

1 MLL4 is required for the first embryonic collective cell migration whereas MLL3
2 is not required until birth

3

4 Deepthi Ashokkumar^{1*}, Qinyu Zhang^{1*}, Christian Much¹, Anita S. Bledau², Jun
5 Fu¹, Konstantinos Anastassiadis², A. Francis Stewart^{1,3}, Andrea Kranz¹

6

7 1. Genomics, Center for Molecular and Cellular Bioengineering,
8 Biotechnology Center, Technische Universität Dresden, Tatzberg 47, 01307
9 Dresden, Germany.

10 2. Stem Cell Engineering, Center for Molecular and Cellular Bioengineering,
11 Biotechnology Center, Technische Universität Dresden, Tatzberg 47, 01307
12 Dresden, Germany.

13 3. Max-Planck-Institute for Cellular Biology and Genetics, Pfotenhauerstr. 105,
14 01307 Dresden, Germany.

15

16 * equal contribution

17 corresponding authors

18 A. Francis Stewart, tel. +49-351-463-40130, fax +49-351-463-40143,
19 francis.stewart@tu-dresden.de

20 Andrea Kranz, BIOTEC, tel. +49-351-463-40152, fax +49-351-463-40143,
21 andrea.kranz@tu-dresden.de

22

23 Running title: *MLL3* and *MLL4* knockouts

24

25

Summary statement

The H3K4 methyltransferases MLL3 and MLL4 have strikingly different null phenotypes during mouse development; MLL3 is required for lung maturation whereas MLL4 is required for anterior visceral endoderm migration.

Abstract

Methylation of histone 3 lysine 4 (H3K4) is a major epigenetic system associated with gene expression. In mammals there are six H3K4 methyltransferases related to yeast Set1 and fly Trithorax, including two orthologs of fly Trithorax-related: MLL3 and MLL4. Exome sequencing has documented high frequencies of *MLL3* and *MLL4* mutations in many types of human cancer. Despite this emerging importance, the requirements of these sister genes in mammalian development have only been incompletely reported. Here we examined the null phenotypes to establish that MLL3 is first required for lung maturation whereas MLL4 is first required for migration of the anterior visceral endoderm (AVE) that initiates gastrulation and is the first collective cell migration in development. This migration is preceded by a columnar to squamous transition in visceral endoderm cells that depends on MLL4. Furthermore, *MLL4* mutants display incompletely penetrant, sex distorted, embryonic haploinsufficiency and adult heterozygous mutants show aspects of Kabuki syndrome, indicating that MLL4 action, unlike MLL3, is dosage dependent. The highly specific and discordant functions of these sister genes argues against their action as general enhancer factors.

Introduction

The lysine methylation status of the histone 3 tail is central to epigenetic regulation pivoting on methylation of lysines at positions 4, 9, 27 and 36. All active Pol II promoters are characterized by trimethylation of histone 3 lysine 4 (H3K4me3) on the first nucleosome in the transcribed region. Dimethylation (H3K4me2) is a general characteristic of transcribed regions whereas monomethylation (H3K4me1) is a general characteristic of active chromatin with peaks on enhancers (Bannister and Kouzarides, 2011).

Mammals have six Set1/Trithorax type H3K4 methyltransferases in three pairs; SETD1A and B (KMT2F and KMT2G), which are homologues of yeast Set1; MLL1 and 2 (KMT2A and KMT2B), which are homologues of *Drosophila* Trithorax; and MLL3 and 4 (KMT2C and KMT2D), which are homologues of *Drosophila* Lost PHD fingers of Trr (*Ltr*) fused to Trithorax-related (*Trr*). All six are found in individual complexes however all six complexes share the same highly conserved scaffold, first reported for yeast Set1C (Roguev et al., 2001), composed of four subunits, WDR5, RBBP5, ASH2L, DPY30 termed 'WRAD' (Lee et al., 2006; Ruthenburg et al., 2007; Ernst and Vakoc, 2012), which surrounds the SET domain and is required for enzymatic activity (Kim et al., 2013; Hsu et al., 2018; Qu et al., 2018). SETD1A apparently conveys most H3K4me3 in most mammalian cell types (Bledau et al., 2014). Similarly the dSet1 homologue conveys most H3K4me3 in most *Drosophila* cell types (Mohan et al., 2011; Ardehali et al., 2011; Hallson et al., 2012). Consequently the Set1 homologues are primarily implicated in trimethylation and general promoter function. In contrast, evidence indicating that MLL3 and 4 are monomethylases (Weirich et al., 2015; Zhang et al., 2015; Li et al., 2016) has triggered their linkage to enhancer function (Lee et al., 2013; Rao and Dou, 2015; Piunti and Shilatifard, 2016). Whether they proceed to catalyze di- and tri- H3K4 methylation remains uncertain (Dhar et al., 2012) and the emergent model relating Set1 activities to promoters and MLL3/4 to enhancers requires further substantiation.

Histone 3 lysine methyltransferases are prominent members of both Trithorax- (Trx-G) and Polycomb-Groups (Pc-G) (Steffen and Ringrose, 2014; Schuettengruber et al., 2017) with the genetic opposition between Trx-G and Pc-G being exerted, in part, by a competition for the methylation status of the

histone 3 tail on key nucleosomes (Schmitges et al., 2011). This opposition is central to epigenetic regulation in development, differentiation, homeostasis and more recently, oncogenesis (Chi et al., 2010; Rao and Dou, 2015; Soshnev et al., 2016) with several Trx-G and Pc-G factors, including the H3K27 methyltransferase, EZH2, implicated as oncogenes or tumor suppressors in a variety of malignancies. *Mll1* was discovered as the major leukemia gene at the 11q23.1 translocation involved in early onset childhood leukemia (Li and Ernst, 2014). The N-terminal half of MLL1 fused to many (now more than 70) different C-terminal partners, including AF4 and AF9 (Slany, 2009; Meyer et al., 2018) is leukemogenic without the need for secondary mutations (Dobson et al., 2000). These MLL1 fusion proteins promote both acute lymphocytic (ALL) and acute myeloid (AML) leukemias, collectively termed mixed lineage leukemias.

Massively parallel sequencing of cancer exomes by the international cancer genome projects revealed somatic mutations in *Mll3* and *Mll4* in almost all cancers analyzed (Rao and Dou, 2015). Inactivating heterozygous mutations have been identified in patients with medulloblastoma, B cell lymphoma, bladder carcinoma, renal carcinoma and colorectal cancer, amongst many other cancers (Morin et al., 2011; Parsons et al., 2011; Pasqualucci et al., 2011). An explanation of these findings is lacking however recent evidence suggests that mutation of *Mll4* promotes defective transcription-coupled DNA repair (Kantidakis et al., 2016).

Exome sequencing also revealed mutations in *Mll4* as the cause of Kabuki syndrome type I (Ng et al., 2010; Li et al., 2011). All *Mll4* Kabuki mutations are apparently *de novo* somatic heterozygous nonsense or frameshift mutations that appear throughout the gene, but most commonly in exon 48. Most of these *Mll4* mutations truncate the protein and all are haploinsufficient (Banka et al., 2012; Bogershausen et al., 2015; Faundes et al., 2019). The less common Kabuki syndrome type 2 is caused by mutations of *Utx*. UTX, which is an H3K27 demethylase, is a subunit of the MLL4 complex (Lee et al., 2006; Lederer et al., 2012; Lederer et al., 2014; Banka et al., 2015).

As for MLL1 and MLL2 (Denissov et al., 2014), MLL3 and MLL4 may have overlapping and redundant functions in mammalian cells (Lee et al., 2013). Notably, the H3K4 methyltransferase activities of MLL3 and MLL4 are dispensable for gene expression in mouse embryonic stem cells (ESCs)

(Dorigi et al., 2017). Similarly, the catalytic activity of the SET domain of Trr (the fly homologue of MLL3/MLL4) is dispensable for development and viability in *Drosophila* (Rickels et al., 2017).

Unlike the other four H3K4 methyltransferases (Yagi et al., 1998; Ernst et al., 2004; Glaser et al., 2006; Jude et al., 2007; Glaser et al., 2009; Andreu-Vieyra et al., 2010; Bledau et al., 2014) and despite their emerging importance for cancer, the roles of *Mll3* and *Mll4* in mammalian development have only been partly described (Ang et al., 2016; Lee et al., 2013; Jang et al., 2017). Here we compare the null phenotypes of these two genes in mouse development.

Material and Methods

Targeting constructs

The targeting constructs for *Mll3* and *Mll4* were generated using Red/ET recombineering (Fu et al., 2010). For *Mll3* an FRT-SA-GT0-T2A-lacZneo-CoTC-FRT-loxP cassette was inserted into intron 48. Additionally, a loxP-rox-PGK-Blasticidin-pA-rox cassette was introduced into intron 49. For *Mll4*, a loxP site was introduced into intron 4 using a loxP-zeo-loxP cassette with subsequent removal in *E. coli* by Cre recombination using pSC101-BAD-Cre-tet (Anastassiadis et al., 2009). Then a loxP-FRT-SA-IRES-lacZneo-pA-FRT cassette was inserted in intron 1 of the gene. The homology arms were 5' 4.6kb/3' 5kb and 5' 4.7kb/3' 4.9kb for the *Mll3* and *Mll4* targeting constructs, respectively.

Gene targeting and generation of conditional knockout mice

R1 ESCs were cultured with FCS-based medium (DMEM + GlutaMAXTM (Invitrogen), 15% FCS (Fetal Calf Serum, PAA), 2 mM L-Glutamine (Invitrogen), 1xnon-essential amino acids (Invitrogen), 1 mM sodium pyruvate (Invitrogen), 0.1 mM β -mercaptoethanol, in the presence of LIF)) on Mitomycin-C inactivated mouse embryonic fibroblasts. Cells (1×10^7) were electroporated with 40 μ g of the linearized targeting construct using standard conditions (250 V, 500 μ F) and selected with 0.2 mg/ml G418. Correct integration in the *Mll3* locus was confirmed by Southern blot analysis using 5' and 3' external probes (4 positive clones out of 36). Correct integration in the *Mll4* locus was confirmed by Southern blot analysis using an internal probe and 5' and 3' external probes (6

positive clones out of 32). To remove the additional selection cassette (PGK-Blasticidin-polyA) downstream of exon 49, correctly targeted *Mll3* clones were electroporated with CAGGS-Dre-IRES-puro expression vector (Anastassiadis et al., 2009) and clones screened by PCR for complete recombination and sensitivity to blasticidin. For both *Mll3* and *Mll4* two correctly targeted ES cell clones were injected into blastocysts. For *Mll3* only one of the clones gave germ line transmission. *Mll3*^{D/+} mice were crossed to the *CAGGS-Flpo* line (Kranz et al., 2010) to generate *Mll3*^{FD/+} mice. Subsequently, those mice were crossed to the *PGK-Cre* line (Lallemand et al., 1998) to produce *Mll3*^{FDC/+} mice. *Mll3*^{D/+} and *Mll3*^{FDC/+} mice were backcrossed at least 15 generations to C57BL/6JOlaHsd mice. Regarding *Mll4* both clones gave rise to several chimeras but only one of these males was able to establish germ line transmission. *Mll4*^{A/+} mice were backcrossed at least 15 generations to CD1 mice. Primers for genotyping are provided in Table S1. All animal experiments were performed according to German law.

Western blot analysis

ESCs were homogenized in buffer E (20 mM HEPES pH 8.0, 350 mM NaCl, 10% glycerol, 0.1% Tween 20, 1 mM PMSF, 1x complete protease inhibitor cocktail) and protein extracts were obtained after three cycles of freezing and thawing. Whole cell extracts were subsequently separated by NuPAGE 3-8% Tris-acetate gel (Invitrogen), transferred to PVDF membranes and probed with an MLL4 specific polyclonal antibody. The antibody was generated by immunizing rabbits with a mixture of three KLH-conjugated synthetic peptides from the central part of the MLL4 protein (QRPRFYVPVSEELHRLAP, NGDEFDLLAYT, KQQLSAQTQRLAPS) (extended data file 1). Antibody information and dilution can be found in Table S3.

Whole-mount X-gal staining and immunostaining

Embryos were dissected, fixed with 0.2% glutaraldehyde and X-gal staining was performed as described (Kranz et al., 2010). Embryos and organs were dissected and fixed with 4% paraformaldehyde overnight. Dehydration and paraffin infiltration was done automatically using the Paraffin-Infiltration-Processor (STP 420, Zeiss). The dehydrated tissues were embedded in paraffin (Paraffin Embedding Center EG116, Leica) and sections were prepared. Antigen retrieval was performed by microwaving slides in 10 mM

citrate buffer (pH 6.0) for 12 min (Microwave RHS 30, Diapath). Immunohistochemical staining was carried out as previously described (Bledau et al., 2014). DAB images were collected with an Olympus WF upright microscope. For immunofluorescence procedures, the sections were permeabilized in 0.5% Triton X-100 in PBS for 10 min, blocked for 1 hour at RT, incubated with primary antibody overnight at 4°C, followed by secondary antibody for 2 hours at RT. Sections were mounted with Mowiol and imaged with Zeiss Scanning confocal microscope LSM/780. Antibody information and dilutions can be found in Table S3.

Whole-mount in situ hybridization

Whole-mount *in situ* hybridization was carried out according to standard procedures (Riddle et al., 1993; Piette et al., 2008). Digoxigenin-labelled riboprobes specific for the following genes were used in this study: *Brachyury* (Herrmann, 1991), *Mox1*, *Hoxb1* and *Wnt1* (Glaser et al., 2006), *Otx2* (Ang et al., 1996), *Nodal*, *Eomes*, *Bmp4* and *Hex* (Norris et al., 2002), *Cer1* (Belo et al., 1997), *Foxa2* (Norris et al., 2002) and *Gsc*, *Lhx1* and *Wnt3* (Liu et al., 1999), *Dkk1* (Stuckey et al., 2011), *Lefty1* (Stuckey et al., 2011), *Shh* and *Tbx6* (Alten et al., 2012). Anti-dig-AP antibody and NBT/BCIP colorimetric signal detection were used for whole-mount *in situ* hybridizations. Embryos were imaged with Nikon SMZ 1500 stereomicroscope.

Whole-mount immunofluorescence for Hex

For whole-mount immunofluorescence, PFA-fixed embryos were permeabilized in 0.5% Triton X-100 in PBS for 1 hour at RT, incubated with anti-Hex antibody (Hoshino et al., 2015) overnight at 4°C followed by goat anti-rabbit IgG-CFL 488 (Santa Cruz) secondary antibody. The embryos were imaged with Zeiss Scanning confocal microscope LSM/780.

Glucose and insulin tolerance test

For glucose tolerance test, mice fasted 16 h before 1.5 mg glucose per g body weight was applied by gavage. For insulin tolerance test, mice were injected intraperitoneally with 0.75 mU insulin per g body weight after 6 h of fasting. For measurement of blood glucose levels blood samples were drawn at 0, 15, 30, 60, 90, and 120 min after administration of the respective solution.

Reverse transcription and quantitative PCR (qRT-PCR) analysis

Total RNA was isolated using Trizol reagent (Sigma-Aldrich) and reverse transcribed using the AffinityScript Multiple Temperature cDNA Synthesis kit (Agilent Technologies). Real-time quantitative PCR was performed with Go Taq qPCR Master Mix (Promega) by Mx3000P QPCR System (Agilent Technologies). Ct values were normalized against *Rpl19*. Primer sequences and length of the amplified products are given in Table S2. Fold differences in expression levels were calculated according to the $2^{-\Delta Ct}$ method (Livak and Schmittgen, 2001).

Results

***Mll3* and *Mll4* are conserved paralogs**

MLL3 and MLL4 are the largest known nuclear proteins at 4903 and 5588 amino acids respectively and their genes clearly arose by duplication. Both genes and proteins have the same architecture based on the positions of splice sites, PHD fingers, HMG box, FYRN/FYRC and SET domains (Fig. 1A; extended data file 1). Except for PHD3 of MLL3, which has been lost from MLL4 (because PHD3 can be found in *Drosophila Lpt* (Lost PHD fingers of *Trr*)) (Mohan et al., 2011; Chauhan et al., 2012), the other five PHD and ePHD fingers share high identity. Both proteins are notable for their extensive regions of low sequence complexity, particularly MLL4 contains several lengthy stretches of glutamine repeats including a patch of 450 amino acids C-terminal to the HMG box with more than 50% glutamines and a 600 amino acid patch with one-third prolines after its second PHD finger (extended data file 1). Both genes encompass more than 50 exons (Fig. 1B, Fig. 1C) spliced to very long mRNAs; 14 kb for *Mll3* and 19 kb for *Mll4*, that are widely expressed in the embryo (Fig. S1C, S2C).

***Mll3* and *Mll4* knockouts die at different developmental stages**

To explore the function of these conserved proteins, we established multipurpose alleles (Testa et al., 2004) for both *Mll3* and *Mll4* by gene targeting. In our multipurpose allele strategy, frameshifting exons are flanked by loxP sites (exon 49 for *Mll3*; Fig. 1B, Fig. S1A, S1B and exons 2-4 for *Mll4*; Fig. 1C, Fig. S2A, Fig. S2B) accompanied by the insertion of a genetrapp stop cassette in the intron upstream of these exons. The stop cassette, which is

flanked by FRT sites, contains a *LacZ* reporter and stops target gene transcription because it includes a 5' splice site, which captures the target gene transcript and a polyadenylation site, which terminates it, thereby - ideally - producing a null allele, termed the 'A' allele (Testa et al., 2004; Skarnes et al., 2011). After FLP recombination to remove the stop cassette, which establishes the 'F' allele and restores wild type expression, subsequent Cre recombination establishes a frame-shifted mRNA in the 'FC' allele that should provoke nonsense mediated mRNA decay (NMD) (Dyle et al., 2019).

The multipurpose allele strategy aims to establish a loxP allele for conditional mutagenesis and also to mutate the target gene in two different ways, either by truncation of the mRNA (A allele) or by NMD (FC allele). So if A/A and FC/FC present the same phenotype, the conclusion that both are null can be established because the A and FC alleles mutate the gene in different ways. Although unlikely, this conclusion is not secure if the two different mutations produce the same hypomorphic or dominant negative phenotypes.

For *Mll3*, in addition to the multipurpose allele strategy we included a rox-flanked blasticidin-selectable cassette to provide for selection of the 3' loxP site. Deletion of the rox-flanked cassette by Dre recombinase established the 'D' allele, which is equivalent to the 'A' allele described above. Subsequent FLP and Cre recombination establish 'FD' and 'FDC' alleles that are equivalent to 'F' and 'FC' alleles respectively.

