



**HAL**  
open science

# HOX epimutations driven by maternal SMCHD1/LRIF1 haploinsufficiency trigger homeotic transformations in genetically wildtype offspring

Shifeng Xue, Thanh Thao Nguyen Ly, Raunak S. Vijayakar, Jingyi Chen, Joel Ng, Ajay S. Mathuru, Frédérique Magdinier, Bruno Reversade

► **To cite this version:**

Shifeng Xue, Thanh Thao Nguyen Ly, Raunak S. Vijayakar, Jingyi Chen, Joel Ng, et al.. HOX epimutations driven by maternal SMCHD1/LRIF1 haploinsufficiency trigger homeotic transformations in genetically wildtype offspring. *Nature Communications*, 2022, 13 (1), 10.1038/s41467-022-31185-8 . hal-03780350

**HAL Id: hal-03780350**





**<https://amu.hal.science/hal-03780350>**

Submitted on 27 Jan 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# HOX epimutations driven by maternal SMCHD1/LRIF1 haploinsufficiency trigger homeotic transformations in genetically wildtype offspring

Shifeng Xue<sup>1,2</sup>, Thanh Thao Nguyen Ly<sup>1,2</sup>, Raunak S. Vijayakar<sup>3</sup>, Jingyi Chen<sup>1</sup>, Joel Ng<sup>1</sup>, Ajay S. Mathuru<sup>2,3,4</sup>, Frederique Magdinier<sup>5</sup> & Bruno Reversade<sup>2,6,7,8</sup>

The body plan of animals is laid out by an evolutionary-conserved HOX code which is colinearly transcribed after zygotic genome activation (ZGA). Here we report that SMCHD1, a chromatin-modifying enzyme needed for X-inactivation in mammals, is maternally required for timely HOX expression. Using zebrafish and mouse *Smchd1* knockout animals, we demonstrate that *Smchd1* haplo-insufficiency brings about precocious and ectopic HOX transcription during oogenesis and embryogenesis. Unexpectedly, wild-type offspring born to heterozygous knockout zebrafish *smchd1* mothers exhibited patent vertebrate patterning defects. The loss of maternal *Smchd1* was accompanied by HOX epi-mutations driven by aberrant DNA methylation. We further show that this regulation is mediated by *Lrif1*, a direct interacting partner of *Smchd1*, whose knockout in zebrafish phenocopies that of *Smchd1*. Rather than being a short-lived maternal effect, HOX mis-regulation is stably inherited through cell divisions and persists in cultured fibroblasts derived from FSHD2 patients haploinsufficient for SMCHD1. We conclude that maternal SMCHD1/LRIF1 sets up an epigenetic state in the HOX loci that can only be reset in the germline. Such an unusual inter-generational inheritance, whereby a phenotype can be one generation removed from its genotype, casts a new light on how unresolved Mendelian diseases may be interpreted.

<sup>1</sup>Department of Biological Sciences, National University of Singapore, Singapore, Singapore. <sup>2</sup>Institute of Molecular and Cell Biology, A\*STAR, Singapore, Singapore. <sup>3</sup>Yale-NUS College, Singapore, Singapore. <sup>4</sup>Department of Physiology, School of Medicine, National University of Singapore, Singapore, Singapore. <sup>5</sup>Aix-Marseille University, INSERM, Marseille Medical Genetics, Marseille, France. <sup>6</sup>Genome Institute of Singapore, A\*STAR, Singapore, Singapore. <sup>7</sup>Department of Paediatrics, School of Medicine, National University of Singapore, Singapore, Singapore. <sup>8</sup>Department of Medical Genetics, KOÇ University, Istanbul, Turkey. ✉email: [dbsxues@nus.edu.sg](mailto:dbsxues@nus.edu.sg); [bruno@reversade.com](mailto:bruno@reversade.com)

Early embryonic development is controlled by the loading of maternal factors into the oocyte during oogenesis. These RNA and proteins, also known as maternal effect genes, support embryogenesis until the zygotic genome is activated for transcription<sup>1</sup>. Maternal effect genes are known for a range of functions including cell cycle regulation, genomic imprinting, pluripotency and cell fate specification<sup>2</sup>. However their effects are usually restricted to embryogenesis and rarely long-lasting.

Structural maintenance of chromosomes flexible hinge domain containing 1 (SMCHD1, MIM614982) is an evolutionarily conserved chromatin-binding protein involved in epigenetic silencing<sup>3</sup>. SMCHD1 was first discovered to be critical for X-inactivation in mammals since homozygous mutant female mice die by mid-gestation while males remain healthy and fertile<sup>4</sup>. Subsequently SMCHD1 was also documented to regulate the expression of mono-allelically expressed and imprinted genes<sup>5</sup>. Germline mutations in *SMCHD1* have been implicated in two very distinct Mendelian human diseases. Haploinsufficiency of *SMCHD1* is responsible for facioscapulohumeral dystrophy type 2 (FSHD2, MIM158901), where the monoallelic depletion of SMCHD1 causes hypomethylation of the D4Z4 repeat region and subsequent expression of *DUX4* which is toxic to muscle lineages<sup>6</sup>. More recently, de novo missense mutations in the GHKL ATPase domain of SMCHD1 are shown to cause a striking craniofacial disorder, Bosma Arhinia Microphthalmia Syndrome (BAMS, MIM603457) which is characterized by the congenital absence of a nose and its associated structures<sup>7,8</sup>.

SMCHD1 expression precedes fertilization and is found in developing oocytes in mice and humans<sup>9–12</sup>. Because *Smchd1* knockout female mice are embryonic lethal, the role of its maternal contribution before the onset of zygotic transcription has been difficult to address. To circumvent this limitation, we turned to zebrafish where the absence of imprinting and X-inactivation permits interrogation of the role of maternal *smchd1*.

Here we show that loss of *smchd1* in zebrafish causes derepression of a large number of *hox* genes resulting in aberrant skeletal patterning. Surprisingly this is caused largely by the maternal pool of *Smchd1* as even wild-type offspring born to heterozygous knockout zebrafish *smchd1* mothers exhibited vertebrate patterning defects. The loss of maternal *Smchd1* was accompanied by aberrant DNA methylation in the *hox* loci. This is conserved in mammals as we further show that *HOX* genes are also misregulated in *Smchd1* haploinsufficient mouse and human fibroblasts derived from FSHD2 patients. This regulation is mediated by *Lrif1*, a direct interacting partner of *Smchd1*, whose knockout in zebrafish phenocopies that of *Smchd1*. We conclude that maternal *Smchd1/Lrif1* set up an epigenetic state in the *hox* loci that can only be reset in the germline. Such an unusual intergenerational inheritance, whereby a phenotype can be one generation away from its genotype, casts a new light on how unresolved Mendelian diseases may be interpreted.

## Results

***MZ smchd1* fish embryos show derepression of *hox* genes.** Using CRISPR/Cas9 technology, we generated three distinct loss-of-function (LoF) *Smchd1* alleles causing out-of-frame mutations in the GHKL ATPase domain (Fig. 1a). Endogenous *smchd1* mRNA is heavily maternally contributed in zebrafish with ubiquitous distribution until gastrulation, after which it becomes restricted to the head and lateral plate mesoderm (Fig. 1b). This large maternal contribution makes zebrafish an ideal system for interrogating the role of maternal *Smchd1*. By RT-QPCR and whole-mount in situ hybridization, all three lines of maternal zygotic (MZ) null embryos showed barely detectable endogenous

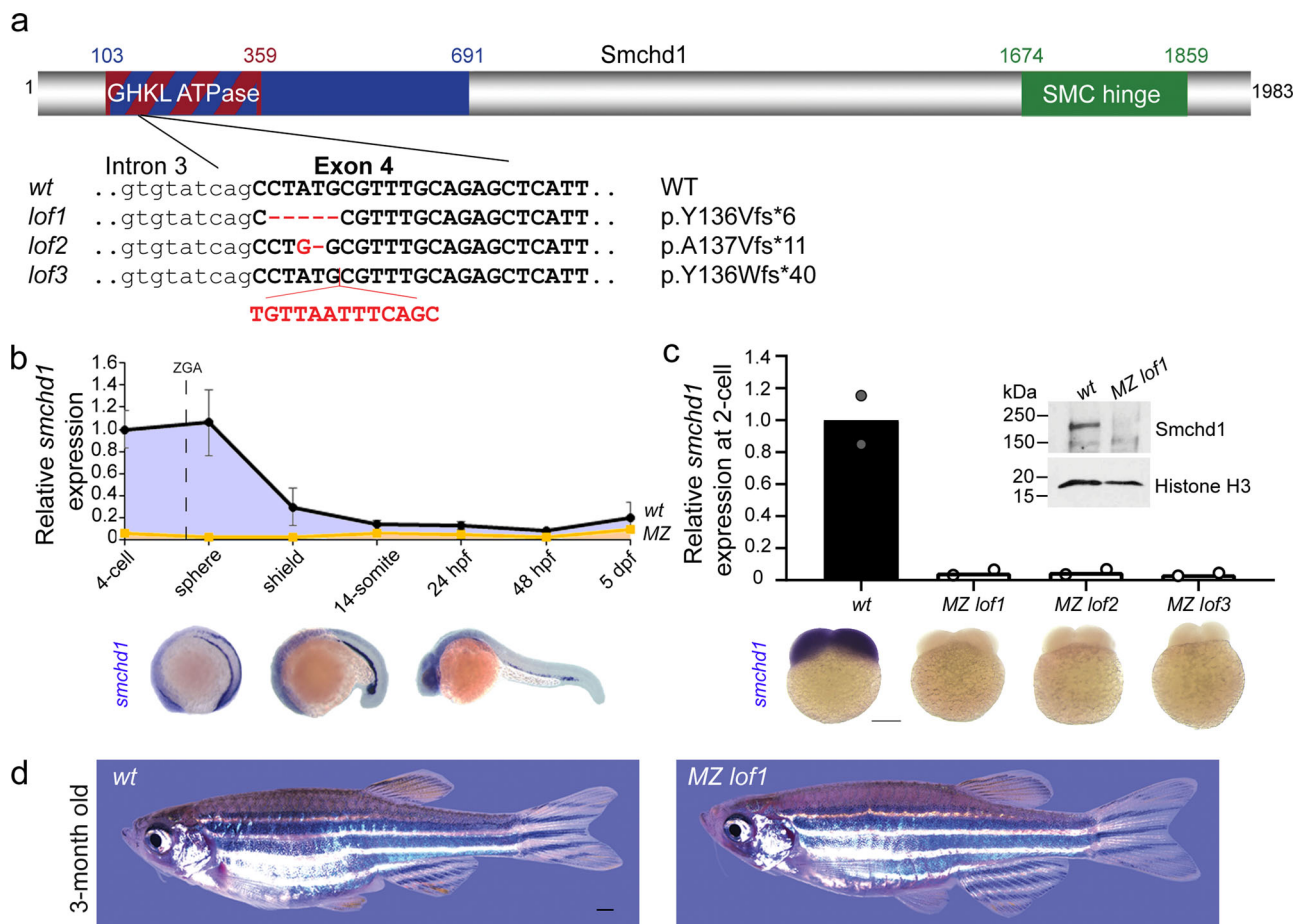
*smchd1* mRNA at the 2-cell stage and protein at the sphere stage (Fig. 1c) indicating that the maternal pool of *smchd1* was degraded through the nonsense-mediated decay pathway. These three independent mutant lines represent the first complete protein-null *smchd1* knockouts in zebrafish.

