



Published in final edited form as:

Nat Cell Biol. 2017 June ; 19(6): 626–638. doi:10.1038/ncb3527.

## Regulation of Mitochondrial Biogenesis in Erythropoiesis by mTORC1-Mediated Protein Translation

Xin Liu<sup>1,6</sup>, Yuannu Zhang<sup>1,2,6</sup>, Min Ni<sup>1,6</sup>, Hui Cao<sup>1</sup>, Robert A.J. Signer<sup>1,3</sup>, Dan Li<sup>1</sup>, Mushan Li<sup>2</sup>, Zhimin Gu<sup>1</sup>, Zeping Hu<sup>1</sup>, Kathryn E. Dickerson<sup>1</sup>, Samuel E. Weinberg<sup>5</sup>, Navdeep S. Chandel<sup>5</sup>, Ralph J. DeBerardinis<sup>1</sup>, Feng Zhou<sup>3,7</sup>, Zhen Shao<sup>2,7</sup>, and Jian Xu<sup>1,7</sup>

<sup>1</sup>Children's Medical Center Research Institute, Department of Pediatrics, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA

<sup>2</sup>Key Laboratory of Computational Biology, Collaborative Innovation Center for Genetics and Developmental Biology, CAS-MPG Partner Institute for Computational Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China

<sup>3</sup>Division of Regenerative Medicine, Department of Medicine, Moores Cancer Center, University of California San Diego, La Jolla, CA 92093, USA

<sup>4</sup>Liver Cancer Institute, Zhongshan Hospital, Key Laboratory of Carcinogenesis and Cancer Invasion, Minister of Education, and Institutes of Biomedical Sciences, Fudan University, Shanghai 200032, China

<sup>5</sup>Department of Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL 60611, USA

### Abstract

Advances in genomic profiling present new challenges of explaining how changes in DNA and RNA are translated into proteins linking genotype to phenotype. Here we compare the genome-scale proteomic and transcriptomic changes in human primary hematopoietic stem/progenitor cells and erythroid progenitors, and uncover pathways related to mitochondrial biogenesis enhanced through post-transcriptional regulation. Mitochondrial factors including TFAM and PHB2 are selectively regulated through protein translation during erythroid specification. Depletion of TFAM in erythroid cells alters intracellular metabolism, leading to elevated histone acetylation, deregulated gene expression, and defective mitochondria and erythropoiesis. Mechanistically, mTORC1 signaling is enhanced to promote translation of mitochondria-associated transcripts through TOP-like motifs. Genetic and pharmacological perturbation of mitochondria or mTORC1

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use: [http://www.nature.com/authors/editorial\\_policies/license.html#terms](http://www.nature.com/authors/editorial_policies/license.html#terms)

<sup>7</sup>Correspondence should be addressed to S.Z., F.Z. or J.X. (zhou\_feng@fudan.edu.cn, shaozhen@picb.ac.cn or jian.xu@utsouthwestern.edu).

<sup>6</sup>These authors contributed equally to this work

**Author Contributions:** J.X. conceived the project. X.L., M.N., H.C., D.L., Z.G. and J.X. performed experiments and analyzed the data. F.Z. performed the proteomic profiling and analyzed the data. Y.Z., Z.S., M.N., M.L., F.Z., and J.X. performed bioinformatic analyses. R.A.J.S. performed HPG incorporation experiments and analyzed the data. Z.H., M.N. and R.J.D. performed the metabolomic profiling and analyzed the data. S.E.W. and N.S.C. contributed the *Tfam* flox mouse strain. J.X. wrote the manuscript and X.L., Y.Z., M.N., K.E.D., F.Z., and Z.S. edited it.

The authors declare no competing financial interests.

specifically impairs erythropoiesis *in vitro* and *in vivo*. Our studies support a mechanism for post-transcriptional control of erythroid mitochondria and may have direct relevance to hematologic defects associated with mitochondrial diseases and aging.

---

## Introduction

Cellular differentiation requires highly coordinated gene expression through transcriptional and post-transcriptional mechanisms. Advances in high-throughput sequencing enabled genome-scale quantification of mRNA for gene expression, yet the extent to which changes in mRNA are translated to protein remains unclear. Mass spectrometry (MS)-based proteomics provide an approach to globally interrogate protein abundance in human tissues<sup>1,2</sup>. We reasoned that comparing the proteomic and transcriptomic profiles might provide insights into the post-transcriptional pathways in developmental specification, which would otherwise have been overlooked using only transcriptome-based analysis.

The development of hematopoietic stem/progenitor cells (HSPCs) into lineage-committed cells serves as a paradigm for studying gene regulatory events governing cellular differentiation<sup>3</sup>. Among them, the specification of erythroid lineage is of particular importance. An exquisite coordination of erythroid maturation is essential for continuous production of red blood cells (RBCs), and their imbalance underlies the pathogenesis of various disorders including  $\beta$ -thalassemia and anemia of chronic disease<sup>4</sup>. Erythroid progenitors and erythroblasts are extraordinarily proliferative, while cell growth decreases dramatically in later maturation. As hemoglobin content reaches a threshold, erythroblasts enucleate to become reticulocytes, exit the bone marrow (BM), clear the transferrin receptors, and extrude the remaining organelles including mitochondria. Little is known about how differentiating erythroid cells coordinate programs of proliferation, maturation and metabolism with the changing environment and cellular processes.

Through an unbiased analysis of proteomic and transcriptomic changes during human erythropoiesis, we uncovered major pathways related to mitochondrial biogenesis enhanced through post-transcriptional pathways. In-depth molecular and genetic studies established a functional link between mitochondrial metabolism and epigenetic regulation required for normal erythropoiesis. Our results support a model in which the erythroid mitochondria are regulated by mTORC1-mediated protein translation and may have direct relevance to hematologic disorders associated with mitochondrial dysregulation.

## Results

### Comparative Proteomic and Transcriptomic Profiling of Erythropoiesis

Human CD34<sup>+</sup> HSPCs were differentiated *ex vivo* into erythroid progenitors (proerythroblasts or ProEs) and late erythroblasts (ERYs)<sup>5</sup> (Fig. 1a). Fetal or adult HSPCs (F0 or A0) and lineage-committed ProEs (F5 or A5) were selected for genome-scale quantification of proteins and mRNAs by MS-based proteomics and RNA-seq analysis, respectively (Fig. 1b). Proteomic profiling using multiplexed isobaric tags (iTRAQ) resulted in identification and quantification of proteins encoded by 14,502 genes, accounting for

72.4% of annotated protein-coding genes<sup>6</sup> (Supplementary Table 1). Among them, 92.0% (13,341 of 14,502) were assigned to corresponding transcripts detected by RNA-seq (Fig. 1c). The proteomic and transcriptomic profiling achieved comparable ranges of quantification (Fig. 1d). A survey of representative genes demonstrated reproducible increases in expression of erythroid-associated markers (CD36, CD71 and CD235a), hemoglobins (HBE1, HBG1, HBB and HBA1), transcription factors (TFs) (GATA1 and KLF1) and enzymes (CA1 and FECH) in ProEs relative to HSPCs. Similarly, the expression of HSPC-associated markers and TFs was significantly decreased in ProEs (Fig. 1e). In total, we identified 1,071 to 2,598 differentially expressed mRNAs (fold change  $\geq 1.5$ ,  $P$ -value  $\leq 0.01$ ; Supplementary Fig. 1a-h and Table 2) and 393 to 1,663 differentially expressed proteins (fold change  $\geq 1.3$ ,  $P$ -value  $\leq 0.01$ ; Supplementary Fig. 1i-p and Table 3) associated with distinct biological processes between fetal or adult HSPCs and ProEs, respectively.

### Post-Transcriptional Control of Proteins Associated with Mitochondrial Biogenesis

To compare mRNA and protein-level expression, we overlaid the transcriptomic and proteomic profiles between adult HSPCs (A0) and ProEs (A5) (Fig. 2a). Strikingly, 1,050 out of 1,549 (67.8%) upregulated proteins were not paralleled by changes in cognate mRNAs. Similarly, 399 out of 785 (50.8%) downregulated proteins had no corresponding change in mRNAs. Similar patterns were observed between fetal HSPCs and ProEs (Supplementary Fig. 1r-t). We also identified differentially expressed mRNAs with no changes in proteins, consistent with regulation through mRNA stability, translation, and/or protein stability.

Hereafter we focused on adult HSPCs and ProEs. We categorized the upregulated genes into three groups (RNA-only, Both, and Protein-only; Fig. 2a; Supplementary Table 4) based on the changes in mRNA and protein expression. By gene ontology (GO), the ‘Both’ genes are highly associated with heme biosynthetic processes that are hallmarks of erythroid differentiation. Strikingly, the most enriched pathways in ‘Protein-only’ genes are related to mitochondrial biogenesis, including ATP biosynthesis, electron transport chain, and oxidative phosphorylation (Fig. 2b; Supplementary Fig. 2a-e). Known mitochondrial genes<sup>7,8</sup> are enriched most significantly in the upregulated ‘Protein-only’ genes (Supplementary Figure 2f,g).

Notably, 25 of 48 proteins related to ‘ATP synthesis coupled electron transport’ were significantly upregulated with little or no changes in mRNAs in ProEs (Fig. 2c). Similarly, 12 of 14 proteins related to ‘mitochondrial ATP synthesis coupled proton transport’ were markedly upregulated without changes in mRNAs (Fig. 2d). These findings were validated by quantitative RT-PCR (qRT-PCR) and Western blot analyses of major components of the mitochondrial H<sup>+</sup>-ATP synthases ATP5A1, ATP5B, ATP5D and ATP5O (Fig. 2e-g). Our results suggest that the expression of mitochondrial proteins is enhanced through post-transcriptional mechanisms during erythropoiesis.

### Mitochondria Are Highly Regulated During Erythroid Differentiation

Mitochondria are critical for heme and iron metabolism, yet their regulation during erythropoiesis remains elusive. We measured mitochondrial mass and activities in CD34<sup>+</sup>

HSPCs, ProEs and late erythroblasts (Fig. 3a). Consistent with enhanced expression of mitochondrial proteins, we observed progressive increases in mitochondrial DNA (mtDNA), mitochondrial mass, membrane potential (MMP) and intracellular ATP during erythroid specification (days 3 to 7) before clearance of mitochondria upon terminal maturation (Fig. 3b-e). By profiling of metabolites involved in major bioenergetic pathways, we observed progressive alterations of the central carbon and other metabolic pathways, such as purine and amino acid metabolism, during erythropoiesis (Fig. 3f,g). Moreover, mtDNA was significantly elevated during erythropoiesis in mouse fetal liver (FL) and BM (R2 and R3 stages<sup>9</sup>; Supplementary Fig. 3).

### TFAM and PHB2 Are Post-Transcriptionally Regulated During Erythropoiesis

We next surveyed annotated human TFs<sup>10,11</sup> and identified 28 TFs whose proteins but not mRNAs were significantly increased during erythropoiesis (Fig. 4a). Notably, Prohibitin 2 (PHB2) and mitochondrial transcription factor A (TFAM) were ranked second and third based on the significance of increased protein expression (Fig. 4a). TFAM is essential for transcription and replication of mitochondrial genome<sup>12,13</sup>, while PHB2 plays a critical scaffolding role on mitochondrial inner membrane. The expression of TFAM or PHB2 proteins and mRNAs was validated by qRT-PCR and Western blot (Fig. 4b-d). TFAM and PHB2 proteins were also increased in mouse FL and BM erythropoiesis (Supplementary Fig. 3). Furthermore, the half-lives of TFAM and PHB2 proteins, as determined by cycloheximide (CHX) chase experiments, were comparable between HSPCs and ProEs, suggesting that the elevated protein expression is not due to increased protein stability (Fig. 4e,f).

### TFAM and PHB2 Are Indispensable for Mitochondria and Erythropoiesis

To determine the role of TFAM and PHB2, we employed shRNA-mediated depletion (Fig. 4g). TFAM or PHB2-deficient ProEs displayed significant decreases in mtDNA, mitochondrial mass, MMP, and ATP compared to control cells expressing the non-targeting shRNA (shNT) (Fig. 4h-j). By measuring the incorporation of the methionine analogue L-homopropargylglycine (HPG)<sup>14</sup>, we observed markedly reduced rate of protein synthesis in TFAM or PHB2-depleted early (CD71<sup>+</sup>CD235a<sup>-</sup>) and late (CD71<sup>+</sup>CD235a<sup>+</sup>) erythroid progenitors (Fig. 4k). TFAM or PHB2 depletion impaired erythroid differentiation, proliferation and increased apoptosis (Fig. 4l-n). TFAM or PHB2-depleted cells appeared morphologically immature with decreased expression of hemoglobins and *GATA1* (Supplementary Fig. 4a-c). We then performed RNA-seq analysis of control, TFAM and PHB2-depleted CD71<sup>+</sup>CD235a<sup>+</sup> cells. By gene-set enrichment analysis (GSEA)<sup>15</sup>, we observed that the mitochondrial<sup>8</sup> and ProE-specific genes (Supplementary Table 2) were significantly downregulated whereas HSPC-specific genes (Supplementary Table 5) were upregulated in TFAM or PHB2-depleted cells (Fig. 4o). These findings demonstrate that TFAM or PHB2 loss impaired the silencing of HSPC genes and the expression of genes required for mitochondria and erythropoiesis.

