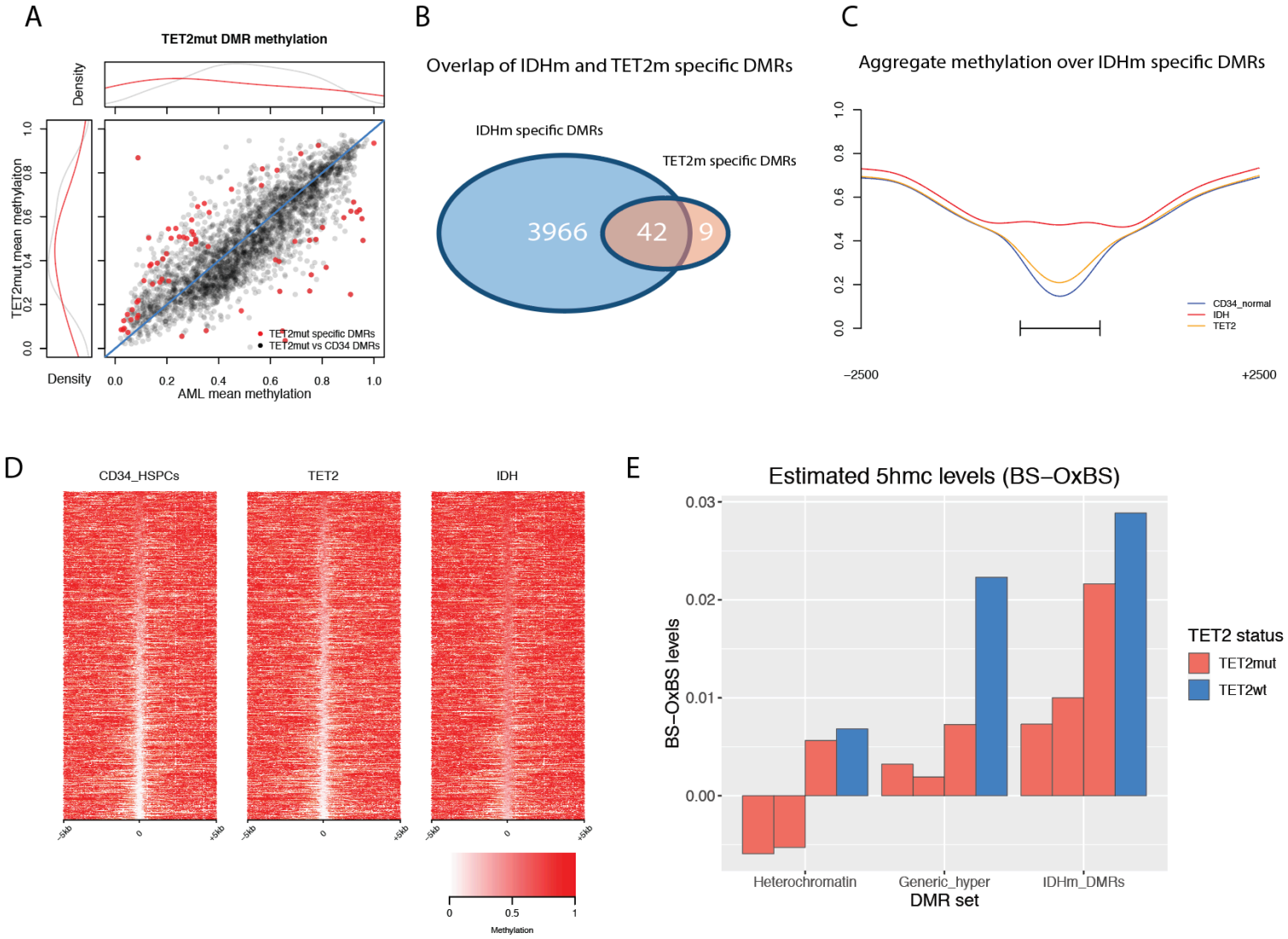

A
TET2mut DMR methylation
Density
TET2mut mean methylation
AML mean methylation
TET2mut specific DMRs
TET2mut vs CD34 DMRs
B
Overlap of IDHm and TET2m specific DMRs
IDHm specific DMRs
TET2m specific DMRs
3966
42
9
C
Aggregate methylation over IDHm specific DMRs
CD34_normal
IDH
TET2
−2500
+2500
D
CD34_HSPCs
TET2
IDH
Methylation
E
Estimated 5hmc levels (BS−OxBS)
BS−OxBS levels
TET2 status
TET2mut
TET2wt
Heterochromatin
Generic_hyper
IDHm_DMRs
DMR set

Figure 4.

