

RESEARCH ARTICLE

STEM CELLS AND REGENERATION

The H3K4 methyltransferase Setd1a is first required at the epiblast stage, whereas Setd1b becomes essential after gastrulation

Anita S. Bledau¹, Kerstin Schmidt¹, Katrin Neumann¹, Undine Hill¹, Giovanni Ciotta², Ashish Gupta², Davi Coe Torres², Jun Fu², Andrea Kranz², A. Francis Stewart^{2,*} and Konstantinos Anastassiadis^{1,*}

ABSTRACT

Histone 3 lysine 4 (H3K4) methylation is a universal epigenetic mark. In mammals, there are six H3K4 methyltransferases related to yeast Set1 and fly Trithorax, including two orthologs of Set1: Setd1a and Setd1b. Here we show that mouse Setd1a is required for gastrulation, whereas Setd1b-deficient embryos survive to E11.5 but are grossly retarded. *Setd1a* knockout embryos implant but do not proceed past the epiblast. Furthermore, Setd1a is not required until the inner cell mass has formed, at which stage it has replaced Mll2 as the major H3K4 methyltransferase. Setd1a is required for embryonic, epiblast and neural stem cell survival and neural stem cell reprogramming, whereas Setd1b is dispensable. Deletion of *Setd1a* in embryonic stem cells resulted in rapid losses of bulk H3K4 methylation, pluripotency gene expression and proliferation, with G1 pileup. Setd1b overexpression could not rescue the proliferation defects caused by loss of Setd1a in embryonic stem cells. The precise developmental requirement for Setd1a suggests that gastrulation is regulated by a switch between the major H3K4 methyltransferases.

KEY WORDS: Setd1a, Setd1b, ESC, Reprogramming, H3K4 methylation

INTRODUCTION

Nucleosomes not only serve to package and organize DNA but also as sites for post-translational modifications that contribute to the status of chromatin. These post-translational modifications, notably acetylation and methylation, modify the physical properties of chromatin and are binding sites for factors and complexes that promote the local chromatin status (Bannister and Kouzarides, 2011; Suganuma and Workman, 2011). Methylations on the histone 3 (H3) tail appear to be primary determinants of chromatin status, as exemplified by methylation at H3 lysine 4 (H3K4) or H3K9, which characterize euchromatic or heterochromatic regions, respectively (Noma et al., 2001). Chromatin is further characterized by the degree of methylation, with mono-, di- and trimethylation of lysines indicating a different status. In particular, all expressed euchromatic genes are marked by trimethylated H3K4 (H3K4me3) on promoter nucleosomes, whereas H3K4me2 is found in transcribed regions (Bannister and Kouzarides, 2011; Suganuma and Workman, 2011).

Yeasts have only one H3K4 methyltransferase (Briggs et al., 2001), originally called the Set1 complex (Set1C) (Roguev et al., 2001; Krogan et al., 2002; Noma and Grewal, 2002; Roguev et al., 2003), and the distinction between mono-, di- and trimethylated H3K4 sites is mediated by regulation of its enzymatic SET domain (Kim et al., 2013) or by opposing demethylases (Seward et al., 2007; Kooistra and Helin, 2012). Yeast Set1C contains eight subunits and was the first linkage between H3K4 methylation and Trithorax group (Trx-G) action (Roguev et al., 2001). This linkage was based on two observations: (1) one of the eight Set1C subunits, Bre2, is the yeast ortholog of the fly Trx-G protein Ash2; (2) among the diversity of SET methyltransferase domains, the SET domain of yeast Set1 is almost identical to that of *Drosophila* Trx. Because yeast Set1C is an H3K4 methyltransferase, it was likely that *Drosophila* Trx and the mammalian orthologs Mll1 and Mll2 would also be H3K4 methyltransferases, as was subsequently proven (Milne et al., 2002).

Drosophila has three proteins carrying the Set1/Trithorax-type SET domain, namely dSet1, Trx and Trithorax-related (Trx), whereas mammals have six (Glaser et al., 2006; Ardehali et al., 2011; Mohan et al., 2011; Hallson et al., 2012; Herz et al., 2012). These six are encoded by three pairs of sister genes, including two orthologs of Set1: Setd1a and Setd1b. All six proteins (Mll1–4, Setd1a,b) have been found in complexes based on the same Ash2, Wdr5, Rbbp5 and Dpy30 scaffold as yeast Set1C (yeast Bre2/Ash2, Swd3, Swd1, Sdc1) (Miller et al., 2001; Roguev et al., 2001; Hughes et al., 2004; Yokoyama et al., 2004; Ruthenburg et al., 2007; Ardehali et al., 2011; Mohan et al., 2011; Hallson et al., 2012). Because H3K4 methylation is the most conserved epigenetic modification, the Ash2 scaffold is the most highly conserved core protein complex involved in epigenetic regulation.

Biochemically, the two mammalian Set1 orthologs have been found in virtually identical complexes as yeast Set1C. However, Setd1a and Setd1b do not coexist in the same complex (Lee et al., 2007). All Set1Cs contain a PHD finger protein called Cxxc1 or Cfp1 in mammals and Spp1 in yeast. This PHD finger binds H3K4me3 (Shi et al., 2007; Eberl et al., 2013). Hence, Set1C is a complex that can both catalyze trimethylation of H3K4 and bind it as well. Potentially, this describes a feed-forward, positively reinforcing loop that could propagate the status of promoter chromatin through cell division and/or add epigenetic momentum to transcriptional decisions.

Given the universal roles played by H3K4 methylation in eukaryotic gene expression, why do yeasts have one, flies three and mammals six Set1/Trithorax-type H3K4 methyltransferases? This question cannot be answered as yet, but recent evidence from flies points to functional specialization as part of the answer. *Drosophila* Set1 is the major bulk H3K4 di- and trimethyltransferase (Ardehali

¹Stem Cell Engineering, Technische Universität Dresden, BioInnovationsZentrum, Tatzberg 47, Dresden 01307, Germany. ²Genomics, Biotechnology Center, Technische Universität Dresden, BioInnovationsZentrum, Tatzberg 47, Dresden 01307, Germany.

*Authors for correspondence (francis.stewart@biotec.tu-dresden.de; konstantinos.anastassiadis@biotec.tu-dresden.de)

et al., 2011; Mohan et al., 2011; Hallson et al., 2012), Trx is restricted to a limited number of regulatory aspects such as homeotic gene regulation (Breen and Harte, 1993; Kuzin et al., 1994), and Trx, which is a co-factor for the Ecdysone receptor (Sedkov et al., 2003), might be a monomethyltransferase with enhancer specificity (Herz et al., 2012). In mice, only the two Trx orthologs Mll1 and Mll2 have been examined sufficiently to compare their roles. Despite ubiquitous expression during development, the two knockouts revealed very different roles. Mll1 is first required at ~E12.5 for definitive hematopoiesis (Yagi et al., 1998; Ernst et al., 2004), whereas Mll2 is widely required shortly after gastrulation from ~E7.5 onwards (Glaser et al., 2006). During development and in the adult neither enzyme is required for bulk H3K4 methylation in any cell type examined. However, Mll2 is the major di- and trimethyltransferase in developing oocytes and during the first cell divisions after fertilization (Andreu-Vieyra et al., 2010).

Although Set1 is the most conserved epigenetic regulator in evolution, the function of neither Setd1a nor Setd1b has been studied in mammalian development before now. Knockout of the Set1C subunit *Cxxc1* (also called *Cfp1*) resulted in peri-implantation lethality in mice (Carlone and Skalnik, 2001). Deletion of *Cxxc1* in embryonic stem cells (ESCs) led to a reduction of H3K4me3 levels particularly at promoters of highly transcribed genes (Clouaire et al., 2012). However, ESCs lacking *Cxxc1* are viable and express pluripotency markers (Carlone et al., 2005). Here we apply targeted and conditional mutagenesis to examine and compare the functional roles of Setd1a and Setd1b in mouse development, ESCs and somatic cell reprogramming. Both proteins are essential. However, despite their virtually identical proteomic profiles, they convey radically different functions.

RESULTS

Setd1a and Setd1b are highly conserved paralogs

As evident from their gene structures, *Setd1a* and *Setd1b* clearly arose by gene duplication in evolution. At least 10 of 19 (*Setd1a*) or 17 (*Setd1b*) exons are bounded by identically placed splice sites, among other conserved features (supplementary material Fig. S1). Both mouse proteins have the same architecture based on an N-terminal RNA recognition motif (RRM) and C-terminal N-SET plus SET/postSET domains with 72% (49/68), 61% (86/140) and 86% (108/125) amino acid sequence identity, respectively, and high overall identity (~38%) and similarity (~68%).

Setd1a and Setd1b knockouts die at different embryonic stages

To explore the function of these highly conserved proteins, we established multipurpose alleles (Testa et al., 2004) for both *Setd1a* and *Setd1b* by gene targeting. In our multipurpose allele strategy, the 5'-most frameshifting exon is flanked by loxP sites (exon 4 for *Setd1a*, Fig. 1A and supplementary material Fig. S2; exon 5 for *Setd1b*, Fig. 1D and supplementary material Fig. S2) and a stop cassette is inserted in the intron upstream of this exon. The stop cassette is flanked by FRT sites and contains a *lacZ* reporter. The 5' splice site of the cassette captures the target gene transcript and the polyadenylation site terminates its transcription, thereby producing a null allele that we term the 'A' allele. Staining of heterozygous *Setd1a*^{A/+} and *Setd1b*^{D/+} embryos for β-galactosidase expression indicated that both genes are expressed ubiquitously in the developing embryo, including the extra-embryonic tissues. Ubiquitous expression was also indicated by qRT-PCR analyses (supplementary material Fig. S2; Fig. 2A).

Removal of the stop cassette by FLP recombination establishes the 'F' allele (for FLP recombined) and restores wild-type expression.

Cre recombination of the F allele generates the 'FC' allele (C for Cre recombined) in which the loxP-flanked exon has been deleted, thereby frameshifting the open reading frame (Fig. 1A). The A and FC alleles are designed to knockout the gene in different ways so that they interrogate the null phenotype twice. If both A/A and FC/FC genotypes provoke the same phenotype, it is likely to be a null. By application of this test to *Setd1a* and *Setd1b*, we conclude that both alleles for both genes are nulls (Tables 1, 2). For technical reasons, a rox-flanked cassette was included next to the 3' loxP site in the *Setd1b* targeting construct. This cassette was removed by Dre recombination to establish the 'D' allele (for Dre recombined), which is operationally the same as an A allele (Fig. 1D). Notably, the *Setd1a* and *Setd1b* null phenotypes are dramatically different. No *Setd1a*^{A/A} or *Setd1a*^{FC/FC} embryos were found after E7.5, whereas *Setd1b*^{D/D} or *Setd1b*^{FDC/FDC} embryos were found at Mendelian ratios until E10.5.

Although *Setd1a*^{A/A} embryos implanted they failed to gastrulate, and antibody staining confirmed the loss of Setd1a protein expression (Fig. 1B,C). The *Setd1a* mutant embryos retained mitotic activity, as detected by immunostaining with anti-H3S10P, and displayed apoptosis as determined by TUNEL assay (supplementary material Fig. S4). By contrast, *Setd1b*^{D/D} embryos continued to develop until the grossly retarded embryos died at ~E11.5; however, growth retardation was observable from E7.5 (Fig. 1E). The presence of all germ layers in *Setd1b* knockout embryos was validated by histological analysis (supplementary material Fig. S3). Loss of Setd1b protein was confirmed by western analysis of whole embryos (Fig. 1F).

Setd1a is required shortly after inner cell mass formation

Both Setd1a and Setd1b are expressed during preimplantation development (Fig. 2A). High expression is detected in the oocyte, which reflects the maternal mRNA stock that declines to a stable level from the 8-cell stage until E8.5 (Fig. 2A). Expression of Setd1a and Setd1b in the blastocyst is mainly nuclear and found in both the inner cell mass (ICM) and the trophectoderm (Fig. 2B). *Setd1a*^{A/A} and *Setd1b*^{D/D} blastocysts were recovered at the expected Mendelian ratio and were indistinguishable from wild-type and heterozygous controls (Table 1). Blastocyst outgrowths from *Setd1b*^{D/+} intercrosses were normal and *Setd1b*^{D/+} and *Setd1b*^{D/D} ESC lines were derived (Table 3). By contrast, although the *Setd1a*^{A/A} blastocysts produced outgrowths we were not able to generate any *Setd1a*^{A/A} ESC lines from a total of 73 picked ICM clumps (Fig. 2D, Table 3). Likewise, we were not able to target the second allele in *Setd1a*^{A/+} ESCs using the successful first allele targeting construct harboring a different (blasticidin) resistance gene (data not shown). These findings suggest an important function(s) of Setd1a in the ICM and ESCs. As Oct4 (Pou5f1 – Mouse Genome Informatics) is the best characterized marker for ICM pluripotency (Nichols et al., 1998), we examined outgrowths from blastocysts isolated from homozygous *Oct4-GFP;Setd1a*^{FC/+} males mated to *Setd1a*^{FC/+} females. Notably, outgrowths from *Setd1a*^{FC/FC} blastocysts and heterozygous siblings both showed GFP expression that was maintained for at least 3 days at the expected Mendelian ratio (Fig. 2C,D). This indicates that the requirement for Setd1a lies shortly after the formation of the ICM and before its transformation into ESCs. This conclusion was supported by inspection of *Oct4-GFP;Setd1a*^{FC/FC} embryos shortly after implantation. At the time when the epiblast still expressed Oct4 in heterozygous siblings, *Setd1a*^{FC/FC} embryos exhibited reduced expression (Fig. 3A), suggesting that the defect arises at or immediately after the establishment of the epiblast.

