

## ARTICLE

# DNMT3A haploinsufficiency causes dichotomous DNA methylation defects at enhancers in mature human immune cells

Jung-Yeon Lim<sup>1</sup>, Sascha H. Duttke<sup>2</sup>, Turner S. Baker<sup>1</sup>, Jihye Lee<sup>1</sup>, Kristyne J. Gambino<sup>1</sup>, Nicholas J. Venturini<sup>1</sup>, Jessica Sook Yui Ho<sup>3</sup>, Simin Zheng<sup>3</sup>, Yesai S. Fstckchyan<sup>3</sup>, Vinodh Pillai<sup>4</sup>, David C. Fajgenbaum<sup>5</sup>, Ivan Marazzi<sup>3</sup>, Christopher Benner<sup>2</sup>, and Minji Byun<sup>1,6</sup>

**DNMT3A** encodes an enzyme that carries out *de novo* DNA methylation, which is essential for the acquisition of cellular identity and specialized functions during cellular differentiation. *DNMT3A* is the most frequently mutated gene in age-related clonal hematopoiesis. As such, mature immune cells harboring *DNMT3A* mutations can be readily detected in elderly persons. Most *DNMT3A* mutations associated with clonal hematopoiesis are heterozygous and predicted to cause loss of function, indicating that haploinsufficiency is the predominant pathogenic mechanism. Yet, the impact of *DNMT3A* haploinsufficiency on the function of mature immune cells is poorly understood. Here, we demonstrate that *DNMT3A* haploinsufficiency impairs the gain of DNA methylation at decommissioned enhancers, while simultaneously and unexpectedly impairing DNA demethylation of newly activated enhancers in mature human myeloid cells. The DNA methylation defects alter the activity of affected enhancers, leading to abnormal gene expression and impaired immune response. These findings provide insights into the mechanism of immune dysfunction associated with clonal hematopoiesis and acquired *DNMT3A* mutations.

## Introduction

*DNMT3A* encodes a DNA methyltransferase that carries out DNA methylation, an evolutionarily conserved epigenetic mechanism critical for mammalian development (Smith and Meissner, 2013). The DNA methylation landscape in somatic cells is stably maintained at the genomic level except for tissue-specific regulatory regions (He et al., 2020). Perturbations of the deposition and removal of DNA methylation at these regions impair lineage fate decisions (Bröske et al., 2009; Trowbridge et al., 2009), indicating the importance of DNA methylation for the acquisition of proper cellular identity and specialized functions.

*DNMT3A* is the most frequently mutated gene in clonal hematopoiesis (CH), which occurs when a clonal population of blood cells carrying advantageous somatic mutations expands faster than others. CH is strongly associated with aging, affecting >10% of individuals older than 65 yr (Genovese et al., 2014; Jaiswal et al., 2014; Xie et al., 2014). *DNMT3A* mutations, together with mutations in *TET2*, account for approximately two thirds of CH (Challen and Goodell, 2020). *DNMT3A* is also recurrently mutated in hematologic malignancies (Yang et al., 2015) and

nonmalignant hematologic conditions, such as idiopathic multicentric Castleman disease (iMCD; Nagy et al., 2018). These observations prompted investigations into the role of *DNMT3A* in hematopoietic stem cells (HSCs), which found that complete *Dnmt3a* deficiency promotes self-renewal and clonal expansion of murine HSCs (Challen et al., 2011; Challen et al., 2014; Jeong et al., 2018).

While the majority of studies on *DNMT3A* have focused on characterizing *DNMT3A*- or *Dnmt3a*-null cells, complete loss of *DNMT3A* is rare in humans. CH-associated *DNMT3A* mutations are typically heterozygous, and the majority of them are single-copy deletions, truncating mutations, or missense mutations affecting functionally critical protein domains (Challen and Goodell, 2020). These observations indicate that haploinsufficiency is the predominant pathogenic mechanism associated with *DNMT3A* mutations. *DNMT3A*-mutated CH carriers do not display gross abnormalities in the number or composition of circulating blood cell subsets (Dawoud et al., 2020). As such, mature immune cells carrying *DNMT3A* mutations can be

<sup>1</sup>Precision Immunology Institute, Icahn School of Medicine at Mount Sinai, New York, NY; <sup>2</sup>Department of Medicine, University of California, San Diego, La Jolla, CA; <sup>3</sup>Department of Microbiology, Icahn School of Medicine at Mount Sinai, New York, NY; <sup>4</sup>Department of Pathology and Laboratory Medicine, The Children's Hospital of Philadelphia, Philadelphia, PA; <sup>5</sup>Department of Medicine, University of Pennsylvania, Philadelphia, PA; <sup>6</sup>Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, NY.

Correspondence to Minji Byun: [minji.byun@mssm.edu](mailto:minji.byun@mssm.edu).

© 2021 Lim et al. This article is distributed under the terms of an Attribution–Noncommercial–Share Alike–No Mirror Sites license for the first six months after the publication date (see <http://www.rupress.org/terms/>). After six months it is available under a Creative Commons License (Attribution–Noncommercial–Share Alike 4.0 International license, as described at <https://creativecommons.org/licenses/by-nc-sa/4.0/>).

detected readily in individuals with CH (Arends et al., 2018; Buscarlet et al., 2018), raising a question about the impact of DNMT3A haploinsufficiency on these cells.

Although DNMT3A haploinsufficiency has been studied in the context of hematopoietic malignancy (Cole et al., 2017; Haney et al., 2016) and neurodevelopmental impairments (Christian et al., 2020), its impact on mature immune cells has not been rigorously examined. Recent studies reported increased all-cause mortality and death due to atherosclerotic cardiovascular disease in CH carriers, including those with DNMT3A mutations (Jaiswal et al., 2014; Jaiswal et al., 2017). Corroborating this finding, CH with mutations in DNMT3A or TET2 is associated with poor outcomes in patients with heart failure (Dorsheimer et al., 2019). For TET2, these associations have been attributed to the proinflammatory nature of TET2-mutated myeloid cells based on murine studies (Fuster et al., 2017; Jaiswal et al., 2017; Sano et al., 2018a).

Although DNMT3A mutations are more common than TET2 mutations in CH, only limited characterization of DNMT3A-mutated myeloid cells has been reported to date (Leoni et al., 2017; Sano et al., 2018a). Importantly, the DNA methylation landscape of DNMT3A-haploinsufficient mature immune cells, which may provide mechanistic insights into the link between DNMT3A mutation-driven CH and the heightened risk of cardiovascular disease, remains poorly characterized. Given the central role that DNA methylation plays in the acquisition and reinforcement of cellular identity, we hypothesized that DNMT3A haploinsufficiency could cause functional impairments of mature immune cells through dysregulated DNA methylation. Here, we tested our hypothesis using genetically defined human macrophage models of DNMT3A haploinsufficiency.

## Results

### Biochemical and functional characterization of DNMT3A haploinsufficiency in patient-derived cells

We studied an adolescent patient diagnosed with iMCD (Hawkins and Pillai, 2015), a rare disorder characterized by episodes of sepsis-like systemic inflammation (Fajgenbaum, 2018). As part of the Castleman Genome Project (Baker et al., 2018), we performed whole-genome sequencing of DNA isolated from the whole blood of the patient and his unaffected parents. A single nucleotide substitution in DNMT3A was detected in the patient's blood DNA. This mutation is predicted to change the Trp297 residue located in the Pro-Trp-Trp-Pro (PWWP) domain to a stop (c.891G>A; p.W297\*; Fig. 1 A). Multiple lines of evidence suggest that this mutation arose somatically: (1) The patient's parents do not carry the mutation; (2) variant allele frequencies of the mutation in circulating blood cell subsets as well as in bone marrow and lymph nodes are between 0.05 and 0.1, far below the level expected for a heterozygous germline variant (0.5); and (3) the mutation was not detectable in hair roots (Fig. 1 B).

We hypothesized that the presumably heterozygous premature stop mutation would cause DNMT3A haploinsufficiency in the affected cells. To test this hypothesis using mutant and WT control cells with isogenic background, we reprogrammed the

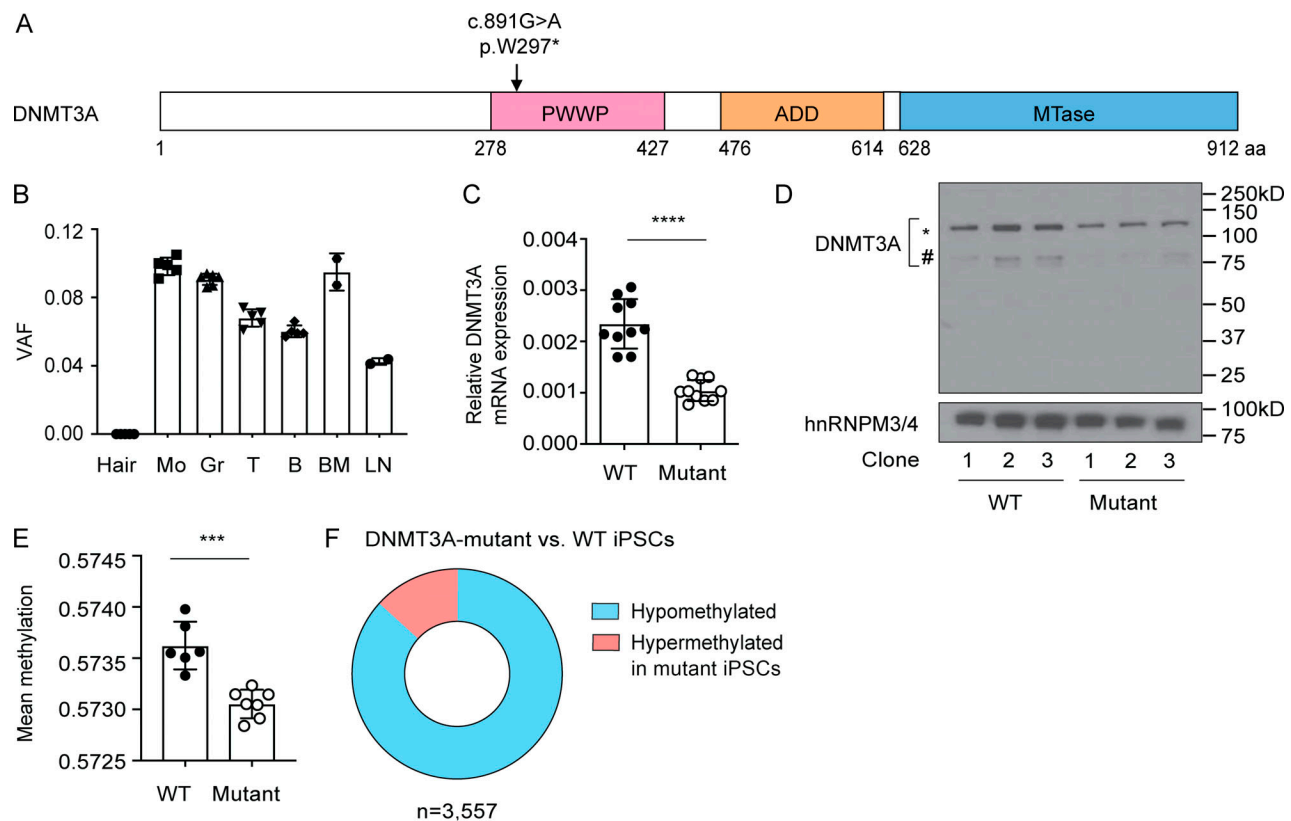
patient's peripheral blood mononuclear cells (PBMCs) to induced pluripotent stem cells (iPSCs). Among the 65 genotyped iPSC clones, 10 were heterozygotes (DNMT3A<sup>W297\*/WT</sup>) and 55 were WT (DNMT3A<sup>WT/WT</sup>). The frequency of mutant clones (15%) corresponded to that of mutation-positive cells in the PBMCs (10–20%) deduced from the variant allele frequency (0.05–0.1). Expression levels of the pluripotency markers were compatible between the mutant and WT iPSCs (Fig. S1). These data indicate that the DNMT3A mutation does not have a marked impact on the reprogramming efficiency or survival of iPSCs under standard culture conditions.

Premature stop mutations typically induce nonsense-mediated mRNA decay (Kervestin and Jacobson, 2012). Accordingly, the level of DNMT3A transcripts was significantly lower in the mutant iPSC clones compared with the WT clones (Fig. 1 C). Consistently, a reduction in the protein level affecting both DNMT3A1 and DNMT3A2 isoforms was observed in the mutant clones (Fig. 1 D). Notably, no bands corresponding to C-terminal-truncated proteins (expected at ~32 kD) were detected using antibodies raised against an N-terminal peptide of DNMT3A (Fig. 1 D). These data indicate that truncated DNMT3A proteins, if made, are not stable and rapidly degraded.

We next assessed the functional impact of the DNMT3A mutation with a microarray that measures DNA methylation levels at >850,000 methylation sites across the human genome (EPIC array; Pidsley et al., 2016). A modest, but reproducible reduction in the mean DNA methylation level was observed in the mutant iPSC clones compared with the WT controls (Fig. 1 E). More than 3,500 sites were significantly differentially methylated between the mutant and WT clones, the majority of which displayed lower methylation in the mutant clones relative to the WT controls (hypomethylated; Fig. 1 F). Together, these results demonstrated a biochemical and functional DNMT3A defect in patient-derived cells carrying a heterozygous truncating mutation in DNMT3A.

### Generation of human embryonic stem cell (hESC) models of DNMT3A haploinsufficiency

While iPSCs are valuable for basic assessment of DNMT3A haploinsufficiency in the native genetic background of patient-derived cells, they are not best suited for tracking cellular differentiation-associated changes in DNA methylation because cell-of-origin-specific DNA methylation patterns can remain in iPSCs (Kim et al., 2011). To interrogate the effect of DNMT3A haploinsufficiency in cells naive to prior differentiation, we generated DNMT3A-haploinsufficient hESCs by carrying out CRISPR-Cas9 gene editing. We chose the hESC line H1 (WA01, NIHhESC-10-0043) for our study because its epigenome has been comprehensively characterized (Bock et al., 2011; Moore et al., 2020; Kundaje et al., 2015). Guide RNAs targeting the PWWP domain were used to introduce frameshift mutations (Fig. 2 A). Three clones with distinct monoallelic frameshift mutations (DNMT3A<sup>+/-</sup>) were selected for downstream analyses (Fig. S2 A). Additionally, three clones with biallelic frameshift mutations (DNMT3A<sup>-/-</sup>) were selected to represent DNMT3A-null controls (Fig. S2 A). Three clones that had undergone the same targeting and selection processes without acquiring



**Figure 1. Biochemical and functional characterization of *DNMT3A* haploinsufficiency in patient-derived cells.** (A) Schematic diagram of the translated product of the *DNMT3A* isoform DNMT3A1 (ENST00000264709.7) with its functional domains. The somatic mutation identified in the iMCD patient is indicated. (B) Variant allele fraction (VAF) of the c.891G>A mutation in various cells and tissues from the patient was determined by digital droplet PCR. Each dot indicates an independent assay run. Error bars represent the SEM. (C) *DNMT3A* mRNA levels were determined by a TaqMan assay. GAPDH was used as an endogenous control. Each dot indicates an iPSC clone. Error bars represent the SEM. Data are representative of three independent experiments. \*\*\*\*,  $P < 0.0001$  by Student's  $t$  test. (D) *DNMT3A* protein levels were determined by immunoblotting using an antibody raised against an N-terminal peptide of *DNMT3A*. The \* and # symbols indicate translated products of *DNMT3A1* and *DNMT3A2* (ENST00000380746.8) isoforms, respectively. A nuclear protein, hnRNPM3/4, was used as a loading control. Three WT and three mutant iPSC clones were tested. Data are representative of three independent experiments. (E) Mean DNA methylation levels of six WT and seven mutant iPSC clones as determined by the EPIC array. A total of 836,845 sites were included in the analysis. Each dot indicates an iPSC clone. Error bars represent the SEM. \*\*\*,  $P < 0.001$  by Student's  $t$  test. (F) Proportions of methylation sites that were significantly hypo- or hypermethylated in mutant iPSCs compared with WT controls. n, the total number of differentially methylated sites. ADD, Atrx-Dnmt3l domain; B, B lymphocytes; BM, bone marrow; Gr, granulocytes; Hair, hair roots; Mo, monocytes; MTase, methyltransferase domain. T, T lymphocytes.

