© 2016. Published by The Company of Biologists Ltd | Development (2016) 143, 4595-4607 doi:10.1242/dev.141465

Setd5 is essential for mammalian development and the co-transcriptional regulation of histone acetylation

Anna B. Osipovich[1,2], Rama Gangula[1], Pedro G. Vianna[1] and Mark A. Magnuson[1,2,*]

ABSTRACT

SET domain-containing proteins play a vital role in regulating gene expression during development through modifications in chromatin structure. Here we show that SET domain-containing 5 (*Setd5*) is divergently transcribed with *Gt(ROSA26)Sor*, is necessary for mammalian development, and interacts with the PAF1 co-transcriptional complex and other proteins. *Setd5*-deficient mouse embryos exhibit severe defects in neural tube formation, somitogenesis and cardiac development, have aberrant vasculogenesis in embryos, yolk sacs and placentas, and die between embryonic day 10.5 and 11.5. *Setd5*-deficient embryonic stem cells have impaired cellular proliferation, increased apoptosis, defective cell cycle progression, a diminished ability to differentiate into cardiomyocytes and greatly perturbed gene expression. SETD5 co-immunoprecipitates with multiple components of the PAF1 and histone deacetylase-containing NCoR complexes and is not solely required for major histone lysine methylation marks. In the absence of *Setd5*, histone acetylation is increased at transcription start sites and near downstream regions. These findings suggest that SETD5 functions in a manner similar to yeast Set3p and *Drosophila* UpSET, and that it is essential for regulating histone acetylation during gene transcription.

KEY WORDS: Setd5, Embryogenesis, Histone acetylation, Transcription, PAF1

INTRODUCTION

Embryogenesis requires dynamic and specific changes in gene expression patterns that involve both genetic and epigenetic mechanisms (Cantone and Fisher, 2013). Histone modifications, which include acetylation, methylation, phosphorylation and monoubiquitylation, represent one such epigenetic mechanism and play an essential role in regulating chromatin accessibility during transcription and genome maintenance (Kouzarides, 2007). Since nucleosomes present a physical barrier to gene transcription by RNA polymerase II (RNAP II), the co-transcriptional nucleosome disassembly and reassembly that occurs is not only highly dynamic but also involves rapid changes in specific histone modifications (Venkatesh and Workman, 2015).

Reversible histone lysine acetylations are installed by histone acetyltransferases (HATs) and removed by histone deacetylases (HDACs) (Verdone et al., 2005). While the acetylation of histones H3 and H4 has long been associated with transcriptionally active genes and open chromatin, more recent studies have established that histone methylations are also associated with actively transcribed genes, including histone H3 methylation at lysines 4, 36 and 79 (H3K4me, H3K36me and H3K79me) and histone H4 methylation at lysine 20 (H4K20me) (Smolle and Workman, 2013). Studies in yeast have shown that these histone modifications occur co-transcriptionally and require coordination between RNAP II, elongation complexes and regulators of histone modification (Tanny, 2014).

Polymerase-associated factor 1 complex (PAF1C), which contains Paf1, Ctr9, Leo1, Cdc73 and Rtf1 as core components, colocalizes to coding regions of genes and is involved in multiple processes during RNAP II transcription. These processes include RNA elongation, coordination between transcriptional activators and mRNA processing machinery, and recruitment and activation of histone modification factors (Crisucci and Arndt, 2011; Jaehning, 2010). Moreover, there is a crosstalk between histone methylation (at H3K4 and H3K36) and histone acetylation in the coding regions of genes, which involves methyl-lysine recognition by multiple HAT and HDAC complexes (Musselman and Kutateladze, 2011). These co-transcriptional processes, and the proteins that modulate them, are less well studied in mammalian species than in lower organisms.

Histone methylation is catalyzed by histone methyltransferases, which mediate the mono-, di- and tri-methylation of lysine ε-amine groups (Del Rizzo and Trievel, 2011). The majority of histone methyltransferases contain an evolutionarily conserved SET domain, which was originally identified based on similarities between Su(var)3-9, Enhancer of zeste and Trithorax proteins (Xiao et al., 2003). Mice that lack various SET domain-containing histone methyltransferases exhibit a wide range of developmental defects. For instance, animals lacking *Setdb1*, *Setd8* (*Kmt5a*) or *Nsd1* exhibit defects prior to implantation (Dodge et al., 2004; Oda et al., 2009; Rayasam et al., 2003); animals without *Mll1* (*Kmt2a*), *Ehmt2* or *Setd2* die during mid-gestation (Hu et al., 2010; Tachibana et al., 2002; Yu et al., 1995); and mice lacking *Setd7*, *Mll3* (*Kmt2c*) or *Suv420h1*/*h2* (*Kmt5b/c*) are born but die shortly thereafter (Kurash et al., 2008; Lee et al., 2006; Schotta et al., 2008).

Our discovery of a lethal phenotype when we deleted a 5.16 kb fragment of *Gt(ROSA)26Sor*, a widely used locus for ubiquitous transgene expression (Soriano, 1999; Zambrowicz et al., 1997), prompted us to explore the function of the adjacent *Setd5*, a previously uncharacterized mammalian SET domain-containing gene. Orthologs of SETD5 include mammalian MLL5 (KMT2E), yeast Set3p and *Drosophila* UpSET, all of which contain a distantly conserved SET domain (Glaser et al., 2006). Interestingly, methyltransferase activity has not been associated with any of the SETD5 orthologs (Pijnappel et al., 2001; Rincon-Arano et al., 2012; Sebastian et al., 2009). However, Set3p, a key component of the yeast SET3C complex, has been shown to modulate gene transcription, repress sporulation and promote cytokinesis by

[1]Center for Stem Cell Biology, Vanderbilt University, Nashville, TN 37232, USA. [2]Department of Molecular Physiology and Biophysics, Vanderbilt University, Nashville, TN 37232, USA.

*Author for correspondence (mark.magnuson@vanderbilt.edu)

A.B.O., 0000-0002-9513-2262; M.A.M., 0000-0002-8824-6499

Received 22 June 2016; Accepted 6 November 2016

Development (2016) 143, 4595-4607 doi:10.1242/dev.141465

altering histone deacetylation (Kim and Buratowski, 2009; Pijnappel et al., 2001; Rentas et al., 2012). Moreover, based on protein homologies, the SET3C complex has been suggested to be the yeast analog of the mammalian NCoR co-repressor complex that interacts with, and mediates the repressive activity of, unliganded nuclear receptors and other transcription factors (Pijnappel et al., 2001).

The mammalian NCOR1/NCOR2 (SMRT) co-repressors, which play a vital role in silencing gene expression, exist in a complex with histone deacetylase 3 (HDAC3) and the WD40 repeat-containing proteins TBL1X and TBL1XR1 (Guenther et al., 2000; Wen et al., 2000; Yoon et al., 2003). Consistent with a broad role in regulating gene expression, the NCoR co-repressor complex has been shown to be essential for the development of multiple organ systems, including the nervous system and heart (Jepsen et al., 2000, 2007), the maintenance of genome integrity (Bhaskara et al., 2010) and for metabolic regulation (Mottis et al., 2013). However, whereas yeast Set3p is a clearly established component of the SET3C complex, no SET domain-containing protein has, to date, been definitively established as part of the mammalian NCoR complex. Indeed, although MLL5 has been suggested to be a part of the NCoR complex, no direct evidence supporting this assertion has been reported (Kittler et al., 2007). Moreover, the phenotype of *Mll5* knockout mice is surprisingly mild, with only mild hematological abnormalities and male sterility reported (Heuser et al., 2009; Yap et al., 2011).

We examined the role of *Setd5* in mouse development and found that mice lacking *Setd5* die by embryonic day (E) 10.5 due to cardiovascular defects, and that *Setd5* is required both for normal cell cycle progression and the regulation of chromatin accessibility during transcription. Furthermore, we show that SETD5 interacts with components of the co-transcriptional PAF1 and the deacetylating NCoR co-repressor complexes and that, in the absence of *Setd5*, histone acetylation is not restricted to transcriptional start regions but is extended towards the 3′ end of the gene. Moreover, mouse embryonic stem cells (mESCs) lacking *Setd5* have a dysregulated transcriptome, including altered expression of myogenesis and vasculogenesis genes. Finally, we show that *Setd5* and *ROSA26* are divergently transcribed from a bidirectional promoter, and that *ROSA26*, a noncoding mRNA, may help maintain *Setd5* expression.

RESULTS

Derivation of a *Setd5^{GFP}* allele

Mice that lack 5.16 kb of DNA containing the promoter and exon 1 of *ROSA26* (Chen et al., 2011) were observed to be homozygous lethal. Examination of the deleted region revealed that the *Setd5* and *ROSA26* genes were in close proximity and in opposite orientation, with their respective transcription start sites (TSSs) separated by only 393 bp (Fig. S1). Thus, we sought to determine whether the *ROSA26* deletion also impaired the expression of *Setd5*. To assess the function and sites of *Setd5* expression in a manner that did not adversely affect *ROSA26* expression, we derived mice that placed GFP coding sequences in the first exon of *Setd5*, upstream of the translation initiation sequence, using an RMCE-based strategy (Fig. S2A,B). The disruption of full-length *Setd5* mRNA and protein expression was confirmed by qPCR, northern and western blot (Fig. S2C-E), confirming that *Setd5^{GFP}* is a null allele.

Setd5 and *ROSA26* are ubiquitously co-expressed from a bidirectional promoter

To identify sites of *Setd5* expression during embryogenesis and in adult animals, we performed fluorescent stereoscopy on E8.0 to E15.5 embryos and on tissues from adult animals that were heterozygous for the *Setd5^{GFP}* allele. Fluorescence was observed in all embryos and adult tissues (Fig. 1B, Fig. S3) in a pattern that closely mimicked that described for *ROSA26* (Soriano, 1999). RT-qPCR analysis confirmed that *Setd5* and *ROSA26* are co-expressed in all tissues (Fig. 1A). By generating double-heterozygous mice (*Setd5^{GFP}*; *ROSA26^{Cherry}*) that express GFP from the *Setd5* promoter and mCherry from *ROSA26* (Chen et al., 2011), we observed very similar expression patterns for both fluorescent proteins (Fig. 1B, Fig. S4).

To confirm the bidirectionality of the *Setd5*/*ROSA26* (SR) promoter, we inserted it into a bidirectional dual luciferase reporter plasmid, which was transfected into a mouse cell line. This revealed robust transcriptional activity in both directions (Fig. 1C), which, together with the tissue co-expression data described above, convincingly indicates that *Setd5* and *ROSA26* are transcribed in opposite directions from the same promoter, making them a co-transcribed divergent transcript pair.

ROSA26 helps to maintain *Setd5* expression

Bidirectional promoters that produce a divergent pair comprising a protein-coding and a noncoding RNA have been suggested to serve as a coupling regulatory unit in which transcriptional activity in the noncoding RNA direction, or the transcript itself, modulates expression of the protein-coding gene (Wei et al., 2011). To test whether the noncoding *ROSA26* transcript modulates *Setd5* expression, we examined the effects of *Setd5* knockout and of several different *ROSA26* knockout alleles (Fig. 2A) on the expression of both transcripts by RT-qPCR analysis of RNA isolated from E9.5 embryos. When *Setd5* mRNA expression was disrupted, *ROSA26* expression was not significantly altered (Fig. 2B). When *ROSA26* mRNA was disrupted at the beginning of the first exon, as in *ROSA26$^{Cherry/+}$* and *ROSA26$^{Cherry/Cherry}$* mice, there was a decrease in *Setd5* expression ($P \leq 0.01$; Fig. 2C). However, in *ROSA26$^{LSL.YFP/LSL.YFP}$* mice, in which the first exon of *ROSA26* is transcribed and spliced but a full-length transcript is not produced, *Setd5* expression was not altered (Fig. 2C). When the entire *ROSA26* transcript and 181 bp of preceding promoter sequence were deleted, as in the *ROSA26$^{228.3TF.GFP}$* allele, *Setd5* transcription was decreased by about half and two-thirds in heterozygotes and homozygotes, respectively, compared with the wild-type allele ($P \leq 0.01$; Fig. 2D). Transfection of a bidirectional dual luciferase reporter driven by a similarly truncated version of the *Setd5*/*ROSA26* promoter (SRΔ181) also exhibited a decrease in luciferase activity in both directions as compared with the full-length promoter ($P \leq 0.01$; Fig. 2E).

Taken together, these results suggest that both the *ROSA26* promoter and first exon must be intact for robust transcription of *ROSA26* to occur and for normal levels of *Setd5* mRNA expression. However, *Setd5* expression does not depend on the presence or absence of a full-length *ROSA26* transcript.

Setd5 is essential for embryonic growth and development

Since live-born *Setd5* null pups were not observed from heterozygous matings, we isolated embryos at E8.5 to 12.5 and observed no differences between wild type and heterozygotes. No viable null embryos were observed after E10.5 (Table S2) and all E10.5 embryos were underdeveloped compared with their heterozygous littermates, with many exhibiting widespread hemorrhage (Fig. S5B-D). Moreover, at E9.5, all null embryos were smaller, had fewer somites, and many had neural tubes that were not closed (Fig. 3A,B). Embryonic patterning defects also

Development (2016) 143, 4595-4607 doi:10.1242/dev.141465

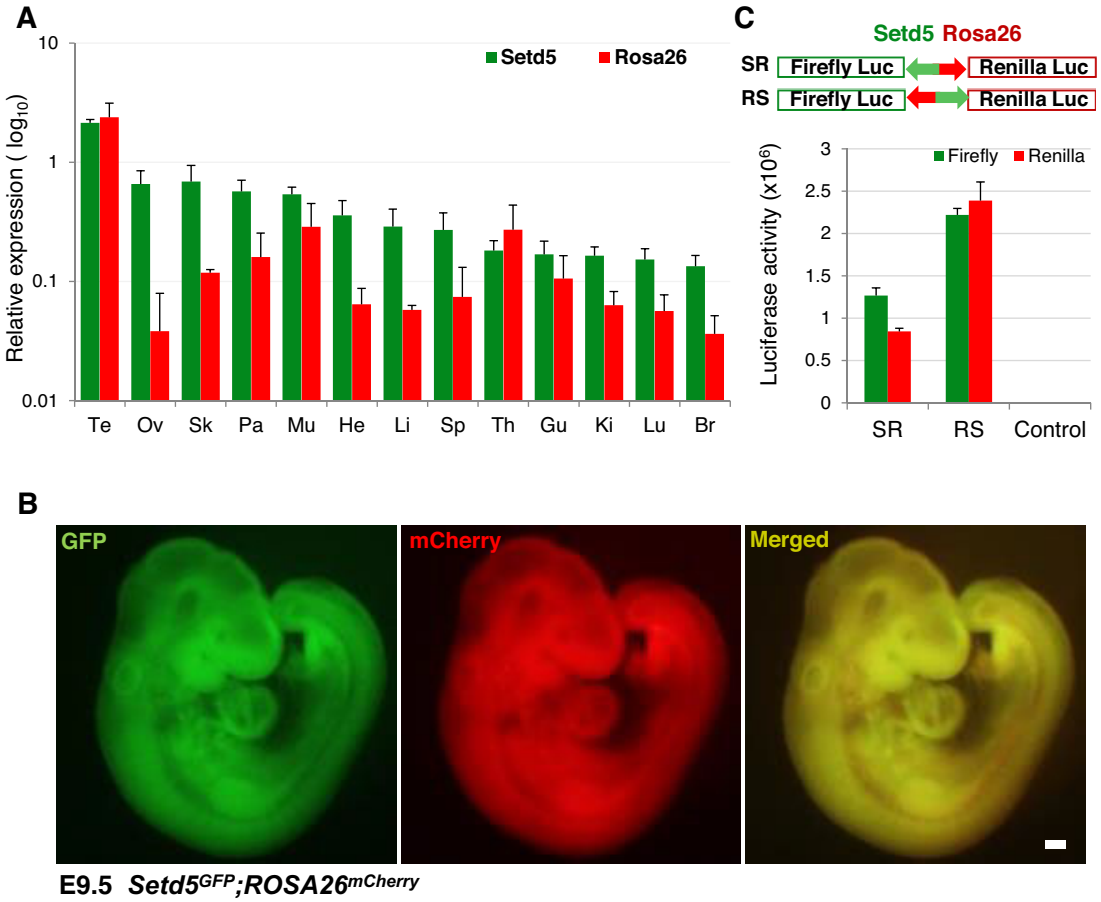

Fig. 1. Co-expression of *Setd5* and *ROSA26* transcripts from a bidirectional promoter. (A) RT-qPCR of *Setd5* and *ROSA26* mRNA levels in adult mouse tissues demonstrates ubiquitous expression of both transcripts. Te, testis; Ov, ovary; Sk, skin; Pa, pancreas; Mu, muscle; He, heart; Li, liver; Sp, spleen; Th, thymus; Gu, gut; Ki, kidney; Lu, lung; Br, brain. (B) Fluorescent imaging of whole-mount E9.5 mouse embryo demonstrates ubiquitous expression and colocalization of GFP and mCherry fluorescent proteins expressed from $Setd5^{GFP}$ and $ROSA26^{Cherry}$ knock-in alleles, respectively. Scale bar: 100 µm. (C) Dual luciferase assay demonstrates bidirectional activity of the *Setd5*/*ROSA26* promoter in direct (SR) and opposite (RS) orientations. n=3; error bars indicate s.e.m.

included massive pericardial effusions (or swelling) and defects in the formation of caudal structures, with the presence of tail effusions, suggesting disrupted blood flow and osmotic imbalance, and pointing to a defect in cardiovascular development. Developmental impairments were already noticeable at E8.5 (Fig. S5A).

To determine the basis of the failure of the *Setd5* null embryos to thrive we examined rates of cellular proliferation and apoptosis in tissues at E9.5 (Fig. 3C-E). BrdU incorporation experiments revealed a low labeling efficiency in knockout embryos (data not shown), possibly due to impaired tissue perfusion. Ki67 staining showed a decrease in cell proliferation rates from 96% to 87% in the heterozygous versus knockout tissues ($P \leq 0.01$). At the same time, the number of apoptotic cells, as measured by the presence of activated caspase 3, was increased from 4% to 25% ($P \leq 0.01$).

Cardiovascular abnormalities in *Setd5* null embryos

We further explored the cardiovascular defects and observed that only about half of *Setd5* null embryos had beating hearts at E9.5. Defects in heart development were further confirmed by whole-mount staining for two markers of cardiac differentiation: MEF2C, a transcription factor important for anterior heart field development (Lin et al., 1997); and myosin heavy chain (MHC), a structural protein enriched in cardiac muscle (Fig. S6A). At E9.5, MEF2C staining of heterozygous embryos showed a well-developed pharyngeal mesoderm/anterior heart field, cardiac muscle in a looped heart tube, and somites. By contrast, *Setd5* null embryos had swollen, underdeveloped hearts with only weak staining in the anterior heart field and myocardium. Similarly, whereas MHC staining of $Setd5^{GFP/+}$ mice showed easily discernible cardiac muscle, and somites and smooth muscle in the dorsal aorta, only the heart region was stained in the knockout embryos. At a histological level, the ventricular chambers exhibited reduced trabeculation in the existing single ventricle, a thinner myocardium layer, a large gap between pericardium and myocardium, and no contact between the myocardium and endocardium (Fig. S6B).

