haematologica
Journal of the Ferrata Storti Foundation

Overlapping DNA methylation changes in enhancers in clonal cytopenia of undetermined significance and myelodysplastic neoplasm patients with *TET2*, *IDH2*, or *DNMT3A* mutations

by Katja Kaastrup, Morten Tulstrup, Jakob W. Hansen, Claudia Schöllkopf, Klas Raaschou-Jensen, Andreas D. Ørskov, Bo Porse, Peter A. Jones, Joachim Weischenfeldt, Linn Gillberg and Kirsten Grønbæk

Received: April 4, 2024.
Accepted: December 17, 2024.

Citation: Katja Kaastrup, Morten Tulstrup, Jakob W. Hansen, Claudia Schöllkopf, Klas Raaschou-Jensen, Andreas D. Ørskov, Bo Porse, Peter A. Jones, Joachim Weischenfeldt, Linn Gillberg and Kirsten Grønbæk. Overlapping DNA methylation changes in enhancers in clonal cytopenia of undetermined significance and myelodysplastic neoplasm patients with TET2, IDH2, or DNMT3A mutations.
Haematologica. 2025 Jan 2. doi: 10.3324/haematol.2024.285466 [Epub ahead of print]

Overlapping DNA methylation changes in enhancers in clonal cytopenia of undetermined significance and myelodysplastic neoplasm patients with *TET2, IDH2,* or *DNMT3A* mutations

Authors

Katja Kaastrup[1,2], Morten Tulstrup[1,2], Jakob W. Hansen[1,2], Claudia Schöllkopf[1], Klas Raaschou-Jensen[3], Andreas D. Ørskov[4], Bo Porse[2,5,6], Peter A. Jones[7], Joachim Weischenfeldt[2,6], Linn Gillberg[1,2,8]*, Kirsten Grønbæk[1,2,5]*

Affiliations

[1]The Epi-/Genome lab, Department of Hematology, Copenhagen University Hospital - Rigshospitalet, Copenhagen, Denmark.

[2]Biotech Research and Innovation Center (BRIC), University of Copenhagen, Copenhagen, Denmark.

[3]Department of Hematology, Odense University Hospital, Odense, Denmark.

[4]Department of Hematology, Zealand University Hospital, Roskilde, Denmark.

[5]Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark.

[6]The Finsen Laboratory, Faculty of Health Sciences, Copenhagen University Hospital - Rigshospitalet, Copenhagen, Denmark.

[7]Department of Epigenetics, Van Andel Institute, Michigan, United States of America.

[8]Department of Biomedical Sciences, University of Copenhagen, Copenhagen, Denmark.

* Equal contribution.

Correspondence

Kirsten Grønbæk. The Epi-/Genome Lab, Department of Hematology, Rigshospitalet
Ole Maaløes Vej 5, 2200 Copenhagen N, Denmark. kirsten.groenbaek@regionh.dk

Author contributions

KK, LG, JWH, JW and KG designed the study. CS, KRJ, ADØ, BP and KG provided study samples. KK handled T cell depletion and DNA isolation. MT handled variant filtration. KK analyzed the data. LG, JWG, JW and KG supervised statistical analyses. KK, MT, LG, JWH, PAJ, JW and KG made interpretations of the results. KK wrote the first draft of the paper. All authors edited the paper and approved the final manuscript.

Disclosures

The study is part of the Danish Research Center for Precision Medicine in Blood Cancers funded by the Danish Cancer Society grant no. R223-A13071 and was supported by a centre grant from the Novo Nordisk Foundation (Novo Nordisk Foundation Centre for Stem Cell Biology, Dan Stem; grant number NNF17CC0027852) and Greater Copenhagen Health Science Partners, and through a grant from Dansk Kræftforskningsfond. KK was supported by a one-year grant from the University of Copenhagen.

Data-sharing statement

The raw and processed data generated in this study as well as other individual-level data are only available under restricted access since these data are considered sensitive personal data according to Danish Law and the European Union General Data Protection Regulation (GDPR) and thus cannot be shared with third parties without prior approval. To access data, an application must be sent to kirsten.groenbaek@regionh.dk. Access can only be granted for research purposes, and only if a data processor or data transfer agreement can be made in accordance with Danish and European law at the given time. The expected timeframe from response until access is granted is ~6 months.

Conflict of interest

KG is on advisory boards for GSK and Nanexa and has received research funding for a different project from Janssen. PAJ is a paid consultant for Zymo Research Corporation whose products were used in the DNA methylation analyses. KK, MT, JWH, CS, KRJ, ADØ, BP, JW and LG have nothing to declare.

The development of myelodysplastic neoplasm (MDS) is multiphasic, and the disease shares several clinical and molecular features with the precursor entities clonal hematopoiesis of indeterminate potential (CHIP) and clonal cytopenia of undetermined significance (CCUS) and the more aggressive acute myeloid leukemia (AML). Common for these is the presence of mutations in genes regulating DNA methylation[1–3]. Notably, two of the most mutated genes in CCUS and MDS are *DNMT3A*, catalyzing de novo methylation, and *TET2*, catalyzing the first steps of active DNA demethylation. We have previously shown that *TET2*-mutated CHIP and CCUS are associated with a distinct pattern of DNA hypermethylation especially enriched in genes related to immune response and leukocyte function[4]. Another study found that *DNMT3A*- and *TET2*-mutated CHIP have distinct and directionally opposing DNA methylation patterns but both likely promoting self-renewal of hematopoietic stem cells (HSCs)[5]. Furthermore, *IDH1* and *IDH2* mutations have been shown to inhibit TET2-mediated demethylation in AML and been associated with a DNA hypermethylation profile similar to *TET2*-mutated AML[6]. The epigenome-wide impact of mutations in *TET2*, *IDH1*, *IDH2*, and *DNMT3A* in CCUS and MDS is, however, not yet fully elucidated nor is it investigated whether the methylation patterns are shared across the range of clonal myeloid disorders.

