

RESEARCH ARTICLE

MLL4 is required after implantation, whereas MLL3 becomes essential during late gestation

Deepthi Ashokkumar^{1,*}, Qinyu Zhang^{1,*}, Christian Much¹, Anita S. Bledau², Ronald Naumann³, Dimitra Alexopoulou⁴, Andreas Dahl⁴, Neha Goveas¹, Jun Fu¹, Konstantinos Anastassiadis², A. Francis Stewart^{1,5,‡} and Andrea Kranz^{1,‡}

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 paralogs 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 that initiates gastrulation in the mouse. This collective cell 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 paralogs in mouse development argues against their action as general enhancer factors.

KEY WORDS: Cell migration, Epigenetic regulation, Exencephaly, Gastrulation, Lung maturation

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 RNA polymerase II (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 orthologous Set1/Trithorax type H3K4 methyltransferases in three paralogous pairs; SETD1A and B (KMT2F and KMT2G), which are homologs of yeast Set1; MLL1 and 2 (KMT2A and KMT2B), which are homologs of *Drosophila* Trithorax; and MLL3 and 4 (KMT2C and KMT2D), which are homologs of *Drosophila* Lost PHD fingers of Trr (*Lpt*) 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 (Miller et al., 2001; Roguev et al., 2001) composed of four subunits, WDR5, RBBP5, ASH2L and DPY30 (Cho et al., 2007; Lee et al., 2006; Ruthenburg et al., 2007; Ernst and Vakoc, 2012) or less precisely ‘COMPASS’, 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 mammalian cells (Bledau et al., 2014). Similarly Set1 conveys most H3K4me3 in most *Drosophila* cell types (Ardehali et al., 2011; Mohan et al., 2011; Hallson et al., 2012). Consequently the Set1 homologs 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; Pinti and Shilatifard, 2016). Whether they proceed to catalyze H3K4 di- and trimethylation 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; Voigt et al., 2012). 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 (AFF1) and AF9 (MLLT3) (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.

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

²Stem Cell Engineering, Center for Molecular and Cellular Bioengineering, Biotechnology Center, Technische Universität Dresden, Tatzberg 47, 01307 Dresden, Germany. ³Transgenic Core Facility, Max Planck Institute of Molecular Cell Biology and Genetics, Pfotenhauerstr. 108, 01307 Dresden, Germany.

⁴DRESDEN-concept Genome Center, Center for Molecular and Cellular Bioengineering, Technische Universität Dresden, Fetscherstr. 105, 01307 Dresden, Germany. ⁵Max Planck Institute of Molecular Cell Biology and Genetics, Pfotenhauerstr. 108, 01307 Dresden, Germany.

*These authors contributed equally to this work

[†]Authors for correspondence (francis.stewart@tu-dresden.de, andrea.kranz@tu-dresden.de)

 D.A., 0000-0002-5862-4882; Q.Z., 0000-0001-8292-938X; C.M., 0000-0002-1644-8200; A.D., 0000-0002-2668-8371; N.G., 0000-0002-2531-5514; K.A., 0000-0002-9814-0559; A.F.S., 0000-0002-4754-1707; A.K., 0000-0002-7481-0220

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, among 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* (*KDM6A*). *UTX*, which is an H3K27 demethylase, is a subunit of the *MLL4* complex (Lee et al., 2006; Lederer et al., 2012, 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) (Dorighi et al., 2017). Similarly, the catalytic activity of the SET domain of Trr (the fly homolog 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; Glaser et al., 2006, 2009; Jude et al., 2007; 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 (Lee et al., 2013; Ang et al., 2016; Jang et al., 2017). Here, we compare the null and heterozygous phenotypes of these two genes in mouse development.

RESULTS

Mll3 and *Mll4* are conserved paralogs

MLL3 and *MLL4* are the largest known nuclear proteins (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; Table S4). Except for PHD3 of *MLL3*, which has been lost from *MLL4* (because PHD3 can be found in *Drosophila Lpt*) (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, in particular *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 (Table S4). Both genes encompass more than 50 exons (Fig. 1B,C) spliced to very long mRNAs (14 kb for *Mll3* and 19 kb for *Mll4*) that are widely expressed in the mouse embryo (Figs S1C, S2C).

Mll3 and *Mll4* knockouts die at different developmental stages

To explore the function of these conserved proteins, we established multipurpose alleles for both *Mll3* and *Mll4* by gene targeting. In our multipurpose allele strategy (Testa et al., 2004), a frameshifting exon(s) is flanked by *loxP* sites (exon 49 for *Mll3*;

Fig. 1B, Fig. S1A, S1B and exons 2–4 for *Mll4*; Fig. 1C, Fig. S2A, S2B) accompanied by the insertion of a genetrap 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 (Fig. S1D) 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 late gestation.

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}* embryos died at birth. As expected, *MLL4* expression was abolished by the intron 1 genetrap cassette insertion in *Mll4^{A/A}* ESCs (Fig. 1D), indicating that the stronger phenotype is the null. However, mRNA (*Mll4* mRNA expression at 68% of wild type in *Mll4^{FC/FC}* ESCs) and truncated protein expression persisted in *Mll4^{FC/FC}* ESCs, indicating that the milder phenotype was hypomorphic. Analysis of this hypomorphic allele is not 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 embryonic day (E) 10.5. Although not investigated, this phenotype was severe and could be the same as *Mll4^{A/A}*. 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^{A/A}*.

The *Mll3* and *Mll4* null phenotypes are strikingly different. Embryos lacking *Mll3* appeared to develop normally until birth

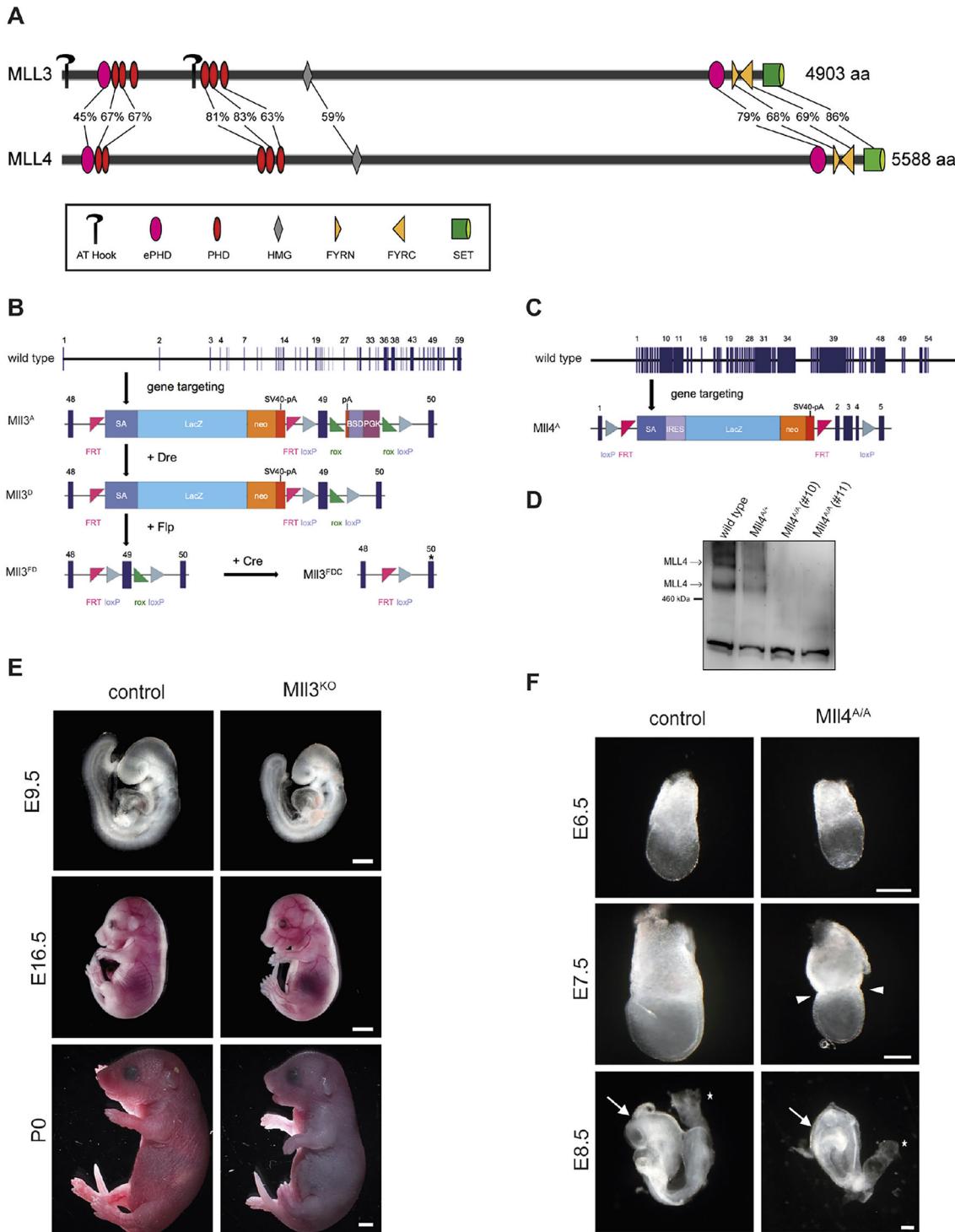


Fig. 1. Allele design and embryonic phenotype of the *Mll3* and *Mll4* mutant embryos. (A) Amino acid sequence alignment of murine MLL3 and MLL4 based on EMBOSS Needle. ePHD, extended PHD finger; FYRN and FYRC, F/Y-rich N- or C-terminus; HMG, high mobility group. Percentages indicate the sequence identities between MLL3 and MLL4 of the indicated domains. Table S4 provides a complete alignment of MLL3 and MLL4 including positions of the targeted cassettes. (B) Diagram of the *Mll3* knockout first allele (*Mll3^A*). The *Mll3* wild-type allele contains 59 coding exons. The targeted allele, *Mll3^A*, is similar to the Dre-recombinant allele, *Mll3^D*, with the PGK-Blasticidin selection cassette removed. The *Mll3^D* allele is converted to conditional (*Mll3^{FDC}*) 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. Asterisk depicts the exon with the premature stop codon. (C) Diagram of the *Mll4* knockout first allele (*Mll4^A*). The *Mll4* wild-type allele contains 54 coding exons. (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^{+/+}* 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 bars: 1 mm for E9.5; 2 mm for E16.5; 2 mm for P0. (F) *Mll4^{A/A}* embryos show growth retardation and a visible constriction is formed at the embryonic/extra-embryonic boundary (marked by arrowheads). Arrows point at the neural fold. Asterisk marks the allantois. Scale bars: 250 µm. BSD, blasticidin; lacZ-neo, β-galactosidase and neomycin resistance gene; pA, polyadenylation signal; PGK, phosphoglycerate kinase-1 promoter; SA, splice acceptor.

Table 1. Genotyping of progeny from *Mll3* intercrosses*Mll3^{D/+} × Mll3^{D/+}* (exon 49 floxed)

Age	+/+		D/+		D/D		Resorbed	Total (litters)
	Live (%)	Dead (%)	Live (%)	Dead (%)	Live (%)	Dead (%)		
Weaned	102 (44.7)	0 (0)	121 (53.1)	0 (0)	5# (2.19)*	0 (0)	0	228 (48)
P0	3 (18.7)	0 (0)	9 (56.3)	0 (0)	0 (0)	4 (25)	0	16 (2)
E18.5	47 (25.1)	1 (0.53)	89 (47.6)	1 (0.53)	49 (26.2)	0 (0)	3	187 (25)
E17.5	7 (38.9)	0 (0)	8 (44.4)	0 (0)	3 (16.7)	0 (0)	1	18 (1)
E16.5	5 (27.8)	0 (0)	11 (61.1)	0 (0)	2 (11.1)	0 (0)	0	18 (2)
E15.5	13 (28.9)	0 (0)	22 (48.9)	0 (0)	10 (22.2)	0 (0)	0	45 (7)
E13.5	2 (25.0)	0 (0)	5 (62.5)	0 (0)	1 (12.5)	0 (0)	2	8 (1)

(*Recovered pups were still in mixed C57BL/6J/OlaHsd-129/Sv background). *P<0.005; χ² test.*Mll3^{FDC/+} × Mll3^{FDC/+}* (exon 49 floxed)

Age	+/+		FDC/+		FDC/FDC		Resorbed	Total (litters)
	Live (%)	Dead (%)	Live (%)	Dead (%)	Live (%)	Dead (%)		
Weaned	54 (45)	0 (0)	66 (55)	0 (0)	0 (0)*	0 (0)	0	120 (31)
P0	6 (33.3)	0 (0)	10 (55.6)	0 (0)	0 (0)	2 (11.1)	0	18 (2)
E18.5	34 (22.7)	0 (0)	85 (57.0)	0 (0)	28 (18.8)	2 (1.3)	1	149 (21)
E17.5	3 (17.6)	0 (0)	9 (52.9)	1 (5.8)	4 (23.5)	0 (0)	0	17 (2)
E16.5	1 (11.1)	0 (0)	6 (66.7)	0 (0)	2 (22.2)	0 (0)	0	9 (1)
E15.5	2 (20.0)	0 (0)	5 (50.0)	0 (0)	3 (30.0)	0 (0)	0	10 (1)
E13.5	5 (26.3)	0 (0)	8 (42.1)	0 (0)	6 (31.6)	0 (0)	0	19 (2)

*P<0.005; χ² test.*Mll3^{D/+} × Mll3^{D/+}* (exon 34 floxed)

Age	+/+		D/+		D/D		Resorbed	Total (litters)
	Live (%)	Dead (%)	Live (%)	Dead (%)	Live (%)	Dead (%)		
Weaned	65 (43.3)	0 (0)	85 (56.67)	0 (0)	0 (0)*	0 (0)	0	150 (16)
P0	6 (37.5)	0 (0)	9 (56.3)	0 (0)	0 (0)	1 (6.3)	0	16 (2)
E18.5	3 (17.6)	0 (0)	13 (76.5)	0 (0)	1 (5.9)	0 (0)	0	17 (3)
E17.5	2 (25.0)	0 (0)	5 (62.5)	0 (0)	1 (12.5)	0 (0)	0	8 (1)
E15.5	2 (28.6)	0 (0)	3 (42.9)	0 (0)	2 (28.6)	0 (0)	0	7 (1)
E13.5	2 (25.0)	0 (0)	4 (50.0)	0 (0)	2 (25.0)	0 (0)	0	8 (1)

*P<0.005; χ² test.

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^{A/A}* embryos was growth retardation starting at E6.5 followed by the appearance of a marked constriction at the embryonic/extrembryonic boundary 1 day later (Fig. 1F). Later in development the *Mll4^{A/A}* embryos displayed abnormal headfolds, absence of somites and heartbeat, and did not turn (Fig. 1F, Table 2). Despite this severe phenotype, *Mll4^{A/A}* embryos displayed no observable change in global mono-, di- or trimethylation of H3K4 (Fig. S3). Furthermore, heterozygous *Mll4* embryos exhibited abnormalities (see below) whereas heterozygous *Mll3* 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* (*Kdm6b*), 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-E).

Considering that the structures of the palate, diaphragm, intercostal muscles, brain stem and heart were normal, and the fact that MLL3 is highly expressed in the lung (Fig. S5A), we focused on defects in the lung as the cause of respiratory failure. Lungs from E15.5 and E18.5 fetuses were examined for morphological differences. At the pseudoglandular stage, E15.5,

Table 2. Embryonic lethality of *Mll4^{A/A}* embryos*Mll4^{A/+} × Mll4^{A/+}*

Age	+/+		A/+		A/A		Exencephaly (% of A/+)	Resorbed	Total (litters)
	Live (%)	Dead (%)	Live (%)	Dead (%)	Live (%)	Dead (%)			
Weaned	17 (58.6)	0 (0)	12 (41.4)	0 (0)	0 (0)	0 (0)*	0 (0)	0	29 (10)
E11.5	6 (19.4)	0 (0)	24 (77.4)	0 (0)	0 (0)	1 (3.2)*	5 (20.80)	3	31 (4)
E10.5	16 (25.4)	0 (0)	32 (50.8)	0 (0)	0 (0)	15 (23.8)*	6 (18.75)	12	63 (7)
E9.5	37 (28.9)	0 (0)	70 (54.7)	0 (0)	0 (0)	21 (16.4)*	14 (20)	23	128 (14)
E8.5	76 (26.1)	0 (0)	153 (52.6)	0 (0)	62 (21.3)	0 (0)	0 (0)	39	291 (29)
E7.5	43 (25.0)	0 (0)	86 (50.0)	0 (0)	43 (25.0)	0 (0)	0 (0)	19	172 (17)
E6.5	14 (35.9)	0 (0)	21 (53.8)	0 (0)	4 (10.3)	0 (0)	0 (0)	12	39 (6)

*P<0.005; χ^2 test.

Mll3^{KO} lungs appeared normal in lobulation pattern and small bifurcations in the distal epithelium (Fig. S5B). Examination of phospho-histone H3 (PH3) during this highly proliferative stage of lung development revealed no significant difference between control and *Mll3^{KO}* lungs (Fig. S5B).

At the saccular stage several *Mll3^{KO}* fetuses, but not all, had smaller lungs than wild-type littermates (Fig. 2A). Hematoxylin and eosin (H&E) staining of E18.5 lung sections showed that *Mll3^{KO}* lung had thickened septae and smaller alveolar spaces (Fig. 2A). Staining with the epithelial marker E-cadherin (CDH1) showed normally developed epithelial lining, and the integrity of the proximal epithelium was confirmed by the tight junctional marker zonula occludens-1 (ZO-1; TJP1) (Fig. S5C). The differentiated proximal epithelium consists of neuroendocrine, ciliated and secretory club cells. Secretoglobin family 1A member 1 (SCGB1A1; CC10), which is a marker for club cells, was reduced in *Mll3^{KO}* lungs (Fig. 2B). 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. Lungs of both genotypes underwent this inhibition of proliferation; however, to a lesser extent in the *Mll3^{KO}* (Fig. S6A). 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}* lungs (Fig. S6B).

To unravel the molecular basis of the lung maturation defect, we performed a global transcriptome analysis on lung samples of E18.5. As expected for an H3K4 methyltransferase, we observed five times more mRNAs down- than upregulated. Analysis of differentially expressed genes (DEGs) at a false discovery rate (FDR) of 5% revealed 70 up- and 354 downregulated genes (Fig. 2C, Table S5). The most highly enriched pathways in the group of downregulated DEGs related to cell motility, vasculature development, regulation of cell differentiation and morphogenesis, and lung development (Fig. 2D). GO term analysis suggests a defect in lung vasculature. Preliminary analysis with antibody staining against CD31 (PECAM1) and α -smooth muscle actin (α SMA) demonstrated the establishment of lung vasculature (Fig. S6C). The functional intactness still needs to be determined.

