



## RESEARCH ARTICLE SUMMARY

## DEVELOPMENT

# Multifaceted SOX2-chromatin interaction underpins pluripotency progression in early embryos

Lijia Li<sup>†</sup>, Fangnong Lai<sup>†</sup>, Xiaoyu Hu<sup>†</sup>, Bofeng Liu<sup>†</sup>, Xukun Lu, Zili Lin, Ling Liu, Yunlong Xiang, Tristan Frum, Michael A. Halbisen, Fengling Chen, Qiang Fan, Amy Ralston, Wei Xie\*

**INTRODUCTION:** During early mammalian development, a totipotent embryo undergoes the first cell fate decision to form a blastocyst that includes inner cell mass (ICM) and trophectoderm. ICM gives rise to epiblast, the origin of future embryonic lineages, and primitive endoderm. Trophectoderm subsequently differentiates to placenta. Pluripotency—the ability for a cell to give rise to all primary embryonic lineages—emerges within ICM and transits through several states: naive, formative, and primed pluripotency. Pioneer transcription factors (TFs), such as OCT4 and SOX2, which can bind and open closed chromatin, are crucial for pluripotency regulation. However, their regulatory circuitry is largely inferred from cultured cells. How the master TFs govern pluripotency progression in vivo remains challenging to study, largely because of the limited research materials from mammalian embryos.

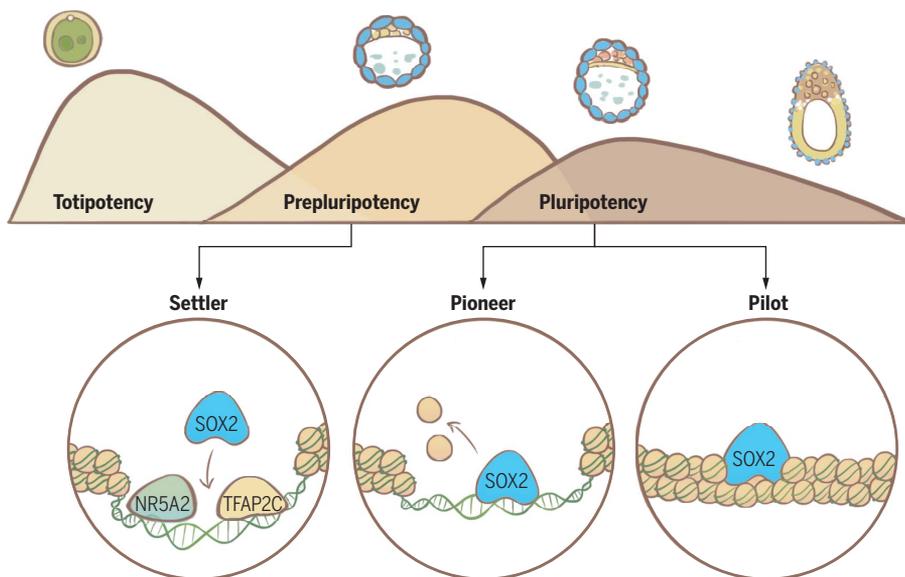
**RATIONALE:** To study TF-gene interactions in early mammalian embryos, we applied CUT&RUN to capture the chromatin binding of SOX2, an early ICM marker, from embryonic day 3.5 (E3.5) to E7.5, which covers the entire progression of pluripotency in mouse embryos. Combined with RNA sequencing (RNA-seq) and assay for transposase-accessible chromatin sequencing (ATAC-seq) in *Sox2* knockout embryos and degron-tagged embryonic stem cells (ESCs), we investigated the roles of SOX2 in gene regulation and enhancer opening during pluripotency progression in vivo and compared this with in vitro conditions.

**RESULTS:** Our data revealed that SOX2 in E3.5 ICM has a regulatory circuitry distinct from that in all other pluripotent states and ESCs. Two massive relocalizations of SOX2-chromatin binding occurred when cells en-

tered naive (E4.5 epiblast) and formative (E5.5 epiblast) pluripotent states, followed by a less dynamic transition to E7.5 ectoderm. These changes were accompanied by a critical role of SOX2 in regulating the ICM-trophectoderm transcription program and the following naive-to-formative pluripotency conversion.

Furthermore, we discovered that SOX2 manifests much more diverse binding modes at enhancers, which include “settler binding,” “pioneer binding,” and “pilot binding,” beyond a simple “pioneer factor.” SOX2 exhibits settler binding in E3.5 ICM, where SOX2 binds pre-accessible enhancers, and its loss does not substantially affect chromatin opening. These pre-accessible enhancers are in part opened by early-stage expressing TFs TFAP2C and NR5A2. Notably, settler binding of SOX2 can still exert impacts on gene expression, especially at sites with strong SOX2 motifs. Pioneer binding of SOX2 becomes widespread in E4.5 epiblast and naive ESCs (2i ESCs), where SOX2 is required for opening naive enhancers, and its binding sites enrich the OCT4-SOX2 motif. Finally, the pilot binding of SOX2 at many formative enhancers in 2i ESCs is insufficient for enhancer opening but poises enhancers for faster opening upon the conversion to formative pluripotency.

**CONCLUSION:** In this work, we dissected the SOX2-governed pluripotency regulatory network in mouse early embryos. These data revealed highly dynamic regulatory circuitry during pluripotency progression in vivo, especially when cells enter naive and formative pluripotency. We also found multifaceted pioneer factor-enhancer interactions that underpin the transition of pluripotency states. Finally, these results also identified a distinct “prepluripotency” state in E3.5 ICM, between totipotency and pluripotency. Prepluripotency features the potency to give rise to both epiblast and primitive endoderm, coexpression of multilineage TFs, and a primitive pluripotency network that lacks interdependence of master pluripotency TFs and their pioneer binding. How these master TFs acquire chromatin opening ability to establish a pluripotency network when cells enter naive pluripotency warrants future investigations. Hence, these data bridge the knowledge gap between in vivo development and in vitro cultured stem cells and pave the way for future studies to understand pluripotency and cell fate decision. ■



**SOX2 regulatory circuitry during pluripotency progression reveals multifaceted master TF-enhancer interaction.** SOX2 regulatory circuitry was dissected during the pluripotency transition in mouse early embryos. The prepluripotency state in E3.5 ICM is proposed to bridge totipotency to pluripotency. Three SOX2-enhancer interaction modes are shown: settler, pioneer, and pilot binding.

The list of author affiliations is available in the full article online.

\*Corresponding author. Email: xieweil21@tsinghua.edu.cn

<sup>†</sup>These authors contributed equally to this work.

Cite this article as L. Li *et al.*, *Science* **382**, eadi5516 (2023).

DOI: 10.1126/science.adi5516

**S** READ THE FULL ARTICLE AT  
<https://doi.org/10.1126/science.adi5516>

## RESEARCH ARTICLE

## DEVELOPMENT

## Multifaceted SOX2-chromatin interaction underpins pluripotency progression in early embryos

Lijia Li<sup>1,2,†</sup>, Fangnong Lai<sup>1,2,†</sup>, Xiaoyu Hu<sup>1,2,†</sup>, Bofeng Liu<sup>1,2,†</sup>, Xukun Lu<sup>1,2</sup>, Zili Lin<sup>3</sup>, Ling Liu<sup>1,2</sup>, Yunlong Xiang<sup>4</sup>, Tristan Frum<sup>5,6</sup>, Michael A. Halbisen<sup>5</sup>, Fengling Chen<sup>1,2</sup>, Qiang Fan<sup>1,2</sup>, Amy Ralston<sup>5</sup>, Wei Xie<sup>1,2,\*</sup>

Pioneer transcription factors (TFs), such as OCT4 and SOX2, play crucial roles in pluripotency regulation. However, the master TF-governed pluripotency regulatory circuitry was largely inferred from cultured cells. In this work, we investigated SOX2 binding from embryonic day 3.5 (E3.5) to E7.5 in the mouse. In E3.5 inner cell mass (ICM), SOX2 regulates the ICM-trophectoderm program but is dispensable for opening global enhancers. Instead, SOX2 occupies preaccessible enhancers in part opened by early-stage expressing TFs TFAP2C and NR5A2. SOX2 then widely redistributes when cells adopt naive and formative pluripotency by opening enhancers or posing them for rapid future activation. Hence, multifaceted pioneer TF–enhancer interaction underpins pluripotency progression in embryos, including a distinctive state in E3.5 ICM that bridges totipotency and pluripotency.

The mammalian embryo undergoes several rounds of cell cleavage after fertilization to give rise to a blastocyst (1, 2). In the mouse, blastocyst is specified into trophectoderm (TE), which eventually contributes to placenta, and inner cell mass (ICM) which subsequently differentiates into primitive endoderm (PrE) and epiblast (1–3). During this process, cells with pluripotency emerge and transit through several states that could be captured *ex vivo*, including the naive, formative, and primed pluripotency (4). Naive pluripotency is represented by the pluripotency state exhibited by naive embryonic stem cells (2i ESCs) that resemble preimplantation epiblast. After implantation, the epiblast cells acutely transit to the formative pluripotency state that manifests competence for both primordial germ cell (PGC) and somatic fate induction (4). This transition can be recapitulated by naive embryonic stem cells differentiating into epiblast-like cells (EpiLCs) (5). Finally, epiblast stem cells (EpiSCs), considered to recapitulate primed pluripotency, could be derived from E5.5 to E8 epiblast or ectoderm (6). Unlike formative pluripotent

cells, EpiSCs have lost the competence to PGC induction (5).