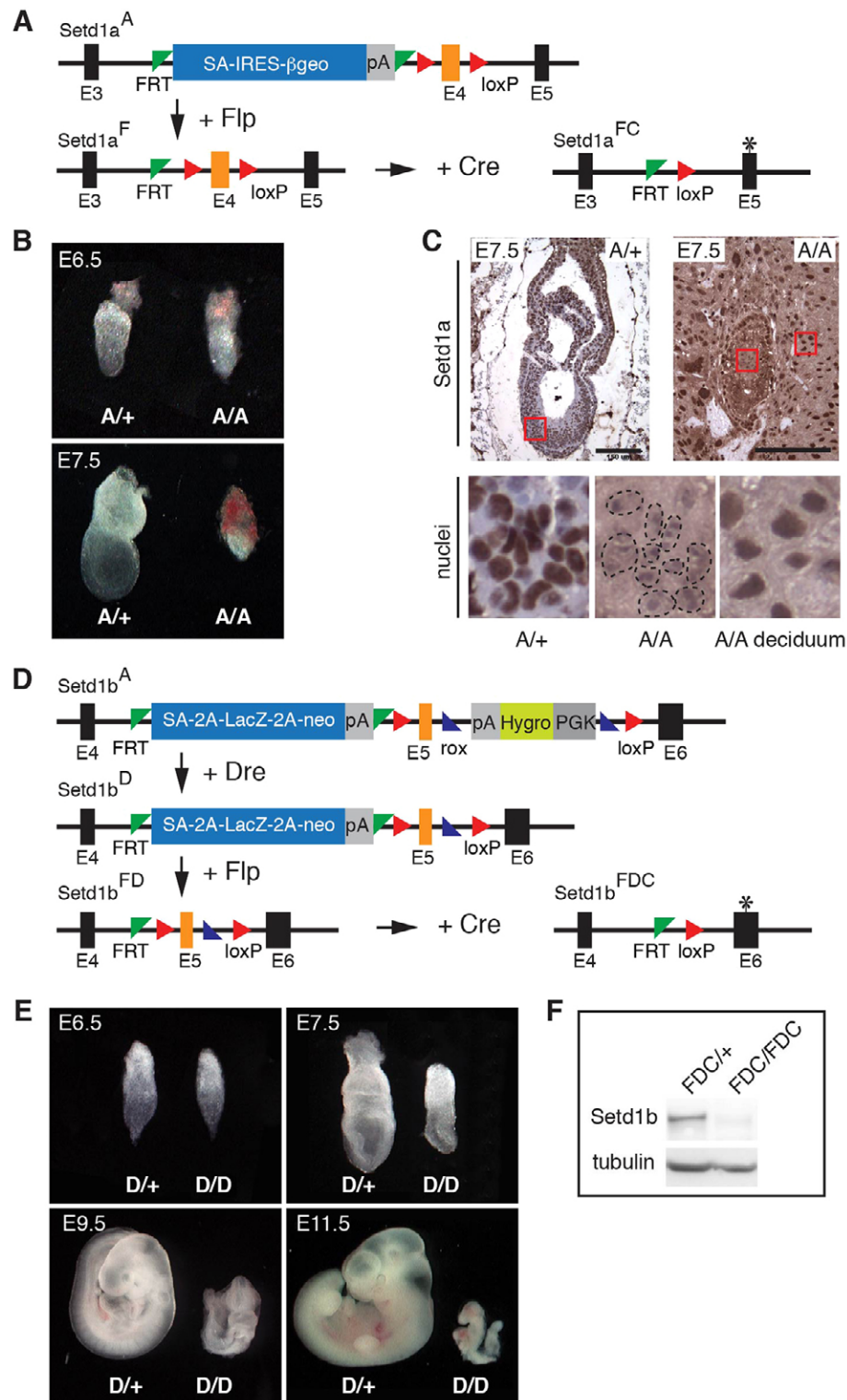


Fig. 1. Allele design and embryonic phenotype of *Setd1a* and *Setd1b* mutant mouse embryos. (A) The *Setd1a* knockout first allele (*Setd1a^A*). This allele is converted to conditional (*Setd1a^F*) upon Flp recombination. Cre recombination leads to excision of the frameshifting exon 4, generating the conditional mutant allele (*Setd1a^{FC}*). (B) Embryos from *Setd1a^{A/+}* intercrosses at E6.5 and E7.5. At E6.5, *Setd1a^{A/A}* embryos are growth retarded compared with littermates. (C) Immunostaining with a *Setd1a*-specific antibody on sagittal sections of *Setd1a^{A/+}* and *Setd1a^{A/A}* embryos at E7.5. Nuclei in the *Setd1a^{A/A}* embryo are devoid of specific staining, confirming loss of protein. The staining is visible in *Setd1a^{A/+}* embryonic cells and in the deciduum of the *Setd1a^{A/A}* embryo. Scale bars: 150 μ m. Red boxes indicate the regions enlarged beneath. (D) The *Setd1b* knockout first allele (*Setd1b^A*). This allele is similar to that resulting after Dre recombination (*Setd1b^D*) to remove the PGK-hygromycin selection cassette. The subsequent recombination steps are identical to the *Setd1a* targeting strategy (A). The frameshifted allele (*Setd1b^{FDC}*) is based on excision of exon 5. (E) Embryos from *Setd1b^{D/+}* intercrosses at E6.5, E7.5, E9.5 and E11.5. Growth retardation is apparent from E7.5 and embryos die before E11.5. (F) Immunoblot on whole cell extracts from *Setd1b* mutant embryos confirms the loss of *Setd1b* protein. For abbreviations see supplementary material Fig. S2.

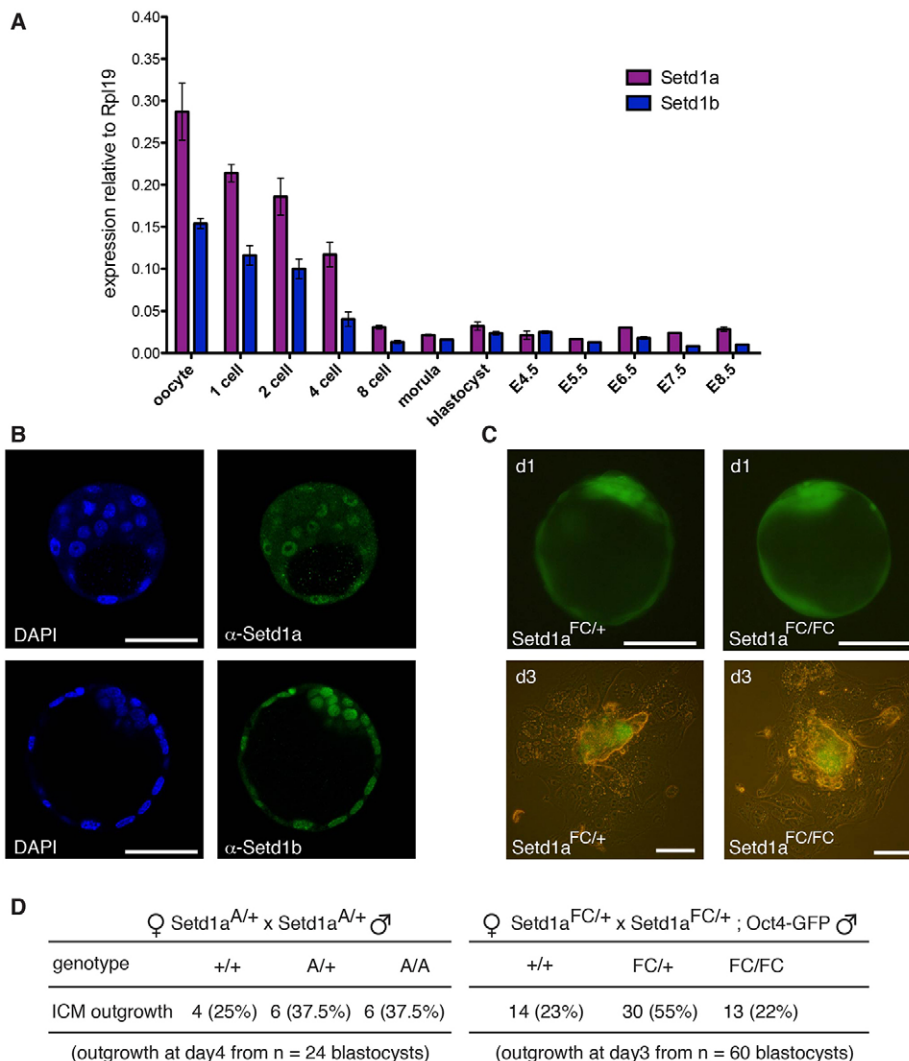


Fig. 2. *Setd1a* mutant blastocysts can make outgrowths but cannot generate ESC lines.

(A) Expression of *Setd1a* and *Setd1b* relative to *Rpl19* analyzed in pools of embryos by qRT-PCR. The number of embryos (*n*) in the pools is as follows: oocyte, 90; 1-cell, 60; 2-cell, 30; 4-cell, 15; 8-cell, 8; morula, 8; E3.5 blastocyst, 4; E4.5 blastocyst, 4; E5.5 to E8.5, 1. (B) Confocal sections of blastocysts immunostained for *Setd1a* or *Setd1b* and with DAPI to visualize nuclei. (C) The ICM in the *Setd1a*^{FC/FC} blastocyst outgrowths expresses Oct4-GFP. (D) Genotypes of outgrowths from *Setd1a*^{A/+} and *Setd1a*^{FC/+} intercrosses. Scale bars: 50 μ m in B and in top row of C; 100 μ m in C bottom row.

We speculated that the early lethality caused by loss of *Setd1a* was due to exhaustion of maternally supplied *Setd1a* protein/mRNA oocyte stockpiles, which might mask an even earlier requirement in development. Therefore, we conditionally mutated *Setd1a* during oocyte development by crossing the *Setd1a* conditional line with the *Gdf9-Cre* line, which excises from early folliculogenesis onwards (Lan et al., 2004). *Setd1a*^{F/F;Gdf9-Cre/+} females [hereafter termed *Setd1a*^{Gdf9} conditional knockout (cKO)] and heterozygous *Setd1a*^{F/+;Gdf9-Cre/+} controls were bred to wild-type males. *Setd1a* mRNA was almost undetectable in *Setd1a*^{Gdf9} cKO oocytes (Fig. 3B). Recombination was complete in all pups that carried the conditional allele and both the number of litters born and litter sizes from *Setd1a*^{Gdf9} cKO females were comparable to those of controls (Fig. 3C). Therefore, unlike the corresponding *Mll2*^{Gdf9} cKO (Andreu-Vieyra et al., 2010), loss of *Setd1a* from developing oocytes did not perturb fertility. Notably, mating of *Setd1a*^{Gdf9} cKO females to *Setd1a*^{FC/+} males recapitulated the null phenotype with no observable increase in severity (Fig. 3D). Furthermore, *Setd1a*^{Gdf9} cKO; *Oct4-GFP* embryos had detectable GFP expression in the epiblast (Fig. 3E). This indicates that the null phenotype revealed at E7.5 is not due to the exhaustion of maternal stores. From these experiments we conclude that there is no direct requirement for maternal *Setd1a* transcripts or protein in the embryo and, moreover, no direct requirement for *Setd1a* until the ICM has formed.

