nature genetics

LETTERS

https://doi.org/10.1038/s41588-019-0418-7

Loss of DUX causes minor defects in zygotic genome activation and is compatible with mouse development

Zhiyuan Chen[1,2,3] and Yi Zhang[1,2,3,4,5*]

How maternal factors in oocytes trigger zygotic genome activation (ZGA) is a long-standing question in developmental biology. Recent studies in 2-cell-like embryonic stem cells (2C-like cells) suggest that transcription factors of the DUX family are key regulators of ZGA in placental mammals[1,2]. To characterize the role of DUX in ZGA, we generated *Dux* cluster knockout (KO) mouse lines. Unexpectedly, we found that both *Dux* zygotic KO (Z-KO) and maternal and zygotic KO (MZ-KO) embryos can survive to adulthood despite showing reduced developmental potential. Furthermore, transcriptome profiling of the MZ-KO embryos revealed that loss of DUX has minimal effects on ZGA and most DUX targets in 2C-like cells are normally activated in MZ-KO embryos. Thus, contrary to the key function of DUX in inducing 2C-like cells, our data indicate that DUX has only a minor role in ZGA and that loss of DUX is compatible with mouse development.

In mammals, early embryonic development is supported first by maternal factors in the egg and later by newly transcribed genes from the zygotic genome. Successful ZGA is essential for embryonic development. In mice, the major wave of ZGA takes place at the 2-cell stage with the activation of thousands of genes and transposable elements, including the ERVL-family retrotransposons[3–6]. Interestingly, ERVL and ERVL-linked genes can also be activated spontaneously in rare and transient embryonic stem (ES) cells known as 2C-like cells[7,8]. As 2C-like cells mimic 2-cell embryos in terms of the expression of 2-cell transient transcripts and have the capacity to contribute to both embryo and extra-embryonic tissues[7], 2C-like cells have been a useful model for understanding totipotency[9] and early embryonic development[10–13]. However, 2C-like cells are not equivalent to 2-cell embryos, as the genes induced in 2C-like cells only represent a subset of the mouse ZGA genes that are activated in 2-cell stage embryos[7].

Dux (also known as *Duxf3*) in mice and its human homolog *DUX4* are double-homeodomain genes that are activated at the onset of ZGA in early embryos[1,2]. In mice, the *Dux* cluster also includes a truncated variant named *Gm4981* (also known as *Duxf4*), which lacks the first homeodomain and is transcribed as early as during oogenesis[2]. In humans, incomplete silencing of *DUX4* causes facioscapulohumeral muscular dystrophy (FSHD)[14] characterized by de-repression of genes and repeats, such as *ZSCAN4* and ERVL, that are only expressed during ZGA in muscle cells in patients with FSHD[15]. In addition, in ES cells, mouse DUX can activate ERVL-family repeats and ERVL-linked genes and is both necessary and sufficient for the ES to 2C-like cell transition[1,2,16]. Furthermore, acute depletion of the *Dux* cluster in zygotes by clustered regularly interspaced short palindromic repeats (CRISPR)–Cas9 injection leads to downregulation of a handful of ZGA genes, as revealed by quantitative PCR with reverse transcription (RT-qPCR)[2]. These results suggest that DUX may have an important role in mouse ZGA and embryonic development[2].

To comprehensively define the role of DUX in ZGA, we attempted to acutely deplete the *Dux*-containing macrosatellite repeats (estimated ~160 kb in C57BL/6 (ref. [17])) by zygotic CRIPSR–Cas9 injection. The pair of single guide RNA (sgRNA) co-injected with the Cas9 mRNA target the same flanking sequences of the *Dux* cluster[2] (Supplementary Fig. 1a). Genotyping of the blastocysts revealed that 25% (9 out of 36) and 5.6% (2 out of 36) of blastocysts carried the mono-allelic and bi-allelic *Dux* cluster deletions, respectively (Supplementary Fig. 1b–d), which is comparable to the reported efficiency for CRIPSR–Cas9-mediated large (>10 kb) genomic fragment deletions[18–22]. As the acute depletion experiment suggests that DUX is not essential for pre-implantation development, we generated mouse lines carrying the *Dux* KO allele to further characterize the role of DUX in mouse development.

To this end, following zygotic CRIPSR–Cas9 injection, 2-cell embryos were transferred to pseudopregnant female mice to obtain live pups. Out of the 87 transferred 2-cell embryos, 27 living pups were obtained and 6 out of the 20 genotyped F_0 mice harbored the *Dux* cluster deletion on one of the alleles (Fig. 1a and Supplementary Fig. 2a,b). When the two F_0 mice (that is, 423 and 426) were backcrossed with wild-type C57BL/6 (WT B6) mice, we obtained 85 (51%) WT and 82 (49%) *Dux* heterozygous (*Dux* Het) mice from a total of 22 litters (Supplementary Fig. 2c). The observed WT/Het ratio is consistent with the expected 50:50 Mendelian frequency, indicating that *Dux* heterozygosity does not impair mouse development.

To determine whether DUX deficiency causes developmental arrest, we genotyped 255 pups from 35 litters of *Dux* Het × Het ($F_1 \times F_1$ and $F_2 \times F_2$) mating pairs (Supplementary Fig. 2c). Contrary to the expectation that *Dux* Z-KO embryos arrest during pre-implantation development owing to ZGA defects, *Dux* Z-KO mice can survive to adulthood without obvious abnormalities, although they are born at a reduced frequency (a birth rate of 18% versus an expected birth rate of 25%, $P = 0.005$) (Fig. 1b,c and Supplementary Fig. 2c). To exclude the possibility that *Dux* copies outside of the deleted macrosatellite repeats may compensate for DUX deficiency, we determined *Dux* RNA levels in the testis, one of the few organs in which *Dux* is expressed in adults[14].