Transcription factors (TFs) play central roles in development by means of precise spatiotemporal regulation of gene expression through action at cis-regulatory elements (7, 8). Key master TFs, such as NANOG, SOX2, and OCT4, are essential to the pluripotency network (9). Among them, SOX2 is the earliest and only pluripotent TF known to be restricted to the inner cells in the mouse morula (16-cell), the ICM progenitors (10). *Sox2* deficiency in mice leads to epiblast formation failure and embryonic lethality shortly after implantation (11). Because of the limited research materials from embryos, the regulatory circuitry governed by master pluripotency factors has largely been inferred from cultured cell models (9, 12, 13), with limited TF regulomes being investigated in mouse blastocysts (e.g., NANOG) (14, 15). Therefore, how these TFs regulate pluripotency and its transition in physiological conditions remains elusive. For example, *Sox2*-deficient embryos can still form blastocysts containing a morphologically normal ICM, which, however, cannot give rise to ESCs. Consistently, ESCs lose pluripotency when SOX2 is depleted (11, 16). Furthermore, OCT4 and SOX2 are considered to be “pioneer factors” that can open inaccessible chromatin and recruit other TFs (17, 18). How they potentially drive such pluripotency transition *in vivo* remains unknown. In this work, we applied CUT&RUN to capture the SOX2-chromatin binding from E3.5 ICM to E7.5 ectoderm, which covers the entire progression of pluripotency. Our results revealed that SOX2 engages enhancers through stage- and context-dependent action modes to regulate development programs. Furthermore, these

data reveal a distinctive state and regulatory circuitry in E3.5 ICM that connects totipotency and pluripotency.

## Genome-wide mapping of SOX2-chromatin binding in mouse early embryos during pluripotency progression

By identifying a SOX2 antibody to allow CUT&RUN analyses using as few as 200 cells (fig. S1, A to C), we examined SOX2-chromatin binding at stages when it is expressed, including E3.5 ICM from early and middle blastocyst, epiblast from E4.5 preimplantation embryos, E5.5 and E6.5 postimplantation embryos, and ectoderm dissected from E7.5 embryos (fig. S2, A and B; Fig. 1, A and B; and materials and methods), which represent states from the onset to the exit of pluripotency. For comparison, we also included 2i ESCs and EpiLCs, which resemble E4.5 epiblast and E5.5 epiblast, respectively (19).

We conducted several analyses to validate these datasets. (i) The CUT&RUN data were reproduced between replicates (Fig. 1C and fig. S2C). (ii) Globally, 82.8 to 92.0% of SOX2 peaks occupied intergenic and intragenic regions away from promoters (fig. S2D), consistent with TFs predominantly binding to enhancers (8). For convenience, we refer to SOX2-bound distal regions as putative enhancers or enhancers hereafter. An examination of known SOX2 target genes *Pou5f1* and *Nanog* (12, 16, 20) revealed SOX2 binding at their known enhancers (Fig. 1D). (iii) We generated *Sox2* maternal and zygotic knockout (mzKO) embryos and found that SOX2 binding was substantially diminished in mutant E3.5 ICM (Fig. 1B and fig. S2E). (iv) Across all stages, SOX2-bound putative enhancers were enriched for SOX2 motif and exhibited distinct motif enrichment compared with assay for transposase-accessible chromatin sequencing (ATAC-seq) (21) (Fig. 1E), which suggests that SOX2 binding is not simply dependent on chromatin accessibility. (v) Finally, by acute degradation of SOX2 in 2i ESCs (22) (fig. S3, A to C), we observed that SOX2-targeted enhancers near ICM-specific genes *Upp1* and *Spic* exhibited SOX2-dependent transcriptional regulatory activity (fig. S3D). Taken together, these analyses suggest that we have identified bona fide SOX2 binding targets in early embryos.

## SOX2 binding in E3.5 ICM manifests a state that is distinct from previously defined pluripotent states and stem cells

Global analysis of SOX2 distal binding sites showed that SOX2 binding in E3.5 ICM and E4.5 epiblast was highly divergent from that at other developmental stages (Fig. 1, C, F, and G), as confirmed by correlation analyses (fig. S4A). SOX2 binding in E3.5 ICM was also distinct from that in 2i ESCs and EpiLCs. An analysis of gene expression and chromatin accessibility

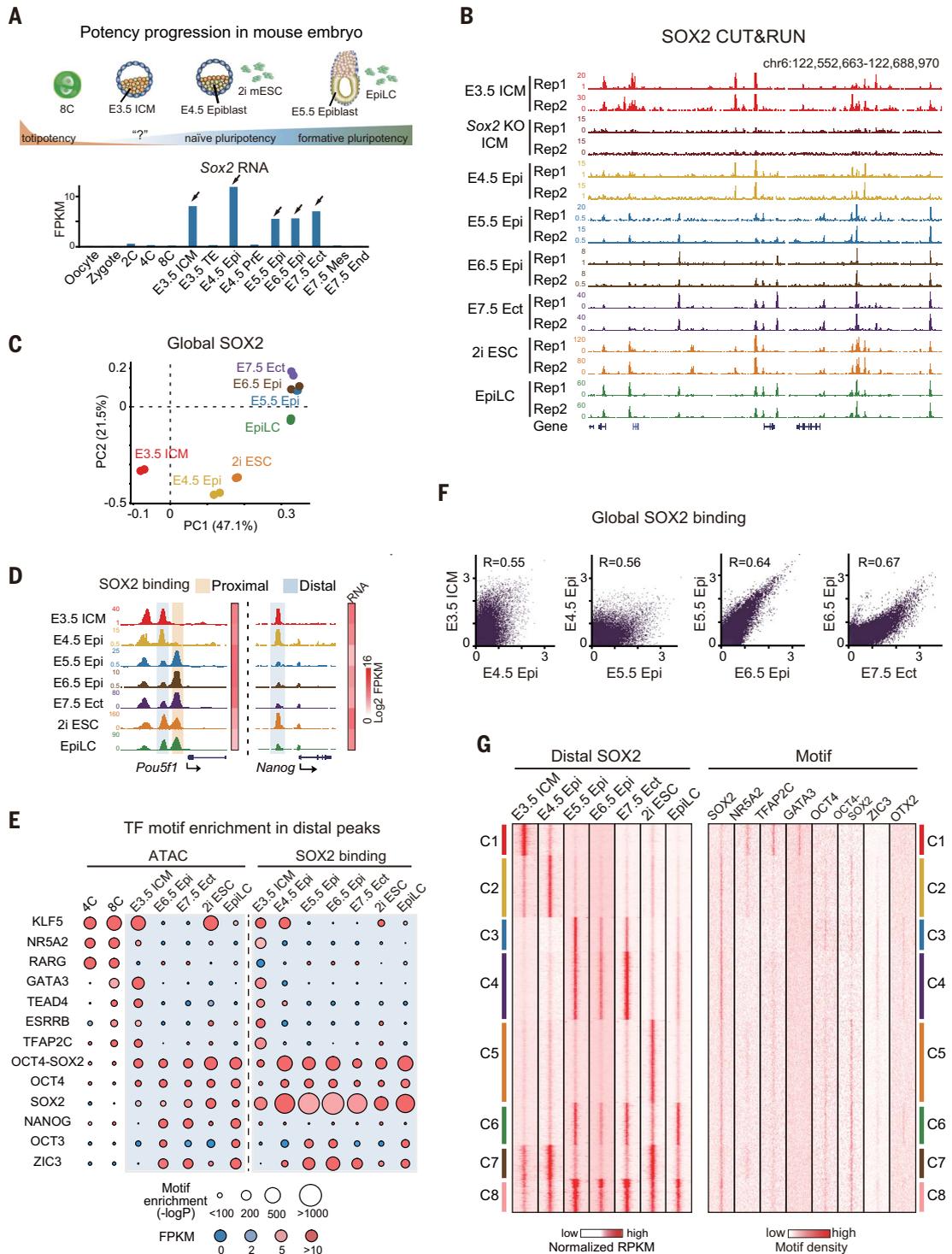
<sup>1</sup>Center for Stem Cell Biology and Regenerative Medicine, MOE Key Laboratory of Bioinformatics, School of Life Sciences, New Cornerstone Science Laboratory, Tsinghua University, Beijing 100084, China. <sup>2</sup>Tsinghua-Peking Center for Life Sciences, Beijing 100084, China. <sup>3</sup>College of Animal Science and Technology College, Beijing University of Agriculture, Beijing 102206, China. <sup>4</sup>Department of Cell Biology and Genetics, School of Basic Medical Sciences, Chongqing Medical University, Chongqing 400016, China. <sup>5</sup>Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, MI 48824, USA. <sup>6</sup>Department of Internal Medicine, Gastroenterology, University of Michigan Medical School, Ann Arbor, MI 48109, USA.

\*Corresponding author. Email: xiewei121@tsinghua.edu.cn

†These authors contributed equally to this work.

### Fig. 1. The dynamics of SOX2-chromatin binding during pluripotency progression in vivo.

**(A)** Developmental stages and related Sox2 expression are shown. Arrows indicate lineages where SOX2 binding was profiled. FPKM, fragments per kilobase of transcript per million mapped fragments. **(B)** UCSC browser view showing SOX2 CUT&RUN signals in E3.5 ICM; Sox2 mzkO ICM; E4.5, E5.5, and E6.5 epiblast (Epi); E7.5 ectoderm (Ect); 2i ESC; and EpiLC (two biological replicates). **(C)** Principal components analysis (PCA) of global SOX2 binding enrichment in mouse early lineages and stem cells. **(D)** UCSC browser views and heatmaps showing SOX2 enrichment and gene expression of representative genes, respectively. Proximal and distal binding sites are shaded. **(E)** TF motifs identified from distal ATAC-seq peaks and SOX2 binding peaks. Sizes of circles indicate levels of  $-\log P$  value. Expression levels of TFs are color coded. **(F)** Scatter plots showing global SOX2 binding correlation between consecutive stages. **(G)** Heatmaps showing the SOX2 binding signals and TF motif densities at stage-specific and shared distal SOX2 binding peaks. C1 to C8, cluster 1 to cluster 8; RPKM, reads per kilobase per million mapped reads.



