Metabolic control of DNA methylation in naive pluripotent cells

Riccardo Betto, Linda Diamante, Valentina Perrera, Matteo Audano, Stefania Rapelli, Andrea Lauria, Danny Incarnato, Mattia Arboit, Silvia Pedretti, Giovanni Rigoni, et al.

HAL Id: hal-03322671
https://hal.science/hal-03322671

Submitted on 14 Oct 2021

Check for updates

Metabolic control of DNA methylation in naive pluripotent cells

Riccardo M. Betto [1], Linda Diamante [1], Valentina Perrera[1,14], Matteo Audano[2], Stefania Rapelli[3,4], Andrea Lauria[3,4], Danny Incarnato[3,5,15], Mattia Arboit[1], Silvia Pedretti[2], Giovanni Rigoni[6,16], Vincent Guerineau[7], David Touboul[7], Giuliano Giuseppe Stirparo [8], Tim Lohoff [8], Thorsten Boroviak [9,10,11,12], Paolo Grumati[13], Maria E. Soriano[6], Jennifer Nichols [8,9], Nico Mitro [2] ✉, Salvatore Oliviero [3,4] ✉ and Graziano Martello [1,6] ✉

Naive epiblast and embryonic stem cells (ESCs) give rise to all cells of adults. Such developmental plasticity is associated with genome hypomethylation. Here, we show that LIF–Stat3 signaling induces genomic hypomethylation via metabolic reconfiguration. $Stat3^{-/-}$ ESCs show decreased α-ketoglutarate production from glutamine, leading to increased Dnmt3a and Dnmt3b expression and DNA methylation. Notably, genome methylation is dynamically controlled through modulation of α-ketoglutarate availability or Stat3 activation in mitochondria. Alpha-ketoglutarate links metabolism to the epigenome by reducing the expression of Otx2 and its targets Dnmt3a and Dnmt3b. Genetic inactivation of Otx2 or Dnmt3a and Dnmt3b results in genomic hypomethylation even in the absence of active LIF–Stat3. $Stat3^{-/-}$ ESCs show increased methylation at imprinting control regions and altered expression of cognate transcripts. Single-cell analyses of $Stat3^{-/-}$ embryos confirmed the dysregulated expression of Otx2, Dnmt3a and Dnmt3b as well as imprinted genes. Several cancers display Stat3 overactivation and abnormal DNA methylation; therefore, the molecular module that we describe might be exploited under pathological conditions.

After fertilization, the zygotic genome is demethylated to establish a blank canvas for embryonic development. DNA methylation occurs on carbon 5 of cytosine (5mC) and is catalyzed by DNA methyltransferases (Dnmts). Ten-eleven translocation (Tet) proteins promote oxidation of 5mC to hydroxymethylcytosine (h5mC)[1,2]. Additional oxidation steps mediated by Tets lead to the conversion of h5mC into unmodified cytosine. Both Dnmts and Tets are dynamically expressed during early development, leading to a local minimum of 5mC at the preimplantation blastocyst stage at embryonic day 3.5 (E3.5)[3–5]. Imprinted genes, expressed mono-allelically in a parent-of-origin fashion, resist this wave of DNA demethylation. Such mono-allelic expression allows tight control of their dosage and is essential for proper embryonic development[6]. How is the expression of Dnmts and Tets controlled in the early embryo? In the embryo, the Jak–Stat pathway is active from E2.5 and E3.5, as shown by phosphorylation of Stat3 as well as transcriptional activation of its targets Socs3 and Tfcp2l1 (refs. [7–11]). Thus, Stat3 represents a good candidate as a regulator of Dnmt and Tet expression.

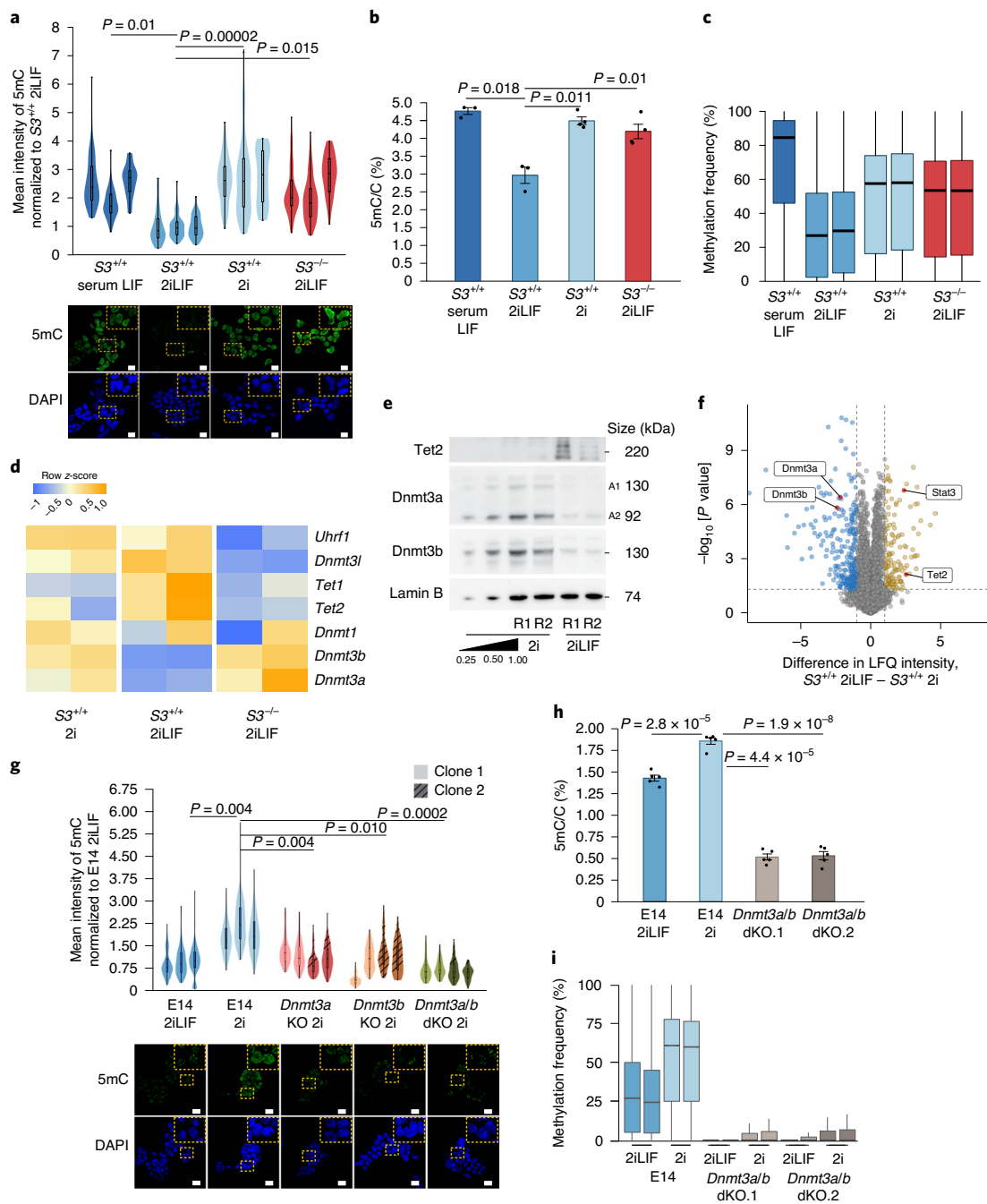Mouse naive embryonic stem cells (ESCs) are derived from preimplantation blastocysts[12] and show genomic hypomethylation only when cultured in the presence of LIF, a ligand of the Jak–Stat pathway, in combination with two inhibitors of the kinases GSK3 and MEK (2iLIF conditions[13–17]). Conversely, ESCs cultured in fetal bovine serum-based medium with LIF (serum LIF conditions[18]) display higher levels of DNA methylation. Such findings indicate that LIF is insufficient to induce genomic hypomethylation in the presence of serum but the requirement of Stat3 to induce hypomethylation in 2iLIF conditions has never been formally tested.

Stat3 represents an ideal regulator of the epigenome, as it both regulates gene expression in the nucleus and cellular metabolism in mitochondria by promoting oxidative phosphorylation (OXPHOS)[19–21]. Several metabolites are known regulators or cofactors of enzymes catalyzing epigenetic modifications[22]. For all these reasons, we genetically tested the role of Stat3 on genome methylation of naive pluripotent cells.

Results

LIF–Stat3 induces hypomethylation in ESCs via Dnmt3a/b regulation in 2i. We measured the levels of 5mC by quantitative immunostaining in ESCs and observed a strong decrease in signal intensity in 2iLIF compared with serum LIF (Fig. 1a), as previously

Fig. 1 | LIF–Stat3 induces hypomethylation in ESCs via Dnmt3a/b regulation. a, Anti-5mC immunofluorescence on $S3^{+/+}$ cells cultured in serum LIF, 2i or 2iLIF and $S3^{-/-}$ cells in 2iLIF. Representative images (bottom) and violin plots (top) showing the distribution of fluorescence intensity of an average of 63 nuclei per sample; $n = 3$ experiments shown as individual violins; statistical analyses were performed on the median intensity values of each sample. **b**, Percentage of 5mC quantified by mass spectrometry. The mean and s.e.m. of $n = 3$ ($S3^{+/+}$ serum LIF and 2iLIF) and $n = 4$ ($S3^{+/+}$ 2i and $S3^{-/-}$ 2iLIF) biological replicates are shown, with each replicate shown as a dot. **c**, DNA methylation at CpG islands measured by RRBS. **d**, Heatmap of transcriptomic data showing the z-score of scaled expression values from RNA-seq; $n = 2$ biological replicates. **e**, Western blot of $S3^{+/+}$ cells in 2i or 2iLIF medium; $n = 2$ biological replicates (R1 and R2). Two isoforms of Dnmt3a were detected (a1 and a2)[73]. Lamin B was used as the loading control. Representative images of $n = 2$ independent experiments. **f**, Proteomic data from $S3^{+/+}$ cells in 2iLIF or 2i medium. The yellow and blue dots indicate proteins that are less or more abundant (>+1 or <–1 difference; $P < 0.05$, as indicated by the dashed line), respectively, in 2iLIF relative to 2i. LFQ, label-free quantification; $n = 5$ biological replicates. See Supplementary Table 3. **g**, Anti-5mC immunofluorescence of E14 cells in 2iLIF and 2i medium as well as Dnmt3a KO, Dnmt3b KO and Dnmt3a/b dKO in 2i. Violin plots of an average of 82 nuclei per sample (top). Independent experiments are shown as violins. Representative images are shown with the insets showing magnified images of the regions in the yellow boxes (bottom). **h**, Percentage of 5mC in E14 cells cultured in 2iLIF and 2i as well as two Dnmt3a/b dKO clones, determined from mass spectrometry data. Mean and s.e.m. of $n = 5$ biological replicates, with each replicate shown as a dot. **i**, RRBS on the indicated samples; $n = 2$ biological replicates. All violin and boxplots indicate the 1st, 2nd and 3rd quartiles; the whiskers indicate the minimum and maximum values. All P values were calculated using a two-tailed unpaired Student's t-test. Scale bars, 20 μm.
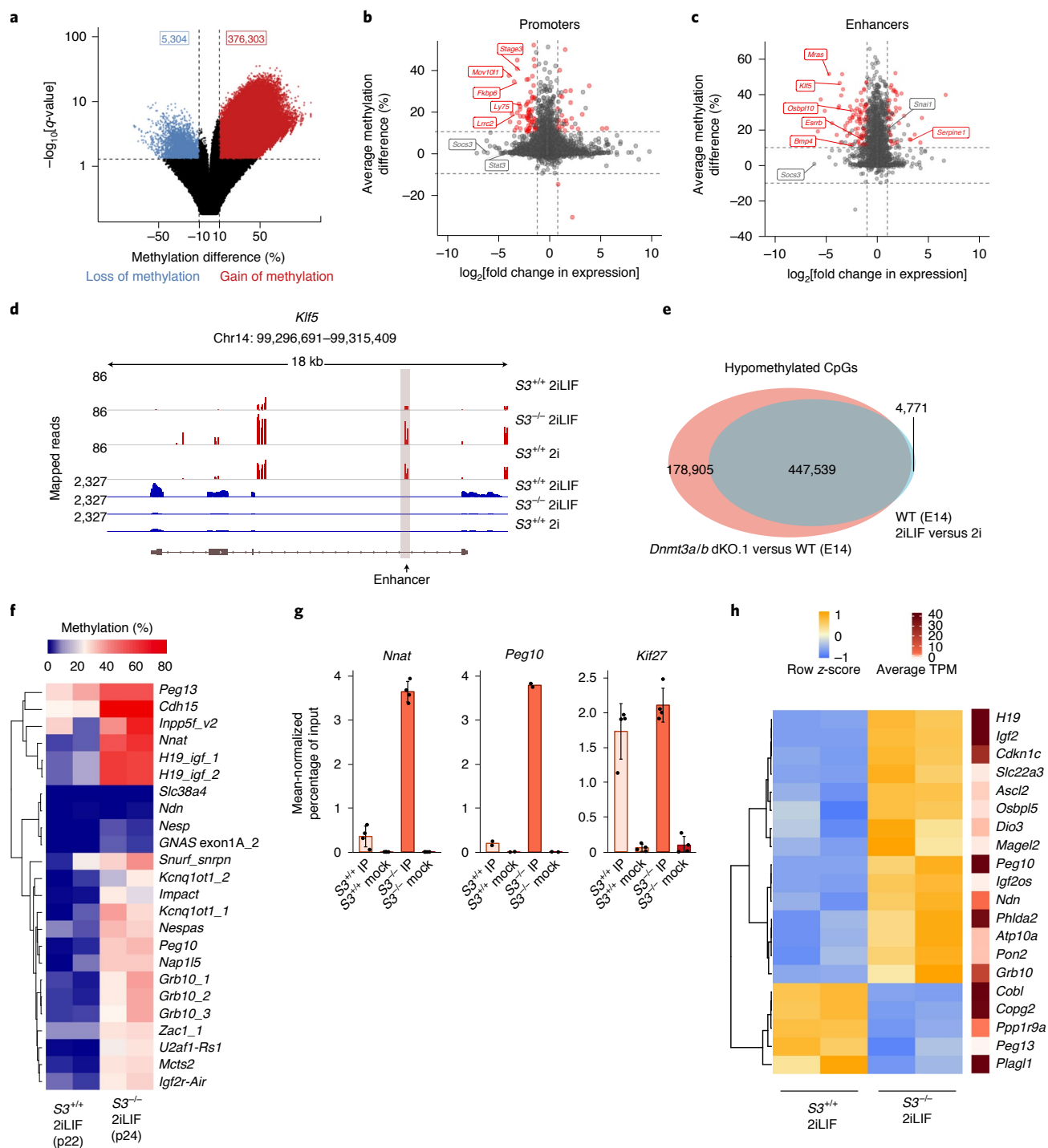
Fig. 2 | Impact of Stat3 on DNA methylation and transcription. a, Distribution of CpG sites with significantly different methylation (q-value < 0.01, difference > +10% or < −10%, as indicated by the dashed lines) between $S3^{−/−}$ and $S3^{+/+}$ cells cultured in 2iLIF, determined from RRBS. **b,c**, Scatter plots showing changes in expression and DNA methylation at active promoters (**b**) or enhancers (**c**) between $S3^{+/+}$ and $S3^{−/−}$ cells cultured in 2iLIF, determined from RNA-seq and RRBS. Red dots, genes for which both changes were statistically significant (q-value < 0.01, methylation difference > +10% or < −10%, shown as horizontal dashed lines, and $\log_2$ fold change in expression > +1 or < −1, shown as vertical dashed lines). See Supplementary Table 2. **d**, Gene tracks showing RRBS and RNA-seq data for $S3^{+/+}$ 2iLIF, $S3^{+/+}$ 2i and $S3^{−/−}$ 2iLIF cells over the $Klf5$ genomic region. One of two representative biological replicates is shown. Schematic of the $Klf5$ gene structure is shown in gray. **e**, Venn diagram of CpG sites whose methylation status is dependent on either LIF (light blue), Dnmt3a/b (red) or both (gray intersection). WT, wild type. **f**, Percentage of DNA methylation at imprinted DMRs from the RRBS data; $n = 2$ biological replicates per sample. See Supplementary Table 2. **g**, MeDIP–qPCR of DMRs and a control region ($Kif27$). Mock immunoprecipitations with a nonspecific IgG antibody served as the negative controls. Mean ± s.d. of $n = 4$ ($Nnat$ and $Kif27$) or mean of $n = 2$ ($Peg10$) experiments, with each replicate shown as a dot; IP, immunoprecipitation **h**, Heatmap showing relative (left) and absolute (right) expression of imprinted genes associated with known DMRs (Fig. 2f) and differentially expressed between $S3^{−/−}$ and $S3^{+/+}$ cells, determined from RNA-seq data. Relative expression (left) is shown as z-scores of the scaled values; absolute expression (right) is indicated as the average transcripts-per-million (TPM) values; $n = 2$ biological replicates per sample. All P values were calculated using a two-tailed unpaired Student's t-test.
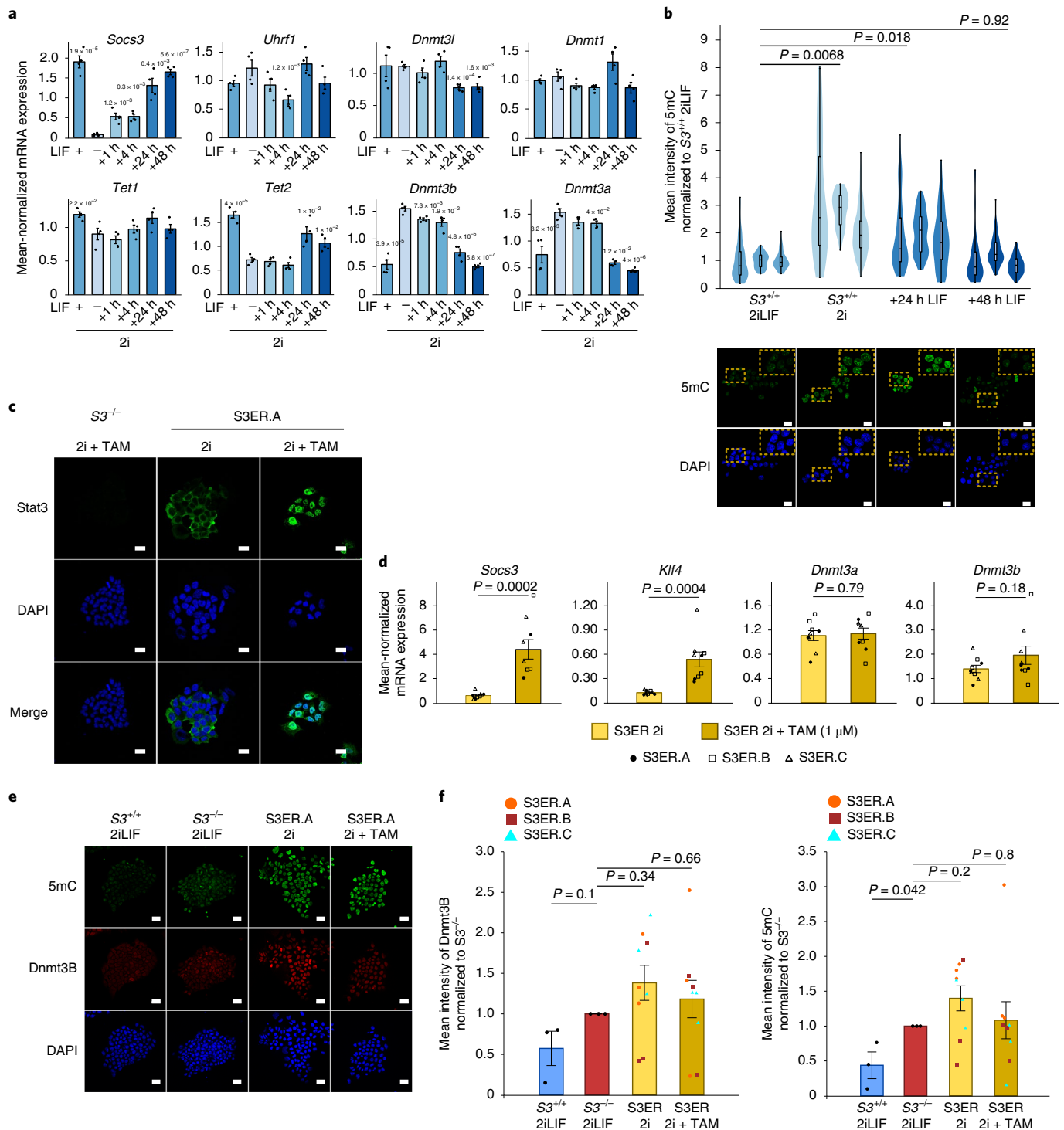
Fig. 3 | Nuclear Stat3 does not regulate Dnmt3a/b and 5mC. a, Gene expression analysis (RT–qPCR) of $S3^{+/+}$ cells stably cultured in 2iLIF (+) or 2i (−), or stimulated acutely with LIF for 1, 4, 24 and 48 h. Mean ± s.e.m. of $n = 4$ experiments, with each replicate shown as a dot. The P values relative to 2i are shown only when $P < 0.05$. **b**, Anti-5mC immunofluorescence following LIF addition for 24 or 48 h to $S3^{+/+}$ cells cultured stably in 2iLIF or in 2i (top). Representative images are shown with the insets showing magnified images of the regions in the yellow boxes (bottom). The violin plots show the distribution of fluorescence intensity of an average of 86 nuclei per sample; the boxplots show the 1st, 2nd and 3rd quartiles; the whiskers indicate the minimum and maximum values; $n = 3$ experiments. P values were calculated using the median values. **c**, Anti-Stat3 immunofluorescence of $S3^{-/-}$ cells in 2i + TAM and a representative clone of S3ER cells in 2i medium or 2i + TAM. Representative images of $n = 3$ independent experiments. **d**, Gene expression analysis (RT–qPCR) of three S3ER clones cultured in 2i with or without TAM. Mean ± s.e.m. of $n = 3$ experiments, with individual data points represented as dots, squares or triangles. **e**, Representative confocal images of $n = 3$ independent experiments are shown for $S3^{+/+}$ and $S3^{-/-}$ cells in 2iLIF medium as well as a representative clone of S3ER cells cultured either in 2i or 2i + TAM and stained with anti-5mC and anti-Dnmt3b antibodies. **f**, Levels of anti-Dnmt3b (left) and anti-5mC (right) immunofluorescence of $S3^{+/+}$ 2iLIF, $S3^{-/-}$ 2iLIF and three S3ER clones in 2i with or without TAM. The mean ± s.e.m. of $n = 3$ experiments per S3ER clone are shown, with individual data points represented as dots, squares or triangles. Scale bars, 20 μm. All P values were calculated using a two-tailed unpaired Student's t-test.

reported[14–17]. Surprisingly, wild-type (*Stat3*[+/+], subsequently denoted *S3*[+/+]) ESCs stably expanded in 2i without LIF or *S3*[−/−] cells in 2iLIF showed significantly higher levels of 5mC than *S3*[+/+] in 2iLIF, comparable to those of *S3*[+/+] cells in serum LIF (Fig. 1a). We performed mass spectrometry to unequivocally identify global 5mC and unmethylated cytosine, and confirmed that only *S3*[+/+] cells in 2iLIF had a reduced fraction of methylated cytosine (Fig. 1b). Reduced representation bisulfite sequencing (RRBS; Fig. 1c and Extended Data Fig. 1a) further confirmed our findings. We concluded that active LIF–Stat3 signaling is required, in combination with 2i, to induce genomic hypomethylation.