Setd1a, but not Setd1b, is essential in ESCs

To define the roles of *Setd1a* and *Setd1b* more precisely, we derived ESC lines from blastocysts carrying conditional *Setd1a* or *Setd1b* alleles combined with the heterozygous *Rosa-Cre-ERT2* (RC⁺) allele. To induce the *Setd1a* or *Setd1b* frameshift mutations, 4-hydroxytamoxifen (4OHT) was added to the medium for 48 hours, by which time both proteins are almost completely lost (Fig. 4A,B). Conditional deletion of *Setd1a* severely affected the growth rate of ESCs, which stopped proliferating 72 hours after 4OHT addition (Fig. 4C). The proliferation defect of *Setd1a*^{FC/FC;RC/+} ESCs could be rescued by a *Setd1a* BAC transgene, which restored protein expression (Fig. 4C,D). The rescue was observed with two different clones that carried either one (1c) or two (2c) copies of the BAC transgene. Cell cycle analysis of *Setd1a*^{F/F;RC/+} ESCs before and after 4OHT induction revealed a reduction in S phase, an apparent increase in G1 and high levels of disintegrating cells (Fig. 4E). By contrast, conditional deletion of *Setd1b* did not affect cellular proliferation, morphology (data not shown) or cell cycle distribution (Fig. 4F).

The reduced cell number upon *Setd1a* deletion in ESCs was accompanied by increased apoptosis, as shown by TUNEL staining (Fig. 4G) and increased annexin V-positive ESCs (Fig. 4H). Notably, the loss of a single copy of *Setd1a* also promoted higher levels of annexin V-positive cells (Fig. 4H). This indication of apparent *Setd1a* haploinsufficiency was supported by the observation that the

Table 1. Embryonic lethality of *Setd1a*^{A/A} and *Setd1a*^{FC/FC} embryos

| Stage | Genotype* | | | | |
|---------------------------------------------------------------|-----------|-----------|-----------|-------------|----------|
| <i>Setd1a</i> ^{A/+} × <i>Setd1a</i> ^{A/+} | +/+ | A/+ | A/A | Resorption* | Total‡ |
| Born pups | 15 (45.5) | 18 (54.5) | 0 | – | 33 (6) |
| E13.5 | 6 (37) | 3 (19) | 0 | 7 (44) | 16 (2) |
| E12.5 | 4 (25) | 10 (62.5) | 0 | 2 (12.5) | 16 (2) |
| E11.5 | 6 (37.5) | 6 (37.5) | 0 | 4 (25) | 16 (2) |
| E10.5 | 5 (26.5) | 9 (47) | 0 | 5 (26.5) | 19 (2) |
| E9.5 | 3 (33.3) | 3 (33.3) | 0 | 3 (33.3) | 9 (1) |
| E8.5 | 12 (20) | 27 (45) | 0 | 21 (35) | 60 (8) |
| E7.5 | 11 (23) | 23 (48) | 14 (29) | 0 | 48 (7) |
| E6.5 | 16 (24) | 39 (57) | 13 (19) | 0 | 68 (9) |
| E3.5 | 16 (35) | 17 (40) | 11 (25) | 0 | 44 (6) |
| <i>Setd1a</i> ^{F/+} × <i>Setd1a</i> ^{F/+} | +/+ | F/+ | F/F | | |
| Born pups | 31 (27) | 62 (54.5) | 21 (18.5) | | 114 (16) |
| <i>Setd1a</i> ^{FC/+} × <i>Setd1a</i> ^{FC/+} | +/+ | FC/+ | FC/FC | | |
| E9.5 | 6 (24) | 11 (44) | 0 | 8 (32) | 25 (3) |
| E8.5 | 2 (12) | 10 (59) | 0 | 5 (29) | 17 (2) |
| E7.5 | 13 (23) | 28 (49) | 16 (28) | 0 | 57 (7) |
| E6.5 | 3 (20) | 9 (60) | 3 (20) | 0 | 15 (2) |

*The percentage of total is shown in parentheses.

‡The number of litters (*n*) is shown in parentheses.

rescued *Setd1a*^{FC/FC;RC/+} ESC clone containing one copy of the *Setd1a* BAC transgene (BAC 1c) showed comparable numbers of annexin V-positive cells to the heterozygously deleted *Setd1a*^{FC/+;RC/+} clone, whereas in the rescued mutant with two copies of the BAC transgene (BAC 2c) annexin V-positive cells returned to wild-type levels. However, we have not yet been able to identify a heterozygous *Setd1a* phenotype in mice.

Due to reduced proliferation, deletion of *Setd1a* led to smaller colonies, which still expressed the classic ESC marker alkaline

phosphatase (AP) (Fig. 4I). However, qRT-PCR evaluation of the ESC pluripotency factors *Oct4*, *Nanog* and *Klf4* revealed strong downregulation that was comparable to, or greater than, the downregulation in the same time frame that accompanies exit from self-renewal promoted by LIF withdrawal (Fig. 4J). Therefore, we looked for upregulation of early lineage markers. Although a 2- to 3-fold increase in *Fgf5*, brachyury (*T*) and *Gata4* mRNA expression was observed, these responses were not equivalent to those promoted by LIF withdrawal (Fig. 4K). These results show that

Table 2. Embryonic lethality of *Setd1b*^{D/D} and *Setd1b*^{FDC/FDC} embryos

| Stage | Genotype* | | | | |
|-----------------------------------------------------------------|-----------|-----------|-----------|-------------|----------|
| <i>Setd1b</i> ^{D/+} × <i>Setd1b</i> ^{D/+} | +/+ | D/+ | D/D | Resorption* | Total‡ |
| Born pups | 29 (30) | 67 (70) | 0 | – | 96 (16) |
| E13.5 | 8 (21) | 14 (37) | 0 | 16 (42) | 38 (4) |
| E12.5 | 10 (23) | 12 (28) | 0 | 21 (49) | 43 (5) |
| E11.5 | 12 (28) | 20 (46.5) | 4 (9.5) | 7 (16) | 43 (5) |
| E10.5 | 22 (25.5) | 31 (36) | 17 (20) | 16 (18.5) | 85 (8) |
| E9.5 | 6 (15) | 21 (54) | 9 (23) | 3 (8) | 39 (5) |
| E8.5 | 14 (33) | 18 (42) | 7 (16) | 4 (9) | 43 (5) |
| E7.5 | 10 (40) | 8 (32) | 7 (28) | 0 | 25 (3) |
| E6.5 | 1 (3) | 18 (62) | 10 (35) | 0 | 29 (3) |
| <i>Setd1b</i> ^{FD/+} × <i>Setd1b</i> ^{FD/+} | +/+ | FD/+ | FD/FD | | |
| Born pups | 31 (18) | 84 (48.5) | 58 (33.5) | | 173 (20) |
| <i>Setd1b</i> ^{FDC/+} × <i>Setd1b</i> ^{FDC/+} | +/+ | FDC/+ | FDC/FDC | | |
| Born pups | 20 (30) | 47 (70) | 0 | – | 67 (13) |
| E13.5 | 4 (27) | 9 (60) | 0 | 2 (13) | 15 (2) |
| E12.5 | 3 (17) | 9 (50) | 0 | 6 (33) | 18 (3) |
| E11.5 | 7 (19) | 19 (51) | 3 (8) | 8 (22) | 37 (4) |
| E10.5 | 8 (15) | 30 (56.5) | 13 (24.5) | 2 (4) | 53 (6) |
| E9.5 | 9 (20) | 20 (43) | 12 (26) | 5 (11) | 46 (5) |
| E8.5 | 3 (12) | 15 (60) | 6 (24) | 1 (4) | 25 (3) |
| E7.5 | 3 (16) | 13 (68) | 3 (16) | 0 | 19 (3) |
| E6.5 | 2 (17) | 6 (50) | 4 (33) | 0 | 12 (2) |

*The percentage of total is shown in parentheses.

‡The number of litters (*n*) is shown in parentheses.

Table 3. Genotypes of ESC lines derived from *Setd1a*^{A/+} and *Setd1b*^{D/+} intercrosses

| Cross | <i>Setd1a</i> ^{A/+} × <i>Setd1a</i> ^{A/+} | | | <i>Setd1b</i> ^{D/+} × <i>Setd1b</i> ^{D/+} | | |
|-----------|-------------------------------------------------------------|----------|--------|-------------------------------------------------------------|----------|---------|
| Genotype | +/+ | A/+ | A/A | +/+ | D/+ | D/D |
| ESC lines | 7 (41%) | 10 (59%) | 0 (0%) | 5 (28%) | 11 (61%) | 2 (11%) |

*n=73 blastocysts; †n=45 blastocysts.

Setd1a is essential for proliferation and maintenance of the transcriptional profile of key pluripotent genes in ESCs in culture, whereas *Setd1b* is dispensable.

Setd1b overexpression cannot rescue the ESC proliferation defects caused by loss of *Setd1a*

Although qRT-PCR results indicated that *Setd1a* and *Setd1b* mRNA levels are similar (Fig. 2A), a recent proteomic analysis of ESCs suggests that *Setd1a* is much more highly expressed than *Setd1b* (van Nuland et al., 2013). Possibly, therefore, the proliferation defects caused by loss of *Setd1a* can be compensated by overexpressing *Setd1b*. To test this proposition we transfected *Setd1a* conditional ESCs with a *Setd1b* BAC. Four stable colonies overexpressing *Setd1b* (2- to 4-fold) were identified and induced for the loss of *Setd1a* by 4OHT administration. No rescue of the proliferation defect was observed (data not shown). Because 4-fold *Setd1b* overexpression might not be enough to rescue the proliferation defect, we made a chimeric *Setd1* BAC. The exonic coding region of the rescuing *Setd1a* BAC (Fig. 4C,D) was replaced

with the exonic coding *Setd1b* region (Fig. 5A). After transfection of the chimeric BAC into *Setd1a*^{F/F} ESCs, strong overexpression of Venus-tagged *Setd1b* from the *Setd1a* promoter and gene context was achieved (Fig. 5B,C). Despite high levels of *Setd1b* expression, no rescue of the ESC proliferation defect caused by loss of *Setd1a* was observed (Fig. 5D,E). This finding implies that *Setd1a* and *Setd1b* regulate different targets in ESCs.

Setd1a is also required for EpiSC and NSC proliferation

To test whether *Setd1a* is required in other cell types apart from ESCs, we established epiblast stem cells (EpiSCs) from *Setd1a*^{F/F;RC/+} ESCs and from those carrying the *Setd1a*-Venus BAC (supplementary material Fig. S5A). As expected, in EpiSCs the expression of *Rex1* (*Zfp42* – Mouse Genome Informatics) was downregulated and the expression of *Oct4* and *Fgf5* upregulated (supplementary material Fig. S5B). Loss of *Setd1a* upon 4OHT administration in EpiSCs resulted in reduced proliferation and downregulation of *Oct4* (supplementary material Fig. S5B,C). A similar proliferation defect was observed with neural stem cells

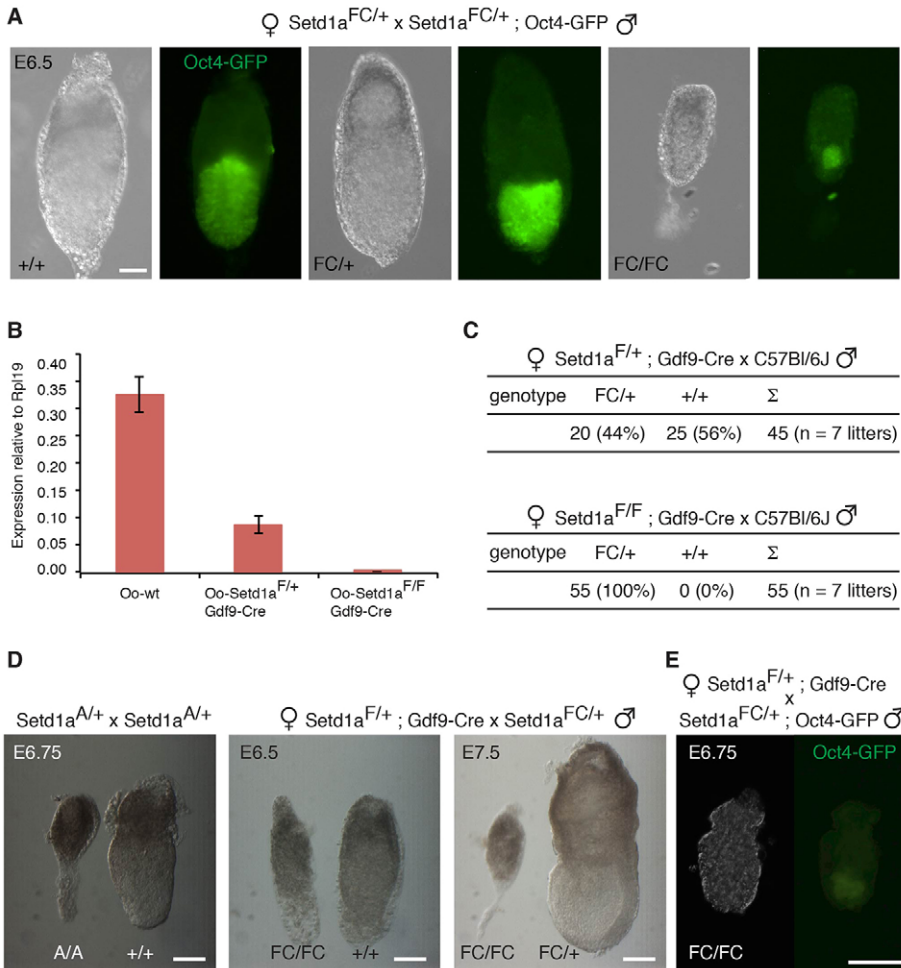


Fig. 3. Depletion of maternal *Setd1a* results in the same phenotype as the conventional knockout. (A) At E6.5, Oct4-GFP staining is still detectable in the epiblast of *Setd1a*^{FC/FC} embryos. (B) *Setd1a*^{F/F}; *Gdf9-Cre* oocytes do not express *Setd1a* mRNA (n=28 oocytes for all three genotypes). Error bars indicate s.d. (C) *Setd1a*^{F/F}; *Gdf9-Cre* mice are fertile and generate progeny of the expected genotypes. (D) *Setd1a*^{FC/FC} embryos, in which *Gdf9-Cre* deletes maternal *Setd1a*, show the same growth retardation phenotype (middle and right) as *Setd1a*^{A/A} embryos (left). (E) At E6.75, Oct4-GFP staining is still detectable in the epiblast of *Setd1a*^{FC/FC}; *Gdf9-Cre* embryos. Scale bars: 100 μm.

