



# Vitamin B<sub>12</sub> is a limiting factor for induced cellular plasticity and tissue repair

Received: 11 October 2022

Accepted: 27 September 2023

Published online: 16 November 2023

Check for updates

Marta Kovatcheva <sup>1</sup>✉, Elena Melendez <sup>1</sup>, Dafni Chondronasiou<sup>1</sup>, Federico Pietrocola<sup>1,2</sup>, Raquel Bernad <sup>1</sup>, Adrià Caballe <sup>1</sup>, Alexandra Junza<sup>3,4</sup>, Jordi Capellades<sup>3,5</sup>, Adrián Holguín-Horcajo <sup>6,7</sup>, Neus Prats<sup>1</sup>, Sylvère Durand<sup>8,9</sup>, Meritxell Rovira <sup>6,7</sup>, Oscar Yanes <sup>3,4</sup>, Camille Stephan-Otto Attolini <sup>1</sup>, Guido Kroemer <sup>8,9,10</sup> & Manuel Serrano <sup>1,11,12</sup>✉

Transient reprogramming by the expression of OCT4, SOX2, KLF4 and MYC (OSKM) is a therapeutic strategy for tissue regeneration and rejuvenation, but little is known about its metabolic requirements. Here we show that OSKM reprogramming in mice causes a global depletion of vitamin B<sub>12</sub> and molecular hallmarks of methionine starvation. Supplementation with vitamin B<sub>12</sub> increases the efficiency of reprogramming both in mice and in cultured cells, the latter indicating a cell-intrinsic effect. We show that the epigenetic mark H3K36me3, which prevents illegitimate initiation of transcription outside promoters (cryptic transcription), is sensitive to vitamin B<sub>12</sub> levels, providing evidence for a link between B<sub>12</sub> levels, H3K36 methylation, transcriptional fidelity and efficient reprogramming. Vitamin B<sub>12</sub> supplementation also accelerates tissue repair in a model of ulcerative colitis. We conclude that vitamin B<sub>12</sub>, through its key role in one-carbon metabolism and epigenetic dynamics, improves the efficiency of *in vivo* reprogramming and tissue repair.

Cellular reprogramming consists of the loss of differentiated cell identity followed by the acquisition of embryonic stem pluripotency, which can be achieved by the simultaneous expression of the transcription factors OCT4, SOX2, KLF4 and MYC (OSKM; in mice encoded by *Pou5f1*, *Sox2*, *Klf4* and *Myc*, respectively)<sup>1</sup>. During recent years, it has become evident that this process involves intermediate states in which cells acquire various degrees of plasticity and differentiation potential, which may have broad implications in regenerative medicine and organ repair (reviewed in ref. 2). Continuous expression of OSKM in

mice can recapitulate full reprogramming to pluripotency, a process that culminates with the generation of teratomas<sup>3</sup>. Interestingly, transient expression of OSKM leads to molecular and physiological features of rejuvenation, including an enhanced capacity for tissue regeneration<sup>4–9</sup>. Nevertheless, *in vivo* reprogramming via OSKM remains a poorly understood process, with low efficiency and high risks, including teratoma and cancer development<sup>3,10,11</sup>. Thus, we sought to unravel new molecular mechanisms of *in vivo* reprogramming that could be harnessed to manipulate cell plasticity and tissue repair.

<sup>1</sup>Institute for Research in Biomedicine (IRB Barcelona), Barcelona Institute of Science and Technology (BIST), Barcelona, Spain. <sup>2</sup>Department of Biosciences and Nutrition, Karolinska Institutet, Huddinge, Sweden. <sup>3</sup>Universitat Rovira i Virgili, Department of Electronic Engineering, IISPV, Tarragona, Spain. <sup>4</sup>CIBER de Diabetes y Enfermedades Metabólicas Asociadas (CIBERDEM), Instituto de Salud Carlos III, Madrid, Spain. <sup>5</sup>Institut d'Investigació Sanitària Pere Virgili (IISPV), Metabolomics Platform, Reus, Spain. <sup>6</sup>Department of Physiological Science, School of Medicine, Universitat de Barcelona (UB), L'Hospitalet de Llobregat, Spain. <sup>7</sup>Pancreas Regeneration: Pancreatic Progenitors and Their Niche Group, Regenerative Medicine Program, Institut d'Investigació Biomèdica de Bellvitge (IDIBELL), L'Hospitalet de Llobregat, Spain. <sup>8</sup>Metabolomics and Cell Biology Platforms UMS AMMICa/UMR 1138, Institut Gustave Roussy, Villejuif, France. <sup>9</sup>Equipe labellisée par la Ligue contre le cancer, Centre de Recherche des Cordeliers, Inserm U1138, Université de Paris, Sorbonne Université, Institut Universitaire de France, Paris, France. <sup>10</sup>Institut du Cancer Paris CARPEM, Department of Biology, Hôpital Européen Georges Pompidou, AP-HP, Paris, France. <sup>11</sup>Catalan Institution for Research and Advanced Studies (ICREA), Barcelona, Spain. <sup>12</sup>Altos Labs, Cambridge Institute of Science, Cambridge, UK. ✉e-mail: [marta.kovatcheva@irbbarcelona.org](mailto:marta.kovatcheva@irbbarcelona.org); [mserrano@altoslabs.com](mailto:mserrano@altoslabs.com)

Given the unique metabolic requirements of *in vitro* reprogramming<sup>12,13</sup>, we hypothesized that unique metabolic requirements may also operate during *in vivo* reprogramming. As a new approach, we considered the gut microbiota as a commensal community in metabolic equilibrium with its host. Indeed, the microbiota is sensitive to perturbations in host physiology, capable of adapting and rewiring itself based on nutrient availability and depletion<sup>14</sup>, a process known as the host–gut microbiota metabolic interaction<sup>15</sup>. We reasoned that analysis and manipulation of the microbiota could provide new insights into the metabolic requirements of *in vivo* reprogramming.

## In vivo reprogramming is dependent on the microbiota

To study modulators of *in vivo* reprogramming, we used a previously described mouse model in which doxycycline drives systemic, inducible OSKM expression<sup>3,9,16,17</sup>. On a short timescale (7 days), OSKM induction causes focal regions of abnormal tissue architecture, correlating with the appearance of rare NANOG-positive cells (a marker of embryonic pluripotency) predominantly in the pancreas, colon and stomach<sup>3</sup>. We first asked whether the microbiota was important for *in vivo* reprogramming by disrupting it with a commonly used, broad-spectrum cocktail of antibiotics (ABX): ampicillin, metronidazole, neomycin and vancomycin<sup>18</sup>. We administered ABX for 3 weeks before and during the 7 days of OSKM induction (Fig. 1a). We noted that mice treated with ABX had very low levels of serum doxycycline (Extended Data Fig. 1a), therefore precluding the induction of OSKM in organs beyond the gastrointestinal tract (Extended Data Fig. 1b). Nevertheless, doxycycline efficiently induced OSKM in the colon and stomach in the presence of ABX (Extended Data Fig. 1b). Strikingly, despite strong transgene induction, reprogramming was significantly reduced in the colon and stomach of ABX-treated mice (Fig. 1b and Extended Data Fig. 1c). Reduction in reprogramming was also reflected in the reduced abundance of SCA1-positive and KRT14-positive cells (Fig. 1b and Extended Data Fig. 1c), markers of early and advanced stages of intermediate *in vivo* reprogramming, respectively<sup>19</sup>. Consistent with low levels of reprogramming, ABX-treated mice lost significantly less weight than mice with normal levels of reprogramming (Extended Data Fig. 1d). These results indicate that the microbiota is critical for the successful reprogramming of tissues *in vivo*.

## In vivo reprogramming causes microbial dysbiosis

Given the profound impact that disruption of the microbiota had on *in vivo* reprogramming, we reasoned that a functional analysis of microbial changes during this process could illuminate previously unknown requirements for reprogramming. To this end, we isolated bacterial DNA from paired stool samples of both OSKM-expressing mice and

wild-type (WT) littermate control mice before and after 7 days of doxycycline treatment, and performed shotgun metagenome sequencing<sup>20</sup> (Extended Data Fig. 2a–c and Supplementary Tables 1 and 2). In both WT and OSKM mice, the microbial diversity as measured by the Shannon index decreased following 7 days of doxycycline treatment, with the most profound loss of diversity occurring in reprogrammed mice (Extended Data Fig. 2a). At a genus level, reprogrammed mice were characterized by a relative expansion of *Chlamydia*, *Bacteroides* and *Alistipes* spp. and a relative contraction of *Muribaculaceae* spp. (Extended Data Fig. 2b). *Muribaculaceae* have been reported to contract during inflammatory colonic injury<sup>21</sup>, which shares features with *in vivo* reprogramming including inflammation and loss of differentiated cell identity<sup>22</sup>. *Alistipes* on the other hand, have been reported to promote colonic interleukin (IL)-6 production<sup>23</sup>, which is an important mediator of *in vivo* reprogramming<sup>16</sup>.

## In vivo reprogramming reduces systemic vitamin B<sub>12</sub> levels

Our whole-genome approach allowed us to investigate changes not only in bacterial species abundance, but also in gene composition and ontology groups, which could uncover pathways relevant to reprogramming. Remarkably, we found that microbial gene modules related to the biosynthesis and metabolism of cobalamin (vitamin B<sub>12</sub>) dominated the bacterial Gene Ontology (GO) groups altered during reprogramming (Fig. 1c and Supplementary Table 2). Under conditions of disrupted cobalamin bioavailability, competition for vitamins can shift the relative abundance of cobalamin-producing and cobalamin-utilizing bacteria in a process referred to as ‘corrinoid remodelling’<sup>14,24</sup>. We found microbial changes consistent with this phenomenon in reprogramming: the few genera of bacteria able to synthesize B<sub>12</sub> (~20 genera)<sup>25</sup> were generally enriched in OSKM mice after 7 days of doxycycline, with *Proteus*, *Escherichia* and *Salmonella* being most significantly enriched among the B<sub>12</sub> synthesizers (Extended Data Fig. 2c and Supplementary Table 2).

The observed changes in the gut microbiota could be indicative of a systemic deficit in B<sub>12</sub>, affecting not only the microbiota but also the entire physiology of the host. To test this, we examined systemic vitamin B<sub>12</sub> levels in the serum during reprogramming, which were significantly reduced in OSKM mice after 7 days of doxycycline administration (Fig. 1d). The liver is one of the organs with the greatest demand for vitamin B<sub>12</sub> (ref. 26) and, as such, is sensitive to B<sub>12</sub> deficiency<sup>27</sup>. In rodents, this manifests as depletion of phosphatidylcholines (PCs)<sup>28</sup>, which are produced in large quantities by the liver in a B<sub>12</sub>-dependent manner. We saw that PCs were significantly reduced in the serum of reprogrammed mice as compared to WT mice treated with doxycycline (Extended Data Fig. 3a). Importantly, the liver of OSKM mice does not exhibit histological changes after 7 days of doxycycline<sup>9</sup>, making the reduction in PCs unlikely to reflect liver dysfunction as a result of reprogramming. The kidney is another organ that is refractory

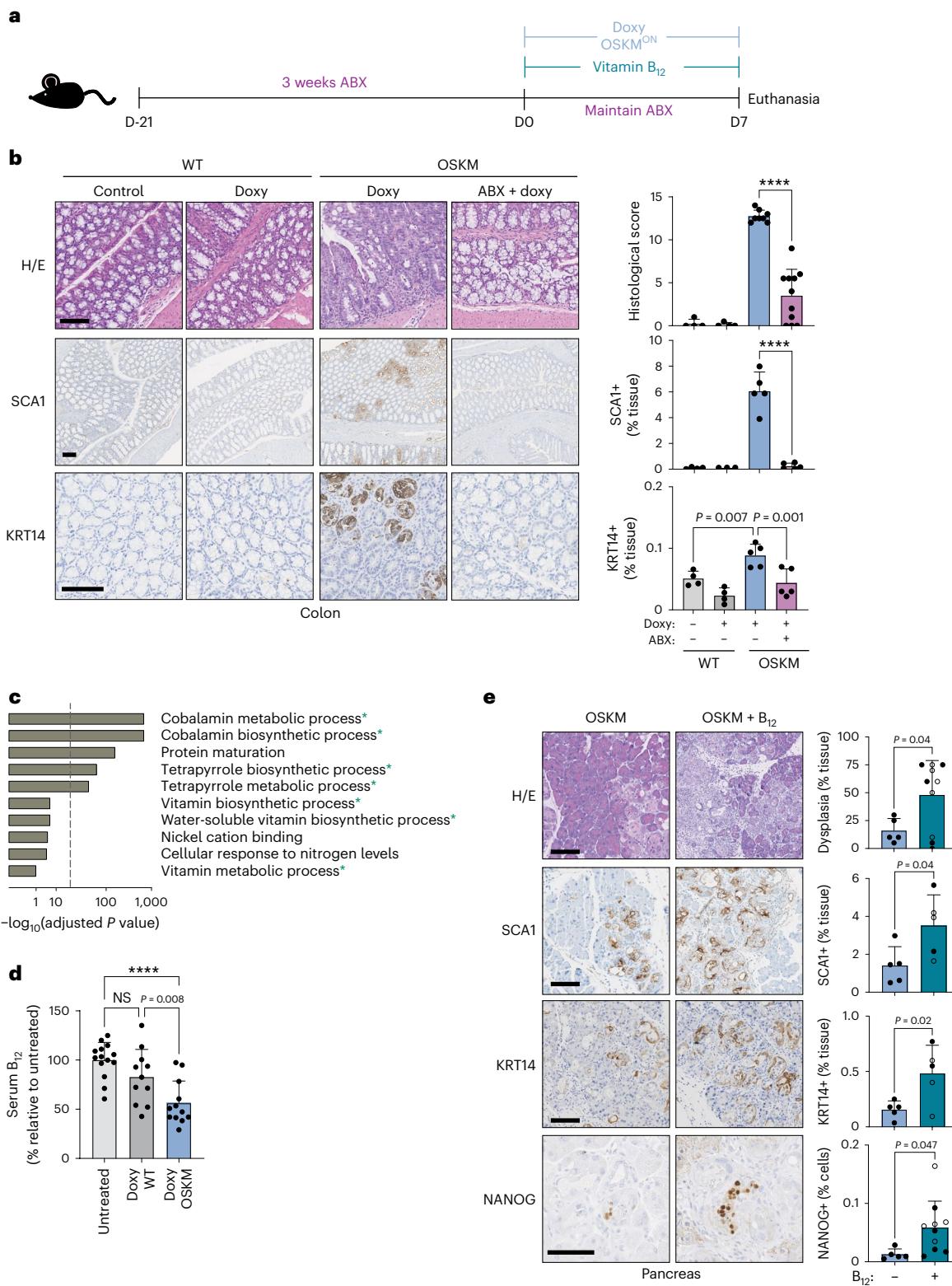
**Fig. 1 | In vivo OSKM reprogramming requires the gut microbiota and is enhanced by vitamin B<sub>12</sub> supplementation.** **a**, Mice were pretreated with an antibiotic cocktail administered in the drinking water for 3 weeks (ABX) before, and during, 7 d of doxycycline administration (doxy), with or without vitamin B<sub>12</sub> supplementation according to the schematic. **b**, Representative histology images and quantification of a blinded histological score, SCA1 staining and KRT14 staining.  $n = 4$  mice (WT; 3 M 1 F),  $n = 8$  (OSKM + doxy; 4 M 4 F),  $n = 11$  (OSKM + doxy + ABX 4 M 7 F); a representative subset of animals was analysed for SCA1 and KRT14. Scale bar, 100  $\mu$ m. **c**, GO pathway analysis of differentially abundant microbial gene signatures in the metagenome sequencing of stool samples. Changes in microbial gene abundance between day 7 and day 0 were compared in a subset of WT ( $n = 4$ ; 2 M 2 F) and OSKM ( $n = 4$ ; 2 M 2 F) mice from **b**. See Supplementary Table 1 for complete gene list. The overlap between GO terms and the 200 most differentially depleted or enriched genes was scored using standard hypergeometric tests and GO terms above a threshold of 30% FDR

are shown (for all GO terms, see Supplementary Table 2). Processes marked with an asterisk are directly related to cobalamin metabolism. Dashed line indicates 5% FDR cut-off. **d**, Serum holoTC (biologically available vitamin B<sub>12</sub>) levels as measured by ADVIA immunoassay in untreated mice or WT and OSKM mice treated with doxycycline for 7 d.  $n = 14$  mice (untreated; 6 M 8 F),  $n = 11$  (doxy WT doxy; 7 M 4 F),  $n = 12$  (doxy OSKM doxy; 6 M 6 F). **e**, OSKM mice received vitamin B<sub>12</sub> supplementation co-administered with doxycycline as indicated and representative images and quantification are shown for the indicated markers in the pancreas. Mice marked by an open circle received both B<sub>12</sub> and folate (B<sub>9</sub>) supplementation (not significant (NS) difference for B<sub>12</sub> versus B<sub>12</sub> + folate; see text for details).  $n = 5$  mice (OSKM; 2 M 3 F),  $n = 10$  mice (OSKM + B<sub>12</sub>; 4 M 6 F); a representative subset of  $n = 5$  animals per group was analysed for SCA1 and KRT14. Scale bars, 100  $\mu$ m. Bar graphs represent the average  $\pm$  s.d.;  $****P < 0.0001$  by two-tailed Student’s *t*-test.

to reprogramming in our mouse model<sup>3</sup>; however, we did observe a significant depletion of vitamin B<sub>12</sub> from the proximal tubules during reprogramming (Extended Data Fig. 3b). The kidney is the primary site of B<sub>12</sub> concentration and storage in rodents, from where it is released for use by other organs upon systemic deficiency<sup>27,29–31</sup>. Collectively, these results suggest that vitamin B<sub>12</sub> becomes systemically depleted during in vivo reprogramming, affecting both the colonic microbiota and the host.

## Vitamin B<sub>12</sub> supplementation improves in vivo reprogramming

Given the systemic reduction of vitamin B<sub>12</sub> during in vivo reprogramming, we hypothesized that B<sub>12</sub> supplementation could enhance reprogramming under normal conditions (that is, in the absence of ABX). Indeed, vitamin B<sub>12</sub> supplementation significantly improved in vivo reprogramming in the pancreas, colon and stomach, as evaluated by the extent of histological dysplasia and SCA1 or KRT14 levels (Fig. 1e)



and Extended Data Fig. 3c–e).  $B_{12}$  also increased the number of NANOG<sup>+</sup> cells, a marker of full pluripotency, in the pancreas (Fig. 1e and Extended Data Fig. 3f).  $B_{12}$  administration did not affect transgene expression (Extended Data Fig. 3g). Even after  $B_{12}$  supplementation, we could not detect histological evidence of reprogramming in the kidney (Extended Data Fig. 3h). However, we did observe a significant increase of vitamin  $B_{12}$  stores within the kidney after supplementation (Extended Data Fig. 3b), indicating that  $B_{12}$  absorption, distribution and storage were occurring normally in the reprogrammed mice.

We also wondered if  $B_{12}$  supplementation could rescue the reprogramming defect of ABX-treated mice. Interestingly,  $B_{12}$  supplementation was able to partially rescue reprogramming in the colon (Extended Data Fig. 3c–e). This supports the concept that an important role of the microbiota during murine reprogramming is to increase the dietary supply of  $B_{12}$  through coprophagy. Another B vitamin that is partly supplied by the microbiota in rodents and humans is vitamin  $B_9$  (folate)<sup>32</sup>, which is functionally related to  $B_{12}$  (ref. 33). However, co-supplementation of  $B_{12}$  and  $B_9$  was indistinguishable from  $B_{12}$  alone (Fig. 1e and Extended Data Fig. 3c,f,g). Collectively, these results demonstrate that vitamin  $B_{12}$  is a limiting factor for *in vivo* reprogramming.

## One-carbon metabolism drives vitamin $B_{12}$ demand during reprogramming

In both humans and mice, vitamin  $B_{12}$  is used as a cofactor by only two enzymes: methionine synthase (MS) and methylmalonyl-CoA mutase (MUT). MS uses  $B_{12}$  as a cofactor to regenerate methionine (Met) from homocysteine (Hcy), forming an integral part of one-carbon (1C) metabolism (Fig. 2a). Met is used to synthesize S-adenosylmethionine (SAM), the universal methyl donor for all methylation reactions<sup>33</sup>. The nuclear-encoded mitochondrial enzyme MUT uses  $B_{12}$  as a cofactor for the catabolism of branched-chain amino acids via isomerization of methylmalonyl-CoA to succinyl-CoA, for use in the tricarboxylic acid cycle (Extended Data Fig. 4a)<sup>26</sup>.

As a first approach to investigate  $B_{12}$ -dependent metabolism during *in vivo* reprogramming, we performed untargeted serum metabolomics. The metabolic pathway with the strongest (by pathway impact) and most significant changes during reprogramming was 'glycine (Gly), serine (Ser), threonine (Thr) metabolism' (Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway map00260; Fig. 2b and Supplementary Table 3). These three amino acids, together with Met, feed 1C metabolism (Fig. 2a). Notably, Thr is known to be critical for the generation of SAM during *in vitro* reprogramming and in the maintenance of pluripotent cells<sup>12,34,35</sup>. In OSKM as compared to WT mice, we saw significant depletion of all four of these amino acids (Gly, Ser, Thr and Met), concomitant with an increase in SAM and an increase in the SAM/Met ratio, progressively over the course of reprogramming (Fig. 2c and Extended Data Fig. 4). The increased SAM/Met ratio was indicative of methylation cycle activation, which occurs during *in vitro* reprogramming<sup>13</sup> and in cultured pluripotent cells<sup>12,34,35</sup>. Neither Hcy nor S-adenosyl-L-homocysteine (SAH), two important 1C metabolites, were detected via untargeted metabolomics, so we performed a separate, targeted serum metabolomic analysis. We found that the SAM/SAH ratio, known as the 'methylation index' because it indicates the methylation capacity of an organism<sup>36</sup>, was significantly increased during *in vivo* reprogramming (Extended Data Fig. 4c). We did not observe changes in Hcy (Extended Data Fig. 4c). Although Hcy accumulation can occur clinically as a result of  $B_{12}$  insufficiency in humans<sup>26</sup>, we observed that induced pluripotent stem (iPS) cells upregulate the expression of genes encoding two main consumers of Hcy: cystathione beta-synthase (CBS, which initiates the transsulfuration pathway) and betaine-homocysteine methyltransferase (BHMT, which synthesizes Met using Hcy and betaine in an MS-independent manner; Fig. 2a and Extended Data Fig. 4d). In support of this, we found that serum betaine levels were

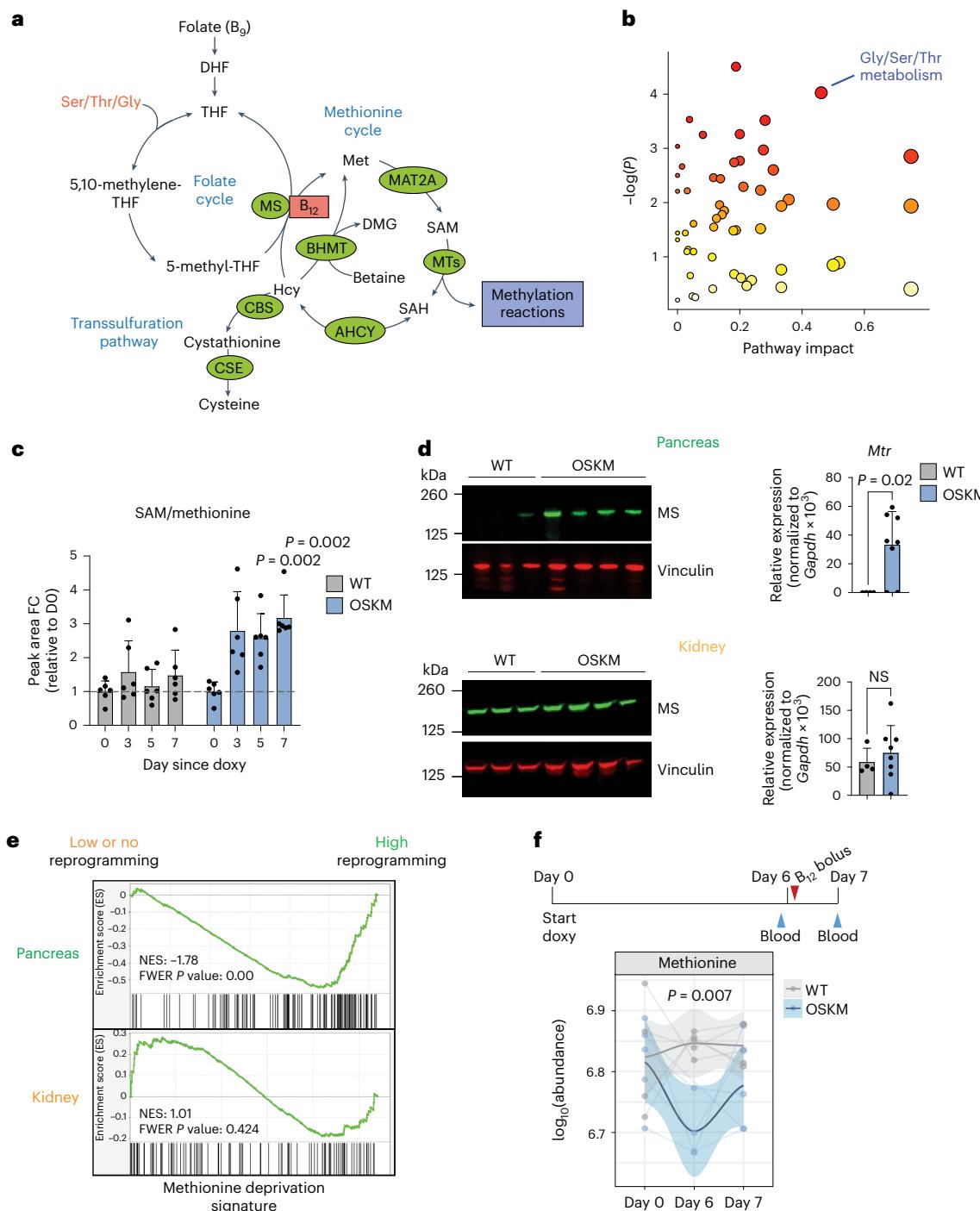
significantly decreased during *in vivo* reprogramming (Extended Data Fig. 4e), which may serve as an additional or alternate source of Met generation.