We investigated whether 58 CCUS and 59 MDS patients with mutations in *TET2*, *IDH2*, or *DNMT3A* have distinct DNA methylation patterns in bone marrow mononuclear cells (MNCs) compared to patients without any of these mutations. Patients were recruited from hematologic departments at three Danish hospitals with approval from The Danish National Committee on Health Research Ethics (ID NVK-1705391). All participants provided written informed consent prior to participation and the study was conducted in accordance with the Danish ethical regulation for work with human participants. Patients donated bone marrow aspirates at time of diagnosis, wherefrom MNCs were isolated using density gradient separation and T-cell depleted using the RoboSep Human CD3 Positive Selection Kit II and the RoboSep Magnetic Cell Separator (StemCell Technologies, Vancouver, Canada). DNA sequencing was performed during diagnostic work-up or using a targeted sequencing panel and Twist Libraries (Twist Bioscience, San Francisco, USA; more information available upon request). Genome-wide DNA methylation levels were measured using Illumina's Infinium MethylationEPIC v1.0 BeadChips at the Genomics Core at the Van Andel Institute (Grand Rapids, Michigan, USA) according to the manufacturer's protocol. Data processing, quality control, and statistical analyses were done using R version 4.2.0

or higher. Data were quality controlled using minfi v.1.36.0[7] and normalized using functional normalization[8]. Probes with detection P-values>0.01, bead count<3 in at least 5% of samples, non-CpG sites, targeting to sex chromosomes, SNPs<5 bp from the target CpG, and probes which previously showed binding to multiple target CpGs[9] were excluded using ChAMP v. 2.20.1[10]. After quality control and filtering, methylation levels of 743,875 CpG sites were available for analysis. ComBat from R package sva 3.38.0 was used for sample plate correction. Differentially methylated positions (DMPs) were identified using linear modeling using Limma v.3.46.0[11] with M-values as outcome and mutation group as explanatory variable adjusted for age, sex, and batch. The wildtype group was defined as patients without mutations in *TET2*, *DNMT3A*, *IDH1*, and *IDH2*. P-values were adjusted for multiple testing using Benjamini and Hochberg's method to control the false discovery rate (FDR). If none of the DMPs were significant after correcting for multiple testing, we selected the threshold for statistical significance to unadjusted P-values<0.001. Associations between dichotomous outcomes were tested using χ^2 test.

The patients were 41-93 years of age (CCUS 54-90 years, MDS 41-93 years), predominantly male, and most MDS patients had lower-risk disease (Supplementary table 1). We first investigated the association between *TET2* mutations and genome-wide methylation levels in CCUS and MDS separately. Patients with mutations in *DNMT3A*, *IDH1*, or *IDH2* were removed from the analyses resulting in 23 CCUS and 20 MDS cases with, and 20 CCUS and 26 MDS cases without, *TET2* mutations (data on 20 CCUS patients from each group have been published previously[4]; Supplementary figure 1A and B). After correcting for multiple testing, we identified 24,515 DMPs in *TET2*-mutated CCUS compared to wildtype (Figure 1A). Almost all DMPs (24,365 DMPs, 99%) were hypermethylated. Similarly, we detected 7,242 DMPs in *TET2*-mutated MDS compared to wildtype with 98% (7,139) DMPs being hypermethylated (Figure 1B).

Next, we investigated the association between *IDH2* mutations and methylation levels. After removal of patients with *TET2*, *DNMT3A*, or *IDH1* mutations, four cases had *IDH2* mutations (one CCUS and three MDS) and 46 were *IDH2* wildtype (Supplementary figure 1C). Like *TET2* mutations, *IDH2* mutations were associated with hypermethylation of 94% (24,259) of the 25,888 DMPs identified after FDR correction (Figure 1C). Since 2-hydroxyglutarate, the metabolic product of mutated IDH2 proteins, inhibits the catalytic activity of TET2 enzymes, we investigated if the methylation profiles detected in *IDH2*-mutant cases were similar to those observed for *TET2* mutations. 98% (23,807 sites) and 97% (6,917 sites) of the *TET2* mutation-associated hypermethylated sites in CCUS and MDS, respectively, were associated with increased methylation

in *IDH2*-mutated cases ($P<2.2x10^{-16}$, Figure 1F and G). *IDH1* mutations were not studied due to a low mutation count (n=1).

Eight CCUS and five MDS cases had mutations in *DNMT3A* and 20 CCUS and 26 MDS cases did not, after removal of cases with *TET2*, *IDH1*, or *IDH2* mutations (Supplementary figure 1D and E). We did not detect any DMPs in *DNMT3A*-mutated CCUS after FDR correction and only 704 DMPs were detected at unadjusted $P<0.001$ (Figure 1D). This may be due to differences in cell populations, lower clonality, and/or number of mutated cases. Similarly, only 1,808 DMPs were detected in *DNMT3A*-mutated MDS at unadjusted $P<0.001$ with 93% (1,681) being hypomethylated (Figure 1E). The larger number of DMPs associated with *TET2* than *DNMT3A* mutations may be due to type and location of the mutation (*DNMT3A* mutations were mainly missense whereas *TET2* mutations were primarily frameshift and nonsense mutations, which have been associated with increased clonal selection[12]), cell type specificity, and/or power since a higher number of patients had *TET2* than *DNMT3A* mutations. In summary, mutations in *TET2* and *IDH2* are associated with a hypermethylation phenotype affecting thousands of CpG sites in both CCUS and MDS, whereas *DNMT3A* mutations are associated with a hypomethylation phenotype, which we only detected in MDS.

