

Metformin reduces the clonal fitness of Dnmt3aR878H hematopoietic stem and progenitor cells by reversing their aberrant metabolic and epigenetic state

Steven Chan (✉ steven.chan@uhnresearch.ca)

Princess Margaret Cancer Centre <https://orcid.org/0000-0002-4354-0731>

Mohsen Hosseini

Princess Margaret Cancer Center

Veronique Voisin

University of Toronto

Ali Chegini

Princess Margaret Cancer Centre

Angelica Varesi

Princess Margaret Cancer Centre

Severine Cathelin

Princess Margaret Cancer Centre

Dhanoop Manikoth Ayyathan

Princess Margaret Cancer Centre

Alex Liu

Princess Margaret Cancer Centre

Yitong Yang

Princess Margaret Cancer Centre

Vivian Wang

Princess Margaret Cancer Centre

Abdula Maher

Princess Margaret Cancer Centre

Eric Grignano

Princess Margaret Cancer Centre

Julie Haines

Angelo D'Alessandro

University of Colorado Anschutz Medical Campus <https://orcid.org/0000-0002-2258-6490>

Kira Young

The Jackson Laboratory

Yiyan Wu

University of Toronto

Martina Fiumara

San Raffaele Telethon Institute for Gene Therapy

Samuele Ferrari

San Raffaele Telethon Institute for Gene Therapy

Luigi Naldini

San Raffaele Telethon Institute for Gene Therapy, IRCCS San Raffaele Scientific Institute

<https://orcid.org/0000-0002-7835-527X>

Federico Gaiti

Princess Margaret Cancer Centre <https://orcid.org/0000-0001-5111-8816>

Shraddha Pai

Donnelly Centre for Cellular and Biomolecular Research

Aaron Schimmer

University Health Network <https://orcid.org/0000-0003-4023-3899>

Gary Bader

Donnelly Centre for Cellular and Biomolecular Research

John Dick

Princess Margaret Cancer Centre, University Health Network

Stephanie Z. Xie

Princess Margaret Cancer Centre, University Health Network

Jennifer Trowbridge

Jackson Laboratory <https://orcid.org/0000-0002-7636-9504>

Biological Sciences - Article

Keywords:

Posted Date: February 6th, 2024

DOI: <https://doi.org/10.21203/rs.3.rs-3874821/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: **Yes** there is potential Competing Interest. S.M.C. has received research funding from the Centre for Oncology and Immunology in Hong Kong, Celgene/BMS, AbbVie Pharmaceuticals, Agios Pharmaceuticals, and Servier Laboratories. F.G. serves as a consultant for S2 Genomics Inc. A.D.S. has received research funding from Takeda Pharmaceuticals, BMS and Medivir AB, and consulting fees/honorarium from Takeda, Novartis, Jazz, and Otsuka Pharmaceuticals. A.D.S. is named on a patent application for the use of DNT cells to treat AML. A.D.S. is a member of the Medical and Scientific

Advisory Board of the Leukemia and Lymphoma Society of Canada. A.D.S. holds the Ronald N. Buick Chair in Oncology Research. J.E.D. has received research funding from Celgene/BMS, and has patents licensed to Trillium Therapeutics/Pfizer.

Metformin reduces the clonal fitness of *Dnmt3a*^{R878H} hematopoietic stem and progenitor cells by reversing their aberrant metabolic and epigenetic state

Mohsen Hosseini¹, Veronique Voisin^{2#}, Ali Chegini^{1,3#}, Angelica Varesi^{1,4#}, Severine Cathelin^{1#}, Dhanoop Manikoth Ayyathan¹, Alex C.H. Liu^{1,3}, Yitong Yang^{1,3}, Vivian Wang^{1,3}, Abdula Maher¹, Eric Grignano¹, Julie A. Reisz⁵, Angelo D'Alessandro⁵, Kira Young⁶, Yiyang Wu³, Martina Fiumara^{7,8}, Samuele Ferrari^{7,8}, Luigi Naldini^{7,8}, Federico Gaiti^{1,3}, Shraddha Pai^{2,9}, Aaron D. Schimmer^{1,3}, Gary D. Bader², John E. Dick^{1,4}, Stephanie Z. Xie¹, Jennifer J. Trowbridge⁶, and Steven M. Chan^{1,3*}

¹Princess Margaret Cancer Centre, Toronto, Ontario, Canada

²Donnelly Centre for Cellular and Biomolecular Research, Toronto, Ontario, Canada

³Department of Medical Biophysics, University of Toronto, Toronto, Ontario, Canada

⁴Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada

⁵Department of Biochemistry and Molecular Genetics, University of Colorado Anschutz Medical Campus, Aurora, CO, USA

⁶The Jackson Laboratory, Bar Harbor, ME, USA

⁷San Raffaele Telethon Institute for Gene Therapy, IRCCS San Raffaele Scientific Institute, Milan, 20132, Italy.

⁸Vita-Salute San Raffaele University, Milan, 20132, Italy.

⁹Ontario Institute for Cancer Research, Toronto, Ontario, Canada

#Contributed equally

*Corresponding Author

Steven M. Chan

Princess Margaret Cancer Research Tower

101 College Street, Room 8-354

Toronto, Ontario

M2N 3L3

Phone: (647) 965-5748

Email: steven.chan@uhn.ca

1 Abstract

2 Clonal hematopoiesis (CH) arises when a hematopoietic stem cell (HSC) acquires a mutation that
3 confers a competitive advantage over wild-type (WT) HSCs, resulting in its clonal expansion.
4 Individuals with CH are at an increased risk of developing hematologic neoplasms and a range of
5 age-related inflammatory illnesses¹⁻³. Therapeutic interventions that suppress the expansion of
6 mutant HSCs have the potential to prevent these CH-related illnesses; however, such interventions
7 have not yet been identified. The most common CH driver mutations are in the DNA
8 methyltransferase 3 alpha (*DNMT3A*) gene with arginine 882 (R882) being a mutation hotspot.
9 Here we show that murine hematopoietic stem and progenitor cells (HSPCs) carrying the
10 *Dnmt3a*^{R878H/+} mutation, which is equivalent to human *DNMT3A*^{R882H/+}, have increased
11 mitochondrial respiration compared with WT cells and are dependent on this metabolic
12 reprogramming for their competitive advantage. Importantly, treatment with metformin, an oral
13 anti-diabetic drug with inhibitory activity against complex I in the electron transport chain (ETC),
14 reduced the fitness of *Dnmt3a*^{R878H/+} HSCs. Through a multi-omics approach, we discovered that
15 metformin acts by enhancing the methylation potential in *Dnmt3a*^{R878H/+} HSPCs and reversing
16 their aberrant DNA CpG methylation and histone H3K27 trimethylation (H3K27me3) profiles.
17 Metformin also reduced the fitness of human *DNMT3A*^{R882H} HSPCs generated by prime editing.
18 Our findings provide preclinical rationale for investigating metformin as a preventive intervention
19 against illnesses associated with *DNMT3A*^{R882} mutation-driven CH in humans.

20 **Main**

21 Mutations in *DNMT3A* are the most common genetic alterations in CH and are found in ~50-60%
22 of CH carriers²⁻⁵. *DNMT3A* encodes a *de novo* DNA methyltransferase that catalyzes transfer of
23 the methyl group from *S*-adenosylmethionine (SAM) to the C-5 position of cytosines in DNA,
24 resulting in 5-methylcytosine (5mC) and production of *S*-adenosylhomocysteine (SAH). *DNMT3A*
25 mutations are classified into those affecting the mutational hotspot at R882 and those affecting
26 other parts of the gene (non-R882)⁶. Although both types of mutations are predicted to reduce
27 methyltransferase activity, *DNMT3A*^{R882} mutations appear to confer a significantly higher risk of
28 progression to acute myeloid leukemia (AML) compared with non-R882 *DNMT3A* mutations^{7,8}.
29 Thus, *DNMT3A*^{R882} mutations represent an important target for preventive intervention.

30 The mutations affecting R882 are almost invariably missense alterations and heterozygous^{9,10}.
31 *DNMT3A*^{R882} mutations have been shown to not only reduce the methyltransferase activity of the
32 mutant protein but also decrease the activity of the WT protein in a dominant negative manner^{10,11}.
33 Consistent with these findings, the differentially methylated regions (DMRs) in human AML cells
34 or peripheral blood cells harboring *DNMT3A*^{R882} mutations are predominantly hypomethylated
35 compared with their WT counterparts^{10,12}.

36 The impact of *DNMT3A* mutations on cell fate decisions of HSCs has previously been studied
37 using genetically modified mouse models. In the *Dnmt3a*^{R878H/+} mouse model, the mutant HSCs
38 are expanded and have a competitive advantage over WT HSCs¹³, thus recapitulating a key
39 functional change associated with the mutation in humans. Here, we employed this model to
40 identify differences in dependencies between *Dnmt3a*^{R878H/+} and WT HSPCs with the goal of
41 targeting such dependencies to selectively suppress the expansion of mutant HSCs.

42 ***Dnmt3a*^{R878H/+} HSPCs exhibit increased mitochondrial respiration**

43 Analysis of publicly available RNA-sequencing (RNA-seq) datasets of primary AML samples
44 revealed an increase in the expression of genes involved in oxidative phosphorylation (OXPHOS)
45 in *DNMT3A*^{R882}-mutated patient samples, but not *DNMT3A*^{non-R882}-mutated samples, compared
46 with *DNMT3A* WT samples (**Extended Data Fig. 1a**). These findings led us to explore if there are
47 potential differences in mitochondrial function between *Dnmt3a*^{R878H/+} and *Dnmt3a*^{+/+} murine

48 HSPCs. We found that *Dnmt3a*^{R878H/+} lineage negative, c-kit positive (LK) cells, which are
49 enriched for HSPCs, possessed higher levels of basal and maximal oxygen consumption rates
50 (OCRs) than *Dnmt3a*^{+/+} LK cells as determined by extracellular flux analysis (**Fig. 1a**). These
51 differences were also observed in unfractionated whole bone marrow (WBM) cells, albeit by a
52 smaller magnitude (**Extended Data Fig. 1b**). Furthermore, the level of mitochondrial reactive
53 oxygen species (ROS) and ratio of mitochondrial transmembrane potential ($\Delta\Psi_m$) to
54 mitochondrial mass (MM) were higher in mutant LK cells than in WT LK cells (**Fig. 1b and 1c**).
55 Together, these findings indicate that the *Dnmt3a*^{R878H} mutation causes metabolic reprogramming
56 in HSPCs resulting in upregulation of OXPHOS.

57 **Inhibition of mitochondrial respiration suppresses the competitive advantage of** 58 ***Dnmt3a*^{R878H/+} HSPCs**

59 We hypothesized that the enhanced mitochondrial respiration in *Dnmt3a*^{R878H} HSPCs is required
60 for their competitive advantage over WT cells. To test this hypothesis, we first established an *in*
61 *vitro* competition assay in which CD45.2⁺ *Dnmt3a*^{R878H/+} and CD45.1⁺ *Dnmt3a*^{+/+} LK cells were
62 mixed at a ~2:3 ratio and cultured in cytokine-supplemented methylcellulose medium for ~10 days,
63 followed by determination of the proportions of CD45.2⁺ and CD45.1⁺ cells (**Extended Data Fig.**
64 **1c**). A parallel competition assay between CD45.2⁺ *Dnmt3a*^{+/+} and CD45.1⁺ *Dnmt3a*^{+/+} LK cells
65 mixed at the same starting ratio served as control. After the culture period, the proportion of
66 CD45.2⁺ *Dnmt3a*^{R878H/+} cells was consistently ~20-30% higher than that of CD45.2⁺ *Dnmt3a*^{+/+}
67 cells in the first passage and ~40-50% higher in the second passage (**Extended Data Fig. 1d**),
68 demonstrating the competitive advantage of the mutant cells. Using this assay, we studied the
69 impact of genetic knockdown of ETC subunits on the competitive advantage of *Dnmt3a*^{R878H/+}
70 HSPCs by transducing the mixed CD45.2⁺ mutant/CD45.1⁺ WT population with lentiviral vectors
71 expressing short-hairpin RNAs (shRNAs) against *Ndufv1* and *Cox15* (**Extended Data Fig. 1e**),
72 which encode critical subunits in complex I and complex IV of the ETC, respectively.
73 Downregulation of these genes reduced the maximal OCR and competitive advantage of CD45.2⁺
74 *Dnmt3a*^{R878H/+} cells (**Fig. 1d and 1e**), indicating that mutant HSPCs are dependent on OXPHOS
75 to outcompete their WT counterparts.

76 To explore the translational relevance of this finding, we tested the impact of metformin, a
77 commonly used oral anti-diabetic drug and pharmacologic inhibitor of Complex I¹⁴, on
78 *Dnmt3a*^{R878H/+} LK HSPCs. Consistent with the genetic knockdown studies, treatment with
79 metformin at a clinically relevant concentration (50 μM)¹⁵ suppressed the competitive advantage
80 of mutant cells *in vitro* (**Fig. 1f**). This effect was rescued by expression of NDI1, a metformin-
81 resistant yeast analog of Complex I^{16,17}, thus confirming that metformin's effect was due to on-
82 target complex I inhibition (**Fig. 1g**). Metformin treatment also selectively reduced the clonogenic
83 potential of *Dnmt3a*^{R878H/+} over *Dnmt3a*^{+/+} LK HSPCs in standard colony-forming unit (CFU)
84 assays (**Extended Data Fig. 1f**).

85 To determine if the effect of metformin was relevant *in vivo* and over a longer treatment period,
86 we conducted a competitive repopulation experiment by mixing CD45.2⁺ *Dnmt3a*^{R878H/+} or
87 *Dnmt3a*^{+/+} WBM cells with CD45.1⁺ *Dnmt3a*^{+/+} WBM cells at a 2:3 ratio and transplanting the
88 mixed cells into lethally irradiated recipients (**Extended Data Fig. 1g**). Five weeks after
89 transplantation, the recipient mice were either left untreated or started on treatment with metformin
90 in their drinking water at 5 mg/mL, a concentration that has previously been shown to result in
91 blood concentrations comparable to those achievable in humans¹⁸. Peripheral blood (PB)
92 chimerism analysis showed a stable ratio of CD45.2⁺ to CD45.1⁺ cells in mice that received
93 CD45.2⁺ *Dnmt3a*^{+/+} control cells and the ratio was not affected by metformin treatment (**Fig. 1h**).
94 In contrast, the ratio of CD45.2⁺ to CD45.1⁺ cells steadily increased over a 7-month period in mice
95 that received CD45.2⁺ *Dnmt3a*^{R878H/+} cells, reflecting their competitive advantage over CD45.1⁺
96 WT cells (**Fig. 1h**). Importantly, metformin abrogated this competitive advantage up to 7 months
97 (**Fig. 1h**). This effect was observed in both the myeloid and lymphoid compartments (**Extended**
98 **Data Fig. 1h**). Our findings collectively indicate that inhibition of mitochondrial respiration is a
99 potential strategy for targeting *DNMT3A*^{R882} mutation-driven CH.

100 **Metformin suppresses the competitive advantage of *Dnmt3a*^{R878H/+} HSCs**

101 Metformin suppressed the long-term competitive advantage of *Dnmt3a*^{R878H/+} donor cells *in vivo*
102 (**Fig. 1h**), reflective of an impact at the HSC level. To provide independent evidence for impact at
103 the HSC level and gain insights into metformin's mechanism of action, we performed single cell
104 RNA-seq (scRNA-seq) analysis on LK-enriched bone marrow (BM) cells from the untreated and
105 metformin-treated recipients at the end of the 7-month treatment period (**Extended Data Fig. 1g**).

