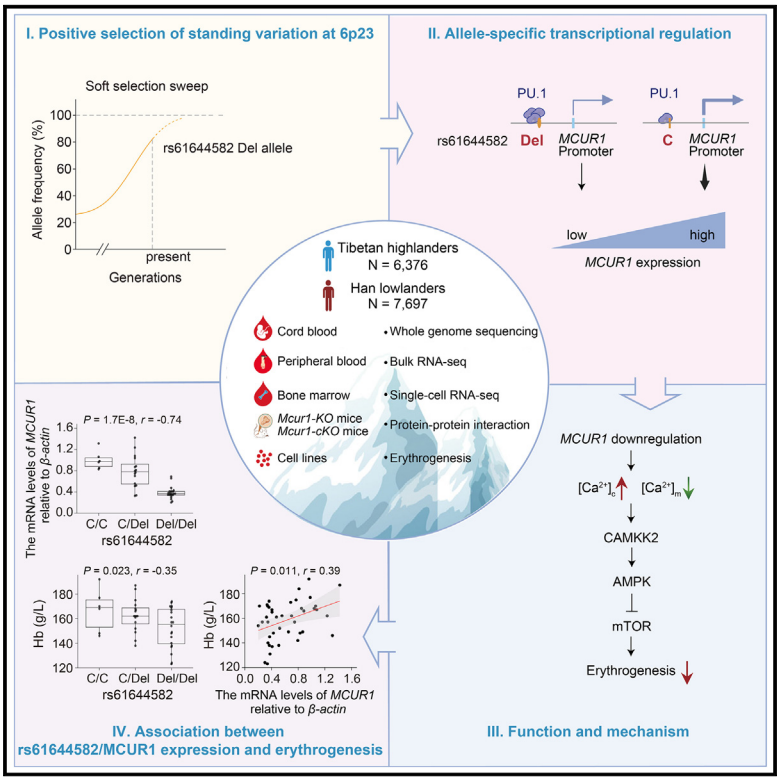
Cell Genomics

Article

A highland-adaptation variant near *MCUR1* reduces its transcription and attenuates erythrogenesis in Tibetans

Graphical abstract



Highlights

- Whole-genome sequencing of Tibetans and Hans reveals multiple regions under selection
- *MCUR1* is the functional gene of the selection signal at 6p23 and facilitates erythropoiesis
- MCUR1 facilitates erythropoiesis through the Ca^{2+}-CAMKK2-AMPK-mTOR pathway
- rs61644582 at 6p23 confers an allele-specific regulation of *MCUR1* expression by PU.1

Authors

Jie Ping, Xinyi Liu, Yiming Lu, ..., Fuchu He, Hongxing Zhang, Gangqiao Zhou

Correspondence

zhanghx08@126.com (H.Z.), zhougq114@126.com (G.Z.)

In brief

Ping et al. performed a whole-genome scan and identified a collection of new regions subject to positive selection in Tibetan genomes. The authors further revealed that MCUR1 is a functional target of the selection signal at chromosome 6p23, and its downregulation weakens the erythropoiesis of Tibetans. These findings highlight that MCUR1-mediated mitochondrial Ca^{2+} homeostasis is a novel regulator of erythropoiesis and deepen our understanding of the genetic mechanisms underlying high-altitude adaptation.

Ping et al., 2025, Cell Genomics *5*, 100782
March 12, 2025
https://doi.org/10.1016/j.xgen.2025.100782

CellPress

Cell Genomics

CellPress
OPEN ACCESS

Article

A highland-adaptation variant near *MCUR1* reduces its transcription and attenuates erythrogenesis in Tibetans

Jie Ping,[1,20] Xinyi Liu,[1,20] Yiming Lu,[1,20] Cheng Quan,[1,20] Pengcheng Fan,[2,3,20] Hao Lu,[1] Qi Li,[1] Cuiling Wang,[1] Zheng Zhang,[1] Mengyu Liu,[1] Shunqi Chen,[1] Lingle Chang,[4] Yuqing Jiang,[5] Qilin Huang,[5] Jie Liu,[6,7] Tana Wuren,[6] Huifang Liu,[6] Ying Hao,[8] Longli Kang,[9,10] Guanjun Liu,[11,12] Hui Lu,[1] Xiaojun Wei,[1] Yuting Wang,[1] Yuanfeng Li,[1] Hao Guo,[13] Yongquan Cui,[13] Haoxiang Zhang,[14] Yang Zhang,[15] Yujia Zhai,[16] Yaoxi He,[17] Wangshan Zheng,[17] Xuebin Qi,[18,19] Ouzhuluobu,[19] Huiping Ma,[2] Linpeng Yang,[2] Xin Wang,[2] Wanjun Jin,[2] Ying Cui,[12] Rili Ge,[6] Shizheng Wu,[6,7] Yuan Wei,[16] Bing Su,[17] Fuchu He,[3] Hongxing Zhang,[3,*] and Gangqiao Zhou[1,4,5,21,*]

[1]State Key Laboratory of Medical Proteomics, Beijing Proteome Research Center, National Center for Protein Sciences (Beijing), Beijing Institute of Radiation Medicine, Beijing 100850, P.R. China
[2]Pharmacy Department, General Hospital of Lanzhou, Lanzhou City 730050, P.R. China
[3]State Key Laboratory of Medical Proteomics, Beijing Proteome Research Center, National Center for Protein Sciences (Beijing), Beijing Institute of Lifeomics, Beijing 102206, P.R. China
[4]Medical College of Guizhou University, Guiyang City 550025, P.R. China
[5]Collaborative Innovation Center for Personalized Cancer Medicine, Center for Global Health, School of Public Health, Nanjing Medical University, Nanjing City 211166, P.R. China
[6]Research Center for High-Altitude Medicine, Qinghai University Medical School, Xining City 810001, P.R. China
[7]Qinghai Provincial People's Hospital, Xining City 810001, P.R. China
[8]College of Life Sciences, Beijing University of Chinese Medicine, Beijing 102488, P.R. China
[9]Key Laboratory for Molecular Genetic Mechanisms and Intervention Research on High-Altitude Disease of Tibet Autonomous Region, School of Medicine, Xizang Minzu University, Xianyang City 712082, P.R. China
[10]Key Laboratory of High-Altitude Environment and Genes Related to Diseases of Tibet Autonomous Region, School of Medicine, Xizang Minzu University, Xianyang City 712082, P.R. China
[11]Henan Provincial People's Hospital, Zhengzhou City 450000, P.R. China
[12]Affiliated Cancer Hospital of Guangxi Medical University, Nanning City 530021, P.R. China
[13]No. 945 Hospital of Joint Logistic Support Force of Chinese PLA, Ya'an City 625000, P.R. China
[14]No. 954 Hospital of Joint Logistic Support Force of Chinese PLA, Shannan City 856000, P.R. China
[15]Medical Center for Human Reproduction, Beijing Chao-Yang Hospital, Capital Medical University, Beijing 100020, P.R. China
[16]Center for Reproductive Medicine, Department of Obstetrics and Gynecology, Peking University Third Hospital, Beijing 100191, P.R. China
[17]State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming City 650223, P.R. China
[18]State Key Laboratory of Primate Biomedical Research, Institute of Primate Translational Medicine, Kunming University of Science and Technology, Kunming 650000, China
[19]Fukang Obstetrics, Gynecology and Children Branch Hospital, Tibetan Fukang Hospital, Lhasa 850000, China
[20]These authors contributed equally
[21]Lead contact
*Correspondence: zhanghx08@126.com (H.Z.), zhougq114@126.com (G.Z.)
https://doi.org/10.1016/j.xgen.2025.100782

SUMMARY

To identify genomic regions subject to positive selection that might contain genes involved in high-altitude adaptation (HAA), we performed a genome-wide scan by whole-genome sequencing of Tibetan highlanders and Han lowlanders. We revealed a collection of candidate genes located in 30 genomic loci under positive selection. Among them, *MCUR1* at 6p23 was a novel pronounced candidate. By single-cell RNA sequencing and comprehensive functional studies, we demonstrated that *MCUR1* depletion leads to impairment of erythropoiesis under hypoxia and normoxia. Mechanistically, *MCUR1* knockdown reduced mitochondrial Ca^{2+} uptake and then concomitantly increased cytosolic Ca^{2+} levels, which thereby reduced erythropoiesis via the CAMKK2-AMPK-mTOR axis. Further, we revealed rs61644582 at 6p23 as an expression quantitative trait locus for *MCUR1* and a functional variant that confers an allele-specific transcriptional regulation of *MCUR1*. Overall, MCUR1-mediated mitochondrial Ca^{2+} homeostasis is highlighted as a novel regulator of erythropoiesis, deepening our understanding of the genetic mechanism of HAA.

INTRODUCTION

The Qinghai-Tibetan plateau, which has an average elevation of over 4,000 m above sea level, is inhospitable to human settlement due to multiple extreme conditions, including low oxygen pressure and cold climate.[1] Nevertheless, the ethnic Tibetans have successfully settled in the plateau for thousands of years and developed a distinctive set of heritable and adaptive physiological traits. Numerous studies have attributed these physiological differences to the genetic adaptation of Tibetans to high altitude.[2] It has been proposed that high-altitude adaptation (HAA) is one of the strongest instances of natural selection acting on humans, which can be detected from a scan of genetic diversity across genomes.[3]

Indeed, many studies have reported several loci subject to positive natural selection, including the *EGLN1* and *EPAS1* loci, both of which are associated with the hypoxia-inducible factor (HIF) pathway that detects and reacts to oxygen supply changes.[3–8] Despite these recent advances, genetic HAA is likely to be a complex process, with a large number of genes involved in the response to not only hypoxia but also other extreme environmental conditions. The loci subject to natural selection might not have been detected in previous studies, possibly due to insufficient statistical power caused by small sample sizes.[3–5] In addition, most of the previous studies have either focused on *a priori* candidate genes or assayed only small portions of the genomes, e.g., through single-nucleotide polymorphism (SNP) genotyping arrays or exome sequencing.[3–5] Therefore, there is likely much yet to be deciphered for genetic HAA.

Here, we performed a genome-wide scan by sequencing the whole genomes of 48 Tibetans residing above 2,500 m and 50 Han individuals living below 100 m. We identified a collection of candidate regions subject to positive selection. Among them, we demonstrated that depletion or downregulation of the mitochondrial calcium uniporter regulator 1 (*MCUR1*) gene at a selection signal on the 6p23 locus reduces erythropoiesis through the Ca^{2+}-CAMKK2-AMPK-mTOR pathway under hypoxia. Further, we demonstrated that rs61644582 is the functional SNP at 6p23, and it confers an allele-specific transcriptional regulation on *MCUR1*. These findings advance our understanding of the genetic mechanisms of HAA and highlight the MCUR1-mediated mitochondrial Ca^{2+} homeostasis as a novel regulator of erythropoiesis.

RESULTS

Whole-genome sequencing of 48 Tibetan highlanders and 50 Han lowlanders

We performed whole-genome sequencing (WGS) for 48 unrelated Tibetan highlanders living above 2,500 m (designated as TIB_discovery population) and 50 unrelated Han lowlanders living below 100 m (designated as HAN_discovery population) in the discovery stage (Table S1). The mean depth of WGS was ~40×, with more than 98% of the genomic regions covered above 10× (Table S2). Using the GATK pipeline[9] and after quality controls, a total of ~10 million SNPs were called and used for subsequent analyses (Tables S3 and S4; supplemental information).

Identification of genomic signatures of positive selection relevant to HAA

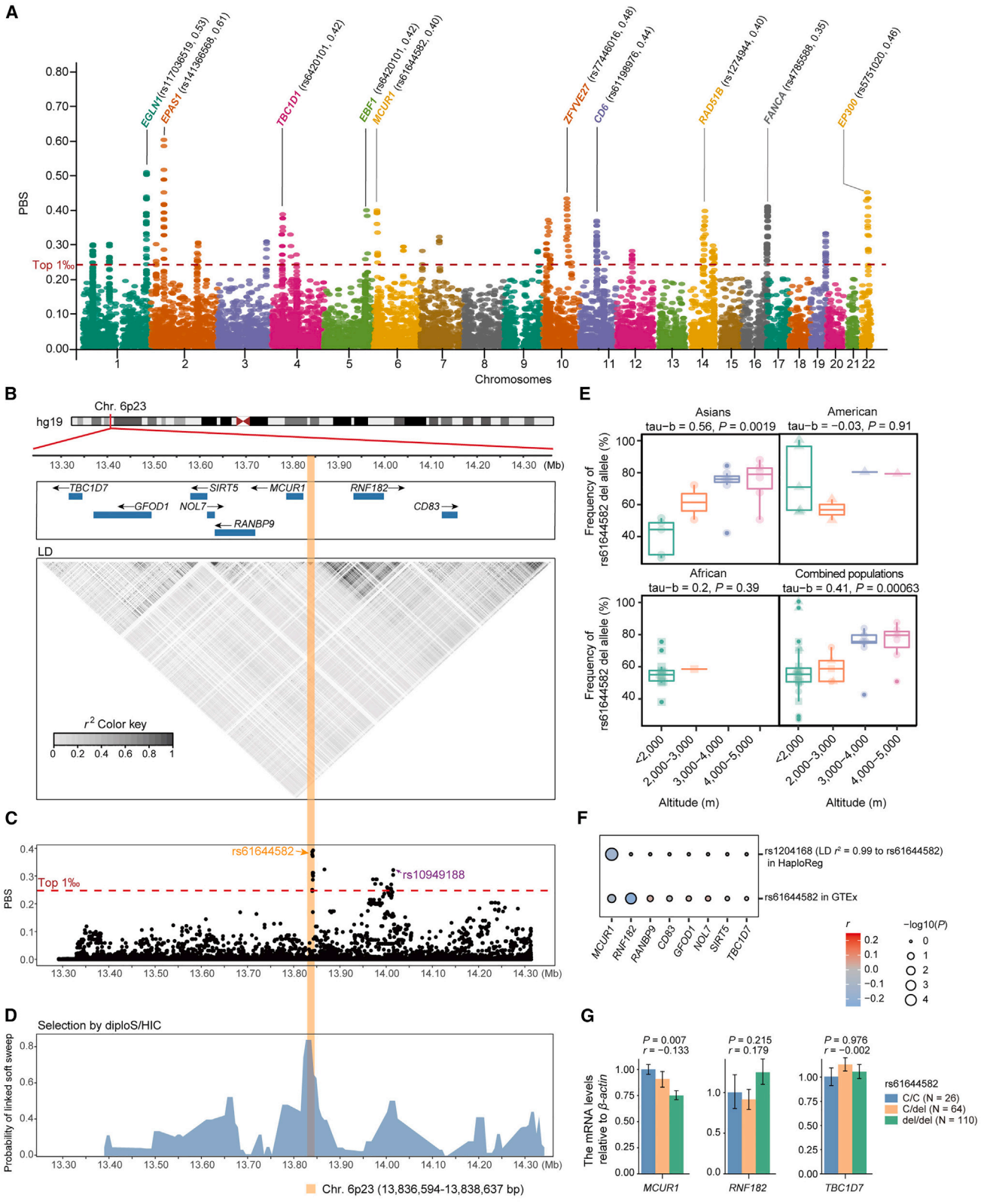
Genetic divergence analyses revealed that the TIB_discovery and HAN_discovery populations not only are genetically closely related but also show recent divergence (Figure S1; Table S5; STAR Methods; supplemental information), which is consistent with previous findings.[4,10] We then sought to identify the genomic signatures of positive selection relevant to HAA by population branch statistics (PBS).[3] The top 1‰ genome-wide value for PBS analysis was set as the significant threshold (with PBS > 0.24 and empirical $p_{PBS} < 0.001$) (Figure S2A). These top 1‰ SNPs were further clustered into 30 independent genomic loci, which contain 43 genes (Table S6; STAR Methods). Among these, 17, including 2 well-known HAA-associated genes, *EPAS1* (PBS = 0.61, ranked first) and *EGLN1* (PBS = 0.53, ranked second), have been reported previously to be relevant to HAA in Tibetans, and the other 26 are new candidates (Tables S7, S8, S9, and S10; Figures S2B–S2G).

These 43 genes are significantly enriched in 32 pathways (all $p_{adj} < 0.05$; Table S10A; supplemental information), among which 12 are replicated in at least 2 of the 18 publicly available datasets (Figure S2E; Table S10B). Notably, the calcium signaling pathway was the most significant candidate ($p = 4.48 \times 10^{-10}$; Figures S2F and S2G), which is consistent with a recent study on HAA in Tibetan chickens.[11] Further, among those genes involved in the calcium signaling pathway, *MCUR1*, which ranks among the top in the newly discovered genes in positive selection scanning, is tagged by the selection signal at chromosome 6p23 (index rs61644582, PBS = 0.40; Figure 1A; Tables S6 and S7). Thus, we investigated this novel 6p23 signal in detail in subsequent studies.

A new positive selection signal relevant to HAA was identified at the 6p23 locus

The index rs61644582 lies in an ~2 kb linkage disequilibrium (LD) block at the 6p23 locus (Figure 1B). This LD block exhibited evidence of highly polarized measure using PBS (all PBS > 0.32 and all $p_{PBS} < 0.001$; Figures 1C and S3A). Notably, several LD-based methods sensitive to the initial selective allele frequency,[12] such as integrated haplotype score (iHS) and cross population extended haplotype homozygosity (XP-EHH) tests, showed no strong evidence of selection at this region (Table S11), suggesting a pattern of recent adaptation, with multiple haplotypes carrying the standing variant prior to the onset of selection.[13,14] Therefore, we further investigated the mutational origin of the beneficial allele of index rs61644582 (i.e., deletion [del] allele). We observed that the rs61644582 del allele has a relatively high frequency across a wide geographic distribution (26.3%–70.6%; Table S12). Meanwhile, we estimated the age of this derived rs61644582 del allele, using a genealogical approach,[15] to be approximately 1,002,650 (95% confidence interval [CI], 911,375–1,099,300) years old, which is much older than the onset age of HAA, which occurred approximately 8,500 years ago[15] (Figure S3B). Furthermore, we trained a deep convolutional neural network implemented in diploS/HIC and then classified genomic regions based on 12 major summary statistics of selective sweeps.[16] As expected, we observed that rs61644582 was closely linked to a significant soft selection sweep in Tibetans (Figures 1D and S3C). Taken together, these findings

CellPress
OPEN ACCESS



(legend on next page)

provide evidence in principle that rs61644582 might initially be a standing variant in Tibetans, and multiple adaptive haplotypes (carrying the beneficial del allele of rs61644582) simultaneously increased to high frequencies under positive selection.

Next, we sought to replicate the candidate HAA signal at 6p23 in independent populations. We genotyped rs61644582 in 942 individuals of the second Tibetan-Han population (designated as replication population 1) and also investigated genotype frequency, fixation index (F_{ST}), and PBS of rs61644582 from a total of 13,851 individuals in publicly available Tibetan-Han populations.[5,6,17,18] The rs61644582 frequencies in Tibetans in the replication stage ranged from 62% to 87%, closely resembling the frequency of ~81% observed in the discovery stage (Figure S3D). Furthermore, these frequencies remained consistently higher than those observed in the Han populations (26%–51%) (Figure S3D). Consistently, the markedly high F_{ST} and PBS of rs61644582 were observed in these Tibetan-Han populations (Figure S3D), therefore confirming this newly identified candidate signal at 6p23. In addition, rs61644582 frequency was significantly positively correlated with altitude in Asians (Kendall's Tau-b = 0.56, p = 0.0019; Figure 1E; Table S12). Although this may be due to its high frequency among Tibetans, it still suggests different selection strengths at rs61644582 at different altitudes and once again supports the contribution of the 6p23 signal to HAA in Tibetans.

In addition to rs61644582, the other 13 SNPs in the ~2 kb LD block at 6p23 also showed evidence of positive selection in this study (all PBS > 0.32 and all p_{PBS} < 0.001) and in two independent studies from Yang et al. (all PBS > 0.09 and all p_{PBS} < 0.05)[6] and Zheng et al. (all PBS > 0.10 and all p_{PBS} < 0.05)[18] (Figure 1C; Table S11). However, these SNPs were in moderate to strong LD with rs61644582 (all r^2 > 0.60; Figure 1B; Table S11), suggesting there exists a single selection signal in this LD block. Notably, within the 1 Mb region surrounding this LD block, there exists another moderately strong positive selection signal (index rs10949188, PBS = 0.32; Figure 1C; Table S6), which is located ~104 kb downstream of the ~2 kb LD block and has been reported previously.[6] However, this signal showed a relatively low possibility of soft sweep (Figure 1D) and showed an extremely low LD to the 6p23 signal (all r^2 < 0.02 with rs61644582; Figure 1B), indicating their independence. Collectively, these results suggest that there exists a novel positive selection signal in the ~2 kb LD block at the 6p23 locus.

MCUR1 is the potential target gene at the 6p23 locus

We next sought to identify the causative gene(s) attributed to the 6p23 signal. A total of eight genes are located within the 1 Mb region surrounding this signal (Figure 1B). Expression quantitative trait locus (eQTL) analyses showed that the index rs61644582 del allele is significantly associated with decreased transcription levels of *MCUR1*, but not the other seven genes, in peripheral blood of the BloodHT12 dataset (n = 1,240, p = 2.2 × 10^{-5}) from HaploReg (v.4.1) (Figure 1F). This eQTL result was replicated in another set of peripheral blood from 670 individuals of the GTEx dataset (p = 0.0042; Figures 1F and S3E) and in peripheral blood mononuclear cells (PBMCs) of the third independent population consisting of 200 Tibetans in this study (designated as TIB_replication 2; p = 0.007; Figure 1G). Taken together, these results suggest that *MCUR1* might be the causative gene targeted by the 6p23 signal.

MCUR1 knockout reduces erythropoiesis in mice under hypoxia

MCUR1 is a critical component of the mitochondrial calcium uniporter complex (MCUC), which is required for mitochondrial Ca^{2+} uptake and maintenance of cellular bioenergetics.[19] However, MCUR1 has not been reported previously to be relevant to HAA. We observed that *MCUR1* is obviously highly expressed in human blood cells according to the BodyMap project 2.0 (Figure S4A). Further, *MCUR1* was highly expressed in erythroid cells compared to the other hematopoietic lineages according to the DMAP dataset[20] (Figure S4B) and two other publicly available single-cell RNA sequencing (scRNA-seq) datasets[21,22] (Figures S4C and S4D), which all came from human bone marrow (BM). More specifically, early erythroid progenitors (EEPs) and committed erythroid progenitors (CEPs) expressed the highest levels of *MCUR1* mRNAs among all erythroid-lineage cells of mouse BM[23] (Figure S4E). Interestingly, in a search of the ARCHS4 libraries, *MCUR1* was significantly associated with the terms "abnormality of cells of the erythroid lineage" and "abnormal number of erythroid precursors."[24] From these results, we thus hypothesize that MCUR1 might play critical roles in erythropoiesis, a well-known phenotype closely relevant to HAA.[8]

We then assessed the effects of MCUR1 on erythropoiesis *in vivo* by generating a whole-body *Mcur1*-knockout (KO) mouse model using the CRISPR-Cas9 system (Figures S4F–S4H). Because the *Mcur1*$^{-/-}$ mice underwent embryonic mortality, the *Mcur1*$^{+/-}$ and *Mcur1*$^{+/+}$ (designated as *Mcur1*-KO and *Mcur1*-WT, respectively) mice were used to assess the erythropoiesis of lineage-negative hematopoietic stem and progenitor cells (HSPCs). Under either normoxic (21% O_2) or hypoxic (11% O_2, equivalent to the oxygen levels at an altitude of ~5,000 m) conditions for 4 weeks, the body weight and several hematological indexes in peripheral blood were similar between *Mcur1*-WT and *Mcur1*-KO mice (Figure S4I). However,

Figure 1. *MCUR1* is the potential functional target at the 6p23 locus relevant to HAA

(A) Population branch statistic (PBS) values of SNPs across the entire human genome. Top 10 loci with their nearest genes are marked by their corresponding colors, with their leading SNPs and PBS values in parentheses.