The multipurpose allele logic worked for *Mll3*. The D/D and FDC/FDC phenotypes were the same (as described below), supporting the conclusion that the knockout phenotype reported here is the null. In further support of this conclusion, targeting to insert the genetrap cassette into *Mll3* intron 33 presented the same phenotype (Table 1) as those described below for exon 49 targeting. Furthermore, a BayGenomics genetrap in *Mll3* intron 9 also provoked neonatal death (Lee et al., 2013). Although no further analysis other than neonatal death was reported, this outcome concords with the other three *Mll3* mutagenic alleles, supporting the conclusion that MLL3 is not required until birth.

For *Mll4*, the *Mll4^{A/A}* and *Mll4^{FC/FC}* phenotypes were different, potentially revealing different aspects of MLL4 function. As described below, *Mll4^{A/A}* embryos are defective before gastrulation whereas *Mll4^{FC/FC}* died at birth. As

expected, MLL4 expression was abolished by the intron 1 genetrap cassette insertion in *Mll4^{Δ/Δ}* ESCs (Fig. 1D), indicating that the stronger phenotype is the null. However, mRNA and truncated protein expression persisted in *Mll4^{FC/FC}* ESCs (data not shown) indicating that the milder phenotype was hypomorphic. Analysis of this hypomorphic allele will not be presented in this paper. Our conclusions about the null *Mll4* phenotype are supported by a briefly described homozygous BayGenomics *Mll4* genetrap in intron 19 (Lee et al., 2013), which also resulted in embryonic lethality before E10.5. Although not investigated, this phenotype was severe and could be the same as *Mll4^{Δ/Δ}*. Notably another *Mll4* allele may also present the same phenotype. Aiming to mutate the methyltransferase activity of MLL4, Jang et al. (2017) mutated three conserved tyrosines to alanines in the SET domain. However they observed MLL4 protein instability and may have inadvertently created a null. Again the embryos were not investigated except for observing severe early embryonic lethality that could be the same as *Mll4^{Δ/Δ}*.

The *Mll3* and *Mll4* null phenotypes are strikingly different. Embryos lacking MLL3 appeared to develop normally until birth whereupon they died because they failed to breathe, although they gasped (Fig. 1E, Table 1). Embryos lacking MLL4 showed abnormalities during gastrulation and died shortly afterwards (Fig. 1F, Table 2). The earliest observable phenotype in *Mll4^{Δ/Δ}* embryos was growth retardation starting at E6.5 followed by the appearance of a marked constriction at the embryonic/extraembryonic boundary one day later (Fig. 1F). Later in development the *Mll4^{Δ/Δ}* embryos displayed abnormal headfolds, absence of somites and heart beat, and did not turn (Fig. 1F, Table 2). Despite this severe phenotype, *Mll4^{Δ/Δ}* embryos displayed no observable change in global mono-, di- or tri-methylation of H3K4 (Fig. S3). Furthermore, heterozygous *Mll4^{Δ/+}* embryos exhibited abnormalities whereas *Mll3^{Δ/+}* embryos were normal.

Loss of MLL3 results in respiratory failure at birth

Both *Mll3^{D/D}* and *Mll3^{FDC/FDC}* mice died at birth and were indistinguishable (Table 1). Hence we now term both *Mll3^{KO}*. To determine the cause of lethality of *Mll3^{KO}* neonates, we followed natural delivery paying attention not to disturb maternal care. *Mll3^{KO}* neonates quickly became cyanotic and died immediately

after birth or were found dead (Fig. 1E, Table 1). These neonates had normal weight and morphology. The hearts of E15.5 and E18.5 *Mll3*^{KO} fetuses showed no morphological abnormalities and were beating indicating normal fetal circulation (Fig. S4A). Neonatal death by asphyxiation can be caused by a defect of the respiratory rhythm generator in the brain stem, which comprises the retrotrapezoid nucleus/parafacial respiratory group (RTN/pFRG) and pattern generator neurons of the ventrolateral medulla named the pre-Bötzinger complex (preBötC). Knockout of *Jmjd3*, which together with the UTX and UTY subunits of the MLL3/4 complexes are the three known H3K27 demethylases (Van der Meulen et al., 2014), die at birth because the preBötC is absent (Burgold et al., 2012). Therefore we looked for, and found, the preBötC in *Mll3*^{KO} perinatal brain stem (Fig. S4B), excluding this explanation for the failure to breathe. Furthermore, the rib cage and palate were intact and the intercostal muscles as well as the diaphragm were normal in *Mll3*^{KO} and littermate fetuses (Fig. S4C and Fig. S4D; data not shown).

Considering that the structures of palate, diaphragm, intercostal muscles, brain stem and heart were normal, we focused on defects in the lung as the cause of respiratory failure. In mouse lung development the pseudoglandular stage (E10.5-E16.5) is characterized by dichotomous branching of the bronchi resulting in smaller bronchioles in the periphery. The distal tips form terminal sacculles at the canalicular stage (E16.5-17.5). In the saccular stage (E17.5-P5), the alveolar epithelium differentiates into type I and type II alveolar epithelial cells (Bird et al., 2015).

Macroscopic analysis of *Mll3*^{KO} E18.5 lungs revealed mostly a normal morphology with five lobes. Several *Mll3*^{KO} E18.5 fetuses, but not all, had smaller lungs than wild type littermates (Fig. 2A). Hematoxylin and eosin staining showed normal proximal conducting airways (Fig. 2A). Staining with the epithelial marker E-cadherin showed normally developed epithelial lining and the integrity of the epithelium was confirmed by the tight junctional marker ZO-1 (Fig. 2B). The differentiated proximal epithelium consists of neuroendocrine, ciliated and secretory (Clara) cells. Cc10, which is a marker for Clara cells, was reduced in both immunofluorescence staining and qRT-PCR in *Mll3*^{KO} lungs (Fig. 2B, Fig. 2D). At E18.5 *Mll3*^{KO} lungs had thickened alveolar walls and smaller alveolar spaces (Fig. 2A) whereas E15.5 lung

morphology was normal (Fig. S5A), suggesting that MLL3 is required for alveolar formation in the saccular phase of lung development. In this phase columnar distal type II alveolar epithelial cells differentiate and give rise to squamous type I alveolar epithelial cells, which line the alveoli and mediate gas exchange (Treutlein et al., 2016). Type II alveolar epithelial cells produce surfactant consisting of phospholipids and the surfactant proteins (SPs) SP-A, SP-B, SP-C and SP-D, which reduce surface tension at the air-liquid interface and prevent alveolar collapse (Woik et al., 2014). In *MLL3^{KO}* E18.5 lungs transcripts for the surfactant proteins were lower than in control littermates (Fig. 2D). mRNA levels for *Abca3* (ATP-binding cassette sub-family A (ABC1), member 3) which encodes a factor required for lamellar body synthesis in type II alveolar epithelial cells showed a slight increase in *MLL3^{KO}* lungs (Fig. 2D). Lysozyme, another marker for type II alveolar epithelial cells, remained unchanged in immunofluorescence staining (Fig. 2C). Aquaporin 5, a characteristic marker for type I alveolar epithelial cells was reduced at both transcript and protein levels (Fig. 2C, Fig. 2D). From these data we conclude a defect(s) in alveolar epithelial cell differentiation and maturation contributed to the respiratory failure.

The specification of the various lung epithelial cell types requires a balance between differentiation and proliferation (Bellusci et al., 1997). The developing lung at the pseudoglandular stage is characterized by high proliferative activity, which drastically decreases during sacculation (Fig. S5B). Lungs of both genotypes underwent this inhibition of proliferation however to a lesser extent in the *MLL3^{KO}* that correlated with increased septum thickness (Fig. S5B). Concomitantly extracellular matrix (ECM) proteins laminin $\alpha 1$ (*Lama1*) and fibronectin were apparently more prevalent in the ECM of the basement membrane in *MLL3^{KO}* lung (Fig. S5C).

Because MLL3 expression was detected in close proximity to pulmonary vessels (Fig. S6A) and several *MLL3^{KO}* lungs were paler than wildtype controls, we examined whether loss of MLL3 impaired pulmonary circulation via improper development of the lung vasculature. However, staining with von Willebrand Factor (vWF) on *MLL3^{KO}* lung sections revealed intact endothelium and staining the smooth muscle surrounding the blood vessels with alpha smooth muscle actin (α SMA) was also apparently normal (Fig. S6B). We conclude vascular

development was unaffected in lungs of *Mll3*^{KO} mice.

Incompletely penetrant neural tube Mll4 haploinsufficiency with sex distortion

The null alleles of *Mll3* and *Mll4* presented very different phenotypes in another way. No impact of the heterozygous *Mll3* knockout was observed (Table 1), whereas the heterozygous *Mll4* knockout presented a phenotype. Neural fold defects were observed in *Mll4*^{A/+} embryos from E9.5 and further throughout gestation. At E9.5, some *Mll4*^{A/+} embryos had failed to close their neural tube. Later in embryogenesis from E10.5 until E16.5, exencephaly was observed with open cranial neural folds displaying an everted and enlarged appearance (Fig. 3A, Fig. 3B). This disorder was inherited with a parent-of-origin distortion. When only the father was carrier of the *Mll4*^A allele, 11% of the heterozygous embryos developed exencephaly, while 50% of the heterozygous embryos were affected if only the mother passed on the allele (Table 3). As expected, when both parents were heterozygous the number of exencephalic heterozygous embryos was in between these two frequencies at 20% (Table 3). Moreover, two-thirds of all exencephalic embryos were female. Later in gestation, exposed neural folds degenerated, producing anencephaly (data not shown). Embryos with such disorder were found dead at P1 as they were either stillborn or died within the first hours after birth, hence the low recovery rate of *Mll4*^{A/+} pups (Table 3).

Because the *Wnt1* gene is only 40 kb 3' of the *Mll4* gene on mouse chromosome 15, and *Wnt1* is typically expressed in the dorsal midline of the developing hindbrain and spinal cord (Parr et al., 1993), we analyzed its' expression using *in situ* hybridization. As seen in Figure 3C, *Wnt1* is expressed normally in the exencephalic embryo so disturbed *Wnt1* expression is not the cause of the phenotype. We also evaluated *Otx2* expression, which is located to the forebrain and midbrain and characterized by a sharp mid-hindbrain boundary (Ang et al., 1994). It was also unaffected by exencephaly (Fig. 3C).

Adult heterozygous Mll4 mice present aspects of Kabuki syndrome

In humans, all *Mll4* mutations associated with Kabuki syndrome are heterozygous. Having observed an embryonic heterozygous phenotype, we

therefore looked for signs of Kabuki syndrome in viable *Mll4*^{A/+} mice after birth. Surviving *Mll4*^{A/+} pups were analyzed for facial, cranial and skeletal abnormalities but none were observed. However we did find indications of the metabolic problems that are associated with Kabuki syndrome. Significant differences in body weight between wild type and *Mll4*^{A/+} mice of both sexes from the age of four weeks to 21 weeks could be observed. *Mll4*^{A/+} mice of both sexes remained 30% smaller throughout adulthood (Fig. 4A) and showed a decreased amount of white adipose tissue (data not shown). Glucose and insulin tolerance tests were performed in reaction to the decreased body weight of the heterozygous mice *Mll4*^{A/+}. Although *Mll4*^{A/+} mice had the same fasting blood glucose levels as wild type littermates, male *Mll4*^{A/+} mice displayed altered glucose tolerance as their blood glucose levels declined faster and reached their initial values earlier (Fig. 4B). Also when performing insulin tolerance tests male *Mll4*^{A/+} mice showed higher insulin sensitivity compared to wild type mice (Fig. 4C). In contrast female *Mll4*^{A/+} mice displayed a slight impairment in glucose tolerance and no change in insulin tolerance tests (Fig. 4B, Fig. 4C).

***Mll4* is required for specification of the embryonic anterior-posterior axis**

In contrast to the incompletely penetrant heterozygous *Mll4* knockout, all *Mll4*^{A/A} embryos displayed a phenotype similar to that observed in knockouts of *Hnf3β* (*Foxa2*), *Otx2* and *Lim1* (*Lhx1*), in which specification of the anterior-posterior (A-P) embryonic axis is disrupted (Kinder et al., 2001). To examine A-P patterning in *Mll4*^{A/A} embryos, we performed whole-mount *in situ* hybridizations for anterior and posterior molecular markers (Fig. 5, Fig. 6) between developmental stages E6.5 to E7.75.

Lefty1, an early anterior marker expressed at the anterior visceral endoderm (AVE) of E6.5 embryos was absent in *Mll4*^{A/A} embryos. Expression of another AVE marker, *Hex*, was restricted to the distal region of *Mll4*^{A/A} embryos. *Otx2* was prominent at the anterior ectoderm and visceral endoderm of control embryos but was detected at the distal epiblast and the overlying visceral endoderm of *Mll4*^{A/A} embryos (Fig. 5A). *Cerberus1* (*Cer1*), a Bmp4, Nodal and Wnt antagonist is normally expressed at the AVE (Piccolo et al., 1999) but only weakly expressed at the distal region in *Mll4*^{A/A} embryos (Fig.

5B). The Wnt antagonist, *Dkk1*, is typically expressed closest to the embryonic/extraembryonic boundary. However, in *Mll4^{A/A}* embryos expression was observed at the distal region (Fig. 5C).

Since the anterior markers were misexpressed at the distal region of *Mll4^{A/A}* embryos (Fig. 5), we expected defective posterior patterning. *Wnt3* and *Brachyury* are the earliest markers of the primitive streak confined exclusively to the posterior side of control embryos. However, in *Mll4^{A/A}* embryos the expression was seen ectopically at the embryonic/extraembryonic boundary (Fig. 6A, Fig. 6B). *Nodal* is essential for the induction and maintenance of the primitive streak (Conlon et al., 1994). Normally *Nodal* is expressed in the posterior embryonic ectoderm, expanding along the length of the primitive streak. *Nodal* did not extend as far distally in *Mll4^{A/A}* embryos. In control embryos, *Bmp4* and *Eomes* marked the posterior and anterior region of the primitive streak, respectively. In addition, they are markers of the extraembryonic chorion and amnion. In *Mll4^{A/A}* embryos both markers were detected at the embryonic/extraembryonic boundary (Fig. 6C). The patterning of posterior markers in *Mll4^{A/A}* embryos indicates that the primitive streak did not extend distally. Consequently the mesodermal wings were absent (Fig. S7).

At the anterior region of the primitive streak is a distinct region called the node from which the anterior mesendoderm derivatives such as head process and notochord develop (Benazeraf and Pourquie, 2013). Since the primitive streak did not extend till the distal tip of the *Mll4^{A/A}* embryos, we examined if the node and the node derivatives were present in *Mll4^{A/A}* embryos. *Goosecoid* (*Gsc*) was distally induced in control embryos but expression was more confined and proximal in *Mll4^{A/A}* embryos. *Foxa2* and *Lhx1* marked the node and its derivatives in control embryos. In *Mll4^{A/A}* embryos expression for both was seen at the proximal epiblast. *Shh* was expressed exclusively in the head process arising from the node in control embryos but was completely absent in *Mll4^{A/A}* embryos (Fig. 6D). Despite the fact that *Mll4^{A/A}* embryos expressed markers of the node they failed to establish the node derivatives.

At E8.5 *Mll4^{A/A}* mutants were characterized by the absence of *Mox1* transcripts confirming the aforementioned lack of somites (Fig. S7). Very weak and discontinuous localization of *Brachyury* verified that the node derivatives were not specified (Fig. S7). We noticed weak expression of *Hoxb1* suggesting

the presence of precursor cells for rhombomere 4 (Fig. S7).

Mll4 is essential for AVE migration

Shortly after implantation, the first asymmetry in normal embryos is evident as a proximal-distal (P-D) axis (Beddington and Robertson, 1999). At the distal tip, AVE cells undergo a transition from columnar to squamous, protrude filipodia and migrate unidirectionally as a collective towards the future anterior of the embryo until they reach the embryonic/extraembryonic boundary (Srinivas et al., 2004). At E6.5 squamous AVE cells form a single-layer reaching the embryonic/extraembryonic boundary indicated by expression of HEX (Fig. 7A). In *Mll4* deficient embryos, AVE cells failed to reach the embryonic/extraembryonic boundary. Notably, HEX expressing cells retained a cuboidal shape and displayed strong apical actin (Fig. 7A) indicating that the first defect in *Mll4*^{A/A} embryos is a failure to undergo the columnar to squamous transition that precedes migration.

Discussion

Despite the similarities between MLL3 and MLL4 including protein architecture, residence in apparently identical protein complexes, ubiquitous expression and high frequency of mutations in almost all human cancers, their null phenotypes in mouse development are dramatically different. MLL4 is indispensable for the establishment of the A-P axis and progression of gastrulation. MLL3 is not required until the final steps of lung development to ensure neonatal breath at birth. It therefore appears that these two conserved sister proteins are required for very different, highly specific functions in mouse development. In this regard, they are similar to MLL1, which is first required for definitive hematopoiesis at E12.5 with apparently little other contribution to the developing embryo (Ernst et al., 2004). The observation that these highly conserved proteins are only required for a few, very specific, developmental functions does not concur with the prevalent model that MLL3 and MLL4 are the transcription co-factors that deposit the universal epigenetic characteristic of enhancers, H3K4me1 (Lee et al., 2013; Rao and Dou, 2015; Piunti and Shilatifard, 2016). This conundrum could find a resolution if the MLLs are embedded in comprehensive functional backup and their individual knockout phenotypes only reveal flaws in this comprehensive backup. Some evidence for this proposition has been acquired (Lee et al., 2013; Denissov et al., 2014; Chen et al., 2017; Chen et al., 2018).

MLL3 and defective respiration

MLL3^{D/D} and *MLL3^{FDC/FDC}* neonates died due to failures in the final steps of lung maturation. MLL3 is not required for patterning of the lung but is required for (i) efficient differentiation of distal lung epithelium when squamous type I alveolar epithelial cells arise from columnar type II alveolar epithelial cells, and (ii) thinning of the mesenchyme possibly due to sustained proliferative activity during the cannalicular/saccular phase. The concomitant occurrence of defects in two nearby cell types suggest an underlining failure of cell signaling during the final maturation of the lung.