106 The cells were collected from the mice transplanted with CD45.2⁺ *Dnmt3a*^{R878H/+} and CD45.1⁺
107 *Dnmt3a*^{+/+} competitor cells and stained with antibody-oligonucleotide conjugates (AOCs) specific
108 for CD45.2 or CD45.1 to identify their donor origin. A total of 22,407 cells from untreated control
109 mice (n=2) were sequenced, 84.6% of which were CD45.2⁺ *Dnmt3a*^{R878H/+} cells (**Fig. 2a**). In
110 comparison, a total of 23,818 cells from metformin-treated mice (n=2) were sequenced, and the
111 proportion of CD45.2⁺ *Dnmt3a*^{R878H/+} cells was significantly less at 56.5% (p<0.0001 by chi-
112 square test) (**Fig. 2a**). To determine which HSPC subsets were affected, we annotated each cell
113 based on their correlation with reference murine HSPC gene sets¹⁹ and identified 11 different
114 hematopoietic subsets (**Fig. 2a**). Metformin treatment reduced the ratio of CD45.2⁺ to CD45.1⁺
115 cells in the HSC cluster as well as myeloid progenitor subsets (**Fig. 2b and 2c**).

116 To corroborate these findings, we transplanted *Dnmt3a*^{R878H/+} or *Dnmt3a*^{+/+} donor WBM cells from
117 sex-matched littermates in a non-competitive manner into lethally irradiated recipients. Five weeks
118 after transplantation, the recipients were either left untreated or started on treatment with
119 metformin for one month (**Extended Data Fig. 2a**). In untreated mice, the number of
120 immunophenotypic HSCs (Lin⁻, c-Kit⁺, Sca-1⁺, CD150⁺, CD48⁻) per femur and the proportion of
121 HSCs in the LK fraction were higher in *Dnmt3a*^{R878H/+} recipients compared with *Dnmt3a*^{+/+}
122 recipients (**Fig. 2d,e**). The expansion of mutant HSCs was associated with a trend towards a higher
123 proportion of HSCs in S/G2/M phase (**Fig. 2f**). Metformin treatment reduced all these parameters
124 in *Dnmt3a*^{R878H/+} recipients to levels comparable to those of untreated *Dnmt3a*^{+/+} recipients (**Fig.**
125 **2d-f**). Our findings demonstrate that metformin treatment suppresses the competitive advantage of
126 *Dnmt3a*^{R878H/+} HSCs.

127 **Metformin acts by increasing the methylation capacity of *Dnmt3a*^{R878H/+} HSPCs**

128 Gene set enrichment analysis (GSEA) of the scRNA-seq data revealed an enrichment of genes
129 associated with OXPHOS in *Dnmt3a*^{R878H/+} HSPCs relative to WT HSPCs and a decrease in
130 expression of these genes with metformin treatment (**Extended Data Fig. 3a**). This unexpected
131 result suggested that metformin could influence mitochondrial respiration not only through a direct
132 inhibition of complex I but also through downregulation of OXPHOS-related genes. To further
133 investigate its mechanism of action, we first studied the impact of *in vivo* metformin treatment on
134 mitochondrial function of HSPCs. We performed extracellular flux analysis on freshly isolated
135 LK-enriched BM cells from the animals that were untreated or treated with metformin for one

136 month (**Extended Data Fig. 2a**). Metformin treatment reduced the basal and maximal OCRs as
137 well as the $\Delta\Psi_m$ of *Dnmt3a*^{R878H/+} LK cells to levels comparable to those of untreated *Dnmt3a*^{+/+}
138 LK cells (**Extended Data Fig. 3b and 3c**). To uncover the impact of metformin on specific
139 metabolic pathways, we performed a mass spectrometry-based metabolomic analysis of the
140 untreated and treated LK cells of both *Dnmt3a* genotypes. This analysis, which focused on
141 metabolites central to energy and redox metabolism, detected 101 named metabolites, of which 14
142 were significantly increased in metformin-treated mutant LK cells compared with untreated mutant
143 cells (**Extended Data Table 1**). Intriguingly, 4 of the 14 upregulated metabolites (reduced
144 glutathione (GSH), taurine, L-cysteate, and dimethylglycine) are involved in one-carbon (1C)
145 metabolism through the methionine cycle (**Fig. 3a and 3b**). Since the methionine cycle generates
146 SAM, these findings suggest that metformin could potentially affect SAM levels and the ratio of
147 [SAM] to [SAH], which is also known as the methylation index, an indicator of cellular
148 methylation potential. To test this hypothesis, we directly measured the intracellular concentrations
149 of SAM and SAH, which were below the detection threshold of the bulk metabolomic analysis.
150 Consistent with our hypothesis, the methylation index was higher in metformin-treated
151 *Dnmt3a*^{R878H/+} LK cells than in untreated cells (**Fig. 3c**), indicative of an increase in their cellular
152 methylation potential. Importantly, the impact of metformin on methylation index was observed
153 only in *Dnmt3a*^{R878H/+} LK HSPCs but not in *Dnmt3a*^{+/+} LK cells (**Fig. 3c**).

154 To determine if the metformin-induced changes in methylation index could be due to alterations
155 in the expression of genes involved in 1C metabolism, we performed bulk RNA-seq analysis of
156 *Dnmt3a*^{+/+} and *Dnmt3a*^{R878H/+} LK-enriched cells that were untreated or treated with metformin for
157 1 month (**Extended Data Fig. 2a**). GSEA of the RNA-seq dataset showed that metformin
158 treatment decreased the expression of genes associated with stemness and OXPHOS (**Extended**
159 **Data Fig. 3d,e**), consistent with our earlier results. Importantly, it also revealed a significant
160 enrichment of genes involved in 1C metabolism in *Dnmt3a*^{R878H/+} LK cells (**Fig. 3d**). To confirm
161 these findings, we performed quantitative RT-PCR to measure the expression of 8 genes that
162 encode enzymes in the folate and methionine cycles (*Shmt2*, *Mthfd2l*, *Shmt1*, *Mthfd1*, *Mthfr*, *Ahcy*,
163 *Cbs*, *Bhmt*) and found that metformin treatment increased their expression in *Dnmt3a*^{R878H/+} LK
164 cells (**Fig. 3e**). Metformin also upregulated expression of 6 of the 8 genes in *Dnmt3a*^{+/+} LK cells,
165 but the magnitude of change was less (**Fig. 3e**). These findings suggest that metformin selectively

166 increases the cellular methylation potential of mutant HSPCs by upregulating the expression of
167 genes involved in 1C metabolism.

168 Based on the above findings, we hypothesized that metformin suppresses the clonal fitness of
169 *Dnmt3a*^{R878H/+} cells by increasing their cellular methylation potential. To test this hypothesis, we
170 investigated the impact of exogenous SAM and SAH on the competitive advantage of mutant LK
171 cells using our *in vitro* assay (**Extended Data Fig. 1c**). The addition of exogenous SAM, which
172 increases the methylation index, was sufficient to reduce the competitive advantage of
173 *Dnmt3a*^{R878H/+} HSPCs (**Fig. 3f**). Conversely, exogenous SAH, which lowers the methylation index,
174 counteracted the suppressive effect of metformin on mutant cells (**Fig. 3f**). To confirm these
175 findings, we inhibited serine hydroxymethyltransferase 2 (SHMT2) activity as an alternative
176 approach to lowering the [SAM]/[SAH] ratio. SHMT2 generates a one-carbon unit (5,10-
177 methylenetetrahydrofolate) that is necessary for SAM synthesis through the folate and methionine
178 cycles (**Fig. 3a**). In line with our hypothesis, both pharmacologic inhibition with SHIN-1, a potent
179 SHMT inhibitor, and genetic knockdown of *Shmt2* expression rescued the suppressive effect of
180 metformin on mutant HSPCs (**Fig. 3g,h and Extended Data Fig. 3f**). Altogether, these findings
181 support a mechanism in which metformin selectively reduces the clonal fitness of *Dnmt3a*^{R878H/+}
182 HSPCs by increasing their cellular methylation potential.

183 **Metformin reverses the aberrant DNA CpG methylation and H3K27me3 profiles in** 184 ***Dnmt3a*^{R878H/+} HSPCs**

185 The mechanism by which *DNMT3A* mutations confer a fitness advantage to mutant HSCs is
186 believed to be mediated through focal DNA hypomethylation secondary to reduced *de novo* DNA
187 methylation activity^{10,20}. In the context of *DNMT3A*^{R882} mutations where a WT copy of the gene
188 remains, the residual DNA methylation activity is estimated to be ~20% of normal but is not
189 absent¹⁰. We hypothesized that the metformin-induced increase in methylation index could
190 augment this activity, resulting in a reversal of the aberrant DNA CpG hypomethylation pattern in
191 mutant cells and consequent decrease in their fitness. A prediction based on this hypothesis is that
192 a further reduction in residual DNMT3A activity should render the mutant cells resistant to the
193 effect of metformin. In line with this prediction, we found that *Dnmt3a*^{R878H/+} LK cells expressing
194 a *Dnmt3a* shRNA to knockdown residual DNMT3A activity maintained their competitive
195 advantage over WT LK cells even in the presence of metformin *in vitro* (**Extended Data Fig.**

196 **4a,b**). Another prediction is that metformin treatment should increase the level of methylation at
197 CpG sites that are differentially hypomethylated in *Dnmt3a*^{R878H/+} cells. To test this hypothesis,
198 we performed reduced representation bisulfite sequencing (RRBS) analysis of LK-enriched BM
199 cells from recipient mice that received *Dnmt3a*^{R878H/+} or *Dnmt3a*^{+/+} WBM cells from sex-matched
200 littermate donors and were either untreated or treated with metformin for one month (**Extended**
201 **Data Fig. 2a**). The RRBS technique, which enriches for CpG-rich regions, was chosen because
202 DNMT3A preferentially catalyzes DNA methylation at CpG dinucleotides. We identified 5,430
203 DMRs in the comparison between untreated *Dnmt3a*^{R878H/+} (n=4) and untreated *Dnmt3a*^{+/+} (n=3)
204 samples. Consistent with prior reports^{10,20}, the majority of the DMRs (n=4,649; 85.6%) were
205 hypomethylated in the untreated *Dnmt3a*^{R878H/+} samples (**Fig. 4a**). In the comparison between
206 metformin-treated *Dnmt3a*^{R878H/+} (n=3) and untreated *Dnmt3a*^{R878H/+} (n=4) samples, we identified
207 3,285 DMRs, 1,923 (58.5%) of which were hypermethylated in the treated samples (**Fig. 4a**). We
208 found 870 overlapping DMRs at the intersection between these two sets (**Extended Data Fig. 4c**).
209 In line with our hypothesis, metformin treatment increased the methylation level at 617 (90.9%)
210 of the 679 hypomethylated DMRs in *Dnmt3a*^{R878H/+} samples (**Fig. 4b and Extended Data Fig.**
211 **4d**). Similar findings were observed in the subsets of DMRs associated with CpG islands and gene
212 promoter regions (**Fig. 4a,b and Extended Data Fig. 4c,d**). These findings demonstrate that
213 metformin treatment can, at least in part, reverse the aberrant DNA CpG hypomethylation pattern
214 in mutant cells.

215 In human CH, the *DNMT3A*^{R882} mutation has previously been reported to result in preferential
216 DNA hypomethylation of targets of the polycomb repressive complex 2 (PRC2)²⁰, which catalyzes
217 the methylation of H3K27. In addition, the *Dnmt3a*^{R878H} mutation was previously found to be
218 associated with a reduction in H3K27me3²¹, indicating its potential influence on another layer of
219 epigenetic regulation. Given that PRC2-mediated methylation activity is also regulated by the ratio
220 of [SAM]/[SAH], we hypothesized that metformin could reverse the aberrant H3K27
221 hypomethylation profile in *Dnmt3a*^{R878H/+} HSPCs. To test this hypothesis, we performed
222 H3K27me3 chromatin immunoprecipitation followed by sequencing (ChIP-seq) analysis of
223 *Dnmt3a*^{R878H/+} and *Dnmt3a*^{+/+} LK-enriched BM cells from the mice that were either untreated or
224 treated with metformin for one month (**Extended Data Fig. 2a**). This analysis revealed a reduction
225 in H3K27me3 levels globally and in the regions surrounding transcription start sites (TSSs) in the
226 untreated *Dnmt3a*^{R878H/+} samples relative to the untreated *Dnmt3a*^{+/+} samples (**Fig. 4c,d**).

227 Consistent with our hypothesis, metformin treatment restored H3K27me3 levels in *Dnmt3a*^{R878H/+}
228 samples to a level comparable to that of untreated *Dnmt3a*^{+/+} samples (**Fig. 4c,d**). To confirm these
229 findings using an orthogonal approach, we measured the level of H3K27me3 by intracellular flow
230 cytometry. Similar to the ChIP-seq results, we found that *Dnmt3a*^{R878H/+} LK cells had lower levels
231 of H3K27me3 than *Dnmt3a*^{+/+} LK cells and metformin treatment restored H3K27me3 levels in
232 mutant cells to levels comparable to those of WT cells (**Fig. 4e**). Altogether, the above findings
233 demonstrate that metformin treatment can reverse the aberrant epigenetic landscape in
234 *Dnmt3a*^{R878H/+} HSPCs.

235 **Metformin decreases the fitness of human *DNMT3A*^{R882H} HSPCs**

236 To explore the relevance of our findings in human CH, we designed and optimized a new prime
237 editing strategy to introduce the R882H mutation into the *DNMT3A* gene in purified CD34⁺ HSPCs
238 from human cord blood (CB) samples as conventional homology-directed repair (HDR)-based
239 CRISPR/Cas9 editing strategies are highly inefficient in human HSPCs²². The Cas nickase (nCas)-
240 based prime editing technique has been shown to cause less cytotoxic/genotoxic stress and edit
241 with higher precision and efficiency in long-term repopulating HSPCs^{22,23}. Using the optimized
242 prime editing strategy, we introduced the *DNMT3A*^{R882H} mutation in 10 HSPC samples from
243 independent donors. As a negative control, we introduced a T>G single nucleotide variant (SNV)
244 in exon 1 of the beta-2-microglobulin (*B2M*) gene which causes a premature stop codon²². The
245 baseline mean variant allele frequency (VAF) on day 3 after prime editing was 9.3% for
246 *DNMT3A*^{R882H} and 51.6% for the *B2M* SNV (**Fig. 5a,b**). On day 4, the edited cell pools were plated
247 in methylcellulose medium to assess the relative fitness of the *DNMT3A*^{R882H} versus *DNMT3A*^{WT}
248 cells in the presence or absence of tumor necrosis factor alpha (TNF α). The impact of TNF α , a
249 proinflammatory cytokine, was studied because it has previously been shown to promote the
250 competitive advantage of *Dnmt3a*^{R878H/+} HSCs²⁴. After an additional 14 days in culture, the mean
251 *DNMT3A*^{R882H} VAF remained stable in the absence of TNF α but increased to 30.9% in the
252 presence of TNF α (**Fig. 5a**), indicative of a relative expansion of the mutant population in a
253 proinflammatory milieu. Metformin treatment significantly prevented the expansion of *DNMT3A*-
254 mutated cells in the presence of TNF α (**Fig. 5a**). Importantly, TNF α and metformin treatment did
255 not affect the VAF of *B2M*-edited cells (**Fig. 5b**), indicating that the observed effects on
256 *DNMT3A*^{R882H} HSPCs were not an artifact of prime editing. These results together support that

257 metformin has the potential to suppress the fitness of *DNMT3A*^{R882}-mutated clones in human CH
258 upon inflammatory stress.