On the other hand, methylmalonic acid (MMA), the substrate of MUT, showed no significant differences between OSKM mice and WT littermate controls treated with doxycycline (Extended Data Fig. 4b). The MMA levels, along with decreases in *Mmut* expression in several organs (Extended Data Fig. 4f), suggested that MUT's enzymatic activity does not become limiting during reprogramming. It is important to note that while serum accumulation of MMA and Hcy are sensitive biomarkers of vitamin  $B_{12}$  deficiency in humans<sup>26</sup>, this is not the case in mice<sup>37,38</sup>.

We next asked whether the metabolic alterations related to 1C metabolism in the serum were caused by changes specifically within those tissues undergoing reprogramming. We first examined expression of MS, which was upregulated at both the protein and RNA levels in the pancreas, colon and stomach of mice undergoing reprogramming, but not in the kidney (Fig. 2d and Extended Data Fig. 5a). Expression of *Cd320*, the main receptor for cellular uptake of  $B_{12}$  (ref. 26), was significantly upregulated in the reprogramming pancreas (Extended Data Fig. 5b). We also examined a gene signature of Met deprivation<sup>39</sup> by gene-set enrichment analysis (GSEA). As a proof of concept, we tested this signature in previously published bulk RNA-sequencing (RNA-seq) data from *in vitro* OSKM reprogramming<sup>40</sup>, and found it was significantly enriched in iPS cells as compared to the mouse embryonic fibroblasts (MEFs) from which the iPS cells were derived (Extended Data Fig. 5c). This is consistent with the fact that ESCs require high Met levels for self-renewal and survival<sup>34</sup>. We then tested this Met deprivation signature in previously published bulk RNA-seq data from *in vivo* reprogramming<sup>16</sup>. At day 7 of doxycycline treatment, Met deprivation was significantly enriched in the pancreas of mice with high levels of reprogramming as compared to mice genetically resistant to reprogramming; in contrast, there was no enrichment of this pathway in the kidney (Fig. 2e). To further validate these results, we analysed a subset of genes from the Met deprivation signature, which were among the most highly enriched by GSEA, in reprogramming tissues by quantitative PCR with reverse transcription (RT-qPCR). A total of 11 genes were assessed and, interestingly, they were broadly upregulated in the pancreas, colon and stomach—but not in the kidney—of mice expressing OSKM (Extended Data Fig. 5d). A subset of these genes was basally high in the colon due to their importance in the stem cell compartment. Importantly,  $B_{12}$  supplementation generally relieved the upregulation of these genes, in support of the idea that limiting  $B_{12}$  levels were driving the Met restriction *in vivo*. In the kidney, these genes were induced with  $B_{12}$  supplementation, likely a feedback response caused by the large influx of  $B_{12}$  storage into the kidney after supplementation (Extended Data Fig. 3b).

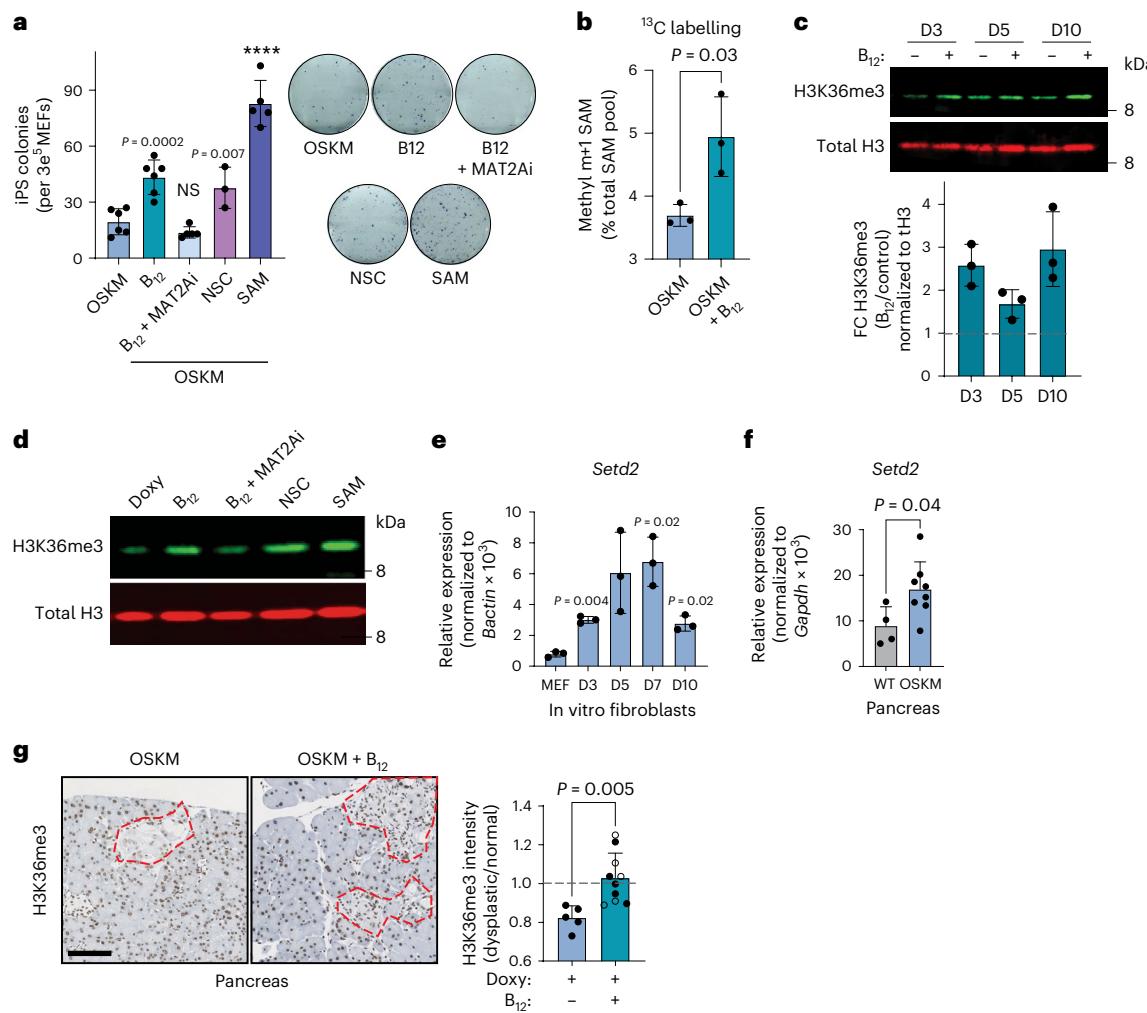
Finally, to ensure that the depletion of serum  $B_{12}$  levels and its associated low levels of Met were not simply caused by defective oral uptake due to the reprogramming of several digestive organs, we administered a large bolus of vitamin  $B_{12}$  (5  $\mu$ g per mouse, 100 times the recommended dietary allowance<sup>41</sup>) on day 6 of doxycycline treatment, 1 day before euthanasia (Fig. 2f). Mice expressing OSKM had significantly higher levels of serum  $B_{12}$  than WT mice following the bolus (Extended Data Fig. 5e,f), which is known to occur in  $B_{12}$ -deficient rodents<sup>27,31</sup>, and further indicated that reprogramming does not compromise oral bioavailability of  $B_{12}$ . Strikingly, the bolus rescued the depletion of Met levels in the serum (Fig. 2f).

Together, these data suggest that tissues undergoing reprogramming are the ones driving the depletion of serum factors that feed 1C metabolism, including Met, serine, glycine, threonine, betaine and vitamin  $B_{12}$ .  $B_{12}$  becomes a limiting factor, as shown by the effects of  $B_{12}$  supplementation in rescuing Met deprivation and promoting reprogramming.



**Fig. 2 | Tissues undergoing in vivo reprogramming exhibit an increased demand of 1C metabolism.** **a**, Summary of the mammalian folate and methionine cycles (1C metabolism) and the transsulfuration pathway. Enzymes are marked in green. Coenzyme vitamin B<sub>12</sub> is marked in red. DHF, dihydrofolate; THF, tetrahydrofolate; MTs, methyltransferases; ACHY, adenosylhomocysteinase; CSE, cystathionine gamma-lyase; DMG, dimethylglycine; Ser, serine; Thr, threonine; Gly, glycine. Figure adapted from ref. 33, Springer Nature Limited. **b**, Changes in metabolic pathways during reprogramming. MetaboAnalyst 4.0<sup>91</sup> was used to assess the annotated metabolites identified in the serum of paired OSKM mice ( $n = 6$ ; 3 M 3 F) at day 5 versus day 0 of doxycycline treatment (serum was collected repeatedly from the same mice). Colour gradient from white to red indicates the  $P$  value; red is most significant. Gly/Ser/Thr metabolism (KEGG map00260) is highlighted. See Supplementary Table 3 for all metabolites, pathways and scores. **c**, Fold change (FC) of SAM/Met ratio detected by mass spectrometry from **b** on the indicated

days.  $P$  values represent significant difference between OSKM and WT mice. **d**, Levels of MS (encoded by *Mtr*) by immunoblot and RT-qPCR in the pancreas (upper) and kidney (lower) from WT ( $n = 4$ ; 3 M 1 F) and OSKM ( $n = 8$ ; 4 M 4 F) mice treated with doxycycline for 7 d. Representative mice are shown in the immunoblot. **e**, Previously published RNA-seq data<sup>16</sup> from the pancreas (highly prone to reprogramming; green) and kidney (refractory to reprogramming; orange) of OSKM-*Cdk2na*<sup>-/-</sup> (low or absent reprogramming) and OSKM-*Tp53*<sup>-/-</sup> (high reprogramming) mice were used to perform GSEA against a published signature (MsigDB: M13537) of Met deprivation<sup>39</sup>. **f**, WT and OSKM mice ( $n = 5$  per group; 5 M) were treated with doxycycline and a bolus of vitamin B<sub>12</sub> as shown in the schematic. Met levels were measured in the indicated serum samples by mass spectrometry. Only  $n = 4$  WT (day 0) and OSKM (day 6) mice are represented, as the blood volume was insufficient. Welch's two-sample *t*-test was used to evaluate differences between groups on day 6. Bar graphs represent the average  $\pm$  s.d.;  $P$  values determined by a two-tailed Student's *t*-test.



**Fig. 3 | Vitamin B<sub>12</sub> supplementation enhances H3K36me3 and has a cell-autonomous role in reprogramming.** **a**, In vitro reprogramming of MEFs with doxycycline-inducible OSKM for 10 d in the presence of doxycycline (OSKM) and/or vitamin B<sub>12</sub>, MAT2Ai, KDM4A/KDM4B inhibition (NSC), or SAM as indicated, cultured in KSR. iPS cell colonies were quantified by alkaline phosphatase staining (left) and representative images are shown (right). Each data point represents MEFs generated from an independent embryo ( $n = 6$  OSKM, B<sub>12</sub>;  $n = 5$  B<sub>12</sub> + MAT2Ai, SAM;  $n = 3$  NSC). **b**, Fraction of total intracellular SAM <sup>13</sup>C-labelled at the methyl m + 1 position (Extended Data Fig. 6d), using <sup>13</sup>C-serine as a precursor. Labelling was initiated at  $t = 72$  h for 6 h. Data from  $n = 3$  independent MEFs are shown. **c**, H3K36me3 dynamics during in vitro reprogramming. MEFs were treated with doxycycline with or without vitamin B<sub>12</sub> as indicated. A representative immunoblot and quantification from  $n = 3$  independent MEFs are shown. **d**, H3K36me3 level correlates with reprogramming efficiency in vitro. MEFs were treated as indicated and H3K36me3 levels were probed in

histone extracts at day 3 after doxycycline treatment. A representative blot from  $n = 2$  independent MEFs with similar results is shown. **e,f**, Expression of the H3K36 trimethyl-transferase *Setd2* during in vitro (**e**) and in vivo pancreatic (**f**) reprogramming. In **e**, *P* values represent significant change from the parental MEFs ( $n = 3$  MEFs). In **f**, samples were collected from WT ( $n = 4$ ; 3 M1 F) or OSKM ( $n = 8$ ; 4 M4 F) mice after 7 d of doxycycline treatment. **g**, H3K36me3 during in vivo reprogramming. Pancreatic tissue from OSKM mice treated with doxy ( $n = 5$ ; 2 M3 F) or doxy + B<sub>12</sub> ( $n = 10$ ; 4 M6 F; Fig. 1) was stained for H3K36me3. Representative images are shown with dysplastic foci demarcated by red dashed lines. The mean nuclear optical density of the H3K36me3 stain is expressed as the ratio between the dysplastic region and the adjacent normal tissue for each mouse. Scale bar, 100  $\mu$ m. Mice that received folate in addition to B<sub>12</sub> are represented by open points. Graphs represent the average  $\pm$  s.d.;  $****P < 0.0001$  by two-tailed Student's *t*-test.

## Vitamin B<sub>12</sub> plays a cell-autonomous role in reprogramming

In vivo, vitamin B<sub>12</sub> deficiency yields a complex phenotype because it impacts multiple cellular processes and organ functions<sup>26</sup>. Therefore, we asked whether the effect of B<sub>12</sub> on reprogramming could be recapitulated in vitro. We first observed, in a previously published bulk RNA-seq dataset of MEFs undergoing in vitro reprogramming<sup>40</sup>, that *Mtr* and *Cd320* were upregulated soon after OSKM induction, remaining high during the early and middle phases of reprogramming, ultimately stabilizing to levels above those measured in MEFs (Extended Data Fig. 6a). This suggests that during in vitro reprogramming there is also a high demand of B<sub>12</sub> and Met, which we

explored further using pharmacological manipulation of the Met cycle. The addition of B<sub>12</sub> significantly increased the efficiency of iPS cell colony formation (Fig. 3a), recapitulating our observations in vivo and demonstrating a cell-intrinsic role for vitamin B<sub>12</sub> in reprogramming. Of note, B<sub>12</sub> supplementation increased the number of successfully formed iPS cell colonies without an obvious effect on the rate of colony formation (Extended Data Fig. 6b), and the improved efficiency was also observed in a doxycycline-free, retroviral-based reprogramming system in WT MEFs (Extended Data Fig. 6c). The B<sub>12</sub>-mediated increase in reprogramming efficiency was prevented by concomitant treatment with a methionine adenosyltransferase 2A inhibitor (MAT2Ai; Fig. 3a); MAT2A is the enzyme immediately downstream of MS, which

converts Met into SAM (Fig. 2a). Moreover, directly supplementing only SAM at a high concentration<sup>34</sup> during reprogramming significantly improved the efficiency of the process, even beyond that of B<sub>12</sub> itself (Fig. 3a).

To directly assess the contribution of B<sub>12</sub> supplementation to SAM generation, we performed stable isotope labelling (SIL) with <sup>13</sup>C-labelled serine. Serine significantly decreases in the serum during in vivo reprogramming (Extended Data Fig. 4b) and can contribute as a methyl donor to 1C metabolism (Extended Data Fig. 6d). We began SIL 72 h after OSKM induction, well before iPS cell colonies are formed. Culturing cells with <sup>13</sup>C-serine did not affect the reprogramming efficiency, nor the capacity of B<sub>12</sub> to enhance reprogramming (Extended Data Fig. 6e). Importantly, B<sub>12</sub> significantly stimulated the incorporation of the <sup>13</sup>C-methyl donor group (m + 1) from serine into SAM (Fig. 3b). Collectively, these data demonstrate that B<sub>12</sub> operates in a cell-intrinsic manner during in vitro reprogramming and that it is a limiting factor for SAM generation and successful reprogramming in vitro.

### H3K36me3 is enhanced by vitamin B<sub>12</sub> during reprogramming

Among the SAM-dependent cellular processes, histone methylation is one of the highest consumers of SAM<sup>42,43</sup>. Histone modifications play a major role in determining cellular identity changes during reprogramming<sup>44,45</sup>, and SAM levels are critical for both mouse<sup>42</sup> and human<sup>34</sup> embryonic stem cell maintenance through appropriate histone modifications. Moreover, some histone modifying enzymes have an affinity for SAM that is in the physiological range of this metabolite and, therefore, their enzymatic output is sensitive to changes in SAM levels<sup>42</sup>. This led us to hypothesize that vitamin B<sub>12</sub>, through SAM, could promote changes in histone methylation during reprogramming. To identify B<sub>12</sub>-mediated histone methylation changes in an unbiased manner, we performed a histone H3 modification array at both early (day 3) and late (day 10) stages of the reprogramming process in vitro. While there were many epigenetic changes associated with vitamin B<sub>12</sub> supplementation, trimethylated histone H3 Lys 36 (H3K36me3) was the histone modification that increased the most, and the only one that increased at both day 3 and day 10 of OSKM induction in MEFs (Extended Data Fig. 6f). We used immunoblotting to confirm the results of the histone array and found that H3K36me3 was enhanced by the addition of vitamin B<sub>12</sub> during reprogramming at all time points tested (Fig. 3c). B<sub>12</sub> also enhanced H3K36me3 in MEFs undergoing reprogramming via retroviral OSKM, in the absence of doxycycline (Extended Data Fig. 6c). H3K36me3 levels correlated with the ability of 1C modulators to promote reprogramming efficiency: MAT2Ai reduced the amount of H3K36me3, while SAM enhanced it (Fig. 3d). We also tested NSC636819 (NSC), an inhibitor of the H3K36me3 (and H3K9me3) demethylases KDM4A/KDM4B<sup>46</sup>. In the absence of B<sub>12</sub> supplementation, NSC enhanced both H3K36me3 levels (Fig. 3d) and reprogramming efficiency to the level of cells treated with B<sub>12</sub> alone (Fig. 3a).

We then tested if the link between B<sub>12</sub>, H3K36me3 and reprogramming observed during in vitro reprogramming of MEFs could also be substantiated during in vivo reprogramming. First, we found that the mRNA levels of *Setd2*, the gene encoding the only known mammalian methyltransferase that catalyses H3K36 trimethylation<sup>47</sup>, were significantly increased during both in vitro reprogramming of MEFs (Fig. 3e) and in vivo reprogramming in the pancreas (Fig. 3f). Immunohistochemistry revealed that the regions of the exocrine pancreas undergoing reprogramming (identified by their abnormal architecture) had reduced levels of H3K36me3 as compared to the surrounding tissue (with normal histology; Fig. 3g). Importantly, in vivo supplementation with B<sub>12</sub> restored H3K36me3 levels within dysplastic regions to parity with those of the neighbouring normal tissue (Fig. 3g). Although we did not detect changes in *Setd2* expression in the colon and stomach with reprogramming (Extended Data Fig. 6g), B<sub>12</sub> also significantly increased H3K36me3 levels in these tissues (Extended Data Fig. 6h), indicating that the SETD2 enzymatic activity is sensitive to B<sub>12</sub> levels. Collectively,

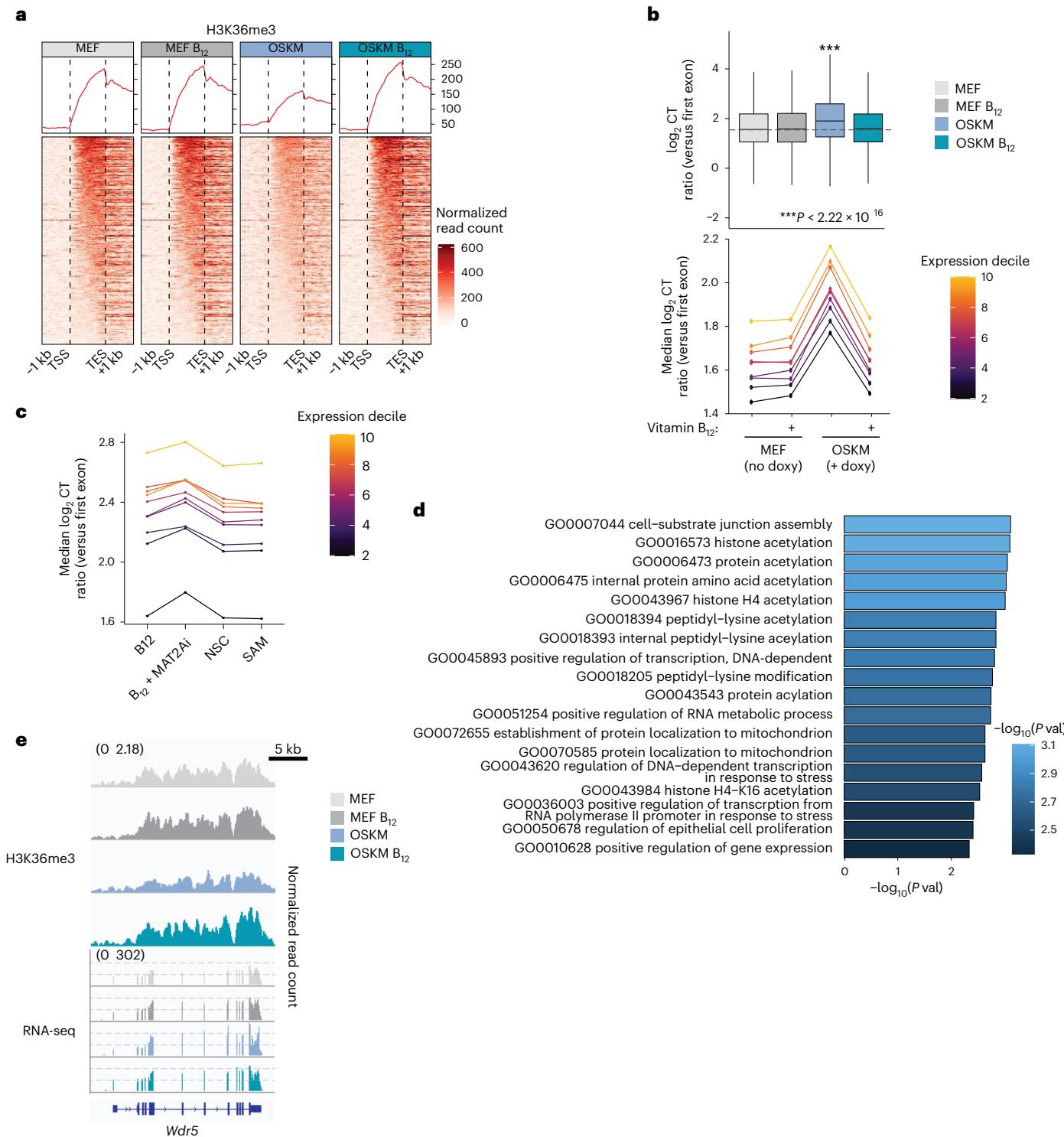
these data strongly suggest that B<sub>12</sub> levels are mechanistically linked to H3K36me3 levels, and that both are positively correlated with the efficiency of reprogramming in vitro and in vivo.

### Vitamin B<sub>12</sub> improves transcriptional fidelity during reprogramming

To address the role of H3K36me3 during reprogramming, we performed (i) chromatin immunoprecipitation followed by sequencing (ChIP-seq) for H3K36me3 and (ii) bulk RNA-seq in MEFs undergoing reprogramming, with and without supplementation of vitamin B<sub>12</sub>. We chose to analyse reprogramming at 72 h because *Setd2* mRNA is already significantly upregulated (Fig. 3e) and B<sub>12</sub> supplementation has increased global H3K36me3 levels (Fig. 3c) at this time point, but iPS cell colonies have not yet formed. After collecting the 72-h samples, we allowed the MEFs from three independent embryos to continue reprogramming up to 10 days, when iPS cells can be seen by light microscopy and stain positive for alkaline phosphatase. We observed that the three MEFs exhibited different reprogramming efficiencies: MEF 1 > MEF 2 > MEF 3 (Extended Data Fig. 7a). This differential efficiency of reprogramming was also reflected in the transcriptomes of these cells, as evaluated by the enrichment of a reprogramming score composed of genes previously reported in MEFs that are ‘poised to reprogram’ at 72 h of OSKM expression<sup>48</sup> (Extended Data Fig. 7b). We focused our subsequent analyses on MEF 1 because of its higher reprogramming efficiency. Consistent with published ChIP data from yeast<sup>49</sup>, mouse<sup>50,51</sup> and human cells<sup>52</sup>, we found the H3K36me3 ChIP signal largely distributed within gene bodies, peaking towards the 3' ends (Fig. 4a and Extended Data Fig. 7c). Interestingly, 72 h of OSKM induction caused a decrease in the amount of H3K36me3 across gene bodies relative to all reads, which was prevented by vitamin B<sub>12</sub> supplementation (Fig. 4a and Extended Data Fig. 7c).