As we have previously found 2,741 hypermethylated CpG sites in *TET2*-mutated CHIP[5], we next investigated whether the same sites were affected by *TET2* mutations across CHIP, CCUS, and MDS. 2,464 CpG sites were available for evaluation after quality control, and 98% (2,458 sites) and 99.6% (2,455 sites) of those were positively associated with *TET2* mutations in CCUS and MDS, respectively ($P<2.2x10^{-16}$, Figure 2A and B). Similarly, 98% (23,973 sites) of the 24,365 sites hypermethylated in *TET2*-mutated CCUS were hypermethylated in *TET2*-mutated MDS ($P<2.2x10^{-16}$, Figure 2C). This indicates that methylation changes associated with *TET2* mutations are affecting many of the same loci irrespective of disease state. Next, we investigated whether *DNMT3A* mutations impact DNA methylation similarly in MDS and CHIP. We found that 85% (2,924 sites) of the CpG sites previously reported hypomethylated in *DNMT3A*-mutated CHIP[5] (3,433 of 3,803 sites available for analysis) were associated with lower methylation in *DNMT3A*-mutated MDS ($P<2.2x10^{-16}$, Figure 2D). In summary, the large overlap of mutation-associated DMPs across different hematopoietic contexts indicates that methylation phenotypes associated with *TET2* and *DNMT3A* mutations are remarkably specific, which is in line with previous findings in CHIP and AML[4–6,13]. Due to the large overlap in CCUS and CHIP, it is unlikely that mutation-associated DMPs in MDS are influenced by blast count or IPSS-R.

Although mutation-specific DMPs distributed across the entire genome, they were significantly enriched at enhancer regions and depleted at CpG islands for both hypermethylated sites associated with *TET2* and *IDH2* mutations and hypomethylated sites associated with *DNMT3A* mutations. Indeed, 11% of hypermethylated sites versus 4% of non-hypermethylated sites detected in *TET2*-mutated CCUS were in enhancer regions ($P<2.2x10^{-16}$, Figure 3A), and only 3% of hypermethylated *TET2*-mutated sites versus 20% of non-hypermethylated sites were in CpG islands ($P<2.2x10^{-16}$). Strikingly similar results were observed for hypermethylated sites associated with *TET2*-mutated MDS and *IDH2*-mutated CCUS and MDS, respectively ($P<2.2x10^{-16}$; Figure 3B). Likewise, 18% of all hypomethylated sites versus 4% of the non-hypomethylated sites associated with *DNMT3A* mutations in MDS were in enhancer regions ($P<2.2x10^{-16}$, Figure 3C) and only 8% of hypomethylated sites were in CpG islands while this was the case for 19% of non-hypomethylated sites ($P<2.2x10^{-16}$). Interestingly, *TET2* mutation-associated hypermethylation was especially enriched in regions annotated as enhancers in monocytes (Figure 3D), whereas *IDH2* mutation-associated hypermethylation was equally present at monocyte and HSC enhancers (Figure 3E). In contrast, *DNMT3A* mutation-associated hypomethylation was enriched in HSC enhancers (Figure 3F). These findings demonstrate that the methylation changes are non-random and context-dependent, implying that the activity of TET2 and DNMT3A may be targeted to specific genomic regions in hematopoietic cells. TET2 has been suggested to have the strongest demethylation effect proximal to bound transcription factors[14] and we have previously shown a strong enrichment of ETS and C/EBP transcription factor binding motifs near *TET2* mutation-associated hypermethylated sites in CHIP and CCUS [4].

Gene ontology (GO) term enrichment analysis of genes in the vicinity of the DMPs revealed strikingly similar results for *TET2* and *IDH2* mutation-associated hypermethylated sites and *DNMT3A* mutation-associated hypomethylated sites with enrichment of genes related to immune system process and leukocyte activation (Supplementary figure 2A-D). Similarly, genes near the 442 DMPs that were common for all three mutations were involved in immune regulation and immune cell activation (Supplementary figure 2E). These findings support a shared target pathway of *TET2*, *IDH2*, and *DNMT3A* mutation-induced methylation changes.

Taken together, our results demonstrate the existence of distinct and non-random DNA methylation changes associated with mutations in *TET2*, *IDH2* and *DNMT3A* in CCUS and MDS. We observed similar mutation-induced methylation changes across a spectrum of myeloid clonal disorders, suggesting that mutation-associated methylation changes may be an early event in leukemogenesis.

Moreover, mutations in *TET2, IDH2,* and *DNMT3A* especially induce methylation changes at enhancer regions but with differential preferences for enhancers according to cell type, suggesting that these mutations affect regulatory regions of genes at various stages of hematopoietic cell differentiation. This is in line with a previous study showing that *DNMT3A* mutations led to DNA hypomethylation in regulatory regions of genes that are related to HSC activity, while *TET2* mutations led to DNA hypermethylation at regulatory regions of genes that downregulate the activity of more differentiated progenitor cells, thus both maintaining a stem cell-like state[5]. Future studies of the genes regulated by these enhancers may shed light on the potential pathogenic impact of methylation changes during leukemogenesis.