How does LIF–Stat3 regulate the levels of 5mC? We measured the expression levels of factors involved in 5mC deposition, maintenance and oxidation, and observed that *S3*[+/+] cells in 2iLIF medium had reduced expression only of de novo methyltransferases *Dnmt3a* and *Dnmt3b* (*Dnmt3a/b*) and increased expression of *Tet2* compared with *S3*[+/+] cells in 2i and *S3*[−/−] cells (Fig. 1d and Extended Data Fig. 1b). Such changes were confirmed at the protein level by western blotting and mass spectrometry (Fig. 1e,f and Extended Data Fig. 1c).

We then investigated whether the hypomethylation observed in 2iLIF medium was dependent on Dnmt3a/b. We generated two independent mutant clones for each genotype of *Dnmt3a* knockout (KO), *Dnmt3b* KO and *Dnmt3a/b* double KO (dKO) ESCs (Extended Data Fig. 1d). Wild-type ESCs (E14) cultured in 2i medium without LIF were hypermethylated, whereas the *Dnmt3a* KO and *Dnmt3b* KO cells displayed a partial reduction in 5mC levels relative to wild-type cells in 2i, and the *Dnmt3a/b* dKO cells cultivated in 2i were hypomethylated (Fig. 1g). Mass spectrometry (Fig. 1h) and RRBS (Fig. 1i and Extended Data Fig. 1e) analyses further confirmed hypomethylation of *Dnmt3a/b* dKO cells in 2i. Furthermore, overexpression of *Dnmt3a* and *Dnmt3b* in *S3*[+/+] cells in 2iLIF led to increased 5mC levels (Extended Data Fig. 1f,g). We concluded that the levels of Dnmt3a/b dictate the DNA methylation status of naive ESCs in 2i.

We also tested whether Tets could have a role in the hypomethylation observed in 2iLIF. *Tet1* and *Tet2* are both robustly expressed in ESCs in 2iLIF (Supplementary Table 1) and have redundant functions[23]. Therefore, we knocked down *Tet1* and *Tet2* simultaneously in *S3*[+/+] 2iLIF and observed no significant changes in 5mC levels (Extended Data Fig. 1h,i), in agreement with previous reports[23,24].

Impact of Stat3 on DNA methylation and transcription. We then investigated which genomic regions showed a DNA methylation profile dependent on Stat3. RRBS analyses of *S3*[+/+] and *S3*[−/−] cells in 2iLIF medium identified 381,607 sites that were differentially methylated, 98.7% of which displayed an increase in methylation in *S3*[−/−] cells (Fig. 2a), whereas the DNA methylation at repetitive elements

was unchanged (Extended Data Fig. 2a). We used H3K4me3 and H3K27ac chromatin immunoprecipitation-sequencing (ChIP–seq) profiles to identify promoters and enhancers in ESCs. We observed increased DNA methylation in *S3*[−/−] cells in 3.6% (323/8,782) of promoters and 36.5% (621/1,701) of enhancers (Fig. 2b,c and Supplementary Table 2), whereas only two of the 8,782 promoters and one of the 1,701 enhancers showed decreased DNA methylation. We intersected transcriptome data comparing *S3*[+/+] and *S3*[−/−] cells and investigated whether the gene associated with each promoter or enhancer was differentially expressed. The gain in DNA methylation at promoters was associated with downregulation of cognate genes in 20.7% (67/323) of the cases and with upregulation in 8% (26/323) of the cases (Fig. 2b). For enhancers, 13.8% (86/621) and 6.7% (42/621) of the cognate genes were significantly downregulated and upregulated, respectively (Fig. 2c). Among these, we noticed the pluripotency factors *Klf5* and *Esrrb*[25,26] (Fig. 2d). We repeated the same analyses comparing *S3*[+/+] in 2i and 2iLIF media (Fig. 2d and Extended Data Fig. 2b–f), and obtained highly comparable results, demonstrating that an absence of LIF or Stat3 had overlapping effects on the transcriptome and 5mC profile of ESCs.

Given that Dnmt3a/b seemed to be functionally relevant for the regulation of 5mC levels downstream of LIF (Fig. 1g–i), we expected similar changes in DNA methylation in response to LIF stimulation or following Dnmt3a/b inactivation. Strikingly, 98.9% of the regions hypomethylated in 2iLIF were also hypomethylated in Dnmt3a/b dKO cells (Fig. 2e and Extended Data Fig. 2g–k), further indicating that Dnmt3a/b are epistatic to LIF–Stat3 for DNA methylation control.

Imprinted genes are organized in clusters under the control of differentially methylated regions (DMRs)[27,28]. *S3*[+/+] cells cultured long term in 2iLIF medium showed reduced levels of DNA methylation at imprinted DMRs, as reported[14,29], whereas *S3*[−/−] cells retained higher levels of DNA methylation at 83.3% (20/24) of the DMRs analyzed (Fig. 2f). These findings were confirmed by methylated DNA immunoprecipitation followed by quantitative PCR (MeDIP–qPCR; Fig. 2g and Extended Data Fig. 3a).

We checked the expression of transcripts associated with DMRs and found that 37% (20/54) were differentially expressed in *S3*[−/−] cells (Fig. 2h and Extended Data Fig. 3b). Considering all expressed imprinted genes, without taking into account DNA methylation information, 50.77% (33/65) were affected by Stat3 inactivation, whereas only 18.77% (2,349/12,510) of all expressed genes were differentially expressed, indicating that imprinted genes are preferentially regulated by Stat3 (Extended Data Fig. 3c; $P = 5.67 \times 10^{-9}$, hypergeometric test).

We concluded that Stat3 regulates DNA methylation at promoters, enhancers and imprinted DMRs, with a concomitant change in expression of a fraction of their associated transcripts.

Fig. 4 | Stat3 controls DNA methylation via metabolic regulation. a, Anti-5mC immunofluorescence on *S3*[+/+] and *S3*[−/−] cells as well as the clones MitoS3.A and MitoS3.B (see Extended Data Fig. 4d–h). Violin plots of an average of 55 nuclei per sample; $n = 3$ experiments. **b**, Percentage of 5mC in the DNA of *S3*[+/+] and *S3*[−/−] cells as well as the MitoS3 clones in 2iLIF, calculated from mass spectrometry data. Mean ± s.d. of $n = 4$ biological replicates. **c**, Imprinted transcripts that were differentially expressed (RNA-seq; q-value < 0.1) between *S3*[−/−] cells and both MitoS3 clones. The mean values of $n = 2$ biological replicates were scaled and are represented as z-scores. **d**, Metabolite abundance, determined using liquid chromatography with tandem mass spectrometry (LC–MS/MS). Mean ± s.e.m. of $n = 5$ biological replicates. **e**, Metabolic-tracing analysis of αKG, OAA and fumarate. Bars represent the percentage of marked carbon after provision of ^{13}C-glucose, -glutamine or -palmitate for 2, 4 and 8 h; mean ± s.e.m. of $n = 6$ biological replicates. **f**, Metabolic-tracing analysis of OAA, citrate and αKG. Bars, percentage of labeled isotopomer 2, 4 and 8 h after exposure to [U-^{13}C$_5$]glutamine. Black circles, ^{13}C$_5$-labeled carbons. Mean ± s.e.m. of $n = 6$ biological replicates. **g,h**, Abundance of αKG, measured by mass spectrometry in *S3*[+/+], *S3*[−/−] and MitoS3 clones (**g**) or in *S3*[+/+] cells cultured in 2iLIF with glutamine, without glutamine or without glutamine and supplemented with 2 mM αKG or 2 mM DM-αKG for 9 d (**h**). Mean ± s.e.m. of $n = 5$ biological replicates. **h**, *S3*[+/+] cells were cultured in 2iLIF with glutamine, without glutamine or without glutamine and supplemented with 2 mM αKG or 2 mM DM-αKG for 9 d. **i**, Anti-5mC immunofluorescence of *S3*[+/+] cells cultured with glutamine, without glutamine or without glutamine and supplemented with 2 mM DM-αKG for 9 d (left). Fluorescence intensity of an average of 96 nuclei per sample; $n = 3$ experiments. Representative images are shown with the insets showing magnified images of the regions in the yellow boxes (right). All violin and boxplots indicate the 1st, 2nd and 3rd quartiles; the whiskers indicate the minimum and maximum values. All P values were calculated using a two-tailed unpaired Student's t-test. **b,d,e,g,h**, Each replicate is shown as a dot. Scale bars, 20 μm.

Stat3 controls DNA methylation via metabolic regulation. Next, we wanted to study the dynamics of LIF-induced effects on Dnmt3a/b and 5mC. We performed quantitative reverse-transcriptase PCR (RT–qPCR) on *S3*[+/+] cells stably cultured in 2i (Fig. 3a) or acutely stimulated with LIF for 1, 4, 24 or 48 h, starting from 2i. The addition of LIF resulted in repression of *Dnmt3a/b*, but only after 24 h. Consistent with this, we observed a mild decrease in 5mC levels

after 24 h, whereas the levels of 5mC at 48 h were as low as those of cells stably cultured in 2iLIF (Fig. 3b).

The slow kinetics observed could indicate that Stat3 does not directly repress the transcription of *Dnmt3a/b*. We interrogated the available Stat3 ChIP–seq data and could not detect binding at enhancers or promoters of *Dnmt3a/b*[30]. Furthermore, we expressed a Stat3 construct fused to an estrogen receptor domain (S3ER)[31],
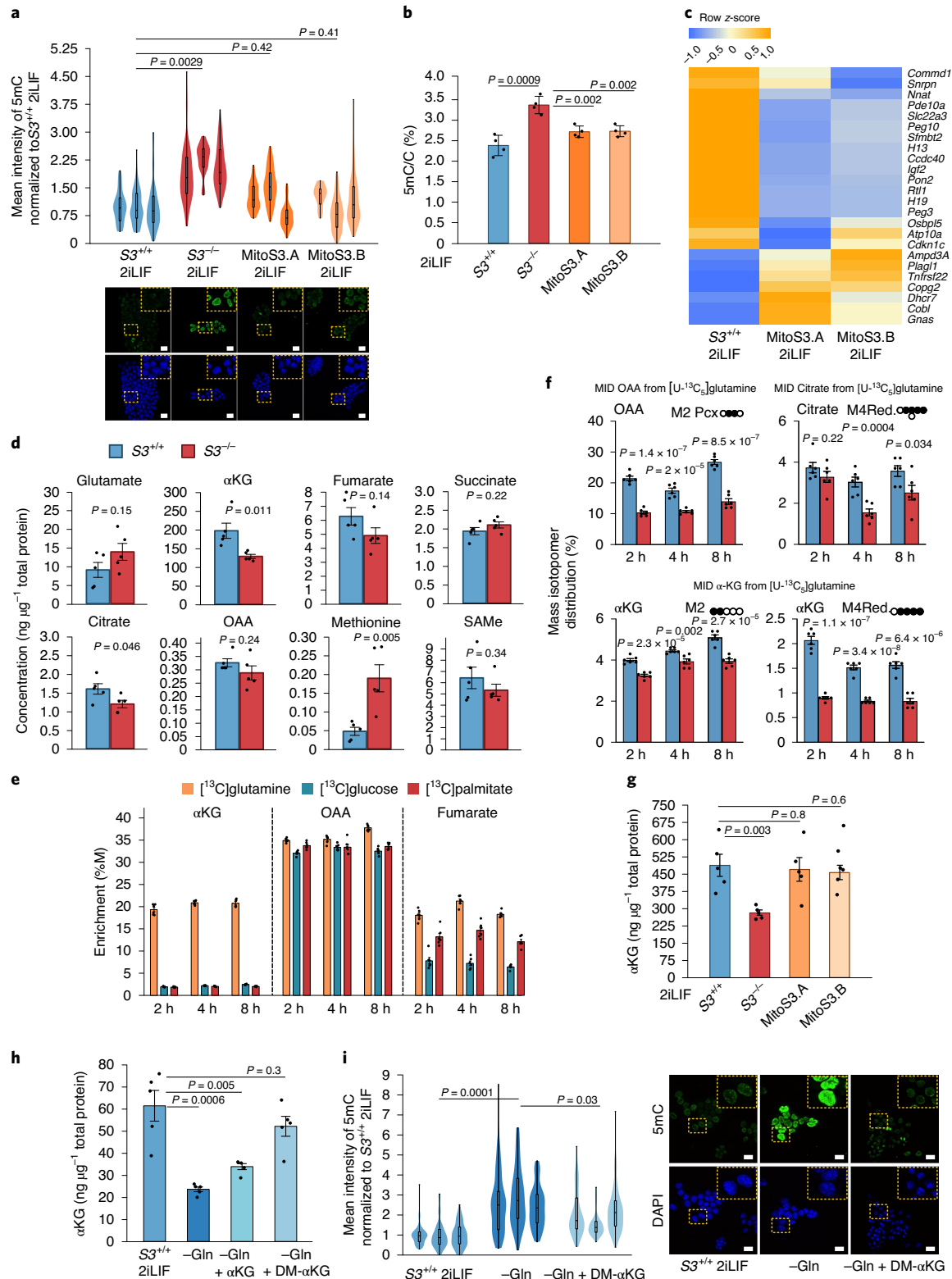
which localizes to the nucleus and activates the direct Stat3 targets *Socs3* and *Klf4* following tamoxifen (TAM) treatment in *S3⁻/⁻* cells (refs. [31,32] and Fig. 3c,d). The levels of *Dnmt3a/b* messenger RNA and protein as well as the levels of 5mC were unchanged (Fig. 3d–f), indicating that nuclear Stat3 does not regulate DNA methylation in ESCs.

Given that Stat3 did not regulate *Dnmt3a/b* expression through a direct transcriptional mechanism, we searched for alternative mechanisms. Global 5mC levels could be affected by passive dilution occurring during genome replication. The LIF–Stat3 axis promotes ESCs proliferation[19]; thus, the reduced 5mC levels in *S3⁺/⁺* cells in 2iLIF medium could be due to enhanced genome replication associated with proliferation. We tested this hypothesis by performing a 5-ethynyl-2′-deoxyuridine (EdU) incorporation assay combined with 5mC immunostaining. For *S3⁻/⁻* cells in 2iLIF or *S3⁺/⁺* cells in 2i, we observed that EdU⁺ cells that underwent genome replication within the last 4 h showed an expected decrease in 5mC compared with EdU⁻ cells (Extended Data Fig. 4a), which never reached the levels of *S3⁺/⁺* in 2iLIF. Moreover, the 5mC levels in *S3⁺/⁺* cells in 2iLIF were unaffected by the EdU status, indicating that differences in cell proliferation can account only in part for the decrease in 5mC induced by LIF–Stat3.

We then hypothesized that LIF–Stat3 might control 5mC levels via the regulation of mitochondrial activity, given that it has previously been reported that *S3⁻/⁻* cells display reduced mitochondrial OXPHOS[19,20]. First, we treated *S3⁺/⁺* cells with inhibitors of complexes I and III of the respiratory chain at concentrations that reduce OXPHOS in ESCs[19]; we observed a strong increase in the 5mC signal (Extended Data Fig. 4b). Importantly, *Dnmt3a/b* dKO cells showed no significant increase in 5mC following inhibition of the respiratory chain (Extended Data Fig. 4c), further indicating that changes in 5mC are dependent on Dnmt3a/b in ESCs. Second, we expressed a Stat3 construct targeted to mitochondria at endogenous levels in *S3⁻/⁻* cells (Extended Data Fig. 4d–g)[19,33]. The two clonal lines, called MitoS3.A and MitoS3.B, showed increased OXPHOS (Extended Data Fig. 4h) and reduced 5mC (Fig. 4a,b) compared with parental *S3⁻/⁻* cells.

Given that expression of imprinted transcripts is linked to 5mC levels on imprinted DMRs, we measured their expression in *S3⁻/⁻* and MitoS3 cells, and found that 25 were differentially expressed (Fig. 4c). We concluded that the hypomethylation observed in 2iLIF is linked to robust OXPHOS of ESCs.

Mitochondrial activity affects the methylation profile of the nuclear genome, which implies that the two organelles communicate. We initially hypothesized that intracellular signaling molecules, such as calcium ions or reactive oxygen species could be implicated, but we did not observe differences in their abundance between *S3⁺/⁺* and *S3⁻/⁻* cells. We then reasoned that mitochondrial activity could influence the abundance of metabolites serving as donors, acceptors or cofactors of DNA methylation and oxidation[22]. Analysis of the steady-state levels of metabolites revealed a decrease in α-ketoglutarate (αKG) in *S3⁻/⁻* cells (Fig. 4d). We also

noticed a marked increase in the methionine levels of *S3⁻/⁻* cells; however, neither *S*-adenosyl methionine (SAMe)—the actual donor of methyl groups to DNA and histones—nor the enzymes involved in methionine/SAMe metabolism—such as MAT2A, MAT2B, AHCY and MTR—differed between *S3⁺/⁺* and *S3⁻/⁻* cells (Supplementary Table 3), indicating that methionine/SAMe metabolism might not be involved in DNA-methylation regulation downstream of Stat3.

We then investigated what carbon source was preferentially used by ESCs to produce αKG and performed metabolic flux analyses with ¹³C-labeled glucose, glutamine and palmitate (see Methods). As reported[34], glutamine represented the main source for production (Fig. 4e). Glutamine is directly converted into glutamate and αKG, which in turn enters the tricarboxylic acid cycle (TCA) for energy production via oxidative metabolism. Analysis of specific isotopomers revealed a decrease in the oxidative glutamine pathway and TCA activity in *S3⁻/⁻* cells (Extended Data Fig. 5a,b; oxaloacetate (OAA) M4 and citrate M4), in line with impaired OXPHOS (Extended Data Fig. 4h). In addition, in the reductive glutamine pathway, glutamine is converted in the cytosol into TCA intermediates and acetyl-CoA, which is diverted to fatty-acid biosynthesis. The OAA obtained from cytosolic citrate cleavage is converted to aspartate or malate and then pyruvate, which feeds the TCA either directly or via conversion into OAA by pyruvate carboxylase. The glutamine reductive pathway was impaired in *S3⁻/⁻* cells. Specifically, we detected a strong decrease in cytosolic OAA M3 and mitochondrial OAA M2 Pcx, citrate M2 and M4, αKG M2 and M4 (Fig. 4f and Extended Data Fig. 5a,b). In agreement with these observations, we found decreased expression of cytoplasmic *isocitrate dehydrogenase 1* (*Idh1*; Extended Data Fig. 5c). These data indicate that reduced αKG levels detected in *S3⁻/⁻* cells are due to decreased flux of carbon from glutamine reductive metabolism re-entering the TCA cycle.

We hypothesized that robust αKG production from glutamine is required for genome hypomethylation. We measured the levels of αKG in cells expressing Stat3 only in mitochondria. Both the MitoS3.A and MitoS3.B clones showed elevated αKG levels, which did not differ significantly from *S3⁺/⁺* cells (Fig. 4g). Both clones also showed reduced 5mC (Fig. 4a,b), further indicating that elevated αKG levels correlate with reduced DNA methylation.

To functionally test our hypothesis, we cultured *S3⁺/⁺* cells in 2iLIF in the absence of glutamine. We first measured the endogenous levels of αKG in cells cultured in the absence of glutamine, and found it to be strongly reduced (Fig. 4h) and observed a robust increase in 5mC (Fig. 4i). Next, we investigated whether restoration of endogenous αKG levels could result in reduced 5mC. We added back a cell-permeable form of αKG (dimethyl 2-oxoglutarate, DM-αKG) and were able to reduce the 5mC levels (Fig. 4h,i). Note that DM-αKG has been reported to stabilize hypoxia-inducible factor 1-α by inducing a pseudohypoxic state[35], but this was not the case in ESCs (Extended Data Fig. 5d). In summary, our results indicate that efficient αKG production from glutamine induces low methylation levels of the nuclear genome.

