SUPPLEMENTAL INFORMATION

Physioxia induced Downregulation of Tet2 in Hematopoietic Stem Cells contributes to Enhanced Self-renewal – Supplementary

SUPPLEMENTARY METHODS

Flow cytometry immunophenotyping. BM cells were incubated with fluorochrome conjugated anti-mouse antibodies (0.2 μg/10^6 cell) in ice cold PBS + 2% BSA for 30 minutes. The 3% O_2 group was stained inside the hypoxia chamber, and the ambient air group was stained in ambient air. All groups were fixed with fresh 2% paraformaldehyde in ice cold PBS for 30 minutes. Intracellular staining was performed using fixation/Permeabilization Solution Kit (BD Biosciences, San Diego, CA) according to manufacturer’s protocol. Data acquisition was performed using a LSRiI flow cytometer (BD Biosciences, San Diego, CA). Data analysis was performed using FlowJo 10.6.2 software. Gating strategy was based on fluorescence-minus-one controls. Absolute numbers of phenotypically-defined HSC and HPC per femur were calculated by multiplying the percentage of each cell population and that of live cells. HSCs were defined as Lin^− c-Kit^+ Sca1^+ CD150^+ CD48^− cells, multipotent progenitors (MPP) cells as Lin^− c-Kit^+ Sca1^+ CD150^+ CD48^− cells, common myeloid progenitors (CMP) as Lin^− c-Kit^+ Sca1^− CD34^+ FcγII/IIIR^{lo} cells, granulocyte-macrophage progenitors (GMP) as Lin^− c-Kit^+ Sca1^− CD34^{hi} FcγII/IIIR^{hi} cells, megakaryocyte-erythrocyte progenitors (MEP) as Lin^− c-Kit^+ Sca1^− CD34^{−/−} FcγII/IIIR^{−/−} cells and common lymphoid progenitors (CLP) as Lin^− c-Kit^{lo} Sca1^{lo} Flt3^{+} IL7R^+ cells.

Colony Forming Unit (CFU) HPC assays. Cells were plated at 5x10^4 cells/mL in 1% methylcellulose culture medium with 30% fetal bovine serum, 1 U/mL recombinant human erythropoietin, 50 ng/mL recombinant mouse stem cell factor, 0.1 mM 2-mercaptoethanol, 2 mM...
L-glutamine, 0.1 mM hemin, and 5% vol/vol pokeweed mitogen mouse spleen cell conditioned medium. Cells were incubated at 5% CO₂ and 5% O₂ in a humidified chamber, and CFU-GM, BFU-E, and CFU-GEMM HPCs were scored at day 6 of incubation and are shown as cells/femur. For high specific activity tritiated thymidine kill assay to estimate the % CFU in S-phase of the cell cycle, BM cells were incubated in IMDM media with control diluent or 50 mCi high specific activity tritiated thymidine for 30 minutes in 5% CO₂ and 5% O₂ in a humidified chamber before plating. Cells were plated in the aforementioned culture and incubation conditions. [1]

**Transplantation to Assess Engraftment of BM Cells and Numbers of Functional HSCs.** Recipient F1 (CD45.1⁺ CD45.2⁺) mice were fed uniprim for one week and irradiated with split dose of 700 followed 400 cGy, 24 hours before transplantation. Lethally irradiated F1 mice were transplanted with 12,500, 25,000, 50,000 limiting dilutions of (C57BL/6 CD45.2⁺) donor cells and 100,000 Boy/J (CD45.1⁺) BM competitor cells. Injections were done intravenously in the hypoxia chamber for the 3% O₂ group and in ambient air for the ambient air control group. Competitor Boy/J cells were collected in air and injected i.v. after the donor cells for both physioxia and ambient air collections of cells. The method of injection inside the hypoxia chamber has been described.[4] Chimerism of the recipient F1 mice in the PB was assessed up to 16 weeks by submandibular vein bleeding. Mice were sacrificed four months after transplantation, and then BM cells were analyzed by flow cytometry for HSCs, and lymphoid and myeloid lineages. Competitive repopulating units (CRUs) frequency was calculated and plotted using ELDA software (bioinf.wehi.edu.au/software/elda/) to calculate functional HSCs. To assess the self-renewal capacity of donor HSCs, we intravenously injected 1.5 x 10⁶ of primary recipient’s BM cells without competitor cells into lethally irradiated secondary F1 (CD45.1⁺ CD45.2⁺) mice recipients. [1]
Analysis of 5-hmC levels in WT and Tet2−/− BM cells. WT and Tet2−/− BM cells were collected in physioxia and ambient air conditions. BM lineage depleted cells and LSK cells were incubated in physioxia and ambient air for 6 hours before DNA extraction. Genomic DNA from Lin-negative and LSK cells derived from WT and Tet2−/− mice was purified by Takarabio NucleoSpin Tissue XS Genomic DNA Purification Kit. 5-hmC level analysis was conducted by Zymo Quest 5-hmC DNA ELISA Kit according to manufacturer’s protocol.