[1]Howard Hughes Medical Institute, Boston Children's Hospital, Boston, MA, USA. [2]Program in Cellular and Molecular Medicine, Boston Children's Hospital, Boston, MA, USA. [3]Division of Hematology/Oncology, Department of Pediatrics, Boston Children's Hospital, Boston, MA, USA. [4]Department of Genetics, Harvard Medical School, Boston, MA, USA. [5]Harvard Stem Cell Institute, Boston, MA, USA. *e-mail: yzhang@genetics.med.harvard.edu

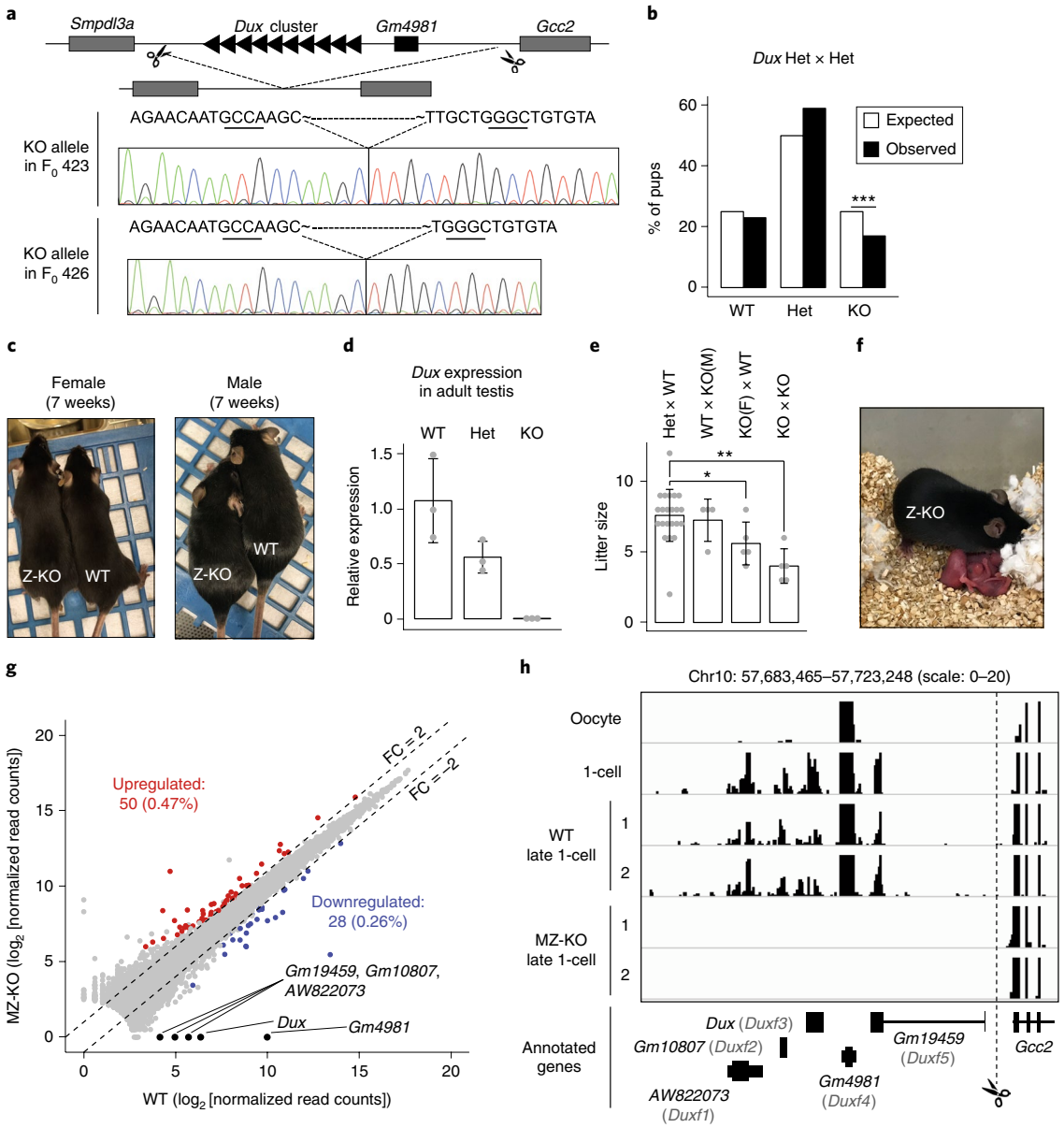

Fig. 1 | Loss of DUX is compatible with mouse development. a, Schematic of the *Dux* cluster in mice and Sanger sequencing results of the KO alleles in the two founder lines. The three underlined nucleotides represent the CRISPR protospacer adjacent motif (PAM) sequences. **b**, Bar graph showing the percentage of pups for each genotype from *Dux* Het × Het crosses. Litter size 7.3 ± 2.5, 35 litters, 255 live pups. *** $P = 0.005$, chi-squared goodness of fit test. **c**, Examples of *Dux* F_2 WT and Z-KO adult mice (analyzed in part **b**). **d**, RT-qPCR results confirming *Dux* KO in adult testis. The expression level of *Dux* in WT adult mouse (9–12 weeks) testis was set as 1.0. Three mice were analyzed for each genotype (denoted as grey dots). Center line and error bars indicate mean and s.d., respectively. **e**, Litter sizes of the indicated crosses. Each grey dot represents a single litter analyzed. Numbers of litters analyzed are 22, 4, 5 and 5 for Het × Het, WT × KO(M), KO(F) × WT and KO × KO mating, respectively. ** $P = 0.0003$, * $P = 0.03$; two-tailed Student's *t* test. Center line and error bars indicate mean and s.d., respectively. **f**, An example of a Z-KO × Z-KO litter with live pups (analyzed in part **e**). **g**, Scatter plot comparing the gene expression levels of *Dux* MZ-KO and WT embryos at late 1-cell stage (~12 hpi). Two RNA-seq replicates were generated for differential gene expression analyses. FC > 2, FDR < 0.05, RPKM > 1. **h**, Genome browser view of RNA-seq signal at the *Dux* cluster in WT and MZ-KO late 1-cell embryos. RNA-seq tracks of oocyte and 1-cell embryos were obtained from ref. [23]. Only uniquely aligned reads were used to generate the RNA-seq tracks.