(21, 23) independently supported the distinctive state of E3.5 ICM (fig. S4B). E3.5 SOX2 binding sites preferentially harbored motifs of NR5A2, GATA, and TFAP2C, which were also enriched in accessible chromatin at the earlier stage [8-cell (8C)] but not the OCT4 and the composite OCT4-SOX2 motifs (Fig. 1E). These data suggest a poised state of E3.5 ICM that co-expresses not only master pluripotency genes

(*Sox2*, *Oct4*, and *Nanog*) but also PrE genes (*Gata6*) and early-stage TFs (*Nr5a2* and *Tfap2c*) (3, 24, 25) (fig. S4C).

SOX2 then underwent substantial redistribution when cells from E3.5 ICM entered naive pluripotency (E4.5 epiblast and 2i ESC) and again when cells entered formative pluripotency (E5.5 epiblast and EpiLC) (Fig. 1, C, F, and G). By contrast, stages from E5.5 epiblast

onward (E6.5 epiblast and E7.5 ectoderm) were clustered closely, which suggests a more gradual formative-to-primed pluripotency transition. Because the pluripotency transition is a continuum, E5.5 epiblast may represent the formative state, whereas E6.5 epiblast may be transitional between formative and primed pluripotency (4). Given their similarity, we mainly focused on E5.5 epiblast in subsequent

analyses. SOX2 binding showed a moderate transition from E6.5 epiblast to E7.5 ectoderm. A small number of binding sites were newly established in E7.5 ectoderm, accompanied by increased ATAC-seq and H3K27ac signals (fig. S4D). Only a small subset (8.8%) of SOX2 distal peaks were shared among all seven cell types (Fig. 1G, “C8”). Genes related to blastocyst and ICM formation were enriched near SOX2 binding sites in E4.5 epiblast (fig. S4E, “C2”). Neural tube patterning- and stem cell maintenance-associated genes were enriched near enhancers shared in E5.5 epiblast, ectoderm, and EpiLCs (fig. S4E, “C6”). These data demonstrate highly dynamic SOX2 binding during pluripotency progression.

### SOX2 is required for the ICM-TE lineage program in E3.5 ICM

To investigate whether SOX2 binding in E3.5 ICM is linked to gene expression, we compared SOX2 binding from E3.5 ICM with that from E4.5 epiblast. E3.5 ICM-specific SOX2 binding preferentially resided near E3.5 ICM-specific genes (e.g., *Gata6*), whereas E4.5 epiblast-specific SOX2 binding tended to occur near E4.5 epiblast-activated genes (e.g., *Left1*) (Fig. 2A, “RNA”). We then performed single-cell RNA sequencing (scRNA-seq) in wild-type (WT) and *Sox2* mzKO E3.5 ICM (Fig. 2, B and C). The coexpression of the epiblast gene *Nanog* and the PrE gene *Gata6* in most cells (fig. S5A) confirmed that the epiblast and PrE segregation has not started in E3.5 ICM (3). The mutant cells were globally separated from WT cells (Fig. 2B), with 510 genes up-regulated and 623 genes down-regulated in the mutant cells (fig. S5B). In total, 239 of the up-regulated genes were TE-enriched genes (such as *Gata3*, *Krt8*, *Eomes*, and *Id2*), whereas 99 of the down-regulated genes were ICM specific (Fig. 2D and fig. S5C). The latter included not only epiblast-enriched genes, such as *Spic*, *Fgf4*, *Utf1*, and *Upp1*, but also PrE genes, such as *Sox17* and *Pdgfra* (Fig. 2E). Notably, the expression of *Pou5f1* and *Nanog* was not significantly affected as previously reported (10) (Fig. 2C and fig. S5A). Therefore, although the blastocyst morphology appears unaffected (10, 11), the ICM-TE transcription program is severely impaired in *Sox2* mutant ICM.

The OCT4 and SOX2-OCT4 composite motif enrichment within SOX2 binding sites was low in E3.5 ICM (Fig. 1E). They became enriched in E4.5 epiblast (Fig. 1G), which indicates that the cooperativity of these two pluripotency factors may only become mature when cells enter naive pluripotency. Although altered ICM-TE expression was also reported in *Oct4* KO E3.5 ICM (26) (fig. S6A), the gene expression changes between *Oct4* KO and *Sox2* KO E3.5 ICM overall showed a weak correlation (correlation coefficient  $r = 0.15$ ) (fig. S6B). Only 96 down-regulated genes (e.g., ICM-specific

genes *Fgf4*, *Sppl1*, and *Spic*) and 95 up-regulated genes (e.g., TE-specific genes *Dppa1*, *Gata3*, and *Krt18*) were shared in *Sox2* and *Oct4* KO ICM (fig. S6C), which indicates that OCT4 and SOX2 have both shared and specific functions at this stage.

### SOX2 binds preaccessible chromatin and is dispensable for enhancer opening in E3.5 ICM

We next investigated how enhancers were affected in *Sox2* mzKO E3.5 ICM. However, ATAC-seq showed few global changes at the SOX2-bound sites (Fig. 3, A and B). A close examination revealed that 75.5% of SOX2-occupied sites in ICM—either E3.5 ICM-specific or E3.5 to E4.5 shared—were already accessible (“preaccessible,” or “preaccess”) in 8C embryos (Fig. 3, A and C, “*Fgf4*”) when *Sox2* expression was still undetectable (Fig. 1A and fig. S2A). SOX2 binding was diminished at preaccessible sites in *Sox2* KO ICM, which confirms that these are bona fide signals (fig. S6D). Although the remaining SOX2-occupied sites became accessible after the 8C stage (opened de novo after *Sox2* expression, or “de novo”) (Fig. 3, A and C), SOX2 was largely dispensable for accessibility of both preaccessible and de novo enhancers (Fig. 3, A and B). In fact, one-third of the de novo sites already showed deoxyribonuclease (DNase) I hypersensitive sites sequencing (DNase-seq) signals (27) in morula at the 16C stage (fig. S6E). By contrast, E4.5 epiblast-specific SOX2-bound enhancers required SOX2 for opening (Fig. 3C, “*Pdgfc*,” and Fig. 3D), which displayed stronger SOX2 motif enrichment compared with E3.5 ICM-specific enhancers (Fig. 3A). Hence, SOX2 preferentially resided at pre-accessible chromatin sites in E3.5 ICM. We referred to this TF binding mode, in which a TF binds enhancers that are preaccessible and its loss does not substantially affect chromatin opening, as “settler binding” (28). SOX2 then adopted an indispensable role in opening E4.5 epiblast-specific enhancers, which we referred to as “pioneer binding.” We note that settler and pioneer here refer to different TF binding modes, which may come from the same TF (e.g., SOX2), rather than different classes of TFs.

We investigated whether such settler binding of SOX2 is associated with gene expression. We found that genes down-regulated (but not the up-regulated genes) in *Sox2* KO E3.5 ICM preferentially resided near enhancers that were bound by SOX2 in both E3.5 ICM and E4.5 epiblast (Fig. 3B, “RNA,” “shared,” and “preaccess”). These enhancers harbored strong SOX2 motif enrichment similar to E4.5 epiblast-specific enhancers (Fig. 3A, “SOX2 motif”). By contrast, SOX2 settler binding at E3.5 ICM-specific enhancers, which exhibited much weaker SOX2 motif enrichment, appeared to have minimal impacts on near-

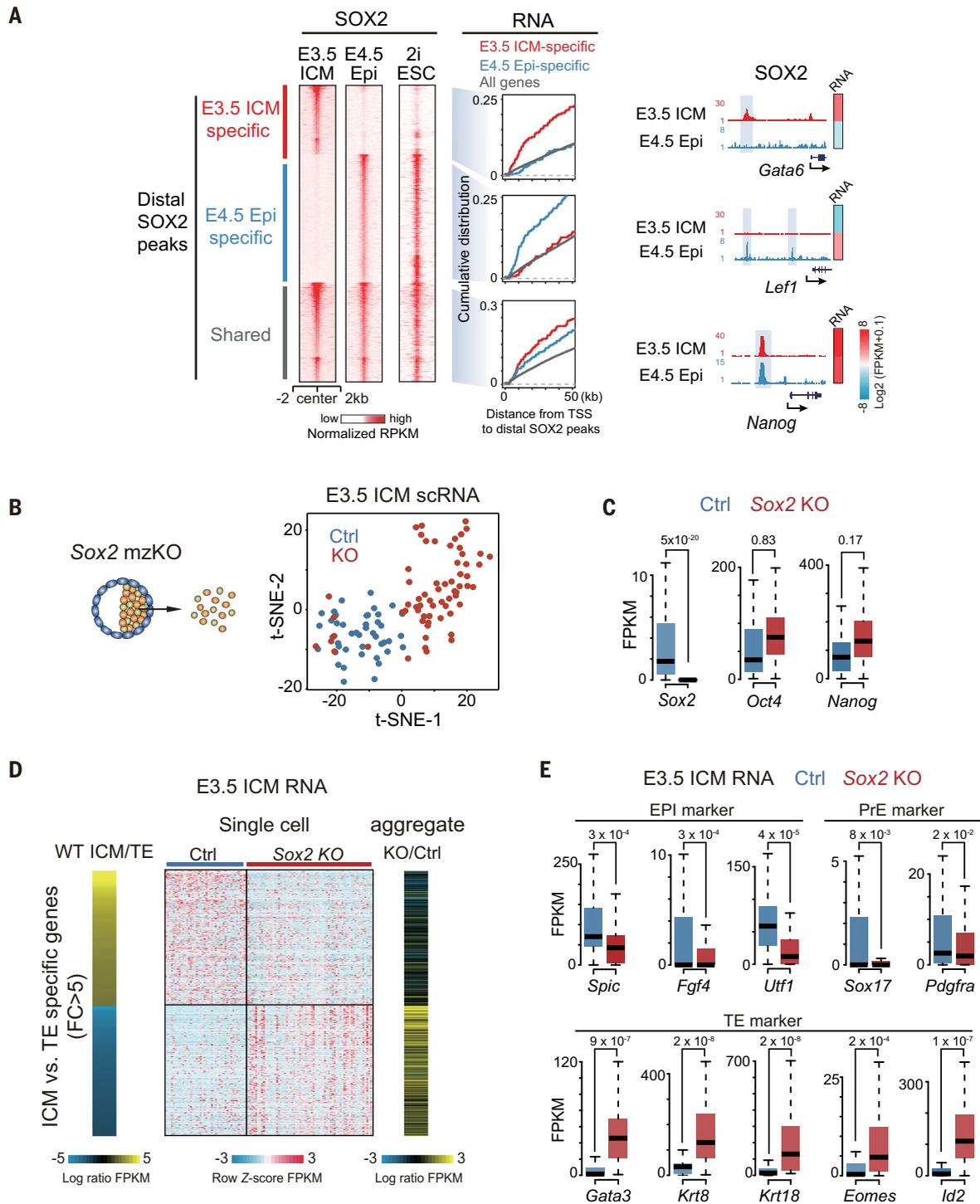
by genes (Fig. 3B, “E3.5 ICM specific” and “pre-access”). Therefore, settler binding of SOX2 at enhancers with strong SOX2 motifs is associated with gene activation.