References

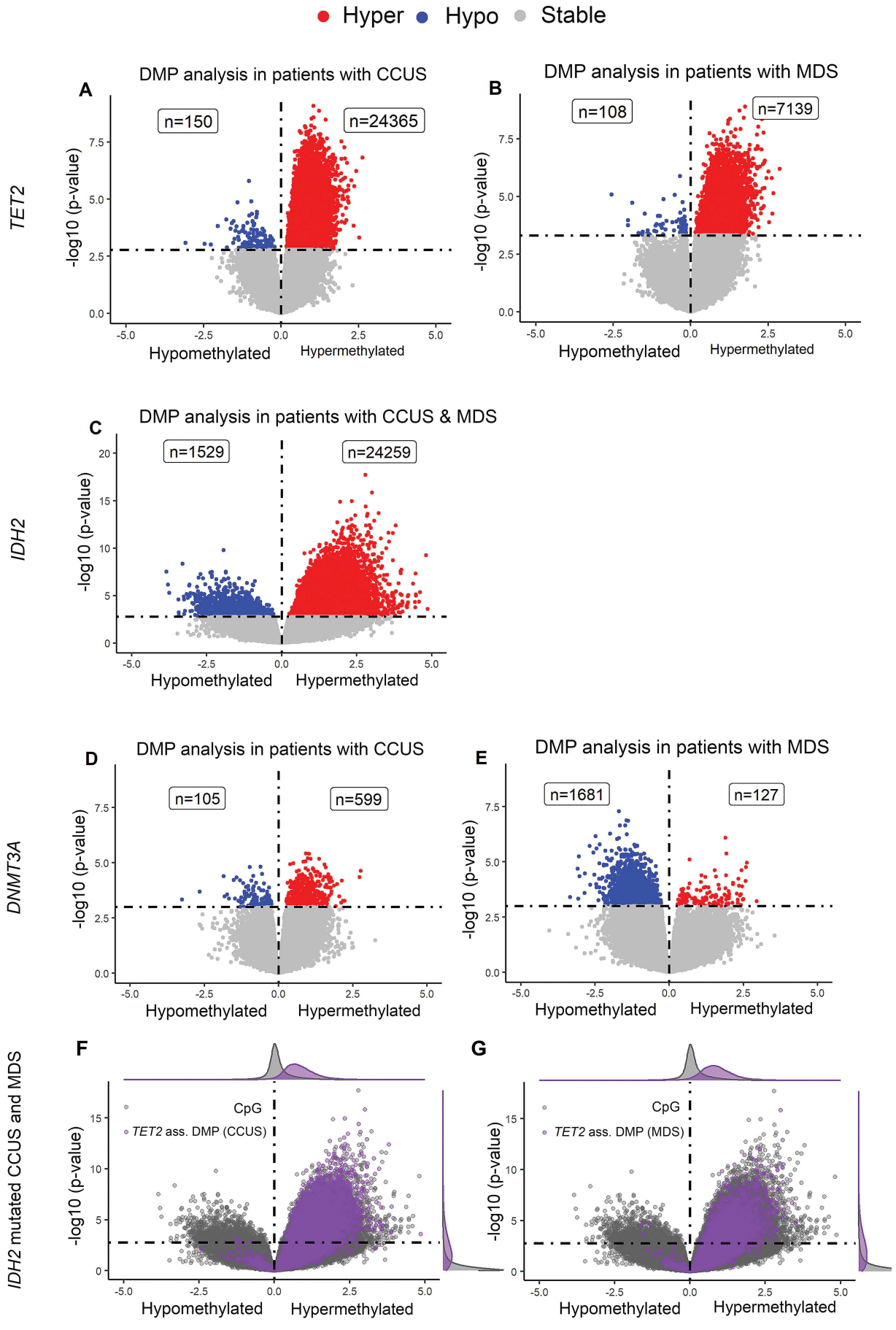
1. Hansen JW, Westman MK, Sjö LD, et al. Mutations in idiopathic cytopenia of undetermined significance assist diagnostics and correlate to dysplastic changes. Am J Hematol. 2016;91(12):1234-1238.
2. Genovese G, Kähler AK, Handsaker RE, et al. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. N Engl J Med. 2014;371(26):2477-2487.
3. Papaemmanuil E, Gerstung M, Malcovati L, et al. Clinical and biological implications of driver mutations in myelodysplastic syndromes. Blood. 2013;122(22):3616-3627.
4. Tulstrup M, Soerensen M, Hansen JW, et al. TET2 mutations are associated with hypermethylation at key regulatory enhancers in normal and malignant hematopoiesis. Nat Commun. 2021;12(1):6061.
5. Uddin M d. M, Nguyen NQH, Yu B, et al. Clonal hematopoiesis of indeterminate potential, DNA methylation, and risk for coronary artery disease. Nat Commun. 2022;13(1):5350.
6. Figueroa ME, Abdel-Wahab O, Lu C, et al. Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. Cancer Cell. 2010;18(6):553-567.
7. Aryee MJ, Jaffe AE, Corrada-Bravo H, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics. 2014;30(10):1363-1369.
8. Fortin JP, Labbe A, Lemire M, et al. Functional normalization of 450k methylation array data improves replication in large cancer studies. Genome Biol. 2014;15(12):503.
9. Nordlund J, Bäcklin CL, Wahlberg P, et al. Genome-wide signatures of differential DNA methylation in pediatric acute lymphoblastic leukemia. Genome Biol. 2013;14(9):r105.
10. Tian Y, Morris TJ, Webster AP, et al. ChAMP: updated methylation analysis pipeline for Illumina BeadChips. Bioinformatics. 2017;33(24):3982-3984.
11. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43(7):e47.
12. Fabre MA, de Almeida JG, Fiorillo E, et al. The longitudinal dynamics and natural history of clonal haematopoiesis. Nature. 2022;606(7913):335-342.
13. Wilson ER, Helton NM, Heath SE, et al. Focal disruption of DNA methylation dynamics at enhancers in IDH-mutant AML cells. Leukemia. 2022;36(4):935-945.
14. Ginno PA, Gaidatzis D, Feldmann A, et al. A genome-scale map of DNA methylation turnover identifies site-specific dependencies of DNMT and TET activity. Nat Commun. 2020;11(1):2680.
15. Roadmap Epigenomics Consortium, Kundaje A, Meuleman W, et al. Integrative analysis of 111 reference human epigenomes. Nature. 2015;518(7539):317-330.