A previous study suggested that zebrafish *smchd1* morphants or crispants display craniofacial anomalies mimicking BAMS<sup>8</sup>. In light of recent debates on whether F0-depleted animals can be used interchangeably for genetic mutants, we examined whether germline *smchd1* knockout zebrafish recapitulate BAMS. No craniofacial abnormalities were seen in: F1 zygotic heterozygous fish (which best recapitulate the *SMCHD1* genotypes of BAMS patients), F2 zygotic homozygous fish or even F3 MZ homozygous *smchd1*-null fish. At 3 months of age, MZ *smchd1* knockout fish are visually indistinguishable from wildtype (*wt*) counterparts (Fig. 1d), did not display any arhinia-associated phenotypes (Supplementary Fig. 1a–c) nor did they exhibit signs of impaired olfaction to ethologically salient attractive and aversive odorants (Supplementary Fig. 1d and Movie 1).

Having genetically and functionally demonstrated that loss-of-function of *Smchd1* in zebrafish does not model BAMS, we sought to understand *Smchd1*'s exact role during embryogenesis. We first performed an unbiased transcriptomic analysis of *wt* and *MZ smchd1<sup>lof1/lof1</sup>* embryos at the 4- to 8-cell stage, when all RNAs in the embryo are maternally supplied, and at sphere stage after zygotic transcription has begun. Gene Ontology analysis of differentially expressed genes at the 4- to 8-cell stage revealed skeletal system development as one of the top categories (Supplementary Fig. 2a–d). Further clustering of the differentially expressed genes at the 4- to 8-cell stage disclosed 2 clusters of genes which were upregulated in the mutant embryos (Fig. 2a). Among these, cluster 5 showed a significant enrichment of genes involved in skeletal system development (Fig. 2b). On closer examination, a conspicuous upregulation of a large number of *hox* genes was recorded at both stages (Fig. 2c, d, Supplementary Fig. 2e, f, 3a, Data 2). *Hox* genes serve as master regulators of anterior posterior (AP) patterning in the embryo and their expression boundaries are tightly controlled to specify the identity of body parts<sup>13</sup>. At 11-somites when most *hox* genes are highly expressed, qRT-PCR analysis found that a series of *hox* transcripts were persistently upregulated (Fig. 2e). For instance the expression of *hoxb2a* is normally restricted to rhombomere 3 and 4 at 10-somites, but appeared to be extended into rhombomere 5 in *MZ smchd1<sup>lof1/lof1</sup>* embryos (Fig. 2f, Supplementary Fig. 3b, c). There was also an anterior expansion of the *hoxc10a* expression boundary (Fig. 2f). These data indicate that loss of SMCHD1 is sufficient to de-repress a vast number of *hox* genes in early development.

***MZ smchd1* fish display vertebral patterning defects.** Zebrafish has undergone a partial genome duplication and carries seven *hox* gene clusters, instead of four in mammals and one in *Drosophila*<sup>14</sup>. Surprisingly adult *MZ smchd1<sup>lof1/lof1</sup>* fish showed patent vertebral patterning defects consistent with homeotic transformations. The observed phenotype (Fig. 2g, h) consisted of a reduction in the number of ribs, caudal vertebrae and occasionally of supraneural vertebrae (Supp. Fig. 4a–d). The future reduction of vertebrae could already be documented in 48 hpf embryos in reduced somite numbers as marked by *xirp2a* (Supplementary Fig. 4e, f). Our results show that the de-repression of multiple *hox* genes caused by the absence of *smchd1* suffices to trigger vertebral patterning defects in zebrafish.

Changes in vertebral numbers have been previously documented with manipulations of the *mir196* family of microRNAs in mouse and zebrafish<sup>15,16</sup>. *mir196* genes are embedded within the



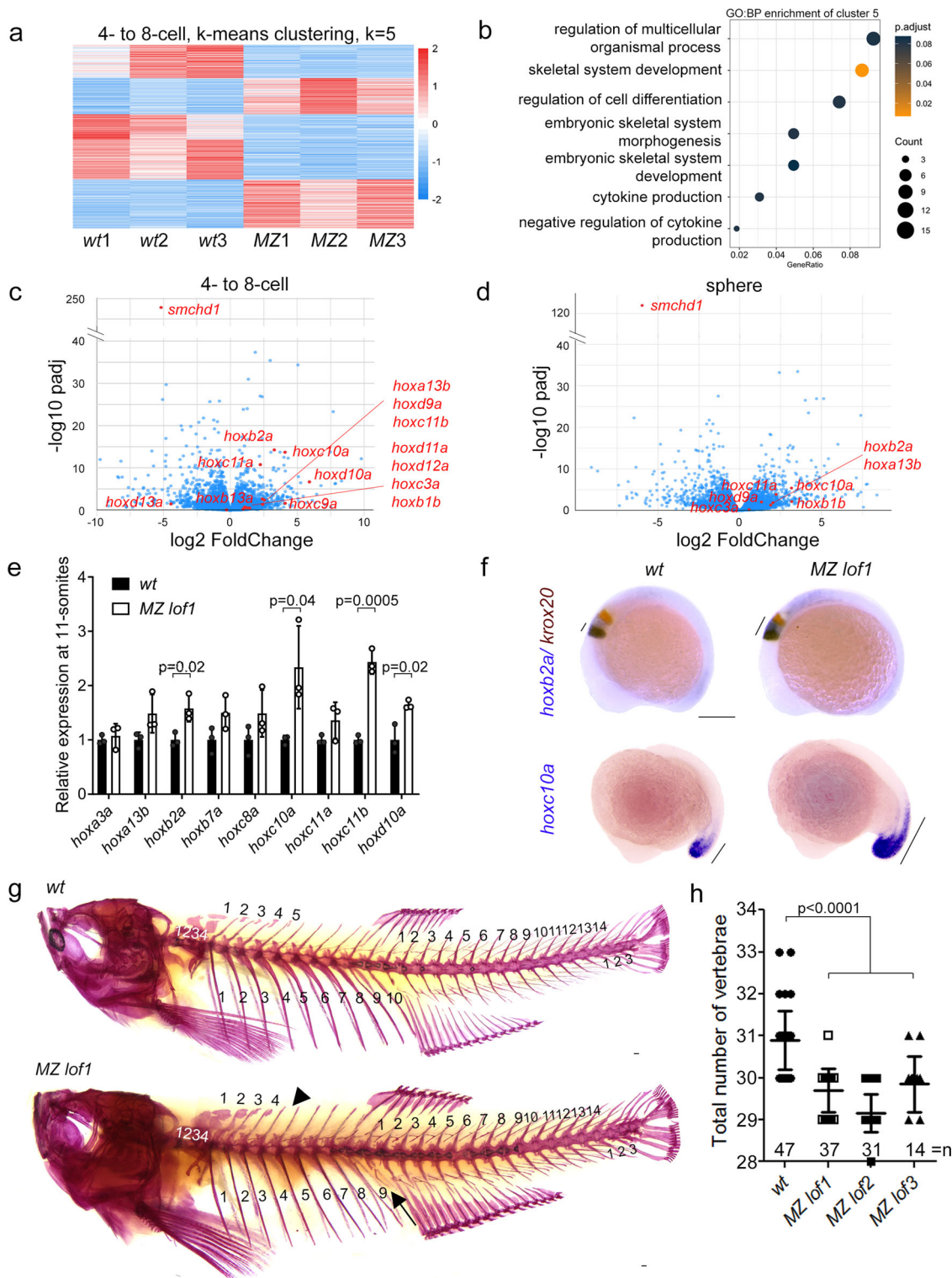
**Fig. 1** Generation of MZ *smchd1* null zebrafish. **a** Three loss-of-function (LoF) alleles in zebrafish were generated by CRISPR/Cas9 injections targeting *smchd1*. **b** Endogenous spatial-temporal expression pattern of *smchd1* in zebrafish embryos as determined by qPCR and WISH. *n* = 3 biological samples (qPCR), 3 independent experiments with at least 20 embryos each (WISH). Data are presented as mean values  $\pm$  SD. **c** Maternal zygotic (MZ) *smchd1*<sup>lof</sup> embryos lack *smchd1* mRNA (qPCR and WISH) and protein (Western blot), suggestive of total nonsense-mediated decay. Scale bar = 200  $\mu$ m. *n* = 2 biological samples (qPCR, Western blot), 3 independent experiments with at least 20 embryos each (WISH). **d** MZ *smchd1*<sup>lof</sup> zebrafish are viable and fertile, and do not model arhinia at 3 months. Scale bar = 1 mm.