Fig. 5 | Alpha-ketoglutarate regulates 5mC mainly by controlling the levels of Dnmt3a/b. a–g, *S3⁺/⁺* and *S3⁻/⁻* cells and two MitoS3 clones, cultured in 2iLIF, were analyzed. **a**, Ratios of αKG/fumarate (left) and αKG/succinate (right). Mean ± s.e.m. of *n* = 5 biological replicates. **b**, Percentages of h5mC. Mean ± s.e.m. of *n* = 4 biological replicates. **c**, Ratio of h5mC/5mC. Mean ± s.e.m. of *n* = 5 biological replicates. **a–c**, Data from mass spectrometry measurements. **d,e**, Expression analysis by RNA-seq (**d**) and qPCR (**e**) of enzymes controlling DNA methylation. **d**, Heatmap showing the *z*-scores from the scaled RNA-seq expression values; *n* = 2 biological replicates. **e**, Mean ± s.e.m. of *n* = 6 RT-qPCR experiments. **f**, Western blot for Dnmt3a, Dnmt3b and Lamin B, which was used as a loading control. Representative images of *n* = 2 independent experiments. **g**, Proteomic analysis. The yellow and blue dots indicate proteins that are more (> +1 difference) or less (< −1 difference) abundant, respectively (*P* < 0.05, as indicated by the dashed lines), in *S3⁻/⁻* cells relative to MitoS3.A cells; *n* = 5 biological replicates. Source data in Supplementary Table 3. **h,i**, Gene expression analysis (RT-qPCR) of epigenetic modifiers (**h**) and imprinted genes (**i**) in *S3⁺/⁺* and *S3⁻/⁻* as well as *S3⁻/⁻* cells cultured in 2iLIF and treated with 2 mM DM-αKG for four passages. Mean ± s.e.m. of *n* = 4 experiments. **j**, Intracellular αKG abundance in *S3⁺/⁺* cells treated with vehicle and *S3⁻/⁻* cells treated with vehicle or 2 mM DM-αKG for 24 h (dark bars) or for three passages (light bars). Mean ± s.e.m. of *n* = 6 biological replicates. **a–c,e,h–j**, Each replicate is shown as a dot. All *P* values were calculated using a two-tailed unpaired Student's *t*-test.

Dnmt3a/b expression is regulated by αKG via the transcription factor Otx2. We next investigated how αKG reduces the levels of 5mC in ESCs. Given that αKG functions as a cofactor for Tet oxidases[22,36,37], $S3^{-/-}$ cells might display increased 5mC due to reduced Tet activity. In addition, it has been recently reported that the abundance of αKG

inversely correlates with the expression levels of $Dnmt3a/b$[38]. Thus, $S3^{-/-}$ cells could show increased $Dnmt3a/b$ expression and 5mC levels because of reduced αKG levels. To investigate the relative contribution of these two possible mechanisms, we took advantage of our MitoS3 cells, where the levels of αKG were rescued to endogenous
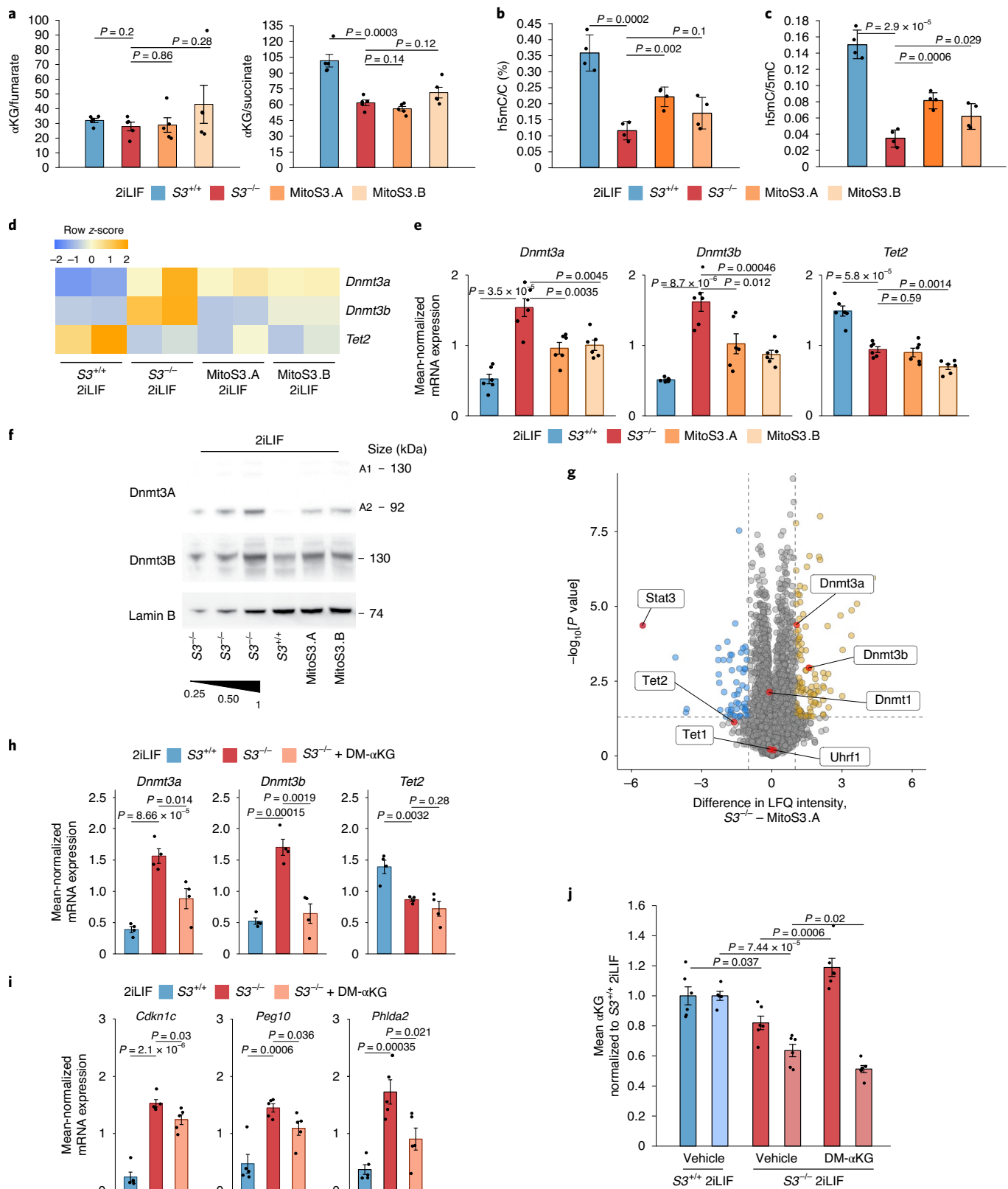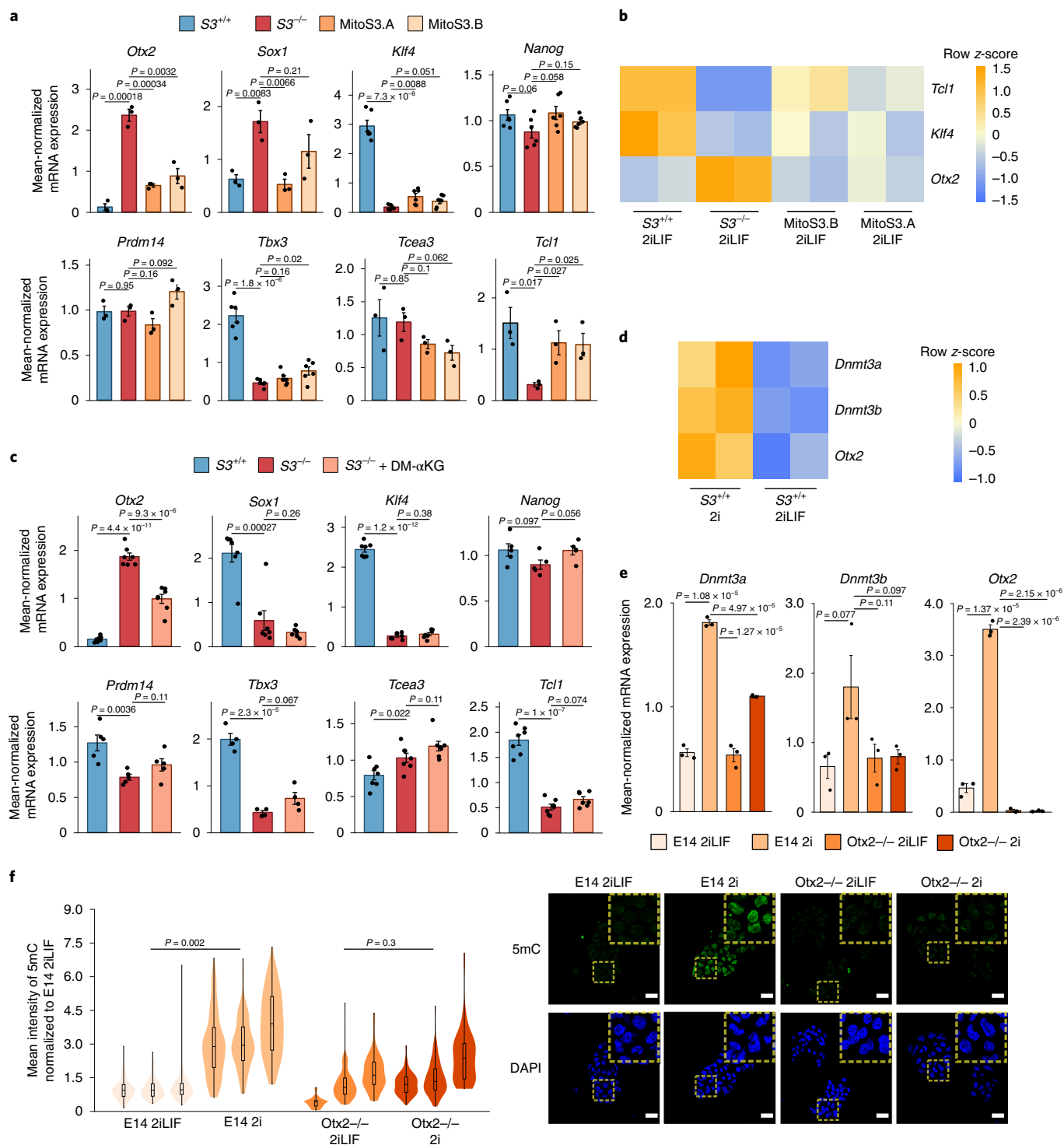
Fig. 6 | Otx2 links αKG to Dnmt3a/b expression. a,b, Expression analysis of $S3^{+/+}$ and $S3^{-/-}$ cells as well as MitoS3 clones, cultured in 2iLIF medium, for potential Dnmt3a/b regulators. **a,** RT–qPCR analyses. Mean ± s.e.m. of $n = 3$ (Otx2, Sox1, Prdm14, Tcea3 and Tcl1) and $n = 6$ (Klf4, Nanog and Tbx3) experiments. **b,** Heatmap of RNA-seq data. Z-scores of scaled expression values; $n = 2$ biological replicates. **c,** RT–qPCR of $S3^{+/+}$ and $S3^{-/-}$ cells as well as $S3^{-/-}$ cells cultured in 2iLIF and treated with 2 mM DM-αKG for three passages. Mean ± s.e.m. of $n = 7$ (Otx2, Sox1, Klf4, Tcea3 and Tcl1), $n = 5$ (Nanog and Prdm14) and $n = 4$ (Tbx3) experiments. **d,** Heatmap of RNA-seq data reporting the expression levels of Dnmt3a, Dnmt3b and Otx2 in $S3^{+/+}$ cells cultured in 2i with or without LIF. The expression levels were scaled and are represented as z-scores; $n = 2$ biological replicates. **e,** RT–qPCR of E14 and $Otx2^{-/-}$ cells stably cultured in 2iLIF or 2i medium. Mean ± s.e.m. of $n = 3$ experiments. **f,** Anti-5mC immunofluorescence of E14 and $Otx2^{-/-}$ cells stably cultured in 2iLIF or 2i. Representative images (right) and violin plots (left) of the fluorescence intensity of an average of 111 nuclei per sample. The boxplots show 1st, 2nd and 3rd quartiles; the whiskers indicate the minimum and maximum values; $n = 3$ experiments. **a,c,e,** Each replicate is shown as a dot. All P values were calculated using a two-tailed unpaired Student's t-test. Scale bars, 20 μm.

levels (Fig. 4g) and 5mC were decreased (Fig. 4a,b), without potentially confounding effects from nuclear Stat3.

Elevated αKG levels or, more precisely, high αKG/succinate and αKG/fumarate ratios are associated with increased Tet activity[34,39–41]. We found no significant differences in the αKG/fumarate ratio, while the αKG/succinate ratio was equally low in $S3^{-/-}$ and MitoS3 cells relative to $S3^{+/+}$ cells (Fig. 5a). We then measured the levels of h5mC and 5mC, and used their ratio as a direct readout of Tet activity, which was low in both $S3^{-/-}$ and MitoS3.A/B cells compared with $S3^{+/+}$ cells (Fig. 5b,c). These results indicate that Tet activity is similarly low in $S3^{-/-}$ and MitoS3 cells, thereby indicating that the differences in 5mC between $S3^{-/-}$ and MitoS3 cells cannot be attributed to Tets (Fig. 4a,b). Such conclusions are in line with the lack of effect on 5mC observed following Tet1/2 knockdown (Extended Data Fig. 1h,i).

Next, we measured the mRNA and protein levels of Dnmt3a/b by RNA sequencing (RNA-seq), RT–qPCR, western blotting and proteomic analyses and found that both the mRNA and protein levels of Dnmt3a/b were reduced in MitoS3.A and MitoS3.B cells compared with $S3^{-/-}$ cells (Fig. 5d–g), indicating that αKG could repress *Dnmt3a/b* expression.

We tested whether αKG negatively regulates *Dnmt3a/b* expression by treating $S3^{-/-}$ cells with DM-αKG and observed a partial reduction in the expression of *Dnmt3a/b* (Fig. 5h) and imprinted genes (Fig. 5i). Such partial effects are probably due to the inability of DM-αKG to stably rescue endogenous αKG levels in $S3^{-/-}$ cells (Fig. 5j).

We concluded that αKG decreases the levels of 5mC through the reduction of *Dnmt3a/b* expression, in agreement with genetic perturbations showing that the levels of Dnmt3a/b dictate 5mC abundance (Fig. 1g–i and Extended Data Fig. 1f,g).

To clarify the molecular mechanism by which αKG controls the expression of *Dnmt3a/b*, we explored literature and analyzed a transcriptomic database[42–44] of a large number of mouse ESC lines in which single transcriptional regulators were either induced or repressed. From this survey, we identified two activators (*Otx2* and *Sox1*)[45,46] and six repressors (*Klf4*, *Nanog*, *Prdm14*, *Tbx3*, *Tcea3* and *Tcl1*)[47–51] of Dnmt3a/b.

We thus checked the expression levels of our candidate regulators in $S3^{+/+}$ and $S3^{-/-}$ cells, and observed that the activators *Otx2* and *Sox1* were upregulated in $S3^{-/-}$ cells, whereas the repressors *Klf4*, *Tbx3* and *Tcl1* were downregulated in $S3^{-/-}$ cells (Fig. 6a). Given that mitochondrial Stat3 expression increased the endogenous αKG levels and reduced *Dnmt3a/b* expression (Fig. 5d,e), we measured the expression of Dnmt3a/b potential regulators in MitoS3 cells and observed that the expression of only *Otx2*, *Klf4* and *Tcl1* was significantly affected by mitochondrial Stat3 (Fig. 6a,b). Finally, we treated $S3^{-/-}$ cells with DM-αKG and only *Otx2* expression was significantly altered (Fig. 6c).

Is Otx2 functionally required for Dnmt3a/b regulation downstream of LIF–Stat3? Wild-type ESCs in 2i medium showed a marked increase in *Otx2* and *Dnmt3a/b* expression (Fig. 6d,e) relative to 2iLIF. This was accompanied by increased 5mC levels (Fig. 6f). If Otx2 is in fact crucial for *Dnmt3a/b* regulation, its

genetic inactivation should render cells unable to upregulate *Dnmt3a/b* in the absence of LIF. We cultured $Otx2^{-/-}$ ESCs[52] in 2i or 2iLIF medium and observed that *Dnmt3a/b* expression, as well as the 5mC levels, were unchanged.

We concluded that Otx2 is regulated by the LIF–Stat3–αKG axis and that Otx2 is genetically required to boost the levels of Dnmt3a/b and 5mC in ESCs.

Mitochondrial Stat3 regulates ESC differentiation. Mitochondrial Stat3 reduces the levels of Otx2, Dnmt3a/b and 5mC, which are all molecular markers of early phases of ESC differentiation[4,23,37,53–58]. Therefore, we hypothesized that mitochondrial Stat3 might stabilize pluripotency and/or slow down differentiation.

We generated transcriptomic data of cells either in 2iLIF medium or undergoing differentiation for 48 h in N2B27 basal medium. The genes that were significantly downregulated in $S3^{-/-}$ cells relative to $S3^{+/+}$ (Fig. 7a, blue) were also downregulated during differentiation of $S3^{+/+}$ cells (Fig. 7b). Several naive pluripotency markers belong to this category (Fig. 7d). Similarly, genes found to be upregulated in $S3^{-/-}$ cells relative to $S3^{+/+}$ (Fig. 7a, orange) were upregulated in $S3^{+/+}$ cells during differentiation (Fig. 7c), including several early differentiation markers and imprinted genes (Fig. 7e). Notably, $S3^{-/-}$ cells displayed a much faster upregulation of early differentiation and imprinted genes after 48 h in N2B27 (Fig. 7e,f and Extended Data Fig. 6a–d). Furthermore, we performed a clonal assay of cells undergoing differentiation for up to 72 h. In 2iLIF, the number of alkaline phosphatase positive (AP+) pluripotent colonies formed by $S3^{-/-}$ cells was slightly reduced relative to $S3^{+/+}$ (Fig. 7g). However, after 24 and 48 h of 2iLIF withdrawal, $S3^{-/-}$ cells formed dramatically fewer AP+ colonies than $S3^{+/+}$ cells. Both cell lines eventually lost clonogenicity after 72 h of 2iLIF withdrawal. Based on the faster transcriptional changes and the faster loss of AP+ colonies, we concluded that $S3^{-/-}$ cells exit more rapidly from the naive pluripotent state.

These results could be explained by the canonical role of LIF and nuclear Stat3 as a transcriptional inducer of naive pluripotency described by several laboratories[9,11,59,60] rather than by mitochondrial Stat3. We therefore investigated whether expression of Stat3 in mitochondria only would affect ESC differentiation. Transcriptionally, we observed that genes downregulated in $S3^{-/-}$ cells were only slightly affected by mitochondrial Stat3 (Fig. 7b,d). Conversely, early-differentiation markers and imprinted genes were strongly reduced in the MitoS3.A/B cells (Fig. 7c,e,f and Extended Data Fig. 7a–d). In clonal assays, mitochondrial Stat3 delayed the exit from naive pluripotency (Fig. 7g). In conclusion, mitochondrial Stat3 represses early differentiation markers and delays differentiation of ESCs.

Stat3 regulates Dnmts and imprinted transcripts in early mouse blastocysts. We then tested the function of Stat3 in the early mouse embryo. We focused our attention on the early blastocysts at E3.5 because Stat3 is active[7,8,10] and the genome is hypomethylated[4,5] in naive pluripotent cells of the inner cell mass (ICM) at

Fig. 7 | Mitochondrial Stat3 regulates ESC differentiation. a–e, Transcriptome analysis by RNA-seq. **a**, Genes that were differentially expressed (log$_2$[fold change (FC)] > +1 or < −1; *q*-value < 0.01, Benjamini–Hochberg adjustment, as indicated by dashed lines) between $S3^{+/+}$ and $S3^{-/-}$ cells; $n = 2$ biological replicates. **b,c**, Expression levels of genes that were down- (**b**) and upregulated (**c**) in $S3^{-/-}$ cells relative to $S3^{+/+}$ cells. Each boxplot shows the 1st, 2nd and 3rd quartiles; the whiskers show the minimum and maximum values. The cells were analyzed in 2iLIF and after 48 h of 2iLIF withdrawal (−48 h). Dashed lines indicate the median expression in $S3^{+/+}$ cells in 2iLIF. Inset: mean log$_2$[FC] relative to $S3^{+/+}$ 2iLIF. **d,e**, Heatmap of makers of naive pluripotency (**d**) as well as early differentiation and imprinted genes (**e**). The expression levels (from RNA-seq data) were scaled and are represented as z-scores; $n = 2$ biological replicates. **f**, Gene expression analysis by RT–qPCR of $S3^{+/+}$ and $S3^{-/-}$ cells as wells as two MitoS3 clones cultured in 2iLIF or without 2iLIF for 24 (−24 h) or 48 h. Mean ± s.e.m. of $n = 3$ experiments, with each replicate shown as a dot. See Extended Data Fig. 6. **g**, Alkaline phosphatase staining of $S3^{+/+}$ and $S3^{-/-}$ cells as well as MitoS3 clones cultured with 2iLIF or without 2iLIF for 24, 48 or 72 h. Representative images (left) and quantification of AP+ colonies (right), relative to $S3^{+/+}$ cells in 2iLIF. Mean ± s.e.m. of $n = 3$ experiments, with each replicate shown as a dot. *P* values were calculated using a two-tailed unpaired Student's *t*-test.