H3K36me3 plays important roles in maintaining transcriptional fidelity<sup>49,52–56</sup>. One such mechanism, active in mammalian stem cells, is to prevent illegitimate transcription initiation outside of promoter regions, at intragenic ‘cryptic’ start sites, particularly within genes that are highly expressed<sup>50,52</sup>. We thus asked whether the amount of cryptic transcription (CT) was changed during reprogramming. CT can be estimated by calculating the ratio of expression from 3' exons (defined as those from the fourth to the penultimate exon) over the expression from the first exon<sup>50</sup> (Supplementary Table 4 and Methods). Consistent with the reduced levels of H3K36me3, global CT significantly increased during OSKM-induced reprogramming as compared to the non-induced MEFs, across genes of all expression levels (Fig. 4b). Remarkably, supplementation with vitamin B<sub>12</sub> during these 72 h of OSKM expression reduced the amount of CT to basal (that is, non-reprogramming) levels (Fig. 4b). This effect was dose dependent with respect to the level of reprogramming in each MEF (Extended Data Fig. 7d). Vitamin B<sub>12</sub> treatment in the non-induced MEFs did not affect CT (Fig. 4b and Extended Data Fig. 7d). These data suggest that the high demand for methylation that occurs during reprogramming causes a deficit in H3K36me3 that results in increased CT, which can be rescued by B<sub>12</sub> supplementation. To further substantiate this concept, we measured reprogramming-associated CT in the presence of pharmacological modulators that aggravate the methylation deficit (MAT2Ai), that alleviate the methylation deficit (SAM), and that stabilize H3K36me3 (NSC). Interestingly, MAT2Ai increased, while SAM and NSC decreased reprogramming-associated CT (Fig. 4c and Extended Data Fig. 7e).

We next asked whether the genes that suffered reprogramming-associated CT had a specific function. We used a hypergeometric test against GO gene-set collections to analyse the function of the top 25% of genes whose CT was induced by OSKM expression, and whose CT could be reduced with vitamin B<sub>12</sub> (Extended Data Fig. 7f). We found that, together with cell-substrate junction assembly, the most highly enriched GO categories were related to histone modification and transcriptional control (Fig. 4d), suggesting that B<sub>12</sub> helps to streamline



**Fig. 4 | Vitamin B<sub>12</sub> supplementation reduces cryptic transcription during in vitro reprogramming.** **a**, H3K36me3 ChIP-seq was performed in MEF1 (Extended Data Fig. 7) at 72 h after the addition of doxycycline and/or vitamin B<sub>12</sub>, as indicated. The normalized ChIP reads in gene bodies relative to all aligned reads are shown for the most highly expressed genes (top tercile). The trace of the ChIP signal is shown, where the y axis represents average coverage across genes. TSS, transcription start site; TES, transcription end site. See also Extended Data Fig. 7c. **b**, The level of CT in MEF1 for all genes for which a score could be calculated is plotted globally, and as a function of gene expression level at  $t = 72$  h, with or without continuous supplementation of vitamin B<sub>12</sub> as indicated. The lower and upper hinges correspond to the first and third quartiles (the 25th and 75th percentiles). The upper and lower whiskers extend from the hinge to the largest and smallest values, respectively, no further than 1.5 times the interquartile range from the hinge. Asterisks represent two-tailed

unpaired Wilcoxon test; no multiple-comparisons adjustment was performed. **c**, OSKM MEFs were cultured with doxycycline and additional compounds for 72 h as indicated and described in Fig. 3a. The median CT ratio for all genes for which a score could be calculated is plotted as a function of gene expression level. Values represent the median from  $n = 2$  independent MEFs; for data from individual MEFs, see Extended Data Fig. 7e. **d**, Functional enrichment analysis was performed on the top 25% of genes whose CT increased between MEF and OSKM conditions, and decreased between OSKM and OSKM + B<sub>12</sub> conditions in MEF1 from **a** and **b**. The top enriched GO biological processes ( $P < 0.005$ ) are shown.  $P$  values were determined by hypergeometric test; no multiple-comparisons adjustment was performed. **e**, H3K36me3 ChIP-seq and RNA-seq tracks from the *Wdr5* gene from MEF1 in the indicated conditions. For all plots, expression quantiles were determined by reads per kilobase per million mapped reads (RPKM) across all samples.

transcription during reprogramming. In support of this, the transcriptional reprogramming score of cells poised to reprogram<sup>48</sup> was improved by B<sub>12</sub> supplementation (Extended Data Fig. 7b). *Wdr5*, a gene whose expression is important for the establishment of pluripotency<sup>57</sup>, was among those genes whose CT was most largely changed, and whose ChIP peaks exemplify these H3K36me3 and CT dynamics during reprogramming (Fig. 4e). Overall, we found that during the early stages of reprogramming, cells undergo a depletion of H3K36me3 along gene bodies that compromises transcriptional fidelity, as reflected by a global increase in CT from expressed genes. Supplementation of vitamin B<sub>12</sub> restores H3K36me3 levels and suppresses CT, improving the reprogramming trajectory of the bulk population.

## Vitamin B<sub>12</sub> promotes tissue repair after DSS-induced colitis

Many processes of tissue injury and repair in adult organs proceed through transient dedifferentiation to a more developmentally primitive state<sup>22,58</sup>. Moreover, transient cell plasticity achieved through OSKM reprogramming has been shown to promote regeneration after acute injury in several organs including the pancreas, muscle<sup>4</sup>, eye<sup>6</sup> and heart<sup>7</sup>. We reasoned that vitamin B<sub>12</sub> administration during an injury repair period may also promote cell plasticity and improve recovery. To address this idea, we used a dextran sodium sulfate (DSS) model of acute ulcerative colitis (Fig. 5a). In this model, successful repair proceeds through natural reprogramming of epithelial cells to a more embryonic-like state, marked by the murine stem cell marker *Ly6a* (encoding SCA1)<sup>22,59,60</sup>. Of note, this marker of tissue repair is also characteristic of the intermediate, plastic state of OSKM reprogramming in vivo in the colon, as well as in the stomach and the pancreas<sup>19</sup>. Interestingly, we found that DSS injury, much like OSKM reprogramming, was associated with a significant depletion of serum vitamin B<sub>12</sub> levels (Fig. 5b). It is important to note that this depletion is unlikely to be attributed to colitis-associated nutrient absorption defects, as B<sub>12</sub> is typically absorbed in the terminal ileum<sup>26</sup>, while DSS primarily affects the distal colon<sup>61</sup>. We asked whether this pathophysiological example of cellular plasticity and B<sub>12</sub> deficiency was also associated with increased CT using a previously published time-course analysis of bulk RNA-seq of the murine colon during DSS injury and repair<sup>62</sup>. We found that colonic CT significantly increased on days 6, 7 and 8 following DSS treatment initiation (Fig. 5c and Extended Data Fig. 8a), concomitant with the transcriptional peak of tissue repair markers like *Ly6a* and *Reg3b*<sup>22</sup> (Extended Data Fig. 8b). We also observed a significant enrichment of the Met deprivation signature on day 6, as compared to uninjured mice (Extended Data Fig. 8c).

### Fig. 5 | Recovery from DSS colitis is improved by OSKM or vitamin B<sub>12</sub>.

**a**, Colitis was induced by DSS, followed by a 48-h pulse of OSKM or continuous administration of vitamin B<sub>12</sub> during recovery, as indicated. MAT2Ai is FIDAS-5, a MAT2A inhibitor, given daily over the indicated time. **b**, Immunoassay of serum holoTC (biologically available vitamin B<sub>12</sub>) in paired mice before (D0) or following (D5) DSS administration ( $n = 15$ ; 15 M); serum was collected repeatedly from the same mice. *P* value was determined by paired two-tailed Student's *t*-test. **c**, CT in mouse colon tissue during DSS injury and recovery was determined from previously published RNA-seq. Data source: [GSE131032](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE131032) (ref. 62). The median CT ratio for all genes for which a score could be calculated is shown. Each time point has 2–3 biological replicates, and symbol shapes indicate the same cage, as reported by the authors. *P* values were compared to the median CT at D0 as computed by two-tailed Student's *t*-test; dashed line represents median CT value at D0. See also Extended Data Fig. 8a. **d**, SCA1 immunohistochemistry in colonic sections on day 9 after DSS with the indicated treatments. Representative images are shown and percentage of SCA1<sup>+</sup> tissue area is quantified. Untreated control ( $n = 6$ ; 3 M 3 F); DSS ( $n = 12$ ; 8 M 4 F); DSS + vitamin B<sub>12</sub> ( $n = 13$ ; 9 M 4 F); DSS + 48 h OSKM ( $n = 7$ ; 7 M); DSS + B<sub>12</sub> + MAT2Ai ( $n = 8$ ; 4 M 4 F). Graphs represent four pooled experiments, each of which had at least  $n = 3$  each DSS and DSS + B<sub>12</sub> controls. *P* values represent the difference as compared to DSS control animals by one-way ANOVA. **e**, Recovery of colonic homeostasis as scored by H&E on

We next asked whether we could enhance or accelerate tissue repair. Following 5 days of DSS administration, we treated recovering mice either with a brief (48 h) pulse of OSKM expression, or with continuous vitamin B<sub>12</sub> supplementation (Fig. 5a). On day 9, 4 days after withdrawal of DSS, we confirmed the previously reported expression of SCA1 in the repairing epithelium, which was further increased with either the OSKM pulse or B<sub>12</sub> supplementation (Fig. 5d). Thus, both approaches could promote tissue reprogramming in the context of acute colitis. Remarkably, both OSKM and B<sub>12</sub> resulted in significantly improved tissue recovery on day 14, as evaluated by blinded histological analysis (Fig. 5e), colon length (Fig. 5f) and mucosal integrity as visualized by PAS staining (Extended Data Fig. 8d). We used fluorescein isothiocyanate-conjugated (FITC) dextran leakage to evaluate functional recovery of intestinal permeability after B<sub>12</sub> administration, and found it was also significantly improved (Extended Data Fig. 8e). Vitamin B<sub>12</sub> and OSKM treatment also accelerated the recovery towards normal homeostatic colonic crypt architecture, as measured by the number of Ki67<sup>+</sup> cells (Extended Data Fig. 8f) and the recovery of *Lgr5* expressing cells (Extended Data Fig. 8g). The increase in SCA1 and the enhanced recovery promoted by B<sub>12</sub> were blunted by the co-administration of a MAT2A inhibitor (FIDAS-5), suggesting that SAM generation is critical for tissue repair (Fig. 5d,e). To extend our findings to a comparable human pathology, we found CT was significantly increased in a previously published bulk RNA-seq dataset<sup>63</sup> of paediatric ulcerative colitis patients as compared to healthy controls (Fig. 5g), suggesting a conserved mechanism in human disease. Thus, echoing our findings during OSKM-driven reprogramming, tissue repair is associated with depletion of B<sub>12</sub> and high levels of CT, and vitamin B<sub>12</sub> supplementation facilitates tissue repair.

## Discussion

In this study, we analysed the murine faecal microbiota during reprogramming to discover that vitamin B<sub>12</sub> is a limiting factor for cellular reprogramming both in vitro and in vivo. We demonstrate that cells undergoing reprogramming experience an elevated demand for vitamin B<sub>12</sub>, associated with increased levels of the B<sub>12</sub>-dependent enzyme MS and increased expression of *Cd320*, which encodes the cellular vitamin B<sub>12</sub> uptake receptor. MS functions to generate Met, which is converted into SAM, the essential 1C donor for all methylation reactions<sup>64</sup>. We report that SAM is central to the improved reprogramming efficiency imparted by vitamin B<sub>12</sub>. One-carbon metabolism and SAM, via its donation of methyl groups, are required for the epigenetic reconfiguration associated with cell reprogramming<sup>45</sup>. Mechanistically, we found that insufficiency of vitamin B<sub>12</sub> during OSKM reprogramming

day 14 after DSS with the indicated treatments (representative images and quantification). Untreated control ( $n = 5$ ; 5 M); DSS ( $n = 17$ ; 11 M 6 F); DSS + vitamin B<sub>12</sub> ( $n = 16$ ; 10 M 6 F); DSS + 48 h OSKM ( $n = 7$ ; 4 M 3 F); DSS + B<sub>12</sub> + MAT2Ai ( $n = 8$ ; 4 M 4 F). Graphs represent five pooled experiments, each of which had at least  $n = 3$  each DSS and DSS + B<sub>12</sub> controls. *P* values represent difference as compared to DSS control animals by one-way ANOVA. **f**, Colon length (caecum to rectum) on day 14 after DSS from mice with the indicated treatments. Untreated control ( $n = 8$ ; 3 M 5 F); DSS ( $n = 13$ ; 9 M 4 F); DSS + vitamin B<sub>12</sub> ( $n = 13$ ; 10 M 3 F); DSS + 48 h OSKM ( $n = 7$ ; 4 M 3 F); DSS + B<sub>12</sub> + MAT2Ai ( $n = 8$ ; 4 M 4 F). *P* values as compared to DSS control animals by one-way ANOVA. **g**, CT in previously published RNA-seq of paediatric human rectal mucosal biopsy samples classified as ulcerative colitis ( $n = 206$ ; 112 M 94 F) or control ( $n = 20$ ; 9 M 11 F). Data source: [GSE109142](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE109142) (ref. 63). Median CT value for all genes for which a score could be calculated is shown, where each dot represents one patient. The lower and upper hinges correspond to the first and third quartiles (the 25th and 75th percentiles), respectively. The upper (lower) whisker extends from the hinge to the largest (smallest) value no further than 1.5 times the interquartile range from the hinge. Data beyond the end of the whiskers are called 'outlying' points and are plotted individually. *P* value calculated by linear model with sex and quantiles as covariates; no multiple-comparisons adjustment was performed. Scale bars, 100  $\mu$ m. Bar graphs represent the average  $\pm$  s.d. \*\*\**P* < 0.0001.

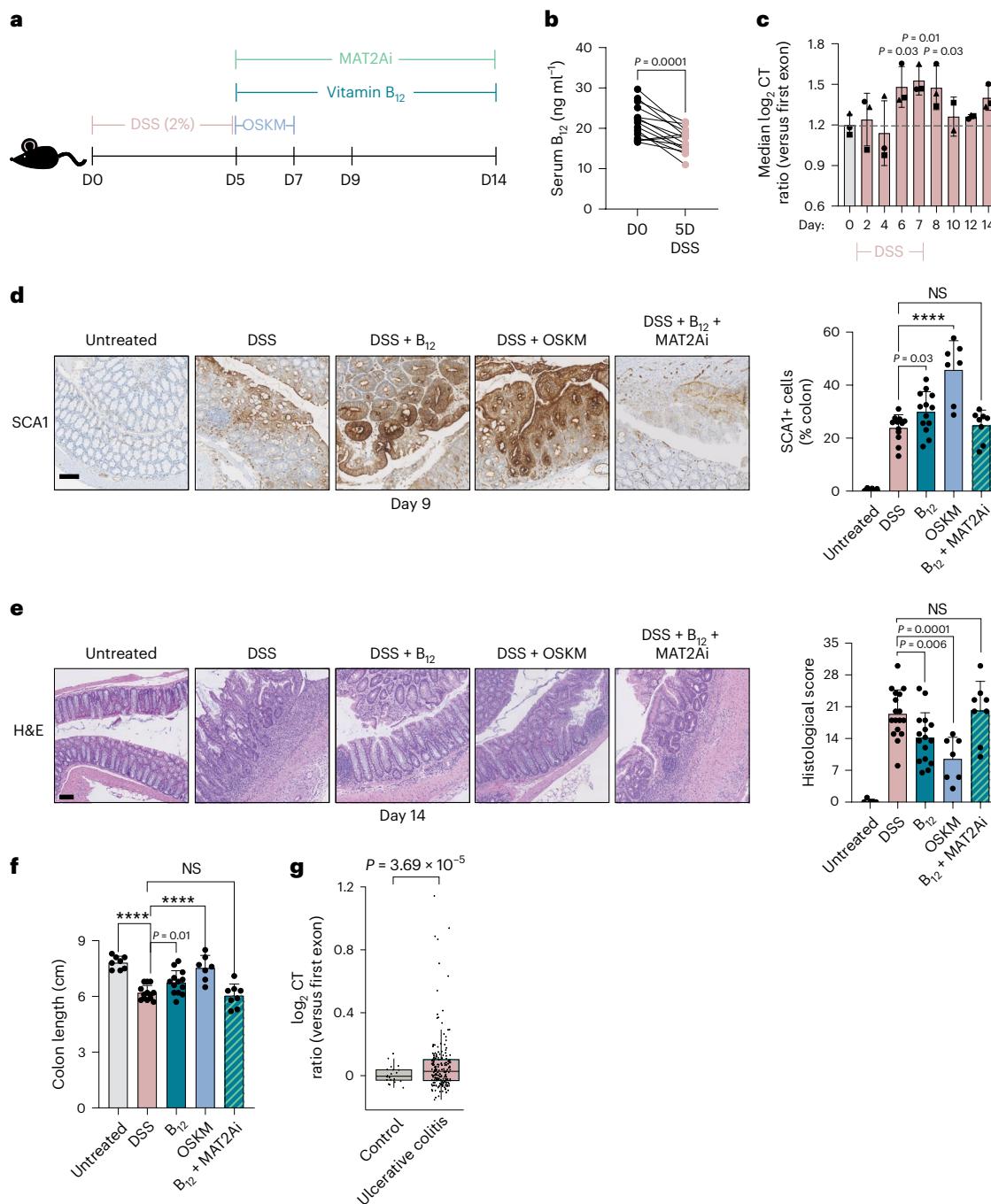
resulted in changes in the bulk levels of many histone methylation marks, including a failure to properly trimethylate H3K36 at transcribed gene bodies, leading to increased levels of CT. CT has negative impacts on cellular and organismal fitness<sup>49,55</sup> and compromises the self-renewal capacity of mammalian stem cells<sup>52</sup>. While we could not directly test the role of H3K36me3 via *Setd2* knockdown because we were technically unable to achieve sustained transcript reduction, supplementation of vitamin B<sub>12</sub> increased H3K36me3 and reduced CT during reprogramming, associated with an enhanced transcriptional reprogramming score and an overall increase in reprogramming efficiency.

It has been estimated that the complete methylation of just 0.1% of nucleosomes would exhaust the entire supply of intracellular SAM<sup>65,66</sup>, rendering the epigenome a critical consumer of SAM levels. SETD2, the H3K36me3 methyltransferase, has a *K<sub>M</sub>* for SAM in the micromolar range<sup>67</sup>, close to the physiological intracellular concentration of SAM.

This makes SETD2 particularly sensitive to fluctuations in intracellular SAM levels, and may explain why we found this particular H3 mark to be the most increased after B<sub>12</sub> supplementation.

Nevertheless, we suspect that the role of B<sub>12</sub> in induced plasticity extends beyond H3K36me3 and CT regulation. Levels of histone methylation are dynamic and inter-dependent<sup>43</sup>. SAM has been described as one of the few 'sentinel metabolites' of the cell, with an estimated 1% of all eukaryotic proteins being SAM-dependent methyltransferases with diverse substrates<sup>65,66</sup>. It will be interesting to explore how the additional histone H3 methyl changes we observed with B<sub>12</sub> supplementation, as well as yet-unidentified changes in methylation of DNA, other proteins, or lipids may contribute to the improved reprogramming efficiency achieved by vitamin B<sub>12</sub>.

*In vivo*, we report that depletion of the microbiota with broad-spectrum antibiotics profoundly limits reprogramming. This was



only partially rescued by supplementation of  $B_{12}$ , indicating that the microbiota plays additional roles in driving this process. We suspect that one such role is enhancing inflammation, particularly IL-6 production, which is critical for *in vivo* reprogramming<sup>16</sup>. Interestingly, we also found that the microbiota is important for oral bioavailability of doxycycline beyond the gastrointestinal tract, highlighting the complex relationship between commensal bacteria, doxycycline-inducible systems and *in vivo* reprogramming. We took advantage of *in vitro* reprogramming systems, including doxycycline-free viral reprogramming, to directly demonstrate the limiting, cell-intrinsic role of  $B_{12}$  in this process. Nevertheless, it will be interesting to further study the relationships between  $B_{12}$ , metabolism, microbiota and doxycycline *in vivo*.

From a biomedical perspective, we have translated our results into a pathological disease model of acute ulcerative colitis, in which intestinal epithelial repair requires dedifferentiation of enterocytes to an embryonic-like state<sup>22</sup>. We have shown that further increasing this dedifferentiation is associated with improved colonic tissue repair, and that vitamin  $B_{12}$  alone can facilitate this process. In future studies, it would be valuable to perform a denser time-course analysis in the presence of vitamin  $B_{12}$  to refine the molecular details of the enhanced repair process. Prior studies have reported a protective role of chronic  $B_{12}$  deficiency (that is, before DSS administration); this was due to a blunted inflammatory response to the DSS insult<sup>68,69</sup>. Indeed,  $B_{12}$  deficiency can have immunomodulatory properties<sup>70</sup>, and it will be interesting to study the impact of acute, post-injury supplementation of  $B_{12}$  on the immune compartment.

We have also shown increased CT in both murine and human samples of ulcerative colitis, suggesting further similarities between OSKM and tissue repair. Consistent with this, *Setd2* deletion in intestinal epithelial cells exacerbates DSS injury, and its expression is lost in a high proportion of individuals with human inflammatory bowel disease<sup>71</sup>. In human inflammatory bowel disease, there is evidence of  $B_{12}$  deficiency<sup>72</sup>, and polymorphisms in the vitamin  $B_{12}$  transporter transcobalamin II have been associated with ulcerative colitis<sup>73</sup>. These findings suggest a therapeutic role for  $B_{12}$  in human disease.

We find it intriguing that there is common upregulation of SCA1 during the intermediate, plastic, phase of OSKM-driven reprogramming<sup>19</sup> and during multiple processes of somatic tissue repair, including the intestine<sup>22,59,60</sup>, pancreas<sup>74</sup>, kidney<sup>75,76</sup> and lung<sup>77,78</sup>. We speculate that similar plasticity-mediated processes are conserved across various adult tissues of limited regenerative capacity, and that vitamin  $B_{12}$  may also improve repair in these diverse settings. Overall, our results advance our current molecular understanding of *in vivo* and *in vitro* reprogramming and highlight the possibility to safely administer  $B_{12}$  for the therapeutic enhancement of cellular plasticity for regenerative medicine, organ injury and repair.

## Methods

### Animal procedures

Animal experimentation at the IRB Barcelona was performed according to protocols approved by the Science Park of Barcelona (PCB) Ethics Committee for Research and Animal Welfare. Mice were housed in a specific pathogen-free facility on a 12-hour light–dark cycle at an ambient temperature of 20–24 °C and humidity of 30–70%. Adult mice were fed *ad libitum* with SAFE R40 pellet diet ([https://safe-lab.com/safe\\_en/](https://safe-lab.com/safe_en/)) containing 0.02 mg per kg body weight vitamin  $B_{12}$ . In general, mice of 8–16 weeks of age of both sexes were treated with 1 mg ml<sup>-1</sup> doxycycline hyclate BioChemica (PanReac, A2951) in the drinking water (supplemented with 7.5% sucrose) for 7 d. Antibiotic treatment was conducted using a broad-spectrum cocktail (1 mg l<sup>-1</sup> each of ampicillin (BioChemica, A0839), neomycin sulfate and metronidazole (Sigma, M1547); 0.5 mg l<sup>-1</sup> vancomycin (Cayman Chemical, CAY-15327) all dissolved in water supplemented with 7.5% sucrose) for 3 weeks before doxycycline initiation and was maintained during doxycycline treatment. Vitamin  $B_{12}$  (Sigma, V2876) supplementation

was provided at 1.25 mg l<sup>-1</sup> and folate supplementation was provided as folic acid (Sigma, F7876) at 40 mg l<sup>-1</sup> in the drinking water, both for 7 d concomitant with doxycycline treatment. For the  $B_{12}$  bolus experiment, mice were administered 5 µg vitamin  $B_{12}$  (Sigma, V2876) dissolved in water by oral gavage on day 6 after the start of doxycycline treatment, and blood samples were taken by submandibular collection just before and 24 h after the bolus. OSKM transgenic mice are the i4F-B strain (derived on a C57/BL6 background and bred in house) described in ref. 3 and are available upon request. WT mice were i4F-B WT littermate controls where specified, or WT C57/BL6J (Charles River France).