Setd5 mutant embryos also exhibited a severe vascular phenotype in the extraembryonic tissues and embryo bodies (Fig. 4A). At E9.5, the knockout embryos had a pale yolk sac with fewer blood vessels. Whole-mount staining for the endothelial marker PECAM1 revealed that whereas heterozygous yolk sacs exhibited a well-structured hierarchical organization of large and small vessels, the yolk sacs of *Setd5* null mice had a very poorly organized and abnormally dilated capillary plexus. Whole-mount staining for CD41 (ITGA2B), a marker for early hematopoietic cells, was observed in blood islands of both heterozygous and knockout yolk sacs, suggesting that definitive hematopoiesis occurred in the null embryos but that blood vessels were distended and contained fewer

Development (2016) 143, 4595-4607 doi:10.1242/dev.141465

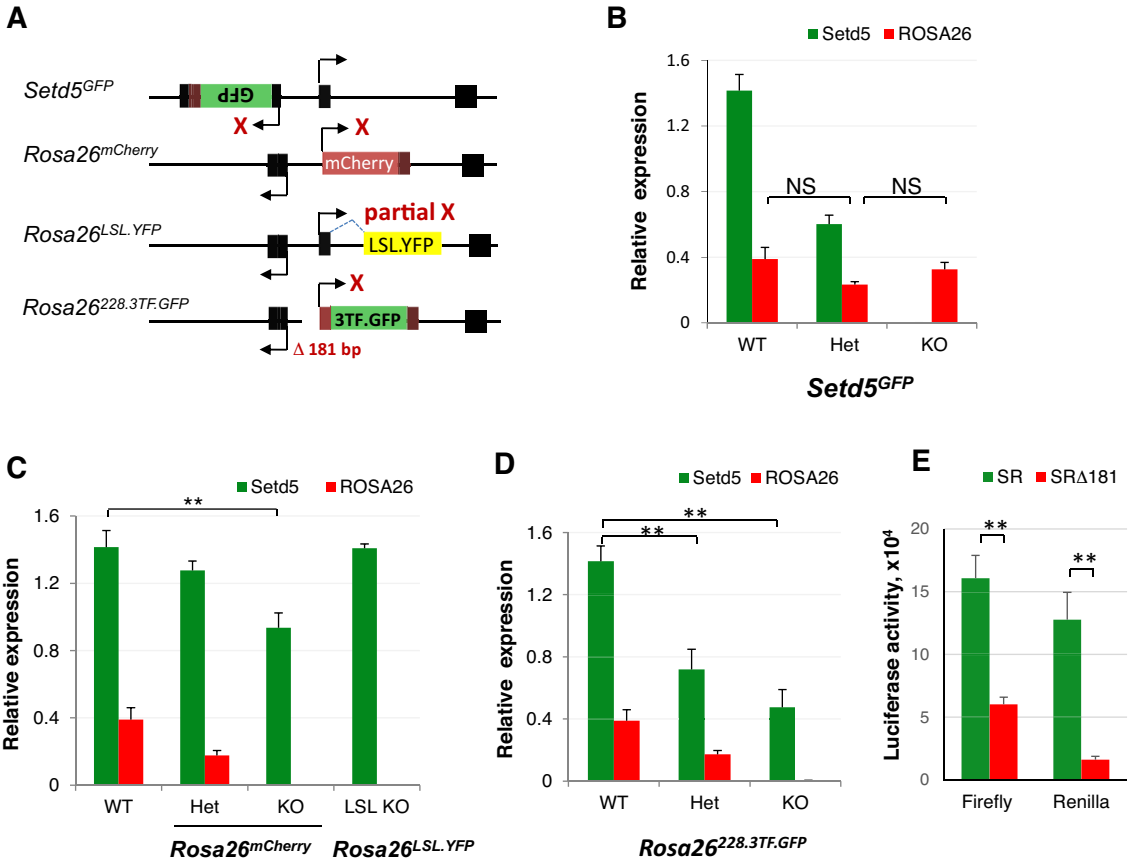

Fig. 2. Deletions in the *ROSA26* promoter or first exon regions decrease the levels of *Setd5* transcript. (A) Alleles used for comparative RT-qPCR analysis of *Setd5* and *ROSA26* mRNA expression levels. In the $Setd5^{GFP}$ allele, the *Setd5* mRNA is disrupted at the beginning of the first exon. In $ROSA26^{Cherry}$, the *ROSA26* mRNA is disrupted at the beginning of the first exon; in $ROSA26^{LSL.YFP}$, the first exon of *ROSA26* is transcribed and spliced but full-length *ROSA26* mRNA is not made; and in $ROSA26^{228.3TF.GFP}$, there is a 181 bp deletion of the promoter region preceding the *ROSA26* TSS and the *ROSA26* transcript is not made. (B) *Setd5* and *ROSA26* mRNA levels in wild-type (WT), $Setd5^{GFP}$ heterozygous (Het) and knockout (KO) embryos. (C) *Setd5* and *ROSA26* mRNA levels in WT, $ROSA26^{Cherry}$ heterozygous (Het) and knockout (KO) and $ROSA26^{LSL.YFP}$ knockout (LSL KO) embryos. (D) *Setd5* and *ROSA26* mRNA levels in WT, $ROSA26^{228.3TF.GFP}$ heterozygous (Het) and knockout (KO) embryos. (E) Dual luciferase assay of bidirectional activity of a truncated *Setd5*/*ROSA26* (SRΔ181) promoter that has a 181 bp deletion similar to the $ROSA26^{228.3TF.GFP}$ mouse allele, compared with the full-length promoter. Luciferase activity is diminished in both directions in SRΔ181. n=3; error bars indicate s.e.m. NS, not significant; $^{**}P \leq 0.01$.

blood cells (Fig. S7). Vascular abnormalities in placental development might also contribute to embryonic lethality. In support of this, Hematoxylin and Eosin (H&E) staining of E9.5 placentas showed that the labyrinthine layer in heterozygous animals was well developed (Fig. 4B), whereas in knockout placentas this layer was significantly thinner, contained only maternal blood vessels with mature erythrocytes, and embryonic blood vessels remained at the periphery in the chorioallantoic region and did not invade the labyrinthine layer.

To test whether the cardiovascular defects might be due to differences in protein expression, we performed immunofluorescent staining of E9.5 wild-type embryos for endogenous SETD5 protein. We found nearly uniform expression of SETD5, with protein localization being predominantly nuclear; in particular, SETD5 expression in $PECAM1^{+}$ and MHC^{+} cells was similar to its expression in other surrounding cells (Fig. S8A,B).

Growth impairment, defects in differentiation potential, changes in global gene expression and in chromatin modification marks in *Setd5* null mESCs

To further characterize the role of *Setd5* in cell growth, we derived wild-type, *Setd5* heterozygous and null mESC lines. Both the $Setd5^{GFP/+}$ and $Setd5^{GFP/GFP}$ lines exhibited strong green fluorescence (Fig. S9A), but the $Setd5^{GFP/GFP}$ cells exhibited slower growth rates than wild-type or heterozygous cells. Measurements of BrdU incorporation (Fig. 5A) revealed a slight decrease in cell proliferation rates in the knockout mESCs compared with the heterozygous cells (from 70% to 58%, $P \leq 0.05$), with no other differences being observed between wild-type and heterozygous cells. FACS analysis of annexin V-stained cells showed a greater than 2-fold increase in apoptosis in $Setd5^{GFP/GFP}$ compared with $Setd5^{GFP/+}$ cells ($P \leq 0.01$; Fig. 5B). These results suggest that both reduced proliferation and elevated apoptosis contribute to the reduced growth rate of *Setd5* null mESCs.

Next, we examined cell cycle progression in the two cell lines by FACS analysis after propidium iodide staining (Fig. 5C). Multiple defects in cell cycle progression were observed in the absence of *Setd5*, with a delay in the G2/M transition being the most pronounced. The knockout cells had shorter G1 and S phases (reduced from 29% to 26% and from 34% to 30% of cells, respectively; $P \leq 0.05$) and an increase in G2 phase (from 30% to 34% of cells; $P \leq 0.01$). Additionally, we observed an increase in the percentage of polyploid cells (3% for wild-type, 6% for $Setd5^{GFP/+}$ and 9% for $Setd5^{GFP/GFP}$ mESCs). This observation was further

Development (2016) 143, 4595-4607 doi:10.1242/dev.141465

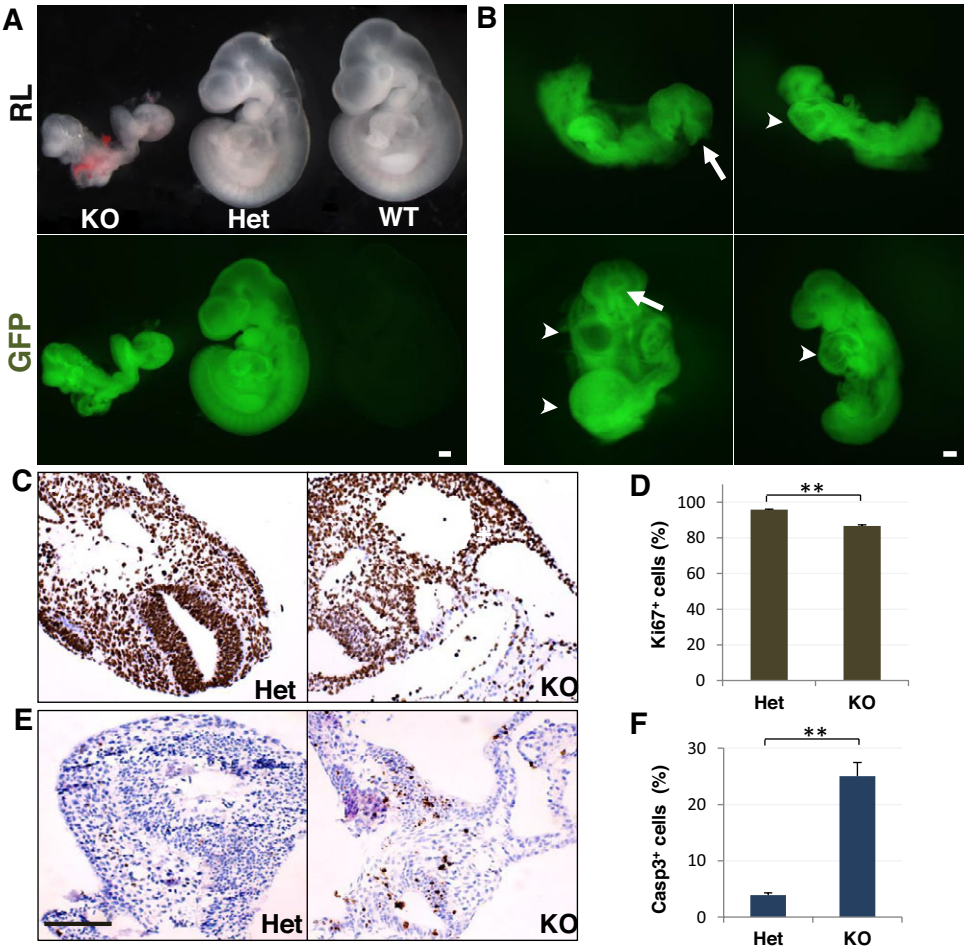

Fig. 3. Phenotype of *Setd5* knockout embryos. (A,B) Gross morphology of *Setd5*^{GFP} wild-type (A), heterozygous (A) and knockout (A,B) embryos at E9.5. Reflected light (RL) and fluorescent (GFP) images show that KO embryos (a range of severity of KO phenotypes at E9.5 is shown in B) are severely underdeveloped, have pericardial and tail effusions (arrowheads) and open neural tubes (arrows). (C,D) Immunohistochemical staining with anti-Ki67 antibodies (C) and its quantification (D) reveal a cell proliferation defect in KO versus Het embryos. (E,F) Immunohistochemical staining with anti-activated caspase 3 antibodies (E) and its quantification (F) demonstrate an increase in the number of apoptotic cells in KO embryos. n=5; error bars indicate s.e.m. $^{**}P \leq 0.01$. Scale bars: 100 µm.

supported by an increase in the percentage of cells displaying visible chromosomes (mitotic index) when nuclei were stained with DAPI (Fig. 5D,E). Although some mitotic cells showed a metaphase plate of chromosomes, many others had two sets of segregating chromosomes suggesting cellular arrest at anaphase stage and failure to complete cytokinesis. In addition, examination of the *Setd5* null mESCs revealed that many had enlarged and misshapen nuclei (Fig. 5D).

To determine whether the absence of *Setd5* affects the differentiation potential of mESCs, we performed directed differentiation towards cardiac progenitors through embryoid body (EB) formation in hanging drops following serum stimulation. Both *Setd5*$^{GFP/+}$ and *Setd5*$^{GFP/GFP}$ mESC lines formed EBs on day 2, grew proportionally in size to day 4, and attached to the plate and formed outgrowths by day 8 of the differentiation protocol (Fig. S9B). However, only a very small number of spontaneously contracting EB outgrowths were observed in knockout cells (3% of total EBs) compared with heterozygous cells (48%). Similarly, staining for the cardiomyocyte marker MHC confirmed that there were fewer MHC-positive cells in *Setd5*$^{GFP/GFP}$ than in *Setd5*$^{GFP/+}$ outgrowths, indicating that the ability of *Setd5* null mESCs to differentiate into cardiomyocytes was drastically reduced (Fig. 5F,G).

To analyze global changes in gene expression profiles in the absence of *Setd5* we performed RNA-seq on RNAs isolated from *Setd5*$^{GFP/+}$ and *Setd5*$^{GFP/GFP}$ mESC lines. Differential expression analysis revealed that of the 22,388 expressed genes in *Setd5* null mESCs, 1292 were upregulated and 1684 were downregulated compared with control cells (adjusted P-value ≤0.01; Table S3). Gene ontology analysis of the top upregulated genes showed an enrichment for genes involved in embryonic morphogenesis and pattern specification processes, including mesodermal transcription factors (*Mixl1*, *Mesp1*, *Gsc*, *T*), blood vessel morphogenic genes (e.g. *Flt1*, *Sox17*), Wnt pathway genes, growth factors and genes linked to apoptosis (Fig. 6A,B). A similar analysis of the top downregulated genes suggested they were involved in myosin complex and contractile fiber function (*Smtnl1*, *Myh13*) and myocyte differentiation (*Meox1*, *Six1*). Genes involved in stem cell maintenance (*Klf4*, *Nanog*, *Sox2*) were also downregulated. These findings indicate that *Setd5* null mESCs have a markedly dysregulated transcriptome and impaired stem cell identity. Interestingly, the expression of *Mesp1* and *Meox1*, two genes that might contribute to the cardiovascular phenotype, were also markedly altered. This finding was validated by RT-qPCR analysis using RNA from mESCs and embryos at E8.5 (Fig. 6C).

To analyze whether *Setd5* has an effect on global chromatin state or is solely required for major histone lysine methylation marks, we performed immunoblot analysis on protein lysates from *Setd5*$^{GFP/+}$ and *Setd5*$^{GFP/GFP}$ mESCs using antibodies that recognize specific modifications (Fig. 7A). H3K4me1, -me2, -me3 and H3K79me3 marks remained similar between the *Setd5*$^{GFP/+}$ and *Setd5*$^{GFP/GFP}$ samples, whereas pan-H3 acetylation, H3K36me1, -me2, -me3, H4K20me3, H3K9me1, -me2 and -me3 marks were slightly increased in cells lacking SETD5. Conversely, H3K27me2 and -me3 were decreased in the *Setd5* knockout mESCs (Fig. 7B). These results suggest that, in the absence of SETD5, there is an imbalance

Development (2016) 143, 4595-4607 doi:10.1242/dev.141465

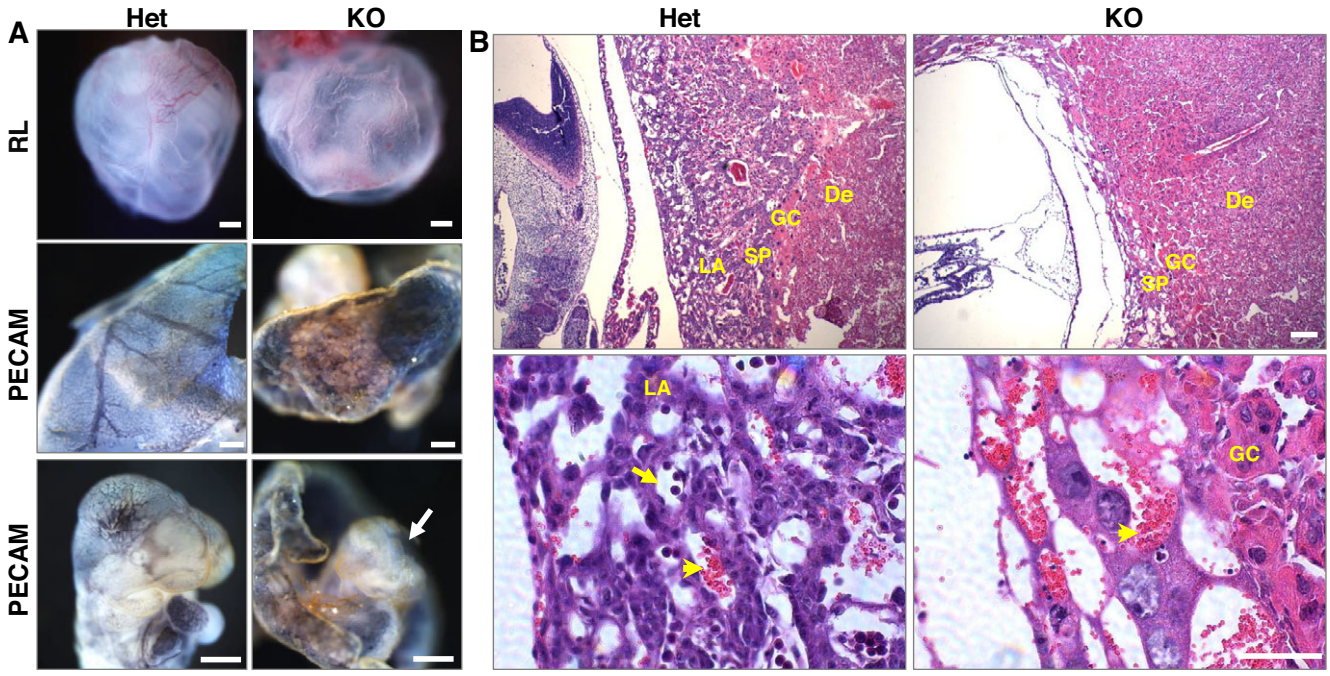

Fig. 4. Vascular abnormalities in *Setd5* knockout embryos. (A) Morphology of yolk sacs from *Setd5* heterozygous (Het) and knockout (KO) embryos at E9.5. Unstained yolk sacs (top row) are pale in appearance and show the absence of large vitelline blood vessels in KO samples, indicating defects in angiogenesis. Whole-mount staining for the endothelial marker PECAM1 revealed a less intricate blood vessel network in KO yolk sacs (middle row) and embryos (bottom row, arrow). (B) Placental defects in KO embryos. H&E-stained sections of E10.5 placentas show that the KO placenta is thinner, with an almost absent labyrinthine layer (top row). Magnified images (bottom row) show that in Het placentas fetal blood vessels with nucleated erythrocytes (arrow) have invaded the labyrinthine layer and are intermixed with maternal blood vessels filled with mature erythrocytes (arrowhead). By contrast, only maternal blood vessels are visible in the KO placenta. LA, labyrinthine layer; SP, spongiotrophoblast layer; GC, giant cells; De, desidua (maternal part). Scale bars: 500 µm in A; 50 µm in B.

in specific chromatin modifications known to play essential roles in both activating and repressing gene expression. However, we failed to observe the total loss of any individual methylation, suggesting that SETD5 may lack methyltransferase activity.

