Metabolic control of DNA methylation in naive pluripotent cells

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

Affiliations:
[1]Department of Molecular Medicine, Medical School, University of Padua, Padua 35121, Italy.
[2]Department of Pharmacological and Biomolecular Sciences (DiSFeB), University of Milan, Milan, Italy.
[3]Department of Life Sciences and Systems Biology, University of Turin, Turin 10123, Italy
[4]Italian Institute for Genomic Medicine (IIGM), Via Nizza 52, Torino 10126, Italy.
[5]Department of Molecular Genetics, Groningen Biomolecular Sciences and Biotechnology Institute (GBB), University of Groningen, Nijenborgh 7, 9747 AG, Groningen, the Netherlands.
[6]Department of Biology, University of Padova, Via U. Bassi 58B, 35121 Padova, Italy
[7]Université Paris-Saclay, CNRS, Institut de Chimie des Substances Naturelles, UPR 2301, 91198, Gif-sur-Yvette, France.
[8]Wellcome-MRC Cambridge Stem Cell Institute, University of Cambridge, Cambridge CB2 1QR, UK.
[9]Department of Physiology, Development and Neuroscience, University of Cambridge, Tennis Court Road, Cambridge CB2 3EG, UK.
[10]Telethon Institute of Genetics and Medicine (TIGEM), Pozzuoli, Italy.

*Correspondence to: nico.mitro@unimi.it (N.M.), salvatore.oliviero@unito.it (S.O.) and graziano.martello@unipd.it (G.M.)

†Current address: Department of Molecular Genetics, Groningen Biomolecular Sciences and Biotechnology Institute (GBB), University of Groningen, Nijenborgh 7, 9747 AG, Groningen, the Netherlands

Abstract

Naive pluripotent epiblast cells of the preimplantation murine embryo and their *in vitro* counterpart, embryonic stem (ES) cells, have the capacity to give rise to all cells of the adult. Such developmental plasticity is associated with a transient global genome hypomethylation. However, it is still unclear how genome methylation is dynamically regulated. Here we show that LIF/Stat3 signaling induces genomic hypomethylation via metabolic reconfiguration. In Stat3-/- ES cells we observed decreased Alpha-ketoglutarate (αKG) production from reductive Glutamine metabolism, leading to increased Dnmt3a/b expression and to a global increase in DNA methylation. Notably, genome methylation is dynamically controlled by simply modulating αKG availability, mitochondrial activity, or Stat3 activation in mitochondria, indicating effective crosstalk between metabolism and the epigenome. Molecularly, αKG reduces the expression of Otx2 and its targets Dnmt3a/b. Genetic inactivation of Otx2 or Dnmt3a/b results in genomic hypomethylation even in the absence of active LIF/Stat3, while Tet1/2 inhibition is inconsequential. Stat3-/- ES cells also show increased methylation at Imprinting Control Regions accompanied by differential expression of cognate imprinted transcripts. Single-cell transcriptome analysis of Stat3-/- embryos confirmed the dysregulated expression of Otx2, Dnmt3a/b and imprinted genes *in vivo*. Our results reveal that the LIF/Stat3 signal bridges the metabolic and epigenetic profiles of naive pluripotent cells, ultimately controlling genome methylation via Dnmt3a/b regulation. A wide range of cancers displays Stat3-overactivation and abnormal DNA methylation, raising the possibility that the molecular module we describe here is exploited under pathological conditions.

Main Text:

After fertilization, the zygotic genome is demethylated in order 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 (5hmC[1,2]). Additional oxidation steps mediated by TETs lead to the conversion of 5hmC into unmodified cytosine. DNMTs and TETs are dynamically expressed during early development, leading to a local minimum of 5mC at the pre-implantation blastocyst stage at E3.5[3–5]. Imprinted genes, expressed monoallelically in a parent-of-origin fashion, resist this wave of DNA demethylation. Such monoallelic expression allows tight control of their dosage and is essential for the proper development of the embryo[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 and transcriptional activation of its targets Socs3 and Tfcp2l1[7–11]. Thus, Stat3 could represent a good candidate as a regulator of Dnmt and Tet expression.

Mouse embryonic stem (ES) cells are in a naïve pluripotent state and share distinguishing molecular features with the preimplantation epiblast[12]. In particular, mES cells show genomic hypomethylation, similarly to the E3.5 blastocyst, but only when cultured in the presence of LIF, a ligand of the Jak/Stat pathway, in combination with two kinase inhibitors of GSK3 and MEK (2iLIF conditions[13–17]). Such hypomethylation was attributed to the downregulation of Dnmt3a/b by the MEK inhibitor[14–17]. Conversely, mES cells 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 not sufficient to induce genomic hypomethylation in

presence of serum, but the requirement of LIF or Stat3 to induce hypomethylation in 2iLIF conditions has never been formally tested.

Stat3 may represent an ideal regulator of the epigenome, considering its capacity to regulate gene expression in the nucleus, together with the ability to control cellular metabolism in mitochondria by promoting oxidative phosphorylation (OXPHOS[19–21]). Several metabolites are known as regulators or cofactors of enzymes catalyzing epigenetic modifications[22]. For all these reasons, we genetically tested the role of LIF/Stat3 axis on genome methylation of naive pluripotent cells.

Results:

LIF/Stat3 induces hypomethylation in mES cells via Dnmt3a/b regulation in 2i

We measured the levels of 5mC by quantitative immunostaining in mES cells and observed a strong decrease in signal intensity in 2iLIF compared to Serum LIF (Fig. 1a), as previously reported[14–17]. We then asked whether LIF and its downstream mediator Stat3 could be required for the decrease in 5mC. Wild-type (S3+/+) mES cells stably expanded in 2i, or Stat3-/- (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 in order to unequivocally identify global 5mC and unmethylated cytosine and confirmed that only S3+/+ cells in 2iLIF showed a reduced fraction of methylated cytosine (Fig. 1b). We further confirmed our findings by Reduced Representation Bisulfite Sequencing (RRBS, Figs. 1c and S1a). We conclude that active LIF/Stat3 signaling is required, in combination with 2i, to induce genomic hypomethylation.

We then asked how LIF/Stat3 could 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 showed reduced expression only of *de novo* methyltransferases Dnmt3a and Dnmt3b and increased expression of Tet2 compared to S3+/+ cells in 2i or to S3-/- cells (Fig. 1d), while the maintenance methylase Dnmt1 was not regulated by LIF. Similar effects induced by LIF were observed also in an independent mES cell line (Rex1-GFPd2-RGd2[23] Fig. S1b) and were confirmed at the protein level by both Western Blot (Fig. 1e) and Mass Spectrometry (Figs. 1f and S1c).

We then asked whether the hypomethylation observed in 2iLIF was dependent on Dnmt3a/b.

We first confirmed in an independent wild-type mES cell line (E14IVc) that culture in 2i led to hypermethylation relative to 2iL (Fig. 1g). We then analyzed two independent mutant clones for each genotype of Dnmt3a KO, Dnmt3b KO, and Dnmt3a/b double KO (dKO) mES cells (Fig. S1d). When cultured in 2i without LIF, Dnmt3a KO and Dnmt3b KO cells displayed a partial reduction of 5mC relative to wild-type cells in 2i, while Dnmt3a/b dKO cells cultivated in 2i were hypomethylated (Fig. 1g). Mass Spectrometry (Fig. 1h) and RRBS (Figs. 1i and S1e) further confirmed that Dnmt3a/b dKO cells in 2i displayed DNA methylation levels even lower than wild-type cells in 2iL. We conclude that the levels of Dnmt3a/b dictate the DNA methylation status of naive ES cells in 2i. Such conclusions were further supported by the overexpression of Dnmt3a and Dnmt3b in S3+/+ cells in 2iLIF, which led to increased 5mC levels (Figs. S1f-g).

We also tested whether 5mC oxidases could have a role in the observed hypomethylation induced by LIF in presence of 2i. Tet1 and Tet2 are both robustly expressed in mES cells in 2iL (Supplementary Table 1) and have redundant functions[24]. Thus, we knocked down Tet1 and Tet2 simultaneously in S3+/+ 2iLIF and observed no significant changes of 5mC (Fig. S1h-i). We conclude that Tet proteins do not appear to regulate 5mC levels in 2iL.

Impact of Stat3 on DNA methylation and transcription

We then asked which genomic regions showed a DNA methylation profile dependent on Stat3. DNA methylation is particularly abundant at repetitive elements, but we could not observe differences in 5mC levels between S3+/+ and S3-/- cells in 2iLIF (Fig. S2a). Interrogation of RRBS data for S3+/+ and S3-/- cells in 2iLIF identifies 381,607 differentially methylated sites, with 98.7% of them displaying gain of methylation in S3-/- cells (Fig. 2a). We used H3K4me3 and H3K27ac profiles to identify promoters and enhancers in mES cells,

144 and measured the levels of DNA methylation at both genomic features. We observed
145 increased DNA methylation in S3-/- cells in 3.6% (323/8782) of promoters and 36.5%
146 (621/1701) of enhancers (Y axes in Fig. 2b-c), while only 2 out of 8782 promoters and 1 out
147 of 1701 enhancers showed decreased DNA methylation. We then intersected transcriptome
148 data, comparing S3+/+ and S3-/- cells, and asked whether the gene associated with each
149 promoter or enhancer was differentially expressed. The gain of DNA methylation at
150 promoters was associated with downregulation of cognate genes in 20.7% (67/323) of the
151 cases and with upregulation in 8% (26/323) of cases (Fig. 2b). For enhancers, we observed
152 13.8% (86/621) and 6.7% (42/621) of cognate genes significantly downregulated or
153 upregulated, respectively. Among significantly regulated genes we noticed the pluripotency
154 factors Klf5 and Esrrb[25,26]. We repeated the same analyses comparing S3+/+ in 2i and in
155 2iLIF (Figs. 2d-g and S2b-c) and obtained highly comparable results, clearly demonstrating
156 that absence of LIF or of Stat3 had overlapping effects on the transcriptome and 5mC profile
157 of ES cells.
158 Given that Dnmt3a/b appeared functionally relevant for the regulation of 5mC levels
159 downstream of LIF (Fig. 1g-i), we expected similar changes in DNA methylation in response
160 to LIF stimulation or upon Dnmt3a/b inactivation. Strikingly, 98.9% of regions
161 hypomethylated in 2iLIF were also hypomethylated in Dnmt3a/b dKO cells (Figs. 2h-i and
162 S2d-g), further indicating that Dnmt3a/b are epistatic to LIF/Stat3 for DNA methylation
163 control.
164
165 Imprinted genes are organized in clusters, under the control of Differentially Methylated
166 Regions (DMRs). Imprinted transcripts are expressed monoallelically in a parent-of-origin
167 fashion, allowing precise regulation of their dosage[27,28]. S3+/+ cells cultured long term in
168 2iLIF showed low levels of DNA methylation at imprinted DMRs, as previously reported[14,29],

while S3-/- cells retained higher levels of DNA methylation at 83.3% (20/24) of DMRs analyzed (Fig. 2j), despite the prolonged culture in the 2iLIF medium. These findings were independently validated by MeDIP-qPCR (Figs. 2k and S2h).

We checked whether aberrant DNA methylation at DMRs affected gene expression and observed that 37% (20/54) of imprinted transcripts expressed in mES cells were differentially expressed in S3-/- cells (Fig. 2l and Fig. S2i). Furthermore, the percentage of differentially expressed imprinted genes, without taking into account DNA methylation information, was 50.77% (33/65), while only 18.77% (2349/12510) of all expressed genes were differentially expressed, indicating that imprinted genes are specifically regulated by Stat3 (Figs. 2m and S2j, *P* value= 5.67 x 10-9, hypergeometric test).

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

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. S3a) or acutely stimulated with LIF for 1, 4, 24 and 48 hours, starting from 2i. The addition of LIF resulted in repression of Dnmt3a and Dnmt3b, however transcriptional changes were slow, requiring 24 hours to reach significance. Consistently, we observed a mild decrease in 5mC levels after 24 hours, while at 48 hours the levels of 5mC were as low as those of cells stably cultured in 2iLIF (Fig. S3b).

The slow kinetics observed would indicate that Stat3 does not directly repress the transcription of Dnmt3a/b. We interrogated available Stat3 ChIP-seq data and could not detect binding at enhancers or promoters of Dnmt3a/b[30]. Further, we expressed in S3-/- cells a Stat3 construct fused to an Estrogen Receptor domain[31], which localizes to the nucleus and

activates direct Stat3 targets Socs3 and Klf4 upon Tamoxifen treatment ([31,32] and Fig. S3c-d). Dnmt3a and Dnmt3b mRNA and protein levels were unaffected (Fig. S3d-f) and, crucially, 5mC levels were unchanged (Fig. S3e-f), indicating the nuclear Stat3 is not sufficient to reduce DNA methylation levels in ES cells in 2i. These results indicate that LIF/Stat3 induces a hypomethylated genome state, characterized by low Dnmt3a and Dnmt3b and that such effects are not explained by a direct transcriptional mechanism, therefore we sought for alternative mechanisms.

First, we thought that global 5mC levels could be affected by passive dilution occurring during genome replication. Given that the LIF/Stat3 axis promotes mES cells proliferation[19] we asked whether reduced 5mC in S3+/+ cells in 2iLIF could be due to enhanced proliferation. We tested this hypothesis by performing the EdU incorporation assay combined with 5mC immunostaining. For S3-/- cells in 2iLIF or S3+/+ cells in 2i, we observed that EdU positive cells that underwent genome replication within the last 4 hours showed an expected decrease in 5mC compared to EdU negative cells (Fig. S4a), which never reached the levels of S3+/+ in 2iLIF. Moreover, 5mC levels in S3+/+ cells in 2iLIF were only marginally affected 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 regulation of mitochondrial activity, given that it has previously been reported that S3-/- cells display reduced mitochondrial OXPHOS[19–21]. We, therefore, tested whether 5mC levels are dependent on changes in OXPHOS. First, we treated S3+/+ cells with inhibitors of Complex I and III of the respiratory chain at concentrations that reduce OXPHOS in mES cells[19] and observed a strong increase in 5mC signal (Fig. S4b). Importantly, Dnmt3a/b KO cells did not show any significant increase in 5mC upon inhibition of the respiratory chain (Fig. S4c), further indicating that changes in 5mC are dependent on Dnmt3a/b in mES cells. Second, we

expressed at endogenous levels a Stat3 construct targeted to mitochondria in S3-/- cells (Fig. S4d-g)[19,33]. The two clonal lines, called MitoS3.A and MitoS3.B, showed increased OXPHOS (Fig. S4h) and reduced 5mC (Fig. 3a-b) compared to parental S3-/- cells.

Given that expression of imprinted transcripts is tightly linked to 5mC levels on imprinted DMRs, we measured their expression in S3-/- and MitoS3 cells and found that 25 were differentially expressed (Fig. 3c), accounting for 38.46% of imprinted transcripts expressed in mES cells. We conclude that the hypomethylation observed in 2iLIF is linked to robust OXPHOS of mES cells.

Our results indicate that mitochondrial activity affects the methylation profile of the nuclear genome, which implies that the two organelles are able to communicate. We initially hypothesized that intracellular signaling molecules, such as Calcium ions or Reactive Oxygen Species (ROS) could be implicated, but we did not observe differences in their abundance between S3+/+ and S3-/- cells, nor effects upon blockade of such signal. We then reasoned that mitochondrial activity could influence the abundance of metabolites, in particular those serving as donors, acceptors or cofactors of DNA methylation and oxidation[22]. Analysis of steady-state levels of metabolites revealed a decrease in Alpha-Ketoglutarate (αKG) in S3-/- cells (Fig. 3d). We also noticed a strong increment in Methionine levels in S3-/- cells, however, nor SAMe, which is the actual donor of methyl groups to DNA and histones, neither the enzymes involved in Methionine/SAMe metabolism, such as MAT2A, MAT2B (Methionine Adenosyltransferase 2A/2B), AHCY (Adenosylhomocysteinase) and MTR (Methionine Synthase) were changed between S3+/+ and -/-, indicating that Methionine/SAMe metabolism might not be involved in DNA methylation regulation downstream of Stat3 (Supplementary Table 4).

We then asked what carbon source was preferentially used by mES cells for the production of αKG and performed metabolic flux analysis with ^{13}C-labelled Glucose, Glutamine or Palmitate by LC-MS/MS (see Methods). As previously reported[34], we observed that Glutamine represents the main source for production (Fig. S5a). 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 (Fig. S5b-c, Oxaloacetate, OAA M4 and Citrate M4), in line with impaired OXPHOS (Fig S4h). Alternatively, in the reductive Glutamine pathway, Glutamine is converted in the cytosol into TCA intermediates and Acetyl-CoA, which is diverted to fatty acid biosynthesis. Conversely, 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 (PCX). S3-/- cells showed impaired Glutamine reductive pathway. Specifically, we detected a strong decrease of cytosolic OAA M3 and mitochondrial OAA M2, Citrate M2 and M4, αKG M2 and M4 (Figs. 3e-f and S5b-c). These data are sustained by decreased expression of cytoplasmic isocitrate dehydrogenase 1 (IDH1, Fig. S5d). Finally, a mild increase of αKG M5 was found in S3-/- cells indicating impaired flux of Glutamine carbons into oxidative and reductive pathways (Figs. 3f and S5b-c).