**DSS-induced colitis.** Mice were treated with 2.5% (wt/vol) DSS, colitis grade (36,000–50,000; MP Biomedicals, MFCD00081551) in drinking water for 5 consecutive days. On day 5, the DSS was removed and drinking water was supplemented with doxycycline hyclate BioChemica (1 mg ml<sup>-1</sup>; PanReac, A2951; with 7.5% sucrose) for 48 h, after which regular water was returned. Mice in the  $B_{12}$  experimental group also received supplementation of vitamin  $B_{12}$  (1.25 mg l<sup>-1</sup>; Sigma, V2876) from the point of DSS removal (that is, day 5) until experimental endpoint. The MAT2Ai group received FIDAS-5 (MedChemExpres, HY-136144) and were dosed with 20 mg per kg body weight per day dissolved in PEG400 by oral gavage as previously described<sup>79</sup>.

**FITC–dextran intestinal permeability assay.** On day 9 (relative to the start of DSS administration), food was withdrawn from mice for 4 h, after which mice were gavaged with FITC–dextran (MW 4,000; Sigma-Aldrich, FD4) at a dose of 44 mg per 100 g of body weight dissolved in PBS. Food restriction was maintained for 3 additional hours, at which point blood was sampled by submandibular vein bleeding. Whole blood was diluted at a ratio of 1:4 in PBS, and 100 µl of blood/PBS mixture from each mouse was loaded into a 96-well plate. Fluorescence intensity was measured on a BioTek Synergy H1 Microplate Reader (excitation 490 nm; emission 520 nm).

### Microbial analysis from murine stool samples

**Sample preparation.** Fresh stool samples were collected directly from mice and snap frozen. gDNA was isolated using a QIAamp DNA Stool Mini Kit (QIAGEN, 51504) according to the manufacturer's protocols.

**Library preparation and sequencing.** Libraries were prepared using the NEBNext Ultra DNA Library Prep Kit for Illumina (E7370L) according to the manufacturer's protocol. Briefly, 50 ng of DNA was fragmented to approximately 400 bp and subjected to end repair plus 'A'-tailing, ligation of NEB adaptor and Uracil excision by USER enzyme. Then, adaptor-ligated DNA was amplified for eight cycles by PCR using indexed primers. All purification steps were performed using AMPure XP Beads (A63881). Final libraries were analysed using an Agilent DNA 1000 chip to estimate the quantity and check size distribution, and were then quantified by qPCR using the KAPA Library Quantification Kit (KK4835, KapaBiosystems) before amplification with Illumina's cBot. Libraries were sequenced (2 × 125 bp) on Illumina's HiSeq 2500.

**Taxonomic analysis.** Reads were aligned to the mm10 genome using STAR 2.7.0a with default parameters<sup>80</sup>. DNA contaminated reads were filtered out from the analysis. The first and final ten bases of the non-contaminated reads were trimmed using DADA2 1.10.1 (ref. 81). Taxonomic assignments were carried out through Kaiju 1.7.0 (ref. 82) using the microbial subset of the NCBI BLAST non-redundant protein database (nr). Resulting sequencing counts were aggregated at genus level. Reads that could not be assigned to any specific genus were classified to the nearest known taxonomic rank (marked by the term \_un). The gut microbial compositional plot displays the relative abundances (percentage) at genus level. Only the 17 most abundant taxa are shown, while the rest were moved to the 'others' category. For all genera, the treatment effect (finish versus start) was compared between OSKM and

control (WT) mice. This was accounted in a model with an interaction term (drug:treatment) using DESeq2 with default options<sup>83</sup>. The paired nature of the experimental design was taken into account in the model as an adjusting factor.

**Functional annotations.** Decontamination from host and trimming was done following the same routines as for the taxonomic analysis. Cleaned sequences for all samples were assembled into contigs using megahit 1.2.4 (ref. 84), and prodigal 2.6.3 (ref. 85) was then used to predict the open reading frames inside the obtained contigs. Protein mapping and KEGG and COG annotations were obtained using the Egg-NOG mapper 2.0.0 (ref. 86). The abundance of the annotated genes was finally measured by counting aligned reads to them via Bowtie2, version 2.2.2, under default parameters<sup>87</sup>. Resulting counts data were aggregated at protein level. The treatment effect (finish versus start) was compared between OSKM and control (WT) mice. This was accounted in a model with an interaction term (drug:treatment) using DESeq2 with default options<sup>83</sup>. The paired nature of the experimental design was considered in the model as an adjusting factor. The top 500 protein hits from the fitted model (nondirectional set) as well as the top 200 positive hits and the top 200 negative hits (directional sets), in all cases ordered by statistical significance, were used to explore enrichment of functional annotations. In this regard, GO terms for bacteria and archaea were considered using the AmiGO 2 GO annotations database<sup>88</sup>, removing from the analysis gene sets with few genes (less than 8) and too many genes (more than 499). Statistically enriched GO terms were identified using the standard hypergeometric test. Significance was defined by the adjusted *P* value using the Benjamini and Hochberg multiple-testing correction. To take into consideration the compositional nature of the data, all DESeq2-based results were complemented with graphical representations of abundance log-ratio (between finish and start matched samples) rankings. This provides a scale invariant way (with regard to the total microbial load) to present the data<sup>89</sup>.

## Metabolomic analysis

**Mouse serum.** Blood was collected via submandibular vein bleed (D0, D2, D4) or intracardiac puncture following deep carbon dioxide anaesthetisation (D7) at approximately 12:00–14:00 h (4–6 h into the light cycle) of each day. Whole blood was spun down for 10 min at 3,381g at 4 °C and supernatant (serum) was separated and stored at –80 °C.

## In vivo time-course metabolomics analysis

**Standard and reagents.** Acetonitrile (Sigma-Aldrich), iso-propanol (Sigma-Aldrich), methanol (Sigma-Aldrich), chloroform (Sigma-Aldrich), acetic acid (Sigma-Aldrich), formic acid (Sigma-Aldrich), methoxyamine hydrochloride (Sigma-Aldrich), MSTFA (*N*-methyl-*N*-(trimethylsilyl) trifluoroacetamide; Sigma-Aldrich), pyridine (Sigma-Aldrich), 3-nitrophenylhydrazine (Sigma-Aldrich), *N*-(3-dimethylaminopropyl)-*N*'-ethylcarbodiimide hydrochloride (EDC; Sigma-Aldrich) and sulfosalicylic acid (Sigma-Aldrich) as previously described<sup>90</sup>.

**Sample preparation serum (lithium heparin).** A volume of 25 µl of serum were mixed with 250 µl a cold solvent mixture with ISTD (methanol/water/chloroform, 9:1:1, –20 °C), into 1.5 ml microtube, vortexed and centrifugated (10 min at 15,000g, 4 °C). The upper phase of supernatant was split into three parts: 50 µl was used for gas chromatography coupled to mass spectrometry (GC–MS) experiments in the injection vial, 30 µl was used for the short-chain fatty acid ultra-high performance liquid chromatography (UHPLC)–MS method, and 50 µl was used for other UHPLC–MS experiments.

**Widely targeted analysis of intracellular metabolites GC coupled to a triple-quadrupole mass spectrometer.** The GC–MS/MS method was performed on a 7890B gas chromatography system (Agilent

Technologies) coupled to a triple-quadrupole 7000C (Agilent Technologies) equipped with a high-sensitivity electronic impact source (EI) operating in positive mode.

**Targeted analysis of bile acids by ion pairing UHPLC coupled to a triple-quadrupole mass spectrometer.** Targeted analysis was performed on an RRLC 1260 system (Agilent Technologies) coupled to a triple-quadrupole 6410 (Agilent Technologies) equipped with an electrospray source operating in positive mode. Gas temperature was set to 325 °C with a gas flow of 12 l min<sup>–1</sup>. Capillary voltage was set to 4.5 kV.

**Targeted analysis of polyamines by ion pairing UHPLC coupled to a triple-quadrupole mass spectrometer.** Targeted analysis was performed on an RRLC 1260 system (Agilent Technologies) coupled to a triple-quadrupole 6410 (Agilent Technologies) equipped with an electrospray source operating in positive mode. The gas temperature was set to 350 °C with a gas flow of 12 l min<sup>–1</sup>. The capillary voltage was set to 3.5 kV.

**Targeted analysis of short-chain fatty acid by ion pairing UHPLC coupled to a 6500+QTRAP mass spectrometer.** Targeted analysis was performed on an RRLC 1260 system (Agilent Technologies) coupled to a 6500 + QTRAP (Sciex) equipped with an electrospray ion source.

**Untargeted analysis of intracellular metabolites by UHPLC coupled to a Q-Exactive mass spectrometer (reversed-phase acetonitrile method).** The profiling experiment was performed with a Dionex Ultimate 3000 UHPLC system (Thermo Scientific) coupled to a Q-Exactive (Thermo Scientific) equipped with an electrospray source operating in both positive and negative mode and full scan mode from 100 to 1,200 *m/z*. The Q-Exactive parameters were: sheath gas flow rate, 55 arbitrary units (a.u.); auxiliary gas flow rate, 15 a.u.; spray voltage, 3.3 kV; capillary temperature, 300 °C; S-Lens RF level, 55 V. The mass spectrometer was calibrated with sodium acetate solution dedicated to low mass calibration.

**MetaboAnalyst.** The peak areas (corrected to quality control) corresponding to each annotated metabolite identified in the serum of reprogrammable mice (*n* = 6 per group) at day 5 and day 7 after doxycycline treatment were converted to log<sub>2</sub> values. Data were represented as log<sub>2</sub> fold change (log<sub>2</sub> FC) values to each mouse at day 0 (before doxycycline administration). Metabolic pathway impact was calculated by Global ANOVA pathway enrichment and Out-degree Centrality Topology analysis through the MetaboAnalyst 4.0 software<sup>91</sup>, using KEGG library (2019) as a reference. The colour gradient from white to red indicates the *P* value, where red is most significant. Bubble size indicates the relative contribution of the detected metabolites in their respective KEGG pathway. Pathway impact scores the centrality of the detected metabolites in the pathway.

**Doxycycline serum analysis.** A total of 30 µl of mouse plasma was acidified with 3 µl solution of 15% phosphoric acid (vol/vol). Afterwards, 42 µl of methyl tert-butyl ether was added and vigorously mixed using a vortex. After 20 min of reequilibration, samples were centrifuged for 10 min at 21,130g at 4 °C. Next, 90 µl of acetonitrile were added to 10 µl of the aqueous phase to facilitate protein precipitation. After another cycle of centrifugation, the supernatant was transferred into a vial before LC–MS analysis.

The extracts were analysed by a UHPLC system coupled to a 6490 triple-quadrupole mass spectrometer (QqQ, Agilent Technologies) with electrospray ion source (LC–ESI–QqQ) working in positive mode. The injection volume was 3 µl. An ACQUITY UPLC BEH HILIC column (1.7 µm, 2.1 × 150 mm, Waters) and a gradient mobile phase consisting of water with 50 mM ammonium acetate (phase A) and acetonitrile (phase B) were used for chromatographic separation. The gradient

was as follows: isocratic for 2 min at 98% B, from 2 to 9 min decreased to 50% B, for 30 s raised to 98%, and finally column equilibrated at 98% B until 13 min. The flow rate was  $0.4 \text{ ml min}^{-1}$ . The mass spectrometer parameters were as follows: drying and sheath gas temperatures, 270 °C and 400 °C, respectively; source and sheath gas flow rates, 15 and  $11 \text{ l min}^{-1}$ , respectively; nebulizer flow, 35 psi; capillary voltage, 3,000 V; nozzle voltage, 1,000 V; and iFunnel HRF and LRF, 130 and 100 V, respectively. The QQQ worked in MRM mode using defined transitions. The transitions for doxycycline and the collision energy (CE(V) were 445 → 428(17), 445 → 98(60).

### Determination of methionine, SAM, SAH and homocysteine in serum

**Sample preparation.** In total, 25  $\mu\text{l}$  of serum was mixed with 25  $\mu\text{l}$  of TCEP and 70  $\mu\text{l}$  of 1% formic acid in methanol. Samples were vortexed and left at  $-20^\circ\text{C}$  for 1 h, centrifuged for 10 min at 21,130g and 4 °C and transferred to glass vials for their analysis by LC-MS.

**LC-MS analysis.** LC-MS was performed with a Thermo Scientific Vanquish Horizon UHPLC system interfaced with a Thermo Scientific Orbitrap ID-X Tribrid Mass Spectrometer.

Metabolites were separated by HILIC chromatography with an InfinityLab Poroshell 120 HILIC-Z 2.7  $\mu\text{m}$ , 2.1 mm × 100 mm column (Agilent Technologies). The mobile phase A was 50 mM ammonium acetate in water, and mobile phase B was acetonitrile. Separation was conducted under the following gradient: 0–2 min, isocratic 90% B; 2–6 min raised to 50% B; 6–7 min, isocratic 50% B; 7–7.2 min, increased to 90% B; 7.2–10.5 min, reequilibration column 90% B. The flow rate was  $0.4 \text{ ml min}^{-1}$ . The injection volume was 5  $\mu\text{l}$ .

Samples were analysed in positive mode in targeted SIM mode and the following setting: isolation window ( $m/z$ ), 4; spray voltage, 3,500 V; sheath gas, 50 a.u.; auxiliary gas, 10 a.u.; ion transfer tube temperature, 300 °C; vaporizer temperature, 300 °C; Orbitrap resolution, 120,000; RF lens, 60%; AGC target, 2e5; maximum injection time, 200 ms.

SAM ( $m/z$  399.145) was monitored from 5–7 min; Met ( $m/z$  150.0583) from 3.2–5.2 min; SAH ( $m/z$  385.1289) from 4–6 min; Hcy ( $m/z$  136.0428) from 3.4–5.5 min, as previously optimized using pure standards.

### Determination of cyanocobalamin in stool

**Sample preparation.** Approximately, 20 mg of dry and pulverized stool samples were mixed with 75  $\mu\text{l}$  of TCEP and 210  $\mu\text{l}$  of 1% formic acid in methanol. Samples were vortexed and subjected to three freeze–thaw cycles using liquid nitrogen. Subsequently, samples were left in ice for 1 h, centrifuged for 10 min at 21,130g and 4 °C and transferred to glass vials for their analysis by LC-MS.

**LC-MS analysis.** LC-MS was performed with a Thermo Scientific Vanquish Horizon UHPLC system interfaced with a Thermo Scientific Orbitrap ID-X Tribrid Mass Spectrometer.

Metabolites were separated by HILIC chromatography with an InfinityLab Poroshell 120 HILIC-Z 2.7  $\mu\text{m}$ , 2.1 mm × 100 mm column (Agilent Technologies). The mobile phase A was 50 mM ammonium acetate in water, and mobile phase B was acetonitrile. Separation was conducted under the following gradient: 0–2 min, isocratic 90% B; 2–6 min raised to 50% B; 6–7 min, isocratic 50% B; 7–7.2 min, increased to 90% B; 7.2–10.5 min, reequilibration column 90% B. The flow was  $0.4 \text{ ml min}^{-1}$ . The injection volume was 5  $\mu\text{l}$ .

Samples were analysed in positive mode in targeted SIM mode and the following setting: isolation window ( $m/z$ ), 4; spray voltage, 3,500 V; sheath gas, 50 a.u.; auxiliary gas, 10 a.u.; ion transfer tube temperature, 300 °C; vaporizer temperature, 300 °C; Orbitrap resolution, 120,000; RF lens, 60%; AGC target, 2e5; maximum injection time, 200 ms. Cyanocobalamin was monitored from ( $m/z$  1355.5747 and  $m/z$  678.291) from 5–5.5 min, as previously optimized using a pure standard.

**Vitamin B<sub>12</sub> serum analysis.** Mouse serum was diluted at a 1:20 ratio in PBS and holotranscobalamin (holoTC) was measured using an ADVIA Centaur Immunoassay System (SIEMENS) with ADVIA Centaur Vitamin B<sub>12</sub> Test Packs (07847260) according to the manufacturer's instructions.

### In vitro SIL experiment

Cell pellets were mixed with 50  $\mu\text{l}$  of TCEP and 140  $\mu\text{l}$  of 1% formic acid in methanol (containing 150  $\mu\text{g l}^{-1}$  of Tryptophan-d5 as internal standard). Samples were vortexed and subjected to three freeze–thaw cycles using liquid nitrogen. Subsequently, samples were left at  $-20^\circ\text{C}$  for 1 h, centrifuged for 10 min at 21,130g and 4 °C and transferred to glass vials for their analysis by LC-MS/MS.

**LC-MS/MS analysis.** Samples were analysed with an UHPLC 1290 Infinity II Series coupled to a QqQ/MS 6490 Series from Agilent Technologies (Agilent Technologies). The source parameters applied operating in positive electrospray ionization (ESI) were gas temperature: 270 °C; gas flow:  $15 \text{ l min}^{-1}$ ; nebulizer: 35 psi; sheath gas heater, 400 a.u.; sheath gas flow, 11 a.u.; capillary, 3,000 V; nozzle voltage: 1,000 V.

The chromatographic separation was performed with an InfinityLab Poroshell 120 HILIC-Z 2.7  $\mu\text{m}$ , 2.1 mm × 100 mm column (Agilent Technologies), starting with 90% B for 2 min, 50% B from minute 2 to 6, and 90% B from minute 7 to 7.2. Mobile phase A was 50 mM ammonium acetate in water, and mobile phase B was acetonitrile. The column temperature was set at 25 °C and the injection volume was 2  $\mu\text{l}$ .

MRM transitions for SAM (RT: 6.1 min) were 399 → 298 (4 V), 399 → 250 (12 V), 399 → 97 (32 V) and 399 → 136 (24 V) for M + 0, and 400 → 299 (4 V), 400 → 251 (12 V), 400 → 97 (32 V), 400 → 137 (24 V), 400 → 250 (12 V) and 400 → 136 (24 V) for M + 1.

### Histology

Samples were fixed overnight at 4 °C with neutral buffered formalin (HT501128-4L, Sigma-Aldrich). Paraffin-embedded tissue sections (2–3  $\mu\text{m}$  in thickness) were air-dried and further dried at 60 °C overnight for immunohistochemical staining.

**Histopathological evaluation.** Sections were stained with haematoxylin and eosin (H&E) for histological evaluation by a board-certified pathologist who was blinded to the experimental groups. Additionally, periodic acid–Schiff staining (AR16592-2, Artisan, Dako, Agilent) was used to visualize mucus-producing cells on 3–4- $\mu\text{m}$  sections of colon that were counterstained with haematoxylin.

In the reprogramming model, the findings were evaluated by focusing mainly on the appearance of hyperplastic and dysplastic changes of the epithelial cells of the digestive mucosa and pancreatic acini. Inflammation and loss of the intestinal goblet cells were also reported. To document the severity and extension, a semi-quantitative grading system was used based on previously used histological criteria:

- Gastric and colon mucosa inflammatory cell infiltrate and multifocal areas of crypt (large intestine) or glandular (stomach) epithelial cell dysplasia were scored from 0 to 5, where 0 indicates absence of lesion and 5 indicates very intense lesions.
- Intestinal crypt hyperplasia: 1, slight; 2, twofold to threefold increase of the crypt length; 3, >threefold increase of the crypt length.
- Goblet cell loss of the mucosa of the large intestine: 1, <10% loss; 2, 10–50% loss; 3, >50% loss.
- Histological total score was presented as a sum of all parameters scored for a given tissue.

In the colitis model, the following parameters were semi-quantitatively evaluated as previously described<sup>92</sup> as follows:

- Inflammation of the colon mucosa: 0, none; 1, slight, 2, moderate; 3, severe.

- Depth of the injury: 0, none; 1, mucosa; 2, mucosa and submucosa; 3, transmural.
- Crypt damage: 0, none; 1, basal and 1/3 damaged; 2, basal and 2/3 damaged; 3, only the surface epithelium intact; 4, entire crypt and epithelium lost.
- Tissue involvement: 0, none; 1, 0–25%; 2, 26–50%; 3, 51–75%; 4, 76–100%.

The score of each parameter was multiplied by the factor of tissue involvement and summed to obtain the total histological score.

### Immunohistochemistry

Immunohistochemistry was performed using a Ventana discovery XT for NANOG and Sca1/Ly6A/E, the Leica BOND RX Research Advanced Staining System for H3K36me3, keratin 14 and vitamin B<sub>12</sub>, and manually for Ki67. Antigen retrieval for NANOG was performed with Cell Conditioning 1 buffer (950-124, Roche) and for Sca1/Ly6A/E with Protease 1 (5266688001, Roche) for 8 min followed with the OmniMap anti-Rat HRP (760-4457, Roche) or OmniMap anti-Rb HRP (760-4311, Roche). Blocking was done with casein (760-219, Roche). Antigen–antibody complexes were revealed with ChromoMap DAB Kit (760-159, Roche). For H3K36me3 and keratin 14, antigen retrieval was performed with BOND Epitope Retrieval 1 (AR9961, Leica) and for vit B<sub>12</sub> with BOND Epitope Retrieval Solution 2 (Leica Biosystems, AR9640) for 20 min, whereas for Ki67, sections were dewaxed as part of the antigen retrieval process using the low pH EnVision FLEX Target Retrieval Solutions (Dako) for 20 min at 97 °C using a PT Link (Dako-Agilent). Blocking was performed with Peroxidase-Blocking Solution at room temperature (RT; S2023, Dako-Agilent) and 5% goat normal serum (16210064, Life technology) mixed with 2.5% BSA diluted in wash buffer for 10 and 60 min at RT. Vitamin B<sub>12</sub> also was blocked with Vector M.O.M. Blocking Reagent (MK-2213, Vector) following the manufacturer's procedures for 60 min. Primary antibodies were incubated for 30, 60 or 120 min. The secondary antibody used was the BrightVision poly HRP-Anti-Rabbit IgG, incubated for 45 min (DPVR-110HPRP, ImmunoLogic) or the polyclonal goat Anti-Mouse at a dilution of 1:100 for 30 min (Dako-Agilent, P0447). Antigen–antibody complexes were revealed with 3,3'-diaminobenzidine (K346811, Agilent or RE7230-CE, Leica). Sections were counterstained with haematoxylin (CS700, Dako-Agilent or RE7107-CE, Leica) and mounted with Mounting Medium, Toluene-Free (CS705, Dako-Agilent) using a Dako Cover-Stainer. Specificity of staining was confirmed by staining with a rat IgG (6-001-F, R&D Systems, Bio-Techne), a Rabbit IgG (ab27478, Abcam) or a mouse IgG1, kappa (Abcam, ab18443) isotype controls. See Supplementary Table 5 for primary antibody details.

**In situ hybridization—RNAscope.** Ready-to-use reagents from RNAscope 2.5LS Reagent Kit-RED (322150, RNAscope, ACD Bio-Techne) were loaded onto the Leica Biosystems BOND RX Research Advanced Staining System according to the user manual (322100-USM). FFPE tissue sections were baked and deparaffinized on the instrument, followed by epitope retrieval (using Leica Epitope Retrieval Buffer 2 at 95 °C for 15 min) and protease treatment (15 min at 40 °C). Probe hybridization, signal amplification, colorimetric detection and counterstaining were subsequently performed following the manufacturer's recommendations.

Hybridization was performed with the RNAscope LS 2.5 Probe - Mm-Lgr5 - *Mus musculus* leucine rich repeat containing G-protein-coupled receptor 5 (312178, RNAscope, ACD Bio-Techne). Control probe used was the RNAscope 2.5 LS Probe - Mm-UBC - *Mus musculus* ubiquitin C (Ubc), as a housekeeping gene (310778, RNAscope - ACD Bio-Techne). The bacterial probe RNAscope 2.5 LS Negative Control Probe\_dapB was used as a negative control (312038, RNAscope - ACD Bio-Techne).

**Image acquisition.** Brightfield images were acquired with a NanoZoomer-2.0 HT C9600 digital scanner (Hamamatsu) equipped with a  $\times 20$  objective. All images were visualized with a gamma correction set at 1.8 in the image control panel of the NDP.view 2 U12388-01 software (Hamamatsu, Photonics).

**Image analysis.** Brightfield images of immunohistochemistry were quantified using QuPath software<sup>93</sup> with standard detection methods. Where the percentage of tissue staining is calculated, pixels were classified as positive and negative using the Thresholder function. Where the percentage of cells is quantified, the Positive Cell Detection function was used.

### Cellular and molecular methods

**Cell culture.** MEFs were cultured in standard DMEM medium with 10% FBS (Gibco, Life Technologies, 10270106) with antibiotics (100 U ml<sup>-1</sup> penicillin–streptomycin; Life Technologies, 11528876). Reprogramming of the doxycycline-inducible 4-Factor (i4F) MEFs with inducible expression of the four Yamanaka factors Oct4, Sox2, Klf4 and cMyc (OSKM) was performed as previously described<sup>3</sup>. Briefly, i4F MEFs were seeded at a density of  $3 \times 10^5$  cells per well in six-well tissue culture plates coated with gelatin and treated with doxycycline (PanReac, A2951) 1 mg ml<sup>-1</sup> continuously to induce expression of the OSKM transcription factors in the presence of 'complete KSR media' (15% (vol/vol) Knockout Serum Replacement (KSR, Invitrogen, 10828028) in DMEM with GlutaMax (Life Technologies, 31966047) basal media, with 1,000 U ml<sup>-1</sup> LIF (Merck, 31966047), non-essential amino acids (Life Technologies, 11140035) and 100  $\mu$ M beta-mercaptoethanol (Life Technologies, 31350010) plus antibiotics (penicillin–streptomycin, Gibco, 11528876)), which was replaced every 48–72 h. After 10 d, iPS cell colonies were scored by alkaline phosphatase staining according to the manufacturer's protocol (APblue membrane substrate detection kit, Sigma, AB0300). Vitamin B<sub>12</sub> (Sigma, V2876; 2  $\mu$ M final), MAT2Ai PF-9366 (MedChemExpress, HY-107778; 2  $\mu$ M final), SAM (S-(5'-adenosyl)-L-methionine iodide, Merck, A4377; 100  $\mu$ M final) and NSC636819 (Sigma-Aldrich, 5.31996; 10  $\mu$ M final) were added continuously to the culture media and replaced every 48–72 h.