### TFAM Is Indispensable for Erythropoiesis *In Vivo*

To establish the *in vivo* function, we generated the erythroid-specific *Tfam* knockout (KO) mice by EpoR-Cre<sup>16</sup>. We found that the embryonic day (E)13.5 *EpoR-Cre<sup>+</sup>;Tfam<sup>fl/fl</sup>*

embryos (hereafter called *Tfam* KO) were pale and anemic compared to littermates (Fig. 5a). *Tfam* protein was absent in CD71<sup>+</sup>Ter119<sup>+</sup> KO erythroid cells and reduced in *EpoR-Cre<sup>+</sup>;Tfam<sup>fl/+</sup>* (heterozygous KO or HET) erythroid cells (Fig. 5b). Consistent with the reduced size (Fig. 5a), the cellularity of *Tfam* KO fetal livers is 45% of the wild-type (WT) controls (Fig. 5c). Erythroid maturation can be analyzed by the expression of CD71 and Ter119 markers (R1 to R5)<sup>9</sup> (Fig. 5d). *Tfam* KO led to substantial increases in R1 and R2 cells and a marked decrease in R3 erythroid progenitors (Fig. 5e,f). *Tfam* HET mice displayed intermediate phenotypes. *Tfam* KO also led to decreased mitochondrial mass and significantly increased apoptosis (Fig. 5g,h; Supplementary Fig. 5).

We then performed RNA-seq of *Tfam* WT and KO CD71<sup>+</sup>Ter119<sup>+</sup> E13.5 erythroid cells. We identified 106 significantly upregulated and 249 downregulated genes in *Tfam* KO (Supplementary Table 6). The downregulated genes are highly associated with mitochondrial metabolism (Fig. 5i,j). The mitochondrial signature genes<sup>8</sup> and genes associated with reactive oxygen species, oxidative phosphorylation, arginine and proline metabolism were markedly downregulated in *Tfam* KO (Fig. 5k). Together these analyses provide compelling evidence that *Tfam* is indispensable for proper control of metabolism and gene expression during erythropoiesis *in vivo*.

### TFAM Loss Leads to Altered Metabolism and Histone Hyperacetylation

Changes in metabolism often result in altered epigenetic regulation<sup>17</sup>. We then measured histone modifications and observed significant increases in the levels of acetylated histones in TFAM-deficient erythroid cells (Fig. 6a; Supplementary Fig. 6a-d). Acetylated histones including H3K27ac are associated with gene activation. By ChIP-seq, we observed that TFAM depletion led to markedly increased H3K27ac at the promoters of the HSPC-specific genes (Fig. 6b,c; Supplementary Fig. 6e), consistent with the increased expression of HSPC genes. In total, 1,214 ChIP-seq peaks showed upregulated H3K27ac in TFAM-deficient erythroid cells (Fig. 6d). The upregulated peaks were highly associated with lymphoid development and hematopoiesis (Fig. 6e), suggesting that histone hyperacetylation results in altered epigenetic regulation.

To establish the causality, we performed metabolomic analysis and identified 32 and 39 significantly upregulated and downregulated metabolites in TFAM-depleted erythroid cells (Fig. 6f,g; Supplementary Fig. 6f,g). Notably,  $\beta$ -hydroxybutyrate ( $\beta$ OHB) was among the most upregulated metabolites. The ketone body  $\beta$ OHB is synthesized from acetyl-CoA by  $\beta$ -hydroxybutyrate dehydrogenase as alternative energy source<sup>18</sup>. Importantly,  $\beta$ OHB and the related butyrate are potent inhibitors of class I and II histone deacetylases (HDACs)<sup>18</sup>. Thus, our results suggest that the increased histone acetylation may be due to HDAC inhibition. Consistently, TFAM depletion led to significantly decreased HDAC activities (Fig. 6h). The level of acetyl-CoA was also lower (Fig. 6i), suggesting that the increased histone acetylation was not due to enhanced histone acetyltransferase activity<sup>19</sup>. Furthermore, human erythroid cells treated with  $\beta$ OHB or butyrate markedly enhanced histone acetylation, increased expression of the HSPC gene *GATA2* and impaired erythroid differentiation in a dose-dependent manner (Fig. j-l).

Together our results suggest that loss of TFAM causes defective mitochondria and metabolism, resulting in increased  $\beta$ OHB and HDAC inhibition. Increased histone acetylation interferes with the developmental silencing of HSPC genes, resulting in impaired erythroid gene expression and differentiation. Hence, our findings provide a molecular link between mitochondrial biogenesis, metabolism and epigenetic regulation required for normal erythropoiesis (Fig. 6m).

### Translation of Mitochondria-Related mRNAs is Enhanced During Erythropoiesis

Increased translation provides an efficient way to control protein production in stem cells<sup>14</sup> and cancer<sup>20,21</sup>. We hypothesized that mitochondria-associated mRNAs are regulated through enhanced protein translation during erythropoiesis. Consistently, the rate of protein synthesis was the lowest in CD34<sup>+</sup> HSPCs and increased by 2.1-fold in differentiated CD34<sup>-</sup> populations (Fig. 7a). Importantly, protein synthesis was increased by 1.8 and 2.5-fold in early (CD71<sup>+</sup>CD235a<sup>-</sup>) and late (CD71<sup>+</sup>CD235a<sup>+</sup>) erythroid progenitors, suggesting that erythroid differentiation from HSPCs is accompanied by enhanced protein translation. Furthermore, we performed polysome profiling and observed increased polysome size, indicating enhanced protein translation in ProEs relative to HSPCs (Fig. 7b). By qRT-PCR analysis of the polysomal fractions, we observed enhanced translation of mitochondria-related mRNAs in ProEs, as illustrated by a shift of these mRNAs towards larger polysomes (Fig. 7c). The translation of *GATA1* but not *ACTB* was also increased, suggesting that erythroid differentiation is accompanied by enhanced translation of a subset of, but not all, mRNAs.

### Enhanced mTORC1 Signaling Regulates Mitochondrial Biogenesis During Erythropoiesis

The serine/threonine kinase mTOR is a major regulator of protein translation<sup>22</sup>. Genetic or pharmacologic perturbation of mTORC1, but not mTORC2, impaired erythroid development in mouse<sup>23,24</sup>. Activated mTORC1 phosphorylates 4EBP1 and S6K to promote protein translation<sup>25</sup>. Consistently, we observed a gradual increase of mTORC1 signaling, as measured by the levels of phosphorylated 4EBP1 and S6K, and mitochondrial proteins during erythroid differentiation (Fig. 7d). Moreover, HSPCs treated with rapamycin or active-site mTOR inhibitors (asTORi) PP242<sup>26</sup> and Torin1<sup>27</sup> decreased 4EBP1 phosphorylation and mitochondrial proteins, mtDNA, MMP, ATP, protein synthesis, and impaired erythroid differentiation (Fig. 7e-k and Supplementary Fig. 7d-f).

To assess the *in vivo* effects, we treated adult mice with PP242 and Torin1 (Fig. 7l). Acute mTORC1 inhibition had profound effects on 4EBP1 phosphorylation, TFAM and PHB2 expression in CD71<sup>+</sup>Ter119<sup>+</sup> erythroid cells (Fig. 7l), and decreased RBCs (Fig. 7m). While the BM cellularity remained unchanged, the frequencies of immature (CD71<sup>+</sup>Ter119<sup>+</sup>) and mature (CD71<sup>-</sup>Ter119<sup>+</sup>) erythroid cells were significantly lower in asTORi-treated mice (Fig. 7n). mTORC1 inhibition led to significant increased R1/R2 progenitors and decreased R3/R4 erythroid cells, similar to *Tfam* KO (Fig. 7o) and mice harboring deletion of the mTOR kinase<sup>28</sup>. Furthermore, mTORC1 inhibition significantly decreased mitochondrial mass and MMP in R2 and R3 cells (Fig. 7p,q), indicating that the erythroid commitment from R2 to R3 stages is selectively sensitive to mTORC1 inhibition.

We next determined the effects of mTORC1 hyperactivation by depleting TSC1 and TSC2, the negative regulators of mTORC1<sup>29</sup>. TSC1 or TSC2 depletion in erythroid cells significantly increased mtDNA, mitochondrial mass, and MMP (Fig. 7r; Supplementary Fig. 7g-i). mTORC1 hyperactivation also impaired erythroid differentiation and increased apoptosis (Supplementary Fig. 7j,k) consistent with previous studies<sup>23</sup>, suggesting that the coordinated control of mTORC1 and mitochondria is essential for erythropoiesis. Similarly, hematopoietic-specific KO of *Pten*, a negative regulator of AKT and mTORC1<sup>30</sup>, significantly augmented mitochondrial mass and impaired erythroid differentiation (Supplementary Fig. 7l,m). Additionally, overexpression of TFAM increased mtDNA and rendered erythroid cells less sensitive to mTORC1 inhibition (Supplementary Fig. 7n-t). Together our results demonstrate that mTORC1 inhibition impaired translation of mitochondria-associated proteins during erythropoiesis.

### Translation of Mitochondria-Associated Proteins is Sensitive to mTORC1 Inhibition

To investigate whether erythroid mitochondria are selectively sensitive to mTORC1 inhibition, we differentiated HSPCs into erythroid cells and granulocytes (Fig. 8a, Supplementary Fig. 8a-c). By polysome profiling, we observed that PP242 markedly inhibited protein synthesis in erythroid cells, as illustrated by decreased polysome content with a concomitant increase in the 80S peak (Fig. 8b). PP242 did not cause complete disassembly of polysomes, indicating that it selectively impaired a subset of mRNAs<sup>31,32</sup>. Notably, PP242 suppressed the translation of mitochondria-associated mRNAs by shifting these mRNAs towards lighter polysome and monosome fractions (Fig. 8c). In contrast, mTORC1 inhibition in granulocytes had minimal impact on polysome profiles or distribution of mitochondrial mRNAs. Granulocyte differentiation remained ostensibly normal upon mTORC1-inhibition or TFAM-depletion. Moreover, granulocyte-specific *Tfam* KO by Mrp8-Cre-ires-GFP<sup>33</sup> had minimal effect on Mac1<sup>+</sup>Gr1<sup>+</sup> granulocytes or the expression of granulocytic genes (Supplementary Fig. 8d-l), suggesting a distinctive requirement for *Tfam* in erythroid versus granulocyte differentiation.

We next performed iTRAQ-based proteomics of PP242-treated erythroid cells and granulocytes (Fig. 8d). Specifically, 935 proteins were significantly downregulated (fold change > 1.3, *P*-value < 0.01) in PP242-treated erythroid cells. Among them, 205 and 245 are Protein-only (Fig. 2a) and mitochondria-associated genes<sup>7,8</sup>, which is significantly more enriched than genomic average. In contrast, among 529 downregulated proteins in PP242-treated granulocytes, only 50 and 33 are Protein-only and mitochondrial genes, respectively. Specifically, the expression of mitochondria-associated proteins was downregulated in PP242-treated erythroid cells but not granulocytes (Fig. 8e). Moreover, mitochondria-associated pathways were among the top enriched GO terms in downregulated proteins in erythroid cells but not granulocytes (Fig. 8f). These findings suggest that mitochondria-associated mRNAs are selectively sensitive to enhanced protein translation in erythroid cells.

### Mitochondria-Associated mRNAs are Regulated by mTORC1

Transcripts regulated by mTORC1 typically have 5' terminal oligopyrimidine (TOP) or TOP-like motifs<sup>32</sup> consisting of at least five pyrimidines within four nucleotides of transcription start site (TSS). By mapping TSS using dbTSS<sup>34</sup>, RefSeq and Ensembl, we

found that *TFAM*, *PHB2*, *ATP5A1*, *ATP5B* and *ATP5D* transcripts contain TOP or TOP-like motifs, whereas *ATP5O* contains a translation initiator of short 5' UTR (TISU) motif found in mTOR-sensitive non-TOP mRNAs<sup>35-37</sup> (Fig. 8g). The Protein-only genes were significantly enriched with TOP-like motifs (Fig. 8h).

To test the role of TOP-like motifs, we engineered *TFAM*, *PHB2* and *ATP5D* 5' UTRs containing the wild-type (WT) or mutated (MUT) TOP-like motifs to a GFP reporter and a TK promoter (Fig. 8i and Supplementary Fig. 8m). Upon integration into erythroid cells, the MUT sequences decreased GFP expression compared to WT. Notably, PP242 treatment led to significant downregulation of GFP containing WT but not MUT 5' UTRs, suggesting that the TOP-like motifs mediate the modulation of protein translation by mTORC1. In contrast, the expression of GFP remained low and unchanged in granulocytes in the presence or absence of PP242.

Our results support a model that mitochondrial biogenesis is coordinately regulated through mTORC1-mediated protein translation during erythropoiesis (Fig. 8j). Specifically, as erythroid cells differentiate, they acquire erythropoietin (Epo) responsiveness, accumulate iron, and become highly proliferative. Both Epo and iron activate mTORC1<sup>23,38</sup>, which phosphorylates 4EBP1 and S6K to promote synthesis of mitochondrial proteins. Stimulating mitochondrial activity engenders not only intracellular metabolism, but also increased ATP to fuel increasing cellular processes such as hemoglobin synthesis. Thus our studies identify an essential role for the mTORC1-mitochondria axis in coordinating protein translation and energy metabolism. Mitochondrial dysfunction leads to a variety of human disorders with impaired erythropoiesis being a severe manifestation in a subset of these diseases. Our findings may have direct relevance to the hematologic defects associated with mitochondrial diseases and aging.