Based on our observations we hypothesized that robust αKG production from Glutamine is required for genome hypomethylation. We measured αKG levels in cells expressing Stat3 only in mitochondria. Both MitoS3.A and MitoS3.B clones showed elevated αKG levels, not significantly different from S3+/+ cells (Fig. 3g). Both clones also showed reduced 5mC (Fig. 3a-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 strongly reduced (Fig. 3h) and observed a robust increase in 5mC (Fig. 3i).

Next, we asked whether restoring endogenous αKG levels could result in reduced 5mC. We added back a cell-permeable form of αKG (DM-αKG) and we were able to reduce 5mC levels (Fig. 3h-i). Of note, DM-αKG has been reported to stabilize Hypoxia-inducible factor 1-alpha (Hif1a) by inducing a pseudohypoxic state[35], but this was not the case in mES cells (Fig. S5e).

In sum, our results indicate that efficient αKG production from Glutamine induces low methylation levels of the nuclear genome.

αKG regulates Dnmt3a/b expression via the transcription factor Otx2

Next, we asked how αKG reduces 5mC levels in mES cells. Alpha-Ketoglutarate functions as a cofactor for Tet oxidases[22,36,37], thus 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 Dnmt3a/b expression levels[38]. Thus, S3-/- cells might show increased Dnmt3a/b expression and 5mC levels, as a consequence of reduced αKG levels. To investigate the relative contribution of these two possible mechanisms, we took advantage of our MitoS3 cells, where αKG levels were rescued to endogenous levels (Fig. 3g) and 5mC were decreased (Fig. 3a-b), with no 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 αKG/succinate ratio appeared equally low in S3-/- and MitoS3

292 cells relative to S3+/+ cells (Fig. 4a). We then measured h5mC and 5mC and used their ratio
293 as a direct readout of Tets activity, which appeared low both in S3-/- and MitoS3.A/B
294 compared to S3+/+ cells (Figs. 4b and S6a). Such results indicate that in S3-/- and MitoS3
295 cells Tets activity is similarly low, therefore it could hardly explain the differences in 5mC
296 between S3-/- and MitoS3 cells (Fig. 3a-b). Such conclusions are in line with the lack of
297 effect on 5mC observed upon Tet1/2 knockdown (Fig S1h-i).
298
299 Next, we measured mRNA and protein levels of Dnmt3a/b by RNAseq, RT-qPCR, Western
300 Blot and proteomics analysis and found that both Dnmt3a/b mRNA and protein levels were
301 reduced in MitoS3.A and MitoS3.B cells compared to S3-/- cells (Figs. 4c-e and S6b),
302 indicating that αKG could repress Dnmt3a/b expression.
303 To directly test whether αKG negatively regulates Dnmt3a/b expression, we treated S3-/-
304 cells with DM-αKG and observed a partial reduction of Dnmt3a/b expression (Fig. 4f) and of
305 imprinted genes (Fig. 4g). Such partial effects are likely due to the fact that DM-αKG is
306 unable to stably rescue endogenous αKG levels in S3-/- cells (Fig. S6c).
307 This set of experiments led us to conclude that αKG decreases 5mC levels via reduction of
308 Dnmt3a/b expression, rather than by increasing Tets activity. Of note, such conclusions are in
309 agreement with our genetic perturbations showing that Dnmt3a/b levels dictate 5mC
310 abundance (Figs. 1g-i and S1f-g).
311 We next decided to clarify the molecular mechanism of how αKG controls the expression of
312 Dnmt3a/b. First, we have explored literature and we have analyzed a database ("NIA Mouse
313 ES Cell Bank"[42–44]) reporting transcriptomic data of a large number of mouse ES cell lines, in
314 which single transcriptional regulators were either induced or repressed. From this survey,
315 we have identified 2 activators (Otx2, Sox1[45,46]) and 6 repressors (Klf4, Nanog, Prdm14,
316 Tbx3, Tcea3, 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, while the repressors Klf4, Tbx3 and Tcl1 were downregulated in S3-/- cells (Fig. 4h). Given that mitochondrial Stat3 expression increased endogenous αKG levels, and reduced 5mC levels (Fig. 3a-b) and Dnmt3a/b expression (Fig. 4c-d), we measured the expression of Dnmt3a/b potential regulators in MitoS3 cells and observed that only Otx2, Klf4 and Tcl1 expression was significantly affected by mitochondrial Stat3 (Fig. 4h). Such results were confirmed by RNA-seq analysis (Fig. S6d). Finally, we asked whether αKG treatment would affect the expression of our putative Dnmt3a/b regulators, and we observed a significant effect only in the case of Otx2 (Fig 4i).

To test whether Otx2 is functionally required for Dnmt3a/b regulation we first cultured wild-type ES cells in 2i or 2iL. In the absence of LIF, Dnmt3a/b expression was increased together with Otx2 expression (~6 fold increase, Figs. 4j and S6e). This was accompanied by the expected increase in 5mC (Fig. 4k). We reasoned that if Otx2 is in fact crucial for Dnmt3a/b regulation in this context, its genetic inactivation should have rendered cells unable to upregulate Dnmt3a/b in the absence of LIF. We cultured Otx2-/- ES cells[52] in 2i or 2iL and observed that Dnmt3a/b expression, as well as 5mC levels, were not significantly changed. We, therefore, conclude that Otx2 is regulated by the LIF/Stat3/αKG axis and that Otx2 is genetically required to boost Dnmt3a/b and 5mC levels in ES cells.

Mitochondrial Stat3 regulates the differentiation of ES cells

Our results indicate that mitochondrial Stat3 regulates DNA methylation in naive ES cells. We then asked whether such regulation has any functional consequence on ES cell behavior. Mitochondrial Stat3 reduces the levels of Otx2, Dnmt3a/b and 5mC, which are all molecular

markers of early phases of mES cell differentiation[4,23,24,37,53–57]. Therefore we hypothesized that mitochondrial Stat3 might stabilize pluripotency and/or slow down differentiation.

To test this hypothesis we generated transcriptomic data of cells either in 2iLIF or undergoing differentiation for 48 hours in the N2B27 basal medium. We first identified the genes significantly downregulated in S3-/- cells relative to S3+/+ in 2iLIF (Fig. 5a, blue), and observed that they were also downregulated during differentiation of S3+/+ cells (Fig. 5b). Several naive pluripotency markers belong to this category (Fig. 5d), including Esrrb, Klf5 and Tet2. Similarly, genes found upregulated in S3-/- cells relative to S3+/+ in 2iLIF (Fig. 5a, orange) were found upregulated in S3+/+ cells during differentiation (Fig. 5c). Among such genes, we found several early differentiation markers, such as Otx2, Lin28b and Pou3f1, as well as the imprinted genes Igf2, Sfmbt2, Cdkn1c and Phlda2 (Fig. 5e). Notably, S3-/- cells display a much faster reduction in naive markers and upregulation of early differentiation and imprinted genes after 48 hours in N2B27 (Figs. 5d-f and S7a-d). Furthermore, we performed a clonal assay of cells undergoing differentiation for up to 72 hours. In 2iLIF, the number of Alkaline phosphatase positive (AP+) pluripotent colonies formed by S3-/- cells was mildly reduced relative to S3+/+ (Fig. 5g), indicating that clonogenicity of S3-/- cells is partially impaired. After 24 hours of 2iLIF withdrawal, S3-/- cells formed fewer AP+ colonies than 2iL, while S3+/+ cells actually display an increase in AP+ colonies. Only after 48h of 2iLIF withdrawal S3+/+ cells showed a mild reduction in AP+ colonies relative to 2iL, while S3-/- cells at the same time point completely lost the capacity to form AP+ colonies. Eventually, after 72 hours of 2iLIF withdrawal, both cell lines lost clonogenicity.

Based on the faster transcriptional changes and the faster loss AP+ colonies we conclude that S3-/- cells exit more rapidly from the naive pluripotent state.

Such results are consistent with the canonical role of LIF and nuclear Stat3 as a transcriptional inducer of naive factors described by several laboratories[9,11,58,59]. In other words, the effects of Stat3 on differentiation could be completely independent from its capacity to control αKG production and 5mC from the mitochondria. Thus, we asked whether the expression of Stat3 in the mitochondria would affect differentiation of S3-/- cells. Transcriptionally, we observed that genes downregulated in S3-/- cells were only mildly affected by mitochondrial Stat3 (Fig. 5b, d). Conversely, early differentiation marker and imprinted genes were strongly reduced in MitoS3.A/B cells (Fig. 5c, e-f). Similar results were obtained when we analyzed only genes associated with genomic features differentially methylated in S3-/- cells (Fig. S8a-d). Finally, in the clonal assays, mitochondrial Stat3 restored clonogenicity in 2iLIF and delayed the exit from naive pluripotency (Fig. 5g).

We conclude that in the absence of Stat3, genes associated with naive pluripotency are prematurely downregulated, while early differentiation genes are overactivated, leading to partially compromised pluripotency and accelerated exit from the naive state. Expression of mitochondrial Stat3 counteracts such effects, specifically by repressing early differentiation genes, indicating that modulation of mES cell metabolism by mitochondrial Stat3 has functional consequences on naive pluripotent cell behavior.

Stat3 regulates Dnmts and imprinted transcripts in the early mouse blastocysts

Based on our in vitro observations, we decided to test the function of Stat3 in the early mouse embryo. We focused our attention on the early blastocysts at E3.5, because at this stage Stat3 is active[7,8,10] and the genome is hypomethylated[4,5] in naive pluripotent cells of the inner cell mass (ICM). At E3.75, ICM cells are specified into pluripotent Epiblast cells (Epi) or into extraembryonic Primitive Endoderm cells (PrE).

Stat3 heterozygous mice were crossed and blastocyst embryos at E3.5 and E3.75 were flushed. Trophectoderm cells were removed by immunosurgery and used to genotype individual embryos. Single ICM, Epi and PrE cells were analyzed by RNA sequencing (Fig. 6a).

At E3.5, global analysis by t-Distributed Stochastic Neighbor Embedding (t-SNE) revealed that the transcriptomes of S3+/+ and S3-/- ICM cells are divergent (Fig. 6b). Differential expression analysis revealed that Stat3 and its targets Socs3 and Tfcp2l1 were downregulated in S3-/- cells together with Nanog and FGF4 (Fig. 6c-d). Notably, Dnmt3a and Dnmt3b were upregulated in S3-/- ICM cells, as observed in vitro (Fig. 1), together with markers of PrE, such as Sox17, Gata4 and Pdgfra (Fig. 6c-d). Such results indicate that S3-/- ICM cells might precociously activate the PrE expression program. To test this hypothesis we generated a list of genes specifically expressed in PrE at E4.5 and found them upregulated in S3-/- ICM cells compared to S3+/+ ICM cells (Fig. S9a).

We then analyzed cells from E3.75 embryos and individual cells could be classified as Epiblast or PrE according to specific markers (Fig. S9b). We also observed a clear separation between S3+/+ and S3-/- cells (Component 2 in Fig. 6e). In S3-/- embryos, Socs3, Tfcp2l1, Nanog and Tet2 were significantly reduced in Epiblast cells at E3.75 (Fig. 6f-g). Interestingly, markers of post-implantation epiblast, such as Utf1, Otx2[7,60]and Dnmt3a/b were also upregulated in S3-/- cells (Fig. 6g). Gene lists associated with Epiblast at E5.5 and E6.5 were also upregulated specifically in E3.75 Epi S3-/-, while genes associated with E3.5 ICM were downregulated (Fig. S9c left panel), further indicating accelerated developmental progression.

We observed that Stat3 regulates imprinted transcripts in vitro (Fig. 2). We analyzed all imprinted transcripts expressed at E3.5 or 3.75 and observed global deregulation (Fig. 6h).

For instance, we observed anticipated expression of Mest and Sfmbt2 or reduced expression of Rhox5 and Pon2 (Fig. 6i).

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

Discussion

Mouse ES cells cultured in 2iLIF display low levels of 5mC and Dnmt3a/b, similarly to naive pluripotent cells of the pre-implantation blastocyst-stage embryo, while cells in Serum LIF show elevated 5mC and Dnmt3a/b. Previous studies reported that the hypomethylation observed in 2iLIF was due to the presence of the MEK inhibitor, one of the two inhibitors used in 2i. MEK inhibition causes upregulation of Prdm14[17,43,44], which, in turn, represses Dnmt3a/b expression[14,15].

Our results indicate that also mitochondrial Stat3 is necessary for Dnmt3a/b downregulation in 2iLIF, given that Stat3 null cells in 2iLIF displayed high 5mC levels, despite the presence of the MEK inhibitor, and expression of a mitochondrially localized Stat3 construct is sufficient to reduce Dnmt3a/b and 5mC levels. We note that Prdm14 expression was not affected by Stat3 (Fig. 4 and 5), overall indicating that Stat3 and Prdm14 are two independent negative regulators of Dnmt3a/b, and they are both genetically required for hypomethylation in 2iLIF.

Previous studies linked the hypomethylation of naive ES cells in 2iLIF to reduced *de novo* DNA methylation activity[14,15,17,48,49,61], in agreement with our results, showing that genetic inactivation or over-expression of Dnmt3a/b resulted in reduced or increased 5mC levels, respectively. In contrast, the maintenance DNA methyltransferase Dnmt1 shows similar mRNA levels, protein levels and activity both in 2iLIF and Serum LIF[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 loci[62].

We here demonstrated that mitochondrial Stat3-dependent reduction of Dnmt3a/b and 5mC levels is crucial to induce genomic hypomethylation in 2iLIF, with no effects on Uhrf1 protein levels (Fig. 4e), indicating that the two mechanisms are independent.

In the embryo, Otx2 and Dnmt3a/b are expressed robustly only after implantation at E5.5[7], while Stat3 is active only in the pre-implantation blastocyst[7]. We showed that genetic inactivation of Stat3 leads to anticipated expression of Otx2 and Dnmt3a/b, altogether indicating 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 embryos relative to maternal tissues.

Stat3 has been shown to act in different cellular compartments, as the nucleus, mitochondria and the Endoplasmic Reticulum[20,21,64]. We showed that mitochondrial Stat3 is critical for repression of Dnmt3a/b, through the control of αKG levels in the cell. Previous work implicated Stat3 in the control of epigenetic modifications during somatic cell reprogramming[32,65]. It would be interesting to test whether this is due to the metabolic activity of mitochondrial Stat3.

The two inhibitors of Gsk3 and Mek have been involved in metabolic rewiring allowing efficient conversion of αKG into Glutamine in mouse ES cells[34]. Moreover, αKG production from Glutamine via Psat1 has been reported to decrease during ES cell differentiation[66]. 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 interconversion of αKG and

Glutamine, in line with the critical roles of both metabolites in fundamental processes such as proliferation, epigenetic regulations and differentiation[34,38,67].

Of note, S3-/- cells show impaired mitochondrial respiration, a condition associated with enhanced reductive Glutamine metabolism in cancer cells[68], 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.

Stat3 regulates imprinted gene expression, which depends on DNA methylation. In ES cells this regulation is due to mitochondrial Stat3 and αKG availability. Importantly, Stat3 inactivation *in vivo* also results in global dysregulation of imprinted genes. We observed that several imprinted genes (e.g. Ddc, Gab1, Commd1, Cobl, Cd81) have been shown to regulate ES cell differentiation[69], suggesting that a balanced expression of imprinted genes could be critical for correct exit from naive pluripotency.

Long term culture of female murine ES cells in 2iLIF has been associated with decreased methylation at imprinted DMRs[15,16,70,71] and reduction of MEK inhibitor concentration allowed to maintain robust methylation at DMRs over extensive culture. Similarly, we showed that in the absence of Stat3 mES cells maintain high methylation levels at DMRs after over 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 ES cells are characterized by bivalent metabolism, a hypomethylated genome and high expression of specific transcription factors and epigenetic modifiers, such as Tet2. Upon differentiation, OXPHOS is decreased, genome methylation is increased, naive specific genes

493 are downregulated and early markers of differentiation are upregulated, including Dnmt3a/b.
494 We propose a model whereby all these molecular processes are elegantly under the control of
495 a single molecule, Stat3. While nuclear Stat3 directly induces and maintains the expression of
496 naive pluripotency factors, mitochondrial Stat3 promotes OXPHOS and αKG production,
497 genome hypomethylation and inhibition of early differentiation markers. Such model explains
498 previous observations, such as the inability of MitoS3 by itself to maintain long term self-
499 renewal[19] and will be useful to test the role of LIF/Stat3 in pluripotent cells of other species
500 and during induction of pluripotency.