**Retroviral reprogramming.** Reprogramming of WT MEFs was performed as previously described<sup>94</sup>. Briefly, HEK-293T (American Type Culture Collection, ATCC-CRL-3216) cells were cultured in DMEM supplemented with 10% FBS and antibiotics (penicillin–streptomycin, Gibco, 11528876). Around  $5 \times 10^6$  cells per 100-mm-diameter dish were transfected with the ecotropic packaging plasmid pCL-Eco (4  $\mu$ g) together with one of the following retroviral constructs (4  $\mu$ g): pMXs-Klf4, pMXs-Sox2, pMXs-Oct4 or pMXs-cMyc (obtained from Addgene) using Fugene-6 transfection reagent (Roche) according to the manufacturer's protocol. The following day, media were changed and recipient WT MEFs to be reprogrammed were seeded ( $1.5 \times 10^5$  cells per well of a six-well plate). Retroviral supernatants (10 ml per plate/factor) were collected serially during the subsequent 48 h, at 12-h intervals, each time adding fresh media to the 293T cells (10 ml). After each collection, supernatant was filtered through a 0.45- $\mu$ m filter, and each well of MEFs received 0.5 ml of each of the corresponding retroviral supernatants (amounting to 2 ml total). Vitamin B<sub>12</sub> supplementation (Sigma, V2876; 2  $\mu$ M final concentration) began on the same day as viral transduction. This procedure was repeated every 12 h for 2 d (a total of four additions). After infection was completed, media were replaced by 'complete KSR media' (see above). Cell pellets were harvested on day 5 (relative to the first infection) and histone extracts were processed for immunoblot as described below. On day 14 (relative to the first infection), iPS cell colonies were scored by alkaline phosphatase staining according to the manufacturer's protocol (APblue membrane substrate detection kit; Sigma, AB0300).

**SIL.** Doxycycline-inducible i4F MEFs were cultured as described in 'Cell culture' above, with 1 mg ml<sup>-1</sup> doxycycline, with or without continuous vitamin B<sub>12</sub> supplementation. At 72 h after the addition of doxycycline, cells were transferred to complete KSR media containing a final concentration of 0.5 mM L-Serine-<sup>13</sup>C<sub>3</sub> (Sigma-Aldrich, 604887). This is the same concentration of unlabelled L-serine normally found in the complete KSR media, and was generated by ordering custom, serine-free DMEM (Life Technologies, ME22803L1) and custom, serine-free non-essential amino acid mixture (Life Technologies, ME22804L1). Six hours after the addition of labelled media, a subset of wells was harvested by scraping in PBS and centrifugation (300g for 5 min); supernatant was removed and pellets were snap frozen. At 72 h after the addition of the labelled media (that is, 6 days into reprogramming), cells still in culture were transferred back to unlabelled complete KSR media, which was changed every 48–72 h. iPS cell colonies were analysed by alkaline phosphatase staining according to the manufacturer's protocol (AP blue membrane substrate detection kit; Sigma, ABO300) on day 10. Doxycycline and vitamin B<sub>12</sub> supplementation were continuous throughout the entire reprogramming protocol, and replenished with every media change (that is, every 48–72 h).

**Histone array.** i4F MEFs were cultured in the presence doxycycline  $\pm 2$   $\mu$ M of vitamin B<sub>12</sub> over 3 or 10 days (culture conditions as described above) and histone extracts were prepared using EpiQuik Total Histone Extraction Kit (EpiGentek, OP-0006-100) according to the manufacturer's instructions. Around 200 ng of total histone extract was used per well in the EpiQuik Histone H3 Modification Multiplex Assay Kit (Colorimetric; EpiGentek, P-3100) according to the manufacturer's instructions.

**Cell lysis and immunoblot.** Histone extracts were prepared using an EpiQuik Total Histone Extraction Kit (EpiGentek, OP-0006-100) according to the manufacturer's instructions and quantified using DC Protein Assay Kit (Bio-Rad, 5000111). Whole-cell extracts were prepared in RIPA buffer (10 mM Tris-HCl, pH 8.0; 1 mM EDTA; 0.5 mM EGTA; 1% Triton X-100; 0.1% sodium deoxycholate; 0.1% SDS; 140 mM NaCl). A total of 10  $\mu$ g of lysate was loaded per lane and hybridized using antibodies against H3K36me3, MS, vinculin, total histone H3 and LI-COR fluorescent secondary reagents (IRDye 800 CW anti-mouse, 926-32210; IRDye 680 CW anti-mouse, 926-68070; IRDye 800 CW anti-rabbit, 926-32211; IRDye 680 CW anti-mouse, 926-68071) all at a dilution of 1:10,000 according to manufacturer's instructions. Immunoblots were visualized on an Odyssey FC Imaging System (LI-COR Biosciences). See Supplementary Table 5 for primary antibody details.

**GSEA.** GSEAPreranked was used to perform a GSEA of annotations from MsigDB M13537, with standard GSEA and leading edge analysis settings. We used the RNA-seq gene list ranked by log<sub>2</sub> fold change, selecting 'gene set' as the permutation method with 1,000 permutations for Kolmogorov-Smirnov correction for multiple testing<sup>95</sup>.

**Selection of genes to measure by qPCR from methionine deprivation signature.** Genes belonging to the leading edge of the GSEA using the Met derivation signature (MsigDB, M13537) in the pancreas of reprogramming mice were selected. These genes were then compared to genes belonging to the leading edge of the same gene signature from i4F MEFs treated with doxycycline in vitro for 72 h, as compared to OSKM MEFs treated with vitamin B<sub>12</sub> (that is, genes in MsigDB M13537 whose upregulation was relieved by B<sub>12</sub> supplementation in vitro). We selected 11 of these genes for which we had qPCR primers available.

**Analysis of mRNA levels by qPCR.** Total RNA was extracted from MEFs with TRIzol (Invitrogen) according to the manufacturer's instructions. Up to 5  $\mu$ g of total RNA was reverse transcribed into cDNA using the iScript Advanced cDNA Synthesis Kit (Bio-Rad, 172-5038; pancreas)

or iScript cDNA Synthesis Kit (Bio-Rad, 1708890; all other organs) for RT-qPCR. Real-time qPCR was performed using GoTaq qPCR Master Mix (Promega, A6002) in a QuantStudio 6 Flex thermocycler (Applied Biosystem) or 7900HT Fast Real-Time PCR System (Thermo Fisher). See Supplementary Table 6 for primer sequences.

**ChIP sample preparation.** i4F MEFs were cultured in the presence or absence of doxycycline  $\pm 2$   $\mu$ M of vitamin B<sub>12</sub> (Merck, V2876) over 3 days in six-well plates (culture conditions as described above). Cells were fixed with 1% (vol/vol) PFA (Fisher Scientific, 50980487) for 2 min and then quenched with 750 mM Tris (PanReac ApplChem, A2264) for 5 min. Cells were washed twice with PBS, scraped, and spun down at 1,200g for 5 min. Pellets were lysed with 100  $\mu$ l (per well) lysis buffer (50 mM HEPES-KOH pH 7.5, 140 mM HCl, 1 mM EDTA pH 8, 1% Triton X-100, 0.1% sodium deoxycholate, 0.1% SDS, protease inhibitor cocktail; Sigma, 4693159001) on ice for 10 min, then sonicated using a Diagenode BioRuptor Pico (Diagenode, B01060010) for ten cycles (30 s on, 30 s off) at 4 °C. Lysates were clarified for 10 min at 8,000g, 1% input samples were reserved, and supernatant was used for immunoprecipitation with Diagenode Protein A-coated Magnetic beads ChIP-seq grade (Diagenode, C03010020-660) and H3K3me3 monoclonal antibody (Cell Signaling Technologies, 4909) with 0.1% BSA (Sigma, 10735094001). The following day, cells were washed once with each buffer: low salt (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH 8.0, 150 mM NaCl), high salt (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH 8.0, 5,000 mM NaCl), LiCl (0.25 M LiCl, 1% NP-40, 1% sodium deoxycholate, 1 mM EDTA, 10 mM Tris-HCl pH 8.0) and eluted in 1% SDS, 100 mM NaHCO<sub>3</sub> buffer. Cross-links were reversed with RNase A (Thermo Fisher, EN0531), proteinase K (Merck, 3115879001) and sodium chloride (Sigma, 71376), and chromatin fragments were purified using QIAquick PCR purification kit (Qiagen, 28104).

**RNA-seq RNA extraction.** i4F MEFs were cultured in the presence or absence of doxycycline and the indicated compounds over 3 days in six-well plates (culture conditions as described above). After 72 h, RNA was extracted using an RNeasy Kit (Qiagen, QIA74106) according to the manufacturer's instructions.

**ChIP-seq.** The concentration of the DNA samples (inputs and immunoprecipitations) was quantified with a Qubit dsDNA HS kit, and fragment size distribution was assessed with the Bioanalyzer 2100 DNA HS assay (Agilent). Libraries for ChIP-seq were prepared at the IRB Barcelona Functional Genomics Core Facility. Briefly, single-indexed DNA libraries were generated from 0.5–1.5 ng of DNA samples using the NEBNext Ultra II DNA Library Prep kit for Illumina (New England Biolabs). Eleven cycles of PCR amplification were applied to all libraries.

The final libraries were quantified using the Qubit dsDNA HS assay (Invitrogen) and quality controlled with the Bioanalyzer 2100 DNA HS assay (Agilent). An equimolar pool was prepared with the 24 libraries and sequenced on a NextSeq 550 (Illumina). 78.9 Gb of SE75 reads were produced from two high-output runs. A minimum of 23.97 million reads were obtained for all samples.

**RNA-seq.** For MEFs 1–3. The concentration of total RNA extractions was quantified with the Nanodrop One (Thermo Fisher), and RNA integrity was assessed with the Bioanalyzer 2100 RNA Nano assay (Agilent). Libraries for RNA-seq were prepared at the IRB Barcelona Functional Genomics Core Facility. Briefly, mRNA was isolated from 1.5  $\mu$ g of total RNA using the kit NEBNext Poly(A) mRNA Magnetic Isolation Module (New England Biolabs). The isolated mRNA was used to generate dual-indexed cDNA libraries using the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (New England Biolabs). Ten cycles of PCR amplification were applied to all libraries.

The final libraries were quantified using the Qubit dsDNA HS assay (Invitrogen) and quality controlled with the Bioanalyzer 2100

DNA HS assay (Agilent). An equimolar pool was prepared with the 12 libraries and submitted for sequencing at the Centre Nacional d'Anàlisi Genòmica (CRG-CNAG). A final quality control by qPCR was performed by the sequencing provider before paired-end 50-nucleotide sequencing on a NovaSeq 6000 S2 (Illumina). Around 77.7 Gb of PES0 reads were produced from three NovaSeq 6000 flow cells. A minimum of 55.7 million reads were obtained for all samples (Extended Data Fig. 7).

**For MEFs 4–5.** Total RNA extractions were quantified with a Nanodrop One (Thermo Fisher), and RNA integrity was assessed with the Bioanalyzer 2100 RNA Nano assay (Agilent). Libraries for RNA-seq were prepared at the IRB Barcelona Functional Genomics Core Facility. Briefly, mRNA was isolated from 1.2 µg of total RNA and used to generate dual-indexed cDNA libraries with the Illumina Stranded mRNA ligation kit (Illumina) and UD Indexes Set A (Illumina). Ten cycles of PCR amplification were applied to all libraries.

Sequencing-ready libraries were quantified using the Qubit dsDNA HS assay (Invitrogen) and quality controlled with the Tapestation HS D5000 assay (Agilent). An equimolar pool was prepared with the 15 libraries for SE75 sequencing on a NextSeq 550 (Illumina). Sequencing output was above 539 million 75-nucleotide single-end reads and a minimum of 28 million reads was obtained for all samples (Extended Data Fig. 7).

### Bioinformatic analysis

**RNA-seq data processing.** All analyses were performed in the R programming language (version 4.0.5)<sup>96</sup> unless otherwise stated. Stranded paired-end reads were aligned to the *Mus musculus* reference genome version mm10 using STAR<sup>80</sup> with default parameters. STAR indexes were built using the ENSEMBL annotation version GRC138.97. SAM files were converted to BAM and sorted using sambamba (version 0.6.7)<sup>97</sup>. Gene counts were obtained with the featureCounts function from the Rsubread package<sup>98</sup> with the gtf file corresponding to ENSEMBL version GRC138.97 and parameters set to: isPairedEnd = TRUE and strandSpecific = 2. Technical replicates were collapsed by adding the corresponding columns in the count matrix.

**Reprogramming score.** We obtained a reprogramming gene signature from published data<sup>48</sup> and selected genes with false discovery rate (FDR) lower than 0.05 and fold change between MEF and d3-EFF larger than 2. The reprogramming score was defined as the average of all genes in the signature after scaling the rlog transformed matrix.

**Computation of cryptic transcript ratios between first and intermediate exons.** Exon counts were generated using the featureCounts function with parameters: isPairedEnd = TRUE, strandSpecific = 2, GTF.featureType = exon, GTF.attrType = transcript\_id, GTF.attrType.extra = gene\_id, allowMultiOverlap = TRUE and useMetaFeatures = FALSE and the same GTF as for gene counts. Technical replicates were collapsed by adding the corresponding counts. For each gene, the longest annotated transcript was selected. Genes with less than four exons of RPKMs lower than  $\exp(-2)$  were discarded from the analysis. Intermediate exons were defined as those from the fourth to the penultimate. A total of 9,365 genes were used to compute the ratio between the intermediate and first exons. Fold changes between untreated and B<sub>12</sub>-treated samples were computed as the ratio between the exon ratios.

**Comparison of cryptic transcript ratios between conditions.** Genes were separated by their expression after transcript length and library size normalization (RPKM). For each sample, we computed the median ratios for genes in each decile.

**Analysis of CT in DSS time course.** Data were accessed from GSE131032. Reads were processed and ratios computed as previously described. log<sub>2</sub> ratios for all transcripts were summarized through the

median by sample. Comparisons between days were performed fitting a linear model to the medians using 'cage' as a covariate. The function glht from the multcomp R package was used to find coefficients and *P* values.

**Functional enrichment in genes with exon ratios affected by vitamin B<sub>12</sub> treatment.** To select genes most affected by the B<sub>12</sub> treatment after reprogramming, we compared ratios between the doxy and MEF conditions and between the doxy and doxy + B<sub>12</sub> conditions. Genes that increased the ratios in the first comparison (upper 25th percentile) and decreased the ratio in the second comparison (bottom 25%) were selected for functional enrichment analysis. A hypergeometric test was performed to find significant overlap between the defined gene set and the Biological Processes GO collection<sup>99</sup>.

**ChIP-seq data processing.** Reads were aligned to the mm10 reference genome with bowtie<sup>100</sup> version 0.12.9 with parameters --n 2 and --m 1 to keep reads with multiple alignments in one position. SAM files were converted to BAM and sorted using sambamba version 0.6.7.

**Heat maps of average coverage in gene bodies.** For each sample, aligned reads were imported into R using the function scanBam from the Rsamtools package<sup>101</sup>. Whole-genome coverage was computed using the coverage function from the IRanges package<sup>102</sup> and binned into 50-bp windows. Gene annotations were imported from Ensembl version GRCm38. The average coverage over gene bodies was computed using the normalizeToMatrix function from the EnrichedHeatmap package<sup>103</sup> with parameters extend = 1,000, mean\_mode = w0 and w = 50. Genes were filtered to coincide with those used in the exon ratio calculation from the RNA-seq data. Rows in the heat map were split by the average RNA-seq RPKM values in all samples.

**Visualization of ChIP tracks.** BAM files were transformed to TDF files using the count function from IGVtools (version 2.12.2)<sup>104</sup> with parameters --z 7, --w 25 and --e 250. Visualization of TDF files was generated using IGV (version 2.9.4)<sup>105</sup>.

**Analysis of human RNA-seq data.** Data were accessed from GSE109142. Reads were processed and ratios computed as previously described except using the ENSEMBL GRCm38.101 human gene annotation and the hg38 genome assembly version. The log<sub>2</sub> ratios for all transcripts were summarized through the median by sample. Comparison between diagnosis status was performed fitting a linear model to the medians with sex and the expression quantiles as covariates. The model was fitted using the lm R function and coefficients and *P* values with the coeff function.

### Statistics and data availability

**Statistical analysis and figure preparation.** Unless otherwise specified, data are presented as the mean  $\pm$  s.d. Statistical analysis was performed by Student's *t*-test or one-way analysis of variance (ANOVA) as indicated, using GraphPad Prism v9.0.0, and specific statistical tests as indicated for each experiment for bioinformatic analyses. *P* values of less than 0.05 were considered as statistically significant. No statistical methods were used to predetermine sample size in the mouse studies, but our sample sizes are similar to those reported in previous publications<sup>3,9,16,17,19</sup>. Animals and data points were not excluded from analysis with the exception of the MEFs that failed to reprogram in the ChIP experiment, which is clearly detailed in the text. Mice were allocated at random to treatment groups, with attempts to balance initial body weight and sex as possible. The investigators were blinded during histological assessment of the mice; other data collection and analysis was not performed blind to the conditions of the experiments. Data distribution was assumed to be normal, but this was not formally tested. Figures were prepared using Illustrator CC 2019 (Adobe).

## Reporting summary

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

## Data availability

All sequencing data are deposited under the following Gene Expression Omnibus accessions: [GSE154149](#), microbial genome analysis from OSKM and WT mice pretreatment and after 7 d of 1 mg ml<sup>-1</sup> treatment with doxycycline in the drinking water; [GSE200578](#), ChIP-seq of H3k36me3 samples of OSKM reprogramming MEFs treated with B<sub>12</sub>; [GSE200579](#), RNA-seq samples of OSKM reprogrammable MEFs treated with or without vitamin B<sub>12</sub>; [GSE232382](#), RNA-seq samples of OSKM reprogrammable MEFs treated with vitamin B<sub>12</sub> and/or various compounds to modulate SAM or histone methylation. Previously published datasets that were used for analysis in the current study are: [GSE131032](#), RNA-seq of time-course analysis of repairing murine epithelium after DSS injury; [GSE109142](#), RNA-seq of human paediatric ulcerative colitis and normal tissue controls; [GSE102518](#), RNA-seq of murine in vitro reprogramming in MEFs of varying genotypes; [GSE77722](#), RNA-seq of murine in vivo reprogramming in mice of varying genotypes. Source data are provided with this paper.

## References

1. Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663–676 (2006).
2. Taguchi, J. & Yamada, Y. In vivo reprogramming for tissue regeneration and organismal rejuvenation. *Curr. Opin. Genet. Dev.* **46**, 132–140 (2017).
3. Abad, M. et al. Reprogramming in vivo produces teratomas and iPS cells with totipotency features. *Nature* **502**, 340–345 (2013).
4. Ocampo, A. et al. In vivo amelioration of age-associated hallmarks by partial reprogramming. *Cell* **167**, 1719–1733 (2016).
5. Doeser, M. C., Schöler, H. R. & Wu, G. Reduction of fibrosis and scar formation by partial reprogramming in vivo. *Stem Cells* **36**, 1216–1225 (2018).
6. Lu, Y. et al. Reprogramming to recover youthful epigenetic information and restore vision. *Nature* **588**, 124–129 (2020).
7. Chen, Y. et al. Reversible reprogramming of cardiomyocytes to a fetal state drives heart regeneration in mice. *Science* **373**, 1537–1540 (2021).
8. Browder, K. C. et al. In vivo partial reprogramming alters age-associated molecular changes during physiological aging in mice. *Nat. Aging* **2**, 243–253 (2022).
9. Chondronasiou, D. et al. Multi-omic rejuvenation of naturally aged tissues by a single cycle of transient reprogramming. *Aging Cell* **21**, e13578 (2022).
10. Ohnishi, K. et al. Premature termination of reprogramming in vivo leads to cancer development through altered epigenetic regulation. *Cell* **156**, 663–677 (2014).
11. Shibata, H. et al. In vivo reprogramming drives Kras-induced cancer development. *Nat. Commun.* **9**, 2081 (2018).
12. Shyh-Chang, N. et al. Influence of threonine metabolism on S-adenosylmethionine and histone methylation. *Science* **339**, 222–226 (2013).
13. Fernández-Arroyo, S. et al. Activation of the methylation cycle in cells reprogrammed into a stem cell-like state. *Oncoscience* **2**, 958–967 (2016).
14. Degnan, P. H., Taga, M. E. & Goodman, A. L. Vitamin B12 as a modulator of gut microbial ecology. *Cell Metab.* **20**, 769–778 (2014).
15. Nicholson, J. K. et al. Host-gut microbiota metabolic interactions. *Science* **336**, 1262–1267 (2012).
16. Mosteiro, L. et al. Tissue damage and senescence provide critical signals for cellular reprogramming in vivo. *Science* **354**, aaf4445 (2016).
17. Melendez, E. et al. Natural killer cells act as an extrinsic barrier for in vivo reprogramming. *Development* <https://doi.org/10.1242/dev.200361> (2022).
18. Kennedy, E. A., King, K. Y. & Baldridge, M. T. Mouse microbiota models: comparing germ-free mice and antibiotics treatment as tools for modifying gut bacteria. *Front. Physiol.* **9**, 1534 (2018).
19. Chondronasiou, D. et al. Deciphering the roadmap of in vivo reprogramming toward pluripotency. *Stem Cell Rep.* **17**, 2501–2517 (2022).
20. Levy, M. et al. Microbiota-modulated metabolites shape the intestinal microenvironment by regulating NLRP6 inflammasome signaling. *Cell* **163**, 1428–1443 (2015).
21. Chang, C. S. et al. Identification of a gut microbiota member that ameliorates DSS-induced colitis in intestinal barrier enhanced Dusp6-deficient mice. *Cell Rep.* **37**, 110016 (2021).
22. Yui, S. et al. YAP/TAZ-dependent reprogramming of colonic epithelium links ECM remodeling to tissue regeneration. *Cell Stem Cell* **22**, 35–49 (2018).
23. Moschen, A. R. et al. Lipocalin 2 protects from inflammation and tumorigenesis associated with gut microbiota alterations. *Cell Host Microbe* **19**, 455–469 (2016).
24. Degnan, P. H., Barry, N. A., Mok, K. C., Taga, M. E. & Goodman, A. L. Human gut microbes use multiple transporters to distinguish vitamin B<sub>12</sub> analogs and compete in the gut. *Cell Host Microbe* **15**, 47–57 (2014).
25. Martens, H., Barg, M., Warren, D. & Jah, J.-H. Microbial production of vitamin B<sub>12</sub>. *Appl. Microbiol. Biotechnol.* **58**, 275–285 (2002).
26. Green, R. et al. Vitamin B<sub>12</sub> deficiency. *Nat. Rev. Dis. Prim.* **3**, 17040 (2017).
27. Kornerup, L. S. et al. Tissue distribution of oral vitamin B<sub>12</sub> is influenced by B<sub>12</sub> status and B<sub>12</sub> form: an experimental study in rats. *Eur. J. Nutr.* **57**, 1459–1469 (2018).
28. Shin, O.-H. et al. Methyl-group donors cannot prevent apoptotic death of rat hepatocytes induced by choline-deficiency. *J. Cell. Biochem.* **64**, 196–208 (1997).
29. Birn, H. et al. Megalin is essential for renal proximal tubule reabsorption and accumulation of transcobalamin-B<sub>12</sub>. *Am. J. Physiol. Ren. Physiol.* **282**, F408–F416 (2002).
30. Scott, J. S. D., Treston, A. M., Bowman, E. P. W., Owens, J. A. & Cooksley, W. G. The regulatory roles of liver and kidney in cobalamin (vitamin B<sub>12</sub>) metabolism in the rat: the uptake and intracellular binding of cobalamin and the activity of the cobalamin-dependent enzymes in response to varying cobalamin supply. *Clin. Science* **67**, 299–306 (1984).
31. Okuda, K. Relationship between intake of vitamin B<sub>12</sub> and its storage by the kidney in the rat. *J. Nutr.* **77**, 131–136 (1962).
32. Rossi, M., Amaretti, A. & Raimondi, S. Folate production by probiotic bacteria. *Nutrients* **3**, 118–134 (2011).
33. Sanderson, S. M., Gao, X., Dai, Z. & Locasale, J. W. Methionine metabolism in health and cancer: a nexus of diet and precision medicine. *Nat. Rev. Cancer* **19**, 625–637 (2019).
34. Shiraki, N. et al. Methionine metabolism regulates maintenance and differentiation of human pluripotent stem cells. *Cell Metab.* **19**, 780–794 (2014).
35. Wang, J. et al. Dependence of mouse embryonic stem cells on threonine catabolism. *Science* **325**, 435–439 (2009).
36. Bravo, A. C. et al. Analysis of S-adenosylmethionine and S-adenosylhomocysteine: method optimisation and profiling in healthy adults upon short-term dietary intervention. *Metabolites* **12**, 373 (2022).
37. Lildballe, D. L., Mutti, E., Birn, H. & Nexo, E. Maximal load of the vitamin B12 transport system: a study on mice treated for four weeks with high-dose vitamin B<sub>12</sub> or cobinamide. *PLoS ONE* **7**, e46657 (2012).