As expected, the level of *Dux* transcript in the testis in Het mice is about half that in WT mice and is undetectable in the testis samples from KO mice (Fig. 1d). This result is consistent with the previous report that the *Dux* cluster on chromosome 10 is the only *Dux* locus in mice[17]. Collectively, these data indicate that loss of zygotic DUX does not arrest mouse development.

To determine whether the truncated DUX variant Gm4981 in oocytes might compensate for the deficiency of DUX in Z-KO embryos, we assessed the development of *Dux* Z-KO × Z-KO offspring, which should lack both zygotic DUX and maternal Gm4981. *Dux* MZ-KO embryos did not show impaired pre-implantation development (Supplementary Fig. 3) and also survived to adulthood without obvious abnormalities (Fig. 1e,f, and Supplementary Fig. 2c). Consistent with the reduced frequency of Z-KO pups in Het × Het crosses, the litter sizes of Z-KO × Z-KO mating pairs are also significantly smaller than those of controls (4.0 ± 1.2 versus 7.6 ± 1.8, $P = 0.0003$). Interestingly, *Dux* Z-KO females showed slightly reduced litter size (5.6 ± 1.5, $P = 0.03$)

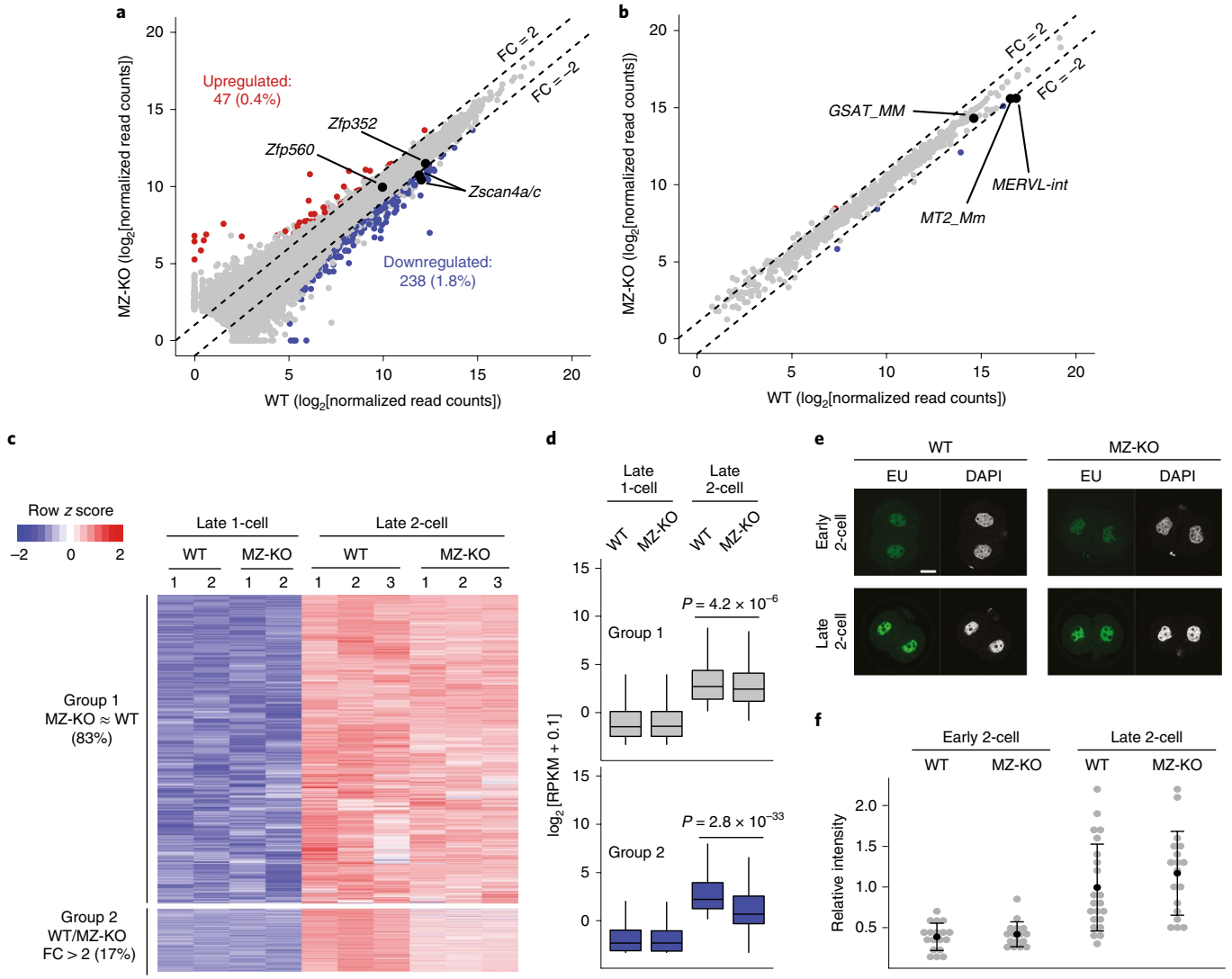

Fig. 2 | Loss of DUX causes minor defects in ZGA. a,**b**, Scatter plots comparing expression levels of the genes (**a**) and repeats (**b**) of late 2-cell *Dux* MZ-KO and WT embryos (~30 hpi). Three RNA-seq replicates were generated for differential gene expression analyses. In **a**, FC > 2, FDR < 0.05, RPKM > 1. In **b**, FC > 2, FDR < 0.05. **c**, Heatmap illustrating the expression levels of 2,906 major ZGA genes at late 1-cell and late 2-cell stages of *Dux* WT and MZ-KO embryos. The ZGA genes were defined using 2-cell/1-cell FC > 5, 2-cell RPKM > 1, FDR < 0.05. Group 1 represents genes that showed similar expression (FC < 2) between WT and MZ-KO 2-cell embryos, while Group 2 represents genes that showed decreased expression (FC > 2 and FDR < 1) in MZ-KO 2-cell embryos. **d**, Box plots illustrating the expression levels of Group 1 ($n = 2,413$) and 2 ($n = 493$) genes in part **c**. Mann–Whitney–Wilcoxon two-sided test was used to calculate the *P* values between WT and MZ-KO data. The middle lines in the boxes represent medians. Box hinges indicate the twenty-fifth and seventy-fifth percentiles, and the whiskers indicate