*Hox* loci and act to post-transcriptionally repress the expression of neighbouring *Hox* genes. Expression of mature *mir196* paralogs was not significantly changed in *smchd1*<sup>lof1</sup> oocytes but was increased approximately 1.5 fold at 11-somites, suggesting that it may also participate in the general dysregulation of the *hox* transcripts (Supplementary Fig. 3d, e). As *mir196*-induced vertebral reductions were previously achieved with an estimated 10–20 fold *mir196* over-expression, we surmise that the modest *mir196* upregulation may not be a main driver of the phenotype in our *smchd1* knockout fish.

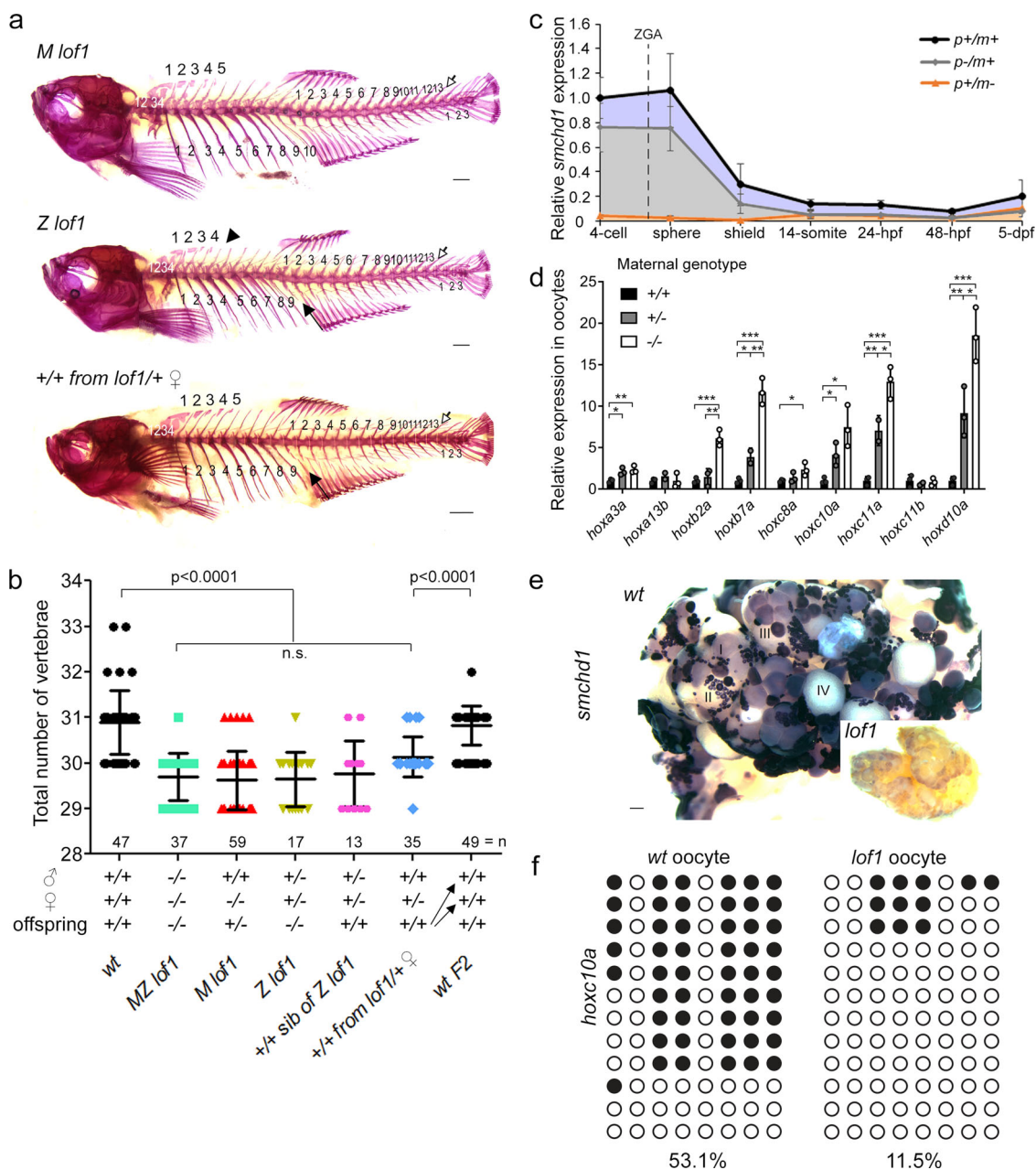
**Skeletal phenotype is caused by maternal *Smchd1* haploinsufficiency.** To assess which of the maternal or zygotic pool of *Smchd1* is responsible for proper AP axial patterning, the skeletons of maternal null *smchd1*<sup>lof1</sup> and zygotic null *smchd1*<sup>lof1</sup> were compared. Both displayed equivalent reductions in the number of vertebrae as MZ *smchd1*<sup>lof1/lof1</sup> (Fig. 3a, b, Supp. Fig. 4a–d) suggesting that both sources may be required. By analysing the levels of *smchd1* in paternally- or maternally-deleted embryos, we found that most, if not all, endogenous *smchd1* up to gastrulation is deposited by the mother (Fig. 3c). This led us to hypothesize that the observed vertebral defects in zygotic null *smchd1*<sup>lof1/lof1</sup> may be the result of insufficient maternal *smchd1* deposition in the egg rather than from a

subsequent loss of zygotic *smchd1*. To answer this, we turned our attention to genetically wildtype *smchd1*<sup>+/+</sup> siblings from the same brood as zygotic null *smchd1*<sup>lof1/lof1</sup> fish, born from the incross of heterozygous *smchd1*<sup>lof1/+</sup> parents. To our surprise, these fish displayed identical vertebral defects to that of zygotic null fish despite being genetically wildtype. We further generated a special clutch of *smchd1*<sup>+/+</sup> offspring resulting from the mating of heterozygous *smchd1*<sup>lof1/+</sup> mutant mothers to outbred wildtype *smchd1*<sup>+/+</sup> males. These fish also displayed reduced vertebral numbers (Fig. 3a, b, Supplementary Fig. 4a–d). These results demonstrate that *Smchd1* is a gene whose fractional reduction during oogenesis can alter the body plan of the future embryo irrespective of its zygotic genotype. This was surprising as the Hox code was not known to be regulated by the maternal genome. Notably, an in-cross of these phenotypically abnormal *smchd1*<sup>+/+</sup> fish produced offspring that were completely normal with a canonical number of vertebrae (Fig. 3a, b, Supplementary Fig. 4a–d). This rescue indicates that no trans-generational inheritance takes place and that once *smchd1* is reset to wildtype levels in the maternal germline, a proper AP patterning ensues in the following generation.

Zebrafish embryos are transcriptionally silent until the 1000-cell stage<sup>1</sup>. To determine whether the observed upregulation of *hox* expression in *smchd1* mutant embryos precedes fertilization or is the result of early zygotic genome activation (ZGA),



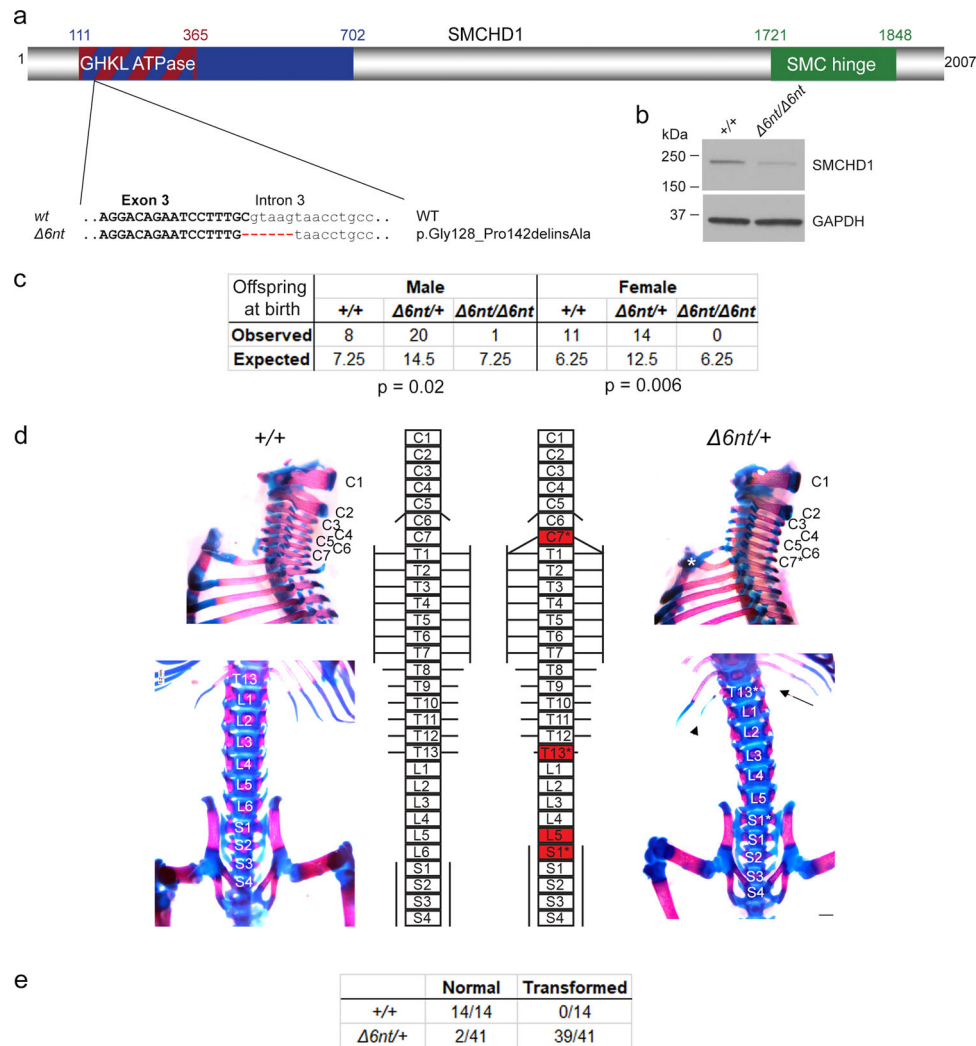
**Fig. 2** MZ *smchd1* fish embryos show derepression of *hox* genes and vertebral patterning defects. **a** K-means clustering of differentially expressed genes at 4–8-cell stage. **b** Gene ontology: biological process (GO:BP) enrichment of cluster 5 from K-means clustering at 4–8-cell stage. RNA-seq of wt and MZ *lof1* embryos show changes in *hox* gene expression at both 4–8-cell stage (**c**) and sphere stage (**d**). **e** *hox* genes show persistent upregulation at 11-somites.  $n = 3$  biological samples,  $p$  values were calculated by 2-tailed unpaired Student's  $t$  test. **f** *hoxb2a* shows expanded expression boundary at 10-somites (top) and *hoxc10a* shows anteriorly expanded expression at 17-somites (bottom). *krox20* marks rhombomeres 3 and 5. **g–h** Adult fish show defects in axial skeletal patterning, including loss of a rib (arrow) and loss of a supraneural vertebra (arrowhead). Scale bar = 200  $\mu\text{m}$ .  $P$  values were calculated by the Kruskal–Wallis test followed by Dunn's Multiple Comparison Test. Bars without  $p$  values are not significant. All data are presented as mean values  $\pm$  SD.