Using published datasets specific for distal lung epithelium [ciliated cells, club cells, alveolar type I and type II epithelial cell (AEC-I and -II)] (Treutlein et al., 2014), gene set enrichment analysis (GSEA) revealed a significant negative correlation exclusively for AEC-I marker genes (Fig. 2E, Table S5). Notably, 66% of the AEC-I marker genes were represented in our dataset (96 out of 145). Quantitative RT-PCR validated the RNA-seq results on all four selected AEC-I marker genes (*Aqp5*, *Clic5*, *Pdpn* and *Hopx*;

Fig. 2F). The top downregulated mRNA was *Aqp5*, encoding a water channel protein. AQP5 staining visualizes the continuous squamous epithelium that lines the air spaces in control lungs at E18.5. However in *Mll3^{KO}* lungs this characteristic marker was significantly reduced (Fig. 2B). *Pdpn*, which is required for effective maturation of AEC-I, was another downregulated AEC-I marker. We conclude that the AEC-I is the most affected cell type in MLL3 zygotic null development.

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 many heterozygous *Mll4* knockouts presented exencephaly. 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,B). This disorder was inherited with a parent-of-origin distortion. When only the father was the carrier of the *Mll4^A* allele, on average 11% of the heterozygous embryos of all dissected developmental stages developed exencephaly, whereas 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 2). Moreover, two-thirds of all exencephalic embryos were female (data not shown). Later in gestation, exposed neural folds degenerated, producing anencephaly (Fig. S8A). These embryos were found dead at P1, 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 whole-mount *in situ* hybridization. As seen in Fig. 3C, *Wnt1* is expressed normally in the exencephalic embryo, so disturbed *Wnt1* expression is not the cause of the phenotype. Furthermore *Otx2* was also normally expressed (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 was observed. However, we did find indications of the metabolic

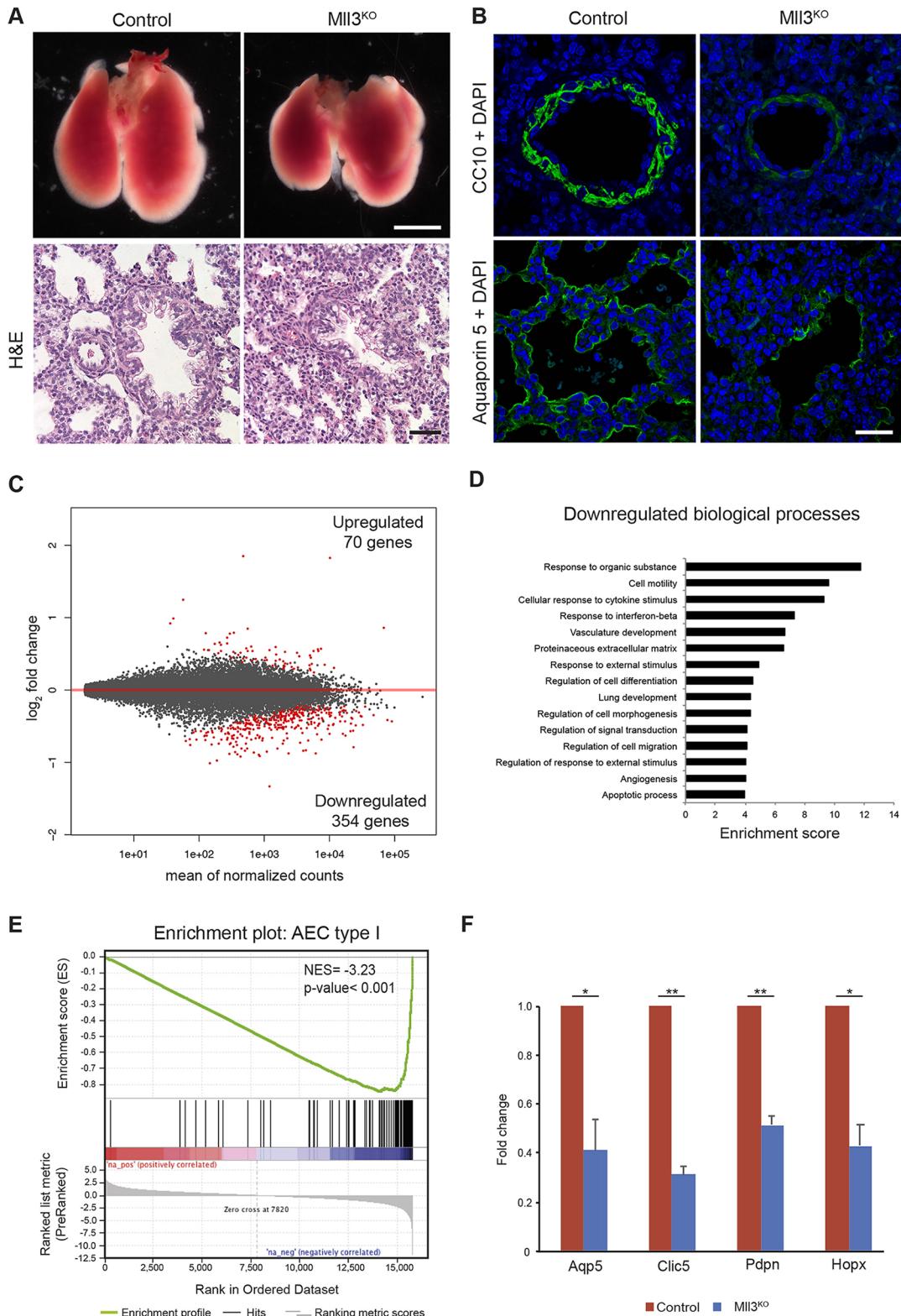


Fig. 2. Loss of MLL3 causes defects in lung maturation. (A) Analysis of gross morphology and H&E staining of $Mll3^{KO}$ and control littermate lung at E18.5. Scale bars: 500 μ m (whole lung); 250 μ m (histology). (B) CC10, a characteristic marker for club cells, is reduced in $Mll3^{KO}$ lung compared with control lung at E18.5. Expression of aquaporin 5, a marker for AEC-I, is reduced in $Mll3^{KO}$ lung compared with control lung at E18.5. Scale bar: 25 μ m. (C) mRNA profiling of $Mll3^{KO}$ and control lung samples at E18.5. MA plot visualizing the \log_2 -fold change differences according to expression level. Red dots represent significant DEGs at a 5% FDR. Grey dots represent genes whose difference in expression is not considered significant at 5% FDR. (D) Enriched terms of biological processes of downregulated DEGs using DAVID GO/BP/FAT database. (E) GSEA shows consistent and significant downregulation of genes from the AEC-I signature gene set in $Mll3^{KO}$ lung compared with control lung samples. NES, normalized enrichment score. (F) qRT-PCR performed on control and $Mll3^{KO}$ lung for selected AEC-I marker mRNAs (*Aqp5*, *Clic5*, *Pdpn*, *Hopx*). One representative mouse pair is shown. Expression levels were normalized to *Rpl19* and plotted as fold change relative to control. Data are mean+s.d. (* P <0.05, ** P <0.01; unpaired Student's *t*-test).

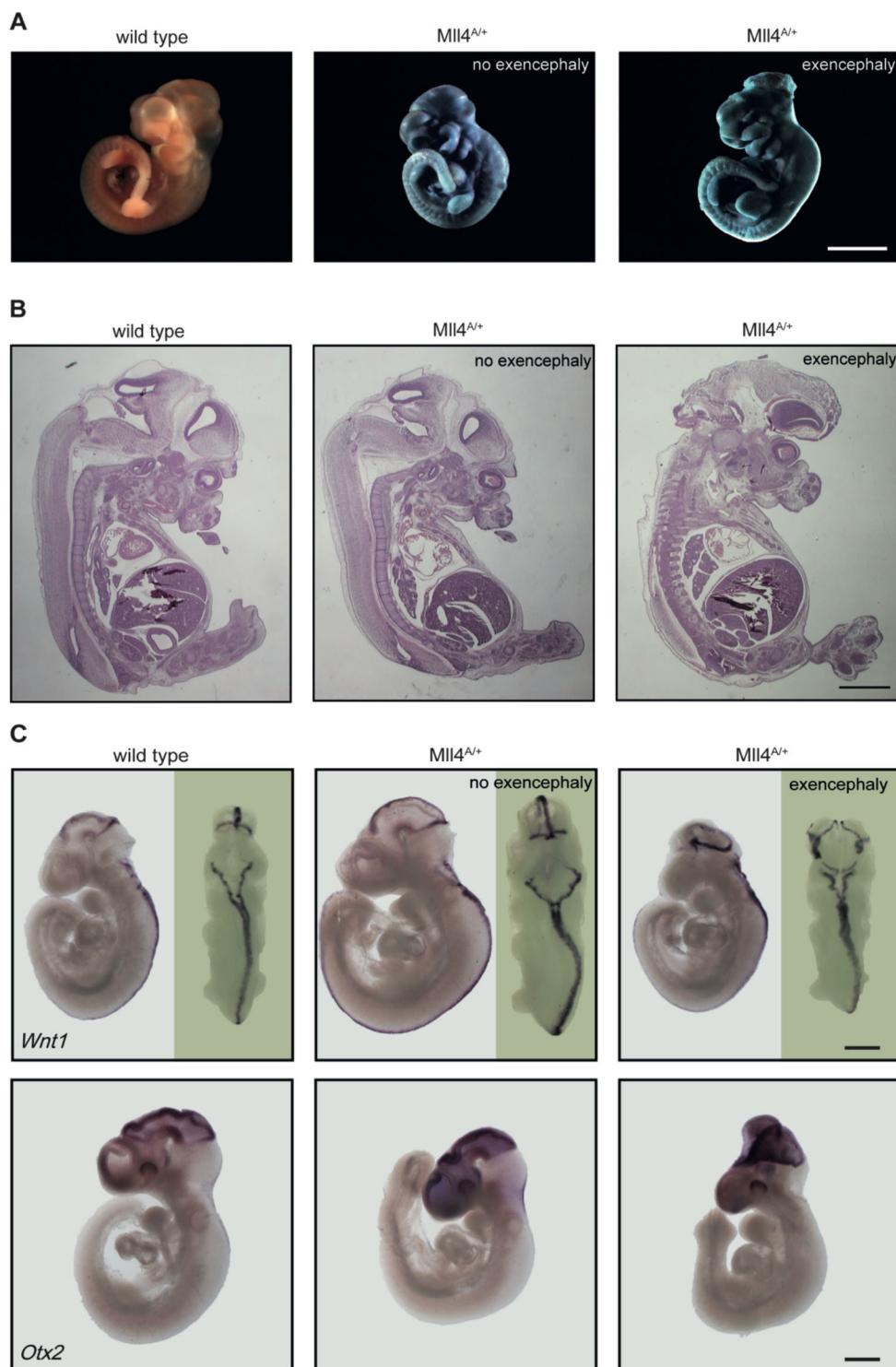


Fig. 3. Neural tube closure defects in *Mll4*^{A/+} embryos. (A) β -galactosidase staining of heterozygous *Mll4*^{A/+} embryos at E10.5. The embryo on the right exhibits exencephaly. Scale bar: 1 mm. (B) H&E-stained sagittal sections of E13.5 embryos show the everted and enlarged exposed neural folds. Scale bar: 1 mm. (C) Whole-mount *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 with wild-type littermates. Scale bars: 250 μ m.

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 4 weeks were observed. *Mll4*^{A/+} mice of both sexes remained 30% smaller throughout adulthood (Fig. 4A) and showed a decreased amount of white adipose tissue (Fig. S8B). 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, insulin tolerance tests reported higher insulin sensitivity in male *Mll4*^{A/+} 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,C).

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 *Foxa2* (*Hnf3b*), *Otx2* and *Lhx1* (*Lim1*), in which specification of the anterior-posterior (A-P) embryonic axis is disrupted (Kinder et al., 2001). To examine A-P

Table 3. Exencephaly of *Mll4^{A/+}* embryos is inherited in a parent-of-origin-dependent manner*♂ Mll4^{A/+} x ♀ CD1*

Age	+/+		A/+		Exencephaly (% of A/+)	Resorbed	Total (litters)
	Live (%)	Dead (%)	Live (%)	Dead (%)			
Weaned	1238 (82.1)	0 (0)	270 (17.9)*	0 (0)	0 (0)	0	1508 (181)
E18.5	29 (53.7)	0 (0)	25 (46.3)	0 (0)	3 (12.0)	0	54 (4)
E16.5	6 (40.0)	0 (0)	9 (60.0)	0 (0)	2 (22.2)	0	15 (1)
E15.5	12 (48.0)	0 (0)	13 (52.0)	0 (0)	0 (0.0)	1	25 (2)
E14.5	9 (75.0)	0 (0)	3 (25.0)	0 (0)	0 (0.0)	0	12 (1)
E13.5	25 (44.6)	0 (0)	31 (55.3)	0 (0)	4 (12.9)	1	56 (4)
E11.5	8 (47.1)	0 (0)	9 (52.9)	0 (0)	1 (11.1)	0	17 (1)
E10.5	43 (52.4)	0 (0)	39 (47.5)	0 (0)	7 (17.9)	2	82 (6)
E9.5	56 (46.6)	0 (0)	64 (53.3)	0 (0)	5 (7.8)	5	120 (9)
Total	188 (49.3)	0 (0)	193 (50.6)	0 (0)	22 (11.4)	9	381 (28)

P<0.005; χ^2 test.♂ CD1 x ♀ Mll4^{A/+}*

Age	+/+		A/+		Exencephaly (% of A/+)	Resorbed	Total (litters)
	Live (%)	Dead (%)	Live (%)	Dead (%)			
Weaned	88 (77.2)	0 (0)	26 (22.8)*	0 (0)	0 (0)	0	114 (20)
E11.5	7 (70.0)	0 (0)	3 (30)	0 (0)	2 (66.7)	2	10 (1)
E10.5	7 (35.0)	0 (0)	13 (65.0)	0 (0)	4 (30.8)	5	20 (2)
E9.5	7 (46.6)	0 (0)	8 (53.3)	0 (0)	6 (75.0)	4	15 (2)
Total	21 (46.6)	0 (0)	24 (53.3)	0 (0)	12 (50)	11	45 (5)

*P<0.005; χ^2 test.

patterning in *Mll4^{A/A}* embryos, we performed whole-mount *in situ* hybridizations for anterior and posterior molecular markers (Figs 5, 6) between developmental stages E6.5 and 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* (*Hhex*), 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). *Bmp4*, *Nodal* and *Wnt* antagonist *cerberus 1* (*Cer1*) 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/extrembryonic boundary; however, in *Mll4^{A/A}* embryos expression was observed at the distal region (Fig. 5C).

As the anterior markers were misexpressed at the distal region of *Mll4^{A/A}* embryos (Fig. 5), we expected defective posterior patterning. *Wnt3* and *brachyury* (*T*) 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/extrembryonic boundary (Fig. 6A,B). *Nodal* is essential for the induction of the primitive streak (Conlon et al., 1994) and is normally expressed in the posterior embryonic ectoderm, expanding along the length of the primitive streak. This failed in *Mll4^{A/A}* embryos. *Bmp4* and *Eomes* mark the posterior and anterior region of the primitive streak, respectively, and in addition mark the extra-embryonic chorion and amnion. In *Mll4^{A/A}* embryos both markers were detected at the embryonic/extrembryonic boundary (Fig. 6C). The patterning of posterior markers in *Mll4^{A/A}* embryos indicates that the primitive streak did not extend distally. Consequently paraxial mesoderm was absent (Fig. S7).

At the anterior region of the primitive streak is a distinct region called the node, from which the axial mesoderm such as head

process and notochord develop (Yamanaka et al., 2007; Benazeraf and Pourquié, 2013). Because the primitive streak did not extend in *Mll4^{A/A}* embryos, we examined whether the node and the node derivatives were present. *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* (*Meox1*) transcripts confirming the lack of somites. Very weak and discontinuous localization of *T* 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 filopodia and migrate unidirectionally as a collective towards the future anterior of the embryo until they reach the embryonic/extrembryonic boundary (Srinivas et al., 2004). At E6.5 squamous AVE cells form a single-layer reaching the embryonic/extrembryonic boundary indicated by expression of HEX (Fig. 7A). In *Mll4*-deficient embryos, AVE cells failed to reach the embryonic/extrembryonic 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.

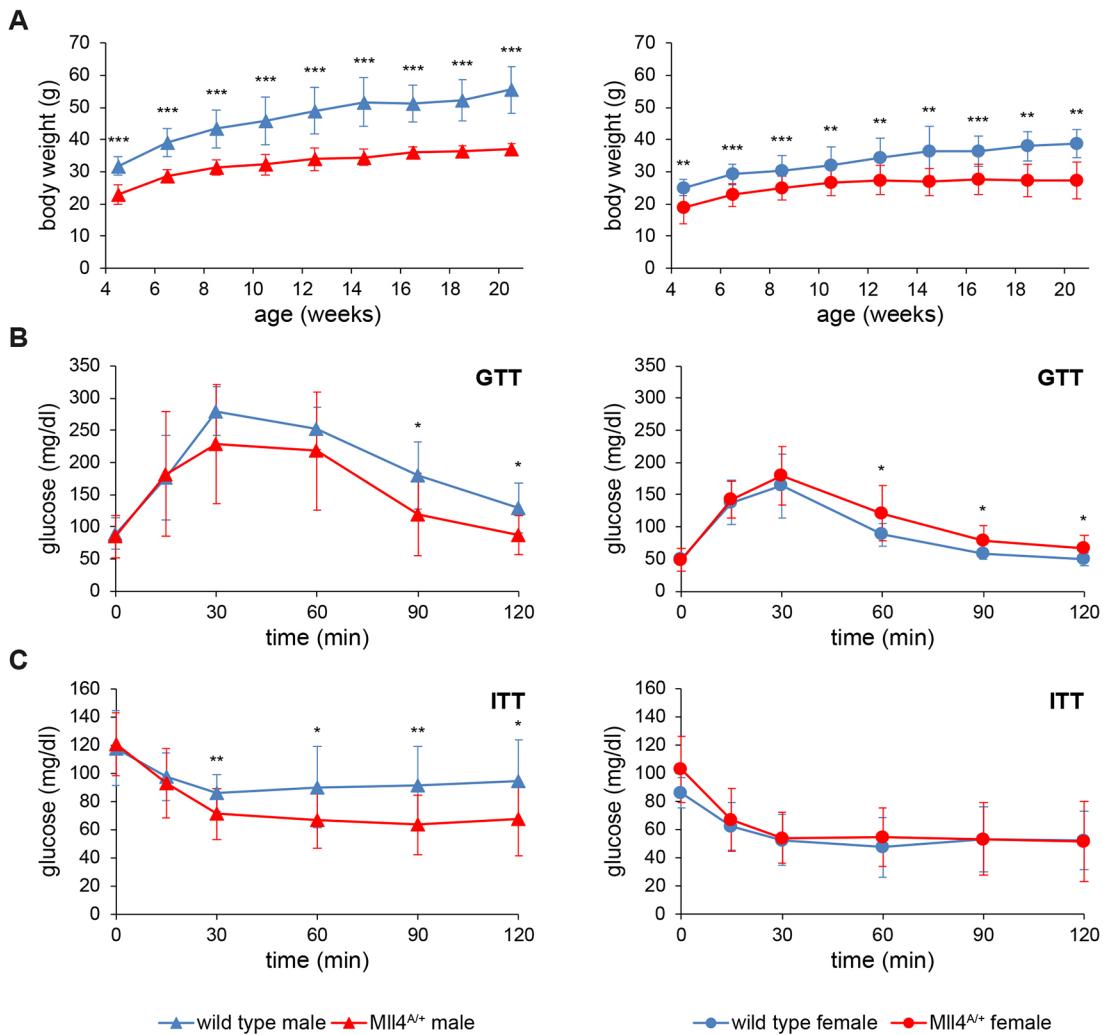


Fig. 4. Decreased body weight and hypoglycemia of *Mll4^{A/+}* mice. (A) Development of body weight over 4–20 weeks for a total of 48 mice (24 males: 15 wild-type and 9 heterozygous mice; 24 females: 16 wild-type and 8 heterozygous mice) is shown. The differences in weight are significant at any developmental stage. Data are mean±s.d. (**P<0.01, ***P<0.001; unpaired Student's 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 9–15 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. Data are mean±s.d. (*P<0.05; unpaired Student's 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. Data are mean±s.d. (*P<0.05, **P<0.01; unpaired Student's t-test).