SETD5 co-immunoprecipitates with components of the PAF1 and NCoR complexes

Given the changes observed in chromatin state and prior studies demonstrating a role for SETD5 orthologs in chromatin regulatory complexes, we sought to determine the protein interaction partners of SETD5 using mass spectrometry. We immunoprecipitated FLAG-tagged SETD5 from transfected HEK 293T cells. MALDI-TOF mass spectrometry analysis detected 153 proteins that were pulled down by the FLAG-SETD5 protein but not from a control eluate (Table S4). Inspection of the most frequently detected proteins revealed that FLAG-SETD5 co-precipitated with four out of five known members of the co-transcriptional PAF1C: PAF1, LEO1, CTR9 and CDC73. Moreover, NCOR1, a member of an HDAC complex, also co-precipitated with FLAG-SETD5. Interestingly, the glycosylating enzyme O-linked N-acetylglucosamine (GlcNAc) transferase (OGT) was also co-precipitated, suggesting that SETD5 might be glycosylated, similar to its ortholog MLL5 (Fujiki et al., 2009). Other co-precipitated proteins included multiple RNAP II subunits and elongation factors, as well as many proteins known to be involved in chromatin remodeling, RNA processing, mitosis and spindle assembly, DNA replication and repair, mitotic checkpoint control and chromosome condensation and cohesion (Table S4).

To validate the key mass spectrometry results we performed anti-FLAG immunoprecipitations using FLAG-SETD5-expressing cell lysates followed by western blot analysis. The results confirmed that SETD5 co-immunoprecipitates with CTR9, PAF1, CDC73, OGT, RUVBL1 and NCOR1 (Fig. 7C). In addition to confirming interactions between SETD5 and NCOR1, we also showed that two other members of the NCoR complex, TBL1X and HDAC3, co-immunoprecipitate with SETD5. Finally, we performed reverse co-immunoprecipitations with antibodies against CTR9, PAF1 and NCOR1 and found, again, that these proteins co-precipitated with FLAG-SETD5 (Fig. 7D). Moreover, endogenous SETD5 protein partially colocalizes with HDAC3 protein in NIH3T3 cells (Fig. S8D). Together, these data indicate that SETD5 exists in close association with multiple members of the co-transcriptional PAF1C as well as with the HDAC-containing NCoR complex.

Setd5 is required for accurate co-transcriptional histone acetylation

Given that SETD5 orthologs have been shown to be involved in co-transcriptional chromatin deacetylation, a process that increases the productive transcription of active genes by suppressing off-target transcription (Kim et al., 2012; Rincon-Arano et al., 2012), we sought to determine whether SETD5 has a similar function. We analyzed the chromatin acetylation status of three constitutively expressed genes using anti-acetylated histone chromatin immunoprecipitation (ChIP) followed by PCR with primers targeting different regions within and downstream of the promoter region (Fig. 8). The genes selected for this analysis were eukaryotic translation elongation factor 1 alpha 1 (*Eef1a1*), hydroxyacyl glutathione hydrolase (*Hagh*) and inositol hexaphosphate kinase 1 (*Ip6k1*). *Eef1a1* is a housekeeping gene, whereas *Hagh* and *Ip6k1* are two metabolic genes that have antisense transcripts near or within the gene locus and whose yeast orthologs were analyzed in studies of Set3p (Kim et al., 2012).

Development (2016) 143, 4595-4607 doi:10.1242/dev.141465

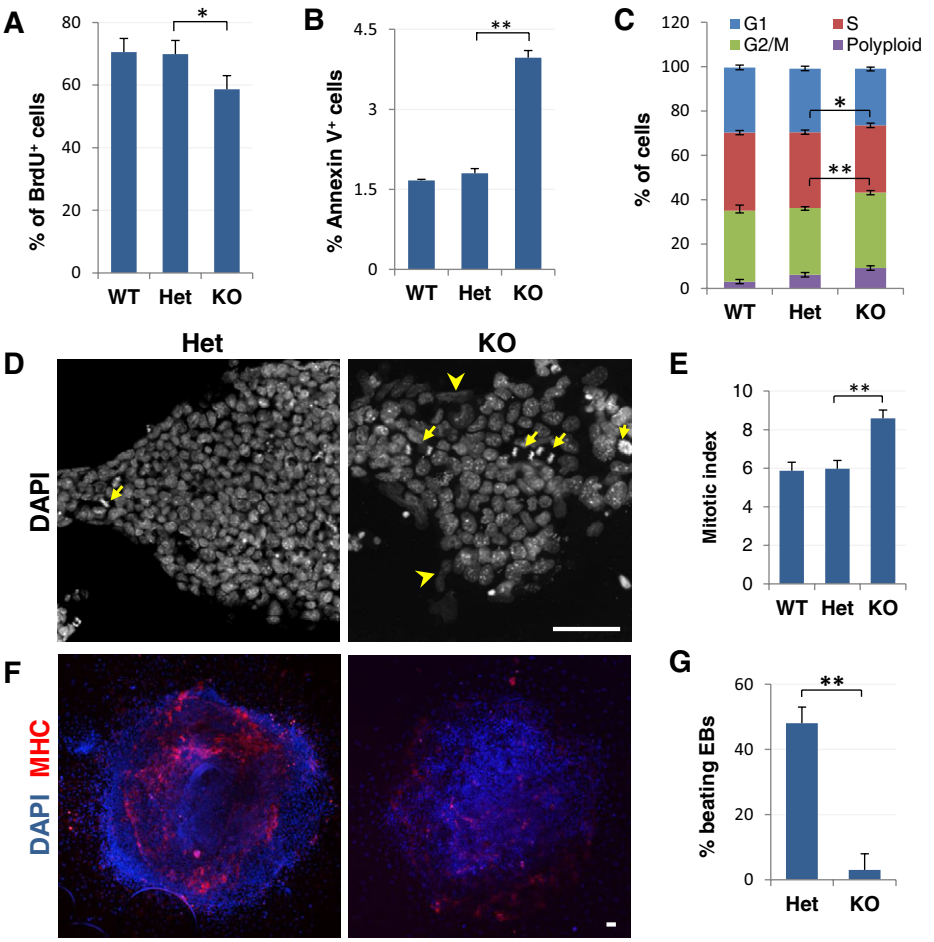

Fig. 5. Phenotypic analysis of *Setd5* knockout mESCs. (A) Proliferation of *Setd5* wild-type (WT), heterozygous (Het) and knockout (KO) mESCs was quantified by counting BrdU-positive cells. KO mESCs show a slight defect in proliferation. (B) The percentage of apoptotic cells was quantified using annexin V staining and FACS. KO cells show a significant increase in apoptosis. (C) The percentage of cells in each stage of the cell cycle was determined by propidium iodide staining and FACS. KO mESCs have fewer cells in G1 and S phase and increased numbers of cells in G2 phase. (D) DAPI staining of DNA shows an increased number of cells arrested in mitosis (with visible chromosomes, arrow) in the KO mESC line. Cells with enlarged, misshapen nuclei are also visible (arrowheads). (E) Mitotic index (the percentage of cells with visible chromosomes) is increased in the KO mESC line. (F) Differentiation of mESCs towards cardiomyocytes. Immunofluorescence staining shows a decreased number of cells positive for myosin heavy chain (MHC) in the KO 8 day embryoid body (EB) outgrowths. (G) Quantification of beating 8 day EB outgrowths, showing that a very small percentage of KO EBs were beating, indicating defects in differentiation potential. n=3 (C,G), n=4 (B) or n=5 (A,E). Error bars indicate s.e.m. *$P\leq0.05$,**$P\leq0.01$. Scale bars: 50 µm.

ChIP experiments were performed using chromatin isolated from *Setd5*$^{GFP/+}$ and *Setd5*$^{GFP/GFP}$ mESC lines, as well as *Setd5*$^{GFP/GFP}$ cells in which *Setd5* expression was rescued through overexpression of FLAG-SETD5 protein under the control of the human *EEF1A1* promoter. In the absence of *Setd5*, chromatin acetylation in all three genes was markedly increased in promoter proximal regions and remained higher in downstream regions than in the control or rescued mESCs (Fig. 8). For example, histone H3 acetylation at *Eef1a1* was higher upstream of the TSS (region I, $P\leq0.05$), at the TSS (region II, $P\leq0.01$) and downstream of exon 1 (region III, $P\leq0.01$) (Fig. 8A,B). The expression of *Eef1a1* mRNA was also increased ($P\leq0.05$; Fig. 8C). Similar differences were observed in both the *Hagh* and *Ip6k1* genes (Fig. 8D,E,G,H). Interestingly, expression of *Fahd1*, the gene divergently transcribed with *Hagh*, was higher in the *Setd5* null mESCs ($P\leq0.01$; Fig. 8F). Similarly, expression of the antisense *Gm38134* transcript, the promoter of which lies within the *Ip6k1* gene, was also increased (Fig. 8I).

Taken together, these data suggest that SETD5 contributes to the recruitment of HDAC to locations near TSSs so as to remove chromatin acetylation marks as RNAP II proceeds to the elongation stage and moves towards the downstream region of genes.

DISCUSSION

In this study we establish that the mouse *Setd5* gene plays a vital role in early development, cell replication and the co-transcriptional regulation of chromatin accessibility, possibly through interactions with protein components of the PAF1 and NCoR complexes.

Setd5 in development

Mouse embryos that lack *Setd5* exhibit severe developmental delay, vascular abnormalities in the embryo and placenta, reduced cellular proliferation and increased apoptosis, and do not survive past E10.5. Although the phenotypic changes might be due to stalled development in the mutants, our findings are consistent, nonetheless, with a widespread impairment in the regulation of gene expression, as might be expected to occur with altered chromatin dynamics during transcription. Similar phenotypes with mid-gestation lethality and vascular abnormalities have been described for knockouts of several other SET domain-containing proteins (Hu et al., 2010; Jones et al., 2008), as well as enzymes regulating histone acetylation such as CBP/p300 (Tanaka et al., 2000; Yao et al., 1998).

Through likely interactions with the PAF1 and NCoR complexes, SETD5 may be involved in modulating gene transcription and, consequently, the development of multiple cell lineages. Differential RNA-seq expression profiling showed that the expression of more than 10% of mapped genes, including lincRNAs, antisense transcripts and pseudogenes, was dependent on *Setd5*. In mESCs lacking *Setd5*, upregulated genes included those important for specifying cellular lineages and for embryonic morphogenesis, gastrulation, pattern specification and regionalization. At the same time, pluripotency genes were downregulated, suggesting that *Setd5* might also contribute to the maintenance of pluripotency, a finding that is consistent with the previously suggested roles for PAF1C in maintaining ESC identity (Ding et al., 2009) and for lineage specification at the blastocyst stage (Zhang et al., 2013).

Development (2016) 143, 4595-4607 doi:10.1242/dev.141465

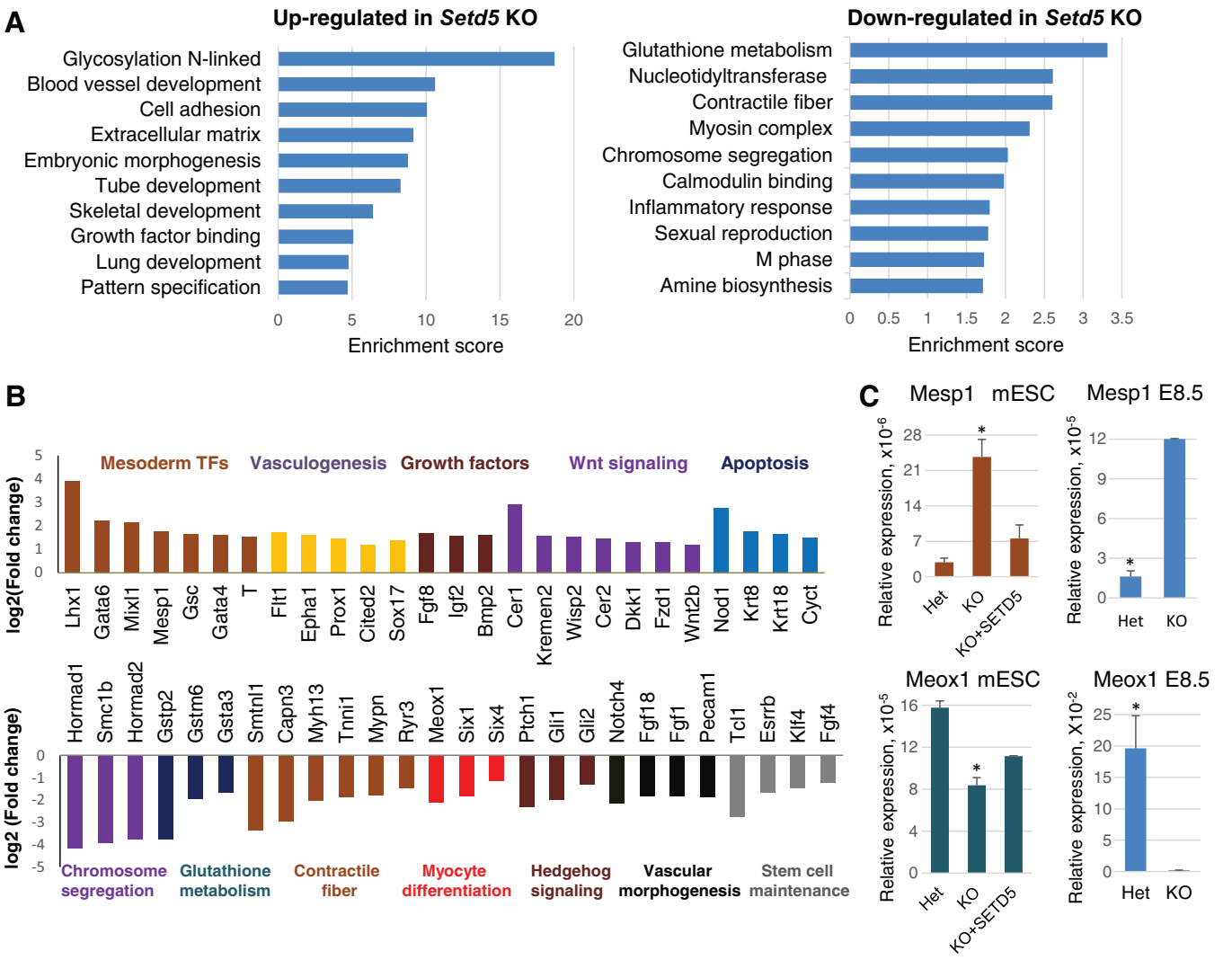

Fig. 6. RNA-seq analysis of *Setd5* knockout mESCs. (A) Gene ontology analysis of top upregulated or downregulated genes (log_2 fold change in KO versus Het ≥1 or ≤−1, respectively) in the *Setd5* KO. (B) Differential expression of select top upregulated and downregulated genes from different functional categories. (C) RT-qPCR analysis of *Mesp1* and *Meox1* expression in *Setd5* heterozygous (Het), knockout (KO) and rescued knockout (KO+SETD5) mESC lines and E8.5 embryos. *n*=3; error bars indicate s.e.m. **P*≤0.05.

Among the top differentially expressed genes in *Setd5*-deficient mESCs were those involved in myocyte and vascular lineage specification and function. For example, *Mixl1*, *Mesp1* and *Flt1* were markedly upregulated, whereas *Meox1* and *Six1* were downregulated, suggesting that their dysregulation might contribute to the cardiovascular phenotype. Similarly, different members of the zebrafish PAF1C have been shown to be important for cardiac and neural crest specification (Nguyen et al., 2010), somite segmentation (Akanuma et al., 2007) and heart tube cardiomyocyte specification (Langenbacher et al., 2011).

Setd5 and the cell cycle

Impaired cell proliferation and survival of *Setd5*-deficient embryonic tissues and mESCs could be due to several factors. First, transcriptome profiling revealed that in the absence of *Setd5* there is an increase in the expression of pro-apoptotic genes and a decrease in expression of genes involved in M-phase chromosome segregation. Second, dysregulation of histone modifications facilitated by PAF1C and NCoR could also play a role. HDAC3 is known to be involved in the control of cell cycle progression and, similar to the defects that we observed in *Setd5* null mESCs, depletion of HDAC3 decreases the number of cells in S phase and leads to an accumulation of cells in G2/M phase (Bhaskara et al., 2008; Wilson et al., 2006).

The increase in cell cycle arrest during mitosis observed in cells lacking *Setd5*, particularly during cytokinesis, is consistent with findings that the NCoR complex localizes near the mitotic spindle, where its HDAC activity enhances mitotic progression (Ishii et al., 2008; Li et al., 2006). Similarly, human PAF1 has been shown to localize to centromeres and promote mitotic spindle formation (Moniaux et al., 2009). Moreover, mass spectrometry analysis indicates that SETD5 is capable of interacting with a number of proteins involved in chromosome cohesion and mitotic spindle assembly at the kinetochore, suggesting the direct involvement of a SETD5-containing complex in these processes.