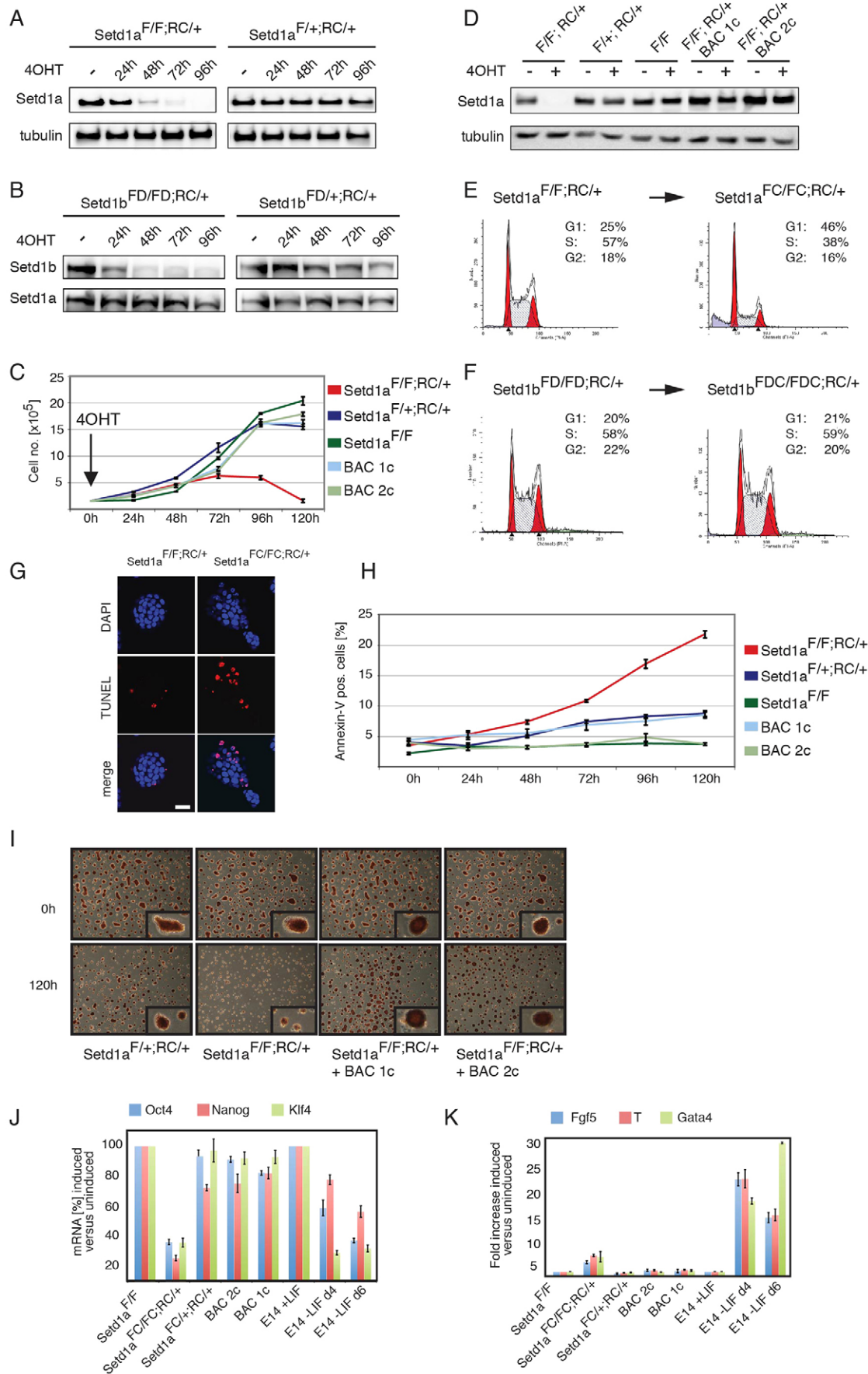


Fig. 4. See next page for legend.

Fig. 4. *Setd1a* deletion in ESCs results in a proliferation defect without obvious signs of differentiation. (A,B) Immunoblot using *Setd1a* (A) or *Setd1b* (B) antibodies on whole cell extracts from ESCs carrying *Rosa26-Cre-ERT2* (RC/+) and conditional homozygous (F/F) or heterozygous (F/+) alleles at different time points after 4OHT addition. Neither protein is detectable 96 hours after 4OHT treatment. Tubulin serves as loading control. *Setd1a* protein levels are unchanged upon *Setd1b* deletion. (C) Proliferation timecourse after 4OHT administration. The proliferation rate of *Setd1a^{FC/FC}* ESCs is dramatically decreased after 72 hours. This defect can be rescued by transfecting a *Setd1a*-containing BAC (BAC 1c has one copy integrated; BAC 2c has two integrated copies). Error bars indicate s.d. of *n*=3 replicates. (D) Immunoblot showing that integration of *Setd1a* BAC transgenes restores *Setd1a* levels in *Setd1a^{F/F;RC/+}* ESCs treated with 4OHT. (E) Cell cycle profile of *Setd1a^{F/F;RC/+}* ESCs before and after 4OHT treatment. Note the increase of G1 phase and decrease of S phase. (F) Cell cycle profile is unchanged in induced *Setd1b* mutant ESCs. (G) *Setd1a* conditional mutant ESCs show increased apoptosis, as detected by TUNEL staining. Nuclei are stained with DAPI. Scale bar: 25 μ m. (H) Quantification (%) of apoptotic annexin V-positive ESCs by flow cytometry at different time points after 4OHT induction. Error bars indicate s.e.m. of *n*=3 replicates. (I) Alkaline phosphatase (AP) staining of *Setd1a* conditional mutant and BAC-rescued ESCs at 0 and 120 hours after 4OHT induction. *Setd1a^{FC/FC}* ESC colonies are smaller but still AP⁺ and they do not show any signs of differentiation. Insets illustrate colonies enlarged. (J) qRT-PCR analysis of the pluripotency genes *Oct4*, *Nanog* and *Klf4*. (K) qRT-PCR analysis of the early differentiation genes *Ggf5*, *brachyury* (*T*) and *Gata4*. Error bars indicate s.e.m. of *n*=3 replicates.

(NSCs) isolated from E15.5 *Setd1a^{F/F;RC/+}* telencephali (supplementary material Fig. S6A,B).

Setd1a is essential for the derivation of induced pluripotent stem cells

We tested the role of *Setd1a* and *Setd1b* in the reprogramming of a somatic state to induced pluripotent stem cells (iPSCs). NSCs were isolated from E15.5 *Rosa-Cre-ERT2/+* telencephali that carried either *Setd1a* or *Setd1b* conditional alleles. The four reprogramming factors (*Oct4*, *Sox2*, *Klf4* and *c-Myc*) were introduced by PiggyBac transposon (Yusa et al., 2009). To define the time window in which *Setd1a* or *Setd1b* function might be essential during the reprogramming process, both heterozygous and homozygous NSCs were treated with 4OHT at different time points before and after adding the four reprogramming factors: (1) 5 days prior to the start of reprogramming (P-1), (2) concomitant with PiggyBac transposition (d0-3), (3) mid-term (d5-8) and (4) late (d10-13) during reprogramming (Fig. 6A). After 14 days, colonies were stained for AP activity and classified into two categories according to size and AP staining intensity. Early removal of *Setd1a* had a strong effect on reprogramming (Fig. 6B,C), which is probably due to its requirement for NSC proliferation (supplementary material Fig. S6A,B). Later removal of *Setd1a* or removal of *Setd1b* did not prevent the efficient formation of iPSC colonies (Fig. 6B-D; supplementary material Fig. S7A,B). Reprogramming quality was confirmed by reactivation of the endogenous pluripotency markers *Nanog* and *SSEA1* (Fig. 6E,F; supplementary material Fig. S7C,D). However, we were not able to establish stable iPSC lines from any of the *Setd1a*-deficient iPSC colonies (Fig. 6E). Together with results from ESCs, EpiSCs and NSCs, this finding underlines the requirement of *Setd1a*, but not *Setd1b*, for stem cell proliferation.

Setd1a deletion leads to a global reduction of H3K4 methylation levels

To evaluate H3K4 methylation after loss of the *Setd1* proteins, we analyzed mono-, di- and trimethylation in embryos and ESCs. Immunostaining of E7.5 *Setd1a^{Δ/Δ}* embryos revealed a 50%

reduction in all three H3K4 methylated species (Fig. 7A,B). This observation was supported by western blots using the *Setd1a* conditional ESCs after addition of 4OHT (Fig. 7C). Whereas total H3, H2B and H3K27me3 levels remained constant, all three methylation marks at lysine 4 showed a significant reduction, which was restored in BAC-rescued ESC lines (BAC 1c and BAC 2c). However, H3K4me3 levels were not restored in *Setd1a* conditional ESCs overexpressing *Setd1b* from the chimeric *Setd1* BAC (data not shown). In contrast to *Setd1a*, deletion of *Setd1b* in ESCs did not affect global H3K4 mono-, di- and trimethylation levels (Fig. 7D). Likewise, we were unable to detect any changes in global H3K4 methylation in E10.5 *Setd1b^{D/D}* embryos (Fig. 7E).

DISCUSSION

Despite highly homologous protein sequences and identical protein complexes, *Setd1a* and *Setd1b* play very different essential roles in mouse development. *Setd1a* is required immediately before gastrulation, shortly after implantation. At this stage and in ESCs, it is the major H3K4 methyltransferase. It is required for the maintenance of pluripotency factors in ESCs as well as for ESC, iPSC, EpiSC and NSC proliferation. By contrast, none of these functions applies to *Setd1b*, which is required after gastrulation during organogenesis.

Because their protein complexes, *Setd1aC* and *Setd1bC*, are otherwise identical (Lee and Skalnik, 2005; Lee et al., 2007), the different functions of *Setd1a* and *Setd1b* must be conveyed by differences in their (1) expression patterns, (2) transient protein-protein interactions or (3) epigenetic interactions. First, all evidence so far indicates that both mRNAs are ubiquitously expressed (Fig. 2; supplementary material Fig. S2). Whether the two proteins are expressed in different patterns, or at very different levels as has been recently suggested (van Nuland et al., 2013), remains to be determined. At least in ESCs, overexpression of *Setd1b* cannot rescue the proliferation defect caused by loss of *Setd1a*. Second, three different protein-protein interactions have been described so far. *Setd1a*, but not *Setd1b*, interacts with host cell factor 1 (*Hcf1*, or *Hcfcl*), whereas *Setd1b*, but not *Setd1a*, interacts with *Rbm15* (Lee and Skalnik, 2008; Lee and Skalnik, 2012). Both *Hcf1* and *Rbm15* are expressed in ESCs (data not shown). It has also been reported that *Setd1a* interacts with β -catenin and binds to the *Tert* promoter. Although β -catenin (*Ctnnb1*)^{-/-} ESCs are viable, the *Tert* promoter loses H3K4me3 and telomerase activity is downregulated (Hoffmeyer et al., 2012). However, we could not detect *Tert* downregulation in our *Setd1a* conditional mutant ESCs (data not shown). Current information does not permit an explanation of the differences between *Setd1a* and *Setd1b* knockout phenotypes based on these different protein-protein interactions. Third, *Setd1a* and *Setd1b* might differentially interact with the epigenetic circuitry. A role for the N-SET domain in the regulation of epigenetic writing was recently assigned to yeast *Set1* (Kim et al., 2013). Also, RNA binding by the RRM domain, which conceptually presents great potential for regulatory specificity, has been described for yeast *Set1* (Trésaugues et al., 2006). Differences between the functions of the N-SET or RRM domains in *Setd1a* and *Setd1b*, in addition to specific protein-protein interactions mediated by the divergent sections of these very similar proteins, are likely to be the source of their different roles in mouse development.