## Discussion

Here we describe the genome-scale comparison of transcriptomic and proteomic changes during human erythropoiesis, and identify mTORC1 as a major modulator of mitochondrial biogenesis. mTORC1 signaling is enhanced through Epo-JAK-AKT signaling<sup>38</sup> or intracellular iron<sup>23</sup> during erythroid differentiation. Elevated mTORC1 increases the translation of mitochondrial mRNAs through the TOP-like motifs, resulting in enhanced mitochondrial activities. mTORC1 also inhibits autophagy, a major pathway in mitochondria degradation<sup>39</sup>. Moreover, mTORC1 promotes mitochondrial function through YY1-PGC1 $\alpha$  and HIF1 $\alpha$  transcriptional complex<sup>40,41</sup>. Hence, it is likely that mTORC1 stimulates mitochondrial function by orchestrating translation and transcription of distinct mitochondria-related genes.

Accumulating evidence indicates that deregulated protein translation underlies many blood disorders. Of note, 5q-syndrome and Diamond-Blackfan anemia (DBA) are caused by ribosomal protein haploinsufficiency<sup>42</sup>, and can be ameliorated by L-leucine which activates mTORC1 signaling<sup>43-45</sup>. Thus, our findings suggest that inhibition of mTOR not only impacts protein translation<sup>45</sup>, but also mitochondrial biogenesis resulting in energy imbalance and metabolic dysfunction. Paradoxically, mTORC1 is viewed as a general



regulator of protein translation. How and why are erythroid cells specifically sensitive to mTORC1? By comparing two hematopoietic lineages, we provide evidence that mitochondrial protein translation and erythropoiesis are selectively sensitive to mTORC1. First, polysome analysis showed increased translation of mitochondria-associated proteins in erythrocytes but not granulocytes. Second, TFAM depletion or mTORC1 inhibition impaired mitochondria and erythropoiesis, whereas granulocytes were largely spared. Third, acute inhibition of mTORC1 in mice impaired erythropoiesis with minimal impact on other hematopoietic lineages. Finally, while *Tfam* KO was incompatible with erythropoiesis, granulocytes remained ostensibly normal.

Our studies suggest that enhanced mitochondrial biogenesis may be a unique feature of erythropoiesis. Developing erythrocytes are characterized by increasing heme and iron-sulfur cluster biosynthesis in the mitochondria. TFAM-deficient erythroid cells failed to silence HSPC-specific genes, increased histone acetylation, and displayed defective metabolism. Most importantly, the ketone body  $\beta$ OHB, a marker of mitochondrial deficiency in human, was significantly upregulated which functions as a potent HDAC inhibitor. Increased  $\beta$ OHB impaired erythroid differentiation and gene expression similar to TFAM-depletion or mTORC1-inhibition. Therefore, our findings provide a molecular link between mitochondrial biogenesis, metabolism and epigenetic regulation indispensable for erythropoiesis. Furthermore, our findings provide a mechanistic explanation for the broader notion that RBCs are uniquely sensitive to deregulated protein translation and/or mitochondria. In light of recent development of inhibitors for protein translation<sup>46</sup>, our study raises the possibility that modulation of mitochondrial protein translation may provide therapeutic opportunities for hereditary anemias and mitochondria-based syndromes. Finally, our results demonstrate how unbiased analysis of proteomic and transcriptomic profiles may identify pathways in cellular regulation. With the increasing availability of genomic and proteomic landscapes, the integrative approach developed here can be extended to the systematic investigation of the gene-protein-pathway networks in health and disease.

## Methods

### Cells and cell culture

Human adult CD34<sup>+</sup> HSPCs were obtained from the Fred Hutchinson Cancer Research Center. Fetal liver CD34<sup>+</sup> HSPCs were isolated as described<sup>1</sup>. ProEs were generated as described<sup>2</sup>. Briefly, CD34<sup>+</sup> HSPCs were cultured in StemSpan SFEM with 1xCC100 cytokine cocktail (StemCell Technologies Inc.) and 2% penicillin-streptomycin for 6 days with media changes every 2~3 days. On day 6, cells were transferred into StemSpan SFEM with 2% penicillin-streptomycin, Epo (1U/ml), IL-3 (5ng/ml), SCF (20ng/ml), dexamethasone (2 $\mu$ M) and  $\beta$ -estradiol (1 $\mu$ M) with media changes every 2~3 days. Primary granulocytes were generated from CD34<sup>+</sup> HSPCs as described<sup>3</sup> with modifications. Briefly, HSPCs were expanded in StemSpan SFEM with 1xCC100 and 2% penicillin-streptomycin for 3 days. Cells were transferred to IMDM with 10% fetal bovine serum (FBS), IL-3 (100ng/ml) and SCF (100ng/ml) for 3 days. Cells were then transferred to IMDM with 10% FBS, IL-3 (100ng/ml), SCF (100ng/ml) and G-CSF (10ng/ml), and maintained at 2~8 $\times$ 10<sup>5</sup> cells/ml with media changes every 2~3 days. Maturing granulocytes (neutrophils) typically

emerge by 7~10 days and can be cultured up to 14 days. For mTORC1 inhibition, cells were treated with control (DMSO), PP242 (2.5 and 10 $\mu$ M), Torin1 (0.25 and 1 $\mu$ M) or rapamycin (10nM, 100nM and 1 $\mu$ M) for 5~7 days. No cell lines used in this study were found in the database of commonly misidentified cell lines that is maintained by ICLAC and NCBI BioSample. All cell lines were tested for mycoplasma contamination.

## Mice

To obtain *Tfam* conditional knockout mice, the *Tfam*<sup>fl/fl</sup> mice<sup>4</sup> were crossed with the EpoR-Cre knock-in mice<sup>5</sup> or the Mrp8-Cre-ires-GFP transgenic mice<sup>6</sup>. To obtain *Pten* conditional knockout mice, the *Pten*<sup>fl/fl</sup> mice<sup>7</sup> were crossed with the *Mx1-Cre* mice<sup>8</sup>. *Pten* deletion was achieved by intraperitoneal (i.p.) administration of five doses of polyI-polyC (0.8 $\mu$ g/g of body weight). PP242 and Torin1 were prepared at 1mg/ml and 2mg/ml in PBS with 5% PEG400 and 5% Tween80, and administrated via i.p. injection daily at 10mg/kg and 20mg/kg for 3 days, respectively. All experiments were performed with the approval of the University of Texas Southwestern Institutional Animal Care and Use Committee.

## Quantitative proteomics analysis

Proteomic analysis was performed as described with modifications<sup>9</sup>. We optimized major components of the platform to achieve genome-scale coverage. First, we used high-pH reversed phase (RP) and strong anion exchange (SAX) separation stages coupled with a narrow-bore extended-length low-pH RP column to achieve temporal separation of peptides in a nanoflow regime. Second, we chose the final dimension column geometry to maintain the integrity of chromatographic separation at ultra-low effluent flow rates to maximize electrospray ionization efficiency. Third, we implemented separation stages in microcapillary format coupled to the mass spectrometer, thus providing automated and efficient transfer of peptides across all separation stages.

**Sample preparation for DEEP SEQ analysis**—Proteins were extracted with 8M urea in 0.1M NH<sub>4</sub>HCO<sub>3</sub>, reduced by 10mM Dithiothreitol (DTT) for 60min, and alkylated by 20mM methyl methanethiosulfonate (MMTS) for 30min. After overnight trypsin digestion at 37°C, the peptides were cleaned using SepPak C18 RP cartridges (Waters Corp.), labeled with 4-plex iTRAQ reagents (AB Sciex), mixed and loaded into an online chromatography.

**Multi-dimension separation and data acquisition**—Nanoscale three dimensional online chromatography consists of first dimension RP column (200 $\mu$ m I.D. capillary packed with 20cm of 5 $\mu$ m dia. XBridge C18 resin), second dimension SAX column (200 $\mu$ m I.D. 20cm of 10 $\mu$ m dia. POROS10HQ resin) and third dimension RP column (25 $\mu$ m I.D. 120cm of 5 $\mu$ m dia. Monitor C18, integrated 1 $\mu$ m dia. emitter tip). The final dimension ran at ~5nl/min with a 600min gradient from 2% A to 50% B (A=0.1% formic acid, B=acetonitrile with 0.1% formic acid). The downstream TripleTOF 5600+ (AB Sciex) was set in information dependent mode (IDA). Top 50 precursors (charge state +2 to +5, >100 counts) in each MS scan (800ms, scan range 350-1500m/z) were subjected to MS/MS (minimum time 140ms, scan range 100-1400m/z). Electrospray voltage was 2.2kV.

**Data processing and protein identification**—MS data was subjected to search against SwissProt database with ProteinPilot V.4.5 (AB Sciex) with the search parameter ‘iTRAQ 4-plex (peptides labeling) with 5600 TripleTOF’. The peptide spectra match (PSM) false discovery rate (FDR) 1% was used to filter peptides. We then summed the intensity of each iTRAQ reporter ion for the peptides assigned to single gene to generate a ratio for each gene.

**Protein quantification**—To determine the significance of the differences in protein expression, we used a statistical model as described<sup>10</sup> with modifications. Specifically, two hypotheses are used: (1) the  $\log_2$  ratio of a non-differentially expressed gene's protein counts in two samples follows normal distribution  $N(0, \sigma^2)$ , in which 0 is the mean and  $\sigma^2$  is the variance that decays exponentially with the gene's average signal intensity<sup>10</sup>; and (2) the majority of genes are non-differentially expressed. In this model, a sliding window of 500 genes moving along the axis of  $\log_{10}(\text{intensity})$  was introduced and the distribution of  $\log_2$  ratio of genes was considered as a mixture of the normal distribution  $N(0, \sigma^2)$  contributed by non-differentially expressed genes and  $\log_2$  ratio far from 0 contributed by differentially expressed proteins. To get the  $\sigma$ , we applied linear regression between the percentiles of the  $\log_2$  ratio against the theoretical percentiles of normal distribution  $N(0, 1)$  from 25% to 75%. The slope was taken as an approximation of the  $\sigma$  of the normal distribution followed by the underlying non-differentially expressed proteins. We fitted an exponential decay function  $\sigma^2 = \beta \times \exp(\alpha \times \log_{10}(\text{average intensity of window}))$  between the  $\sigma^2$  of each window and the average intensity of the proteins. Finally, the  $\sigma^2$  of each gene is calculated from the exponential decay function to measure the technical variations and to determine the significance (*P*-value) of its  $\log_2$  ratio. A protein was considered to be differentially expressed if the following criteria were met in at least two out of the five replicates: (1) iTRAQ intensity  $\geq 300$ ; (2) iTRAQ ratio  $\geq 1.3$ ; and (3) *P*-value  $\leq 0.01$ . However, because peptide ratios measured by MS can be distorted during precursor ion isolation<sup>11</sup>, the variation in some protein expression may be substantially underestimated. The proteomic experiments were performed as five independent proteome measurements and biological triplicates for the erythroid differentiation experiments. We identified 461,962 distinct peptides (FDR 1%) from 12,892,243 spectra. After removing genes with weak quantification signal (iTRAQ intensities  $< 300$ ), we achieved high-quality quantification of proteins encoded by 14,502 genes from 350,843 unique peptide sequences, accounting for 72.4% of annotated RefSeq genes<sup>12</sup>.

### Polysome profiling

Polysome profiling was performed as described with modifications<sup>13</sup>. Briefly, cells were treated with 100 $\mu\text{g/ml}$  cycloheximide (CHX) for 5min, harvested, washed twice with PBS containing 100 $\mu\text{g/ml}$  CHX, and centrifuged at 300 $\times g$  for 5min at 4 $^\circ\text{C}$ . Cells were resuspended in 425 $\mu\text{l}$  hypotonic buffer (5mM Tris-HCl pH7.5, 2.5mM MgCl<sub>2</sub>, 1.5mM KCl, protease inhibitor cocktail) and transferred to a pre-chilled 1.5ml tube. Then 5 $\mu\text{l}$  of 10mg/ml CHX, 1 $\mu\text{l}$  of 1M DTT, and 100U RNase inhibitors were added to cells followed by vortexing for 5sec. Then 25 $\mu\text{l}$  of 10% Triton X-100 and 25 $\mu\text{l}$  of 10% sodium deoxycholate were added to cells followed by vortexing for 5sec. The supernatant were transferred to a pre-chilled 1.5ml tube after centrifugation at 21,000 $\times g$  for 5min at 4 $^\circ\text{C}$ . The same OD

amount of lysates from each sample were loaded onto a 10-50% sucrose gradient, and centrifuged at 35,000rpm for 2h at 4°C using SW40Ti rotor in a Beckman Coulter Optima L-80 ultracentrifuge with no brake. Samples were analyzed on the Bio-Rad BioLogic LP system and BioFrac fraction collector, and chased with 60% (w/v) sucrose with bromophenol blue at 1ml/min. Data were analyzed using the Bio-Rad LP Data View software. Polysome fractions were collected at 0.5 ml/fraction. For mTORC1 inhibition, cells were treated with DMSO or 2.5µM PP242 for 12h.

### Measurement of protein synthesis

Protein synthesis was measured by the incorporation of L-homopropargylglycine (HPG)<sup>14</sup>. 0.5~1×10<sup>5</sup> CD34<sup>+</sup> HSPCs or ProEs were plated in 100µl of methionine-free DMEM with 200µM L-cysteine, 50µM 2-mercaptoethanol, 1mM L-glutamine, 0.1% BSA (Sigma) and appropriate cytokines. Cells were pre-cultured for 45min to deplete endogenous methionine, followed by 1h incubation with 1mM HPG. Cells were washed twice with PBS and fixed 15min in 0.5ml of 1% paraformaldehyde in PBS on ice, followed by PBS wash and permeabilization in 200µl PBS containing 3% FBS and 0.1% saponin for 5min at RT. The Click-iT Cell Reaction Buffer Kit (Life Technologies) was used to conjugate azide to Alexa Fluor 488 or Alexa Fluor 555. After 30min, cells were washed twice in PBS containing 3% FBS and 0.1% saponin, resuspended in PBS containing 4µg/ml of 4',6-diamidino-2-phenylindole (DAPI), and analyzed by flow cytometry.