SETD5 and co-transcriptional regulation of histone acetylation

We also found that SETD5 co-precipitates with four members of the mammalian co-transcriptional PAF1C. RNAP II is known to

Development (2016) 143, 4595-4607 doi:10.1242/dev.141465

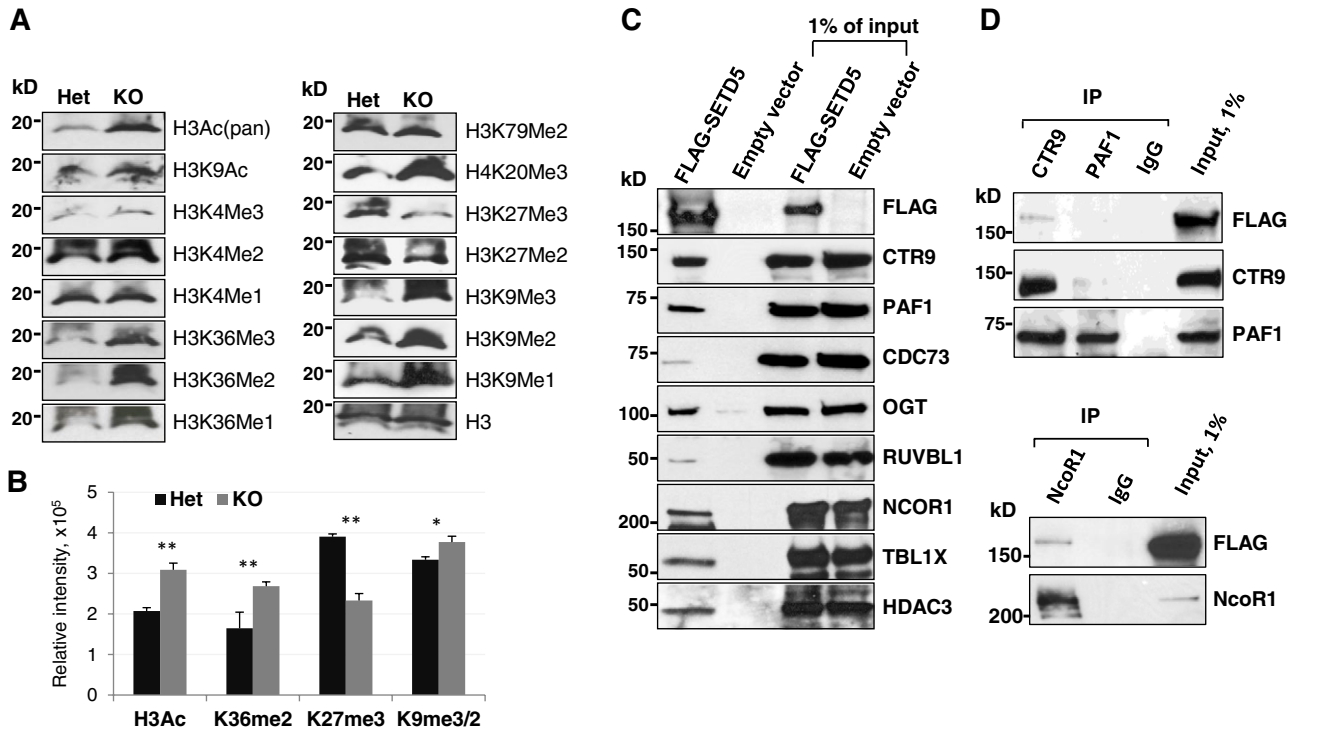

Fig. 7. SETD5 co-immunoprecipitates with members of the PAF1C and NCoR complexes. (A) Western blot of protein lysates from *Setd5* heterozygous (Het) and knockout (KO) mESC lines probed with antibodies against the indicated histone modifications, as well as histone H3 as a loading control. The absence of SETD5 did not result in loss of methylation at any of the tested residues. (B) Quantification of changes in global chromatin marks in *Setd5* heterozygous and knockout mESC lines. n=3; error bars indicate s.e.m. *$P\leq0.05$, **$P\leq0.01$. (C) Anti-FLAG immunoprecipitation of lysates from HEK 293T cells transfected with FLAG-SETD5 and empty FLAG vector. Bound proteins along with 1% inputs were analyzed by immunoblotting with the indicated antibodies. FLAG-SETD5 immunoprecipitates with members of PAF1C (CTR9, PAF1, CDC73), OGT, the chromatin remodeling protein RUVBL1, and members of NCoR protein complex (NCOR1, TBL1X, HDAC3). (D) Reverse immunoprecipitation with antibodies against PAF1, CTR9 and NCOR1 of HEK 293T cells transfected with FLAG-SETD5. FLAG-SETD5 co-precipitates with PAF1, CTR9 and NCOR1 but not with control IgG antibodies.

associate with PAF1C at or near promoters so as to maintain an association during gene transcription, and then to detach near the poly(A) site (Jaehning, 2010). Consistent with the involvement of PAF1C in multiple transcription-related processes, including facilitating the RNAP II transcription elongation machinery and the recruitment of mRNA 3′ end-processing factors, mass spectrometry analysis of SETD5-containing protein complexes identified RNAP II subunits, elongation factors, chromatin remodeling proteins and proteins involved in RNA 3′ end processing.

Our most intriguing finding among the proteomic data is, however, the interaction of SETD5 with the HDAC-containing complex NCoR. This suggests that SETD5 might be necessary for the recruitment of this chromatin-modifying complex to elongating RNAP II through interaction with PAF1C. Indeed, studies in yeast have shown that PAF1C is involved in the regulation of transcription-coupled histone modifications by promoting histone H2BK123 monoubiquitylation and recruitment of histone methyltransferase complexes that methylate histone H3K4, K36 and K79, a subset of modifications primarily associated with transcribed genes (Crisucci and Arndt, 2011; Tanny, 2014). H3K4 and K36 methylation, which modulates histone acetylation by facilitating the recruitment or activity of HAT and HDAC complexes, is also influenced by PAF1C. For instance, the yeast SETD5 ortholog Set3p is recruited near TSSs of genes through binding of its PHD domain to H3K4me2 residues. Moreover, the loss of Paf1, just as with the loss of Set3p, results in increased acetylation close to the TSS at the 5′ ends of genes (Kim and Buratowski, 2009; Chu et al., 2007). However, unlike Set3p, SETD5 appears to lack a methyl lysine-recognizing PHD domain, and might instead be recruited to transcribed genes through direct interaction with PAF1C. In turn, PAF1C may directly associate with RNAP II and/or use its different subunits to recognize native histones or post-translational histone modifications in order to facilitate further histone modification (Chu et al., 2013; Wu and Xu, 2012).

Interestingly, the lack of *Setd5* in mESCs did not lead to complete loss of any of the major histone methylation marks tested. This suggests that SETD5, like its orthologs, might not possess a histone methyltransferase activity. Consistent with SETD5 being recruited with PAF1C and NCoR to actively transcribed genes, *Setd5* null mESCs have slightly increased H3K36 methylation, a co-transcriptional histone modification that is enriched over the coding region of genes. Moreover, we did not observe any global changes in H3K4 methylation, which is typically associated with active promoter regions, suggesting that SETD5/HDAC3 repression might mostly target transcriptional elongation, as does the yeast SET3C complex (Weinberger et al., 2012). In mammals, several HATs and HDACs have been shown to be targeted to the transcribed regions of active genes by phosphorylated RNAP II (Wang et al., 2009).

The *Drosophila* ortholog of SETD5, UpSET, has also been found to be localized to transcriptionally active genes and to be specifically enriched at TSSs (Rincon-Arano et al., 2012). Both UpSET and Set3p have been shown to reset chromatin by removing acetylation marks on histones after the RNAP II machinery has passed into an elongation stage and to restrict active histone

Development (2016) 143, 4595-4607 doi:10.1242/dev.141465

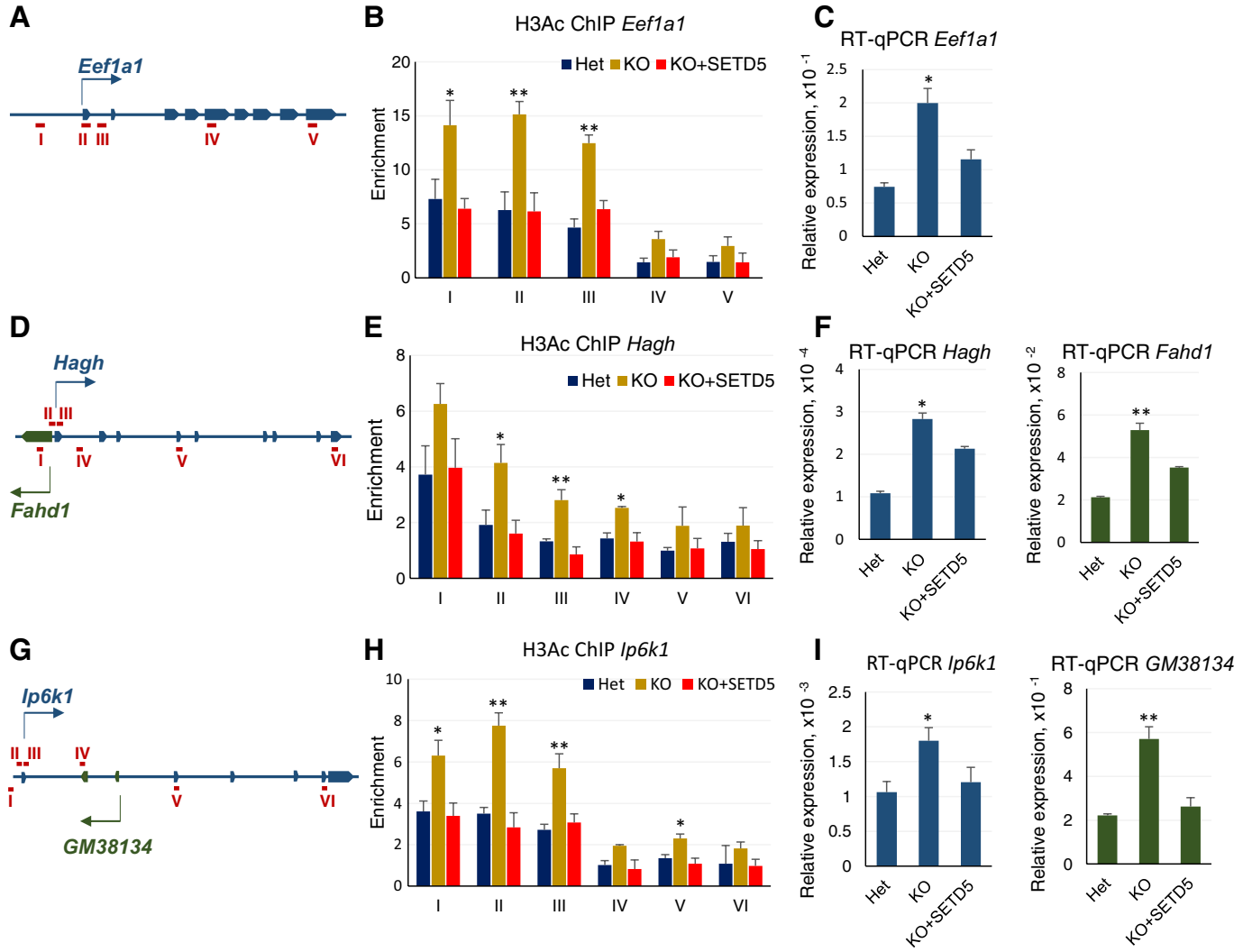

Fig. 8. SETD5 is involved in the regulation of co-transcriptional histone acetylation. ChIP with anti-pan-acetylated histone H3 (H3Ac) antibodies was used to compare levels of histone acetylation within promoter and downstream gene areas. (A) Schematic of the mouse *Eef1a1* gene locus showing the location of gene areas (red) probed by PCR after ChIP. Arrow indicates the direction of transcription. (B) H3Ac ChIP analysis of the *Eef1a1* gene locus in *Setd5* heterozygous (Het), knockout (KO) and rescued knockout (KO+SETD5) mESC lines. (C) RT-qPCR analysis of *Eef1a1* mRNA expression. (D) The *Hagh* gene locus. *Fahd1* is transcribed in the opposite direction. (E) H3Ac ChIP analysis of the *Hagh* gene locus. (F) RT-qPCR analysis of *Hagh* and *Fahd1* mRNA expression. (G) The *Ip6k1* gene locus. *Gm38134* is an antisense transcript located in an intronic region. (H) H3Ac ChIP analysis of *Ip6k1* gene locus. (I) RT-qPCR analysis of *Ip6k1* and *Gm38134* mRNA expression. In all three genes tested, histone acetylation is increased in *Setd5* KO cells beyond the promoter area, thereby increasing mRNA expression levels of neighboring antisense transcripts. n=4; error bars indicate s.e.m. *$P\leq0.05$,**$P\leq0.01$.

modifications to promoter regions (Kim et al., 2012; Rincon-Arano et al., 2012). These repressive actions of UpSET may reduce the probability of unmasking off-target genes or cryptic TSSs. Our data suggest that SETD5 might play a similar role in mammalian cells, since, in the absence of SETD5, there is an increase in chromatin acetylation that has spread from promoter regions downstream into the 5′ to mid regions of genes (Fig. S10).

Setd5 and *ROSA26* as a bidirectional transcript pair

Finally, we have shown that *Setd5* and *ROSA26* are co-expressed from a bidirectional promoter. Although the biological significance of divergent noncoding RNA transcripts is often elusive, it has been proposed that noncoding RNAs may regulate the transcription of the partner gene by helping to maintain a permissive chromatin environment (Kornienko et al., 2013). Such might be the case with the noncoding *ROSA26* transcript, which was discovered as a result of a retroviral-mediated promoter trap but whose functional significance has long been uncertain (Zambrowicz et al., 1997). Disruption of the *ROSA26* transcript by a splice acceptor trap in the intron of the gene, as occurs in *ROSA26*$^{LSL.YFP}$ and other such alleles, did not affect the expression of *Setd5*, suggesting that the noncoding *ROSA26* mRNA does not regulate *Setd5* expression. However, when the *ROSA26* locus is disrupted at the beginning of the first exon, or when adjacent sequences within the bidirectional promoter region are removed, *Setd5* expression is decreased together with *ROSA26* expression. These findings suggest that transcription in the *ROSA26* direction might aid in creating an active and accessible chromatin environment so as to assure strong and ubiquitous expression of the functionally essential *Setd5* gene. Such a notion is supported by a genome-wide modeling study in which divergent noncoding RNAs were predicted to be transcribed from promoters of essential genes in order to help maintain consistent

Development (2016) 143, 4595-4607 doi:10.1242/dev.141465

levels of expression (Wang et al., 2011) and by the finding that chromatin signatures of active transcription are over-represented in regions around bidirectional promoters (Yang and Elnitski, 2008).

Conclusions

Our findings indicate that SETD5 plays a vital role in mammalian embryonic development, cell cycle progression and the co-transcriptional regulation of histone acetylation. Similar to processes that have been described in both yeast and *Drosophila*, the co-transcriptional regulation of histone acetylation by SETD5 may be necessary for the chromatin state of a gene to be reset after the transcription machinery has passed, in order to suppress spurious transcription initiation.

MATERIALS AND METHODS

Mouse lines

The following mouse lines were utilized: *Gt(ROSA)26Sor*$^{tm1(EYFP)Cos}$ (*ROSA26*$^{LSL.YFP}$) (Srinivas et al., 2001), *Gt(ROSA)26Sor*^{tm2Mgn} (*ROSA26*^{mCherry}) (Chen et al., 2011), *ROSA26*$^{228.3TF}$ [unpublished, contains a '−228' deletion of the *ROSA26* promoter similar to that described by Serup et al. (2012)] and *Setd5*^{GFP}. All experimental procedures were reviewed and approved by the Vanderbilt University Institutional Animal Care and Use Committee.

Generation of *Setd5*^{GFP} mice

An exchange vector for performing recombinase-mediated cassette exchange (RMCE) was made beginning with pRosa26.Ex1 and inserting a GFP (Emerald)-SV40 poly(A) cassette (gift from David Piston, Vanderbilt University) between *Sma*I sites located in the first exon of *Setd5*. RMCE was performed as previously described in *ROSA26*^{LCA} mESCs (Chen et al., 2011) to generate *Setd5*$^{GFP/+}$ mESCs, from which mice were then derived. The resulting *Setd5*^{GFP} mouse line was backcrossed to, and maintained in, a heterozygous state on the CD-1 background (≥95%). Genotyping was performed by PCR using the primers detailed in Table S1.

Embryo isolation, RNA analysis and immunostaining

Timed matings were performed, where noon of the day of the vaginal plug was considered to be embryonic day (E) 0.5. Embryos at E8.5-E11.5 were dissected out for whole-mount or section H&E staining or immunostaining and imaging as detailed, including a full description of antibodies, in the supplementary Materials and Methods. RNA was isolated from E9.5 embryos for northern blot analysis with an exon 2-6 fragment of *Setdb5*, or for RT-qPCR using the primers listed in Table S1 with *Actb* as an internal control. For details, see the supplementary Materials and Methods.

BrdU analysis

Cell proliferation was assessed in E9.5 embryos or mESCs by BrdU staining as described in the supplementary Materials and Methods.

Derivation and analysis of mESCs

Setd5$^{GFP/+}$ and *Setd5*$^{GFP/GFP}$ mESCs were derived from blastocysts cultured in standard mESC media containing MEK1 inhibitor. Directed differentiation into cardiomyocytes was achieved by EB formation and subsequent culturing in high-serum media. Cell cycle progression and apoptosis were assessed in mESCs by FACS after propidium iodide or annexin V staining, respectively. For details, see the supplementary Materials and Methods.

RNA-seq

RNA-seq of heterozygous and knockout *Setd5* mESCs was performed on an Illumina HiSeq3000, with differential gene expression and Gene Ontology analysis of the top dysregulated genes as described in the supplementary Materials and Methods.

Dual luciferase assay

The dual luciferase assay and associated constructs used to assess promoter bidirectionality in transfected mESCs are described in the supplementary Materials and Methods.

Immunoblotting and immunoprecipitation

Western blots of protein lysates from embryos or mESC lysates to detect histone modifications and anti-FLAG immunoprecipitations from HEK 293T cells to determine FLAG-SETD5 co-immunoprecipitates are described in the supplementary Materials and Methods.

ChIP

Chromatin from 10^7 mESCs was subject to ChIP analysis as described in the supplementary Materials and Methods.

Mass spectrometry

FLAG peptide eluates from three immunoprecipitation experiments from FLAG-SETD5-transfected and mock transfected cells were combined and analyzed by multidimensional protein identification technology (MudPIT) as previously described (Martinez et al., 2012).

Statistical analysis

P-values were calculated using two-tailed Student's *t*-test for two samples with equal variance.

Acknowledgements

We thank Jean-Philippe Cartaillier for computational analysis of RNA-seq data; Judsen Schneider and Changqing Zhang for deriving the *ROSA26*$^{228.3TF.GFP}$ mice; Katherine Boyer and Anne McGough for technical assistance; Jennifer Stancill and Hannah Clayton for critically reading the manuscript; and Scott Hiebert and Roland Stein for helpful suggestions.

Competing interests

The authors declare no competing or financial interests.

Author contributions

A.B.O. and M.A.M. designed the study and co-wrote the manuscript; A.B.O. performed experiments and analyzed the data. R.G. undertook embryo dissections and western blot analysis; P.G.V. performed RT-qPCR analysis.

Funding

This work was supported by funding from the National Institutes of Health (NIH), National Institute of Diabetes and Digestive and Kidney Diseases to M.A.M. (DK72473 and DK89523). The Vanderbilt transgenic mouse/ESC shared resource was supported by the Division of Cancer Prevention, National Cancer Institute (CA68485) and the National Institute of Diabetes and Digestive and Kidney Diseases (DK20593). Deposited in PMC for release after 12 months.

Data availability

Raw RNA-seq data are available at ArrayExpress under accession number E-MTAB-5189. Mice bearing the *Setd5*^{GFP} allele were deposited into the Mutant Mouse Resource and Research Center (MMRRC) with stock number 036985-JAX.

Supplementary information

Supplementary information available online at
http://dev.biologists.org/lookup/doi/10.1242/dev.141465.supplemental

References

Akanuma, T., Koshida, S., Kawamura, A., Kishimoto, Y. and Takada, S. (2007). Paf1 complex homologues are required for Notch-regulated transcription during somite segmentation. *EMBO Rep.* **8**, 858-863.

Bhaskara, S., Chyla, B. J., Amann, J. M., Knutson, S. K., Cortez, D., Sun, Z.-W. and Hiebert, S. W. (2008). Deletion of histone deacetylase 3 reveals critical roles in S phase progression and DNA damage control. *Mol. Cell* **30**, 61-72.

Bhaskara, S., Knutson, S. K., Jiang, G., Chandrasekharan, M. B., Wilson, A. J., Zheng, S., Yenamandra, A., Locke, K., Yuan, J.-L., Bonine-Summers, A. R. et al. (2010). Hdac3 is essential for the maintenance of chromatin structure and genome stability. *Cancer Cell* **18**, 436-447.

Cantone, I. and Fisher, A. G. (2013). Epigenetic programming and reprogramming during development. *Nat. Struct. Mol. Biol.* **20**, 282-289.