**Fig. 3 Skeletal phenotype is caused by maternal *Smchd1* haploinsufficiency. a, b** Skeletal preparations of fish of different genotypes. Phenotypes include loss of a rib (filled arrow), loss of a supraneural vertebra (arrowhead) and loss of a caudal vertebra (open arrow) Scale bar = 1 mm. *P* values were calculated by the Kruskal-Wallis test followed by Dunn’s Multiple Comparison Test. **c** qPCR shows that there is little to no zygotic *smchd1* expression. **d** qPCR shows that *hox* derepression is already observed in fertilized oocytes and that oocytes from  $\pm$  females show intermediate level of *hox* mis-expression. *n* = 3 biological samples, *p* values were calculated by 2-tailed unpaired Student’s *t* test. \**p* < 0.05, \*\**p* < 0.01, \*\*\**p* < 0.001. All data are presented as mean values  $\pm$  SD. **e** In situ hybridization of *smchd1* shows that *smchd1* is expressed in developing oocytes. Roman numerals denote stages of oocyte development. Scale bar = 200  $\mu$ m. **f** Bisulfite sequencing in oocytes shows hypomethylation of *hoxc10a* locus when *smchd1* is knocked out.

unfertilized oocytes of different *smchd1* genotypes were isolated. The levels of various *hox* genes were inversely correlated to the number of wildtype *smchd1* alleles available during oogenesis: the highest precocious *hox* expression was seen in *smchd1<sup>lof1</sup>* oocytes produced by homozygous *smchd1<sup>lof1/lof1</sup>* mothers while oocytes produced by heterozygous *smchd1<sup>+/lof1</sup>* mothers showed intermediate levels relative to wildtype oocytes (Fig. 3d). These results lend support to the observed maternal *smchd1* haploinsufficiency and correlate with the abundant *smchd1* expression in developing oocytes from the earliest stages of oogenesis (Fig. 3e). Our results point to a very early defect during gametogenesis in the maternal

germline, where *smchd1* depletion in transcriptionally-active immature oocytes is sufficient to cause *hox* de-repression. In mammals, *Smchd1* plays a role in maintaining DNA methylation on the inactive X chromosome and other imprinted genes<sup>4,5,17</sup>. Consistent with being hypomethylated, *smchd1<sup>lof1/lof1</sup>* zebrafish were less susceptible to the effects of the methylation inhibitor 5-Aza-2'-deoxycytidine (5-Aza-dC) (Supplementary Fig. 5a, b). *Hox* genes undergo dynamic DNA methylation changes during early embryogenesis where their methylation levels decrease dramatically from oocyte to ZGA<sup>18,19</sup> (Supp. Fig. 5c). In line with this, we could document the *hoxc10a* locus was markedly de-

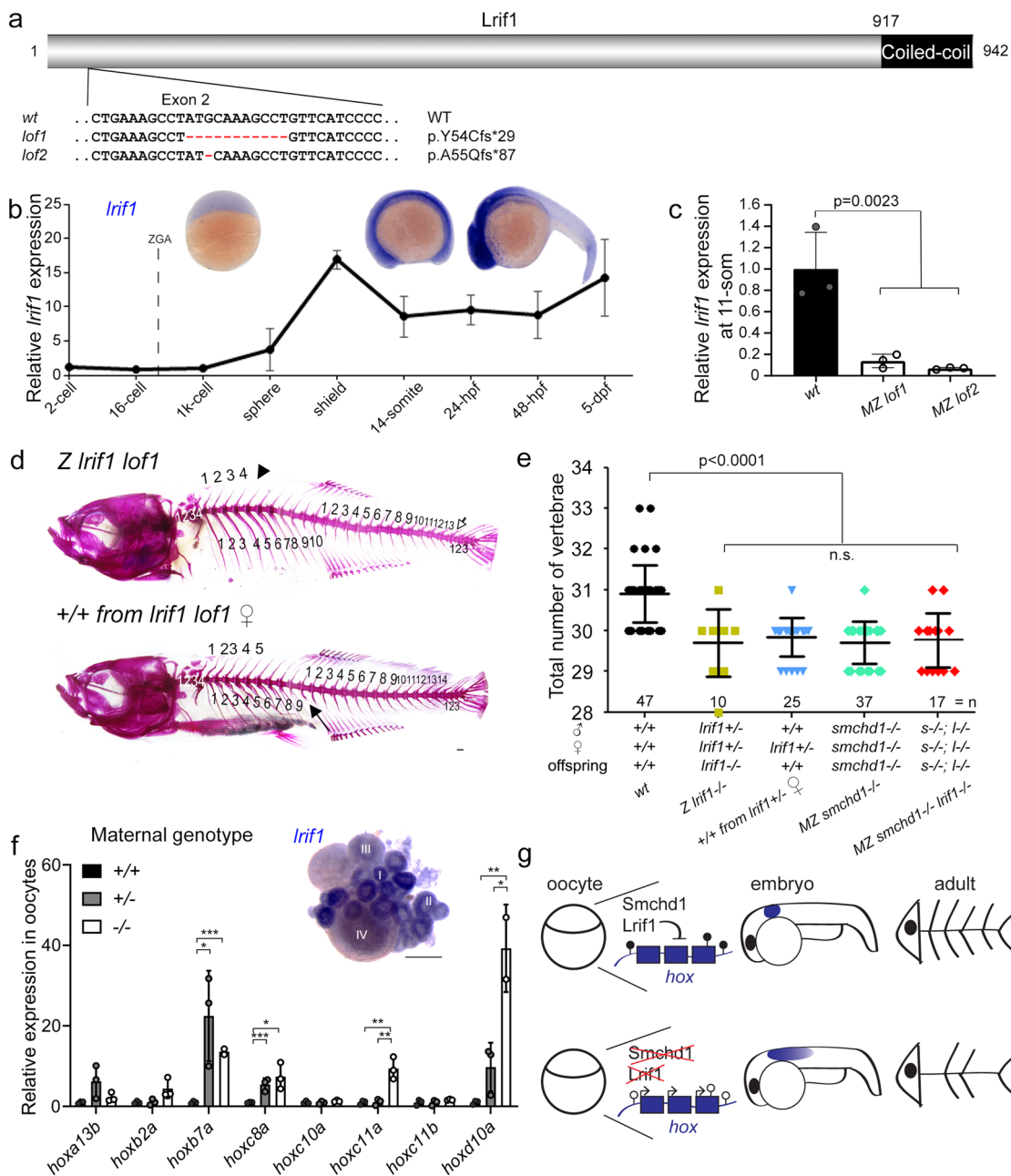


**Fig. 4** *Smchd1* haploinsufficiency in mouse also causes homeotic transformations. **a** A new *Smchd1* loss-of-function allele was generated using CRISPR/Cas9. This produced a 6 nt deletion including the splice donor site of intron 3. This resulted in a number of aberrantly spliced transcripts. The protein variation listed is the smallest in-frame mutation detected, a 14 amino-acid deletion including the catalytic residue of the GHKL ATPase domain p.E147. **b** Western blot from heads of E10.5 embryos show a markedly reduced level of SMCHD1 protein in the  $\Delta 6nt/\Delta 6nt$ , which runs slightly faster than wildtype SMCHD1.  $n =$  at least 5 embryos per genotype. **c** Observed and expected numbers of pups born of each genotype. Axial skeletal patterning analysis of  $+/+$  and  $\Delta 6nt/+$  animals. Homeotic transformations observed in the animals are marked in red in the schematic (**d**). Transformations include C7  $\rightarrow$  T1, T13  $\rightarrow$  L1 and L6  $\rightarrow$  S1. C7 rib is labelled with a \*, the missing T13 rib is marked with an arrow and the hypomorphic T13 rib is labelled with an arrowhead. Number of animals analysed with each phenotype are in (**e**). Animals showing any of the homeotic transformations are considered as transformed. Most  $\Delta 6nt/+$  animals show all of the homeotic transformations marked in (**d**). Scale bar = 500  $\mu$ m.

methylated in *smchd1*<sup>lof1</sup> oocytes as compared to *wt* oocytes (Fig. 3f, Supplementary Fig. 5d, e), suggesting that *Smchd1* helps maintain DNA methylation during oogenesis.