Setd1a in mouse development

We found that *Setd1a* is required for ESC, iPSC, EpiSC and NSC proliferation. Hence, we considered whether it is required for proliferation of all cell types and, if so, then the very early embryonic lethality of the *Setd1a* knockout could report the moment

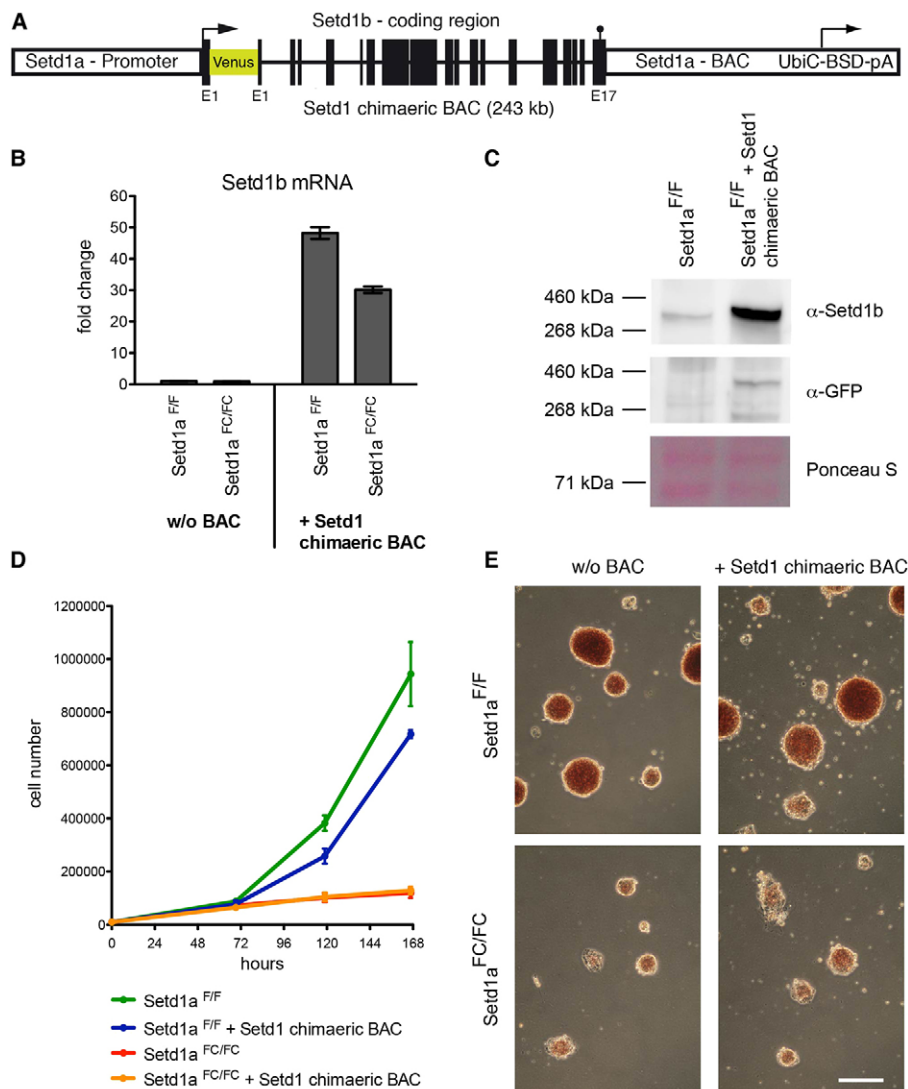


Fig. 5. Setd1b overexpression cannot rescue the ESC proliferation defects caused by loss of Setd1a. (A) The *Setd1a*/Setd1b (*Setd1*) chimeric BAC. In a *Setd1a*-containing BAC, the *Setd1a* gene body was exchanged with an N-terminal-tagged Venus-Setd1b gene body. A ubiquitin C-blasticidin-polyA (UbiC-BSD-pA) selection cassette was inserted next to the bacterial backbone and used for stable selection. (B) Expression of *Setd1b* mRNA in *Setd1a*^{F/F;RC/+} ESCs and in those carrying the chimeric BAC, before and after 4OHT treatment. Error bars indicate s.d. (C) Immunoblots of Setd1b expression in ESCs using Setd1b antibody or anti-GFP. Ponceau staining serves as loading control. (D) The proliferation defect caused by loss of Setd1a in ESCs is not rescued by the chimeric BAC in a timecourse of 168 hours after 4OHT treatment. Error bars indicate s.d. (E) AP staining of *Setd1a* conditional mutant and chimeric BAC ESCs 120 hours after 4OHT induction. Scale bar: 100 μm.

when the oocyte stores of maternal Setd1a are exhausted. However, the removal of Setd1a from oocytes by *Gdf9-Cre* disproved this proposition and showed that Setd1a is not required for early proliferation of the embryo. Deletion of *Setd1a* expression during oogenesis had no impact on fertility or the early embryonic lethal phenotype. By contrast, deletion of *Mll2* using the same *Gdf9-Cre* driver caused complete infertility through loss of ovulation and oocyte degeneration (Andreu-Vieyra et al., 2010). Because the kinetics of recombination-promoted protein loss for these two genes (and also *Setd1b*) in ESCs are the same (Glaser et al., 2009) (Fig. 4), we conclude that Setd1a is not required for oogenesis or cell division in the embryo before ICM formation.

Furthermore, use of an *Oct4-GFP* reporter and attempts to establish Setd1a-deficient ESCs precisely defined the point in development at which Setd1a becomes essential. Our analyses showed that the ICM is established in the absence of Setd1a but not sustained after implantation. *Setd1a* knockout embryos develop primitive endoderm but do not gastrulate. Hatching from cultured blastocysts occurred but further proliferation was impaired and no ESC lines could be established. Hence, Setd1a is required at, or immediately after, formation of the epiblast. Setd1a is also essential in ESCs for proliferation and maintenance of the self-renewal circuitry. However, loss of this circuitry in *Setd1a* knockout ESCs

did not provoke default differentiation, suggesting that Setd1a might be required for ESC transcriptional activity in general, which is a proposition supported by the fact that it is the major ESC H3K4 methyltransferase. These observations are reinforced by our reprogramming experiments from NSCs. Setd1a is required for NSC proliferation and its removal at any stage during a reprogramming timecourse impeded the acquisition of iPSC lines.

At some point in early mouse development, there is a shift of the major regulator of H3K4me2/3 from Mll2 to Setd1a. Mll2 plays this role during oogenesis and the early cleavage stages after fertilization. Thereafter, knockout of *Mll2* had no effect on bulk H3K4 methylation (Lubitz et al., 2007; Glaser et al., 2009) indicating that (an)other H3K4 methyltransferase(s) acquired this role at some point after compaction. Our findings implicate Setd1a, and not Setd1b, as the major H3K4 di- and trimethyltransferase in question and suggest that Setd1a acquires this role at, or immediately before, the epiblast stage when its function becomes required. Whether Setd1a remains the major H3K4 methyltransferase in adults remains to be determined.

H3K4 methyltransferases in mouse development

The embryonic phenotype of the *Setd1b* knockout bears a striking resemblance to that of the *Mll2* knockout. In both cases, the

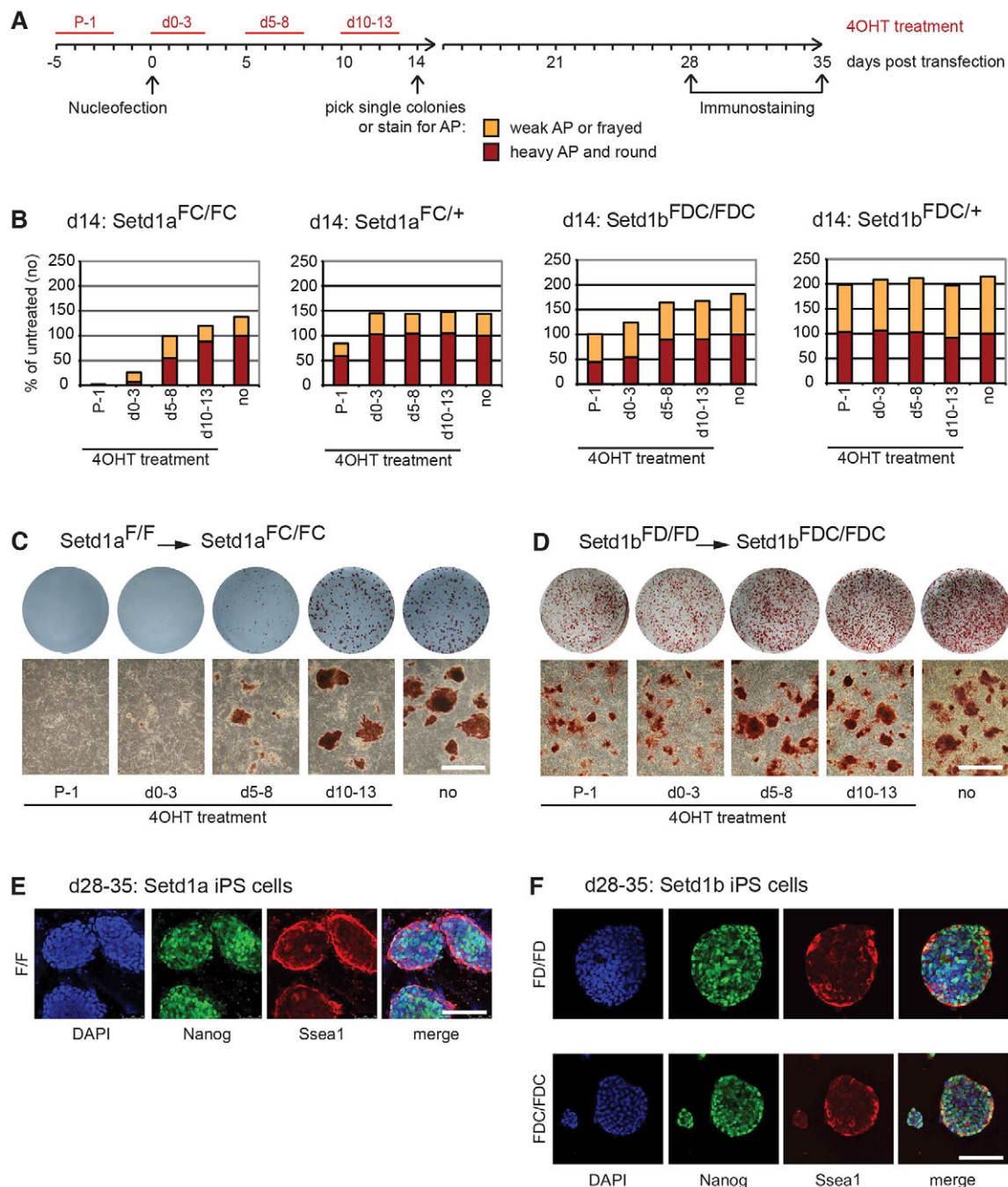


Fig. 6. *Setd1a* knockout NSCs cannot generate iPSCs. (A) The NSC to iPSC reprogramming timescale. NSCs were treated with 4OHT either before (P-1), simultaneously (d0-3), 5 days after (d5-8) or 10 days after transfection of the reprogramming factors. iPSC colonies were stained for AP on day 14 after transfection and classified into weak or strong staining. From each experiment, iPSC colonies were picked, expanded and immunostained for Nanog and SSEA1 between days 28 and 35 after transfection. (B) Reprogramming efficiency shown as a percentage of iPSC colonies derived at different time points of 4OHT induction relative to uninduced cells from the four different conditional lines. (C,D) Micrographs of AP-stained iPSCs from *Setd1a*^{F/F;RC/+} and *Setd1b*^{FD/FD;RC/+} lines. (E,F) iPSCs from *Setd1a*^{F/F;RC/+} and *Setd1b*^{FD/FD;RC/+} generated without or after 4OHT treatment and immunostained for Nanog and SSEA1. Note that *Setd1a*^{FC/FC} iPSC colonies could not be sustained and hence are not shown. Scale bars: 1 mm in C,D; 100 μ m in E,F.

embryos appear normal until E7.5 but further development is retarded. However, all organs and structures appear to arise albeit in an increasingly retarded context and death occurs before E11.5. Neither protein is required for ESC self-renewal. Despite extensive examination we do not know how loss of Mll2 promotes this pleiotropic dysfunction. The fact that *Setd1b* appears to promote a similar phenotype might help solve the puzzle. Of the three other Set1/Trithorax-type H3K4 methyltransferases,

namely Mll1, Mll3 and Mll4, only the knockout phenotype of *Mll1* is known with certainty. It is first required in development, and continues to be required in adults, for definitive hematopoiesis (Yu et al., 1995; Yagi et al., 1998; Ernst et al., 2004; Jude et al., 2007; McMahon et al., 2007; Milne et al., 2010). Hence, Mll1 appears to have a very specialized function, as opposed to the apparently broader requirements for *Setd1a*, *Setd1b* and Mll2.