(B) Linkage disequilibrium (LD) structure of the SNPs at the 6p23 locus in the TIB_discovery population.

(C) Regional plot of PBS values of SNPs at the 6p23 locus in the discovery stage.

(D) Regional plot of probabilities of linked soft selection sweep by diploS/HIC.

(E) Association between rs61644582 deletion allele frequencies and the altitude in Asians, Americans, and Africans and their combined populations.

(F) Association between rs61644582 genotypes and mRNA levels of eight genes surrounding the 6p23 locus in peripheral blood datasets.

(G) Association between rs61644582 genotypes and expression of indicated genes in the peripheral blood of 200 Tibetans (TIB_replication 2).

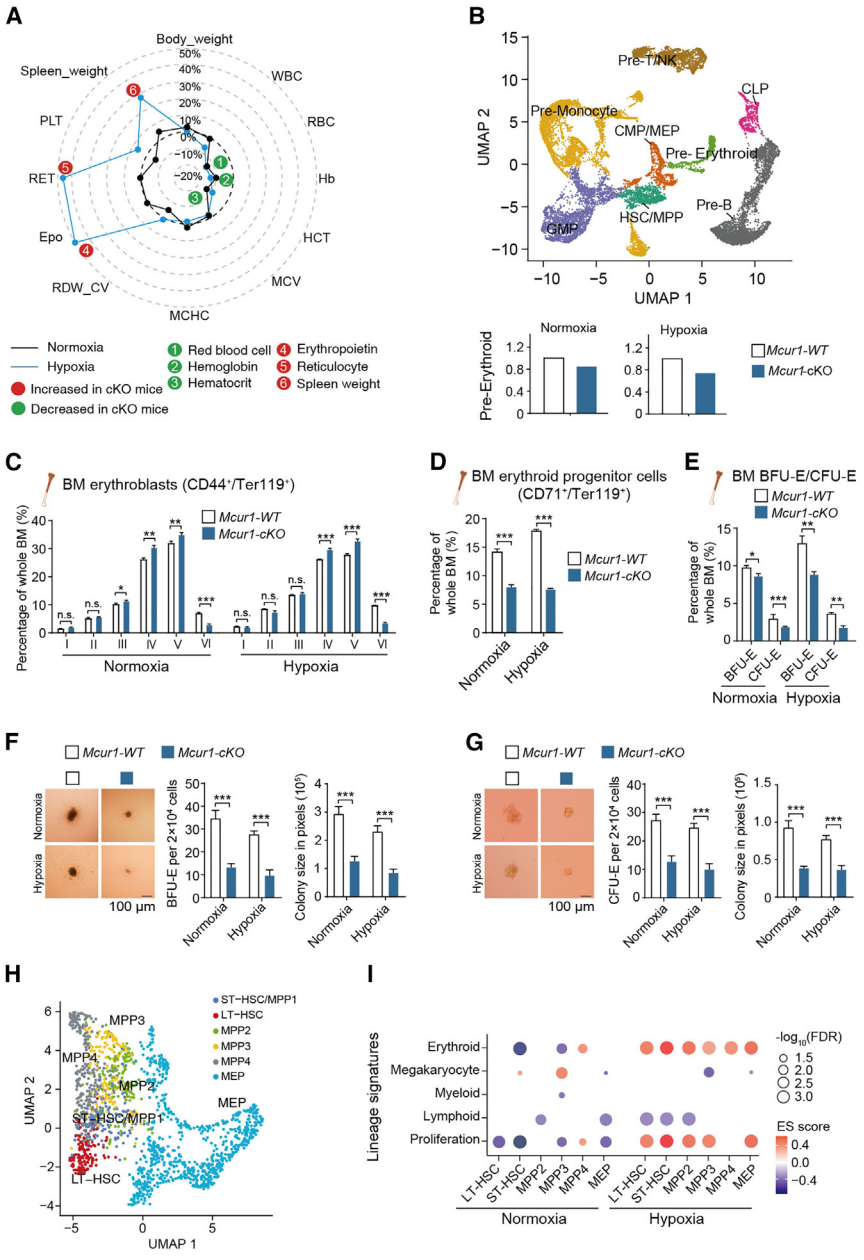


Figure 2. *MCUR1* depletion reduces erythropoiesis in mice

(A) The radar plot of the difference in percentages of body weight, spleen weight, and hematological indexes in peripheral blood when comparing *Mcur1*-WT and *Mcur1*-cKO mice under normoxic or hypoxic conditions.

(B) Uniform manifold approximation and projection (UMAP) plot showing the eight cell clusters identified in cKit$^+$ cells from BM of *Mcur1*-WT and *Mcur1*-cKO mice by scRNA-seq.

(C–E) The fractions of erythroblasts (C), erythroid progenitors (D), and BFU-E/CFU-E cells (E) in bone marrow (BM).

(F and G) The numbers and sizes of the cultured BFU-Es (F) and CFU-Es (G). The representative images of BFU-E/CFU-E are from the 10th/3rd day, respectively.

(H) UMAP plot showing the reclustering of the HSCs/MPPs from (B).

(I) GSEA showing the indicated signatures comparing *Mcur1*-WT with *Mcur1*-cKO mice under hypoxia. The GSEA was performed by weighted Kolmogorov-Smirnov test. FDR, false discovery rate; ES, enrichment score.

Data are shown as the mean ± standard deviation (SD). Data were analyzed by two-tailed t test except where noted otherwise. n.s., not significant. *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$.

compared with *Mcur1*-WT mice, *Mcur1*-KO mice exhibited significantly reduced cell numbers of whole BM (Figure S4J) and red blood cells (RBCs), hemoglobin (Hb) levels, and hematocrit (HCT) and significantly compensatory increased percentages of reticulocyte and erythropoietin (EPO) levels in peripheral whole blood under hypoxia (Figure S4I). Further, fluorescence-activated cell sorting (FACS) assays showed significantly reduced differentiation of erythroblasts and decreased erythroid progenitors, burst-forming unit-erythroids (BFU-Es), and colony-forming unit-erythroids (CFU-Es) in the BM of *Mcur1*-KO mice compared to *Mcur1*-WT mice under hypoxia (Figures S4K–S4M). Together, these findings preliminarily suggest that the heterozygous KO of *Mcur1* leads to erythroid defects *in vivo*.

To further exclude the potential impact of non-hematopoietic cells on observed erythroid defects in the whole body of *Mcur1*-KO mice, we generated a hematopoietic cell-specific *Mcur1* conditional KO (cKO) mouse model, which was confirmed to be homozygous by western blotting assay (Figures S5A–S5C). The hematological analyses of peripheral blood showed that *Mcur1*-cKO mice exhibit the similar erythroid deficiency phenotype as *Mcur1*-KO mice (Figures 2A and S5D). Next, to globally uncover the effect of *Mcur1* KO on the status of HSPCs, we performed scRNA-seq of c-Kit$^+$ cells isolated from BM of *Mcur1*-WT and *Mcur1*-cKO mice under either hypoxic or normoxic conditions. After quality controls, we obtained a total of 35,804 transcriptomes of single cells, which were then clustered and assigned to eight major cell types, including the hematopoietic stem cells and multipotent progenitors (HSCs/MPPs), common progenitors (CPs), and several types of precursor cells (Figures 2B and S5E). Among these major cell types, only the pre-erythroids consistently showed a decrease in cell count in *Mcur1*-cKO mice compared to *Mcur1*-WT mice under both normoxia and hypoxia (Figures 2B and S5F), which was further confirmed by FACS analyses (Figures 2C–2E and S5G–S5J). We also found that the colony number and size of the BFU-Es/CFU-Es cultured *ex vivo* were significantly reduced in *Mcur1*-cKO compared to *Mcur1*-WT mice (Figures 2F and 2G).

For a better resolution of the HSPC population, the cells in the HSC/MPP cluster were reclustered for a second round (Figure 2H), and the impact of *Mcur1* KO on the differentiation spectra of HSPCs was assessed using the specific gene sets. We observed that *Mcur1*-knocked-out long-term HSCs (LT-HSCs), short-term HSCs (ST-HSCs), MPPs (MPP2/3/4), and megakaryocyte/erythroid progenitors (MEPs) exhibit markedly reduced erythroid signature and proliferation signature under hypoxia (Figure 2I). In addition, the trajectory analyses showed that *Mcur1-cKO* HSPCs exhibit significantly lower expression levels of erythropoiesis-related genes (e.g., *Gata1* and *Epo*) along the HSC-MPP-MEP differentiation trajectory (Y_MEP) (Figures S5K and S5L). Together, these data suggest that *Mcur1* depletion reduces the erythroid differentiation preference of HSPCs in mice.

Notably, despite the comparable body weights between *Mcur1*-WT and *Mcur1*-cKO mice (Figure 2A), *Mcur1*-cKO mice showed higher spleen volume and weight compared with *Mcur1*-WT mice (Figures 2A and S5M). Further, the histological analyses of spleen sections from *Mcur1*-cKO mice showed slight splenomegaly with obscured spleen architecture and expansion of the red pulp (Figure S5M). FACS analyses of the spleen showed that compared with *Mcur1*-WT mice, the percentages of erythroid progenitor cells (Figure S5N) and differentiated erythroblasts (Figure S5O) were significantly increased and the percentages of macrophage and monocyte cells were significantly decreased (Figure S5P), while there were no significant changes in megakaryocytes and granulocytes (Figures S5Q and S5R) and megakaryocyte progenitor cells (Figure S5S) in *Mcur1*-cKO mice. Together, given that the spleen is an emergency hematopoietic organ in adults, these results suggest that a decrease in erythropoiesis in BM might lead to the compensatory expansion of the early erythroid compartment in spleen during hypoxia.[25]

MCUR1 knockdown reduces erythropoiesis in human CD34$^+$ HSPCs under hypoxia

Next, we investigated the *ex vivo* effects of MCUR1 on erythropoiesis in CD34$^+$ HSPCs derived from human cord blood and BM under normoxia (21% O_2) or hypoxia (1% O_2). First, in the model of CD34$^+$ HSPCs derived from the human cord blood (Figures 3A and S6A), we observed that *MCUR1* knockdown in HSPCs significantly reduces the proportions of the differentiated erythrocytes upon hypoxia, which were indicated by a decreased number of benzidine-positive cells, and *CD235a* and *HBG1* mRNA levels (two classical markers of erythroid cells) (Figures 3B and 3C). Morphological analyses of primary erythroid cell differentiation also showed a significant increase in proportions of immature erythroblasts (basophilic and polychromatic erythroblasts) and a concomitant decrease in proportions of mature erythroblasts (orthochromatic and erythrocyte) when *MCUR1* was knocked down upon hypoxia (Figure 3D), which was consistent with the decreased percentage of erythroblasts by FACS analyses (Figure 3E). Further, *MCUR1* knockdown significantly reduced the percentage of erythroid progenitor CFU-E (IL3R$^-$GPA$^-$CD34$^-$CD36$^+$) through FACS analyses (Figure 3F) and the colony number and size of cultured BFU-E and CFU-E cells during hypoxia (Figures 3G and 3H). Finally, *MCUR1* knockdown significantly delayed cell growth by inducing apoptosis and reducing the cell-cycle progression of HSPCs (Figures 3I–3K). The effects of *MCUR1* knockdown on erythroid differentiation, cell growth, apoptosis, and cell cycle could be rescued by reexpression of *MCUR1* in human cord blood CD34$^+$ HSPCs (Figure 3).

To minimize the impact of behavioral differences in types of CD34$^+$ HSPCs and ensure the reliability of the results, we also investigated the function of MCUR1 in CD34$^+$ HSPCs derived from human BM and obtained similar results (Figures S7A–S7L). In addition, the inhibitory effects of *MCUR1* knockdown on erythroid differentiation were also observed in normoxic conditions in human cord blood and BM cells (Figures S6B–S6K and S7A–S7L). Accordingly, the *MCUR1* mRNA levels were positively correlated with Hb levels in the peripheral blood of 200 Tibetans from the TIB_replication 2 population ($r = 0.27$, $p = 4.5 \times 10^{-4}$, linear correlation test; Figure S7M). Collectively, these results suggest that *MCUR1* knockdown impairs erythropoiesis in human CD34$^+$ HSPCs under both normoxia and hypoxia.

MCUR1 knockdown reduces erythropoiesis by inhibiting the mTOR pathway

Next, we sought to explore the underlying mechanisms by which *MCUR1* knockdown reduces erythropoiesis. RNA-seq followed by gene set enrichment analysis (GSEA) showed significant enrichment of cellular bioenergetics-associated transcriptional signature and mTOR pathway in *MCUR1*-knocked-down human CD34$^+$ HSPCs under hypoxia following EPO induction (Figures 4A and 4B; Table S13). Intriguingly, the mTOR pathway has previously been shown to be involved in erythropoiesis.[25–27] We therefore hypothesize that *MCUR1* knockdown might reduce erythropoiesis by inhibiting the mTOR pathway during hypoxia.

Indeed, we observed that in human CD34$^+$ HSPCs under hypoxia, *MCUR1* knockdown reduced the phosphorylation levels of mTOR (p-mTOR) and its two downstream targets S6K (p-S6K) and 4EBP1 (p-4EBP1) (Figure 4C). *MCUR1* knockdown also reduced the cells' global protein translation rate (Figures 4D and S8A), which was consistent with the function of mTORC1-mediated ribosome biogenesis.[27] Further, the inhibitory effects of *MCUR1* knockdown on mTOR pathway, erythropoiesis, and cell growth in human CD34$^+$ HSPCs can be eliminated by using a pharmacologic activator of mTOR (MHY1485) (Figures 4C, 4E–4H, S8B, and S8C) or small interfering RNAs (siRNAs) targeting *TSC1* or *TSC2*, two genes encoding upstream kinases that inhibit mTOR activity (Figures S8D–S8J). However, *MCUR1* knockdown did not affect mitochondrial biogenesis-related transcription factors (TFs) YY1, PGC1α, NRF1, and mtTFA[28] (Figures 4C and S8D). In addition, expression of p-ULK1, a key downstream protein involved in mitochondrial autophagy,[29] and two autophagic markers, p62 and LC3, also remained unchanged (Figure 4C). Collectively, these results suggest that *MCUR1* knockdown reduces mTOR-mediated protein translation rate and cell growth, but does not affect mTOR-mediated mitochondrial biogenesis and autophagy, thereby weakening erythropoiesis.

MCUR1 knockdown inhibits the mTOR pathway by inducing AMPK activity

We next asked how *MCUR1* knockdown reduces mTOR signaling. Similar to previous studies in HeLa and HEK293T

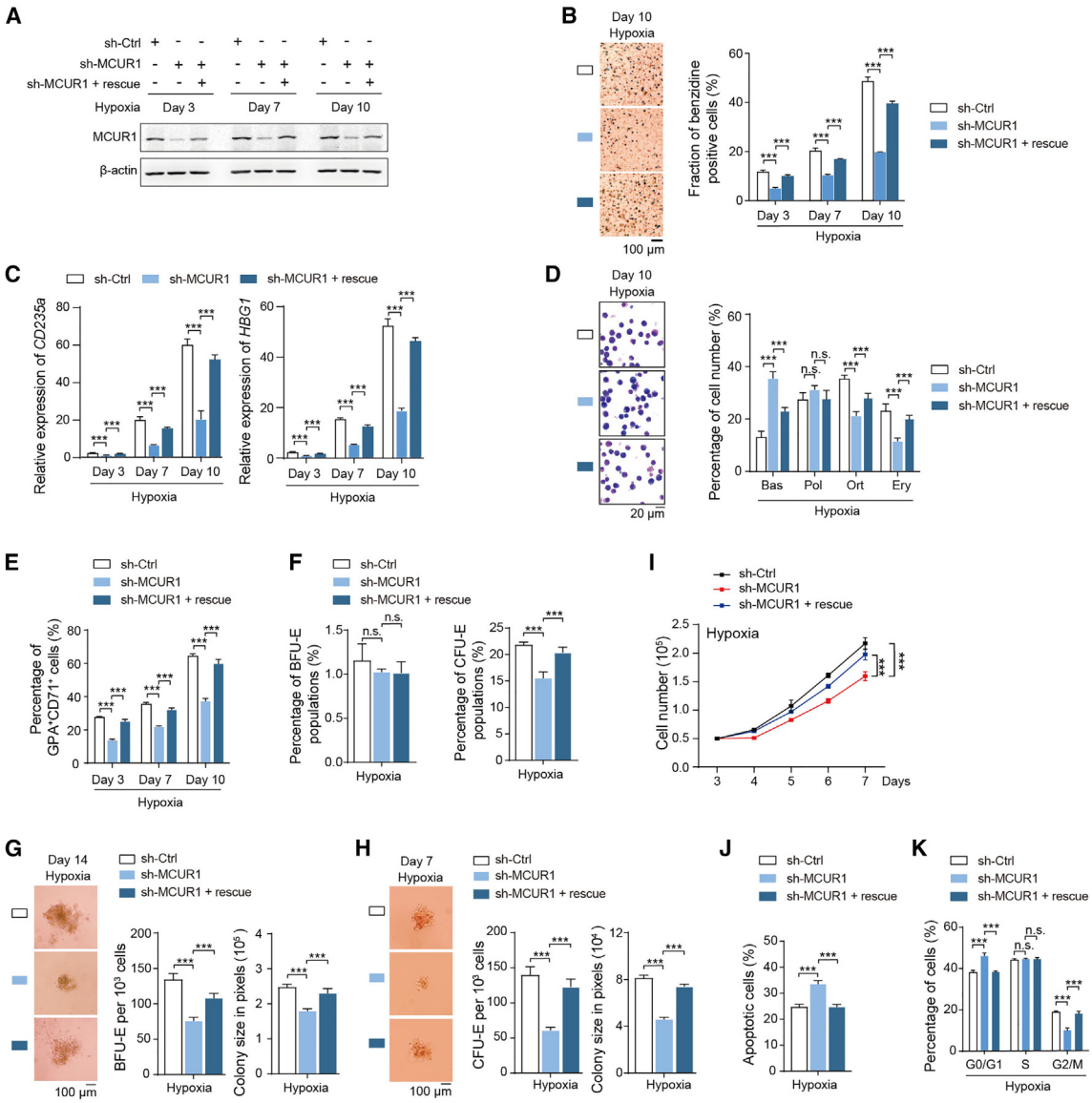


Figure 3. *MCUR1* downregulation reduces erythropoiesis in human cord blood CD34$^+$ HSPCs

(A) Efficiency of knockdown or reexpression of *MCUR1* by western blotting assays.

(B–F) *MCUR1* knockdown reduces the fraction of benzidine-positive cells (B); *CD235a*/*HBG1* mRNA levels (C); percentages of basophilic (Bas), polychromatophilic (Pol), and orthochromatic (Ort) cells and erythroblasts and erythrocytes (Ery) (D); fraction of erythroblasts (E); and percentages of BFU-E (IL3R$^-$GPA$^-$CD34$^+$CD36$^-$) and CFU-E (IL3R$^-$GPA$^-$CD34$^-$CD36$^+$) populations (F). In (B) and (D), left side shows the representative images.

(G and H) *MCUR1* knockdown reduces the number and size of colonies formed of BFU-E (G) and CFU-E (H). Representative images of BFU-E/CFU-E at the 14th/7th day are shown.

(I–K) *MCUR1* knockdown leads to reduced cell growth (I), increased cell apoptosis (J), and decreased cell-cycle progression (K). In (I), the analyses between groups were by two-way ANOVA.

Data are shown as mean ± SD. Data analyses were conducted by two-tailed t test except where noted otherwise. n.s., not significant. $^*p < 0.05$, $^{**}p < 0.01$, and $^{***}p < 0.001$.

cells,[19,30] *MCUR1* knockdown in human CD34$^+$ HSPCs markedly decreased mitochondrial Ca^{2+} ($[Ca^{2+}]_m$) uptake, oxygen consumption rate (OCR), ATP content, mitochondrial membrane potential ($\Delta\Psi m$), and mitochondrial reactive oxygen species (ROS) production (Figures 5A and S9A–S9D), suggesting that MCUR1 also plays essential roles in mitochondrial bioenergetics in hematopoietic cells. It is well known that AMP-activated protein kinase (AMPK) is an evolutionarily conserved energy sensor that regulates mitochondrial homeostasis and cellular bioenergetics.[31] Indeed, we observed that *MCUR1* knockdown in human CD34$^+$ HSPCs can enhance the activation of energy-sensing AMPK, with increased levels of T172-phosphorylated

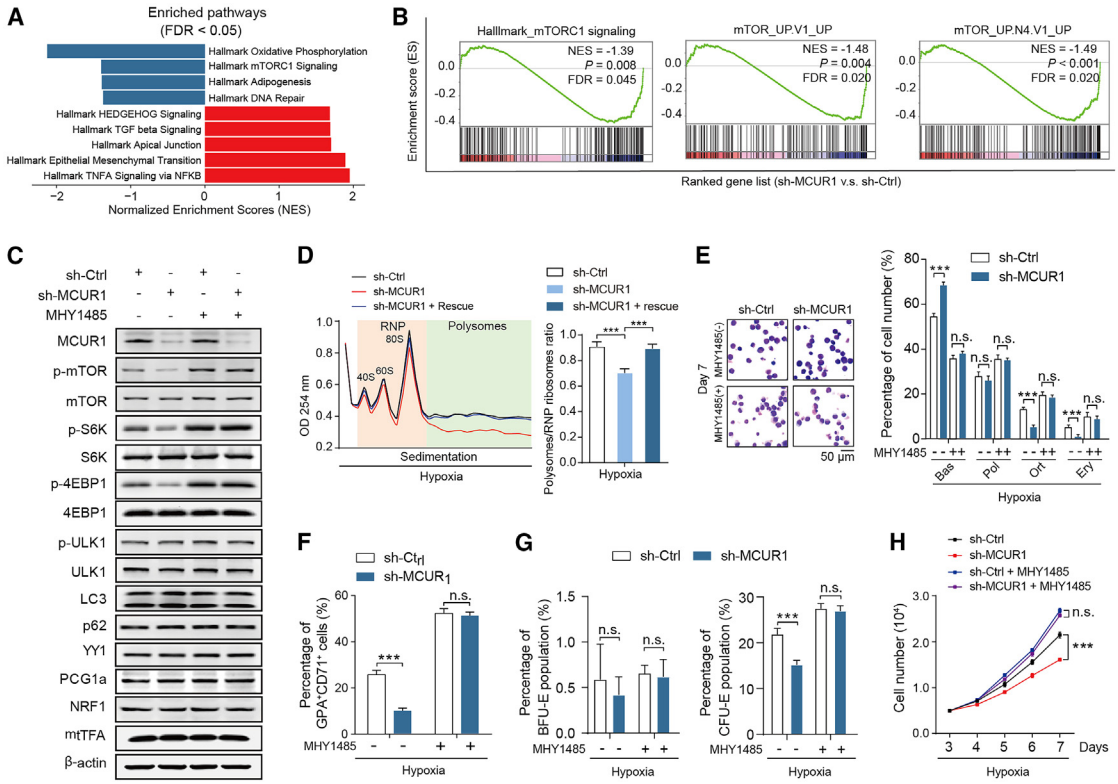


Figure 4. *MCUR1* downregulation reduces erythropoiesis by reducing the activation of the mTOR pathway in human cord blood CD34$^+$ HSPCs

(A) GSEA showing top enriched hallmark pathways after *MCUR1* knockdown in human cord blood CD34$^+$ HSPCs under hypoxia (1% O_2) and EPO treatment (3 U/mL) for 7 days.

(B) GSEA showing the enrichment of mTOR pathways.