38. Beedholm-Ebsen, R. et al. Identification of multidrug resistance protein 1 (MRP1/ABCC1) as a molecular gate for cellular export of cobalamin. *Blood* **115**, 1632–1639 (2010).

39. Kokkinakis, D. M. et al. Mitotic arrest, apoptosis, and sensitization to chemotherapy of melanomas by methionine deprivation stress. *Mol. Cancer Res.* **4**, 575–589 (2006).

40. Zviran, A. et al. Deterministic somatic cell reprogramming involves continuous transcriptional changes governed by myc and epigenetic-driven modules. *Cell Stem Cell* **24**, 328–341 (2019).

41. National Research Council (US) Subcommittee on Laboratory Animal Nutrition. *Nutrient Requirements of Laboratory Animals* (National Academies Press, 1995). <https://doi.org/10.17226/4758>

42. Mentch, S. J. et al. Histone methylation dynamics and gene regulation occur through the sensing of one-carbon metabolism. *Cell Metab.* **22**, 861–873 (2015).

43. Ye, C., Sutter, B. M., Wang, Y., Kuang, Z. & Tu, B. P. A metabolic function for phospholipid and histone methylation. *Mol. Cell* **66**, 180–193 (2017).

44. Chronis, C. et al. Cooperative binding of transcription factors orchestrates reprogramming. *Cell* **168**, 442–459 (2017).

45. van den Hurk, M. et al. Transcriptional and epigenetic mechanisms of cellular reprogramming to induced pluripotency. *Epigenomics* **8**, 1131–1149 (2016).

46. Chu, C. H. et al. KDM4B as a target for prostate cancer: structural analysis and selective inhibition by a novel inhibitor. *J. Med. Chem.* **57**, 5975–5985 (2014).

47. Edmunds, J. W., Mahadevan, L. C. & Clayton, A. L. Dynamic histone H3 methylation during gene induction: HYPB/Setd2 mediates all H3K36 trimethylation. *EMBO J.* **27**, 406–420 (2008).

48. Schwarz, B. A. et al. Prospective isolation of poised iPSC intermediates reveals principles of cellular reprogramming. *Cell Stem Cell* **23**, 289–305 (2018).

49. Li, B. et al. Infrequently transcribed long genes depend on the Set2/Rpd3S pathway for accurate transcription. *Genes Dev.* **21**, 1422–1430 (2007).

50. Neri, F. et al. Intragenic DNA methylation prevents spurious transcription initiation. *Nature* **543**, 72–77 (2017).

51. Barral, A. et al. SETDB1/NSD-dependent H3K9me3/H3K36me3 dual heterochromatin maintains gene expression profiles by bookmarking poised enhancers. *Mol. Cell* **82**, 816–832 (2022).

52. McCauley, B. S. et al. Altered chromatin states drive cryptic transcription in aging mammalian stem cells. *Nat. Aging* **1**, 684–697 (2021).

53. Huang, C. & Zhu, B. Roles of H3K36-specific histone methyltransferases in transcription: antagonizing silencing and safeguarding transcription fidelity. *Biophys. Rep.* **4**, 170–177 (2018).

54. Pu, M. et al. Trimethylation of Lys36 on H3 restricts gene expression change during aging and impacts life span. *Genes Dev.* **29**, 718–731 (2015).

55. Sen, P. et al. H3K36 methylation promotes longevity by enhancing transcriptional fidelity. *Genes Dev.* **29**, 1362–1376 (2015).

56. Carrozza, M. J. et al. Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* **123**, 581–592 (2005).

57. Ang, Y.-S. et al. Wdr5 mediates self-renewal and reprogramming via the embryonic stem cell core transcriptional network. *Cell* **145**, 183–197 (2011).

58. Iismaa, S. E. et al. Comparative regenerative mechanisms across different mammalian tissues. *npj Regen. Med.* **3**, 6 (2018).

59. Ayyaz, A. et al. Single-cell transcriptomes of the regenerating intestine reveal a revival stem cell. *Nature* <https://doi.org/10.1038/s41586-019-1154-y> (2019).

60. Nusse, Y. M. et al. Parasitic helminths induce fetal-like reversion in the intestinal stem cell niche. *Nature* **559**, 109–113 (2018).

61. Perše, M. & Cerar, A. Dextran sodium sulphate colitis mouse model: traps and tricks. *J. Biomed. Biotechnol.* **2012**, 718617 (2012).

62. Czarniewski, P. et al. Conserved transcriptomic profile between mouse and human colitis allows unsupervised patient stratification. *Nat. Commun.* **10**, 2892 (2019).

63. Haberman, Y. et al. Ulcerative colitis mucosal transcriptomes reveal mitochondrialopathy and personalized mechanisms underlying disease severity and treatment response. *Nat. Commun.* **10**, 38 (2019).

64. Mentch, S. J. & Locasale, J. W. One-carbon metabolism and epigenetics: understanding the specificity. *Ann. N. Y. Acad. Sci.* <https://doi.org/10.1111/nyas.12956> (2016).

65. Boon, R. Metabolic fuel for epigenetic: nuclear production meets local consumption. *Front. Genet.* **12**, 768996 (2021).

66. Ye, C. & Tu, B. P. Sink into the epigenome: histones as repositories that influence cellular metabolism. *Trends Endocrinol. Metab.* **29**, 626–637 (2018).

67. Eram, M. S. et al. Kinetic characterization of human histone H3 lysine 36 methyltransferases, ASH1L and SETD2. *Biochim. Biophys. Acta* **1850**, 1842–1848 (2015).

68. Lurz, E. et al. Vitamin B<sub>12</sub> deficiency alters the gut microbiota in a murine model of colitis. *Front. Nutr.* **7**, 83 (2020).

69. Benight, N. M. et al. B-vitamin deficiency is protective against DSS-induced colitis in mice. *Am. J. Physiol. Gastrointest. Liver Physiol.* **301**, G249–G259 (2011).

70. Tamura, J. et al. Immunomodulation by vitamin B<sub>12</sub>: augmentation of CD8<sup>+</sup> T lymphocytes and natural killer (NK) cell activity in vitamin B<sub>12</sub>-deficient patients by methyl-B<sub>12</sub> treatment. *Clin. Exp. Immunol.* **116**, 28–32 (1999).

71. Liu, M. et al. The histone methyltransferase SETD2 modulates oxidative stress to attenuate experimental colitis. *Redox Biol.* **43**, 102004 (2021).

72. Ward, M. G. et al. Prevalence and risk factors for functional vitamin B<sub>12</sub> deficiency in patients with Crohn's disease. *Inflamm. Bowel Dis.* **21**, 2839–2847 (2015).

73. Zheng, S. et al. Association of ulcerative colitis with transcobalamin II gene polymorphisms and serum homocysteine, vitamin B<sub>12</sub>, and folate levels in Chinese patients. *Immunogenetics* **69**, 421–428 (2017).

74. Leinenkugel, G. et al. Sca-1 is a marker for cell plasticity in murine pancreatic epithelial cells and induced by IFN-β in vitro. *Pancreatology* **22**, 294–303 (2022).

75. Dekel, B. et al. Isolation and characterization of nontubular sca-1+lin- multipotent stem/progenitor cells from adult mouse kidney. *J. Am. Soc. Nephrol.* **17**, 3300–3314 (2006).

76. Camarata, T. D., Weaver, G. C., Vasilyev, A. & Arnaout, M. A. Negative regulation of TGFβ signaling by stem cell antigen-1 protects against ischemic acute kidney injury. *PLoS ONE* **10**, e0129561 (2015).

77. Bender Kim, C. F. et al. Identification of bronchioalveolar stem cells in normal lung and lung cancer. *Cell* **121**, 823–835 (2005).

78. Louie, S. M. et al. Progenitor potential of lung epithelial organoid cells in a transplantation model. *Cell Rep.* **39**, 110662 (2022).

79. Zhang, W. et al. Fluorinated N, N-dialkylaminostilbenes repress colon cancer by targeting methionine S-adenosyltransferase 2A. *ACS Chem. Biol.* **8**, 796–803 (2013).

80. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).

81. Callahan, B. J. et al. DADA2: high-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).

82. Menzel, P., Ng, K. L. & Krogh, A. Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nat. Commun.* **7**, 11257 (2016).

83. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

84. Li, D., Liu, C.-M., Luo, R., Sadakane, K. & Lam, T.-W. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **31**, 1674–1676 (2015).

85. Hyatt, D. et al. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**, 119 (2010).

86. Huerta-Cepas, J. et al. EggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res.* **47**, D309–D314 (2019).

87. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).

88. Carbon, S. et al. AmiGO: online access to ontology and annotation data. *Bioinformatics* **25**, 288–289 (2009).

89. Morton, J. T. et al. Establishing microbial composition measurement standards with reference frames. *Nat. Commun.* **10**, 2719 (2019).

90. Viltard, M. et al. The metabolomic signature of extreme longevity: naked mole rats versus mice. *Aging* **11**, 4783–4800 (2019).

91. Pang, Z., Chong, J., Li, S. & Xia, J. MetaboAnalystR 3.0: toward an optimized workflow for global metabolomics. *Metabolites* **10**, 186 (2020).

92. Peng, X. et al. Inhibition of phosphoinositide 3-kinase ameliorates dextran sodium sulfate-induced colitis in mice. *J. Pharmacol. Exp. Ther.* **332**, 46–56 (2010).

93. Bankhead, P. et al. QuPath: open source software for digital pathology image analysis. *Sci. Rep.* **7**, 16878 (2017).

94. Li, H. et al. The Ink4/Arf locus is a barrier for iPS cell reprogramming. *Nature* **460**, 1136–1139 (2009).

95. Subramanian, A. et al. Gene-set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).

96. R Development Core Team *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2020).

97. Tarasov, A., Vilella, A. J., Cuppen, E., Nijman, I. J. & Prins, P. Sambamba: fast processing of NGS alignment formats. *Bioinformatics* **31**, 2032–2034 (2015).

98. Liao, Y., Smyth, G. K. & Shi, W. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Res.* **47**, e47 (2019).

99. Ashburner, M. et al. Gene Ontology: tool for the unification of biology. the Gene Ontology Consortium. *Nat. Genet.* **25**, 25–29 (2000).

100. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).

101. Li, H. et al. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

102. Lawrence, M. et al. Software for computing and annotating genomic ranges. *PLoS Comput. Biol.* **9**, e1003118 (2013).

103. Gu, Z., Eils, R., Schlesner, M. & Ishaque, N. EnrichedHeatmap: an R/Bioconductor package for comprehensive visualization of genomic signal associations. *BMC Genomics* **19**, 234 (2018).

104. Thorvaldsdottir, H., Robinson, J. T. & Mesirov, J. P. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinform.* **14**, 178–192 (2013).

105. Robinson, J. T. et al. Integrative genomics viewer. *Nat. Biotechnol.* <https://doi.org/10.1038/nbt.1754> (2011).

## Acknowledgements

We are grateful to M. I. Muñoz for assistance with the animal protocols, to the IRB Functional Genomics Unit and CRG/CNAG for library preparation and genomic sequencing, and to the IRB (Functional

Genomics, Biostatistics/Bioinformatics, and Histopathology) and PCB (Animal House) core facilities for general research support. M.K. was funded by the Barcelona Institute of Science and technology (BIST) and Asociación Española Contra el Cáncer (AECC; POSTD18020SERR) and supported by the European Molecular Biology Organization (EMBO). F.P. is supported by a Karolinska Institute Starting Grant, Starting Grant from the Swedish Research Council (2019\_02050\_3). G.K. is supported by the Ligue contre le Cancer (équipe labellisée); Agence National de la Recherche (ANR) – Projets blancs; AMMICa US23/CNRS UMS3655; Association pour la recherche sur le cancer (ARC); Cancéropôle Ile-de-France; Fondation pour la Recherche Médicale (FRM); a donation by Elior; Equipex Onco-Pheno-Screen; European Joint Programme on Rare Diseases (EJPRD); European Research Council Advanced Investigator Award (ERC-2021-ADG, ICD-Cancer, grant no. 101052444), European Union Horizon 2020 Projects Oncobiome, Prevalung (grant no. 101095604) and Crimson; Fondation Carrefour; Institut National du Cancer (INCa); Institut Universitaire de France; LabEx Immuno-Oncology (ANR-18-IDEX-0001); a Cancer Research ASPIRE Award from the Mark Foundation; the RHU Immunolife; Seerave Foundation; SIRIC Stratified Oncology Cell DNA Repair and Tumor Immune Elimination (SOCRATE); and SIRIC Cancer Research and Personalized Medicine (CARPEM). This study contributes to the IdEx Université de Paris ANR-18-IDEX-0001. Work in the laboratory of M.S. was funded by the IRB, 'laCaixa' Foundation, and by grants from the Milky Way Research Foundation, the Spanish Ministry of Science co-funded by the European Regional Development Fund (ERDF; SAF2017-82613-R), the European Research Council (ERC-2014-AdG/669622) and the Secretaria d'Universitats i Recerca del Departament d'Empresa i Coneixement of Catalonia (Grup de Recerca consolidat 2017 SGR 282).

## Author contributions

M.K. designed and performed most experiments, contributed to bioinformatic data analysis, performed quantification of immunohistochemistry and co-wrote the paper. E.M., D.C., F.P., R.B., A.H.-H. and M.R. provided general experimental support. A.C. and C.S.-O.A. designed and performed bacterial bioinformatic analysis. C.S.-O.A. designed and performed RNA-seq and ChIP-seq bioinformatics analysis. A.J., J.C., S.D., O.Y., E.M. and G.K. performed metabolomics experiments. E.M., D.C., F.P., A.J., S.D., O.Y. and G.K. helped design and analyse metabolomics experiments. N.P. performed the histopathological study and supervised the histopathological techniques. M.S. designed and supervised the study, secured funding, analysed the data and co-wrote the paper. All authors discussed the results and commented on the paper.

## Competing interests

G.K. holds research contracts with Daiichi Sankyo, Eleor, Kaleido, Lytix Pharma, PharmaMar, Osasuna Therapeutics, Samsara Therapeutics, Sanofi, Tollys and Vascage. G.K. is on the Board of Directors of the Bristol Myers Squibb Foundation France. G.K. is a scientific cofounder of everImmune, Osasuna Therapeutics, Samsara Therapeutics and Therafast Bio. G.K. is on the scientific advisory boards of Hevolution, Institut Servier and Longevity Vision Funds. G.K. is the inventor of patents covering therapeutic targeting of aging, cancer, cystic fibrosis and metabolic disorders. G.K.'s wife, L. Zitvogel, has held research contracts with Glaxo Smyth Kline, Incyte, Lytix, Kaleido, Innovate Pharma, Daiichi Sankyo, Pilege, Merus, Transgene, 9m, Tusk and Roche, was on the Board of Directors of Transgene, is a cofounder of everImmune, and holds patents covering the treatment of cancer and the therapeutic manipulation of the microbiota. G.K.'s brother, R. Kroemer, was an employee of Sanofi and now consults for Boehringer-Ingelheim. M.K. has ongoing or completed research contracts with Galapagos, Rejuveron Senescence Therapeutics and Mesoestetic. D.C. is an employee of Fenix Group International.

M.S. is shareholder of Senolytic Therapeutics, Life Biosciences, Rejuveron Senescence Therapeutics, and Altos Labs. M.S. was consultant until the end of 2022 of Rejuveron Senescence Therapeutics and Altos Labs. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The other authors declare no competing interests.

## Additional information

**Extended data** is available for this paper at  
<https://doi.org/10.1038/s42255-023-00916-6>.

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s42255-023-00916-6>.

**Correspondence and requests for materials** should be addressed to Marta Kovatcheva or Manuel Serrano.

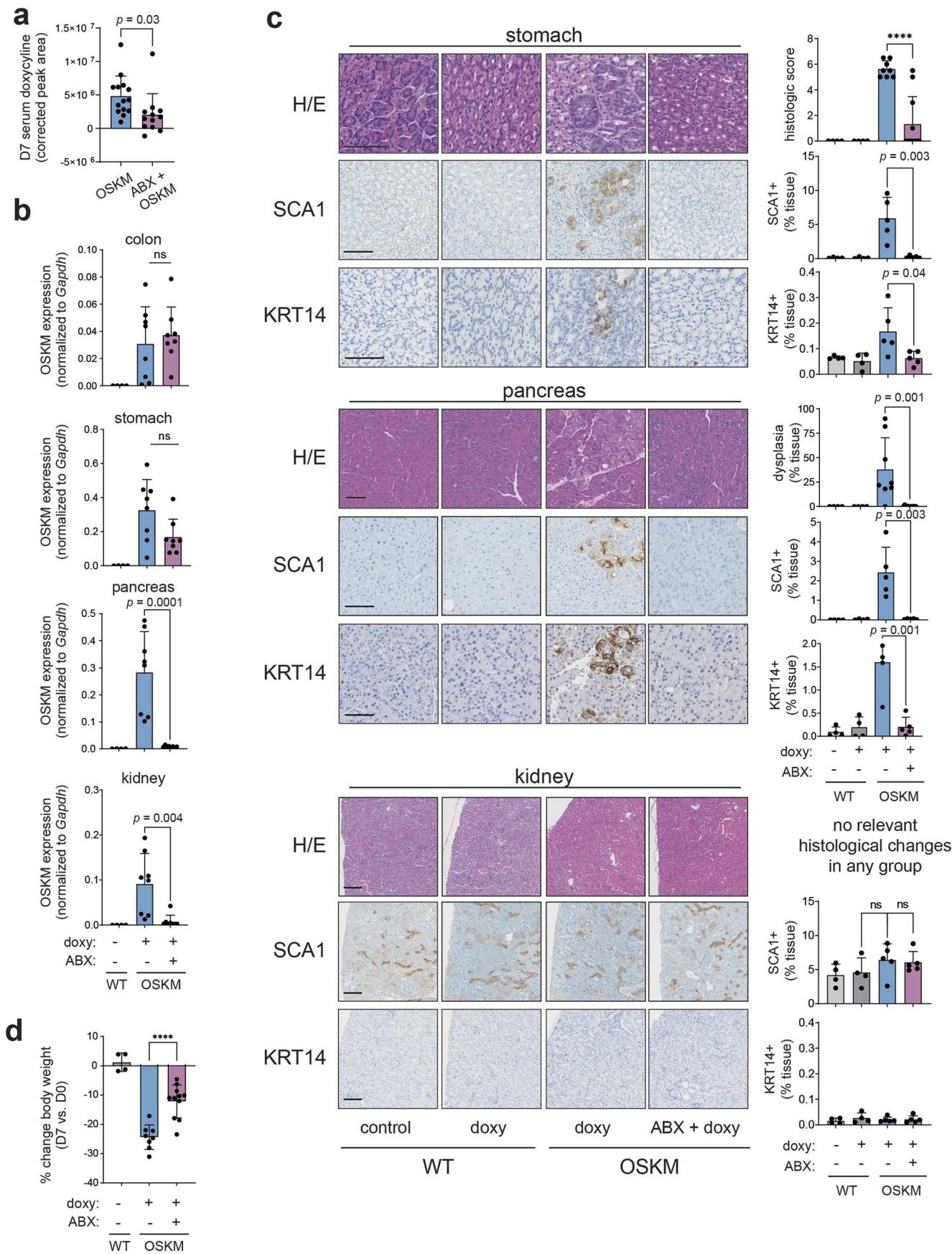
**Peer review information** *Nature Metabolism* thanks Helmuth Gehart and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Yanina-Yasmin Pesch, in collaboration with the *Nature Metabolism* team.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023

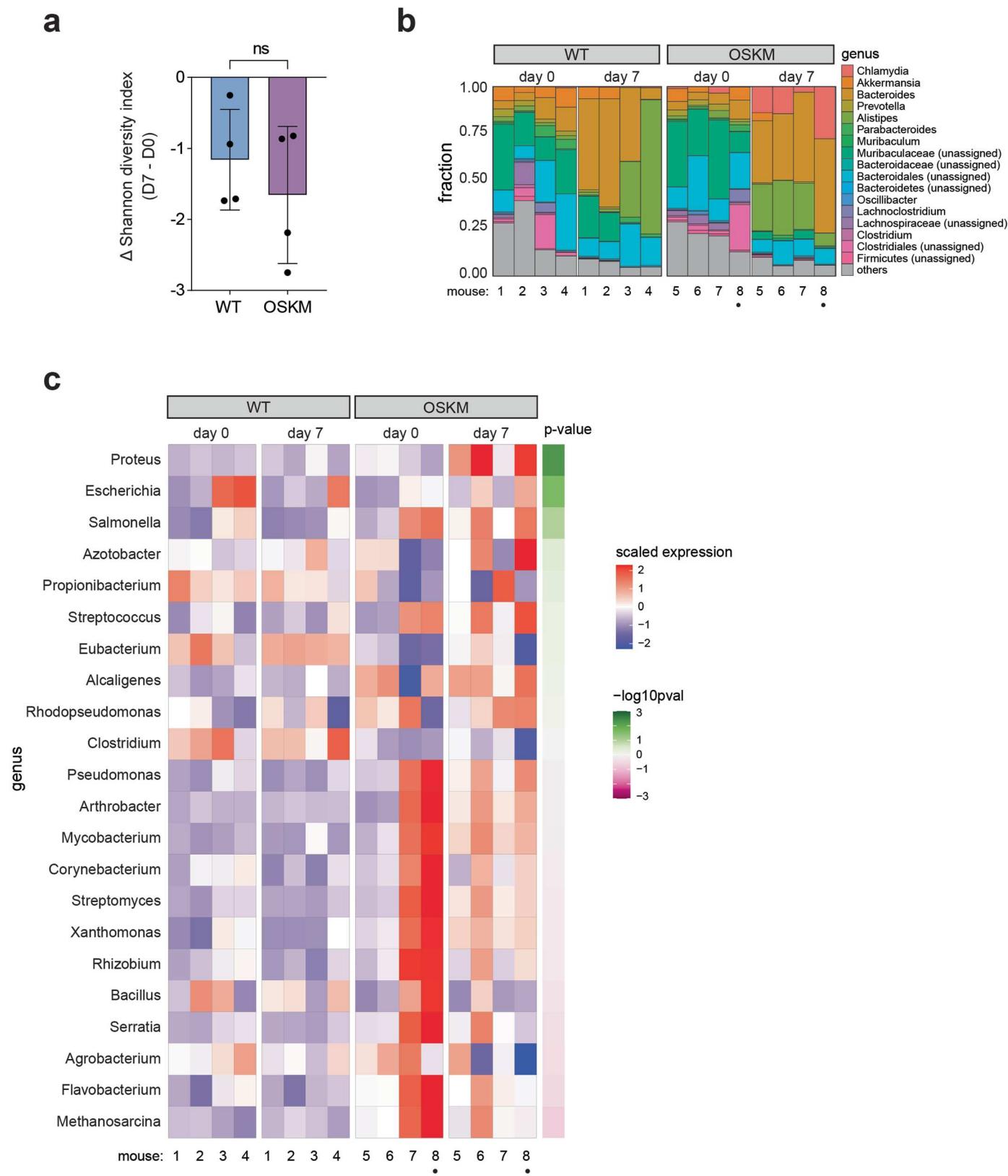


no relevant histological changes in any group

Extended Data Fig. 1 | See next page for caption.