259 **Discussion**

260 Targeting the cell intrinsic mechanisms critical for the selective advantage of mutant HSPCs in
261 CH is a potential strategy for suppressing clonal expansion and lowering the risk of developing
262 CH-related illnesses. Here, we found that upregulation of mitochondrial respiration is a key
263 functional consequence of the *Dnmt3a*^{R878H} mutation and mutant HSPCs are dependent on this
264 metabolic reprogramming to outcompete their WT counterparts. Importantly, this dependency was
265 evident at the level of HSCs. Thus, our findings provide evidence that mitochondrial metabolism
266 is a critical cell intrinsic regulator of clonal fitness in *DNMT3A*^{R882} mutation-driven CH. This
267 notion is consistent with the growing body of evidence demonstrating a role for mitochondrial
268 bioenergetics and dynamics in the regulation of stem cell fate.

269 Our discovery that *Dnmt3a*^{R878H/+} HSPCs are dependent on increased mitochondrial respiration
270 has important therapeutic implications because many components of the ETC are druggable
271 cellular targets. In this study, we focused on the therapeutic potential of metformin, a biguanide
272 widely used in the treatment of diabetes. Although biguanides have been reported to target many
273 cellular proteins, their inhibitory effect on complex I (NADH dehydrogenase) activity is the most
274 well established and supported by structural evidence¹⁴. Indeed, our finding that ectopic expression
275 of the metformin-resistant yeast analog of complex I (NDI1) rendered *Dnmt3a*^{R878H/+} HSPCs
276 insensitive to effects of metformin strongly supports complex I as the main protein target.
277 However, the observed reduction in mitochondrial respiration in metformin-treated *Dnmt3a*^{R878H/+}
278 HSPCs was not due to complex I inhibition alone but also through the downstream downregulation
279 of genes involved in OXPHOS. Results from our multi-omics studies suggest that metformin
280 exerts its downstream effects on gene expression, at least in part, by increasing the methylation
281 potential and consequently, augmenting the activity of DNMT3A, the PRC2 complex, and possibly
282 other SAM-dependent methyltransferases in *Dnmt3a*^{R878H/+} HSPCs. This proposed mechanism is
283 consistent with prior studies demonstrating an association between metformin exposure and an
284 increase in 5mC and H3K27me3 levels in various cellular contexts²⁵⁻²⁷. It is noteworthy that
285 metformin appears to preferentially increase the expression of genes involved in 1C metabolism
286 and methylation index in *Dnmt3a*^{R878H/+} HSPCs over WT cells, indicating a degree of selectivity
287 in its effects. Whether this selectivity is specific for metformin or common across other ETC
288 inhibitors is unclear and warrants further investigations.

289 To explore the relevance of our findings in humans, we optimized a prime editing strategy to
290 introduce the *DNMT3A*^{R882H} mutation into human HSPCs with high editing efficiencies. The prime
291 editing technique has important advantages over Cas9 nuclease-based genome editing strategies
292 that depend on the generation of DNA double-strand breaks (DSBs) which are highly toxic to
293 HSCs. Although prime editing can still induce a small amount of DSBs, it is less genotoxic and
294 can achieve high editing efficiencies in long-term repopulating HSPCs²². Our reported
295 methodology represents an important technical resource for the study of *DNMT3A*^{R882} mutations
296 in human HSPCs.

297 The presence of CH has been shown to be associated with an increased risk of developing not only
298 hematologic malignancies but also a growing list of age-related illnesses. Interventions that
299 effectively lower the risk of these adverse outcomes in CH carriers have the potential to positively
300 impact the health of a large segment of the aging population. However, this goal is not yet possible
301 due to the lack of known interventions that effectively suppress the expansion of mutant clones in
302 CH. The ideal preventive intervention should not only be effective but also easy-to-administer and
303 safe for long term use. Metformin fulfills these criteria and can readily be repurposed as a
304 preventive treatment for *DNMT3A*^{R882}-mutated CH carriers, especially those at high risk of
305 malignant transformation or other CH-related illness. Our findings provide the preclinical rationale
306 for studying this strategy in a prospective clinical trial.

307 **Author Contributions**

308 M.H. and S.M.C. conceived the study and designed the experiments. M.H. performed and analyzed
309 most experiments. V.V., A.C., and S.P. performed the bioinformatics analyses with assistance
310 from Y.W. and A.C.L. F.G., A.D.S. and G.D.B. provided input and supervised the bioinformatics
311 analyses. S.C., D.M.A., A.C.L., Y.Y., V.W., A.M. and E.G. performed experiments. J.A.R. and
312 A.D. performed and analyzed the metabolomics experiments. A.V. performed the prime editing
313 experiments under the supervision of S.Z.X. and J.E.D. M.F. designed the prime editing strategy
314 under the supervision of L.N. and S.F. M.H. and S.M.C. interpreted the data and wrote the
315 manuscript. K.Y. and J.J.T. reviewed and edited the manuscript. All authors provided input in the
316 preparation of the final manuscript.

317 **Acknowledgements**

318 This project was supported by funding from the Princess Margaret Cancer Foundation, a Medicine
319 by Design (Mbd) Award from the University of Toronto, a New Investigator Award from the
320 Leukemia Research Foundation, and a Project Grant (PJT-175186) from the Canadian Institutes
321 of Health Research. J.E.D is supported by funds from the: Princess Margaret Cancer Centre
322 Foundation, , Canadian Institutes for Health Research (RN380110 - 409786), Canadian Cancer
323 Society (grant #703212 (end date 2019), #706662 (end date 2025)), Terry Fox New Frontiers
324 Program Project Grant (Project# 1106), a Canada Research Chair, Princess Margaret Cancer
325 Centre, The Princess Margaret Cancer Foundation, and Ontario Ministry of Health. The authors
326 thank Rachel Culp-Hill, Ph.D. for metabolomics data acquisition. J.A.R. and A.D. acknowledge
327 support from the University of Colorado Cancer Center Support Grant (P30CA046934).

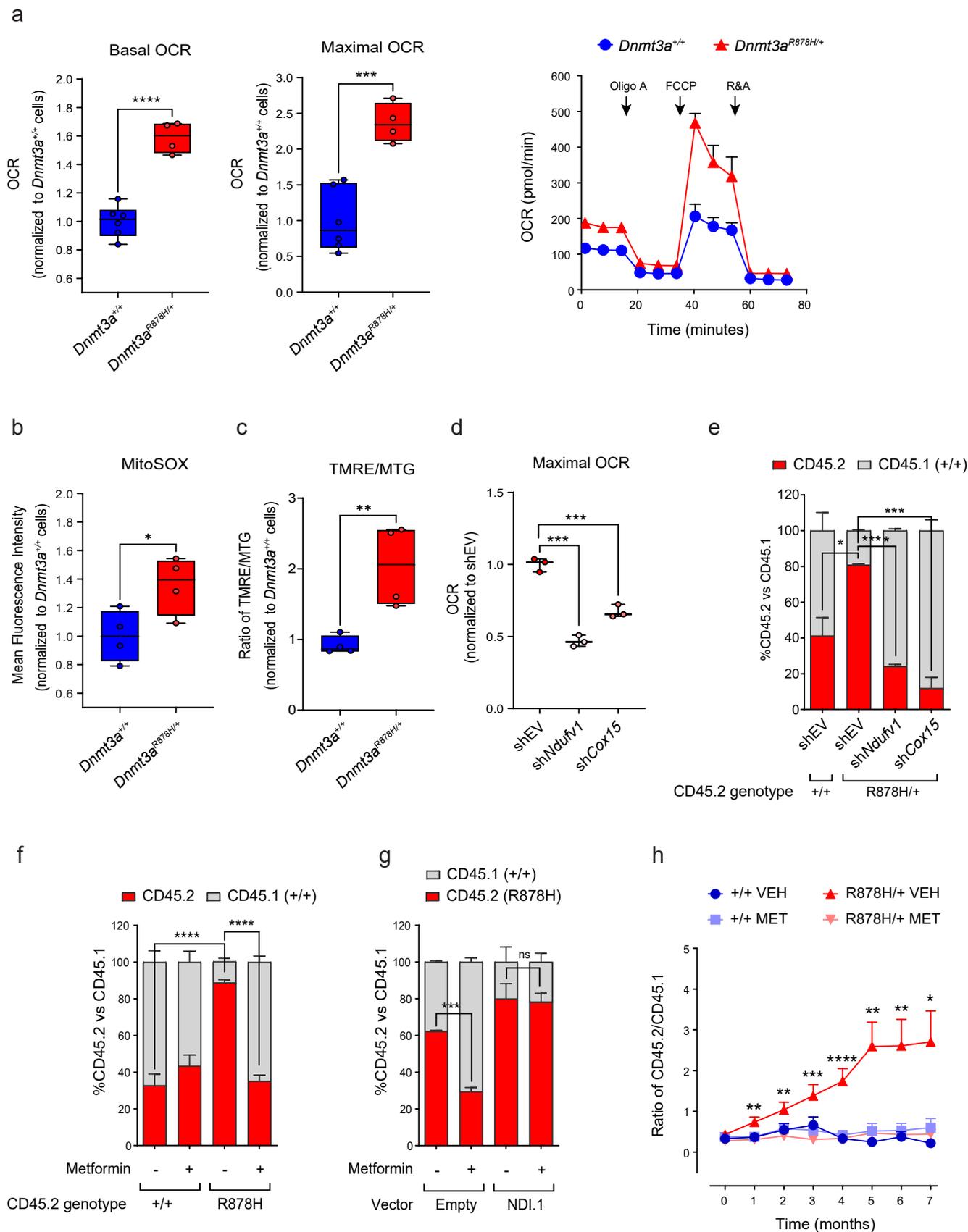
328

329 **Conflicts of Interest**

330 S.M.C. has received research funding from the Centre for Oncology and Immunology in Hong
331 Kong, Celgene/BMS, AbbVie Pharmaceuticals, Agios Pharmaceuticals, and Servier Laboratories.
332 F.G. serves as a consultant for S2 Genomics Inc. A.D.S. has received research funding from
333 Takeda Pharmaceuticals, BMS and Medivir AB, and consulting fees/honorarium from Takeda,
334 Novartis, Jazz, and Otsuka Pharmaceuticals. A.D.S. is named on a patent application for the use
335 of DNT cells to treat AML. A.D.S. is a member of the Medical and Scientific Advisory Board of
336 the Leukemia and Lymphoma Society of Canada. A.D.S. holds the Ronald N. Buick Chair in
337 Oncology Research. J.E.D. has received research funding from Celgene/BMS, and has patents
338 licensed to Trillium Therapeutics/Pfizer.