Figure 1

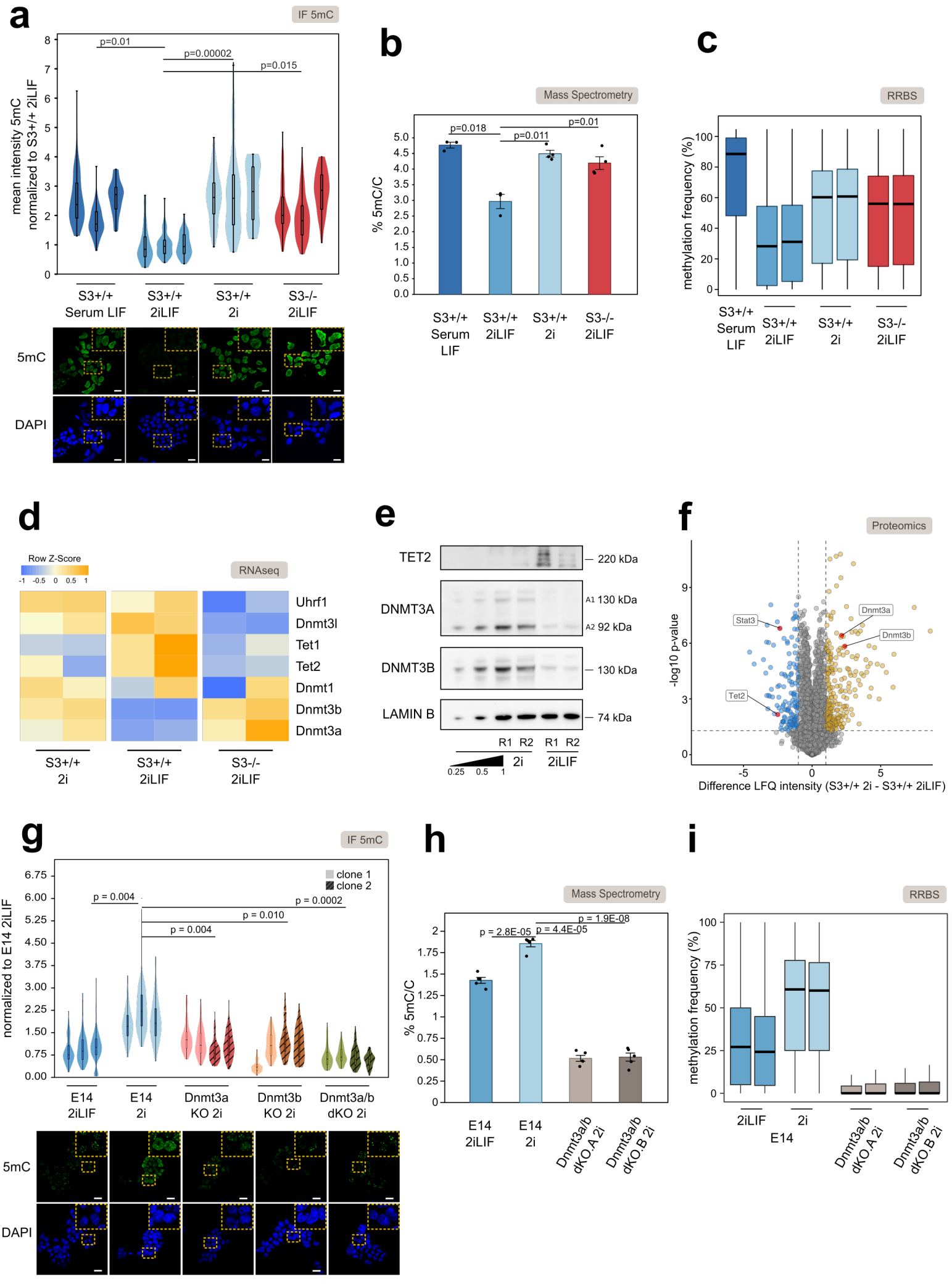
a
IF 5mC
mean intensity 5mC normalized to S3+/+ 2iLIF
p=0.01
p=0.00002
p=0.015
S3+/+ Serum LIF
S3+/+ 2iLIF
S3+/+ 2i
S3-/- 2iLIF
5mC
DAPI
b
Mass Spectrometry
% 5mC/C
p=0.018
p=0.011
p=0.01
S3+/+ Serum LIF
S3+/+ 2iLIF
S3+/+ 2i
S3-/- 2iLIF
c
RRBS
methylation frequency (%)
S3+/+ Serum LIF
S3+/+ 2iLIF
S3+/+ 2i
S3-/- 2iLIF
d
Row Z-Score
RNAseq
Uhrf1
Dnmt3l
Tet1
Tet2
Dnmt1
Dnmt3b
Dnmt3a
S3+/+ 2i
S3+/+ 2iLIF
S3-/- 2iLIF
e
TET2
220 kDa
DNMT3A
A1 130 kDa
A2 92 kDa
DNMT3B
130 kDa
LAMIN B
74 kDa
R1 R2 R1 R2
0.25 0.5 1
2i
2iLIF
f
Proteomics
-log10 p-value
Stat3
Dnmt3a
Dnmt3b
Tet2
Difference LFQ intensity (S3+/+ 2i - S3+/+ 2iLIF)
g
IF 5mC
clone 1
clone 2
mean intensity 5mC normalized to E14 2iLIF
p = 0.004
p = 0.0002
p = 0.010
p = 0.004
E14 2iLIF
E14 2i
Dnmt3a KO 2i
Dnmt3b KO 2i
Dnmt3a/b dKO 2i
5mC
DAPI
h
Mass Spectrometry
% 5mC/C
p = 2.8E-05
p = 4.4E-05
p = 1.9E-08
E14 2iLIF
E14 2i
Dnmt3a/b dKO.A 2i
Dnmt3a/b dKO.B 2i
i
RRBS
methylation frequency (%)
2iLIF
2i
E14
Dnmt3a/b dKO.A 2i
Dnmt3a/b dKO.B 2i

Fig. 1. LIF/Stat3 induced hypomethylation in mES cells via Tet2 and Dnmt3a/b regulation

a, Immunofluorescence on S3+/+ cells cultured in Serum LIF, 2i or 2iLIF and S3-/- cells in 2iLIF stained with anti-5mC antibody (5mC, 5-methylcytosine). Top: Violin plots show the distribution of fluorescence intensity of an average of 63 nuclei per sample, normalized to the mean intensity of S3+/+ 2iLIF. Boxplots show 1st, 2nd and 3rd quartile; Whiskers indicate minimum and maximum values. n = 3 independent experiments are shown as individual violins. Two-tailed unpaired T-test was performed on median intensity values of each sample. Bottom: representative images of the 4 conditions in analysis. Scale bar: 20μm.

b, Liquid chromatography - Mass Spectrometry showing percentages of 5mC in the DNA of S3+/+ cells cultured in Serum LIF, 2i or 2iLIF and S3-/- cells in 2iLIF. 5mC contents are expressed as the percentage of 5mC in the total pool of cytosine. Bars represent means and s.e.m of n = 3 different biological replicates for S3+/+ Serum LIF and S3+/+ 2iLIF and n = 4 different biological replicates for S3+/+ 2i and S3-/- 2iLIF shown as dots. *P* values calculated using two-tailed unpaired T-test.

c, Frequency of DNA methylation at CpG islands measured by RRBS in S3+/+ cells cultured in Serum LIF, 2i or 2iLIF and S3-/- cells in 2iLIF. Each boxplot indicates the 1st, 2nd and 3rd quartile of a biological replicate. Whiskers indicate minimum and maximum value. See also Fig. S1a.

d, Heatmap for 7 genes involved in DNA methylation (Dnmt3a, Dnmt3b, Dnmt1, Dnmt3l, Uhrf1) and methylcytosine oxidation (Tet1, Tet2). RNAseq data derived from S3+/+ and S3-/- cells, expanded in 2i media without LIF or treated with LIF; n = 2 biological replicates are reported for each condition. Expression levels were scaled and represented as z-score. Yellow and blue indicate high and low expression, respectively. See also Fig. S1b.

e, Western blot of S3+/+ cells cultured in 2i or 2iLIF (left). Two biological replicates were loaded for each condition, indicated as R1 and R2. *De novo* methyltransferases DNMT3A(92 and 130kDa) and DNMT3B(130kDa) were less abundant in cells treated with LIF, while TET2 was increased. As previously reported[17], DNMT3A/B were upregulated, while TET2

(220kDa) was downregulated. LAMIN B (74kDa) was used as a loading control. Two isoforms of DNMT3A (92 and 130kDa) were detected, as previously reported[72].

f, Volcano plot of proteomics data, depicting differences in protein abundances between S3+/+ 2i and S3+/+ 2iLIF cells. Each dot represents one protein. The x axis shows the fold change (FC) in protein abundance (measured as LFQ intensity, in log-scale) and the y axis represents the statistical significance (adjusted *P* value, in log-scale). Yellow and blue dots indicate respectively proteins that are less or more abundant (FC > 1 or FC < -1 respectively, *P* value < 0.05) in S3+/+ 2i cells with respect to S3+/+ 2iLIF cells. Similar effects were observed also in S3-/- cells (Fig. S1c). All proteomics data are in Supplementary table 4.

g, Immunofluorescence for 5mC on wild-type mES cell line E14IVc (E14) cultured in 2iLIF and 2i, and on Dnmt3a KO, Dnmt3b KO, and Dnmt3a/b double KO mES cells cultured in 2i without LIF. Two independent mutant clones for each genotype were validated by Western blot (Fig. S1d) and analysed. Top: Violin plots show the distribution of fluorescence intensity of an average of 82 nuclei per sample, normalized to the mean intensity of E14 2iLIF. The independent mutant clones are marked with dark stripe pattern and light colors, respectively. Boxplots show 1st, 2nd and 3rd quartile; Whiskers indicate minimum and maximum values. Independent experiments are shown as individual violins. Two-tailed unpaired T-test was performed on median intensity values of each sample. Bottom: representative images of the 5 conditions in analysis. Scale bar: 20μm.

h, Liquid chromatography - Mass Spectrometry showing percentages of 5mC in the DNA of E14 cells cultured in 2iLIF and 2i, and Dnmt3a/b double KO mES cells in 2i (two independent mutant clones, Dnmt3a/b dKO.A and Dnmt3a/b dKO.B). 5mC contents are expressed as the percentage of 5mC in the total pool of cytosine. Bars represent means and s.e.m. of n = 5 different biological replicates shown as dots. *P* values calculated using two-tailed unpaired T-test.

i, Frequency of DNA methylation at CpG islands measured by RRBS in E14 cells cultured in 2iLIF and 2i, and two clones of Dnmt3a/b double KO mES cells in 2i. Two biological replicates for each sample. Each boxplot indicates the 1st, 2nd and 3rd quartiles. Whiskers indicate minimum and maximum value.

Figure 2

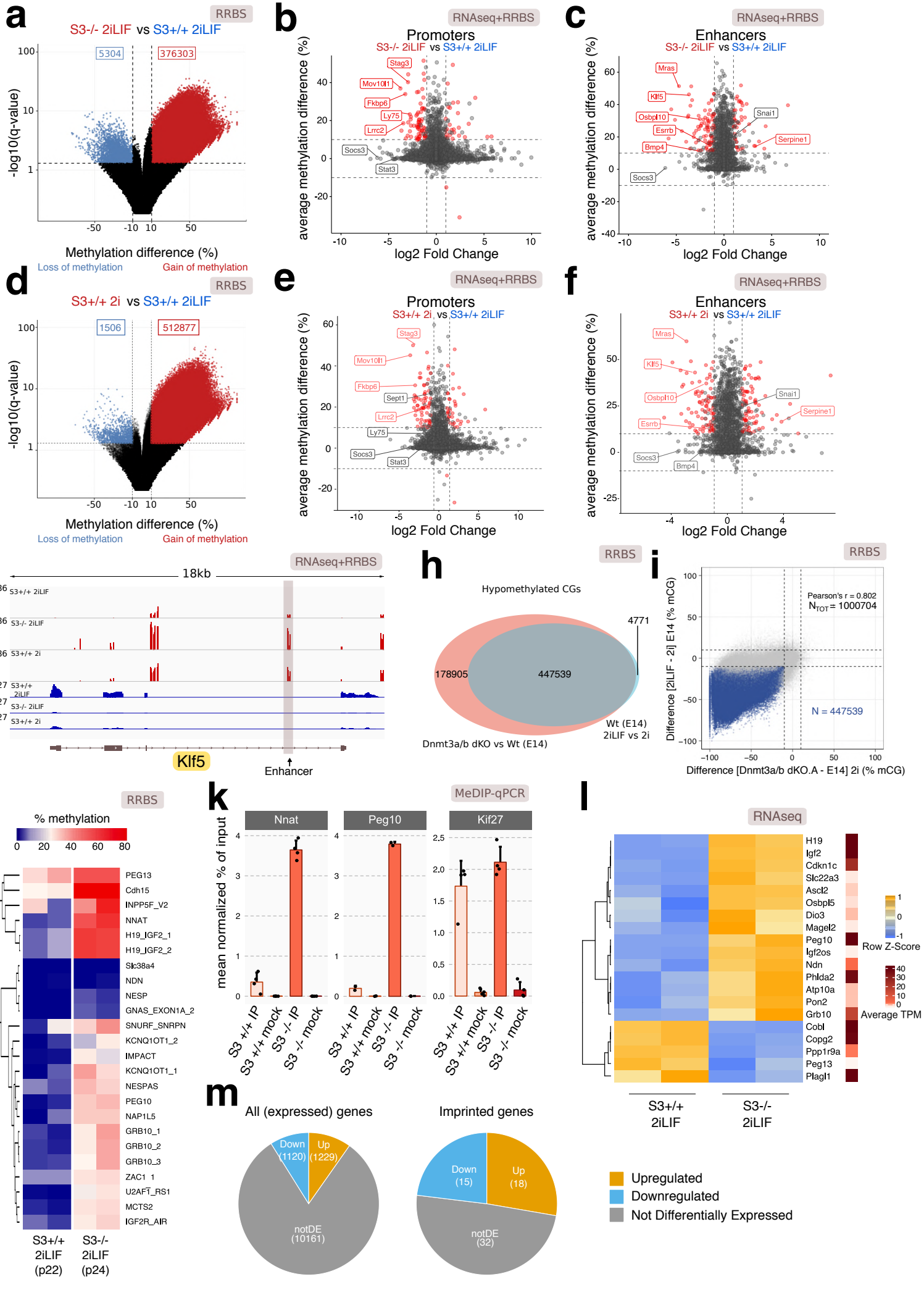
a
RRBS
S3-/- 2iLIF vs S3+/+ 2iLIF
5304
376303
-log10(q-value)
Methylation difference (%)
Loss of methylation
Gain of methylation
b
RNAseq+RRBS
Promoters
S3-/- 2iLIF vs S3+/+ 2iLIF
average methylation difference (%)
log2 Fold Change
Stag3
Mov10l1
Fkbp6
Ly75
Lrrc2
Socs3
Stat3
c
RNAseq+RRBS
Enhancers
S3-/- 2iLIF vs S3+/+ 2iLIF
Mras
Klf5
Osbpl10
Esrrb
Bmp4
Snai1
Serpine1
Socs3
d
RRBS
S3+/+ 2i vs S3+/+ 2iLIF
1506
512877
e
RNAseq+RRBS
Promoters
S3+/+ 2i vs S3+/+ 2iLIF
Sept1
f
RNAseq+RRBS
Enhancers
S3+/+ 2i vs S3+/+ 2iLIF
g
RNAseq+RRBS
18kb
mapped reads
S3+/+ 2iLIF
S3-/- 2iLIF
S3+/+ 2i
Klf5
Enhancer
h
RRBS
Hypomethylated CGs
178905
447539
4771
Dnmt3a/b dKO vs Wt (E14)
Wt (E14) 2iLIF vs 2i
i
RRBS
Pearson's r = 0.802
N_TOT = 1000704
N = 447539
Difference [2iLIF - 2i] E14 (% mCG)
Difference [Dnmt3a/b dKO.A - E14] 2i (% mCG)
j
RRBS
% methylation
PEG13
Cdh15
INPP5F_V2
NNAT
H19_IGF2_1
H19_IGF2_2
Slc38a4
NDN
NESP
GNAS_EXON1A_2
SNURF_SNRPN
KCNQ1OT1_2
IMPACT
KCNQ1OT1_1
NESPAS
PEG10
NAP1L5
GRB10_1
GRB10_2
GRB10_3
ZAC1_1
U2AF1_RS1
MCTS2
IGF2R_AIR
S3+/+ 2iLIF (p22)
S3-/- 2iLIF (p24)
k
MeDIP-qPCR
Nnat
Peg10
Kif27
mean normalized % of input
S3 +/+ IP
S3 +/+ mock
S3 -/- IP
S3 -/- mock
l
RNAseq
H19
Igf2
Cdkn1c
Slc22a3
Ascl2
Osbpl5
Dio3
Magel2
Peg10
Igf2os
Ndn
Phlda2
Atp10a
Pon2
Grb10
Cobl
Copg2
Ppp1r9a
Peg13
Plagl1
Row Z-Score
Average TPM
S3+/+ 2iLIF
S3-/- 2iLIF
m
All (expressed) genes
Down (1120)
Up (1229)
notDE (10161)
Imprinted genes
Down (15)
Up (18)
notDE (32)
Upregulated
Downregulated
Not Differentially Expressed