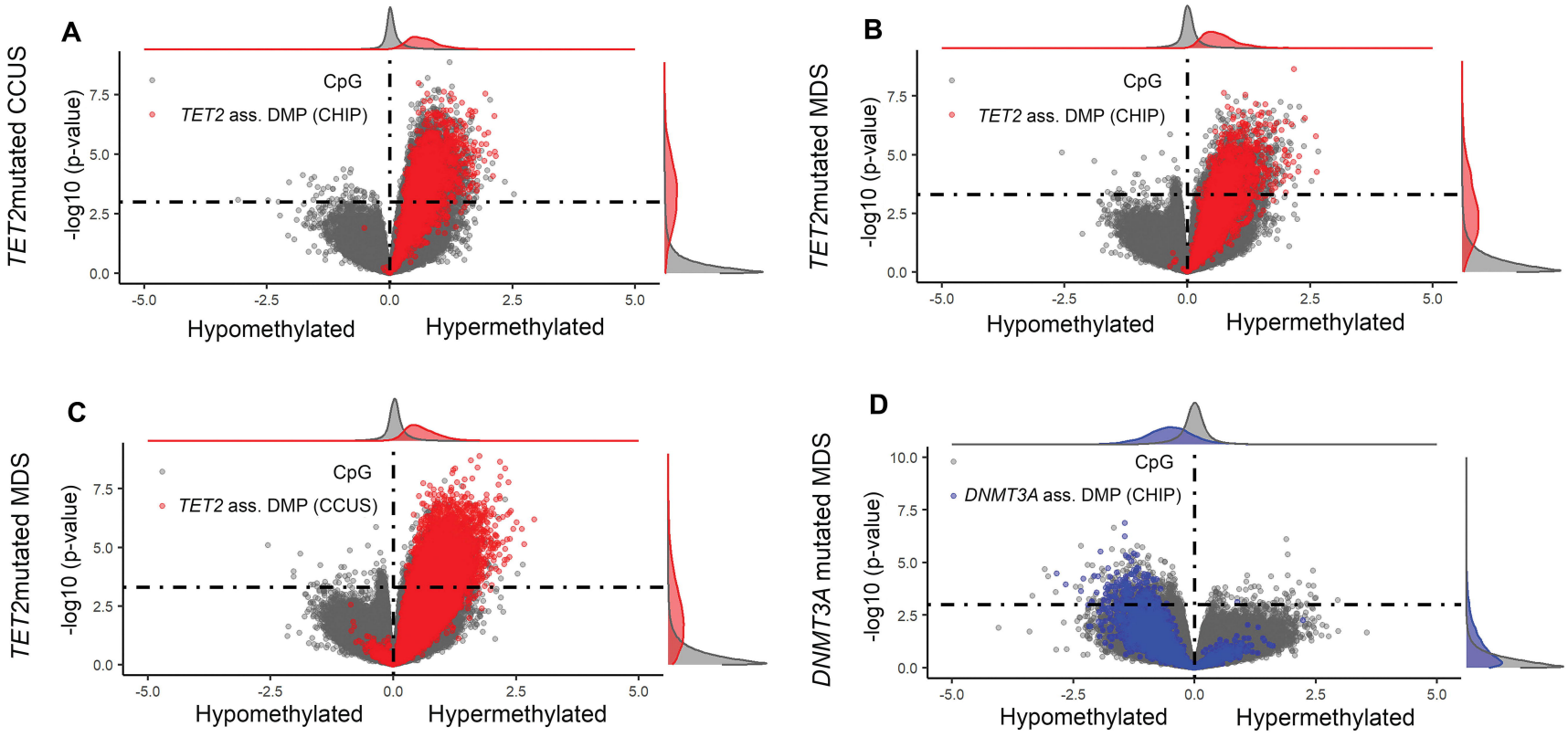
Figure legends

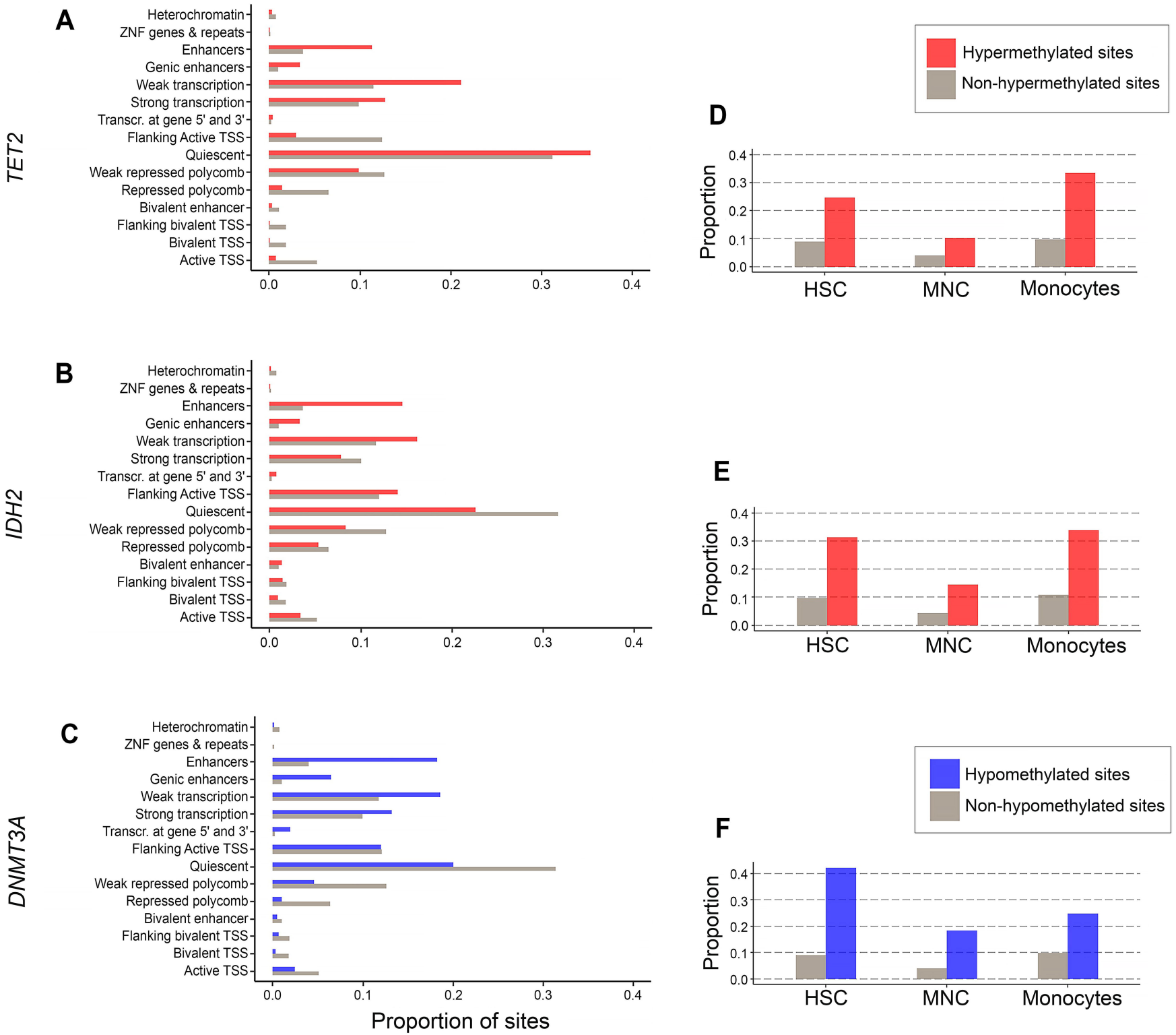
Figure 1. Differential methylation associated with *TET2*, *IDH2*, and *DNMT3A* mutations. Volcano plot of results from DMP analysis of **A**) 23 *TET2*-mutated CCUS compared to 20 *TET2* wildtype CCUS, **B**) 20 *TET2*-mutated MDS compared to 26 *TET2* wildtype MDS, **C**) 4 *IDH2*-mutant compared to 46 *IDH2* wildtype CCUS and MDS, **D**) 8 *DNMT3A*-mutant compared to 20 *DNMT3A* wildtype CCUS, and **E**) 5 *DNMT3A*-mutated MDS compared to 26 *DNMT3A* wildtype MDS. Red and blue dots represent hyper- and hypomethylated CpG sites, respectively, below the significance threshold. Horizontal dashed line represents false discovery rate = 0.05 in A-C) and P<0.001 in D-E). **F and G**) *IDH2*-mutated CCUS and MDS (same as in Figure 1C) with purple dots representing the **F**) 24,365 CpG sites showing *TET2* mutation-associated significant hypermethylation in CCUS and the **G**) 7,139 CpG sites showing *TET2* mutation-associated significant hypermethylation in MDS, respectively. All *IDH2*-mutated cases had cooccurring mutations in *SRSF2*. P-values and effect size estimates in all panels were derived using a linear regression model. All P-values are two-sided and not adjusted for multiple comparisons. CCUS: clonal cytopenia of undetermined significance; MDS: myelodysplastic neoplasms; DMP: differentially methylated position.