339

Figure 1



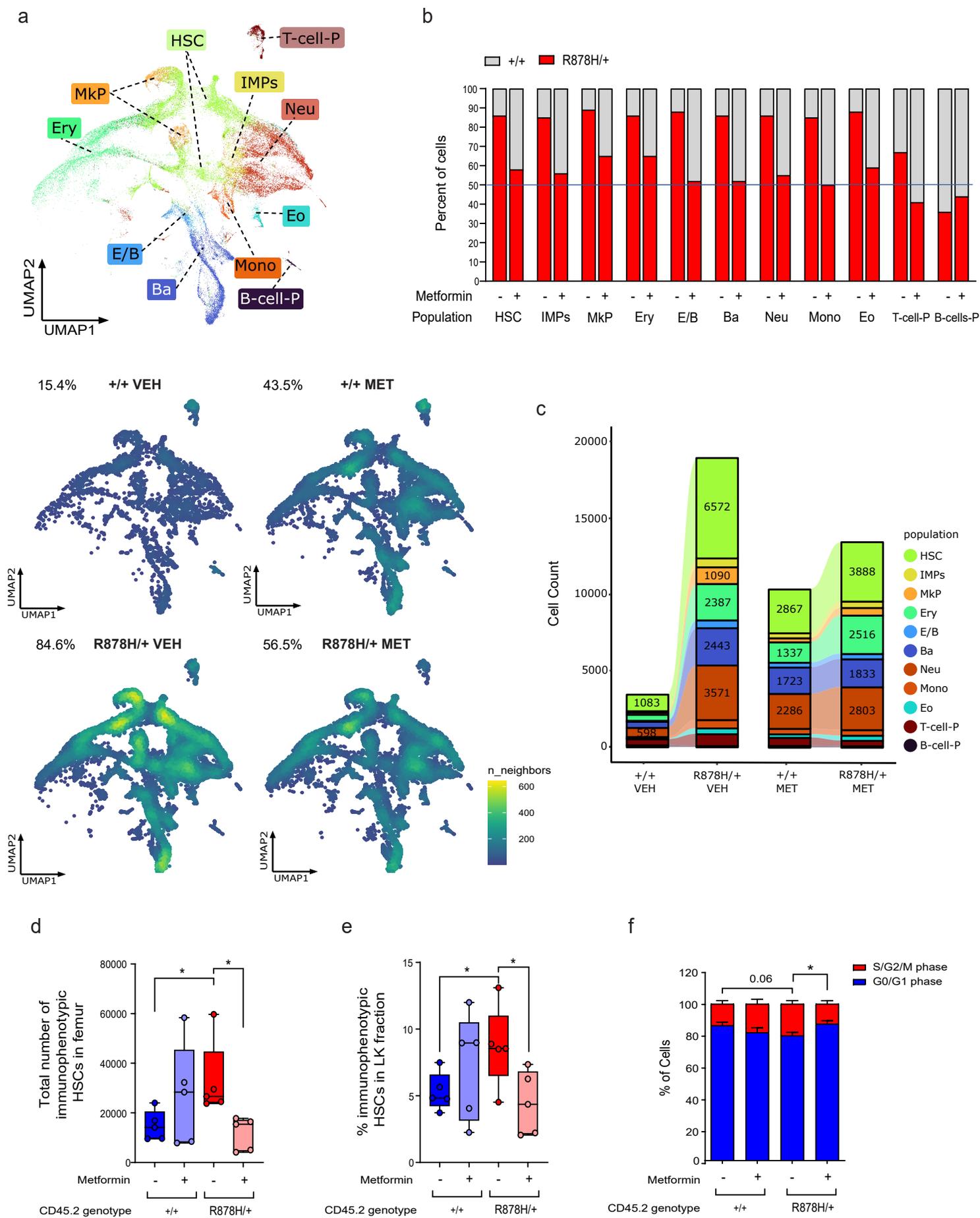
340 **Figure Legends**

341 **Fig. 1 | *Dnmt3a*^{R878H/+} HSPCs have increased mitochondrial respiration compared with**
342 ***Dnmt3a*^{+/+} cells and are dependent on this metabolic reprogramming for their competitive**
343 **advantage. **a**, Left panels, basal and maximal oxygen consumption rates (OCRs) in *Dnmt3a*^{+/+}**
344 **and *Dnmt3a*^{R878H/+} LK HSPCs. Dots represent technical replicates. Right panel, OCRs of LK cells**
345 **of the indicated genotype at baseline and at different time points following treatment with**
346 **oligomycin A (Oligo A), FCCP, and rotenone plus antimycin A (R&A). n=4-6 technical replicates**
347 **for each data point. Representative data of three independent experiments are shown. **b**, Mean**
348 **fluorescence intensity of MitoSOX staining in *Dnmt3a*^{+/+} and *Dnmt3a*^{R878H/+} LK HSPCs. Dots**
349 **represent samples from individual mice. **c**, Ratio of TMRE to MitoTracker Green (MTG) staining**
350 **in *Dnmt3a*^{+/+} and *Dnmt3a*^{R878H/+} LK HSPCs. Dots represent samples from individual mice. **d**,**
351 **Maximal OCR in *Dnmt3a*^{R878H/+} LK HSPCs transduced with an empty shRNA vector control**
352 **(shEV) or a shRNA vector expressing sh*Ndufv1* or sh*Cox15*. Dots represent technical replicates.**
353 ****e**, Proportion of CD45.2⁺ and CD45.1⁺ cells in a competition assay between CD45.2⁺ LK cells of**
354 **the indicated genotype and CD45.1⁺ *Dnmt3a*^{+/+} LK cells. Both populations were transduced with**
355 **the indicated shRNA vectors. n=3 technical replicates. **f**, Proportion of CD45.2⁺ and CD45.1⁺**
356 **in a competition assay between CD45.2⁺ LK cells of the indicated genotype and CD45.1⁺**
357 ***Dnmt3a*^{+/+} LK cells in the absence or presence of metformin at 50μM. n=3 technical replicates.**
358 **Representative data of three independent experiments are shown. **g**, Proportion of CD45.2⁺ and**
359 **CD45.1⁺ cells in a competition assay between CD45.2⁺ *Dnmt3a*^{R878H/+} LK cells and CD45.1⁺**
360 ***Dnmt3a*^{+/+} LK cells in the presence or absence of metformin at 50μM. The CD45.2⁺ *Dnmt3a*^{R878H/+}**
361 **LK cells were transduced with an empty or NDI.1 overexpressing lentiviral vector. n=3 technical**
362 **replicates. **h**, Ratio of CD45.2⁺ to CD45.1⁺ in peripheral blood cells collected from recipient mice**
363 **at the indicated time points after starting treatment with metformin in the drinking water at 5g/L**
364 **(MET) or no treatment (VEH). The mice were transplanted with CD45.1⁺ *Dnmt3a*^{+/+} bone marrow**
365 **cells and CD45.2⁺ bone marrow cells of the indicated genotype 5 weeks prior to starting drug**
366 **treatment. For months 0-4, data are from 3 independent experiments consisting of a total of 21-23**
367 **animals per condition. For months 5-7, data are from 1 experiment consisting of 6-7 animals per**
368 **condition. Statistical significance was calculated in comparison with the untreated (VEH) arm of**
369 **each genotype. In **a** (left panels), **b**, **c**, **d**, the box represents the interquartile range with the median**
370 **indicated by the line inside the box. Whiskers extend to the minimum and maximum values. In **a****

371 (right panel), **e, f, g, h**, data shown are mean \pm SEM. Statistical significance (P values) was
372 calculated using two-sided Student's t-test with * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, and ***
373 $P < 0.0001$. ns, not significant.

374

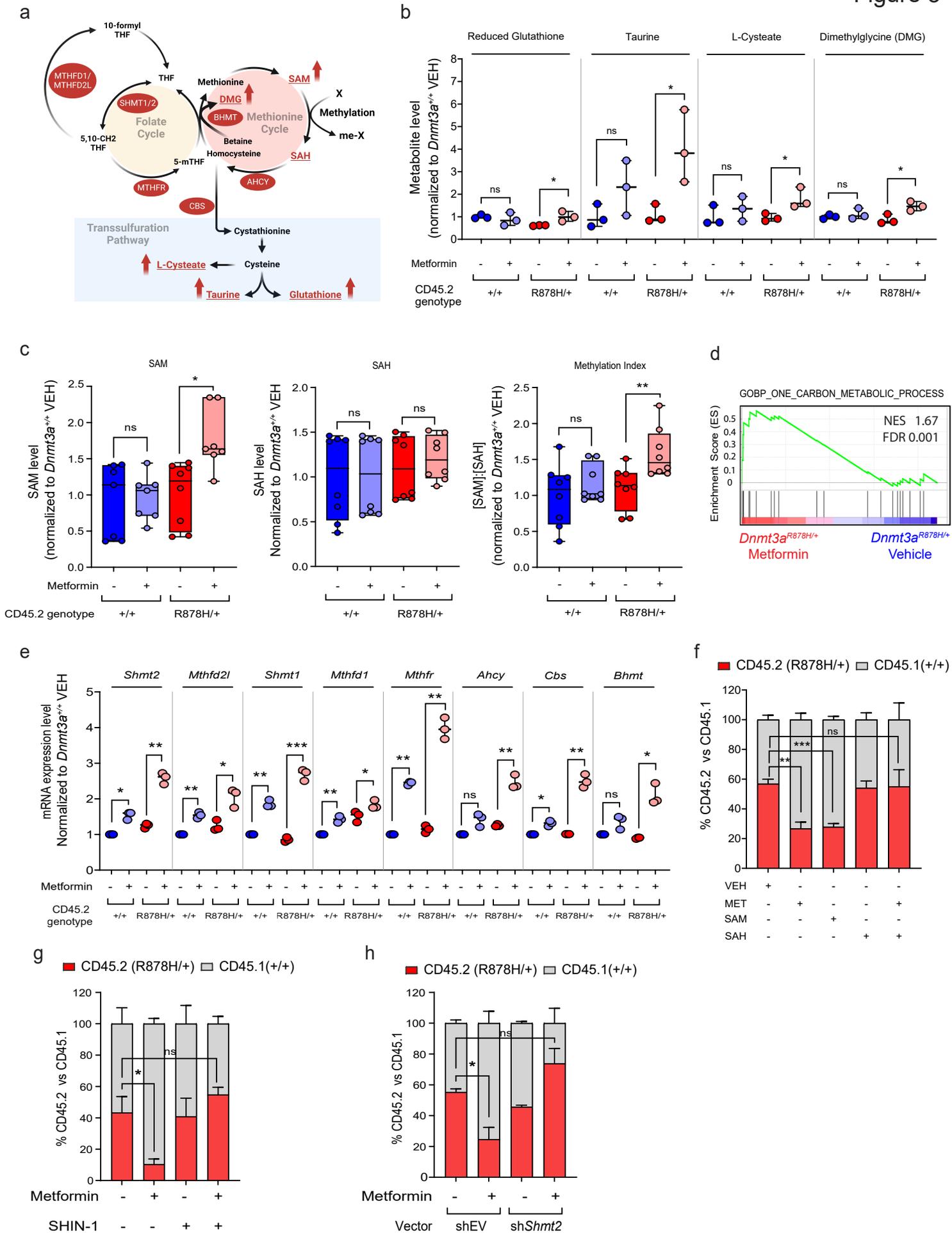
Figure 2



375 **Fig. 2 | Metformin suppresses the competitive advantage of *Dnmt3a*^{R878H/+} HSCs.** **a**, Top,
376 dimensionality reduction using Uniform Manifold Approximation and Projection (UMAP) on all
377 sequenced cells (n=46,225 cells). HSC = Hematopoietic stem cell; IMP = Immature myeloid
378 progenitor; Mono = Monocyte progenitor, Neu = Neutrophil/granulocyte progenitor; E/B =
379 Erythroid/basophil progenitor; Ery = Erythroid progenitor; MkP = Megakaryocyte progenitor; Ba
380 = Basophil progenitor; Eo = Eosinophil progenitor; B-cell-P = B cell progenitor; T-cell-P = T cell
381 progenitor. Bottom, UMAP cell density plots of CD45.1⁺ *Dnmt3a*^{+/+} cells vs. CD45.2⁺
382 *Dnmt3a*^{R878H/+} cells in LK-enriched BM samples collected from mice treated with vehicle (VEH)
383 or metformin (MET). **b**, Proportion of CD45.1⁺ *Dnmt3a*^{+/+} cells vs. CD45.2⁺ *Dnmt3a*^{R878H/+} cells
384 in each HSPC subset from untreated and metformin-treated LK samples. **c**, Sankey diagrams
385 showing the absolute number of sequenced cells in each HSPC subset among CD45.1⁺ *Dnmt3a*^{+/+}
386 vs. CD45.2⁺ *Dnmt3a*^{R878H/+} fractions in LK-enriched BM samples collected from mice treated with
387 vehicle (VEH) or metformin (MET). **d**, Number of immunophenotypic HSCs (Lin⁻, c-Kit⁺, Sca-
388 1⁺, CD150⁺, CD48⁻) in the right femur from mice transplanted with WBM cells of the indicated
389 genotype and treated with or without metformin for 1 month. Dots represent samples from
390 individual mice. **e**, Proportion of immunophenotypic HSCs in the LK fraction collected from mice
391 transplanted with WBM cells of the indicated genotype and treated with or without metformin for
392 1 month. Dots represent samples from individual mice. **f**, Proportion of immunophenotypic HSCs
393 in S/G2/M phase versus G0/G1 phase. Cells were collected from mice transplanted with WBM
394 cells of the indicated genotype and treated with or without metformin for 1 month. n=5 mice per
395 condition. In **d**, **e**, the box represents the interquartile range with the median indicated by the line
396 inside the box. Whiskers extend to the minimum and maximum values. In **f**, data shown are mean
397 ± SEM. Statistical significance (P values) was calculated using two-sided Student's t-test with *
398 P<0.05, ** P<0.01, *** P< 0.001, and **** P<0.0001.

399

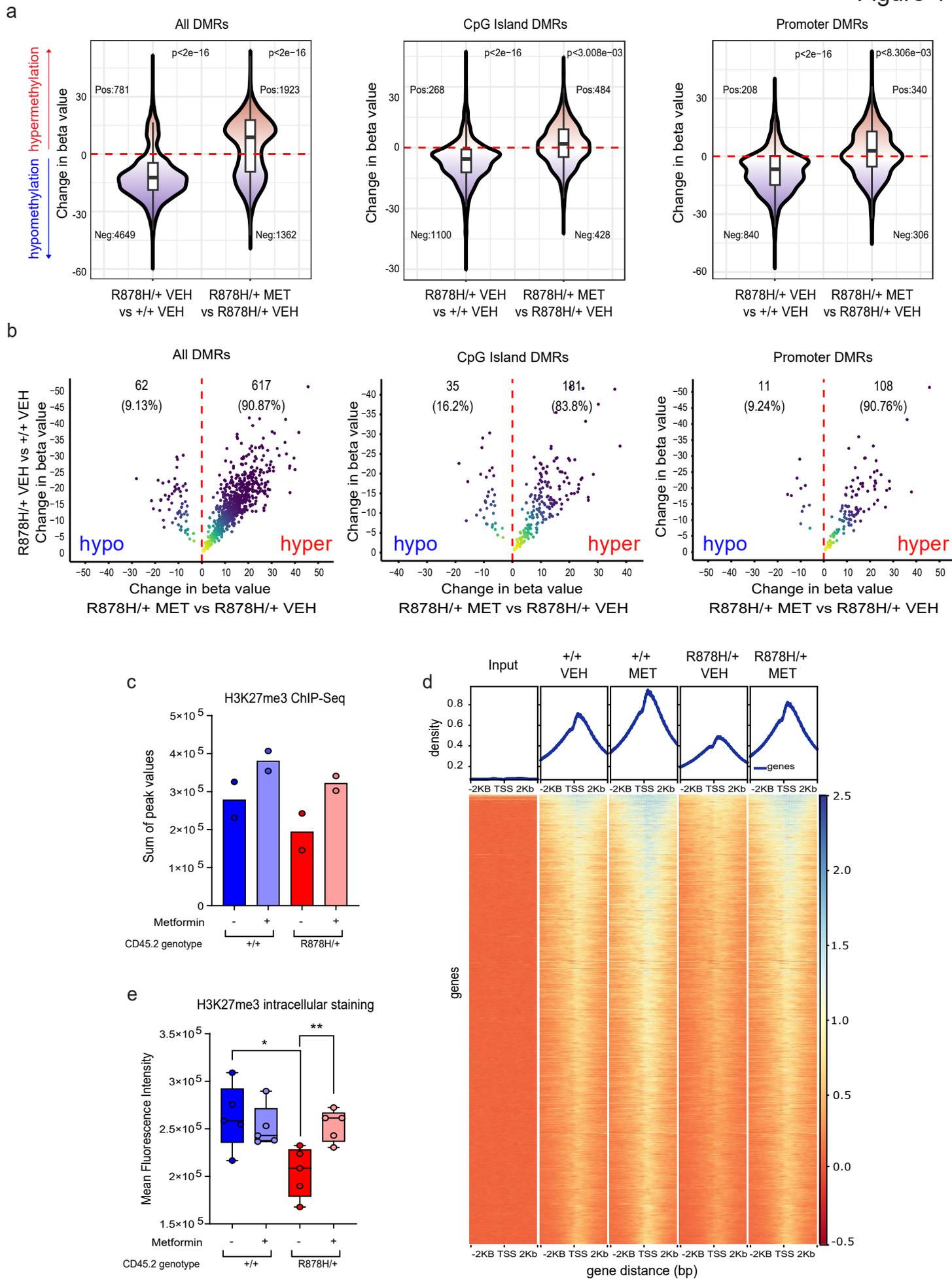
Figure 3



400 **Fig. 3 | Metformin suppresses the fitness of *Dnmt3a*^{R878H/+} HSPCs by enhancing their**
401 **methylation potential. a**, Schematic diagram of the metabolic pathways involved in 1C
402 metabolism. **b**, Levels of the indicated metabolites in LK cells isolated from mice transplanted
403 with BM cells of the indicated genotype. The animals were either left untreated or treated with
404 metformin for 1 month. Dots represent samples from individual mice. **c**, Levels of SAM and SAH
405 and the ratio of [SAM]:[SAH] in LK cells isolated from mice transplanted with BM cells of the
406 indicated genotype. The animals were either left untreated or treated with metformin for 1 month.
407 Dots represent samples from individual mice. **d**, Gene set enrichment plot of bulk RNA-seq data
408 comparing metformin-treated *Dnmt3a*^{R878H/+} LK cells (n=2 biological replicates) versus vehicle-
409 treated *Dnmt3a*^{R878H/+} LK cells (n=2 biological replicates) using the indicated gene set
410 (GO:0006730). **e**, Expression level of the indicated genes by RT-qPCR in LK cells isolated from
411 mice transplanted with BM cells of the indicated genotype. The animals were either left untreated
412 or treated with metformin for 1 month. Dots represent samples from individual mice. **f**, Proportion
413 of CD45.2⁺ and CD45.1⁺ cells in a competition assay between CD45.2⁺ *Dnmt3a*^{R878H/+} LK cells
414 and CD45.1⁺ *Dnmt3a*^{+/+} LK cells in the presence or absence of the indicated compounds. n=4
415 technical replicates. Representation data of 4 independent experiments are shown. **g**, Proportion
416 of CD45.2⁺ and CD45.1⁺ cells in a competition assay between CD45.2⁺ *Dnmt3a*^{R878H/+} LK cells
417 and CD45.1⁺ *Dnmt3a*^{+/+} LK cells in the presence or absence of the indicated compounds. n=4
418 technical replicates. Representation data from 3 independent experiments. **h**, Proportion of
419 CD45.2⁺ and CD45.1⁺ cells in a competition assay between CD45.2⁺ *Dnmt3a*^{R878H/+} LK cells and
420 CD45.1⁺ *Dnmt3a*^{+/+} LK cells in the presence or absence of metformin. Both populations were
421 transduced with the indicated shRNA vectors. n=3 technical replicates. Representation data of 3
422 independent experiments are shown. In **b**, **c**, **e**, the box represents the interquartile range with the
423 median indicated by the line inside the box. Whiskers extend to the minimum and maximum
424 values. In **f**, **g**, **h**, data shown are mean ± SEM. Statistical significance (P values) was calculated
425 using two-sided Student's t-test for all comparisons except for **b** where one-sided Student's t-test
426 was used. * P<0.05, ** P<0.01, *** P< 0.001, and **** P<0.0001. ns, not significant.