***Smchd1* haploinsufficiency in mouse causes homeotic transformations.** To ensure that this inter-generational inheritance also holds true in mammals, we sought to reexamine the effects of the loss of SMCHD1 in mice. We generated a new *Smchd1* knockout allele in the C57BL/6 background using CRISPR/Cas9 (*Smchd1* <sup>$\Delta 6nt$</sup> ). The selected germline c.603\_603 + 5del mutation deletes 6 nucleotides including the splice donor of intron 3. This deleterious allele results in a number of aberrantly spliced isoforms of which the smallest in-frame deletion removes 14 conserved amino acids of the GHKL ATPase domain p.Gly128\_Pro142delinsAla (Fig. 4a, b) which would generate a catalytically-dead enzyme<sup>20</sup>. Similar to the previously generated *Smchd1* mutant mice (*Smchd1*<sup>MommeD1/MommeD1</sup>)<sup>4</sup>, homozygous

females were embryonic lethal while a sublethal phenotype was also observed in mutant males of this new knockout line (Fig. 4c). Previous studies have shown that SMCHD1 regulates long range chromatin interactions at *Hox* clusters<sup>21</sup>. Axial skeletal anomalies were seen in our newly created heterozygous *Smchd1* <sup>$\Delta 6nt$</sup>  mice whereas they have only been reported in homozygous knockout males previously<sup>21</sup>. The observed homeotic transformations were consistent with those previously described including an ectopic rib at C7 and absence of ribs at T13. Additionally, consistent transformation of L6 to S1 was observed (Fig. 4d, e). We were unable to ascertain whether these observed axial transformations were due to a maternal or zygotic loss of gene activity as all *Smchd1* knockout females were embryonic lethal. However, this conundrum has been recently addressed by Benetti and colleagues<sup>22</sup> who have shown, using an oocyte-specific Cre-mediated deletion of maternal of *Smchd1*, that the maternal pool of *Smchd1* has an important role in repressing *Hox* expression and that also leads to vertebral patterning defects in pups born to

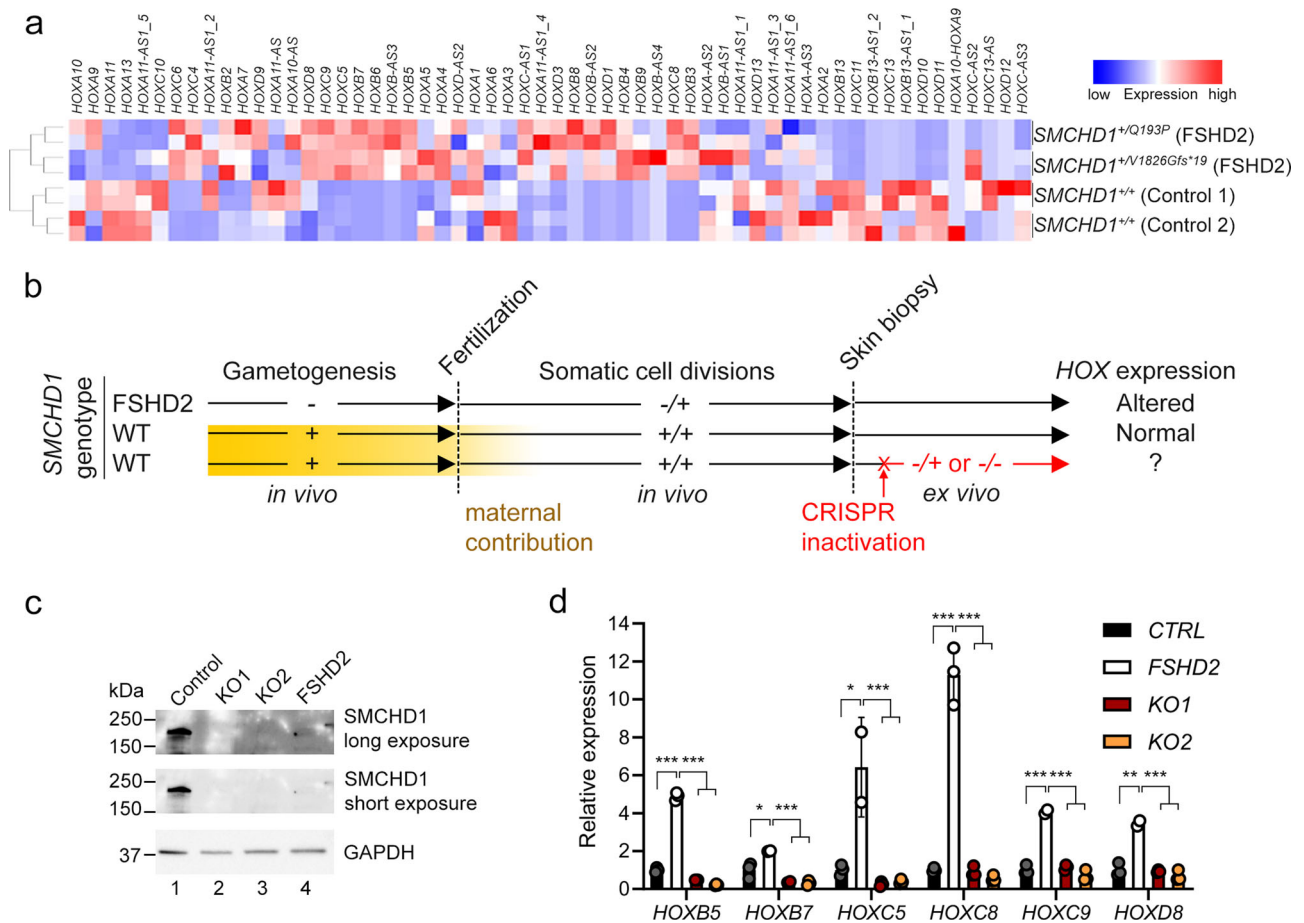


**Fig. 5** Knockout of *Irif1* phenocopies, and does not aggravate, that of *smchd1*. **a** Two loss of function alleles of *Irif1* were generated in zebrafish by CRISPR/Cas9 injections. **b** Endogenous spatial-temporal expression pattern of *Irif1* in zebrafish embryos as determined by qPCR and WISH.  $n = 3$  biological samples (qPCR), 3 independent experiments with at least 20 embryos each (WISH). **c** Maternal zygotic (MZ) *Irif1*<sup>lof</sup> embryos lack *Irif1* mRNA suggestive of total nonsense-mediated decay.  $n = 3$  biological samples.  $P$  values were calculated by 2-tailed unpaired Student’s  $t$  test. **d, e** Axial skeletal patterning defects in adult fish. Phenotypes include loss of a rib (filled arrow), loss of a supraneural vertebra (arrowhead) and loss of a caudal vertebra (open arrow).  $P$  values were calculated by the Kruskal-Wallis test followed by Dunn’s Multiple Comparison Test. **f** qPCR shows that *hox* derepression is detected in unfertilized oocytes from both  $+/-$  and  $-/-$  females.  $P$  values were calculated by 2-tailed unpaired Student’s  $t$  test. Inset shows expression of *Irif1* in Stage I and II oocytes. Scale bars = 200  $\mu$ m.  $n = 3$  biological samples (qPCR), 5 ovaries of adult fish (WISH). **g** Model of *Smchd1/Lrif1*-driven transgenerational inheritance for skeletal patterning. *Smchd1/Lrif1* sets up an epigenetic state during oogenesis to ensure correct expression of *hox* genes during embryogenesis which in turn patterns the vertebrae.

mothers lacking *Smchd1* only during oogenesis. These results are consistent with those made in zebrafish and confirm that *Smchd1* is essential in the maternal germline, before fertilization, to control the expression of *Hox* genes in future embryos. This intergenerational inheritance documented in zebrafish which we propose is driven by *Hox* epi-mutations, may also be at play in higher vertebrate species including mammals.

**Knockout of *Irif1* phenocopies, and does not aggravate, that of *smchd1*.** Ligand-dependent nuclear receptor-interacting factor 1 (*Lrif1*) is the only validated nuclear protein partner of *Smchd1* on chromatin<sup>20,23</sup>. The two nuclear proteins directly bind through *SMCHD1*’s hinge domain and the coiled-coil domain in *LRIF1* (Fig. 5a). In mammals, *LRIF1* is thought to be partially responsible for recruiting *SMCHD1* to the inactive X chromosome<sup>20,23</sup>.





**Fig. 6** Germline, but not adult, *SMCHD1* deficiency leads to *HOX* dysregulation in human cutaneous fibroblasts. **a** RNA-seq performed on controls and FSHD2 patient fibroblasts reveal dysregulations of *HOX* expression when *SMCHD1* was mutated in the germline. **b** Schematic of experimental design to determine whether *HOX* dysregulation is the result of germline or somatic *SMCHD1* haploinsufficiency. *SMCHD1* was CRISPR/ Cas9-inactivated ex vivo in control fibroblasts using two distinct guide RNAs (KO1 and KO2). Yellow box represents maternal contribution of *SMCHD1* which diminishes rapidly post-fertilization. **c** Western blot in fibroblasts cell extract show successful knockout of *SMCHD1*, resulting in protein-null alleles. FSHD2 cells used here are +/V1826Gfs\*19. **d** qPCR shows that, unlike in FSHD2 cells, somatic knockout of *SMCHD1* in adult cells does not lead to the upregulation of *HOX*. *P* values were calculated by 2-tailed unpaired Student's *t* test. \**p* < 0.05, \*\**p* < 0.01, \*\*\**p* < 0.001. Data are presented as mean values ± SD.