Fig. 2. Impact of Stat3 on DNA methylation and transcription

a, Volcano plot showing the significant differentially methylated CpG sites between S3-/- and S3+/+ cells (q-value < 0.01, abs(average methylation difference) > 10%). The x and y axis represent the difference in methylation levels (reported as percentage) and the statistical significance (in log-scale) respectively. Red dots depict hyper-methylated sites in S3-/- cells, blue dots the hypo-methylated ones. The analysis confirmed the overall gain in DNA methylation levels in S3-/- samples, with 376,303 hyper-methylated CpG sites and only 5,304 hypo-methylated sites out of a total of 1,230,955 sites mapped.

b, Scatter plot showing the mutual changes in gene expression (RNAseq) and DNA methylation levels (RRBS) at active promoters between S3+/+ and S3-/- cells. For each gene, the x axis reports the fold change in gene expression and the y axis the average methylation difference; red dots indicate genes for which both changes reached statistical significance (adjusted *P* value < 0.01 for gene expression, q-value < 0.01 for methylation difference). See also Supplementary table 2.

c, Scatter plot showing the mutual changes in gene expression and DNA methylation levels at active enhancers between S3+/+ and S3-/- cells. For each gene, the x axis reports the fold change in gene expression and the y axis the average methylation difference; red dots indicate genes for which both changes reached statistical significance (adjusted *P* value < 0.01 for gene expression, q-value < 0.01 for methylation difference). See also Supplementary table 2.

d, Volcano plot showing the significant differentially methylated CpG sites between S3+/+ 2i and S3+/+ 2iLIF cells, as described in panel a. Similarly to S3-/- cells, the analysis confirmed the overall gain in DNA methylation levels also in S3+/+ 2i samples, with 512,877 hyper-methylated CpG sites and only 1,506 hypo-methylated sites out of a total of 1,327,475 sites mapped. Similar effects of LIF were observed also in an independent wild-type mES cell line (E14) (S2c).

e, Scatter plot showing the mutual changes in gene expression (RNAseq) and DNA methylation levels (RRBS) at active promoters between S3+/+ 2i and S3+/+ 2iLIF cells, as described in panel b.

f, Scatter plot showing the mutual changes in gene expression and DNA methylation levels at active enhancers between S3+/+ 2i and S3+/+ 2iLIF cells, as described in panel c.

g, Gene tracks showing RRBS data (in red) and RNAseq data (in blue) for S3+/+ 2iLIF, S3+/+ 2i and S3-/- 2iLIF cells over a representative genomic region. One representative biological replicate of two is reported for RNAseq and RRBS data.

h, Venn diagram showing number of CpG sites whose methylation status is dependent on LIF (light blue circle) or on Dnmt3a/b (red circle). Gray intersection contains the number of CpG sites that lose DNA methylation both when LIF is added to culture medium or when Dnmt3a/b are genetically deleted. Similar results were obtained in an independent Dnmt3a/b dKO clone (Fig. S2e).

i, Scatter plot comparing the effects of adding LIF (y axis) and deleting Dnmt3a/b (x axis) on the levels of CpG methylation in wild-type E14 cells. For each CpG site, the x axis reports the difference in methylation levels (% mCG) between Dnmt3a/b dKO.A and E14 cultivated in 2i, and the y axis the difference in methylation levels between E14 2iLIF and E14 2i. Grey dots indicate all CpG sites that were commonly covered in at least one technical replicate of each sample with a minimum sequencing depth of 10x; blue dots indicate sites that are hypo-methylated both in the presence of LIF or when Dnmt3a/b are genetically deleted (q-value < 0.01, ΔmCG < -10 %). Similar results were obtained in an independent Dnmt3a/b dKO clone (Fig. S2f).

j, Heatmap reporting percentage of DNA methylation at imprinted Differentially Methylated Regions (DMRs) in S3+/+ and S3-/- cultured in 2iLIF for 22 and 24 passages, respectively. Out of 24 analysed regions, 20 were found significantly more methylated in S3-/- cells compared to S3+/+ cells. n = 2 biological replicates for each sample. See also Supplementary table 2.

k, MeDIP-qPCR analysis of imprinted DMRs in S3+/+ and S3-/- cells. Real-time qPCR was carried out on enriched methylated DNA fractions immunoprecipitated ("IP") from genomic DNA with an antibody specific for 5mC (MeDIP). Negative controls were included, where a non-specific antibody (IgG) was used for immunoprecipitation ("mock"). Enrichment was calculated as % of Input and mean normalised for each experiment. Imprinted DMRs like

Nnat and Peg10 show enrichment in immunoprecipitated fractions, thus higher methylation levels, in S3-/- cells compared to S3+/+ cells; Kif27 was found unchanged in RRBS data, therefore it was included as a control. Bars represent the mean and SD of n = 4 (Nnat, Kif27) or n = 2 (Peg10) independent MeDIP experiments, shown as dots. See also Fig. S2h.

l, Heatmap showing imprinted genes associated with known DMRs (Fig. 2e) and differentially expressed between S3-/- cells and S3+/+ cells in 2iLIF. Expression values were scaled and represented as z-score; absolute expression for each gene is indicated on the right as average TPM values of 2 independent replicates. Yellow and blue indicate high and low expression, respectively. n = 2 biological replicates for each sample. See also Fig. S2i-j.

m, Pie charts showing the number of up- and down-regulated genes (adjusted *P* value < 0.01, abs(log2 FC) > 1) in S3-/- cells with respect to S3+/+ cells among all expressed genes (left), or among all expressed imprinted genes (right).

Figure 3

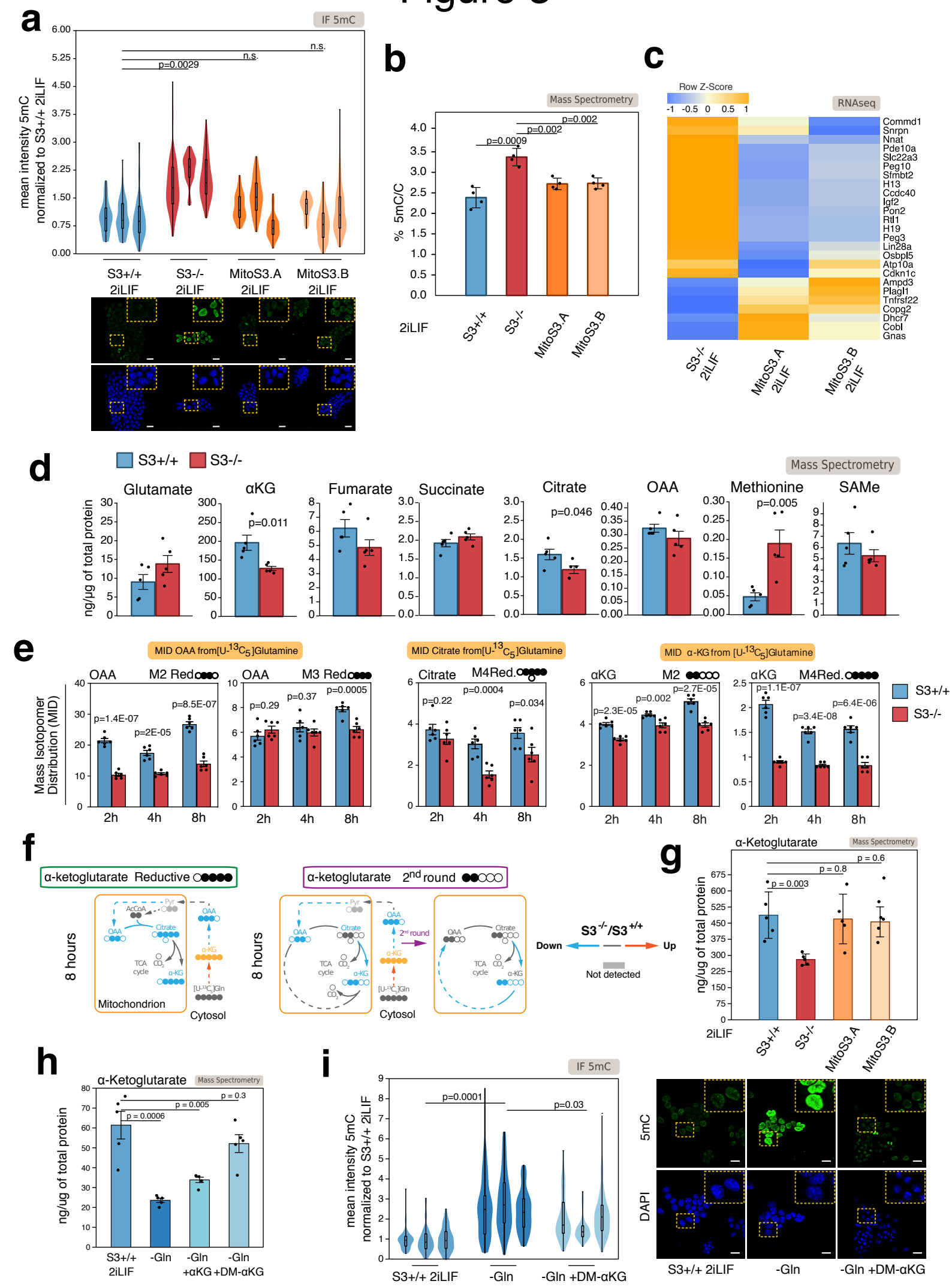
a
IF 5mC
mean intensity 5mC normalized to S3+/+ 2iLIF
p=0.0029
n.s.
n.s.
S3+/+ 2iLIF
S3-/- 2iLIF
MitoS3.A 2iLIF
MitoS3.B 2iLIF
b
Mass Spectrometry
% 5mC/C
p=0.0009
p=0.002
p=0.002
2iLIF
S3+/+
S3-/-
MitoS3.A
MitoS3.B
c
Row Z-Score
RNAseq
Commd1
Snrpn
Nnat
Pde10a
Slc22a3
Peg10
Sfmbt2
H13
Ccdc40
Igf2
Pon2
Rtl1
H19
Peg3
Lin28a
Osbpl5
Atp10a
Cdkn1c
Ampd3
Plagl1
Tnfrsf22
Copg2
Dhcr7
Cobl
Gnas
S3-/- 2iLIF
MitoS3.A 2iLIF
MitoS3.B 2iLIF
d
S3+/+
S3-/-
Mass Spectrometry
Glutamate
αKG
Fumarate
Succinate
Citrate
OAA
Methionine
SAMe
ng/µg of total protein
p=0.011
p=0.046
p=0.005
e
MID OAA from[U-13C5]Glutamine
MID Citrate from[U-13C5]Glutamine
MID α-KG from [U-13C5]Glutamine
Mass Isotopomer Distribution (MID)
OAA M2 Red
OAA M3 Red
Citrate M4Red.
αKG M2
αKG M4Red.
p=1.4E-07
p=2E-05
p=8.5E-07
p=0.29
p=0.37
p=0.0005
p=0.22
p=0.0004
p=0.034
p=2.3E-05
p=0.002
p=2.7E-05
p=1.1E-07
p=3.4E-08
p=6.4E-06
2h
4h
8h
S3+/+
S3-/-
f
α-ketoglutarate Reductive
α-ketoglutarate 2nd round
8 hours
8 hours
Mitochondrion
Cytosol
Cytosol
2nd round
S3-/-/S3+/+
Down
Up
Not detected
g
α-Ketoglutarate
Mass Spectrometry
ng/ug of total protein
p = 0.6
p = 0.8
p = 0.003
2iLIF
S3+/+
S3-/-
MitoS3.A
MitoS3.B
h
α-Ketoglutarate
Mass Spectrometry
ng/ug of total protein
p = 0.3
p = 0.005
p = 0.0006
S3+/+ 2iLIF
-Gln
-Gln +αKG
-Gln +DM-αKG
i
IF 5mC
mean intensity 5mC normalized to S3+/+ 2iLIF
p=0.0001
p=0.03
S3+/+ 2iLIF
-Gln
-Gln +DM-αKG
5mC
DAPI
S3+/+ 2iLIF
-Gln
-Gln +DM-αKG

Fig. 3. Stat3 controls DNA methylation via metabolic regulation

a, 5mC immunofluorescence staining on S3+/+, S3-/- cells and two independent clones, named MitoS3.A and MitoS3.B, where Stat3 is present only in mitochondria (Fig. S4d-h) all cultured in 2iLIF,. Violin plots show the distribution of fluorescence intensity of an average of 55 nuclei per sample, normalized to the mean intensity of S3+/+ 2iLIF. Boxplots show 1st, 2nd and 3rd quartile; n = 3 independent experiments are shown as individual violins. Whiskers indicate minimum and maximum values. Two-tailed unpaired T-test was performed on median intensity values of each sample. Bottom: representative images of the conditions in analysis. Scale bar: 20μm.

b, Mass Spectrometry showing percentages of 5mC in the DNA of S3+/+, S3-/- cells and two MitoS3.A/B clones, cultured in 2iLIF. 5mC content is expressed as the percentage of 5mC in the total pool of cytosines. Bars indicate mean +/- s.e.m. of n = 4 biological replicates, shown as dots. *P* values calculated using two-tailed unpaired T-test.

c, Heatmap reporting differentially expressed imprinted transcripts between S3-/- cells and MitoS3.A/B clones cultured in 2iLIF. Expression values were scaled and represented as z-score. Mean of n = 2 biological replicates is represented for each sample. The genes shown are significantly differentially expressed (adjusted *P* value < 0.1) in both MitoS3 clones relative to S3-/- cells.

d, Barplot showing quantification of individual metabolite abundance measured by Mass Spectrometry; bars indicate mean +/- s.e.m. of n = 5 biological replicates, shown as dots. *P* values were calculated using two-tailed unpaired T-test and are shown only when <0.05.

e, Metabolic tracing analysis of different isotopomers of TCA cycle intermediates (OAA, Citrate and αKG) using [U-$^{13}C_5$]-Glutamine. Barplot represents mass isotopomer distribution (MID, the percentage of labelled isotopomer) at 3 different time points (2h, 4h, 8h). Black circles upper bars represent ^{13}C-labeled carbons. Each bar represents mean +/- s.e.m of n = 6 biological replicates. *P* values were calculated using two-tailed unpaired T-test and are shown only when <0.05.

f, Diagram representing mass isotopomer distribution (MID) of OAA, Citrate and αKG in both oxidative and reductive Glutamine pathways; MID was analysed following 8h of metabolic tracing with [U-$^{13}C_5$]-Glutamine. Orange box indicates the mitochondrion. Color scale outlines the comparison between MID profile in S3-/- cells with respect to S3+/+ for n = 6 biological replicates, where blue color indicates isotopomers (or biochemical pathways) under-represented is S3-/- cells and red color isotopomers or pathways over-represented in S3-/- cells with respect to S3+/+ cells. Full circles represent ^{13}C-labeled carbons. Each isotopomer is corrected for natural isotope abundances.

g, Barplot showing quantification of αKG abundance measured by Mass Spectrometry; bars indicate mean +/- s.e.m. of n = 5 biological replicates, shown as dots. *P* values calculated using two-tailed unpaired T-test.

h, Barplot showing quantification by Mass Spectrometry of αKG abundance in S3+/+ 2iLIF cells cultured with Glutamine, without Glutamine for 9 days or without Glutamine but supplemented with 2mM Alpha-Ketoglutarate (αKG 2mM) or the cell-permeable dimethyl 2-oxoglutarate (DM-αKG 2mM) for 9 days; bars indicate mean +/- s.e.m. of n = 5 biological replicates, shown as dots. Note that absence of Glutamine determines a significant reduction in the levels of αKG, while addition of cell-permeable DM-αKG rescues αKG abundance to levels comparable to S3+/+ 2iLIF with Glutamine. *P* values calculated using two-tailed unpaired T-test.

i, Immunofluorescence staining of 5mC in S3+/+ cells cultured with Glutamine, without Glutamine for 9 days or without Glutamine but supplemented with cell-permeable dimethyl 2-oxoglutarate (DM-αKG 2mM) for 9 days. Left: Violin plots show the distribution of fluorescence intensity of an average of 96 nuclei per sample, normalized to mean intensity of S3+/+ 2iLIF. Boxplots show 1st, 2nd and 3rd quartile; Whiskers indicate minimum and maximum values. n = 3 independent experiments are shown as individual violins. Two-tailed unpaired T-test was performed on median intensity values of each sample. Right: representative images of the 3 conditions in analysis. Scale bar: 20μm.