### Early-stage expressing TFs are responsible for opening preaccessible SOX2 binding sites

We next investigated which TFs might be responsible for opening preaccessible SOX2 binding sites. Both E3.5 ICM-specific and E3.5 ICM-E4.5 epiblast shared SOX2-bound enhancers enriched for motifs of TFAP2C, NR5A2, and GATA (Fig. 4A). *Tfap2c* and *Nr5a2* show highest expression at the 4C to 8C stages (fig. S4C), and both are essential for embryogenesis (29–31). TFAP2C regulates cell polarity starting from the 2C stage before being restricted to extraembryonic lineages in the late blastocyst (32, 33). NR5A2, a pioneer factor (34), is essential for naive pluripotency (together with ESRRB) in 2i ESCs (35) and can regulate the ICM program as early as the 8C stage (21). KO of *Nr5a2* arrested embryos at the morula stage (36, 37).

We then examined the occupancy of TFAP2C and NR5A2 (37) in 8C embryos using CUT&RUN. TFAP2C and NR5A2 extensively resided at pre-accessible sites but not at E4.5 epiblast-specific SOX2 binding sites (Fig. 4, A and B). Among 5578 distal SOX2-bound peaks in E3.5 ICM, 1571 and 1281 sites were already occupied by TFAP2C and NR5A2, respectively, at the 8C stage before *Sox2* expression (Fig. 4C). To determine whether these TFs can open these enhancers, we conducted ATAC-seq in *Tfap2c* mzKO 8C embryos. Depletion of TFAP2C preferentially reduced accessibility of pre-accessible enhancers specifically bound by TFAP2C (Fig. 4D and fig. S7A). Overall, TFAP2C contributed to opening 35% pre-accessible enhancers (SOX2-bound in E3.5 ICM) (Fig. 4E). A similar but less widespread effect was observed for NR5A2 at 8C, as *Nr5a2* knockdown decreased chromatin accessibility of 23% of pre-accessible enhancers (Fig. 4E and fig. S7B). Together, TFAP2C and NR5A2 accounted for opening 49% pre-accessible enhancers (with 9% regulated by both) (Fig. 4E). The rest of the pre-accessible sites may be opened by other early-stage TFs (e.g., GATA) (Fig. 4A).

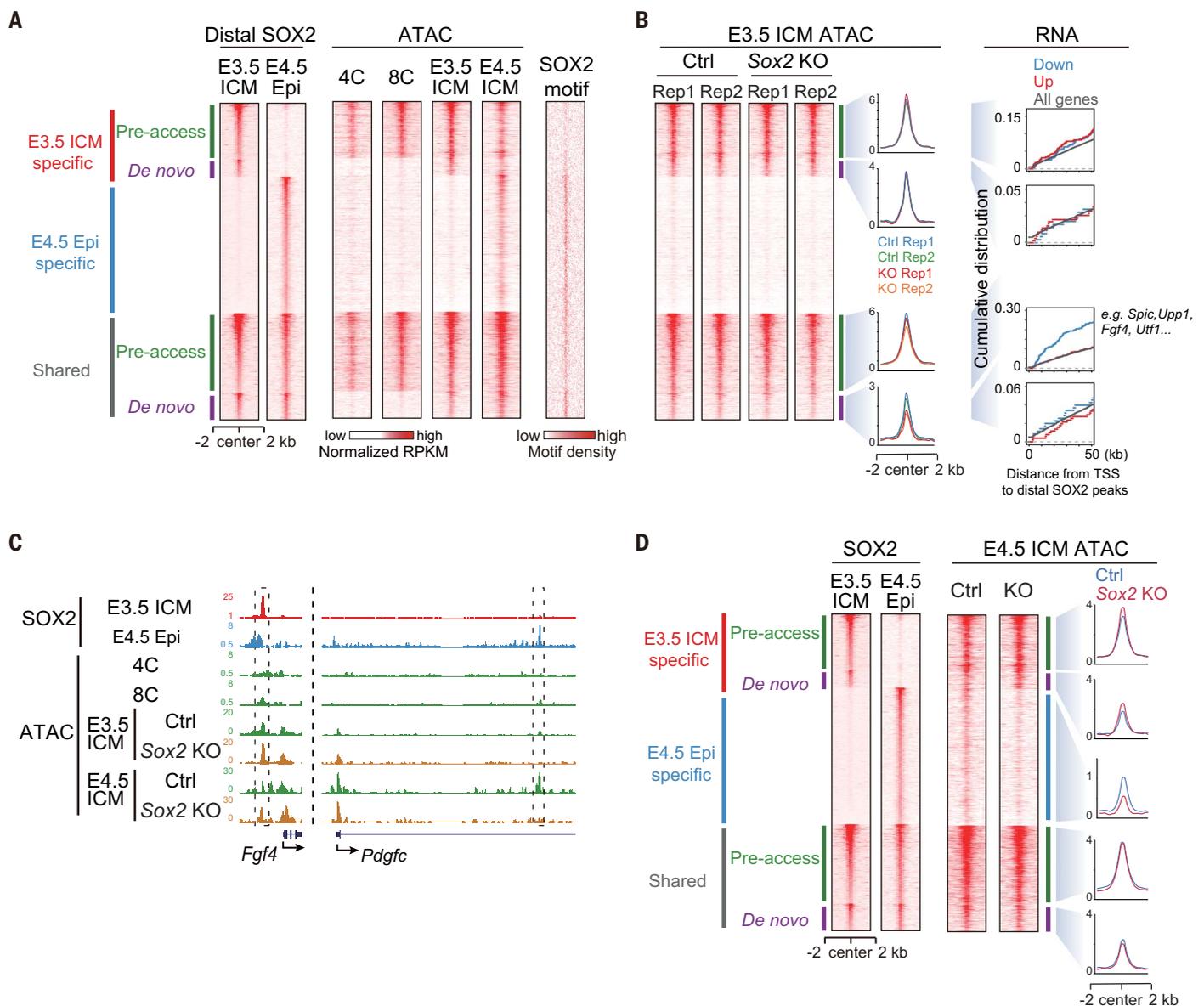
In 8C embryos, NR5A2 was preferentially required for opening chromatin at the 8C-specific binding sites but not regions bound by NR5A2 starting from the 2C stage (fig. S7C, “8C ATAC”), which suggests that these 2C-specific sites were also pre-accessible for NR5A2. Indeed, *Nr5a2* knockdown did not substantially affect global chromatin accessibility in 2C embryos (fig. S7C, “2C ATAC”), resembling SOX2’s role in opening enhancers in E4.5 epiblast but not E3.5 ICM. Therefore, these data indicate that both SOX2 and NR5A2 tend to initially exhibit settler binding at pre-accessible enhancers that are likely opened



**Fig. 2. SOX2 regulates ICM-TE lineage gene expression in E3.5 ICM.**

(A) (Left) Heatmaps showing enrichment of SOX2 binding signal at E3.5 ICM-specific ( $n = 1534$ ), E4.5 epiblast-specific ( $n = 2848$ ), and shared ( $n = 2223$ ) SOX2 binding peaks. (Middle) The cumulative distribution of E3.5 ICM-specific, E4.5 epiblast-specific, and all genes shows the fraction of genes within defined distances ( $x$  axis) between their transcription start sites (TSSs) to the nearest distal SOX2 binding peaks. (Right) UCSC browser views show SOX2 binding enrichment of representative genes. Gene expression is shown in heatmaps. (B) A t-distributed stochastic neighbor embedding (t-SNE) plot for

scRNA-seq of control (blue) and *Sox2* mzKO (red) ICM at E3.5. (C) Box plots showing *Sox2*, *Oct4*, and *Nanog* gene expression in control and *Sox2* mzKO ICM single cells, with  $P$  values ( $t$  test, one-sided) indicated. (D) (Left) Heatmaps showing the fold changes between ICM and TE gene expression. (Middle) The expression in control (Ctrl) and *Sox2* mzKO single cells (row z-score normalized) is mapped. (Right) The aggregated KO/control gene expression ratios (across all single cells) are shown. (E) Box plots showing expression of representative epiblast (EPI), PrE, and TE markers in control and *Sox2* mzKO ICM single cells, with  $P$  values ( $t$  test, one-sided) indicated.



**Fig. 3. SOX2 is dispensable for global enhancer accessibility in E3.5 ICM.** (A) (Left) Heatmaps showing enrichment of SOX2 binding signal, ATAC-seq signals, and SOX2 motif density (number of motifs per base pair) at SOX2 binding peaks. The E3.5 ICM-specific peaks are further clustered into pre-accessible ( $n = 1177$ ) and de novo ( $n = 357$ ) peaks on the basis of their accessibility states in 8C embryos. The shared peaks are also clustered into pre-accessible ( $n = 1658$ ) and de novo ( $n = 565$ ) peaks. (B) (Left) Heatmaps and average plots showing enrichment

of ATAC-seq signals in control and *Sox2* m/zKO E3.5 ICM at SOX2 binding peaks. (Right) The cumulative distributions of down-regulated, up-regulated, and all genes in *Sox2* m/zKO E3.5 ICM show the fraction of genes within defined distances ( $x$  axis) from their TSSs to the nearest SOX2 binding peaks. (C) UCSC browser views showing SOX2 binding (dashed boxes) and ATAC-seq enrichment of representative genes. (D) Heatmaps and average plots showing enrichment of SOX2 binding and ATAC-seq signals in control and *Sox2* m/zKO E4.5 ICM at SOX2 binding peaks.

by other TFs before adopting a more active role in opening enhancers at later stages.