(C) *MCUR1* knockdown reduces the mTOR activity, and activation of mTOR (by MHY1485 [2 μM]) abolishes the decreased mTOR activity by *MCUR1* knockdown.

(D) *MCUR1* knockdown reduces protein translation rate by polysome profiling assays. OD, optical density; RNP, ribonucleoprotein.

(E) The effect of activation of mTOR on the changes in proportions of basophilic (Bas), polychromatophilic (Pol), and orthochromatic (Ort) cells and erythroblasts and erythrocytes (Ery) by *MCUR1* knockdown. Left side shows the representative images.

(F) The effect of activation of mTOR on the decreased fraction of erythroblasts by *MCUR1* knockdown.

(G) The effect of activation of mTOR on the decreased percentages of CFU-E by *MCUR1* knockdown.

(H) The effect of activation of mTOR on the decreased cell growth by *MCUR1* knockdown. The analysis of cell growth between groups was by two-way ANOVA. Data are shown as mean ± SD. Data analyses were conducted by two-tailed t test except where noted otherwise. n.s., not significant. ***$p < 0.001$.

AMPKα (p-AMPKα) (Figure 5B), which is also consistent with the previous findings observed in HeLa and HEK293T cells.[19,30] Several previous studies have shown that AMPK inhibits the mTOR pathway by enhancing the levels of p-TSC2 (at Ser1387) and p-Raptor (at Ser792), two upstream kinases that inhibit mTORC1 activity, and then suppresses cell growth.[32–34] Notably, the finely tuned regulation of AMPK has been shown to be crucial in human adult erythropoiesis.[32–34] Therefore, we hypothesize that *MCUR1* knockdown might reduce the mTOR pathway and then attenuate erythropoiesis by enhancing AMPK activity during hypoxia.

Indeed, we observed that knockdown of *AMPKa1* (encoding AMPKα) significantly reduces the levels of p-TSC2 and p-Raptor, thereby promoting mTOR signaling in human CD34$^+$ HSPCs (Figure 5B). Further, when AMPK activity was reduced by using either the siRNAs targeting *AMPKa1* (Figures 5B–5E and S10A–S10C) or compound C (an AMPK inhibitor) (Figures S10D–S10J), knockdown of *MCUR1* did not affect the levels of p-TSC2, p-Raptor, p-mTOR, p-S6K, or p-4EBP1 or the percentages of erythroid progenitors, proportions of differentiated erythroblasts, or cell growth. Finally, inhibition of AMPK by siRNAs targeting *AMPKa1* (Figures S10K–S10M) or compound C (Figures S10N–S10P) cannot significantly influence $[Ca^{2+}]_m$ uptake or $\Delta\Psi m$, indicating that AMPK is the downstream component of the Ca^{2+} signal. Taken together, these findings suggest that AMPK is required for *MCUR1* knockdown to reduce the mTOR activity and erythropoiesis during hypoxia.

MCUR1 knockdown promotes AMPK activity by inducing CAMKK2 activity

Next, we sought to decipher the underlying mechanism of *MCUR1* knockdown-mediated AMPK activation. We observed

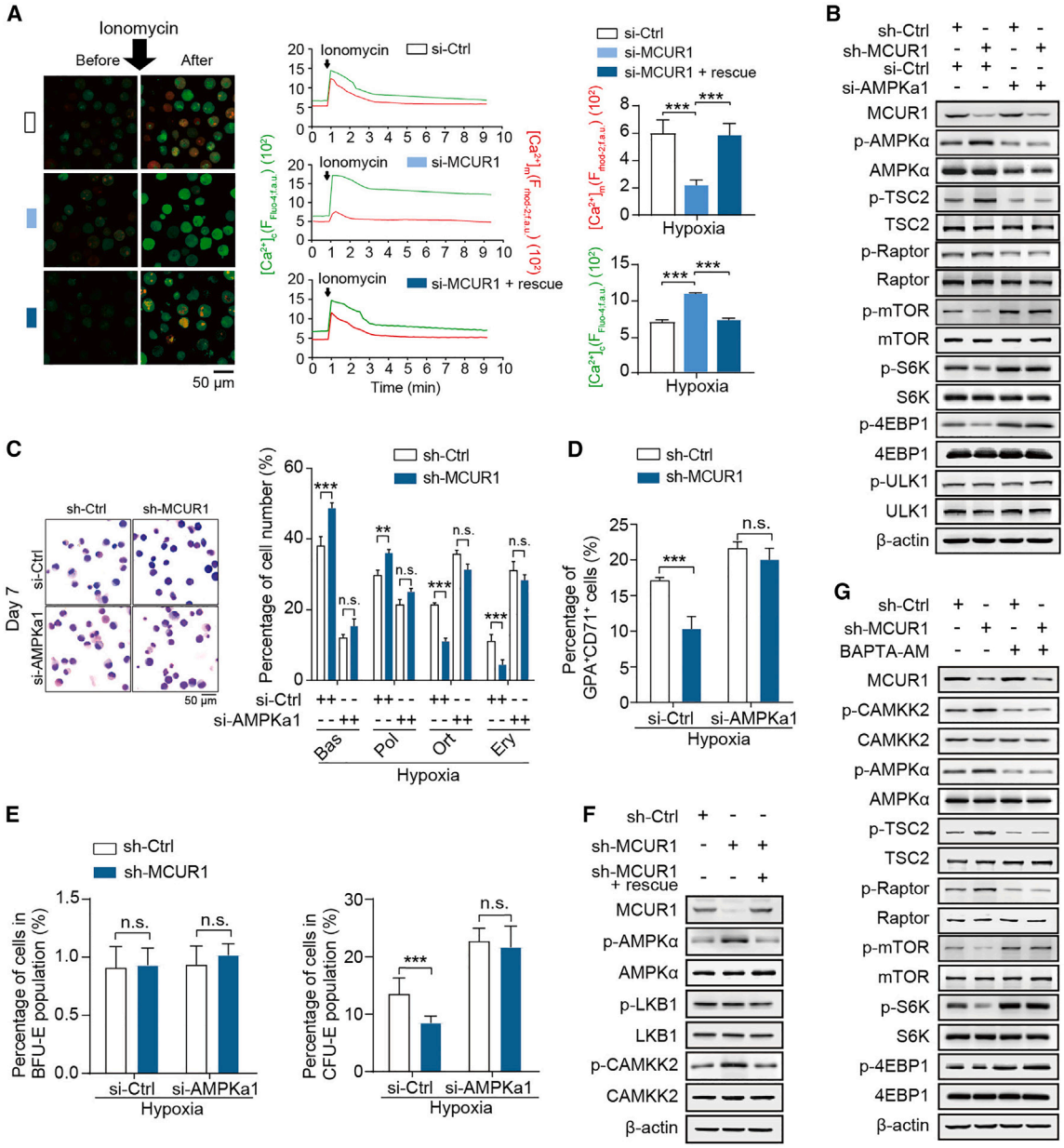


Figure 5. *MCUR1* downregulation reduces activation of the mTOR pathway through the Ca^{2+}-CAMMK2-AMPK axis in human cord blood $CD34^+$ HSPCs

(A) The effects of *MCUR1* knockdown on mitochondrial ($[Ca^{2+}]_m$; red) and cytoplasmic ($[Ca^{2+}]_c$; green) Ca^{2+} responses in human cord blood $CD34^+$ HSPCs before and after ionomycin exposure (2.5 μM), under hypoxia (1% O_2) and EPO treatment (3 U/mL) for 7 days.

(B) Knockdown of *AMPKa1* induces mTOR activity and abolishes the decreased mTOR activity by *MCUR1* knockdown.

(C) The effect of *AMPKa1* knockdown on the changes in percentages of basophilic (Bas), polychromatophilic (Pol), and orthochromatic (Ort) cells and erythroblasts and erythrocytes (Ery) by *MCUR1* knockdown. Left side shows the representative images.

(D) The effect of *AMPKa1* knockdown on the decreased fraction of erythroblasts by *MCUR1* knockdown.

(E) The effect of *AMPKa1* knockdown on the decreased fraction of CFU-E population by *MCUR1* knockdown.

(F) The effects of *MCUR1* knockdown on the activity of two upstream kinases of AMPKα, LKB1 and CAMKK2.

(G) The effects of inhibition of CAMMK2 activity (by BAPTA-AM, an intracellular Ca^{2+} chelator; 1 μM) on the increased AMPKα activity and decreased mTOR activity by *MCUR1* knockdown.

Data are shown as mean ± SD. Data analyses were conducted by two-tailed t test except where noted otherwise. n.s., not significant. $^{**}p < 0.01$ and $^{***}p < 0.001$.

that *MCUR1* knockdown in human CD34$^+$ HSPCs under hypoxia markedly enhances the phosphorylation levels of calcium/calmodulin-dependent protein kinase kinase β (CAMKK2) (Figure 5F), which is an upstream Ca^{2+}-dependent kinase responsible for phosphorylation at Thr172 of AMPKα.[28] However, *MCUR1* knockdown did not affect the phosphorylation levels of another constitutive liver kinase B1 (LKB1) upstream of AMPKα[28] (Figure 5F). In addition, okadaic acid, a pan-phosphatase inhibitor,[28] did not affect the p-AMPKα levels affected by *MCUR1* knockdown (Figure S10Q), therefore also excluding the possibility of phosphatases acting in *MCUR1* knockdown-mediated AMPK activation. Further, we observed that inhibition of *CAMKK2* by siRNAs targeting *CAMKK2* (Figures S11A–S11G) or STO-609 (a CAMKK2 inhibitor) (Figures S11H–S11N) reduced p-AMPKα levels and increased p-mTOR levels, percentages of erythroid progenitors, proportions of erythroblasts, and cell growth. Finally, when *CAMKK2* was knocked down by siRNAs (Figures S11A–S11G) or inhibited by STO-609 (Figures S11H–S11N), *MCUR1* knockdown did not affect the levels of p-AMPKα or p-mTOR, the percentages of erythroid progenitors, the proportions of erythroblasts at different stages of differentiation, or cell growth. Together, these findings suggest that under hypoxia, *MCUR1* knockdown-mediated enhancement of AMPK activity and reduction of mTOR activity and erythropoiesis depend on CAMKK2.

MCUR1 knockdown activates CAMKK2 by inducing cytosolic Ca^{2+} levels

Given that *MCUR1* knockdown reduces $[Ca^{2+}]_m$ uptake (Figure 5A), and knockdown of Ca^{2+}-dependent kinase CAMKK2 decreases p-AMPKα in human CD34$^+$ HSPCs (Figure S11A), we further hypothesized that *MCUR1* knockdown promotes CAMKK2 activity and its downstream AMPK signaling by increasing the levels of cytosolic Ca^{2+} ($[Ca^{2+}]_c$). Indeed, when human CD34$^+$ HSPCs were pre-treated with an intracellular Ca^{2+} chelator, 1,2-bis (2-aminophenoxy)ethane-N,N,N′,N′-tetraacetic acid tetrakis (acetoxymethyl ester) (BAPTA-AM),[35] under hypoxia, knockdown of *MCUR1* did not affect the levels of p-CAMKK2, p-AMPKα, p-TSC2, p-Raptor, p-mTOR, or downstream p-S6K and p-4EBP1 (Figure 5G), therefore suggesting the dependence of $[Ca^{2+}]_c$ levels. Thus, these findings indicate that *MCUR1* knockdown reduces rapid transport of Ca^{2+} into the mitochondrial matrix and then concomitantly increases the $[Ca^{2+}]_c$ levels, which in turn leads to activation of Ca^{2+}-dependent CAMKK2 and downstream AMPK signaling.

Finally, we detected the phosphorylation levels of several proteins in the CAMMK2-AMPK-mTOR axis in BM cells isolated from *Mcur1*-WT and *Mcur1*-cKO mice. Compared with those from *Mcur1*-WT mice, BM cells from *Mcur1*-cKO mice showed markedly increased levels of p-CAMKK2 and p-AMPK and decreased levels of p-mTOR, p-S6K, and p-4EBP1 under both normoxia and hypoxia (Figure S12A). Meanwhile, *Mcur1*-cKO BM cells showed significantly reduced $[Ca^{2+}]_m$ levels and concomitantly increased $[Ca^{2+}]_c$ levels (Figure S12B) and significantly reduced mitochondrial bioenergetics, including reduced OCR and ATP content and concomitantly increased AMP/ATP ratio and reduced mitochondrial $\Delta\Psi m$ and ROS (Figures S12C–S12F). Accordingly, BM cells from *Mcur1*-cKO mice showed markedly reduced protein translation rate (Figure S12G) and significantly decreased cell-cycle progression and increased apoptosis (Figures S12H and S12I). Collectively, these findings suggest that the MCUR1-Ca^{2+}-CAMMK2-AMPK-mTOR axis works in erythropoiesis *in vivo* during hypoxia.

rs61644582 at 6p23 regulates *MCUR1* expression in an allele-specific manner

Next, we sought to discriminate the functional variant(s) that may regulate *MCUR1* expression at the 6p23 locus. There exist 14 SNPs within the ~2 kb LD block at 6p23 (Table S11), and among them, four SNPs (rs1204167, rs1204168, rs1204170, and rs61644582) were predicted to be located at the binding sites of certain TFs using HaploReg (v.4.1). However, among these four SNPs, only rs61644582 was predicted to create an allele-specific binding motif that matches the consensus sequences of two TFs, PU.1 and ELF1 (Figure 6A), both of which were critical transcriptional regulators relevant to erythropoiesis.[36–38] In addition, several epigenetic modification enrichment profiles from the ENCODE project showed that rs61644582 is located in an active enhancer region of multiple cell lines, including two human blood cell lines K562 and GM12878 (Figure 6A). The genome-wide chromosome conformation capture (Hi-C) datasets from these two cell lines also showed that rs61644582 physically interacts with the promoter region of the *MCUR1* gene (Figures S13A and S13B). From these results, we hypothesized that rs61644582 is the causative SNP at the 6p23 signal.

To test this hypothesis, we first investigated whether rs61644582 alters the transcriptional activity of *MCUR1*. Luciferase reporter assays showed that in two types of human blood cell lines, HUDEP2 and K562 cells, both the del and the C allele of rs61644582 repress luciferase expression relative to the basal *MCUR1* promoter, and further, the del allele drives lower luciferase activity than the C allele (both $p < 0.001$; Figure 6B). These data suggest that the motif harboring rs61644582 acts as a repressor in *MCUR1* expression, consistent with the aforementioned finding that the rs61644582 del allele is significantly associated with decreased *MCUR1* expression in PBMCs (Figures 1F and 1G).

We then examined whether rs61644582 could affect binding with a specific TF(s), as suggested by the aforementioned *in silico* predictions (Figure 6A). Indeed, there was overall stronger affinity of nuclear protein binding to the probes containing the del allele than there was for the probes containing the C allele (Figures 6C and S13C, lanes 2–11) by the electrophoresis mobility shift assay (EMSA). In addition, the probes containing the del allele were competed away by excessive unlabeled consensus sequence probes that bind to PU.1 or ELF1 (Figures 6C and S13C, lanes 12 and 13), indicating that PU.1- or ELF1-containing complexes bind at rs61644582. These findings were further confirmed by the observation of a band shift for the del allele in the presence of the antibodies against PU.1 or ELF1 using nuclear extracts from HUDEP2 or K562 cells (Figures 6D and S13D). However, western-blotting EMSA (WEMSA) showed that PU.1, but not ELF1, binds more to the del allele in comparison to C allele (Figures 6E and S13E). Using the chromatin immunoprecipitation (ChIP)-quantitative PCR (qPCR) assay, we confirmed that PU.1 binds at the

CellPress
OPEN ACCESS

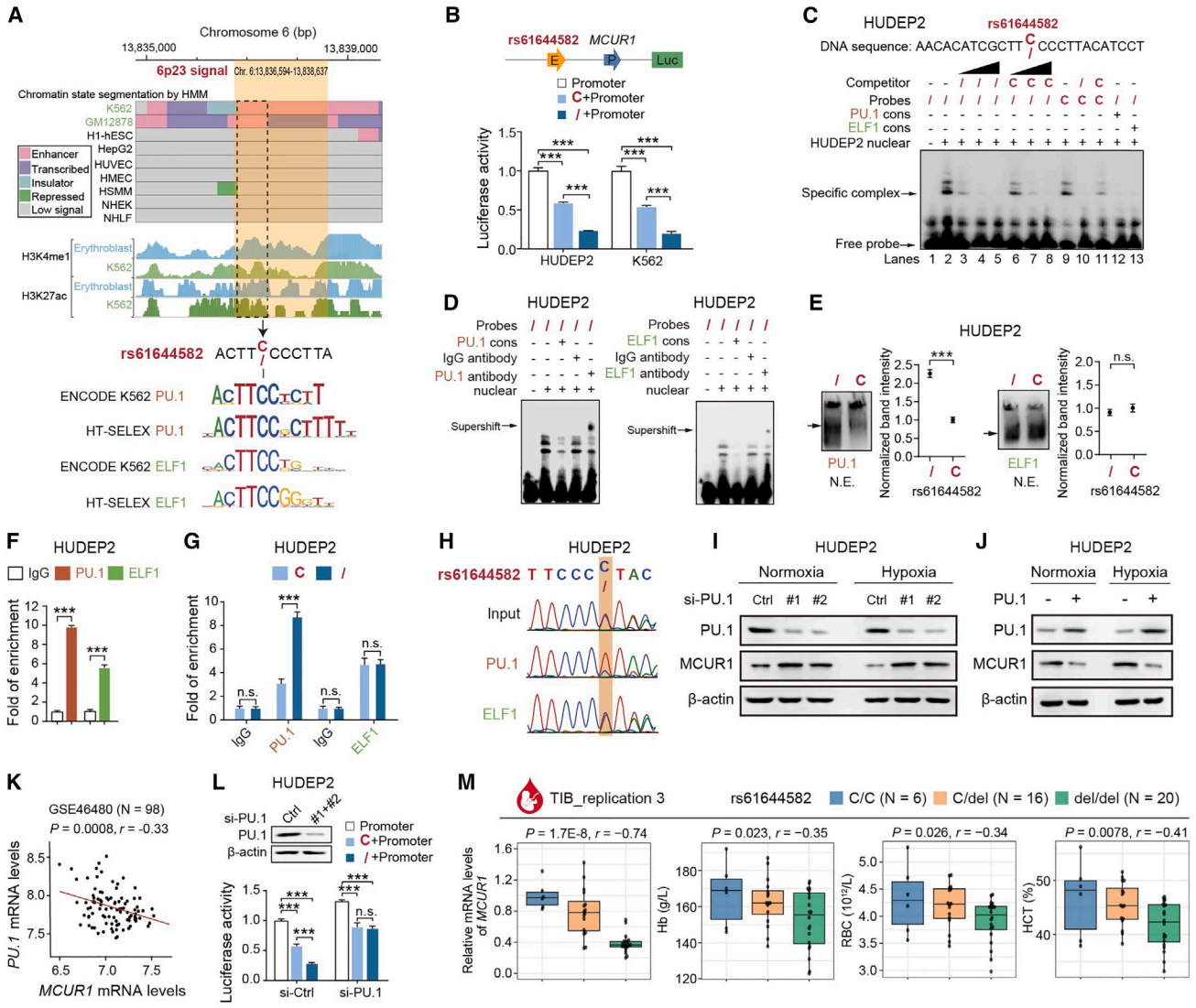


Figure 6. PU.1 preferentially binds to rs61644582 deletion allele and reduces *MCUR1* expression to a greater extent compared to C allele

(A) A comprehensive map around rs61644582 at the 6p23 locus. The epigenetic modification enrichment profiles in erythroblasts and K562 were derived from the BLUEPRINT and ENCODE projects, respectively. C and /, the C and deletion (del) allele of rs61644582, respectively.

(B) The effects of rs61644582 on transcriptional activity of the *MCUR1* promoter by luciferase reporter gene assays. E, enhancer; Luc, luciferase gene; P, promoter.

(C) The binding affinity of del and C alleles of rs61644582 to PU.1 by EMSA. Competition gradients, which are represented by black triangles, correspond to 5-, 50-, and 100-fold excess of cold probes.

(D) The rs61644582 del allele abolishes the binding of a protein complex that is supershifted by the antibody against PU.1 or ELF1.

(E) The binding affinity of del and C alleles of rs61644582 to PU.1 or ELF1 in nuclear extract (N.E.) by WEMSA. The protein quantification was determined using ImageJ.

(F) Chromatin binding of PU.1 or ELF1 at the rs61644582 site by ChIP-qPCR.

(G) PU.1, rather than ELF1, favors binding to the rs61644582 del allele as determined by ChIP-AS-qPCR in HUDEP2 cells, which are heterozygous at rs61644582.

(H) PU.1, rather than ELF1, favors binding to the rs61644582 del allele as determined by ChIP-PCR followed by Sanger sequencing.

(I) *PU.1* knockdown induces MCUR1 expression.

(J) Overexpression of PU.1 reduces MCUR1 expression.

(K) Association between the mRNA levels of *MCUR1* and *PU.1* in PBMCs.

(L) The stronger inhibitory effect of rs61644582 del allele relative to C allele on luciferase reporter activity was abolished when *PU.1* was knocked down. Data are shown as mean ± SD.

(M) Association of rs61644582 genotypes with the mRNA levels of *MCUR1* and several erythrocyte-related indicators in human cord blood samples of 42 Tibetans (TIB_replication 3).

The data in (B), (E–G), and (L) were analyzed using two-tailed t test, while the data in (K) and (M) were analyzed by linear regression analysis. n.s., not significant. ***$p < 0.001$.