### qRT-PCR and RNA-seq analysis

RNA was isolated using RNeasy Plus Mini Kit (Qiagen). qRT-PCR was performed using the iQ SYBR Green Supermix (Bio-Rad). RNA-seq library was prepared using the Truseq RNA Library Prep Kit (Illumina). For RNA-seq in HSPCs and ProEs, the raw reads were aligned to human genome (hg19) by TopHat<sup>15</sup>. Differentially expressed genes were identified by DEseq<sup>16</sup> using fold change 1.5 and *P*-value 0.01. For RNA-seq in fetal liver CD71<sup>+</sup>Ter119<sup>+</sup> erythroid cells, the reads were aligned to mouse genome (mm9) by TopHat<sup>15</sup>. Differentially expressed genes were identified by Cufflinks<sup>17</sup> using fold change 1.5 and *P*-value 0.05.

### Western blot

Western blot was performed as described<sup>2</sup> using the following antibodies: TFAM (7495S, Cell Signaling Technology; sc-23588, Santa Cruz Technology), PHB2 (14085S, Cell Signaling Technology), ATP5A1 (sc-136178, Santa Cruz Technology), ATP5B (sc-16690, Santa Cruz Technology), ATP5D (ab97491, Abcam), ATP5O (ab110276, Abcam, clone 4C11C10D12), 4EBP1 (9644S, Cell Signaling Technology), phos-4EBP1 (Thr37/46) (2855S, Cell Signaling Technology), p70S6K (2708S, Cell Signaling Technology), phos-S6K (Ser235/236) (2211S, Cell Signaling Technology), GATA1 (ab47490, Abcam), GATA2 (ab22849, Abcam), TSC1 (6935, Cell Signaling Technology), TSC2 (4308, Cell Signaling Technology), H3K9ac (9649, Cell Signaling Technology), H3K27ac (8173, Cell Signaling Technology), H3K56ac (4243, Cell Signaling Technology), H4K5ac (ab51997, Abcam), H3K4me3 (9571, Cell Signaling Technology), H3K9me3 (13969, Cell Signaling Technology), H3K27me3 (9733, Cell Signaling Technology), H3K36me3 (4909, Cell Signaling Technology), H3 (4499, Cell Signaling Technology), ACTB (MAB1501,

Millipore, clone C4), and GAPDH (sc-26778, Santa Cruz Biotechnology). Densitometry quantification was performed using ImageJ software. All antibodies were used at 1:1,000 dilutions except phos-4EBP1 (1:500), phos-S6K (1:500) and ACTB (1:2,500).

### ChIP-seq analysis

ChIP-seq was performed as described<sup>2,18</sup> using H3K27ac antibody (ab4729, Abcam) in K562 cells. ChIP-seq libraries were generated using NEBNext ChIP-seq Library Prep Master Mix (NEB). ChIP-seq raw reads were aligned to human genome (hg19) using Bowtie<sup>19</sup>, and peaks were identified using MACS<sup>20</sup>. Differentially enriched ChIP-seq peaks were identified using MANorm<sup>18,21</sup>. MANorm takes the chromosomal coordinates of all peaks and aligned reads in both samples as input. The ( $M$ ,  $A$ ) value of each peak was calculated and plotted, where  $M = \log_2(\text{read density in shTFAM} / \text{read density in shNT})$  and  $A = 0.5 \times \log_2(\text{read density in shTFAM} \times \text{read density in shNT})$ . Robust regression was applied to the ( $M$ ,  $A$ ) values of common peaks and a linear model was derived. The linear model was extrapolated to all peaks for normalization. The normalized  $M$  values represent  $\log_2$ -transformed fold changes of intensities at each peak between two samples. To identify upregulated or downregulated H3K27ac ChIP-seq peaks, we filtered the differentially enriched peaks based on  $M$  values (referred as  $M_{H3K27ac}$ ). For significantly upregulated or downregulated peaks, we kept H3K27ac peaks with  $|M_{H3K27ac}| > 2$ ; for shared H3K27ac peaks we used  $|M_{H3K27ac}| \geq 2$ .

### Mitochondrial DNA, mass and membrane potential

Mitochondrial DNA (mtDNA) was determined using qPCR to measure the ratio of mtDNA versus genomic DNA. The primer sequences are in Supplementary Table 7. Mitochondrial mass was determined using flow cytometry. Briefly,  $0.5 \sim 1 \times 10^6$  cells were incubated with 100nM MitoTracker Green (Invitrogen) and antibodies for cell surface markers for 30min at 37°C, and analyzed on FACSCanto with 488nm excitation. Mitochondrial membrane potential (MMP) was monitored by JC1 staining. Briefly,  $0.5 \sim 1 \times 10^5$  cells were washed with PBS and resuspended in PBS containing 2 $\mu$ M JC1 and incubated at 37°C in 5% CO<sub>2</sub> for 15min. The control cells were incubated with 2 $\mu$ M JC1 with 10 $\mu$ M protonophore FCCP (carbonyl cyanide 4-(trifluoromethoxy) phenylhydrazone) for 15min. Cells were washed, resuspended, and analyzed on FACSCanto.

### Intracellular ATP

Intracellular ATP was determined using the CellTiter-Glo Luminescent Cell Viability Assay (Promega) following manufacturer's protocol.

### Metabolite profiling

Cells ( $3 \sim 5 \times 10^6$ ) were harvested, washed twice with ice-cold saline, overlaid with 500 $\mu$ l of cold methanol/water (50/50, v/v), and subjected to three freeze-thaw cycles. After vortexing, the debris was pelleted by centrifugation at 16,000 $\times g$  and 4°C for 15min and used for protein quantitation. The supernatant was dried using a SpeedVac concentrator, reconstituted in 100 $\mu$ l of 0.03% formic acid, vortex-mixed and centrifuged to remove debris. Targeted metabolite profiling of  $\sim 200$  water soluble endogenous metabolites was performed using a

LC-MS/MS approach. Separation was achieved on a Phenomenex Synergi Polar-RP HPLC column (150×2 mm, 4µm, 80Å) using a Nexera UHPLC system (Shimadzu Corporation, Kyoto, Japan). The mobile phases were 0.03% formic acid in water (A) and 0.03% formic acid in acetonitrile (B). The gradient program was as follows: 0-3min, 100% A; 3-15min, 100%-0% A; 15-21min, 0% A; 21-21.1min, 0%-100% A; 21.1-30min, 100% A. The column was maintained at 35°C. Samples were kept in the autosampler at 4°C with 0.5ml/min flow rate and 10µl injection volume. The mass spectrometer AB QTRAP 5500 (AB SCIEX) was set in multiple reaction monitoring (MRM) mode. Sample analysis was performed in positive/negative switching mode. Declustering potential (DP) and collision energy (CE) were optimized for each metabolite by direct infusion of reference standards. The MRM MS/MS detector conditions were set as follows: curtain gas 30psi; ion spray voltages 1200V (positive) and -1500V (negative); temperature 650°C; ion source gas 150psi; ion source gas 250psi; interface heater on; entrance potential 10V. The MRM transition (*m/z*) and CE (V) of each metabolite were optimized using standard compound. Dwell time for each transition was 3msec. MRM data was acquired using Analyst 1.6.1 software (AB SCIEX). Chromatogram review and peak area integration were performed using MultiQuant software version 2.1 (AB SCIEX). The peak area for each metabolite was normalized against the total ion current of all metabolites. The normalized area values were used as variables for the multivariate and univariate statistical analysis. All multivariate analyses were performed using SIMCA-P (version 13.0.1, Umetrics, Umeå, Sweden). The pre-processed datasets were mean-centered and unit-variance scaled, and evaluated by principal component analysis (PCA). Univariate statistical differences were analyzed using a Student's *t*-test.

### Acetyl-CoA and HDAC assays

Acetyl-CoA was measured using the PicoProbe™ Acetyl-CoA fluorometric assay kit (BioVision K317-100). HDAC activity was measured using the HDAC activity colorimetric assay kit (BioVision K331-100). Specifically, 1.5 million cells were lysed and diluted to 85µl in ddH<sub>2</sub>O in a 96-well plate. 10µl of 10×HDAC assay buffer and 5µl of the HDAC colorimetric substrate were added and mixed thoroughly. The plate was incubated at 37°C for 2h before adding 10µl of lysine developer. The plate was incubated at 37°C for 30min and measured at 405nm. HDAC activity was expressed as the relative O.D. value per mg of protein.

### Lentiviral RNAi

Lentiviral RNAi was performed as described<sup>2</sup>. The shRNAs in the pLKO.1-puro vector for TFAM (sh1: TRCN0000016095; sh2: TRCN0000016093) and PHB2 (sh1: TRCN0000060920; sh2: TRCN0000060921) were obtained from Sigma-Aldrich. The shRNAs in the pGIPZ vector for TSC1 (sh1: V2LHS\_92624) and TSC2 (sh1: V2LHS\_233694; sh2: V2LHS\_93624) were obtained from GE Dharmacon. Sequences of shRNAs are in Supplementary Table 7.

### Flow cytometry

Cells were analyzed using FACSCanto (BD Biosciences) for expression of CD34, CD71, CD235a, or Ter119 conjugated to phycoerythrin (PE), fluorescein isothiocyanate (FITC), or allophycocyanin (APC; BD Pharmingen). Live cells were identified by exclusion of 7-

amino-actinomycin D (7-AAD). Hematopoietic lineages in mouse BM were identified using markers for granulocytes (Gr1<sup>+</sup>Mac1<sup>+</sup>), macrophages (Mac1<sup>+</sup>F4/80<sup>+</sup>), immature (IgM<sup>+</sup>B220<sup>+</sup>) and mature B lymphocytes (IgM<sup>+</sup>B220<sup>+</sup>), immature (CD71<sup>+</sup>Ter119<sup>+</sup>) and mature erythroid cells (CD71<sup>+</sup>Ter119<sup>+</sup>). Data were analyzed using FlowJo software (Ashland, OR).

### Cytospin and hematological indices

Cytopreparations were stained with May-Grunwald-Giemsa as described<sup>22</sup>. Peripheral blood was isolated via the retro-orbital plexus and analyzed on a HEMAVET<sup>®</sup> HV950 (Drew Scientific, Inc.) as described<sup>22</sup>.

### Code availability

Computation codes are available from the corresponding author on request.

### Statistics and reproducibility

Data are represented as mean  $\pm$  standard error of the mean (s.e.m.) or standard deviation (s.d.) of at least three independent experiments as specified in the figure legends. Graphs and statistical analysis were performed using the Microsoft Excel or GraphPad Prism software. Comparisons between groups were performed using the two-sided Student's *t*-test, repeated-measures one-way ANOVA with Dunnett's correction, Fisher Exact *t*-test with Benjamini procedure, or hypergeometric distribution, as denoted in figure legends. The numbers of independent experiments or biological replicate samples and *P*-values (*n.s.* not significant, \**P* < 0.05, \*\**P* < 0.01, \*\*\**P* < 0.001) are provided in individual figures. *P* < 0.05 was considered statistically significant. Panels in Figs 2f, 4c, e, g, 5a, b, d, 6a, j, 7d, e, l, and 8d, e, and Supplementary Figs 3a, c, f, h, 4b, 5a, c, e, 6b, 7d, g, n, and Supplementary Fig. 8a, h show a representative image or flow cytometry of at least three independent experiments or biological replicate samples. Analysis of gene expression was determined using the 2<sup>- $\Delta$</sup>  method with *GAPDH* as the reference gene unless otherwise specified. No statistical method was used to predetermine sample size in animal studies and the experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment.

### Data availability

The RNA-seq datasets for human HSPCs (A0 and F0), ProEs (A5, F5, shNT, shTFAM and shPHB2), E13.5 fetal liver CD71<sup>+</sup>Ter119<sup>+</sup> erythroid cells (*Tfam* WT and KO), and H3K27ac ChIP-seq datasets have been deposited in the Gene Expression Omnibus (GEO) under the accession number GSE86912. Mass spectrometry data have been deposited in ProteomeXchange with the primary accession code PXD006170. Source data for RNA-seq analysis are provided as Supplementary Tables 2, 4, 5 and 6. Source data for proteomic analysis are provided as Supplementary Tables 1, 3 and 4. Statistics source data for Figs 2e, g, 3b-g, 4b, d, f, h-n, 5c, e-h, 6g-l, 7a, c, f-k, m-r, and 8c, i, and Supplementary Figs 3b, d, e, g, i, j, 4c, 6a, c, d, f, 7a-c, e, f, h-m, o-t, and 8c, e-g, i-l are provided in Supplementary Table 8. All other data supporting the findings of this study are available from the corresponding author on request.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

We thank Stuart H. Orkin and Leonard I. Zon for assistance and discussion, Zhipeng Zhou and Yi Liu for assistance with polysome profiling, Ugur Eskiocak, Stacy Yuan, Salma Hasan and Jessica Sudderth for reagents and discussion, and Prashant Mishra for critical reading of the manuscript. We thank Ben van Handel and Hanna Mikkola for providing the fetal CD34<sup>+</sup> cells. This work was supported by NIH grants R35CA197532 (to N.S.C.) and T32HL076139 (to S.E.W.), by the Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institute of Higher Learning (TP2015003) (to F.Z.), by the 100-Talent Program of the Chinese Academy of Sciences (Y516C11851 to Z.S.) and the Science and Technology Commission of Shanghai Municipality (14PJ1410000 to Z.S. and 16ZR1448600 to Y.Z.), by NIH grants K01DK093543, R03DK101665 and R01DK111430, by a Cancer Prevention and Research Institute of Texas (CPRIT) New Investigator award (RR140025), by the American Cancer Society award (IRG-02-196) and the Harold C. Simmons Comprehensive Cancer Center at UT Southwestern, and by an American Society of Hematology Scholar Award (to J.X.).