### SOX2 mediates global enhancer opening in naive pluripotent cells

Although SOX2 was required to open E4.5 epiblast-specific enhancers, the failure to open these enhancers in *Sox2* m/zKO mutants could also be indirectly caused by earlier defective development (Fig. 2D). Therefore, we investigated the dependence of SOX2 for opening

enhancers in 2i ESCs, which resemble E4.5 epiblast (4, 38). We confirmed that E4.5 epiblast-specific, but not E3.5 ICM-specific, enhancers were preferentially bound by SOX2 in 2i ESCs (Fig. 2A). About 67.3% of SOX2 binding sites in E4.5 epiblast were recaptured in 2i ESCs. We then focused on SOX2 binding sites that were shared between 2i ESCs and E4.5 epiblast (Fig. 5A). RNA-seq showed that transcriptome defects were minimal at 12 hours upon SOX2 degradation but became apparent after 24 hours

(fig. S8A). Differentially expressed genes (DEGs) within 48 hours (521 down-regulated and 462 up-regulated) upon SOX2 degradation were identified (fig. S8A). Compared with DEGs in *Sox2* KO E3.5 ICM, only 52 down-regulated genes were shared, including 12 epiblast-specific genes (e.g., *Klf2*, *Fgf4*, and *Etv5*). E3.5 ICM-specific down-regulated genes contained more PrE genes (e.g., *Gata4*, *Pdgfra*, and *Sox17*), whereas ESC-specific down-regulated genes included more epiblast markers (e.g., *Esrrb*, *Klf4*, and *Nanog*)

#### Fig. 4. Early-stage expressing TFs are responsible for opening the preaccessible SOX2 binding sites.

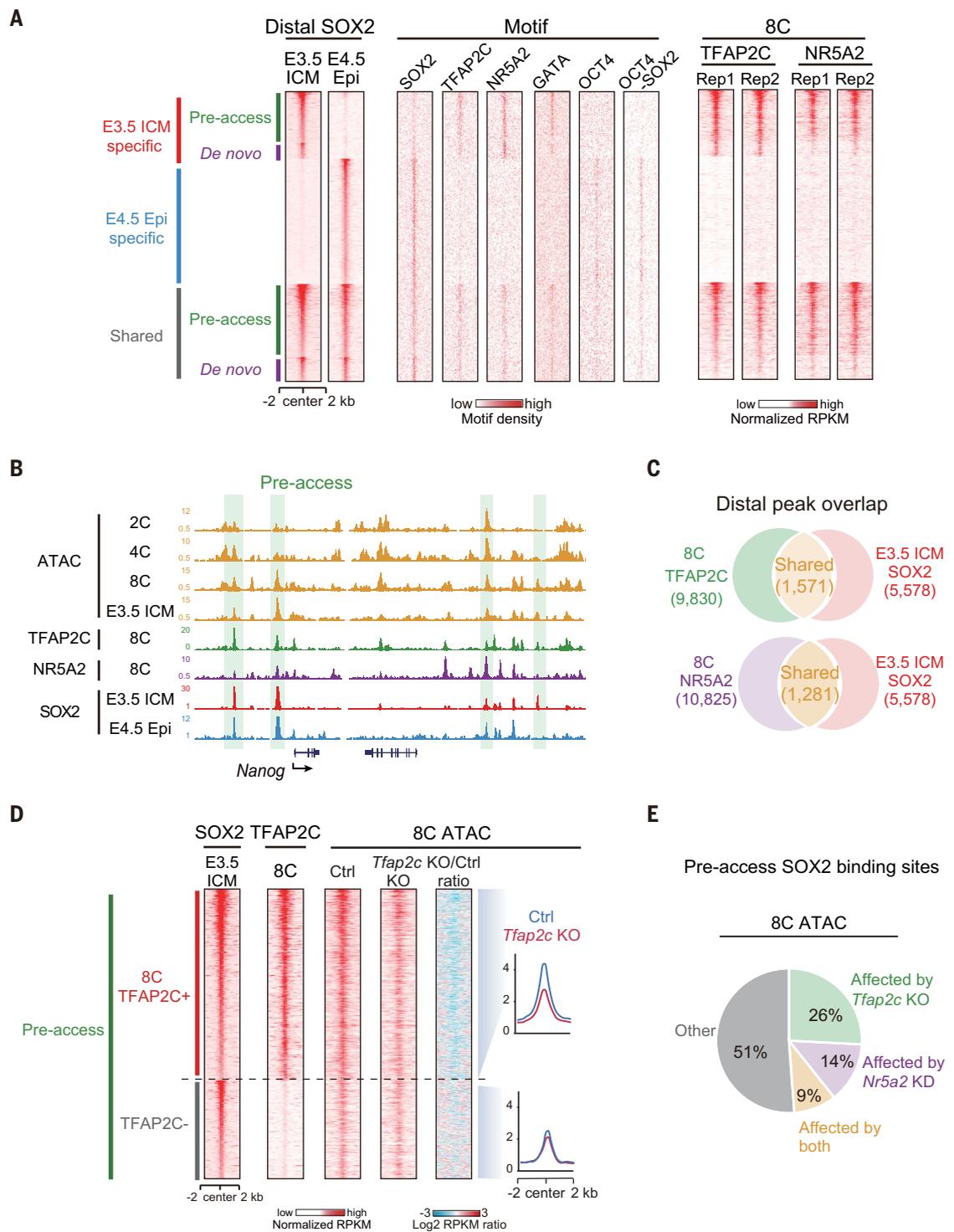
(A) Heatmaps showing enrichment of SOX2 binding signals in E3.5 ICM and E4.5 epiblast, motif densities, and TFAP2C and NR5A2 binding signals in 8C embryos at SOX2 binding peaks.

(B) UCSC browser views showing ATAC-seq enrichment and TFAP2C, NR5A2, and SOX2 binding signals of representative regions. Pre-accessible binding sites are shaded.

(C) Venn diagrams showing the overlap between distal TFAP2C (top) or NR5A2 (bottom) binding peaks in 8C embryos and SOX2 binding peaks in E3.5 ICM.

(D) Heatmaps showing SOX2 binding in E3.5 ICM, TFAP2C binding signals in 8C embryos, ATAC-seq enrichment in control and *Tfap2c* mzkO 8C embryos, and the ratios between *Tfap2c* mzkO and control 8C embryos at the preaccessible E3.5 ICM SOX2 binding peaks. Average plots of ATAC-seq signals are shown.

(E) Pie chart showing the percentages of *Tfap2c* KO affected, *Nr5a2* knockdown (KD) affected, and both affected ATAC-seq signal at the preaccessible SOX2 binding peaks.



(Fig. 5B and fig. S8B). Hence, SOX2 is required for broad expression of pluripotency marker genes in 2i ESCs but not in E3.5 ICM. On the other hand, several key TE markers were up-regulated in both cell types (e.g., *Gata2*, *Id2*, and *Krt8/18*) (>Fig. 5B). These results support a distinct state of E3.5 ICM from naive ESCs, in part reflected by its primitive pluripotency network and broader potency toward both epiblast and PrE.

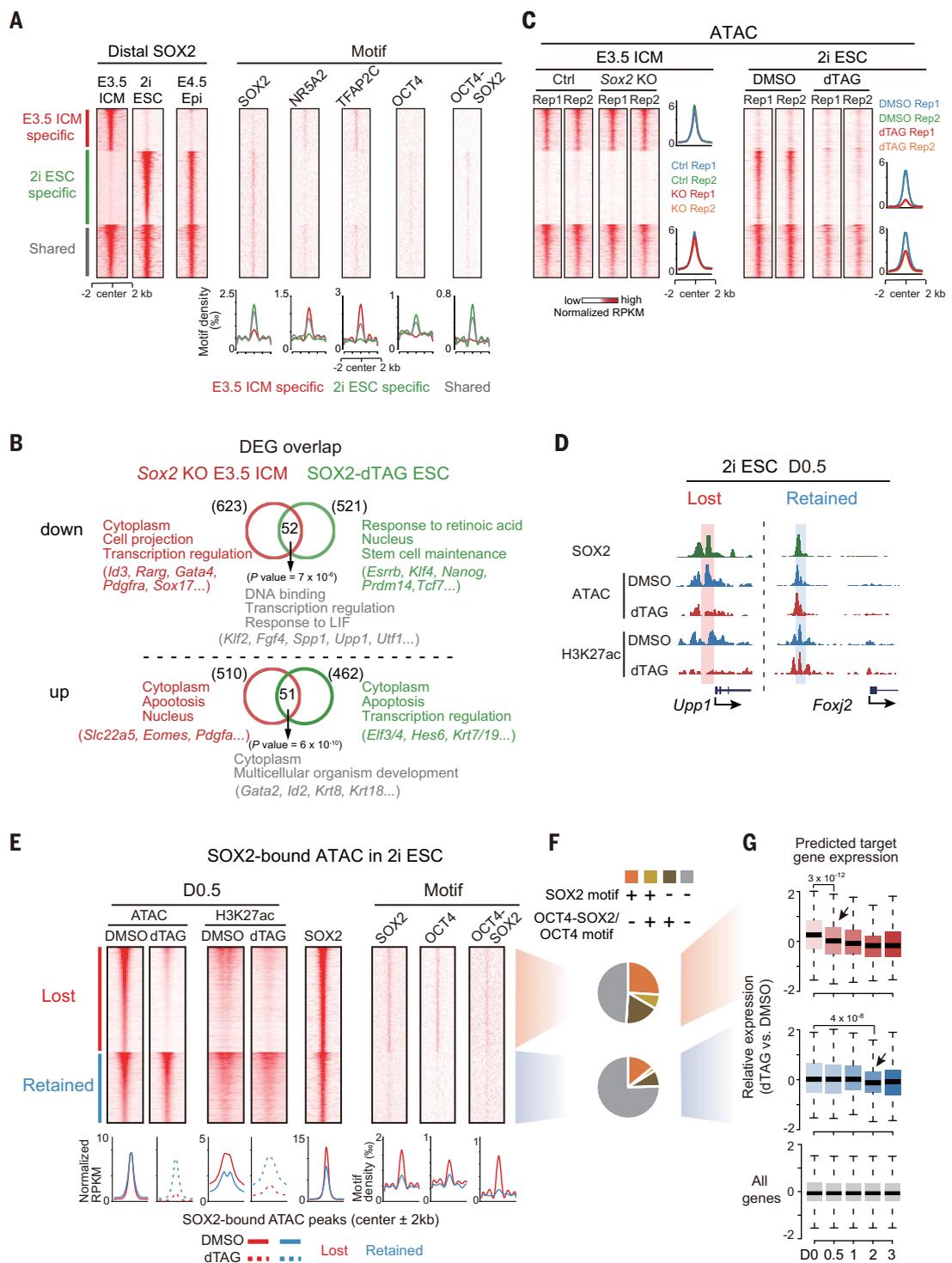
We then performed ATAC-seq and H3K27ac chromatin immunoprecipitation followed by sequencing (ChIP-seq) at day 0.5 (D0.5, or 12 hours) after SOX2 degradation, when transcription perturbation was still small (fig. S8A), to minimize secondary effects. About 59.8% of SOX2-bound ATAC peaks were already lost compared with only 36.8% of SOX2-unbound enhancers (Fig. 5C and fig. S9A;