MLL4 and defective AVE migration

Migration of the AVE is the first collective cell migration in the mouse

embryo and precedes gastrulation (Rossant and Tam, 2009) (Fig. 7B). Before migration extraembryonic visceral endoderm cells change their shape from columnar to squamous, which involves cytoskeletal rearrangements and the projection of filopodia towards the direction of migration (Srinivas et al., 2004). At the migratory front of the cell the Rho GTPase, RAC1, is positioned to regulate actin polymerization. Notably, the *Rac1* knockout phenotype is comparable to *Mill4^{Δ/Δ}* characterized by failed AVE migration and lethality before E9.5 (Sugihara et al., 1998; Migeotte et al., 2010). NAP1, a component of the WAVE complex acts downstream of RAC1 to control actin branching and *Nap1* mutants do not establish the A-P axis due to failed AVE polarization and migration (Rakeman and Anderson, 2006). Considering these similarities, we suggest that MLL4 regulates RAC1 and/or the WAVE complex. Another possibility is that MLL4 acts through interaction with UTX, which was found in an unbiased screen to regulate cell migration (Thieme et al., 2013).

In the absence of MLL4, gene expression associated with the AVE including *Hex*, *Dkk1* and *Cer1* was established but mislocalized at the distal region of the embryo due to the absence of migration. This migration is essential for the correct patterning of both the anterior and posterior of the embryo (Stower and Srinivas, 2014). Consequently the failure of the AVE to reach the anterior embryonic/extraembryonic boundary resulted in failed elongation of the primitive streak towards the distal tip indicated by misexpression of *Brachyury*, *Wnt3*, *Bmp4* and *Eomes*. As a result, node markers *Goosecoid*, *Foxa2* and *Lhx1* were restricted to the posterior-proximal region (Fig. 7B). These failures preceded the absence of mesodermal derivatives emerging from the primitive streak

MLL4 haploinsufficiency

MLL3 and MLL4 also differ regarding their heterozygous phenotypes. In our survey of the six H3K4 methyltransferases in mouse development (Glaser et al., 2006; Glaser et al., 2009; Andreu-Vieyra et al., 2010; Bledau et al., 2014; Denissov et al., 2014; Brici et al., 2017; Chen et al., 2017; Hanna et al., 2018), only the knockout of *Mill4* has presented embryonic haploinsufficiency. Notably *Drosophila Trr* also displays haploinsufficiency (Chauhan et al., 2012). Amongst the various explanations for haploinsufficiency, transcriptional

synergy may be relevant for MLL4. Transcriptional synergy is based on cooperative recruitment of transcription factors to cis regulatory elements to achieve a transcriptional output, which involves a threshold and sigmoidal response to protein concentration (Veitia et al., 2018). Furthermore, a contribution by MLL4 to the stability of decisions based on stochastic choices could explain the incomplete penetrance (Cook et al., 1998). Incomplete penetrance in a rate limiting developmental choice pathway implies that MLL4 does not make the choice rather reduces the error rate by either stabilizing the choice or counter-acting mistakes. A primary role for epigenetic regulation in error reduction concords with our recent findings in yeast where Set1C and the H3K4me3 demethylase Jhd2 act together as a quality control mechanism to ensure symmetrically trimethylated nucleosomes (Choudhury et al., 2019).

The frequency of neural tube defects in *MLL4*^{A/+} embryos was affected by the sex of the mutant parent. The defects were more likely when the null allele was transmitted from the mother. Thus it seems likely that MLL4 contributes to oogenesis as well as neurulation. Alternatively, the sex distortion may relate to the sex specific difference between MLL4 complexes, which in females includes only UTX whereas in males both UTX and UTY are involved.

More than 300 genes have been implicated in neural tube closure defects, which are very common in humans estimated at 1 per 1000 fetuses (Juriloff and Harris, 2018), and epigenetic mechanisms involving DNA and histone methylation have emerged as particularly important (Harris and Juriloff, 2010). Notably, folic acid supplementation during pregnancy, which elevates S-Adenosyl Methionine (SAM) levels, diminishes the probability of neural tube closure defects in certain cases (Greene and Copp, 2005). The accompanying proposition that females are more susceptible to neural tube defects due to the increased requirement for SAM in X-chromosome inactivation (Juriloff and Harris, 2000), may be relevant to our observations of sex distorted neural tube closure defects.

Concordant with *MLL4* haploinsufficiency in the mouse, *de novo* heterozygous mutations of *MLL4* are the primary cause of the rare congenital Kabuki syndrome, which involves mental retardation and a distinctive facial appearance. Further features of this phenotypically variable disorder include postnatal dwarfism, heart and kidney dysfunction, skeletal abnormalities, loss

of hearing, gastrointestinal disorders and metabolic imbalances including hypoglycemia (Banka et al., 2012; Banka et al., 2015; Bogershausen et al., 2015; Yap et al., 2019). A very mild version of Kabuki syndrome has been observed in a mouse model based on a heterozygous, hypomorphic *Mll4* allele (Benjamin et al., 2017). Similarly, the *Mll4*^{A/+} phenotype only presents limited aspects of Kabuki syndrome: reduced body weight, stunted growth and hypoglycemia together with reduced body fat. Nevertheless, these haploinsufficient *Mll4* observations indicate that the amount of expressed MLL4 is critical to its function. *Mll4* haploinsufficiency revealed a requirement for MLL4 function in development after its role in the AVE columnar to squamous transition. *Mll4* conditional mutagenesis also revealed later requirements in heart development, myogenesis and adipogenesis (Lee et al., 2013; Ang et al., 2016). Notably, Ang et al (2016) presented evidence of aortic haploinsufficiency.

Because MLL4 is required for the first collective cell migration in mouse development, and because it is required for the cellular migration processes involved in closure of the neural tube, we suggest that MLL4 is a master regulator of cell migration gene expression programs. Although diverse and as yet only partially documented, the evidence for the various MLL4 functions in mouse development are nevertheless still limited to a small group of specific indications. For MLL3, the indications are even more limited and together, these indications do not offer an explanation for the extraordinary prevalence of *Mll3* and *Mll4* mutations in human cancers. The proposition that the MLL system is deeply embedded in functional redundancy and backup may resolve this conundrum. Testing this proposition requires concerted conditional mutagenesis, which is underway.

Acknowledgements

We thank Mandy Obst, Doris Müller, Isabell Kolbe, Madeleine Walker and Stefanie Weidlich for excellent technical assistance. We also thank the Biomedical Services (BMS) of the Max Planck Institute of Molecular Cell Biology and Genetics, Dresden for the excellent service and technical assistance. We thank Dr. Siddharth Banka (University of Manchester, UK) for

discussions and Dr. Tristan A. Rodriguez (Imperial College, London, UK), Dr. Karin Schuster-Gossler (MH Hannover, Germany), Dr. Janet Rossant (University of Toronto, Canada), Dr. Elizabeth Robertson (University of Oxford, UK), Dr. Dominic Norris (MRC Harwell, UK), Dr. Rachel D. Mullen (MD Anderson, Texas, USA), Dr. Hans Schöler (Max Planck Institut Münster, Germany) and Dr. Andrew P. McMahon (University of Southern California, USA) for providing probes for whole mount *in situ* hybridization and Dr. Go Shioi (RIKEN Kobe, Japan) for providing the anti-Hex antibody and advice. The Advanced Imaging Facility, a core facility of the CMCB Technology Platform at TU Dresden, <http://biotp.tu-dresden.de/facilities/advanced-imaging/> assisted this research.

Funding

This work was supported by funding from the Else Kröner-Fresenius-Stiftung (2012_A300 to A.K. and A.F.S.) the Deutsche Forschungsgemeinschaft (KR 2154/6-1 to A.K. and STE 903/12-1 to A.F.S.), the Deutsche Krebshilfe (110560 to A.K. and A.F.S.) and the Scholarship Program for the Promotion of Early-Career Female Scientists of TU Dresden (to D.A.).

Competing interests statement

The authors declare no competing financial interests.

Author contributions

D.A., Q.Z., A.S.B., C.M., A.F.S., K.A and A.K. designed and performed experiments, D.A., Q.Z, A.K., A.S.B. and C.M. examined the embryonic phenotype, A.S.B., J.F., K.A. generated DNA reagents, A.K., D.A. and A.F.S. wrote the manuscript.

References

- Alten, L., Schuster-Gossler, K., Beckers, A., Groos, S., Ulmer, B., Hegermann, J., Ochs, M. and Gossler, A.** (2012). Differential regulation of node formation, nodal ciliogenesis and cilia positioning by Noto and Foxj1. *Development* **139**, 1276-1284.
- Anastassiadis, K., Fu, J., Patsch, C., Hu, S., Weidlich, S., Duerschke, K., Buchholz, F., Edenhofer, F. and Stewart, A. F.** (2009). Dre recombinase, like Cre, is a highly efficient site-specific recombinase in E. coli, mammalian cells and mice. *Dis Model Mech* **2**, 508-515.
- Andreu-Vieyra, C. V., Chen, R., Agno, J. E., Glaser, S., Anastassiadis, K., Stewart, A. F. and Matzuk, M. M.** (2010). MLL2 is required in oocytes for bulk histone 3 lysine 4 trimethylation and transcriptional silencing. *PLoS Biol* **8**.
- Ang, S. L., Conlon, R. A., Jin, O. and Rossant, J.** (1994). Positive and negative signals from mesoderm regulate the expression of mouse Otx2 in ectoderm explants. *Development* **120**, 2979-2989.
- Ang, S. L., Jin, O., Rhinn, M., Daigle, N., Stevenson, L. and Rossant, J.** (1996). A targeted mouse Otx2 mutation leads to severe defects in gastrulation and formation of axial mesoderm and to deletion of rostral brain. *Development* **122**, 243-252.
- Ang, S. Y., Uebersohn, A., Spencer, C. I., Huang, Y., Lee, J. E., Ge, K. and Bruneau, B. G.** (2016). KMT2D regulates specific programs in heart development via histone H3 lysine 4 di-methylation. *Development* **143**, 810-821.
- Ardehali, M. B., Mei, A., Zobeck, K. L., Caron, M., Lis, J. T. and Kusch, T.** (2011). Drosophila Set1 is the major histone H3 lysine 4 trimethyltransferase with role in transcription. *EMBO J* **30**, 2817-2828.
- Banka, S., Lederer, D., Benoit, V., Jenkins, E., Howard, E., Bunstone, S., Kerr, B., McKee, S., Lloyd, I. C., Shears, D., Stewart, H., White, S. M., Savarirayan, R., Mancini, G. M., Beysen, D., Cohn, R. D., Grisart, B., Maystadt, I. and Donnai, D.** (2015). Novel KDM6A (UTX) mutations and a clinical and molecular review of the X-linked Kabuki syndrome (KS2). *Clin Genet* **87**, 252-258.
- Banka, S., Veeramachaneni, R., Reardon, W., Howard, E., Bunstone, S., Ragge, N., Parker, M. J., Crow, Y. J., Kerr, B., Kingston, H., Metcalfe, K., Chandler, K., Magee, A., Stewart, F., McConnell, V. P., Donnelly, D. E., Berland, S., Houge, G., Morton, J. E., Oley, C., Revencu, N., Park, S. M., Davies, S. J., Fry, A. E., Lynch, S. A., Gill, H., Schweiger, S., Lam, W. W., Tolmie, J., Mohammed, S. N., Hobson, E., Smith, A., Blyth, M., Bennett, C., Vasudevan, P. C., Garcia-Minaur, S., Henderson, A., Goodship, J., Wright, M. J., Fisher, R., Gibbons, R., Price, S. M., D, C. d. S., Temple, I. K., Collins, A. L., Lachlan, K., Elmslie, F., McEntagart, M., Castle, B., Clayton-Smith, J., Black, G. C. and Donnai, D.** (2012). How genetically heterogeneous is Kabuki syndrome?: MLL2 testing in 116 patients, review and analyses of mutation and phenotypic spectrum. *Eur J Hum Genet* **20**, 381-388.
- Bannister, A. J. and Kouzarides, T.** (2011). Regulation of chromatin by histone modifications. *Cell Res* **21**, 381-395.
- Beddington, R. S. and Robertson, E. J.** (1999). Axis development and early asymmetry in mammals. *Cell* **96**, 195-209.

- Bellusci, S., Furuta, Y., Rush, M. G., Henderson, R., Winnier, G. and Hogan, B. L.** (1997). Involvement of Sonic hedgehog (Shh) in mouse embryonic lung growth and morphogenesis. *Development* **124**, 53-63.
- Belo, J. A., Bouwmeester, T., Leyns, L., Kertesz, N., Gallo, M., Follettie, M. and De Robertis, E. M.** (1997). Cerberus-like is a secreted factor with neutralizing activity expressed in the anterior primitive endoderm of the mouse gastrula. *Mech Dev* **68**, 45-57.
- Benazeraf, B. and Pourquie, O.** (2013). Formation and segmentation of the vertebrate body axis. *Annu Rev Cell Dev Biol* **29**, 1-26.
- Benjamin, J. S., Pilarowski, G. O., Carosso, G. A., Zhang, L., Huso, D. L., Goff, L. A., Vernon, H. J., Hansen, K. D. and Bjornsson, H. T.** (2017). A ketogenic diet rescues hippocampal memory defects in a mouse model of Kabuki syndrome. *Proc Natl Acad Sci U S A* **114**, 125-130.
- Bird, A. D., McDougall, A. R., Seow, B., Hooper, S. B. and Cole, T. J.** (2015). Glucocorticoid regulation of lung development: lessons learned from conditional GR knockout mice. *Mol Endocrinol* **29**, 158-171.
- Bledau, A. S., Schmidt, K., Neumann, K., Hill, U., Ciotta, G., Gupta, A., Torres, D. C., Fu, J., Kranz, A., Stewart, A. F. and Anastassiadis, K.** (2014). The H3K4 methyltransferase Setd1a is first required at the epiblast stage, whereas Setd1b becomes essential after gastrulation. *Development* **141**, 1022-1035.
- Bogershausen, N., Tsai, I. C., Pohl, E., Kiper, P. O., Beleggia, F., Percin, E. F., Keupp, K., Matchan, A., Milz, E., Alanay, Y., Kayserili, H., Liu, Y., Banka, S., Kranz, A., Zenker, M., Wiczorek, D., Elcioglu, N., Prontera, P., Lyonnet, S., Meitinger, T., Stewart, A. F., Donnai, D., Strom, T. M., Boduroglu, K., Yigit, G., Li, Y., Katsanis, N. and Wollnik, B.** (2015). RAP1-mediated MEK/ERK pathway defects in Kabuki syndrome. *J Clin Invest* **125**, 3585-3599.
- Brici, D., Zhang, Q., Reinhardt, S., Dahl, A., Hartmann, H., Schmidt, K., Goveas, N., Huang, J., Gahurova, L., Kelsey, G., Anastassiadis, K., Stewart, A. F. and Kranz, A.** (2017). Setd1b, encoding a histone 3 lysine 4 methyltransferase, is a maternal effect gene required for the oogenic gene expression program. *Development* **144**, 2606-2617.
- Burgold, T., Voituron, N., Caganova, M., Tripathi, P. P., Menuet, C., Tusi, B. K., Spreafico, F., Bevengut, M., Gestreau, C., Buontempo, S., Simeone, A., Kruidenier, L., Natoli, G., Casola, S., Hilaire, G. and Testa, G.** (2012). The H3K27 demethylase JMJD3 is required for maintenance of the embryonic respiratory neuronal network, neonatal breathing, and survival. *Cell Rep* **2**, 1244-1258.
- Chauhan, C., Zrally, C. B., Parilla, M., Diaz, M. O. and Dingwall, A. K.** (2012). Histone recognition and nuclear receptor co-activator functions of Drosophila cara mitad, a homolog of the N-terminal portion of mammalian MLL2 and MLL3. *Development* **139**, 1997-2008.
- Chen, Y., Anastassiadis, K., Kranz, A., Stewart, A. F., Arndt, K., Waskow, C., Yokoyama, A., Jones, K., Neff, T., Lee, Y. and Ernst, P.** (2017). MLL2, Not MLL1, Plays a Major Role in Sustaining MLL-Rearranged Acute Myeloid Leukemia. *Cancer Cell* **31**, 755-770.
- Chen, Y., Jones, K. L., Anastassiadis, K., Kranz, A., Stewart, A. F., Grembecka, J., Meyerson, M. and Ernst, P.** (2018). Distinct pathways affected by menin versus MLL1/MLL2 in MLL-rearranged acute myeloid leukemia. *Exp Hematol* **69**, 37-42.

- 765 **Chi, P., Allis, C. D. and Wang, G. G.** (2010). Covalent histone
766 modifications--miswritten, misinterpreted and mis-erased in human cancers. *Nat*
767 *Rev Cancer* **10**, 457-469.
- 768 **Choudhury, R., Singh, S., Arumugam, S., Roguev, A. and Stewart, A. F.**
769 (2019). The Set1 complex is dimeric and acts with Jhd2 demethylation to convey
770 symmetrical H3K4 trimethylation. *Genes Dev* **33**, 550-564.
- 771 **Conlon, F. L., Lyons, K. M., Takaesu, N., Barth, K. S., Kispert, A.,**
772 **Herrmann, B. and Robertson, E. J.** (1994). A primary requirement for nodal in
773 the formation and maintenance of the primitive streak in the mouse.
774 *Development* **120**, 1919-1928.
- 775 **Cook, D. L., Gerber, A. N. and Tapscott, S. J.** (1998). Modeling stochastic
776 gene expression: implications for haploinsufficiency. *Proc Natl Acad Sci U S A* **95**,
777 15641-15646.
- 778 **Denissov, S., Hofemeister, H., Marks, H., Kranz, A., Ciotta, G., Singh, S.,**
779 **Anastassiadis, K., Stunnenberg, H. G. and Stewart, A. F.** (2014). Mll2 is
780 required for H3K4 trimethylation on bivalent promoters in embryonic stem cells,
781 whereas Mll1 is redundant. *Development* **141**, 526-537.
- 782 **Dhar, S. S., Lee, S. H., Kan, P. Y., Voigt, P., Ma, L., Shi, X., Reinberg, D.**
783 **and Lee, M. G.** (2012). Trans-tail regulation of MLL4-catalyzed H3K4
784 methylation by H4R3 symmetric dimethylation is mediated by a tandem PHD of
785 MLL4. *Genes Dev* **26**, 2749-2762.
- 786 **Dobson, C. L., Warren, A. J., Pannell, R., Forster, A. and Rabbitts, T. H.**
787 (2000). Tumorigenesis in mice with a fusion of the leukaemia oncogene Mll and
788 the bacterial lacZ gene. *EMBO J* **19**, 843-851.
- 789 **Dorigi, K. M., Swigut, T., Henriques, T., Bhanu, N. V., Scruggs, B. S.,**
790 **Nady, N., Still, C. D., 2nd, Garcia, B. A., Adelman, K. and Wysocka, J.** (2017).
791 Mll3 and Mll4 Facilitate Enhancer RNA Synthesis and Transcription from
792 Promoters Independently of H3K4 Monomethylation. *Mol Cell* **66**, 568-576.
- 793 **Dyle, M. C., Kolakada, D., Cortazar, M. A. and Jagannathan, S.** (2019).
794 How to get away with nonsense: Mechanisms and consequences of escape from
795 nonsense-mediated RNA decay. *Wiley Interdiscip Rev RNA*, e1560.
- 796 **Ernst, P., Fisher, J. K., Avery, W., Wade, S., Foy, D. and Korsmeyer, S. J.**
797 (2004). Definitive hematopoiesis requires the mixed-lineage leukemia gene. *Dev*
798 *Cell* **6**, 437-443.
- 799 **Ernst, P. and Vakoc, C. R.** (2012). WRAD: enabler of the SET1-family of
800 H3K4 methyltransferases. *Brief Funct Genomics* **11**, 217-226.
- 801 **Faundes, V., Malone, G., Newman, W. G. and Banka, S.** (2019). A
802 comparative analysis of KMT2D missense variants in Kabuki syndrome, cancers
803 and the general population. *J Hum Genet* **64**, 161-170.
- 804 **Fu, J., Teucher, M., Anastassiadis, K., Skarnes, W. and Stewart, A. F.**
805 (2010). A recombineering pipeline to make conditional targeting constructs.
806 *Methods Enzymol* **477**, 125-144.
- 807 **Glaser, S., Lubitz, S., Loveland, K. L., Ohbo, K., Robb, L., Schwenk, F.,**
808 **Seibler, J., Roellig, D., Kranz, A., Anastassiadis, K. and Stewart, A. F.** (2009).
809 The histone 3 lysine 4 methyltransferase, Mll2, is only required briefly in
810 development and spermatogenesis. *Epigenetics Chromatin* **2**, 5.
- 811 **Glaser, S., Schaft, J., Lubitz, S., Vintersten, K., van der Hoeven, F.,**
812 **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.