427

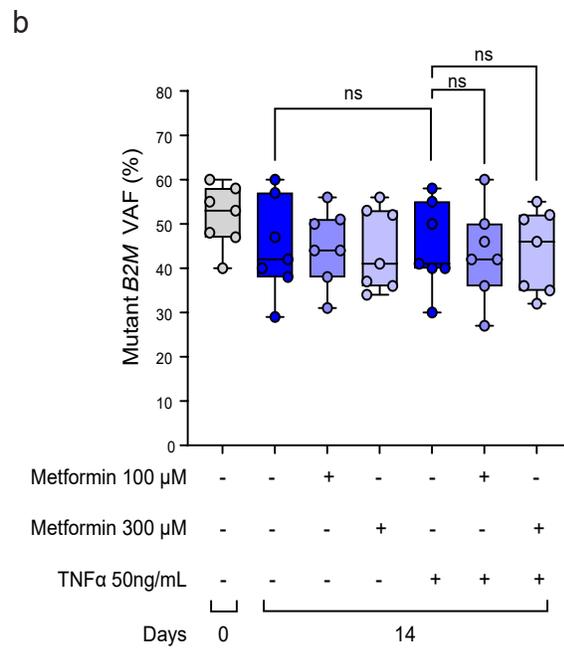
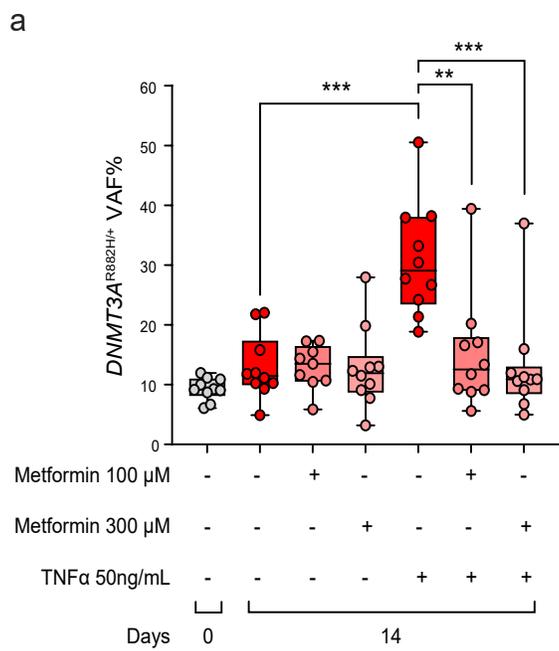
Figure 4



428 **Fig. 4 | Metformin reverses the aberrant DNA CpG methylation and H3K27me3 profiles in**
429 ***Dnmt3a*^{R878H/+} HSPCs. a,** Violin plots of the difference in beta values at all differentially-
430 methylated regions (DMRs), CpG island-associated DMRs, or promoter-associated DMRs in the
431 comparison between untreated *Dnmt3a*^{R878H/+} LK samples versus untreated *Dnmt3a*^{+/+} LK samples
432 (left) and between metformin-treated *Dnmt3a*^{R878H/+} LK samples versus untreated *Dnmt3a*^{R878H/+}
433 LK samples (right). n=3 biological replicates for untreated *Dnmt3a*^{+/+} and metformin-treated
434 *Dnmt3a*^{R878H/+} samples. n=4 biological replicates for metformin-treated *Dnmt3a*^{+/+} and untreated
435 *Dnmt3a*^{R878H/+} samples. The P values adjacent to the plots were calculated using the one-sample
436 Wilcoxon signed rank test determine if the median difference in beta values was significantly
437 different from 0. **b,** Plot showing the change in beta values at overlapping DMRs between
438 metformin-treated *Dnmt3a*^{R878H/+} samples versus untreated *Dnmt3a*^{R878H/+} samples on the X-axis
439 and between untreated *Dnmt3a*^{R878H/+} samples versus untreated *Dnmt3a*^{+/+} samples on the Y-axis.
440 **c,** Sum of peak values from H3K27me3 ChIP-seq analysis of LK HSPC samples collected from
441 mice transplanted with bone marrow cells of the indicated genotype and treated with or without
442 metformin for 1 month. Dots represent samples from individual mice. The means of the 2
443 biological replicates are shown. **d,** Distribution of H3K27me3 signals surrounding (± 2 KB) the
444 transcription start site (TSS) regions with the highest signals (n=10,622) in the indicated samples.
445 **e,** Mean fluorescent intensity of H3K27me3 staining by intracellular flow cytometry of LK HSPCs
446 collected from mice transplanted with bone marrow cells of the indicated genotype and treated
447 with or without metformin for 1 month. Dots represent samples from individual mice. In **e,** the box
448 represents the interquartile range with the median indicated by the line inside the box. Whiskers
449 extend to the minimum and maximum values. Statistical significance (P values) was calculated
450 using two-sided Student's t-test with * P<0.05, ** P<0.01, *** P< 0.001, and **** P<0.0001.

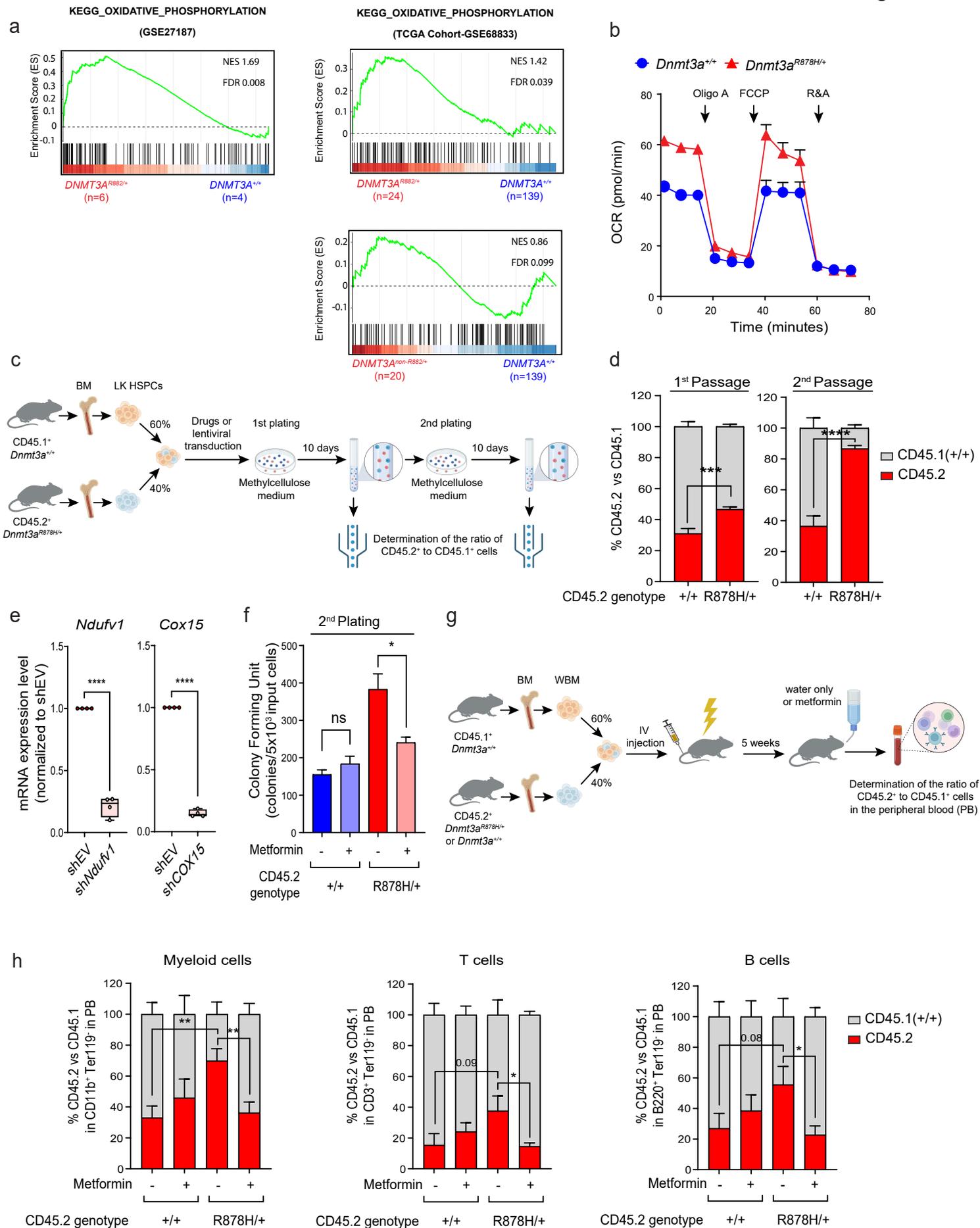
451

Figure 5



452 **Fig. 5 | Metformin decreases the competitive advantage of human *DNMT3A*^{R882H} HSPCs. a,**
453 *DNMT3A*^{R882H} variant allele frequencies (VAFs) of prime edited human HSPCs at baseline (day
454 0) and after 14 days in culture in the presence or absence of TNF α or metformin. **b,** Mutant *B2M*
455 VAFs of prime edited human HSPCs at baseline (day 0) and after 14 days in culture in the presence
456 or absence of TNF α or metformin. Dots represent samples from individual cord blood donors. The
457 box represents the interquartile range with the median indicated by the line inside the box.
458 Whiskers extend to the minimum and maximum values. Statistical significance (P values) was
459 calculated using two-sided Student's t-test with * P<0.05, ** P<0.01, *** P< 0.001, and ***
460 P<0.0001. ns, not significant.

461



462 **Extended Data Fig. 1 | *Dnmt3a*^{R878H/+} HSPCs have increased mitochondrial respiration**
463 **compared with *Dnmt3a*^{+/+} cells and are dependent on this metabolic reprogramming for their**
464 **competitive advantage. a**, Gene set enrichment plots comparing *DNMT3A*^{R882} or *DNMT3A*^{non-}
465 ^{R882} mutated AML samples versus *DNMT3A*^{WT} AML samples using two publicly available gene
466 expression datasets (GSE27187 and GSE68833). **b**, OCRs of whole bone marrow cells of the
467 indicated genotype at baseline and at different time points following treatment with oligomycin A
468 (Oligo A), FCCP, and rotenone plus antimycin A (R&A). n=4-6 technical replicates for each data
469 point. Representative data of 3 independent experiments are shown. **c**, Schematic diagram showing
470 the design of the *in vitro* competition assay. **d**, Proportion of CD45.2⁺ and CD45.1⁺ cells in a
471 competition assay between CD45.2⁺ LK cells of the indicated genotype and CD45.1⁺ *Dnmt3a*^{+/+}
472 LK cells after the 1st passage (n=8 technical replicates) and 2nd passage (n=13 technical replicates).
473 **e**, Expression of the indicated genes in *Dnmt3a*^{+/+} LK cells transduced with an empty shRNA
474 vector control (shEV) or a shRNA vector expressing sh*Ndufv1* or sh*Cox15*. Dots represent
475 technical replicates. **f**, Number of colony forming units in the second plating from *Dnmt3a*^{+/+} or
476 *Dnmt3a*^{R878H/+} LK HSPCs in the absence or presence of metformin. n=3 technical replicates per
477 condition. Representative data of 3 independent experiments are shown. **g**, Schematic diagram
478 showing the design of the *in vivo* competitive repopulation experiment. **h**, Proportion of CD45.2⁺
479 vs. CD45.1⁺ cells in the myeloid (CD11b⁺), T (CD3⁺), and B (B220⁺) cell compartments in
480 peripheral blood cells collected from mice after 4 months of treatment with metformin or vehicle
481 in the experiment shown in Fig. 1h. n=7 mice per condition. In **e**, the box represents the
482 interquartile range with the median indicated by the line inside the box. Whiskers extend to the
483 minimum and maximum values. In **b**, **d**, **f**, **h**, data shown are mean ± SEM. Statistical significance
484 (P values) were calculated using two-sided Student's t-test for all comparisons. * P<0.05, **
485 P<0.01, *** P< 0.001, and **** P<0.0001. ns, not significant.

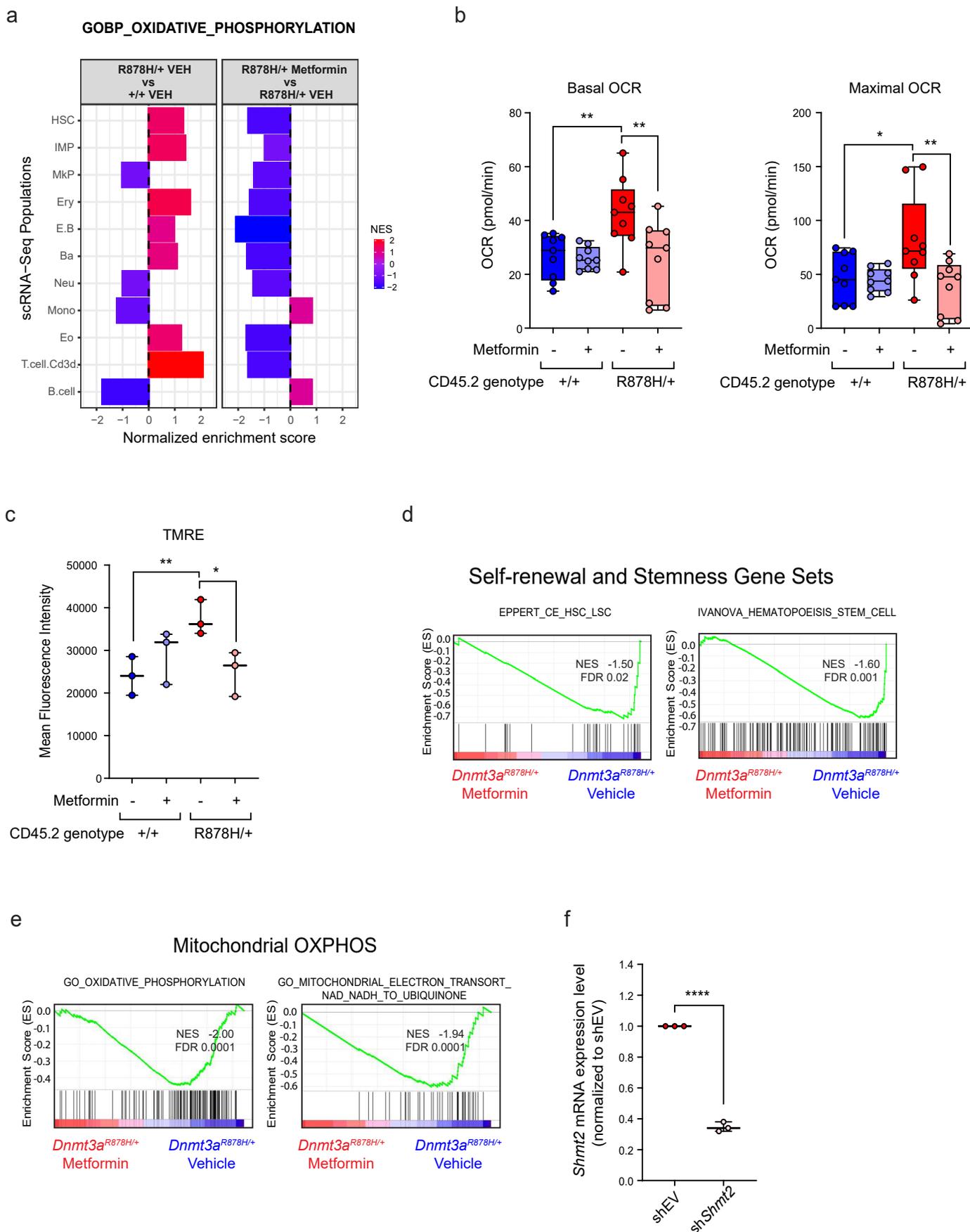
486

a



487 **Extended Data Fig. 2 | Metformin suppresses the competitive advantage of *Dnmt3a*^{R878H/+}**
488 **HSCs. a,** Schematic diagram showing the design of the non-competitive repopulation experiment
489 in which lethally-irradiated recipient mice were transplanted with CD45.2⁺ *Dnmt3a*^{+/+} or
490 *Dnmt3a*^{R878H/+} whole bone marrow (WBM) cells. Five weeks after transplantation, the mice were
491 either left untreated or treated with metformin in the drinking water. After 4 weeks of treatment,
492 BM cells were collected and used for analysis.