To determine if *Smchd1* requires *Lrfl1* to control *hox* expression, we generated *lrfl1* knockout zebrafish (Fig. 5a, c). Unlike *smchd1*, *lrfl1* was almost exclusively zygotically produced with little to no maternal mRNA deposition (Fig. 5b). When bred to homozygosity, zygotic null *lrfl1*<sup>-/-</sup> zebrafish exhibited skeletal abnormalities which were reminiscent of AP patterning defects seen in *smchd1*<sup>-/-</sup> fish. These mutant *lrfl1*<sup>-/-</sup> fish had fewer precaudal and caudal vertebrae (Fig. 5d, e). Despite the absence of measurable maternal *lrfl1* mRNA, we could also document that wildtype *lrfl1*<sup>+/+</sup> offspring, born from *lrfl1*<sup>+/-</sup> mothers crossed to outbred *wt* males, also displayed homeotic transformations (Fig. 5d, e). Likewise, a transcriptional upregulation of *hox* genes was observed in unfertilised oocytes from *lrfl1*<sup>+/-</sup> and *lrfl1*<sup>-/-</sup> females (Fig. 5f), suggestive of maternal haploinsufficiency. Whole-mount in situ hybridisation on ovaries showed that *lrfl1* expression was detected only during the earliest Stages I and II of oogenesis (Fig. 5f). We thus speculate that the *Lrfl1* protein synthesized during early gametogenesis is loaded onto chromatin of maturing oocytes to actively silence *hox* genes before the ZGA. We also generated double maternal zygotic *smchd1*<sup>-/-</sup>; *lrfl1*<sup>-/-</sup> fish, which like single knockouts, presented with comparable skeletal anomalies (Fig. 5e). This combined *Smchd1*/*Lrfl1* deficiency did not further exacerbate vertebral defects, supporting the notion that *Smchd1* and *Lrfl1* act in the same pathway for gene

silencing. This is consistent with the two nuclear proteins working together on chromatin as a single protein complex.

**Germline, but not adult, *SMCHD1* deficiency leads to *HOX* dysregulation in human fibroblasts.** Finally, we set out to examine whether *HOX* misexpression was an ephemeral consequence of maternal *SMCHD1* depletion or rather an epigenetic anomaly that could be inherited through somatic cell divisions. To test this, we turned to *SMCHD1*-haploinsufficient human dermal fibroblasts taken from patients with FSHD2. Transcriptome profiling revealed a patent dysregulation of numerous *HOX* genes in two distinct FSHD2 fibroblasts as compared to those from healthy controls (Fig. 6a). To determine whether this *HOX* dysregulation is the result of haploinsufficiency of *SMCHD1* during human gametogenesis and early embryogenesis, we set out to knock out *SMCHD1* in a control human fibroblast line using CRISPR/Cas9 (Fig. 6b). The somatic inactivation of *SMCHD1* resulted in protein-null cell populations (Fig. 6c) but did not lead to *HOX* dysregulation as was observed in FSHD2 fibroblasts (Fig. 6d). This is reminiscent of previous work showing that somatic *SMCHD1* inactivation does not trigger hypomethylation of D4Z4<sup>24</sup>. Our results indicate that *HOX* misexpression caused by *SMCHD1* haploinsufficiency is set up in the germline, can be stably inherited

through somatic cell divisions and cannot be recapitulated by deleting SMCHD1 in adulthood.

## Discussion

Using zebrafish, we uncovered a maternal role of *Smchd1* that was masked by the mammalian-specific processes of X-inactivation and imprinting<sup>12</sup>. We propose that the ancestral role of *Smchd1* is to regulate developmental genes before being co-opted in mammals for imprinting. Hox genes are zygotic genes that are under multiple layers of transcriptional and post-transcriptional control<sup>25–28</sup>. Here we provide genetic, phenotypic and functional evidence that maternally supplied *Smchd1* and *Lr1f1* are key in preventing premature activation of *hox* genes before fertilization by laying down an epigenetic state in the female germline (Fig. 5g). This parallels mammalian functions of maternal Trim28, which governs the epigenetic state in oocytes to ensure proper development<sup>29</sup>. In zebrafish, *hox* loci are distinctively occupied before ZGA by placeholder nucleosome H2AFV before resolving into bivalent chromatin marked by H3K4me3 and H3K27me3<sup>30,31</sup>. Future work should examine whether *Lr1f1*/*Smchd1* can recognize specific histone variants or modifications.

SMCHD1 belongs to the top 3% of genes most intolerant to mutations in humans<sup>32,33</sup>. Accordingly, removal of one copy of *SMCHD1* is sufficient to cause FSHD2 in humans. Recessive loss-of-function mutations in *LRIF1* can also cause FSHD2<sup>34</sup>. Our findings document misregulation of numerous *HOX* genes in primary fibroblasts taken from FSHD2 patients. This cannot be recapitulated by the somatic inactivation of *SMCHD1*, suggesting FSHD2's aetiology, even if seen as late-onset muscular dystrophy, might have a developmental origin. This also invites the question of whether pervasive *HOX* epi-mutations, driven by insufficient SMCHD1/*LRIF1* activity during gametogenesis, may contribute to the pathogenesis of FSHD2 which exhibits anterior-biased muscle degeneration. Notably, hypoplastic or absent ribs have been reported in some BAMS patients carrying heterozygous missense mutations in *SMCHD1*<sup>35</sup>, drawing a possible link between SMCHD1 and anterior-posterior patterning in humans.

*smchd1* or *Lr1f1* haploinsufficient mothers can produce genetically-wildtype offspring that exhibit homeotic transformations. This mode of inheritance where the offspring's phenotype is one generation removed from its causal parental genotype has been referred to as inter-generational inheritance<sup>36</sup>. Our study may impact the interpretations of unsolved Mendelian diseases, in which the proband's phenotype may be attributed to the parent's genotype through epigenetic inheritance.

## Methods

**CRISPR-mediated zebrafish knockout.** Zebrafish were maintained and used according to the Institutional Animal Care and Use Committee of Biological Resource Centre, A\*STAR, Singapore (IACUC #161172) and the National University of Singapore (IACUC #BR19-1184). Guide RNA against exon 3 of *smchd1* was used with the following targeting sequence, 5'-TACGAGTATTACGCCAGCGA-3'. Guide RNA against exon 1 of *Lr1f1* was used with the following sequence, 5'-GGGATG AACAGGCTTTGCAT-3'. gRNA was synthesized using a MEGA shortscript Kit (Thermo Fisher) according to the manufacturer's protocol and purified with an RNeasy Mini Kit (QIAGEN). Cas9 mRNA was synthesized with an mMESSAGE mMACHINE Kit (SP6) (Thermo Fisher) and a NotI-linearized zebrafish codon-optimized Cas9 construct in pCS2+. The gRNA and Cas9 mRNA were mixed together to a concentration of 375 ng/ $\mu$ l each, and 1 nL of RNA was injected into the yolk of 1-cell AB zebrafish embryos. Zebrafish are genotyped by PCR followed by Sanger sequencing. Primers used for genotyping are in Supplementary Data 1.

**CRISPR-mediated mouse knockout.** All mice procedures were approved by the Institutional Animal Care and Use Committee of Biological Resource Centre, A\*STAR, Singapore (IACUC #201555). All mice used were on the C57BL/6 background. The *Smchd1*  $\Delta$ 6nt mouse was generated by CRISPR/ Cas9 injection into the pronucleus of 1-cell zygotes. Guide RNA against exon 3 was used with the following targeting sequence, 5'-gagcaggttactacgca-3', which was previously used

in Shaw et al. (2017)<sup>8</sup>. Founders were backcrossed with C57BL/6. Genotyping primers are listed in Supplementary Data 1.

**Western blot and antibodies.** Zebrafish embryos at sphere stage were dechorionated using 1 mg/ml pronase and deyolked as previously described<sup>37</sup>. To enrich for nuclear proteins, embryos were lysed in hypotonic lysis buffer (10 mM Tris pH7.5, 10 mM NaCl, 3 mM MgCl<sub>2</sub>, 0.3% NP40, 10% glycerol, 1x Protease inhibitor) for 10 min on ice and centrifuged 800 g for 8 min. The supernatant containing cytoplasmic fraction was discarded. The pellet containing nuclear proteins was resuspended in MWB (10 mM Tris pH: 7.5, 300 mM NaCl, 4 mM EDTA, 1% NP40, 10% glycerol, TURBO DNase, 1x Protease Inhibitor). Nuclear proteins from 300 embryos were loaded per lane. Human fibroblasts and mouse embryo heads were lysed in RIPA buffer (250 mM Tris, pH: 7.5; 150 mM NaCl; 1% NP-40; 0.5% Na deoxycholate, Halt protease inhibitors [Thermo]). Immunoblotting was performed using the following antibodies: anti-SMCHD1 (Atlas HPA039441, 1:500), anti-Histone H3 (abcam ab1791, 1:1000), anti-GAPDH (Santa Cruz SC47724, 1:1000).

In situ hybridization Probes for in situ hybridisation were cloned from zebrafish embryo cDNA using primers in Supplementary Data 1.

**Zebrafish whole mount immunofluorescence.** 33 hpf embryos were dechorionated and fixed for 2 h at room temperature in 4% paraformaldehyde in PBS. The embryos were stored in methanol at  $-20^{\circ}\text{C}$  and then rehydrated in a gradient of PBS/methanol. The embryos were blocked in PBST (PBS, 1% BSA, 1% DMSO and 1% Triton X-100) with 5% goat serum at room temperature for 1 h. Embryos were incubated with anti-GnRH antibodies (Sigma G8294, 1:500) overnight at  $4^{\circ}\text{C}$  and washed in PBST. Secondary antibody used was Alexa fluorophore-labelled anti-rabbit (1:500, Molecular Probes). Embryos were mounted with 2% low melting agarose on slides and imaged with Olympus FV1000 confocal microscope. Terminal nerve length was measured using ImageJ.

**qPCR.** RNA from 30–50 zebrafish embryos at the designated stages or human fibroblast was extracted using RNeasy Mini Kit (Qiagen) with DNase digestion on column. cDNA was synthesized from 1  $\mu$ g of total RNA using the iScript cDNA Synthesis Kit (BioRad). qRT-PCR was performed with Power SYBR Green PCR Master Mix (Thermo Fisher) on QuantStudio 7 (Thermo Fisher). Gene expression was normalized to *actin* for zebrafish samples and *HPRT* for human samples. Data represent mean  $\pm$  SD of biological replicates. A list of primer sequences can be found in Supplementary Data 1.

**miRNA expression analysis.** RNA from 30–50 zebrafish embryos at the designated stages was extracted using Trizol (Thermo Fisher). cDNA was synthesized using Mir-X miRNA First Strand Synthesis kit (Takara Bio) and qPCR performed using TB Green Advantage qPCR Premix (Takara Bio). Gene expression was normalized to *U6* using primers provided with the kit. Data represent mean  $\pm$  SD of biological replicates. Primer sequences can be found in Supplementary Data 1.