the hinge ± 1.5× interquartile range. **e**, The EU staining assay showing the global transcriptional activity in early (~22 hpi) and late (~30 hpi) 2-cell embryos. Scale bar, 20 µm. **f**, Quantification of the EU signal intensity shown in part **e**. The average signal intensity of late 2-cell was set as 1.0. Each grey dot represents a single embryo analyzed. Center dot and error bars indicate mean and s.d., respectively. Two-tailed Student's *t* test was used to compare the signal intensity between WT and MZ-KO (1-cell, $P = 0.58$; 2-cell, $P = 0.29$). The total number of embryos analyzed were 18, 23, 16 and 18 for WT (~22 and ~30 hpi) and MZ-KO (~22 and ~30 hpi), respectively.

when their fertility was tested using WT B6 male mice. The reduction in litter size of *Dux* Z-KO female progeny should occur after implantation as both Z-KO female ovulation and pre-implantation development of MZ-KO embryos appear normal compared to control WT or Het mice (Supplementary Fig. 3). Nevertheless, the fact that both *Dux* Z-KO and MZ-KO can develop to adulthood indicates that DUX and its truncated variant Gm4981 are not essential for mouse development.

As lack of DUX does not arrest mouse development, DUX is unlikely to have a major role in ZGA. To investigate this, we generated late 1-cell and late 2-cell *Dux* MZ-KO embryos by fertilizing F_2 Z-KO oocytes with F_2 Z-KO sperm, then performed RNA-sequencing (RNA-seq). Embryos that were generated by fertilizing F_2 WT oocytes with F_2 WT sperm were used as controls. After confirming data reproducibility (Supplementary Fig. 4), we performed comparative analyses of late 1-cell RNA-seq data sets that revealed that, out of the 10,554 detectable genes (reads per kilobase of transcript per million mapped reads (RPKM) > 1 in either WT or KO), only 50 (0.47%) and 28 (0.26%) were significantly up- and downregulated, respectively, in *Dux* MZ-KO embryos (fold change (FC) > 2 and false discovery rate (FDR) < 0.05) (Fig. 1g and Supplementary Table 1), suggesting that DUX and Gm4981 deficiency has little effect on late 1-cell gene expression. Although it is not feasible to assess the expression level of each *Dux* repeat, owing to the assembly gap at the *Dux* cluster, we note that the annotated *Dux* and the other four known genes (that is, *AW822073/Duxf1*,