Chen, S. X., Osipovich, A. B., Ustione, A., Potter, L. A., Hipkens, S., Gangula, R., Yuan, W., Piston, D. W. and Magnuson, M. A. (2011). Quantification of factors

influencing fluorescent protein expression using RMCE to generate an allelic series in the ROSA26 locus in mice. *Dis. Model. Mech.* **4**, 537-547.
Chu, Y., Simic, R., Warner, M. H., Arndt, K. M. and Prelich, G. (2007). Regulation of histone modification and cryptic transcription by the Bur1 and Paf1 complexes. *EMBO J.* **26**, 4646-4656.
Chu, X., Qin, X., Xu, H., Li, L., Wang, Z., Li, F., Xie, X., Zhou, H., Shen, Y. and Long, J. (2013). Structural insights into Paf1 complex assembly and histone binding. *Nucleic Acids Res.* **41**, 10619-10629.
Crisucci, E. M. and Arndt, K. M. (2011). The roles of the Paf1 complex and associated histone modifications in regulating gene expression. *Genet. Res. Int.* **2011**, 707641.
Del Rizzo, P. A. and Trievel, R. C. (2011). Substrate and product specificities of SET domain methyltransferases. *Epigenetics* **6**, 1059-1067.
Ding, L., Paszkowski-Rogacz, M., Nitzsche, A., Slabicki, M. M., Heninger, A.-K., de Vries, I., Kittler, R., Junqueira, M., Shevchenko, A., Schulz, H. et al. (2009). A genome-scale RNAi screen for Oct4 modulators defines a role of the Paf1 complex for embryonic stem cell identity. *Cell Stem Cell* **4**, 403-415.
Dodge, J. E., Kang, Y.-K., Beppu, H., Lei, H. and Li, E. (2004). Histone H3-K9 methyltransferase ESET is essential for early development. *Mol. Cell. Biol.* **24**, 2478-2486.
Fujiki, R., Chikanishi, T., Hashiba, W., Ito, H., Takada, I., Roeder, R. G., Kitagawa, H. and Kato, S. (2009). GlcNAcylation of a histone methyltransferase in retinoic-acid-induced granulopoiesis. *Nature* **459**, 455-459.
Glaser, S., Schaft, 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.
Guenther, M. G., Lane, W. S., Fischle, W., Verdin, E., Lazar, M. A. and Shiekhattar, R. (2000). A core SMRT corepressor complex containing HDAC3 and TBL1, a WD40-repeat protein linked to deafness. *Genes Dev.* **14**, 1048-1057.
Heuser, M., Yap, D. B., Leung, M., de Algara, T. R., Tafech, A., McKinney, S., Dixon, J., Thresher, R., Colledge, B., Carlton, M. et al. (2009). Loss of MLL5 results in pleiotropic hematopoietic defects, reduced neutrophil immune function, and extreme sensitivity to DNA demethylation. *Blood* **113**, 1432-1443.
Hu, M., Sun, X.-J., Zhang, Y.-L., Kuang, Y., Hu, C.-Q., Wu, W.-L., Shen, S.-H., Du, T.-T., Li, H., He, F. et al. (2010). Histone H3 lysine 36 methyltransferase Hypb/Setd2 is required for embryonic vascular remodeling. *Proc. Natl. Acad. Sci. USA* **107**, 2956-2961.
Ishii, S., Kurasawa, Y., Wong, J. and Yu-Lee, L.-Y. (2008). Histone deacetylase 3 localizes to the mitotic spindle and is required for kinetochore-microtubule attachment. *Proc. Natl. Acad. Sci. USA* **105**, 4179-4184.
Jaehning, J. A. (2010). The Paf1 complex: platform or player in RNA polymerase II transcription? *Biochim. Biophys. Acta* **1799**, 379-388.
Jepsen, K., Hermanson, O., Onami, T. M., Gleiberman, A. S., Lunyak, V., McEvilly, R. J., Kurokawa, R., Kumar, V., Liu, F., Seto, E. et al. (2000). Combinatorial roles of the nuclear receptor corepressor in transcription and development. *Cell* **102**, 753-763.
Jepsen, K., Solum, D., Zhou, T., McEvilly, R. J., Kim, H.-J., Glass, C. K., Hermanson, O. and Rosenfeld, M. G. (2007). SMRT-mediated repression of an H3K27 demethylase in progression from neural stem cell to neuron. *Nature* **450**, 415-419.
Jones, B., Su, H., Bhat, A., Lei, H., Bajko, J., Hevi, S., Baltus, G. A., Kadam, S., Zhai, H., Valdez, R. et al. (2008). The histone H3K79 methyltransferase Dot1L is essential for mammalian development and heterochromatin structure. *PLoS Genet.* **4**, e1000190.
Kim, T. S. and Buratowski, S. (2009). Dimethylation of H3K4 by Set1 recruits the Set3 histone deacetylase complex to 5′ transcribed regions. *Cell* **137**, 259-272.
Kim, T. S., Xu, Z., Clauder-Münster, S., Steinmetz, L. M. and Buratowski, S. (2012). Set3 HDAC mediates effects of overlapping noncoding transcription on gene induction kinetics. *Cell* **150**, 1158-1169.
Kittler, R., Pelletier, L., Heninger, A.-K., Slabicki, M., Theis, M., Miroslaw, L., Poser, I., Lawo, S., Grabner, H., Kozak, K. et al. (2007). Genome-scale RNAi profiling of cell division in human tissue culture cells. *Nat. Cell Biol.* **9**, 1401-1412.
Kornienko, A. E., Guenzl, P. M., Barlow, D. P. and Pauler, F. M. (2013). Gene regulation by the act of long non-coding RNA transcription. *BMC Biol.* **11**, 59.
Kouzarides, T. (2007). Chromatin modifications and their function. *Cell* **128**, 693-705.
Kurash, J. K., Lei, H., Shen, Q., Marston, W. L., Granda, B. W., Fan, H., Wall, D., Li, E. and Gaudet, F. (2008). Methylation of p53 by Set7/9 mediates p53 acetylation and activity in vivo. *Mol. Cell* **29**, 392-400.
Langenbacher, A. D., Nguyen, C. T., Cavanaugh, A. M., Huang, J., Lu, F. and Chen, J.-N. (2011). The PAF1 complex differentially regulates cardiomyocyte specification. *Dev. Biol.* **353**, 19-28.
Lee, S., Lee, D.-K., Dou, Y., Lee, J., Lee, B., Kwak, E., Kong, Y.-Y., Lee, S.-K., Roeder, R. G. and Lee, J. W. (2006). Coactivator as a target gene specificity determinant for histone H3 lysine 4 methyltransferases. *Proc. Natl. Acad. Sci. USA* **103**, 15392-15397.
Li, Y., Kao, G. D., Garcia, B. A., Shabanowitz, J., Hunt, D. F., Qin, J., Phelan, C. and Lazar, M. A. (2006). A novel histone deacetylase pathway regulates mitosis by modulating Aurora B kinase activity. *Genes Dev.* **20**, 2566-2579.
Lin, Q., Schwarz, J., Bucana, C. and Olson, E. N. (1997). Control of mouse cardiac morphogenesis and myogenesis by transcription factor MEF2C. *Science* **276**, 1404-1407.
Martinez, M. N., Emfinger, C. H., Overton, M., Hill, S., Ramaswamy, T. S., Cappel, D. A., Wu, K., Fazio, S., McDonald, W. H., Hachey, D. L. et al. (2012). Obesity and altered glucose metabolism impact HDL composition in CETP transgenic mice: a role for ovarian hormones. *J. Lipid Res.* **53**, 379-389.
Moniaux, N., Nemos, C., Deb, S., Zhu, B., Dornreiter, I., Hollingsworth, M. A. and Batra, S. K. (2009). The human RNA polymerase II-associated factor 1 (hPaf1): a new regulator of cell-cycle progression. *PLoS ONE* **4**, e7077.
Mottis, A., Mouchiroud, L. and Auwerx, J. (2013). Emerging roles of the corepressors NCoR1 and SMRT in homeostasis. *Genes Dev.* **27**, 819-835.
Musselman, C. A. and Kutateladze, T. G. (2011). Handpicking epigenetic marks with PHD fingers. *Nucleic Acids Res.* **39**, 9061-9071.
Nguyen, C. T., Langenbacher, A., Hsieh, M. and Chen, J.-N. (2010). The PAF1 complex component Leo1 is essential for cardiac and neural crest development in zebrafish. *Dev. Biol.* **341**, 167-175.
Oda, H., Okamoto, I., Murphy, N., Chu, J., Price, S. M., Shen, M. M., Torres-Padilla, M. E., Heard, E. and Reinberg, D. (2009). Monomethylation of histone H4-lysine 20 is involved in chromosome structure and stability and is essential for mouse development. *Mol. Cell. Biol.* **29**, 2278-2295.
Pijnappel, W. W. M. P., Schaft, D., Roguev, A., Shevchenko, A., Tekotte, H., Wilm, M., Rigaut, G., Seraphin, B., Aasland, R. and Stewart, A. F. (2001). The S. cerevisiae SET3 complex includes two histone deacetylases, Hos2 and Hst1, and is a meiotic-specific repressor of the sporulation gene program. *Genes Dev.* **15**, 2991-3004.
Rayasam, G. V., Wendling, O., Angrand, P. O., Mark, M., Niederreither, K., Song, L., Lerouge, T., Hager, G. L., Chambon, P. and Losson, R. (2003). NSD1 is essential for early post-implantation development and has a catalytically active SET domain. *EMBO J.* **22**, 3153-3163.
Rentas, S., Saberianfar, R., Grewal, C., Kanippayoor, R., Mishra, M., McCollum, D. and Karagiannis, J. (2012). The SET domain protein, Set3p, promotes the reliable execution of cytokinesis in Schizosaccharomyces pombe. *PLoS ONE* **7**, e31224.
Rincon-Arano, H., Halow, J., Delrow, J. J., Parkhurst, S. M. and Groudine, M. (2012). UpSET recruits HDAC complexes and restricts chromatin accessibility and acetylation at promoter regions. *Cell* **151**, 1214-1228.
Schotta, G., Sengupta, R., Kubicek, S., Malin, S., Kauer, M., Callen, E., Celeste, A., Pagani, M., Opravil, S., De La Rosa-Velazquez, I. A. et al. (2008). A chromatin-wide transition to H4K20 monomethylation impairs genome integrity and programmed DNA rearrangements in the mouse. *Genes Dev.* **22**, 2048-2061.
Sebastian, S., Sreenivas, P., Sambasivan, R., Cheedipudi, S., Kandalla, P., Pavlath, G. K. and Dhawan, J. (2009). MLL5, a trithorax homolog, indirectly regulates H3K4 methylation, represses cyclin A2 expression, and promotes myogenic differentiation. *Proc. Natl. Acad. Sci. USA* **106**, 4719-4724.
Serup, P., Gustavsen, C., Klein, T., Potter, L. A., Lin, R., Mullapudi, N., Wandzioch, E., Hines, A., Davis, A., Bruun, C. et al. (2012). Partial promoter substitutions generating transcriptional sentinels of diverse signaling pathways in embryonic stem cells and mice. *Dis. Model. Mech.* **5**, 956-966.
Smolle, M. and Workman, J. L. (2013). Transcription-associated histone modifications and cryptic transcription. *Biochim. Biophys. Acta* **1829**, 84-97.
Soriano, P. (1999). Generalized lacZ expression with the ROSA26 Cre reporter strain. *Nat. Genet.* **21**, 70-71.
Srinivas, S., Watanabe, T., Lin, C.-S., William, C. M., Tanabe, Y., Jessell, T. M. and Costantini, F. (2001). Cre reporter strains produced by targeted insertion of EYFP and ECFP into the ROSA26 locus. *BMC Dev. Biol.* **1**, 4.
Tachibana, M., Sugimoto, K., Nozaki, M., Ueda, J., Ohta, T., Ohki, M., Fukuda, M., Takeda, N., Niida, H., Kato, H. et al. (2002). G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. *Genes Dev.* **16**, 1779-1791.
Tanaka, Y., Naruse, I., Hongo, T., Xu, M.-J., Nakahata, T., Maekawa, T. and Ishii, S. (2000). Extensive brain hemorrhage and embryonic lethality in a mouse null mutant of CREB-binding protein. *Mech. Dev.* **95**, 133-145.
Tanny, J. C. (2014). Chromatin modification by the RNA polymerase II elongation complex. *Transcription* **5**, e988093.
Venkatesh, S. and Workman, J. L. (2015). Histone exchange, chromatin structure and the regulation of transcription. *Nat. Rev. Mol. Cell Biol.* **16**, 178-189.
Verdone, L., Caserta, M. and Di Mauro, E. (2005). Role of histone acetylation in the control of gene expression. *Biochem. Cell Biol.* **83**, 344-353.
Wang, Z., Zang, C., Cui, K., Schones, D. E., Barski, A., Peng, W. and Zhao, K. (2009). Genome-wide mapping of HATs and HDACs reveals distinct functions in active and inactive genes. *Cell* **138**, 1019-1031.
Wang, G.-Z., Lercher, M. J. and Hurst, L. D. (2011). Transcriptional coupling of neighboring genes and gene expression noise: evidence that gene orientation and noncoding transcripts are modulators of noise. *Genome Biol. Evol.* **3**, 320-331.
Wei, W., Pelechano, V., Jarvelin, A. I. and Steinmetz, L. M. (2011). Functional consequences of bidirectional promoters. *Trends Genet.* **27**, 267-276.

Weinberger, L., Voichek, Y., Tirosh, I., Hornung, G., Amit, I. and Barkai, N. (2012). Expression noise and acetylation profiles distinguish HDAC functions. *Mol. Cell* **47**, 193-202.

Wen, Y.-D., Perissi, V., Staszewski, L. M., Yang, W.-M., Krones, A., Glass, C. K., Rosenfeld, M. G. and Seto, E. (2000). The histone deacetylase-3 complex contains nuclear receptor corepressors. *Proc. Natl. Acad. Sci. USA* **97**, 7202-7207.

Wilson, A. J., Byun, D.-S., Popova, N., Murray, L. B., L'Italien, K., Sowa, Y., Arango, D., Velcich, A., Augenlicht, L. H. and Mariadason, J. M. (2006). Histone deacetylase 3 (HDAC3) and other class I HDACs regulate colon cell maturation and p21 expression and are deregulated in human colon cancer. *J. Biol. Chem.* **281**, 13548-13558.

Wu, J. and Xu, W. (2012). Histone H3R17me2a mark recruits human RNA polymerase-associated factor 1 complex to activate transcription. *Proc. Natl. Acad. Sci. USA* **109**, 5675-5680.

Xiao, B., Wilson, J. R. and Gamblin, S. J. (2003). SET domains and histone methylation. *Curr. Opin. Struct. Biol.* **13**, 699-705.

Yang, M. Q. and Elnitski, L. L. (2008). Diversity of core promoter elements comprising human bidirectional promoters. *BMC Genomics* **9** Suppl. 2, S3.

Yao, T.-P., Oh, S. P., Fuchs, M., Zhou, N.-D., Ch'ng, L.-E., Newsome, D., Bronson, R. T., Li, E., Livingston, D. M. and Eckner, R. (1998). Gene dosage-dependent embryonic development and proliferation defects in mice lacking the transcriptional integrator p300. *Cell* **93**, 361-372.

Yap, D. B., Walker, D. C., Prentice, L. M., McKinney, S., Turashvili, G., Mooslehner-Allen, K., de Algara, T. R., Fee, J., de Tassigny, X. A., Colledge, W. H. et al. (2011). Mll5 is required for normal spermatogenesis. *PLoS ONE* **6**, e27127.

Yoon, H.-G., Chan, D. W., Huang, Z. Q., Li, J., Fondell, J. D., Qin, J. and Wong, J. (2003). Purification and functional characterization of the human N-CoR complex: the roles of HDAC3, TBL1 and TBLR1. *EMBO J.* **22**, 1336-1346.

Yu, B. D., Hess, J. L., Horning, S. E., Brown, G. A. J. and Korsmeyer, S. J. (1995). Altered Hox expression and segmental identity in Mll-mutant mice. *Nature* **378**, 505-508.

Zambrowicz, B. P., Imamoto, A., Fiering, S., Herzenberg, L. A., Kerr, W. G. and Soriano, P. (1997). Disruption of overlapping transcripts in the ROSA beta geo 26 gene trap strain leads to widespread expression of beta-galactosidase in mouse embryos and hematopoietic cells. *Proc. Natl. Acad. Sci. USA* **94**, 3789-3794.

Zhang, K., Haversat, J. M. and Mager, J. (2013). CTR9/PAF1c regulates molecular lineage identity, histone H3K36 trimethylation and genomic imprinting during preimplantation development. *Dev. Biol.* **383**, 15-27.

Supplementary Information.

Materials and Methods.

Antibodies. Antibodies were used at working dilution of 1:1000 unless otherwise noted. The following primary antibodies were purchased from Cell Signaling: rabbit polyclonal anti-Mono-Methyl-Histone H3 (Lys4)(#5326) , anti-Di-Methyl-Histone H3 (Lys4)(#9725), anti-Tri-Methyl-Histone H3 (Lys4)(#9751), anti-Mono-Methyl-Histone H3 (Lys36)(#5928), anti-Di-Methyl-Histone H3 (Lys36)(#2901), anti-Tri-Methyl-Histone H3 (Lys36)(1:500)(#4909), anti-Tri-Methyl-Histone H4 (Lys20)(#5737), anti-Di-Methyl-Histone H3 (Lys79)(#9757), anti-Tri-Methyl-Histone H3 (Lys27)(#9733), anti-Di-Methyl-Histone H3 (Lys27)(#9728), anti-Di-Methyl- Histone H3 (Lys9)(#4658), anti-Mono-Methyl- Histone H3 (Lys9)(#7538), anti- Histone H3(#4499), anti-MEF2C (#5030), anti-NCOR1(#5948), anti-HDAC3(#2632), anti-PAF1(#12883), anti-CTR9(#12619), anti-CDC73(#8126), anti-OGT(#5368) and an anti-RUVBL1(#12300). We also used rabbit polyclonal anti-Histone H3 (acetyl K9+K14+K18+K23+K27) (ab47915, Abcam), anti-TBL1X(13540-1-AP, Proteintech group) and an anti-Ki67 (RM-9106-S1, ThermoFisher (1:500 dilution)) antibodies. In addition, we used a rat monoclonal anti-BrdU (MCA2060T, from AbD Serotech), mouse monoclonal anti-activated CASP3 (C8487, Sigma-Aldrich (1:500 dilution)), rabbit polyclonal anti-SETD5 (#2256.00.02, Strategic Diagnostics (at 1:500 dilution, Western blot), rabbit polyclonal anti-SETD5 (ab 204363, Abcam, 1:200 immunofluorescence), rat monoclonal anti-PECAM 1(#550274), anti-CD41 (#550539, both from BD Biosciences), mouse monoclonal anti-myosin heavy chain (clone MF20) (MAB4470,R&D Systems) and mouse monoclonal anti-FLAG HRP-conjugated (A8592, Sigma-Aldrich (1:500 dilution)) antibodies. The secondary antibodies used were a mouse anti-rabbit light chain specific HRP conjugated (#211-032-171), donkey anti-rabbit HRP conjugated (711-035-152) donkey anti-mouse HRP conjugated (715-035-151) and donkey anti-rat HRP conjugated (712-035-150, all from Jackson ImmunoResearch); donkey-anti-rabbit Alexa-555 conjugated (A-31572, ThermoFisher). ChIP was performed using rabbit anti-Histone H3 (acetyl K9+K14+K18+K23+K27)(ab47915) and normal rabbit IgGs (ab27472, both from Abcam).