Figure 4

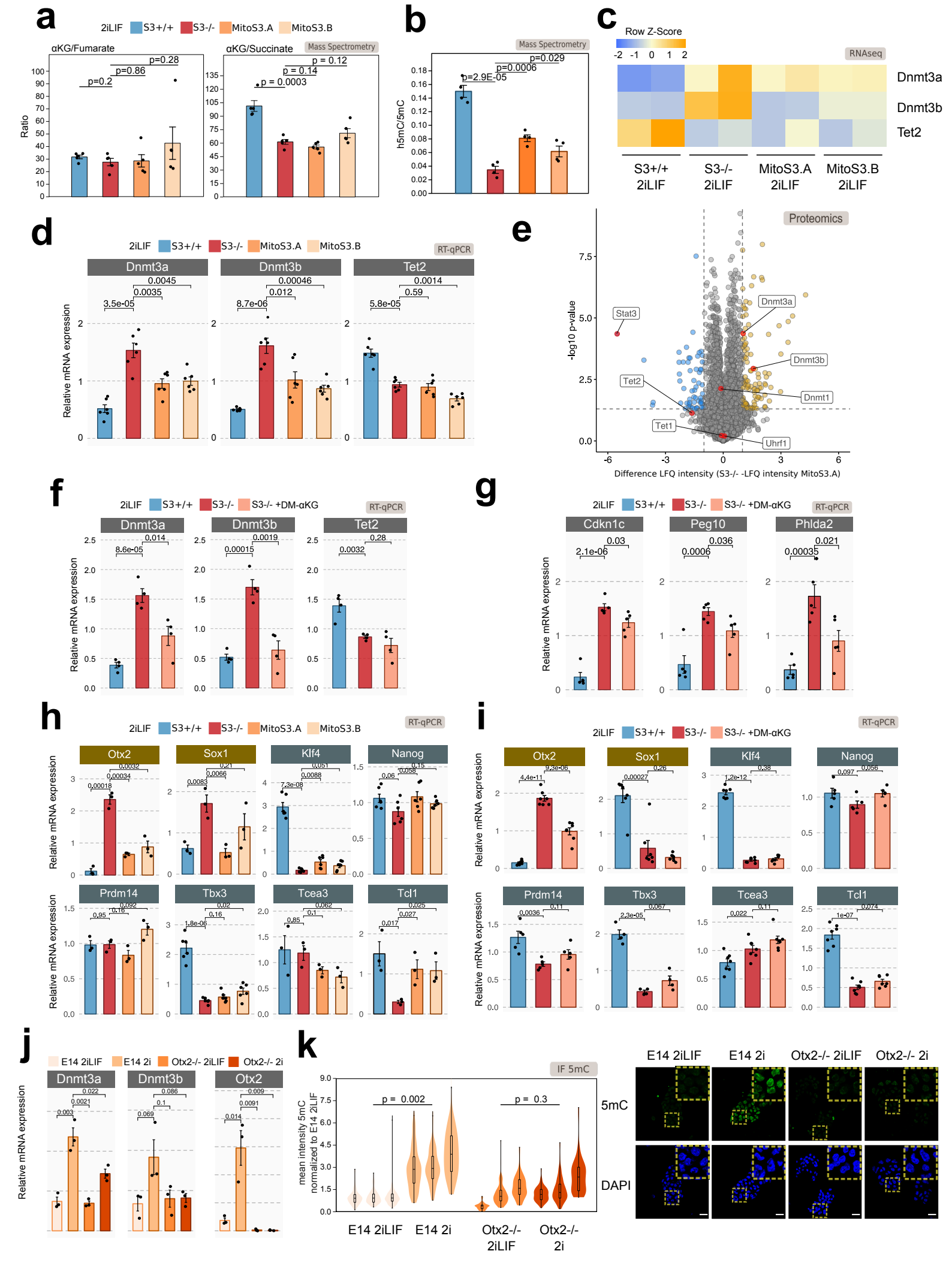
a
2iLIF S3+/+ S3-/- MitoS3.A MitoS3.B
αKG/Fumarate
αKG/Succinate
Mass Spectrometry
Ratio
p=0.2
p=0.86
p=0.28
p = 0.0003
p = 0.14
p = 0.12
b
Mass Spectrometry
h5mC/5mC
p=2.9E-05
p=0.0006
p=0.029
c
Row Z-Score
RNAseq
Dnmt3a
Dnmt3b
Tet2
S3+/+ 2iLIF
S3-/- 2iLIF
MitoS3.A 2iLIF
MitoS3.B 2iLIF
d
2iLIF S3+/+ S3-/- MitoS3.A MitoS3.B
RT-qPCR
Dnmt3a
Dnmt3b
Tet2
Relative mRNA expression
3.5e-05
0.0035
0.0045
8.7e-06
0.012
0.00046
5.8e-05
0.59
0.0014
e
Proteomics
-log10 p-value
Difference LFQ intensity (S3-/- -LFQ intensity MitoS3.A)
Stat3
Dnmt3a
Dnmt3b
Tet2
Dnmt1
Tet1
Uhrf1
f
2iLIF S3+/+ S3-/- S3-/- +DM-αKG
RT-qPCR
Dnmt3a
Dnmt3b
Tet2
Relative mRNA expression
8.6e-05
0.014
0.00015
0.0019
0.0032
0.28
g
2iLIF S3+/+ S3-/- S3-/- +DM-αKG
RT-qPCR
Cdkn1c
Peg10
Phlda2
Relative mRNA expression
2.1e-06
0.03
0.0006
0.036
0.00035
0.021
h
2iLIF S3+/+ S3-/- MitoS3.A MitoS3.B
RT-qPCR
Otx2
Sox1
Klf4
Nanog
Prdm14
Tbx3
Tcea3
Tcl1
Relative mRNA expression
i
2iLIF S3+/+ S3-/- S3-/- +DM-αKG
RT-qPCR
Otx2
Sox1
Klf4
Nanog
Prdm14
Tbx3
Tcea3
Tcl1
Relative mRNA expression
j
E14 2iLIF E14 2i Otx2-/- 2iLIF Otx2-/- 2i
Dnmt3a
Dnmt3b
Otx2
Relative mRNA expression
k
IF 5mC
mean intensity 5mC normalized to E14 2iLIF
p = 0.002
p = 0.3
E14 2iLIF
E14 2i
Otx2-/- 2iLIF
Otx2-/- 2i
E14 2iLIF
E14 2i
Otx2-/- 2iLIF
Otx2-/- 2i
5mC
DAPI

Fig. 4. Alpha-Ketoglutarate regulates 5mC mainly via control of Dnmt3a/b levels

a, Barplot showings the αKG/Fumarate and αKG/Succinate ratios measured by Mass Spectrometry, in S3+/+, S3-/- cells and two MitoS3.A/B clones, cultured in 2iLIF; bars indicate mean +/- s.e.m. of n = 5 biological replicates, shown as dots. *P* values calculated using two-tailed unpaired T-test.

b, Barplot showing h5mC/5mC ratio measured by Mass Spectrometry, in the DNA of S3+/+, S3-/- cells and two MitoS3.A/B clones, cultured in 2iLIF; bars indicate mean +/- s.e.m. of n = 5 biological replicates, shown as dots. *P* values calculated using two-tailed unpaired T-test.

c-d, Mitochondrial Stat3 modulates the expression of the *de novo* DNA methyltransferases. Expression analysis of Dnmt3a, Dnmt3b and Tet2 in S3+/+ and S3-/- cells and two MitoS3.A/B clones in 2iLIF by RNAseq (**c**) and RT-qPCR (**d**).
Heatmap (**c**) reports RNAseq expression values scaled and represented as z-score. Yellow and blue indicate high and low expression, respectively. n = 2 biological replicates for each sample. Bars (**d**) indicate mean +/- s.e.m. of n = 6 different experiments, shown as dots. Expression values were mean-normalised. Beta-actin serves as an internal control. *P* values calculated using two-tailed unpaired T-test relative to S3-/-.

e, Volcano plot of proteomics data, depicting differences in protein abundances between S3-/- and MitoS3.A cells cultivated in 2iLIF. Each dot represents one protein. The x axis shows the fold change (FC) in protein abundance (measured as LFQ intensity, in log-scale) and the y axis represents the statistical significance (adjusted *P* value, in log-scale). Yellow and blue dots indicate respectively proteins that are less or more (FC > 1 or FC < -1 respectively, adjusted *P* value < 0.05) in S3-/- cells with respect to MitoS3.A cells.

f-g, Gene expression analysis by RT–qPCR of epigenetic modifiers (**f**) and imprinted genes (**g**) in S3+/+, S3-/- and S3-/- cells cultured in 2iLIF and treated with 2mM DM-αKG for 4 passages. Beta-actin serves as an internal control. Expression values were mean-normalised. Bars indicate mean and s.e.m. of n = 4 independent experiments, shown as dots. *P* values calculated using two-tailed unpaired T-test relative to S3-/-.

h, Gene expression analysis by RT–qPCR of S3+/+ (blue), S3-/- (red) and two MitoS3.A/B (orange) clones cultured with 2iLIF. Data show expression of genes identified as activators (Gold) and repressors (Grey) of Dnmt3a/b. Data are normalized to the mean of n ≥ 3 independent experiments. 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.

i, Gene expression analysis by RT–qPCR of S3+/+, S3-/- and S3-/- cells cultured in 2iLIF and treated with 2mM DM-αKG for 3 passages. Data show expression of genes identified as activators (Gold) and repressors (Grey) of Dnmt3a/b. Beta-actin serves as an internal control. Expression values were mean-normalised. Bars indicate mean and s.e.m. of n ≥ 4 independent experiments, shown as dots. *P* values calculated using two-tailed unpaired T-test relative to S3-/-.

j, Gene expression analysis by RT–qPCR of E14 and Otx2-/- cells stably cultured in 2iLIF or 2i. Data show expressions of Dnmt3a, Dnmt3b and Otx2. Beta-actin serves as an internal control. Expression values were mean-normalised. Bars indicate mean and s.e.m. of n = 3 independent experiments, shown as dots. *P* values calculated using two-tailed unpaired T-test relative to E14 2i cells.

k, Immunofluorescence on E14 and Otx2-/- cells stably cultured in 2iLIF or 2i stained with anti-5mC antibody. Left: Violin plots show the distribution of fluorescence intensity of an average of 111 nuclei per sample, normalized to the mean intensity of E14 2iLIF. Boxplots show 1st, 2nd and 3rd quartile; Whiskers indicate minimum and maximum values. n = 3 independent experiments are shown as individual violins. Two-tailed unpaired T-test was performed on median intensity values of each sample. Right: representative images of the 4 conditions in analysis. Scale bar: 20μm.

Figure 5

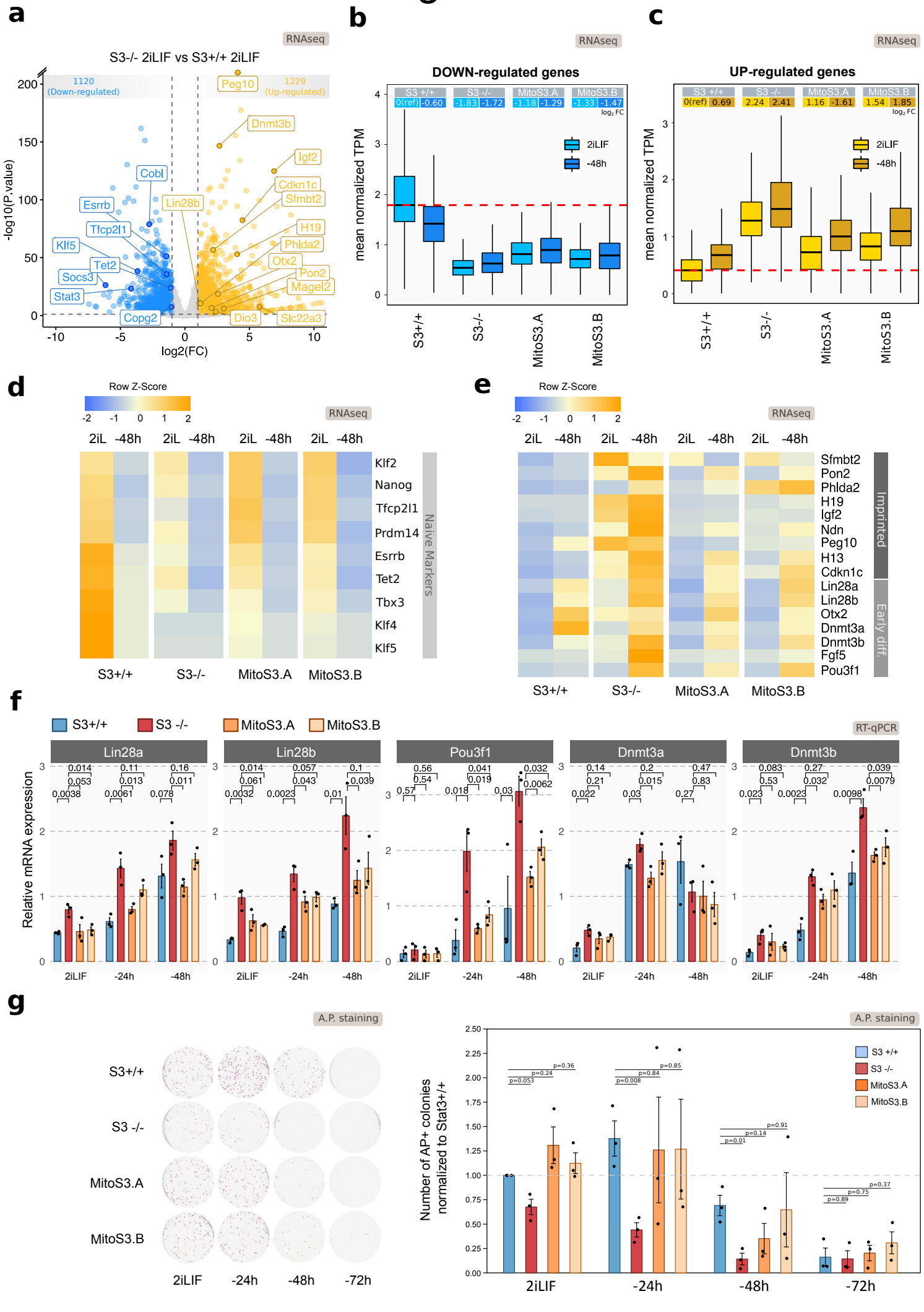
a
RNAseq
S3-/- 2iLIF vs S3+/+ 2iLIF
1120
(Down-regulated)
1229
(Up-regulated)
-log10(P.value)
log2(FC)
Peg10
Dnmt3b
Igf2
Cdkn1c
Sfmbt2
H19
Phlda2
Otx2
Pon2
Magel2
Dio3
Slc22a3
Cobl
Esrrb
Tfcp2l1
Klf5
Tet2
Socs3
Stat3
Copg2
Lin28b
b
RNAseq
DOWN-regulated genes
mean normalized TPM
S3 +/+
S3 -/-
MitoS3.A
MitoS3.B
0(ref) -0.60
-1.83 -1.72
-1.18 -1.29
-1.33 -1.47
log2 FC
2iLIF
-48h
c
RNAseq
UP-regulated genes
mean normalized TPM
0(ref) 0.69
2.24 2.41
1.16 -1.61
1.54 1.85
d
Row Z-Score
RNAseq
2iL -48h
Klf2
Nanog
Tfcp2l1
Prdm14
Esrrb
Tet2
Tbx3
Klf4
Klf5
Naive Markers
e
Row Z-Score
RNAseq
Sfmbt2
Pon2
Phlda2
H19
Igf2
Ndn
Peg10
H13
Cdkn1c
Lin28a
Lin28b
Otx2
Dnmt3a
Dnmt3b
Fgf5
Pou3f1
Imprinted
Early diff.
f
RT-qPCR
Lin28a
Lin28b
Pou3f1
Dnmt3a
Dnmt3b
Relative mRNA expression
2iLIF
-24h
-48h
g
A.P. staining
2iLIF
-24h
-48h
-72h
Number of AP+ colonies normalized to Stat3+/+

Fig. 5. Mitochondrial Stat3 regulates differentiation of ES cells.

a, Volcano plot showing differentially expressed genes between S3+/+ and S3-/- cells. Each dot represents one gene. The x axis shows the fold change (FC) in expression levels (in log-scale) and the y axis represents the statistical significance (adjusted *P* value, in log-scale). Yellow and blue dots indicate respectively transcripts that are up-regulated or down-regulated (log2 FC > 1 or log2 FC < -1 respectively, adjusted *P* value < 0.01) in S3-/- cells with respect to S3+/+ cells.

b, Boxplot reporting expression levels of down-regulated genes in S3-/- cells with respect to S3+/+ cells (Fig. 5a, blue dots). Each boxplot shows 1st, 2nd and 3rd quartile. Whiskers show minimum and maximum values. Y axis represents mean-normalized TPM (Transcripts Per Million) for S3+/+, S3-/- and Mito-Stat3 clones (MitoS3.A and MitoS3.B) in two different conditions: following stable culturing of cells in 2iLIF (light blue) and after 48h of 2iLIF withdrawal from culture medium ("-48h", dark blue). Upper table shows mean log2 FC in gene expression for all analysed conditions, each compared to S3+/+ 2iLIF.

c, Boxplot reporting expression levels of up-regulated genes in S3-/- cells with respect to S3+/+ cells (Fig. 5a, yellow dots). Y axis represents mean-normalized TPM for S3+/+, S3-/- and Mito-Stat3 clones (MitoS3.A and MitoS3.B) in two different conditions: following stable culturing of cells in 2iLIF (light yellow) and after 48h of 2iLIF withdrawal from culture medium ("-48h", dark yellow). Clones expressing only mitochondrial Stat3 (MitoS3.A and MitoS3.B) show a substantial rescue of gene expression levels compared to S3-/- cells, both in 2iLIF and after 48h of LIF withdrawal. Upper table shows mean log2 FC in gene expression for all analysed conditions, each compared to S3 +/+ 2iLIF.

d, Heatmap for 8 pluripotency-associated markers in mES cells. RNAseq data derived from S3+/+, S3-/-, MitoS3.A and MitoS3.B expanded in 2iLIF or without 2iLIF for 48h. Expression levels were scaled and represented as z-score. Yellow and blue indicate high and low expression, respectively. Note that S3-/- cells display reduced levels of naive markers compared to S3+/+ cells. n = 2 biological replicates for each sample.

e, Heatmap for 10 imprinted genes (upper part) and 6 early differentiation markers (lower part) in mouse ES cells. RNAseq data derived from S3+/+ and S3-/- cells and two

independent clones expressing Stat3 localized only in the mitochondria (MitoS3.A and MitoS3.B) expanded in 2iLIF or without 2iLIF for 48h. Expression values were scaled and represented as z-score. Yellow and blue indicate high and low expression, respectively. Note that S3-/- cells display much faster upregulation of early differentiation markers and imprinted genes after 48 hours in N2B27; moreover, mitochondrial Stat3 clones show a rescue in the expression of these transcripts compared to S3-/- cells. n = 2 biological replicates for each sample.

f, Gene expression analysis by RT–qPCR of S3+/+ (blue), S3-/- (red) and two MitoS3.A/B (orange) clones cultured with 2iLIF or without 2iLIF for 24h or 48h ("-24h" or "-48h"). Early differentiation markers are more readily induced in S3-/- and MitoS3.A/B clones rescues this effect. Data are normalized to the mean of n = 3 independent experiments. Beta-actin served as an internal control. Bars indicate mean and s.e.m. of n = 3 independent experiments, shown as dots. *P* values calculated using two-tailed unpaired T-test relative to S3-/-. See also Fig. S4.

g, Alkaline phosphatase (AP) staining in S3+/+, S3-/- and MitoS3.A/B clones cultured with 2iLIF or without 2iLIF for 24h, 48h or 72h. Right: Number of AP-positive colonies, relative to S3+/+ cells in 2iLIF. Mean +/- s.e.m. of n = 3 independent experiments is shown.