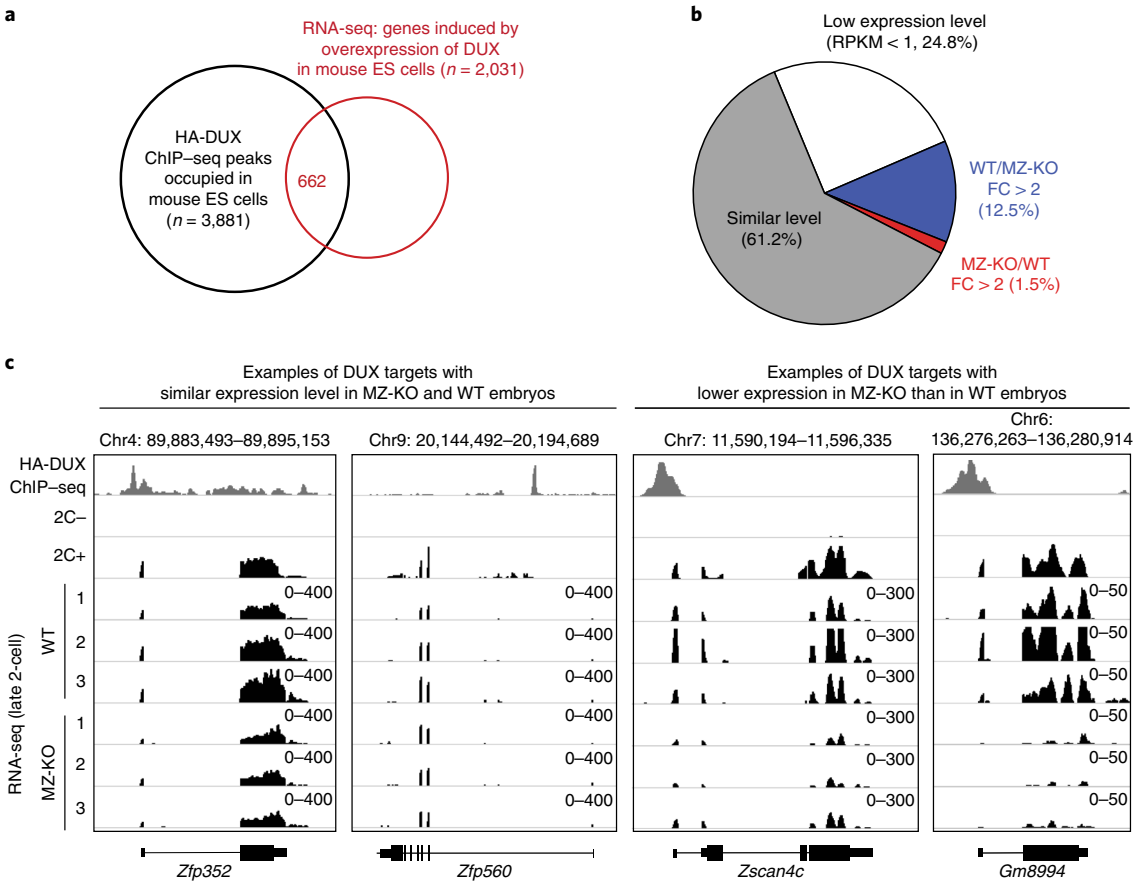

Fig. 3 | The majority of DUX targets identified in 2C-like cells are activated normally in *Dux* MZ-KO 2-cell embryos. a, Identification of known DUX targets by overlapping the genes associated with HA-DUX ChIP–seq peaks and those induced by DUX overexpression in mouse ES cells. Data sets of both HA-DUX ChIP–seq and RNA-seq of 2C-like cells were obtained from ref. [1]. **b**, Expression level changes of known DUX targets in MZ-KO 2-cell embryos. Known DUX targets were defined as genes that are associated with HA-DUX ChIP–seq peaks and are upregulated (FC > 2 and FDR < 0.05) in 2C-like cells. The average of three RNA-seq replicates of WT and MZ-KO late 2-cell embryos were used for the analyses. **c**, Genome browser views illustrating the RNA levels of known DUX targets in 2C-like cells and 2-cell embryos. HA-DUX ChIP–seq and RNA-seq tracks of non-2C (2C-) and 2C-like (2C+) cells were obtained from ref. [1]. Only uniquely aligned reads were used to generate the ChIP–seq and RNA-seq tracks.

Gm10807/Duxf2, *Gm19459/Duxf5*, *Gm4981/Duxf4*) were depleted in the MZ-KO 1-cell embryos (Fig. 1g,h), confirming complete KO of the *Dux* cluster in our mouse lines.

We next analyzed late 2-cell WT and MZ-KO RNA-seq data sets and identified that, out of the 12,960 detectable genes (RPKM > 1 in either WT or MZ-KO), 47 (0.4%) and 238 (1.8%) were significantly up- and downregulated, respectively, in *Dux* MZ-KO embryos (FC > 2 and FDR < 0.05) (Fig. 2a and Supplementary Table 2). Consistent with the few gene expression changes, most repeats also showed comparable expression levels between the two groups (Fig. 2b). To define to what extent DUX contributes to major ZGA, we identified 2,906 major ZGA genes by comparing the transcriptome of WT late 2-cell to late 1-cell embryos (2-cell/1-cell FC > 5, RPKM in WT 2-cell > 1, FDR < 0.05) (Fig. 2c and Supplementary Table 3). Even using a relaxed cutoff to define differential gene expression (FC > 2 and FDR < 1), only 493 (16.9%) ZGA genes showed decreased expression in *Dux* MZ-KO 2-cell embryos (Fig. 2c and Supplementary Table 3). Importantly, these affected genes and repeats still showed higher expression in *Dux* MZ-KO 2-cells than in WT 1-cell embryos (Fig. 2d and Supplementary Table 3), indicating that ZGA of these genes still takes place in *Dux* KO embryos although to a lesser extent. Consistent with this, global transcription levels are comparable between WT and MZ-KO early and late 2-cell embryos as revealed by 5-ethynyl uridine (EU) incorporation assay (Fig. 2e,f). Nevertheless, activation of a small group of ZGA genes was mildly impaired, which may account for the reduced frequency of *Dux* Z-KO mice from F_1 Het × F_1 Het mating pairs and smaller litter sizes of Z-KO × Z-KO mating pairs (Fig. 1b,e). Taken together, these data support that DUX only has a minor role in mouse ZGA.

Owing to the subtle effect of loss of DUX on 2-cell transcriptome, we reasoned that most DUX targets identified in 2C-like cells should be normally activated in *Dux* MZ-KO embryos. Indeed, out of the 662 genes that are activated by exogeneous DUX and also associated with hemagglutinin-tagged DUX (HA-DUX) chromatin immunoprecipitation followed by sequencing (ChIP–seq) peaks[1], only 12.5% showed more than a twofold decrease in *Dux* MZ-KO embryos, while the majority (61.2%) exhibited comparable expression levels in WT and MZ-KO 2-cell embryos (Fig. 3, Supplementary Fig. 5 and Supplementary Table 4). This suggests either that DUX targets that have been identified in 2C-like cells are not targeted by DUX in 2-cell embryos or that factors other than DUX can activate them in 2-cell embryos.