Figure 2. Comparison of results from epigenome-wide analyses of *TET2* and *DNMT3A* mutations in CHIP, CCUS, and MDS. Volcano plot of results from DMP analysis in: **A)** *TET2*-mutated CCUS (same as in Figure 1A) and **B)** *TET2*-mutated MDS (same as in Figure 1B) with red dots indicating the 2,741 sites previously reported hypermethylated in *TET2*-mutated CHIP[4]. **C)** *TET2* mutated MDS (same as in Figure 2B) with red dots representing the 24,365 CpG sites showing *TET2* mutation-associated significant hypermethylation in CCUS. **D)** *DNMT3A*-mutated MDS (same as in Figure 1E) with blue dots indicating the 3,803 CpG sites previously reported hypomethylated in *DNMT3A*-mutated CHIP[5]. Margins display distributions of colored sites versus the rest of CpG sites avaible after quality control (either rest of Infinium HumanMethylation450-sites in case of CHIP or Infinium HumanMethylationEPIC-sites in case of CCUS or MDS). P-values and effect size estimates in all panels were derived using a linear regression model. All P-values are two-sided and not adjusted for multiple comparisons. CHIP: clonal hematopoiesis of indeterminate potential; CCUS: clonal cytopenia of undetermined significance; MDS: myelodysplastic neoplasms; ass.: associated; DMP: differentially methylated position.

Figure 3. Annotation of *TET2, IDH2,* and *DNMT3A* mutation-associated DMPs to chromatin states and cell-specific enhancers. Annotation of interrogated CpG sites and methylation changes to chromatin states of peripheral blood mononuclear cells (A-C). **A)** Proportions of the 24,365 *TET2* mutation-associated hypermethylated sites in CCUS (red bars) and the rest of the sites (grey bars) at different chromatin states. **B)** Proportions of the 24,259 *IDH2* mutation-associated hypermethylated sites in CCUS and MDS (red bars) and the rest of the sites (grey bars) at different chromatin states. **C)** Proportions of the 1,681 *DNMT3A* mutation-associated hypomethylated sites in MDS (blue bars) and the rest of the sites (grey bars) at different chromatin states. DMPs were annotated to genomic regions using the annotation package for Illumina's methylation arrays and to chromatin states using the mononuclear dataset from the Roadmap Epigenomics project[15]. Proportion of DMPs and the rest of available sites interrogated by the EPIC v1.0 array located in regions annotated as enhancers in HSCs, MNCs, or monocytes (**D**-**F**). DMPs are **D)** the 7,139 CpG sites hypermethylated in *TET2*-mutated MDS (95% confidence intervals (CI): 0.32-0.35 and 0.23-0.25 for monocyte and HSC enhancers, respectively), **E)** the 24,259 CpG sites hypermethylated in *IDH2*- mutated CCUS and MDS (95% CI: 0.33-0.34 and 0.31-0.32 for monocyte and HSC enhancers, respectively), and **F)** the 1,681 CpG sites hypomethylated in *DNMT3A*-mutated MDS (95% CI: 0.23-0.27 and 0.40-0.45 for monocyte and HSC enhancers, respectively). CCUS: clonal cytopenia of undetermined significance; DMP: differentially methylated position; HSC: hematopoietic stem cell; MDS: myelodysplastic neoplasms; MNC: mononuclear cell; TSS: transcription start site.


Hyper
Hypo
Stable
TET2
IDH2
DNMT3A
IDH2 mutated CCUS and MDS
A
DMP analysis in patients with CCUS
n=150
n=24365
B
DMP analysis in patients with MDS
n=108
n=7139
C
DMP analysis in patients with CCUS & MDS
n=1529
n=24259
D
DMP analysis in patients with CCUS
n=105
n=599
E
DMP analysis in patients with MDS
n=1681
n=127
F
CpG
TET2 ass. DMP (CCUS)
G
CpG
TET2 ass. DMP (MDS)
-log10 (p-value)
Hypomethylated
Hypermethylated


A
TET2mutated CCUS
-log10 (p-value)
CpG
TET2 ass. DMP (CHIP)
Hypomethylated
Hypermethylated
B
TET2mutated MDS
-log10 (p-value)
CpG
TET2 ass. DMP (CHIP)
Hypomethylated
Hypermethylated
C
TET2mutated MDS
-log10 (p-value)
CpG
TET2 ass. DMP (CCUS)
Hypomethylated
Hypermethylated
D
DNMT3A mutated MDS
-log10 (p-value)
CpG
DNMT3A ass. DMP (CHIP)
Hypomethylated
Hypermethylated


A
B
C
D
E
F
TET2
IDH2
DNMT3A
Heterochromatin
ZNF genes & repeats
Enhancers
Genic enhancers
Weak transcription
Strong transcription
Transcr. at gene 5' and 3'
Flanking Active TSS
Quiescent
Weak repressed polycomb
Repressed polycomb
Bivalent enhancer
Flanking bivalent TSS
Bivalent TSS
Active TSS
0.0
0.1
0.2
0.3
0.4
Proportion of sites
Hypermethylated sites
Non-hypermethylated sites
Hypomethylated sites
Non-hypomethylated sites
Proportion
HSC
MNC
Monocytes

Supplementary data

Supplementary table 1. Clinical characteristics.