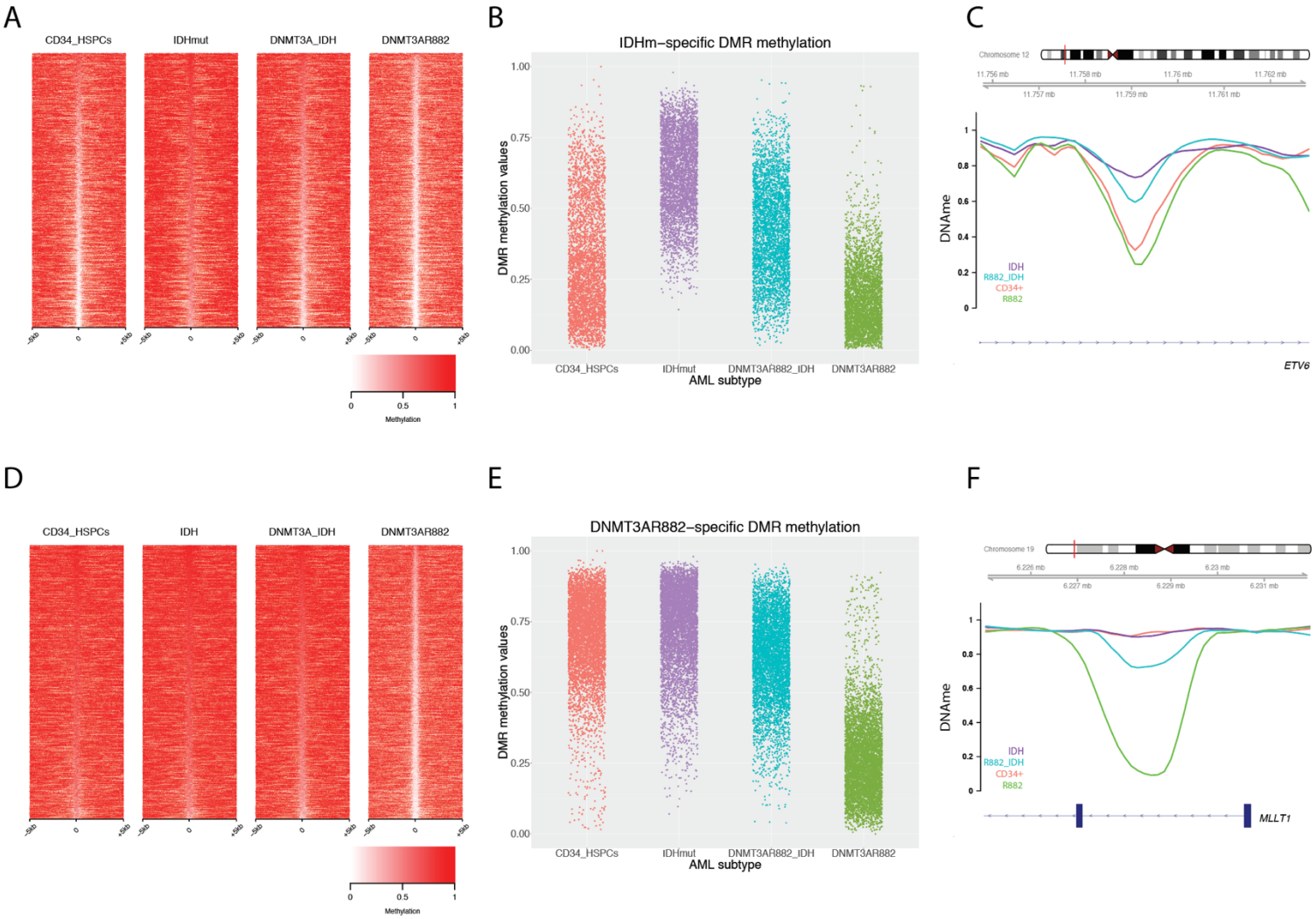
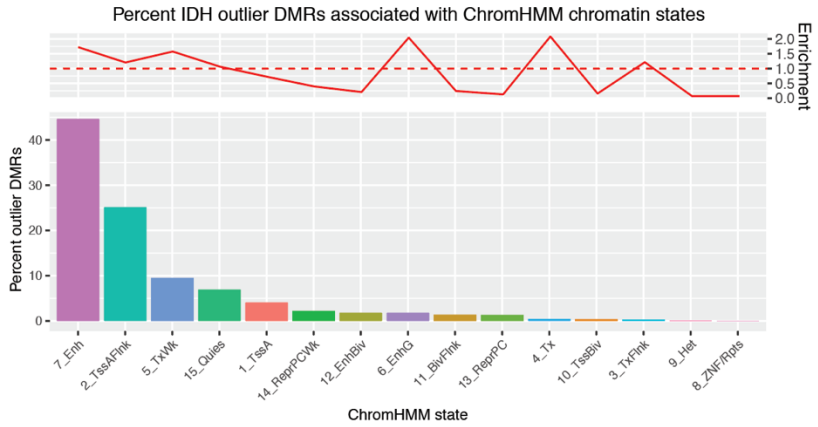
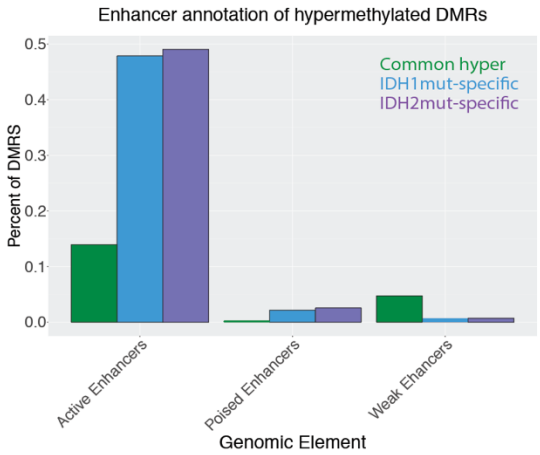