this stage. At E3.75, ICM cells are specified into pluripotent epiblast cells (Epi) or extra-embryonic primitive endoderm cells (PrE). Stat3-heterozygous mice were crossed and blastocysts were collected at E3.5 and E3.75. Trophectoderm cells were removed by immunosurgery and used to genotype individual embryos. Single ICM, Epi and PrE cells were analyzed by RNA-seq (Fig. 8a).
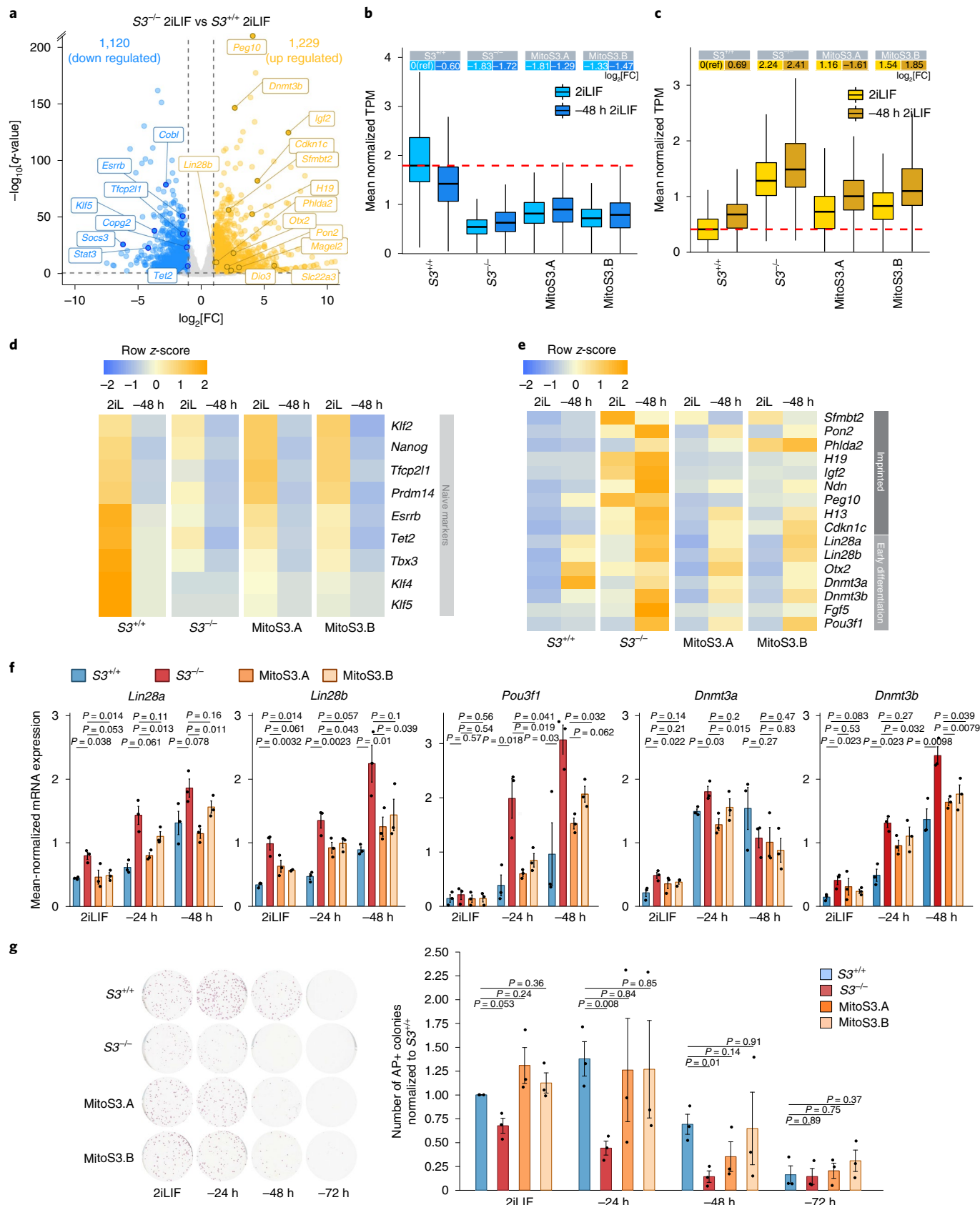
Global analysis by *t*-distributed stochastic neighbor embedding (*t*-SNE) revealed that the transcriptomes of *S3*⁺/⁺ and *S3*⁻/⁻ ICM cells were divergent at E3.5 (Fig. 8b). Differential expression analysis revealed that Stat3 and its targets *Socs3* and *Tfcp2l1* were downregulated in *S3*⁻/⁻ cells (Fig. 8c, Extended Data Fig. 8a and Supplementary Table 4). Notably, *Dnmt3a/b* were upregulated in
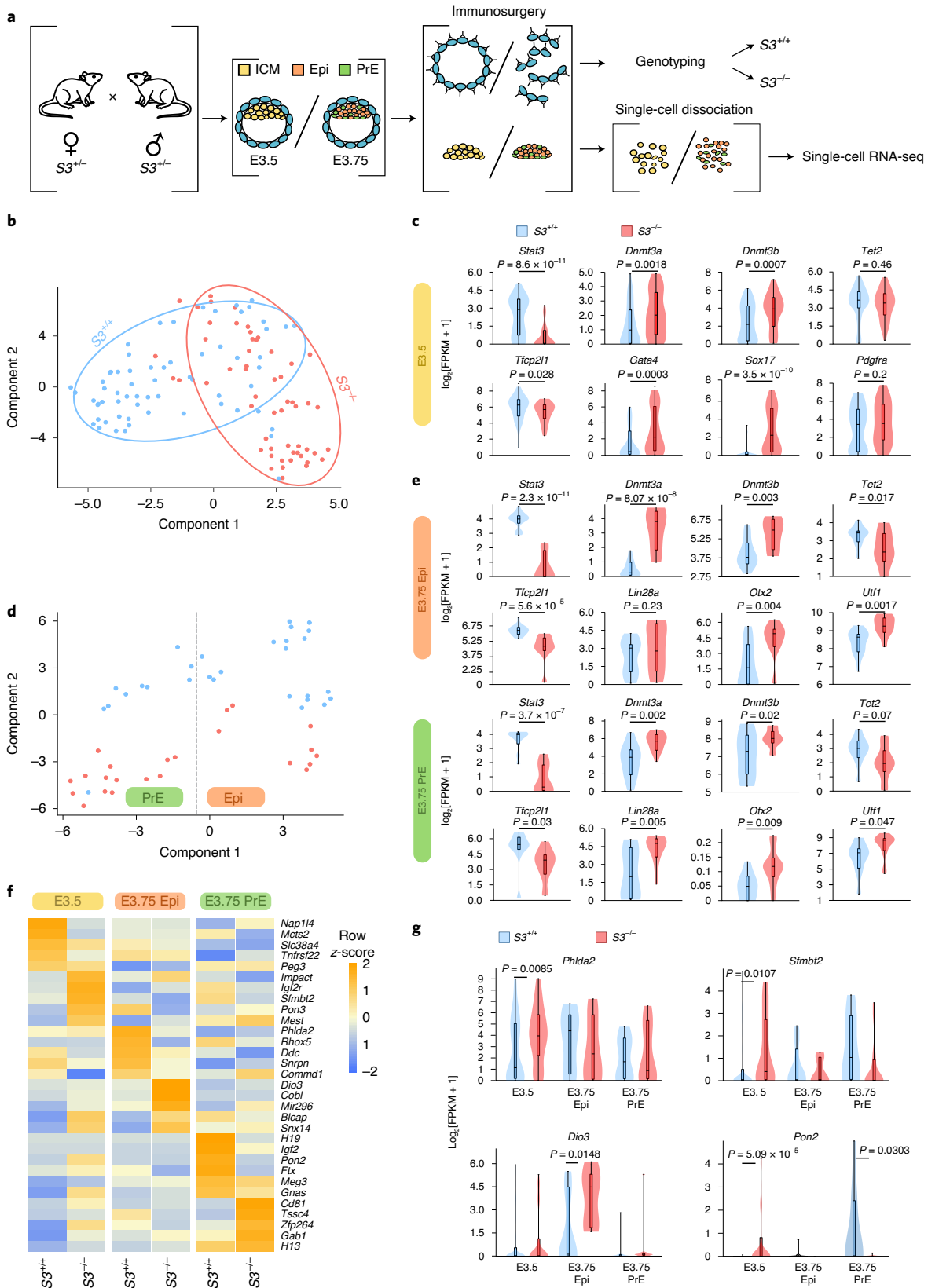
Fig. 8 | Stat3 regulates Dnmts and imprinted transcripts in early mouse blastocysts. a, Outline of the strategy used for the isolation and profiling of single pluripotent cells. A total of 171 cells from 18 embryos were analyzed. See Supplementary Table 4. **b**, Single-cell RNA-seq *t*-SNE based on the whole transcriptome of wild-type (*S3*⁺/⁺) and mutant (*S3*⁻/⁻) ICM cells collected at E3.5. **c**, Distribution of the expression levels at E3.5 for selected markers, determined using single-cell RNA-seq. **d**, Single-cell RNA-seq *t*-SNE based on genome-wide expression of *S3*⁺/⁺ and *S3*⁻/⁻ mouse cells collected at E3.75. **b,d**, Each dot represents a single cell. **e**, Distribution of the gene expression levels of the indicated markers in Epi (top) and PrE (bottom) cells collected at E3.75. **f**, Heatmap reporting average expression levels, determined using single-cell RNA-seq, of imprinted transcripts in three different embryonic populations (E3.5 ICM, E3.75 Epi and E3.5 PrE) from *S3*⁺/⁺ and *S3*⁻/⁻ embryos. The expression values were scaled and are represented as *z*-scores. Only expressed imprinted genes (average fragments per kilobase of transcript per million mapped reads (FPKM) > 1) were analyzed. **g**, Distribution of the expression of imprinted genes. In all violin plots, the boxplots show 1st, 2nd and 3rd quartiles, and the whiskers indicate minimum and maximum values. *P* values were calculated using a two-tailed unpaired Student's *t*-test.

S3⁻/⁻ ICM cells. Moreover, genes that are specifically expressed in PrE, such as *Sox17*, *Gata4* and *Pdgfra*, were precociously activated in *S3*⁻/⁻ ICM cells (Fig. 8c and Extended Data Fig. 8a,b).

Cells from E3.75 embryos could be classified as Epi or PrE according to specific markers (Fig. 8d and Extended Data Fig. 8c). In *S3*⁻/⁻ embryos, *Socs3* and *Tfcp2l1* were significantly reduced in Epi cells at E3.75 (Fig. 8e and Extended Data Fig. 8d). Interestingly, markers of post-implantation epiblasts, such as *Utf1*, *Otx2* (refs. [7,61]) and *Dnmt3a/b*, were also upregulated in *S3*⁻/⁻ cells, whereas genes associated with E3.5 ICM were downregulated (Fig. 8e and Extended Data Fig. 8e), further indicating accelerated developmental progression. Finally, we analyzed all imprinted transcripts expressed at E3.5 and E3.75, and observed global deregulation (Fig. 8f,g).

These results indicate that Stat3 regulates the expression of *Otx2* and its target genes *Dnmt3a/b*, and imprinted transcripts in the preimplantation blastocyst, indicating an altered pace of developmental progression.

Discussion

Mouse ESCs cultured in 2iLIF display low levels of 5mC and Dnmt3a/b, similarly to naive pluripotent cells of the preimplantation blastocyst-stage embryo. The hypomethylation observed in 2iLIF was previously attributed to the MEK inhibitor, which represses *Dnmt3a/b* expression[14,15] via the upregulation of *Prdm14* (refs. [17,42,43]).

Our results indicate that mitochondrial Stat3 is also necessary for Dnmt3a/b downregulation in 2iLIF, given that Stat3-null cells in 2iLIF displayed high Dnmt3a/b levels, despite the presence of the MEK inhibitor, and expression of a mitochondrial Stat3 is sufficient to reduce the Dnmt3a/b and 5mC levels. *Prdm14* expression was unaffected by Stat3 (Fig. 6a), indicating that Stat3 and Prdm14 are two independent negative regulators of Dnmt3a/b in 2iLIF.

Previous studies linked the hypomethylation of naive ESCs in 2iLIF medium to reduced de novo DNA methylation activity[14,15,17,48,49,62], in agreement with our genetic perturbation data. In contrast, the maintenance DNA methyltransferase Dnmt1 shows similar mRNA levels, protein levels and activity both in 2iLIF and serum LIF media[14,15,17], indicating that the hypomethylation in 2iLIF is not due to reduced Dnmt1 activity. However, a recent study found that 2iLIF reduces protein levels of the Dnmt1 cofactor Uhrf1 and its locus-specific recruitment, leading to reduced DNA methylation maintenance on specific H3K9me2-enriched loci[24].

Our results demonstrate that, in addition to reduced Dnmt1 activity, mitochondrial Stat3-dependent reduction of Dnmt3a/b is also crucial to induce genomic hypomethylation in 2iLIF.

In the embryo, *Otx2* and *Dnmt3a/b* are expressed robustly only after implantation at E5.5 (ref. [7]), whereas Stat3 is active only in the preimplantation blastocyst[7]. Genetic inactivation of *Stat3* led to precocious expression of *Otx2* and *Dnmt3a/b*, suggesting that Stat3 is needed to temporally restrict the expression of the post-implantation transcriptional program. Stat3-null embryos fail soon after implantation[63]. It would be interesting to test whether such embryonic

lethality is due to accelerated development of Stat3-null embryos relative to maternal tissues.

Stat3 regulates imprinted gene expression in ESCs and in the early embryo. Several imprinted genes (for example, *Ddc*, *Gab1*, *Commd1*, *Cobl* and *Cd81*) have been shown to regulate ESC differentiation[64], suggesting that a balanced expression of imprinted genes could be critical for correct developmental progression.

Stat3 has been shown to act in different cellular compartments, such as the nucleus, mitochondria and endoplasmic reticulum[20,21,65], and Stat3 has been implicated in the control of epigenetic modifications during somatic cell reprogramming[31,66]. It will be interesting to determine from what organelle Stat3 regulates such processes.

Efficient conversion of αKG into glutamine in mouse ESCs has been attributed to 2i[34]. Moreover, αKG production from glutamine via Psat1 has been reported to decrease during ESC differentiation[67]. Our results complement and expand such studies, showing that glutamine is a major source of αKG production and that Stat3 predominantly regulates the reductive glutamine pathway, overall indicating that multiple metabolic pathways allow the interconversion of αKG and glutamine, in line with the critical roles of both metabolites in fundamental processes such as proliferation, epigenetic regulation and differentiation[34,38,68].

Notably, *S3*⁻/⁻ cells show impaired mitochondrial respiration, a condition associated with enhanced reductive glutamine metabolism in cancer cells[69], suggesting that aberrant activation of Stat3, or its upstream kinases Jaks, observed in several types of cancers might have an impact on glutamine metabolism under pathological conditions.

Long-term culture of female mouse ESCs in 2iLIF medium has been associated with decreased methylation at imprinted DMRs[15,16,70,71] and reduction in the concentration of MEK inhibitor allowed to maintain robust methylation at DMRs over extensive culture[72]. Similarly, we showed that, in the absence of Stat3, ESCs maintain high methylation levels at DMRs after more than 20 passages in 2iLIF, suggesting that tuning LIF–Stat3 activation might be important for the generation and long-term expansion of pluripotent cells with intact imprinting information.

Naive ESCs are characterized by bivalent metabolism, a hypomethylated genome and high expression of specific transcription factors and epigenetic modifiers. Following differentiation, OXPHOS is decreased, genome methylation is increased, naive specific genes are downregulated and early markers of differentiation are upregulated, including *Otx2* and *Dnmt3a/b*. We propose a model (Extended Data Fig. 9) whereby all these molecular processes are elegantly under the control of a single molecule, Stat3. Whereas nuclear Stat3 directly induces and maintains the expression of naive pluripotency factors, mitochondrial Stat3 promotes OXPHOS and αKG production, genome hypomethylation and inhibition of early differentiation markers. This model explains previous observations, such as the inability of MitoS3 to maintain long-term self-renewal on its own[19], and will be useful to test the role of LIF–Stat3 in pluripotent cells of other species and during induction of pluripotency.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41588-020-00770-2.

References

1. Ito, S. et al. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* **466**, 1129–1133 (2010).
2. Ito, S. et al. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* **333**, 1300–1303 (2011).
3. Messerschmidt, D. M., Knowles, B. B. & Solter, D. DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos. *Genes Dev.* **28**, 812–828 (2014).
4. Monk, M., Boubelik, M. & Lehnert, S. Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development. *Development* **99**, 371–382 (1987).
5. Smith, Z. D. et al. A unique regulatory phase of DNA methylation in the early mammalian embryo. *Nature* **484**, 339–344 (2012).
6. Ishida, M. & Moore, G. E. The role of imprinted genes in humans. *Mol. Aspects Med.* **34**, 826–840 (2013).
7. Boroviak, T. et al. Lineage-specific profiling delineates the emergence and progression of naive pluripotency in mammalian embryogenesis. *Dev. Cell* **35**, 366–382 (2015).
8. Do, D. V. et al. A genetic and developmental pathway from STAT3 to the OCT4–NANOG circuit is essential for maintenance of ICM lineages in vivo. *Genes Dev.* **27**, 1378–1390 (2013).
9. Martello, G., Bertone, P. & Smith, A. Identification of the missing pluripotency mediator downstream of leukaemia inhibitory factor. *EMBO J.* **32**, 2561–2574 (2013).
10. Mohammed, H. et al. Single-cell landscape of transcriptional heterogeneity and cell fate decisions during mouse early gastrulation. *Cell Rep.* **20**, 1215–1228 (2017).
11. Ye, S., Li, P., Tong, C. & Ying, Q. L. Embryonic stem cell self-renewal pathways converge on the transcription factor Tfcp2l1. *EMBO J.* **32**, 2548–2560 (2013).
12. Boroviak, T. & Nichols, J. Primate embryogenesis predicts the hallmarks of human naive pluripotency. *Development* **144**, 175–186 (2017).
13. Ying, Q.-L. et al. The ground state of embryonic stem cell self-renewal. *Nature* **453**, 519–523 (2008).
14. Ficz, G. et al. FGF signaling inhibition in ESCs drives rapid genome-wide demethylation to the epigenetic ground state of pluripotency. *Cell Stem Cell* **13**, 351–359 (2013).
15. Habibi, E. et al. Whole-genome bisulfite sequencing of two distinct interconvertible DNA methylomes of mouse embryonic stem cells. *Cell Stem Cell* **13**, 360–369 (2013).
16. Hackett, J. A. et al. Synergistic mechanisms of DNA demethylation during transition to ground-state pluripotency. *Stem Cell Rep.* **1**, 518–531 (2013).
17. Leitch, H. G. et al. Naive pluripotency is associated with global DNA hypomethylation. *Nat. Struct. Mol. Biol.* **20**, 311–316 (2013).
18. Smith, A. G. et al. Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. *Nature* **336**, 688–690 (1988).
19. Carbognin, E., Betto, R. M., Soriano, M. E., Smith, A. G. & Martello, G. Stat3 promotes mitochondrial transcription and oxidative respiration during maintenance and induction of naive pluripotency. *EMBO J.* **35**, 618–634 (2016).
20. Gough, D. J. et al. Mitochondrial STAT3 supports Ras-dependent oncogenic transformation. *Science* **324**, 1713–1716 (2009).
21. Wegrzyn, J. et al. Function of mitochondrial Stat3 in cellular respiration. *Science* **323**, 793–797 (2009).
22. Lu, C. & Thompson, C. B. Metabolic regulation of epigenetics. *Cell Metab.* **16**, 9–17 (2012).
23. Dawlaty, M. M. et al. Loss of Tet enzymes compromises proper differentiation of embryonic stem cells. *Dev. Cell* **29**, 102–111 (2014).
24. von Meyenn, F. et al. Impairment of DNA methylation maintenance is the main cause of global demethylation in naive embryonic stem cells. *Mol. Cell* **62**, 848–861 (2016).
25. Martello, G. et al. Esrrb is a pivotal target of the Gsk3/Tcf3 axis regulating embryonic stem cell self-renewal. *Cell Stem Cell* **11**, 491–504 (2012).
26. Yamane, M., Ohtsuka, S., Matsuura, K., Nakamura, A. & Niwa, H. Overlapping functions of Krüppel-like factor family members: targeting multiple transcription factors to maintain the naive pluripotency of mouse embryonic stem cells. *Development* **145**, dev162404 (2018).
27. Elhamamsy, A. R. Role of DNA methylation in imprinting disorders: an updated review. *J. Assist. Reprod. Genet.* **34**, 549–562 (2017).
28. Ferguson-Smith, A. C. Genomic imprinting: the emergence of an epigenetic paradigm. *Nat. Rev. Genet.* **12**, 565–575 (2011).
29. Hackett, J. A., Kobayashi, T., Dietmann, S. & Surani, M. A. Activation of lineage regulators and transposable elements across a pluripotent spectrum. *Stem Cell Rep.* **8**, 1645–1658 (2017).
30. Sánchez-Castillo, M. et al. CODEX: a next-generation sequencing experiment database for the haematopoietic and embryonic stem cell communities. *Nucleic Acids Res.* **43**, D1117–D1123 (2015).
31. Matsuda, T. et al. STAT3 activation is sufficient to maintain an undifferentiated state of mouse embryonic stem cells. *EMBO J.* **18**, 4261–4269 (1999).
32. Tang, Y. et al. Jak/Stat3 signaling promotes somatic cell reprogramming by epigenetic regulation. *Stem Cells* **30**, 2645–2656 (2012).
33. Peron, M. et al. Mitochondrial STAT3 regulates proliferation of tissue stem cells. Preprint at *bioRxiv* https://doi.org/10.1101/2020.07.17.208264 (2020).
34. Carey, B. W., Finley, L. W. S., Cross, J. R., Allis, C. D. & Thompson, C. B. Intracellular α-ketoglutarate maintains the pluripotency of embryonic stem cells. *Nature* **518**, 413–416 (2015).
35. Hou, P. et al. Intermediary metabolite precursor dimethyl-2-ketoglutarate stabilizes hypoxia-inducible factor-1α by inhibiting prolyl-4-hydroxylase PHD2. *PLoS ONE* **9**, e113865 (2014).
36. Xu, Y. et al. Genome-wide regulation of 5hmC, 5mC, and gene expression by Tet1 hydroxylase in mouse embryonic stem cells. *Mol. Cell* **42**, 451–464 (2011).
37. Chen, T., Ueda, Y., Dodge, J. E., Wang, Z. & Li, E. Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. *Mol. Cell. Biol.* **23**, 5594–5605 (2003).
38. Tischler, J. et al. Metabolic regulation of pluripotency and germ cell fate through α-ketoglutarate. *EMBO J.* **38**, e99518 (2018).
39. Laukka, T. et al. Fumarate and succinate regulate expression of hypoxia-inducible genes via TET enzymes. *J. Biol. Chem.* **291**, 4256–4265 (2016).
40. Xiao, M. et al. Inhibition of α-KG-dependent histone and DNA demethylases by fumarate and succinate that are accumulated in mutations of FH and SDH tumor suppressors. *Genes Dev.* **26**, 1326–1338 (2012).
41. Teslaa, T. & Teitell, M. A. Pluripotent stem cell energy metabolism: an update. *EMBO J.* **34**, 138–153 (2015).
42. Nishiyama, A. et al. Systematic repression of transcription factors reveals limited patterns of gene expression changes in ES cells. *Sci. Rep.* **3**, 5–10 (2013).
43. Correa-Cerro, L. S. et al. Generation of mouse ES cell lines engineered for the forced induction of transcription factors. *Sci. Rep.* **1**, 167 (2011).
44. Nishiyama, A. et al. Uncovering early response of gene regulatory networks in ESCs by systematic induction of transcription factors. *Cell Stem Cell* **5**, 420–433 (2009).
45. Buecker, C. et al. Reorganization of enhancer patterns in transition from naive to primed pluripotency. *Cell Stem Cell* **14**, 838–853 (2014).
46. Yang, S.-H. et al. Otx2 and Oct4 drive early enhancer activation during embryonic stem cell transition from naive pluripotency. *Cell Rep.* **7**, 1968–1981 (2014).
47. Pawlak, M. & Jaenisch, R. De novo DNA methylation by Dnmt3a and Dnmt3b is dispensable for nuclear reprogramming of somatic cells to a pluripotent state. *Genes Dev.* **25**, 1035–1040 (2011).
48. Grabole, N. et al. Prdm14 promotes germline fate and naive pluripotency by repressing FGF signalling and DNA methylation. *EMBO Rep.* **14**, 629–637 (2013).
49. Yamaji, M. et al. PRDM14 ensures naive pluripotency through dual regulation of signaling and epigenetic pathways in mouse embryonic stem cells. *Cell Stem Cell* **12**, 368–382 (2013).
50. Dan, J. et al. Roles for Tbx3 in regulation of two-cell state and telomere elongation in mouse ES cells. *Sci. Rep.* **3**, 3492 (2013).
51. Palamarchuk, A. et al. Tcl1 protein functions as an inhibitor of de novo DNA methylation in B-cell chronic lymphocytic leukemia (CLL). *Proc. Natl Acad. Sci. USA* **109**, 2555–2560 (2012).
52. Acampora, D., Giovannantonio, L. G. D. & Simeone, A. Otx2 is an intrinsic determinant of the embryonic stem cell state and is required for transition to a stable epiblast stem cell condition. *Development* **140**, 43–55 (2013).
53. Kalkan, T. et al. Tracking the embryonic stem cell transition from ground state pluripotency. *Development* **144**, 1221–1234 (2017).
54. Koh, K. P. et al. Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. *Cell Stem Cell* **8**, 200–213 (2011).
55. Meissner, A. et al. Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* **454**, 766–770 (2008).
56. Tahiliani, M. et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324**, 930–935 (2009).