Figure 6

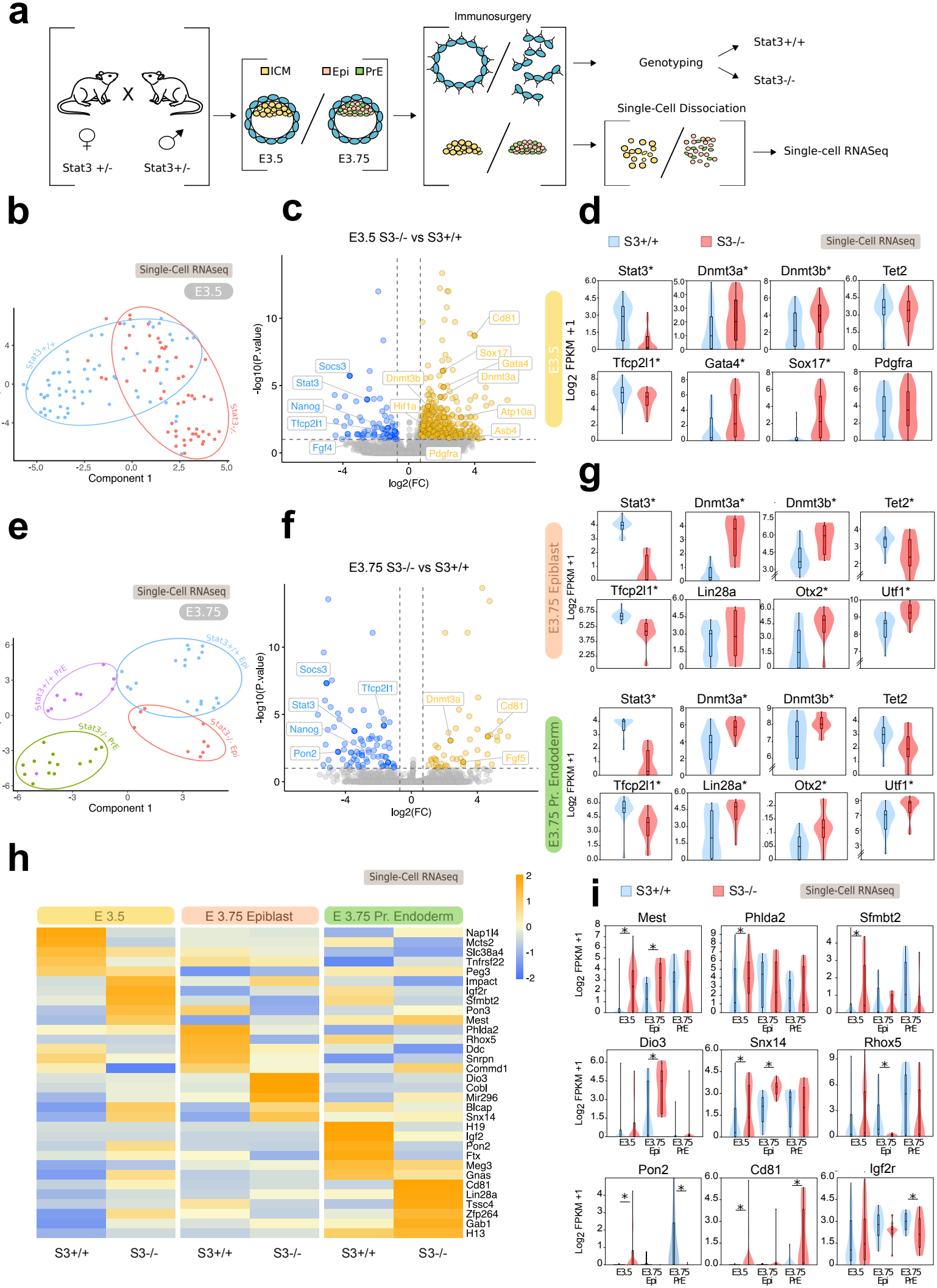
a
Stat3 +/-
Stat3+/-
ICM
Epi
PrE
E3.5
E3.75
Immunosurgery
Genotyping
Stat3+/+
Stat3-/-
Single-Cell Dissociation
Single-cell RNASeq
b
Single-Cell RNAseq
E3.5
Component 1
Component 2
c
E3.5 S3-/- vs S3+/+
-log10(P.value)
log2(FC)
d
S3+/+
S3-/-
Single-Cell RNAseq
E3.5
Log2 FPKM +1
e
Single-Cell RNAseq
E3.75
f
E3.75 S3-/- vs S3+/+
g
E3.75 Epiblast
E3.75 Pr. Endoderm
h
Single-Cell RNAseq
E 3.5
E 3.75 Epiblast
E 3.75 Pr. Endoderm
i
S3+/+
S3-/-
Single-Cell RNAseq
Log2 FPKM +1

Fig. 6. Stat3 regulates Dnmts and imprinted transcripts in the early mouse blastocysts.

a, Schematic outline of the single-cell isolation from embryos and profiling. A total of 171 cells from 18 embryos were analysed (See Supplementary table 3).

b, t-Distributed Stochastic Neighbor Embedding (t-SNE) based on whole transcriptome of wild-type (S3+/+) and mutant (S3-/-) cells collected at embryonic day E3.5; each dot represents a single cell.

c, Volcano plot of genes differentially expressed between S3-/- and S3+/+ cells at E3.5. Each dot represents one gene. The x axis shows the fold change in expression levels (in log-scale) and the y axis represents the statistical significance (adjusted *P* value, in log-scale). Red and blue dots indicate respectively transcripts that are upregulated or downregulated (log2 FC > 0.7 or log2 FC < -0.7 respectively, adjusted *P* value < 0.1) in S3-/- cells relative to S3+/+ cells. See also Supplementary table 3.

d, Violin plots showing the distribution of expression levels (log2 FPKM+1) for selected markers in S3+/+ (blue) and S3-/- (red) cells collected at E3.5. Boxplots show 1st, 2nd and 3rd quartile. Whiskers indicate minimum and maximum values. Note that Stat3 and its target Tfcp2l1 are downregulated in S3-/- cells at E3.5 while post implantation epiblast markers (i.e. Dnmt3a, Dnmt3b) and PrE markers (i.e. Sox17, Pdgfra) are upregulated in S3-/-. * = *P* value <0.05, two-tailed unpaired T-test.

e, t-SNE based on genome-wide expression of S3+/+ and S3-/- mouse cells collected at embryonic day E3.75; each dot represents a single cell. Clustering resolves distinct sample groups by embryonic cells type (Epiblast - Epi and Primitive Endoderm - PrE) along the first dimension, and by genotype (S3+/+ and S3-/-) along the second dimension.

f, Volcano plot of genes differentially expressed between S3-/- and S3+/+ cells at E3.75. Each dot represents one gene. The x axis shows the fold change in expression levels (in log-scale) and the y axis represents the statistical significance (adjusted *P* value, in log-scale). Red and blue dots indicate respectively transcripts that are upregulated or downregulated (log2 FC > 0.7 or log2 FC < - 0.7 respectively, adjusted *P* value < 0.1) in S3-/- cells with respect to S3+/+ cells. See also Supplementary table 3.

g, Violin plots showing the distribution of gene expression levels (log2 FPKM+1) for S3+/+ (blue) and S3-/- (red) ES cells collected at E3.75 . Boxplots show 1st, 2nd and 3rd quartile. Whiskers indicate minimum and maximum values. * = *P* value <0.05, two-tailed unpaired T-test.

h, Heatmap reporting average expression levels of imprinted transcripts in three different embryonic populations (E3.5 ICM, E3.75 Epi, E3.5 PrE) from S3+/+ and S3-/- embryos. Expression values were scaled and represented as z-score. Yellow and blue indicate high and low expression, respectively. Only expressed imprinted genes (average FPKM >1) we analysed.

i, Violin plots showing the distribution of expression levels (log2 FPKM+1) of imprinted genes for S3+/+ (blue) and S3-/- (red) ES cells collected at E3.5, 3.75 Epi and 3.75 PrE. Boxplots show 1st, 2nd and 3rd quartile. Whiskers indicate minimum and maximum values. * = *P* value <0.05, two-tailed unpaired T-test.

Acknowledgements

The authors thank Harry Leitch for critical reading of the manuscript, and Marco Montagner, Sirio Dupont and the Martello Laboratory for discussions and suggestions. We thank Paolo Martini for help with Next Generation sequencing data analyses and Diego De Stefani for technical help with mitochondrial assays. We thank Antonio Simeone for the Otx2 null ES cells. We thank Federico Caicci and Alberto Dinarello for electron microscopy acquisition and technical support. G.M.'s Laboratory is supported by grants from the Giovanni Armenise–Harvard Foundation, the Telethon Foundation (TCP13013) and an ERC Starting Grant (MetEpiStem). S.O. Laboratory is supported by grants from Associazione Italiana Ricerca sul Cancro (AIRC -IG 2017 Id. 20240). N.M. Laboratory is supported by grants from the Giovanni Armenise–Harvard Foundation, Intramural Transition grant from Università degli Studi di Milano, European Foundation for the Study of Diabetes/Lilly Programme 2015, MIUR Progetto Eccellenza to Department of Pharmacological and Biomolecular Sciences (DiSFeB), University of Milan, Milan, Italy.

Author contributions:

G.M., N.M. and S.O. designed the study; R.M.B. and L.D. performed ES cell culture, molecular characterization and functional assays and visualization; V.P. and L.D. performed MeDIP qPCR; R.M.B., V.P. and S.R. performed RRBS; R.M.B. and S.R. performed Western Blots; A.L. and D.I. performed RRBS integrated analysis; M.A. and L.D. performed RNAseq analysis; R.M.B., S.P. and M.A. performed metabolomic analyses; R.M.B., V.P., V.G. and D.T. performed nucleotide Mass Spectrometry; P.G. performed proteomics; M.E.S. and G.R. performed mitochondrial and nuclear fractionations; L.D. and G.G.S. performed single-cell RNAseq analysis; T.L., T.B. and J.N performed embryo dissection and single-cell RNAseq

library preparation; G.M. wrote the manuscript with inputs from all authors; G.M., N.M. and S.O. supervised the study.

Competing interests:

The authors declare no competing interests.

Data and materials 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-sequencing and RRBS process data, used in Figures 1d, 2a-c, 2g-h, 3g, 4b-g, 5a-e, 6b-i, S2c-g, S5 are reported in Supplementary tables 1, 2 and 6 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, 4e , S1c are reported in Supplementary tables 4 and ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD020385. Additional data that support the findings of this study, such as Mass Spectrometry measurements and reagents are available from the corresponding authors upon reasonable request.

Methods

Cell lines and culture routine

Wild type or Stat3 KO (described previously in [13,63] and provided by A. Smith's Lab) mouse ES cell lines were routinely cultured without feeders on gelatine-coated plates (0.2% gelatine, Sigma-Aldrich, cat. G1890). Media was changed every 2 days and cells were passaged when approaching confluency (every 2-3 days); to passage, cells were replated at required density following dissociation with either Accutase (Thermo-Fisher, cat. A11105-01) or Trypsin (Thermo-Fisher, cat. 15090-046).

All cells were maintained at 37°C in humidified incubators with 5% CO2.

Media and supplements

Cells were grown under two different culture conditions, prepared as follows:

2i / 2iLIF: serum-free KSR (Knockout serum replacement) 10% (Life Technologies, cat. 10828-028) - based medium in GMEM (Sigma-Aldrich, cat. G5154) supplemented with 1% FBS (Sigma-Aldrich, cat. F7524), 100 mM 2- mercaptoethanol (Sigma-Aldrich, cat. M7522), 1×MEM non-essential amino acids (Invitrogen, cat. 1140-036), 2mM L-Glutamine, 1mM sodium Pyruvate (both from Invitrogen), and with small-molecule inhibitors PD (1 μM, PD0325901), CH (3 mM, CHIR99021) from Axon (cat. 1386 and 1408) and LIF (100 units/ml, produced in-house).

Serum LIF: GMEM (Sigma-Aldrich, cat. G5154) supplemented with 10% FBS (Sigma-Aldrich, cat. F7524), 100 mM 2-mercaptoethanol (Sigma-Aldrich, cat. M7522), 1× MEM non-essential amino acids (Invitrogen, cat. 1140-036), 2mM L-Glutamine, 1mM Sodium Pyruvate (both from Invitrogen), and 100 units/ml LIF.

Generation of mutant cell lines

Dnmt3b -/- ESCs Knockout production was performed using TALEN technology as described in Neri et al. [73]. In brief, cells were transfected with the two TALEN constructs targeting Exon 17 of murine Dnmt3b and after 16 hours were seeded as a single cell. After 1 week, clones were screened by western blot analyses. Positive clones were analyzed by genomic sequencing of the TALEN target.

Dnmt3a-/- and Dnmt3a/b-/- ESCs generation were performed by CRISP-R/ Cas9 method. E14 Wild-type cells and Dnmt3b-/- were co-transfected with Cas9 construct and the two RNA guide construct targeting Exon 19 (FW gRNA exn19= CACCgACCGCCTCCTGCATGATGCGCGG, REV gRNA exn19= aaacCCGCGCATCATGCAGGAGGCGGTc) of murine Dnmt3a. After 16 hours single cells sorting was carried out. Ten days later, clones were screened by western blot analysis. Positive clones were analyzed by genomic sequencing.

For DNA transfection, we used Lipofectamine 2000 (Life Technologies, cat. 11668-019) and performed reverse transfection. For one well of a 12-well plate, we used 3 ul of transfection reagent, 1 ug of plasmid DNA, and 150,000 cells in 1 ml of KSR medium. The medium was changed after overnight incubation.

Stable transgenic ESCs lines expressing sh-TET1/2, sh-Scramble or MLS-Stat3-NES were generated by transfecting cells with PiggyBac transposon plasmids CAG-sh-TET1/2 *(targets seq: TET1'=CTCATCTACTTCTCACCTAGTG, TET1''=AAGAGAACCTGGTGCATCAGA, TET2'=AGCTCTGAACAGTATTCAAAGC, TET2''=A TAGGACTATAATGTATAGATA)* CAG-sh-Scramble *(targetseq:Scramble_miR30-shRNA = ACCTAAGGTTAAGTCGCCCTCG)*

or CAG-MLS-Stat3-NES (seq MLS=
GTGGACGAGATGACCAAGAAGTTCGGCACCCTGACCATCCACGACACCGAGAAG

derived from [21]; seq NES=
GTGGACGAGATGACCAAGAAGTTCGGCACCCTGACCATCCACGACACCGAGAAG
)with piggyBac transposase expression vector pBase. Selection for transgenes was applied, and stable clones were selected in 2iLIF conditions.

STAT3ER plasmid transfection previously described in Takahiko Matsuda et al. 1999 [31] was performed with a 1ug of linearized plasmid (enzyme PbuI). Plasmid encondes for the entire coding region of mouse STAT3 followed by the modified ligand-binding domain (G525R) of mouse estrogen receptor under control of CAG promoter (pCAGGS vector).

Dnmt3a1/2 costructs for Overexpression experiments were obtained by PCR amplification of the entire coding region (Dnmt3a1 or Dnmt3a2) and cloned into the XbaI–NotI site of pEF6/V5-His vector (Invitrogen). Dnmt3b construct was obtained by PCR amplification and cloned into pEF6/V5-His vector (Invitrogen) previously described in Neri et al. 2017 Nature[73].