A
CD34_HSPCs
IDHmut
DNMT3A_IDH
DNMT3AR882
-5kb
0
+5kb
0
0.5
1
Methylation
B
IDHm−specific DMR methylation
DMR methylation values
1.00
0.75
0.50
0.25
0.00
CD34_HSPCs
IDHmut
DNMT3AR882_IDH
DNMT3AR882
AML subtype
C
Chromosome 12
11.756 mb
11.758 mb
11.76 mb
11.762 mb
11.757 mb
11.759 mb
11.761 mb
DNAme
IDH
R882_IDH
CD34+
R882
ETV6
D
CD34_HSPCs
IDH
DNMT3A_IDH
DNMT3AR882
-5kb
0
+5kb
0
0.5
1
Methylation
E
DNMT3AR882−specific DMR methylation
DMR methylation values
1.00
0.75
0.50
0.25
0.00
CD34_HSPCs
IDHmut
DNMT3AR882_IDH
DNMT3AR882
AML subtype
F
Chromosome 19
6.226 mb
6.228 mb
6.23 mb
6.227 mb
6.229 mb
6.231 mb
DNAme
IDH
R882_IDH
CD34+
R882
MLLT1

Figure 5.

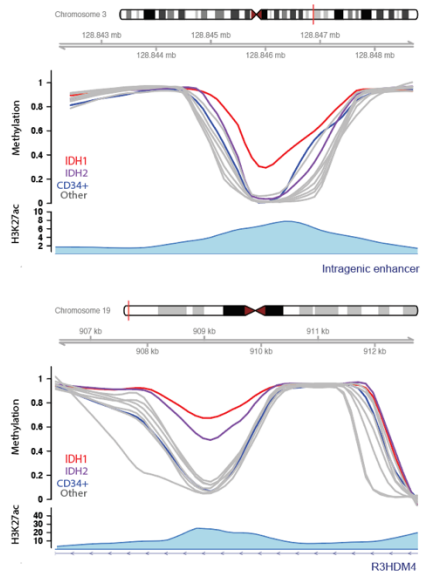
A



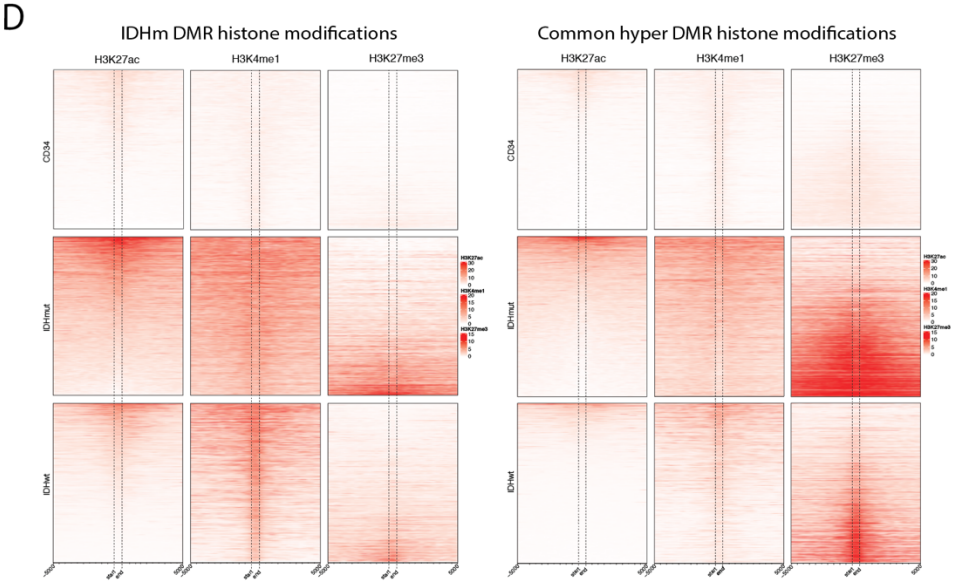
B



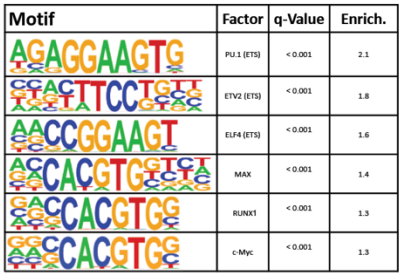
C



D



E

Motif	Factor	q-Value	Enrich.
	PU.1 (ETS)	< 0.001	2.1
	ETV2 (ETS)	< 0.001	1.8
	ELF4 (ETS)	< 0.001	1.6
	MAX	< 0.001	1.4
	RUNX1	< 0.001	1.3
	c-Myc	< 0.001	1.3