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 paralogs 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 prevailing 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 a system of extensive functional redundancy and backup. If so, their individual knockout phenotypes mainly reveal flaws in the backup system rather than their ongoing activities. Some evidence for the functional backup proposition has been acquired (Lee et al., 2013; Denissov et al., 2014; Chen et al., 2017, 2018).

MLL3 and defective respiration

Mll3^{D/D} and *Mll3^{FDC/FDC}* neonates died owing to failures in the final steps of lung maturation. MLL3 is not required for patterning of the lung but is required for efficient differentiation of distal lung epithelium when squamous alveolar type I epithelial cells arise from columnar alveolar type II epithelial cells, and thinning of the mesenchyme possibly due to sustained proliferative activity during the canalicular/saccular phase. The concomitant occurrence of defects in two nearby cell types suggest an underlying failure of cell signaling during the final maturation of the lung.

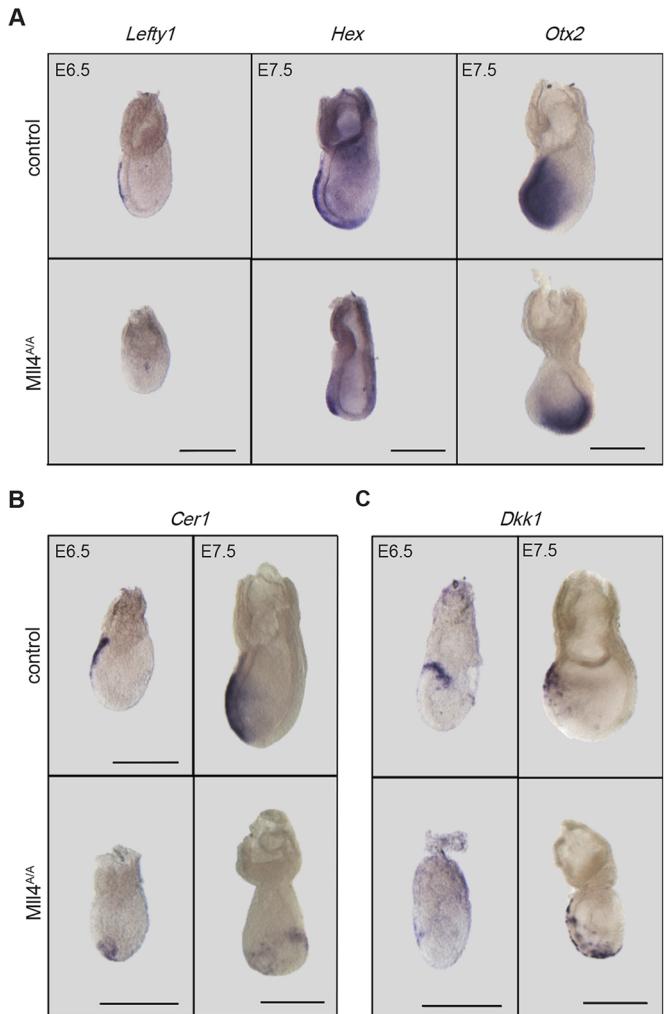


Fig. 5. *Mll4^{AA/A}* embryos show defective patterning of the AVE from E6.5 to E7.5. (A) *Lefty1* ($n=5/5$), *Hex* ($n=5/6$), *Otx2* ($n=2/3$) expression in control and *Mll4^{AA/A}* embryos between E6.5 and E7.75. (B,C) *Cer1* ($n=4/4$ for E6.5 and $n=5/5$ for E7.5) (B) and *Dkk1* ($n=2/2$ for E6.5 and $n=2/4$ for E7.5) (C) expression in control and *Mll4^{AA/A}* embryos between E6.5 and E7.5. Numbers (n) refers to mutant embryos showing the represented expression. Where possible, embryos are oriented with the anterior to the left. Scale bars: 250 μ m.

MLL4 and defective AVE migration

Collective cell migration of the AVE in the mouse embryo precedes gastrulation (Rossant and Tam, 2009) (Fig. 7B). Before migration, extra-embryonic 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 with *Mll4^{AA/A}*, characterized by failed AVE migration and lethality before E9.5 (Sugihara et al., 1998; Migeotte et al., 2010). NAP1 (NCKAP1), a component of the WAVE complex, acts downstream of RAC1 to control actin branching, and *Nap1* mutants do not establish the A-P axis owing 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,

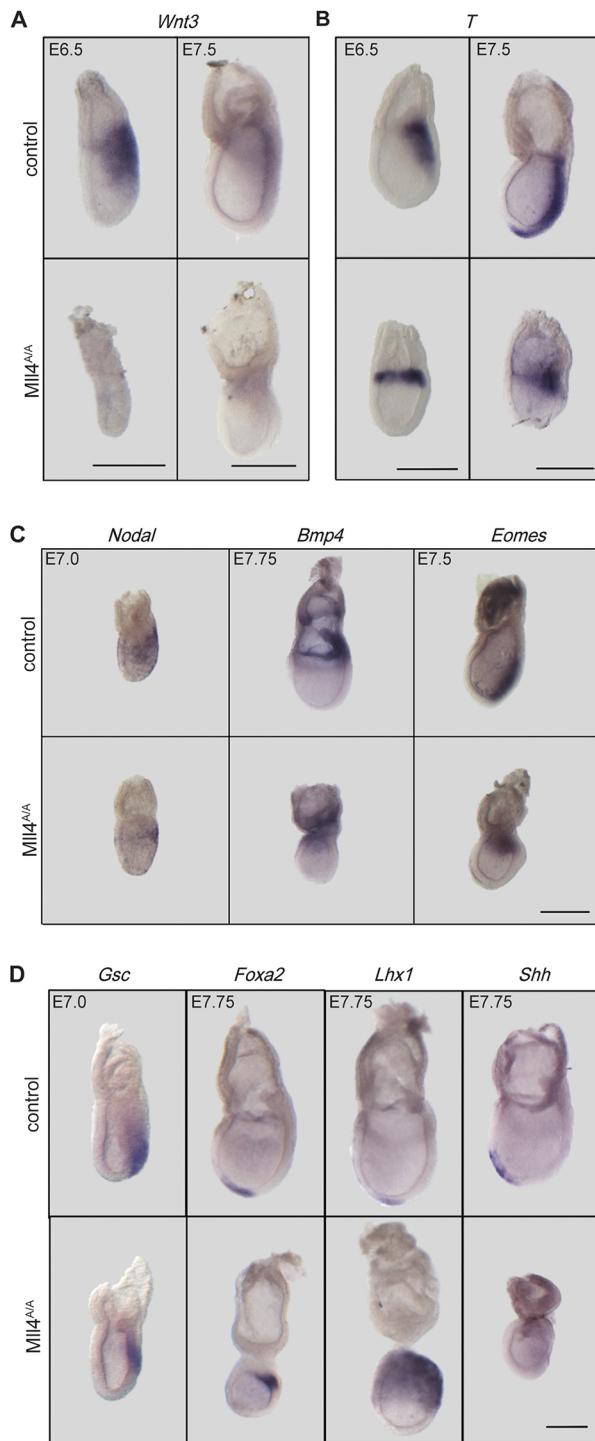


Fig. 6. Ectopic expression of primitive streak and node markers in *Mll4^{AA/A}* embryos. (A,B) *Wnt3* ($n=1/1$ for E6.5 and $n=2/4$ for E7.5) (A) and *T* ($n=1/1$ for E6.5 and $n=3/3$ for E7.5) (B) marks the length of primitive streak (PS) as it forms in control embryos from E6.5 to E7.5, but in *Mll4^{AA/A}* embryos they are expressed at the embryonic/extra-embryonic boundary and do not extend to the distal tip. (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^{AA/A}* embryos the expression is restricted to the embryonic/extra-embryonic boundary. (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^{AA/A}* embryos. The node derivative marker *Shh* ($n=3/3$) is completely absent in *Mll4^{AA/A}* embryos. Numbers (n) refers to mutant embryos showing the represented expression. Where possible, embryos are oriented with the anterior to the left. Scale bars: 250 μ m.

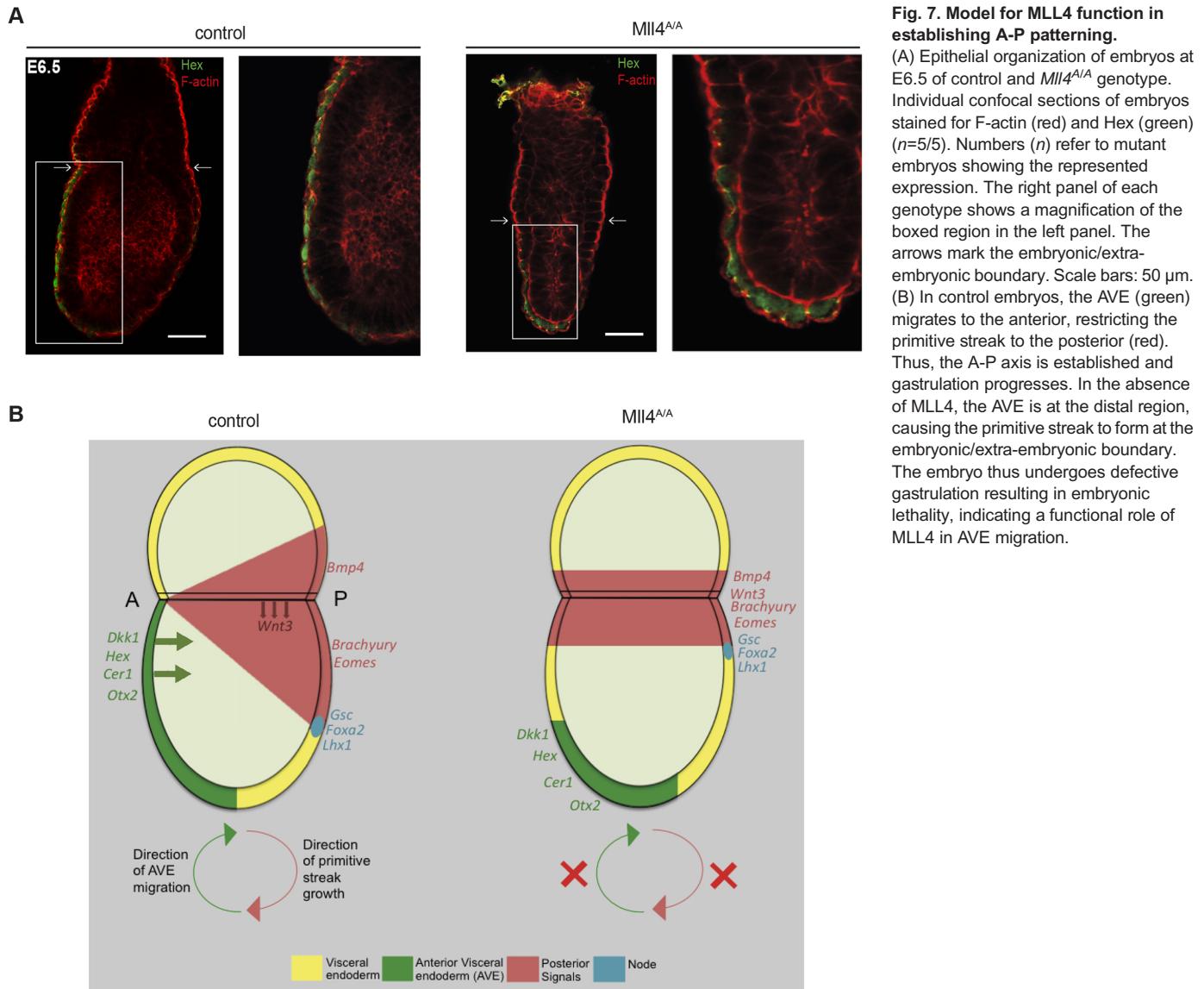


Fig. 7. Model for *MLL4* function in establishing A-P patterning.

(A) Epithelial organization of embryos at E6.5 of control and *Mll4^{AA}* genotype. Individual confocal sections of embryos stained for F-actin (red) and Hex (green) ($n=5/5$). Numbers (n) refer to mutant embryos showing the represented expression. The right panel of each genotype shows a magnification of the boxed region in the left panel. The arrows mark the embryonic/extra-embryonic boundary. Scale bars: 50 μ m. (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/extra-embryonic boundary. The embryo thus undergoes defective gastrulation resulting in embryonic lethality, indicating a functional role of *MLL4* in AVE migration.

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 owing 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/extra-embryonic boundary resulted in failed elongation of the primitive streak towards the distal tip, indicated by misexpression of *T*, *Wnt3*, *Bmp4* and *Eomes*. As a result, node markers *Gsc*, *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, 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

Mll4 has presented embryonic haploinsufficiency. Notably, *Drosophila Trr* also displays haploinsufficiency (Chauhan et al., 2012). Among the various explanations for haploinsufficiency, transcriptional synergy may be relevant for *MLL4*. Transcriptional synergy is based on co-operative 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 fixing the stability of decisions made by stochastic choices could explain the incomplete penetrance (Cook et al., 1998). Incomplete penetrance in a developmental choice indicates that *MLL4* does not make the choice but rather reduces the error rate by either stabilizing the choice or counteracting mistakes. A primary role for epigenetic regulation in choice stabilization and 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^{+/+}* 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; in females this 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-adenosylmethionine (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, 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 crucial to its function. *Mll4* haploinsufficiency reveals 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 one of the earliest collective cell migrations 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.

MATERIAL AND METHODS

Targeting constructs

The targeting constructs for *Mll3* and *Mll4* were generated using Red/ET recombinering (Fu et al., 2010). For *Mll3* an FRT-SA-GT0-T2A-lacZneo-CoTC-FRT-loxP cassette was inserted into intron 48. In addition, 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' 4.6 kb/3' 5 kb and 5' 4.7 kb/3' 4.9 kb for the *Mll3* and *Mll4* targeting constructs, respectively.

Gene targeting and generation of conditional knockout mice

Gene targeting in R1 ESCs was performed as described (Bledau et al., 2014). The 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. After germline transmission, *Mll3^{D/+}* mice were crossed to *CAGGs-Flopo* (Kranz et al., 2010) to generate *Mll3^{FD/+}* mice and then crossed to *PGK-Cre* (Lallemand et al., 1998) to produce *Mll3^{FDC/+}* mice. *Mll3^{D/+}* and *Mll3^{FDC/+}* mice were backcrossed to *C57BL/6JOLAhsd* mice (>15 generations). *Mll4^{A/+}* mice were backcrossed to CD1 mice (>15 generations). Primers for genotyping are provided in Table S1. All animal experiments were performed in accordance with German animal welfare legislation, and were approved by the relevant authority: the Landesdirektion Dresden.

Western blot, whole-mount X-gal staining and immunostaining

ESCs were homogenized in buffer E [20 mM HEPES (pH 8.0), 350 mM NaCl, 10% glycerol, 0.1% Tween 20, 1 mM PMSF, 1× 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 (QRPRFYPVSEELHRLAP, NGDEFDLLAYT, KQQLSAQTQRLAPS) (Table S4).

Embryos were dissected, fixed with 0.2% glutaraldehyde and X-gal stained as previously described (Kranz et al., 2010). Embryos and organs were dissected and fixed with 4% paraformaldehyde (PFA) overnight. Dehydration and paraffin infiltration utilized the Paraffin-Infiltration-Processor (STP 420, Zeiss). Dehydrated tissues were embedded in paraffin (Paraffin Embedding Center EG116, Leica) and sectioned. Antigen retrieval was performed by microwaving slides in 10 mM citrate buffer (pH 6.0) for 12 min (Microwave RHS 30, Diaphat). Immunohistochemical staining was performed as previously described (Bledau et al., 2014). Images were collected using an Olympus WF upright microscope. For immunofluorescence, sections were permeabilized in 0.5% Triton X-100 in PBS for 10 min, blocked for 1 h at room temperature (RT), incubated with primary antibody overnight at 4°C, followed by secondary antibody for 2 h at RT. Sections were mounted with Mowiol and imaged with Zeiss Laser Scanning Confocal Microscope LSM/780. Antibody information and dilutions are in Table S3.

Whole-mount *in situ* hybridization

Standard procedures were employed (Riddle et al., 1993; Piette et al., 2008). Digoxigenin-labelled riboprobes were: *T* (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* and *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 using a Nikon SMZ 1500 stereomicroscope.

Whole-mount immunofluorescence for HEX

PFA-fixed embryos were permeabilized in 0.5% Triton X-100 in PBS for 1 h 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 Biotechnology) secondary antibody. The embryos were imaged with Zeiss Laser 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.