## References

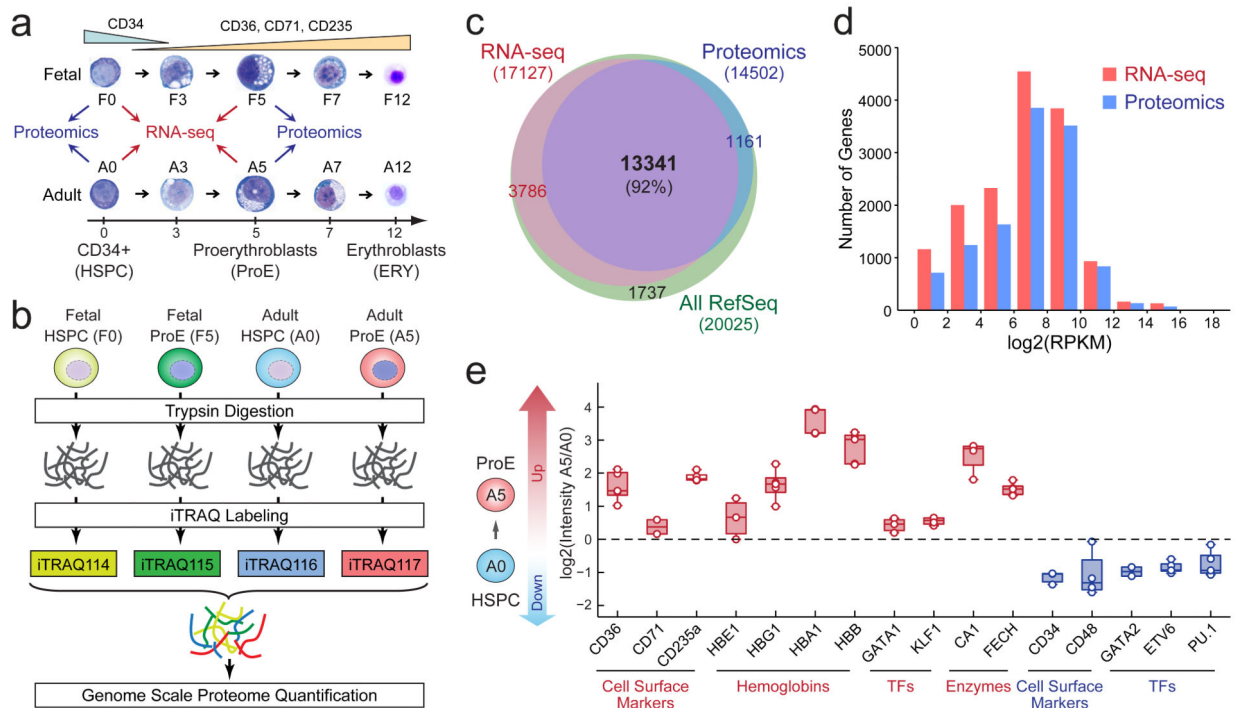
1. Kim MS, et al. A draft map of the human proteome. *Nature*. 2014; 509:575–581. DOI: 10.1038/nature13302 [PubMed: 24870542]
2. Wilhelm M, et al. Mass-spectrometry-based draft of the human proteome. *Nature*. 2014; 509:582–587. DOI: 10.1038/nature13319 [PubMed: 24870543]
3. Orkin SH, Zon LI. Hematopoiesis: an evolving paradigm for stem cell biology. *Cell*. 2008; 132:631–644. [PubMed: 18295580]
4. Rivella S. Ineffective erythropoiesis and thalassemias. *Current opinion in hematology*. 2009; 16:187–194. DOI: 10.1097/MOH.0b013e32832990a4 [PubMed: 19318943]
5. Xu J, et al. Combinatorial assembly of developmental stage-specific enhancers controls gene expression programs during human erythropoiesis. *Developmental Cell*. 2012; 23:796–811. DOI: 10.1016/j.devcel.2012.09.003 [PubMed: 23041383]
6. Pruitt KD, et al. RefSeq: an update on mammalian reference sequences. *Nucleic acids research*. 2014; 42:D756–763. DOI: 10.1093/nar/gkt1114 [PubMed: 24259432]
7. Szklarczyk D, et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic acids research*. 2015; 43:D447–452. DOI: 10.1093/nar/gku1003 [PubMed: 25352553]
8. Mootha VK, et al. Integrated analysis of protein composition, tissue diversity, and gene regulation in mouse mitochondria. *Cell*. 2003; 115:629–640. [PubMed: 14651853]
9. Zhang J, Socolovsky M, Gross AW, Lodish HF. Role of Ras signaling in erythroid differentiation of mouse fetal liver cells: functional analysis by a flow cytometry-based novel culture system. *Blood*. 2003; 102:3938–3946. [PubMed: 12907435]
10. Babu MM, Luscombe NM, Aravind L, Gerstein M, Teichmann SA. Structure and evolution of transcriptional regulatory networks. *Current opinion in structural biology*. 2004; 14:283–291. DOI: 10.1016/j.sbi.2004.05.004 [PubMed: 15193307]
11. Vaquerizas JM, Kummerfeld SK, Teichmann SA, Luscombe NM. A census of human transcription factors: function, expression and evolution. *Nature reviews Genetics*. 2009; 10:252–263. DOI: 10.1038/nrg2538
12. Larsson NG, et al. Mitochondrial transcription factor A is necessary for mtDNA maintenance and embryogenesis in mice. *Nature genetics*. 1998; 18:231–236. DOI: 10.1038/ng0398-231 [PubMed: 9500544]
13. Wang J, et al. Dilated cardiomyopathy and atrioventricular conduction blocks induced by heart-specific inactivation of mitochondrial DNA gene expression. *Nature genetics*. 1999; 21:133–137. DOI: 10.1038/5089 [PubMed: 9916807]
14. Signer RA, Magee JA, Salic A, Morrison SJ. Haematopoietic stem cells require a highly regulated protein synthesis rate. *Nature*. 2014; 509:49–54. DOI: 10.1038/nature13035 [PubMed: 24670665]



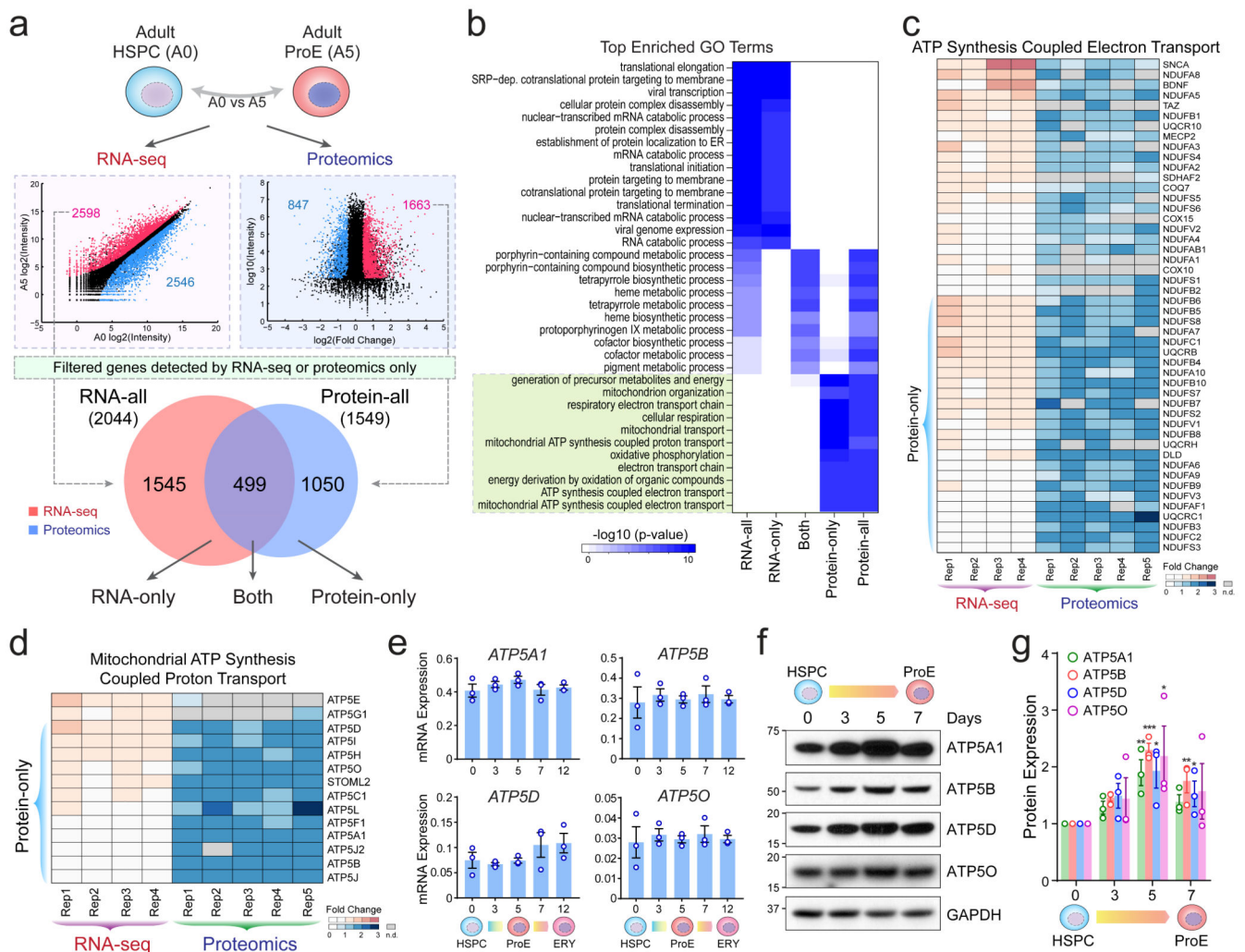
15. Subramanian A, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America*. 2005; 102:15545–15550. [PubMed: 16199517]
16. Heinrich AC, Pelanda R, Klingmuller U. A mouse model for visualization and conditional mutations in the erythroid lineage. *Blood*. 2004; 104:659–666. [PubMed: 15090451]
17. Kaelin WG Jr, McKnight SL. Influence of metabolism on epigenetics and disease. *Cell*. 2013; 153:56–69. DOI: 10.1016/j.cell.2013.03.004 [PubMed: 23540690]
18. Shimazu T, et al. Suppression of oxidative stress by beta-hydroxybutyrate, an endogenous histone deacetylase inhibitor. *Science (New York, NY)*. 2013; 339:211–214. DOI: 10.1126/science.1227166
19. Wellen KE, et al. ATP-citrate lyase links cellular metabolism to histone acetylation. *Science (New York, NY)*. 2009; 324:1076–1080. DOI: 10.1126/science.1164097
20. Larsson O, et al. Distinct perturbation of the transcriptome by the antidiabetic drug metformin. *Proceedings of the National Academy of Sciences of the United States of America*. 2012; 109:8977–8982. DOI: 10.1073/pnas.1201689109 [PubMed: 22611195]
21. Morita M, et al. mTORC1 controls mitochondrial activity and biogenesis through 4E-BP-dependent translational regulation. *Cell metabolism*. 2013; 18:698–711. DOI: 10.1016/j.cmet.2013.10.001 [PubMed: 24206664]
22. Zoncu R, Efeyan A, Sabatini DM. mTOR: from growth signal integration to cancer, diabetes and ageing. *Nature reviews Molecular cell biology*. 2011; 12:21–35. DOI: 10.1038/nrm3025 [PubMed: 21157483]
23. Knight ZA, Schmidt SF, Birsoy K, Tan K, Friedman JM. A critical role for mTORC1 in erythropoiesis and anemia. *eLife*. 2014; 3(e0)doi: 10.7554/eLife.01913
24. Magee JA, et al. Temporal changes in PTEN and mTORC2 regulation of hematopoietic stem cell self-renewal and leukemia suppression. *Cell stem cell*. 2012; 11:415–428. DOI: 10.1016/j.stem.2012.05.026 [PubMed: 22958933]
25. Roux PP, Topisirovic I. Regulation of mRNA translation by signaling pathways. *Cold Spring Harbor perspectives in biology*. 2012; 4
26. Feldman ME, et al. Active-site inhibitors of mTOR target rapamycin-resistant outputs of mTORC1 and mTORC2. *PLoS biology*. 2009; 7:e38. [PubMed: 19209957]
27. Thoreen CC, et al. An ATP-competitive mammalian target of rapamycin inhibitor reveals rapamycin-resistant functions of mTORC1. *The Journal of biological chemistry*. 2009; 284:8023–8032. DOI: 10.1074/jbc.M900301200 [PubMed: 19150980]
28. Guo F, et al. Mouse gene targeting reveals an essential role of mTOR in hematopoietic stem cell engraftment and hematopoiesis. *Haematologica*. 2013; 98:1353–1358. DOI: 10.3324/haematol.2012.080424 [PubMed: 23716557]
29. Lamb RF, et al. The TSC1 tumour suppressor hamartin regulates cell adhesion through ERM proteins and the GTPase Rho. *Nature cell biology*. 2000; 2:281–287. DOI: 10.1038/35010550 [PubMed: 10806479]
30. Yilmaz OH, et al. Pten dependence distinguishes haematopoietic stem cells from leukaemia-initiating cells. *Nature*. 2006; 441:475–482. DOI: 10.1038/nature04703 [PubMed: 16598206]
31. Hsieh AC, et al. The translational landscape of mTOR signalling steers cancer initiation and metastasis. *Nature*. 2012; 485:55–61. DOI: 10.1038/nature10912 [PubMed: 22367541]
32. Thoreen CC, et al. A unifying model for mTORC1-mediated regulation of mRNA translation. *Nature*. 2012; 485:109–113. DOI: 10.1038/nature11083 [PubMed: 22552098]
33. Passegue E, Wagner EF, Weissman IL. JunB deficiency leads to a myeloproliferative disorder arising from hematopoietic stem cells. *Cell*. 2004; 119:431–443. DOI: 10.1016/j.cell.2004.10.010 [PubMed: 15507213]
34. Suzuki A, et al. DBTSS as an integrative platform for transcriptome, epigenome and genome sequence variation data. *Nucleic acids research*. 2015; 43:D87–91. DOI: 10.1093/nar/gku1080 [PubMed: 25378318]
35. Elfakess R, et al. Unique translation initiation of mRNAs-containing TISU element. *Nucleic acids research*. 2011; 39:7598–7609. DOI: 10.1093/nar/gkr484 [PubMed: 21705780]