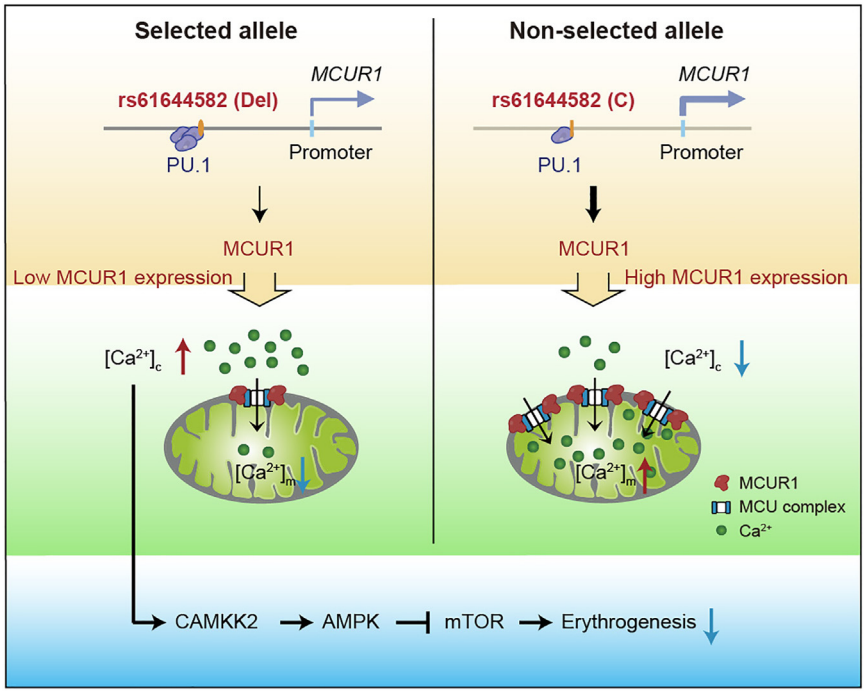


Figure 7. A model of the allele-specific transcriptional regulation of *MCUR1* by rs61644582 and the function and underlying mechanism of MCUR1 in regulating erythropoiesis

The positively selected beneficial rs61644582 del allele has a stronger binding ability to the transcription factor PU.1 relative to the C allele, therefore resulting in relatively lower expression of *MCUR1*. Downregulation of MCUR1 reduces the mitochondrial Ca^{2+} uptake and then concomitantly increases the cytosolic Ca^{2+} levels, thereby attenuating erythropoiesis via the CAMKK2-AMPK-mTOR axis.

rs61644582 site *in vivo* in HUDEP2 and HEK293T cells, which were confirmed to be heterozygous at this SNP site (Figures 6F and S13F). Further, ChIP followed by allele-specific qPCR (AS-qPCR) (Figures 6G and S13G) or by PCR and Sanger sequencing assays (Figures 6H and S13H) consistently showed that PU.1, but not ELF1, is preferentially recruited to the rs61644582 del allele compared to the C allele. Collectively, these results reveal del-allele-specific binding of PU.1 at a distal regulatory element encompassing rs61644582 and its inhibitory effect on the activity of the *MCUR1* promoter, therefore consistent with the previous reports on the function of PU.1 as a transcription repressor[39] and its well-known role as a repressor in erythropoiesis.[36–38]

Next, we assessed whether the effects of rs61644582 on *cis* regulation of *MCUR1* expression are dependent on PU.1-chromatin binding. Knockdown of *PU.1* in HUDEP2 and K562 cells greatly enhanced MCUR1 protein levels (Figures 6I and S13I), and conversely, overexpression of PU.1 had the opposite effect (Figures 6J and S13J). Consistently, the expression levels of *PU.1* were significantly negatively associated with *MCUR1* expression levels in PBMCs from 98 individuals residing at high altitude (GEO: GSE46480; $r = -0.33$, $p = 0.0008$; Figure 6K). Further, we found that knockdown of *PU.1* eliminates the allelic difference in activation of luciferase at the rs61644582 del- or C-allele-containing region (Figures 6L and S13K). Together, these findings suggest that the allele-specific regulation of rs61644582 on *MCUR1* expression is dependent on PU.1.

rs61644582 correlates with erythropoiesis-related phenotypes in Tibetans

Finally, we sought to assess whether the functional rs61644582 is physiologically associated with erythropoiesis-related phenotypes in Tibetans living at high altitude. To this end, we recruited a cohort of 42 Tibetan newborns (designated as TIB_replication 3; Table S1) and collected their cord blood. Genotyping using Sanger sequencing showed that there were 6 individuals with rs61644582 C/C, 16 with C/del, and 20 with del/del genotypes. Consistent with the eQTL results, the rs61644582 del allele was significantly associated with lower *MCUR1* mRNA levels in $CD34^+$ HSPCs isolated from cord blood (Figure 6M). Further, the cord blood with the del allele had significantly reduced RBC count, HCT, and Hb level (Figure 6M); reduced percentage of erythroblasts (Figure S13L) and proportions of differentiated erythroblasts (Figure S13M); and reduced colony number and size of cultured BFU-E or CFU-E based on $CD34^+$ HSPCs (Figure S13N) compared to those without the del allele. Accordingly, the cord blood $CD34^+$ HSPCs with the del allele had markedly increased levels of p-CAMKK2 and p-AMPK, as well as decreased levels of p-mTOR compared to those without the del allele (Figures S14A and S14B). Similarly, low *MCUR1* mRNA or protein levels mediated by the rs61644582 del allele were consistently correlated with these erythropoiesis-related phenotypes and levels of p-CAMKK2, p-AMPK, and p-mTOR (Figures S14C–S14G). In addition, the significant associations between the rs61644582 del allele and decreased Hb levels were also observed in multiple Tibetan populations (Table S14). Collectively, these findings reveal the causal effect of the rs61644582 del allele on reducing *MCUR1* expression by enhancing PU.1-chromatin binding, suggesting that rs61644582 and *MCUR1* are the plausible causative SNP and gene, respectively, therefore explaining the HAA (e.g., reduced erythropoiesis) relevant to the positive selection signal at 6p23 in Tibetans (Figure 7).

DISCUSSION

Mitochondrial Ca^{2+} homeostasis is responsible for the regulation of numerous cellular functions, including energy metabolism and cell growth and death. Several components of the protein machinery governing mitochondrial Ca^{2+} homeostasis have recently been elucidated, including MCU and its regulators MICU1/2/3 and MCUR1.[31,40] *MICU1* has been identified as the candidate

HAA-associated gene in Ethiopian highlanders of African origin.[41] Interestingly, we here identified *MCUR1* as another candidate in Tibetan highlanders of Chinese origin, suggesting parallel evolution for HAA existing in high-altitude populations of different ancestries. In addition, recent studies have revealed that several other members of MCUC can regulate mitochondrial Ca^{2+} homeostasis in specific cell/tissue types during hypoxia.[31] Here, we observed that *MCUR1* is specifically highly expressed in erythroid cells, and its depletion reduces erythropoiesis by attenuating mitochondrial Ca^{2+} uptake during hypoxia *in vitro* and *in vivo*. MCUR1 acts as a positive modulator of MCUC, whereas MICU1 exhibits a dual character on regulation of MCUC opening, both as a "gatekeeper" at low cytosolic Ca^{2+} concentrations and as a cooperative activator at high cytosolic Ca^{2+} concentrations.[42] Therefore, their functions in the modulation of MCUC exhibit some differences and complexity, and their interactions deserve further analyses. To our best knowledge, this is the first study establishing the mechanistic link between the MCUC-mediated Ca^{2+} dynamics and erythropoiesis.

AMPK is a highly conserved metabolic energy sensor allowing for adaptive changes in cell growth and differentiation.[31] In addition, AMPK is essential to regulation of breathing during hypoxia and thus oxygen and energy distribution to the body.[30] Recent studies have revealed several aspects of the effects of AMPKα on erythropoiesis, such as inducing O_2 release from erythrocytes[43] and controlling proliferation and survival of erythroblasts.[32,44] Here, we revealed a new role for AMPKα, which is activated in the condition of *MCUR1* downregulation, in suppressing differentiation preference from EEP to terminal mature erythroblast.

A major undertaking of genetic studies is to determine functional variants of candidate genes. A few successful examples are an intronic mutation of *EPAS1* and two missense mutations at exon 1 of *EGLN1* suggesting that Tibetan versions of *EPAS1* and *EGLN1* push toward the same direction by blunting the response to hypoxia.[8,45] Here, we identified a high-frequency functional variant rs61644582, which is located in an intergenic region and distally regulates transcription of *MCUR1* located ~20 kb downstream. Similar to *EPAS1* and *EGLN1* variants, rs61644582 contributes to HAA of Tibetans by blunting erythropoietic response to hypoxia. Furthermore, we revealed that this variant regulates erythropoiesis through the PU.1-MCUR1-Ca^{2+}-CAMKK2-AMPK-mTOR axis, adding a new layer of regulation of erythropoiesis in Tibetans.

Downregulation of PU.1 is a key step to allow erythroid progenitors to undergo terminal differentiation.[37] PU.1 upregulates the molecules that block erythroid differentiation, such as Fli1 and c-myb, and downregulates the molecules that promote erythroid differentiation, such as Gfi-1b, EpoR, and Klf1.[46] Here, we demonstrated that PU.1 downregulates *MCUR1* expression in an rs61644582 allele-specific manner and then mitigates cell growth and erythroid differentiation of HSPCs in erythropoiesis via the Ca^{2+}-CAMKK2-AMPK-mTOR axis. These findings add new knowledge of mechanisms by which PU.1 functions as the well-known antagonist of erythropoiesis.

The selected beneficial allele during human local adaptation can arise from a *de novo* mutation or be previously polymorphic (i.e., standing variation) or introduced by admixture with other populations.[47] Here, based on the inferred demographic model, we showed that the rs61644582 del allele is more ancient than the beginning of its adaptation to the Qinghai-Tibetan plateau environment, suggesting a selection sweep from a standing variant in Tibetans. Thus, our findings add a new example to the list of standing variants that have been selected during the recent evolutionary history of humans, similar to those at the *PDE10A* and *BDKRB2* genes in the adaptation to diving in the Bajau people[48] and the *LCT* gene in adult lactose digestion in Europeans.[49]

In summary, we here present a genome-wide scan for novel regions/genes subject to local positive selection in Tibetans during HAA. Further, we conducted a comprehensive study on the novel top candidate, *MCUR1* at 6p23, and elucidated that MCUR1-mediated mitochondrial Ca^{2+} homeostasis is a novel regulator for erythropoiesis. The other candidates are also undoubtedly very promising for future biological studies. After determining the functional relevance of these candidates, it might have potential implications for improving the prevention and treatment of high-altitude-driven diseases.

Limitations of the study

This study still has several limitations and further investigations are required. First, whether the evolution of *MCUR1* and its causative variant(s) has occurred in other populations living at the other high-altitude areas in China (e.g., permanent residents at the Yunnan-Guizhou Plateau) and in the world (e.g., Ethiopian highlanders and the South American Andes highlanders) is not clear yet and worth further investigation. Second, additional design and assessment of exposure factors for gene-environment interactions may contribute to a better understanding of HAA, which needs to be considered in future studies. Third, given that *MCUR1* depletion impairs erythropoiesis under both normoxic and hypoxic conditions, the potential applications of MCUR1-mediated erythropoiesis may help improve the prevention and treatment of specific erythropoiesis-related disorders and high-altitude-driven diseases, such as anemia and high-altitude polycythemia, which are also worth exploration in the future.

RESOURCE AVAILABILITY

Lead contact

Requests for further information and resources and reagents should be directed to and will be fulfilled by the lead contact, Gangqiao Zhou (zhougq114@126.com).

Materials availability

This study did not generate unique reagents.

Data and code availability

Raw data of whole-genome sequencing and bulk and single-cell RNA-seq generated in this study are available in the National Genomics Data Center (NGDC)-National Center for Bioinformatics (CNCB), China, under accession no. PRJCA022898. This study did not generate any new code.

ACKNOWLEDGMENTS

We thank all the subjects participating in this study. This work was supported by grants from the National Natural Science Foundation of China (32300487,

U24A20750, 81730055, 81125017, 81222027, and 31540081), Shanghai Municipal Science and Technology Major Project (2017SHZDZX01), Beijing Institute of Radiation Medicine Innovation Fund (BIOX0105), and Open Project Program of the State Key Laboratory of Medical Proteomics (SKLP-O201510).

AUTHOR CONTRIBUTIONS

G.Z. was the principal investigator who conceived the study and obtained financial support. G.Z., J.P., and Hongxing Zhang designed the study. H.G. and J.P. were responsible for the recruitment of the TIB_discovery population. Hongxing Zhang, J.P., Ying Cui, Yongquan Cui, G.L., and F.H. were responsible for the recruitment of the HAN_discovery and HAN_replication 1 populations. J.P., C.Q., and Y. Li performed the analyses of positive selection and evolutionary history. L.K. and Haoxiang Zhang were responsible for the recruitment of the TIB_replication 1 and TIB_replication 4 populations. S.W., R.G., and J.L. were responsible for the collection of whole-blood samples of the TIB_replication 2 population from the People's Hospital at Xi'ning City, Qinghai Province. Y. Zhang, T.W., Y. Wei, and B.S. were responsible for the collection of cord blood samples of the TIB_replication 3 population with the help of Y. Wang, Y. Zhai, H. Liu, Y. He, X.Q., and Ouzhuluobu. W.Z. helped with the PBS analysis. J.P., Y. Lu, Hao Lu, Y.J., and Q.H. performed the analyses of single-cell and bulk transcriptomic data. J.P. performed SNP genotyping assays in replication stage. X.L., L.C., and Q.L. helped perform the cell culture and differentiation experiments and *in vitro* functional experiments with the help of C.W., Z.Z., M.L., S.C., X. Wei, Y. Wang, and Hui Lu. P.F., Y. Hao, X.L., L.C., H.M., L.Y., X. Wang, and W.J. performed *in vivo* functional experiments. J.P. conducted data management and statistical analyses. G.Z., Haoxiang Zhang, and J.P. interpreted the results and drafted the manuscript. G.Z. approved the final version of the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Human subjects
  - Animals
  - Cell lines
- METHOD DETAILS
  - WGS, SNP detection and data quality control
  - SNPs genotyping in replication stage
  - Phenotypic analyses of mice
  - scRNA-seq of BM cells from *Mcur1-WT and Mcur1-cKO* mice
  - siRNAs and shRNAs
  - Construction of MCUR1 overexpression vector
  - qRT-PCR assays
  - Western blotting assays
  - Flow cytometry analyses of human $CD34^+$ HSPCs
  - Benzidine staining assays
  - Cells growth assays
  - Cell cycle and cell apoptosis assays
  - BFU-E and CFU-E colony-forming assays for human $CD34^+$ HSPCs
  - Cytospin assays
  - Polysome profiling assays
  - Measurements of the mitochondrial bioenergetics
  - ChIP assays
  - EMSA and WEMSA
  - Luciferase reporter gene assays
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Population genetic structure analyses
  - Genetic drift sharing analyses
  - Coalescent simulations
  - Natural selection signals
  - Estimation of the mutation age of rs61644582
  - Estimation of strength and time of selection signals surrounding *MCUR1*, *EGLN1* and *EPAS1* loci
  - Expression enrichment analyses of HAA-associated genes
  - Pathway enrichment analyses of HAA-associated genes
  - Associations between rs61644582 and mRNA levels of nearby genes, altitudes and erythropoiesis-related phenotypes
  - Publicly available *MCUR1* expressions in multiple tissues and cell lines
  - Publicly available *MCUR1*-associated phenotypes prediction
  - RNA-seq and GSEA analyses
  - Functional annotation of SNPs at 6p23 locus
  - Other statistical analyses

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.xgen.2025.100782.

Received: April 15, 2024
Revised: September 3, 2024
Accepted: February 3, 2025
Published: March 4, 2025

REFERENCES

1. Mathioudakis, A.G., Chatzimavridou-Grigoriadou, V., Corlateanu, A., and Vestbo, J. (2017). Procalcitonin to guide antibiotic administration in COPD exacerbations: a meta-analysis. Eur. Respir. Rev. *26*, 160073. https://doi.org/10.1183/16000617.0073-2016.
2. He, Y., Zheng, W., Guo, Y., Yue, T., Cui, C., Wu, T., Ouzhuluobu, Zhang, H., Zhang, H., Liu, K., et al. (2023). Deep phenotyping of 11,880 highlanders reveals novel adaptive traits in native Tibetans. iScience *26*, 107677. https://doi.org/10.1016/j.isci.2023.107677.
3. Yi, X., Liang, Y., Huerta-Sanchez, E., Jin, X., Cuo, Z.X.P., Pool, J.E., Xu, X., Jiang, H., Vinckenbosch, N., Korneliussen, T.S., et al. (2010). Sequencing of 50 human exomes reveals adaptation to high altitude. Science *329*, 75–78. https://doi.org/10.1126/science.1190371.
4. Peng, Y., Yang, Z., Zhang, H., Cui, C., Qi, X., Luo, X., Tao, X., Wu, T., Ouzhuluobu, B., Basang, et al. (2011). Genetic variations in Tibetan populations and high-altitude adaptation at the Himalayas. Mol. Biol. Evol. *28*, 1075–1081. https://doi.org/10.1093/molbev/msq290.
5. Simonson, T.S., Yang, Y., Huff, C.D., Yun, H., Qin, G., Witherspoon, D.J., Bai, Z., Lorenzo, F.R., Xing, J., Jorde, L.B., et al. (2010). Genetic evidence for high-altitude adaptation in Tibet. Science *329*, 72–75. https://doi.org/10.1126/science.1189406.
6. Yang, J., Jin, Z.B., Chen, J., Huang, X.F., Li, X.M., Liang, Y.B., Mao, J.Y., Chen, X., Zheng, Z., Bakshi, A., et al. (2017). Genetic signatures of high-altitude adaptation in Tibetans. Proc. Natl. Acad. Sci. USA *114*, 4189–4194. https://doi.org/10.1073/pnas.1617042114.
7. Lou, H., Lu, Y., Lu, D., Fu, R., Wang, X., Feng, Q., Wu, S., Yang, Y., Li, S., Kang, L., et al. (2015). A 3.4-kb Copy-Number Deletion near EPAS1 Is Significantly Enriched in High-Altitude Tibetans but Absent from the Denisovan Sequence. Am. J. Hum. Genet. *97*, 54–66. https://doi.org/10.1016/j.ajhg.2015.05.005.
8. Lorenzo, F.R., Huff, C., Myllymäki, M., Olenchock, B., Swierczek, S., Tashi, T., Gordeuk, V., Wuren, T., Ri-Li, G., McClain, D.A., et al. (2014). A genetic mechanism for Tibetan high-altitude adaptation. Nat. Genet. *46*, 951–956. https://doi.org/10.1038/ng.3067.
9. McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., and DePristo, M.A. (2010). The Genome Analysis Toolkit: a MapReduce framework for

analyzing next-generation DNA sequencing data. Genome Res. *20*, 1297–1303. https://doi.org/10.1101/gr.107524.110.

10. Xu, S., Li, S., Yang, Y., Tan, J., Lou, H., Jin, W., Yang, L., Pan, X., Wang, J., Shen, Y., et al. (2011). A genome-wide search for signals of high-altitude adaptation in Tibetans. Mol. Biol. Evol. *28*, 1003–1011. https://doi.org/10.1093/molbev/msq277.
11. Wang, M.S., Li, Y., Peng, M.S., Zhong, L., Wang, Z.J., Li, Q.Y., Tu, X.L., Dong, Y., Zhu, C.L., Wang, L., et al. (2015). Genomic Analyses Reveal Potential Independent Adaptation to High Altitude in Tibetan Chickens. Mol. Biol. Evol. *32*, 1880–1889. https://doi.org/10.1093/molbev/msv071.
12. Wang, M., Huang, X., Li, R., Xu, H., Jin, L., and He, Y. (2014). Detecting recent positive selection with high accuracy and reliability by conditional coalescent tree. Mol. Biol. Evol. *31*, 3068–3080. https://doi.org/10.1093/molbev/msu244.
13. Garud, N.R., Messer, P.W., and Petrov, D.A. (2021). Detection of hard and soft selective sweeps from Drosophila melanogaster population genomic data. PLoS Genet. *17*, e1009373. https://doi.org/10.1371/journal.pgen.1009373.
14. Garud, N.R., Messer, P.W., Buzbas, E.O., and Petrov, D.A. (2015). Recent selective sweeps in North American Drosophila melanogaster show signatures of soft sweeps. PLoS Genet. *11*, e1005004. https://doi.org/10.1371/journal.pgen.1005004.
15. Albers, P.K., and McVean, G. (2020). Dating genomic variants and shared ancestry in population-scale sequencing data. PLoS Biol. *18*, e3000586. https://doi.org/10.1371/journal.pbio.3000586.
16. Kern, A.D., and Schrider, D.R. (2018). diploS/HIC: An Updated Approach to Classifying Selective Sweeps. G3 (Bethesda) *8*, 1959–1970. https://doi.org/10.1534/g3.118.200262.
17. Jeong, C., Witonsky, D.B., Basnyat, B., Neupane, M., Beall, C.M., Childs, G., Craig, S.R., Novembre, J., and Di Rienzo, A. (2018). Detecting past and ongoing natural selection among ethnically Tibetan women at high altitude in Nepal. PLoS Genet. *14*, e1007650. https://doi.org/10.1371/journal.pgen.1007650.
18. Zheng, W., He, Y., Guo, Y., Yue, T., Zhang, H., Li, J., Zhou, B., Zeng, X., Li, L., Wang, B., et al. (2023). Large-scale genome sequencing redefines the genetic footprints of high-altitude adaptation in Tibetans. Genome Biol. *24*, 73. https://doi.org/10.1186/s13059-023-02912-1.
19. Mallilankaraman, K., Cárdenas, C., Doonan, P.J., Chandramoorthy, H.C., Irrinki, K.M., Golenár, T., Csordás, G., Madireddi, P., Yang, J., Müller, M., et al. (2012). MCUR1 is an essential component of mitochondrial Ca2+ uptake that regulates cellular metabolism. Nat. Cell Biol. *14*, 1336–1343. https://doi.org/10.1038/ncb2622.
20. Novershtern, N., Subramanian, A., Lawton, L.N., Mak, R.H., Haining, W.N., McConkey, M.E., Habib, N., Yosef, N., Chang, C.Y., Shay, T., et al. (2011). Densely interconnected transcriptional circuits control cell states in human hematopoiesis. Cell *144*, 296–309. https://doi.org/10.1016/j.cell.2011.01.004.
21. He, S., Wang, L.H., Liu, Y., Li, Y.Q., Chen, H.T., Xu, J.H., Peng, W., Lin, G.W., Wei, P.P., Li, B., et al. (2020). Single-cell transcriptome profiling of an adult human cell atlas of 15 major organs. Genome Biol. *21*, 294. https://doi.org/10.1186/s13059-020-02210-0.
22. Hay, S.B., Ferchen, K., Chetal, K., Grimes, H.L., and Salomonis, N. (2018). The Human Cell Atlas bone marrow single-cell interactive web portal. Exp. Hematol. *68*, 51–61. https://doi.org/10.1016/j.exphem.2018.09.004.
23. Tusi, B.K., Wolock, S.L., Weinreb, C., Hwang, Y., Hidalgo, D., Zilionis, R., Waisman, A., Huh, J.R., Klein, A.M., and Socolovsky, M. (2018). Population snapshots predict early haematopoietic and erythroid hierarchies. Nature *555*, 54–60. https://doi.org/10.1038/nature25741.
24. Lachmann, A., Torre, D., Keenan, A.B., Jagodnik, K.M., Lee, H.J., Wang, L., Silverstein, M.C., and Ma'ayan, A. (2018). Massive mining of publicly available RNA-seq data from human and mouse. Nat. Commun. *9*, 1366. https://doi.org/10.1038/s41467-018-03751-6.
25. Knight, Z.A., Schmidt, S.F., Birsoy, K., Tan, K., and Friedman, J.M. (2014). A critical role for mTORC1 in erythropoiesis and anemia. Elife *3*, e01913. https://doi.org/10.7554/eLife.01913.
26. Malik, N., Dunn, K.M., Cassels, J., Hay, J., Estell, C., Sansom, O.J., and Michie, A.M. (2019). mTORC1 activity is essential for erythropoiesis and B cell lineage commitment. Sci. Rep. *9*, 16917. https://doi.org/10.1038/s41598-019-53141-1.
27. Liu, X., Zhang, Y., Ni, M., Cao, H., Signer, R.A.J., Li, D., Li, M., Gu, Z., Hu, Z., Dickerson, K.E., et al. (2017). Regulation of mitochondrial biogenesis in erythropoiesis by mTORC1-mediated protein translation. Nat. Cell Biol. *19*, 626–638. https://doi.org/10.1038/ncb3527.
28. Garcia, D., and Shaw, R.J. (2017). AMPK: Mechanisms of Cellular Energy Sensing and Restoration of Metabolic Balance. Mol. Cell. *66*, 789–800. https://doi.org/10.1016/j.molcel.2017.05.032.
29. Zhu, H., Foretz, M., Xie, Z., Zhang, M., Zhu, Z., Xing, J., Leclerc, J., Gaudry, M., Viollet, B., and Zou, M.H. (2014). PRKAA1/AMPKalpha1 is required for autophagy-dependent mitochondrial clearance during erythrocyte maturation. Autophagy *10*, 1522–1534. https://doi.org/10.4161/auto.29197.
30. Evans, A.M., Mahmoud, A.D., Moral-Sanz, J., and Hartmann, S. (2016). The emerging role of AMPK in the regulation of breathing and oxygen supply. Biochem. J. *473*, 2561–2572. https://doi.org/10.1042/BCJ20160002.
31. Zhao, H., Li, T., Wang, K., Zhao, F., Chen, J., Xu, G., Zhao, J., Li, T., Chen, L., Li, L., et al. (2019). AMPK-mediated activation of MCU stimulates mitochondrial Ca(2+) entry to promote mitotic progression. Nat. Cell Biol. *21*, 476–486. https://doi.org/10.1038/s41556-019-0296-3.
32. Fang, X., Shen, F., Lechauve, C., Xu, P., Zhao, G., Itkow, J., Wu, F., Hou, Y., Wu, X., Yu, L., et al. (2018). miR-144/451 represses the LKB1/AMPK/mTOR pathway to promote red cell precursor survival during recovery from acute anemia. Haematologica *103*, 406–416. https://doi.org/10.3324/haematol.2017.177394.
33. Shimobayashi, M., and Hall, M.N. (2014). Making new contacts: the mTOR network in metabolism and signalling crosstalk. Nat. Rev. Mol. Cell Biol. *15*, 155–162. https://doi.org/10.1038/nrm3757.
34. Kishton, R.J., Barnes, C.E., Nichols, A.G., Cohen, S., Gerriets, V.A., Siska, P.J., Macintyre, A.N., Goraksha-Hicks, P., de Cubas, A.A., Liu, T., et al. (2016). AMPK Is Essential to Balance Glycolysis and Mitochondrial Metabolism to Control T-ALL Cell Stress and Survival. Cell Metab. *23*, 649–662. https://doi.org/10.1016/j.cmet.2016.03.008.
35. Perez-Liebana, I., Juaristi, I., Gonzalez-Sanchez, P., Gonzalez-Moreno, L., Rial, E., Podunavac, M., Zakarian, A., Molgo, J., Vallejo-Illarramendi, A., Mosqueira-Martin, L., et al. (2022). A Ca(2+)-Dependent Mechanism Boosting Glycolysis and OXPHOS by Activating Aralar-Malate-Aspartate Shuttle, upon Neuronal Stimulation. J. Neurosci. *42*, 3879–3895. https://doi.org/10.1523/JNEUROSCI.1463-21.2022.
36. Dispirito, J.R., Fang, B., Wang, F., and Lazar, M.A. (2013). Pruning of the adipocyte peroxisome proliferator-activated receptor gamma cistrome by hematopoietic master regulator PU.1. Mol. Cell Biol. *33*, 3354–3364. https://doi.org/10.1128/MCB.00599-13.
37. Rothenberg, E.V., Hosokawa, H., and Ungerbäck, J. (2019). Mechanisms of Action of Hematopoietic Transcription Factor PU.1 in Initiation of T-Cell Development. Front. Immunol. *10*, 228. https://doi.org/10.3389/fimmu.2019.00228.
38. Gottgens, B., Broccardo, C., Sanchez, M.J., Deveaux, S., Murphy, G., Gothert, J.R., Kotsopoulou, E., Kinston, S., Delaney, L., Piltz, S., et al. (2004). The scl +18/19 stem cell enhancer is not required for hematopoiesis: identification of a 5' bifunctional hematopoietic-endothelial enhancer bound by Fli-1 and Elf-1. Mol. Cell Biol. *24*, 1870–1883. https://doi.org/10.1128/mcb.24.5.1870-1883.2004.
39. Soldner, F., Stelzer, Y., Shivalila, C.S., Abraham, B.J., Latourelle, J.C., Barrasa, M.I., Goldmann, J., Myers, R.H., Young, R.A., and Jaenisch, R. (2016). Parkinson-associated risk variant in distal enhancer of alpha-synuclein modulates target gene expression. Nature *533*, 95–99. https://doi.org/10.1038/nature17939.