Reverse transcription and real-time quantitative PCR (qRT-PCR) analysis

Total RNA was isolated using Trizol (Sigma-Aldrich) and reverse transcribed using the AffinityScript Multiple Temperature cDNA Synthesis kit (Agilent Technologies). qRT-PCR was performed with GoTaq qPCR Master Mix (Promega) using an 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).

mRNA expression profiling

Total RNA from E18.5 lung samples of control ($n=3$) and *Mll3*^{KO} ($n=3$) mice was purified using Trizol (Sigma-Aldrich) and quality ensured by using Bioanalyzer (Agilent) with the RNA 6000 Nano Kit (Agilent). mRNA was isolated from 1 µg total RNA by poly-dT enrichment using the NEBNext Poly(A) mRNA Magnetic Isolation Module (New England Biolabs) according to the manufacturer's instructions. Samples were then directly subjected to the workflow for strand-specific RNA-seq library preparation (Ultra Directional RNA Library Prep II, New England Biolabs). Resulting libraries were pooled in equimolar quantities for 75 bp single read sequencing on Illumina NextSeq500. FastQC (Babraham Bioinformatics) and RNA-SeQC (v1.1.8.1) (DeLuca et al., 2012) were used to perform a basic quality control on the resulting reads. Fragments were then aligned to the mouse genome (GRCm38/mm10) using GSAP (2019-06-10) (Wu and Watanabe, 2005; Wu and Nacu, 2010) and a table of read counts per gene was created based on the overlap of the uniquely mapped reads with the Ensembl Gene annotation v. 98 for mm10, using Feature Counts (v1.6.3) (Liao et al., 2014). Normalization of the raw read counts based on the library size and testing for differential gene expression between the different conditions was performed using the DESeq2 R package (v. 1.24.0) (Anders and Huber, 2010). Genes with an adjusted *P*-value (*padj*) ≤ 0.05 were considered as significantly differentially expressed, accepting this way a maximum of 5% false discoveries. To identify enrichment for particular biological processes associated with the DEGs, the DAVID GO/BP/FAT database (Huang da et al., 2009) was used. Gene set enrichment analysis was performed using GSEA software from the Broad Institute (Subramanian et al., 2005).

Acknowledgements

We thank Mandy Obst, Doris Müller, Isabell Kolbe, Madeleine Walker and Stefanie Weidlich for excellent technical assistance. We also thank the Biomedical Services of the Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Germany, 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, TX, USA), Dr Hans Schöler (Max Planck Institute for Molecular Biomedicine, Münster, Germany) and Dr Andrew P. McMahon (University of Southern California, CA, 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 Technische Universität Dresden, <http://biotp.tu-dresden.de/facilities/advanced-imaging/> assisted this research.

Competing interests

The authors declare no competing or financial interests.

Author contributions

Conceptualization: D. Ashokkumar, Q.Z., C.M., A.S.B., J.F., K.A., A.F.S., A.K.; Methodology: D. Ashokkumar, Q.Z., C.M., A.S.B., R.N., A.D., N.G., J.F., K.A., A.K.; Validation: D. Ashokkumar, Q.Z., A.K.; Formal analysis: D. Ashokkumar, Q.Z., D. Alexopoulou, N.G., A.F.S., A.K.; Investigation: D. Ashokkumar, Q.Z., C.M., A.S.B., J.F., K.A., A.K.; Resources: R.N., A.D., J.F., K.A.; Data curation: D. Alexopoulou; Writing - original draft: D. Ashokkumar, Q.Z., A.F.S., A.K.; Writing - review & editing: D. Ashokkumar, A.F.S., A.K.; Visualization: D. Ashokkumar, Q.Z., C.M.; Supervision: K.A., A.F.S., A.K.; Funding acquisition: D. Ashokkumar, A.F.S., A.K.

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 Technische Universität Dresden (to D. Ashokkumar).

Data availability

RNA-seq data have been deposited in GEO under accession number GSE146915.

Supplementary information

Supplementary information available online at <http://dev.biologists.org/lookup/doi/10.1242/dev.186999.supplemental>

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. doi:10.1242/dev.072728
- 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. doi:10.1242/dmm.003087
- Anders, S. and Huber, W. (2010). Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106. doi:10.1186/gb-2010-11-10-r106
- Andreu-Vieyra, C. V., Chen, R., Agno, J. E., Glaser, S., Anastassiadis, K., Stewart, A. F. and Matzuk, M. M. (2010). MLL2 is required in oocytes for bulk histone 3 lysine 4 trimethylation and transcriptional silencing. *PLoS Biol.* **8**, e1000453. doi:10.1371/journal.pbio.1000453
- 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. doi:10.1242/dev.132688
- 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. doi:10.1038/emboj.2011.194
- Banka, S., Veeramachaneni, R., Reardon, W., Howard, E., Bunstone, S., Ragge, N., Parker, M. J., Crow, Y. J., Kerr, B., Kingston, H. et al. (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. doi:10.1038/ejhg.2011.220
- Banka, S., Lederer, D., Benoit, V., Jenkins, E., Howard, E., Bunstone, S., Kerr, B., McKee, S., Lloyd, I. C., Shears, D. et al. (2015). Novel KDM6A (UTX) mutations and a clinical and molecular review of the X-linked Kabuki syndrome (KS2). *Clin. Genet.* **87**, 252-258. doi:10.1111/cge.12363
- Bannister, A. J. and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. *Cell Res.* **21**, 381-395. doi:10.1038/cr.2011.22
- Beddington, R. S. and Robertson, E. J. (1999). Axis development and early asymmetry in mammals. *Cell* **96**, 195-209. doi:10.1016/S0092-8674(00)80560-7
- 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 neuralizing activity expressed in the anterior primitive endoderm of the mouse gastrula. *Mech. Dev.* **68**, 45-57. doi:10.1016/S0925-4773(97)00125-1
- Benazeraf, B. and Pourquié, O. (2013). Formation and segmentation of the vertebrate body axis. *Annu. Rev. Cell Dev. Biol.* **29**, 1-26. doi:10.1146/annurev-cellbio-101011-155703
- Benjamin, J. S., Pilarowski, G. O., Carosso, G. A., Zhang, L., Huso, D. L., Goff, L. A., Vernon, H. J., Hansen, K. D. and Björnsson, H. T. (2017). A ketogenic diet rescues hippocampal memory defects in a mouse model of Kabuki syndrome. *Proc. Natl. Acad. Sci. USA* **114**, 125-130. doi:10.1073/pnas.1611431114
- Bledau, A. S., Schmidt, K., Neumann, K., Hill, U., Ciotta, G., Gupta, A., Torres, D. C., Fu, J., Kranz, A., Stewart, A. F. et al. (2014). The H3K4 methyltransferase Setd1a is first required at the epiblast stage, whereas Setd1b becomes essential after gastrulation. *Development* **141**, 1022-1035. doi:10.1242/dev.098152
- Bogershausen, N., Tsai, I. C., Pohl, E., Kiper, P. O., Beleggia, F., Percin, E. F., Keupp, K., Matchan, A., Milz, E., Alanay, Y. et al. (2015). RAP1-mediated MEK/ERK pathway defects in Kabuki syndrome. *J. Clin. Invest.* **125**, 3585-3599. doi:10.1172/JCI80102
- Brici, D., Zhang, Q., Reinhardt, S., Dahl, A., Hartmann, H., Schmidt, K., Goveas, N., Huang, J., Gahurova, L., Kelsey, G. et al. (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. doi:10.1242/dev.143347

- Burgold, T., Voituron, N., Caganova, M., Tripathi, P. P., Menuet, C., Tusi, B. K., Spreafico, F., Bevengut, M., Gestreau, C., Buontempo, S. et al. (2012). The H3K27 demethylase JMJD3 is required for maintenance of the embryonic respiratory neuronal network, neonatal breathing, and survival. *Cell Rep.* **2**, 1244-1258. doi:10.1016/j.celrep.2012.09.013
- Chauhan, C., Zraly, 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. doi:10.1242/dev.076687
- Chen, Y., Anastassiadis, K., Kranz, A., Stewart, A. F., Arndt, K., Waskow, C., Yokoyama, A., Jones, K., Neff, T., Lee, Y. et al. (2017). MLL2, not MLL1, plays a major role in sustaining MLL-rearranged acute myeloid leukemia. *Cancer Cell* **31**, 755-770.e756. doi:10.1016/j.ccr.2017.05.002
- 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. doi:10.1016/j.exphem.2018.10.001
- Chi, P., Allis, C. D. and Wang, G. G. (2010). Covalent histone modifications—miswritten, misinterpreted and mis-erased in human cancers. *Nat. Rev. Cancer* **10**, 457-469. doi:10.1038/nrc2876
- Cho, Y. W., Hong, T., Hong, S., Guo, H., Yu, H., Kim, D., Guszcynski, T., Dressler, G. R., Copeland, T. D., Kalkum, M. et al. (2007). PTIP associates with MLL3- and MLL4-containing histone H3 lysine 4 methyltransferase complex. *J. Biol. Chem.* **282**, 20395-20406. doi:10.1074/jbc.M701574200
- Choudhury, R., Singh, S., Arumugam, S., Roguev, A. and Stewart, A. F. (2019). The Set1 complex is dimeric and acts with Jhd2 demethylase to convey symmetrical H3K4 trimethylation. *Genes Dev.* **33**, 550-564. doi:10.1101/gad.322222.118
- Conlon, F. L., Lyons, K. M., Takaesu, N., Barth, K. S., Kispert, A., Herrmann, B. and Robertson, E. J. (1994). A primary requirement for nodal in the formation and maintenance of the primitive streak in the mouse. *Development* **120**, 1919-1928.
- Cook, D. L., Gerber, A. N. and Tapscott, S. J. (1998). Modeling stochastic gene expression: implications for haploinsufficiency. *Proc. Natl. Acad. Sci. USA* **95**, 15641-15646. doi:10.1073/pnas.95.26.15641
- DeLuca, D. S., Levin, J. Z., Sivachenko, A., Fennell, T., Nazaire, M. D., Williams, C., Reich, M., Winckler, W. and Getz, G. (2012). RNA-SeQC: RNA-seq metrics for quality control and process optimization. *Bioinformatics* **28**, 1530-1532. doi:10.1093/bioinformatics/bts196
- Denissov, S., Hofemeister, H., Marks, H., Kranz, A., Ciotta, G., Singh, S., Anastassiadis, K., Stunnenberg, H. G. and Stewart, A. F. (2014). Mll2 is required for H3K4 trimethylation on bivalent promoters in embryonic stem cells, whereas Mll1 is redundant. *Development* **141**, 526-537. doi:10.1242/dev.102681
- Dhar, S. S., Lee, S. H., Kan, P. Y., Voigt, P., Ma, L., Shi, X., Reinberg, D. and Lee, M. G. (2012). Trans-tail regulation of MLL4-catalyzed H3K4 methylation by H4R3 symmetric dimethylation is mediated by a tandem PHD of MLL4. *Genes Dev.* **26**, 2749-2762. doi:10.1101/gad.203356.112
- Dobson, C. L., Warren, A. J., Pannell, R., Forster, A. and Rabbitts, T. H. (2000). Tumorigenesis in mice with a fusion of the leukaemia oncogene Mll and the bacterial lacZ gene. *EMBO J.* **19**, 843-851. doi:10.1093/emboj/19.5.843
- Dorighi, K. M., Swigut, T., Henriques, T., Bhanu, N. V., Scruggs, B. S., Nady, N., Still, C. D., II, Garcia, B. A., Adelman, K. and Wysocka, J. (2017). Mll3 and Mll4 facilitate enhancer RNA synthesis and transcription from promoters independently of H3K4 monomethylation. *Mol. Cell* **66**, 568-576.e564. doi:10.1016/j.molcel.2017.04.018
- Dyle, M. C., Kolakada, D., Cortazar, M. A. and Jagannathan, S. (2019). How to get away with nonsense: Mechanisms and consequences of escape from nonsense-mediated RNA decay. *Wiley Interdiscip. Rev. RNA* **11**, e1560. doi:10.1002/wrna.1560
- Ernst, P. and Vakoc, C. R. (2012). WRAD: enabler of the SET1-family of H3K4 methyltransferases. *Brief Funct. Genomics* **11**, 217-226. doi:10.1093/bfgp/els017
- Ernst, P., Fisher, J. K., Avery, W., Wade, S., Foy, D. and Korsmeyer, S. J. (2004). Definitive hematopoiesis requires the mixed-lineage leukemia gene. *Dev. Cell* **6**, 437-443. doi:10.1016/S1534-5807(04)00061-9
- Faundes, V., Malone, G., Newman, W. G. and Banka, S. (2019). A comparative analysis of KMT2D missense variants in Kabuki syndrome, cancers and the general population. *J. Hum. Genet.* **64**, 161-170. doi:10.1038/s10038-018-0536-6
- Fu, J., Teucher, M., Anastassiadis, K., Skarnes, W. and Stewart, A. F. (2010). A recombining pipeline to make conditional targeting constructs. *Methods Enzymol.* **477**, 125-144. doi:10.1016/S0076-6879(10)77008-7
- Glaser, S., Schaft, J., Lubitz, S., Vintersten, K., van der Hoeven, F., Tufteland, K. R., Aasland, R., Anastassiadis, K., Ang, S. L. and Stewart, A. F. (2006). Multiple epigenetic maintenance factors implicated by the loss of Mll2 in mouse development. *Development* **133**, 1423-1432. doi:10.1242/dev.02302
- Glaser, S., Lubitz, S., Loveland, K. L., Ohbo, K., Robb, L., Schwenk, F., Seibler, J., Roellig, D., Kranz, A., Anastassiadis, K. et al. (2009). The histone 3 lysine 4 methyltransferase, Mll2, is only required briefly in development and spermatogenesis. *Epigenetics Chromatin* **2**, 5. doi:10.1186/1756-8935-2-5
- 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. doi:10.1002/ajmg.c.30051
- Hallson, G., Hollebakken, R. E., Li, T., Syrzycka, M., Kim, I., Cotsworth, S., Fitzpatrick, K. A., Sinclair, D. A. and Honda, B. M. (2012). dSet1 is the main H3K4 di- and tri-methyltransferase throughout Drosophila development. *Genetics* **190**, 91-100. doi:10.1534/genetics.111.135863
- 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. doi:10.1038/s41594-017-0013-5
- 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. doi:10.1002/bdra.20676
- 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. doi:10.1016/j.ydbio.2015.03.023
- 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.e1109. doi:10.1016/j.cell.2018.06.038
- Huang da, W., Sherman, B. T. and Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44-57. doi:10.1038/nprot.2008.211
- 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. doi:10.1016/j.jmb.2016.12.016
- 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. doi:10.1016/j.stem.2007.05.019
- Juriloff, D. M. and Harris, M. J. (2000). Mouse models for neural tube closure defects. *Hum. Mol. Genet.* **9**, 993-1000. doi:10.1093/hmg/9.6.993
- 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**, E22. doi:10.3390/jdb6030022
- Kantidakis, T., Saponaro, M., Mitter, R., Horswell, S., Kranz, A., Boeing, S., Aygun, O., Kelly, G. P., Mattheyses, N., Stewart, A. et al. (2016). Mutation of cancer driver MLL2 results in transcription stress and genome instability. *Genes Dev.* **30**, 408-420. doi:10.1101/gad.275453.115
- 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. doi:10.1016/j.molcel.2013.01.034
- 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. doi:10.1002/dvg.20641
- 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. doi:10.1023/A:1008868325009
- 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. doi:10.1016/j.ajhg.2011.11.021
- 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. doi:10.1002/ajmg.a.36442
- Lee, S., Lee, D. K., Dou, Y., Lee, J., Lee, B., Kwak, E., Kong, Y. Y., Lee, S. K., Roeder, R. G. and Lee, J. W. (2006). Coactivator as a target gene specificity determinant for histone H3 lysine 4 methyltransferases. *Proc. Natl. Acad. Sci. USA* **103**, 15392-15397. doi:10.1073/pnas.0607313103
- Lee, J. E., Wang, C., Xu, S., Cho, Y. W., Wang, L., Feng, X., Baldridge, A., Sartorelli, V., Zhuang, L., Peng, W. et al. (2013). H3K4 mono- and dimethyltransferase MLL4 is required for enhancer activation during cell differentiation. *Elife* **2**, e01503. doi:10.7554/elife.01503.027
- Li, B. E. and Ernst, P. (2014). Two decades of leukemia oncprotein epistasis: the MLL1 paradigm for epigenetic deregulation in leukemia. *Exp. Hematol.* **42**, 995-1012. doi:10.1016/j.exphem.2014.09.006
- Li, Y., Bögershausen, N., Alanay, Y., Simsek Kiper, P. Ö., Plume, N., Keupp, K., Pohl, E., Pawlik, B., Rachwalski, M., Milz, E. et al. (2011). A mutation screen in patients with Kabuki syndrome. *Hum. Genet.* **130**, 715-724. doi:10.1007/s00439-011-1004-y
- Li, Y., Han, J., Zhang, Y., Cao, F., Liu, Z., Li, S., Wu, J., Hu, C., Wang, Y., Shuai, J. et al. (2016). Structural basis for activity regulation of MLL family methyltransferases. *Nature* **530**, 447-452. doi:10.1038/nature16952