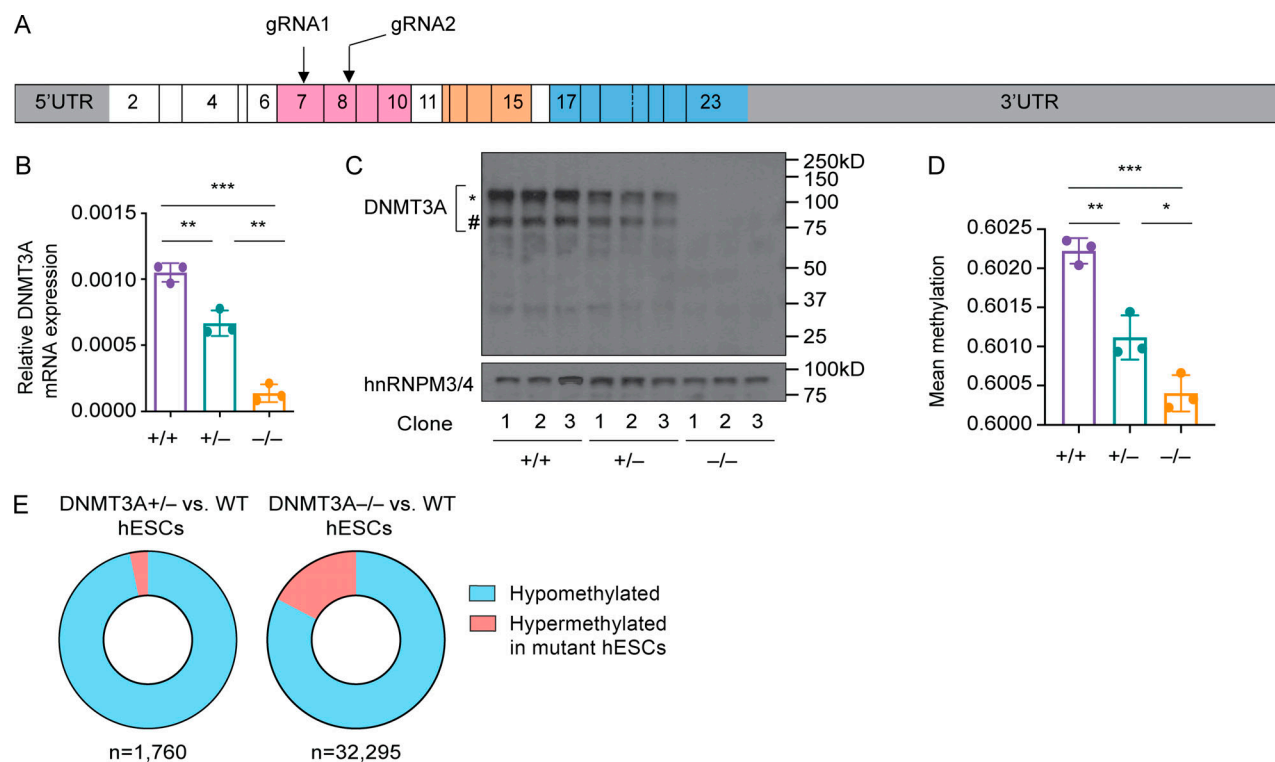
mutations in the targeted loci were used as WT controls (*DNMT3A*<sup>+/+</sup>) throughout the study.

The *DNMT3A*<sup>+/-</sup> and *DNMT3A*<sup>-/-</sup> hESC clones expressed similar levels of pluripotency markers compared with the WT clones (Fig. S2 B), consistent with a previous report (Liao et al., 2015). *DNMT3A* mRNA levels were significantly lower in the *DNMT3A*<sup>+/-</sup> and *DNMT3A*<sup>-/-</sup> clones relative to the WT controls, presumably due to nonsense-mediated mRNA decay (Fig. 2 B). Immunoblotting confirmed decreased and undetectable levels of *DNMT3A* proteins in the *DNMT3A*<sup>+/-</sup> and *DNMT3A*<sup>-/-</sup> clones, respectively (Fig. 2 C). DNA methylation analysis with the EPIC array demonstrated a *DNMT3A* dosage-dependent decrease in the global mean of DNA methylation (Fig. 2 D). The majority of differentially methylated sites exhibited lower methylation in the *DNMT3A*<sup>+/-</sup> or *DNMT3A*<sup>-/-</sup> clones than in the WT controls (Fig. 2 E). Together, these results recapitulate the observations in the patient-derived iPSCs and validate genome-edited hESCs as a robust model for *DNMT3A* haploinsufficiency.

### ***DNMT3A* haploinsufficiency causes dichotomous DNA methylation defects in mature macrophages**

To gain insights into the impact of *DNMT3A* haploinsufficiency in mature human immune cells, we employed hESC-derived macrophages as an experimental model. Macrophages were differentiated using an established embryoid body (EB)-based method with minor modifications (Buchrieser et al., 2017; van Wilgenburg et al., 2013). EBs were sequentially exposed to different cytokines that resemble hematopoietic specification and myeloid expansion, followed by directed terminal differentiation of nonadherent myeloid progenitors (MPs) to mature macrophages (Fig. 3 A). Expression of cell surface markers was examined in MPs and macrophages to confirm the mature phenotype of the latter (Fig. 3 B). Notably, the expression of *DNMT3A* is gradually increased over the course of differentiation, in contrast to the sharply diminished expression of *DNMT3B* (Fig. 3 C).

Using this method, we differentiated macrophages from WT, *DNMT3A*-haploinsufficient, and -null hESCs.



**Figure 2. Generation of hESC models of *DNMT3A* haploinsufficiency.** (A) Schematic diagram of the *DNMT3A1* transcript. Positions of the guide RNAs (gRNA) used for CRISPR-Cas9 targeting are indicated. Boxes indicate exons, and the numbers within the boxes are the exon numbers. Exons encoding the protein domains shown in Fig. 1 A are shaded with the matching colors. (B) *DNMT3A* mRNA levels in hESC clones with the indicated *DNMT3A* genotype were determined with a TaqMan assay. GAPDH was used as an endogenous control. Each dot indicates a clone. Error bars represent the SEM. Data are representative of five independent experiments. \*\*,  $P < 0.01$ ; \*\*\*,  $P < 0.001$  by Student's *t* test. (C) *DNMT3A* protein levels were assessed by immunoblotting using hnRNPM3/4 as a loading control. The \* and # symbols indicate *DNMT3A1* and *DNMT3A2* isoforms, respectively. Data are representative of three independent experiments. (D) Mean DNA methylation levels of WT (*DNMT3A*<sup>+/+</sup>), *DNMT3A*<sup>+/-</sup>, and *DNMT3A*<sup>-/-</sup> hESC clones were determined with the EPIC array. A total of 821,449 sites were included in the analysis. Each dot indicates a clone. Error bars represent the SEM. \*,  $P < 0.05$ ; \*\*,  $P < 0.01$ ; \*\*\*,  $P < 0.001$  by Student's *t* test. (E) Proportions of methylation sites that were significantly hypo- or hypermethylated in mutant hESCs compared with WT controls. *n*, the total number of differentially methylated sites. UTR, untranslated region.

*DNMT3A*-haploinsufficient and -null macrophages did not display gross defects in cell surface marker expression or the ability to uptake zymosan and modified low-density lipoproteins (Fig. S3, A and B). Yet, the EPIC array analysis revealed a large number of differentially methylated sites in these cells (Fig. 3 D and Fig. S3 C). Compared to the predominantly hypomethylated phenotype of *DNMT3A*-mutated hESCs (Fig. 2 E), it was remarkable that one third to one half of differentially methylated sites were hypermethylated in *DNMT3A*-mutated macrophages (Fig. 3 D and Fig. S3 C). To determine whether these methylation defects were established before or after the terminal differentiation to macrophages, we tested the effect of knocking down *DNMT3A* in fully differentiated macrophages. WT macrophages transfected with a pool of *DNMT3A*-targeting siRNAs demonstrated a significant and sustained reduction in *DNMT3A* transcripts and proteins for at least 5 d compared with macrophages transfected with nontargeting control siRNAs (Fig. 3 E). However, none of the sites examined by the EPIC array were significantly differentially methylated between *DNMT3A*-targeting siRNA-treated and control siRNA-treated macrophages. Consistently, their mean methylation levels across all

examined sites as well as the hypo- and hypermethylated sites in *DNMT3A*<sup>+/-</sup> hESC-derived macrophages were comparable (Fig. 3 F). These results refute the hypothesis that *DNMT3A* haploinsufficiency dynamically modulates DNA methylation in terminally differentiated macrophages and indicate that the methylation defects in *DNMT3A*-haploinsufficient macrophages are established during differentiation.

To determine at which differentiation stage the DNA methylation defects appear, we additionally profiled DNA methylation of WT and *DNMT3A*-haploinsufficient EBs and MPs. While the number of hypomethylated sites in *DNMT3A*-haploinsufficient cells increased gradually over the course of differentiation, the number of hypermethylated sites increased sharply between the MP and the macrophage stages (Fig. 3 G). Furthermore, hypomethylation affected sites that gain methylation during differentiation, whereas hypermethylation affected sites that lose methylation during MP-to-macrophage transition (Fig. 3 H). Together, these results indicated that DNA methylation defects in *DNMT3A*-haploinsufficient macrophages were tightly linked to differentiation-associated DNA methylation gains and losses.

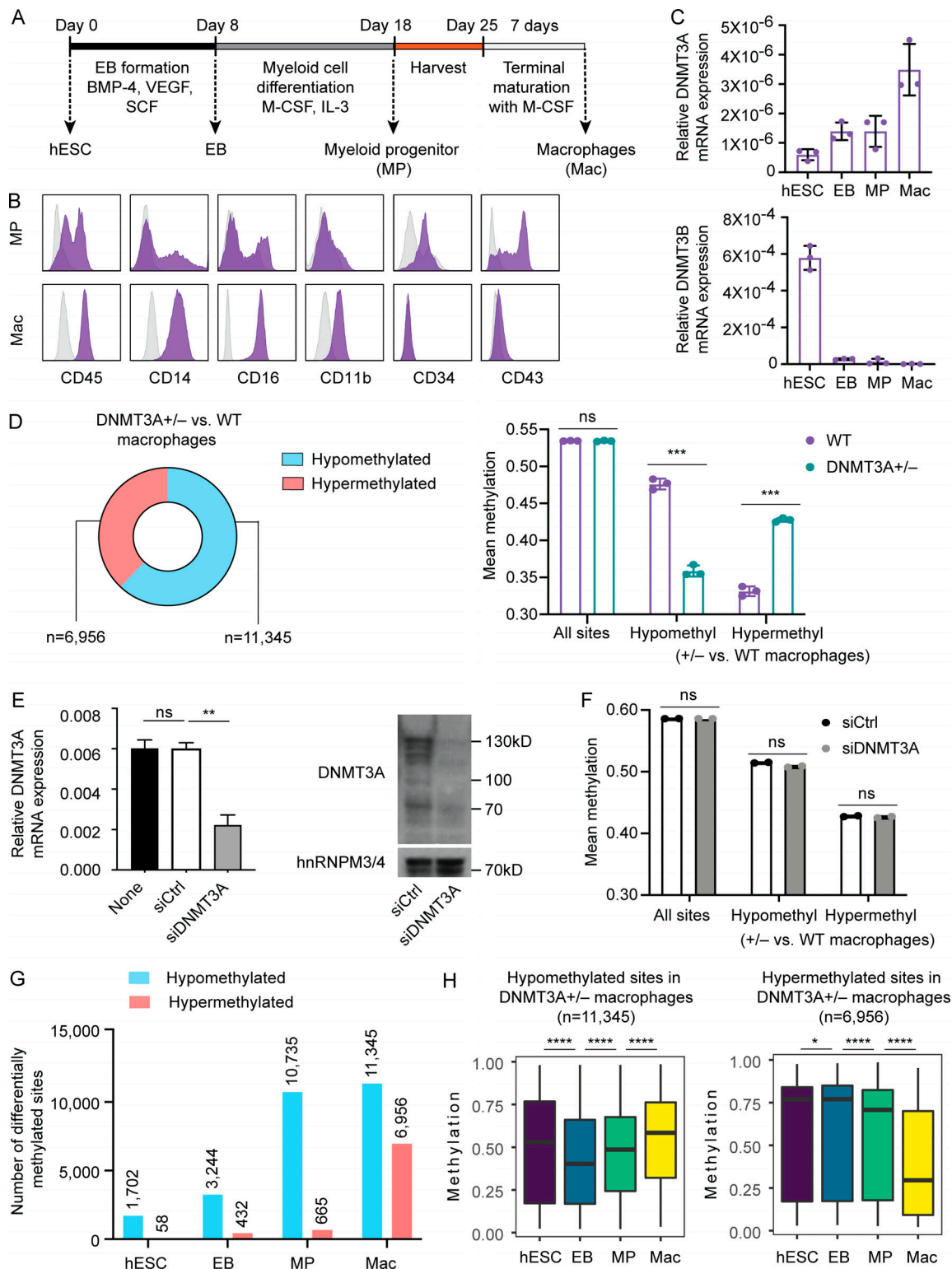


Figure 3. ***DNMT3A* haploinsufficiency causes dichotomous DNA methylation defects in mature macrophages.** (A) Schematic of the macrophage differentiation protocol. (B) WT H1 hESC-derived MPs harvested at day 18 and macrophages harvested at day 32 were examined for expression of the indicated cell surface markers with flow cytometry. Gray histograms indicate isotype controls. Data are representative of two independent experiments. (C) Relative mRNA levels of *DNMT3A* and *DNMT3B* at the indicated differentiation stages were assessed by TaqMan gene expression assays. Eukaryotic 18S ribosomal RNA was used as an endogenous control. Each dot indicates a WT hESC clone and cells derived from it. Error bars represent the SEM. Data are representative of two independent experiments. (D) DNA methylation of macrophages differentiated from three WT, three *DNMT3A*<sup>+/-</sup>, and three *DNMT3A*<sup>-/-</sup> hESC clones were determined with the EPIC array. Proportions and numbers of hypo- or hypermethylated sites (left) and mean methylation (right) in *DNMT3A*-haploinsufficient macrophages compared with WT controls are shown. Each dot in the bar graph indicates a clone. Error bars represent the SEM. ns, P > 0.05; \*\*\*, P < 0.001 by Student's *t* test. See Fig. S3 C for *DNMT3A*-null macrophage data. (E) *DNMT3A* levels in WT hESC-derived macrophages transfected with the indicated siRNAs

were measured with a TaqMan gene expression assay (left) and immunoblotting (right) 5 d after transfection. siCtrl, a pool of nontargeting siRNAs; siDNMT3A, a pool of DNMT3A-targeting siRNAs. GAPDH was used as an endogenous control for the TaqMan assay. hnRNP3/4 was used as a loading control for immunoblotting. Error bars represent the SEM. Data are representative of two independent experiments. ns,  $P > 0.05$ ; \*\*,  $P < 0.01$  by Student's  $t$  test. **(F)** DNA methylation in WT hESC-derived macrophages transfected with the indicated siRNAs for 5 d was assessed using the EPIC array. Mean methylation levels across the same methylation sites in D are shown. Each dot represents an independent transfection experiment. ns,  $P > 0.05$  by Student's  $t$  test. **(G)** The numbers of hypo- or hypermethylated sites in DNMT3A-haploinsufficient cells compared with their WT counterparts (e.g., DNMT3A-haploinsufficient EB vs. WT EB) are shown. **(H)** Methylation levels of the hypo- or hypermethylated sites, as defined in D, in hESCs, EBs, MPs, and macrophages (all WT). \*,  $P < 0.05$ ; \*\*\*\*,  $P < 0.0001$  by paired  $t$  test. BMP-4, bone morphogenetic protein 4; M-CSF, macrophage colony-stimulating factor; SCF, stem cell factor; VEGF, vascular endothelial growth factor.

### Cell-type-specific enhancers are abnormally methylated in DNMT3A-haploinsufficient macrophages

Differentiation-associated DNA methylation gains and losses occur primarily at enhancers (Luo et al., 2018; Stadler et al., 2011). While the EPIC array provides comprehensive coverage of promoter regions, its coverage of enhancers is sparse (Pidsley et al., 2016). To obtain a single basepair resolution methylation map of all regulatory elements, including enhancers, we performed whole-genome bisulfite sequencing (WGBS) of WT, DNMT3A-haploinsufficient, and DNMT3A-null hESC-derived macrophages. More than 23 million CpG sites covered at least five times in each of all nine samples (three independent clones per genotype) were included in the differential methylation analysis, representing a 27-fold increase over the total number of methylation sites analyzed by the EPIC array.

The number of hypomethylated CpG sites and the degree of methylation loss were larger in DNMT3A-null macrophages than DNMT3A-haploinsufficient macrophages (Fig. 4 A). Highly methylated sites were the main target of hypomethylation in both DNMT3A-haploinsufficient and -null macrophages (Fig. 4 A). In comparison, while the number of hypermethylated sites and the degree of methylation gain were similar between DNMT3A-haploinsufficient and DNMT3A-null macrophages, lowly methylated sites were more frequently hypermethylated in DNMT3A-haploinsufficient macrophages than in DNMT3A-null macrophages (Fig. 4 A).