Embryo isolation, whole mount staining and imaging. Timed matings were performed where noon of the day of identification of a vaginal plug was considered to be embryonic (E) day 0.5. Embryos at E8.5 to 11.5 were dissected out and imaged using a Leica MZ 16 FA fluorescent stereoscope equipped with a QImaging RETIGA 4000R camera and genotyped using embryonic or yoke sac tissue. For whole mount staining embryos were fixed in 4% paraformaldehyde for 2 hours, permeabilized in 1% Triton X-100/ PBS for 2 hours, blocked in staining buffer (3% BSA, 1%Triton X-100 in PBS) with 0.3% H2O2 at 4°C overnight, stained with primary antibodies in staining buffer overnight, washed with 1%Triton X-100 in PBS and stained with HRP-conjugated secondary antibodies overnight. Detection of HRP activity was done using DAB Peroxidase Substrate Kit (Vector Labs).

Northern blot and RT-qPCR analysis. RNA was isolated from embryos at E9.5 using TRIzol Reagent (Thermo Fisher). Northern blot analysis was performed using standard procedures. 5 μg aliquots of total RNAs were separated by 1% agarose formaldehyde gel electrophoresis in MOPS buffer (Thermo Fisher), blotted to a nylon membrane, and probed with a ^{32}P-labeled probe (652 bp EcoRI fragment containing exons 2-6 of *Setd5*). For RT-qPCR, cDNA was prepared using a High Capacity cDNA Reverse Transcription Kit (Life Technologies). 2 ng of cDNA were used in real-time qPCR with Power SYBR Green PCR master mix (Thermo Fisher) using ABI 7900HT Real Time PCR system (Applied Biosystems). Primer sequences are listed in Table S1. Relative expression was determined from at least three independent assays by $2^{-\Delta Ct}$ method (Winer et al., 1999). Mouse *Actb* gene was used as an endogenous control.

Immunohistochemistry and immunofluorescence. Embryos were dissected at E9.5, fixed in 4% paraformaldehyde, embedded in paraffin, cut into 5 μm serial sections and select sections were stained with haematoxylin and eosin stain (Sigma). To assess cell proliferation in embryos, pregnant females were injected IP with 100 μg of 5-bromodeoxyuridine (BrdU, Sigma) per gram of body weight and sacrificed after 3 hours. For BrdU staining, rehydrated slides were treated with 2 M HCl/0.3% Triton X-100 for 30 min then neutralized with 0.1 M sodium borate for 10 min. Otherwise, after heat antigen retrieval in sodium citrate buffer (pH 6), slides were washed with PBS,

blocked with staining buffer (3% BSA in PBST (0.02% Tween in PBS)) for 1 hour at room temperature, and then incubated with primary antibodies in the staining buffer at 4°C overnight. After washing with PBST the HRP-conjugated secondary antibodies were applied in the staining buffer for 2 hours at room temperature. Following final washes with PBST, slides were processed with DAB Peroxidase Substrate Kit**,** counterstained with haematoxylin, dehydrated and mounted in Cytoseal mounting media (Thermo Fisher). Mouse ES cells were grown in 8-well chamber slides (Thermo Fisher) and, for proliferation assay, incubated with 20 μM of BrdU for 30 min, then washed with PBS and fixed with 4% paraformaldehyde for 15 min. Immunofluorescence staining was performed as described above using fluorochrome-conjugated secondary antibodies. Slides were mounted with ProLong Gold antifade with DAPI reagent (Thermo Fisher). Images were acquired using an Axioplan2 fluorescent microscope (Zeiss) with 20X and 40X objectives and QImaging RETIGA EXi camera. Confocal images were taken with Zeiss LSM 510 META confocal laser scanning microscope using Zeiss imaging software. For display, images were converted into JPG format and processed using Adobe Photoshop software. Cell counts were done using ImageJ software. For tissues, step sections (50 μm apart) from at least 3 different animals of the same age and genotype were counted. At least 1000 cells were counted for each genotype in each experiment.

Dual luciferase assay. A 439 bp DNA fragment containing the Setd5-Rosa26 promoter region (Fig. S1) was PCR-amplified and cloned into pLucRluc (Polson et al., 2011) in both orientations thereby generating pLucRluc-SR and pLucRLuc-RS, respectively. Similarly, pLucRluc-SRΔ181, was made to mimic the 181 bp deletion that occurs in the ROSA26228.3TF mice. These plasmids were transfected into mouse ES cells together with pCMV-β-Gal (Thermo Fisher,) at a 2:1 ratio using Xfect Mouse Embryonic Stem Cell Transfection Reagent (Clontech). Cells were lysed, and luciferase activity was measured by using the Dual-Luciferase reporter Assay System (Promega) and a BioTek Microplate Reader. Luciferase activity was normalized to β-galactosidase activity detected in parallel using β-Galactosidase Assay System (Promega).

ES cell line derivation. Wild type (WT), heterozygous (Het) and knockout (KO) *Setd5* mES cell lines were derived from inner cells mass outgrowths obtained from blastocysts cultured in mouse ES cell media (Dulbecco's modified Eagle's medium supplemented with 15% FBS, 1X nonessential amino acids, 1X penicillin–streptomycin, 1× L-glutamine (all from Thermo Fisher)) containing MEK1 inhibitor PD98059 (Cell Signaling) following published protocol (Meissner et al., 2009). All mES cells were cultured on a feeder layer of mitotically inactivated mouse embryonic fibroblasts (MEFs) in mES cell culture medium supplemented with 1000 U/ml of leukemia inhibitory factor (ESGRO/LIF, Millipore) according to standard protocols. To obtain KO(+SETD5) ES cell line, KO cells were transfected with pEf1a-FLAG-SETD5 vector using Xfect Mouse Embryonic Stem Cell Transfection Reagent (Clontech).

RNA-seq analysis. RNA isolates (N=3) from heterozygous and knockout genotype of *Setd5* ES cells were subjected to RNA quality analysis and sequenced. Paired-end sequencing (75 bp length) was performed on an Illumina HiSeq3000 genome analyzer. Read alignment to the mouse genome mm10 (GRCm38) was done by Spliced Transcripts Alignment to a Reference (STAR) (Dobin et al., 2013) software using GENCODE comprehensive gene annotations (Release M8) as a reference. Differential gene expression analysis was done using DESeq2 (Love et al., 2014). Gene Ontology analysis was performed using DAVID Bioinformatics Resources v.6.7 (Huang da et al., 2009).

Directed differentiation of mouse ES cells. Cardiac differentiation in embryoid bodies (EBs) was induced in ES cell culture medium with 20% FBS without LIF at a final concentration of 5 $\times 10^4$ cells/ml following ES cell dissociation and suspension. Hanging drops (20 µl) were plated on the inside lids of low attachment dishes. After 48 hours, EBs were transferred in 10 ml of medium to low attachment dishes. At day 4, the EBs were plated on tissue culture dishes, allowed to adhere and beating was scored at day 8. A portion of the EBs were plated on 8-well chamber slides coated with fibronectin (Sigma) at day 7, and fixed and processed for immunostaining at day 8.

Cell cycle and apoptosis assays. For cell cycle distribution analysis, mESCs were trypsinized, washed with PBS, resuspended in 300 µl of PBS and fixed by addition of 100% ice cold ethanol and incubating at 4C overnight. Cells were washed the next

day with PBS, incubated for 30 min in 1 ml of PBS containing 10 μg/ml of ribonuclease A (Sigma) and 20 μg/ml of propidium iodide. The cell suspension was then analyzed for DNA content on a LSRII/Fortessa flow cytometer (BD Biosciences) using FACS DIVA (BD Biosciences) software. For quantitative analysis of apoptosis, mESCs were trypsinized, washed with PBS and incubated with 5 μl of Alexa 555-conjugated annexin-V (Thermo Fisher) in 100 μl of annexin binding buffer (10 mM HEPES/NaOH, pH 7.4; 140 mM NaCl; 2.5 mM $CaCl_2$) for 30 min at 37°C in the dark. The uptake of DAPI (Sigma), as measured by flow cytometry, was used to assess cell viability. The percentage of apoptotic cells was calculated on the basis of the proportion of Annexin V positive/DAPI negative cells.

Immunoblotting. Embryos or mES cells were lysed in RIPA buffer (10 mM TrisHCl, pH 7.5; 140 mM NaCl; 1 mM EDTA; 1% Nonidet P-40; 0.1% sodium deoxycholate; 0.1% SDS) containing protease inhibitors (300 μg/ml phenylmethylsulfonyl fluoride [PMSF], 1X protease inhibitor cocktail (Sigma)). Samples (5 μg of total protein) in 1X Laemmli sample buffer were resolved on a gradient 4-20% SDS-PAGE and transferred to polyvinylidene difluoride (PVDF) membranes (Millipore). Membranes were blocked with TBST (50 mM TrisHCl; 150 mM NaCl; 0.02 % Tween 20, pH 7.5) containing 5% non-fat dry milk and incubated with primary antibodies overnight. After washing with TBST, the membranes were then incubated with a secondary HRP-conjugated antibody, and the signals were detected with Western Lightning Plus ECL chemiluminescence kit (Perkin Elmer). Each Western blot experiment was done at least 3 times.

Immunoprecipitation. A FLAG-SETD5 expression plasmid was made by cloning *Setd5* coding sequences (from clone 30536831, ThermoFisher) into pCMV-FLAG-6a (Sigma). HEK293T cells were grown in DMEM supplemented with 10% FBS and antibiotics and transfected using PolyFect Reagent (Qiagene). After 48 hours, cells were lysed in RIPA buffer containing protease inhibitors and the lysates were cleared by centrifugation. The soluble fraction of lysates (2 mg of total protein) was used for immunoprecipitations with 40 μl of anti-FLAG agarose beads (Sigma) at 4°C overnight. FLAG-bound beads were washed four times in RIPA buffer and proteins were eluted with 200 μg/ml of FLAG peptide (Sigma) in 10 mM TrisHCl, pH 7.5, for 30 min at 4°C.

For reverse immunoprecipitations, lysates were incubated with 40 μl of Protein G magnetic beads (Cell Signaling) pre-coupled to 10 μg of specific or control Ig isotype antibodies. Eluates in 1X Laemmli sample buffer were resolved by 7.5% SDS-PAGE, and processed for immunoblotting with the indicated antibodies. Each IP experiment was done at least 3 times.

Chromatin Immunoprecipitation. 10^7 mESCs were cross-linked with 1% formaldehyde in PBS for 8 min, quenched with 125 mM glycine and washed with PBS. Cells were lysed in 100μl of ChIP lysis buffer (50 mM Tris-HCL, pH 8.0, 10 mM EDTA, 1% SDS, 1X protease inhibitor mix (Sigma), 1 mM PMSF, 20 mM sodium butyrate), sonicated for 25 cycles (Bioraptor Pico, Diagenode, Denville, NJ, USA) and lysates diluted to 1 ml with RIPA buffer (10 mM Tris-HCL, pH 7.5, 140 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 1% Triton X-100, 0.1% (wt/vol) SDS, 1X protease inhibitor mix (Sigma), 1 mM PMSF, 20 mM sodium butyrate). 20 μg of sonicated chromatin were incubated with 10 μg of antibodies coupled to Protein G magnetic beads (Cell Signaling) overnight at 4°C. The magnetic beads were washed 5 times with RIPA buffer and once with TE buffer (10 mM Tris, pH 8.0, 1 mM EDTA). After washing, bound DNA was eluted at 65°C in elution buffer (20 mM Tris-HCl, pH 7.5, 5 mM EDTA, 50 mM NaCl, 1% SDS, 50 μg/ml Proteinase K) overnight. After cross-linking reversal, the immunoprecipitated DNA was purified by Qiagen columns and eluted in 30 μL of TE buffer. Enrichment at target promoters was determined by real-time PCR with Power SYBR Green PCR master mix (Life Technologies) using primers specific for different areas of target genes (**Table S1**). Relative fold enrichment at different sites was calculated by using the $2^{-\Delta Ct}$ method using data from 4 different ChIP experiments.

References.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15-21.

Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc *4*, 44-57.

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

Meissner, A., Eminli, S., and Jaenisch, R. (2009). Derivation and manipulation of murine embryonic stem cells. Methods in molecular biology *482*, 3-19.

Winer, J., Jung, C.K., Shackel, I., and Williams, P.M. (1999). Development and validation of real-time quantitative reverse transcriptase-polymerase chain reaction for monitoring gene expression in cardiac myocytes in vitro. Analytical biochemistry *270*, 41-49.

Figure S1.

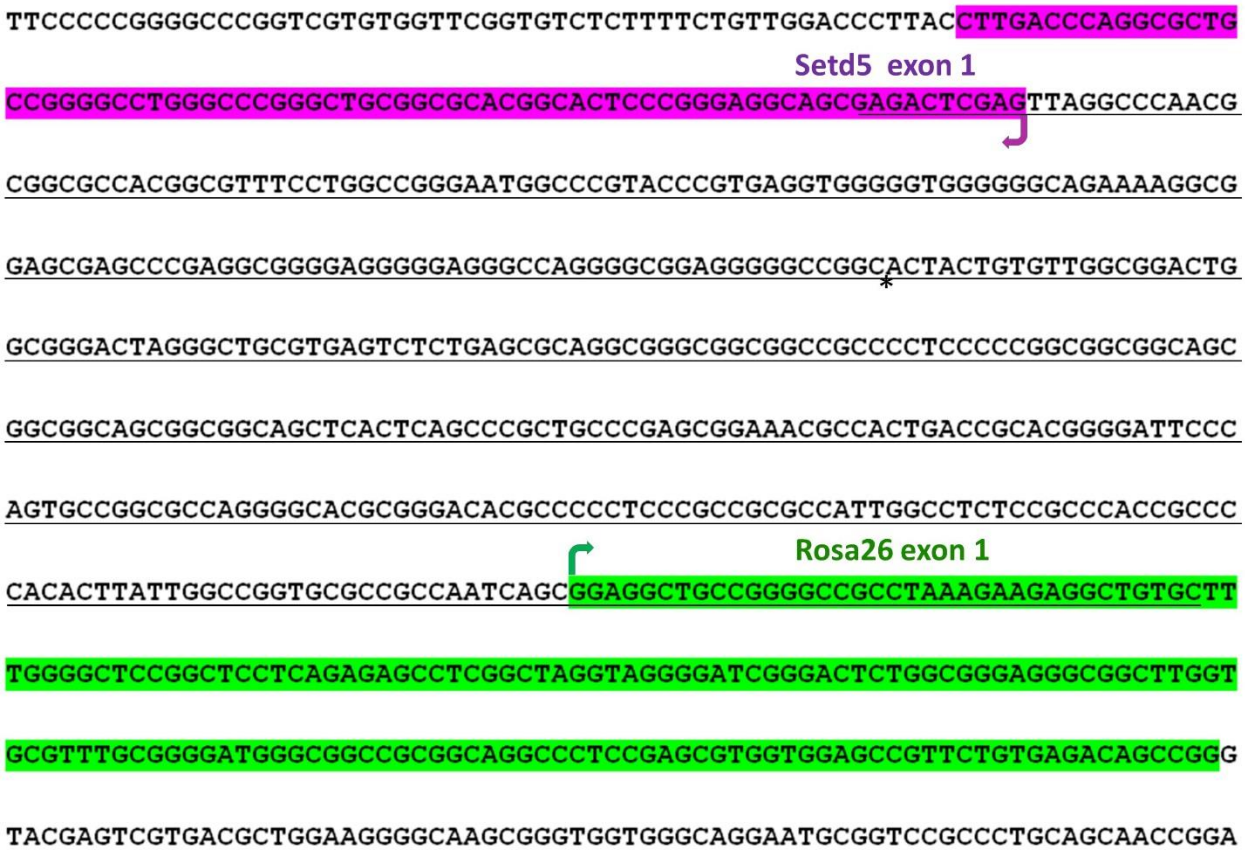

Figure S1. *Setd5-ROSA26* (SR) promoter region. The DNA sequence of the anti-sense strand (*ROSA26* direction) of *Setd5-ROSA26* promoter region is shown. Highlighted are *Setd5* exon1 (pink) and *ROSA26* exon 1 (green) according to the latest annotation of mouse genome at NCBI. The sequence of the portion of the SR promoter cloned into luciferase reporter vector to study bi-directional activity is underlined. Asterisk denotes the start position of ROSA26 promoter deletion in *ROSA26*$^{-228.3TF.GFP}$ mice and in SRΔ181 promoter in luciferase reporter.

Figure S2.

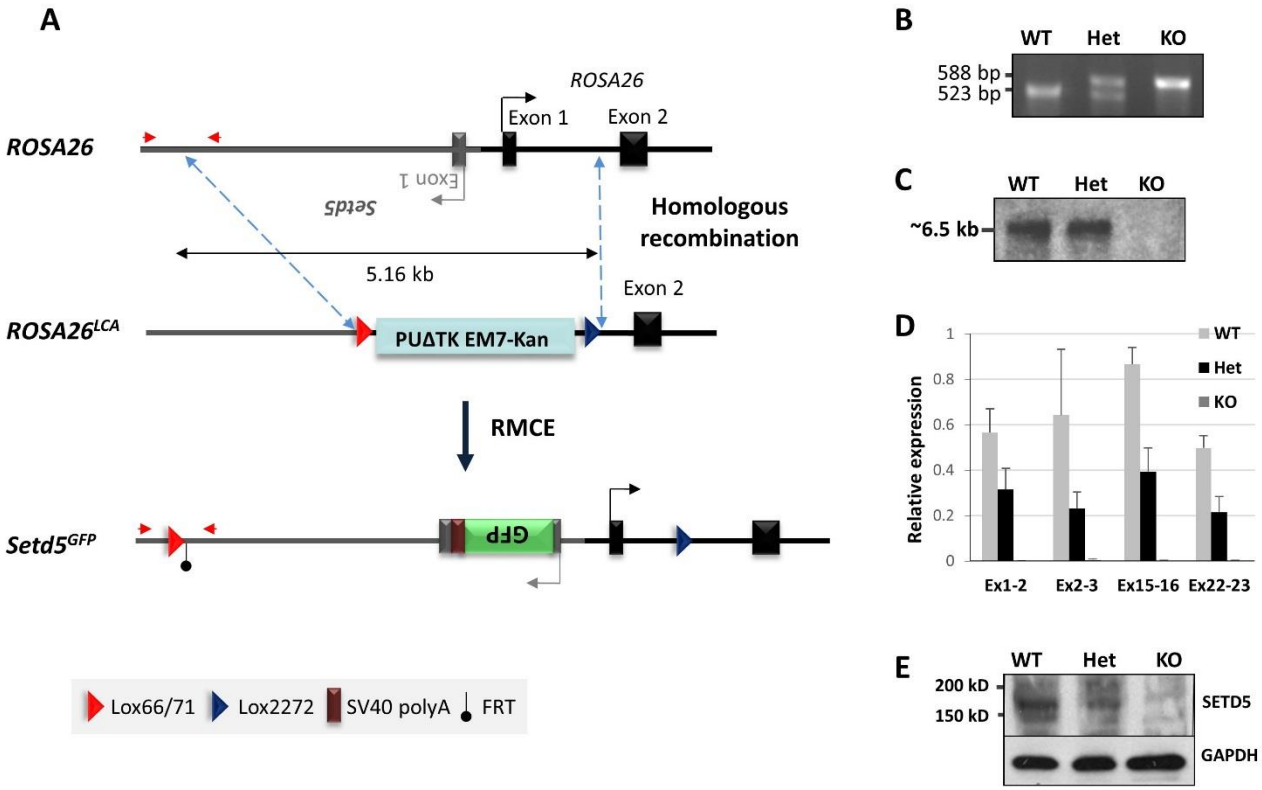

Figure S2. Generation of *Setd5*^{GFP} allele. (**A**) Schematic representation of the strategy used to make *Setd5*^{GFP} allele. In the first step, *the Setd5/ROSA26* locus was modified using homologous recombination to make *ROSA26*^{LCA} allele thereby removing 5.16 kb of genomic DNA and replacing it with the selection cassette flanked by heterotypic LoxP sites. In the second step, *Recombinase-Mediated Cassette Exchange* (RMCE) was used to derive the *Setd5*^{GFP} knock-in allele where deleted portions of the locus were re-introduced together with the insertion of a GFP-polyA cassette into the first exon of *Setd5* gene. The positions of genotyping primers are indicated by red arrows. (**B**) PCR genotyping of *Setd5*^{GFP} wild type (WT), heterozygous (Het) and knockout (KO) embryos. The wild-type and targeted alleles are represented by 523 and 588 base pair (bp)

bands, respectively. (**C**) Northern blot analysis of total RNA from WT, Het and KO embryos probed with *Setd5* cDNA probe (exons 2-6) shows complete disruption of *Setd5* mRNA (6.5kb) in the KO sample. (**D**) RT-qPCR analysis of total RNA from WT, Het and KO embryos confirmes abscence of *Setd5* transcripts in knockout tissues (N=3). (**E**) Western blot analysis of protein lysates of WT, Het and KO embryos probed with anti-S*ETD5* antibodies shows the band corresponding to the predicted SETD5 protein size (~158 kb) is absent in KO sample. Error bars: SEM.