57. Tsumura, A. et al. Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b. *Genes Cells* **11**, 805–814 (2006).

58. Zhou, W. et al. HIF1α induced switch from bivalent to exclusively glycolytic metabolism during ESC-to-EpiSC/hESC transition. *EMBO J.* **31**, 2103–2116 (2012).

59. Bourillot, P. Y. et al. Novel STAT3 target genes exert distinct roles in the inhibition of mesoderm and endoderm differentiation in cooperation with Nanog. *Stem Cells* **27**, 1760–1771 (2009).

60. Niwa, H., Ogawa, K., Shimosato, D. & Adachi, K. A parallel circuit of LIF signalling pathways maintains pluripotency of mouse ES cells. *Nature* **460**, 118–122 (2009).

61. Boroviak, T., Loos, R., Bertone, P., Smith, A. & Nichols, J. The ability of inner-cell-mass cells to self-renew as embryonic stem cells is acquired following epiblast specification. *Nat. Cell Biol.* **16**, 516–528 (2014).

62. McLaughlin, K. et al. DNA methylation directs polycomb-dependent 3D genome re-organization in naive pluripotency. *Cell Rep.* **29**, 1974–1985 (2019).

63. Takeda, K. et al. Targeted disruption of the mouse Stat3 gene leads to early embryonic lethality. *Proc. Natl Acad. Sci. USA* **94**, 3801–3804 (1997).

64. Leeb, M., Dietmann, S., Paramor, M., Niwa, H. & Smith, A. Genetic exploration of the exit from self-renewal using haploid embryonic stem cells. *Cell Stem Cell* **14**, 385–393 (2014).

65. Avalle, L. et al. STAT3 localizes to the ER, acting as a gatekeeper for ER-mitochondrion Ca²⁺ fluxes and apoptotic responses. *Cell Death Differ.* **26**, 932–942 (2019).

66. Wang, L. et al. JAK/STAT3 regulated global gene expression dynamics during late-stage reprogramming process. *BMC Genomics* **19**, 183 (2018).

67. Hwang, I.-Y. et al. Psat1-dependent fluctuations in α-ketoglutarate affect the timing of ESC differentiation. *Cell Metab.* **24**, 494–501 (2016).

68. Zhang, J. et al. LIN28 regulates stem cell metabolism and conversion to primed pluripotency. *Cell Stem Cell* **19**, 66–80 (2016).

69. Mullen, A. R. et al. Reductive carboxylation supports growth in tumour cells with defective mitochondria. *Nature* **481**, 385–388 (2012).

70. Choi, J. et al. Prolonged Mek1/2 suppression impairs the developmental potential of embryonic stem cells. *Nature* **548**, 219–223 (2017).

71. Yagi, M. et al. Derivation of ground-state female ES cells maintaining gamete-derived DNA methylation. *Nature* **548**, 224–227 (2017).

72. Gretarsson, K. J. & Hackett, J. A. *Dppa2* and *Dppa4* counteract de novo methylation to establish a permissive epigenome for development. *Nat. Struct. Mol. Biol.* **27**, 706–716 (2020).

73. Chen, T., Ueda, Y., Xie, S. & Li, E. A novel Dnmt3a isoform produced from an alternative promoter localizes to euchromatin and its expression correlates with active de novo methylation. *J. Biol. Chem.* **277**, 38746–38754 (2002).

Methods

Cell lines and culture routine. Wild-type or Stat3-KO (described previously[13,19,63] and provided by A. Smith's laboratory) mouse ESC lines were routinely cultured without feeders on gelatin-coated plates (0.2% gelatin; Sigma-Aldrich, cat. no. G1890). The media were changed every 2 d and the cells were passaged when approaching confluency (every 2–3 d); to passage, the cells were replated at the required density following dissociation with either accutase (Thermo Fisher, cat. no. A11105-01) or trypsin (Thermo Fisher, cat. no. 15090-046). All cells were maintained at 37 °C in humidified incubators with 5% CO_2.

Media and supplements. Cells were cultured under two different culture conditions, prepared as follows.

2i/2iLIF: 10% Serum-free KnockOut serum replacement (Life Technologies, cat. no. 10828-028)-based medium in GMEM medium (Sigma-Aldrich, cat. no. G5154) supplemented with 1% fetal bovine serum (Sigma-Aldrich, cat. no. F7524), 100 mM 2-mercaptoethanol (Sigma-Aldrich, cat. no. M7522), 1× MEM nonessential amino acids (Invitrogen, cat. no. 1140-036), 2 mM L-glutamine, 1 mM sodium pyruvate (both from Invitrogen) as well as the small-molecule inhibitors PD0325901 (1 μM) and CHIR99021 (3 mM) from Axon (cat. nos 1386 and 1408) and LIF (100 units ml⁻¹; produced in-house).

Serum LIF: GMEM medium (Sigma-Aldrich, cat. no. G5154) supplemented with 10% fetal bovine serum (Sigma-Aldrich, cat. no. F7524), 100 mM 2-mercaptoethanol (Sigma-Aldrich, cat. no. M7522), 1× MEM nonessential amino acids (Invitrogen, cat. no. 1140-036), 2 mM L-glutamine, 1 mM sodium pyruvate (both from Invitrogen) and 100 units ml⁻¹ LIF.

LIF induction. For the LIF induction experiments, ESCs cultured in 2i without LIF for at least four passages were plated in 2i. Twenty-four hours after plating, LIF was added for the indicated time periods (24 or 48 h) and the cells were then fixed for immunofluorescence (described previously[19]).

Treatments. For inhibition of the respiratory chain, cells were treated acutely with 100 nM of the complex I inhibitor rotenone (Sigma-Aldrich, cat. no. R8875) and 200 nM of the complex III inhibitor antimycin (Sigma-Aldrich, cat. no. A8674), as described previously[19].

For studies on glutaminolysis, cells were cultured in KnockOut serum replacement-based medium prepared as described earlier but without the addition of glutamine. Exogenous DM-αKG used for treatments was added to the culture medium at the indicated concentration and absorbed by cells as it is membrane-permeable (Sigma-Aldrich, cat. no. 349631-5G).

Clonal assay. The ability of single ESCs to form pluripotent colonies was assessed through a clonal assay. The cells were harvested by trypsinization and plated at clonal density; to do this, they were counted and diluted to obtain a final number of 600 cells per well. The cells were cultured for 4–5 d before they were fixed and stained for alkaline phosphatase.

Differentiation assay. Cells were cultured with or without 2iLIF for 24, 48 and 72 h. After 72 h, the cells were detached and replated at clonal density in 2iLIF. Surplus cells were conserved for gene expression analysis. Finally, the cells were stained for alkaline phosphatase after 5 d to evaluate the number of pluripotent cells. For AP staining, the cells were fixed with a citrate-acetone-formaldehyde solution and stained using an alkaline phosphatase kit (Sigma-Aldrich, cat. no. 86R-1KT). Plates were scanned using a Nikon scanner and scored manually.

Immunofluorescence and staining. For 5mC staining, cells were fixed in 4% formaldehyde (Sigma-Aldrich, cat. no. F8775) for 10 min, washed in PBS and then treated for 15 min with NH_4Cl. Next, the cells were permeabilized through 0.5% PBST treatment (PBS and 0.5% Triton X-100 (Sigma-Aldrich, cat. no. 93443)) for 1 h and 2 N HCl was added for 45 min to denature the DNA. The cells were blocked for 1 h in 5% horse serum (Thermo Fisher, cat. no. 16060122) with 0.3% PBST and then incubated overnight at 4 °C with anti-5mC primary antibody (Eurogentec, cat. no. BI-MECY-0500; Supplementary Table 5) diluted in 2% horse serum with 0.3% PBST. After washing with 0.1% PBST, the cells were incubated with secondary antibody (Alexa Fluor 488, 1:500, donkey anti-mouse IgG; Thermo Fisher, A-21202) for 45 min at room temperature. The nuclei were stained with the mounting medium Fluoroshield containing DAPI (Sigma-Aldrich, cat. no. F6057).

For EdU staining, cells were exposed to an EdU (nucleoside analogue of thymidine incorporated into DNA during active DNA synthesis) pulse of 4 h before fixation in 4% formaldehyde for 10 min; the samples were then processed according to the manufacturer's instructions (Life Technologies).

For Atad3 and Stat3 colocalization staining, cells were fixed for 10 min in 4% formaldehyde, washed in PBS, and blocked and permeabilized for 1 h in 5% horse serum with 0.3% PBST. The cells were incubated overnight at 4 °C with primary antibodies (Supplementary Table 5). After washing with 0.1% PBST, the cells were incubated with secondary antibodies (Alexa Fluor 568, 1:500, donkey anti-mouse IgG; Thermo Fisher, A-11036 and Alexa Fluor 488; 1:500, goat anti-rabbit IgG; Thermo Fisher, A-21206) for 30 min at room temperature. The nuclei were stained with mounting medium Fluoroshield containing DAPI.

Images were acquired using a Leica SP5 or a Zeiss LSM 700 confocal microscope equipped with a CCD camera. The fluorescence intensity was quantified using the freely available software Fiji (http://fiji.sc/Fiji).

RNA-seq. Total RNA was isolated using a Total RNA purification kit (Norgen Biotek, cat. no. 37500) and sequenced using an Illumina HiSeq4000, in 150-bp, paired-end format.

The reads were aligned to the mouse transcriptome (*Mus musculus* transcriptome generated by rsem-prepare-reference with ENSEMBL93 GTF) and mouse genome (GRCm38.p6) using HISAT2 (v. 2.1.0).

The gene expression levels were quantified using RSEM (v. 1.3.1) using transcriptome alignments. Genome alignment was used to create bigWig files using deeptools (v. 3.2.1).

Genes were sorted based on the average expression calculated in a total of 18 samples and a final expression matrix was generated excluding genes that had an average expression lower than 22.88 raw counts; after applying this filter, we obtained the expression values of 12,510 genes.

All RNA-seq analyses were carried out in R environment (v. 3.5.3) with Bioconductor (v. 3.7).

We computed differential expression analysis using the DESeq2 R package (v. 1.24.0)[74]; DESeq2 performs the estimation of size factors, the estimation of dispersion for each gene and fits a generalized linear model. Transcripts with absolute value of $\log_2[FC] > 1$ and a q-value (P_{adj}) < 0.01 (Benjamini–Hochberg adjustment) were considered significant and defined as differentially expressed for the comparison in the analysis.

Heatmaps were made using transcripts-per-million values with the pheatmap function from the pheatmap R package (v.1.0.12; distance = 'correlation', scale = 'row') on differentially expressed genes or selected markers.

Volcano plots were computed with $\log_2[FC]$ and $-\log_{10}[q\text{-}value]$ from DESeq2 differential analysis output using the ggscatter function from the ggpubr R package (v. 0.2).

See also Supplementary Table 1.

Metabolite analysis by mass spectrometry. Cells were cultured in six-well plates, harvested in ice-cold PBS and centrifuged at 500g for 3 min at 4 °C. The pellets were then resuspended in 250 μl methanol/acetonitrile 1:1 containing [U-¹³C₆] glucose and [U-¹³C₅]glutamine, each at 1 ng μl⁻¹ (internal standards; Sigma-Aldrich, cat. no. 389374) and centrifuged at 20,000g for 5 min at 4 °C. The supernatants were then passed through a regenerated cellulose filter, dried and resuspended in 100 μl methanol for subsequent analysis.

Metabolomic data analyses were performed on an API-4000 triple quadrupole mass spectrometer (Sciex) coupled with an HPLC system (Agilent), a CTC PAL HTS autosampler (PAL System) and on an API3500 instrument (Sciex). All of the methods were arranged by setting multiple ion monitoring with pure commercial standards to confirm the identity of all metabolites.

Quantification of different metabolites was performed using a LC–MS/MS method using a C18 column (Biocrates) for amino acids and SAMe, and a cyano-phase LUNA column (50 mm × 4.6 mm, 5 μm; Phenomenex). Methanolic samples were analyzed by a 10-min run in positive- (amino acids and SAMe) and 5-min run in negative-ion (all other metabolites) mode with specific multiple reaction monitoring transitions. Amino-acid quantification was performed through the previous derivatization. Briefly, 50 μl of 5% phenyl isothiocyanate) in 31.5% ethanol and 31.5% pyridine in water were added to 10 μl of each sample. The mixtures were then incubated with phenyl isothiocyanate solution for 20 min at room temperature, dried under N_2 flow and suspended in 100 μl of 5 mM ammonium acetate in methanol/H_2O 1:1. The mobile phases for positive-ion-mode analysis (amino acids and SAMe) were: phase A, 0.2% formic acid in water and phase B, 0.2% formic acid in acetonitrile. The mobile phase for negative-ion-mode analysis (all other metabolites) was: phase A, water and phase B, 2 mM ammonium acetate in methanol. The gradient was 90% phase B for all of the analyses, with a flow rate of 500 μl min⁻¹. MultiQuantTM software (version 3.0.2) was used for the data analysis and peak review of chromatograms. Quantitative evaluation of all metabolites was performed based on calibration curves with pure standards; the data were then normalized to the total protein content.

Metabolic flux analysis. For metabolic-tracing analyses, cells were exposed for 24 h to 1 mM [U-¹³C₆]glucose (Sigma-Aldrich, 389374), 2 mM [U-¹³C₅]glutamine (Sigma-Aldrich, cat. no. 605166) or 100 μM [U-¹³C₁₆]palmitate (Sigma-Aldrich, cat. no. 605573). The cells were harvested in ice-cold PBS and centrifuged at 500g for 3 min at 4 °C. The pellets were then resuspended in 250 μl methanol/acetonitrile 1:1 and centrifuged at 20,000g for 5 min at 4 °C. The supernatants were then passed through a regenerated cellulose filter, dried under N_2 flow and resuspended in 100 μl methanol for subsequent analysis. Metabolomic data analyses were performed on an API-4000 triple quadrupole mass spectrometer (Sciex) coupled with a HPLC system (Agilent) and CTC PAL HTS autosampler (PAL System). The identity of all metabolites was confirmed using pure standards. Quantification of different metabolites was performed using a LC–MS/MS method using a cyano-phase LUNA column (50 mm × 4.6 mm, 5 μm; Phenomenex) Methanolic samples were analyzed by a 5-min run in negative-ion (metabolites) mode. The

mobile phases for the negative-ion-mode analysis were: phase A, 2 mM ammonium acetate in methanol and phase B, water. The gradient was 90% phase A throughout the analysis, with a flow rate of 500 μl min⁻¹. The MultiQuantTM software (version 3.0.2) was used for data analysis and peak review of chromatograms. The samples were analyzed after 8 h of ^{13}C-labeling to ensure that isotopic equilibrium was reached, as previously shown in ESCs cultured in 2iLIF by Carey and colleagues[34]. All detected ^{13}C-labeled metabolites were corrected for natural isotope abundances.

RRBS. The RRBS was performed as previously described[75]. Briefly, 500 ng DNA was digested at 37 °C with 200 U MspI restriction endonuclease (NEB). The digested DNA was then end repaired, dA-tailed and ligated to methylated adapters, using an Illumina TruSeq DNA sample prep kit following the manufacturer's instructions. The adapter-ligated DNA was loaded on a 2% agarose gel and a fraction from 200 to 400 bp was recovered. The purified DNA was then subjected to bisulfite conversion using an EpiTect bisulfite kit (Qiagen). The bisulfite-converted DNA was finally enriched by 15 cycles of PCR using Kapa HiFi HotStart uracil (Roche).

RRBS data processing and analysis. After quality controls, the sequencing reads were mapped to the mouse genome reference (mm10/GRC.m38.p6) with BSMAP (v2.89)[76] using the RRBS mode (parameters: -s 12 -D C-CGG -v 0.01 --n 1). The levels of CpG methylation were extracted from the aligned reads as the ratio of the number of Cs over the total number of Cs and Ts using the methratio.py script. The CpG methylation ratios from both strands were combined (parameters: --g). For downstream analysis, the CpG sites that were commonly covered in at least one technical replicate of each sample with a minimum sequencing depth of 10× were retained. All samples were processed identically.

Statistical analyses were conducted in the R software environment. Differential methylation analysis at single-nucleotide resolution was performed for each comparison (that is, S3$^{+/+}$ 2iLIF versus S3$^{−/−}$ 2iLIF, S3$^{+/+}$ 2i versus S3$^{+/+}$ 2iLIF, Dnmt3a/b dKO.1 and dKO.2 2i versus E14 2i, and 2iLIF versus 2i E14) using the methylKit R/Bioconductor package[77], exploiting the logistic-regression approach for testing replicates (calculateDiffMeth function with default parameters). CpG sites with an absolute methylation difference ≥0% and q-value ≤0.01 were considered to be differentially methylated. Correlation analysis between the effect of LIF and Dnmt3a/b on CpG methylation was performed on the methylation difference of each condition with respect to wild-type cells cultured in 2i medium using the cor.test R function.

For the study of DNA methylation levels on regulatory elements, ChIP–seq data of histone marks (H3K27ac and H3K4me3) generated in E14 ESCs were retrieved from ENCODE (https://www.encodeproject.org/). Active promoters and enhancers were defined from processed peaks data as follows: promoters, H3K4me3 peaks in a 2-kb window centered in the transcription start site of annotated genes (GENCODE release M20); and enhancers, distal H3K27ac peaks (more than 1-kb up/downstream of the nearest transcription start site).

Differential methylation analysis (calculateDiffMeth function, q-value ≤0.05 and methylation difference ≥ 10%) was performed on these regions (that is, testing all of the covered CpGs overlapping with the ChIP–seq peaks, with 200 bp of flanking region) for the comparisons S3$^{+/+}$ 2iLIF versus S3$^{−/−}$ 2iLIF and S3$^{+/+}$ 2i versus S3$^{+/+}$ 2iLIF. These results were then integrated with RNA-seq data. After performing differential expression analysis, the fold change in gene expression levels was visualized against the average changes in the DNA methylation levels of the promoters/enhancers. Similar analyses were conducted on a manually curated list of imprinted DMRs (see also Supplementary Table 2).

Single-cell RNA-seq analysis of S3$^{−/−}$ and S3$^{+/+}$ embryos. Immunosurgery and single-cell dissociation was performed as described[7]. The method for single-cell RNA-seq and library preparation was previously described[78]. A total of 171 cells from 18 embryos were analyzed.

The experiments were performed in accordance with European Union guidelines for the care and use of laboratory animals and under authority of UK governmental legislation. The use of animals in this project was approved by the ethical review committee of the University of Cambridge, and relevant Home Office licenses are in place.

Mus musculus GRCm38.87 gene annotation and mm10 genome sequence were downloaded from Ensembl (https://www.ensembl.org/index.html). All reads were aligned using Spliced transcripts alignment to a reference (STAR, v. 2.5.2b)[78]. Alignments to gene loci were quantified using htseq-count (v. 0.11.0)[79] based on the annotation from Ensembl 87. Principal-component-analysis outliers were computed and removed. Mouse embryo data for the E4.5, E5.5 and E6.5 stages were compiled from earlier studies[10,80].

The principal component analysis was based on log$_2$[FPKM] values computed using the Bioconductor package DESeq (v. 1.42.0)[81], custom scripts and FactoRmineR package (v. 2.3)[82]. The Monocle package (v. 2.18.0) was used to compute t-SNE clustering. Differential expression analysis was performed using the DESeq2 package (v. 1.24.0) following pre-processing of raw counts with the zinbwave package (v. 1.12.0), which allows extraction of low-dimensional signal from zero-inflated data—as described[83]—that fits individual error models for the assessment of differential expression between groups of cells. Fractional identity

between E3.5/E3.75 S3$^{+/+}$ and S3$^{−/−}$ cells, and embryo stages (E4.5 Epi, E5.5 Epi and E6.5 Epi) was computed using the R package DeconRNASeq (v. 1.32.0)[84], which makes use of quadratic programming to estimate the proportion of distinctive types of tissue. The average expression of the embryo stages was used as a 'signature' dataset (see also Supplementary Table 4).

Proteomics. All of the experiments were performed in a labeling-free setting. For each sample, 50 mg total cellular protein extract were precipitated overnight at 4 °C in acetone, and then reduced and alkylated in a solution of 6 M guanidine–HCl, 5 mM TCEP and 55 mM chloroacetamide. Peptides were obtained by digesting the proteins with LysC (WAKO) for 3 h at 37 °C and with the endopeptidase sequencing-grade trypsin (Promega) overnight at 37 °C. The collected peptide mixtures were concentrated and desalted using the Stop-and-go extraction (STAGE) technique[85].