Overall, our results demonstrate that mouse ZGA genes, including many exogenous DUX targets identified in mouse ES cells, can be activated in *Dux* MZ-KO embryos and therefore that loss of DUX does not arrest mouse development. It is possible that other transcription factors and/or chromatin remodelers have a redundant role in 2-cell embryos for successful ZGA. Our results in mice

seem to be in direct contrast to observations in ES cells, in which DUX is essential for the entry of ES cells into the 2C-like state[1,2]. Therefore, despite the simplicity of the 2C-like state, caution should be taken in using the ES cell system to study the totipotent state as there are fundamental differences between the in vitro 2C-like cell state and 2-cell embryos.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of code and data availability and associated accession codes are available at https://doi.org/10.1038/s41588-019-0418-7.

Received: 25 October 2018; Accepted: 11 April 2019;
Published online: 27 May 2019

References

1. Hendrickson, P. G. et al. Conserved roles of mouse DUX and human DUX4 in activating cleavage-stage genes and MERVL/HERVL retrotransposons. *Nat. Genet.* **49**, 925–934 (2017).
2. De Iaco, A. et al. DUX-family transcription factors regulate zygotic genome activation in placental mammals. *Nat. Genet.* **49**, 941–945 (2017).
3. Hamatani, T., Carter, M. G., Sharov, A. A. & Ko, M. S. Dynamics of global gene expression changes during mouse preimplantation development. *Dev. Cell* **6**, 117–131 (2004).
4. Wang, Q. T. et al. A genome-wide study of gene activity reveals developmental signaling pathways in the preimplantation mouse embryo. *Dev. Cell* **6**, 133–144 (2004).
5. Zeng, F., Baldwin, D. A. & Schultz, R. M. Transcript profiling during preimplantation mouse development. *Dev. Biol.* **272**, 483–496 (2004).
6. Svoboda, P. et al. RNAi and expression of retrotransposons MuERV-L and IAP in preimplantation mouse embryos. *Dev. Biol.* **269**, 276–285 (2004).
7. Macfarlan, T. S. et al. Embryonic stem cell potency fluctuates with endogenous retrovirus activity. *Nature* **487**, 57–63 (2012).
8. Zalzman, M. et al. Zscan4 regulates telomere elongation and genomic stability in ES cells. *Nature* **464**, 858–863 (2010).
9. Lu, F. & Zhang, Y. Cell totipotency: molecular features, induction, and maintenance. *Natl Sci. Rev.* **2**, 217–225 (2015).
10. Percharde, M. et al. A LINE1-nucleolin partnership regulates early development and ESC identity. *Cell* **174**, 391–405 e19 (2018).
11. Ishiuchi, T. et al. Early embryonic-like cells are induced by downregulating replication-dependent chromatin assembly. *Nat. Struct. Mol. Biol.* **22**, 662–671 (2015).
12. Rodriguez-Terrones, D. et al. A molecular roadmap for the emergence of early-embryonic-like cells in culture. *Nat. Genet.* **50**, 106–119 (2018).
13. Eckersley-Maslin, M. A. et al. MERVL/Zscan4 network activation results in transient genome-wide DNA demethylation of mESCs. *Cell Rep.* **17**, 179–192 (2016).
14. Snider, L. et al. Facioscapulohumeral dystrophy: incomplete suppression of a retrotransposed gene. *PLoS Genet.* **6**, e1001181 (2010).
15. Geng, L. N. et al. DUX4 activates germline genes, retroelements, and immune mediators: implications for facioscapulohumeral dystrophy. *Dev. Cell* **22**, 38–51 (2012).
16. Whiddon, J. L., Langford, A. T., Wong, C. J., Zhong, J. W. & Tapscott, S. J. Conservation and innovation in the DUX4-family gene network. *Nat. Genet.* **49**, 935–940 (2017).
17. Clapp, J. et al. Evolutionary conservation of a coding function for D4Z4, the tandem DNA repeat mutated in facioscapulohumeral muscular dystrophy. *Am. J. Hum. Genet.* **81**, 264–279 (2007).
18. Fujii, W., Kawasaki, K., Sugiura, K. & Naito, K. Efficient generation of large-scale genome-modified mice using gRNA and CAS9 endonuclease. *Nucleic Acids Res.* **41**, e187 (2013).
19. Han, J. et al. Efficient in vivo deletion of a large imprinted lncRNA by CRISPR/Cas9. *RNA Biol.* **11**, 829–835 (2014).
20. Inoue, K. et al. The rodent-specific microRNA cluster within the Sfmbt2 gene is imprinted and essential for placental development. *Cell Rep.* **19**, 949–956 (2017).
21. Wang, L. et al. Large genomic fragment deletion and functional gene cassette knock-in via Cas9 protein mediated genome editing in one-cell rodent embryos. *Sci. Rep.* **5**, 17517 (2015).
22. Zhang, L. et al. Large genomic fragment deletions and insertions in mouse using CRISPR/Cas9. *PLoS ONE* **10**, e0120396 (2015).
23. Wu, J. et al. The landscape of accessible chromatin in mammalian preimplantation embryos. *Nature* **534**, 652–657 (2016).

Acknowledgements

We would like to thank A. Inoue for training Z.C. to manipulate mouse embryos and for advice on generating the *Dux* KO mouse lines. We acknowledge X. Fu, C. Zhang, X. Wu and W. Zhang for helpful discussion. We thank N. Djekidel for his advice on bioinformatic analyses. This project was supported by the US National Institutes of Health (NIH) (R01HD092465) and Howard Hughes Medical Institute. Y.Z. is an Investigator of the Howard Hughes Medical Institute.