Figure S3

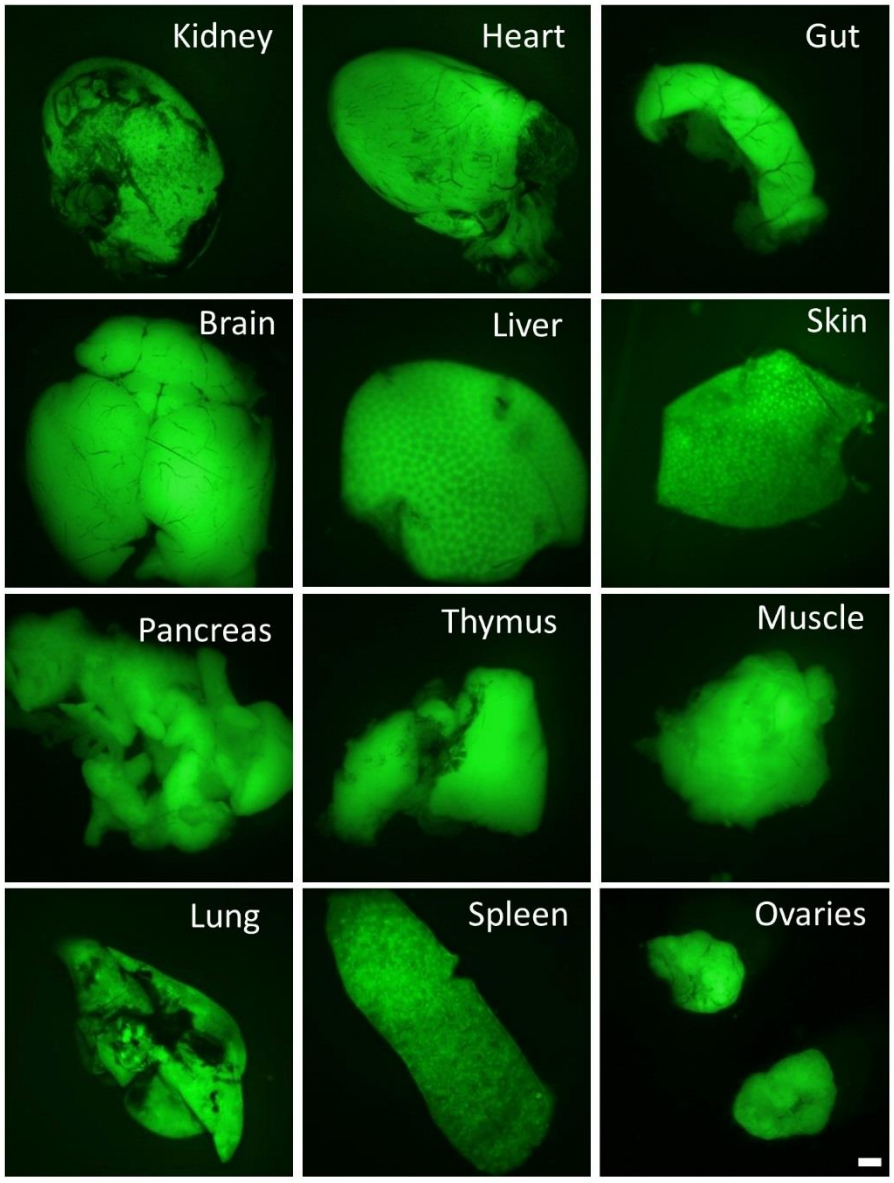

Figure S3. Expression of GFP in adult tissues of *Setd5*^{GFP} mice.

Stereoscope images of green fluorescence show ubiquitous distribution of GFP reporter expression in the adult tissues of *Setd5*^{GFP} mice. Scale bar: 1mm.

Figure S4

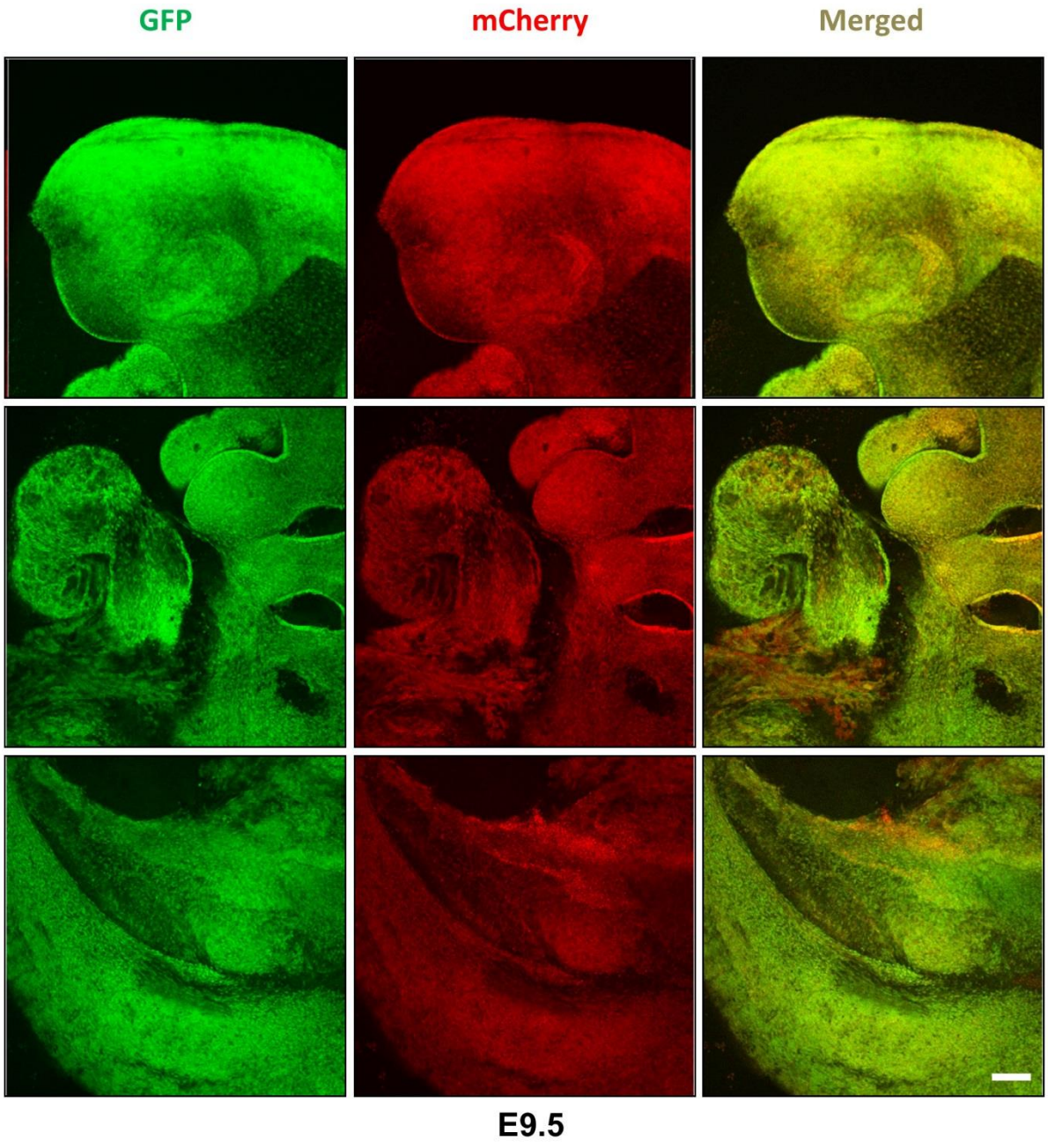

Figure S4. Co-expression of fluorescent proteins driven by Setd5 and ROSA26 promoters. Confocal imaging of direct fluorescence of *Setd5*^{GFP} and *ROSA26*^{mCherry} double heterozygous embryo (E9.5) shows co-localization of green (GFP) and red (mCherry) fluorescent proteins in different cell types. Scale bar: 100μm.

Figure S5

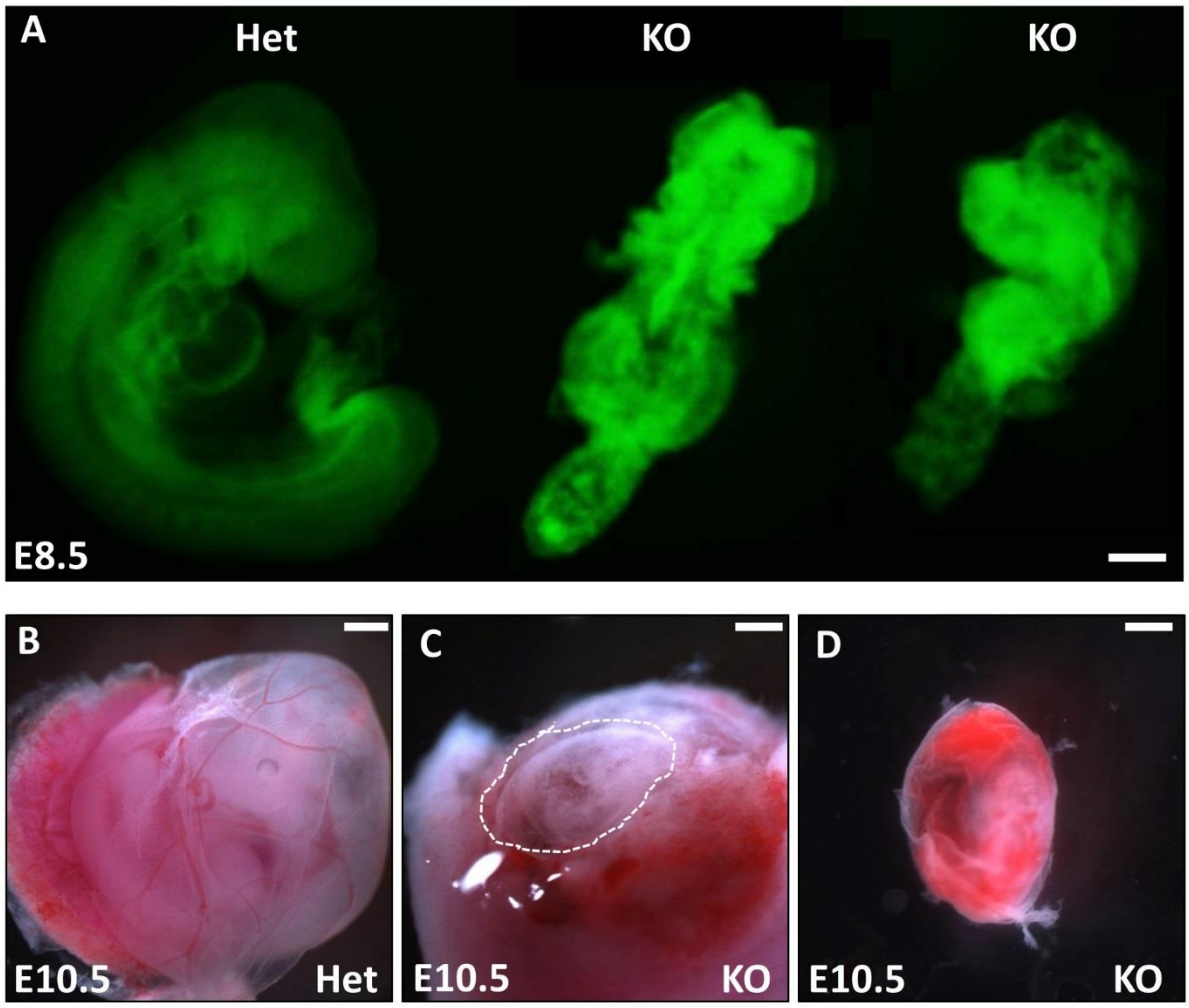

Figure S5. Phenotype of *Setd5* knockout embryos at E8.5 and E10.5. (**A**) Fluorescent (GFP) image shows gross morphology of *Setd5*^{GFP} heterozygous (Het) and knockout (KO) embryos at E8.5. KO embryos appear to be underdeveloped. (**B**) Stereoscope images taken at the same magnification in reflected light show morphology of *Setd5*^{GFP} Het and KO embryos in yolk sacs at E10.5. Dashed outline demarcates the size of KO yolk sac. (**C**) Many surviving KO embryos have widespread hemorrhaging in a yolk sac at E10.5. Scale bar: Scale bars: 500μm.

Figure S6

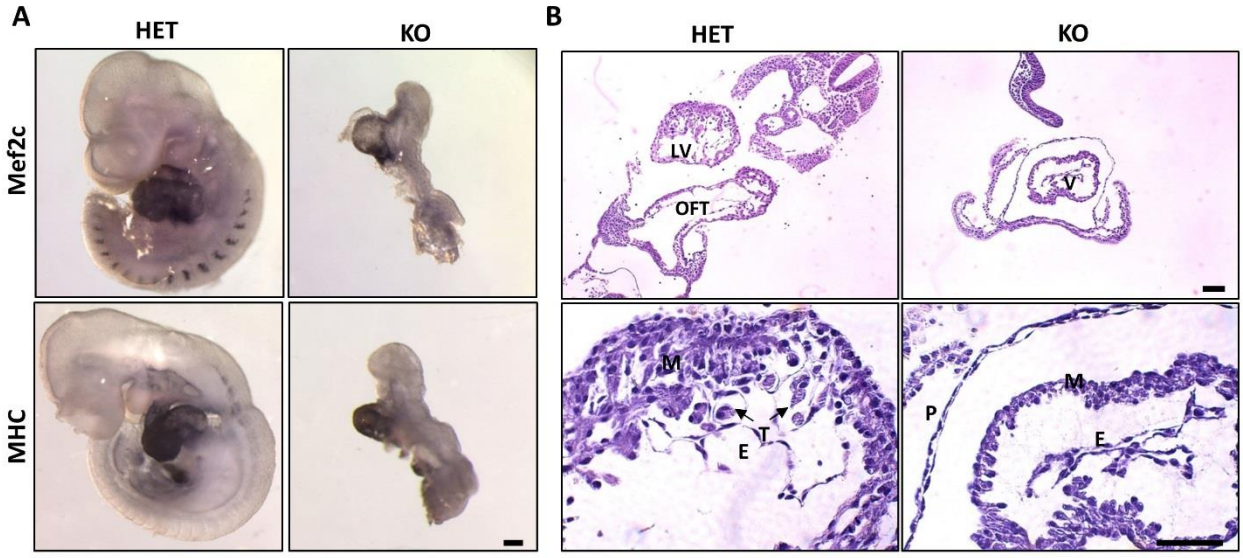

Figure S6. Cardiac phenotype of *Setd5* knockout embryos at E9.5.
(**A**) Whole mount staining with cardiac specific markers MEF2C and myosin heavy chain (MHC) shows strong staining in looped heart in *Setd5*^{GFP} heterozygous (Het) embryos, while in the knock-out (KO) embryos weaker expression is evident in the heart. Scale bar: 100μm. (**B**) H&E stained transverse sections of *Setd5*^{GFP} Het and KO hearts. The KO heart did not undergo looping and has only one ventricle (V) (top panel). Moreover, KO hearts have thinner myocardium layer with reduced trabeculation and gaps between pericardium, myocardium, and endocardium. LV, left ventricle; OFT, outflow tract; M, myocardium; P, pericardium; E, endocardium; T, trabeculae. Scale bar: Scale bar: 50μm.

Figure S7

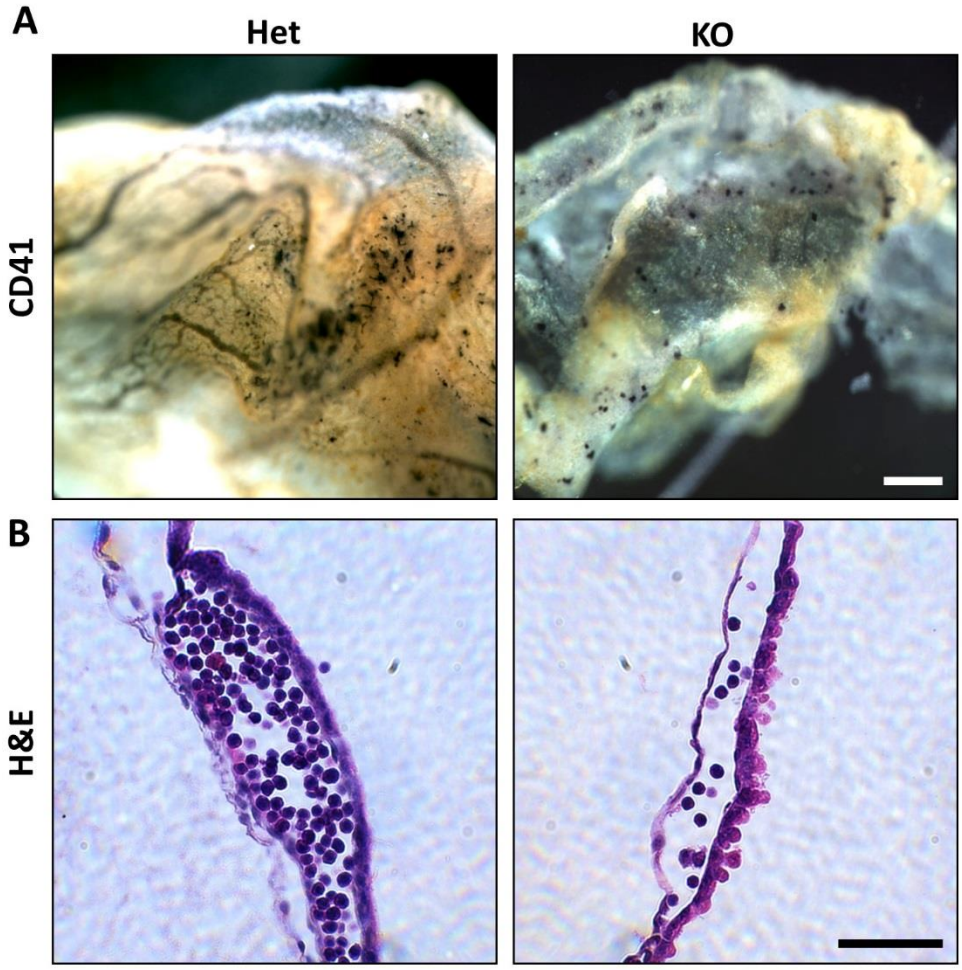

Figure S7. Hematopoiesis in Setd5 knockout embryos at E9.5. (**A**) Whole mount staining with definitive hematopoiesis marker CD41 demonstrate visible blood islands in both Setd5 heterozygous (Het) and knockout (KO) embryos. Scale bar: 500μm. (**B**) H&E stained cross-section through the yolk sac blood vessel shows the presence of blood cells including mature erythrocytes. Scale bar: 50μm.