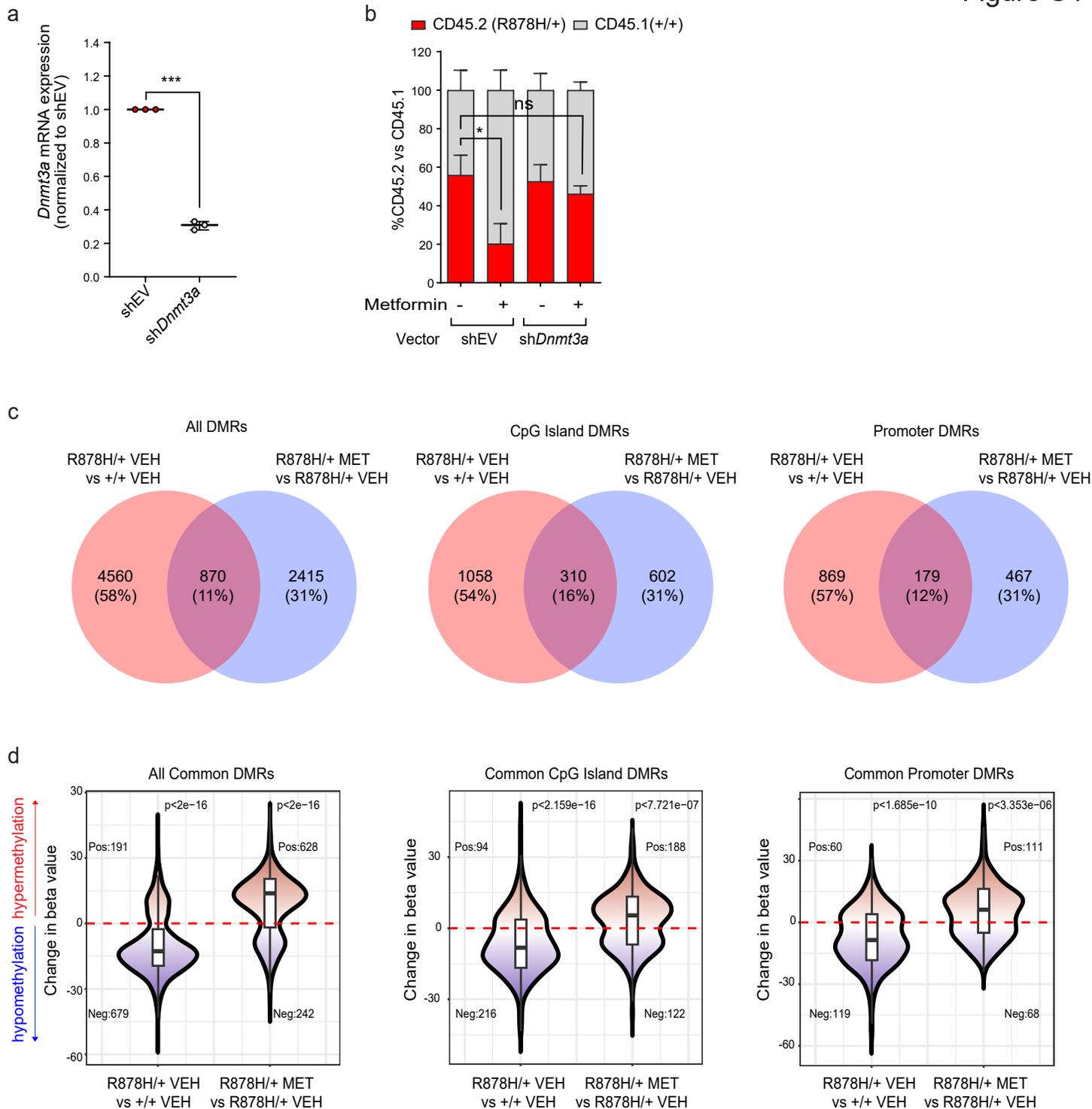
493



494 **Extended Data Fig. 3 | Metformin suppresses the fitness of *Dnmt3a*^{R878H/+} HSPCs by**
495 **enhancing their methylation potential. a**, Normalized enrichment scores (NES) for the indicated
496 gene set (GO: 0006119) in each HSPC subset using scRNA-seq gene expression data. **b**, Basal and
497 maximal OCRs in LK HSPCs collected from mice transplanted with WBM cells of the indicated
498 genotype and treated with or without metformin for 1 month. Dots represent samples from
499 individual mice. **c**, Mean fluorescence intensity of TMRE staining in LK HSPCs collected from
500 mice transplanted with WBM cells of the indicated genotype and treated with or without
501 metformin for 1 month. Dots represent samples from individual mice. **d**. Gene set enrichment plots
502 of bulk RNA-seq data comparing metformin-treated *Dnmt3a*^{R878H/+} LK cells (n=2 biological
503 replicates) versus vehicle-treated *Dnmt3a*^{R878H/+} LK cells (n=2 biological replicates) using the
504 indicated self-renewal and stemness-related gene sets. **e**. Gene set enrichment plots of bulk RNA-
505 seq data comparing metformin-treated *Dnmt3a*^{R878H/+} LK cells (n=2 biological replicates) versus
506 vehicle-treated *Dnmt3a*^{R878H/+} LK cells (n=2 biological replicates) using the indicated
507 mitochondrial function-related gene sets. **f**, Expression of *Shmt2* in *Dnmt3a*^{+/+} LK cells transduced
508 with an empty shRNA vector control (shEV) or a shRNA vector expressing sh*Shmt2*. n=3 technical
509 replicates per condition. In **b**, **c**, **f**, the box represents the interquartile range with the median
510 indicated by the line inside the box. Whiskers extend to the minimum and maximum values.
511 Statistical significance (P values) was calculated using two-sided Student's t-test for all
512 comparisons. * P<0.05, ** P<0.01, *** P< 0.001, and **** P<0.0001.

513

Figure S4



514 **Extended Data Fig. 4 | Metformin reverses the aberrant DNA CpG methylation and**
515 **H3K27me3 profiles in *Dnmt3a*^{R878H/+} HSPCs. a,** Expression of *Dnmt3a* in *Dnmt3a*^{+/+} LK cells
516 transduced with an empty shRNA vector control (shEV) or a shRNA vector expressing sh*Dnmt3a*.
517 n=3 technical replicates per condition. **b,** Proportion of CD45.2⁺ and CD45.1⁺ cells in a
518 competition assay between CD45.2⁺ *Dnmt3a*^{R878H/+} LK cells and CD45.1⁺ *Dnmt3a*^{+/+} LK cells in
519 the presence or absence of metformin. Both populations were transduced with the indicated
520 shRNA vectors. n=4-6 technical replicates. Representation data of 3 independent experiments are
521 shown. **c,** Venn diagram showing the overlap of DMRs between metformin-treated *Dnmt3a*^{R878H/+}
522 samples versus untreated *Dnmt3a*^{R878H/+} samples and between untreated *Dnmt3a*^{R878H/+} samples
523 versus untreated *Dnmt3a*^{+/+} samples. **d,** Violin plots of the difference in beta values at the
524 overlapping DMRs in the comparison between untreated *Dnmt3a*^{R878H/+} LK samples versus
525 untreated *Dnmt3a*^{+/+} LK samples (left) and between metformin-treated *Dnmt3a*^{R878H/+} LK samples
526 versus untreated *Dnmt3a*^{R878H/+} LK samples (right). The P values adjacent to the plots were
527 calculated using the one-sample Wilcoxon signed rank test to determine if the median difference
528 in beta values was significantly different from 0. In **a**, the box represents the interquartile range
529 with the median indicated by the line inside the box. Whiskers extend to the minimum and
530 maximum values. In **b**, data shown are mean ± SEM. Statistical significance (P values) were
531 calculated using two-sided Student's t-test for all comparisons. * P<0.05, ** P<0.01, *** P< 0.001,
532 and *** P<0.0001. ns, not significant.

533 **Extended Table 1 |** Metabolite levels in metabolomics analysis

534 **Extended Table 2 |** List of RNA sequences used for prime editing

535 **Extended Table 3 |** List of primer and oligonucleotides sequences used in the study

536 **Extended Table 4 |** List of antibodies used in the study

537 **Extended Table 5 |** List of chemicals and reagents used in the study

538 **Materials and Methods**

539 **Human cord blood samples**

540 Cord blood (CB) CD34⁺ HSPCs were obtained with informed consent from Trillium Health, Credit
541 Valley and William Osler Hospitals according to procedures approved by the University Health
542 Network (UHN) Research Ethics Board. The mononuclear cells (MNC) were separated by
543 centrifugation using Lymphoprep medium and then dissolved using ammonium chloride. The
544 CD34⁺ cells were then isolated using the CD34 Microbead kit and purified using LS columns,
545 following the directions provided by the manufacturer (Miltenyi). The CD34⁺ CB cells were
546 preserved in a solution containing 50% PBS, 40% fetal bovine serum (FBS), and 10% DMSO at
547 temperatures of -80°C or -150°C.

548 **Primary cell culture**

549 CB CD34⁺ HSPCs were thawed via dropwise addition of X-VIVOTM-10 media with 50% FBS and
550 DNaseI (200 µg/ml). Cells were centrifuged at 350g for 10min at room temperature and seeded at
551 the concentration of 5x10⁵ cells per ml in serum-free StemSpan medium (StemCell Technologies)
552 supplemented with 2% glutamine, 100 ng/mL hSCF (R&D), 100 ng/mL hFlt3-L, 20 ng/mL hTPO,
553 1 µM SR1 (StemCellTechnologies) and 50nM UM171 (MedChemExpress LLC). Cells were
554 cultured in a 5% CO₂ humidified atmosphere at 37 °C.

555 **mRNA in vitro transcription**

556 The PE3max plasmid was used to synthesize the mRNA encoding nCas9-RT through *in vitro*
557 transcription, as described in Fuimara *et al*²³. Briefly, plasmid was linearized with SpeI (New
558 England Biolabs) and purified by phenol-chloroform extraction. mRNAs were transcribed in vitro
559 (5X MEGAscript T7 kit, Thermo Fisher), capped with 8 mM CleanCapAG (Trilink), purified
560 (RNeasy Plus Mini Kit, Qiagen) and quality assessed by capillary electrophoresis (4200
561 TapeStation System, Agilent)²³. mRNAs were resuspended in RNase free water and stored at -80
562 °C.

563 **Gene editing of human HSPCs and analysis**

564 A total of 1x10⁵–5x10⁵ cells were rinsed with PBS and subjected to electroporation using the P3
565 Primary Cell 4D-Nucleofector X Kit and Nucleofector 4D device, with program EO-100 from

566 Lonza after 1 or 3 days of culture. The electroporation mixture consisted of 180 pM sgRNA from
567 Synthego, 270 pM pegRNA from IDT, and 12µg PE3max mRNA. Following electroporation, cells
568 were allowed to recuperate for 3 minutes at room temperature and subsequently maintained in
569 culture according to prior findings. Three days following electroporation, CD34⁺ cells were
570 harvested to obtain genomic gDNA for molecular investigation. The sgRNA sequences can be
571 found in (Supplementary Excel Table 2). Standard pegRNA for the *B2M* positive control as
572 described in Fuimara *et al*²³. Engineered pegRNA (epgRNA) targeting *DNMT3A* with a
573 protective linker and motif at the 3' end from guide degradation, were designed with pegFinder
574 (<http://pegfinder.sidichenlab.org/>) and pegLIT (<https://peglit.liugroup.us/>)²⁸.

575 **Human clonal competition assay**

576 A clonal competition assay was conducted 24 hours after the editing procedure by placing 1000
577 cells per milliliter in a methylcellulose-based medium (MethoCultOptimum H4034, StemCell
578 Technologies). The medium was supplemented with 10 ng/mL of hIL-6, and 10 ng/mL of hFlt3L.
579 Each condition was subjected to four technical duplicates. After a period of two weeks following
580 the plating process, the cells were collected and obtained for molecular analysis.

581 **Molecular analysis**

582 The genomic DNA (gDNA) was extracted using the QIAamp DNA Micro Kit (Quiagen) from the
583 pellet of in vitro cultured cells, or with QuickExtract (Epicentre) from cultivated cells in
584 MethoCultOptimum H4034, following the instructions provided by the manufacturer. The
585 efficiency of B2M PE was assessed using Sanger sequencing (performed by Eurofins Scientific)
586 and the EditR software (available at <http://baseeditr.com>). To adapt EditR for B2M prime editing,
587 we used as input the sequence TGGCCTTAGCTGTGCTCGC and selected the reverse
588 complement orientation option as described in Fuimara *et al*²³. The efficiency of DNMT3A R882H
589 PE was assessed by ddPCR. QX200 Droplet Digital PCR System was used to examine 10–50 ng
590 of gDNA for in vitro samples and 4 µl of gDNA for colonies in accordance with the manufacturer's
591 instructions. The VAF was calculated as the number of FAM-positive droplets divided by total
592 droplets containing a target. The primers and probes are enumerated in (Supplementary Table 3).