Author contributions

Z.C. and Y.Z. conceived the project. Z.C. designed and performed experiments. Z.C. analyzed sequencing datasets. Z.C. and Y.Z. interpreted the data and wrote the manuscript.

Competing interests

The authors declare no competing financial interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41588-019-0418-7.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to Y.Z.

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

Methods

Generation of *Dux* KO mice. All animal experiments were performed in accordance with the protocols of the Institutional Animal Care and Use Committee at Harvard Medical School. For superovulation, B6D2F1 (BDF1) female mice (6–8 weeks old) (Jackson Laboratory, 100006) were injected interperitoneally with 7.5 IU of pregnant mare serum gonadotropin (PMSG, Millipore) on day 1 and human chorionic gonadotropin (hCG, Millipore) on day 3 (44–48 h after PMSG injection). For in vitro fertilization (IVF), the oocytes collected 12–16 h after hCG injection were inseminated with the activated spermatozoa collected from the caudal epididymis of BDF1 males (9–10 weeks old) in human tubal fluid (HTF) medium supplemented with 10 mg ml^{-1} bovine serum albumin (BSA, Sigma). The spermatozoa were capacitated by pre-incubation in HTF medium for 1 h. At 2 h post IVF (hpi), *Cas9* mRNA (100 ng µl^{-1}) and sgRNA (50 ng µl^{-1} each) were injected into cytoplasm of fertilized eggs using a Piezo impact-driven micromanipulator (Primer Tech). Following injection, zygotes were cultured in HTF medium for another 4 h and then cultured in KSOM (Millipore) at 37 °C under 5% CO_2 with air. At approximately 24 h hpi, 2-cell embryos were transferred into oviducts of surrogate ICR strain mothers. The synthesis of *Cas9* mRNA and sgRNA was carried out as described previously[24]. The sgRNA sequences were the same as reported previously[2].

To genotype blastocysts, each embryo collected at 120 hpi was lysed in 8 µl lysis buffer (50 mM Tris-HCl (pH 8.0), 0.5% Triton, 400 µg ml^{-1} Proteinase K) (Sigma) at 60 °C for 1 h. Following heat inactivation at 90 °C for 5 min, 2 µl of lysis buffer containing genomic DNA was used as template for nested PCR. The primers used for genotyping are included in Supplementary Table 6 (WT allele 268 bp and KO allele ~320 bp). For both rounds of PCR, the following program was used: initial denaturation, 5 min at 95 °C, 30 cycles of 30 s at 95 °C, 30 s at 60 °C and 30 s at 72 °C; final extension, 5 min at 72 °C.

To genotype colonies, a mouse tail tip was lysed in the same lysis buffer (70 µl) at 60 °C overnight and the supernatants were used as template for PCR (only inner primers were used, WT allele 268 bp, KO allele 322 bp and 318 bp for lines 423 and 426, respectively).

RNA-seq. For embryos collected for RNA-seq (that is, *Dux* $F_2 \times F_2$), IVF was performed as described above, except that the micro-injection steps were omitted. Late 1-cell and late 2-cell were collected at approximately 12 and 30 hpi, respectively. For each biological replicate, 11–13 embryos were pooled for RNA-seq analyses. Specifically, the embryos were incubated briefly in Acidic Tyrode's Solution (Millipore) to remove zona pellucida and then washed three times in 0.2% BSA in PBS prior to library construction.

RNA-seq libraries were prepared as described previously[25]. In brief, SMARTer Ultra Low Input RNA cDNA preparation kit (Clontech, 643890) was used for reverse transcription and cDNA amplification (11 cycles). cDNA was then fragmented, adaptor-ligated and amplified using a Nextera XT DNA Library Preparation Kit (Illumina) according to the manufacturer's instructions. Single-end 100-bp sequencing was performed on a HiSeq 2500 sequencer (Illumina). A summary of the generated data sets can be found in Supplementary Table 5.

RNA-seq analyses. RNA-seq reads were first trimmed to remove adaptor sequences and low-quality bases using Trimgalore (version 0.4.5). Reads (> 35 bp) were aligned to mm9 reference genome using HISAT2 (version 2.1.0)[26] with default parameters and RPKM values for each gene were computed using Cufflinks (version 2.2.1)[27]. For differential gene and repeat expression analyses, TEtranscripts (version 1.5.1)[28] was used to generate read counts for genes (uniquely aligned reads only) and repeats (including both unique- and multi-aligned reads), and DESeq (ref. [29]) was used to compute the FDR using the 'nbinomTest' function.

For the comparative analyses of WT and MZ-KO late 1-cell and 2-cell embryos, only genes with both FC > 2 and FDR < 0.05 were considered as differentially expressed (Supplementary Tables 1,2). For the determination of whether known DUX targets or major ZGA genes were affected in KO embryos, a more relaxed criterion that only considers fold change (FC > 2 and FDR < 1) was used (Supplementary Tables 3,4).

Detection of RNA synthesis by EU incorporation. Early (~21 hpi) and late (~29 hpi) 2-cell embryos were incubated in KSOM supplemented with 500 µM EU (Invitrogen) for 1 h prior to fixation in 3.7% paraformaldehyde (Sigma). Following permeabilization in PBS containing 0.5% Triton X-100 (Sigma), embryos were stained using a Click-iT RNA Alexa Fluor 488 Imaging Kit (Invitrogen). Fluorescence was detected using a laser scanning confocal microscope (Zeiss LSM800) and the images were acquired using Axiovision software (Carl Zeiss). Signal intensity of nuclei and cytoplasm of two blastomeres were acquired and the cytoplasmic signal was subtracted from the nuclei signal as background. The averaged signal intensity of the WT late 2-cell (~30 hpi) was set as 1.0.