40. Fan, M., Zhang, J., Tsai, C.-W., Orlando, B.J., Rodriguez, M., Xu, Y., Liao, M., Tsai, M.-F., and Feng, L. (2020). Structure and mechanism of the mitochondrial Ca2+ uniporter holocomplex. Nature *582*, 129–133. https://doi.org/10.1038/s41586-020-2309-6.

41. Scheinfeldt, L.B., Soi, S., Thompson, S., Ranciaro, A., Woldemeskel, D., Beggs, W., Lambert, C., Jarvis, J.P., Abate, D., Belay, G., and Tishkoff, S.A. (2012). Genetic adaptation to high altitude in the Ethiopian highlands. Genome Biol. *13*, R1. https://doi.org/10.1186/gb-2012-13-1-r1.

42. Patron, M., Checchetto, V., Raffaello, A., Teardo, E., Vecellio Reane, D., Mantoan, M., Granatiero, V., Szabò, I., De Stefani, D., and Rizzuto, R. (2014). MICU1 and MICU2 finely tune the mitochondrial Ca2+ uniporter by exerting opposite effects on MCU activity. Mol. Cell. *53*, 726–737. https://doi.org/10.1016/j.molcel.2014.01.013.

43. Liu, H., Zhang, Y., Wu, H., D'Alessandro, A., Yegutkin, G.G., Song, A., Sun, K., Li, J., Cheng, N.Y., Huang, A., et al. (2016). Beneficial Role of Erythrocyte Adenosine A2B Receptor-Mediated AMP-Activated Protein Kinase Activation in High-Altitude Hypoxia. Circulation *134*, 405–421. https://doi.org/10.1161/CIRCULATIONAHA.116.021311.

44. Ladli, M., Richard, C., Aguilar, L.C., Ducamp, S., Bondu, S., Sujobert, P., Tamburini, J., Lacombe, C., Azar, N., Foretz, M., et al. (2019). Finely-tuned regulation of AMP-activated protein kinase (AMPK) is crucial for human adult erythropoiesis. Haematologica *104*, 907–918. https://doi.org/10.3324/haematol.2018.191403.

45. Peng, Y., Cui, C., Yang, L., He, Y., Ouzhuluobu, Zhang, H., Yang, D., Bianbazhuoma, Zhang, Q., He, Y., et al. (2017). Down-Regulation of EPAS1 Transcription and Genetic Adaptation of Tibetans to High-Altitude Hypoxia. Mol. Biol. Evol. *34*, 818–830. https://doi.org/10.1093/molbev/msw280.

46. Wontakal, S.N., Guo, X., Will, B., Shi, M., Raha, D., Mahajan, M.C., Weissman, S., Snyder, M., Steidl, U., Zheng, D., and Skoultchi, A.I. (2011). A large gene network in immature erythroid cells is controlled by the myeloid and B cell transcriptional regulator PU.1. PLoS Genet. *7*, e1001392. https://doi.org/10.1371/journal.pgen.1001392.

47. Rees, J.S., Castellano, S., and Andrés, A.M. (2020). The Genomics of Human Local Adaptation. Trends Genet. *36*, 415–428. https://doi.org/10.1016/j.tig.2020.03.006.

48. Ilardo, M.A., Moltke, I., Korneliussen, T.S., Cheng, J., Stern, A.J., Racimo, F., de Barros Damgaard, P., Sikora, M., Seguin-Orlando, A., Rasmussen, S., et al. (2018). Physiological and Genetic Adaptations to Diving in Sea Nomads. Cell *173*, 569–580.e15. https://doi.org/10.1016/j.cell.2018.03.054.

49. Myles, S., Bouzekri, N., Haverfield, E., Cherkaoui, M., Dugoujon, J.M., and Ward, R. (2005). Genetic evidence in support of a shared Eurasian-North African dairying origin. Hum. Genet. *117*, 34–42. https://doi.org/10.1007/s00439-005-1266-3.

50. Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics *26*, 589–595. https://doi.org/10.1093/bioinformatics/btp698.

51. Patterson, N., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y., Genschoreck, T., Webster, T., and Reich, D. (2012). Ancient admixture in human history. Genetics *192*, 1065–1093. https://doi.org/10.1534/genetics.112.145037.

52. Alexander, D.H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. Genome Res. *19*, 1655–1664. https://doi.org/10.1101/gr.094052.109.

53. Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. USA *102*, 15545–15550. https://doi.org/10.1073/pnas.0506580102.

54. Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M., and Lee, J.J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience *4*, 7. https://doi.org/10.1186/s13742-015-0047-8.

55. Pers, T.H., Karjalainen, J.M., Chan, Y., Westra, H.J., Wood, A.R., Yang, J., Lui, J.C., Vedantam, S., Gustafsson, S., Esko, T., et al. (2015). Biological interpretation of genome-wide association studies using predicted gene functions. Nat. Commun. *6*, 5890. https://doi.org/10.1038/ncomms6890.

56. Reimand, J., Arak, T., Adler, P., Kolberg, L., Reisberg, S., Peterson, H., and Vilo, J. (2016). g:Profiler-a web server for functional interpretation of gene lists (2016 update). Nucleic Acids Res. *44*, W83–W89. https://doi.org/10.1093/nar/gkw199.

57. Straube, J., Eifert, T., Vu, T., Janardhanan, Y., Haldar, R., von Eyss, B., Cooper, L., Bruedigam, C., Ling, V.Y., Cooper, E., et al. (2023). Cre recombinase expression cooperates with homozygous FLT3 internal tandem duplication knockin mouse model to induce acute myeloid leukemia. Leukemia *37*, 741–750. https://doi.org/10.1038/s41375-023-01832-0.

58. Herrejon Chavez, F., Luo, H., Cifani, P., Pine, A., Chu, E.L., Joshi, S., Barin, E., Schurer, A., Chan, M., Chang, K., et al. (2023). RNA binding protein SYNCRIP maintains proteostasis and self-renewal of hematopoietic stem and progenitor cells. Nat. Commun. *14*, 2290. https://doi.org/10.1038/s41467-023-38001-x.

59. Kurita, R., Suda, N., Sudo, K., Miharada, K., Hiroyama, T., Miyoshi, H., Tani, K., and Nakamura, Y. (2013). Establishment of immortalized human erythroid progenitor cell lines able to produce enucleated red blood cells. PLoS One *8*, e59890. https://doi.org/10.1371/journal.pone.0059890.

60. Lu, D., Lou, H., Yuan, K., Wang, X., Wang, Y., Zhang, C., Lu, Y., Yang, X., Deng, L., Zhou, Y., et al. (2016). Ancestral Origins and Genetic History of Tibetan Highlanders. Am. J. Hum. Genet. *99*, 580–594. https://doi.org/10.1016/j.ajhg.2016.07.002.

61. Fan, J., Guang, H., Zhang, H., Chen, D., Ding, L., Fan, X., Xue, F., Gan, Z., Wang, Y., Mao, S., et al. (2018). SIRT1 Mediates Apelin-13 in Ameliorating Chronic Normobaric Hypoxia-induced Anxiety-like Behavior by Suppressing NF-kappaB Pathway in Mice Hippocampus. Neuroscience *381*, 22–34. https://doi.org/10.1016/j.neuroscience.2018.04.013.

62. Ferrari, M., Jain, I.H., Goldberger, O., Rezoagli, E., Thoonen, R., Cheng, K.H., Sosnovik, D.E., Scherrer-Crosbie, M., Mootha, V.K., and Zapol, W.M. (2017). Hypoxia treatment reverses neurodegenerative disease in a mouse model of Leigh syndrome. Proc. Natl. Acad. Sci. USA *114*, E4241–E4250. https://doi.org/10.1073/pnas.1621511114.

63. Ema, H., Morita, Y., Yamazaki, S., Matsubara, A., Seita, J., Tadokoro, Y., Kondo, H., Takano, H., and Nakauchi, H. (2006). Adult mouse hematopoietic stem cells: purification and single-cell assays. Nat. Protoc. *1*, 2979–2987. https://doi.org/10.1038/nprot.2006.447.

64. Dong, F., Hao, S., Zhang, S., Zhu, C., Cheng, H., Yang, Z., Hamey, F.K., Wang, X., Gao, A., Wang, F., et al. (2020). Differentiation of transplanted haematopoietic stem cells tracked by single-cell transcriptomic analysis. Nat. Cell Biol. *22*, 630–639. https://doi.org/10.1038/s41556-020-0512-1.

65. Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., 3rd, Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. Cell *184*, 3573–3587.e29. https://doi.org/10.1016/j.cell.2021.04.048.

66. Qiu, X., Mao, Q., Tang, Y., Wang, L., Chawla, R., Pliner, H.A., and Trapnell, C. (2017). Reversed graph embedding resolves complex single-cell trajectories. Nat. Methods *14*, 979–982. https://doi.org/10.1038/nmeth.4402.

67. Dover, G.J., Chan, T., and Sieber, F. (1983). Fetal hemoglobin production in cultures of primitive and mature human erythroid progenitors: differentiation affects the quantity of fetal hemoglobin produced per fetal-hemoglobin-containing cell. Blood *61*, 1242–1246.

68. Mukhopadhyay, P., Rajesh, M., Haskó, G., Hawkins, B.J., Madesh, M., and Pacher, P. (2007). Simultaneous detection of apoptosis and mitochondrial superoxide production in live cells by flow cytometry and confocal microscopy. Nat. Protoc. *2*, 2295–2301. https://doi.org/10.1038/nprot.2007.327.

69. Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A., and Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. Nat. Genet. *38*, 904–909. https://doi.org/10.1038/ng1847.

70. Quan, C., Li, Y., Liu, X., Wang, Y., Ping, J., Lu, Y., and Zhou, G. (2021). Characterization of structural variation in Tibetans reveals new evidence of high-altitude adaptation and introgression. Genome Biol. *22*, 159. https://doi.org/10.1186/s13059-021-02382-3.

71. Gutenkunst, R.N., Hernandez, R.D., Williamson, S.H., and Bustamante, C.D. (2009). Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. PLoS Genet. *5*, e1000695. https://doi.org/10.1371/journal.pgen.1000695.

72. Scally, A. (2016). The mutation rate in human evolution and demographic inference. Curr. Opin. Genet. Dev. *41*, 36–43. https://doi.org/10.1016/j.gde.2016.07.008.

73. Hsieh, P., Vollger, M.R., Dang, V., Porubsky, D., Baker, C., Cantsilieris, S., Hoekzema, K., Lewis, A.P., Munson, K.M., Sorensen, M., et al. (2019). Adaptive archaic introgression of copy number variants and the discovery of previously unknown human genes. Science *366*, eaax2083. https://doi.org/10.1126/science.aax2083.

74. Portik, D.M., Leaché, A.D., Rivera, D., Barej, M.F., Burger, M., Hirschfeld, M., Rödel, M.O., Blackburn, D.C., and Fujita, M.K. (2017). Evaluating mechanisms of diversification in a Guineo-Congolian tropical forest frog using demographic model selection. Mol. Ecol. *26*, 5245–5263. https://doi.org/10.1111/mec.14266.

75. Kelleher, J., Etheridge, A.M., and McVean, G. (2016). Efficient Coalescent Simulation and Genealogical Analysis for Large Sample Sizes. PLoS Comput. Biol. *12*, e1004842. https://doi.org/10.1371/journal.pcbi.1004842.

76. Holsinger, K.E., and Weir, B.S. (2009). Genetics in geographically structured populations: defining, estimating and interpreting F(ST). Nat. Rev. Genet. *10*, 639–650. https://doi.org/10.1038/nrg2611.

77. Gormley, P., Anttila, V., Winsvold, B.S., Palta, P., Esko, T., Pers, T.H., Farh, K.H., Cuenca-Leon, E., Muona, M., Furlotte, N.A., et al. (2016). Meta-analysis of 375,000 individuals identifies 38 susceptibility loci for migraine. Nat. Genet. *48*, 856–866. https://doi.org/10.1038/ng.3598.

78. GTEx Consortium (2013). The Genotype-Tissue Expression (GTEx) project. Nat. Genet. *45*, 580–585. https://doi.org/10.1038/ng.2653.

79. Yang, D., Jang, I., Choi, J., Kim, M.S., Lee, A.J., Kim, H., Eom, J., Kim, D., Jung, I., and Lee, B. (2018). 3DIV: A 3D-genome Interaction Viewer and database. Nucleic Acids Res. *46*, D52–D57. https://doi.org/10.1093/nar/gkx1017.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit monoclonal anti-4EBP1	Cell Signaling Technology	Cat#9644; RRID: AB_2097841
Rabbit monoclonal anti-AMPKα	Cell Signaling Technology	Cat#5831; RRID: AB_10622186
Rabbit monoclonal anti-CAMKK2	Cell Signaling Technology	Cat#16810; RRID: AB_2798771
BV421 anti-Human-CD34 (561)	BioLegend	Cat#343609; RRID: AB_11147951
PE anti-Human-CD36 (CB38)	BD Biosciences	Cat#555455; RRID: AB_395848
PE-cy7 anti-Human-CD71 (OKT9)	Thermo Fisher Scientific	Cat#25-0719-42; RRID: AB_2573366
APC anti-Human-GPA (HIR2)	Thermo Fisher Scientific	Cat#17-9987-42; RRID: AB_2043823
PE-cy7 anti-Human-IL3R (6H6)	Thermo Fisher Scientific	Cat#25-1239-42; RRID: AB_1257136
Mouse monoclonal anti-ELF1	Santa Cruz Biotechnology	Cat#sc-373772; RRID: AB_10916710
Rabbit polyclonal anti-Human MCUR1	Sigma-Aldrich	Cat#SAB2100356; RRID: AB_10604334
Rabbit monoclonal anti-LC3A/B	Cell Signaling Technology	Cat#12741; RRID: AB_2617131
Rabbit monoclonal anti-LKB1	Cell Signaling Technology	Cat#3047; RRID: AB_2198327
Rabbit polyclonal anti-Mouse MCUR1	Sigma-Aldrich	Cat#SAB2700722; RRID: AB_3105939
Rabbit monoclonal anti-mTOR	Cell Signaling Technology	Cat#2983; RRID: AB_2105622
Rabbit polyclonal anti-NRF1	Proteintech	Cat#12482-1-AP; RRID: AB_2282876
Rabbit polyclonal anti-p62	Proteintech	Cat#55274-1-AP; RRID: AB_11182278
Rabbit polyclonal anti-p70 S6 kinase	Cell Signaling Technology	Cat#9202; RRID: AB_331676
Mouse monoclonal anti-PGC1α	Proteintech	Cat#66369-1-Ig; RRID: AB_2828002
Rabbit monoclonal anti-phospho-4EBP1 (Thr37/46)	Cell Signaling Technology	Cat#2855; RRID: AB_560835
Rabbit monoclonal anti-phospho-AMPKα (Thr172)	Cell Signaling Technology	Cat#2535; RRID: AB_331250
Rabbit polyclonal anti-phospho-CAMKK2 (Ser495)	Cell Signaling Technology	Cat#16737; RRID: AB_2798769
Rabbit monoclonal anti-phospho-LKB1 (Ser428)	Cell Signaling Technology	Cat#3482; RRID: AB_2198321
Rabbit monoclonal anti-phospho-mTOR (Ser2448)	Cell Signaling Technology	Cat#5536; RRID: AB_10691552
Rabbit polyclonal anti-phospho-p70 S6 kinase (Thr389)	Cell Signaling Technology	Cat#9205; RRID: AB_330944
Rabbit monoclonal anti-phospho-Raptor (Ser792)	Cell Signaling Technology	Cat#89146; RRID: AB_2934061
Rabbit monoclonal anti-phospho-TSC2 (Ser1387)	Cell Signaling Technology	Cat#23402; RRID: AB_2798864
Rabbit monoclonal anti-phospho-ULK1 (Ser757)	Cell Signaling Technology	Cat#14202; RRID: AB_2665508
Rabbit monoclonal anti-PU.1	Cell Signaling Technology	Cat#2258; RRID: AB_2186909
Rabbit polyclonal anti-Raptor	Proteintech	Cat#20984-1-AP; RRID: AB_11182390
Rabbit polyclonal anti-mtTFA	Proteintech	Cat#19998-1-AP; RRID: AB_10638429
Rabbit monoclonal anti-TSC1	Cell Signaling Technology	Cat#6935; RRID: AB_10860420
Rabbit monoclonal anti-TSC2	Cell Signaling Technology	Cat#4308; RRID: AB_10547134
Rabbit polyclonal anti-ULK1	Proteintech	Cat#20986-1-AP; RRID: AB_2878783
Mouse monoclonal anti-YY1	Proteintech	Cat#66281-1-Ig; RRID:AB_2737053
Mouse monoclonal anti-β-actin	Proteintech	Cat#60008-1-Ig; RRID:AB_2289225
Goat anti-Mouse IgG, HRP conjugated	CWBIO	Cat#CW0102; RRID: AB_2814710

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Goat anti-Rabbit IgG, HRP conjugated	CWBIO	Cat#CW0103; RRID: AB_2814709
Chemicals, peptides, and recombinant proteins		
BAPTA-AM	Thermo Fisher Scientific	Cat#B6769
Benzidine	Sigma-Aldrich	Cat#B3383
Compound C	Selleck	Cat#S7840
Dihydrorhodamine 123	Thermo Fisher Scientific	Cat#D632
Dulbecco's modified Eagle's medium	HyClone-GE	Cat#SH30022.01B
EDTA-free protease inhibitor cocktail	Roche	Cat#04693132001
EPO	R&D Systems	Cat#287-TC
Fetal bovine serum	HyClone-GE	Cat#SH30071.04
Flt3-Ligand	Peprotech	Cat#300-19
Fluo-4	Thermo Fisher Scientific	Cat#F23917
IL-3	Peprotech	Cat#200-03
Ionomycin	Thermo Fisher Scientific	Cat#124222
MitoSox Red	Thermo Fisher Scientific	Cat#M36008
Penicillin and Streptomycin	Wisent	Cat#450-201-EL
Phosphatase Inhibitor Cocktail	Sigma-Aldrich	Cat#P2850
Poly-L-lysine	Sigma-Aldrich	Cat#P4707
Puromycin	Sigma-Aldrich	Cat#P9620
Recombinant ELF1 protein	Abnova	Cat#H00001997-P01
Recombinant PU.1 protein	Abnova	Cat#H00006688-P01
Rhod-2	Thermo Fisher Scientific	Cat#R1245MP
RIPA buffer	EMD-Millipore	Cat#20-188
RPMI 1640	Gibco	Cat#A1049101
SCF	Peprotech	Cat#300-07
Serum-free medium	Stem Cell Technologies	Cat#09650
STO-609	Selleck	Cat#S8274
Tetramethylrhodamine, methyl ester	Thermo Fisher Scientific	Cat#T668
TPO	Peprotech	Cat#300-18
Critical commercial assays		
AMP-Glo luminescent assay kit	Promega	Cat#V5011
APC-conjugated Annexin V/7-AAD Apoptosis Detection Kit	Thermo Fisher Scientific	Cat#88-8007-74
Blood DNA extraction kit	Qiagen	Cat#51106
CellTiter-Glo luminescent assay kit	Promega	Cat#G7570
Chromium Single Cell 3' Library, Gel Bead & Multiplex Kit and Chip Kit	10× Genomics	Cat#CG000204
Dual-Luciferase® Reporter Assay System	Promega	Cat#E1910
High-capacity cDNA reverse transcription kit	Life Technologies	Cat#4368814
IQ SYBR Green Supermix kit	Bio-Rad	Cat#170-8862
LightShift chemiluminescent electrophoresis mobility shift assay kit	Thermo Fisher Scientific	Cat#20148
X-tremegene HP DNA Transfection Reagent	Roche	Cat#6366244001
Magna ChIP™ A/G One-Day Chromatin Immunoprecipitation Kit	Millipore	Cat#17-10086
QuickChange site-directed mutagenesis kit	Stratagene	Cat#200513
riboFECT Transfection Kit	Ribobio	Cat#C10511-1