- Liao, Y., Smyth, G. K. and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930. doi:10.1093/bioinformatics/btt656
- 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. doi:10.1038/11932
- Livak, K. J. and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* **25**, 402-408. doi:10.1006/meth.2001.1262
- Meyer, C., Burmeister, T., Groger, D., Tsaur, G., Fechner, L., Renneville, A., Sutton, R., Venn, N. C., Emerenciano, M., Pombo-de-Oliveira, M. S. et al. (2018). The MLL recombinome of acute leukemias in 2017. *Leukemia* **32**, 273-284. doi:10.1038/leu.2017.213
- 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. doi:10.1371/journal.pbio.1000442
- Miller, T., Krogan, N. J., Dover, J., Erdjument-Bromage, H., Tempst, P., Johnston, M., Greenblatt, J. F. and Shilatifard, A. (2001). COMPASS: a complex of proteins associated with a trithorax-related SET domain protein. *Proc. Natl. Acad. Sci. USA* **98**, 12902-12907. doi:10.1073/pnas.231473398
- Mohan, M., Herz, H. M., Smith, E. R., Zhang, Y., Jackson, J., Washburn, M. P., Florens, L., Eisenberg, J. C. and Shilatifard, A. (2011). The COMPASS family of H3K4 methylases in Drosophila. *Mol. Cell. Biol.* **31**, 4310-4318. doi:10.1128/MCB.06092-11
- 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. et al. (2011). Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. *Nature* **476**, 298-303. doi:10.1038/nature10351
- Ng, S. B., Bigham, 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. et al. (2010). Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. *Nat. Genet.* **42**, 790-793. doi:10.1038/ng.646
- 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. et al. (2011). The genetic landscape of the childhood cancer medulloblastoma. *Science* **331**, 435-439. doi:10.1126/science.1198056
- Pasqualucci, L., Trifunovic, V., Fabbri, G., Ma, J., Rossi, D., Chiarenza, A., Wells, V. A., Grunn, A., Messina, M., Elliot, O. et al. (2011). Analysis of the coding genome of diffuse large B-cell lymphoma. *Nat. Genet.* **43**, 830-837. doi:10.1038/ng.892
- Piccolo, S., Agius, E., Leyns, L., Bhattacharya, 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. doi:10.1038/17820
- 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. doi:10.1038/nprot.2008.103
- Piunti, A. and Shilatifard, A. (2016). Epigenetic balance of gene expression by Polycomb and COMPASS families. *Science* **352**, aad9780. doi:10.1126/science.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.e1112. doi:10.1016/j.cell.2018.07.020
- 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. doi:10.1242/dev.02473
- Rao, R. C. and Dou, Y. (2015). Hijacked in cancer: the KMT2 (MLL) family of methyltransferases. *Nat. Rev. Cancer* **15**, 334-346. doi:10.1038/nrc3929
- 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. et al. (2017). Histone H3K4 monomethylation catalyzed by Trf and mammalian COMPASS-like proteins at enhancers is dispensable for development and viability. *Nat. Genet.* **49**, 1647-1653. doi:10.1038/ng.3965
- 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. doi:10.1016/0092-8674(93)90626-2
- 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. doi:10.1093/emboj/20.24.7137
- Rossant, J. and Tam, P. P. (2009). Blastocyst lineage formation, early embryonic asymmetries and axis patterning in the mouse. *Development* **136**, 701-713. doi:10.1242/dev.017178
- 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. doi:10.1038/nrm2298
- Schmitges, F. W., Prusty, A. B., Faty, M., Stutzer, A., Lingaraju, G. M., Aiwasian, J., Sack, R., Hess, D., Li, L., Zhou, S. et al. (2011). Histone methylation by PRC2 is inhibited by active chromatin marks. *Mol. Cell* **42**, 330-341. doi:10.1016/j.molcel.2011.03.025
- Schuetzengruber, 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. doi:10.1016/j.cell.2017.08.002
- Skarnes, W. C., Rosen, B., West, A. P., Koutsourakis, M., Bushell, W., Iyer, V., Mujica, A. O., Thomas, M., Harrow, J., Cox, T. et al. (2011). A conditional knockout resource for the genome-wide study of mouse gene function. *Nature* **474**, 337-342. doi:10.1038/nature10163
- Slany, R. K. (2009). The molecular biology of mixed lineage leukemia. *Haematologica* **94**, 984-993. doi:10.3324/haematol.2008.002436
- 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. doi:10.1016/j.molcel.2016.05.004
- Srinivas, S., Rodriguez, T., Clements, M., Smith, J. C. and Beddington, R. S. (2004). Active cell migration drives the unilateral movements of the endoderm. *Development* **131**, 1157-1164. doi:10.1242/dev.01005
- 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. doi:10.1038/nrm3789
- 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**, 20130546. doi:10.1098/rstb.2013.0546
- 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. doi:10.1242/dev.063537
- Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S. et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* **102**, 15545-15550. doi:10.1073/pnas.0506580102
- Sugihara, K., Nakatsuji, N., Nakamura, K., Nakao, K., Hashimoto, R., Otani, H., Sakagami, H., Kondo, H., Nozawa, S., Aiba, A. et al. (1998). Rac1 is required for the formation of three germ layers during gastrulation. *Oncogene* **17**, 3427-3433. doi:10.1038/sj.onc.1202595
- 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. doi:10.1002/gen.20012
- Thieme, S., Gyurfas, T., Richter, C., Ozhan, G., Fu, J., Alexopoulou, D., Muders, M. H., Michalk, I., Jakob, C., Dahl, A. et al. (2013). The histone demethylase UTX regulates stem cell migration and hematopoiesis. *Blood* **121**, 2462-2473. doi:10.1182/blood-2012-08-452003
- Treutlein, B., Brownfield, D. G., Wu, A. R., Neff, N. F., Mantalas, G. L., Espinoza, F. H., Desai, T. J., Krasnow, M. A. and Quake, S. R. (2014). Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. *Nature* **509**, 371-375. doi:10.1038/nature13173
- Van der Meulen, J., Speleman, F. and Van Vlierberghe, P. (2014). The H3K27me3 demethylase UTX in normal development and disease. *Epigenetics* **9**, 658-668. doi:10.4161/epi.28298
- Veitia, R. A., Cabreut, S. and Birchler, J. A. (2018). Mechanisms of Mendelian dominance. *Clin. Genet.* **93**, 419-428. doi:10.1111/cge.13107
- Voigt, P., LeRoy, G., Drury, W. J., III, Zee, B. M., Son, J., Beck, D. B., Young, N. L., Garcia, B. A. and Reinberg, D. (2012). Asymmetrically modified nucleosomes. *Cell* **151**, 181-193. doi:10.1016/j.cell.2012.09.002
- 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. doi:10.1186/s13148-015-0075-3
- Wu, T. D. and Nacu, S. (2010). Fast and SNP-tolerant detection of complex variants and splicing in short reads. *Bioinformatics* **26**, 873-881. doi:10.1093/bioinformatics/btq057
- Wu, T. D. and Watanabe, C. K. (2005). GMAP: a genomic mapping and alignment program for mRNA and EST sequences. *Bioinformatics* **21**, 1859-1875. doi:10.1093/bioinformatics/bti310
- 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. doi:10.1182/blood.V92.1.108.413k11_108_117
- Yamanaka, Y., Tamplin, O. J., Beckers, A., Gossler, A. and Rossant, J. (2007). Live imaging and genetic analysis of mouse notochord formation reveals regional morphogenetic mechanisms. *Dev. Cell* **13**, 884-896. doi:10.1016/j.devcel.2007.10.016
- Yap, K. L., Johnson, A. E. K., Fischer, D., Kandikatla, P., Deml, J., Nelakudit, V., Halbach, S., Jeha, G. S., Burrage, L. C., Bodamer, O. et al. (2019). Congenital

hyperinsulinism as the presenting feature of Kabuki syndrome: clinical and molecular characterization of 9 affected individuals. *Genet. Med.* **21**, 233-242. doi:10.1038/s41436-018-0013-9

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. doi:10.1016/j.str.2015.07.018

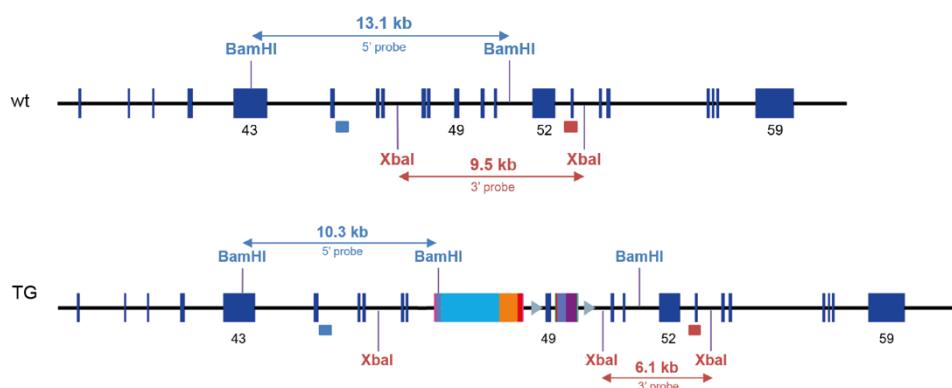
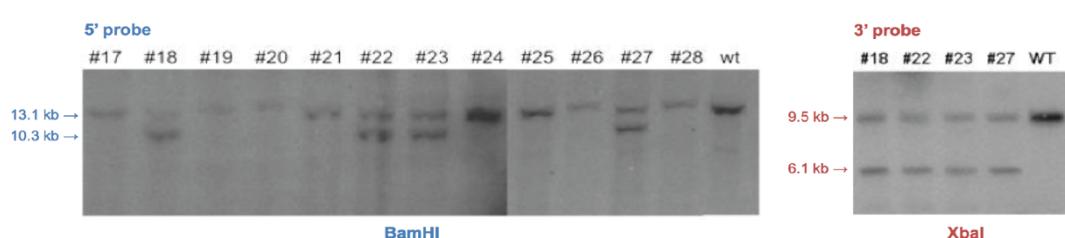
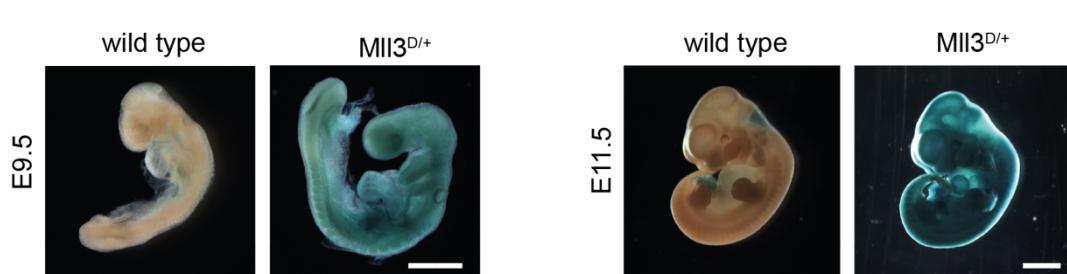
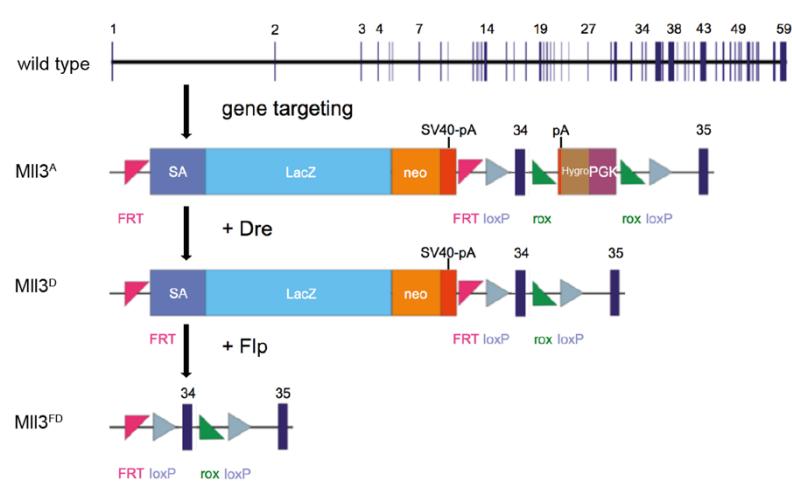
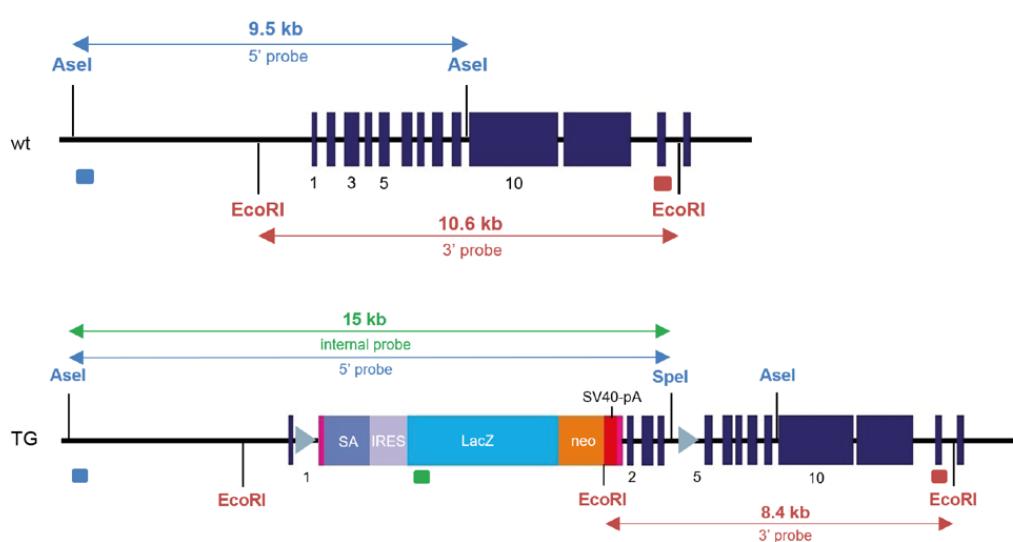
Figure S1**A****B****C****D**

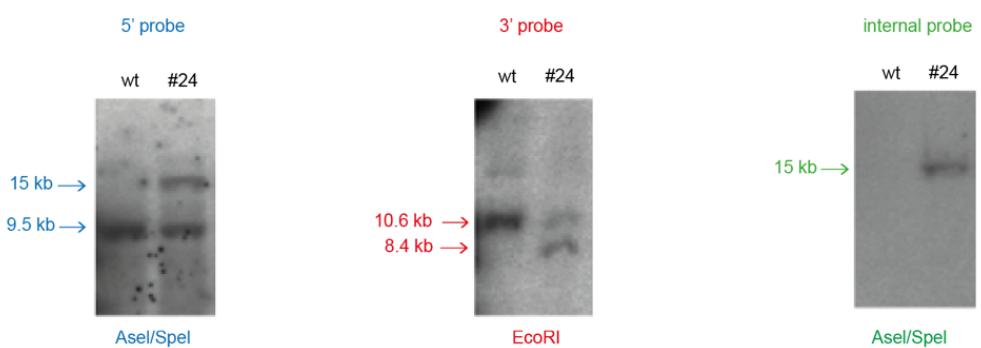
Fig. S1. Gene targeting strategy for *MII3*. (A) Schematic representation of the Southern strategy to identify correctly targeted events in the *MII3* locus using 5' (light blue box) and 3' (red box) external probes. Blue boxes with numbers underneath indicate exons. (B) Southern blot analysis using 5' and 3' external probes. Clones #18, #22, #23 and #27 carried the correctly targeted event. (C) Wild type and *MII3*^{D/+} embryos at E9.5 and E11.5 were stained for β-galactosidase. Scale bar 500 μm. (D) Diagram of the *MII3* knockout first allele (*MII3*^A) with floxed exon 34. Instead of the PGK-Blasticidin-PolyA a PGK-Hygromycin-PolyA cassette was inserted. For abbreviations see Fig. 1B.

Figure S2

A



B



C

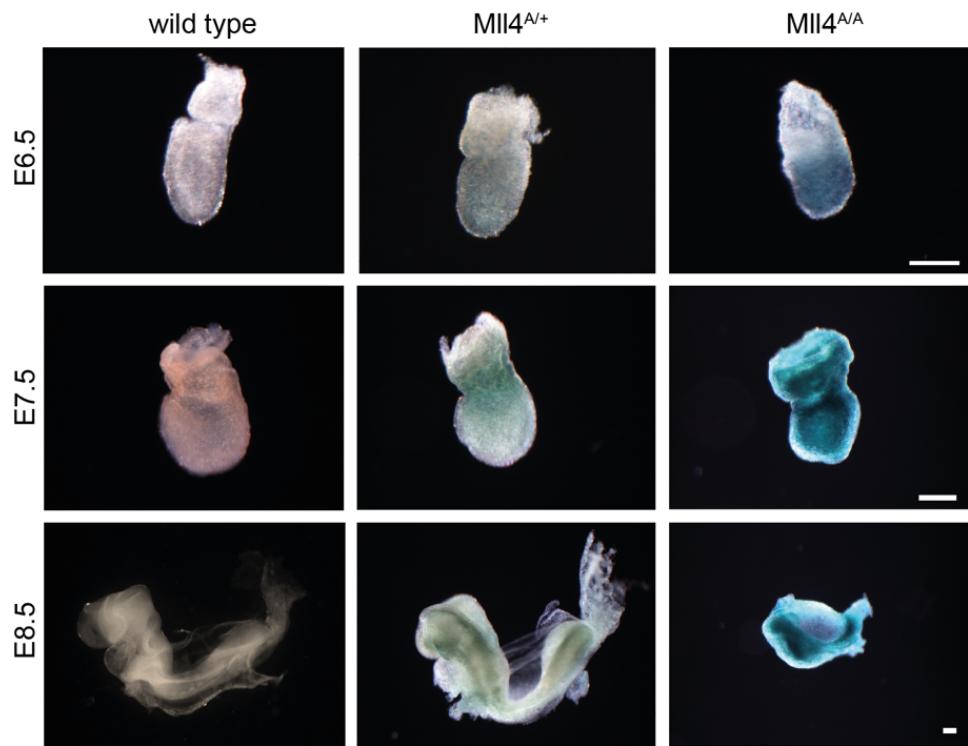


Fig. S2. Gene targeting strategy for *Mll4*. (A) Schematic representation of the Southern blot strategy to identify correctly targeted events in the *Mll4* locus using 5' (light blue box), 3' (red box) external probes and the internal LacZ (green box) probe. Blue rectangles with numbers underneath indicate 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*^{A/+} embryos at E6.5, E7.5 and E8.5 were stained for β -galactosidase. Scale bar 250 μ M.

Figure S3

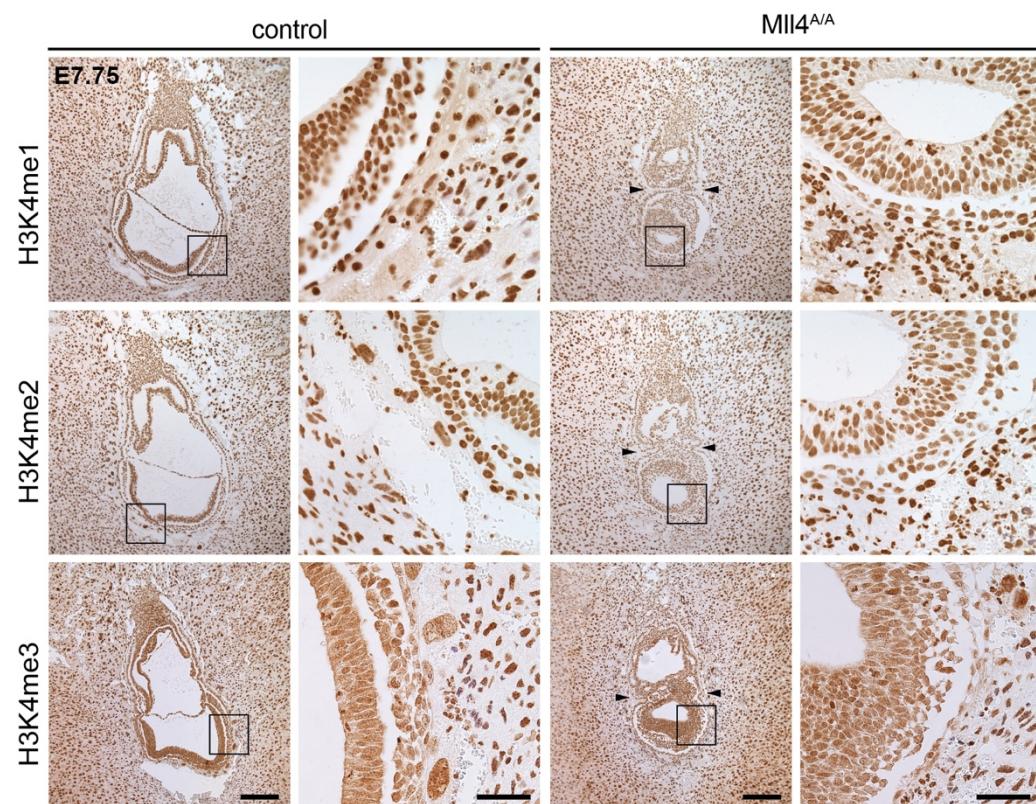


Fig. S3. No global changes in H3K4 methylation levels. Immunohistochemistry on sections of wild type and *Mll4^{A/A}* E7.75 embryos *in utero* with H3K4me1-, H3K4me2- and H3K4me3-specific antibodies. The boxed region is magnified on the right. Scale bars 50 μ m and 200 μ m, respectively.