Greene, N. D. and Copp, A. J. (2005). Mouse models of neural tube defects: investigating preventive mechanisms. *Am J Med Genet C Semin Med Genet* **135C**, 31-41.

Hallson, G., Hollebakk, R. E., Li, T., Syrzycka, M., Kim, I., Cotsworth, S., Fitzpatrick, K. A., Sinclair, D. A. and Honda, B. M. (2012). dSet1 is the main H3K4 di- and tri-methyltransferase throughout Drosophila development. *Genetics* **190**, 91-100.

Hanna, C. W., Taudt, A., Huang, J., Gahurova, L., Kranz, A., Andrews, S., Dean, W., Stewart, A. F., Colome-Tatche, M. and Kelsey, G. (2018). MLL2 conveys transcription-independent H3K4 trimethylation in oocytes. *Nat Struct Mol Biol* **25**, 73-82.

Harris, M. J. and Juriloff, D. M. (2010). An update to the list of mouse mutants with neural tube closure defects and advances toward a complete genetic perspective of neural tube closure. *Birth Defects Res A Clin Mol Teratol* **88**, 653-669.

Herrmann, B. G. (1991). Expression pattern of the Brachyury gene in whole-mount TWis/TWis mutant embryos. *Development* **113**, 913-917.

Hoshino, H., Shioi, G. and Aizawa, S. (2015). AVE protein expression and visceral endoderm cell behavior during anterior-posterior axis formation in mouse embryos: Asymmetry in OTX2 and DKK1 expression. *Dev Biol* **402**, 175-191.

Hsu, P. L., Li, H., Lau, H. T., Leonen, C., Dhall, A., Ong, S. E., Chatterjee, C. and Zheng, N. (2018). Crystal Structure of the COMPASS H3K4 Methyltransferase Catalytic Module. *Cell* **174**, 1106-1116.

Jang, Y., Wang, C., Zhuang, L., Liu, C. and Ge, K. (2017). H3K4 Methyltransferase Activity Is Required for MLL4 Protein Stability. *J Mol Biol* **429**, 2046-2054.

Jude, C. D., Climer, L., Xu, D., Artinger, E., Fisher, J. K. and Ernst, P. (2007). Unique and independent roles for MLL in adult hematopoietic stem cells and progenitors. *Cell Stem Cell* **1**, 324-337.

Juriloff, D. M. and Harris, M. J. (2000). Mouse models for neural tube closure defects. *Hum Mol Genet* **9**, 993-1000.

Juriloff, D. M. and Harris, M. J. (2018). Insights into the Etiology of Mammalian Neural Tube Closure Defects from Developmental, Genetic and Evolutionary Studies. *J Dev Biol* **6**.

Kantidakis, T., Saponaro, M., Mitter, R., Horswell, S., Kranz, A., Boeing, S., Aygun, O., Kelly, G. P., Matthews, N., Stewart, A., Stewart, A. F. and Svejstrup, J. Q. (2016). Mutation of cancer driver MLL2 results in transcription stress and genome instability. *Genes Dev* **30**, 408-420.

Kim, J., Kim, J. A., McGinty, R. K., Nguyen, U. T., Muir, T. W., Allis, C. D. and Roeder, R. G. (2013). The n-SET domain of Set1 regulates H2B ubiquitylation-dependent H3K4 methylation. *Mol Cell* **49**, 1121-1133.

Kinder, S. J., Tsang, T. E., Ang, S. L., Behringer, R. R. and Tam, P. P. (2001). Defects of the body plan of mutant embryos lacking Lim1, Otx2 or Hnf3beta activity. *Int J Dev Biol* **45**, 347-355.

Kranz, A., Fu, J., Duerschke, K., Weidlich, S., Naumann, R., Stewart, A. F. and Anastassiadis, K. (2010). An improved Flp deleter mouse in C57Bl/6 based on Flpo recombinase. *Genesis* **48**, 512-520.

Lallemand, Y., Luria, V., Haffner-Krausz, R. and Lonai, P. (1998). Maternally expressed PGK-Cre transgene as a tool for early and uniform activation of the Cre site-specific recombinase. *Transgenic Res* **7**, 105-112.

Lederer, D., Grisart, B., Digilio, M. C., Benoit, V., Crespin, M., Ghariani, S. C., Maystadt, I., Dallapiccola, B. and Verellen-Dumoulin, C. (2012). Deletion of KDM6A, a histone demethylase interacting with MLL2, in three patients with Kabuki syndrome. *Am J Hum Genet* **90**, 119-124.

Lederer, D., Shears, D., Benoit, V., Verellen-Dumoulin, C. and Maystadt, I. (2014). A three generation X-linked family with Kabuki syndrome phenotype and a frameshift mutation in KDM6A. *Am J Med Genet A* **164A**, 1289-1292.

Lee, J. E., Wang, C., Xu, S., Cho, Y. W., Wang, L., Feng, X., Baldridge, A., Sartorelli, V., Zhuang, L., Peng, W. and Ge, K. (2013). H3K4 mono- and dimethyltransferase MLL4 is required for enhancer activation during cell differentiation. *Elife* **2**, e01503.

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 U S A* **103**, 15392-15397.

Li, B. E. and Ernst, P. (2014). Two decades of leukemia oncoprotein epistasis: the MLL1 paradigm for epigenetic deregulation in leukemia. *Exp Hematol* **42**, 995-1012.

Li, Y., Bogershausen, N., Alanay, Y., Simsek Kiper, P. O., Plume, N., Keupp, K., Pohl, E., Pawlik, B., Rachwalski, M., Milz, E., Thoenes, M., Albrecht, B., Prott, E. C., Lehmkuhler, M., Demuth, S., Utine, G. E., Boduroglu, K., Frankenbusch, K., Borck, G., Gillessen-Kaesbach, G., Yigit, G., Wiczorek, D. and Wollnik, B. (2011). A mutation screen in patients with Kabuki syndrome. *Hum Genet* **130**, 715-724.

Li, Y., Han, J., Zhang, Y., Cao, F., Liu, Z., Li, S., Wu, J., Hu, C., Wang, Y., Shuai, J., Chen, J., Cao, L., Li, D., Shi, P., Tian, C., Zhang, J., Dou, Y., Li, G., Chen, Y. and Lei, M. (2016). Structural basis for activity regulation of MLL family methyltransferases. *Nature* **530**, 447-452.

Liu, P., Wakamiya, M., Shea, M. J., Albrecht, U., Behringer, R. R. and Bradley, A. (1999). Requirement for Wnt3 in vertebrate axis formation. *Nat Genet* **22**, 361-365.

Livak, K. J. and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**, 402-408.

Meyer, C., Burmeister, T., Groger, D., Tsaour, G., Fechina, L., Renneville, A., Sutton, R., Venn, N. C., Emerenciano, M., Pombo-de-Oliveira, M. S., Barbieri Blunck, C., Almeida Lopes, B., Zuna, J., Trka, J., Ballerini, P., Lapillonne, H., De Braekeleer, M., Cazzaniga, G., Corral Abascal, L., van der Velden, V. H. J., Delabesse, E., Park, T. S., Oh, S. H., Silva, M. L. M., Lund-Aho, T., Juvonen, V., Moore, A. S., Heidenreich, O., Vormoor, J., Zerkalenkova, E., Olshanskaya, Y., Bueno, C., Menendez, P., Teigler-Schlegel, A., Zur Stadt, U., Lentès, J., Gohring, G., Kustanovich, A., Aleinikova, O., Schafer, B. W.,

Kubetzko, S., Madsen, H. O., Gruhn, B., Duarte, X., Gameiro, P., Lippert, E., Bidet, A., Cayuela, J. M., Clappier, E., Alonso, C. N., Zwaan, C. M., van den Heuvel-Eibrink, M. M., Izraeli, S., Trakhtenbrot, L., Archer, P., Hancock, J., Moricke, A., Alten, J., Schrappe, M., Stanulla, M., Strehl, S., Attarbaschi, A., Dworzak, M., Haas, O. A., Panzer-Grumayer, R., Sedek, L., Szczepanski, T., Caye, A., Suarez, L., Cave, H. and Marschalek, R. (2018). The MLL recombinome of acute leukemias in 2017. *Leukemia* **32**, 273-284.

Migeotte, I., Omelchenko, T., Hall, A. and Anderson, K. V. (2010). Rac1-dependent collective cell migration is required for specification of the anterior-posterior body axis of the mouse. *PLoS Biol* **8**, e1000442.

Mohan, M., Herz, H. M., Smith, E. R., Zhang, Y., Jackson, J., Washburn, M. P., Florens, L., Eissenberg, J. C. and Shilatifard, A. (2011). The COMPASS family of H3K4 methylases in Drosophila. *Mol Cell Biol* **31**, 4310-4318.

Morin, R. D., Mendez-Lago, M., Mungall, A. J., Goya, R., Mungall, K. L., Corbett, R. D., Johnson, N. A., Severson, T. M., Chiu, R., Field, M., Jackman, S., Krzywinski, M., Scott, D. W., Trinh, D. L., Tamura-Wells, J., Li, S., Firme, M. R., Rogic, S., Griffith, M., Chan, S., Yakovenko, O., Meyer, I. M., Zhao, E. Y., Smailus, D., Moksa, M., Chittaranjan, S., Rimsza, L., Brooks-Wilson, A., Spinelli, J. J., Ben-Neriah, S., Meissner, B., Woolcock, B., Boyle, M., McDonald, H., Tam, A., Zhao, Y., Delaney, A., Zeng, T., Tse, K., Butterfield, Y., Birol, I., Holt, R., Schein, J., Horsman, D. E., Moore, R., Jones, S. J., Connors, J. M., Hirst, M., Gascoyne, R. D. and Marra, M. A. (2011). Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. *Nature* **476**, 298-303.

Ng, S. B., Bigam, A. W., Buckingham, K. J., Hannibal, M. C., McMillin, M. J., Gildersleeve, H. I., Beck, A. E., Tabor, H. K., Cooper, G. M., Mefford, H. C., Lee, C., Turner, E. H., Smith, J. D., Rieder, M. J., Yoshiura, K., Matsumoto, N., Ohta, T., Niikawa, N., Nickerson, D. A., Bamshad, M. J. and Shendure, J. (2010). Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. *Nat Genet* **42**, 790-793.

Norris, D. P., Brennan, J., Bikoff, E. K. and Robertson, E. J. (2002). The Foxh1-dependent autoregulatory enhancer controls the level of Nodal signals in the mouse embryo. *Development* **129**, 3455-3468.

Parr, B. A., Shea, M. J., Vassileva, G. and McMahon, A. P. (1993). Mouse Wnt genes exhibit discrete domains of expression in the early embryonic CNS and limb buds. *Development* **119**, 247-261.

Parsons, D. W., Li, M., Zhang, X., Jones, S., Leary, R. J., Lin, J. C., Boca, S. M., Carter, H., Samayoa, J., Bettegowda, C., Gallia, G. L., Jallo, G. I., Binder, Z. A., Nikolsky, Y., Hartigan, J., Smith, D. R., Gerhard, D. S., Fults, D. W., VandenBerg, S., Berger, M. S., Marie, S. K., Shinjo, S. M., Clara, C., Phillips, P. C., Minturn, J. E., Biegel, J. A., Judkins, A. R., Resnick, A. C., Storm, P. B., Curran, T., He, Y., Rasheed, B. A., Friedman, H. S., Keir, S. T., McLendon, R., Northcott, P. A., Taylor, M. D., Burger, P. C., Riggins, G. J., Karchin, R., Parmigiani, G., Bigner, D. D., Yan, H., Papadopoulos, N., Vogelstein, B., Kinzler, K. W. and Velculescu, V. E. (2011). The genetic landscape of the childhood cancer medulloblastoma. *Science* **331**, 435-439.

Pasqualucci, L., Trifonov, V., Fabbri, G., Ma, J., Rossi, D., Chiarenza, A., Wells, V. A., Grunn, A., Messina, M., Elliot, O., Chan, J., Bhagat, G., Chadburn, A., Gaidano, G., Mullighan, C. G., Rabadan, R. and Dalla-Favera, R. (2011).

Analysis of the coding genome of diffuse large B-cell lymphoma. *Nat Genet* **43**, 830-837.

Piccolo, S., Agius, E., Leyns, L., Bhattacharyya, S., Grunz, H., Bouwmeester, T. and De Robertis, E. M. (1999). The head inducer Cerberus is a multifunctional antagonist of Nodal, BMP and Wnt signals. *Nature* **397**, 707-710.

Piette, D., Hendrickx, M., Willems, E., Kemp, C. R. and Leyns, L. (2008). An optimized procedure for whole-mount in situ hybridization on mouse embryos and embryoid bodies. *Nat Protoc* **3**, 1194-1201.

Piunti, A. and Shilatifard, A. (2016). Epigenetic balance of gene expression by Polycomb and COMPASS families. *Science* **352**, aad9780.

Qu, Q., Takahashi, Y. H., Yang, Y., Hu, H., Zhang, Y., Brunzelle, J. S., Couture, J. F., Shilatifard, A. and Skiniotis, G. (2018). Structure and Conformational Dynamics of a COMPASS Histone H3K4 Methyltransferase Complex. *Cell* **174**, 1117-1126.

Rakeman, A. S. and Anderson, K. V. (2006). Axis specification and morphogenesis in the mouse embryo require Nap1, a regulator of WAVE-mediated actin branching. *Development* **133**, 3075-3083.

Rao, R. C. and Dou, Y. (2015). Hijacked in cancer: the KMT2 (MLL) family of methyltransferases. *Nat Rev Cancer* **15**, 334-346.

Rickels, R., Herz, H. M., Sze, C. C., Cao, K., Morgan, M. A., Collings, C. K., Gause, M., Takahashi, Y. H., Wang, L., Rendleman, E. J., Marshall, S. A., Krueger, A., Bartom, E. T., Piunti, A., Smith, E. R., Abshiru, N. A., Kelleher, N. L., Dorsett, D. and Shilatifard, A. (2017). Histone H3K4 monomethylation catalyzed by Trr and mammalian COMPASS-like proteins at enhancers is dispensable for development and viability. *Nat Genet* **49**, 1647-1653.

Riddle, R. D., Johnson, R. L., Laufer, E. and Tabin, C. (1993). Sonic hedgehog mediates the polarizing activity of the ZPA. *Cell* **75**, 1401-1416.

Roguev, A., Schaft, D., Shevchenko, A., Pijnappel, W. W., Wilm, M., Aasland, R. and Stewart, A. F. (2001). The *Saccharomyces cerevisiae* Set1 complex includes an Ash2 homologue and methylates histone 3 lysine 4. *EMBO J* **20**, 7137-7148.

Rossant, J. and Tam, P. P. (2009). Blastocyst lineage formation, early embryonic asymmetries and axis patterning in the mouse. *Development* **136**, 701-713.

Ruthenburg, A. J., Li, H., Patel, D. J. and Allis, C. D. (2007). Multivalent engagement of chromatin modifications by linked binding modules. *Nat Rev Mol Cell Biol* **8**, 983-994.

Schmitges, F. W., Prusty, A. B., Faty, M., Stutzer, A., Lingaraju, G. M., Aiwazian, J., Sack, R., Hess, D., Li, L., Zhou, S., Bunker, R. D., Wirth, U., Bouwmeester, T., Bauer, A., Ly-Hartig, N., Zhao, K., Chan, H., Gu, J., Gut, H., Fischle, W., Muller, J. and Thoma, N. H. (2011). Histone methylation by PRC2 is inhibited by active chromatin marks. *Mol Cell* **42**, 330-341.

Schuettengruber, B., Bourbon, H. M., Di Croce, L. and Cavalli, G. (2017). Genome Regulation by Polycomb and Trithorax: 70 Years and Counting. *Cell* **171**, 34-57.

Skarnes, W. C., Rosen, B., West, A. P., Koutsourakis, M., Bushell, W., Iyer, V., Mujica, A. O., Thomas, M., Harrow, J., Cox, T., Jackson, D., Severin, J., Biggs, P., Fu, J., Nefedov, M., de Jong, P. J., Stewart, A. F. and Bradley, A.