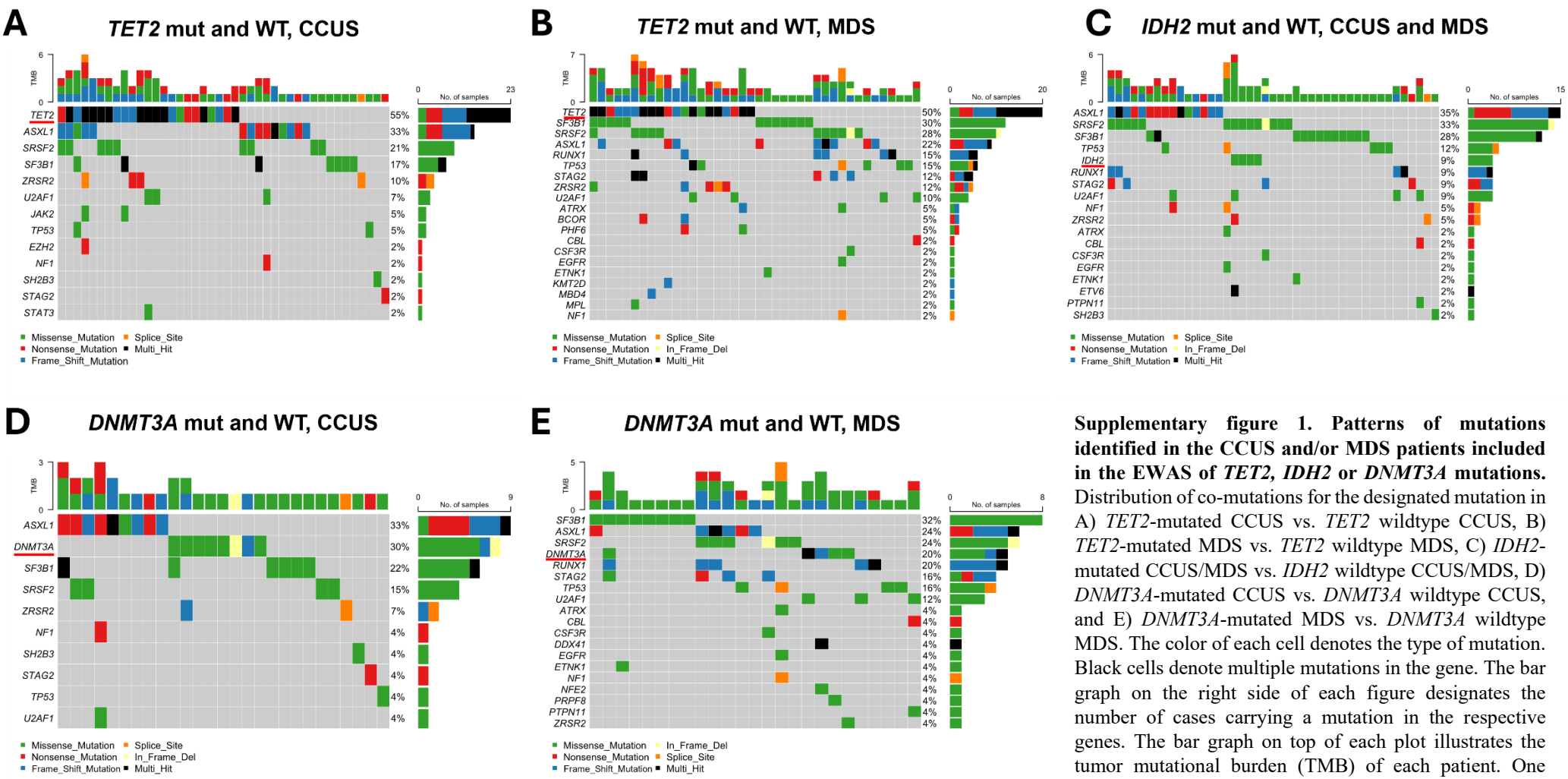
Characteristic	CCUS, n = 58 [1]	MDS, n = 59 [1]	P-value [2]
Age	75.0 (54.0, 90.0)	74.0 (41.0, 93.0)	>0.9
Sex			0.3
Female	12 (21%)	17 (29%)	
Male	46 (79%)	42 (71%)	
Hgb (mmol/L)	7.5 (4.8, 10.0)	6.3 (3.0, 9.0)	**<0.001**
ANC (cells/L)	2.2 (0.4, 11.1)	1.9 (0.3, 6.7)	0.07
Missing	1	0	
Platelets (cells/L)	135.0 (20.0, 539.0)	126.0 (25.0, 666.0)	0.6
Missing	1	0	
IPSS-R category			>0.9
Very low		13 (25%)	
Low		27 (51%)	
Intermediate		8 (15%)	
High		5 (9.4%)	
Very high		0 (0%)	
Missing		6	
Number of mutations			**0.003**
0	1 (1.7%) [3]	6 (10%)	
1	27 (47%)	10 (17%)	
2	12 (21%)	16 (27%)	
≥ 3	18 (31%)	27 (46%)	

[1] Median (range); n (%)