F

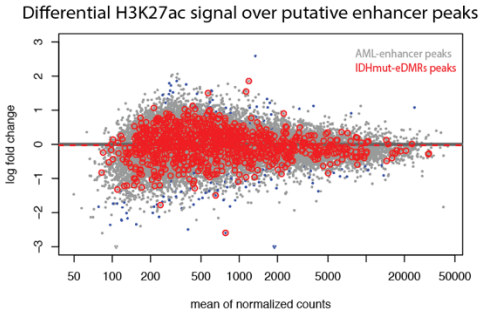
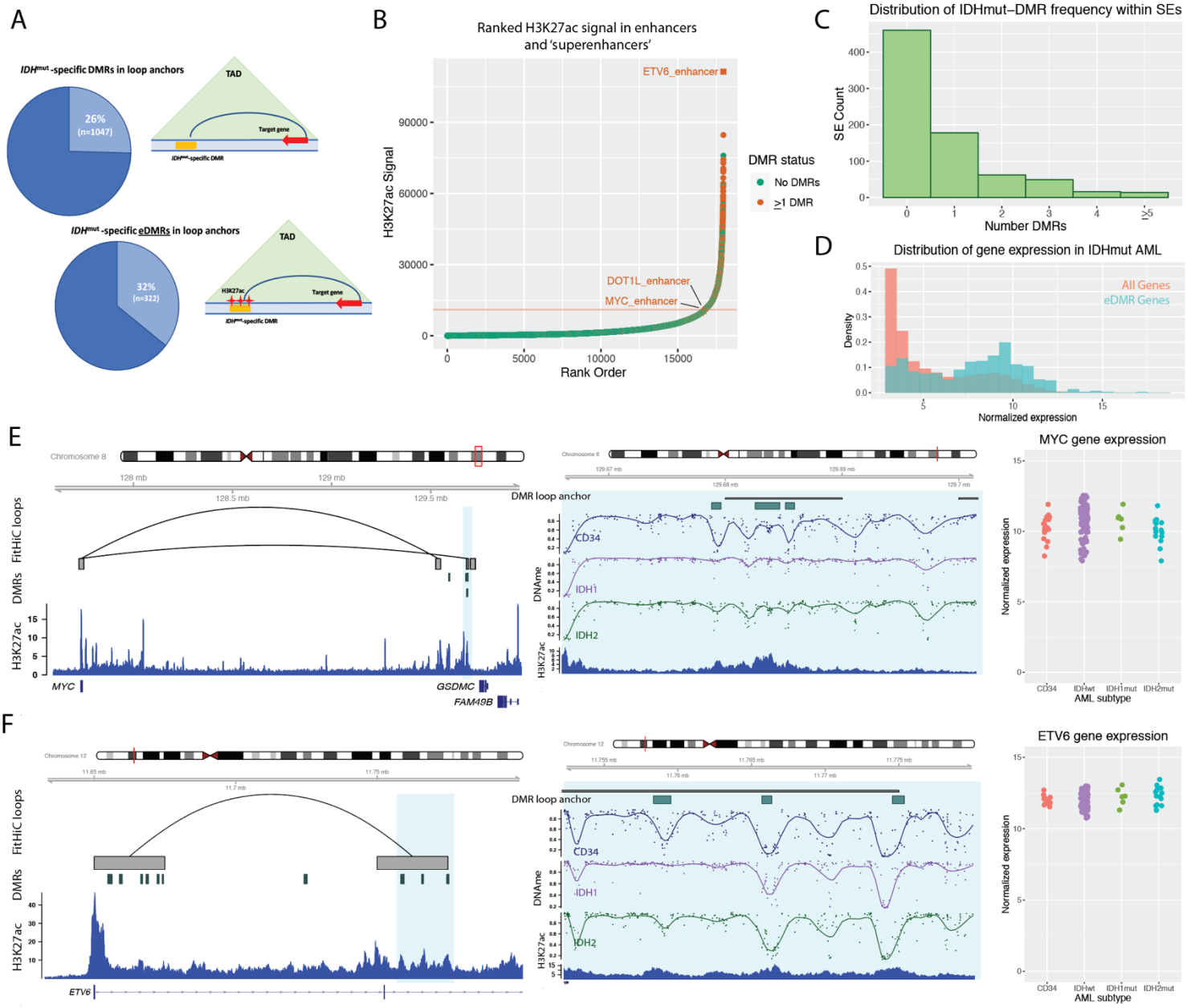


Figure 6.


A
IDHmut -specific DMRs in loop anchors
26%
(n=1047)
TAD
Target gene
IDHmut-specific DMR
IDHmut -specific eDMRs in loop anchors
32%
(n=322)
TAD
H3K27ac
Target gene
IDHmut-specific DMR
B
Ranked H3K27ac signal in enhancers
and 'superenhancers'
ETV6_enhancer
DOT1L_enhancer
MYC_enhancer
H3K27ac Signal
Rank Order
DMR status
No DMRs
≥1 DMR
C
Distribution of IDHmut–DMR frequency within SEs
SE Count
Number DMRs
D
Distribution of gene expression in IDHmut AML
All Genes
eDMR Genes
Density
Normalized expression
E
Chromosome 8
128 mb
129 mb
128.5 mb
129.5 mb
FitHiC loops
DMRs
H3K27ac
MYC
GSDMC
FAM49B
DMR loop anchor
CD34
IDH1
IDH2
DNAme
MYC gene expression
Normalized expression
CD34
IDHwt
IDH1mut
IDH2mut
AML subtype
F
Chromosome 12
FitHiC loops
DMRs
H3K27ac
ETV6
DMR loop anchor
CD34
IDH1
IDH2
DNAme
ETV6 gene expression
Normalized expression
CD34
IDHwt
IDH1mut
IDH2mut
AML subtype