**Skeletal analysis.** Zebrafish of 2 months or older were euthanized in an ice bath. The skin, eyes and internal organs of the fish are removed before fixing in 4% PFA at RT for 2 days. Fish are washed with water and moved into a saturated sodium borate solution for 0.5 days. Fish are then bleached in 0.45% H<sub>2</sub>O<sub>2</sub> in 1% KOH for 1 h under light, then kept in 1% KOH overnight. Bones are stained in 0.1% alizarin red in 1% KOH in the dark overnight at RT. Skeletons are cleared in 0.05% trypsin in 35% saturated sodium borate overnight followed by 1% KOH overnight. Clearing is repeated until bones are clearly seen. Skeletons are then passed through a glycerol-1% KOH series and stored in 80% glycerol in 1% KOH.

For craniofacial cartilage staining of larval zebrafish, 5 dpf zebrafish were fixed overnight in 4% PFA then dehydrated through a methanol series. Larval fish are stained overnight in Alcian Blue Stain (0.1% Alcian Blue, 70% Ethanol, 30% Glacial Acetic Acid) followed by 3 washes in 100% Ethanol for 5 min each. Larval fish are then rehydrated through an ethanol series into PBST and digested in 0.05% trypsin for 2 h at room temperature. Tissue is cleared in 1% KOH for 10 min. The skeletons are passed through a glycerol-1% KOH series and stored in 80% glycerol in 1% KOH.

P0 mouse pups were deskked, eviscerated and fixed in 95% ethanol overnight. They were then transferred to acetone overnight to remove fats. Pups were stained with alcian blue solution (0.05% Alcian Blue, 80% Ethanol, 20% Glacial Acetic Acid) for 24 hours. They were washed in 70% Ethanol over 6–8 h. They were cleared in 1% KOH until the tissues are visibly cleared, then stained in Alizarin red solution (50  $\mu$ g/ml Alizarin red in 1% KOH) in the dark for 6 h. The skeletons were de-stained in 1% KOH for 1 day. They are then passed through a glycerol-1% KOH series and stored in 80% glycerol in 1% KOH.

**RNA-seq analysis.** 3 biological samples each of *wt* and *MZ lof1* embryos at 4–8-cell and sphere stages were used in the RNA-seq. The quantity and integrity of RNA were measured using the Bioanalyzer 2100 (Agilent Technologies). Samples were poly-A enriched and sequenced (paired-end) on HiSeq4000 (Illumina). Reads were mapped onto the danRer10 zebrafish reference genome and the genes were annotated using the *Danio rerio* annotation from Ensembl release 89. Genes with a normalized mean expression  $\leq 10$  and ribosomal RNA genes were removed from further analysis.

DESeq2 was used to perform differential gene expression analysis. Genes with a FDR value below 0.05, mean expression > 10 and fold change > 2 were classified as differentially expressed (DEGs). K-means clustering was performed using silhouette width and within-sum-of-squares metrics. DEGs were enriched against Gene Ontology Biological Processes (GO:BP), Gene Ontology Molecular Function (GO:MF), and Kyoto Encyclopaedia of Genes and Genomes (KEGG) ontologies. Gene Set Enrichment Analysis (GSEA) was performed on all genes expressed. RNA seq methods in primary fibroblast are described in Laberthionniere et al.<sup>38</sup>

**DNA methylation analysis.** Unfertilized oocytes were squeezed out from anesthetized adult female fish. Oocytes were activated in egg water for 10 min. Eggs that failed to activate or had lysed were discarded. Healthy eggs were collected into tubes and lysed in a chilled cell lysis solution (10 mM Tris pH8, 10 mM NaCl, 0.5% NP40) by passing them through a 20 G needle. Nuclei were collected by centrifugation at 3500 g for 5 min at 4 °C. The supernatant was removed and the nuclei pellet lysed in nuclear lysis buffer (50 mM Tris pH8, 10 mM EDTA, 1% SDS, 10 µl Proteinase K, 2 µl RNase A/ml) at 55 °C for 2 h. DNA was extracted by phenol chloroform extraction. DNA was digested again with RNase A at 37 °C for 3 h, followed by phenol chloroform extraction. DNA from 500 eggs was subjected to bisulfite conversion using MethylEdge Bisulfite Conversion System (Promega). PCR was performed with HotStarTaq (Qiagen) and PCR product TA-cloned into pGEM-T Easy (Promega). 16 clones were picked per genotype and sequenced with T7 primer. Efficiency of bisulfite conversion was over 99% as observed by the conversion of non-CpG cytosines.

**Zebrafish odour tests.** In total, 2–4 weeks old larval fish were gently introduced into the observation tank (75 mm × 30 mm × 30 mm (L × W × H) with 40 ml of water. Video recordings started after 3–5 min of acclimation. After 5 min of video recording, 0.5 ml of 100 µM glycochenodeoxycholic acid (GCDA) at one end of the tank and a similar volume of tank water from the other end were delivered. Recording continued for the next 5 minutes. Fish were tracked online and tracked data was analyzed as described previously<sup>39</sup>. In total, 20% area of the tank from the delivery port was defined as the stimulus zone.

**Schreckstoff preparation and behavioural data analysis** was performed as described previously<sup>40</sup> with minor modifications. Briefly, *Schreckstoff* was prepared from euthanized zebrafish by introducing 7 to 10 shallow lesions with a Sharp point knife (22.5° stab). Fish were then immersed into 2 ml of fish water for 1 to 2 min and rocked on a rocker. The 2 ml crude extract was then centrifuged at 20,000 × g, filtered, and heated overnight at 95 °C.

Adult subject fish were transferred from the fish facility and gently introduced into the observation tank. Video recordings started after 5–7 min of acclimation. 250 µl of *Schreckstoff* was introduced via an automated delivery valve after 120 s of video recording (Pre) at 1 ml/minute. Recording continued for the next 120 s including the delivery period (Post). For analyzing the position and speed of the fish swimming, videos were re-digitized at 15 fps and fish position was tracked automatically using the “track objects” algorithm in MetaMorph. 6.3. Automated algorithms were used to compute all parameters described from the XY position data of the subject.

**5-Aza-dC treatment.** 8-cell stage embryos were treated with various concentrations of DNA methylation inhibitor 5-Aza-2'-deoxycytidine (5-Aza-dC) for 48 h then washed to remove the inhibitor. Embryos were observed for survival and developmental defects. Dead embryos were removed daily.

**Cell culture.** Written informed consent of individuals were received prior to collection of samples and subsequent analysis for research purposes according to the ethical approvals of the local Institutional Review Boards in Singapore (A\*STAR IRB 2019-087, NUS IRB N-20-054E) and France (La Timone Children's Hospital). All skin biopsies except for one control (foreskin) were taken from the forearms of individuals. Primary fibroblasts were cultured in DMEM (Life Technologies) supplemented with 10% foetal bovine serum (Hyclone). gRNAs targeting SMCHD1 were cloned into LentiCRISPRv2 puro (a gift from Brett Stringer, Addgene plasmid # 98290). Two guide RNAs targeting 5'-CTTACCTCACTATGACACAC-3' (KO1) and 5'-TTGTATTACAACAACAAGT-3' (KO2) were used. Lentiviruses were packaged with third generation packaging plasmids. Following viral transduction, cells were selected in 1 µg/ml puromycin.

**Statistics.** All graphs represent means ± standard deviation. *P* values were calculated by the Kruskal-Wallis test followed by Dunn's Multiple Comparison Test, unless otherwise stated. \**p* < 0.05, \*\**p* < 0.01, \*\*\**p* < 0.001.

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

## Data availability

The RNA-seq data produced in this study were submitted to GEO under the following accession number: [GSE173462](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE173462), [GSE174604](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE174604). Source data are provided as a Source Data file. Source data are provided with this paper.

Received: 15 July 2021; Accepted: 7 June 2022;