**Extended Data Fig. 1 | Broad-spectrum antibiotic treatment inhibits in vivo reprogramming.** (a-d) Mice were treated with doxycycline (doxy) and antibiotics (ABX) as described in Fig. 1a. (a) Serum was harvested from OSKM mice after 7 days of doxycycline treatment (OSKM n = 14; 5 M 9 F), or after ABX + doxycycline treatment (ABX + OSKM n = 12; 3 M 9 F) and doxycycline levels were analyzed by mass spectrometry. The total area of the doxycycline peak corrected for background levels for each mouse is shown. (b-d) Representative subsets of mice from Fig. 1b n = 4 mice (WT; 3 M 1 F), n = 8 (OSKM + doxy; 4 M 4 F), n = 11 (OSKM + doxy + ABX 4 M 7 F) were used for all analyses (b) Tissues were harvested

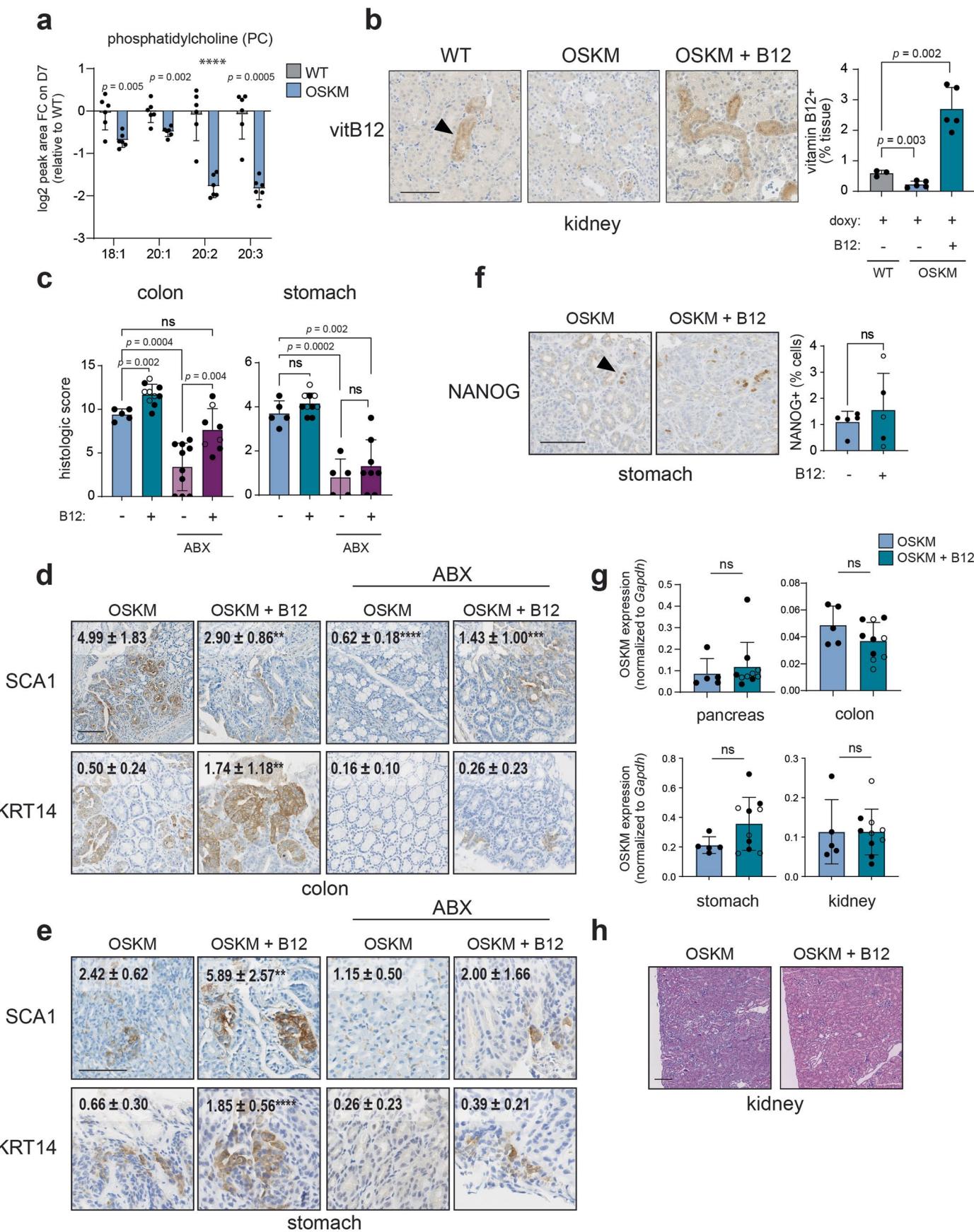
on day 7 post doxycycline initiation and the expression of the OSKM transgene was assessed by RT-qPCR. (c) Representative histology images of the indicated treatments and markers are shown (left) and quantified (right). Histologic score for stomach dysplasia and inflammation was assigned by a blinded pathologist. No relevant morphologic changes were observed in the kidney cortex and medulla in any of the mice. (d) Body weights of mice from the indicated treatment groups, expressed as a percent change on day 7 post-doxycycline as compared to pre-doxycycline (day 0). Scale bars are 100  $\mu$ m. Graphs represent average  $\pm$  SD; \*\*\*p < 0.0001 by two-tailed Student's t-test.



Extended Data Fig. 2 | See next page for caption.

**Extended Data Fig. 2 | Metagenomic analysis of fecal bacteria during *in vivo* reprogramming.** (a-c) Mice were treated with doxycycline for 7 days as described in Fig. 1a. Stool from subset of WT ( $n = 4$ ; 2 M 2 F) and OSKM ( $n = 4$ ; 2 M 2 F) mice was analyzed before (pre; day 0) and after seven days (post; day 7) of doxycycline (stool was collected repeatedly from the same mice). (a) The change in the Shannon diversity index on day 7 as compared to day 0 in the indicated treatment groups. Negative values indicate reduced diversity post-doxycycline. Graph represents average  $\pm$  SD; ns  $p > 0.05$  by two-tailed Student's t-test. (b) Microbial stool diversity at the genus level, as categorized

by 16 S rRNA sequencing. The 17 genera with the highest abundance are shown. “Unassigned” refers to bacterial species that could not be assigned a specific genus. “Others” includes assigned bacteria with low counts, as well as archaea. See Supplementary Table 1 for complete list. Mouse 8, marked with a dot (●), had a large proportion of host (murine) DNA contamination in the stool sample. (c) Normalized abundance of genera capable of de novo synthesis of vitamin B12 in the indicated conditions. Genera are ordered by significance in the day7/day0 (OSKM) vs day7/day0 (WT) interaction, and raw p-values (by two-tailed Wald test (DESeq2)) for this score are shown.

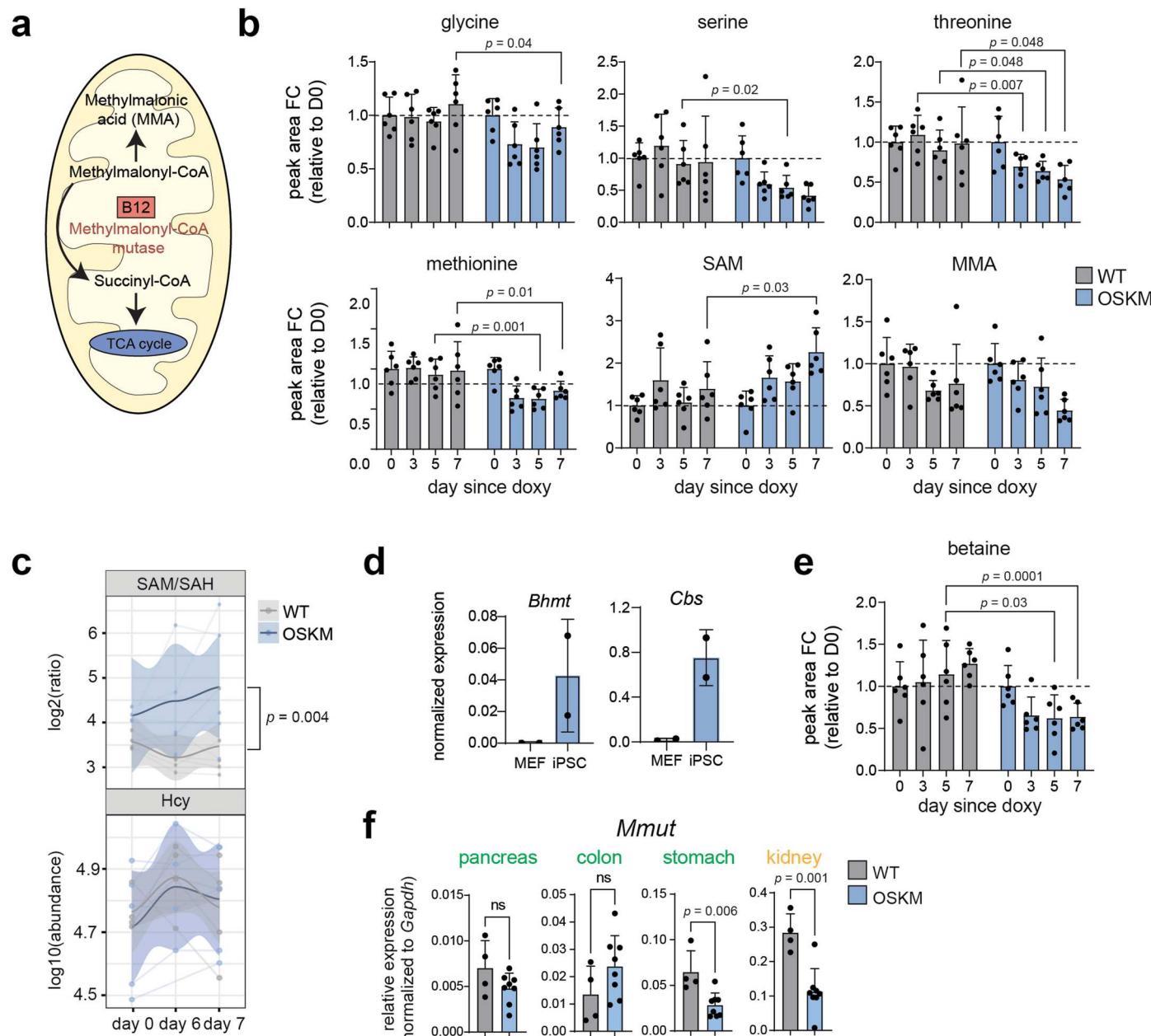


Extended Data Fig. 3 | See next page for caption.

**Extended Data Fig. 3 | Vitamin B12 deficiency limits in vivo reprogramming.**

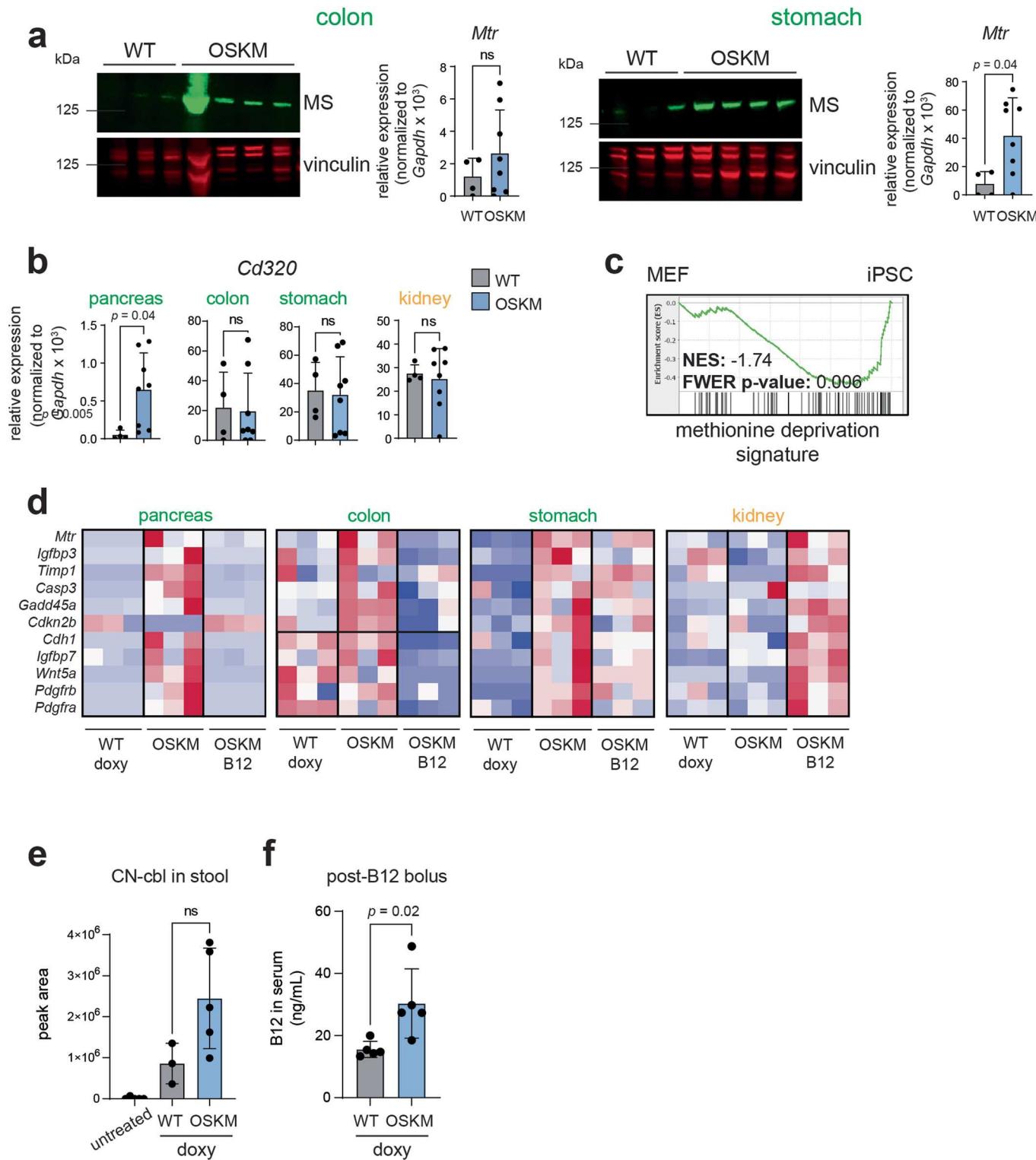
(a) Abundance of phosphatidylcholine (PC) species in the serum during reprogramming. WT and OSKM mice ( $n = 6$  per group; 3 M 3 F) were treated with doxycycline for 7 days and paired serum samples were taken pre- (D0) and post (D7) doxycycline for metabolic analysis by mass spectrometry. The relative abundance of the indicated PC moieties are given as a log2 fold change (FC) as compared to WT on D7. P-values represent the comparison between WT and OSKM groups by two-tailed Student's t-test for each PC moiety. (b) Vitamin B12 immunohistochemistry in the kidney. WT and OSKM mice were treated with doxycycline with or without vitamin B12 as indicated for 7 days. Representative images of kidneys stained for vitamin B12 are shown (left) and total staining is quantified (right). WT doxy ( $n = 3$ ; 2 M 1 F); OSKM doxy ( $n = 5$ ; 5 M) and OSKM doxy + B12 ( $n = 5$ ; 1 M 4 F). (c-h) Analysis of reprogramming in mice treated as described in Fig. 1a, with treatments as indicated. (c) Histologic score of reprogramming was assigned by a blinded histopathologist in the colon (left) and stomach (right).  $n = 5$  mice (OSKM 2 M 3 F and OSKM + ABX, stomach 5 F),

$n = 10$  (OSKM + B12 4 M 6 F and OSKM + ABX, colon 10 F),  $n = 8$  (OSKM + ABX + B12 3 M 5 F). (d-e) Immunohistochemistry of SCA1 and KRT14 in the colon (d) and stomach (e). Representative images and quantification (average  $\pm$  SD; inset) from a subset of  $n = 5$  mice per condition are shown. Asterisks represent comparison to the OSKM condition by one-way ANOVA; \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , \*\*\*\* $p < 0.0001$ . (f-h) Analysis of mice from Fig. 1e treated with doxycycline, with or without vitamin B12 supplementation as indicated. (f) NANOG immunohistochemistry in the stomach of  $n = 5$  representative mice per group. (g) Expression of the OSKM transgene assessed by RT-qPCR in the indicated tissues on D7 post doxy.  $n = 5$  mice (OSKM; 2 M 3 F),  $n = 10$  mice (OSKM + B12; 4 M 6 F). (h) Representative images of H/E staining of the kidney. No relevant histologic changes were observed in the kidney of any of the mice. Scale bars are 100  $\mu$ m. Mice marked by an open circle received both B12 and folate supplementation (n.s. difference for B12 vs B12 + folate; see text for details). Graphs represent average  $\pm$  SD; p-values are by two-tailed Student's t-test unless otherwise indicated.



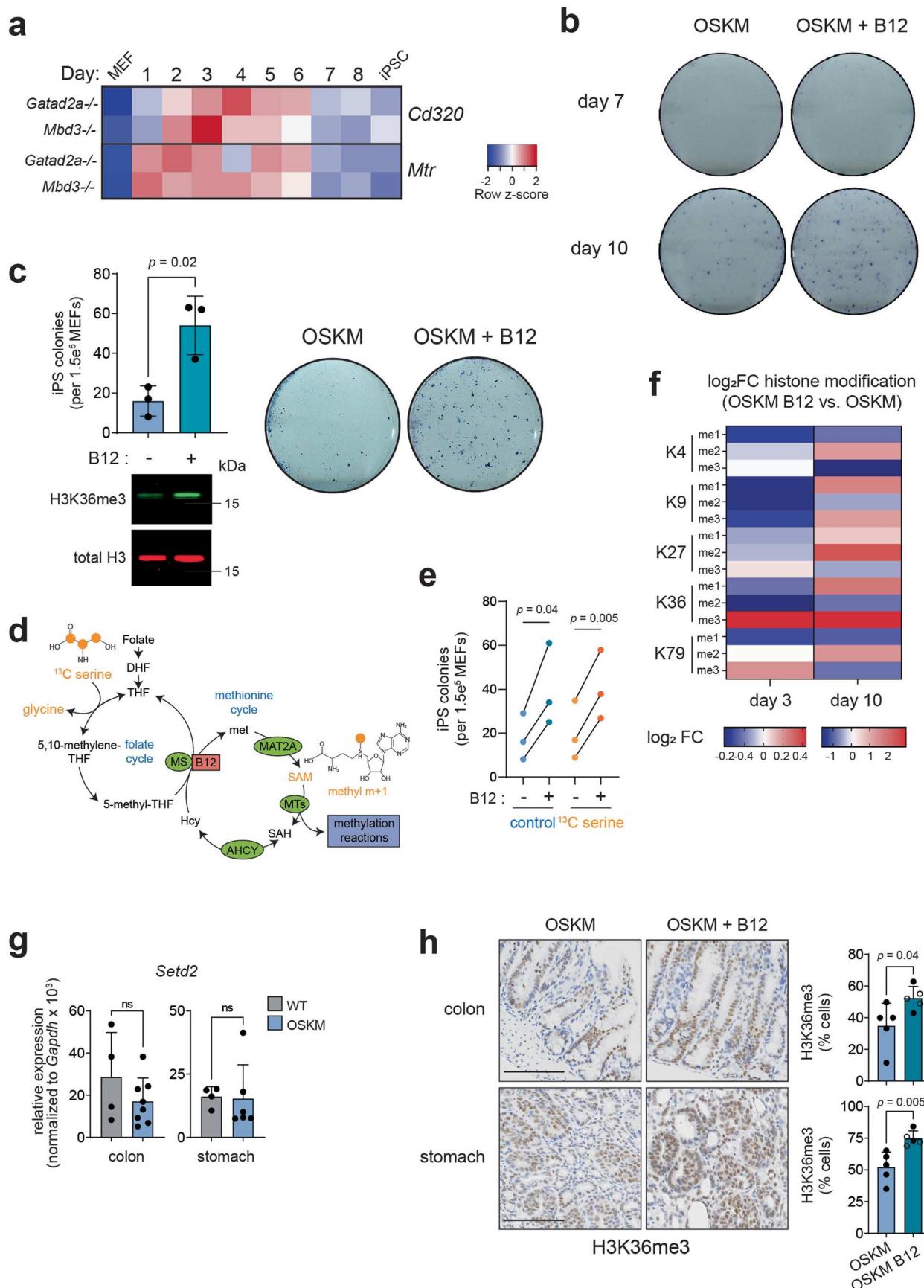
**Extended Data Fig. 4 | Vitamin B12 and 1-carbon (1 C) metabolism markers during in vivo reprogramming.** (a) Schematic of methylmalonyl-coA mutase (MUT) enzymatic activity in the mitochondria. Figure adapted from ref. 26, Springer Nature Limited. TCA: tricarboxylic acid. (b) Metabolites of the Gly/Ser/Thr metabolism pathway (KEGG map00260) identified in the serum of WT and OSKM mice ( $n = 6$  mice per genotype; 3 M 3 F) during a time course of doxycycline administration. Fold change (FC) of the peak area of the indicated metabolites as detected by mass spectrometry is shown relative to (D0) for the respective groups (that is WT and OSKM). P-values represent significant difference between OSKM and WT mice on the indicated day. Serum was collected repeatedly from the same mice. (c) WT and OSKM mice ( $n = 5$  per group; 5 M) were treated with doxycycline and given a bolus of vitamin B12 on day 6, as schematized in Fig. 2f. Serum (repeated measurements) was analyzed by mass spectrometry prior to

doxycycline (day 0), prior to the B12 bolus (day 6), and 24 hours after the bolus (day 7). The ratio of SAM/SAH abundance (upper) and total abundance of homocysteine (Hcy, lower) are shown. P-value represents comparison of WT vs. OSKM by two-way ANOVA. (d) Expression of Bhmt and Cbs in previously published RNA sequencing data from  $n = 2$  paired parental MEFs and induced pluripotent stem cells (iPSC), generated by doxycycline-induced OSKM expression. Data source: 40. (e) Serum analysis of betaine in the mice described in panel (b). (f) Expression of methylmalonyl-coA mutase (Mmut) in WT ( $n = 4$ ) and OSKM ( $n = 8$ ) mice from Fig. 1b. Mice were treated with doxycycline for 7 days and RT-qPCR was performed in tissues that do (pancreas, stomach, colon; green) and do not (kidney; orange) exhibit histologic signs of reprogramming. Bar graphs represent average  $\pm$  SD; p-values by two-tailed Student's t-test unless otherwise indicated.



**Extended Data Fig. 5 | Reprogramming induces a methionine deprivation signature that is relieved by vitamin B12.** (a) Immunoblot (left) and RT-qPCR (right) to measure levels of methionine synthase (encoded by *Mtr*) in the colon and stomach of WT (n = 4, 3 M1 F) and OSKM (n = 8, 4 M 4 F) mice from Fig. 1b treated with doxycycline for 7 days. Representative mice are shown in the immunoblot. (b) Expression of *Cd320* measured by RT-qPCR in the indicated tissues from WT and OSKM mice described in panel (a). (c) Previously published RNAseq data of iPS cells and their parental MEFs (data source: 40) was used to perform gene set enrichment analysis against a published signature of methionine deprivation (msigdb: M13537; ref. 39). (d) RT-qPCR analysis in the indicated tissues from a subset of mice from Fig. 1. The genes were derived from the methionine deprivation signature shown in (c); see Methods for details.

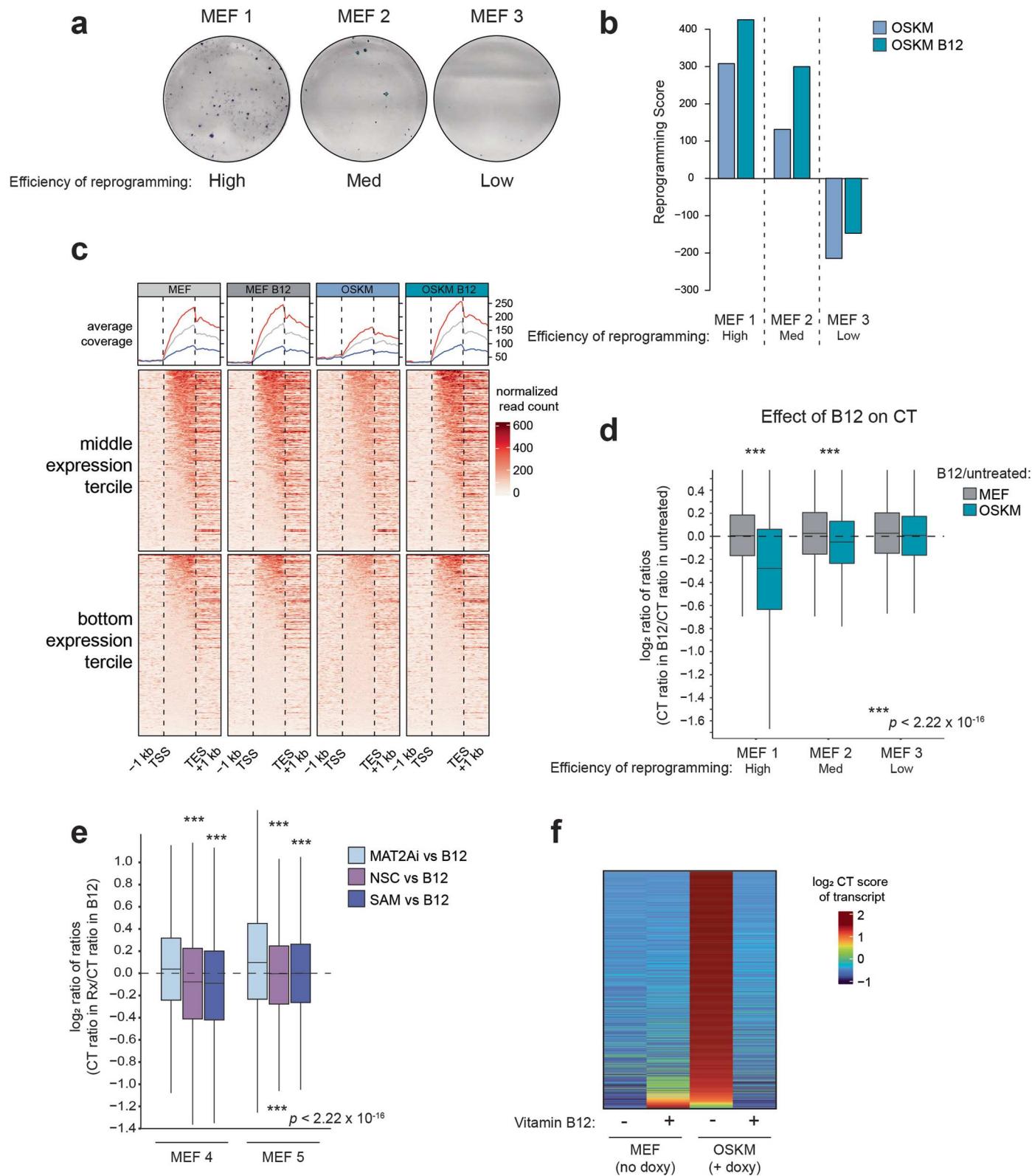
(e-f) Analysis of vitamin B12 levels in the WT and OSKM mice receiving doxycycline and the vitamin B12 bolus, as schematized in Fig. 2f. (e) Cyanocobalamin (CN-cbl) levels in the stool measured by mass spectrometry. Mice were individualized immediately after the bolus and all the fecal matter was collected from the cage 24 hours later. Untreated mice received neither the B12 bolus nor doxycycline. n = 5 mice (untreated), n = 3 (WT doxy), n = 5 (OSKM doxy), all males. (f) Vitamin B12 levels in the serum (holotranscobalamin, holoTC) measured by ADVIA Immunoassay on day 7, 24 hours after the B12 bolus, n = 5 mice (all males) per group. Color of labels indicate tissues that are prone (pancreas, colon, stomach; green) and refractory (kidney; orange) to reprogramming. Bar graphs represent average  $\pm$  SD; p-values by two-tailed Student's t-test.