In agreement with the EPIC array data (Fig. 3 H), CpGs that gain methylation upon macrophage differentiation (macrophage-gain sites: methylation in macrophages > methylation in hESCs) were hypomethylated in DNMT3A-haploinsufficient and -null macrophages, whereas CpGs that lose methylation upon differentiation (macrophage-loss sites: methylation in macrophages < methylation in hESCs) were hypermethylated (Fig. 4 B). To dissect this phenotype further, we investigated whether the differentially methylated sites were associated with cell-type-specific chromatin states. Because uniformly processed macrophage chromatin states were not available, chromatin states of primary blood CD14<sup>+</sup> monocytes were used as a proxy (Kundaje et al., 2015). The hypermethylated sites in DNMT3A-haploinsufficient macrophages were enriched with enhancer-like chromatin states (TssAFlnk, Enh, TxFlnk, EnhG) of monocytes but not those of hESCs (Fig. 4, C and D; and Fig. S4, A and B). Consistently, the binding sites of transcription factors (TFs) active in macrophages, such as PU.1 and RUNX, were enriched near ( $\pm 100$  bp) the hypermethylated sites (Fig. 4 E). In comparison, the hypomethylated sites were enriched with enhancer-like chromatin states of hESCs (Fig. 4, C and D; and Fig. S4, A and B) and were

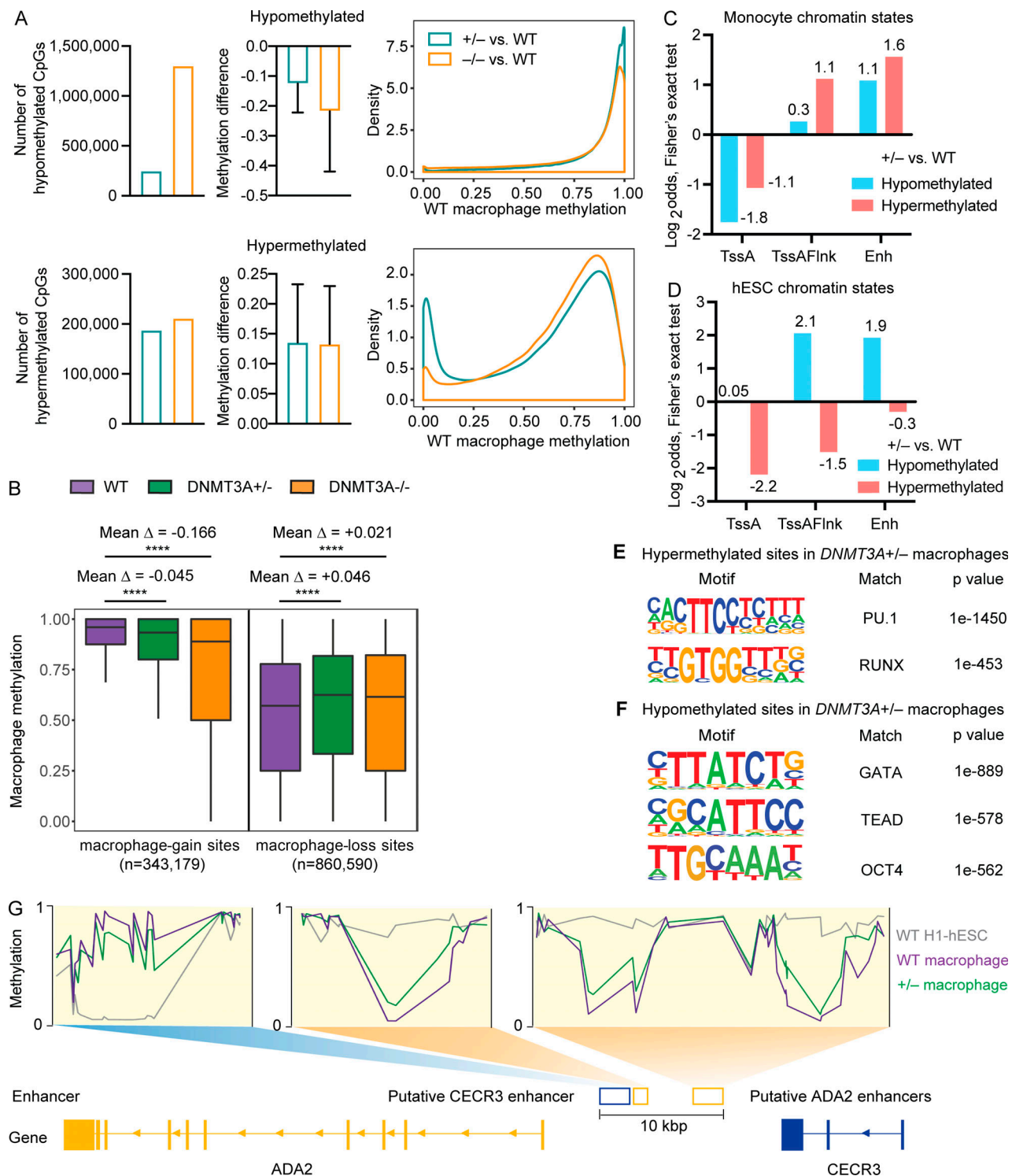
located in proximity to the binding sites of TFs active in ESCs and early hematopoietic progenitors, such as OCT4, TEAD, and GATA (Fig. 4 F).

An intergenic region between *ADA2* and *CECR3* on chromosome 22, containing several putative enhancers annotated by ENCODE (Moore et al., 2020) and GeneHancer (Fishilevich et al., 2017), exemplifies these findings (Fig. 4 G). *ADA2*, encoding adenosine deaminase 2, is expressed exclusively in myeloid cells, such as monocytes and macrophages (Iwaki-Egawa et al., 2006; Zavialov et al., 2010), whereas *CECR3* encodes a long noncoding RNA of unknown function with undetectable expression in myeloid cells (Choi et al., 2019). DNMT3A-haploinsufficient macrophages display hypomethylation compared with WT macrophages at the putative *CECR3* enhancer, which gains methylation during macrophage differentiation (Fig. 4 G, left). Conversely, they display hypermethylation at the putative *ADA2* enhancers, which undergo demethylation during differentiation (Fig. 4 G, center and right).

### DNMT3A defects cause DNA hypermethylation at cell-type-specific enhancers in primary myeloid cells and hESC-derived neurons

An overgrowth intellectual disability syndrome known as Tatton-Brown-Rahman syndrome (TBRS; Online Mendelian Inheritance in Man accession no. 615879) is caused by germline heterozygous mutations in *DNMT3A* (Tatton-Brown et al., 2014). The spectrum of TBRS-associated *DNMT3A* variants overlaps with that of CH-associated variants (Shen et al., 2017), encompassing missense variants affecting functional domains, truncating variants, and the dominant-negative Arg882 variants (p.Arg882His and p.Arg882Cys; Russler-Germain et al., 2014). We thus hypothesized that mature immune cells from a TBRS patient would display DNA methylation defects similar to what we observed in DNMT3A-mutated hESC-derived macrophages.

A previous study examined DNA methylation of monocytes and neutrophils from a 9-yr-old TBRS patient with a p.Arg882His mutation and his unaffected 13-yr-old sibling who is WT for *DNMT3A* (Spencer et al., 2017). The same study examined DNA methylation of CD34<sup>+</sup> HSCs and progenitor cells, monocytes, and neutrophils isolated from healthy donors (Spencer et al., 2017). We used these publicly available WGBS datasets to identify CpGs that gain methylation during differentiation to monocytes (monocyte-gain sites: methylation in monocytes > methylation in CD34<sup>+</sup> cells) and those that lose methylation (monocyte-loss sites: methylation in monocytes < methylation in CD34<sup>+</sup> cells). As expected, monocyte-loss sites were enriched with monocyte enhancers and binding sites of TFs important for



**Figure 4. Cell-type-specific enhancers are abnormally methylated in DNMT3A-haploinsufficient macrophages.** (A) WGBS was performed for macrophages differentiated from three WT, three DNMT3A<sup>+/-</sup> and three DNMT3A<sup>-/-</sup> hESC clones. The number of significantly differentially methylated CpGs (left), the average difference in the methylation level (middle), and the distribution of WT macrophage methylation levels at hypo- and hypermethylated sites in DNMT3A<sup>+/-</sup> and DNMT3A<sup>-/-</sup> macrophages are shown. Error bars represent the SEM. (B) Methylation levels of WT, DNMT3A<sup>+/-</sup>, and DNMT3A<sup>-/-</sup> macrophages are compared at macrophage-gain sites and macrophage-loss sites. Mean  $\Delta$ , mean differences in the methylation level. \*\*\*\*,  $P < 0.0001$  by paired t test. (C) Two-sided Fisher's exact test results (Log<sub>2</sub>odds) for enrichment/depletion of monocyte chromatin states in hypo- and hypermethylated sites in DNMT3A<sup>+/-</sup> macrophage are shown ( $P < 0.05$  for all tests). TssA, active TSS; TssAFlnk, flanking active TSS; Enh, enhancers. Actual values used for plotting are shown near each column. (D) Same as C for hESC chromatin states. (E) Top two significant TF binding motifs enriched near ( $\pm 100$  bp) the hypermethylated sites in DNMT3A<sup>+/-</sup> macrophage. (F) Top three significant TF binding motifs enriched near ( $\pm 100$  bp) the hypomethylated sites in DNMT3A<sup>+/-</sup> macrophage. (G) Schematic of a locus on chromosome 22 containing two genes, ADA2 (chr22:17,179,304–17,221,989) and CECR3 (chr22:17,256,859–17,266,733). DNA

methylation levels at the three intergenic enhancers in WT and *DNMT3A*<sup>+/-</sup> macrophages as well as WT H1-hESCs are shown. Putative *CECR3* enhancer: GeneHancer GH22J017233 (chr22:17,233,077–17,236,254). Putative *ADA2* enhancers: GeneHancer GH22J017236 (chr22:17,236,596–17,237,929) and GeneHancer GH22J017241 (chr22:17,241,402–17,244,449). All coordinates are from the GRCh38 assembly.

monocyte differentiation (Fig. 5, A and B). Comparison of the TBRS patient's monocytes and his sibling's monocytes revealed a modest overall hypomethylation in the patient's cells (Fig. 5 C, left), in agreement with the previous report (Spencer et al., 2017). A focused analysis of monocyte-gain sites also showed hypomethylation in the patient's monocytes (Fig. 5 C, center). In contrast, at the monocyte-loss sites, the patient's monocytes displayed significant hypermethylation compared with the sibling's monocytes (Fig. 5 C, right). Similarly, the TBRS patient's neutrophils displayed higher methylation levels than his sibling's neutrophils at neutrophil-loss sites, which are enriched with neutrophil enhancers and binding sites of TFs important for neutrophil differentiation (Fig. 5, D–F). Together, these results provide evidence that *DNMT3A* mutations cause DNA hypermethylation at cell-type-specific enhancers in primary myeloid cells.

We wondered whether this phenomenon was specific to the hematopoietic lineage or could affect other cell lineages in a similar manner. A previous study of human motor neurons differentiated from WT and *DNMT3A*-null hESCs provided an opportunity to examine this issue (Ziller et al., 2018). Notably, this study used another hESC line, HUES64, and clones derived from it (Ziller et al., 2018). By analyzing publicly available WGBS data from this study (National Center for Biotechnology Information Gene Expression Omnibus [NCBI GEO] accession no. GSE90553), we identified CpGs that gain methylation (neuron-gain sites: methylation in motor neurons > methylation in hESCs) during hESC-to-motor neuron differentiation and those that lose methylation (neuron-loss sites: methylation in motor neurons < methylation in hESCs). We then compared the methylation levels of WT and *DNMT3A*-null hESC-derived motor neurons at these sites. *DNMT3A*-null motor neurons displayed significantly lower methylation at neuron-gain sites than WT controls (Fig. 5 G). In contrast, they displayed significantly higher methylation at neuron-loss sites (Fig. 5 G). Moreover, we observed enrichment of neuron TF binding sites near the hypermethylated sites and ESC TF binding sites near the hypomethylated sites (Fig. 5 H). These results indicate that *DNMT3A* defects cause similar DNA methylation defects across at least two independent cell lineages (hematopoietic and neuronal) and support an unexpected role for *DNMT3A* in promoting DNA demethylation at cell-type-specific enhancers during cellular differentiation.

### DNA methylation defects alter enhancer activity

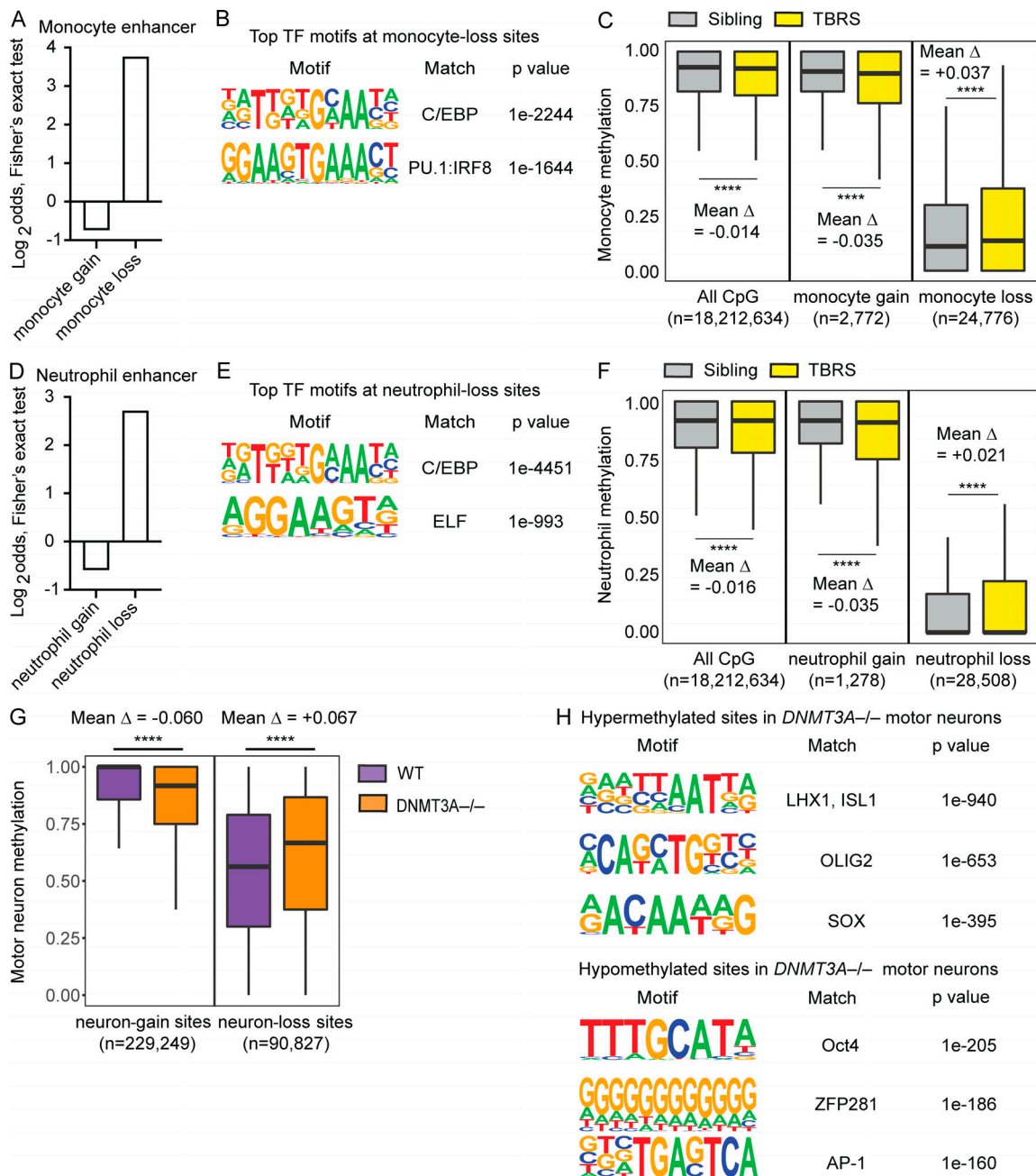
DNA methylation at CpG-rich promoters is associated with repression of gene expression (Deaton and Bird, 2011; Smith and Meissner, 2013). In comparison, the impact of DNA methylation on the activity of enhancers, which typically have low CpG density, is less well understood. We thus investigated whether the DNA methylation defects at enhancers in *DNMT3A*-haploinsufficient macrophages were associated with altered

enhancer activity. The presence of enhancer RNAs (eRNAs) is a specific indicator of enhancer activity, outperforming predictions based on histone modification (Andersson et al., 2014; Tippens et al., 2020). Hence enhancer activity can be assessed by quantifying eRNAs using a method such as capped small RNA sequencing (csRNA-seq; Duttke et al., 2019). The csRNA-seq is based on the observation that RNA polymerase II initially creates a 20–60-nucleotide RNA with an m7G-5' cap structure before stalling during transcription initiation (Gu et al., 2012; Nechaev et al., 2010). By isolating short RNAs from total RNA using size selection and enriching for 5'-capped RNAs, all transcription start sites (TSSs), including those from eRNAs and other rapidly processed or degraded transcripts, can be mapped genome wide and quantified (Duttke et al., 2019).

Application of the csRNA-seq to hESC-derived macrophages identified TSS clusters predominantly corresponding to TssA (promoters), TssAFlnk (flanking promoters), or Enh (enhancer) chromatin states of monocytes (Fig. 6 A; Kundaje et al., 2015), confirming that the csRNA-seq could robustly detect transcription initiation events at regulatory chromatin regions. We grouped the csRNA-seq TSS clusters into four bins based on the degree and the direction of DNA methylation changes during hESC-to-macrophage differentiation (Fig. 6 B): two groups that lose DNA methylation during differentiation (bin1 and bin2), a group with minimal methylation changes (bin3), and another that gains methylation during differentiation (bin4). Chromatin state analyses indicated that bin1 and bin2 were enriched with enhancers, whereas bin3 was enriched with promoters (Fig. 6 C). Roughly one half of detectable TSS clusters were members of bin1 or bin2, and the vast majority of the other half were members of bin3. Less than 2% of total TSS clusters fell into bin4 (Fig. 6 D).