LIF induction

For LIF induction experiments, ES cells cultured in 2i without LIF for at least 4 passages were plated in 2i. Twenty-four hours after plating, LIF was added for the indicated amount of time (24 and 48 hours), cells were then fixed for immunofluorescence.

Treatments

For inhibition of the respiratory chain, cells were treated acutely with 100nM complex I inhibitor Rotenone (Sigma-Aldrich, cat. R8875) and 200nM complex III inhibitor Antimycin (Sigma-Aldrich, cat. A8674).

For studies on glutaminolysis, cells were cultured in KSR-based medium prepared as described above, but without the addition of Glutamine. Exogenous DM-αKG (dimethyl 2-oxoglutarate) used for treatments was added to culture medium at the indicated concentration and absorbed by cells as it is membrane-permeable (Sigma-Aldrich, cat. 349631-5G).

Clonal assay

The ability of single mES cells to form pluripotent colonies was assessed through a clonal assay. 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/well. Cells were grown for 4-5 days before they were fixed and stained for Alkaline Phosphatase.

Differentiation assay

Cells were cultured with or without 2iLIF for 24, 48 and 72 hours. After 72 hours cells were detached and replated at clonal density in 2iLIF. Cells surplus were conserved for gene expression analysis. Finally, cells were stained with Alkaline Phosphatase after 5 days to evaluate the number of pluripotent cells. For AP staining, cells were fixed with a citrate-acetone-formaldehyde solution and stained using the Alkaline Phosphatase kit (Sigma-Aldrich, cat. 86R-1KT). Plates were scanned using a Nikon Scanner and scored manually.

Alkaline Phosphatase Staining

Fixation solution: 65% Acetone, 25% Citrate (provided with kit), 8% Formaldehyde

Staining solution: Alkaline Phosphatase (AP) kit (Sigma-Aldrich, cat. 86R-1KT) according to the manufacturer's protocol.

Culture medium was removed from adherent cells and they were fixed with fixation solution. Plates were then washed with H2O and the staining solution was added for 5 minutes in the dark. Then plates were washed again with H2O and dried.

Colonies were scored manually using an optical microscope, discriminating between undifferentiated (AP-positive), mixed or differentiated (AP-negative).

Immunofluorescence and stainings

For 5mC staining, cells were fixed in 4% formaldehyde (Sigma-Aldrich, cat. F8775) for 10 minutes, then washed in PBS and treated for 15 minutes with NH4Cl. Next, cells were permeabilized with 1h PBST 0.5% treatment (PBS, 0.5% Triton X-100, Sigma-Aldrich, cat. 93443) and 2N HCl was added for 45 minutes to denature the DNA. Cells were blocked for 1 hour in 5% horse serum (HS) with 0.3% PBST (Thermo-Fisher, cat. 16060122) and then incubated overnight at 4°C with anti-5mC primary antibody (Eurogentec, cat. BI-MECY-0500, Supplementary table 5) diluted in 2% HS with 0.3% PBST. After washing with PBST 0.1%, cells were incubated with secondary antibody (Alexa Fluor 488 donkey anti-mouse, Life Technologies, cat. A21202) for 45 minutes at room temperature (RT). Nuclei were stained with mounting medium Fluoroshied containing DAPI (Sigma-Aldrich, cat. F6057).

For EdU staining, cells were exposed to an EdU pulse of 4 hours before fixation in 4% formaldehyde for 10 minutes; samples were then processed according to manufacturer's instructions (Life Technologies).

For Atad3 and Stat3 colocalization staining, cells were fixed for 10 minutes in 4% formaldehyde, washed in PBS and blocked and permeabilized for 1 hour in 5% horse serum (HS) with 0.3% PBST. Cells were incubated overnight at 4°C with primary antibodies (

Supplementary table 5). After washing with 0.1% PBST, cells were incubated with secondary antibodies (Alexa, Life Technologies) for 30 minutes at RT. Nuclei were stained with mounting medium Fluoroshied containing DAPI.

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

Oxygen consumption assay (Seahorse Assay)

Oxygen consumption was measured using the Seahorse XF24 (Seahorse Bioscience). For this, 20 hours before the analysis both S3+/+ and S3-/- cells were seeded in a 24-well cell culture plate (Seahorse Bioscience) coated with laminin (Sigma-Aldrich, cat. L2020) at a density of 100,000 cells per well in KSR media supplemented with 2i + LIF. It is crucial to have an evenly plated monolayer of cells to obtain reliable measurements. Cells were maintained in a 5% CO2 incubator at 37°C, and 1 hour before the experiment, the cells were washed and incubated in 600 µl of DMEM (Sigma-Aldrich, cat. D5030-10X1L) with 2mM Glutamine, 1mM NaPy, 25 mM glucose, 3 mg/L phenol red and 143 mM NaCl, with pH 7.4 at 37°C in a non-CO2 incubator.

During the experiment, oxygen concentration was measured over time periods of 2 min at 5 minutes intervals, consisting of a 3-min mixing period and 2 minutes waiting period. Oxygen consumption rate (OCR) is measured before and after the addition of inhibitors to derive several parameters of mitochondrial respiration. Initially, cellular OCR is measured in basal conditions to derive the basal mitochondrial respiration; next, 200 nM mitochondrial uncoupler FCCP (carbonyl cyanide-p- trifluoromethoxyphenyl-hydrazon) is automatically added to the medium to maximize Electron Transport Chain (ETC) function, in order to derive maximal respiratory capacity. Next, Antimycin A and Rotenone - inhibitors of

complex III and I - are released into the medium to block ETC, revealing the non-mitochondrial respiration.

Gene expression analysis by quantitative PCR with reverse transcription

Total RNA was isolated using a Total RNA Purification kit (Norgen Biotek, cat. 37500), and complementary DNA (cDNA) was made from 500 ng using M-MLV Reverse Transcriptase (Invitrogen, cat. 28025-013) and dN6 primers (Invitrogen). For real-time PCR, SYBR Green Master mix (Bioline, cat. BIO-94020) was used. Three technical replicates were carried out for all quantitative PCR. An endogenous control (beta-actin) was used to normalize expression. Primers are detailed in Supplementary table 6.

RNAseq

Total RNA was isolated using Total RNA Purification kit (Norgen Biotek, cat. 37500) and sequenced using an Illumina HiSeq4000, in 150-base, paired-end format.

Reads were aligned to mouse transcriptome (Mus musculus transcriptome generated by rsem-prepare-reference with ENSEMBL93 GTF) and mouse genome (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).

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

All RNAseq 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 log2-fold change > 1 and with an adjusted *P* value (P_{adj}) < 0.01 (Benjamini-Hochberg adjustment) were considered significant and defined as differentially expressed (Differentially Expressed Genes = DEG) for the comparison in the analysis.

Heatmaps were made using TPM values with the pheatmap function from pheatmap R package (v.1.0.12, distance = 'correlation', scale = 'row') on DEGs or selected markers.

Volcano plots were computed with log2fold change and −log10Padj from DESeq2 differential analysis output using ggscatter function from ggpubr R package (v. 0.2).

Western blot

Cells were washed in PBS and harvested with lysis F-buffer (10 mM TrisHCl pH7, 50 mM NaCl, 30 mM Sodium pyrophosphate tetrabasic, 50 mM NaF, 1% Triton X-100 Buffer). In order to obtain protein lysates, extracts were exposed to ultrasound in a sonicator (Diagenode Bioruptor) for 3 pulses. Cellular extracts were centrifuged for 10 minutes at 4°C (max speed) to remove the insoluble fraction. Extracts were quantified using bicinchoninic acid (BCA) assay (BCA protein assay kit; catalog no. 23225; Pierce). Samples were boiled at 95°C for 5 minutes in 1X Sample Buffer (50mM Tris HCl pH 6.8, 2% SDS, 0.1% Bromophenol Blue, 10% glycerol, 2% 2-mercaptoethanol).

Each sample was loaded in a commercial 4-12% MOPS acrylamide gel (Life Technologies; BG04125BOX/BG00105BOX) and electrophoretically transferred on a PVDF membrane (Millipore; IPFL00010) in a Transfer solution (50mM Tris, 40mM glycine, 20% methanol, 0.04% SDS). Membranes were then saturated with 5% Non-Fat Dry Milk powder (BioRad; 170-6405- MSDS) in TBST (8g NaCl, 2.4g Tris, 0.1% Tween20/liter, pH 7.5) for 1 hour at room temperature and incubated overnight at 4 °C with the primary antibody (Supplementary table 5) diluted in a range of 0,5-1% milk powder (depending on antibody) in TBST.

Membranes were then incubated with secondary antibodies conjugated with a peroxidase, diluted in 0,1% or 0,5% milk in TBST. Pico SuperSignal West chemiluminescent reagent (Thermo Scientific, cat. 34078) was used to incubate membranes and chemiluminescence from the interaction between peroxidase and substrate present in the commercial reagent was digitally acquired by ImageQuant LAS 4000.

Mitochondria and nuclear isolation

Nuclear and mitochondrial isolation was performed as indicated in Rosner M. et al. 2013 and Frezza C. et al. 2007[75,76], respectively, with some modifications. Briefly, cells 4×10^7 collected in PBS were centrifuged at 600 x g, 5 min, and the pellet, resuspended in 2 ml of Isolation buffer (IBc), was homogenized with a Dounce homogenizer and Teflon pestle by 20 strokes (x4) on ice. Then, the homogenate was centrifuged at 600 x g, 5 min, 4°C, and supernatant (SN) and pellet were collected to proceed separately with mitochondria and nuclei isolation. SN was centrifuged twice at 40 x g and pellets were conserved to proceed with nuclei isolation; SNs were further centrifuged sequentially at 600 x g and 1200 x g for 5 min at 4°C, to further eliminate debris. Finally, the resulting SN was centrifuged at 7500 x g to obtain the mitochondrial pellet that was finally washed twice with IBc and centrifuged at 9000 x g. For nuclei isolation, the pellet obtained after the first centrifugation at 40 x g was resuspended in nuclear isolation buffer (20mM TRIS pH 7.5, 50 mM ß-mercaptoethanol, 0.1mM EDTA, 2mM MgCl2, 1mM PMSF) supplemented with protease inhibitor, and incubated in sequence 2 min at RT and 10 min on ice, to then proceed with centrifugation at 600 x g, 4°C. The obtained pellet was resuspended in 400 µl of the same buffer with addition of 1% NP40. Nuclei were pelleted at 500 x g, 4°C, and washed 3 times with the same buffer. Finally, mitochondrial and nuclear pellets were lysed in RIPA buffer.

Metabolites analysis by Mass Spectrometry

Cells were grown in 6-well plates, harvested in ice-cold PBS and centrifuged at 500g for 3 minutes at 4°C. Pellets were then resuspended in 250µl methanol/acetonitrile 1:1 containing [U-$^{13}C_6$]-Glucose and [U-$^{13}C_5$]-Glutamine each at 1ng/µl (internal standards, Sigma-Aldrich, cat. 389374) and centrifuged at 20,000g for 5 minutes at 4°C. Supernatants were then passed through a regenerated cellulose filter, dried and resuspended in 100µl of MeOH for subsequent analysis.

Metabolomic data were performed 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). All the methods have been arranged by setting multiple ion monitoring (MRM) with pure commercial standards in order to confirm the identity of all metabolites.

Quantification of different metabolites was performed with a liquid chromatography/tandem Mass Spectrometry (LC-MS/MS) method using a C18 column (Biocrates) for amino acids and SAMe and cyano-phase LUNA column (50mm x 4.6mm, 5µm; Phenomenex). Methanolic samples were analyzed by a 10 minutes run in positive (amino acids and SAME) and 5 minutes run in negative (all other metabolites) ion mode with specific multiple reaction monitoring (MRM) transitions. Amino acids quantification was performed through the previous derivatization. Briefly, 50µl of 5% phenyl isothiocyanate (PITC) in 31.5% EtOH and 31.5% pyridine in water were added to 10µl of each sample. Mixtures were then incubated with PITC solution for 20 min at RT, dried under N2 flow and suspended in 100µl of 5mM ammonium acetate in MeOH/H2O 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 B: 2 mM ammonium acetate in MeOH. The gradient

was 90% B for all the analysis with a flow rate of 500µl/min. MultiQuantTM software (version 3.0.2) was used for data analysis and peak review of chromatograms. Quantitative evaluation of all metabolites was performed based on calibration curves with pure standards, then data were normalized on total protein content.

Metabolic flux analysis

For metabolic tracing analyses, cells were exposed for 24h to [U-$^{13}C_6$]-Glucose 1mM (Sigma-Aldrich, 389374) or [U-$^{13}C_5$]-Glutamine 2mM (Sigma-Aldrich, cat. 605166) or [U-$^{13}C_{16}$]-Palmitate 100µM (Sigma-Aldrich, cat. 605573). Cells were harvested in ice-cold PBS and centrifuged at 500g for 3 minutes at 4°C. Pellets were then resuspended in 250µl methanol/acetonitrile 1:1 and spun at 20,000g for 5 min at 4°C. Supernatant were then passed through a regenerated cellulose filter, dried under N2 flow and resuspended in 100µl of MeOH for subsequent analysis. Metabolomic data 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 with a liquid chromatography/tandem Mass Spectrometry (LC-MS/MS) method using a cyano-phase LUNA column (50mm x 4.6mm, 5µm; Phenomenex) Methanolic samples were analyzed by a 5 min run in negative (Metabolites) ion mode. The mobile phases for negative ion mode analysis was phase A: 2 mM ammonium acetate in MeOH and phase B: water. The gradient was 90%A for all the analysis with a flow rate of 500µl/min. MultiQuantTM software (version 3.0.2) was used for data analysis and peak review of chromatograms. Samples were analyzed after 8 hours of ^{13}C-labelling to ensure that isotopic equilibrium was reached, as previously shown in ES cells cultured in 2iLIF by Carrey and colleagues. All detected ^{13}C-labelled metabolites were corrected for natural isotope abundances.

Nucleoside preparation for Mass Spectrometry

DNA was extracted using Puregene core kit A, then measured with a Nanodrop spectrophotometer. 50µg DNA were passed through the Microcon YM-10 centrifugal filtration cartridge (Millipore, cat. no. 42407, MRCPRT010) 10KDa columns two times. The first time 50µg of DNA were solubilized into 500µL of double-distilled water, then concentrated to about 30 µL by spinning the columns at 13900g for 25 minutes. The second time, the 30µL of recovered DNA were solubilized into 500µL of 1X digestion buffer and then concentrated to about 15µL by spinning at 13900g for 35 minutes.

After the 2 steps, the DNA concentration was measured at the Nanodrop spectrophotometer. The DNA was then digested to nucleosides, at 37°C for 6-7 hours, using a mix containing 2U Antarctic Phosphatase (stock solution is 5U/µl) (New England Biolabs, M0289S), 3mU Snake venom phosphodiesterase I (stock solution is 1mU/µl) (Crotalus adamanteus venom, Sigma-Aldrich, P3243-1VL), 2.5U Benzonase (stock solution is 250U/µl) (Sigma-Aldrich E1014-5KU), in 3.4µl volume of enzyme mix + 1.6µl of double distilled water + 5µl 2X digestion buffer (20mM Tris Hcl pH 7.9 100mM NaCl, 20mM MgCl2) + 5µl of DNA (7,5 – 10µg) in 1X digestion buffer. After the digestion 1µg of undigested genomic DNA and 1µg of digested DNA were loaded on a 2% gel, in order to confirm the complete digestion of the genomic DNA. 12µl of digested nucleosides were provided for MS analysis to CNRS at Gif-sur-Yvette [77].

Mass Spectrometric Analysis of Total Nucleosides

Analysis of the nucleoside digests of DNA by HPLC was performed with a U-3000 HPLC system (Thermo-Fisher). An Accucore RP-MS (2.1 mm X 100 mm, 2.6 µm particle, Thermo-Fisher) column was used at a flow rate of 200 µL/min and a fixed temperature at 30°C. Mobile phases were 5 mM ammonium acetate, pH 5.3 (buffer A) and 40% aqueous

acetonitrile (Buffer B). A multilinear gradient was used with only minor modifications from that described previously [78]. The injection volume was fixed at 6 µL.

A LTQ orbitrap Mass Spectrometer (Thermo-Fisher) equipped with an electrospray ion source was used for the HPLC-MS identification and quantification of nucleosides. Mass Spectra were recorded in the positive ion mode over an *m/z* range of 100-1000 with a capillary temperature set at 300°C, spray voltage at 4.5 kV and sheath gas, auxiliary gas and sweep gas at 40, 12 and 7 arbitrary units, respectively.

Calibration curves were generated using a mixture of synthetic standards of 2'-Deoxycytidine (2dC)(Sigma-Aldrich), 5-Methyl-2'-deoxycytidine(5-mdC) and 5-hydroxymethyl-2'-deoxycytidine (5-hmdC) (Bertin-Pharma) in the ranges of 10-100 injected pmol for 2dC, 0.4-4 injected pmol for 5-mdC and 0.5-10 injected pmol for 5-hmdC. Each calibration point was injected in triplicate. Extracted Ion Chromatograms (EIC) of base peaks of the following signals: 2dC (*m/z* 228.08-228.12), 5-mdC (*m/z* 242.10-242.13), and 5-hmdC (*m/z* 258.08-258.12), were used for quantification. In all cases, coefficients of variations for peak areas were always below 15%. Experimental data (peak area *versus* injected quantity) were fitted with a linear regression model for each compound leading to coefficient of determination (R^2) values better than 0.97. Accuracies were calculated for each calibration point and were always better than 15%.