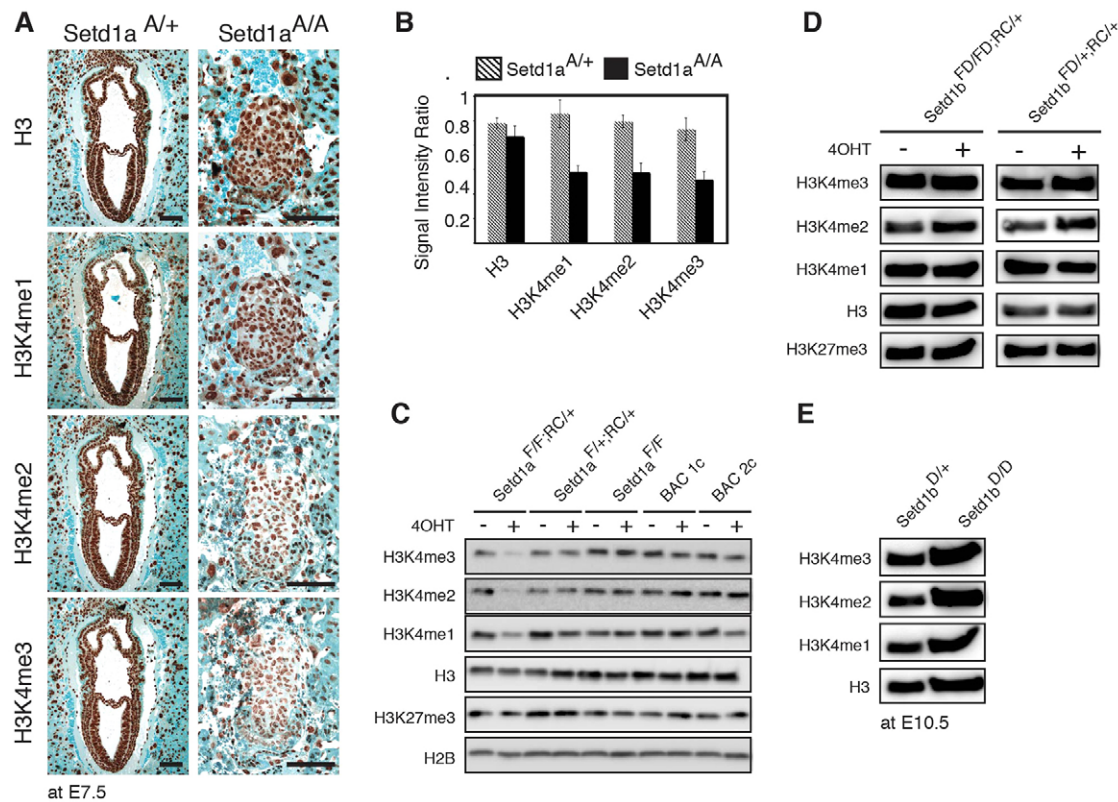


Fig. 7. Global reduction of H3K4 methylation levels in *Setd1a* but not *Setd1b* knockout ESCs and embryos. (A) Immunohistochemistry on sections of *Setd1a*^{A/+} and *Setd1a*^{A/A} embryos with H3, H3K4me1, H3K4me2 and H3K4me3 antibodies. (B) Quantification of the methylation levels by measuring the signal intensity of the nuclei in the embryo relative to the intensity in the neighboring deciduum. Error bars indicate s.d. of 20 nuclei per region from *n*=6 different embryo sections of each genotype. (C) Immunoblot on whole cell extracts from *Setd1a* conditional mutant and BAC-rescued ESCs using antibodies against H3K4me3, H3K4me2, H3K4me1, H3, H3K27me3 and H2B. Note the decrease in bulk H3K4 methylation in *Setd1a*^{FDC/FDC} ESCs. (D) Immunoblot on whole cell extracts from *Setd1b*^{FDC/FDC} and *Setd1b*^{FDC/RC+} ESCs using antibodies against H3K4me3, H3K4me2, H3K4me1, H3 and H3K27me3. (E) Immunoblot on whole cell extracts from *Setd1b*^{D/+} and *Setd1b*^{D/D} embryos at E10.5. Scale bars: 100 μ m.

Several subunits of Set1C have been knocked down or out. In the Ash2 scaffold, knockdown of *Wdr5* in ESCs resulted in reduced self-renewal, increased differentiation and global loss of H3K4me3 (Ang et al., 2011). However, knockdown of *Dpy30* and *Rbbp5* had no effect on self-renewal, but decreased differentiation into the neural lineage and a global reduction of H3K4me3 levels (but not H3K4me1 and H3K4me2) were observed (Jiang et al., 2011). These *Dpy30* and *Rbbp5* knockdown phenotypes are similar to the *Mll2* knockout in ESCs, which showed delayed exit from self-renewal and impaired neuronal differentiation (but no change in H3K4 methylation) (Lubitz et al., 2007). Possibly, the discrepancy with the *Wdr5* knockdown phenotype is due to the presence of *Wdr5* in several other protein complexes [e.g. ATAC (Nagy and Tora, 2007; van Nuland et al., 2013)]. Embryos homozygous for an *Ash2l* hypomorphic allele showed a more severe phenotype than that of the *Setd1a* knockout, consistent with the inclusion of Ash2l in other H3K4 methyltransferase complexes. No *Ash2l* hypomorphic embryos implanted, and attempts to derive ESCs from blastocysts failed (Stoller et al., 2010). By contrast, knockdown of *Ash2l* induced ESC differentiation (Wan et al., 2013).

Knockout of the Setd1C-specific subunit *Cxxc1* caused an early embryonic lethality apparently similar to that of the *Setd1a* knockout (Carlone and Skalnik, 2001). However, *Cxxc1* knockout ESCs are viable but cannot undergo differentiation and there is no global change in H3K4me3 levels (Carlone et al., 2005; Lee and Skalnik, 2005; Tate et al., 2009). But there was a repositioning of H3K4me3

away from active promoters to new ectopic sites indicating that *Cxxc1* restricts Setd1Cs to active promoters (Tate et al., 2009; Tate et al., 2010; Clouaire et al., 2012). Because *Setd1a* is responsible for bulk H3K4 methylation in ESCs, we predict that it is the enzyme that is being mispositioned. If so, it will be interesting to examine whether these ectopic sites are subject to other aspects specific to *Setd1a*, such as Hcf1 activity (Lee and Skalnik, 2008).

MATERIALS AND METHODS

Targeting constructs and BAC transgenes

The targeting constructs for *Setd1a* and *Setd1b* were generated using recombineering (Fu et al., 2010). For *Setd1a*, the engrailed-intron-splice-acceptor-IRES-lacZ-neomycin-polyA cassette flanked by FRT sites for FLP recombination was inserted in intron 3 of the gene. The critical exon 4, which upon deletion results in a frameshift and a premature stop codon, was flanked by loxP sites. For targeting the second allele of *Setd1a* we generated a similar targeting construct by replacing the neomycin resistance gene with blasticidin resistance. For *Setd1b*, the engrailed-intron-splice-acceptor-2A-lacZ-2A-neomycin-polyA cassette flanked by FRT sites was inserted in intron 4 and the critical exon 5 was flanked by loxP sites. The 3' loxP was accompanied by a PGK-hygromycin-polyA cassette flanked by rox sites for Dre recombination. The homology arms were 5 kb 5' and 4.9 kb 3' for *Setd1a* and 2.3 kb 5' and 6.4 kb 3' for *Setd1b*. Bacterial artificial chromosomes (BACs) containing *Setd1a* (clone name: RP23-188K7) and *Setd1b* (clone name: bMQ-33K9) genes were modified by recombineering to introduce an N-terminal Venus tag cassette directly after the second codon (Ciotta et al., 2011). The *Setd1a/Setd1b* chimeric BAC was made by

swapping the *Setd1a* gene body with that of *Setd1b* in the *Setd1a* BAC. All three BACs contained a ubiquitin C-blasticidin-polyA (UbiC-BSD-pA) selection cassette inserted close to the bacterial backbone.

Gene targeting and generation of conditional knockout mice

R1 ESCs were cultured with fetal calf serum (FCS)-based medium [DMEM + GlutaMAXTM (Invitrogen), 15% FCS (PAA), 2 mM L-glutamine (Invitrogen), 1x non-essential amino acids (Invitrogen), 1 mM sodium pyruvate (Invitrogen), 0.1 mM β -mercaptoethanol, in the presence of 1000 units LIF (Chemicon) per ml] on mitomycin-C inactivated mouse embryonic fibroblasts. Cells (1×10^7) were electroporated with 40 μ g linearized targeting construct using standard conditions (250 V, 500 μ F) and selected with 0.2 mg/ml G418. Colonies were screened for correct targeted events by Southern blot hybridization using an internal probe and 5' and 3' external probes. For removing the selection cassettes, correctly targeted *Setd1a* clones were electroporated with a CAG-Flpo-IRES-puro expression vector (Kranz et al., 2010). For both *Setd1a* and *Setd1b*, two correctly targeted ESC clones were injected into blastocysts and the resulting chimeras were subsequently bred to C57BL/6 mice. *Setd1b*^{+/+} mice were crossed to the *CAGGs-Dre* line (Anastassiadis et al., 2009) to remove the additional selection cassette (PGK-Hygro) downstream of the 3' loxP site generating *Setd1b*^{D/+} mice. Further, *Setd1a*^{+/+} and *Setd1b*^{D/+} mice were crossed to the *hACTB-Flpe* (Rodríguez et al., 2000) and to the *CAGGs-Flpo* (Kranz et al., 2010) line to generate *Setd1a*^{F/+} and *Setd1b*^{F/D/+} mice, respectively. Subsequently, those mice were crossed to either the *PGK-Cre* line (Lallemand et al., 1998) to produce *Setd1a*^{FC/+} and *Setd1b*^{FDC/+} mice or to the *Rosa26-Cre-ERT2* (RC) line (Seibler et al., 2003) to generate conditional, tamoxifen-inducible *Setd1a*^{F/+;RC/+} and *Setd1b*^{F/D/+;RC/+} mice. The *Oct4-GFP* mice (Szabó et al., 2002) were kindly provided by Hans Schöler and the *Gdf9-Cre* mice (Lan et al., 2004) by Rolf Jessberger. Primers for genotyping are provided in supplementary material Table S1. All animal experiments were performed according to German law.

Whole-mount X-gal staining and immunohistochemistry

Embryos were dissected and fixed in 4% paraformaldehyde (PFA), washed in PBS, and X-gal staining was performed as described (Glaser et al., 2006). For whole-mount immunohistochemistry, blastocysts or E6.5 embryos were fixed in 3% PFA for 30 minutes, washed with PBS/BSA, permeabilized with 0.2% Triton X-100 in PBS/BSA for 45 minutes and blocked with 3% FCS, 0.1% Tween 20 in PBS for 1 hour at room temperature. Incubation with the primary antibody [anti-phospho-H3S10 (1:200; #06-570, Upstate), anti-Setd1a (1:50; A300-289, Bethyl), anti-Setd1b (1:50; A302-281A, Bethyl)] was for 1 hour in blocking buffer followed by washing and incubation with the secondary antibody (1:1000; goat anti-rabbit IgG Alexa488; A11034, Molecular Probes). For immunohistochemistry, 1 μ m paraffin sections were rehydrated, treated with 3% hydrogen peroxide and antigen unmasking was performed in 10 mM citrate buffer (pH 6.0) for 45 minutes. Sections were incubated in blocking serum (5% goat serum) for 1 hour at 37°C with the primary antibody, followed by incubation with the biotinylated secondary antibody and the streptavidin/peroxidase complex (ABC Elite Reagent). Sections were developed using the peroxidase substrate solution (liquid DAB substrate) and counterstained with Hematoxylin and Eosin (H&E) or Light Green.

Derivation of ESC lines

All conditional ESC lines (*Setd1a* and *Setd1b*, *F/F;RC/+* and *F/+;RC/+*) were isolated from blastocyst outgrowths in the presence of 1 μ M MEK inhibitor PD0325901 in ESC medium containing Serum Replacement. Established conditional ESC lines were transferred into FCS-containing ESC medium in the presence of the MEK inhibitor PD0325901 (1 μ M) and the GSK3 inhibitor CH99021 (3 μ M), together known as 2i (Ying et al., 2008). Cre-mediated recombination was induced with 1 μ M 4-hydroxytamoxifen (4OHT). Two conditional *Setd1a*^{F/F;RC/+} ESC lines were transfected with BACs via lipofection (Ciotta et al., 2011) and clones were selected in the presence of 4 μ g/ml blasticidin (BSD). The copy numbers of the integrated BACs were determined by quantitative PCR.