CITE-seq and cell hashing. (Lin− Ckit+ Sca1+) cells were enriched inside the hypoxia chamber using mouse direct lineage cell depletion kit, EasySep™ release mouse biotin positive selection kit and EasySep™ mouse SCA1 positive selection kit consecutively. LSK enriched cells were divided into two groups; group 1 was incubated inside the hypoxia chamber, and group 2 was equilibrated in ambient air for 6 hours. Dead and apoptotic cells were removed using dead cell removal kit. Cells were blocked with TruStain FcX™ PLUS (anti-mouse CD16/32). Cell multiplexing and labeling was performed according to the manufacturer’s (Biolegend) instructions. Surface protein fixation was done by treating the cells with freshly prepared cell membrane-impermeable crosslinker (DTSSP, 3,3’- dithiobis(sulfosuccinimidyl propionate). Each biological sample was labeled with a specific Total seq hashtag antibody; WT1, WT2, WT3, Tet2−/− 1, Tet2−/− 2, and Tet2−/− 3 were labeled with Total-seq hashtang 1-6 respectively. All samples were pooled in equal numbers into two groups (ambient air (~21% O2) and physioxia (3% O2). Labeled cells were loaded to a 10X Genomics Chromium Next GEM chip G for reverse transcription and complementary DNA inside gel bead-in-emulsion. Full-length cDNA was amplified, and sequencing libraries were normalized for sequencing. The constructed library was sequenced on Illumina HiSeq4000.

CITE-seq analysis. Cite-seq data were processed by Cell Ranger 6.1(https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/what-
The scRNA-seq data was mapped to mm10 genome and the expression status of protein markers were separately computed by using Cite-Seq Count (https://github.com/Hoohm/CITE-seq-Count). In total, 30603 cells having matched RNA-seq and antibody profiles were identified. A K-mean clustering was applied to assess the labels of the cells (I) belonging to the four-sample source and (II) different expression status of the protein markers. We identified 6413, 6257, 10175, and 7758 cells belonging to Tet2<sup>-</sup> physioxia, Tet2<sup>-</sup> ambient air, WT physioxia, and WT ambient air conditions, respectively. Cell types were annotated based on the antibody profiling of surface proteins. In this study, we specifically focused on LSK cells with high expression of CD150. In total, we identified 596, 419, 374, and 315 CD150 high cells in the Tet2<sup>-</sup> physioxia, Tet2<sup>-</sup> air, WT physioxia and WT- air conditions, respectively. Cell clustering analysis of the scRNA-seq data was computed and visualized using Seurat R package <sup>(2)</sup> with default parameters. <sup>(3)</sup> Differential gene expression analysis was conducted by using left truncated mixture Gaussian model <sup>(4)</sup> with p value <0.005 as the significant cutoff. Pathway enrichment analysis was conducted using hypergeometric test and single-sample gene set enrichment analysis (ssGSEA) against MsigDB v6.0 canonical pathways <sup>(5)</sup> Gene Ontology <sup>(6)</sup>, and gene signatures of hematopoietic cell types collected in our past studies. <sup>(7)</sup>

The metabolic flux analysis was performed using a graph neural network-based method for sample-wise metabolic rate prediction utilizing scRNA-seq, called scFEA (single-cell flux estimation analysis).<sup>(8)</sup> scFEA models metabolic flux at the cell level using changes in the expression levels of enzymes in the cell based on two assumptions: (1) the overall influx of each metabolite is approximately equal to its total outflux; and (2) the changes in metabolites outflux can be modeled as a (non-linear) function of changes in the expression levels of genes involved in metabolites production. We began by reconstructing mouse major bioenergetic pathways, which includes glycolysis, TCA cycle, glutaminolysis pathways, glucose, glutamine, and glutamate transport, as well as other metabolic reactions (Fig. S. 4A). scFEA was applied to all cell samples in our scRNA-seq data in order to predict the metabolic flux on a cell-by-cell basis.
As previously stated, the relative abundance of intermediate substrates in each sample may be estimated using the difference between the projected in- and out-flux values for each substrate.(8)