Consistent with the observation made with genome-wide CpGs (Fig. 4 B), TSS clusters containing CpG sites that lose methylation during differentiation (bin1 and bin2) were hypermethylated in *DNMT3A*-haploinsufficient macrophages, and those that gain methylation during differentiation (bin4) were hypomethylated (Fig. 6 E). Because TSS clusters belonging to bin1 or bin2 far outnumber those belonging to bin4 (Fig. 6 D), these data indicate that the majority of transcriptionally active enhancers are hypermethylated in *DNMT3A*-haploinsufficient macrophages compared with WT macrophages. This observation underscores the importance of considering cell-type-specific activity of regulatory elements in determining cell-type-specific downstream effects of DNA methylation defects.

We next investigated the relationship between DNA methylation and enhancer activity. There were more hypermethylated TSS clusters (based on the average methylation level of all CpGs within a TSS cluster) with reduced activity than those with enhanced activity (bin1; Fig. 6 F). Conversely, there were more hypomethylated TSS clusters with enhanced activity

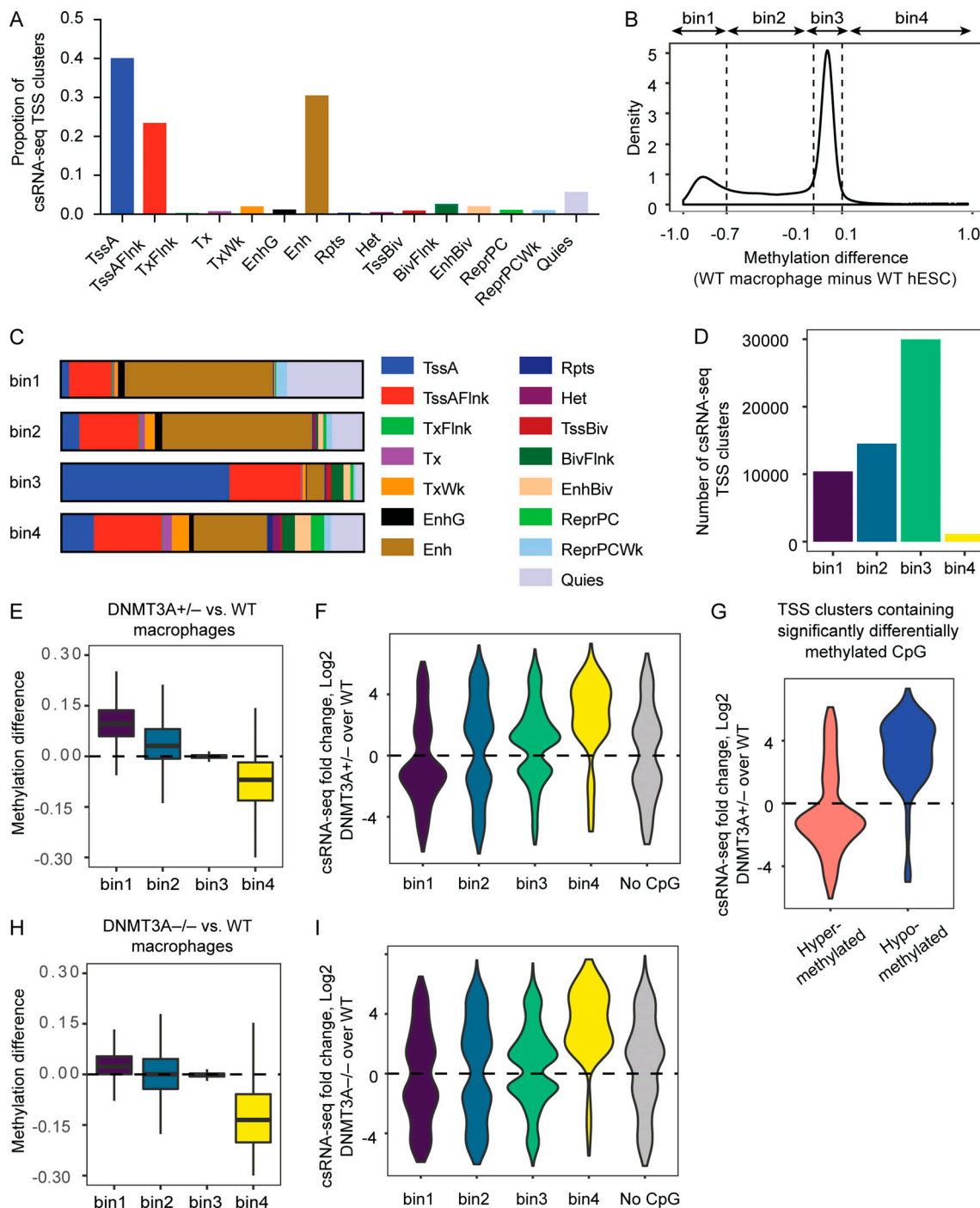


**Figure 5. DNMT3A defects cause DNA hypermethylation at cell-type-specific enhancers in primary myeloid cells and hESC-derived neurons.** Results of analyzing publicly available datasets: A–F (Spencer et al., 2017) and G and H (Ziller et al., 2018). (A) Two-sided Fisher's exact test results (Log<sub>2</sub>odds) for enrichment/depletion of the monocyte enhancer chromatin state in monocyte-gain and monocyte-loss sites ( $P < 0.05$  for both tests). (B) Top two significant TF binding motifs enriched near ( $\pm 100$  bp) the monocyte-loss sites. (C) Methylation levels of monocytes from a TBRS patient and his sibling are compared at all examined CpG, monocyte-gain, and monocyte-loss sites. The number of CpG sites included in each analysis and mean differences in the methylation level ( $\Delta$ ) are shown. \*\*\*\*,  $P < 0.0001$  by paired t test. (D) Same as A for neutrophils. (E) Top two significant TF binding motifs enriched near ( $\pm 100$  bp) the neutrophil-loss sites. (F) Same as C for neutrophils. (G) Methylation levels of motor neurons differentiated from WT or DNMT3A<sup>-/-</sup> HUES64-hESCs are compared at neuron-gain and neuron-loss sites. The number of CpG sites included in each analysis and mean differences in the methylation level ( $\Delta$ ) are shown. \*\*\*\*,  $P < 0.0001$  by paired t test. (H) Top three significant TF binding motifs enriched near ( $\pm 100$  bp) the hyper- or hypomethylated sites in DNMT3A<sup>-/-</sup> motor neurons.

than those with reduced activity (bin4; Fig. 6 F). Since each csRNA-seq TSS cluster ( $\sim 150$  bp) can contain several CpGs that are independently regulated, we additionally examined whether csRNA-seq TSS clusters containing at least one significantly hyper- or hypomethylated CpG displayed differential activity. Consistent with the results based on the average methylation,

we observed an inverse relationship between CpG methylation and enhancer activity (Fig. 6 G).

Compared with DNMT3A-haploinsufficient macrophages, DNMT3A-null macrophages demonstrated milder hypermethylation of the bin1 and bin2 TSS clusters and more pronounced hypomethylation of the bin4 TSS clusters (Fig. 6 H).



**Figure 6. DNA methylation defects alter enhancer activity.** (A) Proportions of csRNA-seq TSS clusters in hESC-derived macrophage corresponding to indicated monocyte chromatin states. (B) DNA methylation levels of WT hESCs and WT hESC-derived macrophages were determined for each csRNA-seq TSS cluster (average of all CpGs within a given TSS cluster), and their differential is plotted as a density plot. (C) csRNA-seq TSS clusters grouped as in B were annotated based on monocyte chromatin states. (D) The number of csRNA-seq TSS clusters belonging to each bin is shown. (E) Differences in the methylation levels between DNMT3A-haploinsufficient macrophages and WT controls for each TSS-cluster bin. Positive y-axis values indicate hypermethylation in DNMT3A-haploinsufficient macrophages. (F) csRNA-seq TSS clusters differentially expressed ( $P < 0.05$ ) in DNMT3A-haploinsufficient macrophages compared with WT macrophages are grouped based on their bin membership. The No CpG group represents TSS clusters devoid of CpGs. Positive y-axis values indicate increased expression/activity. (G) csRNA-seq fold changes of TSS clusters containing at least one CpG site significantly hyper- or hypomethylated in DNMT3A-haploinsufficient macrophages are plotted, regardless of their bin membership. (H) Same as E for DNMT3A-null macrophages. (I) Same as F for DNMT3A-null macrophages. BivFlnk, flanking bivalent TSS/Enh; EnhBiv, bivalent enhancer; EnhG, genic enhancers; Het, heterochromatin; Quies, quiescent; ReprPC, repressed polycomb; ReprPCWk, weak repressed polycomb; Rpts, ZNF genes and repeats; TssBiv, bivalent/poised TSS; TxFlnk, transcribed at gene 5' and 3'; Tx, strong transcription; TxWk, weak transcription.

Considering that absolute methylation levels at TSSs are typically low, this result was consistent with the reduced representation of lowly methylated sites among hypermethylated sites in *DNMT3A*-null macrophages compared with *DNMT3A*-haploinsufficient macrophages (Fig. 4 A). Consistently, only a modest bias toward reduced activity was observed for the bin1 TSS clusters, whereas more exaggerated enhancement of activity was observed for the bin4 clusters (Fig. 6 I). Notably, these results provide an example in which a phenotype of *DNMT3A*-haploinsufficient cells cannot be readily deduced from that of *DNMT3A*-null cells, cautioning against the use of gene ablation models for the investigation of haploinsufficiency-driven pathologies.

### ***DNMT3A* haploinsufficiency causes complex alterations in gene expression**

Based on the dichotomous DNA methylation defects at thousands of enhancers and the associated alterations in enhancer activity, we anticipated that *DNMT3A* haploinsufficiency would have a complex impact on gene expression. Transcriptomic analyses revealed only a modest correlation between differential enhancer activity and differential expression of genes nearest to the enhancer (Fig. S5 A). This is not surprising, as (1) enhancer action often occurs over a long genomic distance and, therefore, the nearest-gene approach is suboptimal for linking an enhancer with its target genes (Marsman and Horsfield, 2012), and (2) factors other than transcription initiation events (measured by csRNA-seq), such as transcript degradation rates, also contribute to the steady-state transcript level (measured by RNA-seq). Nevertheless, we identified several loci, including the *TLR5* locus in chromosome 1, that exhibit a correlation among enhancer DNA hypermethylation, decreased enhancer activity, and reduced gene expression (Fig. S5 B).

The elevated risk of atherosclerotic cardiovascular disease in *DNMT3A*-mutated CH carriers (Jaiswal et al., 2017) prompted us to examine whether *DNMT3A* haploinsufficiency alters the expression of inflammatory response genes. Dozens of genes involved in myeloid leukocyte activation (Fig. S5 C) were among the hundreds of genes differentially expressed between *DNMT3A*-haploinsufficient macrophages and WT controls at steady state (Fig. 7 A). Expression levels of several genes encoding chemokines, such as CXCL family members and S100A8/A9 heterodimer, were elevated in *DNMT3A*-haploinsufficient macrophages compared with WT controls (Fig. 7 A). Simultaneously, expression levels of genes encoding essential immune signaling molecules, such as TLR family members, phosphatidylinositol 3-kinases, tyrosine kinases, and ADA2, whose loss-of-function mutations are associated with a systemic autoinflammatory syndrome (Navon Elkan et al., 2014; Zhou et al., 2014), were reduced in *DNMT3A*-haploinsufficient macrophages (Fig. 7 A).

To test whether altered expression of these genes could affect an immune response, we focused on *TLR4*, one of the TLR genes underexpressed in *DNMT3A*-haploinsufficient macrophages (Fig. 7 A). *TLR4* is critical for the recognition of LPS, which induces a robust inflammatory response in macrophages (Medzhitov, 2001). Accordingly, thousands of genes were induced or repressed in WT macrophages upon LPS stimulation

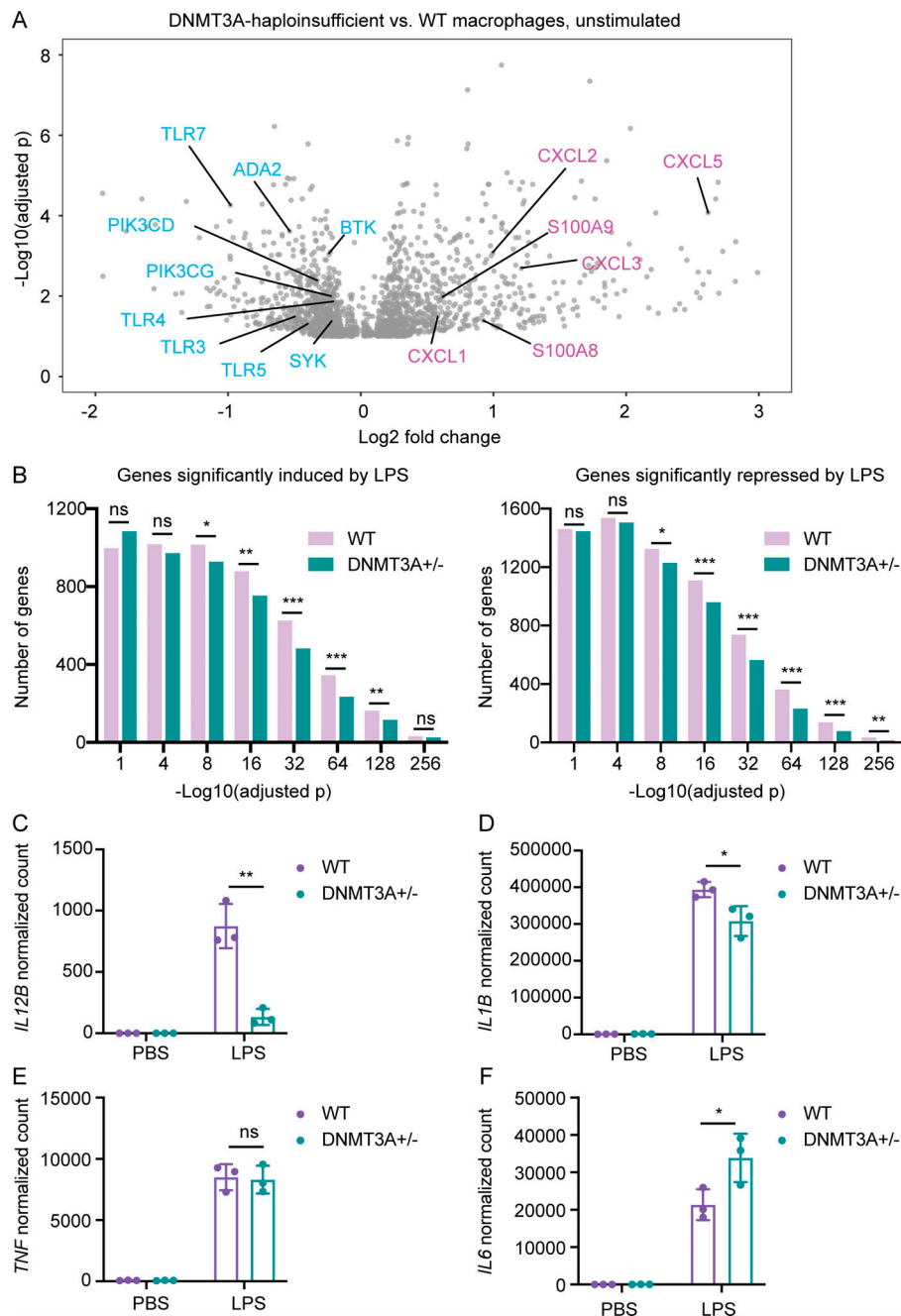
(Fig. 7 B). In comparison, the strength of overall LPS response was diminished in *DNMT3A*-haploinsufficient macrophages as evidenced by reduced numbers of genes significantly induced or repressed by LPS (Fig. 7 B and Fig. S5 D). While expression levels of some LPS-induced cytokine genes, such as *IL12B* and *IL1B*, were lower in *DNMT3A*-haploinsufficient macrophages compared with controls, expression levels of others, such as *TNF* and *IL6*, were not significantly different or even higher in *DNMT3A*-haploinsufficient macrophages (Fig. 7, C–F). Together, these data demonstrate that *DNMT3A* haploinsufficiency is associated with complex alterations in gene expression, and it can cause immune dysfunction in a context-specific manner.