RNA isolation, reverse transcription and quantitative PCR. Total RNA was isolated from testis of adult mice (9–12 weeks old) using Trizol Reagent (Invitrogen) according to the manufacturer's instructions. Following RQ1 DNase (Promega) treatment, RNA was used as template to synthesize cDNA with the use of SuperScriptIII First-Strand Synthesis System (Invitrogen). To ensure no genomic DNA contamination, a minus reverse-transcriptase control were also included. A SYBR green gene expression assay (Invitrogen) was used to determine *Dux* transcript abundance in a ViiA 7 Real-Time PCR System (ThermoFisher Scientific). The threshold cycles were normalized to the housekeeping gene *Gapdh* and the relative abundance in each sample was calculated using the comparative C_T method. The primers used are included in Supplementary Table 6.

Statistical analyses and data visualization. All statistical analyses were performed with R (http://www.r-project.org/). Pearson's *r* co-efficient was computed using the 'cor' function. Figure 2c was generated using the R function 'heatmap.2'. Smoothed scatter plots (Supplementary Fig. 4) were generated with the R function 'smoothScatter' and all other plots were generated using the ggplot2 package. The RNA-seq and ChIP–seq bigwig tracks were generated with uniquely aligned reads using deeptools (version 3.0.2)[30] with the following parameters '--skipNonCoveredRegions --binSize 10 --scaleFactor 1/DESeq's sizeFactor'. The bigwig tracks were visualized in the Integrative Genomic Viewer genome browser[31].

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

Data availability

All RNA-seq data sets that were generated in this study have been deposited in the Gene Expression Omnibus under accession number GSE121746. Oocyte and 1-cell RNA-seq data were obtained from a previous publication[23]. HA-DUX ChIP–seq data and *Dux* overexpression RNA-seq data in mouse ES cells were downloaded from a previous report[1].

References

24. Wang, H. et al. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* **153**, 910–918 (2013).
25. Inoue, A., Chen, Z., Yin, Q. & Zhang, Y. Maternal *Eed* knockout causes loss of H3K27me3 imprinting and random X inactivation in the extraembryonic cells. *Genes Dev.* **32**, 1525–1536 (2018).
26. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* **12**, 357–360 (2015).
27. Trapnell, C. et al. Transcript assembly and quantification by RNA-seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28**, 511–515 (2010).
28. Jin, Y., Tam, O. H., Paniagua, E. & Hammell, M. TEtranscripts: a package for including transposable elements in differential expression analysis of RNA-seq datasets. *Bioinformatics* **31**, 3593–3599 (2015).
29. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010).
30. Ramirez, F., Dundar, F., Diehl, S., Gruning, B. A. & Manke, T. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.* **42**, W187–W191 (2014).
31. Robinson, J. T. et al. Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).

nature research

Corresponding author(s): Yi Zhang

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	RNA-sequencing reads were generated at Illumina HiSeq 2500 platform (illumina). Signal intensity of EU-staining were acquired using Axiovision software (Carl Zeiss). Quantitative PCR was performed in a ViiA 7 Real-Time PCR System (ThermoFisher Scientific).
Data analysis	RNA-seq analyses: RNA-seq reads were first trimmed to remove adaptor sequences and low-quality bases using Trimgalore (version 0.4.5). Reads (>35bp) were aligned to mm9 reference genome using HISAT2 (version 2.1.0) with default parameters and RPKM values for each gene were computed using Cufflinks (version 2.2.1). For differential gene/repeats expression analyses, TEtranscripts (version 1.5.1) was used to generate read counts for genes (uniquely aligned reads only) and repeats (including both unique- and multi-aligned reads) and DESeq package was used to compute false discovery rate using the 'nbinomTest' function. Pearson's r co-efficient was computed using 'cor' function. Heatmap was generated using the R function 'heatmap.2'. Smoothed scatter plots were generated with the R function 'smoothScatter' and all other plots such as dot plots, bar plots, and box plots were generated using ggplot2 package. The RNA-seq bigwig tracks were generated with uniquely aligned reads using deeptools (version 3.0.2). The bigwig tracks were visualized in the Integrative Genomic Viewer genome (IGV) browser. Other analyses: Chi-square good of fitness test was performed for assessing whether mice with different genotypes fit the expected Mendelian ratio. Student t-test (two-sided) was used to assess whether the litter sizes, oocytes number, 1-cell embryos number, and EU signal intensity are significantly different between groups.

Fisher's exact test (two-sided) was used to compare the blastocyst stage embryos between groups.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

--All RNA-seq data sets generated in this study have been deposited in the Gene Expression Omnibus under accession number GSE121746.
--Figure 1, 2, and 3, and Figure S4 and S5 have associated raw data.
--There are no restrictions for data availability

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample sizes were determined without statistical measures, but based on widely accepcted sample sizes in relevant publications within this field of research. See Figures legends for each experiment.
Data exclusions	No data were excluded for analyses.
Replication	All attempts of replication were successful. RNA-seq experiments included two replicates for 1-cell embryos and three replicates for 2-cell embryos. The reproducibility between replicates were assessed by Pearson correlation.
Randomization	Selection of mice for collection of embryos and RNA-seq analyses within either WT or KO group was random.
Blinding	Blinding is not practicable in our study as mice need to be assigned to WT or KO groups based on their genotypes.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☒	☐ Unique biological materials
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☒	☐ Human research participants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	Mouse, B6/DBA background, 6-10 weeks, females/males
Wild animals	This study did not involve wild animals

Field-collected samples	This study did not involve samples collected from the field.