**Bulk RNA-seq library preparation and data analysis.** LSK cells were enriched and processed as described above. RNA extraction was performed inside the hypoxia chamber for the physioxia group and in an ambient air. Total RNA was extracted using Qiagen RNeasy Plus Micro Kit. RNA with Agilent Bioanalyzer RIN of > 7 were selected for library preparation. Library preparation was performed using New England Biolabs Single Cell/Low Input RNA Library Prep Kit for Illumina. Paired-end RNA-sequencing (RNA-seq) was performed on Illumina HiSeq4000 by IU School of Medicine Medical Genomics core. Each group/condition consists of four matched biological replicates and ~30 million reads were obtained per sample. The pipeline employed for data analysis included FastQC for quality control, Cutadapt for adapter trimming and excluding low-quality reads, STAR for genome indices creation and alignment of reads to the GRCm38 mouse genome, HTSeq to assign reads to genes, and DESeq2 to detect differentially expressed genes.

**Statistical analysis.** Statistical analysis was performed in GraphPad Prism 9. Two-way ANOVA with a post hoc Tukey’s multiple comparison test was used for statistical analysis between four groups. Two-tailed Students t-test was used for analysis between two groups. Poisson statistical analysis was performed for the limiting dilution assays. Wilcoxon rank sum test was used for CITE-seq gene expression and scFEA estimation analysis. P values significance levels: *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001. Error bars represent SEM unless stated otherwise.
### REAGENTS AND MATERIALS

#### List of fluorophore conjugated antibodies used in the surface protein phenotyping studies

| Antibody (Anti-mouse) | Fluorophore      | Clone                  | Vendor   |
|------------------------|------------------|------------------------|----------|
| Mouse lineage cocktail | FITC             | 145-2C11; RB6-8C5; RA3-6B2; Ter-119; M1/70 | Biolegend |
| Ly6A/E (Sca-1)         | PE/Dazzle 594    | D7                     |          |
| CD117 (c-Kit)          | Biotin           | 2B8                    |          |
| CD34                   | PE               | RAM34                  |          |
| CD135                  | BV421            | A2F10.1                |          |
| CD16/32                | TruStain FcX™ PLUS | S17011E                |          |
| CD150                  | APC              | TC15-12f12.2           |          |
| CD45.1                 | Alexa Flour 700  | A20                    |          |
| CD45.1                 | FITC             | A20                    |          |
| CD45.1                 | PE               | A20                    |          |
| CD45.2                 | APC              | 104                    |          |
| CD3                    | Alexa Flour 594  | 17A2                   |          |
| Donkey anti-rabbit IgG| FITC             | Poly4064               |          |
| CD3                    | FITC             | 17A2                   |          |
| Ly-6G/Ly-6C (Gr-1)     | BV421            | RB6-8C5                |          |
| CD45R/B220             | PerCP-Cyanine 5.5 | RA3-6B2                |          |
| Antibody (Anti-mouse) | Barcode                     | Cat No | Vendor       |
|------------------------|----------------------------|--------|--------------|
| CD11b PE- Cyanine7     | M1/70                      |        | BD Bioscience|
| CD150 PerCP/Cyanine5.5 | TC15-12f12.2               |        |              |
| CD48 PE/Cy7            | HM48-1                     |        |              |
| CD127 BV711            | SB199                      |        |              |
| CD16/32 PerCP-Cyanine 5.5 | 2.4G2                   |        |              |
| CD117 (c-Kit) APC-H7   | 2B8                       |        |              |
| Tet2 Unconjugated      | D6C7K                      |        | Cell Signaling|

### List of barcoded antibodies used in the cell hashing and CITE-seq study

| CITE-seq/Cell Hashing | Barcode                     | Cat No | Vendor       |
|-----------------------|----------------------------|--------|--------------|
| TotalSeq™-A0130 (Sca-1) | TTCCTTTTCTACGCA            | 108151 | Biolegend    |
| TotalSeq™-A0012 cKit (CD117) | TGCATGTCATCGGTG        | 105843 |              |
| TotalSeq™-A0203 CD150  | CAACGCCTAGAAACC           | 115945 |              |
| TotalSeq™-A0429 CD48  | AGAACCGCCGTAGTT           | 103447 |              |
| TotalSeq™-A0098 CD135 | GTAGCAAGATTCAAG           | 135316 |              |
| TotalSeq™-A0091 Mouse IgG2a, κ | CTCCTACCTAAACTG        | 400285 |              |
| TotalSeq™-A0241 Armenian IgG | CCTGTCATTAAGACT     | 400973 |              |
| TotalSeq™-A0301 Hashtag 1 | ACCCACCAGTAAGAC          | 155801 |              |
| TotalSeq™-A0302 Hashtag 2 | GGTCGAGAGCATTCA          | 155803 |              |
| TotalSeq™-A0303 Hashtag 3 | CTTGCCGCATGTCAT          | 155805 |              |
| TotalSeq™-A0304 Hashtag 4 | AAAGCATTCTTCACG          | 155807 |              |
| TotalSeq™-A0305 Hashtag 5 | CTTGTCTTTTGAG            | 155809 |              |
### List of reagents