### **Discussion**

DNA methylation defects involving both hypo- and hypermethylation have been observed in cells with complete deficiency of human *DNMT3A* or its murine orthologue *Dnmt3a* (Challen et al., 2011; Jeong et al., 2018; Jeong et al., 2014; Liao et al., 2015; Yagi et al., 2020; Ziller et al., 2018) as well as in *Dnmt3a*-haploinsufficient murine bone marrow cells, blood cells from a TBRS patient, and B lymphoblastoid cell lines from a germline mosaic *DNMT3A* R771Q mutation carrier (Cole et al., 2017; Spencer et al., 2017; Tovy et al., 2020). These studies mainly focused on characterizing the scope and consequence of DNA hypomethylation. Comparatively, the DNA hypermethylation phenotype received little attention, in part because *DNMT3A* has been best characterized as a DNA methyltransferase, and, therefore, hypermethylation in *DNMT3A*-deficient cells is unexpected. While it has been speculated that hypermethylation may be caused by a compensatory mechanism in the complete absence of *DNMT3A/Dnmt3a* or a consequence of malignant transformation, our data demonstrate that DNA hypermethylation can occur in *DNMT3A*-haploinsufficient cells in the absence of malignant transformation.

Based on previous works as well as data reported in this study, we propose that *DNMT3A* plays a dual role in remodeling DNA methylation at enhancers during cell differentiation by (1) performing de novo DNA methylation of decommissioned stem cell/progenitor enhancers and (2) facilitating DNA demethylation of newly activated mature cell enhancers. *DNMT3A* haploinsufficiency impairs both these processes, resulting in hypomethylation of stem cell/progenitor enhancers and hypermethylation of mature cell enhancers. These DNA methylation defects suppress (hypermethylation) or enhance (hypomethylation) the activity of enhancers, ultimately driving complexly altered gene expression and functional impairment of the affected mature immune cells.

Notably, our data identify a previously unrecognized role of *DNMT3A* in facilitating active DNA demethylation. Emerging evidence nominates TET2 as the key enzyme that carries out DNA demethylation of enhancers (Hon et al., 2014; Rasmussen et al., 2019; Sardina et al., 2018). While *DNMT3* and TET family members have been shown to compete at somatic enhancers in pluripotent cells (Charlton et al., 2020), they may have a cooperative interaction during cellular differentiation. It has been shown that *DNMT3A* is recruited to active enhancers in human



**Figure 7. DNMT3A haploinsufficiency causes complex alterations in gene expression.** (A) A volcano plot showing genes differentially expressed between DNMT3A-haploinsufficient macrophages and WT controls at steady state. Each gray dot represents a gene with adjusted  $P < 0.1$ . Representative genes overexpressed (pink) or underexpressed (blue) in DNMT3A-haploinsufficient macrophages compared with WT controls are labeled. Results are based on a pooled analysis of two independent RNA-seq experiments. Macrophages derived from three independent DNMT3A<sup>+/-</sup> and three WT hESC clones were included in each RNA-seq experiment. (B) Macrophages stimulated with either PBS or LPS for 4 h were subjected to RNA-seq. The numbers of significantly (adjusted  $P < 0.1$ ) LPS-induced or LPS-repressed genes are plotted in histogram plots. Results are based on a pooled analysis of two independent RNA-seq experiments. ns,  $P > 0.05$ ; \*,  $P < 0.05$ ; \*\*,  $P < 0.01$ ; \*\*\*,  $P < 0.001$  by Fisher's exact test. (C) Normalized counts of *IL12B* transcripts in hESC-derived macrophages stimulated with PBS or LPS for 4 h. Macrophages differentiated from three DNMT3A<sup>+/-</sup> and three WT hESC clones were included in the RNA-seq experiment. Each dot in the bar plot indicates a clone. A representative result of two independent RNA-seq experiments is shown. Error bars represent the SEM. \*\*,  $P < 0.01$  by Student's *t* test. (D) Same as C for *IL1B*. \*,  $P < 0.05$  by Student's *t* test. (E) Same as C for *TNF*. ns,  $P > 0.05$  by Student's *t* test. (F) Same as C for *IL6*. \*,  $P < 0.05$  by Student's *t* test.

epidermal stem cells, and DNMT3A knockdown results in a specific reduction in 5-hydroxymethylcytosine levels at active enhancers (Rinaldi et al., 2016). As 5-hydroxymethylcytosine is an essential intermediate of TET2-mediated DNA demethylation (Wu and Zhang, 2017), these data suggest that DNMT3A may assist in TET2-mediated conversion of 5-methylcytosine to 5-hydroxymethylcytosine and, ultimately, demethylation. Supporting this scenario, knockdown of TET2 increases DNA methylation uniquely at the DNMT3A-bound enhancers in human epidermal stem cells (Rinaldi et al., 2016). While future studies are needed to determine the precise interplay between DNMT3A and TET2 in enhancer DNA demethylation, it is noteworthy that DNMT3A and TET2 are frequently mutated in CH in a not mutually exclusive manner (Genovese et al., 2014;

Jaiswal et al., 2014; Xie et al., 2014). Similar to DNMT3A, TET2 mutations associated with CH are loss-of-function mutations, raising the question of how loss-of-function mutations in genes encoding proteins with seemingly antagonistic enzymatic activity lead to similar clinical outcomes. The dependency of TET2-mediated DNA demethylation on DNMT3A can provide an answer to this puzzling question.

Understanding whether and how DNMT3A haploinsufficiency impacts function of the affected immune cells has high public health significance because of (1) the high prevalence of DNMT3A mutations in age-related CH and the world's rapidly aging population and (2) the association of CH with increased mortality and atherosclerotic cardiovascular disease (Jaiswal et al., 2014; Jaiswal et al., 2017). While prior studies reported that

*Tet2*-deficient murine macrophages are hyperinflammatory (Fuster et al., 2017; Jaiswal et al., 2017; Sano et al., 2018a; Sano et al., 2018b), a study of germline *TET2* mutation carriers did not find evidence of enhanced proinflammatory cytokine and chemokine production by *TET2*-haploinsufficient human macrophages (Kaasinen et al., 2019). Our data suggest that *DNMT3A* haploinsufficiency alters gene expression in a complex manner, involving both over- and underexpression of many inflammatory response genes. This complexity, which is consistent with the essential role of *DNMT3A* for remodeling DNA methylation at thousands of enhancers during cellular differentiation, cautions against drawing conclusions based on results obtained under limited experimental conditions. Further investigations are needed to determine whether CH-associated *TET2* and *DNMT3A* mutations promote a hyperinflammatory state of primary human myeloid cells.

While hESC-derived macrophages provide a robust genetically defined model system for molecular mechanistic studies, significant differences in cell ontology and gene expression profiles exist between these cells and primary human myeloid cells. This is an important limitation of our study. Nevertheless, the findings reported here can serve as the foundation for the epigenetic, transcriptomic, and functional investigation of primary mature immune cells with *DNMT3A* mutations. The discovery of DNA hypermethylation at cell-type-specific enhancers in primary myeloid cells from a TBRS patient is one such example. We anticipate that follow-up studies of primary mature immune cells from CH carriers will further elucidate the mechanism of immune dysregulation associated with aging and acquired mutations in *DNMT3A* as well as other CH-driver genes.

## Materials and methods

### Whole-genome sequencing and analysis

Whole-genome sequencing and analysis was performed as previously described (Baker et al., 2018). Briefly, genomic DNA was extracted from granulocytes isolated with Ficoll density gradient centrifugation of whole blood from the iMCD patient and his parents using PAXgene Blood DNA Kit (QIAGEN). Samples were sequenced to 30× average coverage with 150-bp paired-end reads using an Illumina HiSeq X instrument. Raw sequence data were aligned to the human genome build GRCh38 with BWA-MEM, duplicates were marked with Picard, and base quality recalibration was performed with GATK. Single nucleotide variants and small insertions and deletions were jointly called with GATK HaplotypeCaller. Variants were annotated with Variant Effect Predictor (McLaren et al., 2016) followed by the creation of a GEMINI database (Paila et al., 2013).

### Digital droplet PCR

T cell, B cell, and monocyte subsets were sorted from the patient's PBMCs based on the cell surface expression of CD3, CD19, and CD14, respectively, using a FACSAria II sorter (BD Biosciences). Genomic DNA was extracted from each sorted cell populations using Zymo Quick-DNA Kit. Genomic DNA extracted from the pellet fraction of the Ficoll gradient was used to estimate variant allele frequency in granulocytes. Genomic DNA

was isolated from formalin-fixed paraffin-embedded bone marrow and lymph node tissue slides using Zymo Quick-DNA FFPE Miniprep Kit. The following primers and probes were used for allele-specific digital droplet PCR: universal forward primer: 5'-ACTGTGTAATGATTTCTGCT-3'; universal reverse primer: 5'-GCCACCAGGAGAAGC-3'; WT *DNMT3A*-specific probe: 5'-/5HEX/AGTTTCCCCCACACCA/3IABkFQ/-3' and mutant-specific probe: 5'-/56-FAM/AGTTTCCCTCACACCA/3IABkFQ/-3' (the reference and the mutated nucleotides are in bold). Primers and probes were custom synthesized (Integrated DNA Technologies). Digital droplet PCR was performed with reagents, supplies, and the QX200 instrument from Bio-Rad, following the manufacturer's instruction.

### Generation of patient-specific iPSCs

Human iPSCs were generated from the patient's PBMCs using integration-free Sendai virus vectors (CytoTune-iPS 2.0 Sendai Reprogramming Kits; Thermo Fisher Scientific) in collaboration with the Mount Sinai Pluripotent Stem Cell Core. iPSCs were cultured with StemFlex media (Gibco) on Cultrex (Amsio) in feeder-free condition. Cells were passaged with Accutase (STEMCELL Technologies) and plated in StemFlex media containing 1 μmol/liter thiazovivin (Selleck Chemicals). Clearance of the Sendai virus in the iPSC clones was confirmed by performing semiquantitative PCR specific to Sendai virus as described in the manufacturer's manual (Life Technologies). Expression of PSC markers was analyzed with LSRFortessa (BD Biosciences). Antibodies used were SSEA4, OCT3/4, NANOG, and SOX2 (Invitrogen). All antibodies were used as recommended by the manufacturer.

### Generation of genome-edited hESC clones

hESC line H1 (WA01) was obtained from WiCell and maintained in feeder-free condition as described for iPSCs. H1 cells were transfected with plasmids containing *Streptococcus pyogenes* Cas9, EGFP, and a guide RNA (a gift of Eirini Papapetrou, Icahn School of Medicine at Mount Sinai, New York, NY) using the Neon Transfection System (Thermo Fisher Scientific). The following guide RNAs were used to target the PWWP domain of *DNMT3A*: 5'-GCTACCACGCTGAGCCCGT-3' and 5'-TGGGTCA TGTGGTTCGGAGA-3'. After 48 h, EGFP<sup>+</sup> cells were sorted with FACSAria and plated on irradiated mouse embryonic fibroblast feeder layer. Colonies were genotyped by amplifying genomic segments flanking the target sites and carrying out Sanger sequencing. Clones with mono- or biallelic frameshift mutations underwent an additional round of single-cell cloning to ensure clonality.

### Quantitative RT-PCR measurements of *DNMT3A* and *DNMT3B* transcripts

Total RNA was extracted using Quick-RNA Kit (Zymo Research). Total RNA (500 ng) was reverse transcribed at 25°C for 10 min at 37°C for 2 h followed by 85°C for 5 min (Applied Biosystems). Multiplex quantitative PCR was performed using the following TaqMan gene expression assays: *DNMT3A* (Hs00173377), *DNMT3B* (Hs00171876), *GAPDH* (4326317E), and 18S rRNA (Hs99999901\_s1), all from Thermo Fisher Scientific. Relative

mRNA expression levels of *DNMT3A* and *DNMT3B* were obtained using the  $2^{-\Delta C_t}$  method, using *GAPDH* or 18S rRNA levels for normalization.

### DNMT3A immunoblotting

One million iPSCs or ESCs were lysed in 2× Laemmli sample buffer (Alfa Aesar) with  $\beta$ -mercaptoethanol and boiled at 95°C for 10 min. After removal of the insoluble fraction by centrifugation at 10,000 ×g for 10 min, protein samples were separated by SDS gel electrophoresis and transferred to a polyvinylidene difluoride membrane (Immobilon-P). Membranes were stained with DNMT3A antibody (Cell Signaling Technology; 3598S) at a dilution of 1:1,000 or hnRNP3/4 antibody (Bethyl Laboratories; A303-910A-T) at a dilution of 1:5,000 at 4°C overnight. After washing with Tris-buffered saline and Tween 20, HRP-conjugated anti-rabbit IgG (Cell Signaling Technology; 7074S) were added at a dilution of 1:10,000 and incubated for 1 h at room temperature. After washing, bands were detected using an enhanced chemiluminescence detection kit (Millipore) and Hyperfilm-ECL reagents (Amersham).

### Differentiation of hESCs to macrophages

hESCs were washed with PBS and harvested by incubating the cells for 5 min at 37°C with Accutase. Cells were counted, washed with PBS, and resuspended at a final concentration of  $1.25 \times 10^5$  cells/ml in EB media containing mTeSR1 (STEMCELL Technologies), 50 ng/ml BMP-4 (R&D Systems), 20 ng/ml stem cell factor (R&D Systems), and 50 ng/ml vascular endothelial growth factor (R&D Systems). 100  $\mu$ l of cell suspension in EB media supplemented with 10  $\mu$ mol/liter thiazovivin (Selleck Chemicals) were added per well of a 96-well ultra-low adherence plate (Costar 7007), centrifuged at 100 ×g for 3 min, and incubated for 8 d. EBs were fed at days 1 and 2 by aspirating 50  $\mu$ l of used media and adding 50  $\mu$ l of fresh EB media. At day 4, 100  $\mu$ l of fresh EB media was added. EBs were fed at day 6 by aspirating 100  $\mu$ l of used media and adding 100  $\mu$ l of fresh EB media. At day 8, EBs were transferred into a 100-mm tissue culture plate (96 EBs per plate; USA Scientific) and resuspended in monocyte/macrophage differentiation media I consisting of X-VIVO 15 (Lonza) supplemented with 100 ng/ml M-CSF (R&D Systems), 25 ng/ml IL-3 (R&D Systems), 2 mM glutamax (Gibco), 100 U/ml penicillin and 100 mg/ml streptomycin (Gibco), and 55  $\mu$ M  $\beta$ -mercaptoethanol (VWR International). Two thirds of the media was changed every 5 d. Nonadherent monocytes/macrophages were harvested from the supernatant every 5 d from day 10 to day 25 and cultivated on a new tissue culture plate for 7–10 d in differentiation media II consisting of X-VIVO 15 supplemented with 50 ng/ml M-CSF, 2 mM glutamax, 100 U/ml penicillin and 100 mg/ml streptomycin, and 55  $\mu$ M  $\beta$ -mercaptoethanol for further maturation.

### Macrophage characterization

hESC-derived macrophages were harvested using a cell scraper and resuspended in PBS for analysis. Expression of macrophage markers was evaluated with flow cytometry using LSRFortessa (BD Biosciences). Antibodies used were CD45, CD11b, CD14, and CD16 (eBioscience) and CD34 and CD43 (BioLegend). All

antibodies were used as recommended by the manufacturer. For zymosan phagocytosis assay, hESC-derived macrophages were plated on tissue culture plates in RPMI 1640 (Gibco) with 10% heat-inactivated human serum (Sigma), 2 mM glutamax (Gibco), and 100 U/ml penicillin and 100 mg/ml streptomycin (Gibco) for 1 d. On the day of the analysis, negative control cells were pretreated with 10  $\mu$ M cytochalasin D for 1 h. After pretreatment, all wells were fed with media containing two particles per cell of Alexa Fluor 488-labeled zymosan (Thermo Fisher Scientific) and incubated for 30 min at 37°C. Macrophages were detached using a cell scraper in PBS, and membrane-bound zymosan was quenched in 250 mg/ml trypan blue (Gibco) for 5 min. Stained macrophages were analyzed with LSRFortessa (BD Biosciences). For the modified low-density lipoprotein uptake assay, macrophages were incubated for 4 h with DiI-labeled acetylated low-density lipoprotein (DiI-AcLDL 5  $\mu$ g/ml; Invitrogen) in RPMI 1640 containing 2 mM glutamax (Gibco) and 100 U/ml penicillin and 100 mg/ml streptomycin (Gibco) at 37°C. DiI-AcLDL uptake was analyzed by flow cytometry with LSRFortessa (BD Biosciences). DiI-AcLDL uptake is represented as a percentage, with the untreated control indicated as 100%.

### DNMT3A silencing with siRNAs

A pool of nontargeting siRNAs and a pool of *DNMT3A*-targeting siRNAs were purchased from Dharmacon. WT hESC-derived macrophages were plated in 24-well plates at  $1.2 \times 10^5$  cells per well for 24 h before transfection. Lipid-siRNA complexes were prepared by incubating 10 nM of siRNA with lipofectamine RNAiMAX (Life Technologies) in the volume recommended by the manufacturer. Lipid-siRNA complexes were added to the wells in a final volume of 500  $\mu$ l of serum-free RPMI 1640. After incubation for 24 h, cells were washed and reincubated in RPMI 1640 containing 10% heat-inactivated human serum (Sigma). Total RNA was extracted 4 d after (5 d after the addition of siRNAs) for quantitative RT-PCR. Genomic DNA was extracted at the same time for the DNA methylation analysis with the EPIC array.