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
SuperSignal™ West Pico Chemiluminescent Substrate Kit	Thermo Fisher Scientific	Cat#34080
XF Cell Mito stress kit	Seahorse Bioscience	Cat#103015-100
Mouse Erythropoietin ELISA kit	R&D Systems	Cat#MEP00B
RNeasy mini kit	Qiagen	Cat#74104
QIAquick PCR purification kit	Qiagen	Cat#28006
Deposited data		
Raw data of whole genome sequencing, bulk RNA sequencing, single-cell RNA sequencing	This paper	NGDC-CNCB ID: PRJCA022898 (https://ngdc.cncb.ac.cn/gsa/s/f72vu02z)
Human reference genome build hg19	UCSC	https://genome.ucsc.edu/cgi-bin/hgTracks?db=hg19
1000 Genomes Project	The International Genome Sample Resource	http://www.internationalgenome.org/data
SNP datasets from 3,008 Tibetans in Tibet and 7,284 subjects of East Asian ancestry	Yang et al.[6]	http://cnsgenomics.com/data/yang_et_al_2017_pnas.html
SNP datasets from 1,104 Tibetans	Jeong et al.[17]	NCBI-SRA ID: PRJNA420511 (https://www.ncbi.nlm.nih.gov/sra/?term=PRJNA420511)
SNP datasets from 1,001 Tibetans	Zheng et al.[18]	Apply to the corresponding author
GTEx gene expression	The Genotype-Tissue Expression (GTEx) project	https://gtexportal.org/home/datasets
ENCODE	National Human Genome Research Institute	https://www.encodeproject.org
Experimental models: Cell lines		
HUDEP2	RIKEN BioResource Center	RCB4557
HEK293T	CCTCC	GDC0187
K562	CCTCC	GDC0037
Experimental models: Organisms/strains		
Mcur1-KO mice	This paper	N/A
Mcur1-cKO mice	This paper	N/A
Oligonucleotides		
Primers for SNP genotyping, see Table S15	This paper	N/A
Primers for real-time quantitative PCR, see Table S15	This paper	N/A
Primers and probes for ChIP-PCR/qPCR/AS-qPCR and EMSA assays, see Table S15	This paper	N/A
siRNAs and shRNAs, see Table S15	This paper	N/A
Primers for gene clone, see Table S15	This paper	N/A
Recombinant DNA		
Plasmid: pAD-MCS-IRES-mcherry-MCUR1	This paper	N/A
Plasmid: pGL3-promoter vector	Promega	Cat#E1761
Plasmid: pRL-SV40 plasmid	Promega	Cat#E2231
Software and algorithms		
BWA	Li and Durbin.[50]	http://bio-bwa.sourceforge.net/
GATK	McKenna et al.[9]	https://software.broadinstitute.org/gatk/
ADMIXTOOLS	Patterson et al.[51]	https://github.com/DReichLab/AdmixTools
ADMIXTURE	Alexander et al.[52]	https://dalexander.github.io/admixture/
GSEA	Subramanian et al.[53]	http://software.broadinstitute.org/gsea

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
PLINK	Chang et al.[54]	http://zzz.bwh.harvard.edu/plink
DEPICT	Pers et al.[55]	https://www.broadinstitute.org/depict
g:profiler	Reimand et al.[56]	http://biit.cs.ut.ee/gprofiler/index.cgi
ARCHS4	Lachmann et al.[24]	http://amp.pharm.mssm.edu/archs4/
FlowJo	TreeStar	https://www.flowjo.com/
cellSens Standard	Olympus	https://lifescience.evidentscientific.com.cn/zh/software/cellsens/
ImageJ	NIH	https://imagej.net/software/imagej/
Human genome dating server	Albers and McVean.[15]	https://human.genome.dating
R package	CRAN	https://cran.r-project.org

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Human subjects

In this study we recruited seven independent populations, which include a total of 1,263 native ethnic Tibetan individuals and 320 native ethnic Han individuals (Table S1). All the participants were self-reported unrelated subjects, which was confirmed by genetic relatedness test using the '–genome' in PLINK (v1.9).[54]

Populations at discovery stage used for scanning for positive selection signals

A total of 48 unrelated ethnic Tibetans (designated as TIB_discovery population) and 50 unrelated individuals of Han nationality (designated as HAN_discovery population) were recruited for whole-genome sequencing (WGS). These 48 TIB_discovery individuals, who live in five counties (including Batang, Daofu, Kangding, Xinduqiao and Litang) at Sichuan province located in southwest China, at > 2,600 meters (m) over sea level, were recruited during a physical examination program at community conducted from June 2010 to August 2010. The male/female ratio and the mean age (standard deviation [s.d.]) of these Tibetans are 1.50 and 35.31 (12.31) years old, respectively. These 50 HAN_discovery individuals, who live in Nanning City, Guangxi Province located in southern China, at < 100 m over sea level, were recruited in a regular physical examination from July 2010 to August 2010 at the Guangxi Cancer Hospital (Nanning City, China). The male/female ratio and the mean age (s.d.) of the Han individuals are 1.50 and 41.22 (6.39) years old, respectively. The male/female ratio and age were well-matched between the TIB_discovery and HAN_discovery populations (P = 0.84 and 0.068, respectively). The response rates for the TIB_discovery and HAN_discovery populations were both > 95%.

Populations at replication stage used for replicating the positive selection signals at 6p23 locus

An independent population, which includes a total of 672 Tibetans (designated as TIB_replication 1 population) and 270 Han nationality subjects (designated as HAN_replication 1 population) was used for genotyping the SNPs and replicating the positive selection signal at 6p23 locus that was identified at the discovery stage (index SNP rs61644582). The TIB_ replication 1 subjects were recruited during a physical examination program conducted from August 2011 to October 2012 at Lhasa City, Tibet Autonomous Region, China by collaborators from Xizang Minzu University (Xianyang City, Shaanxi Province, China). They live in Lhasa City in Tibet autonomous region located in southwest China, at average altitude ~3,600 m. The male/female ratio and the mean age (s.d.) of them are 1.42 and 39.47 (16.95) years old, respectively. The HAN_replication 1 individuals were recruited from August 2011 to October 2012 at Ya'an hospital, Ya'an county (at altitude < 400 m) in Sichuan province located in southwest China, who are all males and the mean age (s.d.) of them are 21.53 (2.71) years old. The response rates for the TIB_replication 1 and HAN_replication 1 populations were both over 95%. Additionally, to bolster the statistical power for the replication of the 6p23 signal, we collected the frequencies and F_{ST} values of SNPs at 6p23 locus from a previous study [6], which consists of 3,008 Tibetans in Tibet autonomous region and 7,284 subjects of East Asian ancestry (designated as Yang et al. study). The male/female ratio of the Tibetans is 0.59 and the mean age of them is not available, and the male/female ratio and the mean age of the East Asians are also not available in that study.

Populations used for SNP-expression association analyses

Two independent Tibetan populations (designated as TIB_replication 2 population and TIB_replication 3, respectively) were recruited for assessing whether a SNP serves as an expression quantitative trait locus (eQTL) with specific gene(s) expression levels (*i.e.*, genotype-expression association analyses). The TIB_replication 2 contains a total of 200 Tibetan peripheral blood samples, which were collected at over 4,000 m above sea level at Dari county, Qinghai province in northwest China by the Qinghai Provincial People's Hospital (Xi'ning City, China). The male/female ratio and the mean age (s.d.) of the Tibetans are 0.96 and 29.2 (6.54) years old, respectively. The TIB_replication 3 contains a total of 42 Tibetan cord blood samples, which were collected from August 2023 to September 2023 by our lab in Tibetan Fukang Hospital at Lhasa City, Tibet autonomous region. The male/female ratio of these newborns are 1.47. Their parents were born and grew up at Lhasa City, Tibet autonomous region.

Populations used for SNP-erythropoiesis-related phenotypes association analyses

Two independent Tibetan population (designated as TIB_replication 3 and TIB_replication 4, respectively) were recruited for this analysis. The TIB_replication 3 population were already described as mentioned in the aforementioned section. The TIB_replication 4 population contains 301 unrelated Tibetans, who live over 3,650 m above sea level and were recruited from April 2020 to May 2020 by our lab during a physical examination program at the community conducted in Shannan City, Tibet autonomous region. The male/female ratio and the mean age (s.d.) of these Tibetans are 0.58 and 55.2 (13.79) years old, respectively. Additionally, the HAN_replictaion 1 population and other nine publicly available datasets derived from populations of different ancestries were also used for the SNP genotype-hemoglobin (Hb) levels association analyses (Table S14).

Genomic DNA was extracted from 5 mL peripheral blood of all subjects and cord blood of TIB_replication 3 using the blood DNA extraction kit (#51106; Qiagen, Valencia, CA). Additionally, total RNAs were also extracted from peripheral blood mononuclear cells (PBMCs) of TIB_replication 2 subjects and cord blood of TIB_replication 3 using TRIZOL Reagent (Invitrogen, CA, USA). The cord blood (from TIB_replication 3) and peripheral blood (from TIB_replication 4 and HAN_replictaion 1) were used for measurement of the erythropoiesis-related phenotypes, including RBC count and Hb levels, *etc*.

The protocol of this study was reviewed and approved by the Ethical Committees of Beijing Institute of Radiation Medicine (approval ID: AF/SC-08/02.150) and Peking University Third Hospital Medical Science Research (approval ID: M2023666), as well as the Internal Review Board of Kunming Institute of Zoology, Chinese Academy of Sciences (approval ID: SMKX-SQ-20200414-083-02). The research scheme is in accordance with the Regulations of the People's Republic of China on the Administration of Human Genetic Resources. Written informed consent was obtained from each participant, and personal information on demographic and other factors were collected by structured questionnaire.

Animals

The *Mcur1* whole body knockout (*Mcur1-KO*) mice were generated on a C57BL/6 background using the CRISPR/Cas9 system supplied by Biogle Co Ltd. (Hangzhou City, China). The small guide RNA (sgRNA) targeting exon 1 of the mouse *Mcur1* gene (5'-GGAACGCGCTGGGCCGTCTG-3') were cloned into pX330 (Addgene, MA, USA). pCAG-EGxxFP was used to examine the efficiency of the target DNA cleavage by the sgRNA and Cas9. The pX330 plasmids containing each sgRNA sequence were injected into the pronuclear stage eggs. *Mcur1*$^{-/-}$ mice that harbored a 71 base pairs (bp) deletion in exon 1 of the *Mcur1* gene die before birth, but *Mcur1*$^{+/-}$ mice were obtained which appeared to be outwardly normal, and were designated as *Mcur1-KO* mice in this study. The genotypes of *Mcur1* in all mice were analyzed by PCR with the following primers: Forward, 5'-CACCTTCCCTTGCTCTCCTG-3', and Reverse, 5'-GAGACCCTCTATCACCTGCG-3'.

Further, the *Mcur1* conditional KO (*Mcur1-cKO*) mouse model was constructed on the C57BL/6 background in Shanghai Model Organisms Center (China). In the generation of *Mcur1* flox mice, we designed the loxP region to specifically target the exons 3 - 4 of *Mcur1* gene. Thus, a loxP element (lox2272-loxP-pA-IRES-En2SA-lox2272-loxP) was inserted into intron 2 and intron 4 of *Mcur1* gene using two sgRNAs, respectively. The sgRNA1 targeting intron 2 was 5'-CTAGCCTGAAAAAGACAATTTGG-3', and the other sgRNA2 targeting intron 4 was 5'-CATCAAACAGAGGTCTAGGATGG-3'. The genotype of *Mcur1* flox mice were confirmed by PCR with primers: Forward, 5'-GTTAGGGGGTGAAACTGGGG-3', and Reverse, 5'-CCTGCAAGTGTTTCAGTGGC-3'. The sizes of the PCR products from the floxed and WT alleles are 367 and 313 bp, respectively. To obtain the hematopoiesis-specific *Mcur1-cKO* mice, Mx1 was used according to the guideline from previous studies.[57,58]

The effect of the *Mcur1* (conditional) knockout on mice was validated through Western blotting assays by using the antibodies against MCUR1 (#A08547-1; BosterBio, USA) and β-actin (#60008-1-Ig; Proteintech, China). All the animal experiments were performed with the approval of the Institutional Animal Care and Use Committee of Beijing Institute of Radiation Medicine (Beijing, China).

Cell lines

The human erythroleukemic cell line K562 was maintained in RPMI 1640 medium (GIBCO, NY, USA) supplemented with 10% fetal bovine serum (FBS; HyClone, CA, USA), 100 U/mL penicillin and 100 μg/mL streptomycin. The cells were incubated at 37 °C in a humidified incubator containing 5% CO_2. The human embryonic kidney cell line HEK293T were grown in Dulbecco's modified Eagle's medium (DMEM; HyClone, CA, USA) with L-glutamine supplemented with 10% FBS, 100 U/mL penicillin and 100 μg/mL streptomycin at 37 °C and 5% CO_2. The human umbilical cord blood-derived erythroid progenitor 2 (HUDEP2) cell, which is an immortalized CD34$^+$ hematopoietic stem and progenitor cells (HSPCs)-derived erythroid progenitor cell line,[59] was expanded in serum-free medium (Stem Cell Technologies, Vancouver, Canada) supplied with doxycycline (1 μg/mL), dexamethasone (10^{-6} M), stem cell factor (SCF; 50 ng/mL), erythropoietin (EPO; 3 U/mL), 1% L-glutamine and 2% penicillin/streptomycin at 37 °C in 5% CO_2.

According to the institutional guidelines and after informed consent, the umbilical cord bloods and bone marrow (BM) aspirates from healthy individuals were obtained at the Fifth Medical Center of Chinese PLA (Beijing, China). Human CD34$^+$ HSPCs were then purified from these cord bloods and BM aspirates by using the magnetic-activated cell sorting magnetic beads system (Miltenyi Biotec, Bergisch-Gladbach, Germany), and the purity of them was > 95%. Human CD34$^+$ HSPCs were cultured on Retronectin-coated (Pan Vera, WI, USA) plates in serum-free medium (Stem Cell Technologies, Vancouver, Canada) containing growth factors (IL-3 [25 ng/mL], interleukin-6 [10 ng/mL], Flt-3 ligand [100 ng/mL], SCF [50 ng/mL] and thrombopoietin [100 ng/mL]) at 37°C in 5% CO_2. To assess the erythropoiesis, human CD34$^+$ HSPCs were further cultured in erythroid differentiation conditions (SCF, 5 ng/mL; IL-3, 5 ng/mL; and EPO, 3 U/mL) at 37 °C in 5% CO_2.

METHOD DETAILS

WGS, SNP detection and data quality control

Whole-genome sequencing (WGS) was performed for the 98 participants in the TIB_discovery (n = 48) and HAN_discovery (n = 50) populations at discovery stage at Wuxi AppTec (Wuxi City, Jiangsu Province, China) using Illumina's Hiseq X Ten platform to obtain a mean, per-sample depth of ~40× (Table S2). To avoid the influence of reads with artificial bias, such as low-quality paired reads that mainly result from base-calling duplicates and adapter contamination, the following several types of read were removed [60]: (a) reads with ≥ 10% unidentified nucleotides; (b) reads with < 10 nucleotides aligned to the adapter; (c) reads with > 10% nucleotides not matching with the adapter; (d) reads with > 50% nucleotides having phred quality < 5; and (e) putative PCR duplicates generated by PCR amplification in the library construction process (*i.e.*, read 1 and read 2 of two paired-end reads that were completely identical). Finally, an average of 119.8 Gigabytes (Gb) of data was retained for assembly, and among which ~86% of the bases achieved at least a base quality score of 30 (Q30) (Table S2). We further genotyped the 50 HAN_discovery individuals using the Illumina Human Omni ZhongHua-8 BeadChip at CapitalBio Corporation (Beijing, China), and then the concordance between the genotypes determined by Illumina SNP arrays and those determined by WGS was evaluated using the Kappa test. A high concordance rate between them was achieved (99.8%; $P < 2.2 \times 10^{-16}$, Kappa test; Table S4).

We conducted analysis of variant calling using the bioinformatics platform of the National Center for Protein Sciences (Beijing). In brief, we simultaneously aligned the sequencing reads of these 98 participants to the human reference genome (UCSC Genome Browser hg19) by Burrows-Wheeler Aligner (BWA, v0.5.9).[50] Next, we adjusted the alignments *via* several steps [9], including Genome Analysis Toolkit (GATK, v2.8-1) indel realignment, Picard read duplicate marking, and GATK quality score recalibration modules.[9] Then, the SNPs were called and filtered *via* the GATK UnifiedGenotyper under default parameter settings. As a result, the sequences were free of mapping bias due to similar base quality, mean depth and mapping coverage for all subjects (Table S2). By applying identity-by-descent test, none of the individuals has hidden relatedness.

SNPs genotyping in replication stage

In the replication stage, three SNPs at *MCUR1* (rs61644582) loci were genotyped by Sanger sequencing. PCR amplifications were performed in 25 μL reactions using the KOD-Plus-Neo high-fidelity thermostable DNA polymerase (Toyobo, Japan) on a BioRad c1000 Touch Thermal Cycler (Bio-Rad, USA). The primers for Sanger sequencing of these SNPs were provided in Table S15. Particularly notably, for the public available datasets not containing the rs61644582 genotypes, such as Yang et al. and Zheng et al. datasets, the rs1204175 was used as the proxy to rs61644582 in replication studies (Figure S3D), due to high LD between them based on the TIB_discovery population ($r^2 = 0.98$).

Phenotypic analyses of mice

A total of 28 randomly selected male *Mcur1-KO* mice (14 *Mcur1*$^{+/+}$ and 14 age-matched *Mcur1*$^{+/-}$ mice, respectively) were used for phenotypic analyses. The *Mcur1*$^{+/+}$ and *Mcur1*$^{+/-}$ mice (8-week-old) were placed in normoxic (21% O_2; 7 *Mcur1*$^{+/+}$ and 7 *Mcur1*$^{+/-}$ mice, respectively) or hypoxic (11% O_2, representing 5,000 m above sea level; 7 *Mcur1*$^{+/+}$ and 7 *Mcur1*$^{+/-}$ mice, respectively) environment in a polycarbonate hypoxic chamber (Coy Laboratory Products, Michigan, USA) as previously described.[61,62] In brief, the mice were placed in a hypobaric hypoxic chamber with inspired oxygen (FIO_2) of 0.11 for 23 h per day, and continued for 4 weeks. Therefore, the mice were exposed to normoxic conditions for 1 h per day, during which the cage were cleaned and the water and food were replenished. A standard light-dark cycle of 12 h-12 h light exposure was used. The oxygen level in the chamber is controlled automatically by balancing the air with nitrogen (N_2). After exposure to prolonged hypoxia for 4 weeks, mice were first equilibrated by room air for 1 hour (h) in order to avoid acute vasomotor responses, and were then anesthetized with urethane (1.4 mg/g). Body weights and spleen weights of all the mice were recorded and the experiment was terminated at the 28th day. We performed the same phenotypic analyses on *Mcur1-cKO* mice as in *Mcur1-KO* mice.

Hematology analyses and serum EPO ELISA

Peripheral blood was collected in heparinized glass capillaries (50 μL) by tail vein bleeding, and immediately analyzed for Hb concentration by photometric method using HemoCue Hb 201$^+$ analyzer (Hemo Cue, Angelholm, Sweden), and RBC count, hematocrit (HCT) level and other hematological parameters by auto hematology analyzer (Mindary BC-3000 plus, Shenzhen City, China). Reticulocytes were assayed using the Retic-Count reticulocyte reagent system (BD Biosciences, CA, USA). Serum EPO levels were measured using a mouse Erythropoietin ELISA kit (#MEP00B; R&D Systems, MN, USA) according to the manufacturer's instructions.

HE staining

For hematoxylin and eosin (HE) staining, the mice were perfused with phosphate buffered saline (PBS) by cardiac perfusion techniques, and 4% PFA was used for tissue fixation. For spleen histology, sections of spleens (12 mm) were used for HE staining. For BM histology analyses, sternal bones were fixed for > 24 h in 10% neutral buffered formalin at room temperature and demineralized in 12% EDTA for 2 weeks. The slides of spleen or sternums were treated at 60 °C for 30 minutes (min) and were then rinsed with water. After washing, the slides were stained with hematoxylin for 3 min. Then, the slides were washed sufficiently with water, immersed in 95% ethanol for 5 seconds (s), and counterstained with eosin for 10 s. The slides were then dehydrated in a gradient concentration of ethanol and dehydrated with xylene for 3 - 5 min. Finally, the slides were mounted with neutral balsam. Digital

images were obtained using a Nikon Eclipse E400 microscope (Nikon, Tokyo, Japan) equipped with a digital camera and analyzed using Spot Advanced software (model 2.1.1; Diagnostic Instruments, MI, USA).

Flow cytometry and cell sorting

Whole BM was isolated by flushing and crushing pelvic and hind leg bones with PBS supplemented with 2% heat-inactivated FBS (GIBCO, NY, USA) and penicillin-streptomycin (GIBCO, NY, USA). Then, the BM was lysed on ice with RBC lysis solution (Invitrogen, CA, USA) and washed in PBS with 2% FBS. Single-cell suspensions of spleen were prepared by pressing tissue through a 70 μM cell strainer followed by RBC lysis. Then the cell numbers of BM in per femur were counted by the Vi-CELL cell viability analyzer (Beckman Coulter, CA, USA). For peripheral blood, 250 - 500 μL of peripheral blood was diluted in 2 mL of RPMI media (GIBCO, NY, USA) and carefully pipetted onto 3 mL of lympholyte solution in a 15-mL tube. Single-cell suspensions from BM and spleen were stained with panels of fluorochrome-conjugated antibodies (Table S15). The following lineage markers were used: anti-mouse Ter119, Gr-1, CD11b, B220, CD3, CD4 and CD8. The staining and enrichment procedures for flow cytometry have been previously described.[63] Cells sorting were performed on a FACSAria flow cytometer (BD Biosciences, CA, USA). Data were analyzed by FlowJo software (v10.0; TreeStar, Oregon, USA). Results were obtained from three independent experiments and each experiment was done in triplicate.

BFU-E and CFU-E colony-forming assays

The cells from whole BM were harvested, subjected to RBC lysis, and re-suspended in IMDM with 10% FBS and 5% penicillin-streptomycin. Then, the cells from BM were plated on BFU-E culture mediums (M3436; Stem Cell Technologies, Vancouver, Canada) or CFU-E culture mediums (M3334; Stem Cell Technologies, Vancouver, Canada) according to the manufacturer's instructions at 37°C in a humidified atmosphere with 21% O_2 or 1% O_2 without opening the cell incubator during the period of hypoxia exposure, and counted at the 10th day for BFU-E and the 3rd day for CFU-E, respectively. Digital images of every colony were obtained by using an Olympus IX73 inverted microscope (Olympus, Tokyo, Japan), and colony sizes were measured by the area of each individual colony (pixel count) using a pixel count algorithm provided in ImageJ software (NIH). Results were obtained from three independent experiments and each experiment was done in triplicate.

scRNA-seq of BM cells from *Mcur1-WT and Mcur1-cKO* mice

Single-cell RNA sequencing (scRNA-seq)

For scRNA-seq, BM cells were obtained from the *Mcur1-WT* or *Mcur1-cKO* mice (12-week-old) under normoxia or hypoxia as previously described.[64] To obtain enough number of cells in library construction for scRNA-seq, for each group, BM cells from 3 mice were mixed into one sample. All cell suspensions were filtered through the 70 μm cell strainers and cell numbers were determined using the Vi-CELL cell viability analyzer (Beckman Coulter, CA, USA). To enrich c-Kit$^+$ cells, BM cells were stained with c-Kit-APC780 followed by anti-APC-conjugated microbeads. Cells were then separated by LS columns (Miltenyi Biotec, Bergisch-Gladbach, Germany) at 4 °C. The enriched c-Kit$^+$ cells were immediately converted to barcoded scRNA-seq libraries by using the Chromium Single Cell 3' Library, Gel Bead & Multiplex Kit and Chip Kit (#CG000204; 10× Genomics, USA), aiming for an estimated 10,000 cells per library and following the manufacturer's instructions. Samples were processed using kits pertaining to V3.1 barcoding chemistry of 10× Genomics. A single sample is always processed in a single well of a PCR plate, allowing all cells from a sample to be treated with the same master mix and in the same reaction vessel. All the samples were processed in parallel in the same thermal cycler. Then, the generated scRNA-seq libraries were sequenced on a NovaSeq sequencer (Illumina, USA).