Published online: 23 June 2022

## References

- Vastenhouw, N. L., Cao, W. X. & Lipshitz, H. D. The maternal-to-zygotic transition revisited. *Development* **146**, dev161471 (2019).
- Mitchell, L. E. Maternal effect genes: Update and review of evidence for a link with birth defects. *Hum. Genet. Genomics Adv.* **3**, 100067 (2022).
- Jansz, N., Chen, K., Murphy, J. M. & Blewitt, M. E. The epigenetic regulator SMCHD1 in development and disease. *Trends Genet.* **33**, 233–243 (2017).
- Blewitt, M. E. et al. SmcHD1, containing a structural-maintenance-of-chromosomes hinge domain, has a critical role in X inactivation. *Nat. Genet.* **40**, 663–669 (2008).
- Gendrel, A.-V. et al. Epigenetic functions of smchd1 repress gene clusters on the inactive X chromosome and on autosomes. *Mol. Cell. Biol.* **33**, 3150–3165 (2013).
- Lemmers, R. J. L. F. et al. Digenic inheritance of an SMCHD1 mutation and an FSHD-permissive D4Z4 allele causes facioscapulohumeral muscular dystrophy type 2. *Nat. Genet.* **44**, 1370–1374 (2012).
- Gordon, C. T. et al. De novo mutations in SMCHD1 cause Bosma arhinia microphthalmia syndrome and abrogate nasal development. *Nat. Genet.* **49**, 249–255 (2017).
- Shaw, N. D. et al. SMCHD1 mutations associated with a rare muscular dystrophy can also cause isolated arhinia and Bosma arhinia microphthalmia syndrome. *Nat. Genet.* **49**, 238–248 (2017).
- Deng, Q., Ramsköld, D., Reinius, B. & Sandberg, R. Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells. *Science* **343**, 193–196 (2014).
- Yan, L. et al. Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. *Nat. Struct. Mol. Biol.* **20**, 1131–1139 (2013).
- Midic, U. et al. Novel key roles for structural maintenance of chromosome flexible domain containing 1 (Smchd1) during preimplantation mouse development. *Mol. Reprod. Dev.* **85**, 635–648 (2018).
- Wanigasuriya, I. et al. is a maternal effect gene required for genomic imprinting. *Elife* **9**, e55529 (2020).
- Wellik, D. M. Chapter 9 Hox Genes and Vertebrate Axial Pattern. *Current Topics in Developmental Biology* 257–278 [https://doi.org/10.1016/s0070-2153\(09\)88009-5](https://doi.org/10.1016/s0070-2153(09)88009-5) (2009).
- Duboule, D. The rise and fall of Hox gene clusters. *Development* **134**, 2549–2560 (2007).
- He, X. et al. miR-196 regulates axial patterning and pectoral appendage initiation. *Dev. Biol.* **357**, 463–477 (2011).
- Wong, S. F. L. et al. Independent regulation of vertebral number and vertebral identity by microRNA-196 paralogs. *Proc. Natl Acad. Sci. USA* **112**, E4884–E4893 (2015).
- Mould, A. W. et al. Smchd1 regulates a subset of autosomal genes subject to monoallelic expression in addition to being critical for X inactivation. *Epigenetics Chromatin* **6**, 19 (2013).
- Potok, M. E., Nix, D. A., Parnell, T. J. & Cairns, B. R. Reprogramming the maternal zebrafish genome after fertilization to match the paternal methylation pattern. *Cell* **153**, 759–772 (2013).
- Jiang, L. et al. Sperm, but not oocyte, DNA methylome is inherited by zebrafish early embryos. *Cell* **153**, 773–784 (2013).
- Brideau, N. J. et al. Independent Mechanisms Target SMCHD1 to Trimethylated Histone H3 Lysine 9-Modified Chromatin and the Inactive X Chromosome. *Mol. Cell. Biol.* **35**, 4053–4068 (2015).
- Jansz, N. et al. Smchd1 regulates long-range chromatin interactions on the inactive X chromosome and at Hox clusters. *Nat. Struct. Mol. Biol.* **25**, 766–777 (2018).
- Benetti, N. et al. Maternal SMCHD1 regulates Hox gene expression and patterning in the mouse embryo. <https://doi.org/10.1101/2021.09.08.459528>.
- Nozawa, R.-S. et al. Human inactive X chromosome is compacted through a PRC2-independent SMCHD1-HBix1 pathway. *Nat. Struct. Mol. Biol.* **20**, 566–573 (2013).
- Dion, C. et al. SMCHD1 is involved in de novo methylation of the DUX4-encoding D4Z4 macrosatellite. *Nucleic Acids Res.* **47**, 2822–2839 (2019).
- Yekta, S., Tabin, C. J. & Bartel, D. P. MicroRNAs in the Hox network: an apparent link to posterior prevalence. *Nat. Rev. Genet.* **9**, 789–796 (2008).
- Kondrashov, N. et al. Ribosome-mediated specificity in Hox mRNA translation and vertebrate tissue patterning. *Cell* **145**, 383–397 (2011).
- Noordermeer, D. et al. The dynamic architecture of Hox gene clusters. *Science* **334**, 222–225 (2011).
- Schumacher, A., Faust, C. & Magnuson, T. Positional cloning of a global regulator of anterior-posterior patterning in mice. *Nature* **384**, 648 (1996).

29. Messerschmidt, D. M. et al. Trim28 is required for epigenetic stability during mouse oocyte to embryo transition. *Science* **335**, 1499–1502 (2012).
30. Murphy, P. J., Wu, S. F., James, C. R., Wike, C. L. & Cairns, B. R. Placeholder nucleosomes underlie germline-to-embryo DNA methylation reprogramming. *Cell* **172**, 993–1006.e13 (2018).
31. Vastenhouw, N. L. et al. Chromatin signature of embryonic pluripotency is established during genome activation. *Nature* **464**, 922–926 (2010).
32. Petrovski, S., Wang, Q., Heinzen, E. L., Allen, A. S. & Goldstein, D. B. Genic intolerance to functional variation and the interpretation of personal genomes. *PLoS Genet.* **9**, e1003709 (2013).
33. Lek, M. et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285–291 (2016).
34. Hamanaka, K. et al. Homozygous nonsense variant in associated with facioscapulohumeral muscular dystrophy. *Neurology* **94**, e2441–e2447 (2020).
35. Delaney, A. et al. Insight Into the Ontogeny of GnRH Neurons From Patients Born Without a Nose. *J. Clin. Endocrinol. Metab.* **105**, 1538–1551 (2020).
36. Perez, M. F. & Lehner, B. Intergenerational and transgenerational epigenetic inheritance in animals. *Nat. Cell Biol.* **21**, 143–151 (2019).
37. Link, V. et al. Identification of regulators of germ layer morphogenesis using proteomics in zebrafish. *J. Cell Sci.* **119**, 2073–2083 (2006).
38. Laberthonnière, C. et al. AKT Signaling Modifies the Balance between Cell Proliferation and Migration in Neural Crest Cells from Patients Affected with Bosma Arhinia and Microphthalmia Syndrome. *Biomedicines* **9**, 751 (2021).
39. Krishnan, S. et al. The right dorsal habenula limits attraction to an odor in zebrafish. *Curr. Biol.* **24**, 1167–1175 (2014).
40. Schirmer, A., Jesuthasan, S. & Mathuru, A. S. Tactile stimulation reduces fear in fish. *Front. Behav. Neurosci.* **7**, 167 (2013).

## Acknowledgements

We are grateful to all members of the Reversade and Xue lab for support. We thank Keerthika, Lydia Ng, Glenn Koh for technical assistance, IMCB-MMD (formerly IMB-KOre unit) for generating the knockout mice and P. Ingham for feedback. B.R. is indebted to Davor Solter, for his mentorship, acumen and inspiration. This work was supported by NMRC/OFYIRG/0062/2017 and NUS PYP startup grant to S.X. Work in A.S.M.'s lab was supported by the Ministry of Education, Singapore (MOE) and Yale-NUS College through IG19-BG106, MOE-T2EP30220-0020, and an SUG. Work in F.M.'s lab was funded by “Association Française contre les Myopathies” (AFM; TRIM-RD, MoThARD) and Fondation Maladies Rares. B.R. is an investigator of the National Research Foundation (NRF, Singapore), Branco Weiss Foundation (Switzerland) and an EMBO Young Investigator, and is supported by an inaugural Use-Inspired Basic Research (UIBR) central fund from the Agency for Science & Technology and Research (A\*STAR) in Singapore.

## Author contributions

S.X. and B.R. designed the study and obtained funding. S.X., T.T.N.L., J.C. and J.N. performed the experiments. A.S.M. performed and analyzed olfactory experiments. A.S.M. and R.S.V. analyzed RNA-seq data. F.M. provided patient material and data. S.X. and B.R. wrote the manuscript with input from all authors.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-022-31185-8>.

**Correspondence** and requests for materials should be addressed to Shifeng Xue or Bruno Reversade.

**Peer review information** *Nature Communications* thanks Gavin Kelsey, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

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



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

© The Author(s) 2022

## Terms and Conditions

Springer Nature journal content, brought to you courtesy of Springer Nature Customer Service Center GmbH (“Springer Nature”).

Springer Nature supports a reasonable amount of sharing of research papers by authors, subscribers and authorised users (“Users”), for small-scale personal, non-commercial use provided that all copyright, trade and service marks and other proprietary notices are maintained. By accessing, sharing, receiving or otherwise using the Springer Nature journal content you agree to these terms of use (“Terms”). For these purposes, Springer Nature considers academic use (by researchers and students) to be non-commercial.

These Terms are supplementary and will apply in addition to any applicable website terms and conditions, a relevant site licence or a personal subscription. These Terms will prevail over any conflict or ambiguity with regards to the relevant terms, a site licence or a personal subscription (to the extent of the conflict or ambiguity only). For Creative Commons-licensed articles, the terms of the Creative Commons license used will apply.

We collect and use personal data to provide access to the Springer Nature journal content. We may also use these personal data internally within ResearchGate and Springer Nature and as agreed share it, in an anonymised way, for purposes of tracking, analysis and reporting. We will not otherwise disclose your personal data outside the ResearchGate or the Springer Nature group of companies unless we have your permission as detailed in the Privacy Policy.

While Users may use the Springer Nature journal content for small scale, personal non-commercial use, it is important to note that Users may not:

1. use such content for the purpose of providing other users with access on a regular or large scale basis or as a means to circumvent access control;
2. use such content where to do so would be considered a criminal or statutory offence in any jurisdiction, or gives rise to civil liability, or is otherwise unlawful;
3. falsely or misleadingly imply or suggest endorsement, approval, sponsorship, or association unless explicitly agreed to by Springer Nature in writing;
4. use bots or other automated methods to access the content or redirect messages
5. override any security feature or exclusionary protocol; or
6. share the content in order to create substitute for Springer Nature products or services or a systematic database of Springer Nature journal content.

In line with the restriction against commercial use, Springer Nature does not permit the creation of a product or service that creates revenue, royalties, rent or income from our content or its inclusion as part of a paid for service or for other commercial gain. Springer Nature journal content cannot be used for inter-library loans and librarians may not upload Springer Nature journal content on a large scale into their, or any other, institutional repository.

These terms of use are reviewed regularly and may be amended at any time. Springer Nature is not obligated to publish any information or content on this website and may remove it or features or functionality at our sole discretion, at any time with or without notice. Springer Nature may revoke this licence to you at any time and remove access to any copies of the Springer Nature journal content which have been saved.

To the fullest extent permitted by law, Springer Nature makes no warranties, representations or guarantees to Users, either express or implied with respect to the Springer nature journal content and all parties disclaim and waive any implied warranties or warranties imposed by law, including merchantability or fitness for any particular purpose.

Please note that these rights do not automatically extend to content, data or other material published by Springer Nature that may be licensed from third parties.

If you would like to use or distribute our Springer Nature journal content to a wider audience or on a regular basis or in any other manner not expressly permitted by these Terms, please contact Springer Nature at

[onlineservice@springernature.com](mailto:onlineservice@springernature.com)