- (2011). A conditional knockout resource for the genome-wide study of mouse gene function. *Nature* **474**, 337-342.
- Slany, R. K.** (2009). The molecular biology of mixed lineage leukemia. *Haematologica* **94**, 984-993.
- Soshnev, A. A., Josefowicz, S. Z. and Allis, C. D.** (2016). Greater Than the Sum of Parts: Complexity of the Dynamic Epigenome. *Mol Cell* **62**, 681-694.
- Srinivas, S., Rodriguez, T., Clements, M., Smith, J. C. and Beddington, R. S.** (2004). Active cell migration drives the unilateral movements of the anterior visceral endoderm. *Development* **131**, 1157-1164.
- Steffen, P. A. and Ringrose, L.** (2014). What are memories made of? How Polycomb and Trithorax proteins mediate epigenetic memory. *Nat Rev Mol Cell Biol* **15**, 340-356.
- Stower, M. J. and Srinivas, S.** (2014). Heading forwards: anterior visceral endoderm migration in patterning the mouse embryo. *Philos Trans R Soc Lond B Biol Sci* **369**.
- Stuckey, D. W., Clements, M., Di-Gregorio, A., Senner, C. E., Le Tissier, P., Srinivas, S. and Rodriguez, T. A.** (2011). Coordination of cell proliferation and anterior-posterior axis establishment in the mouse embryo. *Development* **138**, 1521-1530.
- Sugihara, K., Nakatsuji, N., Nakamura, K., Nakao, K., Hashimoto, R., Otani, H., Sakagami, H., Kondo, H., Nozawa, S., Aiba, A. and Katsuki, M.** (1998). Rac1 is required for the formation of three germ layers during gastrulation. *Oncogene* **17**, 3427-3433.
- Testa, G., Schaft, J., van der Hoeven, F., Glaser, S., Anastassiadis, K., Zhang, Y., Hermann, T., Stremmel, W. and Stewart, A. F.** (2004). A reliable lacZ expression reporter cassette for multipurpose, knockout-first alleles. *Genesis* **38**, 151-158.
- Thieme, S., Gyarmas, T., Richter, C., Ozhan, G., Fu, J., Alexopoulou, D., Muders, M. H., Michalk, I., Jakob, C., Dahl, A., Klink, B., Bandola, J., Bachmann, M., Schrock, E., Buchholz, F., Stewart, A. F., Weidinger, G., Anastassiadis, K. and Brenner, S.** (2013). The histone demethylase UTX regulates stem cell migration and hematopoiesis. *Blood* **121**, 2462-2473.
- Treutlein, B., Lee, Q. Y., Camp, J. G., Mall, M., Koh, W., Shariati, S. A., Sim, S., Neff, N. F., Skotheim, J. M., Wernig, M. and Quake, S. R.** (2016). Dissecting direct reprogramming from fibroblast to neuron using single-cell RNA-seq. *Nature* **534**, 391-395.
- Van der Meulen, J., Speleman, F. and Van Vlierberghe, P.** (2014). The H3K27me3 demethylase UTX in normal development and disease. *Epigenetics* **9**, 658-668.
- Veitia, R. A., Caburet, S. and Birchler, J. A.** (2018). Mechanisms of Mendelian dominance. *Clin Genet* **93**, 419-428.
- Wang, C., Lee, J. E., Lai, B., Macfarlan, T. S., Xu, S., Zhuang, L., Liu, C., Peng, W. and Ge, K.** (2016). Enhancer priming by H3K4 methyltransferase MLL4 controls cell fate transition. *Proc Natl Acad Sci U S A* **113**, 11871-11876.
- Weirich, S., Kudithipudi, S., Kycia, I. and Jeltsch, A.** (2015). Somatic cancer mutations in the MLL3-SET domain alter the catalytic properties of the enzyme. *Clin Epigenetics* **7**, 36.

Woik, N., Dietz, C. T., Schaker, K. and Kroll, J. (2014). Kelch-like ECT2-interacting protein KLEIP regulates late-stage pulmonary maturation via Hif-2alpha in mice. *Dis Model Mech* **7**, 683-692.

Yagi, H., Deguchi, K., Aono, A., Tani, Y., Kishimoto, T. and Komori, T. (1998). Growth disturbance in fetal liver hematopoiesis of Mll-mutant mice. *Blood* **92**, 108-117.

Yap, K. L., Johnson, A. E. K., Fischer, D., Kandikatla, P., Deml, J., Nelakuditi, V., Halbach, S., Jeha, G. S., Burrage, L. C., Bodamer, O., Benavides, V. C., Lewis, A. M., Ellard, S., Shah, P., Cody, D., Diaz, A., Devarajan, A., Truong, L., Greeley, S. A. W., De Leo-Crutchlow, D. D., Edmondson, A. C., Das, S., Thornton, P., Waggoner, D. and Del Gaudio, D. (2019). Congenital hyperinsulinism as the presenting feature of Kabuki syndrome: clinical and molecular characterization of 9 affected individuals. *Genet Med* **21**, 233-242.

Zhang, Y., Mittal, A., Reid, J., Reich, S., Gamblin, S. J. and Wilson, J. R. (2015). Evolving Catalytic Properties of the MLL Family SET Domain. *Structure* **23**, 1921-1933.

Figure legends

Fig. 1 Allele design and embryonic phenotype of the *Mll3* and *Mll4* mutant embryos.

(A) Amino acid sequence alignment of murine MLL3 and MLL4 based on EMBOSS Needle. FYRN and FYRC = F/Y-rich N- or C-terminus, HMG = high mobility group, ePHD = extended PHD finger. (B) Diagram of the *Mll3* knockout first allele (*Mll3^A*). The *Mll3* wild type (wt) allele contains 59 coding exons. The targeted allele, *Mll3^A*, is similar to the Dre recombined allele, *Mll3^D*, with the PGK-Blasticidin selection cassette removed. The *Mll3^D* allele is converted to conditional (*Mll3^{FD}*) upon FLP recombination. Cre recombination leads to excision of the frameshifting exon 49 generating the conditional mutant allele (*Mll3^{FDC}*). Blue rectangles with numbers on top indicate exons. The star depicts the exon with the premature stop codon. SA = splice acceptor, BSD = blasticidin, PGK = phosphoglycerate kinase-1 promoter, pA = polyadenylation signal, lacZ-neo = β -galactosidase and neomycin resistance gene. (C) Diagram of the *Mll4* knockout first allele (*Mll4^A*). The *Mll4* wild type (wt) allele contains 57 coding exons. For abbreviations see Fig. 1B. (D) Loss of MLL4 protein confirmed by Western blot on two *Mll4^{A/A}* ESC clones. Immunoblot analysis of extracts from wild type and *Mll4^{A/+}* ESC clones detected a signal larger than 460 kDa in agreement with the estimated molecular weight of 600 kDa for MLL4 protein. (E) Dissections from *Mll3^{FDC/+}* intercrosses at different stages of development. The homozygous embryos are not distinguishable from their wild type and heterozygous littermates. Scale bar for E9.5 1 mm, for E16.5 2 mm and for P0 2 mm. (F) *Mll4^{A/A}* embryos show growth retardation and a visible constriction is formed at the embryonic/extraembryonic boundary (marked by arrowheads). Arrows point at the neural fold. Asterisk marks the allantois. Scale bar 250 μ m.

Fig. 2. Loss of MLL3 causes defects in lung development.

(A) Analysis of gross morphology and H&E staining of *Mll3^{KO}* and control littermate lung at E18.5. Scale bar 500 μ m (whole lung) and 250 μ m (histology).

(B) E-cadherin expression in the proximal lung epithelium and ZO-1, a tight junction marker, remain unchanged in control and *Mll3*^{KO} littermates. However, expression of cc10 is reduced in *Mll3*^{KO} when compared to control littermates. Scale bar 25 μ m. (C) Lysozyme expression, a characteristic marker for type II alveolar epithelial cells is comparable in *Mll3*^{KO} and control littermate lung. Expression of Aquaporin 5, a marker for type I alveolar epithelial cells is reduced in *Mll3*^{KO} when compared to control littermates. Scale bar 25 μ m. (D) qRT-PCR analysis of proximal and distal airway cell type markers between control and *Mll3*^{KO} lungs at E18.5. Expression levels were normalized to *Rpl19* and plotted as fold change relative to control. Mean \pm s.d. is shown (*P < 0.05 and **P < 0.01 as calculated by unpaired t-test).

Fig. 3. Neural tube closure defects in *Mll4*^{A/+} embryos.

(A) β -galactosidase staining of heterozygous *Mll4*^{A/+} embryos at E10.5. The embryo on the right exhibits exencephaly. Scale bar 1 mm. (B) H&E-stained sagittal sections of E13.5 embryos show the everted and enlarged exposed neural folds. Scale bar 1 mm. (C) *In situ* hybridization analysis for *Otx2* and *Wnt1* expression at E9.5. The expression of these markers was not affected in exencephalic *Mll4*^{A/+} embryos compared to wild type littermates. Scale bar 250 μ m.

Fig. 4. Decreased body weight and hypoglycemia of *Mll4*^{A/+} mice.

(A) Development of body weight over 4-20 weeks for a total of 48 mice (24 males: 15 wild type and 9 heterozygous mice; 24 females: 16 wild type and 8 heterozygous mice) is shown. The differences in weight are significant at any developmental stage. Mean \pm s.d. is shown (**P < 0.01 and ***p < 0.001 as calculated by unpaired t-test). (B) For the glucose tolerance test (GTT), in total 30 mice (14 males: 7 wild type and 7 heterozygous mice; 16 females: 8 wild type and 8 heterozygous mice) at an age of 9 to 15 weeks have fasted for 16 h, until a glucose solution (1.5 mg/g body weight) was orally applied. Blood glucose levels were measured over a period of 120 min. Mean \pm s.d. is shown (*P < 0.05 as calculated by unpaired t-test). (C) Insulin tolerance tests (ITT) after 6 h of fasting were performed with 45 mice (30 males: 14 wild type and 16

heterozygous mice; 15 females: 7 wild type and 8 heterozygous mice) between the age of 12 and 45 weeks. Insulin (0.75 mU/g body weight) was injected intraperitoneally. Blood glucose levels were measured within a time frame of 120 min. Mean \pm s.d. is shown (*P<0.05 and **p <0.01 as calculated by unpaired t-test).

Fig. 5. *Mll4*^{A/A} embryos show defective patterning of the AVE from E6.5 to E7.5.

(A) *Lefty1* (n=5/5), *Hex* (n=5/6), *Otx2* (n=2/3) expression in control and *Mll4*^{A/A} embryos between E6.5 and E7.75. All embryos are oriented with the anterior to the left. Scale bar 250 μ m. (B) (C) *Cer1* (n=4/4 for E6.5 and n=5/5 for E7.5) and *Dkk1* (n=2/2 for E6.5 and n=2/4 for E7.5) expression in control and *Mll4*^{A/A} embryos between E6.5 and E7.5. All embryos are oriented with the anterior to the left. Scale bar 250 μ m.

Fig. 6. Ectopic expression of primitive streak and node markers in *Mll4*^{A/A} embryos.

(A) (B) *Wnt3* (n=1/1 for E6.5 and n=2/4 for E7.5) and *Brachyury* (n=1/1 for E6.5 and n=3/3 for E7.5) marks the length of primitive streak (PS) as it forms in control embryos from E6.5 to E7.5, but in *Mll4*^{A/A} embryos they are expressed at the embryonic/extraembryonic boundary and do not extend till the distal tip. All embryos are oriented with the anterior to the left. Scale bar 250 μ m. (C) In control embryos, *Nodal* (n=4/7), *Bmp4* (n=3/3) and *Eomes* (n=6/7) are expressed at the PS, posterior PS and anterior PS, respectively. In *Mll4*^{A/A} embryos the expression is restricted to the embryonic/extraembryonic boundary. All embryos are oriented with the anterior to the left. Scale bar 250 μ m. (D) *Gsc* (n=2/2), *Foxa2* (n=3/4) and *Lhx1* (n=3/4) are expressed at the node and node derivatives of control embryos but are proximally expressed in *Mll4*^{A/A} embryos. The node derivative marker *Shh* (n=3/3) is completely absent in *Mll4*^{A/A} embryos. All embryos are oriented with the anterior to the left. Scale bar 250 μ m.

Fig. 7. Model for MLL4 function in establishing A-P patterning.

(A) Epithelial organization of embryos at E6.5 of control and *MLL4^{Δ/Δ}* genotype. Individual confocal sections of embryos stained for F-actin (red) and Hex (green) (n=5/5). The boxed region is magnified on the right. Scale bar 50 μm. (B) In control embryos, the AVE (green) migrates to the anterior, restricting the primitive streak to the posterior (red). Thus, the A-P axis is established and gastrulation progresses. In the absence of MLL4, the AVE is at the distal region, causing the primitive streak to form at the embryonic/extraembryonic boundary. The embryo thus undergoes defective gastrulation resulting in embryonic lethality indicating a functional role of MLL4 in AVE migration.

Fig. S1. Gene targeting strategy for *MLL3*.

(A) Schematic representation of the Southern strategy to identify correctly targeted events in the *MLL3* locus using 5' (light blue box) and 3' (red box) external probes. Blue boxes with numbers underneath indicate exons. (B) Southern blot analysis using 5' and 3' external probes. Clones #18, #22, #23 and #27 carried the correctly targeted event. (C) Wild type and *MLL3^{D/+}* embryos at E9.5 and E11.5 were stained for β-galactosidase. Scale bar 500 μm. (D) Diagram of the *MLL3* knockout first allele (*MLL3^A*) with floxed exon 34. Instead of the PGK-Blasticidin-polyA a PGK-Hygromycin-PolyA cassette was inserted. For abbreviations see Fig. 1B.

Fig. S2. Gene targeting strategy for *MLL4*.

(A) Schematic representation of the Southern blot strategy to identify correctly targeted events in the *MLL4* locus using 5' (light blue box), 3' (red box) external probes and the internal LacZ (green box) probe. Blue rectangles with numbers underneath indicate exons. (B) Southern blot analysis using 5' and 3' external probes and LacZ internal probe. Only clone #24 is shown. (C) Wild type and *MLL4^{Δ/+}* embryos at E6.5, E7.5 and E8.5 were stained for β-galactosidase. Scale bar 250 μm.

Fig. S3. No global changes in H3K4 methylation levels.

Immunohistochemistry on sections of wild type and *Mll4^{Δ/Δ}* E7.75 embryos *in utero* with H3K4me1-, H3K4me2- and H3K4me3-specific antibodies. The boxed region is magnified on the right. Scale bars 50 μm and 200 μm, respectively.

Fig. S4. No morphological differences in *Mll3* mutant fetuses.

(A) Four chamber view of the fetal heart at E15.5 (top), morphology of the fetal heart at E18.5 (down). ra = right atrium, la = left atrium, rv = right ventricle, lv = left ventricle, se = intraventricular septum, ao = aorta. Scale bar 500 μm. (B) Overview of the brain stem at E15.5 (top) and nuclear enriched region of pre-Bötzinger complex and nucleus ambiguus (down) stained with Nissl. The size of these two regions was comparable among all genotypes. M = medulla, nA = nucleus ambiguus, preBötC = pre-Bötzinger complex. (C) Morphology of the rib cage after Alizarin red/Alcian blue staining at E18.5. Scale bar 5 mm. (D) Ventral view of the upper portion of the head showing the roof of the mouth at E18.5. The palate is intact in fetuses of different genotype.

Fig. S5. Increased proliferation and extensive deposition of extracellular matrix proteins in *Mll3^{KO}* lungs.

(A) H&E staining on sections of E15.5 lungs showed no morphological differences between control and *Mll3^{KO}* littermates. Sox9 expression is seen in the distal epithelial tips of control and *Mll3^{KO}* lungs. Scale bar 250 μm. (B) No difference in proliferation was observed between control and *Mll3^{KO}* lungs of E15.5 fetuses after PH3 staining (top). At E18.5 an increase in proliferation is observed in *Mll3^{KO}* lungs (bottom). Scale bar 250 μm. PH3 positive cells from E18.5 lung sections from control (n=3) and *Mll3^{KO}* (n=3) lungs were counted. Four defined areas (each 1.8 mm²) were analysed. Mean ± s.d. is shown (**p < 0.01 as calculated by unpaired t-test). (C) Extracellular matrix protein deposition of fibronectin and Lama1 in the basal lamina of *Mll3^{KO}* lung is more compared to control littermate. Scale bar 25 μm.

Fig. S6. Normal lung vasculature at E18.5.

(A) β -galactosidase staining reflecting MLL3 expression on whole lung (top) and lung sections (bottom). MLL3 expression in close proximity to pulmonary vessels demonstrated by β -galactosidase staining on wildtype and *MLL3^{D/D}* lung sections. Scale bar 500 μ m (whole lung) and 250 μ m (lung section). (B) Expression of vWF in endothelial cells of the large blood vessels in control and *MLL3^{KO}* lungs. Smooth muscle surrounding the large blood vessels marked by expression of α SMA in control and *MLL3^{KO}* lungs. Scale bar 25 μ m.

Fig. S7. Defective gastrulation after loss of MLL4.

Whole mount *in situ* hybridization of *Tbx6* (mesodermal wing marker) at E7.75, *Mox1* (somite marker) at E8.5, *Brachyury* at E8.5, and *Hoxb1* at E9.0 on embryos from *MLL4^{A/+}* intercrosses. Arrows point towards the prospective rhombomere 4. Scale bars 250 μ m.