Single-cell gene expression quantification, quality control and batch correction

The Cell Ranger software (version 4.0.0; 10× Genomics, USA) was used to perform sample demultiplexing, barcode processing and single-cell 3' counting. The fastq files for each sample were processed with the *count* function in Cell Ranger, which was used to align the reads to mouse genome (build mm10) and quantify the gene expression levels in single cells. To filter out low-quality cells for each sample, the cells that had either fewer than 500 or over 5,000 expressed genes were removed. To filter out dead or dying cells, the cells that had over 20% unique molecular identifiers (UMIs) derived from mitochondrial genome were further removed. Gene expression in single cells was normalized using *SCTransform* function in R package Seurat (v4.0) [65] and the anchor-based batch correction method was employed to merge samples from different groups.

Cells clustering and hematopoietic cell clusters annotation

Principle component analysis (PCA) was performed for dimensionality reduction and 30 principal components (PCs) were then used for uMAP visualization and cell clustering. For cell clustering, we used the *FindClusters* function in Seurat (v4.0), which implements the shared nearest neighbor (SNN) modularity optimization-based clustering algorithm. Twenty-eight cell clusters were identified after integrating *Mcur1-WT* and *Mcur1-cKO* cells from mice under hypoxia or normoxia conditions. Cell clusters with extremely low nUMI count and high fraction of doublets (two cells encapsulated in a single droplet) were assigned as low-quality clusters and were then excluded from further analyses. Major cell types were annotated using the canonical marker genes of hematopoietic cell types.[64] The hematopoietic stem cells and multipotent progenitors (HSCs/MPPs) were compiled and re-clustered for a second round using the same procedure as described above, and the clusters annotation was performed using a set of previously reported canonical marker genes of HSPCs, and was further confirmed by comparing them with the well-annotated cell clusters in the Atlas of Mouse Blood Cells dataset.[64] For the fairness of comparison of cellular abundances between different samples, we first normalized the counts of pre-erythroid cells with the overall cell counts identified in the corresponding samples. Then, we calculated the relative abundances of pre-erythroid cells in *Mcur1*-cKO samples to those in *Mcur1*-WT samples.

Identification of marker genes for cell clusters

To identify the marker genes for major cell types and cell subpopulations (clusters), we contrasted cells from a major type/subpopulation to all the other cells using the *FindAllMarkers* function of Seurat (v4.0), which identifies differentially expressed genes between two groups of cells using a Wilcoxon rank-sum test. *P* values were then corrected using Bonferroni correction based on the total number of genes in the dataset. Marker genes were defined as those genes with an adjusted *P* value < 0.01, a detectable percentage higher than 25% in that cluster, and an average expression level in one cluster at least 2-fold higher than that in the other clusters.

Lineage score calculation

Gene set enrichment analysis (GSEA) was performed to identify up-/down-regulated lineage-specific gene modules in HSPCs between *Mcur1-WT* and *Mcur1-cKO* mice under normoxic or hypoxic condition. Gene modules with enrichment score (ES) > 0 and false discovery rate (FDR) < 0.05 were defined as up-regulated modules, while ES < 0 and FDR < 0.05 were defined as down-regulated modules. Lineage signature gene sets, including erythroid, megakaryocytic, myeloid, lymphoid and proliferation, were obtained from previous study.[64]

Trajectory analysis and gene expression trend fitting

We used Monocle 3 (v.3.0) [66] to construct the trajectory tree of all groups of cells, which indicated the distinct lineage differentiation potential for each group of cells. Through integration analysis of the differentiation trajectory and erythroid-specific expressed genes, we further showed the differentially expressed gene in HSPCs between *Mcur1-WT* and *Mcur1-cKO* mice along the HSC-MPP-megakaryocyte/erythroid progenitors (MEP) differentiation trajectory. Gene expression trend versus pseudotime was fitted using Local Polynomial Regression model implemented in *stats* R package.

siRNAs and shRNAs

The small interfering RNAs (siRNAs) targeting *MCUR1*, *TSC1*, *TSC2*, *AMPKa1*, *CAMKK2*, *PU.1* and a non-targeting control siRNA were synthesized by RiboBio (Guangzhou City, China). The cells were transfected with 50 nM siRNAs using riboFECT Transfection Kit (#C10511-1; RiboBio, Guangzhou City, China) according to the manufacturer's instructions.

Lentiviral shRNA constructs (pLVshRNA-puro-EGFP) coding a scramble sequence (sh-Ctrl) and two independent shRNAs targeting *MCUR1* were obtained from Inovogen Tech (Beijing, China). Lentiviruses were produced with the gene-specific lentiviral shRNA constructs, psPAX2 and PMD2.G (Addgene, MA, USA). For lentiviral infection, human $CD34^+$ HSPCs at day 4 of differentiation were incubated with lentiviruses for 12 h before washing the excess virus. Forty-eight hours after lentiviral infection, the cells were selected with 1 μg/mL puromycin until the end of the culture period. All the constructs used in this study were confirmed by DNA Sanger sequencing. The sequences of all siRNAs and shRNAs are listed in Table S15.

Construction of MCUR1 overexpression vector

For rescue assays, five silent point mutations were introduced into the region targeted by *MCUR1* shRNAs (at nucleotides 497-515 based on NCBI NM_001031713, new sequence: 5'- TT-AAC-ACA-CAC-GCA-TTA-GT-3') to generate an *MCUR1* construct resistant to sh-MCUR1. The silent point mutations were generated using the QuickChange site-directed mutagenesis kit (#200513; Stratagene, CA, USA). Then, the full-length of cDNAs encoding this mutant *MCUR1* were cloned into adenovirus vector pAD-MCS-IRES-mcherry (Syngentech, Beijing, China). All the mutant *MCUR1* plasmids were confirmed by DNA sequencing.

qRT-PCR assays

RNA samples were isolated using the RNeasy mini kit (#74104; Qiagen, Hilden, Germany) and were reversely transcribed into cDNAs with the high-capacity cDNA reverse transcription kit (#4368814; Life Technologies, CA, USA). Quantitative real-time PCR assays were performed using iQ SYBR Green Supermix (#170-8862; Bio-Rad, CA, USA) on an iQ5 real-time PCR detection system (Bio-Rad, CA, USA). Each gene was assessed at least in triplicate using the $^{\Delta\Delta}$Ct method, and the *β-actin* (*ACTB*) was used as an internal control. Results were obtained from three independent experiments and each experiment was done in triplicate. The sequences of the primers were designed using Primer3 (v0.4.0) and listed in Table S15.

Western blotting assays

Western blotting assays were performed according to standard protocols. Cells were lysed in RIPA buffer plus EDTA-free protease inhibitor cocktail (#04693132001; Roche, Basel, Switzerland) and subjected to sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). Proteins were transferred to BioTrace™ NT Nitrocellulose Transfer Membrane (#P/N66485; PALL, USA) and incubated with antibodies against Human MCUR1 (SAB2100356; Sigma, USA), Mouse MCUR1 (SAB2700722; Sigma, USA), mTOR (#2983; Cell Signaling Technology, USA), phosphorylated mTOR (p-mTOR, Ser2448; #5536; Cell Signaling Technology, USA), p70 S6 kinase (S6K; #9202; Cell Signaling Technology, USA), phosphorylated p70 S6 kinase (p-S6K, Thr389; #9205; Cell Signaling Technology, USA), 4EBP1 (#9644; Cell Signaling Technology, USA), phosphorylated 4EBP1 (p-4EBP1, Thr37/46; #2855; Cell Signaling Technology, USA), ULK1 (#20986-1-AP; Proteintech, China), phosphorylated ULK (p-ULK1, Ser757; #14202; Cell Signaling Technology, USA), LC3 (#12741; Cell Signaling Technology, USA), p62 (#55274-1-AP; Proteintech, China), PGC1α (#66369-1-Ig; Proteintech, China), NRF1 (#12482-1-AP; Proteintech, China), YY1 (#66281-1-Ig; Proteintech, China), mtTFA (#19998-1-AP; Proteintech, China), AMPKα (#5831; Cell Signaling Technology, USA), phosphorylated AMPKα (p-AMPKα, Thr172; #2535; Cell Signaling Technology, USA), TSC1 (#6935; Cell Signaling Technology, USA), TSC2 (#4308; Cell Signaling Technology,

USA), phosphorylated TSC2 (p-TSC2, Ser1387; #23402; Cell Signaling Technology, USA); Raptor (#20984-1-AP; Proteintech, China), phosphorylated Raptor (p-Raptor, Ser792; #89146; Cell Signaling Technology, USA), LKB1 (#3047; Cell Signaling Technology, USA), phosphorylated LKB1 (p-LKB1, Ser428; #3482; Cell Signaling Technology, USA), CAMKK2 (#16810; Cell Signaling Technology, USA), phosphorylated CAMKK2 (p-CAMKK2, Ser495; #16737; Cell Signaling Technology, USA), PU.1 (#2258; Cell Signaling Technology, USA) and β-actin (#60008-1-Ig; Proteintech, China). Horseradish peroxidase (HRP)-conjugated Goat anti-Mouse IgG (#CW0102; CWBIO, China) or Goat anti-Rabbit IgG (#CW0102; CWBIO, China) secondary antibodies were used. Immunoreactive bands on the membrane were detected using SuperSignal™ West Pico Chemiluminescent Substrate kit (#34080; Thermo Scientific, MA, USA) and Western Blotting Detection System (Bio-Rad, CA, USA). The quantification of proteins in SDS-PAGE gel bands was measured by gray scanning analyses using ImageJ software (NIH).

Flow cytometry analyses of human CD34$^+$ HSPCs

For examining *in vitro* erythropoiesis, human CD34$^+$ HSPCs (isolated from cord bloods or BM aspirates) taken from culture at the 3rd, 7th and 10th day were washed in PBS and re-suspended in 50 mL of PBS containing 1% of FBS (Gibco, NY, USA). Then the human CD34$^+$ HSPCs were sorted for GPA$^+$CD71$^+$ erythroblasts, and BFU-E or CFU-E erythroid progenitor cells. Briefly, the cells were stained with the appropriate dilution of the antibodies and incubated for 30 min at 4 °C in the dark. Quadrant/gating in dot plots was determined, based on the isotype staining. Antibodies used in this study were as follows: CD34-BV421 (#343609; Biolegend, San Diego, CA, USA), IL-3R-PEcy7 (#25-1239-42; Thermo Scientific, MA, USA), CD36-PE (#555455; BD Biosciences, CA, USA), CD71-PEcy7 (#25-0719-42; Thermo Scientific, MA, USA) and GPA-APC (#17-9987-42; Thermo Scientific, MA, USA). Electronic compensation was done using unstained samples. The detailed sorting strategies were shown in Table S15. Stained cells were monitored by FACSAria flow cytometer (BD Biosciences, CA, USA) and analyzed with FlowJo software (v10.0; TreeStar, Oregon, USA). Results were obtained from three independent experiments and each experiment was done in triplicate.

Benzidine staining assays

Benzidine staining was used for detecting the hemoglobin-positive cells. Benzidine (Sigma, MO, USA) stock solution containing 3% benzidine and 90% glacial acetic acid was used to prepare a fresh working solution comprising a ratio of 1:1:5 for benzidine stock solution, H_2O_2 and ddH_2O, respectively. Induced human CD34$^+$ HSPCs (1.0×10^6) were centrifuged at 300 × g for 5 min. The cell pellet was re-suspended in 500 μL PBS and 100 μL of the benzidine working solution was added. After a 5 min incubation period, cells were centrifuged at 300 × g for 5 min and re-suspended in 500 μL PBS. The benzidine-positive cells, *i.e.*, the stained blue cells, were counted under a light microscope (Olympus IX73, Olympus, Tokyo, Japan) using a Bright-Line hemocytometer (Hausser, PA, USA), and were imaged using the cellSens Standard software (Olympus, Tokyo, Japan). Results were obtained from three independent experiments and each experiment was done in triplicate.

Cells growth assays

Human CD34$^+$ HSPCs were seeded in 24-well plates at the density of 1×10^5 cells per well with 50 μL of complete culture medium. The culture medium is changed daily. Following 3, 4, 5, 6 and 7 days, respectively, the cells were harvested, diluted by trypan blue working solution and then counted with an automated cell counter TC10™ (Bio-Rad, CA, USA), to allow for growth curve construction. Results were obtained from three independent experiments and each experiment was done in triplicate.

Cell cycle and cell apoptosis assays

Cells (1×10^6) were trypsinized and re-suspended to generate single-cell suspensions. For cell cycle analysis, we first fixated cells with ethanol and stained DNA with propidium iodide (PI) (Sigma-Aldrich, St. Louis, MO, USA) containing RNase A (Sigma-Aldrich), and the samples were then analyzed by flow cytometry (BD Biosciences, CA, USA). For apoptosis analysis, the cells were stained with APC-conjugated Annexin V/7-AAD Apoptosis Detection Kit (#88-8007-74; Thermo Scientific, MA, USA), as suggested by the manufacturer. Then, the apoptosis ratio was measured by flow cytometry. Results were obtained from three independent experiments and each experiment was done in triplicate.

BFU-E and CFU-E colony-forming assays for human CD34$^+$ HSPCs

Human CD34$^+$ HSPCs sorted from the cord bloods or BM were diluted at a density of 200 cells in 1 mL of MethoCult® H4434 classic medium for BFU-E colony-forming assays, and in 1 mL of MethoCult® H4330 medium for CFU-E colony-forming assays (Stem Cell Technologies, Vancouver, Canada), respectively, and incubated at 37°C in a humidified atmosphere with 1% O_2 or 21% O_2. The cells culture environment for hypoxia cultures (1% O_2) were achieved using a Forma Series II Water-Jacketed CO_2 Incubator (#3131; Thermo Fisher scientific, MA, USA). Culture environment was maintained at 1% O_2, 5% CO_2 and 94% N_2 without opening the cell incubator during the period of hypoxia exposure as previously described.[8] The BFU-E and CFU-E colonies were defined according to the criteria described previously by Dover *et al*.[67] Briefly, CFU-E colonies were counted and imaged on the 7th day, while BFU-E colonies were counted and imaged on the 14th day by using an Olympus IX73 inverted microscope (Olympus, Tokyo, Japan). Colony sizes were measured by the area of each individual colony (pixel count) using a pixel count algorithm provided in ImageJ software (NIH). Results were obtained from three independent experiments and each experiment was done in triplicate.

Cytospin assays

A total of 1 × 10^5 induced human CD34$^+$ HSPCs (isolated from cord bloods or BM aspirates) in 200 μL were used for cytospin preparation on coated slides by using the Thermo Scientific Shandon 4 Cytospin. The slides were stained with May-Grünwald solution (Sigma, MO, USA) for 5 min, rinsed in 40 mM Tris buffer for 90 s, and subsequently stained with Giemsa solution (Sigma, MO, USA) for 15 min. The overall trend of RBC maturation is from large and pale nucleus to darker and smaller nucleus, and then to nucleus loss; The cytoplasm gradually increases; The size gradually decreases; The cytoplasm changes from dark blue (filled with RNA) to grayish (a mixture of RNA and hemoglobin) and then to reddish (filled with hemoglobin, without RNA). The morphological characteristics of Basophilic erythroblasts (Bas) are large size, high nucleus/cytoplasm ratio, and with slight nuclear condensation. Polychromatic erythroblasts (Pol) exhibit a decrease in cell size, lighter-grayish cytoplasm and additional nuclear condensation. Orthochromatic erythroblasts (Ort) exhibit dark-opaque nucleus, gray-red cytoplasm, and the nucleus has become pyknotic. Erythrocytes (Ery) have no nucleus and their cytoplasm is orange-red. The cells were imaged and the numbers of the four types of cell were counted using an Olympus IX73 inverted microscope (Olympus, Tokyo, Japan). The percentages of each type of cells to the total were calculated. Results were obtained from three independent experiments and each experiment was done in triplicate.

Polysome profiling assays

Briefly, human CD34$^+$ HSPCs (isolated from the cord bloods or BM aspirates) or mouse whole BM cells were treated with 100 μg/mL cycloheximide (CHX) for 5 min. Harvested cells were washed with ice-cold PBS containing 100 μg/mL CHX. Cells were lysed in 500 μL hypotonic buffer (5 mM Tris-HCl pH 7.5, 2.5 mM $MgCl_2$, 1.5 mM KCl, protease inhibitor cocktail) containing CHX (100 μg/mL) and Triton-X100 (1% v/v). The supernatant was transferred to a pre-chilled 1.5 mL tube after centrifugation at 21,000 g for 5 min at 4 °C. Aliquots of the same OD number of lysates from each sample were loaded onto a 10% ∼ 50% sucrose gradient, and centrifuged at 35,000 revolutions per minute (r.p.m.) for 2 h at 4 °C using a SW40Ti rotor in a Beckman Coulter Optima L-80 ultra-centrifuge with no brake (Beckman Coulter, CA, USA). After centrifugation, gradients were fractionated and the optical density at 254 nm was continuously recorded using a BioFrac fraction collector (CA, USA), and chased with 60% (w/v) sucrose with bromophenol blue at 1 mL/min. Data were analyzed using the Bio-Rad LP Data View software. Polysome fractions were collected at 0.5 mL per fraction. Polysome/monosome (P/M) ratios of the traced profiles were used as a measure of protein translation rates, which were measured by the areas below the ribonucleoproteins (RNP) ribosomes and polysomes peaks. Results were obtained from three independent experiments and each experiment was done in triplicate.

Measurements of the mitochondrial bioenergetics

Measurements of the cytoplasmic and mitochondrial Ca^{2+} concentration

Human cord blood-derived CD34$^+$ HSPCs or mouse whole BM cells were cultured on poly-L-lysine-coated glass bottom Petri-dishes upon normoxic or hypoxic conditions for 48 h, and were then loaded with 5 μM rhod-2 (30 min; Thermo Scientific, MA, USA) and 5 μM Fluo-4 (30 min; Thermo Scientific, MA, USA) for measurements of the mitochondrial Ca^{2+} ($[Ca^{2+}]_m$) and cytoplasmic Ca^{2+} ($[Ca^{2+}]_c$) concentration, respectively, in extracellular medium (ECM; 120 mM NaCl, 5 mM KCl, 1 mM KH_2PO_4, 0.2 mM $MgCl_2$, 0.1 mM EGTA, 20 mM HEPES, pH 7.4). The glass bottom Petri-dishes were mounted in an open perfusion microincubator (PDMI-2; Harvard Apparatus, USA) at 37 °C and imaged. After 1 min of baseline recording, the agonist ionomycin (2.5 μM; MCE, NJ, USA) was added and confocal images were recorded every 10 s (510 Meta; Carl Zeiss, Jena, Germany) at 488 nm (for $[Ca^{2+}]_c$) and 568 nm (for $[Ca^{2+}]_m$) excitation using a 40× objective. Ionomycin is a Ca^{2+} ionophore, which can enhance the plasma membrane Ca^{2+} permeability while leaving mitochondrial membranes intact, resulting in rapid increase of $[Ca^{2+}]_c$ to trigger the mitochondrial Ca^{2+} uptake. Images were analyzed and quantified using ImageJ software (NIH).

Measurement of the mitochondrial respiration

Human cord blood-derived CD34$^+$ HSPCs or mouse whole BM cells were seeded in poly-L-lysine-coated 96-well Seahorse culture plates at the density of 1 × 10^4 cells. After 24 h, oxygen consumption rate (OCR) was measured using the XF Cell Mito stress kit (#103015-100; Seahorse Bioscience, MA, USA) at 37°C in an XF96 extracellular flux analyzer (Seahorse Bioscience, MA, USA). OCR was determined at four levels: (1) with no additions; (2) adding oligomycin (1 μM); (3) adding carbonyl cyanide 4-(trifluoromethoxy) phenylhydrazone (FCCP, 300 nM); and (4) adding rotenone (100 nM) plus antimycin A (100 nM). Results were obtained from three independent experiments and each experiment was done in triplicate.

Measurements of the cellular AMP and ATP levels

Total cellular ATP levels in human cord blood-derived CD34$^+$ HSPCs or mouse whole BM cells were measured using the CellTiter-Glo luminescent assay kit (#G7570; Promega, WI, USA). AMP levels were measured using the AMP-Glo luminescent assay kit (#V5011; Promega, WI, USA). Luminescence was measured using a microplate reader (Infinite M1000 PRO; Tecan, NC, USA). Results were obtained from three independent experiments and each experiment was done in triplicate.

Measurement of the mitochondrial reactive oxygen species (ROSs)

Human cord blood-derived CD34$^+$ HSPCs or mouse whole BM cells were plated on poly-L-lysine-coated glass coverslips. After 24 h, cells were loaded with the mitochondrial superoxide sensitive fluorophore MitoSOX Red (2 μM; Thermo Scientific, MA, USA) and Dihydrorhodamine 123 (Rhod123; 2.5 μg/mL; Thermo Scientific, MA, USA) in William's E medium (without serum) at 37°C for 30 min. Cells were then washed and imaged using a Carl Zeiss 510 confocal microscope with a 40× oil immersion objective at 561 nm as described previously.[68] MitoSOX fluorescence was then quantified using ImageJ software and plotted as arbitrary units using

GraphPad Prism package (v6.0; GraphPad Software, CA, USA). Results were obtained from three independent experiments and each experiment was done in triplicate.

TMRM staining for the mitochondrial membrane potential ($\Delta\Psi m$)

Human cord blood-derived $CD34^+$ HSPCs or mouse whole BM cells were placed on poly-L-lysine-coated glass coverslips. On the next day, cells were stained with tetramethylrhodamine, methyl ester (TMRM; 50 nM; Thermo Scientific, MA, USA) and Rhod123 (2.5 μg/mL; Thermo Scientific, MA, USA) for 30 min at 37°C. Images were obtained using a Carl Zeiss 510 confocal microscope with a 40× oil objective at excitations of 561 nm for TMRM and 488 nm for Rhod 123, respectively. Then, the TMRM fluorescence intensity was quantified using ImageJ software and plotted as arb unit using GraphPad Prism package (v6.0; GraphPad Software, CA, USA). Results were obtained from three independent experiments and each experiment was done in triplicate.

ChIP assays

We tested whether rs61644582 could affect the binding with specific PU.1 or ELF1 by chromatin immunoprecipitation (ChIP) assays. The human HUDEP2 and HEK293T cell lines, both of which were confirmed to be heterozygous (C/del genotype) at rs61644582, were crosslinked and processed using the ChIP Assay Kit (17-10086; Millipore, CA, USA). Briefly, approximately 1×10^7 cells were cross-linked with 1% formaldehyde, and cells nuclear extracts (N.E.) were sonicated on a Bioruptor Plus sonication system (Diagenode, Liege, Belgium). Chromatin lysate was pre-cleared using Dynabead protein A (Invitrogen, CA, USA) and subjected to immunoprecipitation using antibody against PU.1 (#2258; Cell Signaling Technology, MA, USA) or ELF1 (#sc-373772; Santa Cruz Biotechnology, CA, USA), or control IgG (#2729; Cell Signaling Technology, MA, USA). DNA-protein complex was precipitated using Dynabead protein A, eluted in washing buffers, and treated using proteinase K and RNase A in turn to reverse cross-links. DNA was then purified using the QIAquick PCR purification kit or Mini-Elute PCR purification kit (#28006; Qiagen, Hilden, Germany), and pre-immunoprecipitated DNA from each sample was used as input control.

For ChIP-qPCR analysis, the relative fold of enrichment was assessed by calculating the immunoprecipitation efficiency above fragment-specific background (IgG control) followed by normalization to the occupancy level observed in input sample. ChIP followed by allele-specific qPCR (ChIP-AS-qPCR) analysis was performed similarly to ChIP-qPCR. The difference of these two assays was that the primers were designed for allele-specific amplification of the rs61644582 region with a C or del allele in the DNA samples from ChIP. In ChIP followed by PCR and Sanger sequencing assays, PCR was performed on the target DNA fragments with rs61644582 C/del genotype, followed by sequencing. All the primer pairs used in ChIP assays were shown in Table S15. Results were obtained from three independent experiments and each experiment was done in triplicate.