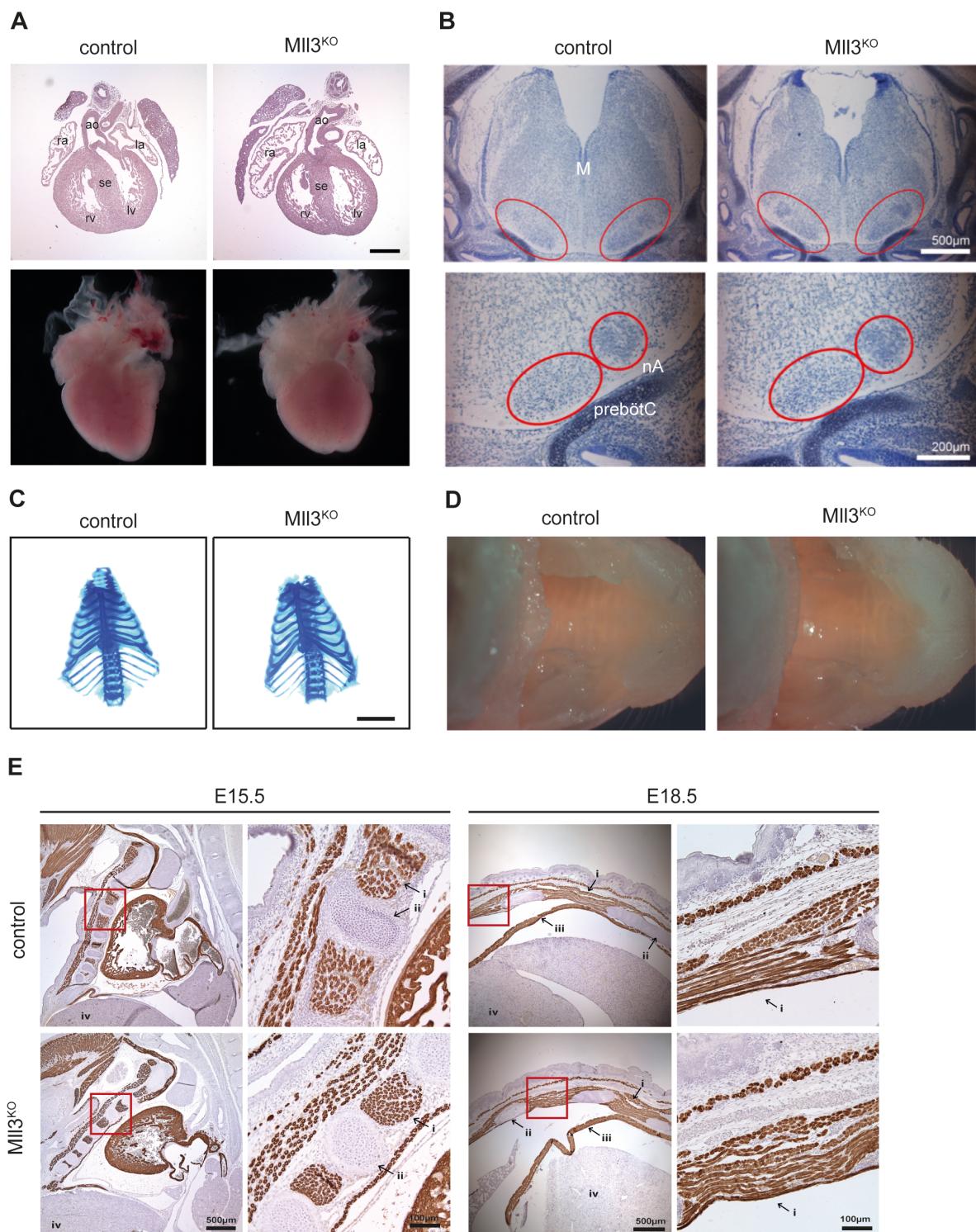
Figure S4

Fig. S4. No morphological differences in *MII3* 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. (E) MyHC immunostaining demonstrates normally developed intercostal muscles and diaphragm between control and *MII3^{KO}* littermates. The boxed region is magnified on the right. i, intercostal muscle; ii, rib cage; iii, diaphragm; iv, liver.

Figure S5

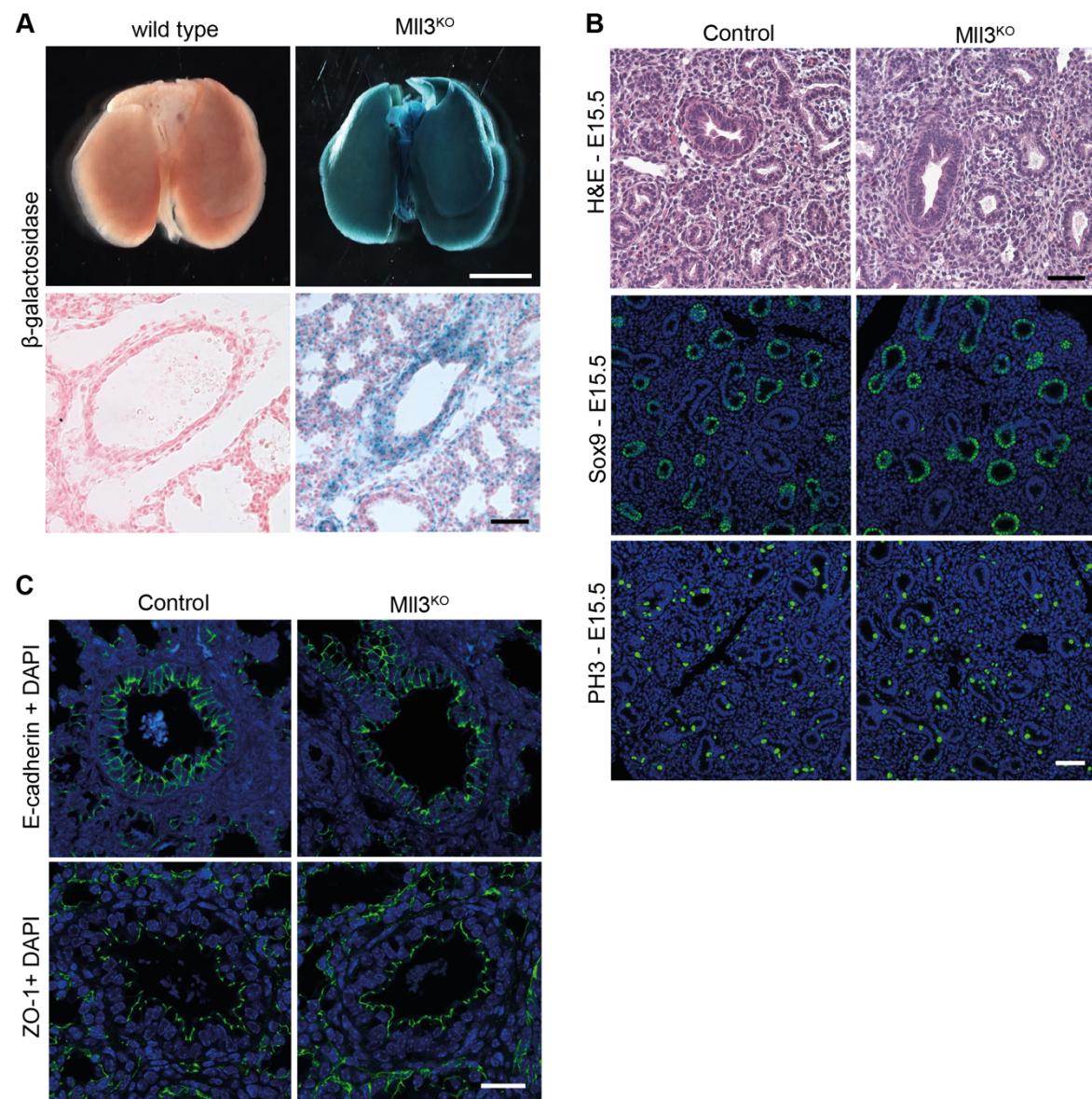
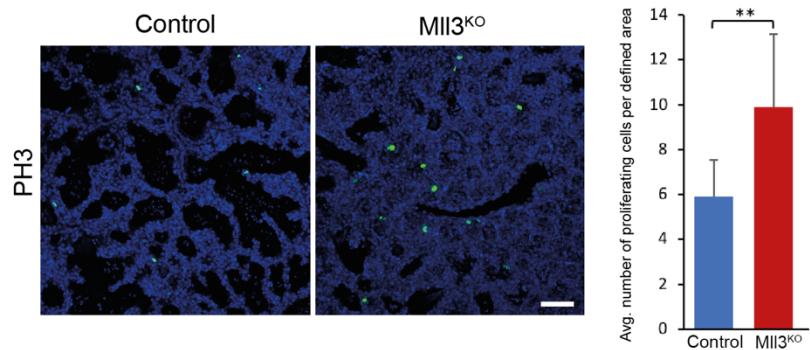


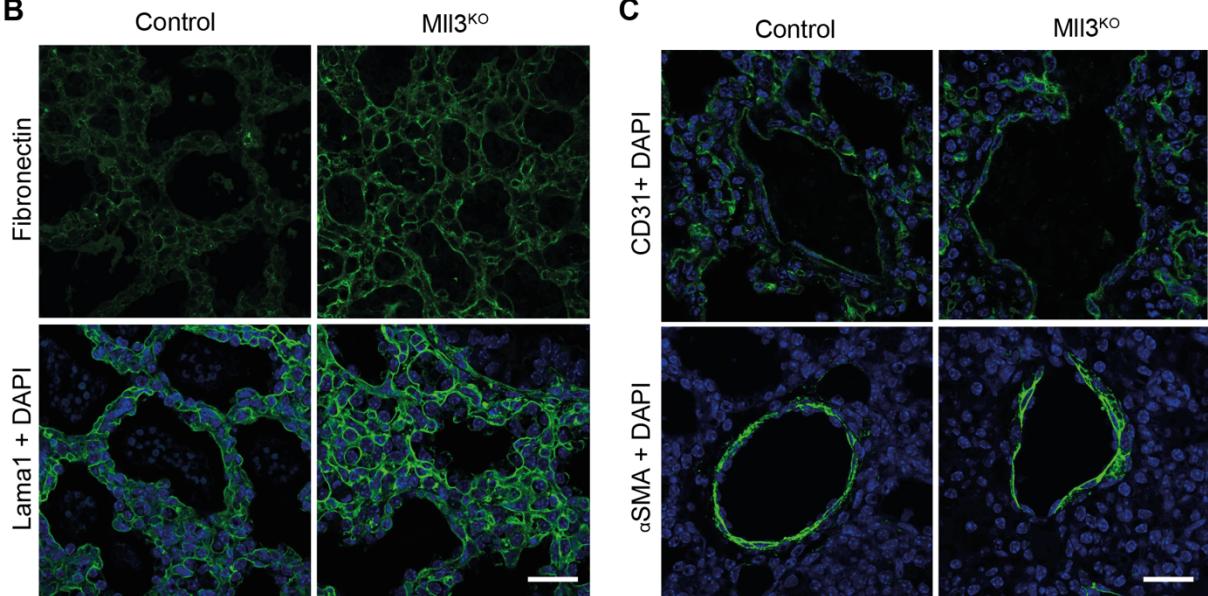
Fig. S5. Morphology of the developing lung is not altered. (A) β -galactosidase staining reflecting MLL3 expression on whole lung (top) and lung sections (bottom) at E18.5. Scale bar 500 μ m (whole lung) and 250 μ m (lung section). (B) H&E staining on sections of E15.5 lungs showed no morphological differences between control and *MII3*^{KO} littermates. Sox9 expression is seen in the distal epithelial tips of control and *MII3*^{KO} lungs. No difference in proliferation was observed between control and *MII3*^{KO} lungs of E15.5 fetuses after PH3 staining. Scale bar 250 μ m. (C) E-cadherin expression in the proximal lung epithelium and ZO-1, a tight junction marker, remain unchanged in control and *MII3*^{KO} littermates at E18.5. Scale bar 25 μ m.

Figure S6

A



B



C

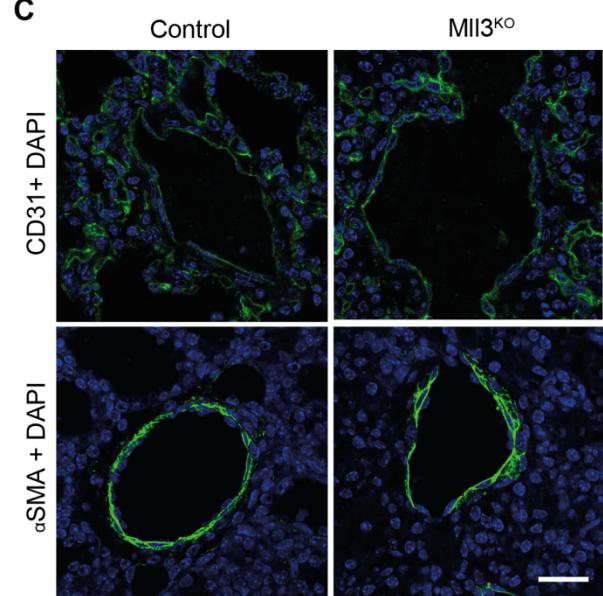


Fig. S6. Increased proliferation and extensive deposition of extracellular matrix proteins in MII3^{KO} lungs. (A) At E18.5 an increase in proliferation is observed in MII3^{KO} lungs. Scale bar 250 μ m. PH3 positive cells from E18.5 lung sections from control (n=3) and MII3^{KO} (n=3) lungs were counted. Four defined areas (each 1.8 mm²) were analysed. Mean \pm s.d. is shown (**p < 0.01 as calculated by unpaired t-test). (B) Extracellular matrix protein deposition of fibronectin and Lama1 in the basal lamina of MII3^{KO} lung is more compared to control littermate. Scale bar 25 μ m. (C) Expression CD31 in endothelial cells of the large blood vessels and capillaries in control and MII3^{KO} lungs. Smooth muscle surrounding the large blood vessels marked by expression of α SMA in control and MII3^{KO} lungs. Scale bar 25 μ m.

Figure S7

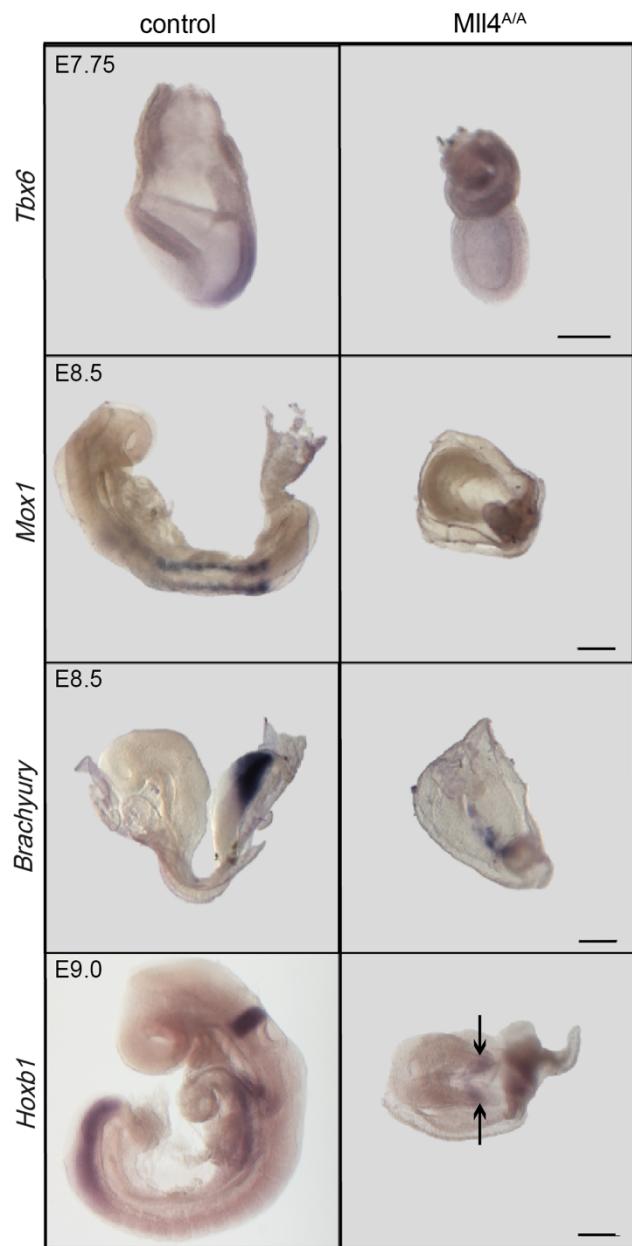


Fig. S7. Defective gastrulation after loss of MLL4. Whole mount *in situ* hybridization of *Tbx6* (paraxial mesoderm marker) at E7.75, *Mox1* (somite marker) at E8.5, *Brachyury* at E8.5, and *Hoxb1* at E9.0 on embryos from *Mll4^{A/+}* intercrosses. Arrows point towards the prospective rhombomere 4. Scale bars 250 μ m.

Figure S8

A

Mll4^{A/+}



B

wild type

Mll4^{A/+}

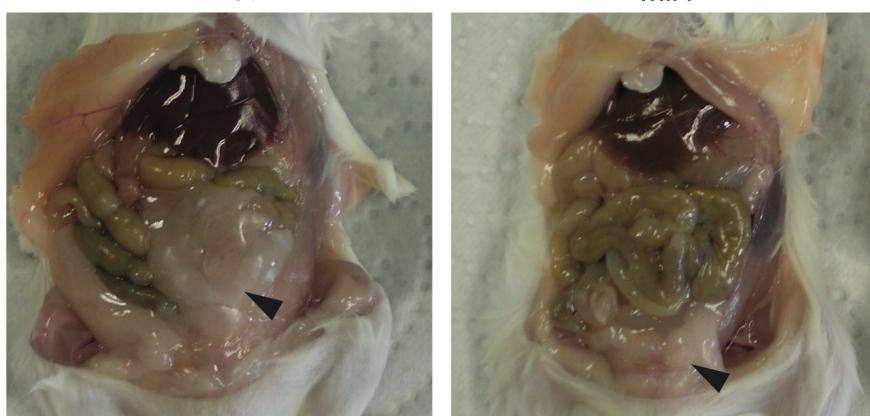


Fig. S8. Anencephaly and decreased fat tissue in *Mll4^{A/+}* mice. (A) During the late stages of gestation, exposed neural folds degenerate in the *Mll4^{A/+}* embryos with exencephaly, resulting in anencephaly at P1. Fetuses with such disorder were found dead at P1. (B) A seven-week-old female *Mll4^{A/+}* mouse shows less white adipose tissue (arrow head) than a female wild-type littermate.