Table 1. Genotyping of progeny from MII3 intercrosses

MIl3^{D/+} x MII3^{D/+} (exon 49 floxed)					
Age	+/+	D/+	D/D	Total (n = litter)	Litter size
P0	18.75%	56.25%	25% [†]	16 (2)	8.0
E18.5	25.3%	43.4%	31.3%	77 (10)	7.5
E17.5	10.0%	80.0%	10.0%	10 (1)	10.0
E16.5	10.0%	70.0%	20.0%	10 (1)	10.0
E13.5	25.0%	62.5%	12.5%	8 (1)	8.0
E11.5	42.9%	57.1%	0.0%	7 (1)	7.0
E10.5	17.7%	64.7%	17.7%	17 (2)	8.5
E7.5	11.1%	44.4%	44.4%	18 (2)	9.0
MIl3^{FDC/+} x MII3^{FDC/+} (exon 49 floxed)					
Age	+/+	FDC/+	FDC/FDC	Total (n = litter)	Litter size
P0	33.3%	55.6%	11.1% [†]	18 (2)	9.0
E18.5	20.0%	56.9%	23.1%	65 (9)	7.2
E17.5	17.6%	58.8%	23.5%	17 (2)	8.5
E16.5	11.1%	66.7%	22.2%	9 (1)	9.0
E13.5	22.2%	44.4%	33.3%	18 (2)	9.0
E11.5	11.1%	66.7%	22.2%	9 (1)	9.0
E10.5	50.0%	50.0%	0.0%	8 (1)	8.0
E9.5	22.2%	44.4%	33.3%	9 (1)	9.0
MIl3^{D/+} x MII3^{D/+} (exon 34 floxed)					
Age	+/+	D/+	D/D	Total (n = litter)	Litter size
P0	37.5%	56.3%	6.3% [†]	16 (2)	8.0
E18.5	17.6%	76.5%	5.9%	17 (3)	5.7
E17.5	25%	62.5%	12.5%	10 (1)	10.0
E15.5	28.6%	42.9%	28.6%	7 (1)	7.0
E13.5	25.0%	50%	25.0%	8 (1)	8.0

† indicates dead fetuses

Table 2. Embryonic lethality of MII4^{A/A} embryos

MII4 ^{A/+} X MII4 ^{A/+}							
Age	+/+	A/+	A/A	Resorbed	Exen- cephaly	Total (n = litter)	Litter size
Weaned	59.1 %	40.9 %	0.0 %			22 (8)	2.8
E11.5	10.8 %	48.6 %	0.0 %	40.5 %	10.8 %	37 (4)	9.3
E10.5	21.5 %	44.6 %	13.8 % [†]	20.0%	7.7 %	65 (6)	10.8
E9.5	23.0 %	45.9 %	16.2 % [†]	14.9 %	9.5 %	148 (14)	10.6
E8.5	22.9 %	45.9 %	17.8 %	13.4 %	-	314 (27)	11.6
E7.5	22.1 %	46.3 %	18.9 %	12.6 %	-	95 (7)	13.6
E6.5	18.8 %	50.0 %	12.5 %	18.8 %	-	16 (2)	8.0

[†] indicates dead embryos

Table 3. Exencephaly of MII4^{A/+} embryos is inherited in a parent-of-origin dependent manner

♂ MII4 ^{A/+} X ♀ CD1							
Age	+/+	A/+	A/A	Resorbed	Exen- cephaly	Total (n = litter)	Litter size
Weaned	82.1 %	17.9 %				1508 (181)	8.3
E18.5	53.7 %	46.3 %	0.0 %	0.0 %	12.0 %	54 (4)	13.5
E16.5	40.0 %	60.0 %	0.0 %	0.0 %	22.2 %	15 (1)	15.0
E15.5	46.2 %	50.0 %	0.0 %	3.8 %	0.0 %	26 (2)	13.0
E14.5	75.0 %	25.0 %	0.0 %	0.0 %	0.0 %	12 (1)	12.0
E13.5	43.9 %	54.4 %	0.0 %	1.8 %	12.9 %	57 (4)	14.3
E11.5	47.1 %	52.9 %	0.0 %	0.0 %	11.1 %	17 (1)	17.0
E10.5	51.2 %	46.4 %	0.0 %	2.4 %	17.9 %	84 (6)	14.0
E9.5	44.8 %	51.2 %	0.0 %	4.0 %	7.8 %	125 (9)	13.9
E8.5	23.1 %	76.9 %	0.0 %	0.0 %	-	13 (1)	13.0

♂ CD1 X ♀ MII4 ^{A/+}							
Age	+/+	A/+	A/A	Resorbed	Exen- cephaly	Total (n = litter)	Litter size
Weaned	77.2 %	22.8 %	0.0 %			114 (20)	5.7
E11.5	58.3 %	25.0 %	0.0 %	16.7 %	66.7 %	12 (1)	12.0
E10.5	28.0 %	52.0 %	0.0 %	20.0 %	30.8 %	25 (2)	12.5
E9.5	36.8 %	42.1 %	0.0 %	21.1 %	75.0 %	19 (2)	9.5

Table S1. Primers for genotyping

Primer pairs	Sequence (5' – 3')	Product in bp
Genotyping of MII3D knockout		
MIl3 LoxP1	CAAATCCTCCCAGATTAGCAAC	361 (wt);
MIl3 LoxP2	CAACAATGCACTATCTATGGCTTC	499 (D)
Genotyping of MII3FD		
MIl3 flp se	GAGAGTGAGTGAGAGAAAACC	765 (wt);
MIl3 ex49 as	AGGATTTGCAGATCCTGGTGG	933 (FD)
Genotyping of MII3 wild type		
MIl3 ex49 se	CCACCAGGATCTGCAAATCCT	374 (wt);
MIl3 LoxP2	CAACAATGCACTATCTATGGCTTC	
Genotyping of MII3FDC knockout		
MIl3 flp se	GAGAGTGAGTGAGAGAAAACC	1106 (wt);
MIl3 LoxP2	CAACAATGCACTATCTATGGCTTC	1423 (FD); 298 (FDC)
Genotyping of Cre		
19Cre se	GCCTGCATTACCGGTGCGATGCAACGA	700 (Cre)
20Cre as	GTGGCAGATGGCGCGGCAACACCATT	
Genotyping of MII4 knockout (embryos)		
MIl4_5lox	ACTATGTGGTGTGGACAAGGCC	169 (wt);
MIl4_3lox	CGCTGCAAGGATGGGCAGAGCC	297 (targeted)
Genotyping of MII4 knockout (tails)		
MIl4_se	CCAGAACTATGTGGTGTGGAC	140 (wt);
MIl4_as	GCTTAGCCAAGGTTAGCCACT	268 (targeted)
Sex genotyping		
Ube1x se	TGGTCTGGACCCAAACGCTGTCCACA	217 (female);
Ube1x as	GGCAGCAGCCATCACATAATCCAGATG	198 (male)

Table S2: Primers for qRT-PCR

Primer pairs	Sequence (5' – 3')	Product size (bp)
Abca3 se	ATCACCTCAGGGGATGCCTTT	150
Abca3 as	TCGGAGCCGTGCATACATAAC	
Aqp5 se	GTGGTCATGAATCGGTTTCAGC	167
Aqp5 as	CAGTCCTCCTCTGGCTCATAT	
Rpl19 se	CCGAGAGAACACCAAAAGGA	147
Rpl19 as	CTTCTCAGGCATCCGAGCATT	
cc10 se	CTGAAGAGACTGGTGGATACC	156
cc10 as	GCAGTGACAAGGCTTTAGCAG	
SP-A se	GAGAACATGGAGACAAGGGAG	146
SP-A as	CATGGATCCTTGCAAGCTGAG	
SP-B se	CAGTGAACAGGCTATGCCACA	211
SP-B se	TCAGAGGTGTGGGGTTTGGAA	
SP-C se	AGCAAAGAGGTCCTGATGGAG	143
SP-C as	CCTACAATCACCGACGACAACG	
SP-D se	CCCTAGTCATGTGTAGCCCAA	174
SP-D as	CAGCAGAGCCATTCTCTCCTT	

Table S3: Antibodies for Western blot (WB), immunofluorescence (IF) and immunohistochemical (IHC) staining

Antibodies	Company, Cat. Number	Dilution
Primary antibodies		
Rabbit anti-Hex	Dr. Go Shioi	1:2000 IF
Rabbit anti-cc10	Santa Cruz, sc-25554	1:100 IF
Rabbit anti-MII4	Diagenode, C15310100	1:500 WB
Rabbit anti-H3K4me1	Diagenode, C15410037	1:2000 IHC
Rabbit anti-H3K4me2	Diagenode, pAb-035-050	1:2500 IHC
Rabbit anti-H3K4me3	Abcam, ab8580	1:1000 IHC
Rabbit anti-phospho-H3S10	Upstate, 06-570	1:500 IF
Rabbit anti-Aquaporin 5	Abcam, ab78486	1:800 IF
Rabbit anti-Fibronectin	Abcam, ab23750	1:500 IF
Rabbit anti-Lama1	Sigma, L9393	1:100 IF
Rabbit anti- α SMA	Abcam, ab5694	1:100 IF
Rabbit anti-Lysozyme	Dako, A0099	1:500 IF
Rabbit anti-Sox9	Millipore, AB5535	1:500 IF
Rabbit anti-von Willebrand Factor	Abcam, ab9378	1:100 IF
Mouse anti-ZO-1	Thermo Scientific, 33-9100	1:500 IF
Rat anti-E-cadherin	Invitrogen, 13-1900	1:50 IF
Secondary antibodies		
Biotinylated goat anti-Rabbit IgG	Vector, BA-1000	1:500 IHC
Goat anti-Rabbit IgG-HRP	Thermo Scientific, 31460	1:20000 WB
Goat anti-Rabbit IgG-CFL 488	Santa Cruz, sc-362262	1:500 IF
Goat anti-mouse IgG Alexa Fluor 488	Invitrogen, A11029	1:500 IF
Donkey anti-rat IgG Alexa Fluor 488	Invitrogen, A21208	1:500 IF