see Fig. 5D for example). SOX2-dependent enhancers showed relatively stronger SOX2 binding and SOX2 motif enrichment (Fig. 5, E and F). Enhancers with both SOX2 and OCT4 motifs showed the highest SOX2 binding and SOX2 dependency (fig. S9B). Moreover, the putative target genes of the SOX2-dependent enhancers (materials and methods) were down-regulated at an earlier time point (D0.5) upon

### Fig. 5. Distinct roles of SOX2 in gene expression and enhancer regulation in E3.5 ICM and 2i ESCs.

(A) Heatmaps showing enrichment of SOX2 binding signals and TF motif densities at E3.5 ICM-specific, 2i ESC-specific (also E4.5 epiblast FPKM > 1), and shared (also E4.5 epiblast FPKM > 1) distal SOX2 binding peaks, with the average plots of TF motif densities shown below. (B) Venn diagrams showing the overlap of down- and up-regulated genes between Sox2 KO (versus control) E3.5 ICM and dTAG [versus dimethyl sulfoxide (DMSO)-treated] 2i ESCs. Gene ontology results and example genes are shown. P values (hypergeometric distribution) are shown.

(C) Heatmaps showing enrichment of ATAC-seq signals in control and Sox2 mZKO E3.5 ICM and in DMSO- and dTAG-treated (for 12 hours) 2i ESCs. The average plots of ATAC-seq signal are shown. (D) UCSC browser views showing enrichment of SOX2 binding, ATAC-seq, and H3K27ac signals in DMSO- and dTAG-treated 2i ESCs of representative genes. The lost and retained ATAC-seq peaks are shaded. (E) Heatmaps showing enrichment of ATAC-seq, H3K27ac, SOX2 binding signals, and TF motif densities at lost and retained ATAC-seq peaks in DMSO- or dTAG-treated 2i ESCs. The average enrichment is shown below. (F) Pie charts showing the percentages of peaks with SOX2 motifs only, both SOX2 and OCT4-SOX2/OCT4 motifs, OCT4-SOX2/OCT4 motifs, and neither motif at the peaks. (G) The relative expression of predicted target genes (row z-score normalized) of the ATAC-seq peaks and all genes (control), with P values (t test, two-sided) indicated on top. The arrows indicate the time when a significant decrease of gene expression is detected.



SOX2 loss compared with targets of SOX2-independent enhancers, which were down-regulated around day 2 (Fig. 5G), supporting more direct SOX2 impacts. SOX2-independent enhancers showed comparable ATAC-seq and even stronger H3K27ac signals compared with SOX2-dependent enhancers in WT cells (Fig. 5E), which suggests that they may be opened by

other TFs, although the motif analysis did not reveal obvious candidates (fig. S9C). In sum, these data suggest that in the naive pluripotent state, SOX2 opens enhancers—preferentially those with SOX2 or OCT4 motifs—in a pioneer binding mode. Together with the dispensability of SOX2 for opening global enhancers in E3.5 ICM, this result is in line with the finding

that *Sox2*-null embryos can give rise to ICM but not to ESCs (17).

#### The second acute global binding transition underlies the essential role of SOX2 for formative pluripotency induction

*Sox2*-null embryos die shortly after implantation (17). Coincidentally, SOX2 binding sites

showed a second major relocalization from E4.5 epiblast to E5.5 epiblast (fig. S10A), which suggests an involvement of SOX2 during the naive-to-formative pluripotency transition. In this process, whereas pluripotency genes *Sox2* and *Oct4* continue to be expressed, naive pluripotency genes (e.g., *Nanog*, *Tbx3*, and *Tbx20*) are repressed, and postimplantation epiblast genes (e.g., *Sall2*, *Fgf5*, and *Fgf15*) are activated (5). Accordingly, the motifs of SOX2 and OCT4 were enriched at both E4.5 epiblast- and E5.5 epiblast-specific SOX2-bound enhancers. By contrast, E5.5 epiblast-specific sites were enriched for motifs of ZIC3 and OTX2 (fig. S10A)—two TFs up-regulated during formative pluripotency induction, which mediate the naive-to-formative transition (39–42).

Because early lethality precludes the study of SOX2's role during the naive-to-formative pluripotency transition in vivo, we used the 2i ESC-to-EpiLC conversion ex vivo model (5, 43). SOX2 binding in EpiLCs recapitulated that in E5.5 epiblast, with its binding sites enriched for the motifs of OTX2 and ZIC3 (fig. S10A). In fact, the co-occupancy of SOX2 with OCT4, ZIC3, and OTX2 was observed in EpiLCs (fig. S10B). After SOX2 depletion, global gene expression transition from naive-to-formative pluripotency was severely impaired (Fig. 6, A and B). About 57.3% (425 of 742) naive genes failed to be properly repressed (fig. S11A, left, "Down dependent"), and 78.2% (453 of 579) formative genes showed defective activation, including marker genes *Pou3f1*, *Fgf15*, *Dnmt3a/b*, *Zic3*, and *Otx2* (fig. S11A, left, "Up dependent," and fig. S11, B and C). The differentiation defects are likely a result of a systematic failure of the transcription program because they cannot be rescued by reintroducing *Otx2* or *Zic3* alone (fig. S11D). Supporting a direct role in gene activation, SOX2 in EpiLCs preferentially occupied enhancers near SOX2-dependent formative genes (fig. S11A, right). To probe how enhancers were globally affected upon the loss of SOX2, we performed ATAC-seq and H3K27ac ChIP-seq during the ESC-to-EpiLC transition. The decommissioning of ESC-specific enhancers (indicated by the loss of ATAC-seq peaks) upon differentiation was largely unaffected by SOX2 depletion (fig. S12A). However, 85% of SOX2-bound, newly established EpiLC-specific enhancers (compared with 61% of SOX2-unbound enhancers) failed to be properly established (Fig. 6C). To further validate the direct function of SOX2 at these enhancers, we tested SOX2-bound enhancers near *Otx2* and *Zic3*—two SOX2-dependent formative genes in EpiLCs (fig. S12B). All five enhancers that we tested drove strong reporter activities, with four being SOX2 dependent (fig. S12B). In sum, SOX2 is essential for the naive-to-formative pluripo-

tency conversion and is required for the activation of formative enhancers.

### Prebinding of SOX2 is insufficient to open enhancers but correlates with faster future enhancer opening

We hypothesized that SOX2-dependent formative-specific enhancers were activated by SOX2 through pioneer binding. However, we found that 54% of these enhancers were already prebound by SOX2 in 2i ESCs, despite still being inaccessible when assayed by ATAC-seq (Fig. 6D, "*Dnmt3b*," and Fig. 6E, "prebinding"). SOX2 binding at these sites further increased upon differentiation to EpiLCs. Prebinding of SOX2 at these enhancers was also observed in E4.5 epiblast in vivo (fig. S13A, "SOX2 in vivo"). The remaining 46% acquired SOX2 binding during differentiation, consistent with pioneer binding (Fig. 6D, "*Fgf15*," and Fig. 6E, "pioneer binding"). We investigated whether these enhancers with SOX2 prebinding and pioneer binding exhibit functional differences. Predicted target genes of both groups were activated with similar kinetics upon 2i ESC-to-EpiLC differentiation, in a SOX2-dependent manner (Fig. 6D, "Lost"). Thus, both SOX2 prebinding and pioneer binding were required for gene activation. By contrast, SOX2 was not essential to activate targets of SOX2-independent formative enhancers (Fig. 6D, "Retained").

We then sought to identify features that distinguish prebinding and pioneer binding enhancers. Both classes were depleted of H3K4me1 in ESCs, which indicates that they were not the classic "poised" enhancers (44) (fig. S13A). However, pioneer binding sites were enriched for both the SOX2 motif and the OTX2 motif, suggesting cooperative binding upon formative pluripotency induction (Fig. 6, F and G). Prebinding sites were more enriched for the SOX2 motif, which raises the possibility that such a strong motif is sufficient to recruit SOX2 in 2i ESCs, which may lower the threshold of enhancer activation to compensate for the weak OTX2 motif. We found that prebinding enhancers became accessible faster than pioneer binding enhancers during formative induction (Fig. 6E and fig. S13B), an observation also reproduced when analyzing H3K27ac (fig. S13C). Finally, OTX2 binding showed comparable increases in the two groups (fig. S13A, "OTX2"). We speculate that the increased binding of OTX2 in the prebinding enhancer group lacking its motif could be facilitated by other factors, such as SOX2. These data indicate that although prebinding of SOX2 is insufficient to open formative enhancers, it may poise those with weak formative TF motifs for faster future opening.