593 **Mouse model and *in vivo* experiments**

594 All *in vivo* experiments were performed in accordance with institutional guidelines approved by
595 the University Health Network Animal care committee. C57BL/6J mice, also referred to as
596 B6.CD45.2, and B6.SJL-Ptprca Pepcb/BoyJ mice, known as B6.CD45.1, were obtained from The
597 Jackson Laboratory and held in the same facility for the duration of the study. The *Dnmt3a*<sup>fl-
598 R878H/+</sup> mice (JAX stock #032289) were crossbred with B6. CgTg(Mx1-cre)1Cgn/J mice (JAX
599 stock #003556) and genotyped as described by Jackson Laboratory. All mice were female, and
600 experiments initiated at 8–12 weeks of age. Mice were injected five times (once every other day)
601 via intraperitoneal (IP) injection with 15 mg/kg high molecular weight polyinosinic-polycytidylic
602 acid (pIpC) (Sigma-Aldrich ref: P1530) to induce Mx1-Cre recombinase expression. Before and
603 after pIpC administration, genomic DNA was extracted from PB cells for recombination PCR.
604 Primers used for genotyping are enumerated in (Supplementary Table 3). In addition, RNA was
605 extracted, and cDNA synthesized from whole BM cells for Sanger sequencing to verify mutant
606 allele expression.

607 CD45.2 *Dnmt3a*^{R878/+} or *Dnmt3a*^{+/+} BM cells (1E6 cells per mouse) were resuspended in Opti-
608 MEM medium and transplanted by tail vein injection into 10 weeks old female CD45.2 *Dnmt3a*
609 ^{+/+} recipient mice conditioned with 12Gy of irradiation. For *in vivo* competition experiments,
610 CD45.2 *Dnmt3a*^{R878/+} or *Dnmt3a*^{+/+} BM cells were mixed with CD45.1 *Dnmt3a*^{+/+} BM cells from
611 sex and age matched donor at a 1:2 ratio prior to transplantation. After five weeks following the
612 transplantation, each group was subdivided before starting treatment with metformin at 5mg/ml in
613 drinking water. Drinking water was replaced twice a week for the indicated period.

614 **Isolation of HSPCs from murine bone marrow**

615 The bone marrow (BM) cells were enriched for Lin⁻Kit⁺ cells (HSPCs) using the EasySep mouse
616 hematopoietic progenitor isolation kit (StemCell Technologies, Cat# 19856), followed by further
617 enrichment using the c-KIT positive enrichment kit (StemCell Technologies, Cat# 18757).

618 **Murine *in vitro* competition assays**

619 Competition experiments were conducted using 96-well flat-bottom tissue culture plates (Corning,
620 Ref#351172). The bone marrow cells from 4–5-month-old mice were harvested after 5 weeks of
621 post pIpC injection. HSPCs were enriched from BM cells as described above. The CD45.2
622 *Dnmt3a*^{+/+} HSPCs and CD45.2 *Dnmt3a*^{R878H/+} HSPCs (competing cells), were combined with

623 CD45.1 *Dnmt3a*^{+/+} HSPCs in a proportion of 40% and 60%, respectively. The cell mixture was
624 added to mouse MethoCult™ GF M3434 medium (StemCell Technologies, Cat # 3434) with a
625 density of 200 cells per well, treatment was administered as indicated. Cells were incubated at
626 37°C with 5% CO₂ for 10 days.

627 **Colony-forming unit (CFU) assays**

628 A total of 3x10³ murine HSPC enriched cells were mixed with 1.1 mL of MethoCult™ GF M3434
629 medium, and metformin was administered at 100uM. Subsequently, the cell suspension was
630 transferred to the wells of a 6-well tissue culture plate. After 10 days of culture, the colony
631 formation was examined, with the potential to replat the cells as indicated.

632 **Flow cytometry**

633 All flow cytometry analyses were conducted utilizing a Beckman Coulter CytoFLEX instrument,
634 Cells were stained for 30 min at 4 °C with antibodies (listed in Supplementary Table 4) at the
635 suggested dilutions in 100ul of FACS buffer (HBSS supplemented with 2% FBS and 0.1% sodium
636 azide) and wash once prior to flow cytometry analysis. FCS files analysis was performed with
637 FlowJo™ V10 software.

638 **Cell cycle assay**

639 Murin enriched HSPC were stained for surface markers as described above prior to asses cell cycle
640 via intracellular flow cytometry staining. The BD Cytofix/Cytoperm fixation and permeabilization
641 reagent was utilized in accordance with the instructions provided by the manufacturer. Briefly,
642 cells that had been fixed and permeabilized were stained for one hour at room temperature with a
643 KI67 antibody diluted 1:200 in HBSS supplemented with 2% FBS and 0.1% sodium azide. The
644 cells were subsequently rinsed in FACS buffer and subjected to staining with DAPI diluted 1:1,000
645 in PBS containing 1% FBS and 50mM EDTA at room temperature for 30 minutes. An IgG isotype
646 antibody was tested for evaluating the level of isotype control staining.

647 **Quantification of mitochondrial reactive oxygen species (ROS) levels, mitochondrial** 648 **membrane potential ($\Delta\Psi$ m), and mitochondrial mass**

649 The levels of mitochondrial ROS, mitochondrial membrane potential ($\Delta\Psi$ m), and mitochondrial
650 mass in freshly isolated mouse BM cells or HSPCs were assessed through flow cytometry utilizing

651 MitoSOX™ Red reagent, tetramethylrhodamine ethyl ester (TMRE), and MitoTracker Green
652 (MTG) probes (chemical reagents listed in Supplemental Table 5). The reagents were added
653 directly to the cells in FACS buffer, resulting in final concentrations of 5 μ M, 100 nM, and 200
654 nM, respectively. The cells were thereafter placed in an incubator set at a temperature of 37°C for
655 20 minutes. After the incubation period, the cells were washed with PBS 1X and then stained with
656 Annexin V conjugated with Alexa Fluor 647 or Sytox™ Blue.

657 **Seahorse Analyzer-Mitostress test**

658 All tests were carried out using the XF96 Extracellular Flux Analyzer from Seahorse Bioscience
659 (North Billerica, MA). The sensor cartridge was hydrated overnight in a non-CO₂ incubator using
660 the calibration buffer medium supplied by Seahorse Biosciences (200 μ l of buffer per well) on the
661 day prior to the assay. The wells of Seahorse XFe96 microplates were coated the following day at
662 4°C with a 40 μ l solution of Cell-Tak (Corning; Cat#354240) containing 22.4 μ g/ml. The Cell-
663 Tak-coated Seahorse microplate wells were subsequently rinsed with distilled water.

664 For cell plating, all cells were seeded at a density of 3×10^5 cells per well on Seahorse XFe96
665 microplates, using XF base minimal DMEM media containing 11 mM glucose, 2 mM glutamine,
666 and 1 mM sodium pyruvate. Following cell seeding, 180 μ l of XF base minimal DMEM medium
667 was added to each well, and the plate was centrifuged at 100 g for 5 minutes. Following a one-
668 hour incubation of the seeded cells at 37°C in a non-CO₂ incubator, the oxygen consumption rate
669 (OCR) and extracellular acidification rate (ECAR) were evaluated under the baseline and in
670 response to 1 μ M Oligomycin, 1 μ M carbonylcyanide-4-(trifluoromethoxy)-phenylhydrazone
671 (FCCP), and 1 μ M Antimycin and Rotenone (all from Sigma-Aldrich) using the XFe96 analyzer.

672 **RNA expression by RT-qPCR**

673 mRNA was extracted from cells using Qiagen RNeasy plus kit and quantified on a Nanodrop
674 spectrophotometer. Reverse transcription and quantitative PCR were performed at once using
675 Luna® Universal One-step RT-qPCR buffer and enzyme (NEB #E3005S). All qPCR experiments
676 were done on a Bio-Rad CFX touch real-time PCR detection system.

677 **Bulk RNA Sequencing (RNA-seq)**

678 RNA extraction from enriched murine HSPC was performed using the RNeasy Plus Mini Kit
679 (QIAGEN, Cat #74136) following the instructions provided by the manufacturer. RNA samples
680 were processed by Novogene Corporation in Sacramento, USA for sequencing analysis. Libraries
681 were generated using the NEBNext Ultra RNA Library Prep Kit for Illumina, and sequencing was
682 performed using the NovaSeq 6000 S4 with PE150 BP sequencing system. The readings were
683 mapped to the mm10 reference genome using the STAR (v2.5) program. The HTSeq v0.6.1
684 software was utilized to tally the number of reads that were aligned to each individual gene.

685 **Cellular Indexing of Transcriptomes and Epitopes by Sequencing (CITE-seq):**

686 We performed CITE-seq, a single-cell multi-omics technology that measures RNA and protein
687 expression simultaneously in single cells. The dataset was derived from an in vivo competition
688 experiment after 1 month of treatment, with 2 animals per treatment groups. Murin HSPC were
689 enriched and then CD45.1 + cells (*Dnmt3a*^{+/+}) and CD45.2 + cells (*Dnmt3a*^{R879/+}) were marked
690 using TotalSeq™-B antibodies (listed in Supplementary Table 4).

691 CITE-seq library generation was performed using 10x Genomics Chromium Single Cell 3'
692 v3.1_CellSurfaceProtein_RevD kit and the Novaseq 6000 sequencing system. 20000 cells for each
693 of the 4 samples were sequenced. The scRNA reads were aligned against the mouse reference
694 sequence mm10 using Cell Ranger (v6.1.2). The filtered Cell Ranger output was then used in the
695 Seurat package v4 from Satija lab for further processing. Cells with more than 500 and less than
696 8000 number of genes per cells and less than 15% of mitochondrial genes were kept for the
697 analysis. Fast integration using reciprocal PCA (RPCA) was used to find anchors across datasets
698 to integrate the 4 samples. Normalization, variance stabilization and selection of 3000 top variable
699 features of the molecular count data were performed using SCTransform followed by dimension
700 reduction by PCA and UMAP embedding using the top 30 principal components. Hematopoietic
701 subtypes were assigned to each cell using the AddModuleScore function and gene sets specific to
702 murine hematopoietic populations as previously defined¹⁹. The top gene set enrichment score was
703 establishing the selected annotation for each single cell. As additional filtering steps, the
704 maturing erythroblastic cells expressing low level of Kit and Ptpnc (CD45) were removed from
705 the analysis and scDblFinder 1.16.0 was run on the Cell Ranger raw output of each individual
706 sample to identify potential cell doublets. A total of 46225 cells were kept for downstream
707 analyses.

708 The CD45.1 and CD45.2 sequencing antibody derived tags (ADTs) were log normalized. ADTs
709 with a normalized value greater than 6 were identified as outlier points and removed from the
710 analysis. Because the experiment is a competition assay, each sample contains a mix of cells
711 expressing CD45.1 or CD45.2 at their surface, the ratio between CD45.1 and CD45.2 normalized
712 ADTs was therefore used to label each cell as wild type (CD45.1 > CD45.2) or mutant (CD45.1 <
713 CD45.2). To perform gene differential expression estimation, RNA counts from each defined
714 hematopoietic cell population labelled as mutant and wildtype were transformed in a Single Cell
715 Experiment object and aggregated for each of the metformin or vehicle samples using the sum of
716 the counts in the R package scuttle (v1.0.4).

717 **Differential expression and Gene Set Enrichment Analysis (GSEA)**

718 For both bulk RNA-seq and scRNA data, R package edgeR 3.36.0 was used to fit a generalized
719 linear model to estimate differential expression between groups. All genes were ranked from the
720 top up-regulated ones to down-regulated using the $\text{sign}(\log\text{FC}) * -\log_{10}(\text{pvalue})$ formula. GSEA
721 was performed using software from <https://www.gsea-msigdb.org/gsea/index.jsp> and
722 ClusterProfiler 4.2.2 fgsea embedded functions using the rank gene files and 2 defined Gene
723 Ontology Biological Process (GO BP) gene sets: oxidative phosphorylation (oxphos)
724 (<https://www.informatics.jax.org/go/term/GO:0006119>) and 1-carbon metabolism
725 (https://www.informatics.jax.org/vocab/gene_ontology/GO:0006730).

726 **Metabolomics analysis**

727 Murine HSPCs were collected, washed with PBS, and pelleted by centrifugation. The cell pellets
728 were then snap frozen and metabolites analyzed at the University of Colorado School of Medicine
729 Metabolomics Core. Metabolites were extracted from frozen cells pellets at a concentration of
730 2×10^6 cells/mL using a cold 5:3:2 methanol:acetonitrile:water solvent and quantified on a Thermo
731 Vanquish UHPLC coupled to a Thermo Q Exactive mass spectrometer in a positive and negative
732 ion modes (separate runs) exactly as described previously^{29,30}. Signals were annotated and
733 integrated using Maven in conjunction with the KEGG database and an in-house standard library
734 as reported³¹.

735 **Measurement of S-adenosylmethionine and S-adenosylhomocysteine concentrations**

736 The levels of *S*-adenosylmethionine (SAM) and *S*-adenosylhomocysteine (SAH) were quantified
737 using the SAM and SAH Combo ELISA Kit developed by Cell Biolabs (Cat# STA-671-C). This
738 enzyme immunoassay kit is specifically designed to accurately detect and measure SAH and SAM
739 in cell lysate samples. In brief, for sample preparation, 3×10^6 snap-frozen mouse HSPCs were
740 thawed and sonicated in 1 mL of cold PBS 1X on ice. Following that, the homogenized samples
741 underwent a 15-minute centrifugation at 10,000 g at 4°C. The supernatant obtained was carefully
742 collected, kept on ice, and aliquoted for use in the assay. Any excess supernatant that was not
743 utilized immediately was stored at a temperature of -80°C. The quantification of *S*-
744 adenosylmethionine (SAM) and *S*-adenosylhomocysteine (SAH) concentrations was performed in
745 accordance with the experimental protocols outlined in the manual provided by Cell Biolabs.

746 **Lentiviral production and transduction**

747 Oligonucleotides used to generate shRNA vectors were chemically synthesized by Integrated DNA
748 Technologies, annealed, and ligated into the BbsI site in the DECIPHER shRNA expression
749 lentiviral vector (Cellecta; pRSI9-U6-(sh)-UbiC-TagRFP-2A-Puro) (Addgene plasmid# 28289).
750 The TagRFP sequence has been replaced by the BFP sequence PCR -amplified and cloned into
751 XbaI and BamHI restriction sites. The NDI1 coding sequence was PCR-amplified from PMXS-
752 NDI1 (Addgene plasmid # 72876) and cloned into pLVX-EF1a-IRES-ZsGreen1 (Clontech, Cat #
753 631982) using EcoRI and SpeI restriction sites. The oligonucleotide and primers sequences are
754 shown in (Supplementary Table 3).

755 For lentiviral production HEK293T cells were grown in RPMI 1640 medium (Wisent #350-035-
756 CL) supplemented with 10% fetal bovine serum (FBS, Wisent #080-150) and 1% GlutaPlus
757 (Wisent #609-066-EL). Cells were seeded into 150mm tissue culture plates at the density of 7×10^6
758 cells/plate the day before transfection. On the day of transfection, cells were co-transfected with
759 lentiviral vectors, psPAX2 (Addgene#12259), and pCMV-VSVG (Cell Biolabs Part No.320023)
760 plasmids using jetPRIME transfection reagents (Polyplus #CA89129-924) according to the
761 manufacturer protocol. Viral particles were collected 48 and 72h post transfection and resuspended
762 in HBSS (Gibco #14170112) +25mM HEPES (Thermo Fisher #15630-080). The combined
763 supernatant was centrifuged at 450 x g and filtered through a 0.2 mm PES filter (Thermo Fisher
764 Scientific, Cat # 564-0020). The filtered supernatant (40 mL) was mixed with 10 mL of PBS
765 containing 20% (w/v) PEG 8000 (Sigma, Cat # 89510- 1KG-F), incubated overnight at 4°C, and

766 centrifuged at 3,700 rpm for 30 mins at 4°C. The pellet containing lentiviral particles was
767 resuspended in 2 mL of HBSS (Thermo Fisher Scientific, Cat # 14170-112) with 25 mM HEPES
768 (Thermo Fisher Scientific, Cat # 15630-080), aliquoted, and store at -80°C.