Cell proliferation and cell cycle analysis

The proliferation count assay is described elsewhere (Lubitz et al., 2007). For cell cycle analysis, asynchronously growing ESCs were fixed with ethanol and stained with 1 μ g/ml propidium iodide containing 0.1 mg/ml RNase A (both from Sigma-Aldrich). Cells were analyzed by flow cytometry, and G1, S, and G2/M cell cycle distribution was determined using ModFit LT software (Verity Software House, Topsham, ME, USA).

Alkaline phosphatase staining and apoptosis assays

ESCs were fixed in 4% formaldehyde, permeabilized with 0.5% Triton X-100/PBS, and stained with 25 mM Tris-maleate pH 9.0 (Sigma), 8 mM MgCl₂, 0.4% (w/v) 1-naphthyl phosphate (Sigma) and 1% (w/v) Fast Red TR (Sigma). To detect apoptotic cells, terminal deoxynucleotidyltransferase-mediated dUTP-biotin nick end labeling (TUNEL) staining was performed with the *In Situ* Cell Death Detection Kit (Roche) and ESCs were counterstained with DAPI. TUNEL staining on whole-mount embryos was performed accordingly. To determine the levels of apoptosis, ESCs were stained with APC-labeled annexin V according to the manufacturer's instructions (BD Biosciences). Cells were subjected to flow cytometry and analyzed by FACSDiva software (BD Biosciences).

Differentiation of ESCs to EpiSCs

ESCs were passaged using accutase (PAA), washed twice in DMEM/F12 (Gibco) and 1×10^4 cells per cm² were seeded on fibronectin-coated (10 ng/ml in PBS, Millipore) 6-well plates or glass coverslips in EpiSC medium [50% DMEM/F12, 50% Neurobasal medium (Gibco), 0.5 \times B27 supplement (Gibco), 0.5 \times self-made N2 supplement, 2 mM glutamate, 100 μ M β -mercaptoethanol, 12 ng/ml recombinant human FGF2 (Protein Facility, MPI-CBG), 20 ng/ml recombinant human activin A (MPI-CBG)]. Stable EpiSC lines were established after three to eight passages.

Derivation and differentiation of NSCs to neurons and astrocytes

Mouse NSCs were isolated from E15.5 *Setd1a* and *Setd1b* (*F/F;RC/+* and *F/+;RC/+*) forebrains and cultured in NSC medium [Euromed-N (Biozol), 1 \times N2 (self-made), 2 mM L-glutamine, 10 ng/ml recombinant FGF2 (MPI-CBG), 10 ng/ml EGF (Peprotech), supplemented with 0.5 \times B27]. NSCs were passaged using accutase, washed in DMEM/F12 and 1×10^5 cells per cm² were seeded on laminin-coated (2 μ g/ml, Roche) glass coverslips for subsequent fixation and permeabilization for immunofluorescence using mouse anti-nestin [1:20; Rat-401, Developmental Studies Hybridoma Bank (DSHB)] or mouse anti-Pax6 (1:20, DSHB) antibodies.

Reverse transcription and quantitative PCR (qRT-PCR) analysis

Total RNA from oocytes, embryos or cells was isolated using Trizol reagent (Ambion) and reverse transcribed using the SuperScript III First-Strand Synthesis System (Invitrogen). Real-time quantitative PCR was performed with Go Taq qPCR Master Mix (Promega) on a Mx3000P multiplex PCR instrument (Agilent). Ct values were normalized against *Gapdh* or *Rpl19*. Primer sequences and the lengths of the amplified products are given in supplementary material Table S1. Fold differences in expression levels were calculated according to the 2^{- $\Delta\Delta$ Ct} method (Livak and Schmittgen, 2001).

Western blot analysis

ESCs were homogenized in buffer E (20 mM HEPES pH 8.0, 350 mM NaCl, 10% glycerol, 0.1% Tween 20, 1 μ g/ml pepstatin A, 0.5 μ g/ml leupeptin, 2 μ g/ml aprotinin, 1 mM PMSF) and protein extracts were obtained after three cycles of freezing and thawing. Whole cell extracts were separated by SDS-PAGE (3-8% Tris-acetate gel for Setd1 proteins, 12% Tris-glycine gel for histone proteins), transferred to PVDF membranes and probed with primary antibodies: Setd1a (1:500; A300-289, Bethyl), Setd1b (1:500; A302-281A, Bethyl), alpha tubulin (1:10,000; 1878-1, Epitomics), H3 (1:3500; ab1791, Abcam), H3K4me1 (1:2500; ab889, Abcam), H3K4me2 (1:2500; ab7766, Abcam), H3K4me3 (1:1000; 07-473, Millipore), H3K27me3 (1:2500; 07-449, Millipore) and ubiquitylated H2B (1:2500; ab1790, Abcam).

NSCs and reprogramming

Mouse NSCs were cultured as described above. Cre-mediated recombination was induced with 0.1 μ M 4OHT. For reprogramming, 8×10^6 NSCs were nucleofected with circular PiggyBac-OSKM and HA-mPB transposase plasmids (Yusa et al., 2009) using the Mouse Neural Stem Cell Nucleofector Kit (Lonza). From day 2 onwards, medium was changed to FCS-based ESC medium. Fourteen days after nucleofection, induced pluripotent stem cell colonies from each well were picked onto mitomycin-C-treated mouse embryonic fibroblasts and expanded for verification of complete recombination by genotyping. Subsequent analysis comprised immunofluorescence staining for Nanog (1:200, sc1000, Calbiochem) and SSEA1 (1:10, MC-480, DSHB). Residual colonies on reprogramming plates were stained for alkaline phosphatase expression.

Acknowledgements

We thank the Biomedical Services (BMS) of the Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, for excellent service; Mandy Obst for technical assistance; Hans Schöler (MPI, Muenster, Germany) for *Oct4-GFP* mice; Rolf Jessberger (MTZ, TU Dresden, Germany) for *Gdf9-Cre* mice; and Allan Bradley and the Wellcome Trust Sanger Institute for kindly providing the PiggyBac-OSKM and HA-mPB transposase plasmids. Nestin, Pax6 and SSEA1 antibodies were obtained from the Developmental Studies Hybridoma Bank developed under the auspices of the NICHD and maintained by The University of Iowa, Department of Biological Sciences, Iowa City, IA 52242, USA.

Competing interests

The authors declare no competing financial interests.

Author contributions

A.S.B., K.S., K.N., A.K., A.F.S. and K.A. designed the experiments. A.S.B., K.S., A.K. and K.A. examined the embryonic phenotype. A.S.B., K.S., K.N., U.H. and K.A. generated and analyzed conditional ESCs, NSCs and EpiSCs and performed the reprogramming experiments. A.S.B., A.G., G.C., D.C.T. and J.F. generated DNA reagents. A.S.B., A.K., A.F.S. and K.A. wrote the manuscript.

Funding

This work was supported by funding from the Deutsche Forschungsgemeinschaft (DFG), priority program SPP1356 [AN 649/1-2 to K.A.; and STE 903/1-2 to A.F.S.].

Supplementary material

Supplementary material available online at <http://dev.biologists.org/lookup/suppl/doi:10.1242/dev.098152/-DC1>

References

- Anastassiadis, K., Fu, J., Patsch, C., Hu, S., Weidlich, S., Duereschke, K., Buchholz, F., Edenhofer, F. and Stewart, A. F. (2009). Dre recombinase, like Cre, is a highly efficient site-specific recombinase in *E. coli*, mammalian cells and mice. *Dis. Model. Mech.* **2**, 508-515.
- Andreu-Vieyra, C. V., Chen, R., Agno, J. E., Glaser, S., Anastassiadis, K., Stewart, A. F. and Matzuk, M. M. (2010). MLL2 is required in oocytes for bulk histone 3 lysine 4 trimethylation and transcriptional silencing. *PLoS Biol.* **8**, e1000453.
- Ang, Y. S., Tsai, S. Y., Lee, D. F., Monk, J., Su, J., Ratnakumar, K., Ding, J., Ge, Y., Dang, H., Chang, B. et al. (2011). Wdr5 mediates self-renewal and reprogramming via the embryonic stem cell core transcriptional network. *Cell* **145**, 183-197.
- Ardehali, M. B., Mei, A., Zobeck, K. L., Caron, M., Lis, J. T. and Kusch, T. (2011). Drosophila Set1 is the major histone H3 lysine 4 trimethyltransferase with role in transcription. *EMBO J.* **30**, 2817-2828.
- Bannister, A. J. and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. *Cell Res.* **21**, 381-395.
- Breen, T. R. and Harte, P. J. (1993). Trithorax regulates multiple homeotic genes in the bithorax and Antennapedia complexes and exerts different tissue-specific, parasegment-specific and promoter-specific effects on each. *Development* **117**, 119-134.
- Briggs, S. D., Bryk, M., Strahl, B. D., Cheung, W. L., Davie, J. K., Dent, S. Y., Winston, F. and Allis, C. D. (2001). Histone H3 lysine 4 methylation is mediated by Set1 and required for cell growth and rDNA silencing in *Saccharomyces cerevisiae*. *Genes Dev.* **15**, 3286-3295.
- Carlone, D. L. and Skalnik, D. G. (2001). CpG binding protein is crucial for early embryonic development. *Mol. Cell Biol.* **21**, 7601-7606.
- Carlone, D. L., Lee, J. H., Young, S. R., Dobrota, E., Butler, J. S., Ruiz, J. and Skalnik, D. G. (2005). Reduced genomic cytosine methylation and defective cellular differentiation in embryonic stem cells lacking CpG binding protein. *Mol. Cell Biol.* **25**, 4881-4891.
- Ciotto, G., Hofemeister, H., Maresca, M., Fu, J., Sarov, M., Anastassiadis, K. and Stewart, A. F. (2011). Recombineering BAC transgenes for protein tagging. *Methods* **53**, 113-119.
- Clouaire, T., Webb, S., Skene, P., Illingworth, R., Kerr, A., Andrews, R., Lee, J. H., Skalnik, D. and Bird, A. (2012). Cfp1 integrates both CpG content and gene activity for accurate H3K4me3 deposition in embryonic stem cells. *Genes Dev.* **26**, 1714-1728.
- Eberl, H. C., Spruijt, C. G., Kelstrup, C. D., Vermeulen, M. and Mann, M. (2013). A map of general and specialized chromatin readers in mouse tissues generated by label-free interaction proteomics. *Mol. Cell* **49**, 368-378.
- Ernst, P., Fisher, J. K., Avery, W., Wade, S., Foy, D. and Korsmeyer, S. J. (2004). Definitive hematopoiesis requires the mixed-lineage leukemia gene. *Dev. Cell* **6**, 437-443.
- Fu, J., Teucher, M., Anastassiadis, K., Skarnes, W. and Stewart, A. F. (2010). A recombineering pipeline to make conditional targeting constructs. *Methods Enzymol.* **477**, 125-144.
- Glaser, S., Schaff, J., Lubitz, S., Vintersten, K., van der Hoeven, F., Tufteland, K. R., Aasland, R., Anastassiadis, K., Ang, S. L. and Stewart, A. F. (2006). Multiple epigenetic maintenance factors implicated by the loss of Mll2 in mouse development. *Development* **133**, 1423-1432.
- Glaser, S., Lubitz, S., Loveland, K. L., Ohbo, K., Robb, L., Schwenk, F., Seibler, J., Roellig, D., Kranz, A., Anastassiadis, K. et al. (2009). The histone 3 lysine 4 methyltransferase, Mll2, is only required briefly in development and spermatogenesis. *Epigenetics Chromatin* **2**, 5.
- Hallson, G., Hollebakken, R. E., Li, T., Syrzycka, M., Kim, I., Cotsworth, S., Fitzpatrick, K. A., Sinclair, D. A. and Honda, B. M. (2012). dSet1 is the main H3K4 di- and tri-methyltransferase throughout Drosophila development. *Genetics* **190**, 91-100.
- Herz, H. M., Mohan, M., Garruss, A. S., Liang, K., Takahashi, Y. H., Mickey, K., Voets, O., Verrijzer, C. P. and Shilatfard, A. (2012). Enhancer-associated H3K4 monomethylation by Trithorax-related, the Drosophila homolog of mammalian Mll3/Mll4. *Genes Dev.* **26**, 2604-2620.
- Hoffmeyer, K., Raggioli, A., Rudloff, S., Anton, R., Hierholzer, A., Del Valle, I., Hein, K., Vogt, R. and Kemler, R. (2012). Wnt/ β -catenin signaling regulates telomerase in stem cells and cancer cells. *Science* **336**, 1549-1554.
- Hughes, C. M., Rozenblatt-Rosen, O., Milne, T. A., Copeland, T. D., Levine, S. S., Lee, J. C., Hayes, D. N., Shanmugam, K. S., Bhattacharjee, A., Biondi, C. A. et al. (2004). Menin associates with a trithorax family histone methyltransferase complex and with the hox8 locus. *Mol. Cell* **13**, 587-597.
- Jiang, H., Shukla, A., Wang, X., Chen, W. Y., Bernstein, B. E. and Roeder, R. G. (2011). Role for Dpy-30 in ES cell-fate specification by regulation of H3K4 methylation within bivalent domains. *Cell* **144**, 513-525.
- Jude, C. D., Climer, L., Xu, D., Artinger, E., Fisher, J. K. and Ernst, P. (2007). Unique and independent roles for MLL in adult hematopoietic stem cells and progenitors. *Cell Stem Cell* **1**, 324-337.
- Kim, J., Kim, J. A., McGinty, R. K., Nguyen, U. T., Muir, T. W., Allis, C. D. and Roeder, R. G. (2013). The n-SET domain of Set1 regulates H2B ubiquitylation-dependent H3K4 methylation. *Mol. Cell* **49**, 1121-1133.
- Kooistra, S. M. and Helin, K. (2012). Molecular mechanisms and potential functions of histone demethylases. *Nat. Rev. Mol. Cell Biol.* **13**, 297-311.
- Kranz, A., Fu, J., Duereschke, K., Weidlich, S., Naumann, R., Stewart, A. F. and Anastassiadis, K. (2010). An improved Flp deleter mouse in C57Bl/6 based on Flpo recombination. *Genesis* **48**, 512-520.
- Krogan, N. J., Dover, J., Khorrami, S., Greenblatt, J. F., Schneider, J., Johnston, M. and Shilatfard, A. (2002). COMPASS, a histone H3 (Lysine 4) methyltransferase required for telomeric silencing of gene expression. *J. Biol. Chem.* **277**, 10753-10755.
- Kuzin, B., Tiliib, S., Sedkov, Y., Mizrokh, L. and Mazo, A. (1994). The Drosophila trithorax gene encodes a chromosomal protein and directly regulates the region-specific homeotic gene fork head. *Genes Dev.* **8**, 2478-2490.
- Lallemant, Y., Luria, V., Haffner-Krausz, R. and Lonai, P. (1998). Maternally expressed PGK-Cre transgene as a tool for early and uniform activation of the Cre site-specific recombinase. *Transgenic Res.* **7**, 105-112.
- Lan, Z. J., Xu, X. and Cooney, A. J. (2004). Differential oocyte-specific expression of Cre recombinase activity in GDF-9-iCre, Zp3cre, and Mx2Cre transgenic mice. *Biol. Reprod.* **71**, 1469-1474.
- Lee, J. H. and Skalnik, D. G. (2005). CpG-binding protein (CXXC finger protein 1) is a component of the mammalian Set1 histone H3-Lys4 methyltransferase complex, the analogue of the yeast Set1/COMPASS complex. *J. Biol. Chem.* **280**, 41725-41731.
- Lee, J. H. and Skalnik, D. G. (2008). Wdr82 is a C-terminal domain-binding protein that recruits the Setd1A Histone H3-Lys4 methyltransferase complex to transcription start sites of transcribed human genes. *Mol. Cell Biol.* **28**, 609-618.
- Lee, J. H. and Skalnik, D. G. (2012). Rbm15-Mkl1 interacts with the Setd1b histone H3-Lys4 methyltransferase via a SPOC domain that is required for cytokine-independent proliferation. *PLoS ONE* **7**, e42965.
- Lee, J. H., Tate, C. M., You, J. S. and Skalnik, D. G. (2007). Identification and characterization of the human Set1B histone H3-Lys4 methyltransferase complex. *J. Biol. Chem.* **282**, 13419-13428.
- Livak, K. J. and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. *Methods* **25**, 402-408.
- Lubitz, S., Glaser, S., Schaff, J., Stewart, A. F. and Anastassiadis, K. (2007). Increased apoptosis and skewed differentiation in mouse embryonic stem cells lacking the histone methyltransferase Mll2. *Mol. Biol. Cell* **18**, 2356-2366.
- McMahon, K. A., Hiew, S. Y., Hadjir, S., Veiga-Fernandes, H., Menzel, U., Price, A. J., Kioussis, D., Williams, O. and Brady, H. J. (2007). Mll has a critical role in fetal and adult hematopoietic stem cell self-renewal. *Cell Stem Cell* **1**, 338-345.