| Reagent                                                        | Cat No        | Vendor            |
|----------------------------------------------------------------|--------------|------------------|
| αKG (Octyl-α-KG)                                               | SML2205      | Sigma            |
| Dead cell removal kit                                         | 130-090-101  | Miltenyi B       |
| Direct mouse lineage cell depletion kit                        | 130-110-470  | Miltenyi B       |
| DTSSP (3,3'-dithiobis(sulfosuccinimidyl propionate))          | 21578        | Thermo F.        |
| EasySep™ Mouse SCA1 Positive Selection Kit                    | 18756        | Stemcell.T       |
| EasySep™ Release Mouse Biotin Positive Selection K             | 17655        | Stemcell. T      |
| Fetal Bovine Serum                                            | 35010CV      | Corning          |
| Fixation/Permeabilization Solution Kit                        | 555028        | BD Bioscience    |
| Hemin                                                          | H9039        | Sigma            |
| Is cove’s Modified Dulbecco’s Medium                           | 51471C       | Sigma            |
| L-glutamine                                                   | G7513        | Sigma            |
| Mercaptoethanol                                                | M6250        | Sigma            |
| New England Single Cell/Low Input RNA Library Prep            | E6420S       | NE Biolabs       |
| NucleoSpin Tissue XS Genomic DNA Purification Kit             | 740901       | Takara Bio       |
| Pierce™ 16% Formaldehyde (w/v), Methanol-free                 | 28906        | Thermo F.        |
| Plerixafor/AMD3100                                            | S3013        | Selleckchem      |
| Quest 5-hmC DNA ELISA Kit                                     | D5425        | Zymo R           |
| Recombinant Human Erythropoietin                              | 287-TC-500   | R&D S.           |
| Recombinant mouse stem cell factor                            | 455-MC-050   | R&D S.           |
SUPPLEMENTAL FIGURES

S.1 (A) Proportion of cell subsets classified by surface protein expression status of CD150/CD117/CD48 cell subtypes, (B) Gene expression status of Slamf1 in CD150 expressing WT LSK cells under physioxia and ambient air. (C) Gene set enrichment analysis of MAPK signaling pathway genes in WT LSK-CD150^hi^ cells. (D) Heatmap of averaged expression level of selected HSCs differentiation and (E) apoptosis genes in WT LSK-CD150^hi^ cells under physioxia and ambient air.

S.2 (A) Flow cytometry gating strategy for single cells suspension. (B, C) Second replicate of limiting dilution engrafting assay using WT and Tet2^-^ cells collected and processed in 3% O_2 vs. ambient air (~21% O_2). 5 recipient mice per group are shown. Donor (CD45.2^+) cells were injected i.v. at doses of 25,000, and 50,000, 100,000 into lethally irradiated F1 host mice (CD45.1^+^CD45.2^+) in air or physioxia. Competitor Boy/J BM cells (CD45.1^+) were injected i.v. at 100,000 cell doses in air. Analysis has been described in Figure 3. Data are presented as mean± SEM. *p < 0.05, **p < 0.01, ***p < 0.001, analyzed by analyzed by Two-way ANOVA with a post hoc Tukey's multiple comparison test.

S.3 (A) CITE-seq and cell hashing workflow for WT and Tet2^-^ BM LSK cells under physioxia and ambient air. (B) Proportion of cell subsets classified by surface protein expression status of WT and Tet2^-^ CD150/CD117/CD48 cell subtypes, (C) Gene expression status of Slamf1 in CD150 expressing WT and Tet2^-^ LSK cells under physioxia and ambient air. (D) Intracellular flow
cytometry assessment of Tet2 in WT and Tet2−/− LSK cells under physioxia and ambient air. (E) 5-hmC levels in WT and Tet2−/− BM lineage depleted cells under physioxia and ambient air. Data are presented as mean±SEM. *p < 0.05, **p < 0.01, ***p < 0.001, analyzed by Two-way ANOVA with a post hoc Tukey’s multiple comparison test.