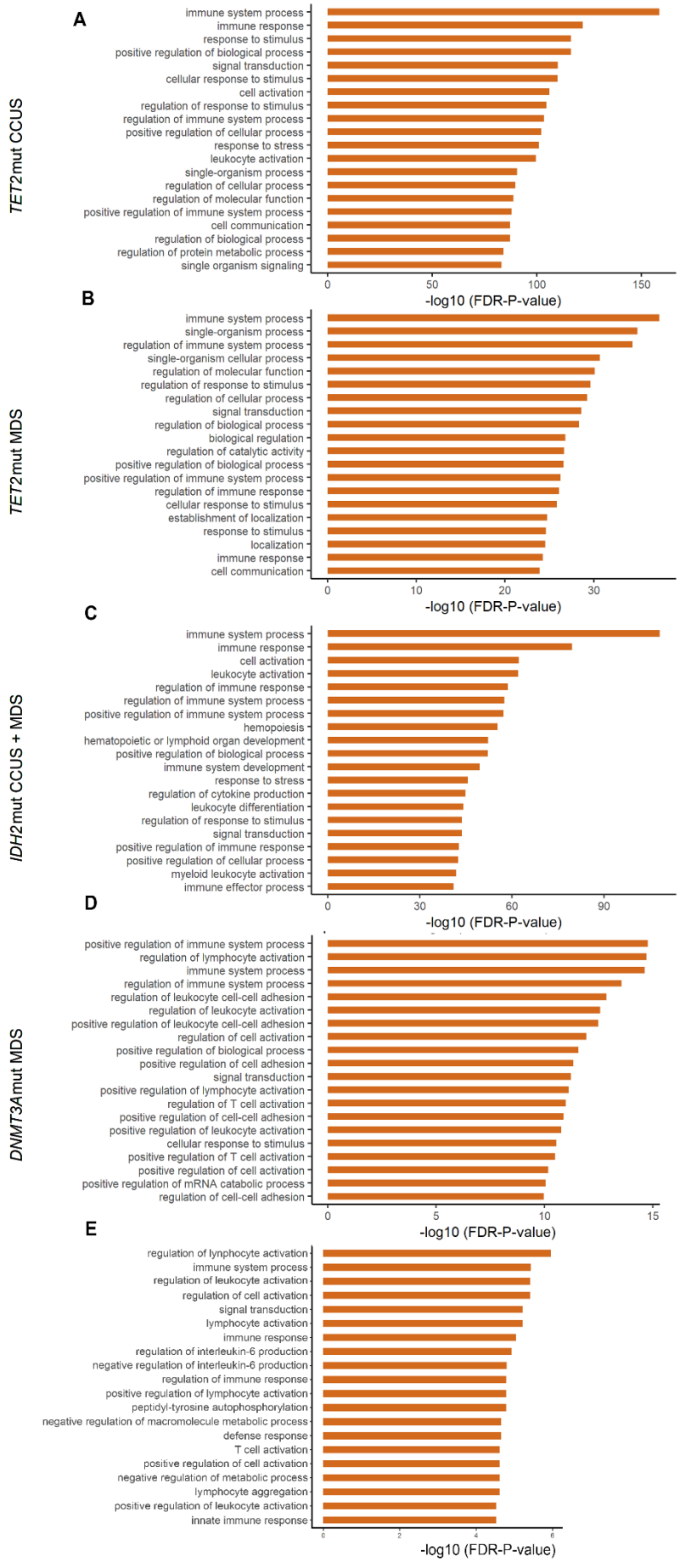
[2] Wilcoxon rank sum test; Pearson's Chi-squared test; Fisher's exact test

[3] One patient with CCUS has no somatic mutations but present with trisomy 15

Hgb: hemoglobin; ANC: absolute neutrophil count; IPSS-R, Revised International Prognostic Scoring System; CCUS: clonal cytopenia of undetermined significance; MDS: myelodysplastic neoplasms.



Supplementary figure 1. Patterns of mutations identified in the CCUS and/or MDS patients included in the EWAS of *TET2, IDH2* or *DNMT3A* mutations. Distribution of co-mutations for the designated mutation in A) *TET2*-mutated CCUS vs. *TET2* wildtype CCUS, B) *TET2*-mutated MDS vs. *TET2* wildtype MDS, C) *IDH2*-mutated CCUS/MDS vs. *IDH2* wildtype CCUS/MDS, D) *DNMT3A*-mutated CCUS vs. *DNMT3A* wildtype CCUS, and E) *DNMT3A*-mutated MDS vs. *DNMT3A* wildtype MDS. The color of each cell denotes the type of mutation. Black cells denote multiple mutations in the gene. The bar graph on the right side of each figure designates the number of cases carrying a mutation in the respective genes. The bar graph on top of each plot illustrates the tumor mutational burden (TMB) of each patient. One CCUS patient (with trisomy 15) and six MDS patients has no mutations of the examined genes and are not included in the plots. The plots are generated using matfools, Bioconductor. EWAS: epigenome-wide association study; CCUS: clonal cytopenia of undetermined significance; MDS, myelodysplastic neoplasms; mut, mutation; TMB, tumor mutational burden; WT, wildtype.



Supplementary figure 2. Enrichment of Gene Ontology terms for genes near *TET2* and *IDH2* mutation-associated DMPs in CCUS and MDS (A, B and C), *DNMT3A* mutation-associated DMPs in MDS (D), and the 442 shared DMPs associated with *TET2*-mutated CCUS and MDS, *IDH2*-mutated CCUS and MDS, and *DNMT3A* mutated CHIP and MDS (E). Top 20 Gene Ontology terms are presented for genes near *TET2* mutation-associated hypermethylated sites in CCUS (24,365 sites; **A**) and MDS (7,139 sites; **B**), *IDH2* mutation-associated hypermethylated sites in CCUS and MDS (24,259 sites; **C**) *DNMT3A* mutation-associated hypomethylated sites in MDS (1,681 sites; **D**) and DMPs common for all three mutations (442 sites, **E**). P-values derive from hypergeometric tests performed using rGREAT v. 1.28.0 and are FDR-corrected. CCUS: clonal cytopenia of undetermined significance; MDS: myelodysplastic neoplasms; FDR: false discovery rate.