- Miller, T., Krogan, N. J., Dover, J., Erdjument-Bromage, H., Tempst, P., Johnston, M., Greenblatt, J. F. and Shilatifard, A. (2001). COMPASS: a complex of proteins associated with a trithorax-related SET domain protein. *Proc. Natl. Acad. Sci. USA* **98**, 12902-12907.
- Milne, T. A., Briggs, S. D., Brock, H. W., Martin, M. E., Gibbs, D., Allis, C. D. and Hess, J. L. (2002). MLL targets SET domain methyltransferase activity to Hox gene promoters. *Mol. Cell* **10**, 1107-1117.
- Milne, T. A., Kim, J., Wang, G. G., Stadler, S. C., Basur, V., Whitcomb, S. J., Wang, Z., Ruthenburg, A. J., Elenitoba-Johnson, K. S., Roeder, R. G. et al. (2010). Multiple interactions recruit MLL1 and MLL1 fusion proteins to the HOXA9 locus in leukemogenesis. *Mol. Cell* **38**, 853-863.
- Mohan, M., Herz, H. M., Smith, E. R., Zhang, Y., Jackson, J., Washburn, M. P., Florens, L., Eissenberg, J. C. and Shilatifard, A. (2011). The COMPASS family of H3K4 methylases in Drosophila. *Mol. Cell. Biol.* **31**, 4310-4318.
- Nagy, Z. and Tora, L. (2007). Distinct GCN5/PCAF-containing complexes function as co-activators and are involved in transcription factor and global histone acetylation. *Oncogene* **26**, 5341-5357.
- Nichols, J., Zevnik, B., Anastasiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Schöler, H. and Smith, A. (1998). Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* **95**, 379-391.
- Noma, K. and Grewal, S. I. (2002). Histone H3 lysine 4 methylation is mediated by Set1 and promotes maintenance of active chromatin states in fission yeast. *Proc. Natl. Acad. Sci. USA* **99** Suppl. 4, 16438-16445.
- Noma, K., Allis, C. D. and Grewal, S. I. (2001). Transitions in distinct histone H3 methylation patterns at the heterochromatin domain boundaries. *Science* **293**, 1150-1155.
- Rodríguez, C. I., Buchholz, F., Galloway, J., Sequerra, R., Kasper, J., Ayala, R., Stewart, A. F. and Dymecki, S. M. (2000). High-efficiency deleter mice show that FLPe is an alternative to Cre-loxP. *Nat. Genet.* **25**, 139-140.
- Roguev, A., Schaft, D., Shevchenko, A., Pijnappel, W. W., Wilm, M., Aasland, R. and Stewart, A. F. (2001). The *Saccharomyces cerevisiae* Set1 complex includes an Ash2 homologue and methylates histone 3 lysine 4. *EMBO J.* **20**, 7137-7148.
- Roguev, A., Schaft, D., Shevchenko, A., Aasland, R., Shevchenko, A. and Stewart, A. F. (2003). High conservation of the Set1/Rad6 axis of histone 3 lysine 4 methylation in budding and fission yeasts. *J. Biol. Chem.* **278**, 8487-8493.
- Ruthenburg, A. J., Allis, C. D. and Wysocka, J. (2007). Methylation of lysine 4 on histone H3: intricacy of writing and reading a single epigenetic mark. *Mol. Cell* **25**, 15-30.
- Sedkov, Y., Cho, E., Petruk, S., Cherbass, L., Smith, S. T., Jones, R. S., Cherbass, P., Canaani, E., Jaynes, J. B. and Mazo, A. (2003). Methylation at lysine 4 of histone H3 in ecdysone-dependent development of Drosophila. *Nature* **426**, 78-83.
- Seibler, J., Zevnik, B., Küter-Luks, B., Andreas, S., Kern, H., Hennek, T., Rode, A., Heimann, C., Faust, N., Kauselmann, G. et al. (2003). Rapid generation of inducible mouse mutants. *Nucleic Acids Res.* **31**, e12.
- Seward, D. J., Cubberley, G., Kim, S., Schonewald, M., Zhang, L., Tripet, B. and Bentley, D. L. (2007). Demethylation of trimethylated histone H3 Lys4 in vivo by JARID1 JmjC proteins. *Nat. Struct. Mol. Biol.* **14**, 240-242.
- Shi, X., Kachirskaja, I., Walter, K. L., Kuo, J. H., Lake, A., Davrazou, F., Chan, S. M., Martin, D. G., Fingerman, I. M., Briggs, S. D. et al. (2007). Proteome-wide analysis in *Saccharomyces cerevisiae* identifies several PHD fingers as novel direct and selective binding modules of histone H3 methylated at either lysine 4 or lysine 36. *J. Biol. Chem.* **282**, 2450-2455.
- Stoller, J. Z., Huang, L., Tan, C. C., Huang, F., Zhou, D. D., Yang, J., Gelb, B. D. and Epstein, J. A. (2010). Ash2l interacts with Tbx1 and is required during early embryogenesis. *Exp. Biol. Med.* **235**, 569-576.
- Suganuma, T. and Workman, J. L. (2011). Signals and combinatorial functions of histone modifications. *Annu. Rev. Biochem.* **80**, 473-499.
- Szabó, P. E., Hübner, K., Schöler, H. and Mann, J. R. (2002). Allele-specific expression of imprinted genes in mouse migratory primordial germ cells. *Mech. Dev.* **115**, 157-160.
- Tate, C. M., Lee, J. H. and Skalnik, D. G. (2009). CXXC finger protein 1 contains redundant functional domains that support embryonic stem cell cytosine methylation, histone methylation, and differentiation. *Mol. Cell. Biol.* **29**, 3817-3831.
- Tate, C. M., Lee, J. H. and Skalnik, D. G. (2010). CXXC finger protein 1 restricts the Setd1A histone H3K4 methyltransferase complex to euchromatin. *FEBS J.* **277**, 210-223.
- Testa, G., Schaft, J., van der Hoeven, F., Glaser, S., Anastasiadis, K., Zhang, Y., Hermann, T., Stremmel, W. and Stewart, A. F. (2004). A reliable lacZ expression reporter cassette for multipurpose, knockout-first alleles. *Genesis* **38**, 151-158.
- Trésaugues, L., Dehé, P. M., Guérois, R., Rodríguez-Gil, A., Varlet, I., Salah, P., Pamblanco, M., Luciano, P., Quevillon-Cheruel, S., Sollier, J. et al. (2006). Structural characterization of Set1 RNA recognition motifs and their role in histone H3 lysine 4 methylation. *J. Mol. Biol.* **359**, 1170-1181.
- van Nuland, R., Smits, A. H., Pallaki, P., Jansen, P. W., Vermeulen, M. and Timmers, H. T. (2013). Quantitative dissection and stoichiometry determination of the human SET1/MLL histone methyltransferase complexes. *Mol. Cell. Biol.* **33**, 2067-2077.
- Wan, M., Liang, J., Xiong, Y., Shi, F., Zhang, Y., Lu, W., He, Q., Yang, D., Chen, R., Liu, D. et al. (2013). The trithorax group protein Ash2l is essential for pluripotency and maintaining open chromatin in embryonic stem cells. *J. Biol. Chem.* **288**, 5039-5048.
- Yagi, H., Deguchi, K., Aono, A., Tani, Y., Kishimoto, T. and Komori, T. (1998). Growth disturbance in fetal liver hematopoiesis of Mll-mutant mice. *Blood* **92**, 108-117.
- Ying, Q. L., Wray, J., Nichols, J., Battle-Morera, L., Doble, B., Woodgett, J., Cohen, P. and Smith, A. (2008). The ground state of embryonic stem cell self-renewal. *Nature* **453**, 519-523.
- Yokoyama, A., Wang, Z., Wysocka, J., Sanyal, M., Aufiero, D. J., Kitabayashi, I., Herr, W. and Cleary, M. L. (2004). Leukemia proto-oncoprotein MLL forms a SET1-like histone methyltransferase complex with menin to regulate Hox gene expression. *Mol. Cell. Biol.* **24**, 5639-5649.
- Yu, B. D., Hess, J. L., Horning, S. E., Brown, G. A. and Korsmeyer, S. J. (1995). Altered Hox expression and segmental identity in Mll-mutant mice. *Nature* **378**, 505-508.
- Yusa, K., Rad, R., Takeda, J. and Bradley, A. (2009). Generation of transgene-free induced pluripotent mouse stem cells by the piggyBac transposon. *Nat. Methods* **6**, 363-369.