Table S1. Primers for genotyping

Primer pairs	Sequence (5' – 3')	Product in bp
Genotyping of MII3D knockout		
MII3 LoxP1	CAAATCCTCCCAGATTAGCAAC	361 (wt);
MII3 LoxP2	CAACAATGCACTATCTATGGCTTC	499 (D)
Genotyping of MII3FD		
MII3 flp se	GAGAGTGAGTGAGAGAAAACC	765 (wt);
MII3 ex49 as	AGGATTGAGATCCTGGTGG	933 (FD)
Genotyping of MII3 wild type		
MII3 ex49 se	CCACCAGGATCTGCAAATCCT	374 (wt);
MII3 LoxP2	CAACAATGCACTATCTATGGCTTC	
Genotyping of MII3FDC knockout		
MII3 flp se	GAGAGTGAGTGAGAGAAAACC	1106 (wt);
MII3 LoxP2	CAACAATGCACTATCTATGGCTTC	1423 (FD); 298 (FDC)
Genotyping of Cre		
19Cre se	GCCTGCATTACCGGTGATGCAACGA	700 (Cre)
20Cre as	GTGGCAGATGGCGCGGAAACACCATT	
Genotyping of MII4 knockout (embryos)		
MII4_5lox d	ACTATGTGGTGTGGACAAGGCC	169 (wt);
MII4_3lox d	CGCTGCAAGGATGGGCAGAGCC	297 (targeted)
Genotyping of MII4 knockout (tails)		
MII4_se	CCAGAACTATGTGGTGTGGAC	140 (wt);
MII4_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)
Aqp5 se	GTGGTCATGAATCGGTTAGC	167
Aqp5 as	CAGTCCTCCTCTGGCTCATAT	
Clic5 se	GCCAGAGGAGATTGACACCAA	159
Clic5 as	GCTGGGATGTCATAGTTCGG	
Hopx se	GGTGGAGATCCTGGAGTACAA	170
Hopx as	ATCTGCATTCCGAAGGCAAGC	
Pdpn se	CGTCGGAGGGATCTTCATTGT	123
Pdpn as	AGGGCAAGTTGGAAGCTCTCT	
Rpl19 se	CCGAGAGAACACCAAAAGGA	147
Rpl19 as	CTTCTCAGGCATCCGAGCATT	

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
Mouse anti-ZO-1	Thermo Scientific, 33-9100	1:500 IF
Mouse anti-MyHC	Hybridoma Bank, A4.1025	1:20 IHC
Rat anti-E-cadherin	Invitrogen, 13-1900	1:50 IF
Rat anti-CD31	BD Biosciences, 01951A	1:100 IF
Secondary antibodies		
Biotinylated goat anti-rabbit IgG	Vector, BA-1000	1:500 IHC
Biotinylated goat anti-mouse IgG	Vector, BA-9200	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

Table S4. Alignment of murine MLL3 and MLL4

M113	MSSEEDRSAEQQQPPPAPPEPGAPAPSAAADKRPRGRPRKDASPQQRARKKPRSRGK	60
M114	-----MDSQKPPAED-----	10
M113	STVEDEDSMGLETTETENIVETEIKEQSVEEDAETEVDSKQPVSAALQRSVSEESANSL	120
M114	--KDSDA <u>P</u> ADGLAAPEKPGATEPDLPILCIGEVSVPGSG--GSRPQ	52
M113	VSVGVEAKIS <u>E</u> QLCAF C YC G E K SSL G Q D L K Q F R V T P W K D Q SN K IDDDNSSGTC	180
M114	KPPHDCSRGPARRCALNCGE P GLHG Q REL Q RFELPS---DWPGFPVPV--SGGNNSGPC	106
M113	EKIQNYAPRKORG R KERPPQOQSAVSCVSVSTQTACEDQAGKLWDELSLVGLPDAIDVQA	240
M114	EAVL-----PKEDASQIGFPEGLTPAH	128
M113	ePHD1	
M113	LFDSIT <u>G</u> TCWAHHRCVEWSLGICQMEPLLVNVDKAVVSGSTERCAFCKHLGATIKCEEK	300
M114	LGEPGG H CAWWHC A WSAGVG W QEGPELCGV D KAIFSG I SORCSH C ARE G AS V PCRSPG	188
M113	CTQMYHYPCAAGAGTFQDFSHFFLLCPEHIDQAPERS <u>S</u> KE <u>D</u> ANCAVCDSPGDLLDQFFCT	360
M114	CSRLYHFPCATASGSFLSMKTLQLLCPEHSDGAA--HLE <u>A</u> RCACVCEGPG G OLCDLLFC <u>S</u>	246
M113	PHD	
M113	CGQHYHGMCLDIATPLKRA G WOC P EC R VCQNCK O SGEDSKMLVCDTCDKG Y HTFCLQP?	420
M114	CGHHYHGACLD T ALT A RK R AS W OC P EC R VC Q SCR K PGND S KMLVC E TC D KG Y HTFCLKP?	306
M113	PHD	
M113	MKSVP <u>TNG</u> WKCKNCRICIECGTRS-----STQWHHNCLICDTCYQQQDNLCP-----	467
M114	MEDLPAHSWKCK C RCLCRACGAGSAELNPNSEWFENYSLCHRCHKAQGSQPTSVAEQH?	366
M113	-FCGKCYHPEL <u>Q</u> KD-----MLHCNM-----CKRWVHLECDKPTDQEL <u>S</u> Q <u>L</u>	507
M114	AVCS <u>R</u> LSPP <u>E</u> GEIPIDAPD A LYVAC Q QPKGGHVTS Q PKEL A PL Q CEAK P l <u>G</u> RA <u>G</u> T <u>Q</u> L	426
M113	KEDYICM Y <u>C</u> KHL-----GAEIDPLHPGNEVEMPELPT <u>D</u> YASGM-----	545
M114	EAQ L EAPLHE E MP <u>L</u> PP <u>E</u> ES <u>P</u> LSPP <u>E</u> EE-SPT <u>S</u> PP <u>E</u> AS <u>R</u> LS <u>S</u> PP <u>E</u> ES <u>P</u> LS <u>S</u> PP <u>E</u> SSPF	485
M113	-EIEGTEDE-----VVFLEQTVNKVDSDH <u>Q</u> -----CR-----P-----	572
M114	SPLEGCPP <u>S</u> PAL <u>D</u> T <u>P</u> LSPP <u>E</u> AS <u>P</u> LSPP <u>E</u> ES <u>P</u> LSPP <u>E</u> EL <u>P</u> SSPP <u>E</u> AS <u>R</u> LS <u>S</u> PP <u>E</u> ES <u>P</u>	545
M113	-GIVPD <u>A</u> Q V Q-----Y T EE <u>P</u> -----Q K SN <u>P</u> LES <u>P</u> D <u>T</u> V <u>G</u> LI <u>T</u> S ESSDN <u>K</u> M <u>N</u> P <u>D</u> -----	613
M114	MS <u>S</u> PP <u>E</u> SP <u>M</u> S <u>P</u> PP <u>E</u> AS <u>R</u> LS <u>F</u> P <u>F</u> E <u>E</u> PL <u>S</u> PP <u>E</u> PL <u>S</u> PP <u>E</u> AS <u>R</u> LS <u>S</u> PP <u>E</u> PD <u>S</u> PM <u>S</u> PP <u>E</u>	605
M113	-L-----	614
M114	DSP <u>M</u> S <u>P</u> PP <u>E</u> VS <u>R</u> FL <u>P</u> L <u>P</u> V <u>L</u> SH <u>L</u> SP <u>L</u> PE <u>V</u> RS <u>R</u> LS <u>S</u> PP <u>E</u> PL <u>S</u> PP <u>E</u> PD <u>S</u> PA <u>S</u> PP <u>E</u> AS <u>R</u> LS <u>P</u>	665
M113	-ANEIA-----HEVDTEKTEM <u>L</u> SKGRHVCEEDQNE-----	643
M114	PP <u>E</u> D <u>S</u> P <u>A</u> PP <u>E</u> AS <u>R</u> LS <u>R</u> P <u>E</u> D <u>S</u> P <u>A</u> PP <u>E</u> DS <u>L</u> V <u>S</u> LP <u>M</u> E <u>E</u> PL <u>S</u> PL <u>P</u> E <u>E</u> EL <u>R</u> IC <u>P</u> OE <u>E</u> P <u>Y</u>	725
M113	-DRME-VTENIE-----VLPHQTIV <u>P</u> Q <u>E</u> D <u>LL</u> -----	667
M114	LSP <u>Q</u> P <u>E</u> E <u>P</u> RL <u>C</u> P <u>Q</u> PE <u>E</u> PL <u>S</u> P <u>Q</u> SE <u>E</u> P <u>C</u> L <u>S</u> P <u>V</u> U <u>V</u> EP <u>G</u> P <u>S</u> Q <u>E</u> E <u>P</u> PH <u>L</u> S <u>P</u> V <u>P</u> Q <u>E</u> PH <u>L</u> S <u>P</u> Q <u>E</u>	785
M113	-LSEDSEVA-----SKEL-----SPP <u>K</u> S <u>A</u> P <u>E</u> T <u>A</u> PE-----	692
M114	EP <u>H</u> L <u>S</u> P <u>O</u> PO <u>O</u> L <u>H</u> L <u>S</u> P <u>H</u> SE <u>E</u> P <u>C</u> L <u>S</u> P <u>M</u> PE <u>E</u> P <u>C</u> L <u>S</u> P <u>O</u> PE <u>E</u> L <u>N</u> G <u>P</u> PO <u>A</u> P <u>E</u> P <u>P</u> E <u>P</u> SO <u>S</u> AP <u>K</u> E <u>L</u>	845
M113	ALLSPHSERSLS-----CKEPLLTERV-----	714
M114	SLFSPSGE <u>P</u> PL <u>P</u> P <u>M</u> GE <u>P</u> AL <u>S</u> E <u>P</u> GE <u>P</u> PL <u>S</u> PL <u>P</u> E <u>E</u> PL <u>S</u> L <u>S</u> GE <u>P</u> V <u>L</u> S <u>P</u> OL <u>M</u> P <u>D</u> PL <u>PP</u> PL <u>S</u>	905
M113	-QEEMEQ K ENSEFS-----TGC <u>V</u> DEF <u>E</u> MT <u>L</u> -----	737
M114	PI <u>I</u> P <u>A</u> A <u>P</u> P <u>A</u> L <u>S</u> PL <u>G</u> E <u>L</u> E <u>P</u> FG <u>A</u> K <u>G</u> D <u>S</u> D <u>P</u> E <u>S</u> PL <u>A</u> API <u>E</u> TP <u>I</u> S <u>P</u> P <u>E</u> AN <u>C</u> T <u>D</u> P <u>E</u> P <u>V</u> P <u>P</u> MI	965
M113	-AVD <u>S</u> CD-----KD <u>S</u> SC <u>Q</u> G <u>D</u> KY-VELP-----A <u>E</u> E-----	760
M114	LPP <u>S</u> PG <u>S</u> PL <u>G</u> P <u>A</u> SP <u>I</u> L <u>M</u> ER <u>L</u> PP <u>P</u> C <u>S</u> PL <u>P</u> H <u>S</u> L <u>P</u> PP <u>T</u> PP <u>S</u> H <u>C</u> SP <u>P</u> AL <u>P</u> L <u>S</u> V <u>P</u> S <u>P</u> LS <u>P</u> V <u>Q</u> K	1025
M113	ESTFSSATDLNKADVS- <u>S</u> S <u>S</u> T <u>I</u> L <u>C</u> S <u>D</u> -----LP <u>S</u> C <u>D</u> ML <u>H</u> G <u>Y</u> P-----	795
M114	AVDVSDE <u>A</u> E <u>L</u> HE <u>E</u> MT <u>D</u> K <u>G</u> P <u>E</u> P <u>C</u> PA <u>L</u> E <u>P</u> R <u>A</u> T <u>S</u> PL <u>P</u> S <u>P</u> LG <u>D</u> L <u>S</u> C <u>P</u> A <u>P</u> S <u>P</u> A <u>P</u> AL <u>D</u> D <u>F</u> S <u>G</u> L <u>G</u> E	1085
M113	-PAFNSAAGSI-----MPTTYISVTP-----KIGMGKPAIT-----	825
M114	DTAPLDGTQMSG S LAGELKG S PVLLDPEELTPVTPMEVYG E CK Q AG Q GS P CEE Q EEPG	1145
M113	-KR-----KFS-----PGRPRSKQG-AWSNH----NTVS-----	848
M114	APMAPMPPTLI <u>K</u> S <u>D</u> IV <u>N</u> E <u>I</u> S <u>N</u> L <u>S</u> Q <u>G</u> D <u>A</u> S <u>A</u> S <u>F</u> PG <u>S</u> E <u>P</u> LL <u>G</u> S <u>P</u> D <u>E</u> GG <u>G</u> S <u>L</u> S <u>M</u> E <u>L</u> GV <u>S</u> D <u>V</u> S	1205

M113	PP-----SWAP----DTSEG-REIFKPRQLSGSAIWSIKVGRGSGFP	885
M114	PARDEGSLRLCTDSLPETDDSLLCDTGTATSGGKAEGDKGRRRSPARSRIKQGRSSSFP	1265
M113	GKRRPRGAGLSSGRGRSKLKSGIGAVVL--PGVSAADISSNKDEEENSMHNTVVLFSS	943
M114	GRRRPRGAAHGGGRGRARLKSTTSSVETLVADIDSSPSKEEEEEDDTMQNTVVLFSN PHD	1325
M113	SDKFTLCQDMCVVCGSGFGQGAEGRLLACSQCGQCYHPYCVSICKTVVLSKGWRCLCT'7	1003
M114	TDKFVLMQDMCVVCGSGFGRAEGHLLACSQCSQCYHPYCNSKICKVMLLKGWRCECI'7	1385
M113	CEACGKATDPGRLLCDDCDISYHTYCLDPPLQTVPKGWKCKWCVWCRHCGATSAGLRC	1063
M114	CEVCQASDPSRLLCDDCDISYHTYCLDPPLTVPKGWKCKWCVSCMOCGAASPGFHCPHD	1445
M113	EWQNNTYQCAPCASISSCPVCCRNYREEDILILQCRQCDRWMAHVQNLNTEEEVENVAD	1123
M114	EWQNSYTHCGPCASIWTCPVCHAPYVEEDILLIQCRHCERWMHAGCESLFTEDEVQAADI	1505
M113	GFDCSMCRPYMPVSNVPSDCCDSSLVAQIVTKVKELDPKTYTQDGVCILTESGMSQLQS	1183
M114	GFDCVSCQPYVVVKPVVPVAPP---EL---VPVKV-EPEPQFFRFEGVWLTTETGMAVRN	1558
M113	LTVTAPRRKRTPKLKLKIINQNSVAVLQTPPDIQSEHSDRGEMDD-----SREGEIMD	1237
M114	LTMSPLHKRRQR-RGRILGLPGEAGLEGSEPSDALGPDDKKDGDLDTDDLKGEVVGEQME	1617
M113	CDGKSESSPEREAGDETKGIEGTDAIKKRKRKPYRPGIGGMVVRQRSRTGQGKAKRSVV	1297
M114	CEIKLEG---PASPDVELGKEETEESKKRKRKPYRPGIGGMVVRQRSHTRVKRGPAQ	1673
M113	RKDSSGSISEQLPSRDDGWREQLPDTLVDE---PVSVAEN-TDKIKKRYRKRNKLEETFAE	1353
M114	-----VLSGDGQPDEVMPADLPAEGSVEQSLAEGDEKKQORRARKKSKEQHLC	1724
M113	PAYLQEAFFGKDLLTSRQNKLVDNLSEAAQLSFKTGFLLPSSDPLLSSSTSAPGT	1413
M114	PAYLQEAFFGKDLLSRKALFAVGVG---RPGFGLGASKPRAD-----	1765
M113	QGTADDPLADISEVLNTDDDILGIISDDLAKSVDHSDIGPT-TADASSLPQPGVSQSSRP	1472
M114	-GGSDR--KELMTAMHKGDDGP----DVA---DEESHGPEGTADLPGLEGGVKASPVP	1814
M113	LT----EEQLDGILSPELDKMTDGAILEGKLYKIPELGGKDVEDLFTAVALSPATTQAPL	1528
M114	SDPEKPGTPGEGVLSSLDLDRIPTE-----ELPKMESKDLQQLFKDVLGSEREQHLGC	1866
M113	PQPPPPQOLLPMHNQDVFSRMPILMNGLIGSPHLPHNSLPPGSGLGTFPAIAQSPYTDV	1588
M114	RGTPLG-----EGRTSLQRPFLQG-----GLALGSLPSSSPMDSYPLCQSPFLDSR	1913
HMG		
M113	DKSPAFNAIASDPNSSWA-----PTTPSMEGENDTLSNAQRSTLKWEKEEALGEM	1639
M114	ERGGFFSPEPGEPDSPWTGSGGTTPSTPTTPEGEGDGLSYNQRLQWEKDEELGQ	1973
M113	TVAPVLYTNINFNLKEEFPDWTRVKQIAKLRKASSQERAPYVQKARDNRAALRINK'7	1699
M114	TISPVLYANINFNLKQDYPDWSSRCKQIMKLWRKVPAAKDYLOKAKDNRAAHRISK'7	2033
M113	QMSNDSMKRQQOO-DSIDPSSRIDSDLFKDPLKQRESEHEQEWEKFROQMROKSKQQAKIE	1758
M114	OKQAESQISKQAKMDIAR--KTDRP-----ALHLRI---PSQ-PGALG	2071
M113	ATQKLEQVKNEQQQQQQQQQQQQQQQQLASQHLLVAPGSDTPSSG-----AQSPITPQAGNG	1814
M114	-----SPPPAAAPTIF--LGSPPTPAGLSTSADGFLKPPAGT-	2106
M113	NVSPAQTFHKDLFSKHLPGTPASTPSDGVFVKPQPPPPSTPSRIPVQESLSQSNSQPP	1874
M114	--VPGPDSPGELFLKLPPQVPAQVPSQDPFGL-----AP-----	2138
M113	SPQMFSPGSSHSPRSPVDPYAKMVGTPRPPPGHSF-----PRRNS	1916
M114	---AYAPEPRFSAAAPTYPPYPSPTGAPAQPPMLGTTTRPGTGQPGFHTTPPGTPRHQP	2195
M113	VTP----VENCVPLSSVPRPIH--MNETSATR---PSPAR-----DLCASSMTNSDPY	1960
M114	STPDPFLKPRCPSDLNIALVPESPGVAGGKASEPLSPPFGESRKSLEVKEELGASSPG	2255
M113	AKPPDT----P--RP-----MMTDQFSKPFSLPRSPVISEQSTKGPLTTGT-----	2000
M114	YGPVNLCVDSPSAGPHLGGLELKAPDVFKAPLT---PRA---SQVEPQSPGLGLRAQE	2308
M113	-----SDHFTKPSRTDAFQRQRLPDPYAGPSLTPAPLGNGPFKTPLHPPPSQDPYGSVSQ	2056
M114	PPPAQALAPSPPSHPDVFRSGPYPDGYAQPLTPRQPPPESCC-----AP	2355
M113	TSRRLSVDPYERPALTPRVDNFHSQSNDPYSHPLTPHPAMTESFTHASRAFPQPGTI	2116
M114	PPRSLPSDPFSRVPASPO-----SQSSSQSPLTPRPLSAEAFCP-----PV-----	2397