### SOX2 prebinds germ layer enhancers in E5.5 epiblast

The fact that SOX2 can prebind and possibly poise enhancers for future activation prompted

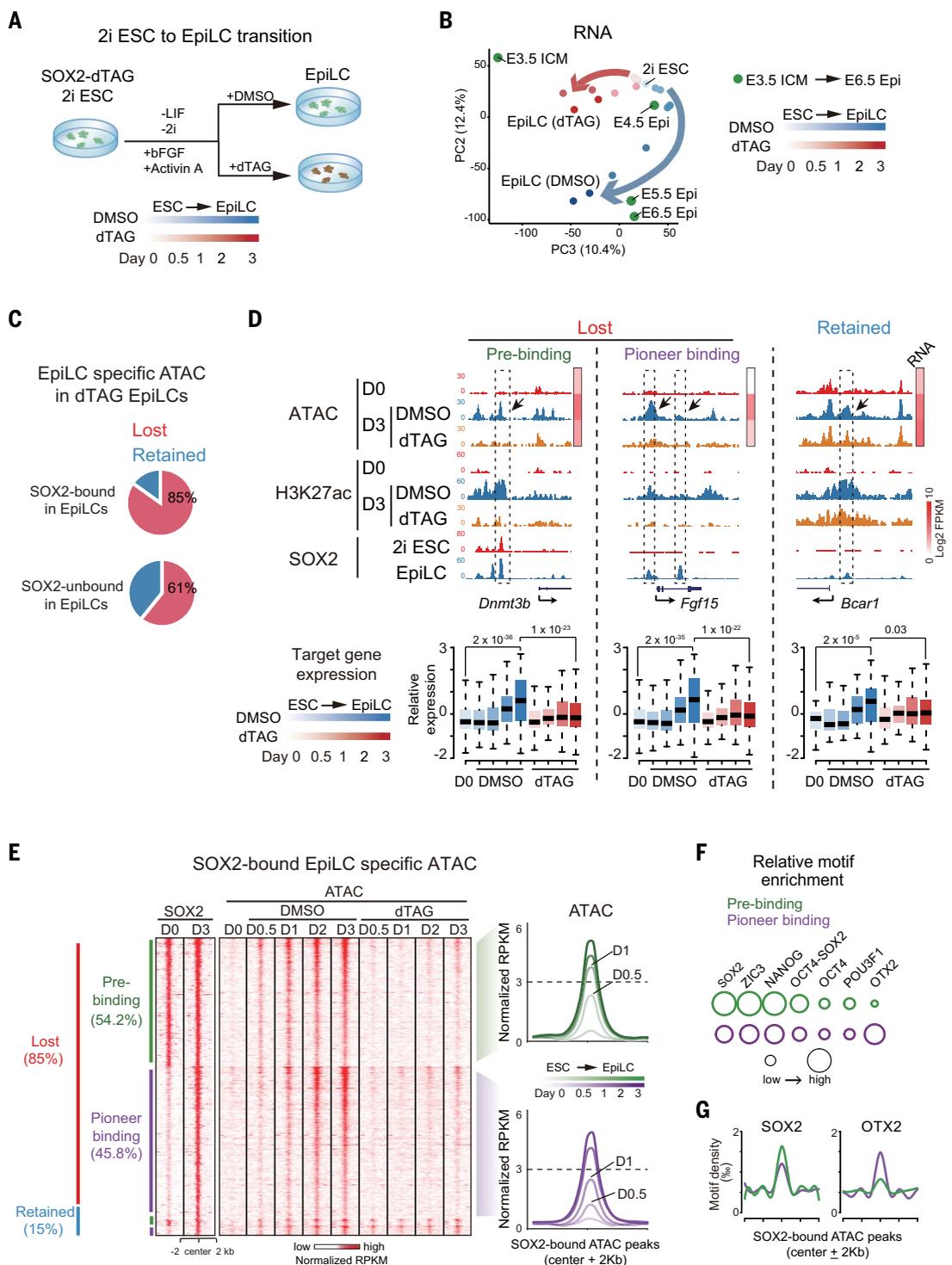
us to explore whether this observation can be extended to other developmental processes. During gastrulation, SOX2 is required to drive neural ectoderm differentiation (45). Many enhancers in ectoderm are already primed in mouse epiblast (23, 46), consistent with the model that ectoderm is a default differentiation lineage from epiblast (47). By identifying the putative enhancers specific to epiblast and three germ layers using ATAC-seq data (23), we found that SOX2 occupied not only epiblast-specific but also 45% of ectoderm-specific enhancers in E5.5 epiblast (fig. S14, A and B). Moreover, SOX2 preferentially resided near both epiblast-specific and ectoderm-specific genes in E5.5 epiblast (fig. S14C). Upon the transition to ectoderm, SOX2 binding at ectoderm-specific enhancers was strengthened, whereas SOX2 binding at epiblast-specific enhancers was lost (fig. S14A). Hence, SOX2 prebinds a subset of developmentally regulated enhancers, which supports the notion that SOX2 functions as a lineage specifier toward ectoderm during gastrulation, and that formative pluripotency installs competence for somatic lineage specification (4). The prebinding of SOX2 resembles the reported binding of pioneer TFs and nonpioneer "bookmarking" TFs to regulatory elements before gene activation (48), but this binding does not immediately create open chromatin before receiving further differentiation cues. Therefore, we referred to such prebinding as "pilot binding," which further demonstrates the flexibility and versatility of pioneer factors in different cellular contexts. In fact, a quantitative analysis of enhancers for their SOX2 binding and chromatin accessibility across developmental stages revealed that different modes of SOX2-chromatin interactions are under constant transition (Fig. 7A). These distinct binding actions likely depend on cell type-specific cooperative TFs as well as genetic and epigenetic contexts.

### Discussion

Tremendous progress has been achieved to understand the molecular circuitry underlying pluripotency regulation using stem cell models. How master TFs guide pluripotency progression in vivo remains poorly understood. In this work, by profiling the chromatin binding of SOX2 in mouse early embryos, we found a chromatin state and transcription circuitry for E3.5 ICM that differs from all other pluripotent states. The potency of E3.5 ICM exceeds pluripotency because it can give rise to epiblast and PrE. Moreover, E3.5 ICM shows distinct transcriptome and chromatin accessibility (fig. S4B), which likely reflect the coexpression of pluripotency factors, early-stage TFs, and extraembryonic lineage TFs. The pluripotency master TF-mediated regulatory network appears still at a primitive stage in E3.5 ICM, as supported by multiple

### Fig. 6. Depletion of SOX2 impedes the naive-to-formative pluripotency transition.

**(A)** Schematic showing differentiation of SOX2-dTAG 2i ESCs to EpiLCs with DMSO or dTAG treatment. **(B)** PCA showing RNA-seq of cells from 2i ESCs to EpiLCs at day 0 to day 3 with DMSO (blue) or dTAG (red) treatment. **(C)** Pie charts showing the percentages of lost and retained peaks at EpiLC-specific distal ATAC-seq peaks with or without SOX2 binding. **(D)** (Top) UCSC browser views and heatmaps showing SOX2 binding signals, ATAC-seq, H3K27ac enrichment, and gene expression of representative genes in 2i ESCs (D0) and DMSO- and dTAG-treated EpiLCs (D3). Arrows and dashed boxes indicate lost or retained ATAC-seq peaks. (Bottom) Box plots show the relative expression (row z-score normalized) of predicted target genes of peaks during the EpiLC transition, with *P* values (*t* test, two-sided) indicated on top. **(E)** (Left) Heatmaps showing enrichment of SOX2 binding and ATAC-seq signals at the lost and retained EpiLC-specific ATAC-seq peaks during the EpiLC transition with DMSO or dTAG treatment. The lost peaks are further clustered into SOX2 prebinding and pioneer binding sites in 2i ESCs. (Right) Line charts show the average ATAC-seq enrichment at the SOX2 prebinding or pioneer binding EpiLC-specific ATAC-seq peaks. **(F)** TF motifs identified from EpiLC-specific ATAC-seq peaks with SOX2 prebinding or pioneer binding. Sizes of circles indicate levels of  $-\log P$  values. **(G)** Density plots showing corresponding TF motif density at SOX2 prebinding or pioneer binding EpiLC-specific ATAC-seq peaks.