Instruments for LC–MS/MS analysis consisted of a NanoLC 1200 coupled via a nano-electrospray ionization source to the quadrupole-based Q Exactive HF benchtop mass spectrometer[86]. Peptide separation was carried out according to their hydrophobicity on a PicoFrit column, 75-mm internal diameter, 8-μm tip, 25-mm bed packed with Reprosil-PUR, C18-AQ, 1.9-mm particle size and 120-Å pore size (New Objective, Inc., cat. no. PF7508-250H363) using a binary buffer system consisting of solutions A (0.1% formic acid) and B (80% acetonitrile and 0.1% formic acid). Runs of 120 min, after loading, were used for the proteome samples, with a constant flow rate of 300 nl min⁻¹. After sample loading, the run was started at 5% buffer B for 5 min, followed by a series of linear gradients from 5 to 30% B in 90 min, then a 10-min step to reach 50% and a 5-min step to reach 95%. This last step was maintained for 10 min.

The Q Exactive HF settings were as follows: mass spectrometry spectra were acquired using 3×10^6 as an AGC target, a maximal injection time of 20 ms and a 120,000 resolution at 200 m/z.

The mass spectrometer was operated in a data-dependent Top20 mode with subsequent acquisition of higher-energy collisional dissociation fragmentation MS/MS spectra of the top-20 most-intense peaks. For the MS/MS spectra, the resolution was set to 15,000 at 200 m/z, the AGC target to 1×10^5, maximum injection time to 20 ms and the isolation window to 1.6 Th. The intensity threshold was set to 2.0×10^4 and dynamic exclusion to 30 s.

All of the acquired RAW files were processed using MaxQuant (1.6.2.10) and the implemented Andromeda search engine. For protein assignment, spectra were correlated with the UniProt mouse database (v. 2019) including a list of common contaminants. Searches were performed with tryptic specifications and default settings for mass tolerances for MS and MS/MS spectra. Carbamidomethyl at cysteine residues was set as a fixed modification, whereas oxidations at methionine and acetylation at the amino terminus were defined as variable modifications. The minimal peptide length was set to seven amino acids and the false-discovery rate for proteins and peptide-spectrum matches to 1%. For LFQ, the minimum ratio count was set to one. The match-between-run feature with a time window of 1 min was used. For further analysis, the Perseus software (1.6.2.3) was used and first filtered for contaminants and reverse entries as well as proteins that were only identified by a modified peptide. The LFQ of analyzed samples were log-transformed, and the five biological replicates of each sample type were grouped. Proteins were filtered for a minimum of four out of five replicates in at least one group. Missing values were replaced by random numbers that were drawn from a normal distribution. A two-sample Student's t-test was performed using false-discovery rate = 0.05. P < 0.05 was considered statistically significant. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD020385 (see also Supplementary Table 3).

Statistics and reproducibility. For each dataset, the sample size n refers to the number independent experiments or biological replicates, as stated in the figure legends. Biological replicates indicate when a cell line was exposed to a given treatment multiple times and the samples were harvested, processed and analyzed all at once, as in the case of RNA-seq or proteomics experiments. Independent experiments indicate when one or more cell lines were treated, harvested and analyzed multiple independent times, as in the case of the immunofluorescence and quantitative PCR assays. All P values were calculated using an unpaired two-tailed Student's t-test and have been indicated as their numerical values in each plot, unless stated otherwise; the P values were not calculated for datasets with n < 3. For both bulk and single-cell RNA-seq, the P values were calculated using the Wald test. We used the χ^2 test for RRBS and Student's t-test for proteomics. Adjusted P values (q-values) were calculated in the case of multiple testing. For the RNA-seq data, we used the Benjamini–Hochberg adjustment; for RRBS, we used the SLIM method and for proteomics, we used false-discovery rate. Either the Excel or R software were used for the statistical analyses. The error bars indicate the s.e.m. or s.d., as stated in the figure legends.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Bulk and single-cell RNA-seq and RRBS data generated during the current study are available at the Gene Expression Omnibus (GEO) repository under the

accession numbers GSE133926 and GSE134450. All RNA-seq and RRBS process data used in Figs. 1d,i, 2a–e,h, 4c, 5d, 6b,d, 7a–e, 8a–g and Extended Data Figs. 1a,b,e, 2b–k, 3b,c, 5c, 7a–d, 8a–f are reported in Supplementary Tables 1, 2 and 4. The RNA-seq data of Rex1-GFPd2 cells can be found at GEO under the accession number GSE111694. Mass spectrometry proteomic data of the following samples: $S3^{+/+}$ cells in 2i; $S3^{+/+}$, $S3^{−/−}$, MitoS3.A and MitoS3.B cells in 2iLIF used in Figs. 1f, 5g, and Extended Data Fig. 1c are reported in Supplementary Table 3 and on ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD020385. The primers and oligonucleotide sequences are in Supplementary Tables 6–8. Additional data that support the findings of this study, such as mass spectrometry measurements, analysis pipelines and reagents, are available from the corresponding authors on reasonable request. Source data are provided with this paper.

Code availability
All software and bioinformatic tools used in the present study are publicly available.

References
74. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
75. Neri, F., Incarnato, D., Krepelova, A., Parlato, C. & Oliviero, S. Methylation-assisted bisulfite sequencing to simultaneously map 5fC and 5caC on a genome-wide scale for DNA demethylation analysis. *Nat. Protoc.* **11**, 1191–1205 (2016).
76. Xi, Y. & Li, W. BSMAP: whole genome bisulfite sequence MAPping program. *BMC Bioinform.* **10**, 232 (2009).
77. Akalin, A. et al. methylKit: A comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biol.* **13**, R87 (2012).
78. Boroviak, T. et al. Single cell transcriptome analysis of human, marmoset and mouse embryos reveals common and divergent features of preimplantation development. *Development* **145**, dev167833 (2018).
79. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
80. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
81. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010).
82. Lê, S., Josse, J. & Husson, F. FactoMineR: an R package for multivariate analysis. *J. Stat. Softw.* **25**, 1–18 (2008).
83. Risso, D., Perraudeau, F., Gribkova, S., Dudoit, S. & Vert, J.-P. A general and flexible method for signal extraction from single-cell RNA-seq data. *Nat. Commun.* **9**, 284 (2018).
84. Gong, T. & Szustakowski, J. D. DeconRNASeq: a statistical framework for deconvolution of heterogeneous tissue samples based on mRNA-Seq data. *Bioinformatics* **29**, 1083–1085 (2013).
85. Rappsilber, J., Ishihama, Y. & Mann, M. Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Anal. Chem.* **75**, 663–670 (2003).
86. Michalski, A. et al. Mass spectrometry-based proteomics using Q Exactive, a high-performance benchtop quadrupole Orbitrap mass spectrometer. *Mol. Cell. Proteomics* **10**, M111.011015 (2011).

Author contributions
G.M., N.M. and S.O. designed the study. R.M.B. and L.D. performed the ESC culture, molecular characterization and functional assays and visualization. V.P. and L.D. performed the MeDIP–qPCR. R.M.B., V.P. and S.R. performed the RRBS. R.M.B. and S.R. performed the western blots. A.L. and D.I. performed the RRBS integrated analysis. M. Arboit and L.D. performed the RNA-seq analyses. R.M.B., S.P. and M. Audano performed the metabolomic analyses. R.M.B., V.P., V.G. and D.T. performed the nucleotide mass spectrometry. P.G. performed the proteomics. M.E.S. and G.R. performed the mitochondrial and nuclear fractionations. L.D. and G.G.S. performed the single-cell RNA-seq analyses. T.L., T.B. and J.N. performed the embryo dissections and single-cell RNA-seq library preparations. G.M. wrote the manuscript with input from all authors. G.M., N.M. and S.O. supervised the study.

Competing interests
The authors declare no competing interests.

Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41588-020-00770-2.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41588-020-00770-2.

Correspondence and requests for materials should be addressed to N.M., S.O. or G.M.

Peer review information *Nature Genetics* thanks Heather Christofk, Christian Frezza, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permissions information is available at www.nature.com/reprints.

a RRBS

b RNAseq

c Proteomics

d

e RRBS

f qPCR

g IF 5mC

h qPCR

i IF 5mC
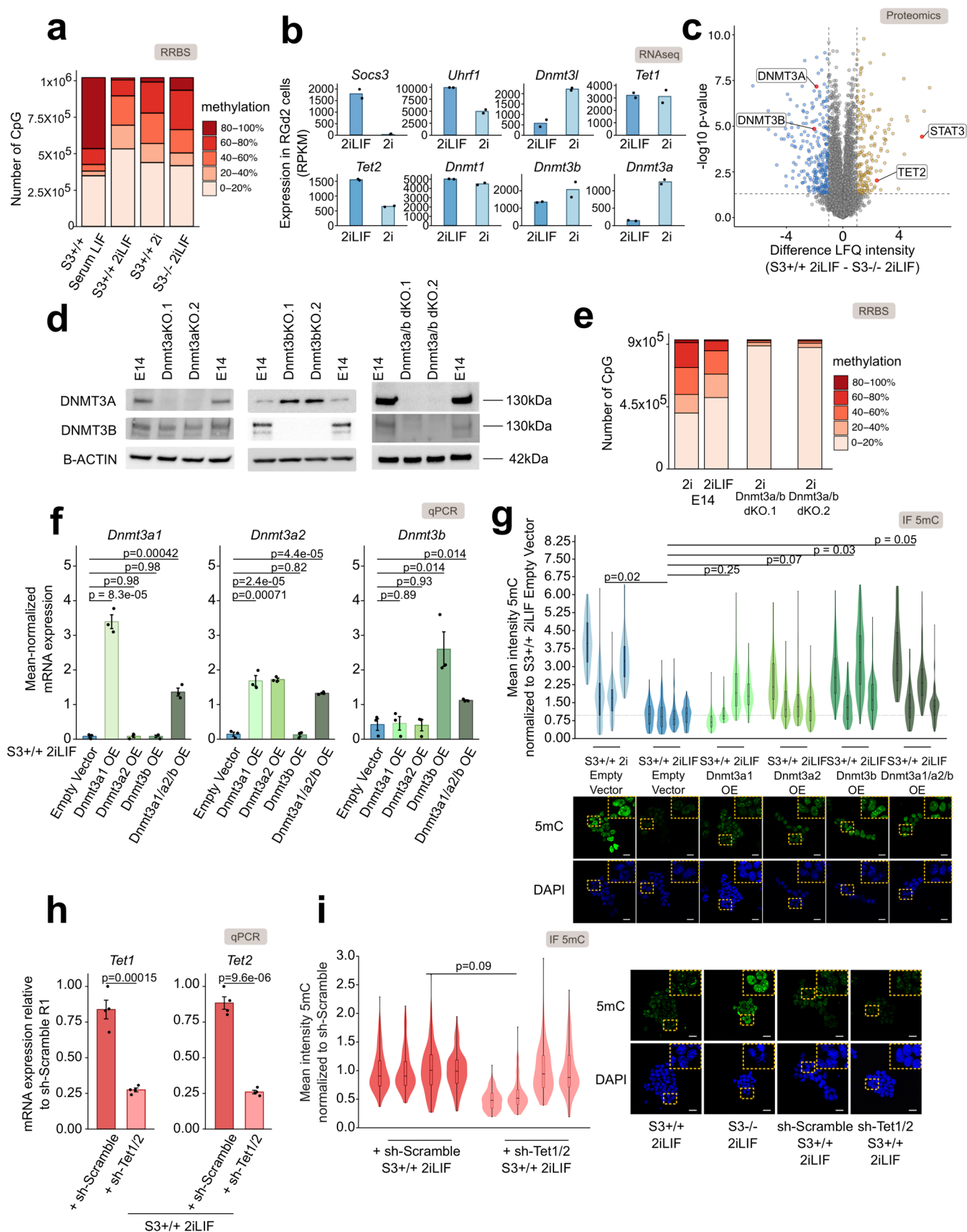
Extended Data Fig. 1 | See next page for caption.

Extended Data Fig. 1 | Dnmt3a/b control 5mC levels downstream of LIF–Stat3. a, Distribution of DNA methylation levels at CpG islands in S3+/+ cells cultured in Serum LIF, 2i or 2iLIF and S3-/- cells in 2iLIF. **b**, Gene expression analysis in RGd2 cells in 2i with or without LIF. Socs3 was used as a control of LIF–Stat3 activation. Bars: mean of n = 2 biological replicates.**c**, Proteomics data from S3+/+ and S3-/- cells in 2iLIF. Yellow and blue dots indicate respectively proteins that are more or less abundant (difference > 1 or < -1, p-value < 0.05) in S3+/+ relative to S3-/-. n = 5 biological replicates. **d**, Western blot of E14 mES cells and of Dnmt3a KO, Dnmt3b KO, and Dnmt3a/b double KO cells (two clones for each mutant genotype) in Serum LIF. Two E14 samples are loaded on the right and left for each KO cell line. B-ACTIN used as a loading control. Representative images of n = 2 independent experiments. **e**, Distribution of DNA methylation levels at CpG islands in E14 mES cells in 2i or 2iLIF and two clones of Dnmt3a/b double KO mES cells in 2i.**f**, Gene expression analysis of S3+/+ cells cultured in 2iLIF transiently expressing an Empty Vector, Dnmt3a (two isoforms, Dnmt3a1 and Dnmt3a2 – as previously identified in[73]), Dnmt3b, or the three genes simultaneously (Dnmt3a1/a2/b OE). Bars: mean ± s.e.m. of n = 3 independent experiments, shown as dots. **g**, Anti-5mC immunofluorescence on S3+/+ in 2i and 2iLIF, and Dnmt3a1 OE, Dnmt3a2 OE, Dnmt3b OE, and Dnmt3a1/a2/b triple OE cells in 2iLIF. Violin plots of an average of 89 nuclei per sample. n = 4 experiments. **h**, Gene expression analysis of S3+/+ cells in 2iLIF stably expressing shRNA to knock-down Tet1 and Tet2 simultaneously or a scrambled control shRNA. Bars: mean ± s.e.m. of n = 4 experiments. **i**, Anti-5mC immunofluorescence on S3+/+ in 2iLIF transiently expressing a scrambled control shRNA and shRNA against Tet1/Tet2. Violin plots of an average of 78 nuclei per sample. n = 4 experiments.All violin and boxplots indicate the 1st, 2nd and 3rd quartiles, with whiskers indicating minimum and maximum value. All p-values calculated by two-tailed unpaired T-test. Scale bars: 20μm.

a

MeDIP q-PCR

Legend: S3+/+ IP · S3+/+ mock · S3-/- IP · S3-/- mock

IAP | *IAPEY3* | *LINE L1* | *Maj Sat* | *Min Sat* | *SINE B1* | *SINE B2*

y-axis: mean normalized % of input

b

RRBS

S3+/+ 2i vs S3+/+ 2iLIF

1506 | 512877

y-axis: -log10(q-value)
x-axis: Methylation difference (%)
Loss of methylation | Gain of methylation

c

RNAseq+RRBS

Promoters

S3+/+ 2i vs S3+/+ 2iLIF

Labels: Stag3, Mov10l1, Fkbp6, Sept1, Lrrc2, Ly75, Socs3, Stat3

y-axis: average methylation difference (%)
x-axis: log2 Fold Change Expression

d

RNAseq+RRBS

Enhancers

S3+/+ 2i vs S3+/+ 2iLIF

Labels: Mras, Klf5, Osbpl10, Esrrb, Socs3, Bmp4, Snai1, Serpine1

y-axis: average methylation difference (%)
x-axis: log2 Fold Change Expression

e

RNAseq

R = 0.89 , p < 2.2e-16

Labels: Igf2, H19, Peg10, Dnmt3b, Nnat, Fgf5, Pou3f1, Dnmt3a, Cdkn1c, Prdm14, Nanog, Otx2, Lin28b, Tfcp2l1, Tbx3, Lin28a, Stat3, Phlda2, Tet2, Socs3, Tet1, Esrrb

y-axis: log2 FC(S3+/+ 2i vs S3+/+ 2iLIF)
x-axis: log2 FC(S3-/- vs S3+/+ 2iLIF)

f

RRBS

2i vs 2iLIF - E14

355 | 481279

y-axis: -log10(q-value)
x-axis: methylation difference (%)

g

RRBS

Pearson's r = 0.802
N_TOT = 1000704
N = 447539

y-axis: difference [2iLIF - 2i] E14 (% mCG)
x-axis: difference [Dnmt3a/b dKO.1 - E14] 2i (% mCG)

h

RRBS

2i - Dnmt3a/b dKO.1 vs E14

643479 | 275

y-axis: -log10(q-value)
x-axis: methylation difference (%)

i

RRBS

Pearson's r = 0.803
N_TOT = 1012664
N = 453771

y-axis: difference [2iLIF - 2i] E14 (% mCG)
x-axis: difference [Dnmt3a/b dKO.2 - E14] 2i (% mCG)

j

RRBS

2i - Dnmt3a/b dKO.2 vs E14

646241 | 344

y-axis: -log10(q-value)
x-axis: methylation difference (%)

k

RRBS

Hypomethylated CpGs

7168

175736 | 453771

Dnmt3a/b dKO.2 vs wild-type | 2iLIF vs 2i
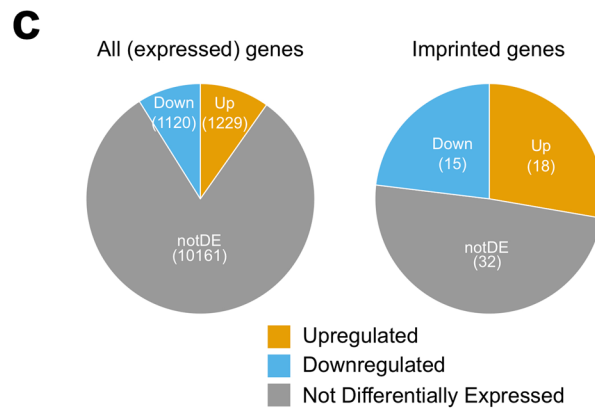
Extended Data Fig. 2 | See next page for caption.

Extended Data Fig. 2 | Impact of LIF–Stat3 and Dnmt3a/b on genome methylation. a, MeDIP-qPCR of repetitive elements in S3+/+ and S3-/- cells. Mock immunoprecipitations with a non-specific IgG antibody served as negative controls. Mean of 2 experiments, shown as dots. **b**, Volcano plot showing the significant differentially methylated CpG sites (q-value < 0.01, difference > 10% or < 10%) between S3+/+ 2i and S3+/+ 2iLIF cells, out of 1,327,475 detected sites. **c, d** Scatter plot showing the mutual changes in expression and DNA methylation at active promoters (**c**) and enhancers (**d**) between S3+/+ 2i and S3+/+ 2iLIF cells. Red dots: genes for which both changes were statistically significant (q-value < 0.01). **e**, Scatter plot comparing effects on transcription of the absence of LIF (S3+/+ 2i) and the absence of Stat3 (S3-/- 2iLIF). Pearson's correlation coefficient (R) and corresponding p-value are indicated in the panel. **f**, Volcano plot showing the significant differentially methylated CpG sites (q-value < 0.01, difference > 10% or < -10%) between E14 2i and E14 2iLIF cells, out of 1,084,350 detected sites. **g**, CpG methylation changes caused by LIF addition (y axis) or by Dnmt3a/b deletion (x axis, Dnmt3a/b dKO.1). Dots indicate all CpG sites covered (sequencing depth >10x) in at least one technical replicate of each sample; blue dots: hypomethylated sites (q-value < 0.01, difference < -10 %). **h**, Volcano plot showing the significant differentially methylated CpG sites between Dnmt3a/b dKO.1 and E14 cells cultivated in 2i.**i**, CpG methylation changes caused by LIF (y axis) or by Dnmt3a/b deletion (x axis, Dnmt3a/b dKO.2), as described in panel g. **j**, Volcano plot showing the significant differentially methylated CpG sites between Dnmt3a/b dKO.2 and E14 cells cultivated in 2i. **k**, Venn diagram of CpG sites whose methylation status is dependent on either LIF (light blue) or Dnmt3a/b (red) or on both (grey intersection), for an independent mutant Dnmt3a/b dKO clone (Dnmt3a/b dKO.2).

a

MeDIP q-PCR
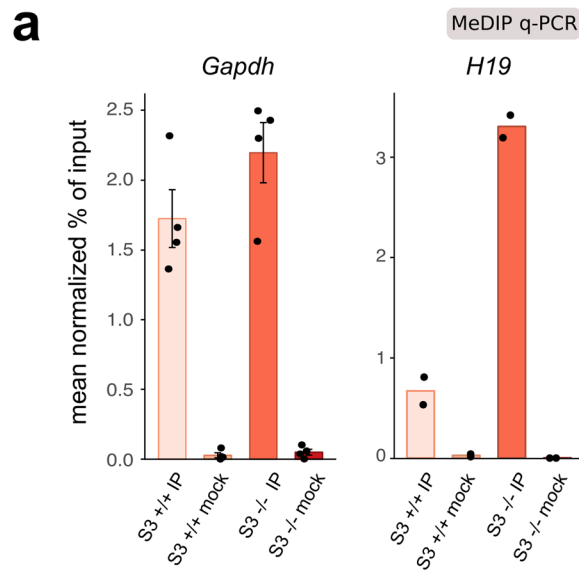
b

DMR	gene	methylation status of S3-/- cells (RRBS)	gene expression (expected)	gene expression (observed)
PEG13 chr15: 72806335 - 72811649	Peg13	+		
H19/IGF2 chr7: 142580263 - 142582519 chr7:142580772 - 142582022	H19	+		
	Igf2	+		
	Igf2os	+		
SNURF/SNRPN PWS-IC chr7:60004992 - 60005415	Ndn	+		
	Magel2	+		
	Atp10a	+		
KCNQ1OT1 KvDMR1, LIT1 chr7:143295155 - 143295492 chr7: 143295599 - 143295622	Ascl2	+		
	Cdkn1c	+		
	Phlda2	+		
	Osbpl5	+		
PEG10 chr6:4747209-4747507	Pon2	+		
	Ppp1r9a	+		
	Peg10	+		
GRB10 chr11:12025555 - 12026332 chr11:12025528 - 12026024 chr11:120225482 - 12025787	Cobl	+		
	Grb10	+		
Zac1 chr10:13090470 - 13090798	Plagl1	+		
IGF2R/AIRN chr17:12741297 - 12742707	Slc22a3	+		