M113	SRSASQDPYSQPPGTPRPLIDSYSQTSGTARSNPDPSQPPGTPRPNTIDPYSQQPPTPR	2176
M114	-----TPRFQSPDPYSRPPSRPQ--SRDPFAPLHKPPR	2428
M113	PSPQTDMFVSSVANQRH--TDPYTHH--LGPPR-PGISVPYSQPPAVPRPRTSEGFTRPS	2231
M114	PQPPEVAFKAGPLAHTPLGAGGFPAALPSGPAGEHLAKVPSGQOPTNFARSPGTGTFVGTP	2488
M113	SARPALMPNQDPFLQAAQNRVPGLPGLIRPPDTCQSQTPrPP-----GPGRIDTFTHAS	2285
M114	SP----M--RFTF-----PQGVGEPSLKPPVPQGPLSPHGINSHFGPGTLGKPQST	2535
M113	SSAVR-DPYD--QPPV--TPRHSESFGTTSQVVHDLVDRPVGSEGNFSTSS-----	2332
M114	NYAVATGNFHPSGSPLGPNSGPTGEGYGLSPLRPASV-LPPPAPDGSLPYLTHGASQRVG	2594
M113	-NLPVSS---QGQQFSSVSQQLPVPVTSGTDQNTVNMSQADTEKLRQRQKLREIIIQO	2388
M114	ITSPVEKREDPGATMSS-SSLA--TPELSSAQDAGISSLSQTELEKQRQRQRLRELLIRO	2651
M113	QQQKKIASRQEKGPOQDTAVVPHPVPLPHWQPESSINQAF-----RPPPPYPGSTRSPVIP--	2443
M114	QIQRNTRQEKEATA--AAAAGAVGPPGNWGAEPSSPAFEQLSRGQTPFTGSQDRSSIVGL	2709
M113	-----PLGPRYAVFPKDQRGYPYPPEV-AGMGMRRPHG-----FRFGFPAGHGPM	2487
M114	PASKLGGPTLPGAFSSDDRILARPLPPATPSSMDMNSRQLVGGSQAFYQRTPYPGSLPLQ	2769
M113	SQDRFHVPQOI-----QGSGIPPHIRRPMMSME	2514
M114	QQQQQQQQQQQQQQQQQQQQQQQQQQQLWQQQQQQQQQQQQAAAAAAATSMRLAMSAR	2829
M113	MPRPSNNP---PLN-NPVGLPQHFPPQGLPVQQHNILGQAFIELRHRAPDGRSRLPFAA	2569
M114	FPSTPGPELGRQALGSPLAGIPTRLPGPAEPVPG-PAGPAQFIELRHNVQKGLGP----	2883
M113	SPSSVIESPSHPRHGNFLPRPDFPGPRHTDPIRQPSQCLSNOQLPVHPNLEQVPPSQ---	2626
M114	-----GVSPFPQGPP--QR-----PRFYPVSEELHRLAPEGLRGL	2917
M113	-EQGHPAHQSSIV-MRPLNHPLSGE-FSEAP-LST-STPAETSPDNLEIAGQSSAGLEEK	2681
M114	AVPGLPSQKPSALPAPELNNSLHQTPHAKGPALASGLELVSRRPPSNTELSRPPLALEAGK	2977
M113	LDSDDPSVKELDVKDLEGVEVKDLDDEDLENLNLDTEDKGK-DDLDLTLNLETNDPNLDD	2740
M114	LPCEDPELD-----DDFDAHKALEDEELAHLGLGVDVAKGDELGTLENLETNDPHLDD	3032
M113	LLRSGEFDIIAYTDPELDLGDKKSMFNEELDLNVPIDDKLDNQCA--SVEPKTRDQGDKT	2798
M114	LLNGDEFDLLAYTDP ELDTGDKKDIFNEHLRLVESANEKAEREALLRGVEPVSLGPEERP	3092
M113	MVLEDKDLPLQKSSVSSEIK-----TEALSPYSK-EEIQS-----E	2833
M114	PPAPDNSEP-RLTSVLPEVKPKVEEGGRHPSPCQFTINTPKVEPAPPATSLISLGLKPGQT	3151
M113	IKNHDDSRGDADTACSQAASAQTNHSDRGKTALLT-TDQDMLEKRCNCNQENAGPVVSAIQ-	2891
M114	VMGTRDTRGGVTGSFPSSGHT---AEKGPFGATGGTPAHLNP---SSLSPGAASSLLE	3205
M113	-----GSTPLPAR-----DVMNSCDITGSTPVL-SSLLSNEKCDDSDIRPSGSSPP	2936
M114	KFELESGALTLPSPGHAAAGDELDKMESSLVASELPILLIEDLLEHEKKELQK-KQQLSAQT	3264
M113	SLPISPSTH-----GSSLPPTLIVPPSPLLDNTVNSNTVVPRINHAFSQGVPVNPFGI	2990
M114	VLPAQQQQQQQQQQQQQQQHTLLPTPGPAQALPLP--HEPGPPQQLALGIGSTRQPLG	3322
M113	QGQSSVN---HNLGTGKPTNQTVPLTNQSTSMSGPQQLMIPQTLAQONRERPLLEEQPL	3047
M114	QSMVPIQPPAHALOQR-LAPS VAMVSQNQGHMLSGQQAGQ--TGLVPQOSSQ-----PV	3372
M113	LLQDLDQERO-EQQQORQMQAMIRQRSEPFNIDFD-----AITDPIMKAKMVALKG	3101
M114	LAQKPMPSAMPASCMKPKQQLAMQQQQLANSFFPDTDLDFAAEDIIDPIAKAKMVALKG	3432
M113	NKVMAQNSLGMPPMVMS-RFPFMGPSVAGTQNNDGQTLP-QAVAQDGSIITHQISRPNPP	3159
M114	KKVMAQGSIGVAPGMNRQOVSLILAQLSGSGSDLQNHVAPGSGQERNAGDPAQPRPNPP	3492
M113	NFGPGFVNDSQRKQYEEWLQETOQOLLQMQQKYLEEQIGAHRKSKKALSAKORTAKKAGRE	3219
M114	TFAQGVINREADQRQYEEWLHTQOLLQMQQLKVLEEQIGVHRKSRKALCAKORTAKKAGRE	3552
M113	FPEEDAEQLKHVTQOSMVQKOLEQIRKQQKEHAELIEDYRIKQQQQQQCALA-----	3273
M114	FPEADEAKLKLVTQOSKIQKQLDQVRKQQKEHTNLMAEYRNKQQQQQQQQQQQQQHS	3612
M113	-----PPILMPGVQPQPPPLVP-GATS--LT-----MSQPNFPMVP	3305
M114	AVLAVSPSONPRVLTKLPGQOLPAHGLQPPQAPPGGQAGGLRLPPGGMVLPQSGGPFLN	3672

M113	---QOLHQOHTAVISGHTSPARM--P-----	3327
M114	<u>TTLAQQQQQQHSGVAGSLTGPPGSFFPGNLALRSLGPDSRLLQEROLQLOQOORMQLAQKL</u>	3732
M113	-----SLPGWQSNSASAHLPLNPPRIOPPIAQLSLK-----	3358
M114	<u>QQQQQQQQQQQQQQHLLGQVAIQQQQGPGVQNQ</u> ---ALGPKPQGLLPPSNHQGLLVQQL	3788
M113	---TC-----TPAPGTV-----SSANPQNGPPPRVEFDDNNPFSE-----	3390
M114	<u>SPQQSQGSQGLLGPQAQTVLQQQQQQQHSGALGPQGPHRQVLMTQSRVLSSPQLAQQGH</u>	3848
M113	-----SFQERERKERLREQQER-----QR-----	3409
M114	<u>SLMGHRLLTAAQQQQQQQQQQQQQQQQQQQQQGSMTGLSQLQOGMMSHGGQPKMSAQ</u>	3908
M113	-----VQLM-----QEVDRQRA-----LQORMEMEQHCL	3433
M114	<u>ALGSLQQQQQQQLQQQQMLQQQQQLQQQQQQQLQQQQQQQLQQQQQQQLQQQQQQQL</u>	3968
M113	M-----GAELAN---RTPVSQMPFYGSDRPCDFLQPPRP	3464
M114	<u>QHQQQQQQQLQQQQQLQQQQQQQLHLLQQQLHQQQLQLQQQQMGLLNQNR</u> --TLLSPQQQ	4026
M113	LQOSPQHQ-----QQIGPVLQQQNVOQQGSVNSPPNQTFMQTN-EQ--RQV	3506
M114	<u>QQQQQQQQQQQQQQQQQQQQQVTLGPGLPVKPLQHFSSSGALGPTLLTGKEQNNAET</u>	4086
M113	GPPSFVPDSPS-----ASGGSPNFHSVKPGHGNL--PGSS-----FQQSPL	3545
M114	ALPSEVTEGPSTHQGGPPAVGTAPEPMSEVEPGEVKPISGDSQLLLQVQSAQSATSVQL	4146
M113	RPPFTPILPGTS-PVA-----NSNVPCGQDPAVT	3573
M114	<u>QPPLR-LPGQPQPVNLLHTAGGGSHGQQLGSGSSSEPAVPHLLAQPSVSLGEQPGP-</u>	4203
M113	QGQNYSGSSQSL-----IQLYS---DII-	3593
M114	MAQNLLGSQQPLGLDRPIQNNNTGSQPPKGSPAPQSGQGPPGAGVMPTVGQLRAQLQGVLA	4263
M113	-----PEEKKKKR---TRK-----KKKDDDAESGKAPSTP--HSDCAAPLTPG	3632
M114	<u>KNPQLRHLSQQQQQLQALLMQRQLQOSQAVRQTPFPQEPGTQPSPLQGLLGCQ-PQPGG</u>	4322
M113	LSETTSTPA-----VSSPSELPO-----QRQEPVEPVPVPT	3664
M114	FSVSQTGPLQELGAGSRPQGPRLPVPQGALSTGPVLGPAHPTPPPSSPQEPKRPSQLPS	4382
M113	PNVS--A---GQPCIE-SENKLPNSEFIKETSNQQTHVNAEADKPSVETPNKTEEIKLEK	3718
M114	PSAQLTPTHPGTPKPGPALELPPGRVSPAALQADTFFGKGLGPWDPSDNLTEAQKPEQ	4442
M113	AETQPSQEDTKVEEKTGNKIKD-----IVAGP-VSSIQCPSPHPVGTPT---	3760
M114	SSLVPG---HLDQVNGQVVHEPSQLSIKQEPREECALGAQTVKREANGE PAGAPGTSN	4498
M113	-----TKGDTGNELLKHLLKNKKASSLLTQKPEGTLOSSDESSTKDGLIEKQSPAEGL <u>Q</u>	3814
M114	HLLLAGSRSEAGHLLLQKLLRAKNVQLGAGRGP EGL-----	4534
M113	TLGAQMCGFGGGNSQL-PKTDGASENKKQ-----RSKRTQRTGEKAAPRSK---	3860
M114	--RAEINGHIDSKLQGLEQKLOQTSSNKEDAATRKPLPAKPKRVQKTSDRLPSRKKLK	4592
M113	-----KRKKDEEEKQAMY	3873
M114	EDGVRANEALLKQLQELSOLPLTEPTITANFSLFAPFGSGCLVSGQSQLRGAFGSGALH	4652
M113	SSDSFTHLKQONNLSNPPTPPASLPPPTPPMACQKMANGFATTEELAGKAGVILVSHEVA	3933
M114	TGPDYYSQQLTKNNLNSNPPTPPSSLPPPTPPSVQOKMVNGVTPSDELGERPKDTASAQ--	4710
M113	RALGPKFQLPFRPQDDLLARATAQGPKTVDPVSPASLPTPPHNNHEEL <u>RIQDHYGRDTPD</u>	3993
M114	-----DSEGALRDAAEVKSLSLDLLAALPTPPHNQTEDVRMESD-EDSDSPD	4754
M113	SFVPSSSPESVVGVEVNKYPDLISLVKEEPPEPVPSPIIPILPSISGK <u>NSESRR</u> --NDIK-	4050
M114	SIVPASSPESILGEAIPRFQQLGSGRWEQDNRALSPVPIIIPRTGIP <u>VFPDTKPYGVLDL</u>	4814
M113	TEPGTLFFTSPFGSSPNGPRSGLISVAITLHPTAA <u>ENISSVVAFSDLLHVRI</u> PNSYEVS	4110
M114	EVPGKLP-AW---EKGKGSEVSVMLTVSAAA <u>AKNLNGVMVAVAELL</u> SMKI PNSYEVL	4869
M113	NAPDVPPMGLVSSHRVNPSLEYRQHLLLRLGPPPGSANPPRLATSYRLKQPNVBFPPPTS <u>NG</u>	4170
M114	FPDGPARAGLEPKK-----GE-----AEG	4888
M113	LSGYKDSSHGPAEGASLRPQWCCHCKVVIILGSGVRSKCDLTUVNKGSRENTKRMEKDIV	4230
M114	<u>PGGKEKG--LSGKGPDGTGPWLKQFDAVLPGYTLKSQLDILSLLKQESPAP</u> E-----	4939
M113	FCSNNCFILYSSAAQAKNSDNKESLPSLPQSPMKEPSKAFHQYSNNISTLDVHCLPQFQ-	4289

M114	-----	-SIQHSYTINYNSNLDVRLSAPP-	4962
M113	-----	-EKVSPPASPPISFPPAFEAAKVESKPD E LVTVKLK	4325
M114	EEPSPPPSPLAPSPASP PAEPMVELQ A ERPAEPPPIPSPPLA-----	-SSPE-----SARPK	5013
M113	PRLRTPVVGLED C RPLNKWWRGMWKWKS I HIVIPKGTFKPPCEDEIIDEFLKKLGTC L KP	4385	
M114	PRARPPEESED S RP P RLKKWKGVRWKRLRLLT I QKGSGHQEDEREVAEFMEQFGTALRP ePHD2	5073	
M113	DPVPKD C KCCFCHEEGDGLTDGPARLLNLNDL D LWVHLNCALWSTEVYETQAGALINVE	4445	
M114	SKVPRDN R CCFCHEEGDGT D GATDGPARLLNLNDL D LWVHLNCALWSTEVYETQGGALMNVE	5133	
M113	ALRRGLQM K CVFCHKTGATSGCHRFRCTNIYHFTCATKAQCMMFKDKTMLCP H PKPGI H	4505	
M114	ALHRG L LT K CSLC Q RTGATSSCNRMRCPNVHFACAIRAKCMFFKD K TMLCP V HK I KGPC FYRN	5193	
M113	EQQLSYFAVFRRVYVQRDEV Q IASI V QRGERDHTFRVGS L IFHTIG Q LLPQQM A FHS P	4565	
M114	EQELSSFAVFRRVYIERDEV K QIAS T I V QRGERLHMFRVGG L VFAIG Q LLP H OMAD F HS A FYRC	5253	
M113	KALFPVGYEASRLYWSTRYANRR C RYL C SI E KDGR P VF V IRIVEQ G HEDLV L SDSS P K	4625	
M114	TALYPVGYEATRIYWSLRNNR C CYRC C SI E NNGR P EF V IKVIEQ G LEDLV F TDAS P K	5313	
M113	VWDKILEPVACVRKKSEMLQLFPAYLK G EDLFGLTVSAVARIAESLPG V EACENYTF R G	4685	
M114	VWNRII E PVAAMR K ADM L RF P EYL K GEEL G LT V HAVLRIAESLPG V ESCONYL F R G	5373	
M113	RNP I M E L P LA N PTGC A RS E PKMSAHVKRF V L R PHTLNST T STS K S F Q S TV T G E LNAPYS K	4745	
M114	RH P I M E L P LI N PTGC A RS E PK I LTHY K -----R P HTLNST T SM S KAY Q ST F TGETNT P YS K SET	5429	
M113	QFVHSKSSQYRRMKTEW K SNVYLARSRI Q GLGLYAARDIEKHTMVIEYIG T IIRNEVAN R	4805	
M114	QFVHSKSSQYRR L TEW K NNVYLARSRI Q GLGLYA A KDLEKHTMVIEYIG T IIRNEVAN R	5489	
M113	KEKLYE S QNRGVYMFRMDNDHV I DA T LTGG P A R Y INHSCAPNCVAEVVT F ERGH K I I S ;	4865	
M114	REKIYE E Q N RGIY M FRIN N EH V DA T LTGG P A R Y INHSCAPNCVAEV V TFD K E D K I I I S ;	5549	
M113	N RRI Q GEEL C DY K FD F ED D D Q H K I P CH C G A V N CR K WM N 4904		
M114	S RR I P KG E EL T D Y Q FD F ED D D Q H K I P CH C G A V N CR K WM N 5588		

MLL3	start	end	domain	MLL4	start	end	domain	Database
246	329	ePHD1		134	217	ePHD1		CDD
342	388	PHD		228	274	PHD		Smart
390	434	PHD		276	320	PHD		Smart
465	517	PHD						Smart
946	1002	PHD		1334	1384	PHD		Smart
1003	1049	PHD		1385	1431	PHD		Smart
1080	1131	PHD		1462	1513	PHD		Smart
1639	1703	HMG		1969	2037	HMG		Smart
4395	4499	ePHD2		5083	5189	ePHD2		CDD
4554	4597	FYRN		5242	5285	FYRN		Smart
4603	4690	FYRC		5291	5378	FYRC		Smart
4764	4886	SET		5448	5570	SET		Smart

Alignment of murine MLL3 and MLL4. Exons alternate in black and blue. The conserved PHD fingers, HMG boxes, FYR-N/FYR-C and SET domain regions are boxed. Letters in red color: residue overlaps splice site. Note in MLL4, regions highly dense in proline are underlined in black (exon 10-11) and several stretches with glutamine repeats are underlined in purple (exon 39). The insertion sites for the FRT-SA-GT0-T2A-lacZneo-CoTC-FRT-loxP and loxP-rox-PGK-Blasticidin-pA-rox cassettes are marked by purple triangles above the MLL3 sequence (as shown in Fig 1B). The insertion sites for the FRT-GT1-T2A-lacZneo-pA-CoTC-FRT-loxP and loxP-rox-PGK-Hygromycin-pA-rox cassettes are marked by red triangles above the MLL3 sequence (as shown in Fig S1C). The insertion sites for the loxP-zeo-loxP and loxP-FRT-SA-IRES-lacZneo-pA-FRT cassettes are marked by green triangles beneath the MLL4 sequence (as shown in Fig. 1C). Three KLH-conjugated synthetic peptides from the central part of the MLL4 protein (QRPRFYVPSEEELHRLAP, NGDEFDLLAYT, KQQLSAQTQRLAPS) were used to raise a rabbit polyclonal antibody and are marked in orange. After raising the antibody, an updated gene annotation revealed that the third peptide was not continuous but interrupted by 72 a.a.s.

Table S5. Expression profiling on whole lung at E18.5

[Click here to Download Table S5](#)