36. Gandin V, et al. nanoCAGE reveals 5' UTR features that define specific modes of translation of functionally related MTOR-sensitive mRNAs. *Genome research*. 2016; 26:636–648. DOI: 10.1101/gr.197566.115 [PubMed: 26984228]
37. Sinvani H, et al. Translational tolerance of mitochondrial genes to metabolic energy stress involves TISU and eIF1-eIF4GI cooperation in start codon selection. *Cell metabolism*. 2015; 21:479–492. DOI: 10.1016/j.cmet.2015.02.010 [PubMed: 25738462]
38. Zhang X, et al. FOXO3-mTOR metabolic cooperation in the regulation of erythroid cell maturation and homeostasis. *American journal of hematology*. 2014; 89:954–963. DOI: 10.1002/ajh.23786 [PubMed: 24966026]
39. Fleming A, Noda T, Yoshimori T, Rubinsztein DC. Chemical modulators of autophagy as biological probes and potential therapeutics. *Nature chemical biology*. 2011; 7:9–17. DOI: 10.1038/nchembio.500 [PubMed: 21164513]
40. Cunningham JT, et al. mTOR controls mitochondrial oxidative function through a YY1-PGC-1 $\alpha$  transcriptional complex. *Nature*. 2007; 450:736–740. DOI: 10.1038/nature06322 [PubMed: 18046414]
41. Duvel K, et al. Activation of a metabolic gene regulatory network downstream of mTOR complex 1. *Molecular cell*. 2010; 39:171–183. DOI: 10.1016/j.molcel.2010.06.022 [PubMed: 20670887]
42. Narla A, Ebert BL. Ribosomopathies: human disorders of ribosome dysfunction. *Blood*. 2010; 115:3196–3205. DOI: 10.1182/blood-2009-10-178129 [PubMed: 20194897]
43. Jaako P, et al. Dietary L-leucine improves the anemia in a mouse model for Diamond-Blackfan anemia. *Blood*. 2012; 120:2225–2228. DOI: 10.1182/blood-2012-05-431437 [PubMed: 22791294]
44. Payne EM, et al. L-Leucine improves the anemia and developmental defects associated with Diamond-Blackfan anemia and del(5q) MDS by activating the mTOR pathway. *Blood*. 2012; 120:2214–2224. DOI: 10.1182/blood-2011-10-382986 [PubMed: 22734070]
45. Chung J, et al. The mTORC1/4E-BP pathway coordinates hemoglobin production with L-leucine availability. *Science signaling*. 2015; 8:ra34. [PubMed: 25872869]
46. Skrtic M, et al. Inhibition of mitochondrial translation as a therapeutic strategy for human acute myeloid leukemia. *Cancer cell*. 2011; 20:674–688. DOI: 10.1016/j.ccr.2011.10.015 [PubMed: 22094260]
47. Van Handel B, et al. The first trimester human placenta is a site for terminal maturation of primitive erythroid cells. *Blood*. 2010; 116:3321–3330. DOI: 10.1182/blood-2010-04-279489 [PubMed: 20628147]
48. Gupta D, Shah HP, Malu K, Berliner N, Gaines P. Differentiation and characterization of myeloid cells. *Current protocols in immunology* / edited by John E Coligan ... [et al]. 2014; 104 Unit 22F. 25.
49. Kuhn R, Schwenk F, Aguet M, Rajewsky K. Inducible gene targeting in mice. *Science (New York, N.Y.)*. 1995; 269:1427–1429. [PubMed: 7660125]
50. Zhou F, et al. Genome-scale proteome quantification by DEEP SEQ mass spectrometry. *Nature communications*. 2013; 4:2171.
51. Zhang Y, et al. A robust error model for iTRAQ quantification reveals divergent signaling between oncogenic FLT3 mutants in acute myeloid leukemia. *Molecular & cellular proteomics : MCP*. 2010; 9:780–790. DOI: 10.1074/mcp.M900452-MCP200 [PubMed: 20019052]
52. Karp NA, et al. Addressing accuracy and precision issues in iTRAQ quantitation. *Molecular & cellular proteomics : MCP*. 2010; 9:1885–1897. DOI: 10.1074/mcp.M900628-MCP200 [PubMed: 20382981]
53. Dowling RJ, et al. mTORC1-mediated cell proliferation, but not cell growth, controlled by the 4E-BPs. *Science (New York, NY)*. 2010; 328:1172–1176. DOI: 10.1126/science.1187532
54. Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics (Oxford, England)*. 2009; 25:1105–1111. DOI: 10.1093/bioinformatics/btp120
55. Anders S, Huber W. Differential expression analysis for sequence count data. *Genome biology*. 2010; 11:R106. [PubMed: 20979621]
56. Trapnell C, et al. Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nature biotechnology*. 2013; 31:46–53. DOI: 10.1038/nbt.2450

57. Huang J, et al. Dynamic Control of Enhancer Repertoires Drives Lineage and Stage-Specific Transcription during Hematopoiesis. *Developmental Cell*. 2016; 36:9–23. DOI: 10.1016/j.devcel.2015.12.014. [PubMed: 26766440]
58. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome biology*. 2009; 10:R25. [PubMed: 19261174]
59. Zhang Y, et al. Model-based analysis of ChIP-Seq (MACS). *Genome biology*. 2008; 9:R137. [PubMed: 18798982]
60. Shao Z, Zhang Y, Yuan GC, Orkin SH, Waxman DJ. MAnorm: a robust model for quantitative comparison of ChIP-Seq data sets. *Genome biology*. 2012; 13:R16. [PubMed: 22424423]
61. Xu J, et al. Correction of sickle cell disease in adult mice by interference with fetal hemoglobin silencing. *Science (New York, NY)*. 2011; 334:993–996. DOI: 10.1126/science.1211053



**Figure 1. Comparative Proteomic and Transcriptomic Analyses of Human Erythropoiesis**  
**(a)** Fetal and adult-stage CD34<sup>+</sup> HSPCs were differentiated into ProEs *ex vivo*. Cells at matched stages of differentiation were collected for RNA-seq ( $n = 4$  independent experiments) and proteomic ( $n = 5$  independent experiments) analyses. **(b)** Schematic of iTRAQ-based quantitative proteomic analysis. **(c)** Venn diagram shows overlap between identified mRNA transcripts and proteins. **(d)** Overlay of RNA-seq and proteomic data demonstrates that the proteomic analysis achieved a comparable range of quantification as the RNA-seq-based transcriptomic profiling. RPKM, reads per kilobase of transcript per million reads.  $n = 4$  independent RNA-seq experiments and  $n = 5$  independent proteomic experiments. **(e)** Changes in expression of representative proteins between adult-stage HSPCs (A0) and ProEs (A5) are shown. Each color circle represents the measurement from an independent experiment. Boxes show the mean of the data and quartiles ( $n = 4$  independent RNA-seq experiments and  $n = 5$  independent proteomic experiments). Whiskers show the minimum and maximum of the data.



**Figure 2. Comparative Transcriptomic and Proteomic Analyses Revealed Post-Transcriptional Control of Mitochondrial Pathways**

(a) The correlation between RNA and protein-level expression changes in adult-stage HSPCs (A0) and ProEs (A5) is shown. Differentially expressed RNA transcripts and proteins (red, upregulated; blue, downregulated) were accessed for overlap. After filtering genes detected exclusively by RNA-seq or proteomics, 2,044 and 1,549 significant upregulated RNAs (RNA-all) and proteins (Protein-all) were identified, respectively. The differentially expressed genes were further divided into three groups based on significant changes at RNA (RNA-only), protein (Protein-only), and both levels, respectively. (b) GO enrichment analysis of protein and RNA expression changes. The green box highlights the top enriched GO terms for Protein-only genes. (c) Expression heatmap is shown for genes associated with ‘ATP synthesis coupled electron transport’. The Protein-only genes are shown on the bottom. *n.d.* not detected. (d) Expression heatmap is shown for genes associated with ‘Mitochondria ATP synthesis coupled proton transport’. (e) mRNA expression of representative Protein-only genes (*ATP5A1*, *ATP5B*, *ATP5D* and *ATP5O*) was measured by qRT-PCR in HSPCs and differentiating erythroid cells. Results are mean  $\pm$  s.e.m. ( $n = 3$  independent experiments). (f) Western blot analysis of representative Protein-

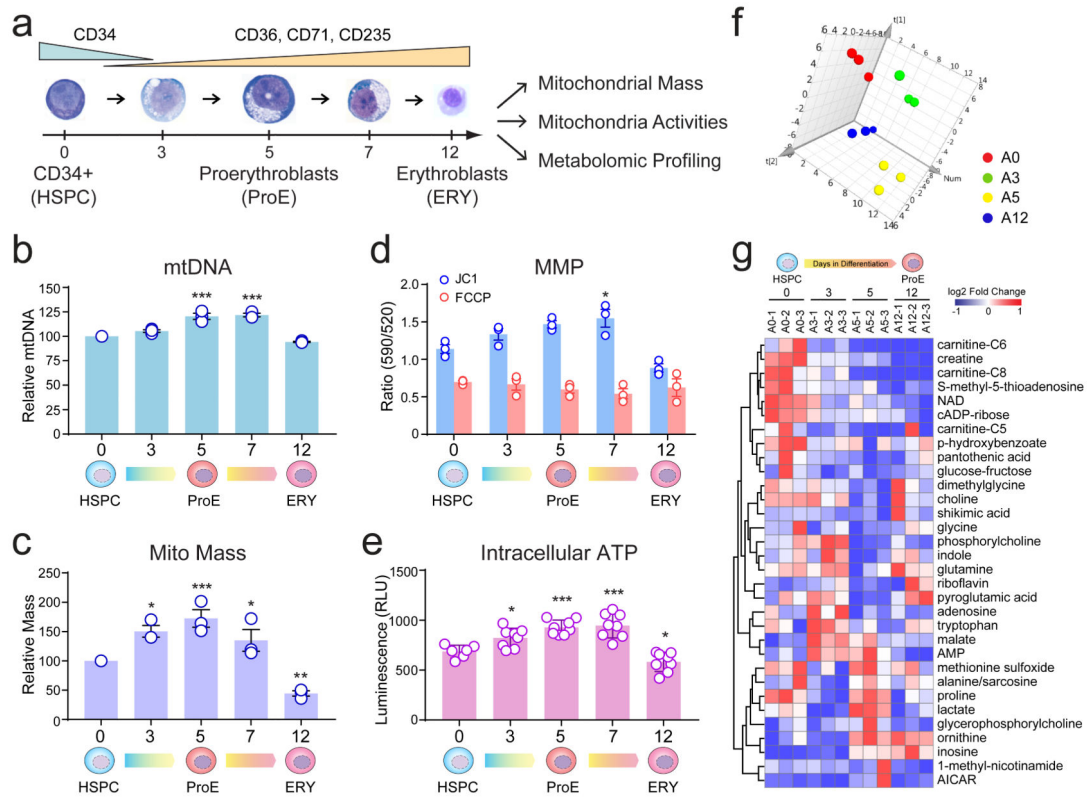
only genes. **(g)** Quantification of Western blot analysis. Results are mean  $\pm$  s.e.m. ( $n = 3$  independent experiments). Differences relative to HSPCs (day0) were assessed using a repeated-measures one-way ANOVA followed by Dunnett's test for multiple comparisons.  $*P < 0.05$ ,  $**P < 0.01$ ,  $***P < 0.001$  relative to HSPCs (day0) were considered significant. See Statistics Source Data in Supplementary Table 8. Unprocessed original scans of blots are shown in Supplementary Fig. 9.