### DNA methylation array and analysis

Genomic DNA was extracted from iPSCs, hESCs, hESC-derived EBs, MPs, and macrophages using Quick-DNA Kit (Zymo Research). DNA methylation levels were assessed with Infinium MethylationEPIC BeadChips (Illumina). We followed a Bioconductor workflow using multiple R packages for analyzing the methylation array data (Maksimovic et al., 2016). Briefly, the raw data were imported with minfi and normalized with the preprocessQuantile method (Aryee et al., 2014). Cross-reactive probes (Chen et al., 2013) as well as probes that have failed in one or more samples were filtered out. A linear regression modeling using the limma package was used to identify the differentially methylated probes (Ritchie et al., 2015). CpG sites with Benjamini-Hochberg-adjusted  $P < 0.05$  were considered significantly differentially methylated between groups.

### WGBS and analysis

Genomic DNA was isolated from hESC-derived macrophages using Quick-DNA Kit (Zymo Research). Bisulfite conversion,

library preparation, and sequencing of the libraries were performed by Genewiz. Briefly, 100–200 ng of genomic DNA was fragmented using a Covaris S2. The sheared DNA was purified using the DNA Clean & Concentrator Kit (Zymo Research) per the manufacturer's recommendations. Bisulfite conversion of DNA was conducted using the EZ DNA Methylation-Gold Kit (Zymo Research). Bisulfite-converted DNA was immediately processed for generating WGBS libraries using the Accel-NGS Methyl-Seq DNA Library Kit (Swift BioSciences), following the manufacturer's protocol. The libraries were sequenced on an Illumina HiSeq 2500 sequencer as 150-bp paired-end reads. Quality- and adaptor-trimmed raw sequences were aligned to GRCh38/hg38 primary genome and deduplicated using Bismark (Krueger and Andrews, 2011). CpG methylation data extracted using the bismark\_methylation\_extractor was modified to construct a bsseq object (Hansen et al., 2012), which was used for statistical testing with the DSS package (Feng et al., 2014) DMLtest with smoothing. Only CpG sites covered at least five times in all nine samples (three clones per group;  $n = 23,450,022$ ) were included in the DMLtest analysis. CallDML function with  $p.\text{threshold} = 0.05$  was used to identify differentially methylated sites.

To determine macrophage-gain and -loss sites, the ENCODE project WGBS data for H1-hESC was used (ENCSR617FKV, processed methylation state at CpG). For the chromatin state analyses, hg38-lifted 15-coremark mnemonics files for H1 (E003) and primary monocytes (E029) downloaded from the Roadmap Epigenomics project portal were used. For the TF binding motif analyses, findMotifsGenome.pl function of HOMER (Heinz et al., 2010) was used with the -size 200 option.

For the analysis of the TBR5 patient vs. his sibling's WGBS data (Spencer et al., 2017), bigwig files for coverage and methylation levels for the patient's and the sibling's monocytes, neutrophils, and T cells were downloaded from <https://wustl.app.box.com/s/6fvifntz6golh4vu52epkfqxyd0vqmk>. Only CpGs covered at least five times in all six samples ( $n = 18,212,634$ ) were included in the analysis shown in Fig. 5. From the same URL, WGBS-based bigwig files for coverage and methylation levels for CD34-positive progenitor cells (six samples), monocytes (two samples), and neutrophils (three samples) from healthy donors were downloaded to determine monocyte-gain/loss sites and neutrophil-gain/loss sites. Because the coverages of individual CD34-positive samples were generally low (less than 1 million CpGs covered at least five times), data were collapsed for each cell type (CD34, monocytes, neutrophils) using the collapseBSseq function of the bsseq package. Only CpGs covered at least 10 times in collapsed CD34 and collapsed monocytes ( $n = 18,245,172$ ) were used to determine monocyte-gain and -loss sites using the DSS package DMLtest analysis (smoothing=TRUE). Likewise, only CpGs covered at least 10 times in collapsed CD34 and collapsed neutrophils ( $n = 28,958,116$ ) were used to determine neutrophil-gain and -loss sites using the DSS package DMLtest analysis (smoothing=TRUE). For the analysis of WGBS data of DNMT3A-deficient hESC-derived motor neurons (Ziller et al., 2018), the following data were downloaded from NCBI GEO under accession no. GSE90553: WGBS\_hESC\_KO3A\_D0 (two samples), WGBS\_

hESC\_WT\_D0 (two samples), WGBS\_hESC\_KO3A\_D14 (three samples), and WGBS\_hESC\_WT\_D14 (four samples). Only CpGs covered at least five times in all 11 samples ( $n = 6,905,467$ ) were included in the DSS package DMLtest analysis with smoothing.

### csRNA-seq and analysis

csRNA-seq was performed as described in Duttke et al. (2019). Total RNA was extracted from unstimulated hESC-derived macrophages using Quick-RNA Kit (Zymo Research). Small RNAs of ~20–60 nt were size selected from 0.4–1 µg of total RNA by denaturing gel electrophoresis. A 10% input sample was taken aside, and the remainder was enriched for 5'-capped RNAs with 3'-OH. Monophosphorylated RNAs were selectively degraded by Terminator 5'-Phosphate-Dependent Exonuclease (Lucigen). Subsequent 5' dephosphorylation by CIP (New England Biolabs [NEB]) followed by decapping with RppH (NEB) augments Cap-specific 5' adapter ligation by T4 RNA ligase 1 (NEB). The 3' adapter was ligated using truncated T4 RNA ligase 2 (NEB) without prior 3' repair to select against degraded RNA fragments. Following cDNA synthesis, libraries were amplified for 14 cycles and sequenced on the Illumina NextSeq 500 instrument as 75-bp single-end reads. Sequencing reads were trimmed for 3' adapter sequences (AGATCGGAAGAGCACACGTCT) using HOMER ("homerTools trim") and aligned using STAR with default parameters. TSS clusters were defined using HOMER's findcsRNATSS.pl tool that automates the discovery of strand-specific regions enriched with csRNA-seq reads found within 150 bp with a minimum read depth of 7 reads per  $10^7$  aligned reads and greater than twofold reads per bp than the surrounding 10 kb. This step eliminates loci with minimal numbers of supporting reads or regions with high levels of diffuse signal. Short RNA input libraries and total RNA-seq were used to filter out likely false-positive TSS clusters. TSS clusters with at least three samples with a count of five or more were included for the differential expression analysis by DESeq2. TSS clusters with  $P < 0.05$  was considered as significantly differentially expressed. For the TF binding motif analyses, the findMotifsGenome.pl tool of HOMER was used with -size -150,50 option.

### RNA-seq and analysis

Total RNA was extracted from PBS-stimulated or LPS-stimulated (100 ng/ml for 4 h) hESC-derived macrophages using Quick-RNA Kit (Zymo Research). RNA-seq libraries were generated using Illumina TruSeq RNA Library Preparation Kits. Pools were sequenced on the Illumina HiSeq 2500 sequencer as 150-bp paired-end reads. RNA-seq data were analyzed using Bioconductor packages Salmon (Patro et al., 2017), tximeta (Love et al., 2020), and DESeq2 (Love et al., 2014), following the recommended RNA-seq workflow (Love et al., 2015). Only genes with at least three samples with a count of  $\geq 10$  were included for the differential expression analysis by DESeq2. Two independent experiments were pooled with batch effect correction by adding batch in the design formula of DESeq2. Genes with Benjamini-Hochberg-adjusted  $P < 0.05$  were considered as significantly differentially expressed. Shrunk log2-fold changes based on the apeglm method (Zhu et al., 2019) was used for some of the plots.

## Genomic and transcriptomic data

The EPIC array, WGBS, csRNA-seq, and RNA-seq data are available through NCBI GEO under accession no. GSE168722. The whole-genome sequence of the iMCD patient and his parents are deposited in NCBI dbGaP under accession no. phs001706.v1.p1.

## Study approval

This study was approved by the institutional review boards of the Icahn School of Medicine at Mount Sinai and Washington University School of Medicine in St. Louis. Written informed consent was received from the iMCD patient and his parents before inclusion in the study.

## Online supplemental material

**Fig. S1** supplements **Fig. 1** and shows the expression levels of pluripotency markers in WT and mutant iPSC clones derived from the iMCD patient. **Fig. S2** supplements **Fig. 2** and shows CRISPR/Cas9-edited hESC clones' *DNMT3A* genotype and their expression levels of pluripotency markers. **Fig. S3** supplements **Fig. 3** and shows phenotypic characterization of macrophages derived from WT or *DNMT3A*-mutated hESCs. **Fig. S4** supplements **Fig. 4** and shows enrichment/depletion of cell type-specific chromatin states for differentially methylated sites in *DNMT3A*-null macrophages. **Fig. S5** supplements **Fig. 7** and shows transcriptomic characterization of *DNMT3A*-haploinsufficient macrophages.

## Acknowledgments

The authors thank Eirini Papapetrou for sharing plasmids and protocols for CRISPR/Cas9-based genome editing of hESCs and the Flow Cytometry Core and Pluripotent Stem Cell Engineering Core at Icahn School of Medicine at Mount Sinai for their contributions.

This work was funded by Icahn School of Medicine at Mount Sinai, Castleman Disease Collaborative Network, Chan Zuckerberg Initiative Single-Cell Analysis of Inflammation Program, and National Institutes of Health (NIH) R01HL153974 to M. Byun; NIH K99GM135515 to S.H. Duttke; NIH U19AI135972 and GM134366 to C. Benner; the Burroughs Wellcome Fund, Chan Zuckerberg Initiative Neurodegeneration Challenge Network, Irma T. Hirsch Trust Young Investigator fellowship, and NIH U01AI150748 and R01AI143840 to I. Marazzi.

Author contributions: J.-Y. Lim designed experiments, performed iPSC and hESC characterization, generated and characterized hESC-derived macrophages, analyzed the data, and wrote the manuscript. S.H. Duttke performed csRNA-seq. T.S. Baker performed digital droplet PCR and designed the guide RNAs for CRISPR-Cas9 targeting. J. Lee, K.J. Gambino, and N.J. Venturini generated CRISPR-Cas9-targeted hESCs. J.S.Y. Ho, S. Zheng, and Y.S. Fstchyan analyzed the data. V. Pillai and D.C. Fajgenbaum provided the iMCD patient's biopsy specimen. I. Marazzi and C. Benner analyzed and interpreted the data. M. Byun conceptualized the study, provided supervision, designed experiments, analyzed and interpreted the data, and wrote the manuscript. All authors edited the manuscript.

Disclosures: D.C. Fajgenbaum reported grants from EUSA Pharma and non-financial support from Pfizer outside the

submitted work; in addition, D.C. Fajgenbaum had a patent for "Methods of Treating Idiopathic Multicentric Castleman Disease with JAK1/2 inhibition" pending (no licensee) and a patent for "Treatment of Castleman Disease" pending (no licensee). No other disclosures were reported.

Submitted: 23 December 2020

Revised: 4 March 2021

Accepted: 2 April 2021

## References

- Andersson, R., C. Gebhard, I. Miguel-Escalada, I. Hoof, J. Bornholdt, M. Boyd, Y. Chen, X. Zhao, C. Schmidl, T. Suzuki, et al. 2014. An atlas of active enhancers across human cell types and tissues. *Nature*. 507:455–461. <https://doi.org/10.1038/nature12787>
- Arends, C.M., J. Galan-Sousa, K. Hoyer, W. Chan, M. Jäger, K. Yoshida, R. Seemann, D. Noerenberg, N. Waldhueter, H. Fleischer-Notter, et al. 2018. Hematopoietic lineage distribution and evolutionary dynamics of clonal hematopoiesis. *Leukemia*. 32:1908–1919. <https://doi.org/10.1038/s41375-018-0047-7>
- Aryee, M.J., A.E. Jaffe, H. Corrada-Bravo, C. Ladd-Acosta, A.P. Feinberg, K.D. Hansen, and R.A. Irizarry. 2014. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics*. 30:1363–1369. <https://doi.org/10.1093/bioinformatics/btu049>
- Baker, T.S., K.J. Gambino, L. Schrieffer, J.Y. Lim, K.M. Steinberg, D.C. Fajgenbaum, A. Martín García-Sancho, and M. Byun. 2018. A novel FAS mutation with variable expressivity in a family with unicentric and idiopathic multicentric Castleman disease. *Blood Adv.* 2:2959–2963. <https://doi.org/10.1182/bloodadvances.2018023911>
- Bock, C., E. Kiskinis, G. Versteppen, H. Gu, G. Boulting, Z.D. Smith, M. Ziller, G.F. Croft, M.W. Amoroso, D.H. Oakley, et al. 2011. Reference Maps of human ES and iPS cell variation enable high-throughput characterization of pluripotent cell lines. *Cell*. 144:439–452. <https://doi.org/10.1016/j.cell.2010.12.032>
- Bröske, A.M., L. Vockentanz, S. Kharazi, M.R. Huska, E. Mancini, M. Scheller, C. Kuhl, A. Enns, M. Prinz, R. Jaenisch, et al. 2009. DNA methylation protects hematopoietic stem cell multipotency from myeloid restriction. *Nat. Genet.* 41:1207–1215. <https://doi.org/10.1038/ng.463>
- Buchrieser, J., W. James, and M.D. Moore. 2017. Human induced pluripotent stem cell-derived macrophages share ontogeny with MYB-independent tissue-resident macrophages. *Stem Cell Reports*. 8:334–345. <https://doi.org/10.1016/j.stemcr.2016.12.020>
- Buscariet, M., S. Provost, Y.F. Zada, V. Bourgoin, L. Mollica, M.P. Dubé, and L. Busque. 2018. Lineage restriction analyses in CHIP indicate myeloid bias for TET2 and multipotent stem cell origin for *DNMT3A*. *Blood*. 132:277–280. <https://doi.org/10.1182/blood-2018-01-829937>
- Challen, G.A., and M.A. Goodell. 2020. Clonal hematopoiesis: mechanisms driving dominance of stem cell clones. *Blood*. 136:1590–1598.
- Challen, G.A., D. Sun, M. Jeong, M. Luo, J. Jelinek, J.S. Berg, C. Bock, A. Vasanthakumar, H. Gu, Y. Xi, et al. 2011. *Dnmt3a* is essential for hematopoietic stem cell differentiation. *Nat. Genet.* 44:23–31. <https://doi.org/10.1038/ng.1009>
- Challen, G.A., D. Sun, A. Mayle, M. Jeong, M. Luo, B. Rodriguez, C. Mallaney, H. Celik, L. Yang, Z. Xia, et al. 2014. *Dnmt3a* and *Dnmt3b* have overlapping and distinct functions in hematopoietic stem cells. *Cell Stem Cell*. 15:350–364. <https://doi.org/10.1016/j.stem.2014.06.018>
- Charlton, J., E.J. Jung, A.L. Mattei, N. Bailly, J. Liao, E.J. Martin, P. Giesselmann, B. Brändl, E.K. Stamenova, F.J. Müller, et al. 2020. TETs compete with *DNMT3* activity in pluripotent cells at thousands of methylated somatic enhancers. *Nat. Genet.* 52:819–827. <https://doi.org/10.1038/s41588-020-0639-9>
- Chen, Y.A., M. Lemire, S. Choufani, D.T. Butcher, D. Grafodatskaya, B.W. Zanke, S. Gallinger, T.J. Hudson, and R. Weksberg. 2013. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics*. 8:203–209. <https://doi.org/10.4161/epi.23470>
- Choi, J., T.M. Baldwin, M. Wong, J.E. Bolden, K.A. Fairfax, E.C. Lucas, R. Cole, C. Biben, C. Morgan, K.A. Ramsay, et al. 2019. Haemopedia RNA-seq: a database of gene expression during haematopoiesis in mice and