Figure S8.

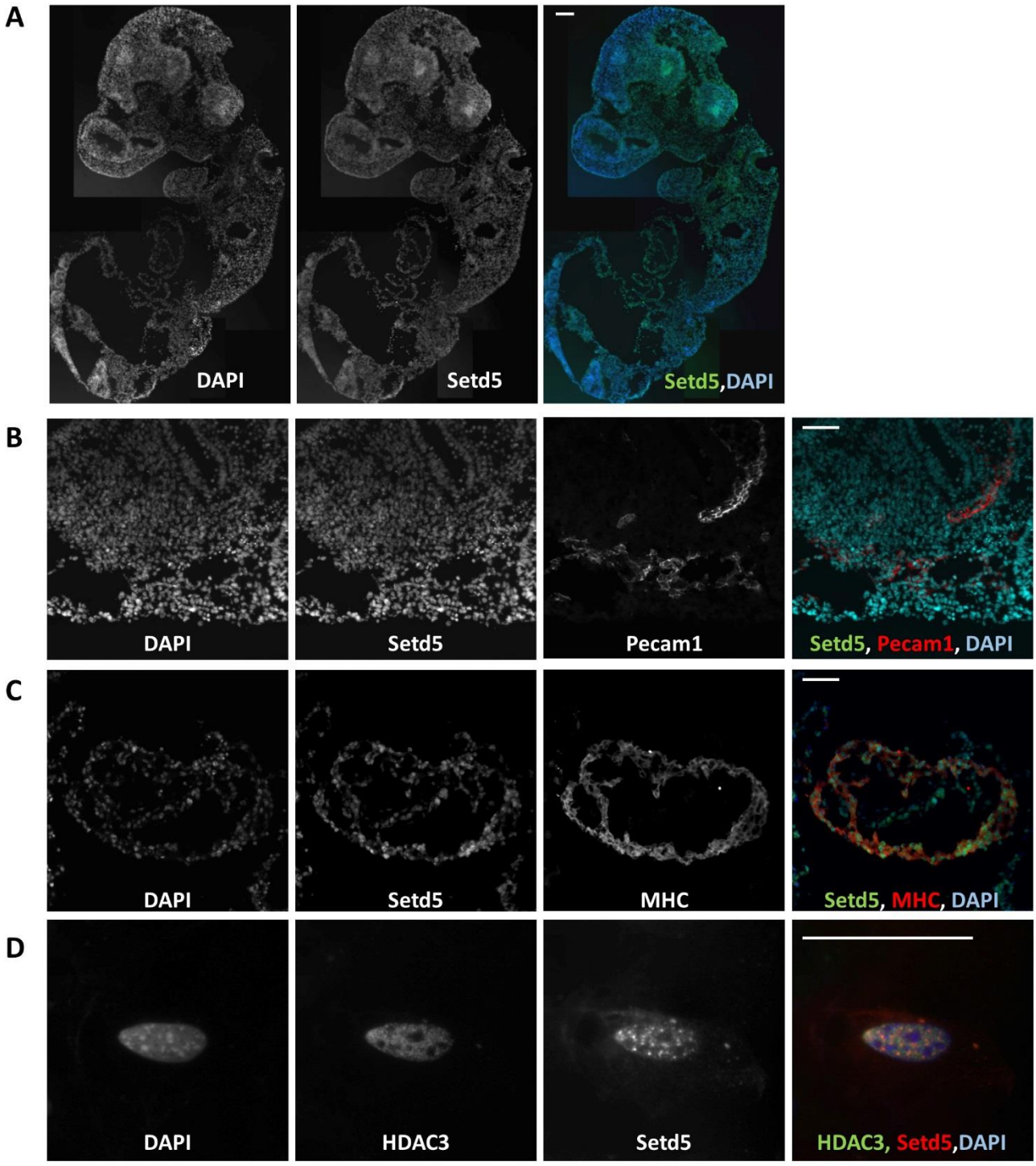

Figure S8. Immunofluorescent staining for endogenous SETD5 protein. (**A**) SETD5 protein is expressed at relatively similar levels in every cell in E9.5 wild type embryo.

Several images are stitched together to reveal sagittal view of the whole embryo. SETD5 expression in PECAM1 positive cells (**B**) and MHC positive cells (**C**) is not significantly different from surrounding tissues. (**D**) SETD5 is localized to nucleus but not nuclei of NIH3T3 cells and is partially co-localized with HDAC3. Scale bar: 50μm.

Figure S9

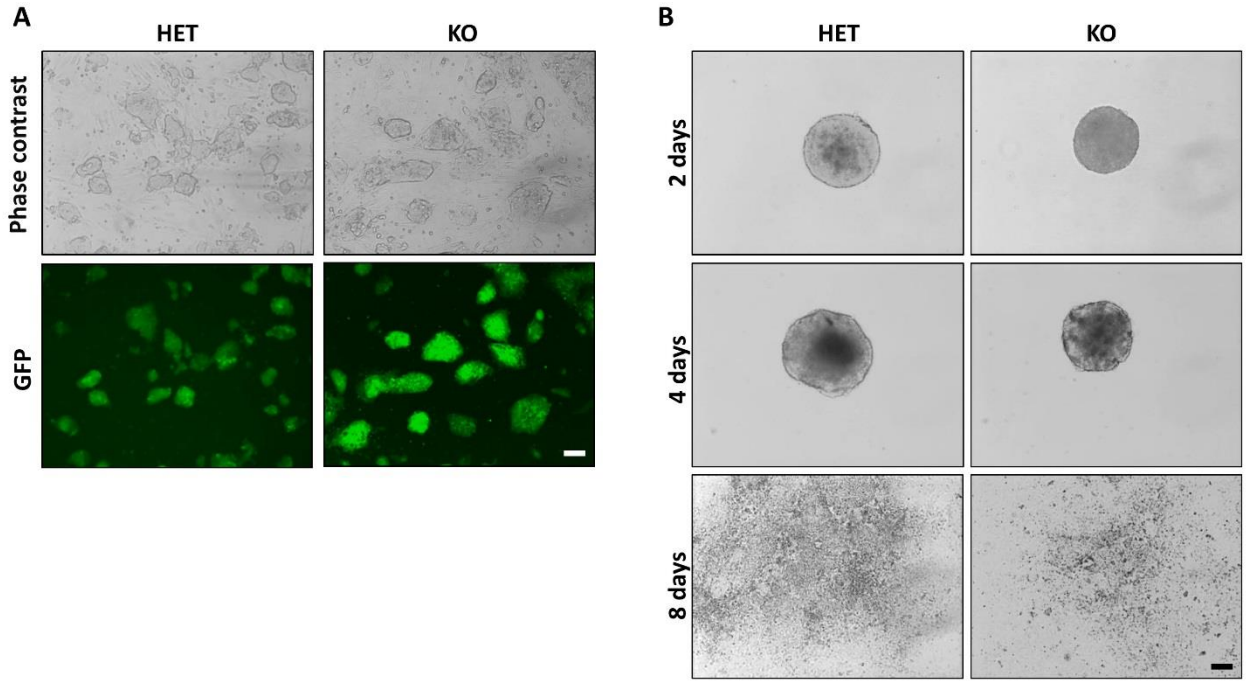

Figure S9. Phenotype of *Setd5* ES cell lines. (A) Microscope images of *Setd5*^{GFP} heterozygous (Het) and knockout (KO) ES cell lines taken with phase contrast (top panel) and green fluorescence (GFP) filter. Both cell lines display typical ES cell line morphology. (B) Differentiation of *Setd5*^{GFP} ES cells though EB formation. Appearance of EBs after 2 days (top), 4 days (middle) and EB outgrowths at 8 days of the differentiation protocol (bottom). Both, heterozygous and KO embryos formed EBs. Scale bar: 100μm.

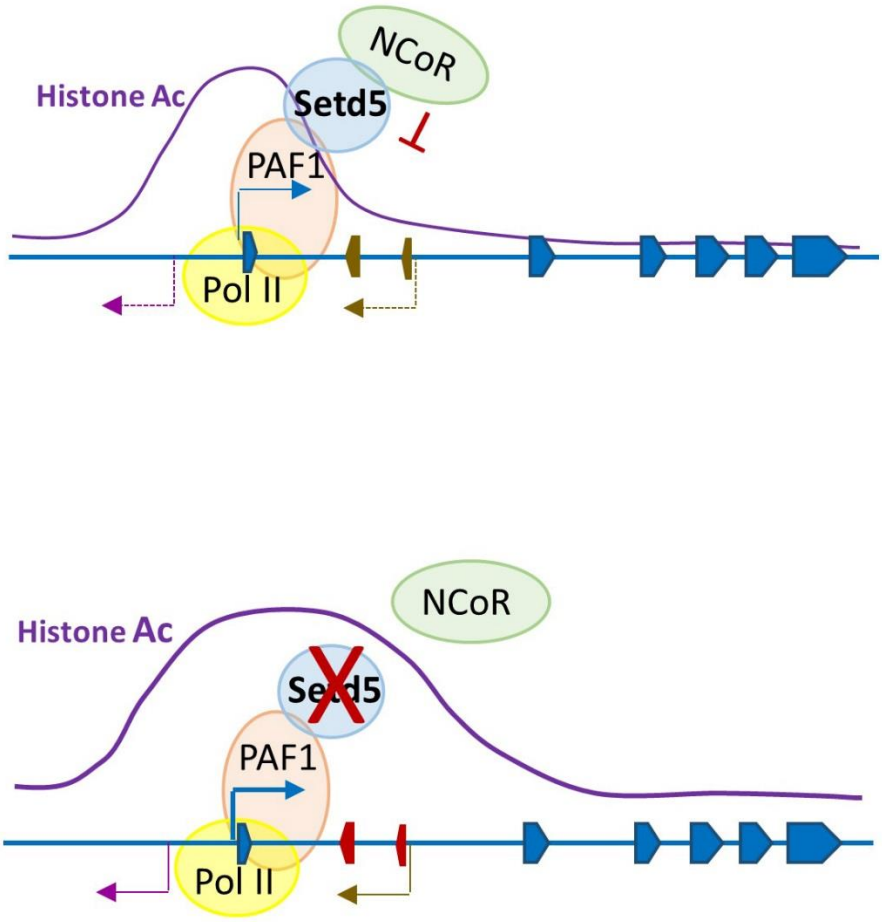

Figure S10. Model for *Setd5* involvement in co-transcriptional regulation of histone acetylation. SETD5 protein may facilitate recruitment of histone deacetylase-containing NCoR complex to PAF1 elongation complex thereby restricting histone acetylation and chromatin accessibility around promoter/first exon regions. In the absence of SETD5, histone acetylation is increased around promoter and further into a gene region, possibly, due to failed recruitment of NCoR complex to PAF1. Changes in the chromatin landscape increase expression of antisence genes located in a locus.

Table S1. PCR primers used in the study.

Primer	Sequence (5' to 3')	PCR product size	Application
Rosa26.S1 Hygro5'	AGACTTATCTACCTCATAGGTG ACGAGACTAGTGAGACGTGCTACT	649 bps	*Rosa26* RMCE screening
Rosa26.S11 Rosa26.S2	CGTGCTGAGCCAGACCTCCAT TCACAAGCAATAATAACCTGTAGT	Targeted: 588 bps Wild type: 523 bps	Genotyping *of Setd5^{GFP}*, *ROSA26$^{EN.Cherry}$*, and ROSA26228.3TF mice
Setd5D1 Setd5R1	CTGGGTCAAGGATTCCTCATG GACGTCCAAGACAGCAGAAGG	126 bps (exons 1-2)	RT-qPCR of *Setd5*
Setd5Ex2-3D Setd5Ex2-3R	GGCTGTCTTGGACGTCATGAG CTGCTGGACTAGCCTCCACAG	116 bps (exons 2-3)	RT-qPCR of *Setd5*
Setd5Ex15-16D Setd5 Ex15-16R	GATGGCACATTCAGCTCCTG CTAGTCTCTTGGGGCACAGATG	92 bps (exons 15-16)	RT-qPCR of *Setd5*
Setd5Ex22-23D Setd5Ex22-23R	GAGCCATTGAGCTCAGCACTC GGACGAACTCTGCTGAAGGAG	126 bps (exons 22-23)	RT-qPCR of *Setd5*
SRpromF SRpromR	GAGACTCGAGTTAGGCCCAAC GCACAGCCTCTTCTTTAGGC	439 bps (SR promoter)	Cloning of Setd5 - Rosa26 promoter
SRpromF SRprom∆181R	GAGACTCGAGTTAGGCCCAAC GCCGGGGAGGGCGGC	222 bps (SR∆181 promoter)	Cloning of Setd5 - Rosa26 promoter
R26mRNA-D R26mRNA-R	TCCTCAGAGAGCCTCGGCTAG CATCATGCCTCTGCTTGCTTC	160 bps (exons 1-2)	RT-qPCR of ROSA26
Ef1aPr1-F Ef1aPr1-R	GAATCAACCGGCATTGGATG CTGGGAGAGGAACACAATGTTG	132bps (Area I)	ChIP *Eef1a1*
Ef1a1ex-F Ef1a1ex-R	GAACGGTATATAAGTGCGGCAG GGGAATGCTCGCAGCTAATC	171bps (Area II)	ChIP *Eef1a1*
Ef1a1in-F Ef1a1in-R	GTCATGGTTGGGGAGGAATG CTGGTTGCTTCGGGAAAAAC	103bps (Area III)	ChIP *Eef1a1*
Ef1a4ex-F Ef1a4ex-R	CCTGATTGTTGCTGCTGGTG GACTGTATGGTGGCTCGGTG	152bps (Area IV)	ChIP *Eef1a1*
Ef1a8ex-F Ef1a8ex-R	GTCGCTTTGCTGTTCGTGAC ACTGGGGTGGCAGGTGTTAG	154bps (Area V)	ChIP *Eef1a1*
Ip6kPr1-F Ip6kPr1-R	CAGGAACTTAACGGTGCTCTGAG CACAGCCGTCTGTGCTACCTC	127bps (Area I)	ChIP *Ip6k1*
Ip6kPr2-F Ip6kPr2-R	CTCCACTGCTAGCAAACCTGAG CTTCGCGTTGATTCGTCAAC	127bps (Area II)	ChIP *Ip6k1*
Ip6kEx1-F Ip6kEx1-R	GTTGACGAATCAACGCGAAG CACTGGGTACGGATCAACAAC	164bps (Area III)	ChIP *Ip6k1*
Ip6kIn1-F Ip6kIn1-R	CTGCTTGCTCTGGCCCATAG CTGACTGCTCTCCTGGAGGTC	156bps (Area IV)	ChIP *Ip6k1*
Ip6kEx2-F Ip6kEx2-R	GACTCTGGCCCAGAGTTCCTC GACGTGCTCTGCTCACACTTC	136 bps (Area V)	ChIP *Ip6k1*
Ip6kEx5-F Ip6kEx5-R	GACTCTGGCCCAGAGTTCCTC GACGTGCTCTGCTCACACTTC	155bps (Area VI)	ChIP *Ip6k1*
HaghPr1-F HaghPr1-R	CATTAACACCGGTGAGCCCTC CTGCGTGGGGAGGAACTATG	121bp (Area I)	ChIP *Hagh1*, RT-qPCR Fadh1
HaghPr2-F HaghPr2-R	CACTGGTTGGCTGTGAACTCTATG CAGTGTCGTCCTCCTACTGCTC	127bps (Area II)	ChIP *Hagh1*
HaghEx1-F HaghEx1-R	GAGCAGTAGGAGGACGACACTG GACAGGCTCCGGAGACACAG	148bps (Area III)	ChIP *Hagh1*
HaghIn1-F	CAGGGTTGTGACAGAGCACTTC	127bps (Area IV)	ChIP *Hagh1*

HaghIn1-R	CACCGAAGGATTGTCTTGGAG		
HaghEx4-F HaghEx4-R	CTGGTGGGAACGAGAAGCTG CTGCAGTGTGGAGAGGTGTG	110bps (Area V)	ChIP *Hagh1*
HaghEx8-F HaghEx8-R	GACTGTGCAACAGCATGCTG CAGGCCAGTCGGTTAGAAGTC	137bps (Area VI)	ChIP *Hagh1*
HaghRT-F HaghRT-R	GACGCTGGTGTTGCGGACTC GATAGTCCCAGCAGGGCCTG	140bps (exons 1-2)	ChIP *Hagh1*
Ef1a1ex-F Ef1aRT-R	GAACGGTATATAAGTGCGGCAG GTGGTGGACTTGCCGGAATC	171bps (exons 1-2)	RT-qPCR *Eef1a1*
GM38RT -R Ip6kIn1-F	ATGGGCTGGTGAGATGGCTC CTGCTTGCTCTGGCCCATAG	189bps (exons 1-2)	RT-qPCR *GM38134*
Ipk6RT-F Ipk6RT-R	GATCCGTACCCAGTGGGCAG GAGGCCAATGGTCACAGTCTG	212bps (exons 1-2)	RT-qPCR *Ip6k1*
Mesp1-F Mesp1-R	TGTACGCAGAAACAGCATCC TTGTCCCCTCCACTCTTCAG	144bps (exons 1-2)	RT-qPCR *Mesp1*
Meox1-F Meox1-R	GAGTTGAAGGTTAGGAAGTGGC TTCCAAGGTCCACGTATCTC	92bp (exon 3)	RT-qPCR *Meox1*
Actb-F Actb-R	ACGATGCTCCCCGGGCTGCATTC TCTCTTGCTCTGGGCCTCGTCACC	115bp (exons2-3)	RT-qPCR *Actb*

Table S2. Analysis of *Setd5*^{GFP} matings at weaning and during embryogenesis.

	Genotype				
	+/+		+/−		−/−
Weaning	27		42		0
	39%	:	61%	:	
E9.5	48		103		43(6)
	25%	:	53%	:	22%
E10.5	12		21		3(8)
	33%	:	58%	:	8%
E11.5	8		14		(10)
	36%	:	63%	:	
Expected	25%	:	50%	:	25%

The first row for each age shows the number of viable offspring of different genotypes derived from matings of animals heterozygous for *Setd5*. The number of dead/resorbed embryos are indicated in parentheses. The second row (shaded grey) shows the percentage of embryos of each genotype of the total live embryos analyzed for each age. Expected Mendelian ratios are shown in the bottom row.

Table S3.

Click here to Download Table S3

Table S4.

Click here to Download Table S4