769 For lentiviral transductions, non-TC-treated 24 well plates were coated with 20 mg/mL of
770 Retronectin (Takara, Cat # T100B) for 2 hours at room temperature followed by aspiration and
771 blocking with PBS containing 2% (w/v) BSA (Wisent Bioproducts, Cat # 800-096-EG) for 30 min
772 at room temperature. After aspiration of the blocking buffer, the concentrated virus suspension
773 was added to wells. The plates were then centrifuged at 3,700 rpm for two hours at 4°C to allow
774 virus binding. Following centrifugation, unbound virus was aspirated, and 0.5 to 1x10⁶ AML cells
775 were added. The plates were then transferred to a 37°C incubator to initiate lentiviral infection.

776 **Reduced-Representation Bisulfite Sequencing (RRBS) and analysis**

777 Genomic DNA extracted from LK-enriched BM cell samples were submitted to Novogene
778 (Sacramento, CA) for RRBS. Briefly, genomic DNA was digested with MspI, and the resulting
779 fragments were end-repaired, A-tailed, and ligated with methylated adapters. After size-selection,
780 bisulfite conversion was performed using the EZ DNA Methylation-Gold™ Kit (Zymo Research).
781 The bisulfite-converted DNA was PCR-amplified to enrich for adapter-ligated fragments. RRBS
782 libraries were quality-checked and sequenced on an Illumina HiSeq/NovaSeq platform, generating
783 paired-end reads of 150bp nucleotides.

784 Raw data were trimmed to remove adapter and low-quality bases using Trimmomatic-0.36,
785 followed by quality control assessment with FastQC (v0.11.5). Trimmed reads were aligned to
786 mouse reference genome from Ensembl (GRCm38/mm10) and duplicated reads were removed.
787 DNA methylation calls were extracted from the aligned reads as CpG coverage files. Differentially
788 methylated regions (DMRs) were identified using the open-source R package methylKit
789 (v1.26.0)^{32,33}. CpG sites on unmapped genome assembly contigs were removed, and remaining
790 CpG sites were filtered to exclude CpGs with <10× coverage PCA analysis in R. We used
791 methylKit to perform pairwise comparisons to identify DMRs between untreated *Dnmt3a*^{R878H/+}
792 vs. untreated *Dnmt3a*^{+/+} samples, and between metformin-treated *Dnmt3a*^{R878H/+} and untreated
793 *Dnmt3a*^{R878H/+} samples. To this end, the genome was tiled into 500bp non-overlapping bins. To
794 calculate DMR p-values, a logistic regression test was used methylKit. P-values were adjusted for
795 multiple testing (i.e., q-value) via the SLIM method³⁴. DMRs with p-value < 0.01 were used for

796 downstream analysis. CpG islands were annotated by using the University of California Santa Cruz
797 (UCSC) (<https://genome.ucsc.edu/index.html>) database with using plyranges R package
798 (v1.20.0)³⁵. Promoters were defined as 1kb upstream and 150bp downstream around the
799 transcription start site (TSS) and annotated by ChIPseeker R package (v1.36.0)^{36,37}.

800 **H3K27me3 Chromatin-immunoprecipitation sequencing (ChIP-seq)**

801 LK-enriched BM cells were fixed with 1% formaldehyde for 15 min according to the Active Motif
802 ChIP cell fixation protocol. Fixed cell pellets were submitted to Active Motif (Carlsbad, CA) for
803 ChIP-Seq. Briefly, 15ug chromatin and 4ul of antibody against H3K27me3 (Active Motif cat#
804 39155) were used to immunoprecipitated genomic DNA regions of interest. Illumina base-cell data
805 were processed and demultiplexed using bcl2fastq2 v2.20 and low-quality bases with Phred scores
806 less than 33 were trimmed. 75 bp single-end sequence reads were subsequently mapped to the
807 genome through BWA v0.7.12 algorithm with default settings. Low quality reads were filtered out
808 and PCR duplicates were removed. Aligned sequencing reads, or tags, were extended to 200bp
809 from the 3' end, followed by dividing the genome into 32bp bins and counting the number of
810 fragments in each bin. The resulting histograms (genomic “signal maps”) were stored in bigWig
811 files. Peak locations were determined using the MACS algorithm (v2.1.0) with a cutoff of p-value
812 =1e-7. Peaks that were on the ENCODE blacklist of known false ChIP-Seq peaks were removed.
813 18556 peaks that were identified in at least one sample with a cutoff p-value of 1e-7 were all
814 merged in a common matrix and the total number of present peaks as well as averaged peak values
815 were calculated and plotted for each sample and condition using the R package ggplot2. A t-test
816 was used to assess difference in mean between each condition. ChIP-Seq profiles +-2kb of the
817 transcription start site (TSS) of 10,622 transcripts representing unique genes was created using the
818 plotHeatmap function of DeepTools 3.5.1.

819 **Quantification and Statistical analysis**

820 Statistical analyses were carried out according to the specifications detailed in the figure legends
821 using GraphPad Prism v10 (GraphPad Software, La Jolla, CA). The figure legends contain details
822 regarding the quantity of experimental repetitions or animals in each group. Statistical significance
823 was defined as P values < 0.05. In the context of GSEA analysis, statistical significance was
824 established using P values below 0.05 and FDR values below 0.05. This rigorous approach was
825 employed to ascertain statistical significance in the GSEA results.

826 **Data availability**

827 Raw and processed data from the bulk RNA-seq, scRNA-seq, RRBS, and CHIP-seq experiments
828 will be submitted to a publicly available repository (GEO).

829 **Code availability**

830 All the codes used for bioinformatics analysis will be submitted to a publicly available repository
831 (GitHub).

832

833 **References**

- 834 1 Jaiswal, S. *et al.* Clonal Hematopoiesis and Risk of Atherosclerotic Cardiovascular Disease. *N Engl J Med* **377**, 111-121, doi:10.1056/NEJMoa1701719 (2017).
835
- 836 2 Jaiswal, S. *et al.* Age-related clonal hematopoiesis associated with adverse outcomes. *N Engl J Med* **371**, 2488-2498, doi:10.1056/NEJMoa1408617 (2014).
837
- 838 3 Genovese, G. *et al.* Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N Engl J Med* **371**, 2477-2487, doi:10.1056/NEJMoa1409405 (2014).
839
- 840 4 Bick, A. G. *et al.* Inherited causes of clonal haematopoiesis in 97,691 whole genomes. *Nature* **586**, 763-768, doi:10.1038/s41586-020-2819-2 (2020).
841
- 842 5 Buscarlet, M. *et al.* DNMT3A and TET2 dominate clonal hematopoiesis and demonstrate benign phenotypes and different genetic predispositions. *Blood* **130**, 753-762, doi:10.1182/blood-2017-04-777029 (2017).
843
844
- 845 6 Venugopal, K., Feng, Y., Shabashvili, D. & Guryanova, O. A. Alterations to DNMT3A in Hematologic Malignancies. *Cancer Res* **81**, 254-263, doi:10.1158/0008-5472.CAN-20-3033 (2021).
846
- 847 7 Young, A. L., Tong, R. S., Birmann, B. M. & Druley, T. E. Clonal hematopoiesis and risk of acute myeloid leukemia. *Haematologica* **104**, 2410-2417, doi:10.3324/haematol.2018.215269 (2019).
848
- 849 8 Jawad, M. *et al.* DNMT3A R882 Mutations Confer Unique Clinicopathologic Features in MDS Including a High Risk of AML Transformation. *Front Oncol* **12**, 849376, doi:10.3389/fonc.2022.849376 (2022).
850
851
- 852 9 Larsson, C. A., Cote, G. & Quintas-Cardama, A. The changing mutational landscape of acute myeloid leukemia and myelodysplastic syndrome. *Mol Cancer Res* **11**, 815-827, doi:10.1158/1541-7786.MCR-12-0695 (2013).
853
854
- 855 10 Russler-Germain, D. A. *et al.* 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, doi:10.1016/j.ccr.2014.02.010 (2014).
856
857
- 858 11 Kim, S. J. *et al.* A DNMT3A mutation common in AML exhibits dominant-negative effects in murine ES cells. *Blood* **122**, 4086-4089, doi:10.1182/blood-2013-02-483487 (2013).
859
- 860 12 Smith, A. M. *et al.* Functional and epigenetic phenotypes of humans and mice with DNMT3A Overgrowth Syndrome. *Nat Commun* **12**, 4549, doi:10.1038/s41467-021-24800-7 (2021).
861
- 862 13 Loberg, M. A. *et al.* Sequentially inducible mouse models reveal that Npm1 mutation causes malignant transformation of Dnmt3a-mutant clonal hematopoiesis. *Leukemia* **33**, 1635-1649, doi:10.1038/s41375-018-0368-6 (2019).
863
864
- 865 14 Bridges, H. R. *et al.* Structural basis of mammalian respiratory complex I inhibition by medicinal biguanides. *Science* **379**, 351-357, doi:10.1126/science.ade3332 (2023).
866
- 867 15 LaMoia, T. E. & Shulman, G. I. Cellular and Molecular Mechanisms of Metformin Action. *Endocr Rev* **42**, 77-96, doi:10.1210/endrev/bnaa023 (2021).
868
- 869 16 Seo, B. B. *et al.* Molecular remedy of complex I defects: rotenone-insensitive internal NADH-quinone oxidoreductase of *Saccharomyces cerevisiae* mitochondria restores the NADH oxidase activity of complex I-deficient mammalian cells. *Proc Natl Acad Sci U S A* **95**, 9167-9171, doi:10.1073/pnas.95.16.9167 (1998).
870
871
872
- 873 17 Wheaton, W. W. *et al.* Metformin inhibits mitochondrial complex I of cancer cells to reduce tumorigenesis. *Elife* **3**, e02242, doi:10.7554/eLife.02242 (2014).
874
- 875 18 Dowling, R. J. *et al.* Metformin Pharmacokinetics in Mouse Tumors: Implications for Human Therapy. *Cell Metab* **23**, 567-568, doi:10.1016/j.cmet.2016.03.006 (2016).
876
- 877 19 Izzo, F. *et al.* DNA methylation disruption reshapes the hematopoietic differentiation landscape. *Nat Genet* **52**, 378-387, doi:10.1038/s41588-020-0595-4 (2020).
878

879 20 Nam, A. S. *et al.* Single-cell multi-omics of human clonal hematopoiesis reveals that DNMT3A R882
880 mutations perturb early progenitor states through selective hypomethylation. *Nat Genet* **54**,
881 1514-1526, doi:10.1038/s41588-022-01179-9 (2022).

882 21 Dai, Y. J. *et al.* Conditional knockin of Dnmt3a R878H initiates acute myeloid leukemia with mTOR
883 pathway involvement. *Proc Natl Acad Sci U S A* **114**, 5237-5242, doi:10.1073/pnas.1703476114
884 (2017).

885 22 Fiumara, M. *et al.* Genotoxic effects of base and prime editing in human hematopoietic stem cells.
886 *Nat Biotechnol*, doi:10.1038/s41587-023-01915-4 (2023).

887 23 Chen, P. J. & Liu, D. R. Prime editing for precise and highly versatile genome manipulation. *Nat*
888 *Rev Genet* **24**, 161-177, doi:10.1038/s41576-022-00541-1 (2023).

889 24 SanMiguel, J. M. *et al.* Distinct Tumor Necrosis Factor Alpha Receptors Dictate Stem Cell Fitness
890 versus Lineage Output in Dnmt3a-Mutant Clonal Hematopoiesis. *Cancer Discov* **12**, 2763-2773,
891 doi:10.1158/2159-8290.CD-22-0086 (2022).

892 25 Cuyas, E. *et al.* Metformin regulates global DNA methylation via mitochondrial one-carbon
893 metabolism. *Oncogene* **37**, 963-970, doi:10.1038/onc.2017.367 (2018).

894 26 Cuyas, E. *et al.* Metformin directly targets the H3K27me3 demethylase KDM6A/UTX. *Aging Cell*
895 **17**, e12772, doi:10.1111/accel.12772 (2018).

896 27 Garcia-Calzon, S. *et al.* DNA methylation partially mediates antidiabetic effects of metformin on
897 HbA1c levels in individuals with type 2 diabetes. *Diabetes Res Clin Pract* **202**, 110807,
898 doi:10.1016/j.diabres.2023.110807 (2023).

899 28 Nelson, J. W. *et al.* Engineered pegRNAs improve prime editing efficiency. *Nat Biotechnol* **40**, 402-
900 410, doi:10.1038/s41587-021-01039-7 (2022).

901 29 Gehrke, S. *et al.* Red Blood Cell Metabolic Responses to Torpor and Arousal in the Hibernator
902 Arctic Ground Squirrel. *J Proteome Res* **18**, 1827-1841, doi:10.1021/acs.jproteome.9b00018
903 (2019).

904 30 Nemkov, T., Reisz, J. A., Gehrke, S., Hansen, K. C. & D'Alessandro, A. High-Throughput
905 Metabolomics: Isocratic and Gradient Mass Spectrometry-Based Methods. *Methods Mol Biol*
906 **1978**, 13-26, doi:10.1007/978-1-4939-9236-2_2 (2019).

907 31 Nemkov, T., Hansen, K. C. & D'Alessandro, A. A three-minute method for high-throughput
908 quantitative metabolomics and quantitative tracing experiments of central carbon and nitrogen
909 pathways. *Rapid Commun Mass Spectrom* **31**, 663-673, doi:10.1002/rcm.7834 (2017).

910 32 Akalin, A. *et al.* methylKit: a comprehensive R package for the analysis of genome-wide DNA
911 methylation profiles. *Genome Biol* **13**, R87, doi:10.1186/gb-2012-13-10-r87 (2012).

912 33 Akalin, A. *et al.* Base-pair resolution DNA methylation sequencing reveals profoundly divergent
913 epigenetic landscapes in acute myeloid leukemia. *PLoS Genet* **8**, e1002781,
914 doi:10.1371/journal.pgen.1002781 (2012).

915 34 Wang, H. Q., Tuominen, L. K. & Tsai, C. J. SLIM: a sliding linear model for estimating the proportion
916 of true null hypotheses in datasets with dependence structures. *Bioinformatics* **27**, 225-231,
917 doi:10.1093/bioinformatics/btq650 (2011).

918 35 Lee, S., Cook, D. & Lawrence, M. plyranges: a grammar of genomic data transformation. *Genome*
919 *Biol* **20**, 4, doi:10.1186/s13059-018-1597-8 (2019).

920 36 Wang, Q. *et al.* Exploring Epigenomic Datasets by ChIPseeker. *Curr Protoc* **2**, e585,
921 doi:10.1002/cpz1.585 (2022).

922 37 Yu, G., Wang, L. G. & He, Q. Y. ChIPseeker: an R/Bioconductor package for ChIP peak annotation,
923 comparison and visualization. *Bioinformatics* **31**, 2382-2383, doi:10.1093/bioinformatics/btv145
924 (2015).

925

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Dnmt3aPaperMHSupplementarytables.xlsx](#)