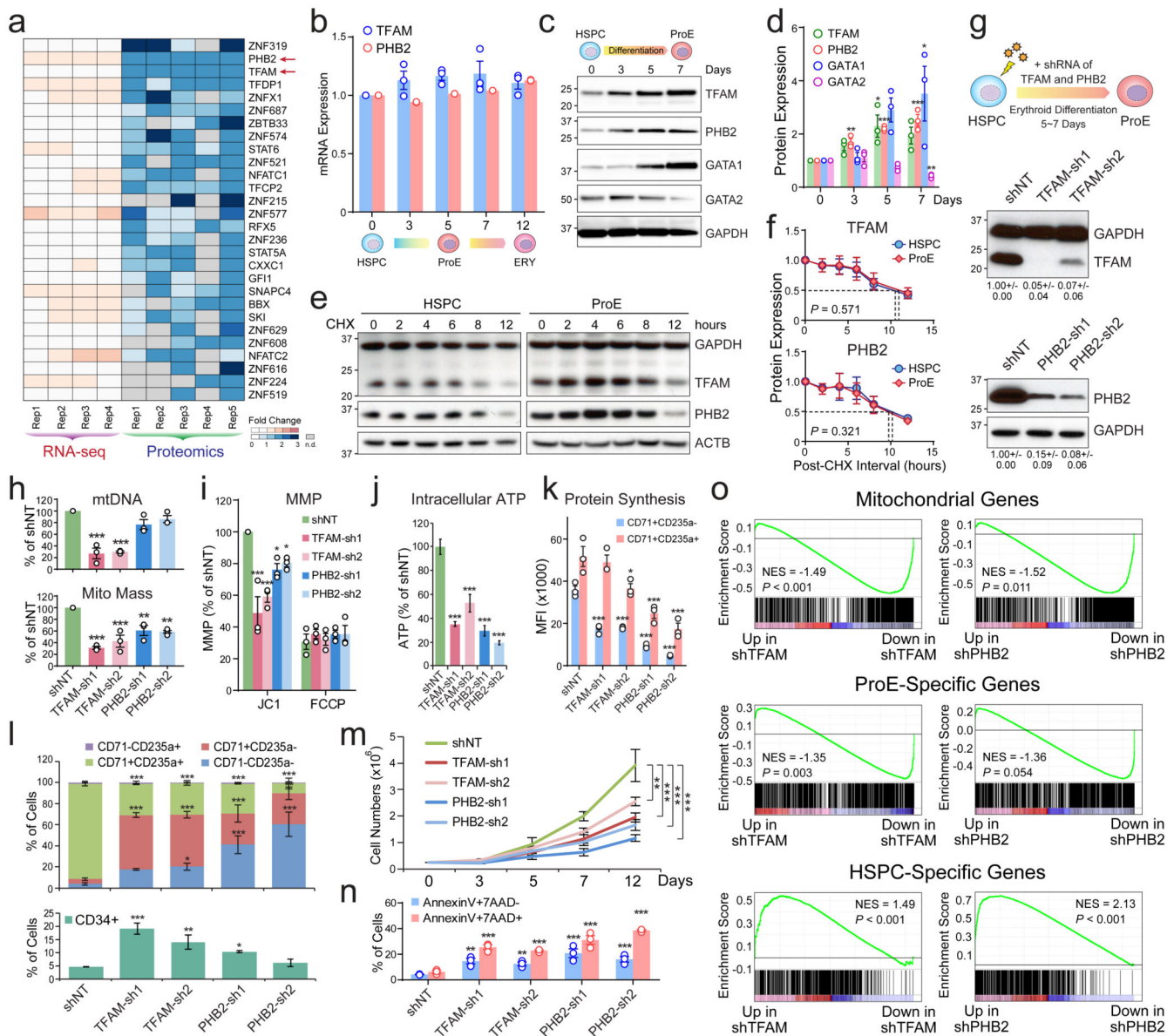
Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript



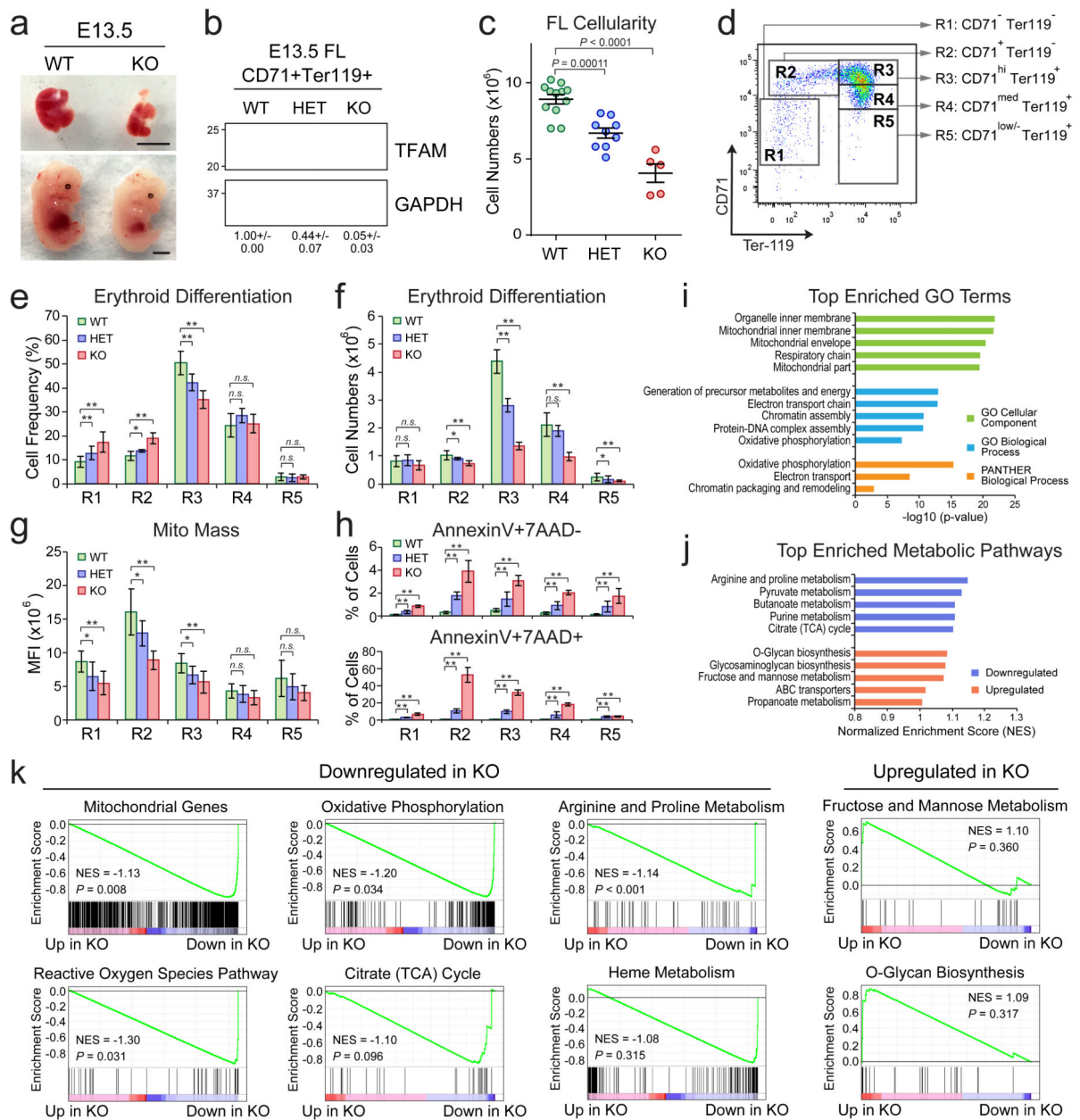


**Figure 4. TFAM and PHB2 Are Post-Transcriptionally Regulated and Indispensable for Mitochondria and Erythropoiesis**

(a) Heatmap is shown for transcription factors significantly upregulated at the protein but not mRNA levels. The genes are ranked based on the significance of fold changes in protein expression between HSPCs (A0) and ProEs (A5). *n.d.* not detected. (b) qRT-PCR analysis of *TFAM* and *PHB2* in HSPCs and differentiating erythroid cells. (c) Western blot of *TFAM*, *PHB2*, *GATA1* and *GATA2* in HSPCs and erythroid cells. (d) Quantification of Western blot analysis. (e) The protein half-lives of *TFAM* and *PHB2* were determined by CHX chase experiments. (f) Quantification of *TFAM* and *PHB2* half-lives in HSPCs and ProEs. The dashed lines indicate the calculated half-lives for both proteins. Results are mean  $\pm$  s.e.m. of  $n=3$  independent experiments (**b,d,f**). Differences relative to HSPCs (day0) were assessed using a repeated-measures one-way ANOVA followed by Dunnett's test for multiple



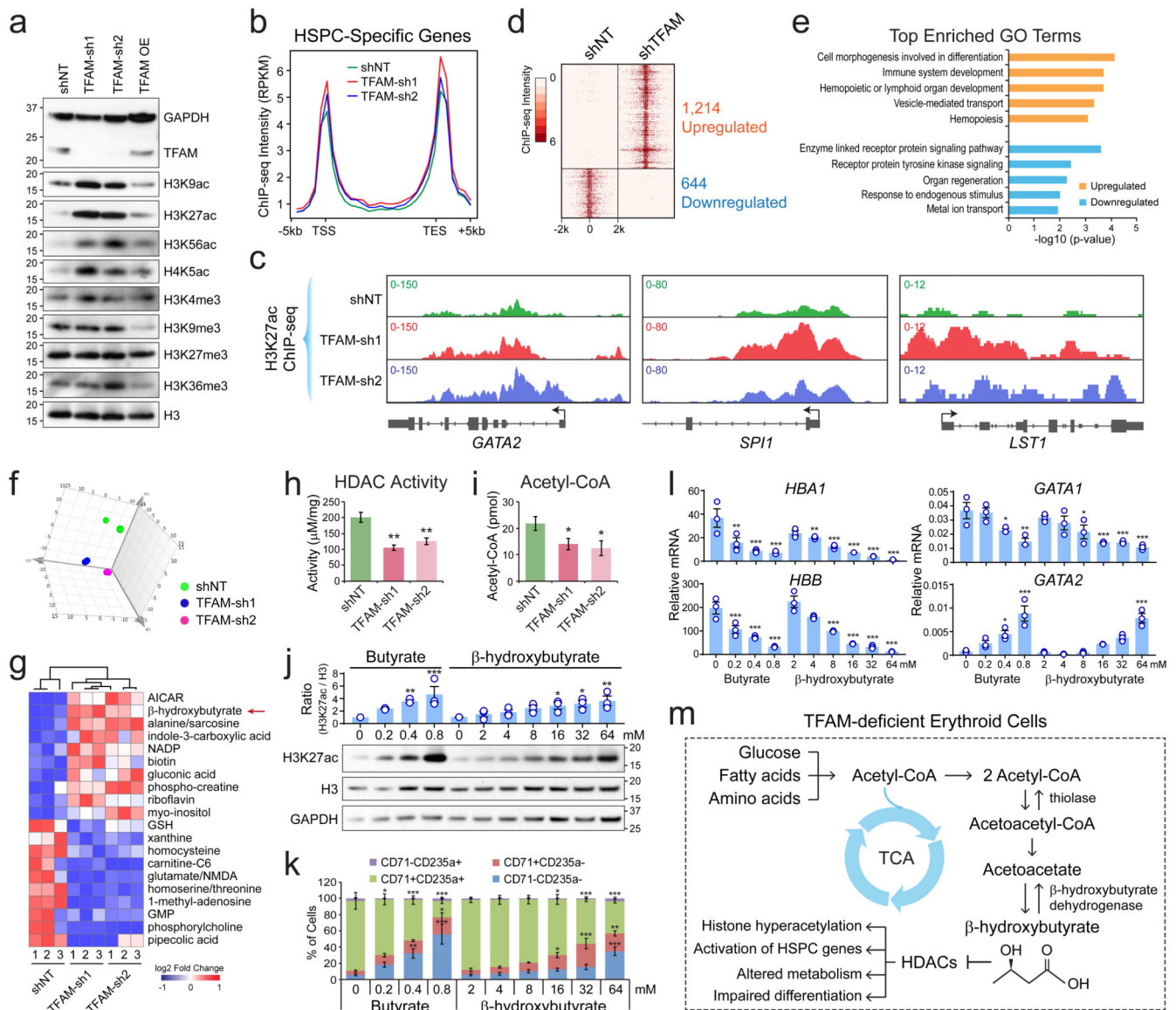
comparisons. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$  (**b,d,f**). (**g**) Schematic of lentiviral shRNA-mediated depletion of TFAM or PHB2 in HSPCs, followed by erythroid differentiation (*top*). Validation of TFAM or PHB2 depletion by Western blot (*bottom*). Cells transduced with non-targeting shRNA (shNT) were analyzed as controls. The quantified protein expression of TFAM or PHB2 from three independent experiments is shown as mean  $\pm$  s.e.m. on the *bottom*. (**h**) Depletion of TFAM significantly decreased mtDNA (*top*). TFAM and PHB2 depletion significantly decreased mitochondrial mass (*bottom*). (**i-n**) Depletion of TFAM or PHB2 significantly decreased MMP (**i**), intracellular ATP (**j**), protein synthesis in CD71<sup>+</sup>CD235a<sup>-</sup> and CD71<sup>+</sup>CD235a<sup>+</sup> erythroid progenitors (**k**), erythroid differentiation measured by the expression of CD34, CD71 and CD235a (**l**), proliferation (**m**), and increased apoptosis by AnnexinV and 7-AAD staining (**n**). Results are mean  $\pm$  s.e.m. of  $n=3$  independent experiments (**h,i,k-n**) or mean  $\pm$  s.d. of  $n=8$  independent measurements from three experiments (**j**). Differences relative to shNT were assessed using a repeated-measures one-way ANOVA followed by Dunnett's test for multiple comparisons; \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$  (**h-n**). (**o**) GSEA analysis of mitochondrial genes, ProE or HSPC-specific genes using the RNA-seq of shTFAM or shPHB2 relative to shNT, respectively. See Statistics Source Data in Supplementary Table 8.