- humans. *Nucleic Acids Res.* 47(D1):D780–D785. <https://doi.org/10.1093/nar/gky1020>
- Christian, D.L., D.Y. Wu, J.R. Martin, J.R. Moore, Y.R. Liu, A.W. Clemens, S.A. Nettles, N.M. Kirkland, T. Papouin, C.A. Hill, et al. 2020. DNMT3A haploinsufficiency results in behavioral deficits and global epigenomic dysregulation shared across neurodevelopmental disorders. *Cell Rep.* 33: 108416. <https://doi.org/10.1016/j.celrep.2020.108416>
- Cole, C.B., D.A. Russler-Germain, S. Ketkar, A.M. Verdoni, A.M. Smith, C.V. Bangert, N.M. Helton, M. Guo, J.M. Klco, S. O’Laughlin, et al. 2017. Haploinsufficiency for DNA methyltransferase 3A predisposes hematopoietic cells to myeloid malignancies. *J. Clin. Invest.* 127:3657–3674. <https://doi.org/10.1172/JCI93041>
- Dawoud, A.A.Z., W.J. Tapper, and N.C.P. Cross. 2020. Clonal myelopoiesis in the UK Biobank cohort: ASXL1 mutations are strongly associated with smoking. *Leukemia*. 34:2660–2672. <https://doi.org/10.1038/s41375-020-0896-8>
- Deaton, A.M., and A. Bird. 2011. CpG islands and the regulation of transcription. *Genes Dev.* 25:1010–1022. <https://doi.org/10.1101/gad.203751>
- Dorsheimer, L., B. Assmus, T. Rasper, C.A. Ortmann, A. Ecke, K. Abou-El-Ardat, T. Schmid, B. Brüne, S. Wagner, H. Serve, et al. 2019. Association of mutations contributing to clonal hematopoiesis with prognosis in chronic ischemic heart failure. *JAMA Cardiol.* 4:25–33. <https://doi.org/10.1001/jamacardio.2018.3965>
- Duttkie, S.H., M.W. Chang, S. Heinz, and C. Benner. 2019. Identification and dynamic quantification of regulatory elements using total RNA. *Genome Res.* 29:1836–1846. <https://doi.org/10.1101/gr.253492.119>
- Fajgenbaum, D.C. 2018. Novel insights and therapeutic approaches in idiopathic multicentric Castleman disease. *Blood*. 132:2323–2330. <https://doi.org/10.1182/blood-2018-05-848671>
- Feng, H., K.N. Conneely, and H. Wu. 2014. A Bayesian hierarchical model to detect differentially methylated loci from single nucleotide resolution sequencing data. *Nucleic Acids Res.* 42:e69. <https://doi.org/10.1093/nar/gku154>
- Fishilevich, S., R. Nudel, N. Rappaport, R. Hadar, I. Plaschkes, T. Iny Stein, N. Rosen, A. Kohn, M. Twik, M. Safran, et al. 2017. GeneHancer: genome-wide integration of enhancers and target genes in GeneCards. *Database (Oxford)*. 2017:bax028. <https://doi.org/10.1093/database/bax028>
- Fuster, J.J., S. MacLauchlan, M.A. Zuriaga, M.N. Polackal, A.C. Ostriker, R. Chakraborty, C.L. Wu, S. Sano, S. Muralidharan, C. Rius, et al. 2017. Clonal hematopoiesis associated with TET2 deficiency accelerates atherosclerosis development in mice. *Science*. 355:842–847. <https://doi.org/10.1126/science.aag1381>
- Genovese, G., A.K. Kähler, R.E. Handsaker, J. Lindberg, S.A. Rose, S.F. Bakhoum, K. Chambert, E. Mick, B.M. Neale, M. Fromer, et al. 2014. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N. Engl. J. Med.* 371:2477–2487. <https://doi.org/10.1056/NEJMoa1409405>
- Gu, W., H.C. Lee, D. Chaves, E.M. Youngman, G.J. Pazour, D. Conte Jr., and C.C. Mello. 2012. CapSeq and CIP-TAP identify Pol II start sites and reveal capped small RNAs as C. elegans piRNA precursors. *Cell*. 151: 1488–1500. <https://doi.org/10.1016/j.cell.2012.11.023>
- Haney, S.L., G.M. Upchurch, J. Opavsky, D. Klinkebiel, R.A. Hlady, S. Roy, S. Dutta, K. Datta, and R. Opavsky. 2016. DNMT3a is a haploinsufficient tumor suppressor in CD8+ peripheral T cell lymphoma. *PLoS Genet.* 12: e1006334. <https://doi.org/10.1371/journal.pgen.1006334>
- Hansen, K.D., B. Langmead, and R.A. Irizarry. 2012. BSmooth: from whole genome bisulfite sequencing reads to differentially methylated regions. *Genome Biol.* 13:R83. <https://doi.org/10.1186/gb-2012-13-10-r83>
- Hawkins, J.M., and V. Pillai. 2015. TAFRO syndrome or Castleman-Kojima syndrome: a variant of multicentric Castleman disease. *Blood*. 126:2163. <https://doi.org/10.1182/blood-2015-07-662122>
- He, Y., M. Hariharan, D.U. Gorkin, D.E. Dickel, C. Luo, R.G. Castanon, J.R. Nery, A.Y. Lee, Y. Zhao, H. Huang, et al. 2020. Spatiotemporal DNA methylome dynamics of the developing mouse fetus. *Nature*. 583: 752–759. <https://doi.org/10.1038/s41586-020-2119-x>
- Heinz, S., C. Benner, N. Spann, E. Bertolino, Y.C. Lin, P. Laslo, J.X. Cheng, C. Murre, H. Singh, and C.K. Glass. 2010. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell*. 38:576–589. <https://doi.org/10.1016/j.molcel.2010.05.004>
- Heinz, S., L. Texari, M.G.B. Hayes, M. Urbanowski, M.W. Chang, N. Givarkes, A. Rialdi, K.M. White, R.A. Albrecht, L. Pache, et al. 2018. Transcription Elongation Can Affect Genome 3D Structure. *Cell*. 174:1522–1536.e22. <https://doi.org/10.1016/j.cell.2018.07.047>
- Hon, G.C., C.X. Song, T. Du, F. Jin, S. Selvaraj, A.Y. Lee, C.A. Yen, Z. Ye, S.Q. Mao, B.A. Wang, et al. 2014. 5mC oxidation by Tet2 modulates enhancer activity and timing of transcriptome reprogramming during differentiation. *Mol. Cell*. 56:286–297. <https://doi.org/10.1016/j.molcel.2014.08.026>
- Iwakawa, S., T. Yamamoto, and Y. Watanabe. 2006. Human plasma adenosine deaminase 2 is secreted by activated monocytes. *Biol. Chem.* 387:319–321. <https://doi.org/10.1515/BC.2006.042>
- Jaiswal, S., P. Fontanillas, J. Flannick, A. Manning, P.V. Grauman, B.G. Mar, R.C. Lindsley, C.H. Mermel, N. Burt, A. Chavez, et al. 2014. Age-related clonal hematopoiesis associated with adverse outcomes. *N. Engl. J. Med.* 371:2488–2498. <https://doi.org/10.1056/NEJMoa1408617>
- Jaiswal, S., P. Natarajan, A.J. Silver, C.J. Gibson, A.G. Bick, E. Shvartz, M. McConkey, N. Gupta, S. Gabriel, D. Ardissino, et al. 2017. Clonal hematopoiesis and risk of atherosclerotic cardiovascular disease. *N. Engl. J. Med.* 377:111–121. <https://doi.org/10.1056/NEJMoa1701719>
- Jeong, M., D. Sun, M. Luo, Y. Huang, G.A. Challen, B. Rodriguez, X. Zhang, L. Chavez, H. Wang, R. Hannah, et al. 2014. Large conserved domains of low DNA methylation maintained by Dnmt3a. *Nat. Genet.* 46:17–23. <https://doi.org/10.1038/ng.2836>
- Jeong, M., H.J. Park, H. Celik, E.L. Ostrander, J.M. Reyes, A. Guzman, B. Rodriguez, Y. Lei, Y. Lee, L. Ding, et al. 2018. Loss of Dnmt3a immortalizes hematopoietic stem cells in vivo. *Cell Rep.* 23:1–10. <https://doi.org/10.1016/j.celrep.2018.03.025>
- Kaasinen, E., O. Kuismin, K. Rajamäki, H. Ristolainen, M. Aavikko, J. Kondein, S. Saarinen, D.G. Berta, R. Katainen, E.A.M. Hirvonen, et al. 2019. Impact of constitutional TET2 haploinsufficiency on molecular and clinical phenotype in humans. *Nat. Commun.* 10:1252. <https://doi.org/10.1038/s41467-019-09198-7>
- Kervestin, S., and A. Jacobson. 2012. NMD: a multifaceted response to premature translational termination. *Nat. Rev. Mol. Cell Biol.* 13:700–712. <https://doi.org/10.1038/nrm3454>
- Kim, K., R. Zhao, A. Doi, K. Ng, J. Untermaier, P. Cahan, H. Huo, Y.H. Loh, M.J. Aryee, M.W. Lensch, et al. 2011. Donor cell type can influence the epigenome and differentiation potential of human induced pluripotent stem cells. *Nat. Biotechnol.* 29:1117–1119. <https://doi.org/10.1038/nbt.2052>
- Krueger, F., and S.R. Andrews. 2011. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics*. 27:1571–1572. <https://doi.org/10.1093/bioinformatics/btr167>
- Kundaje, A., W. Meuleman, J. Ernst, M. Bilenky, A. Yen, A. Heravi-Moussavi, P. Kheradpour, Z. Zhang, J. Wang, M.J. Ziller, et al. Roadmap Epigenomics Consortium. 2015. Integrative analysis of 111 reference human epigenomes. *Nature*. 518:317–330. <https://doi.org/10.1038/nature14248>
- Leoni, C., S. Montagner, A. Rinaldi, F. Bertoni, S. Polletti, C. Balestrieri, and S. Monticelli. 2017. Dnmt3a restrains mast cell inflammatory responses. *Proc. Natl. Acad. Sci. USA*. 114:E1490–E1499. <https://doi.org/10.1073/pnas.1616420114>
- Liao, J., R. Karnik, H. Gu, M.J. Ziller, K. Clement, A.M. Tsankov, V. Akopian, C.A. Gifford, J. Donaghey, C. Galonska, et al. 2015. Targeted disruption of DNMT1, DNMT3A and DNMT3B in human embryonic stem cells. *Nat. Genet.* 47:469–478. <https://doi.org/10.1038/ng.3258>
- Love, M.I., W. Huber, and S. Anders. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15: 550. <https://doi.org/10.1186/s13059-014-0550-8>
- Love, M.I., S. Anders, V. Kim, and W. Huber. 2015. RNA-Seq workflow: gene-level exploratory analysis and differential expression. *F1000Res*. 4:1070. <https://doi.org/10.12688/f1000research.7035.1>
- Love, M.I., C. Soneson, P.F. Hickey, L.K. Johnson, N.T. Pierce, L. Shepherd, M. Morgan, and R. Patro. 2020. Tuxedo: Reference sequence checksums for provenance identification in RNA-seq. *PLoS Comput. Biol.* 16: e1007664. <https://doi.org/10.1371/journal.pcbi.1007664>
- Luo, C., P. Hajkova, and J.R. Ecker. 2018. Dynamic DNA methylation: in the right place at the right time. *Science*. 361:1336–1340. <https://doi.org/10.1126/science.aat6806>
- Maksimovic, J., B. Phipson, and A. Oshlack. 2016. A cross-package Bioconductor workflow for analysing methylation array data. *F1000 Res*. 5: 1281. <https://doi.org/10.12688/f1000research.8839.2>
- Marsman, J., and J.A. Horsfield. 2012. Long distance relationships: enhancer-promoter communication and dynamic gene transcription. *Biochim. Biophys. Acta*. 1819:1217–1227. <https://doi.org/10.1016/j.bbagr.2012.10.008>
- McLaren, W., L. Gil, S.E. Hunt, H.S. Riat, G.R. Ritchie, A. Thormann, P. Flicek, and F. Cunningham. 2016. The Ensembl Variant Effect Predictor. *Genome Biol.* 17:122. <https://doi.org/10.1186/s13059-016-0974-4>
- Medzhitov, R. 2001. Toll-like receptors and innate immunity. *Nat. Rev. Immunol.* 1:135–145. <https://doi.org/10.1038/35100529>

- Moore, J.E., M.J. Purcaro, H.E. Pratt, C.B. Epstein, N. Shores, J. Adrian, T. Kauli, C.A. Davis, A. Dobin, R. Kaul, et al. ENCODE Project Consortium. 2020. Expanded encyclopaedias of DNA elements in the human and mouse genomes. *Nature*. 583:699–710. <https://doi.org/10.1038/s41586-020-2493-4>
- Nagy, A., A. Bhaduri, N. Shahmarvand, J. Shahryari, J.L. Zehnder, R.A. Warnke, T. Mughal, S. Ali, and R.S. Ohgami. 2018. Next-generation sequencing of idiopathic multicentric and unicentric Castleman disease and follicular dendritic cell sarcomas. *Blood Adv.* 2:481–491. <https://doi.org/10.1182/bloodadvances.2017009654>
- Navon Elkan, P., S.B. Pierce, R. Segel, T. Walsh, J. Barash, S. Padeh, A. Zlotogorski, Y. Berkun, J.J. Press, M. Mukamel, et al. 2014. Mutant adenosine deaminase 2 in a polyarteritis nodosa vasculopathy. *N. Engl. J. Med.* 370:921–931. <https://doi.org/10.1056/NEJMoal307362>
- Nechaev, S., D.C. Fargo, G. dos Santos, L. Liu, Y. Gao, and K. Adelman. 2010. Global analysis of short RNAs reveals widespread promoter-proximal stalling and arrest of Pol II in *Drosophila*. *Science*. 327:335–338. <https://doi.org/10.1126/science.1181421>
- Paila, U., B.A. Chapman, R. Kirchner, and A.R. Quinlan. 2013. GEMINI: integrative exploration of genetic variation and genome annotations. *PLOS Comput. Biol.* 9:e1003153. <https://doi.org/10.1371/journal.pcbi.1003153>
- Patro, R., G. Duggal, M.I. Love, R.A. Irizarry, and C. Kingsford. 2017. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods*. 14:417–419. <https://doi.org/10.1038/nmeth.4197>
- Pidsley, R., E. Zotenko, T.J. Peters, M.G. Lawrence, G.P. Risbridger, P. Molloy, S. Van Dijk, B. Muhlhäuser, C. Stirzaker, and S.J. Clark. 2016. Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. *Genome Biol.* 17:208. <https://doi.org/10.1186/s13059-016-1066-1>
- Rasmussen, K.D., I. Berest, S. Keßler, K. Nishimura, L. Simón-Carrasco, G.S. Vassiliou, M.T. Pedersen, J. Christensen, J.B. Zaugg, and K. Helin. 2019. TET2 binding to enhancers facilitates transcription factor recruitment in hematopoietic cells. *Genome Res.* 29:564–575. <https://doi.org/10.1101/gr.239277.118>
- Rinaldi, L., D. Datta, J. Serrat, L. Morey, G. Solanas, A. Avgustinova, E. Blanco, J.I. Pons, D. Matallanas, A. Von Kriegsheim, et al. 2016. Dnmt3a and Dnmt3b associate with enhancers to regulate human epidermal stem cell homeostasis. *Cell Stem Cell*. 19:491–501. <https://doi.org/10.1016/j.stem.2016.06.020>
- Ritchie, M.E., B. Hipson, D. Wu, Y. Hu, C.W. Law, W. Shi, and G.K. Smyth. 2015. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43:e47. <https://doi.org/10.1093/nar/gkv007>
- Russler-Germain, D.A., D.H. Spencer, M.A. Young, T.L. Lamprecht, C.A. Miller, R. Fulton, M.R. Meyer, P. Erdmann-Gilmore, R.R. Townsend, R.K. Wilson, et al. 2014. The R882H DNMT3A mutation associated with AML dominantly inhibits wild-type DNMT3A by blocking its ability to form active tetramers. *Cancer Cell*. 25:442–454. <https://doi.org/10.1016/j.ccr.2014.02.010>
- Sano, S., K. Oshima, Y. Wang, Y. Katanasaka, M. Sano, and K. Walsh. 2018a. CRISPR-mediated gene editing to assess the roles of Tet2 and Dnmt3a in clonal hematopoiesis and cardiovascular disease. *Circ. Res.* 123:335–341. <https://doi.org/10.1161/CIRCRESAHA.118.313225>
- Sano, S., K. Oshima, Y. Wang, S. MacLauchlan, Y. Katanasaka, M. Sano, M.A. Zuriaga, M. Yoshiyama, D. Goukassian, M.A. Cooper, et al. 2018b. Tet2-mediated clonal hematopoiesis accelerates heart failure through a mechanism involving the IL-1 $\beta$ /NLRP3 inflammasome. *J. Am. Coll. Cardiol.* 71:875–886. <https://doi.org/10.1016/j.jacc.2017.12.037>
- Sardina, J.L., S. Collombet, T.V. Tian, A. Gómez, B. Di Stefano, C. Berenguer, J. Brumbaugh, R. Stadholders, C. Segura-Morales, M. Gut, et al. 2018. Transcription factors drive Tet2-mediated enhancer demethylation to reprogram cell fate. *Cell Stem Cell*. 23:905–906. <https://doi.org/10.1016/j.stem.2018.11.001>
- Shen, W., J.M. Heeley, C.M. Carlston, R. Acuna-Hidalgo, W.M. Nillesen, K.M. Dent, G.V. Douglas, K.L. Levine, P. Bayrak-Toydemir, C.L. Marcelis, et al. 2017. The spectrum of DNMT3A variants in Tatton-Brown-Rahman syndrome overlaps with that in hematologic malignancies. *Am. J. Med. Genet. A*. 173:3022–3028. <https://doi.org/10.1002/ajmg.a.38485>
- Smith, Z.D., and A. Meissner. 2013. DNA methylation: roles in mammalian development. *Nat. Rev. Genet.* 14:204–220. <https://doi.org/10.1038/nrg3354>
- Spencer, D.H., D.A. Russler-Germain, S. Ketkar, N.M. Helton, T.L. Lamprecht, R.S. Fulton, C.C. Fronick, M. O’Laughlin, S.E. Heath, M. Shinawi, et al. 2017. CpG island hypermethylation mediated by DNMT3A is a consequence of AML progression. *Cell*. 168:801–816.e13. <https://doi.org/10.1016/j.cell.2017.01.021>
- Stadler, M.B., R. Murr, L. Burger, R. Ivanek, F. Lienert, A. Schöler, E. van Nimwegen, C. Wirbelauer, E.J. Oakeley, D. Gaidatzis, et al. 2011. DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature*. 480:490–495. <https://doi.org/10.1038/nature10716>
- Tatton-Brown, K., S. Seal, E. Ruark, J. Harmer, E. Ramsay, S. Del Vecchio Duarte, A. Zachariou, S. Hanks, E. O’Brien, L. Aksglaede, et al. Childhood Overgrowth Consortium. 2014. Mutations in the DNA methyltransferase gene DNMT3A cause an overgrowth syndrome with intellectual disability. *Nat. Genet.* 46:385–388. <https://doi.org/10.1038/ng.2917>
- Tippens, N.D., J. Liang, A.K. Leung, S.D. Wierbowski, A. Ozer, J.G. Booth, J.T. Lis, and H. Yu. 2020. Transcription imparts architecture, function and logic to enhancer units. *Nat. Genet.* 52:1067–1075. <https://doi.org/10.1038/s41588-020-0686-2>
- Tovy, A., J.M. Reyes, M.C. Gundry, L. Brunetti, H. Lee-Six, M. Petljak, H.J. Park, A.G. Guzman, C. Rosas, A.R. Jeffries, et al. 2020. Tissue-biased expansion of DNMT3A-mutant clones in a mosaic individual is associated with conserved epigenetic erosion. *Cell Stem Cell*. 27:326–335.e4. <https://doi.org/10.1016/j.stem.2020.06.018>
- Trowbridge, J.J., J.W. Snow, J. Kim, and S.H. Orkin. 2009. DNA methyltransferase 1 is essential for and uniquely regulates hematopoietic stem and progenitor cells. *Cell Stem Cell*. 5:442–449. <https://doi.org/10.1016/j.stem.2009.08.016>
- van Wilgenburg, B., C. Browne, J. Vowles, and S.A. Cowley. 2013. Efficient, long term production of monocyte-derived macrophages from human pluripotent stem cells under partly-defined and fully-defined conditions. *PLoS One*. 8:e71098. <https://doi.org/10.1371/journal.pone.0071098>
- Wu, X., and Y. Zhang. 2017. TET-mediated active DNA demethylation: mechanism, function and beyond. *Nat. Rev. Genet.* 18:517–534. <https://doi.org/10.1038/nrg.2017.33>
- Xie, M., C. Lu, J. Wang, M.D. McLellan, K.J. Johnson, M.C. Wendl, J.F. McMichael, H.K. Schmidt, V. Yellapantula, C.A. Miller, et al. 2014. Age-related mutations associated with clonal hematopoietic expansion and malignancies. *Nat. Med.* 20:1472–1478. <https://doi.org/10.1038/nm.3733>
- Yagi, M., M. Kabata, A. Tanaka, T. Ukai, S. Ohta, K. Nakabayashi, M. Shimizu, K. Hata, A. Meissner, T. Yamamoto, and Y. Yamada. 2020. Identification of distinct loci for de novo DNA methylation by DNMT3A and DNMT3B during mammalian development. *Nat. Commun.* 11:3199. <https://doi.org/10.1038/s41467-020-16989-w>
- Yang, L., R. Rau, and M.A. Goodell. 2015. DNMT3A in haematological malignancies. *Nat. Rev. Cancer*. 15:152–165. <https://doi.org/10.1038/nrc3895>
- Zavialov, A.V., E. Gracia, N. Glaichenhaus, R. Franco, A.V. Zavialov, and G. Lauvau. 2010. Human adenosine deaminase 2 induces differentiation of monocytes into macrophages and stimulates proliferation of T helper cells and macrophages. *J. Leukoc. Biol.* 88:279–290. <https://doi.org/10.1189/jlb.1109764>
- Zhou, Q., D. Yang, A.K. Ombrello, A.V. Zavialov, C. Toro, A.V. Zavialov, D.L. Stone, J.J. Chae, S.D. Rosenzweig, K. Bishop, et al. 2014. Early-onset stroke and vasculopathy associated with mutations in ADA2. *N. Engl. J. Med.* 370:911–920. <https://doi.org/10.1056/NEJMoal307361>
- Zhu, A., J.G. Ibrahim, and M.I. Love. 2019. Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences. *Bioinformatics*. 35:2084–2092. <https://doi.org/10.1093/bioinformatics/bty895>
- Ziller, M.J., J.A. Ortega, K.A. Quinlan, D.P. Santos, H. Gu, E.J. Martin, C. Gálonska, R. Pop, S. Maidl, A. Di Pardo, et al. 2018. Dissecting the functional consequences of de novo DNA methylation dynamics in human motor neuron differentiation and physiology. *Cell Stem Cell*. 22:559–574.e9. <https://doi.org/10.1016/j.stem.2018.02.012>