MEDIP

Genomic DNA was extracted with phenol-chloroform, resuspended in TE buffer containing 20µg/ml RNAse A (Thermo-Fisher, cat. EN0531) and passed through a needle 10 times to reduce its viscosity, then measured at the Nanodrop spectrophotometer.

40µg of DNA were resuspended in 130µl of TE, transferred to a microtube (microtubes AFA fiber pre-slit snap cap 6x16mm, Covaris) and sonicated with the Covaris S2 (Duty cycle 10%,

Intensity 5, Cycles burst 200, 45 seconds per cycle; 3 cycles to have a distribution of size between 100 and 600, 4 cycles to have a distribution of size between 100 and 400 and 5 cycles to have a distribution of size between 100 and 300).

10µg of sonicated DNA were diluted in 1,125ml TE and denatured for 10 minutes at 100°C in a thermoblock, then quickly cooled on ice for additional 10 minutes. 450µl (= 4µg tot) of denatured DNA were distributed in two low binding tubes with 51µl of 10X IP buffer (100 mM Na-Phosphate pH 7.0, 1.4 M NaCl, 0.5 % Triton X-100), plus 10µl of antibody anti-5mC (Supplementary table 5) (IP sample) or IgG (mock control) were added. The tubes were left rotating with overhead shaking for 2 hours at 4°C. The leftover (= 225µl) is the Input material (50% of Input), to be used in the quantitative PCR.

Dynabeads Protein G (Thermo-Fisher, cat. 10003D) were prepared by taking 40µl per each sample, then washed twice for 5 minutes in 800µl of 0.1% BSA in PBS and finally resuspended in 40µl of 1X IP buffer.

After 2 hours, Dynabeads Protein G were added to the IP and mock samples. Samples were left rotating at 4°C with overhead shaking for additional 2 hours.

The beads were then separated using the magnetic stand and washed 3 times for 10 minutes in 1X IP buffer; the supernatant was removed and trashed (unbound fraction). Finally the beads were resuspended in 250µl of proteinase K digestion buffer (50 mM Tris pH 8.0, 10 mM EDTA, 0.5 % SDS) containing 3.5µl of proteinase K (20mg/µl). The samples were incubated overnight at 50°C in a shaking thermoblock (500rpm). The day after the beads were separated with a magnetic rack and the supernatant was saved.

The DNA contained in the supernatant fraction was purified using Qiaquick PCR Purification kit (Qiagen, cat. 2816) and eluted in 30 µl. The saved Input material (50% Input) was re-purified and concentrated using Qiaquick kit; elution was done in a volume of 30µl. Primers are detailed in Supplementary table 7.

Reduced Representation Bisulfite Sequencing (RRBS)

RRBS was performed as previously described in[79]. Briefly, 500ng of DNA was digested at 37°C with 200U of MspI restriction endonuclease (NEB). Digested DNA was than end repaired, dA-tailed, and ligated to methylated adapters, using the Illumina TruSeq DNA Sample Prep Kit, following manufacturer's instructions. Adapter-ligated DNA was loaded on 2% agarose gel and a fraction from 200 to 400 bp was recovered. Purified DNA was then subjected to bisulfite conversion using the EpiTect Bisulfite Kit (Qiagen). 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, sequencing reads were mapped to mouse genome reference (mm10/GRC.m38.p6) with BSMAP (v2.89)[80] using RRBS mode (parameters: -s 12 -D C-CGG -v 0.01 –n 1). CpG methylation levels were extracted from aligned reads as the ratio of the number of Cs over the total number of Cs and Ts using the methratio.py script. 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 10x were retained. All samples were processed identically.

Statistical analyses were conducted within the R software environment. Differential methylation analysis at single nucleotide resolution was performed for each comparison (i.e. S3+/+ 2iLIF vs S3 -/- 2iLIF; S3+/+ 2i vs S3 +/+ 2iLIF; Dnmt3a/b dKO.A and dKO.B 2i vs E14 2i; 2iLIF vs 2i E14) using the methylKit R/Bioconductor package[81], exploiting the logistic regression approach for testing replicates (calculateDiffMeth function with default parameters). CpG sites with absolute methylation difference >= 10% and q-value <= 0.05 were considered as 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 2i-cultured wild type cells 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 mES cells were retrieved from ENCODE (https://www.encodeproject.org/). Active promoters and enhancers were defined from processed peaks data as following:

Promoters: H3K4me3 peaks in a 2kb window centered in the TSS of annotated genes (GENCODE release M20);

Enhancers: distal H3K27ac peaks (more than 1kb up/downstream the nearest TSS).

Differential methylation analysis (calculateDiffMeth function, q-value <=0.05, methylation difference >= 10%) was performed on these regions (i.e., testing all the covered CpG overlapping with the ChIP-seq peaks, with 200bp of flanking region) for the comparisons: S3+/+ 2iLIF vs S3 -/- 2iLIF and S3+/+ 2i vs 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 DNA methylation levels of promoters/enhancers. Similar analyses were conducted on a manually curated list of imprinted DMRs.

Single-cell RNAseq analysis of Stat3-/- and +/+ embryos

Immunosurgery and single-cell dissociation was performed as described in Boroviak 2015 Dev Cell[7]. The method for single-cell RNA-seq and Library preparation was previously described in Boroviak 2018 Dev [82]. A total of 171 cells from 18 embryos were analysed. Experiments were performed in accordance with EU guidelines for the care and use of laboratory animals, and under authority of UK governmental legislation. Use of animals in

this project was approved by the ethical review committee for 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 [82]. Alignments to gene loci were quantified with htseq-count [83] based on annotation from Ensembl 87. PCA outliers were computed and removed. Mouse embryo for E4.5, E5.5 and E6.5 stages were compiled from earlier studies[10,84]. Principal component analysis was based on Log2 FPKM values computed with the Bioconductor package DESeq [85], custom scripts and FactoRmineR package [86]. Differential expression analysis was performed with scde [87], that fits individual error models for 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 R package DeconRNASeq [88] 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 "signature" dataset. See also Supplementary table 3.

Proteomics

All the experiments have been performed in a labeling free setting. For each sample, 50mg of total cellular protein extract were precipitate over-night at 4°C in acetone, then reduced and alkylated in a solution of 6M Guanidine-HCl, 5mM TCEP, and 55mM chloroacetamide. Peptides were obtained digesting proteins with LysC (WAKO) for 3h at 37°C and with the endopeptidase sequencing-grade Trypsin (Promega) overnight at 37°C. Collected peptide mixtures were concentrated and desalted using the Stop and Go Extraction (STAGE) technique [89].

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 [90]. Peptide separation was carried out according to their hydrophobicity on a PicoFrit column, 75mm ID, 8Um tip, 250mm bed packed with Reprosil-PUR, C18-AQ, 1.9mm particle size, 120 Angstrom pore size (New Objective, Inc., cat. PF7508-250H363), using a binary buffer system consisting of solution A: 0.1% formic acid and B: 80% acetonitrile, 0.1% formic acid. Runs of 120 min, after loading, were used for proteome samples, with a constant flow rate of 300nl/min. After sample loading, run start at 5% buffer B for 5min, followed by a series of linear gradients, from 5% to 30% B in 90min, then a 10 min step to reach 50% and a 5 min step to reach 95%. This last step was maintained for 10 min.

Q Exactive HF settings: MS spectra were acquired using 3E6 as an AGC target, a maximal injection time of 20ms and a 120,000 resolution at 200m/z.

The mass spectrometer operated in a data-dependent Top20 mode with subsequent acquisition of higher-energy collisional dissociation (HCD) fragmentation MS/MS spectra of the top 20 most intense peaks. Resolution, for MS/MS spectra, was set to 15,000 at 200m/z, AGC target to 1E5, max injection time to 20ms and the isolation window to 1.6Th. The intensity threshold was set at 2.0 E4 and Dynamic exclusion at 30 seconds.

All 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, while oxidations at methionine, acetylation at the N-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 label free quantification (LFQ), minimum ratio count was set as 1. 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 Ratios were logarithmized, grouped and filtered for min. valid number (min. 4 in at least one group). Missing values have been replaced by random numbers that are drawn from a normal distribution. Two sample t-test was performed using FDR=0.05. *P* values < 0.05 were 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.

Electron microscopy and DAB staining

Cells were fixed in a 24 wells plate with 4% Paraformaldehyde in PBS (pH 7,4) for 30 min. at RT. After fixation cells were washed 5 times with PBS (5 min. each) blocked and permeabilized with 5% normal goat serum and 0,1% saponin in PBS for 30 min, and then incubated with primary antibody anti-Stat3 O.N at 4°C. in PBS 5% normal goat serum and 0,05% saponin. After 5 washing with PBS, (5 min each) cells were incubated with HRP-conjugated Fab fragments of the secondary antibody for 2h. RT. After 5 washing cells were incubated in the DAB solution (0.01gr DAB in 20ml TRIS-HCl buffer plus 30% H2O2 solution just before use). Subsequently the samples were postfixed with 1% osmium tetroxide plus potassium ferrocyanide 1% in 0.1M sodium cacodylate buffer for 1 hour at 4°. After three water washes, samples were dehydrated in a graded ethanol series and embedded in an epoxy resin (Sigma-Aldrich). Ultrathin sections (60-70 nm) were obtained with an Ultrotome V (LKB) ultramicrotome, counterstained with uranyl acetate and lead citrate and viewed with

a Tecnai G2 (FEI) transmission electron microscope operating at 100 kV. Images were captured with a Veleta (Olympus Soft Imaging System) digital camera.

Statistics and reproducibility

For each dataset, sample size n refers to experimental or biological replicates, as stated in the figure legends. All *P* values were calculated using the unpaired two-tailed T-test and indicated as their numerical values in each plot; *P* values were not calculated for datasets with $n < 3$. Adjusted *P* values (q-values) were calculated in the case of multiple testing. Either Excel or R software were used for statistical analysis. Error bars indicate the standard error of the mean (s.e.m.) or the standard deviation (SD), as stated in the figure legends.

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. *Dev. Camb. Engl.* **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 naïve 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.* (2013) doi:10.1016/j.stemcr.2013.11.010.
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 RasDependent 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. Kalkan, T. *et al.* Tracking the embryonic stem cell transition from ground state pluripotency. *Dev. Camb. Engl.* **144**, 1221–1234 (2017).
24. Dawlaty, M. M. *et al.* Loss of Tet Enzymes Compromises Proper Differentiation of Embryonic Stem Cells. *Dev. Cell* **29**, 102–111 (2014).
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 naïve pluripotency of mouse embryonic stem cells. *Development* **145**, (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**, D1117–23 (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. *bioRxiv* 2020.07.17.208264 (2020) doi:10.1101/2020.07.17.208264.
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.* e99518 (2018) doi:10.15252/embj.201899518.
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**, 1–6 (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. U. S. A.* **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. 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).
54. Meissner, A. *et al.* Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* **454**, 766–770 (2008).
55. Tahiliani, M. *et al.* Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1. *Science* **324**, 930–935 (2009).
56. 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).
57. 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).
58. 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).
59. 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).
60. 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.

1538 *Nat. Cell Biol.* **16**, 516–528 (2014).
1539 61. McLaughlin, K. *et al.* DNA Methylation Directs Polycomb-Dependent 3D Genome Re-
1540 organization in Naive Pluripotency. *Cell Rep.* **29**, 1974-1985.e6 (2019).
1541 62. von Meyenn, F. *et al.* Impairment of DNA Methylation Maintenance Is the Main Cause of
1542 Global Demethylation in Naive Embryonic Stem Cells. *Mol. Cell* **62**, 848–861 (2016).
1543 63. Ballou, R. H. Logística empresarial: transportes, administração de materiais e
1544 distribuição física. **94**, 388 (1993).
1545 64. Avalle, L. *et al.* STAT3 localizes to the ER, acting as a gatekeeper for ER-mitochondrion
1546 Ca2+ fluxes and apoptotic responses. *Cell Death Differ.* **26**, 932–942 (2019).
1547 65. Wang, L. *et al.* JAK/STAT3 regulated global gene expression dynamics during late-stage
1548 reprogramming process. *BMC Genomics* **19**, 183 (2018).
1549 66. Hwang, I.-Y. *et al.* Psat1-Dependent Fluctuations in α-Ketoglutarate Affect the Timing of
1550 ESC Differentiation. *Cell Metab.* **24**, 494–501 (2016).
1551 67. Zhang, J. *et al.* LIN28 Regulates Stem Cell Metabolism and Conversion to Primed
1552 Pluripotency. *Cell Stem Cell* **19**, 66–80 (2016).
1553 68. Mullen, A. R. *et al.* Reductive carboxylation supports growth in tumour cells with
1554 defective mitochondria. *Nature* **481**, 385–388 (2012).
1555 69. Leeb, M., Dietmann, S., Paramor, M., Niwa, H. & Smith, A. Genetic exploration of the
1556 exit from self-renewal using haploid embryonic stem cells. *Cell Stem Cell* **14**, 385–393
1557 (2014).
1558 70. Choi, J. *et al.* Prolonged Mek1/2 suppression impairs the developmental potential of
1559 embryonic stem cells. *Nature* **548**, 219–223 (2017).
1560 71. Yagi, M. *et al.* Derivation of ground-state female ES cells maintaining gamete-derived
1561 DNA methylation. *Nature* **548**, 224–227 (2017).
1562 72. Chen, T., Ueda, Y., Xie, S. & Li, E. A Novel Dnmt3a Isoform Produced from an
1563 Alternative Promoter Localizes to Euchromatin and Its Expression Correlates with
1564 Activede Novo Methylation. *J. Biol. Chem.* **277**, 38746–38754 (2002).
1565 73. Neri, F. *et al.* Intragenic DNA methylation prevents spurious transcription initiation.

Nature **543**, 72–77 (2017).

74. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, (2014).
75. Frezza, C., Cipolat, S. & Scorrano, L. Measuring mitochondrial shape changes and their consequences on mitochondrial involvement during apoptosis. *Methods Mol. Biol. Clifton NJ* **372**, 405–420 (2007).
76. Rosner, M., Schipany, K. & Hengstschläger, M. Phosphorylation of nuclear and cytoplasmic pools of ribosomal protein S6 during cell cycle progression. *Amino Acids* **44**, 1233–1240 (2013).
77. Quinlivan, E. P. & Gregory, J. F. DNA digestion to deoxyribonucleoside: A simplified one-step procedure. *Anal. Biochem.* **373**, 383–385 (2008).
78. Pomerantz, S. C. & McCloskey, J. A. [44] Analysis of RNA hydrolyzates by liquid chromatography-mass spectrometry. in *Methods in Enzymology* vol. 193 796–824 (Academic Press, 1990).
79. 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).
80. Xi, Y. & Li, W. BSMAP: whole genome bisulfite sequence MAPping program. *BMC Bioinformatics* **10**, 232 (2009).
81. Akalin, A. *et al.* methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biol.* **13**, R87 (2012).
82. Boroviak, T. *et al.* Single cell transcriptome analysis of human, marmoset and mouse embryos reveals common and divergent features of preimplantation development. *Dev. Camb. Engl.* **145**, (2018).
83. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
84. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).

85. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010).
86. Lê, S., Josse, J. & Husson, F. FactoMineR: An R Package for Multivariate Analysis. *J. Stat. Softw.* **25**, 1–18 (2008).
87. Kharchenko, P. V., Silberstein, L. & Scadden, D. T. Bayesian approach to single-cell differential expression analysis. *Nat. Methods* **11**, 740–742 (2014).
88. 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).
89. 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).
90. Michalski, A. *et al.* Mass Spectrometry-based Proteomics Using Q Exactive, a High-performance Benchtop Quadrupole Orbitrap Mass Spectrometer. *Mol. Cell. Proteomics* **10**, (2011).
91. Habib, M. *et al.* DNA Global Hypomethylation in EBV-Transformed Interphase Nuclei. *Exp. Cell Res.* **249**, 46–53 (1999).
92. Neri, F. *et al.* Single-Base Resolution Analysis of 5-Formyl and 5-Carboxyl Cytosine Reveals Promoter DNA Methylation Dynamics. *Cell Rep.* **10**, 674–683 (2015).
93. Su, Y.-T. *et al.* Monoubiquitination of Filamin B Regulates Vascular Endothelial Growth Factor-Mediated Trafficking of Histone Deacetylase 7. *Mol. Cell. Biol.* **33**, 1546–1560 (2013).
94. He, J. *et al.* The AAA+ protein ATAD3 has displacement loop binding properties and is involved in mitochondrial nucleoid organization. *J. Cell Biol.* **176**, 141–146 (2007).
95. Kang, Y. *et al.* Tim29 is a novel subunit of the human TIM22 translocase and is involved in complex assembly and stability. *eLife* **5**, (2016).