Extended Data Fig. 6 | See next page for caption.

**Extended Data Fig. 6 | The impact of vitamin B12 administration on in vitro reprogramming.** (a) Changes in gene expression of the vitamin B12 uptake receptor Cd320 and Mtr (encoding MS) during a previously published time course of iPSC generation by OSKM expression. iPSCs were derived from MEFs of two different genetic backgrounds (Gata2a-/- and Mbd3-/-) that both yield high-efficiency reprogramming. Each row is an independent embryo of the indicated genotype. Data source: 40 (b) Vitamin B12 supplementation does not accelerate iPSC formation. OSKM MEFs were treated with doxycycline with or without vitamin B12 supplementation as indicated, and iPSC colony formation was assessed by alkaline phosphatase staining after 7 and 10 days of treatment. Representative images are shown of MEFs derived from a single embryo;  $n = 3$  independent embryos were tested. (c) Vitamin B12 enhances doxycycline-free retroviral reprogramming. Wild type MEFs were transduced with retroviral vectors encoding Oct4, Sox2, Klf4, and cMyc. Samples were collected on day 5 post-transduction for histone extraction and immunoblot (lower). iPSC colony formation was evaluated by alkaline phosphatase (AP) staining on day 14 (upper). Immunoblot and AP images are representative from  $n = 3$  independent MEFs. (d-e) Stable isotope labeling (SIL) with fully labeled 13C-serine during in vitro reprogramming. (d) Schematic of 1C metabolism indicating how fully labeled 13C-serine contributes to methyl m + 1 labeling of SAM. Orange points indicate 13C atoms. Figure adapted from ref. 33, Springer Nature Limited. (e) SIL does not affect reprogramming efficiency. Doxycycline-inducible OSKM MEFs were cultured in normal iPS media with doxycycline without or without vitamin B12 supplementation as indicated. 13C-serine labeled media was administered as described in Methods and reprogramming was quantified by iPS colonies evaluated by AP staining on day 10. Quantification of colonies is shown for the indicated treatments. Each point represents an individual MEF clone. P-values

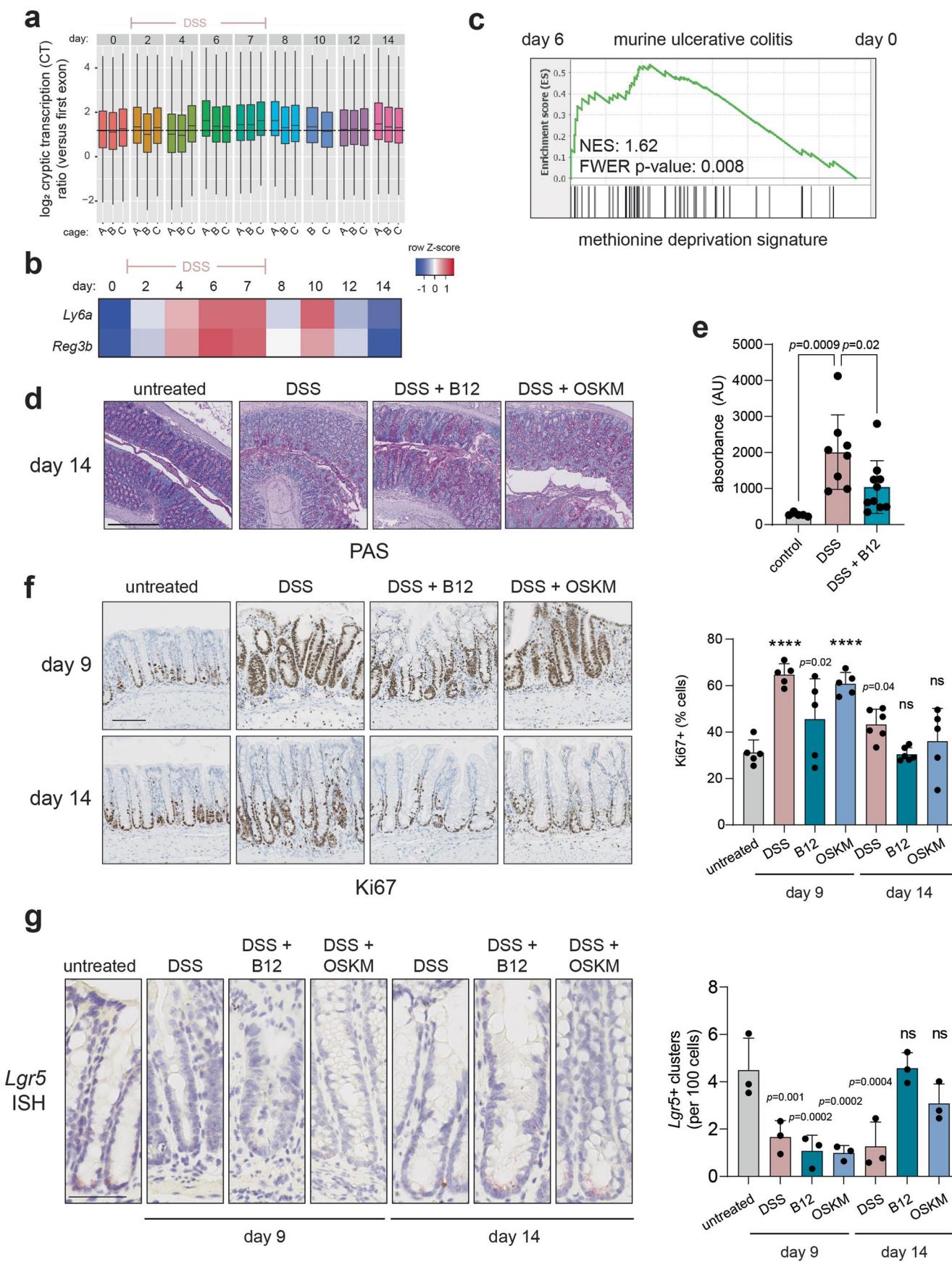
represent comparison of the OSKM vs OSKM + B12 condition for each media type. (f) Changes in histone methylation during in vitro reprogramming with vitamin B12 supplementation. Histone extracts were prepared from OSKM MEFs treated with doxycycline for 3 and 10 days, with or without vitamin B12 supplementation. Levels of the indicated histone methylation modifications were measured by colorimetric multiplex assay and the relative abundance (fold change, FC) in cells reprogrammed in the presence as compared to absence of vitamin B12 is shown. Note that MEFs derived from independent embryos were used for day 3 and day 10 analyses. (g) Expression of Setd2, which encodes the H3K36 trimethyl-transferase in the colon (left) and stomach (right) from WT and OSKM mice after 7 days of doxycycline treatment. Mice are from Fig. 1b  $n = 4$  mice (WT),  $n = 8$  (OSKM colon),  $n = 6$  (OSKM stomach). (h) H3K36me3 during in vivo reprogramming in the colon and stomach of OSKM mice with and without B12 supplementation. Representative images are shown (left) and quantified (right) in a subset ( $n = 5$  for each group) of mice from Fig. 1e. Scale bar is 100  $\mu$ m. Mice that received folate in addition to B12 are represented by open points. Graphs represent average  $\pm$  SD. P-values by two-tailed Student's t-test, used for day 3 and day 10 analyses. (g) Expression of Setd2, which encodes the H3K36 trimethyl-transferase in the colon (left) and stomach (right) from WT and OSKM mice after 7 days of doxycycline treatment. Mice are from Fig. 1b  $n = 4$  mice (WT),  $n = 8$  (OSKM colon),  $n = 6$  (OSKM stomach). (h) H3K36me3 during in vivo reprogramming in the colon and stomach of OSKM mice with and without B12 supplementation. Representative images are shown (left) and quantified (right) in a subset ( $n = 5$  for each group) of mice from Fig. 1e. Scale bar is 100  $\mu$ m. Mice that received folate in addition to B12 are represented by open points. Graphs represent average  $\pm$  SD. P-values by two-tailed Student's t-test.



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

**Extended Data Fig. 7 | Reprogramming and transcriptional dynamics across MEF clones.** (a) The reprogramming efficiency of three independent MEFs was measured by the formation of iPS colonies as stained with alkaline phosphatase 10 days after continuous addition of doxycycline to the media. (b) The Reprogramming Score was calculated using transcriptional data from MEFs expressing doxycycline-inducible OSKM for 72 h, whose cell surface markers indicated they were poised for high-efficiency reprogramming<sup>48</sup>. See Methods for details. (c) H3K36me3 chromatin immunoprecipitation (ChIP) sequencing was performed in Clone 1, 72 h after the addition of doxycycline and/or vitamin B12, as indicated. The normalized ChIP reads in gene bodies relative to all aligned reads are shown for the middle and bottom tercile in terms of gene expression (as determined by RNA-seq RPKM levels across all condition) are shown (lower). The trace of the ChIP signal is shown, where the y-axis represents average coverage across genes (upper; blue = bottom tercile; grey = middle tercile; red = top tercile, shown for reference; see Fig. 4a for top tercile). TSS – transcription start site; TES – transcription end site. (d-e) The ratio of cryptic transcription (CT). (d) CT in B12-treated cells as compared to non-B12 is shown for the median

CT ratio for all genes for which a score could be calculated is plotted. Individual MEFs ( $n = 3$  total) are shown, separated into MEF (no doxycycline) and OSKM (72 h of doxycycline treatment to induce OSKM expression) conditions. Asterisks represent the comparison between the MEF and OSKM bars. (e) OSKM cells were treated with doxycycline and the indicated compounds for 72 h, as described in Fig. 4c. The ratio of CT is plotted as a ratio of the CT in the B12 condition for each treatment, in  $n = 2$  independent MEFs. Asterisks represent the comparison between the indicated treatment and the MAT2Ai treatment. The lower and upper hinges correspond to the first and third quartiles (the 25th and 75th percentiles). The upper (lower) whisker extends from the hinge to the largest (smallest) value no further than  $1.5 * \text{IQR}$  from the hinge. P-values were determined by two-tailed unpaired Wilcoxon test no multiple comparisons adjustment (f) Genes whose CT increases upon reprogramming (top 25% comparing OSKM vs. MEF) and decrease with vitamin B12 supplementation (top 25% comparing OSKM B12 vs. OSKM) in Clone 1, used for the hypergeometric analysis in Fig. 4d. The CT ratio of individual genes is shown for each condition.



**Extended Data Fig. 8 | See next page for caption.**

**Extended Data Fig. 8 | Recovery from DSS colitis is improved by OSKM or vitamin B12.** (a-c) Previously published time course of bulk colonic RNAseq during DSS injury and recovery Data source: GSE131032<sup>33</sup>. Mice were treated with 2.5% DSS for 7 days as indicated. (a) CT analysis for all genes for which a score could be calculated are shown. Each boxplot represents one biologic replicate (that is, mouse), with the cages indicated as described by the authors of the original study. Dashed line represents median CT value at D0. (b) RNA-sequencing expression of the indicated genes. Values represent the scaled log2FC as compared to D0 across all samples as computed by the authors of the original study. (c) Gene set enrichment analysis of a published signature of methionine deprivation (msigdb: M13537 ref. 39) in samples from day 6 vs day 0 of DSS treatment as computed by the authors of the original study. (d-g) Recovery from DSS injury of mice in Fig. 5, with treatment as schematized in Fig. 5a; a representative subset of mice was used for analyses. (d) Periodic acid-Schiff

(PAS) staining, which marks mucous produced by differentiated goblet cells, in colonic sections of mice on day 14. n = 5 mice from each group were analyzed with similar results; representative images are shown. (e) Intestinal barrier permeability as measured by FITC-dextran leak assay. Fluorescence intensity was measured in the serum of mice with the indicated treatments on day 9. n = 5 mice (control), n = 8 (DSS), n = 10 (DSS + B12). (f) Ki67 was quantified within the colonic epithelium (right) and representative immunohistochemistry images are shown (left). P-values indicate comparison to the untreated control. Scale bar is 100  $\mu$ m. n = 6 mice (DSS D14), n = 5 all other groups. (g) Analysis of Lgr5 by single molecule in situ hybridization (ISH). The number of Lgr5+ clusters per 100 cells in the colonic crypts was quantified from at least 20 crypts per mouse (right) from n = 3 mice and representative images are shown (left). Asterisks indicate comparison to the untreated control. Scale bar is 50  $\mu$ m. Bar graphs represent average  $\pm$  SD; \*\*\*p < 0.0001 by one-way ANOVA.

## Reporting Summary

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

### Statistics

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

n/a Confirmed

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

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

#### Data collection

Histology: NanoZoomer-2.0 HT C9600 digital scanner (Hamamatsu)  
 bacterial sequencing: HiSeq 2500 (Illumina)  
 transcobalamin levels: ADVIA Centaur Immunoassay System (SIEMENS)  
 serum metabolomics: 7890B gas chromatography, RRLC 1260 system, 6490 triple-quadrupole mass spectrometer (Agilent Technologies), Dionex Ultimate 3000 UHPLC system, Orbitrap ID-X Tribrid Mass Spectrometer (Thermo Scientific) as specified in methods  
 immunoblots: LiCor Odyssey FC  
 qPCR: LifeTechnologies Quant Studio 6 Flex System  
 C13 stable isotope labeling metabolomics: UHPLC 1290 Infinity II Series coupled to a QqQ/MS 6490 Series from Agilent Technologies  
 ChIP-seq: NextSeq550 (Illumina)  
 RNA-seq: NovaSeq6000 S2 (Illumina) and NextSeq550 (Illumina)  
 FITC-dextran leak assay: ioTek Synergy H1 Microplate Reader

#### Data analysis

General statistical analyses: GraphPad Prism 9.0.0 with the statistical tests as indicated in each figure.  
 bacterial analysis: Microbiome data analyses were performed in the R programming language version 4.0.5. Reads that aligned to the mm10 genome, STAR (v2.7.0a), were filtered out. Taxonomic assignments were carried out through Kaiju (v1.7.0). Annotations were extracted from the microbial subset of the NCBI BLAST non-redundant protein database (nr). Cleaned sequences were assembled into contigs using megahit (v1.2.4). Open reading frames from the obtained contigs were predicted with prodigal (v2.6.3). Protein mapping as well as Kegg and COG annotations were obtained using the eggNOG mapper (v2.0.0). Alignment to the assembled transcriptome was performed with Bowtie2 (v2.2.2). Sam files were sorted and converted to BAM format using samtools (v0.6.7).  
 Serum metabolic pathway impact: Global ANOVA pathway enrichment and Out-degree Centrality Topology analysis through the MetaboAnalyst 4.0 software, using KEGG library (2019) as reference.

All RNASeq analyses were performed in the R programming language version 4.0.5. Reads were aligned to the mm10 mouse genome using STAR (v2.7.10a). Annotations were extracted from ENSEMBL (vGRC138.97). ChIPseq data was aligned using Bowtie (v0.12.9). Sam files were sorted and converted to BAM format using sambamba (v0.6.7). IGVtools (v2.9.4) was used to generate alignment visualizations. Histology was quantified using QuPath 0.4.3.

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

## Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

All sequencing data are deposited under the following GEO accessions:

GSE154149 Microbial genome analysis from OSKM and WT mice pre-treatment and after 7 days of 1 mg/mL treatment with doxycycline in the drinking water.

GSE200578 ChIP-Seq of H3K36me3 samples of OSKM reprogramming mouse embryonic fibroblasts treated with B12

GSE200579 RNA-Seq samples of OSKM reprogrammable MEFs treated with or without vitamin B12

GSE232382 RNA-Seq samples of OSKM reprogrammable MEFs treated with vitamin B12 and/or various compounds to modulate SAM or histone methylation

Previously published datasets that were used for analysis in this study are:

GSE131032 RNA-Seq of time course of repairing murine epithelium after DSS injury

GSE109142 RNA-Seq of human pediatric ulcerative colitis and normal tissue controls

GSE102518 RNA-Seq of murine in vitro reprogramming in MEFs of varying genotypes

GSE77722 RNA-Seq of murine in vivo reprogramming in mice of varying genotypes

## Human research participants

Policy information about [studies involving human research participants and Sex and Gender in Research](#).

Reporting on sex and gender

n/a

Population characteristics

n/a

Recruitment

n/a

Ethics oversight

n/a

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Field-specific reporting

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

Life sciences

Behavioural & social sciences

Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](http://nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

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

Sample size

No statistical methods were used to predetermine sample size in the mouse studies. As requested by Reviewer 2, for the revision, we conducted a power analysis using the calculator published by the Chinese University of Hong Kong (<http://www.lasec.cuhk.edu.hk/sample-size-calculation.html>) to assess and appropriately power the analysis of the data presented in Fig. 1d

Data exclusions

In our ChIPseq (and corresponding RNAseq), we used n=3 independent MEFs. As clearly explained in the text and shown in the figures, two clones had low-to-no inherent reprogramming capacity (biologic variability) so we selected the n=1 clone with high efficiency for ChIP.

Replication

All replicate values (as indicated in Figure Legends) represent biological replicates. For qPCR experiments, biological replicates (i.e. independent mice, MEFs) were used and indicated in the n for each figure, and each sample was run in technical triplicate. For in vitro experiments, replicates were generally performed on the same day with n=2-3 cell lines as indicated, each derived from independent mouse embryos; independent experiments always included OSKM (doxy) and OSKM + B12 controls. For the fecal bacterial analysis, stool samples from biologic replicates (i.e. independent mice) were collected on the same days. For the serum metabolomic analysis, two independent

experiments were performed (i.e. untargeted metabolomics and targeted 1C metabolomics); samples for each experiment were collected on the same days/times from independent mice, and time was maintained across experiments. In vivo reprogramming experiments were done in the span of months-to-years as the project evolved; in general, independent panels represent cohorts of mice that were treated together and sacrificed together. For the DSS experiments, results represent data from pooled experiments as indicated in the figure legend, with each experiment having at least n=3 DSS and DSS + B12 control mice. All attempts at replication were successful except as indicated regarding the variability of reprogramming in the ChIP experiment.

Randomization	Mice were randomized into experimental groups based on age (i.e. always trying to maintain consistent 8-16 week age) and, when possible, initial body mass.
Blinding	All histological scoring was conducted by a blinded histopathologist.

## Reporting for specific materials, systems and methods

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

### Materials & experimental systems

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern

### Methods

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

## Antibodies

### Antibodies used

IHC: NANOG D2A3 (8822, CellSignaling), Sca1/Ly6A/E [E13 161-7] (ab51317, Abcam), H3K36me3 (Lys36) (D5A7) XP (4909, CellSignaling), Ki67 (ab15580, Abcam), Vitamin B12 (Sigma, V9505). Primary antibodies were diluted 1:100 for NANOG, 1:2000 for Ki67, 1:1000 for Ly6a/Sca1, 1:400 for H3K36me3, 1: 2000 for Keratin 14 and 1:750 for Vit B12.

ChIP: H3K36me3 (Lys36) (D5A7) XP (4909, CellSignaling)

Immunoblot: H3K36me3 (Lys36) (D5A7) XP (4909, CellSignaling) 1:1000, methionine synthase (Abcam#ab66039) 1:1000, vinculin (SCBT#sc-73614) 1:5000, total histone H3 (CellSignaling #3638) 1:1000; LI-COR fluorescent secondary reagents (IRDye 800 CW anti mouse #926-32210; IRDye 680 CW anti mouse #926-68070; IRDye 800 CW anti rabbit #926-32211; IRDye 680 CW anti mouse #926-68071) all used at 1:10 000

### Validation

all IHC antibodies were validated by our histopathology facility, and IHC protocols include a secondary-only control. ChIP antibody has been validated by the manufacturer for use in mouse cells.

## Eukaryotic cell lines

### Policy information about [cell lines](#) and [Sex and Gender in Research](#)

#### Cell line source(s)

all primary mouse embryonic fibroblasts (MEFs) were generated in-house from E13.5 embryos. Cells were genotyped to determine whether they were OSKM transgenic. Cells were not sexed. HEK293T cells were purchased from the ATCC.

#### Authentication

all MEFs were made in-house from fresh embryos and were not authenticated further. 293T were purchased from the ATCC and come with a certificate of analysis.

#### Mycoplasma contamination

all MEFs were derived from embryos in-house and were used in experiments as P1; thus, no mycoplasma testing was carried out on these cells. Mycoplasma testing on all cell lines in culture is carried out once per month, and any positive cells are discarded so that no cell lines within the lab are mycoplasma positive.

#### Commonly misidentified lines (See [ICLAC](#) register)

HEK293T do not appear in the ICLAC register

## Animals and other research organisms

### Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research, and [Sex and Gender in Research](#)

#### Laboratory animals

C57/BL6 mice and OSKM mice (which are on a C57/BL6 background) aged 8-16 weeks of mixed gender were used in all experiments.

Laboratory animals	OSKM mice refer to the i4F-B strain in Abad et al Science 2013.
Wild animals	study did not involve wild animals
Reporting on sex	Sex was not considered in study design and animals of both sexes were used in all experiments, keeping as close to 50/50 as possible, as litters allowed; specific male (M) and female (F) ratios for each in vivo experiment are indicated in the figure legends next to the total n (ex. n=5; 3M 2F). We have not done sex-based analysis, as separating sex would limit the sample size.
Field-collected samples	study did not involve field collected sample
Ethics oversight	Animal experimentation at the IRB Barcelona was performed according to protocols approved by the Science Park of Barcelona (PCB) Ethics Committee for Research and Animal Welfare.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## ChIP-seq

### Data deposition

- Confirm that both raw and final processed data have been deposited in a public database such as [GEO](#).
- Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links  
*May remain private before publication.*

<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE200578>

Files in database submission

GSM6038017\_0142\_2022\_0071\_2022\_MACS2\_broad\_peaks\_peaks.txt.gz  
GSM6038018\_0143\_2022\_0072\_2022\_MACS2\_broad\_peaks\_peaks.txt.gz  
GSM6038021\_0144\_2022\_0073\_2022\_MACS2\_broad\_peaks\_peaks.txt.gz  
GSM6038022\_0145\_2022\_0074\_2022\_MACS2\_broad\_peaks\_peaks.txt.gz

RAW data has been deposited to the Sequence Read Archive (SRA) and will be made public after publication

Genome browser session  
(e.g. [UCSC](#))

[http://genome.ucsc.edu/s/yocamille/MKovatcheva\\_2022](http://genome.ucsc.edu/s/yocamille/MKovatcheva_2022)

### Methodology

Replicates	Only one of three MEFs was analyzed in ChIP due to poor inherent reprogramming efficiency of others, as described in the paper. Input and IP samples were generated for each condition.
Sequencing depth	reads processed 331910925 uniquely aligned 233201616 failed to align 32857274 multialigned 65852035  Reads were 75bps long and single end
Antibodies	H3K36me3 (D5A7) (CellSignaling #4909)
Peak calling parameters	MACS2 v. 2.2.6 was used to call peaks using parameters: callpeak --broad -g 3209286105 with INPUT sample as control
Data quality	MACS2peaks (qvalue < 0.1) dox_noB12 116853 MEF_noB12 91216 MEF_B12 106819 dox_noB12 118386
Software	Aligner: bowtie v1.2.3 with parameters: -n 2 -p 30 -m 1 --phred33-quals Sam to bam and sorting: sambamba v0.6.7 Peak caller: MACS2 v. 2.2.6 with parameters: callpeak --broad -g 3209286105