Supplemental material

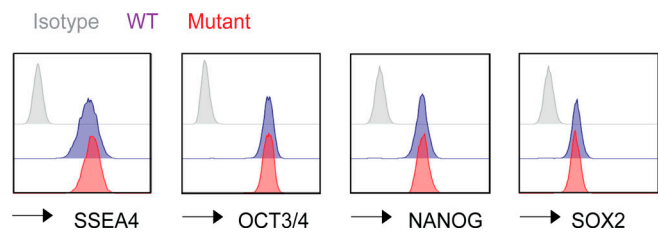


Figure S1. **Expression levels of pluripotency markers in the iMCD patient-derived iPSCs.** Expression levels of the indicated pluripotency markers in WT and mutant iPSC clones were determined with flow cytometry. Data for a representative clone for each genotype are shown.

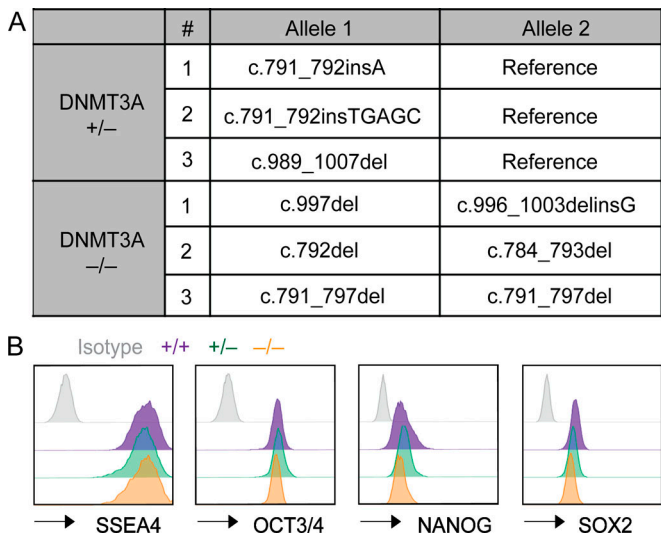


Figure S2. **DNMT3A genotypes and expression levels of pluripotency markers in genome-edited hESCs.** (A) Genotypes of DNMT3A-haploinsufficient and -null hESCs generated by CRISPR-Cas9 genome editing. (B) Expression levels of the indicated pluripotency markers were assessed with flow cytometry. Data for a representative clone for each DNMT3A genotype are shown.

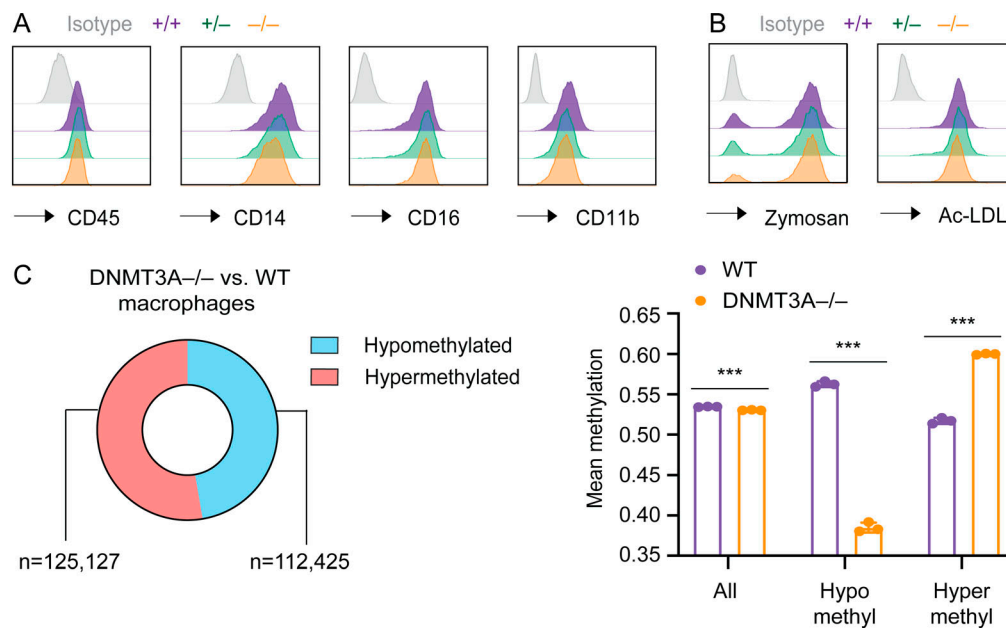


Figure S3. **Characterization of DNMT3A $^{+/-}$  and DNMT3A $^{-/-}$  hESC-derived macrophages.** (A) Expression levels of the indicated cell surface markers were assessed with flow cytometry in macrophages derived from hESC clones with indicated DNMT3A genotype. (B) The ability to uptake fluorescently labeled zymosan and acetylated low-density lipoprotein (AcLDL) was assessed in the same set of cells as in A with flow cytometry. (C) Proportions and numbers of hypo- or hypermethylated sites (left) as well as mean methylation levels (right) in DNMT3A-null hESC-derived macrophages in comparison with WT controls are shown. Error bars represent the SEM. \*\*\*,  $P < 0.001$  by Student's  $t$  test.

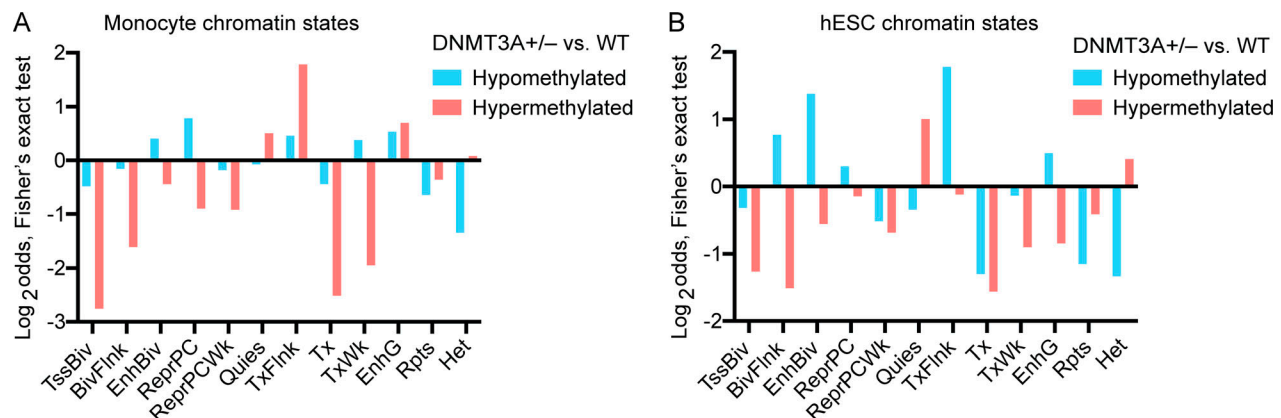
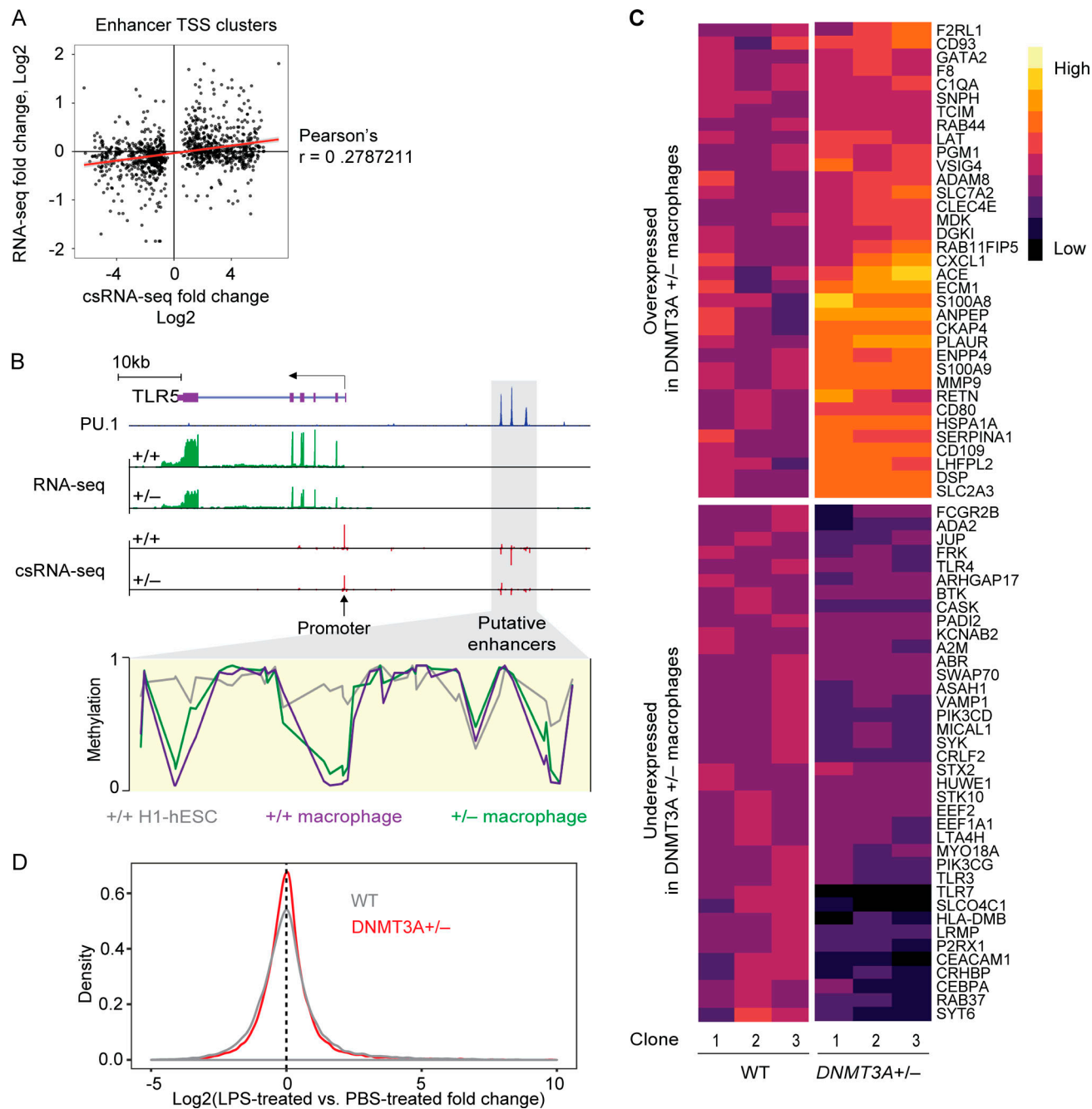


Figure S4. **Association of cell-type-specific chromatin states with differentially methylated sites in DNMT3A-haploinsufficient macrophages.** (A) Two-sided Fisher's exact test results (Log<sub>2</sub>odds) for enrichment/depletion of monocyte chromatin states are shown ( $P < 0.05$  for all tests). BivFlnk, flanking bivalent TSS/Enh; EnhBiv, bivalent enhancer; EnhG, genic enhancers; Het, heterochromatin; Quies, quiescent; ReprPC, repressed polycomb; ReprPCWk, weak repressed polycomb; Rpts, ZNF genes and repeats; TssBiv, bivalent/poised TSS; TxFlnk, transcribed at gene 5' and 3'; Tx, strong transcription; TxWk, weak transcription. (B) Same as A for hESC chromatin states.



**Figure S5. Transcriptomic characterization of DNMT3A-haploinsufficient macrophages.** (A) Comparison of eRNA fold-changes (*DNMT3A*<sup>+/-</sup> over WT macrophages) as measured with csRNA-seq and mRNA fold-changes (*DNMT3A*<sup>+/-</sup> over WT macrophages) of nearest genes as measured with RNA-seq. (B) Schematic of a locus on chromosome 1 containing TLR5 (chr1:223,109,404–223,143,248). PU.1 chromatin immunoprecipitation sequencing data of blood monocyte-derived macrophages is from a published study (Heinz et al., 2018). RNA-seq and csRNA-seq data from hESC-derived macrophages with indicated *DNMT3A* genotype are shown in green and red, respectively. DNA methylation levels at putative TLR5 enhancers (gray box) in WT and *DNMT3A*<sup>+/-</sup> macrophages as well as WT H1-hESCs are shown. (C) A heatmap of genes with Gene Ontology term “myeloid leukocyte activation” (GO:0002274), expressed differentially between WT and *DNMT3A*-haploinsufficient macrophages. Each column indicates macrophages differentiated from an independent hESC clone. (D) The LPS-treated vs. PBS-treated gene expression fold-change distributions in WT and *DNMT3A*-haploinsufficient macrophages are shown in a density plot. The vertical dotted line indicates genes with no detectable expression difference between LPS- and PBS-treated conditions.