Figure 1

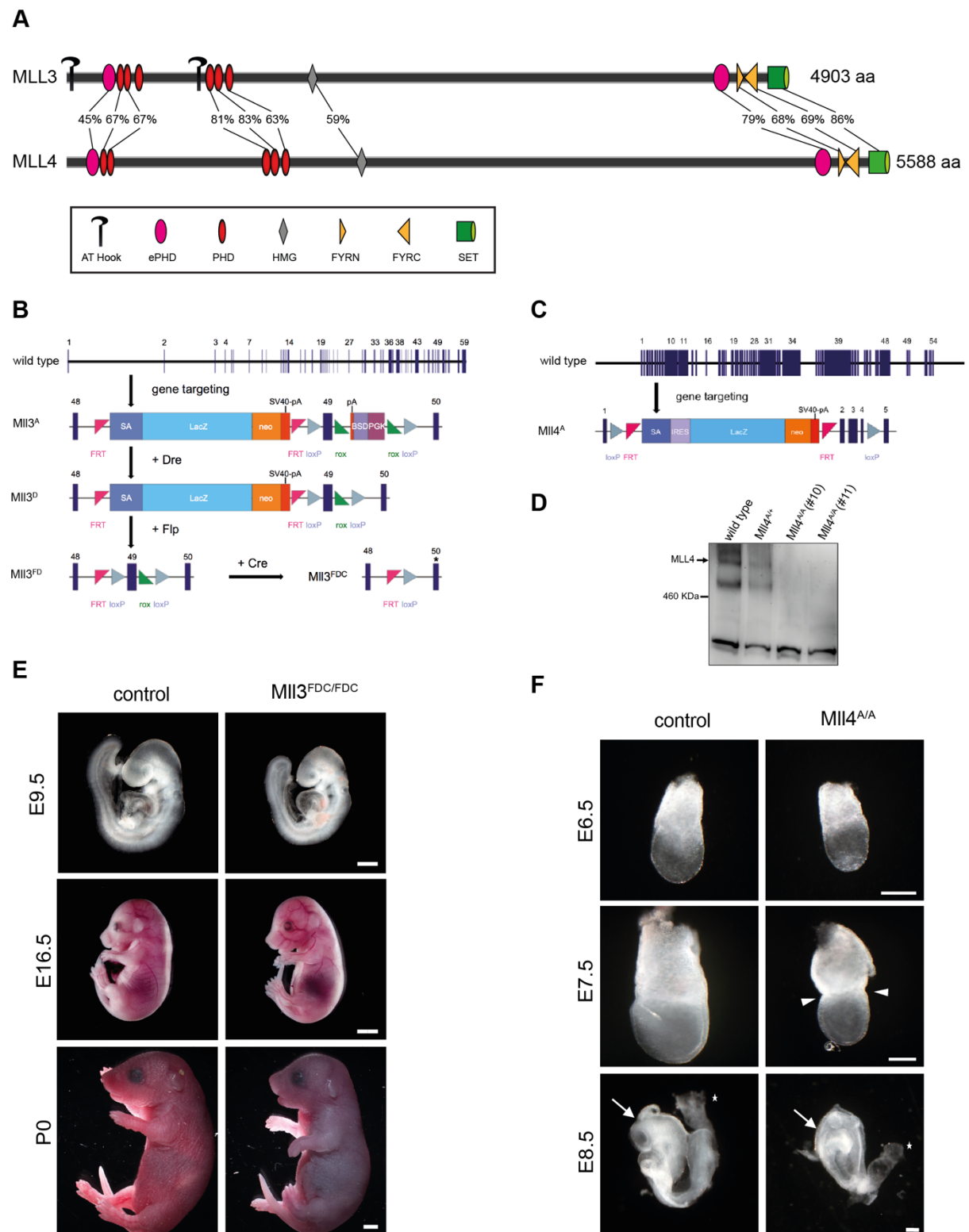


Figure 2

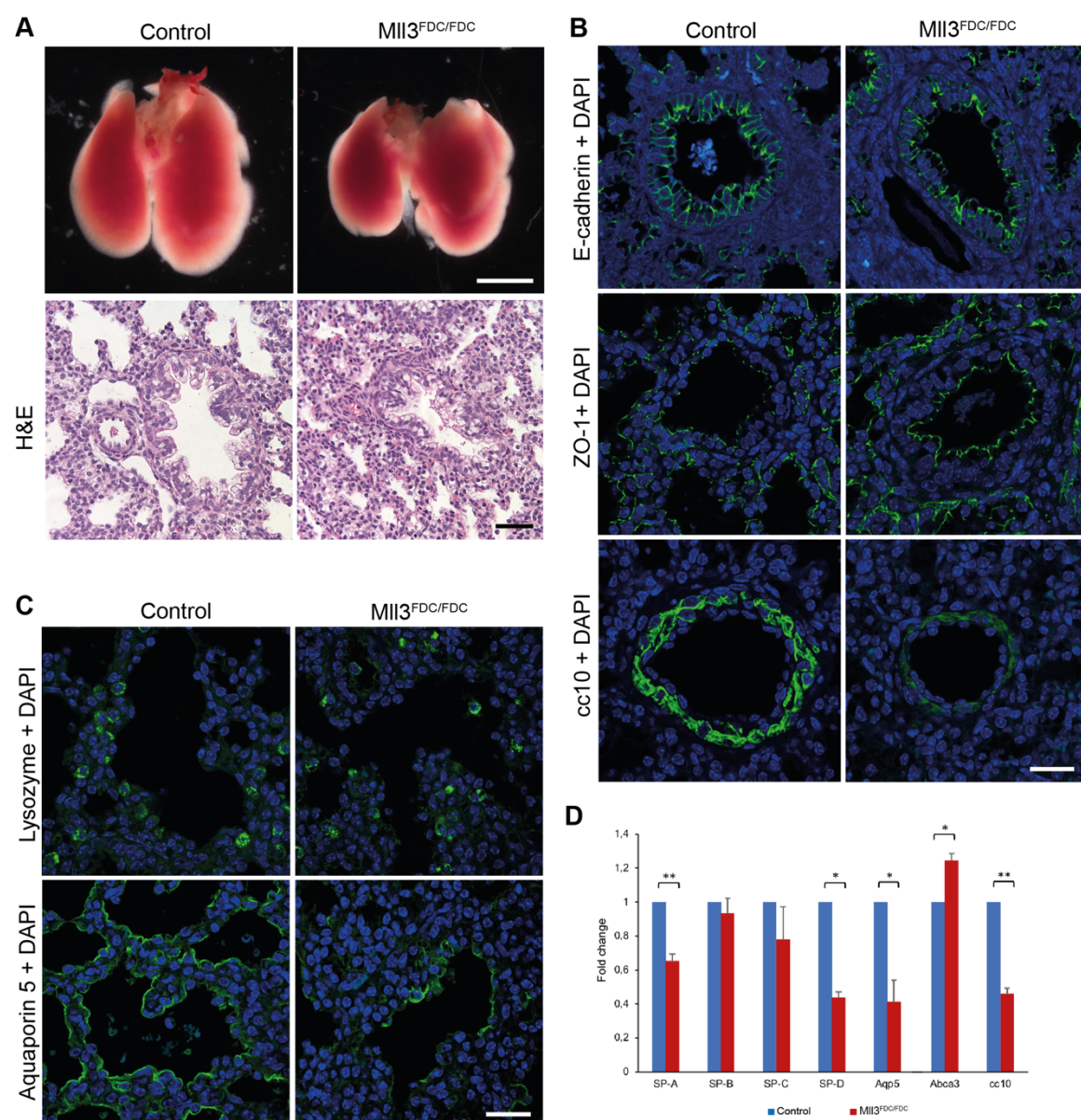


Figure 3

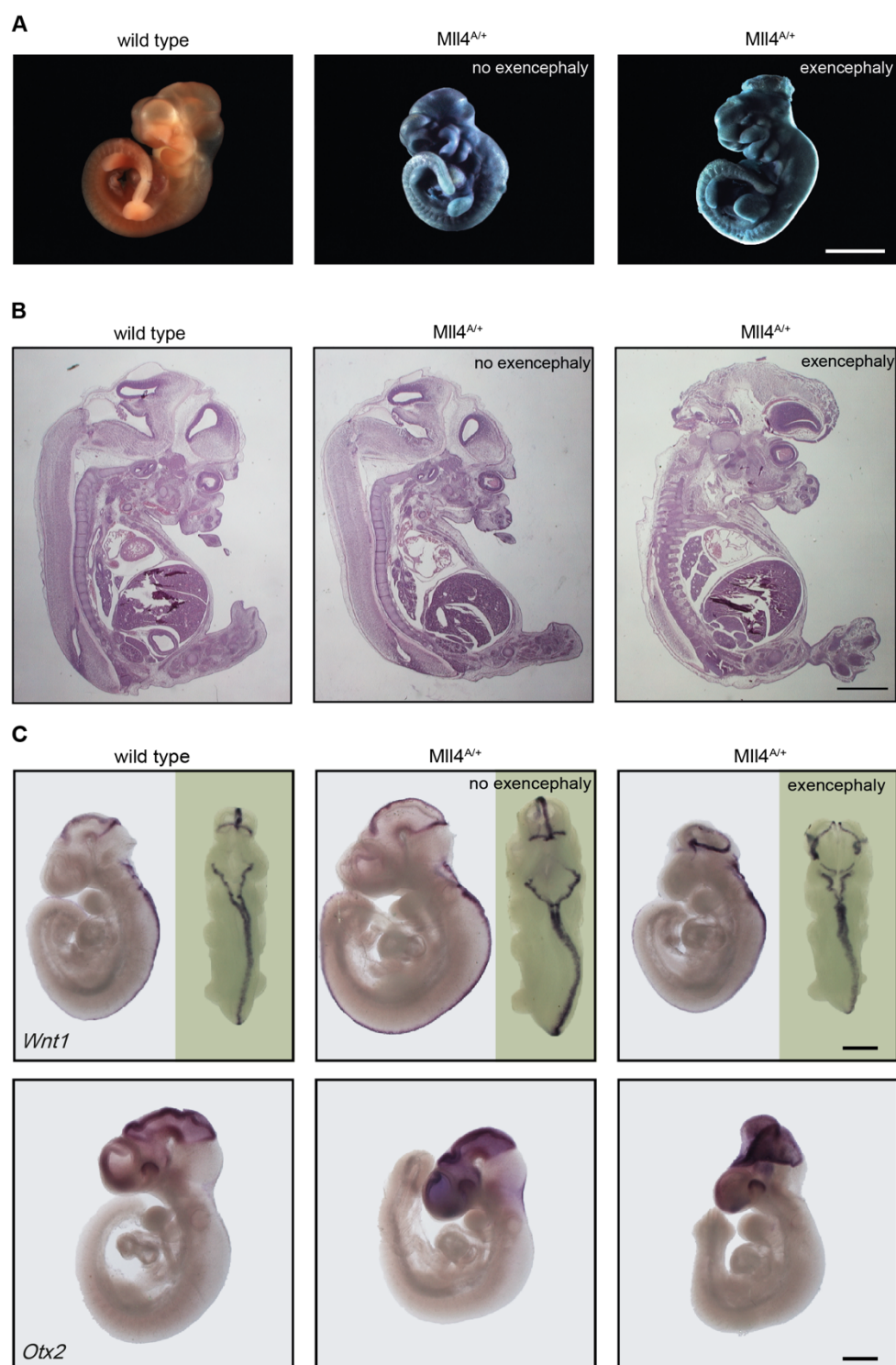


Figure 4

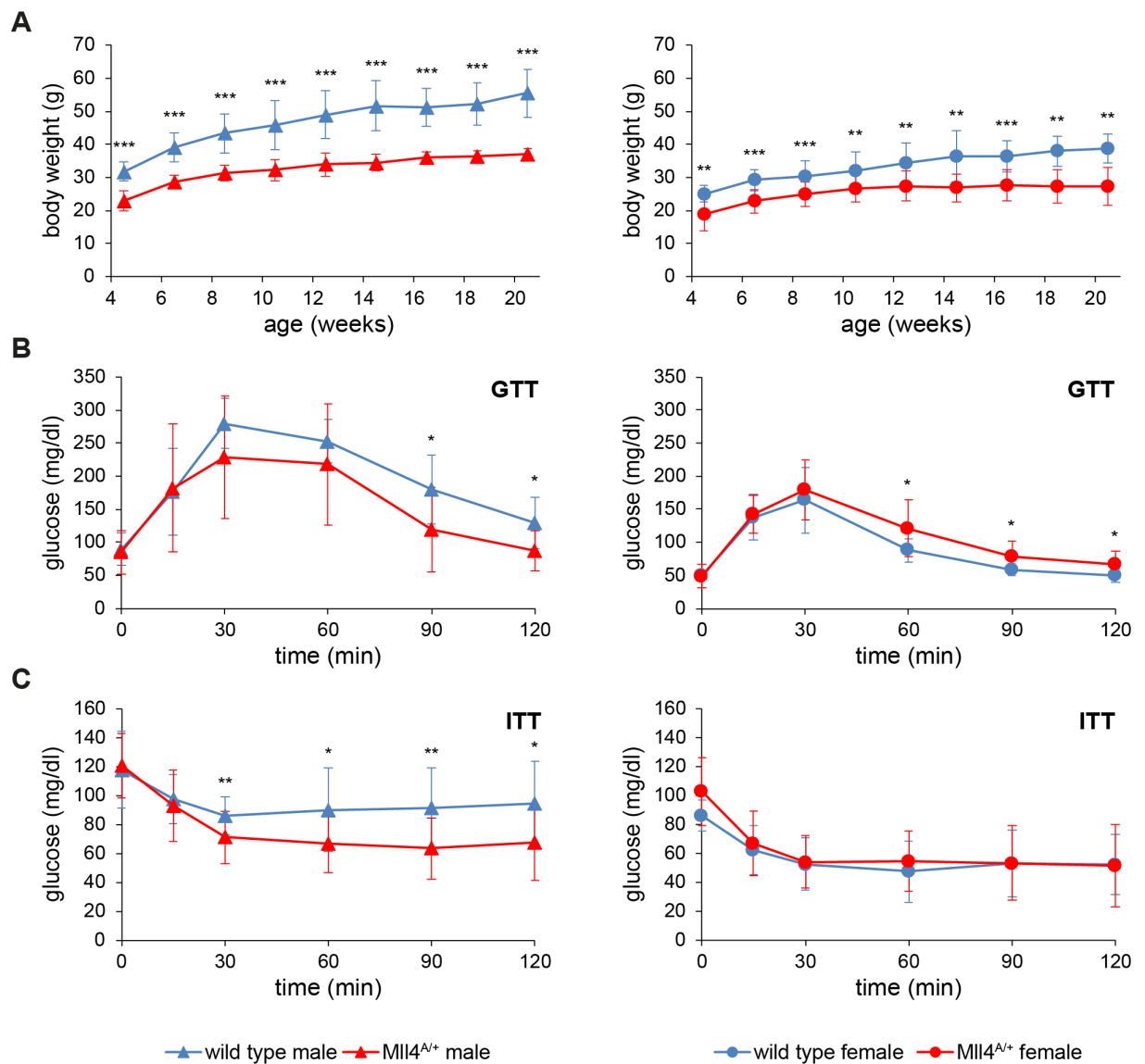


Figure 5

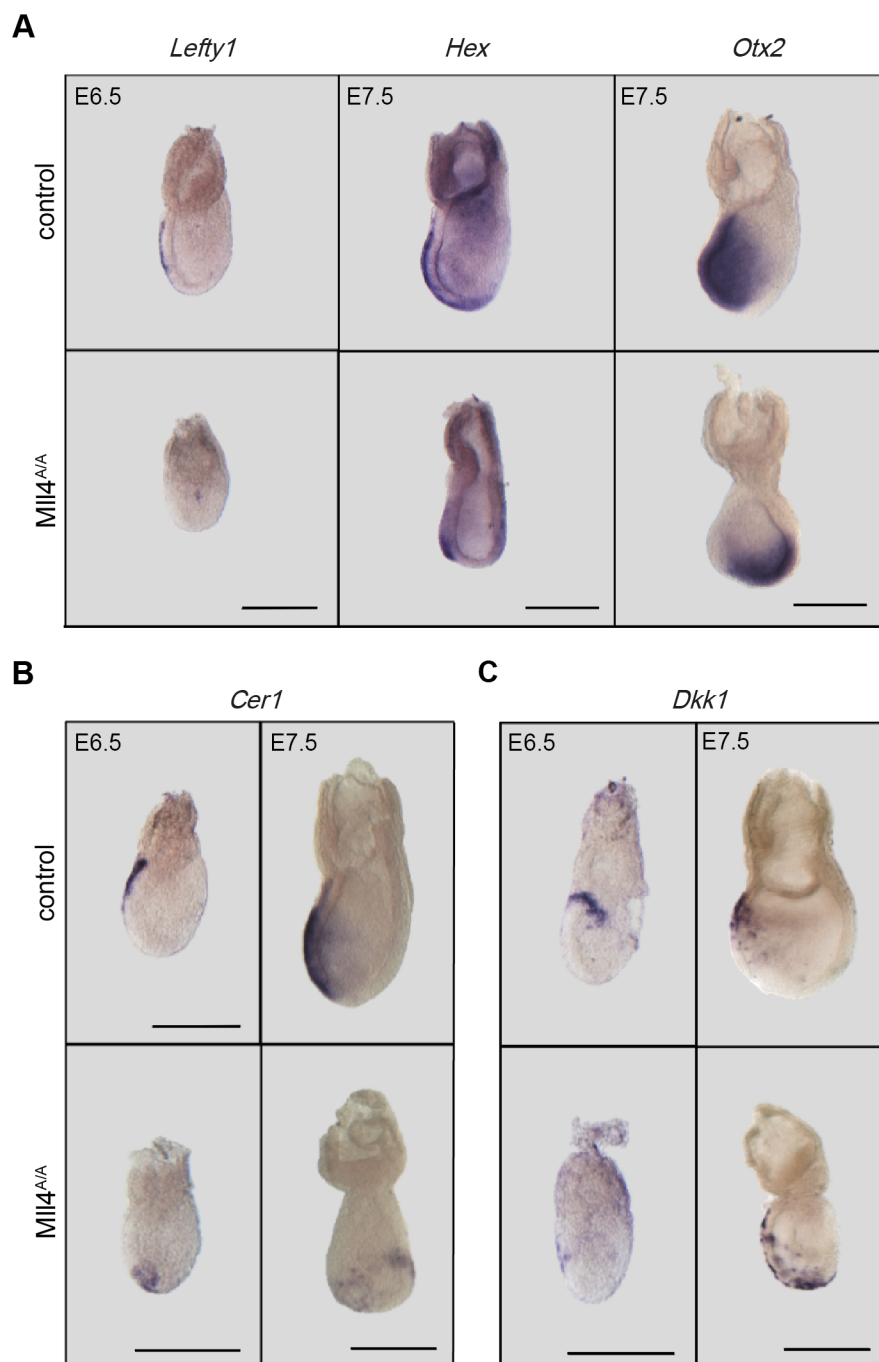


Figure 6

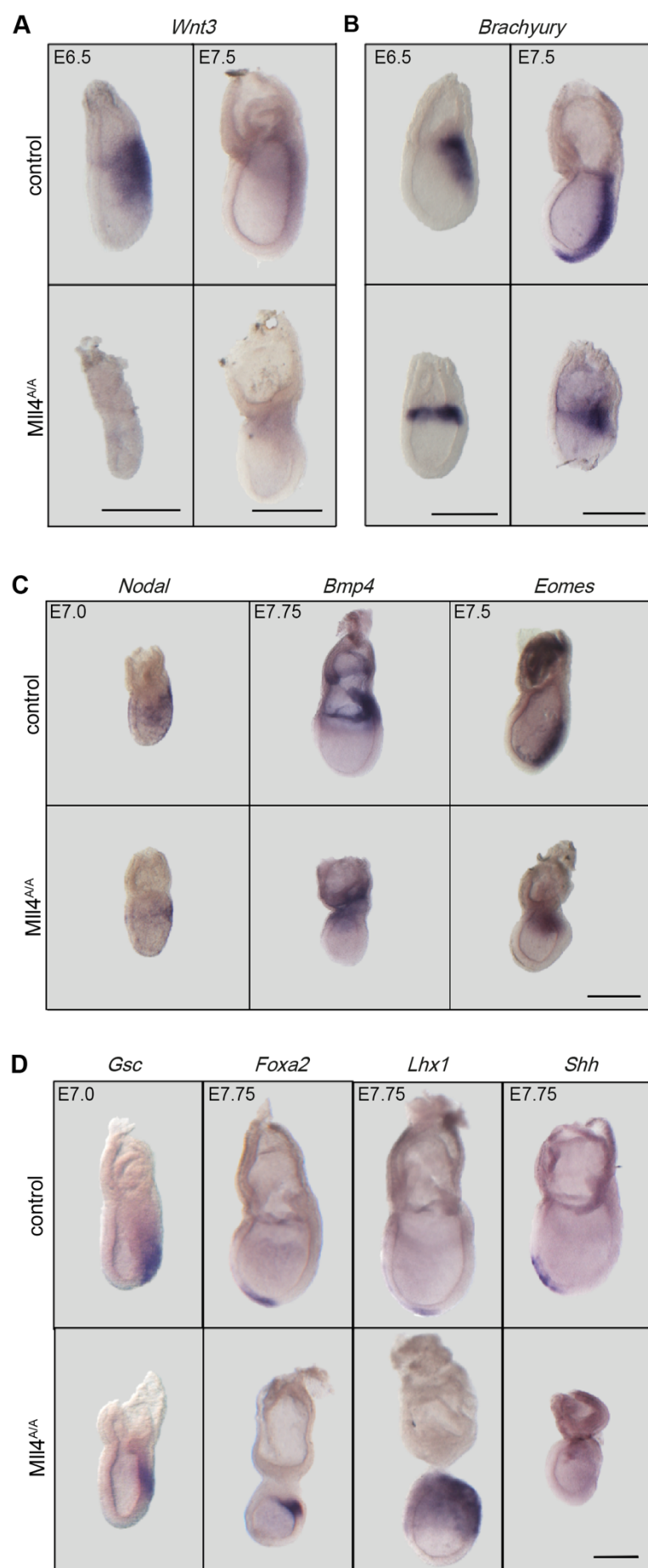


Figure 7

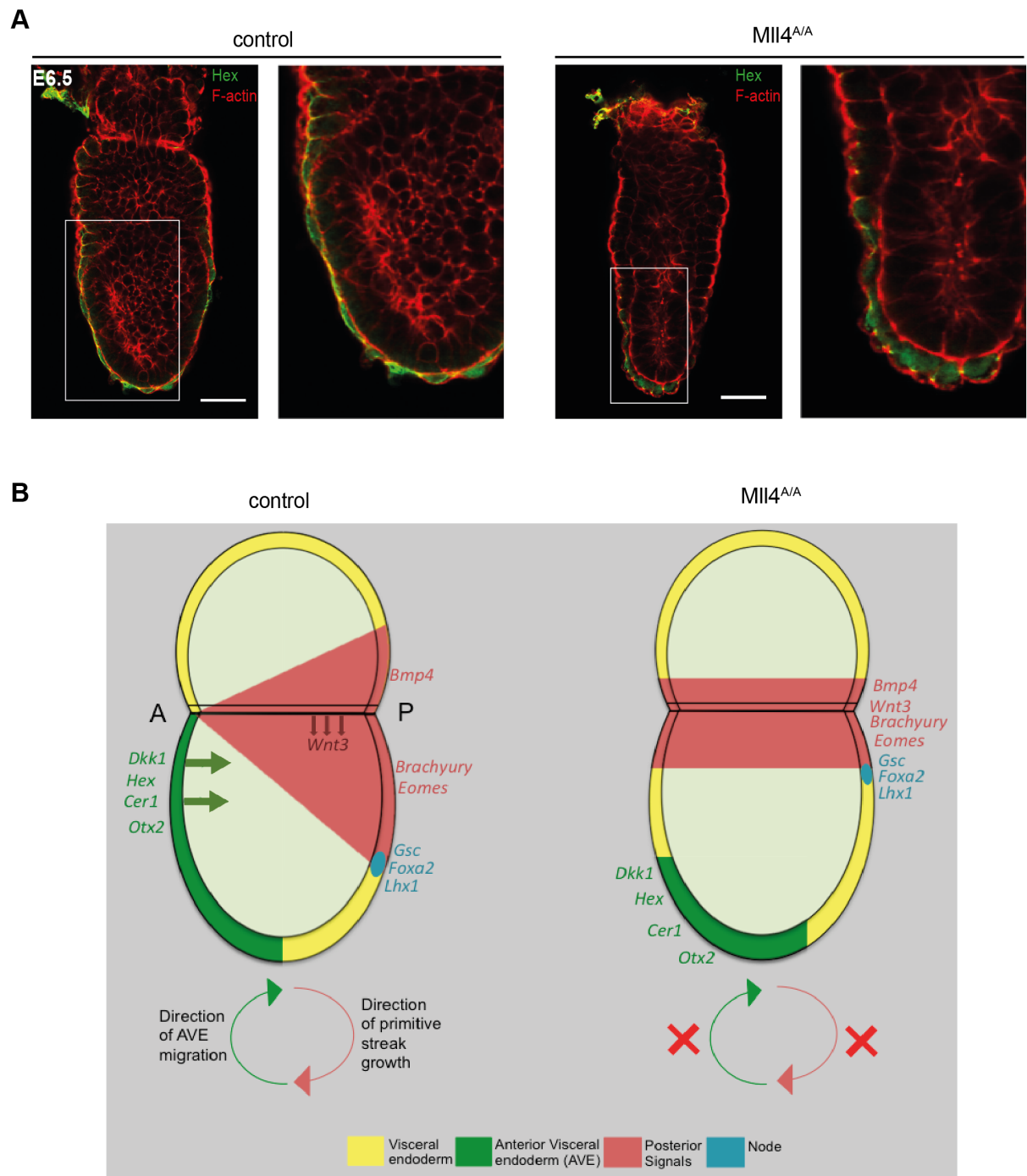


Figure S1