S.4 (A) Graphical representation of the scFEA glycolysis, glutamine/glutamate and TCA cycle flux pathways. (B) Flow cytometry plot representation of LT-HSC numbers under ambient air, physioxia and physioxia with α-KG. (C) Number of Tet2−/− LT-HSC after 6 hours incubation with α-KG under physioxia and ambient air. (D) Number of Tet2−/− HPCs after 6 hours incubation with α-KG under physioxia and ambient air. Data are presented as mean±SEM. *p < 0.05, **p < 0.01, ***p < 0.001, analyzed by Two-way ANOVA with a post hoc Tukey’s multiple comparison test.

SUPPLEMENTAL TABLES

S.1 Table. Differential gene expression statistics for multiple comparisons between WT-LSK-CD150hi-physioxia, WT LSK-CD150hi-air, Tet2−/− LSK-CD150hi-physioxia, Tet2−/− LSK-CD150hi-air. (Hypoxia/H = Physioxia, Normoxia/N = ambient air).

S.2 Table. Pathway enrichment analysis for multiple comparisons between WT-LSK-CD150 hi -physioxia, WT-LSK-CD150 hi -air, Tet2−/− LSK-CD150 hi -physioxia, Tet2−/− LSK-CD150 hi -air. (Hypoxia/H = Physioxia, Normoxia/N = ambient air).
**S.3 Table.** Estimated 2OG level in LSK-CD150^{hi}-Physioxia and air cells. (Hypoxia/H = Physioxia, Normoxia/N = ambient air).

**S.4 Table.** Estimated 2HG level in LSK-CD150^{hi}-Physioxia and air cells. (Hypoxia/H = Physioxia, Normoxia/N = ambient air).

**SUPPLEMENTAL REFERENCES**

1. Cooper, SH. Capitano, ML, and Broxmeyer HE., et al, *Experimental mouse models of mouse and human hematopoietic stem cell transplantation*, in *Methods in Molecular Biology; Hematopoietic Stem Cells*, a.H.J. In L. M. Pelus, Editor. 2022, Springer Nature.

2. Butler, A., et al., *Integrating single-cell transcriptomic data across different conditions, technologies, and species*. Nat Biotechnol, 2018. 36(5): p. 411-420.

3. Hao, Y., et al., *Integrated analysis of multimodal single-cell data*. Cell, 2021. 184(13): p. 3573-3587 e29.

4. Wan, C., et al., *LTMG: a novel statistical modeling of transcriptional expression states in single-cell RNA-Seq data*. Nucleic Acids Res, 2019. 47(18): p. e111.

5. Liberzon, A., et al., *Molecular signatures database (MSigDB) 3.0*. Bioinformatics, 2011. 27(12): p. 1739-40.

6. Subramanian, A., et al., *Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles*. Proc Natl Acad Sci U S A, 2005. 102(43): p. 15545-50.

7. Lu, X., et al., *SSMD: a semi-supervised approach for a robust cell type identification and deconvolution of mouse transcriptomics data*. Brief Bioinform, 2021. 22(4).

8. Alghamdi, N., et al., *A graph neural network model to estimate cell-wise metabolic flux using single-cell RNA-seq data*. Genome Res, 2021. 31(10): p. 1867-1884.
A: Air
P: Physioxia

A. Pie charts showing WT phoxsioxia and WT air.

B. Bar chart showing expression levels for WT phoxsioxia and WT air.

C. Enrichment score for KEGG MAPK Signaling.

D. Heat map showing expression levels of genes in WT phoxsioxia and WT air.

E. Heat map showing expression levels of genes in WT air and WT phoxsioxia.
25,000 donor cell

WT

Tet2<sup>-/-</sup>

50,000 donor cell

WT

Tet2<sup>-/-</sup>

100,000 donor cell

WT

Tet2<sup>-/-</sup>

Donor (CD45.2) % in PB

Donor cell dose x 1000

25

50

100

Donor (CD45.2) % in BM

S2
**A**

3% O₂

10x Genomics library

RNA

Antibody derived tags (ADT)

21% O₂

10x Genomics library

RNA

Antibody derived tags (ADT)

**B**

|        | Tet2⁻/⁻ physioxia-1 | Tet2⁻/⁻ air-1 | WT physioxia-1 | WT air-1 |
|--------|----------------------|---------------|---------------|----------|
| CD150/CD117 | 52.12% | 23.37% | 18.67% | 10.38% |
| CD150/CD117/CD48 | 19.27% | 15.64% | 18.08% | 12.89% |
| CD150H | 5.98% | 16.43% | 41.23% | 10.38% |
| CD150H/CD48 | 0.89% | 0.89% | 0.89% | 0.89% |

**C**

**Slamf1**

Normalized Expression

**D**

Isotype + 2ry antibody controls

WT-air

WT-physioxia

Tet2⁻/⁻ air

Tet2⁻/⁻ physioxia

**E**

* ns

****

S3