EMSA and WEMSA

For the electrophoresis mobility shift assay (EMSA), the synthetic double-stranded and 3' biotin-labeled oligonucleotides containing the rs61644582 del or C allele (Table S15) were incubated with nuclear extracts (N.E.) from HUDEP2 or K562 cells at 25 °C for 20 min using the LightShift chemiluminescent EMSA kit (#20148; Thermo Scientific, MA, USA). The reaction mixture was separated on 6% PAGE, and the products were detected by Stabilized Streptavidin-Horseradish Peroxidase Conjugate. Unlabeled oligonucleotides at 100-fold molar excess were added to the reaction for competition. We confirmed the identity of the DNA-binding protein in assays using the antibodies specific to PU.1 (#2258; Cell Signaling Technology, MA, USA) or non-specific rabbit IgG (#2729; Cell Signaling Technology, MA, USA).

For the Western blotting-EMSA (WEMSA), probes containing the rs61644582 del or C allele (Table S15) were incubated with N.E. from HUDEP2 or K562 cells, or purified full-length recombinant PU.1 protein (100 ng; Abnova, Taipei, China) and purified full-length recombinant ELF1 protein (100 ng; Abnova, Taipei, China) along with the same binding buffer as described for EMSAs. The samples were incubated at 4°C for 20 min before they were run on DNA retardation gels at 250 V on ice for 20 min. The samples were then transferred on ice to a nitrocellulose membrane for Western blotting (30 V for 1 h). Following transfer, the membrane was blocked with 5% milk in TBST for 1 h with rocking at room temperature before primary antibody was applied (1:1,000 dilution in 5% milk) overnight at 4°C with rocking. The primary antibodies used for WEMSAs included antibodies against PU.1 (#2258; Cell Signaling Technology, MA, USA) and ELF1 (#sc-373772; Santa Cruz Biotechnology, CA, USA). The membrane was washed with TBST before addition of secondary antibody (#CW0102; CWBIO, Jiangsu, China) for 1 h at room temperature and was washed before substrate was added for imaging.

EMSA and WEMSA membranes were all imaged using the Amersham Imager 600 (GE Healthcare, USA). WEMSA bands were quantified using ImageJ software to assess the differences in binding between the two oligonucleotide probes (del and C allele) on each membrane. Each experiment was then normalized to the lower intensity and plotted according to protein to display the differences in binding based on alleles. Results were obtained from three independent experiments and each experiment was done in triplicate.

Luciferase reporter gene assays

The 2,100 bp genomic region surrounding the rs61644582 (Chr.6:13,836,590-13,838,690 bp; hg19) was amplified and inserted into the upstream of SV40 promoter in the pGL3-promoter vector (#E1761; Promega, WI, USA). Site-directed mutagenesis was employed to obtain either the C or del allele at the rs61644582 site. The SV40 promoter was subsequently replaced by *MCUR1* promoter region (-350 bp to +100 bp relative to the transcriptional start site). The primers were shown in Table S15. Then, we seeded 0.5×10^6

HUDEP2 or 1.5 × 10^6 K562 cells per well in 12-well plates, and transfected them with luciferase reporter plasmids using X-tremegene HP DNA Transfection Reagent (#6366244001; Roche, Basel, Switzerland). The pRL-SV40 plasmid (#E2231; Promega, WI, USA) was co-transfected as a negative control. After 48 h of incubation, the cells were collected and analyzed for luciferase activity using the Dual-Luciferase Reporter Assay System (#E1910; Promega, WI, USA). Results were obtained from three independent experiments and each experiment was done in triplicate.

QUANTIFICATION AND STATISTICAL ANALYSIS

Population genetic structure analyses

Principal component analyses

To investigate the populations' genetic structure and relatedness, we performed a series of principal component analyses (PCAs) by EIGENSTRAT (v3.0) software.[69] We evaluated the genetic structures for the TIB_discovery (n = 48) and the HAN_discovery subjects (n = 50) at the discovery stage and the TIB_Lu (n = 38) and HAN_Lu subjects (n = 39) from Lu *et al*,[60] along with the reference populations in the 1000 Genomes Project, which includes African (AFR, n = 246), European (EUR, n = 379), Admixed Mixed American (AMR, n = 181), South Asian (SAS, n = 489) and East Asian (EAS, n = 504). To eliminate the effects of linkage disequilibrium (LD) across SNPs on population structure, a total of 17,784 independent autosomal SNPs with minor allele frequency > 0.05 and r^2 < 0.001 between each other were selected from ~10 million SNPs and were used for the PCAs. Consistent with previous studies, the PCAs indicated that the Tibetans are genetically more similar to the Han populations (including HAN_discovery subjects, CHBs and CHSs) than the other ethnic populations, suggesting that Han population is an appropriate lowland population for subsequent comparative genomics analyses.[5] For batch effects filtering, we assessed the concordance of SNP calls from Chinese Han populations between HAN_discovery, CHB and CHS from the IGSR collection that are used for population structure analysis. We first exclude the SNPs with a missing rate > 0.05 for any collections. Then, we counted the number of homozygous (1/1), heterozygous (1/0), reference (0/0) and missing sites (./.) for each SNP, and performed a chi-squared test between all pairs of collections. *P* values were corrected by the Benjamini-Hochberg (BH) method. All SNPs that failed in more than one comparison ($P < 0.05$) were defined as batch effect and were excluded from the subsequent analyses.

ADMIXTURE analyses

ADMIXTURE, which is a model-based ancestry-estimation method,[52] was used to infer the individual genetic-ancestry coefficients by conditioning on a specific number (K) of "ancestral populations". Here, we used the same populations and set of SNPs (n = 17,784) as those used in PCAs. We ran ADMIXTURE with a random seed for the merged dataset from K = 2 to K = 10 with default parameters in ten replicates for each K. The K = 8 model with the lowest cross-validation error was selected as the best-fitting model. The TIB_discovery subjects showed no signs of significant shared ancestry with the Europeans, American or Africans, but did show high degree of shared ancestry with the individuals from the East Asian populations.

Runs of homozygosity (ROH) analyses

ROH is an indicator of genomic autozygosity, estimated for each individual using PLINK v 1.9 with the default parameters.[54] The same set of SNPs (n = 17,784) as those in the PCAs were used in ROH analyses.

Genetic drift sharing analyses

We performed genetic drift sharing analyses using the outgroup *f3* statistics,[51] and YRI (Yoruba in Ibadan, Nigeria) from the 1000 Genomes Project were used as the outgroup. We assumed that no admixture had occurred in a tree with topology (Yoruba; A, B), where the expected value was proportional to the shared genetic history between the A and B populations. That is, the larger the *f3* value, the greater the genetic relatedness between the two populations. ADMIXTOOLS [51] with the qp3pop program was employed to calculate the outgroup *f3* statistics in the form of *f3* (YRI; TIB, X), where TIB represents the Tibetans, and X represents a non-Yoruba population such as East Asian, Central Asian/Siberian, or South Asian populations.

Coalescent simulations

We introduced the best-fit demographic model for Tibetan and Han populations from our recent publication[70] by a diffusion approximation method of demographic inference,[71] with a mutation rate of 1.5 × 10^{-8} per site per generation[72] and a generation time of 29 years.[73] Briefly, we first tested three alternative demographic models with symmetric migrations to describe the simplified evolution paths. The diffusion approximation method of ∂a∂i was utilized to analyze the joint site frequency spectrum for Tibetans and Hans. Our analysis involved nine consecutive rounds of optimizations (with 100 replicates) following the dadi_pipeline workflow.[74] In each round, multiple replicates were executed, and the parameters estimated from the replicate with the highest log-likelihood were used to seed searches in the subsequent rounds. Parameter optimization was carried out using the derivative-based BFGS algorithm. Then, we used msprime [75] to perform whole-genome coalescent simulations around rs61644582 (± 1 Mb). To account for the parameter uncertainties, we simulated 1,000 replicates of the best-fit demographic model by randomly sampling values from the confidence intervals of each parameter, assuming that they had a multivariate normal distribution.

Natural selection signals

The population branch statistic (PBS) statistic was used to detect SNPs with positive selection signals by measuring high polarized divergence of allele frequency between TIB_discovery and HAN_discovery populations with respect to an outgroup CEU (Utah residents with northern and western European ancestry) from the 1000 Genomes Project (Phase 3). The detailed method has been described previously by Yi et al. (2010)[3]. In brief, we first calculated fixation index (F_{ST}) of each SNP between each pair of these three populations by using VCFtools[76], and then measured the PBS statistic by the formula from Yi et al. (2010)[3]. We defined the top 1‰ SNPs with PBS $\geq$ 0.24 as the genome-wide significance threshold to determine the regions of interest. The empirical *P* values of the PBS were computed by the rank statistic across the total number of SNPs included in the scan. These top SNPs can be clustered into 30 candidate loci targeted by positive selection. Then, the HAA-associated genes at a candidate locus were identified according to at least one of the following evidences: (1) containing coding variants that were in strong LD with the tag SNP; (2) biologically related to HAA based on manual literature review; (3) containing a cis-eQTL that were in strong LD with the tag SNP; and (4) nearest the tag SNP. Based on these lines of evidence, a total of 43 candidate genes were identified in the 30 loci potentially targeted by positive selection.

Meanwhile, diploS/HIC[16] was used to accurately detect and classify selective sweeps in genomic regions surrounding 6p23 locus through a deep convolutional neural network. We trained the model with eleven summary statistics that were calculated from simulation data, and then classified a genomic region into following one of the five classes: hard selection sweep, soft selection sweep, linked to a hard selection sweep, linked to a soft selection sweep, and selectively neutral region unlinked to a selection sweep.[16]

Estimation of the mutation age of rs61644582

The mutation age of rs61644582 was obtained from the pre-calculated mutation age estimates based on the 1000 Genomes Project populations (Phase 3) from the human genome dating server (https://human.genome.dating/) [15]. Because rs61644582 was not available in this human genome dating server dataset, rs1204168, which is in strong LD with rs61644582 (r^2 = 0.99, in Asians from 1000 Genomes Project), was used as its proxy.

Estimation of strength and time of selection signals surrounding *MCUR1*, *EGLN1* and *EPAS1* loci

For each positive selection signal, we estimated the strength (s), time (t), and frequency of the index allele at the divergence time (f) by simulating allele frequency trajectories in Tibetans using the "make sample with selection" (MSMS) software. We assumed the index allele to be segregating prior to the divergence time due to the presence of this allele in Han Chinese. Across a 3-D grid of values of these 3 parameters, we simulated 1,000 trajectories at each grid-point. To estimate the selection on a particular allele, we took a bin around its estimated allele frequencies in Tibetans by using a Clopper-Pearson 99% confidence interval (CI) on the allele's true population frequencies.[48] For alleles simulated under neutrality (*i.e.*, s = 0), we observed no hits in the bins for each locus under any combinations of (t, f), suggesting that the allele frequency discrepancy is unlikely under neutral evolution in the inferred demographic model ($P < 0.001$). Specifically, for estimating the selection at 6p23 locus (*MCUR1*), we fixed f = 0.26 for rs61644582 del allele in HAN_discovery population, and 99% CI bins with 0.49 (0.36 - 0.62). For estimating the selection at *EPAS1* locus, we fixed f = 0.01 for index rs76347095, and 99% CI bins with 0.49 (0.36 - 0.62). For estimating the selection at *EGLN1* locus, we fixed f = 0.01 for index rs186996510, and 99% CI bins with 0.58 (0.45 - 0.71). The results were shown in Figure S3B.

Expression enrichment analyses of HAA-associated genes

We used the Data-driven Expression Prioritized Integration for Complex Traits (DEPICT)[55] analyses to test whether the HAA-associated genes are significantly highly expressed in specific type(s) of tissue or cell. DEPICT includes the gene expression profiles of 209 types of tissue or cell from 37,427 human samples. In brief, for each type of tissue or cell, the DEPICT method performs *t* tests by comparing the tissue-specific expression levels of HAA-associated genes with those of genes randomly selected. Then, the empirical enrichment *P* values were computed by repeatedly sampling the random sets of genes (matched to the actual HAA-associated genes by genes density) from the entire genome, and the empirical mean (s.d.) of the enrichment statistic's null distribution were estimated. The *P* values of tissue- or cell-type enrichment analyses were calculated empirically using 500 permutations for bias adjustment and 20 replications for false discovery rate (FDR) estimation. Finally, 9 types of tissue or cell were identified in which our candidate HAA-associated genes were significantly enriched by DEPICT analyses (all $P < 0.05$, student's *t* test; Table S8).

To replicate these enrichment findings in independent sample sets, we also collected 18 publicly available genome datasets (Table S9). These genome datasets were from 2 human populations living on the Qinghai-Tibetan Plateau, 1 human population living on the Andean Altiplano Highland, 1 human population living on the Ethiopian Highland, and 14 animal populations living on the Qinghai-Tibetan Plateau, respectively. For the Tibetan human datasets, we applied the same pipeline as in the discovery stage of this study, and performed a regular quality control procedure on raw data and subsequently obtained the candidate HAA-associated genes. For those datasets from 2 human populations living on non-Tibet plateaus and 14 types of Tibetan animals, we directly obtained the HAA-associated genes annotated in the original papers. If these HAA-associated genes from Tibetan animal studies have homologous genes in humans, we then mapped them to the homologous human genes to obtain the final HAA-associated genes (Table S9). Finally, for each population, DEPICT with the default parameters was then used to investigate whether the HAA-associated genes derived from this population were significantly highly expressed in specific type(s) of tissue or cell.

We also performed the expression enrichment analyses in tissues by another method based on the data from the GTEx project and permutation test, as previously described.[77] In brief, the GTEx project (v8) contained the bulk RNA sequencing data of 8,555

samples.[78] A total of 52,577 transcripts were retained for subsequent analyses if they have HUGO Gene Nomenclature Committee (HGNC) identities (IDs) (n = 20,932). Those transcripts were further excluded if they have a non-coding RNA IDs (remaining n = 18,359). Then, we ranked all the transcripts by transcripts per million (TPM) across all samples, and generated 1,000 permutations of each credible set gene list by selecting a random transcript for each entry in the credible set within ± 100 ranks of the transcript for that gene. For each sample, the TPM values were converted into ranks for that transcript, and sums of ranks within each tissue were computed for each gene. We calculated the enrichment *P* values for each tissue by taking the total number of instances when the gene list of interest had a lower sum of ranks than the permuted sum of ranks that was divided by the total number of permutations. These candidate HAA-associated genes in the present study and the abovementioned HAA-associated genes from 18 publicly available genome datasets were analyzed (Table S9).

Pathway enrichment analyses of HAA-associated genes

To investigate the pathways potentially involved in HAA, we first used g:profiler[56] to map the candidate HAA-associated ge4nes to the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. Those pathways with less than 5 members or more than 500 members were excluded. Benjamini-Hochberg FDR was used to correct the *P* values, and the pathways were considered to be significantly enriched when the corrected $P < 0.05$. A total of 32 pathways displayed significant enrichment (all adjusted $P < 0.05$). Then, by using the same pipeline analyses, we sought to replicate these 32 pathways in the 18 aforementioned publicly available datasets. Finally, 12 ones out of these 32 pathways remained significant enrichment in at least one dataset (all corrected $P < 0.05$; Table S10). The Cytoscape tool was then used to construct the overlap relationship of genes in these significantly enriched pathways (Figures S2F and S2G).

Associations between rs61644582 and mRNA levels of nearby genes, altitudes and erythropoiesis-related phenotypes

Association between rs61644582 and mRNA levels of nearby genes

We evaluated the associations between SNP genotypes and mRNA levels of nearby genes within 1 Mb of the 6p23 locus based on datasets from HaploReg (v4.1) and GTEx (v8) databases, respectively. Because the genotype data of rs61644582 were not available in the HaplorReg datasets, the rs1204168 (in HaploReg) was used as the proxy to rs61644582 due to strong LD between them in Asians from 1000 Genomes Project (LD $r^2 = 0.99$). The genotypes T/T, T/C, and C/C of rs1204168 correspond to the genotypes C/C, C/del, and del/del of rs61644582, respectively. Given the significant eQTLs of *RNF182* and *TBC1D7* shown in GTEx database, we assessed the associations between rs61644582 and mRNA levels of these two genes as well as *MCUR1* in another independent Tibetan population (*i.e.*, above-mentioned TIB_replication 2), and finally linked the rs61644582 with the *MCUR1* gene. We also assessed the association between rs61644582 and mRNA levels of *MCUR1* in the cord blood samples from TIB_replication 3 population to replicate the eQTL findings of *MCUR1*. These Tibetans were genotyped using Sanger sequencing, and their *MCUR1* mRNA levels were quantified using qRT-PCR assays and then log2 transformed. The associations between the genotypes of rs61644582 and the expression levels of the corresponding genes were evaluated using a linear regression model. *P* values were considered to be significant when below 0.05.

Association between rs61644582 and altitudes

The frequencies of the rs61644582 del allele were obtained from this study and several previous studies (Table S12). The altitude information of most of the populations has been reported in the original public papers. For those populations which do lack altitude information in the original paper, we obtained it by searching for the Google map of their recruitment location. The correlations between the frequency of rs61644582 del allele and altitude were performed by Kendall rank test. We observed that the frequency of the rs61644582 del allele is significantly associated with an increase of altitude in Asian populations, but not in populations from other continents (Figure 1F). *P* values were considered to be significant when below 0.05.

Associations between rs61644582 and erythropoiesis-related phenotypes

Three in-house populations (*i.e.*, TIB_replication 3, TIB_replication 4 and HAN_replictaion 1) and several other publicly available populations were used for this analysis (Table S14). For the in-house populations, the cord bloods (from TIB_replication 3) or peripheral bloods (from TIB_replication 4 and HAN_replictaion 1) were sampled. The rs61644582 was genotyped by Sanger sequencing, and the hematological phenotypes were measured by photometric method using HemoCue Hb 201$^+$ analyzer (Hemo Cue, Angelholm, Sweden) and auto hematology analyzer (Mindary BC-3000 plus, Shenzhen City, China). Additional erythropoiesis-related physiological phenotypes in cord bloods, such as erythroblasts (GPA$^+$CD71$^+$), were measured by FACS sorting assays, and BFU-E/CFU-E was determined by colony-forming assays. The associations were evaluated by using a linear regression model corrected for age and sex when available. Especially, the analyses in TIB_replication 3 were corrected for age of their mothers and sex of newborns. *P* values were considered to be significant when below 0.05. For the publicly available populations, the associations were derived from the original papers, or re-estimated when the data are available by the same analyses as that performed in the in-house populations.

Associations between MCUR1 mRNA levels and erythropoiesis-related phenotypes

TIB_replication 3 was used for this analysis. *MCUR1* mRNA levels and erythropoiesis-related phenotypes were measured, respectively, as the description in the aforementioned section of "associations between rs61644582 and erythropoiesis-related phenotypes". The associations were evaluated by using a linear regression model.

Publicly available *MCUR1* expressions in multiple tissues and cell lines

The data of MCUR1 expression in multiple tissues were obtained from the BodyMap 2.0 project (GSE30611) and DMAP dataset [20]. In the BodyMap 2.0 project, the MCUR1 expressions from 16 types of human tissue can be retrieved in GEPIA (http://gepia.cancer-pku.cn/detail.php?gene=MCUR1). Additionally, to reveal the potential target cell(s) in which MCUR1 plays its roles in hematopoietic differentiation, we searched for the expression levels of MCUR1 from the DMAP dataset, which contains 211 mRNA expression profiles of 38 types of cell lines of hematopoietic cell development lineage by the Affymetrix Human Genome U133 Plus 2.0 arrays (Affymetrix, CA, USA). *P* values were obtained from the original paper.[20]

Publicly available *MCUR1*-associated phenotypes prediction

We searched for the potential phenotypes relevant to *MCUR1* gene in humans and mice that were predicted by all RNA-seq and ChIP-seq sample and signature search (ARCHS4).[24] ARCHS4 integrated a large number of RNA-seq datasets from 72,363 mouse and 65,429 human samples in the Sequence Read Archive (SRA) and the Gene Expression Omnibus (GEO); thus, ARCHS4 provides predictions of phenotypes for genes in humans and mice based on prior knowledge combined with co-expression data. Based on ARCHS4, we found that *MCUR1* is relevant to specific HAA phenotype(s), *e.g.*, erythropoiesis.

RNA-seq and GSEA analyses

Total RNAs were extracted from the cultured human $CD34^+$ HSPCs using TRIZOL Reagent (Invitrogen, CA, USA), which stably expressed sh-*MCUR1* or sh-Ctrl under hypoxia (1% O_2) and 7 days of EPO stimulation (3 U/mL). Endogenous *MCUR1* was knocked-down by pooled shRNAs targeting *MCUR1* (sh-MCUR1, including sh-MCUR1-1 and sh-MCUR1-2 [1:1]) in HSPCs. RNA-seq was performed using Illumina NovaSeq6000 at Berry Genomics Co. Ltd. (Beijing, China), and a mean of 40 million paired-end reads per sample was obtained. The Kallisto (v0.43.1) was used to estimate the gene expression levels based on the novel idea of pseudoalignment for rapidly determining the compatibility of reads with targets, without the need for alignment. The expression levels for each gene were normalized to TPM to facilitate the comparison of transcript levels among samples. The relative abundances of transcripts were measured by the normalized RNA-seq fragment counts.

Then, gene set enrichment analysis (GSEA)[53] was performed based on the paired groups (*MCUR1* knockdown *vs.* control) of the RNA-seq data, with the genes ranking by their signal-to-noise ratios. The gene sets enriched by *MCUR1* knockdown were identified by the weighted Kolmogorov-Smirnov test on the basis of the Molecular Signatures Database (MSigDB, v6.2). The tests were permuted 1,000 times to compute the FDR, and the FDR *q* values less than 0.05 were considered to be statistically significant.

Functional annotation of SNPs at 6p23 locus

HaploReg (v4.1) was used to search for SNPs which are tagged by rs61644582 (with LD $r^2 > 0.8$ in East Asian population from the 1000 Genomes Project). HaploReg was also used to predict whether or not a non-coding variant was within a motif binding with transcription factors (TFs), and the effects of the variant on binding strength of TFs. To investigate whether the candidate SNPs at 6p23 locus interacted physically with the promoter regions of nearby genes, the data of high-through chromosome conformation capture (3C) from two types of human blood cell line (K562 and GM12878) [79] were used. We checked the promoter regions of five genes located within 200 Kb surrounding the rs61644582 (including *SIRT5*, *NOL7*, *RANBP9*, *MCUR1* and *RNF182*), and observed there indeed exist chromatin interactions between the promoter regions of *MCUR1* and DNA fragments containing the rs61644582.

Other statistical analyses

Fisher's exact test or χ^2 test was used for analyses of contingency tables depending on the sample sizes, for analyses of difference in allele frequencies between Tibetans and Hans, and for independence tests among SNPs in genomic regions subject to positive selection at 6p23 locus. *P* values were calculated by two-sided Student's *t* test for means of age, activities of reporter genes, genes mRNA levels. The two-way Analysis of Variance (ANOVA) test was used for analyses on cells growth experiments. The estimate of variation within each group of data was carried out by F-test. Pearson's test was used to evaluate the correlation coefficiency (*r*) and the two-tailed *P* values of the correlation between the mRNA levels of *MCUR1* and those of *PU.1*. Multiple testing correction was performed using the Benjamini-Hochberg FDR. All statistical analyses were performed using R package (https://www.R-project.org). Data are shown as mean ± s.d.. Results were obtained from three independent experiments and each experiment was done in triplicate. *P* values less than 0.05 were considered to be statistically significant unless otherwise specified.