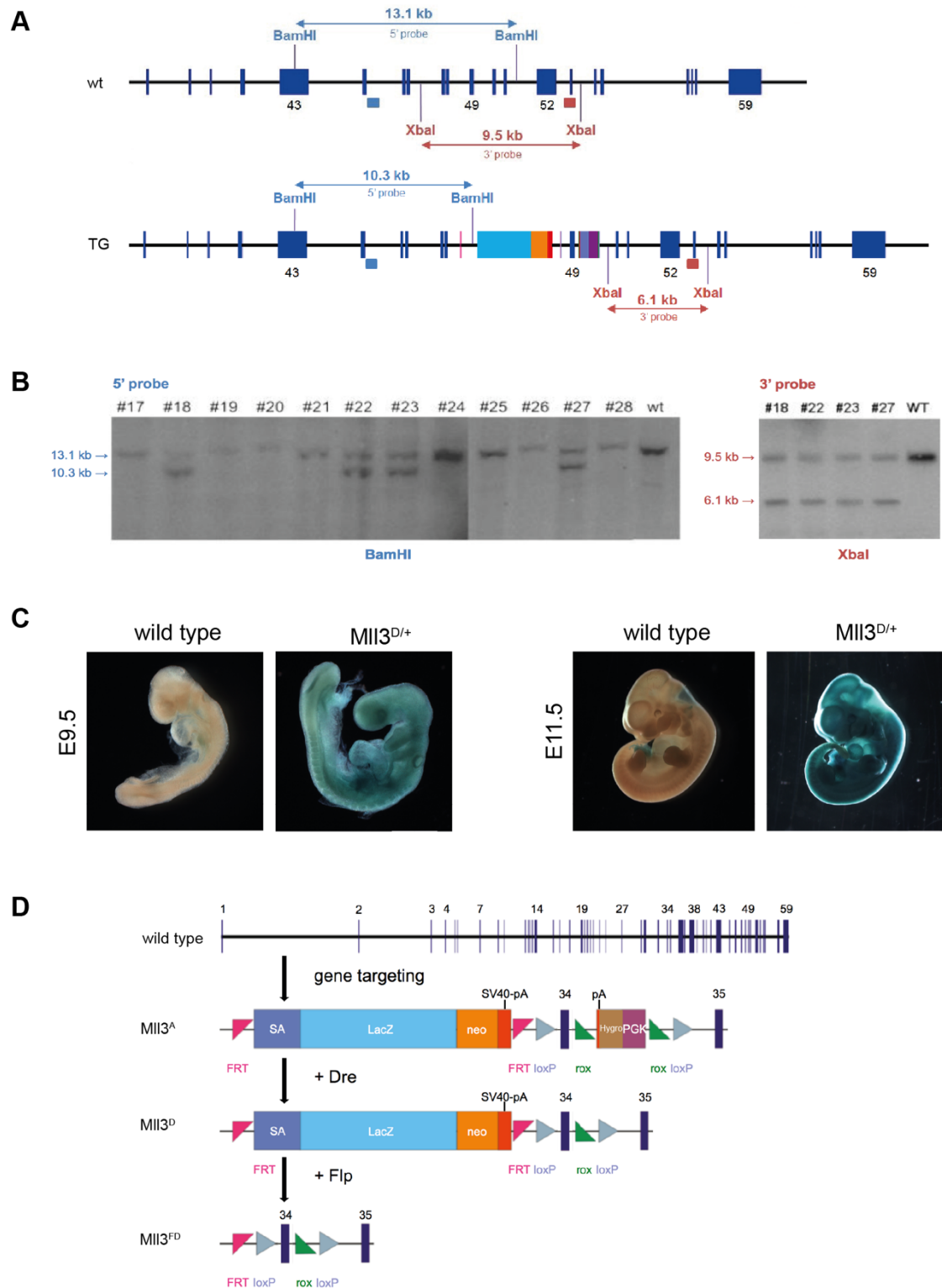
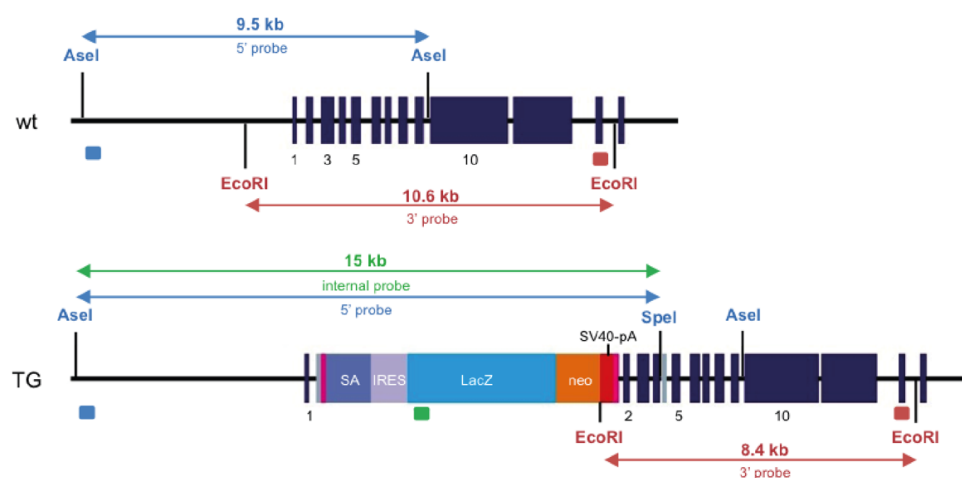
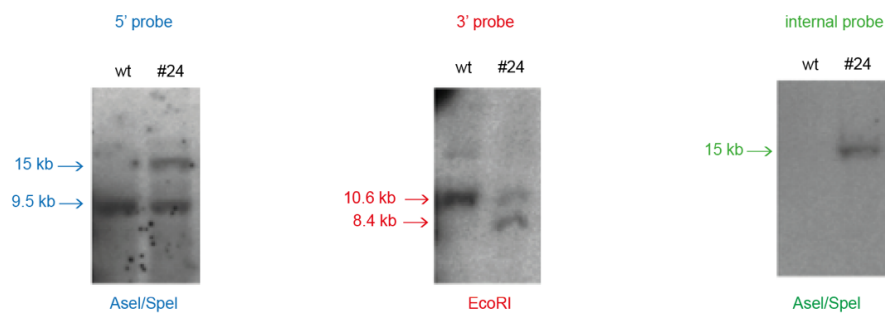


Figure S2

A



B



C

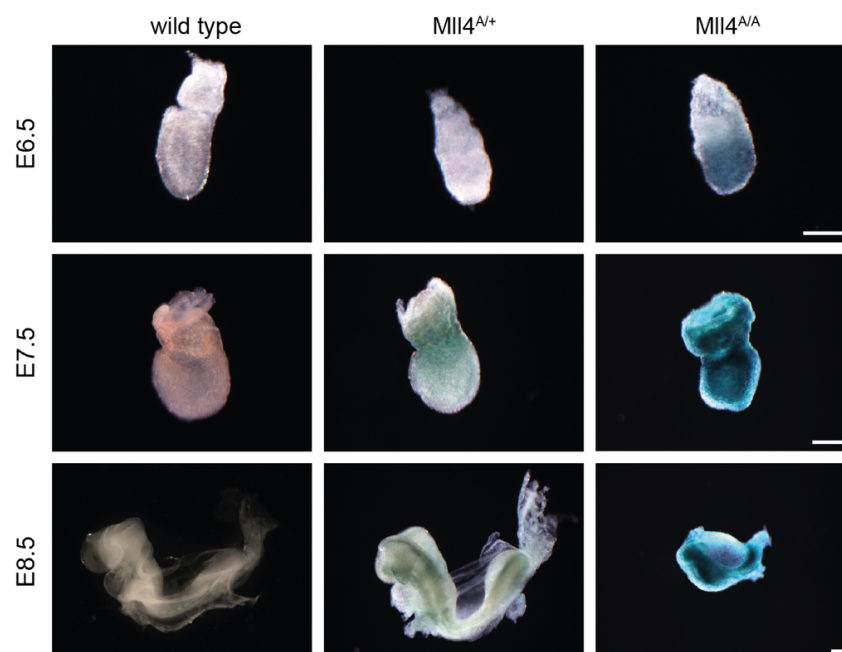


Figure S3

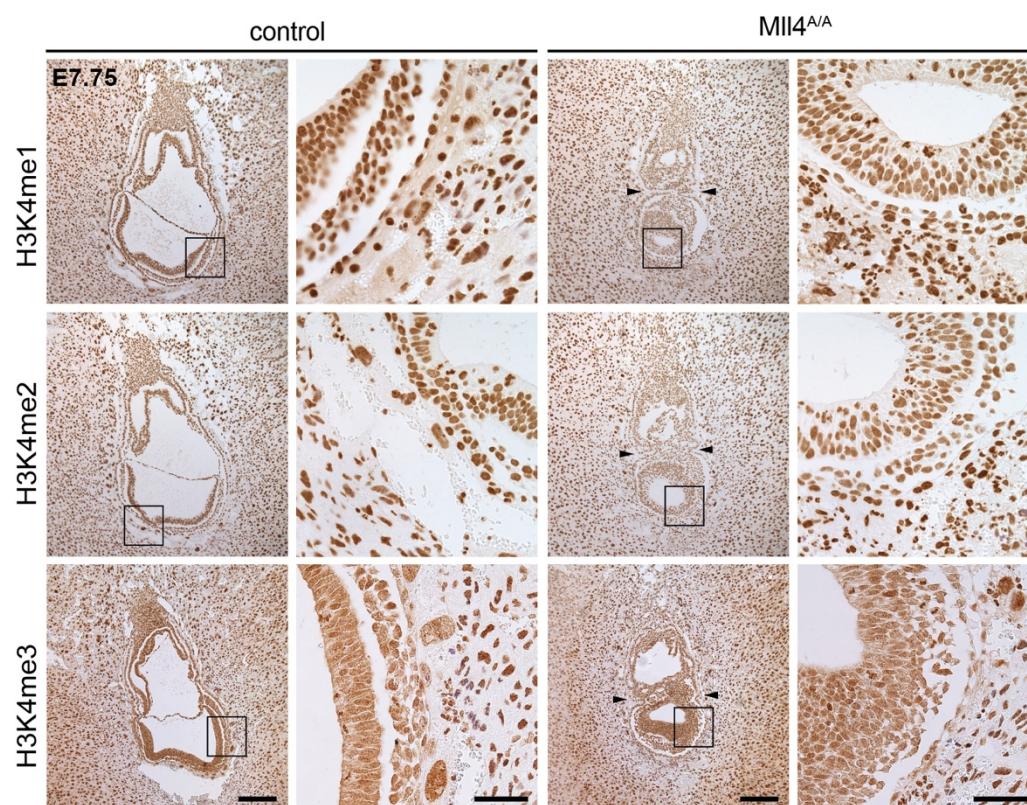


Figure S4

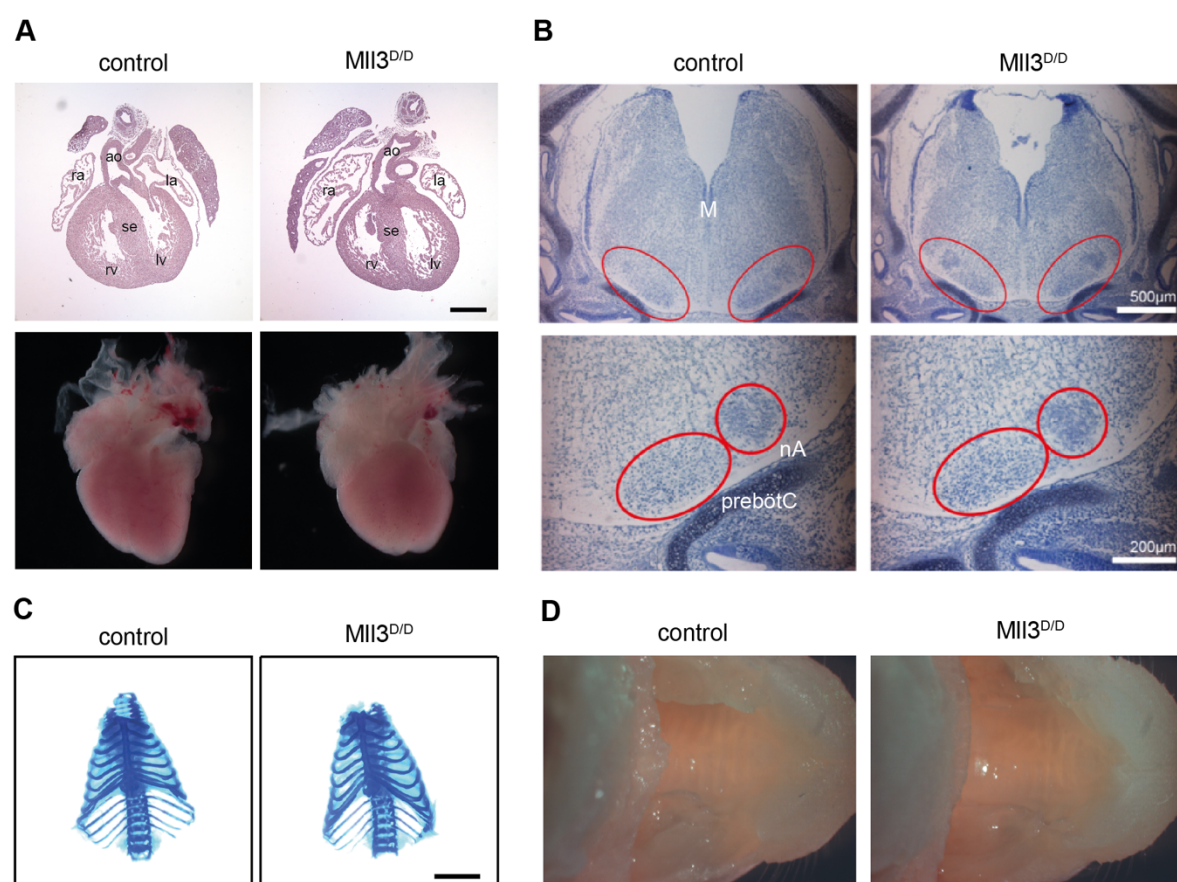


Figure S5

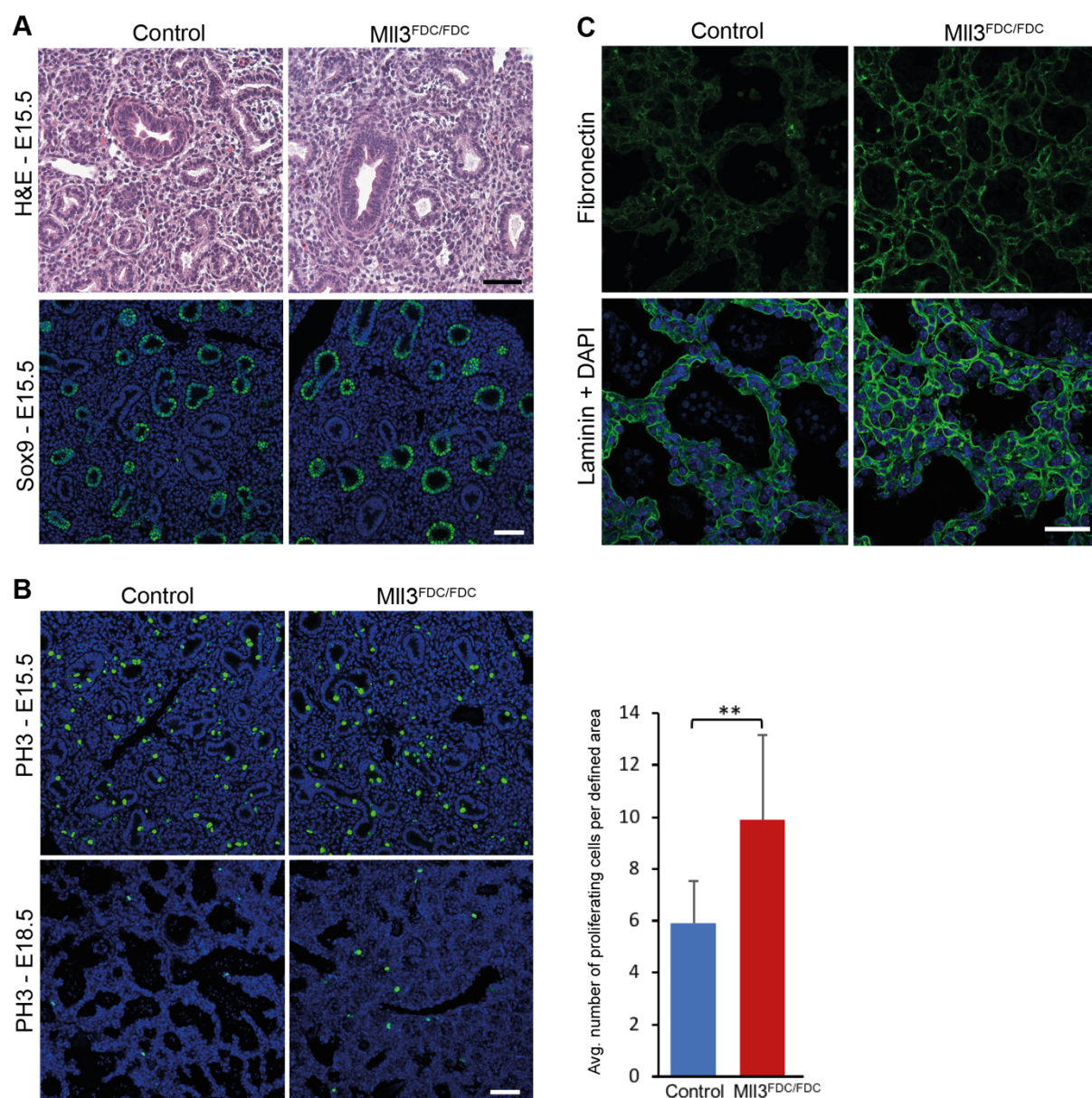


Figure S6

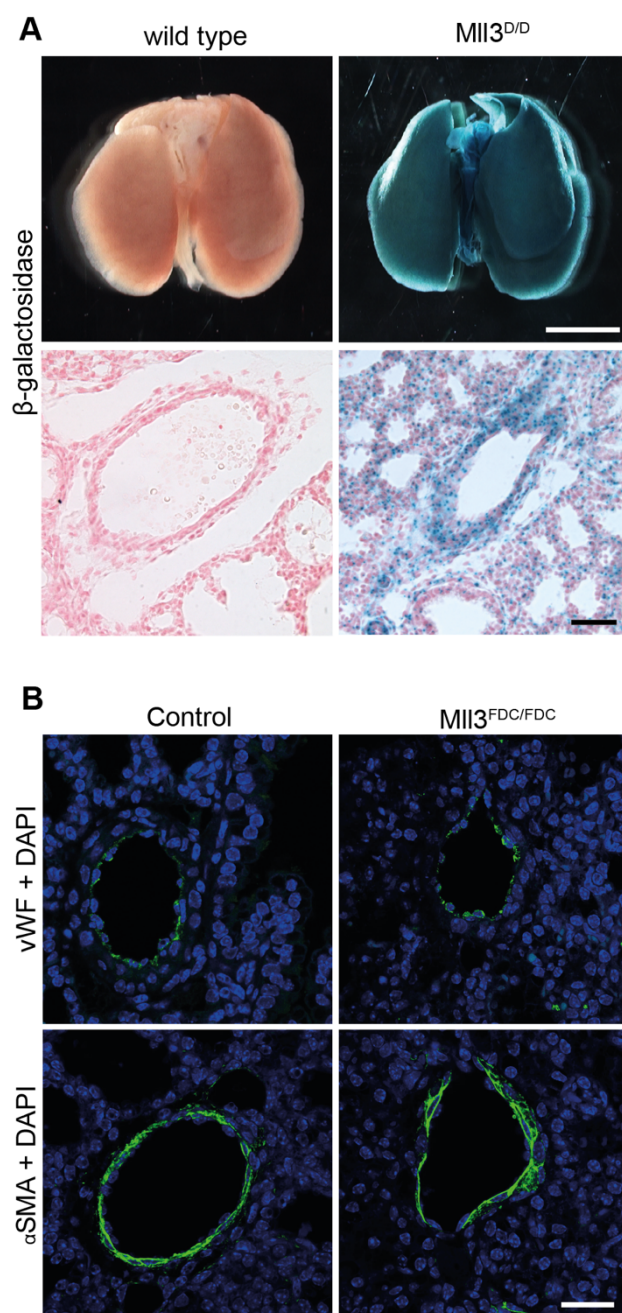


Figure S7