■ Upregulated ■ Downregulated

c

All (expressed) genes — Down (1120), Up (1229), notDE (10161)

Imprinted genes — Down (15), Up (18), notDE (32)

■ Upregulated
■ Downregulated
■ Not Differentially Expressed
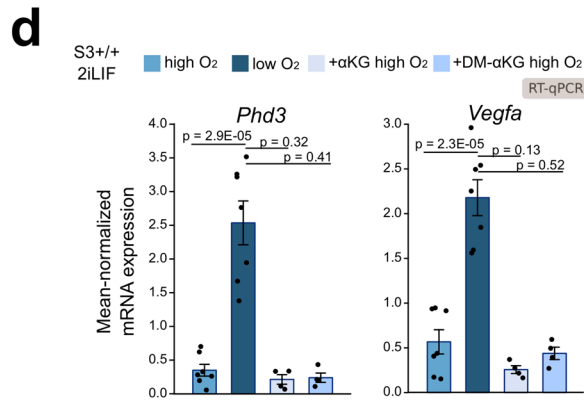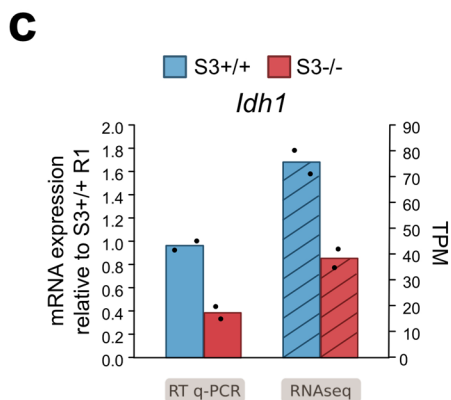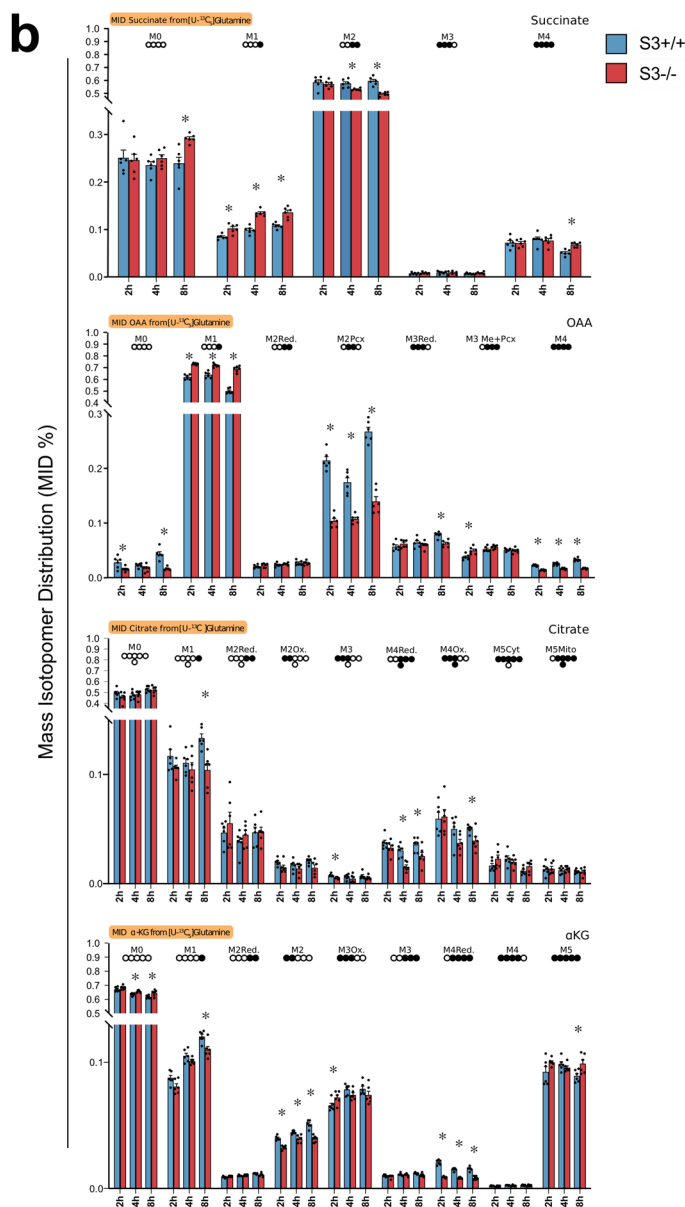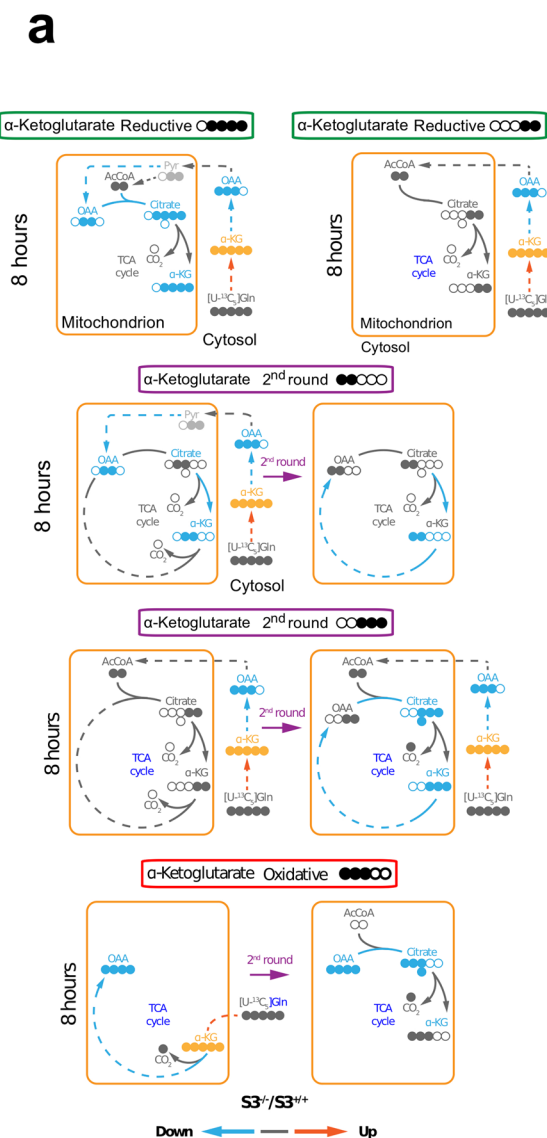
Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | Effect of Stat3 inactivation on imprinted transcripts. a, MeDIP-qPCR of two control regions for DNA methylation change (Gapdh as negative and H19 as positive control). Mock immunoprecipitations with a non-specific IgG antibody served as negative controls. Mean ± s..e.m of 4 experiments for Gapdh and mean of 2 experiments for H19, shown as dots. **b**, Differentially Methylated Regions and associated imprinted genes. Table reports Differentially Methylated Regions (DMRs) analyzed by RRBS with associated coordinates (reference genome: mm10); information about these DMRs in mouse genome was collected from three different databases (WAMIDEX https://atlas.genetics.kcl.ac.uk/; MouseBook - Imprinting Loci https://www.mousebook.org/; Geneimprint http://geneimprint.com/site/genes-by-species.Imprinted genes whose expression is controlled by the same DMR are grouped accordingly and indicated in the second column of the table. For each imprinted gene, fourth column reports the expected effect on gene expression - either upregulation or downregulation - caused by methylation deposition at the associated DMR (data from literature). The last column of the table shows observed expression levels (RNAseq data) of each imprinted gene in S3-/- cells, were hypermethylation was detected at the corresponding DMR (see Fig. 2f). **c**, Pie charts showing the number of up- and down-regulated genes (q-value < 0.01, Benjamini-Hochberg adjustment and log2 FC > 1 or < -1) in S3-/- cells with respect to S3+/+ cells among all expressed genes (left), or among all expressed imprinted genes (right).
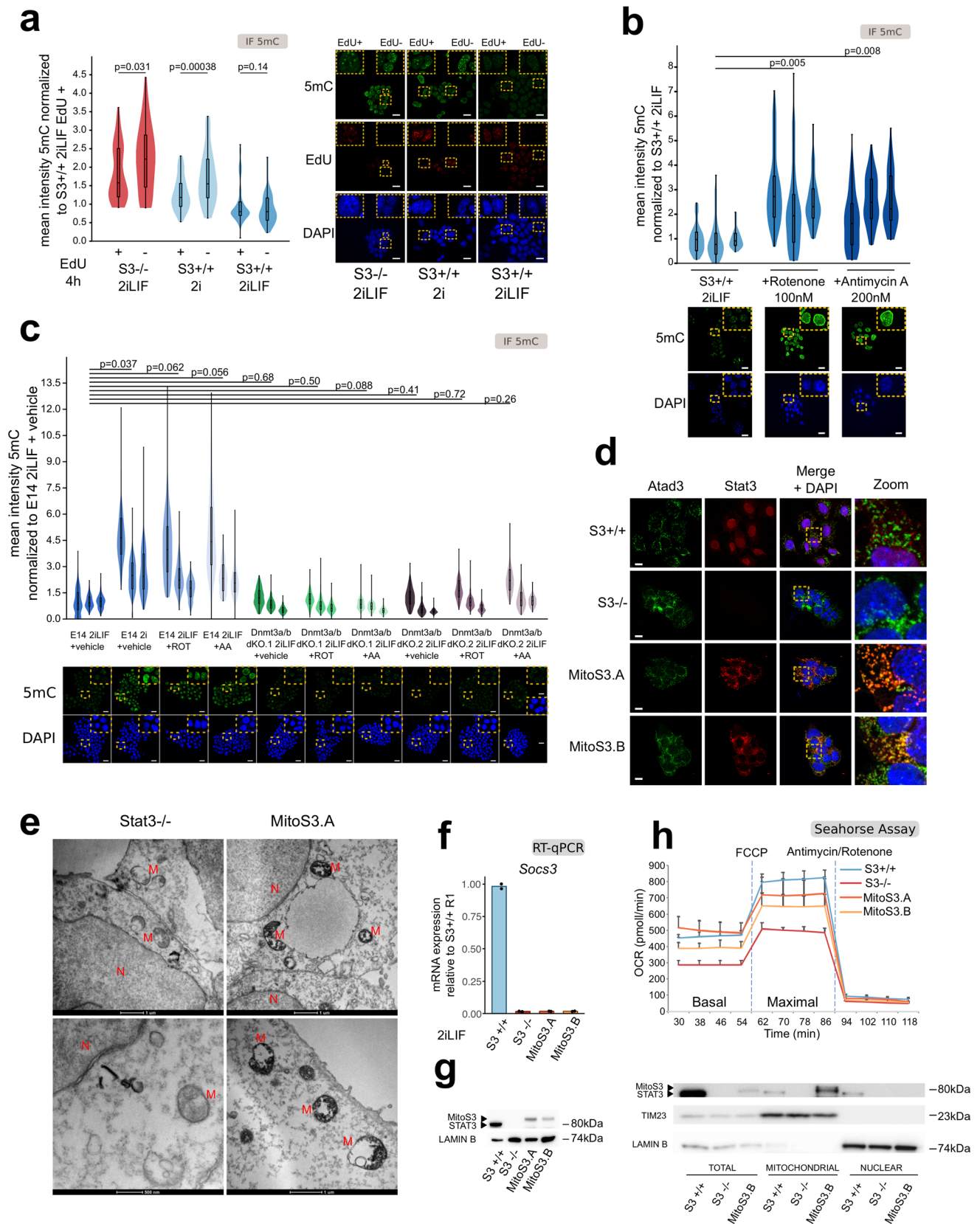
Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | Modulating mitochondrial activity affects 5mC levels. a, Anti-5mC immunofluorescence on S3+/+ cells in 2i or 2iLIF and S3-/- cells in 2iLIF treated with EdU for 4 h. Violin plots show the distribution of an average of 67 nuclei per sample; one representative experiment is shown for each condition. **b**, Anti-5mC immunofluorescence on S3+/+ cells treated with Rotenone or Antimycin A. Violin plots show the distribution of an average of 74 nuclei per sample. 3 independent experiments shown as individual violins. **c**, Anti-5mC immunofluorescence on E14 in 2iLIF and 2i, and on Dnmt3a/b double KO cells in 2iLIF, treated with Vehicle, Rotenone or Antimycin A. Violin plots show the distribution of an average of 183 nuclei per sample. 3 experiments shown as individual violins. **d**, Confocal images of S3+/+, S3-/- cells and MitoS3.A/B clones stained with anti-Stat3 and anti-Atad3 antibodies. Representative images of 3 independent experiments. **e**, Electron Microscopy images of STAT3 protein stained by Diaminobenzidine photooxidation method, in S3-/- and MitoS3.A cells. Representative images of 2 experiments. M, mitochondria; N, nucleus. **f**, Expression analysis of Socs3 in S3+/+, S3-/- cells and MitoS3.A/B clones in 2iLIF. Bars: mean of n = 2 experiments, shown as dots. **g**, (Left) Western blot of S3+/+, S3-/- cells and MitoS3.A and MitoS3.B clones in 2iLIF. LAMIN B: loading control. (Right) Western blot of total lysates or mitochondrial and nuclear fractions. The nuclear protein LAMIN B and mitochondrial marker TIM23 confirmed successful nuclear and mitochondrial isolation. Representative images of 2 independent experiment. **h**, Oxygen consumption rate measured by Seahorse extracellular flux assay of S3+/+, S3-/- and MitoS3.A/B clones cultured in 2iLIF. Mean and S.D. of n = 5 biological replicates is shown. All violin and boxplots indicate the 1st, 2nd and 3rd quartiles, with whiskers indicating minimum and maximum value. All p-values calculated by two-tailed unpaired T-test. Scale bars: 20μm.
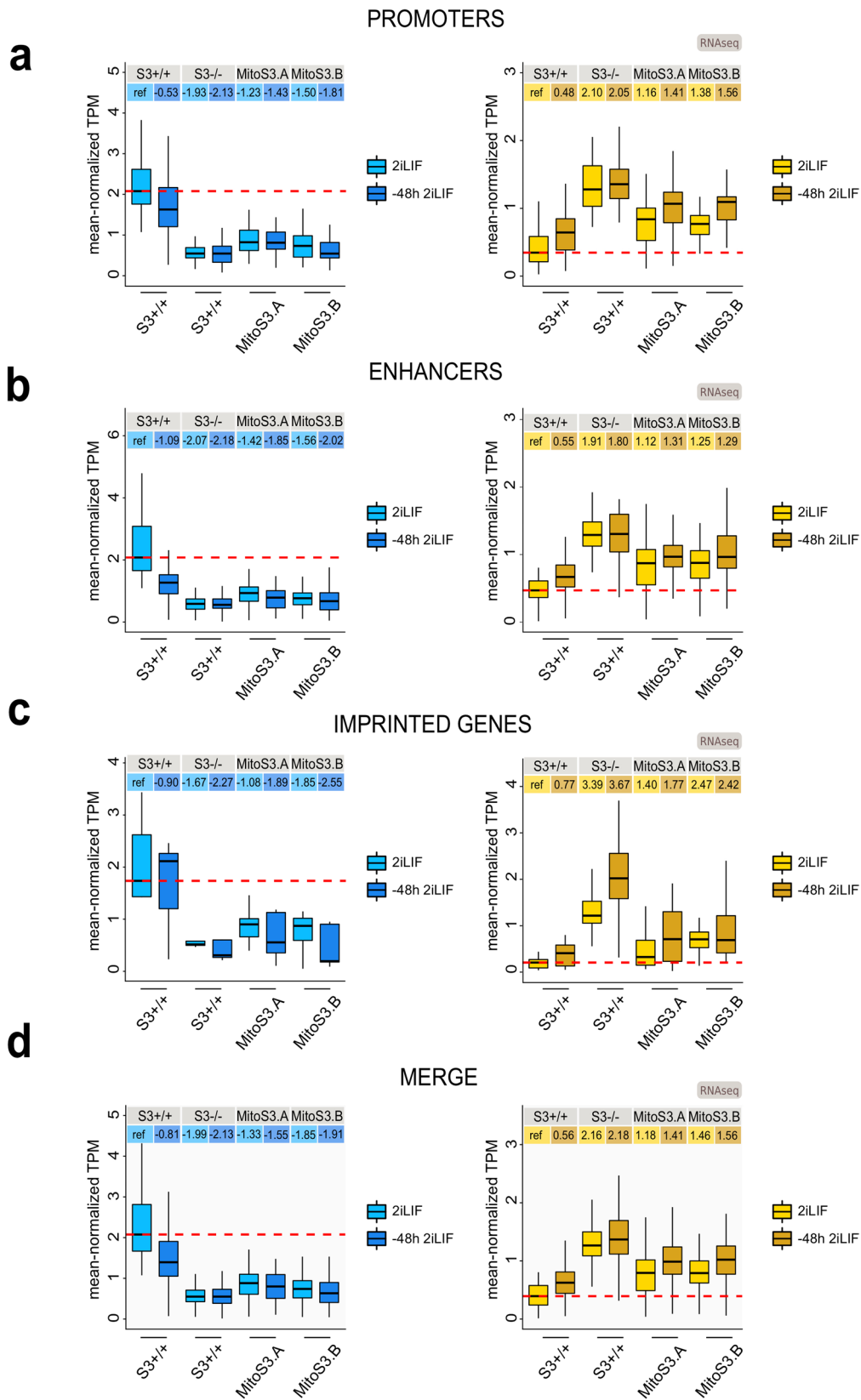
Extended Data Fig. 5 | See next page for caption.

Extended Data Fig. 5 | Metabolic reconfiguration following Stat3 deletion. a, Diagram representing mass isotopomer distribution (MID) of OAA, Citrate and αKG in both oxidative and reductive Glutamine pathways; MID was analysed following 8 h of metabolic tracing with [U-^{13}C$_5$]-Glutamine. Orange box: mitochondrion. Full circles: ^{13}C-labeled carbons. Color scale outlines the comparison between MID profile in S3-/- relative to S3+/+for n = 6 biological replicates; blue: isotopomers (or biochemical pathways) under-represented is S3-/- cells; red: isotopomers or pathways over-represented in S3-/-cells. Each isotopomer is corrected for natural isotope abundances. **b**, Metabolic tracing analysis of different isotopomers of TCA cycle intermediates (Succinate, Oxaloacetate, Citrate and αKG) using [U-^{13}C$_5$]-Glutamine. Barplot represents mass isotopomer distribution (MID %) at 3 different time points (2 h, 4 h, 8 h). Black circles: ^{13}C-labeled carbons. Bars: mean ± s.e.m of n = 6 biological replicates. * p-value<0.05, two-tailed unpaired T-test. Each isotopomer is corrected for natural isotope abundances. **c**, Expression of IDH1, measured by RT q-PCR and RNAseq in S3+/+ and S3-/- cells. Bars: mean of n = 2 biological replicates. **d**, Gene expression analysis of two Hif1a targets in S3+/+ cells cultured in 2iLIF in normoxia (high O$_{2-21}$%), in hypoxia (low O$_2$, 21%), or in normoxia with the addition of αKG or DM-αKG. Bars: indicate mean ± s.e.m. of n = 7 experiments for high O$_2$ and low O$_2$, and n = 4 for treatments with αKG and DM-αKG, shown as dots.
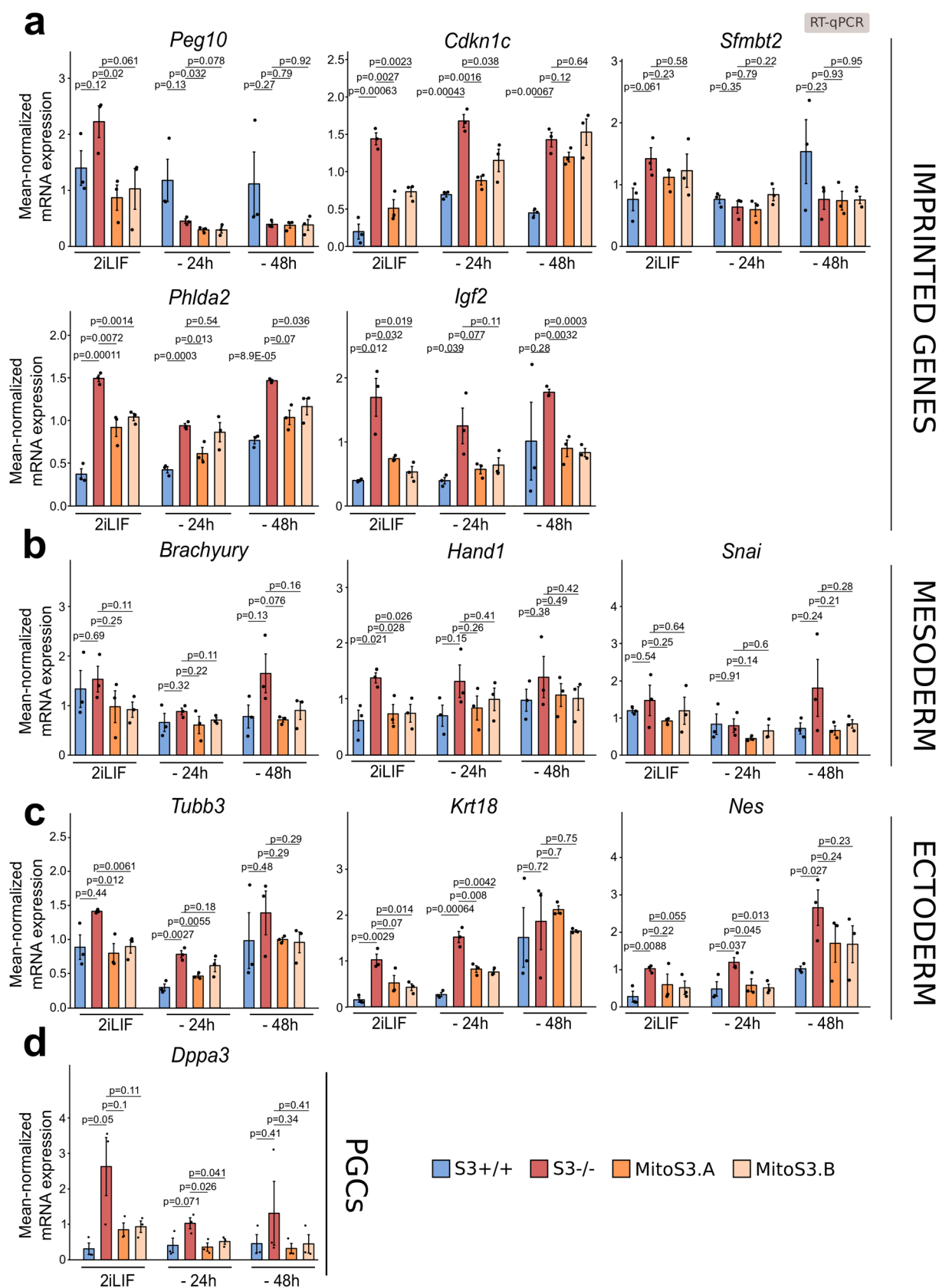
Extended Data Fig. 6 | Mitochondrial Stat3 slows down ESC differentiation. Gene expression analysis by RT–qPCR of S3+/+, S3-/- and MitoS3.A/B clones cultured with 2iLIF or without 2iLIF for 24 h or 48 h. Data show expression of Imprinted genes **a**, Mesoderm **b**, Ectoderm **c**, and PGCs **d**, markers that are more readily induced in S3-/- and MitoS3.A/B clones rescues this effect. Beta-actin served as an internal control. Bars indicate mean ± s.e.m. of n = 3 independent experiments, shown as dots. Two-tailed unpaired T-test relative to S3-/- for each time point.