**Figure 5. TFAM and PHB2 Are Required for Erythroid Development In Vivo**

(a) Representative pictures of E13.5 *Tfam* WT and KO embryos (*bottom*) and fetal livers (*top*). Scale bar, 0.5cm. (b) Western blot of TFAM expression in WT, HET and KO CD71<sup>+</sup>Ter119<sup>+</sup> E13.5 fetal liver erythroid cells. The quantified protein expression of TFAM is shown on the bottom as mean ± s.e.m of four independent samples for each genotype. (c) Total cell numbers in control (WT;  $n = 12$ ), *Tfam* HET ( $n = 9$ ) and KO ( $n = 5$ ) E13.5 fetal livers. (d) Erythroid differentiation was assessed by expression of CD71 and Ter119. (e, f) The frequency and number of R1 to R5 cells in *Tfam* WT ( $n = 16$ ), HET ( $n = 6$ ) and KO ( $n = 4$ ) E13.5 fetal livers. (g) Mitochondrial mass was determined by MitoTracker Green in *Tfam* WT ( $n = 16$ ), HET ( $n = 6$ ) and KO ( $n = 4$ ) E13.5 erythroid cells. (h) Cell apoptosis in

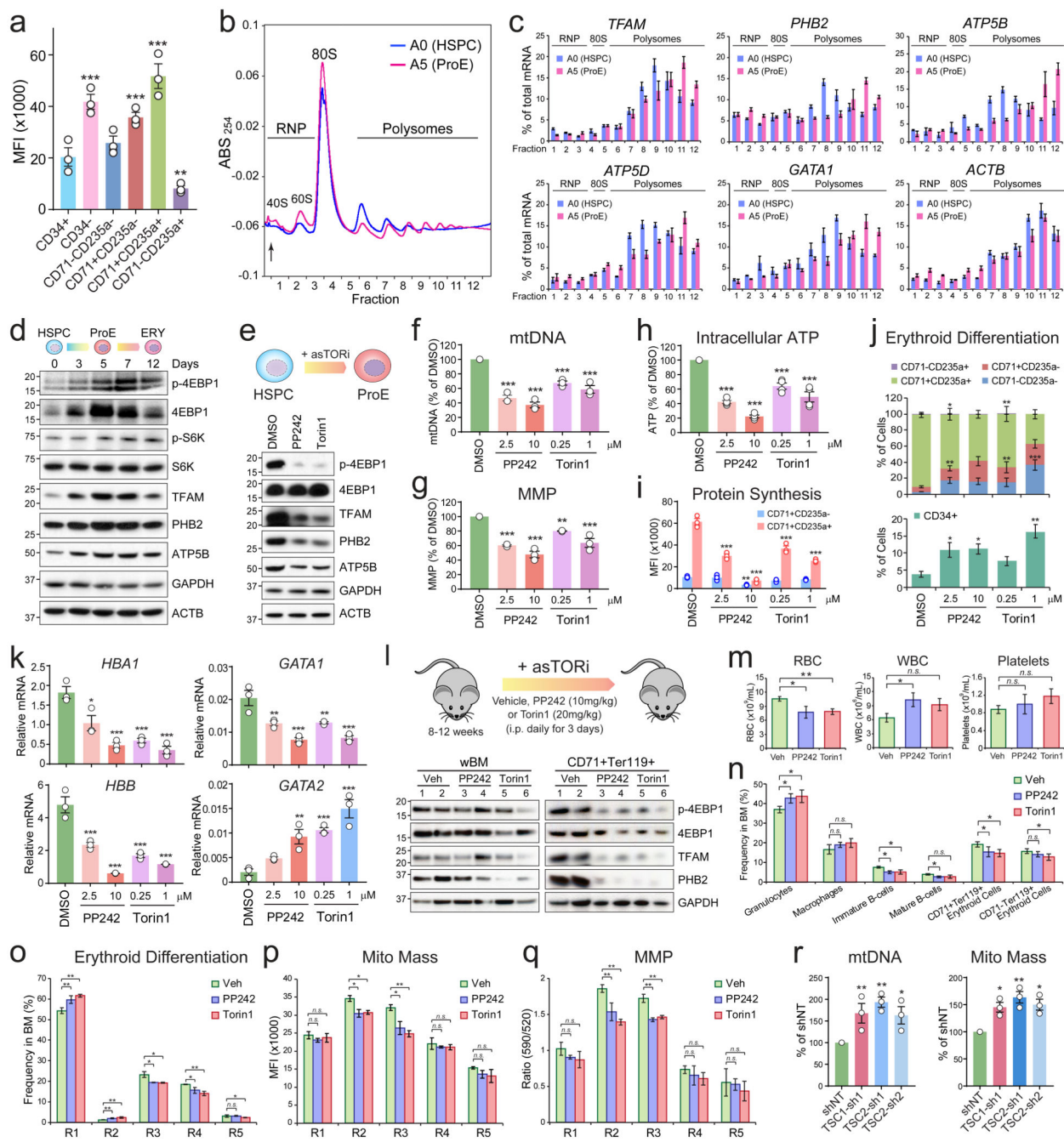
*Tfam* WT ( $n = 6$ ), HET ( $n = 5$ ) and KO ( $n = 4$ ) E13.5 erythroid cells. Results are mean  $\pm$  s.d. (**c**, **e-h**) and analyzed by one-way ANOVA (**c**) or a two-tailed *t*-test (**e-h**) to evaluate the differences between WT and KO (or HET). \* $P < 0.05$ , \*\* $P < 0.01$  were considered significant. (**i**) Top enriched gene ontology (GO) terms in significantly downregulated genes in *Tfam* KO erythroid cells. (**j**) Top enriched metabolic pathways in downregulated or upregulated genes in *Tfam* KO erythroid cells. (**k**) GSEA analysis of mitochondrial genes and other indicated signature genes downregulated or upregulated in *Tfam* KO erythroid cells. See Statistics Source Data in Supplementary Table 8. Unprocessed original scans of blots are shown in Supplementary Fig. 9.



**Figure 6. TFAM Depletion Leads to Altered Metabolism and Histone Hyperacetylation in Erythroid Cells**

(a) Western blot of histone modifications in human erythroid (K562) cells transduced with TFAM shRNAs, control (shNT), or overexpression (OE) constructs. Quantification of Western blot is shown in Supplementary Fig. 6a. (b) The distribution of H3K27ac intensity at HSPC-specific genes. The average ChIP-seq intensity (RPKM) is shown between 5kb upstream of transcription start site (TSS) and downstream of transcription end site (TES). (c) H3K27ac ChIP-seq density plots are shown for three representative HSPC-specific genes. (d) Heatmap is shown for differentially enriched H3K27ac ChIP-seq peaks. 1,214 upregulated and 644 downregulated H3K27ac peaks were identified with >4-fold changes in intensity between shTFAM and shNT. (e) Top enriched GO terms in genes associated with the upregulated or downregulated H3K27ac peaks. (f) Principle component analysis (PCA) of metabolomic profiles in K562 erythroid cells transduced with control or TFAM shRNAs.

*n* = 3 biological replicates per group. **(g)** Heatmap of the top 10 most upregulated or downregulated metabolites between control and TFAM-depleted erythroid cells. **(h)** HDAC activity in erythroid cells transduced with TFAM or control shRNAs. **(i)** Acetyl-CoA level in erythroid cells transduced with TFAM or control shRNAs. Results are mean ± s.e.m. of *n*=7 **(h)** or *n*=5 **(i)** independent measurements. Differences relative to shNT were analyzed by a two-tailed *t*-test. \**P* < 0.05, \*\**P* < 0.01. **(j)** Western blot of H3K27ac and H3 in human primary erythroid cells treated with butyrate or β-hydroxybutyrate for 12h. Quantification of Western blot is shown on the *top*. Results are mean ± s.e.m. of *n*=3 experiments. **(k)** Erythroid differentiation was assessed by CD71 and CD235a. Results are mean ± s.e.m. of *n*=3 experiments. **(l)** mRNA expression of hemoglobin genes (*HBA1* and *HBB*), *GATA1* and *GATA2*. Results are mean ± s.e.m. of *n*=3 experiments. Differences relative to control cells (0mM) were assessed using a repeated-measures one-way ANOVA with Dunnett's test for multiple comparisons; \**P* < 0.05, \*\**P* < 0.01, \*\*\**P* < 0.001 **(j-l)**. **(m)** Schematic of changes in metabolism and histone acetylation in TFAM-deficient erythroid cells. See Statistics Source Data in Supplementary Table 8. Unprocessed original scans of blots are shown in Supplementary Fig. 9.



**Figure 7. mTORC1 Regulates Mitochondrial Biogenesis in Erythroid Cells**

(a) Protein synthesis was assessed by HPG incorporation. Results are mean  $\pm$  s.e.m. of  $n=3$  experiments. Differences relative to CD34<sup>+</sup> were assessed using a repeated-measures one-way ANOVA with Dunnett's test for multiple comparisons. \*\* $P < 0.01$ , \*\*\* $P < 0.001$ . (b) Polysome profiling of HSPCs and ProEs. A representative profile is shown where the arrow represents the start of fraction. The positions of ribonucleoproteins (RNP), 80S ribosomes and polysomes are indicated. ABS<sub>254</sub>, absorbance at 254nm. (c) qRT-PCR analysis of polysome fractions in HSPCs and ProEs. Results are mean  $\pm$  s.d. of  $n=4$  measurements and

shown as the percentage (%) of total mRNA of all fractions combined. **(d)** Western blot analysis in HSPCs and erythroid cells. **(e)** HSPCs treated with control (DMSO), PP242 (2.5 $\mu$ M) or Torin1 (0.25 $\mu$ M) during erythroid differentiation for 5~7 days were analyzed by Western blot. **(f-k)** mTORC1 inhibition decreased mtDNA **(f)**, MMP **(g)**, ATP **(h)**, protein synthesis **(i)**, and impaired erythroid maturation **(j)** and gene expression **(k)**. Results are mean  $\pm$  s.e.m. of  $n=3$  experiments **(f,g,i-k)** or mean  $\pm$  s.d. of  $n=4$  experiments **(h)**. Differences relative to DMSO were assessed using a repeated-measures one-way ANOVA; \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$  **(f-k)**. **(l)** Schematic of mTOR inhibition. Western blot was performed using wBM or CD71<sup>+</sup>Ter119<sup>+</sup> erythroid cells. **(m)** Blood counts in mice treated with Veh ( $n=7$ ), PP242 ( $n=5$ ) or Torin1 ( $n=5$ ). **(n)** Frequencies of BM hematopoietic lineages in mice treated with Veh ( $n=6$ ), PP242 ( $n=5$ ) or Torin1 ( $n=5$ ). **(o)** Erythroid differentiation was assessed by CD71 and Ter119. **(p)** Mitochondrial mass was determined by MitoTracker Green. **(q)** MMP was determined by JC1 staining. Results are mean  $\pm$  s.e.m. of Veh ( $n=4$ ), PP242 ( $n=3$ ) or Torin1 ( $n=3$ ), and analyzed by a two-tailed  $t$ -test; \* $P < 0.05$ , \*\* $P < 0.01$  **(m-q)**. **(r)** Depletion of TSC1 or TSC2 in erythroid cells increased mtDNA and mitochondrial mass. Results are mean  $\pm$  s.e.m. of  $n=3$  experiments. Differences relative to shNT were assessed using a repeated-measures one-way ANOVA. \* $P < 0.05$ , \*\* $P < 0.01$ . See Statistics Source Data in Supplementary Table 8. Unprocessed original scans of blots are shown in Supplementary Fig. 9.



**Figure 8. mTORC1 Selectively Regulates Protein Translation of Mitochondria-Associated mRNAs in Erythroid Cells**

(a) Schematic of *ex vivo* differentiation of human primary HSPCs into ProEs or granulocytes. (b) A representative profile is shown for polysome profiling of human erythroid cells treated with control (DMSO) or PP242 (2.5 $\mu$ M) for 12h. (c) qRT-PCR analysis of indicated transcripts in cells treated with DMSO or PP242. Results are mean  $\pm$  s.d. of  $n=4$  independent measurements and shown as the percentage (%) of total mRNA of all fractions combined. (d) The number of downregulated proteins (fold change 1.3,  $P$



0.01), Protein-only, or mitochondrial proteins were determined by iTRAQ-based proteomics in ProEs or granulocytes treated with PP242 (2.5 $\mu$ M) or control (DMSO) for 12h. *P*-values were calculated by hypergeometric distribution using All ( $n=935$  and 529 in ProEs and granulocytes, respectively), Protein-only ( $n=205$  and 50) and Mitochondria ( $n=245$  and 33) genes. *n.s.* not significant. (e) The changes in the levels of indicated proteins were determined by iTRAQ-based proteomics in ProEs or granulocytes treated with PP242. Results are shown for one representative proteomic experiment. (f) Top enriched GO terms of proteins downregulated in PP242-treated ProEs or granulocytes. (g) The most frequent TSS of the indicated transcripts were annotated using dbTSS, RefSeq and Ensembl databases. The sequences of TOP, TOP-like or TISU motifs were shown as the *pink* color. (h) The frequency of TOP-like motifs in Protein-only, RNA-only or all genes. *P*-values were calculated by hypergeometric distribution using All ( $n=4,228$  out of 6,624 total detected proteins), Protein-only ( $n=418$  out of 594 total detected proteins) and RNA-only ( $n=363$  out of 612 total detected proteins) genes. (i) Mutation of TOP-like motifs at the 5' UTRs of *TFAM*, *PHB2* and *ATP5D* impaired GFP reporter expression in human primary erythroid cells but not granulocytes. Cells treated control (DMSO) or PP242 (2.5 $\mu$ M) for 12h were analyzed. Results are mean  $\pm$  s.e.m. of  $n=3$  independent experiments and analyzed by a two-tailed *t*-test. \* $P < 0.05$ , \*\* $P < 0.01$ , *n.s.* not significant. See Statistics Source Data in Supplementary Table 8. (j) Model of mTORC1-mediated post-transcriptional control of mitochondrial biogenesis during erythropoiesis.