Extended Data Fig. 7 | See next page for caption.

**Extended Data Fig. 7 | Stat3 regulates transcripts associated to differentially methylated genomic features. a**, Left: boxplot reporting expression levels of genes down-regulated in S3-/- cells relative to S3+/+ cells (Fig. 7a, blue dots) and differentially methylated at promoter regions (Fig. 2b). Right: boxplot reporting expression levels of genes up-regulated in S3-/- cells with respect to S3+/+ cells (Fig. 7a, yellow dots) and with differential methylation at promoter regions (Fig. 2b). Each boxplot shows 1st, 2nd and 3rd quartile. Whiskers shows minimum and maximum values. Y axis represents mean-normalized TPM values for S3+/+, S3-/- and MitoS3.A and MitoS3.B) in two different conditions: following stable culturing of cells in 2iLIF (light color) and after 48 h of 2iLIF withdrawal from culture medium (dark color). **b**, Boxplot reporting expression levels of genes down-regulated (left) or up-regulated (right) in S3-/- cells relative to S3+/+ cells and differentially methylated at enhancer regions (Fig. 2c), as described above. **c**, Boxplot reporting expression levels of genes down-regulated (left) or up-regulated (right) in S3-/- cells relative to S3+/+ cells and differentially methylated at the associated DMR (Fig. 2f), as described above. **d**, Merge of genes contained in boxplots shown in Extended Data Fig. 7a–c.
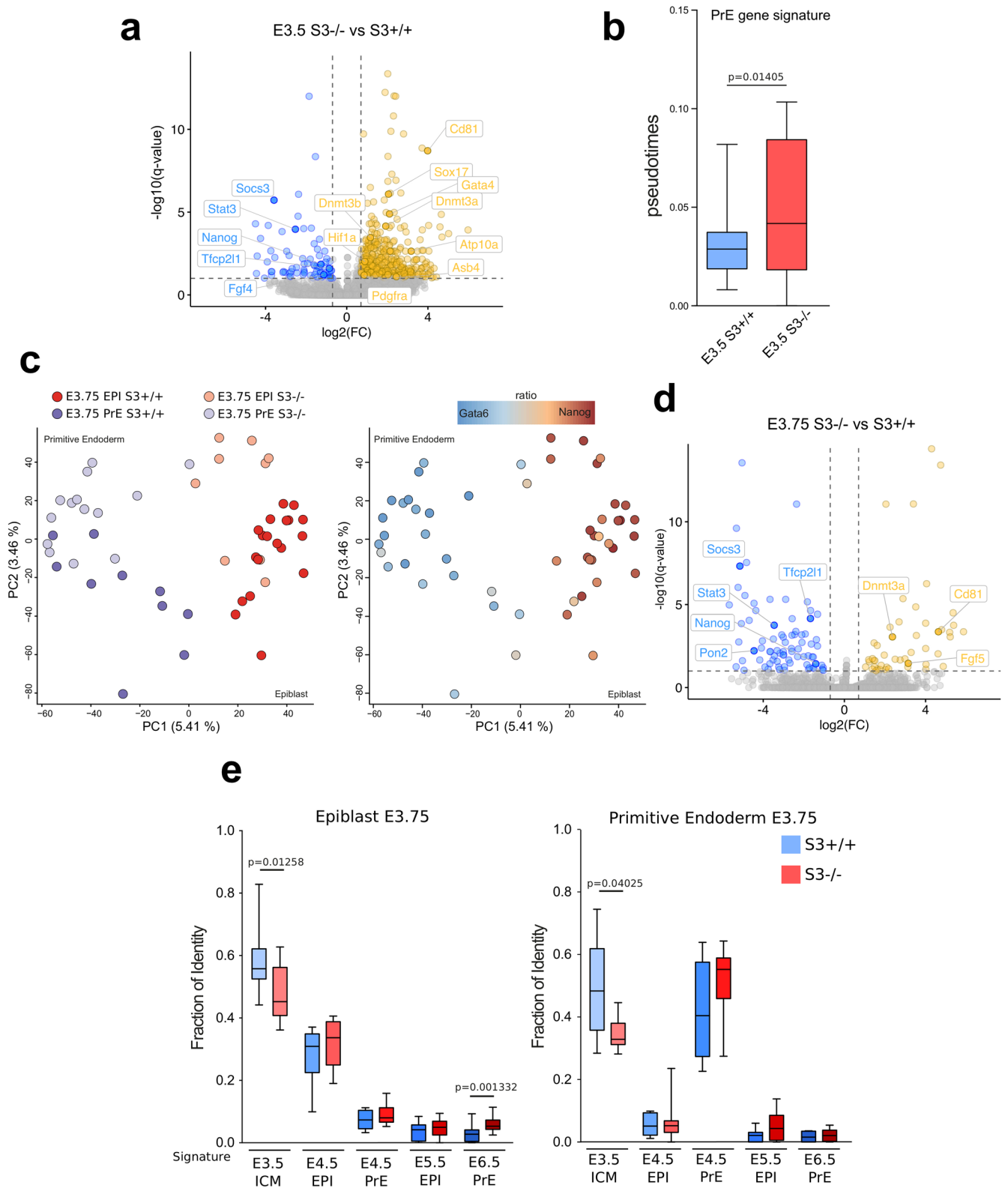
Extended Data Fig. 8 | See next page for caption.

Extended Data Fig. 8 | Accelerated progression of Stat3 null embryos. a, Volcano plot of genes differentially expressed between S3-/- and S3+/+ cells at E3.5. Red and blue dots indicate respectively transcripts that are upregulated or downregulated (log2 FC > 0.7 or FC < -0.7 respectively, q-value < 0.1) in S3-/- cells relative to S3+/+ cells. **b**, Diffusion pseudotime of E3.5 cells and the PrE gene signature computed with R package "destiny" (http://bioinformatics.oxfordjournals.org/content/32/8/1241) using all the expressed genes as input list. P-value calculated with two-tailed Mann-Whitney test. **c**, PCA plot computed with all the expressed genes. Colors represent different lineages/genotypes (left panel) or ratio between Gata6 and Nanog expression (right panel). **d**, Volcano plot of genes differentially expressed between S3-/- and S3+/+ cells at E3.75. Red and blue dots indicate respectively transcripts that are upregulated or downregulated (log2 FC > 0.7 or FC < - 0.7 respectively, q-value < 0.1) in S3-/- cells relative to S3+/+ cells. **e**, Fraction of identity between E3.75 EPI(left panel)/PrE(right panel) and E3.5 ICM, E4.5 EPI, E4.5 PrE, E5.5 EPI and E6.5 EPI stages computed with all the expressed genes. P-value calculated with two-tailed Mann-Whitney test. All boxplot shows 1st, 2nd and 3rd quartile. Whiskers show minimum and maximum values.
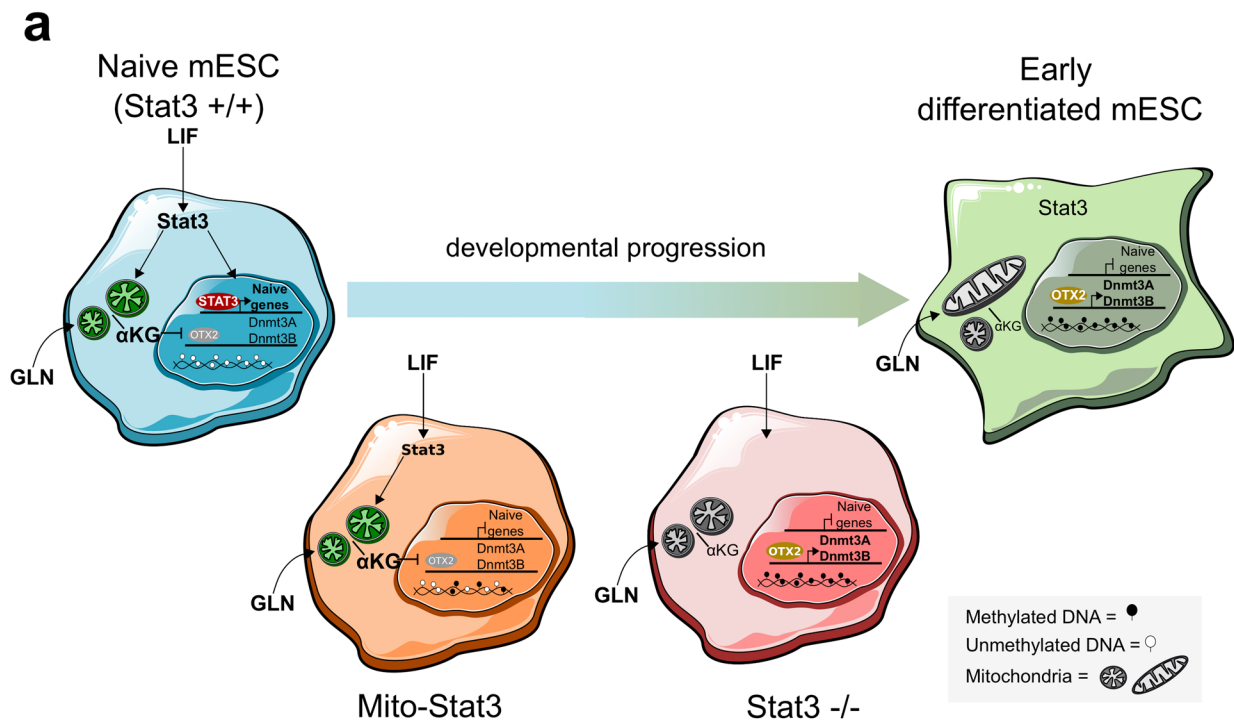
Extended Data Fig. 9 | Diagram summarizing how mitochondrial Stat3 affects nuclear transcription.

	Corresponding author(s):	Graziano Martello
	Last updated by author(s):	Oct 9, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

| Data collection | For qPCR data collection the Applied Biosystems™ QuantStudio™ 6 Flex Real-Time PCR System version 1.3 was used.
For immunostaining data collection the Leica SP5 LAS AF and the Zeiss LSN700 were used.
Western Blots were digitally acquired by ImageQuant LAS 4000.
For bulk and single-cell RNAseq, Illumina HiSeq4000 was used.
For RRBS, Illumina NextSeq 500 was used.
Mass Spectrometry data of total nucleosides were acquired on a LTQ orbitrap Mass Spectrometer (Thermo-Fisher) coupled to a U-3000 HPLC system (Thermo-Fisher).
Data on metabolites were acquired on an API-4000 triple quadrupole mass spectrometer (Sciex) coupled with a HPLC system (Agilent) and CTC PAL HTS autosampler (PAL System) and on API3500 (Sciex) coupled with UHPLC system and HTS autoampler (Sciex).
Data on proteins were acquired by a NanoLC 1200 coupled via a nano-electrospray ionization source to the quadrupole-based Q Exactive HF benchtop mass spectrometer.
Electron microscopy images were acquired with a Tecnai G2 (FEI) transmission electron microscope operating at 100 kV and captured with a Veleta (Olympus Soft Imaging System) digital camera.
Oxygen consumption was measured using the Seahorse XF24 (Seahorse Bioscience). |
|---|---|
| Data analysis | A detailed description of all software, including software version and parameters used, for the analysis of RNA-sequencing data, RRBS data and single-cell RNA-sequencing data has been provided in the Methods section.
Image analysis was performed with FIJI version 2.0.0.
All statistical analyses were carried out in R environment (v. 3.5.2) with Bioconductor (v. 3.7).
For bulk RNAseq data processing reads were aligned to mouse transcriptome reference (Mus musculus transcriptome generated by rsem-prepare-reference with ENSEMBL93 GTF) and mouse genome (mm10/GRCm38.p6) using HISAT2 (v. 2.1.0). Gene expression levels were quantified with RSEM (v. 1.3.1) using transcriptome alignments. Genome alignment were used to create bigWig files using deeptools (v. 3.2.1). Differential expression analysis was computed using the DESeq2 R package (v. 1.24.0).
For RRBS data processing, sequencing reads were mapped to mouse genome reference (mm10/GRC.m38.p6) with BSMAP (v. 2.89) using RRBS |

mode. Differential methylation analysis at single nucleotide resolution was performed using the methylKit R/Bioconductor package (v. 1.16.0), exploiting the logistic regression approach for testing replicates (calculateDiffMeth function with default parameters).
For metabolites analysis, MultiQuantTM software (v. 3.0.2, Sciex) was used for data analysis and peak review of chromatograms.
For proteomics analysis, all acquired RAW files were processed using MaxQuant (v. 1.6.2.10) and the implemented Andromeda search engine; for protein assignment, spectra were correlated with the UniProt mouse database (v. 2019) including a list of common contaminants. For further analysis, the Perseus software (v.1.6.2.3) was used.
For single-cell RNAseq data processing, Mus musculus GRCm38.87 gene annotation and mm10 genome sequence were used. All reads were aligned STAR (v. 2.5.2b). Alignments to gene loci were quantified with htseq-count (v. 0.11.0), based on annotation from Ensembl 87. Principal component analysis was based on Log2 FPKM values computed with the Bioconductor package DESeq (v. 1.42.0), custom scripts and FactoMineR package (v. 2.3). t-SNE clustering was computed with monocle package (v. 2.18.0). Differential expression analysis was performed with DESeq2 package (v. 1.24.0) following pre-processing of raw counts with zinbwave package (v. 1.12.0). Fractional identity between E3.5/E3.75 S3 +/+ and S3 -/- cells and embryo stages (E4.5 EPI, E5.5 EPI and E6.5 EPI) was computed using R package DeconRNASeq (v. 1.32.0).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
  - Accession codes, unique identifiers, or web links for publicly available datasets
  - A list of figures that have associated raw data
  - A description of any restrictions on data availability

Bulk and single-cell RNA sequencing data and RRBS data generated during the current study are available via the Gene Expression Omnibus (GEO) repository under the accession numbers GSE133926 and GSE134450. All RNA-seq and RRBS process data, used in
Figures 1d, 1i, 2a-e, 2h, 4c, 5d, 6b, 6d, 7a-e, 8a-g, Extended Data Fig. 1a-b, Extended Data Fig.1e, Extended Data Fig. 2b-k, Extended Data Fig. 3b-c, Extended Data Fig. 5c, Extended Data Fig. 7a-d, Extended Data Fig. 8a-f are reported in Supplementary Tables 1, 2 and 4 and via the Gene Expression Omnibus (GEO) repository under the accession numbers GSE133926. RNA sequencing data of Rex1-GFPd2 cells can be found at GEO under accession number GSE111694. Mass spectrometry proteomics data of following samples: S3+/+ cells in 2i; S3+/+, S3-/-, MitoS3.A and MitoS3.B cells in 2iLIF; used in Figures 1f, 5g, Extended Data Fig. 1c are reported in Supplementary Table 3 and ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD020385. Primers and oligonucleotides sequences are present in Supplementary Tables 6-7-8. Additional data that support the findings of this study, such as Mass Spectrometry measurements, analysis pipelines and reagents are available from the corresponding authors upon reasonable request.
The source data underlying Figs. 1a-b,e,g-h; 2g; 3a-b,d,f; 4a-b,d-e,g-i; 5a-c,e-f,h-j; 6a,c,e-f; 7f-g and Extended Data Figs. 1b,f-i; 2a; 3a;4a-c,f-h; 5b-d; 6a-d are provided as Source Data files.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences   ☐ Behavioural & social sciences   ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample size calculation was not performed. Sample size was chosen based on standards in the field (Ying et al. 2008 Nature; Song et al. 2019 Molecular Cell; Carey et al. 2015 Nature; Boroviak et al. 2018 Dev. Camb. Engl.) and is indicated in figure legends.
Data exclusions	We presented all data obtained, also in the case of negative results.
Replication	All key experiments were repeated between two and six times independently, as indicated in figure legends. Each independent qPCR experiments were performed with three technical replicates. For Next-Generation sequencing experiments we analysed at least 2 independent biological replicates for each sample. Mass Spectrometry analyses of nucleosides were performed on at least 3 biological replicates, Mass Spectrometry analyses of proteins were performed on at least 5 biological replicates, while Mass Spectrometry analyses of metabolites were performed on at least 5 biological replicates and similar results were obtained in at least two independent experiments.
Randomization	Randomization was not relevant. All cell lines or biological samples were analysed or treated in the same manner.
Blinding	We applied blinding for RNA-seq and RRBS processing and data analysis. Investigators were not blind to cell line identities during all other analyses as knowledge of identity was necessary in order to perform the analysis, but key results were independently replicated by two different operators.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Please see Table 5 for the list of antibodies, product codes, dilutions used for Immunostaining and Western Blot. Antibodies for immunofluorescence: 5mC (1:250, EUROGENTEC, BI-MECY); Dnmt3b (1:500, ABCAM, ab122932); Stat3 (1:100, Cell Signalling, cat. 9139); Atad3 (1:100, AB-Biotechnologies, cat. 224485); Secondary Antibody Alexa Fluor 488 (1:500, ThermoFisher, A-21202), Donkey anti-Mouse IgG (H+L); Secondary Antibody Alexa Fluor 488 (1:500, ThermoFisher, A-21206), Goat anti-Rabbit IgG (H+L); Secondary Antibody Alexa Fluor 568 (1:500, ThermoFisher, A-11036), Donkey anti-Mouse IgG (H+L). Antibodies for Western Blot: Tet2 (1:500, ABCAM, Ab94580); Dnmt3a (1:500, Novus Biological, NB120-13888), Dnmt3b (1:1000, Novus Biological, NB100-56514), Lamin B (1:1000, Santa Cruz Biotechnologies, cat. 6216), Stat3 (1:1000, Cell Signalling, cat. 9139), TIM23 (1:1000, BD Biosciences, cat. 611223) Antibodies for MEDIP: 5mC (10 μg , EUROGENTEC, BI-MECY); Rabbit IgG (10 μg Sigma-Aldrich, 12-370).
Validation	All antibodies chosen for this work have been previously used and validated in literature and by manufacturers. Antibodies were used following manufacturers guidelines and were further validated in our laboratory by using relevant controls or, when possible, independent techniques. See Table 5 for all relevant details. Antibodies for immunofluorescence: 5mC staining quantification was confirmed by Mass Spectrometry for 5mC; 5mC antibody was also previously validated in M. Habib et al. 1999 Experimental Cell Research. Dnmt3b antibody was previously validated in Nakanishi et al., 2019 Cell. Stat3 staining is absent in Stat3-/- cells; Stat3 antibody was also previously validated in Carbognin et al., 2016 EMBOJ. Atad3 antibody was previously validated in He et al., 2007 Journal of Cell Biology. Antibodies for Western Blot: Tet2, Dnmt3a and Dnmt3b antibodies were previously validated in Neri et al., 2015 Cell Reports. Lamin B was previously validated in Su et al., 2013 Molecular Cell Biology. Stat3 signal is absent in Stat3-/- cells; Stat3 antibody was also previously validated in Carbognin et al., 2016 EMBOJ. TIM23 antibody was previously validated in Kang et al. 2016 eLife. Antibodies for MEDIP: 5mC antibody for immunoprecipitation was previously validated in Weber 2005 Nature Genetics.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	Stat3+/+, Stat3-/- and E14IVC mouse ES cells were generated and kindly provided by Austin Smith's and Jennifer Nichols' laboratories. Dnmt3a KO, Dnmt3b KO, Dnmt3a/b dKO cell lines were generated by Salvatore Oliviero's Lab from E14IVC cells. Otx2-/- cell line was generated and kindly provided by Antonio Simeone's lab. Stable transgenic ESCs lines expressing MLS-Stat3-NES were generated by transfecting Stat3-/- cells with CAG-MLS-Stat3-NES and with piggyBac transposase expression vector pBase. Stable transgenic ESCs lines expressing sh-TET1/2 and sh-Scramble were generated by transfecting Stat3+/+ cells with PiggyBac transposon plasmids CAG-sh-TET1/2 or CAG-sh-Scramble and with piggyBac transposase expression vector pBase. Selection for transgenes was applied, and stable clones were selected in 2iLIF conditions.
Authentication	Stat3+/+, Stat3-/-, E14IVC, Dnmt3a KO, Dnmt3b KO, Dnmt3a/b dKO were regularly validated by PCR, Western Blot and immunostaining. MLS-Stat3-NES clones were authenticated by PCR, Western Blot, immunostaining and electron microscopy. Otx2-/-, sh-TET1/2 and sh-Scramble cells were validated by qPCR.
Mycoplasma contamination	Cells were routinely tested for Mycoplasma contamination. All cell lines were Mycoplasma negative.
Commonly misidentified lines (See ICLAC register)	The cell lines used are not listed in ICLAC.