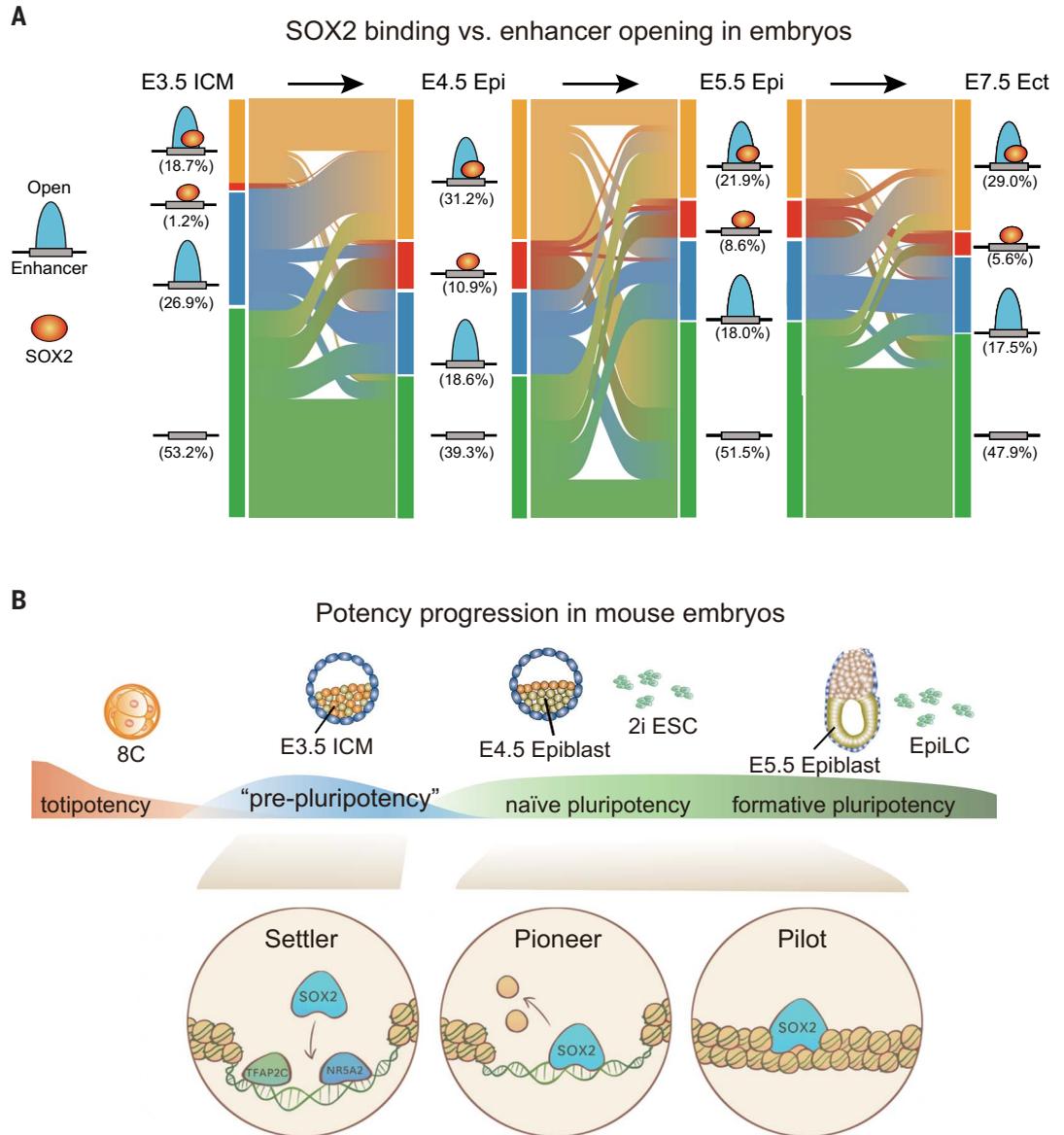


pieces of evidence. (i) SOX2's binding peaks are less enriched for motifs of SOX2, OCT4, and OCT4-SOX2 but are more enriched for the motifs of early-stage TFs (such as NR5A2, TFAP2C, and GATA) (Fig. 1E). SOX2 binds pre-accessible chromatin in part opened by these early-stage TFs (Fig. 4E). (ii) The OCT4-SOX2 motif enrichment in SOX2-bound sites is strong in 2i ESCs and E4.5 epiblast but not in E3.5

ICM, which raises the possibility that the cooperative function of OCT4 and SOX2 may only become dominant after entering naive pluripotency. Accordingly, although both OCT4 and SOX2 promote ICM-specific genes and repress TE-specific genes in E3.5 ICM (26, 49, 50), their targets appear to differ (fig. S6B). Moreover, the transcriptional interdependence of master pluripotency TFs OCT4, SOX2, and

NANOG in ESCs (12, 16, 51, 52) was also not observed in E3.5 ICM. Unlike that in ESCs (10), *Sox2* KO in E3.5 ICM did not affect expression of *Oct4* or *Nanog* (Fig. 2C), and *Oct4* KO did not affect *Nanog* expression and only partially down-regulates *Sox2* (26, 53, 54) (fig. S6A). (iii) SOX2 is globally dispensable for enhancer opening in E3.5 ICM, whereas it is essential for opening enhancers genome-wide in 2i ESCs and E4.5

**Fig. 7. SOX2-chromatin binding in early development and its multifaceted interaction modes with enhancers.** (A) Alluvial diagrams showing the dynamics states of enhancers for their SOX2 binding and accessibility (based on ATAC-seq) in embryos. The percentages of each class of SOX2-chromatin interactions (combination of SOX2 occupancy and chromatin accessibility) within all enhancers (pooled from all stages examined) are shown for each stage. (B) A model illustrating the multifaceted SOX2 interaction modes with enhancers during the pluripotency transition. Three SOX2 binding modes are proposed: settler, pioneer, and pilot binding. The settler binding refers to SOX2 binding pre-accessible enhancers, and its depletion does not substantially affect chromatin opening, as exemplified by most SOX2 binding in E3.5 ICM. The pioneer binding occurs in E4.5 epiblast or 2i mESCs, where SOX2 is required to establish or maintain the enhancer accessibility. The pilot binding of SOX2 at many formative enhancers in 2i mESCs is insufficient for enhancer opening but may help poise enhancers for faster opening upon conversion to formative pluripotency. E3.5 ICM is in a prepluripotency state to bridge totipotency to pluripotency, featured by coexpression of multilineage markers (epiblast, PrE, and early-stage TFs), expanded potency toward epiblast and PrE, and a primitive pluripotency network.



epiblast. A larger role of pluripotency factors in ESCs is consistent with them being required for mouse ESC derivation (50, 55) and maintenance (16). Hence, the E3.5 ICM exhibits a distinct, "prepluripotency" state, featuring the potential to give rise to both epiblast and PrE, coexpression of multilineage TFs, and a primitive pluripotency network.

Pioneer TFs are believed to bind and open inaccessible chromatin, leading to the subsequent recruitment of additional TFs (17). We found that SOX2 manifests more diverse roles at enhancers beyond a simple pioneer factor, which include settler binding, pioneer binding, and pilot binding. SOX2 exhibits settler binding in E3.5 ICM, where SOX2 binds pre-accessible enhancers, and its loss does not substantially affect chromatin opening (Fig. 7B). We speculate that the relatively short expression period and lack of cooperation with

OCT4 or other cofactors may disable its pioneering binding function. Notably, such settler binding can still exert impacts on gene expression, especially at sites with SOX2 motifs (Fig. 3B). It is possible that the SOX2 motif may increase the residence time of SOX2, which in turn promotes gene expression, for example by increasing promoter-enhancer interactions (8, 56). It also remains to be investigated whether some settler binding may help sequester excess SOX2 from other binding sites to prevent premature activation of later-stage genes. Widespread pioneer binding is then observed in E4.5 epiblast and 2i ESCs, where SOX2 is required for naive enhancers opening (Fig. 3D and Fig. 5C). Finally, the pilot binding of SOX2 at many formative enhancers in 2i ESCs is insufficient for enhancer opening but likely helps enhancers with weak formative TF motifs achieve faster opening upon differentiation (Fig.

6E). We propose that such multifaceted—rather than a universal pioneering—chromatin interacting modes may also hold true for other pioneer TFs to allow precise yet adaptable responses to developmental cues beyond pluripotency regulation.

#### Materials and methods summary

A detailed materials and methods section is provided in the supplementary materials. All animals were cared for according to the guidelines of the Institutional Animal Care and Use Committee of Tsinghua University. Embryos were collected from superovulated females crossed with males. To generate *Sox2* mzkO embryos, *Sox2*<sup>fllox/fllox</sup>, *Zp3-Cre* females and *Sox2*<sup>fllox/fllox</sup>, *Stra8-Cre* males were used for breeding. Immunosurgery was performed as reported previously (57) to remove TE and isolate ICM. ICMs were then incubated in TrypLE

and dissociated by repetitive pipetting using a Pasteur pipette. For scRNA-seq, individual E3.5 ICM cells were transferred into single-cell lysis buffer following the Smart-seq2 protocol, as described previously (58). E4.5 blastocysts were flushed from the uterus after human chorionic gonadotropin (hCG) injection at 114 to 116 hours. Given that SOX2 was present only in epiblast but not in PrE at E4.5 (10), we profiled SOX2 binding using the entire ICM because the signals were expected to arise exclusively from epiblast cells. E5.5 to E7.5 embryo tissues were collected as previously described (23, 59, 60).

CUT&RUN was conducted following the published protocol (61) with some modifications. The fresh samples were resuspended and bound with concanavalin-coated magnetic beads. After incubation with SOX2 antibody for 2 to 3 hours at 4°C, the samples were incubated with protein A-micrococcal nuclease (pA-MNase) for 1 hour. The STAR ChIP-seq for H3K27ac and miniATAC-seq were performed as previously described (62, 63).

To construct SOX2-dTAG ESCs, the sequence encoding FKBP<sup>F36V</sup>-GFP was fused to the C terminus of the endogenous *Sox2* locus. SOX2-FKBP proteins were depleted by adding dTAG<sup>V</sup>-1 into the medium. Time-course experiments were performed by inducing protein degradation and collecting the samples at different time points. Naive mESCs (2i mESCs) were cultured in the N2B27 medium supplemented with PD0325901, Chir99021, and LIF. To induce EpiLC differentiation, 2i ESC cells were plated on tissue culture dishes pretreated with matrigel in N2B27-based medium supplemented with 1% knockout serum replacement (KSR), basic fibroblast growth factor (bFGF), and activin A.

## REFERENCES AND NOTES

- J. Rossant, P. P. L. Tam, New Insights into Early Human Development: Lessons for Stem Cell Derivation and Differentiation. *Cell Stem Cell* **20**, 18–28 (2017). doi: [10.1016/j.stem.2016.12.004](https://doi.org/10.1016/j.stem.2016.12.004); PMID: 28061351
- M. Zernicka-Goetz, S. A. Morris, A. W. Bruce, Making a firm decision: Multifaceted regulation of cell fate in the early mouse embryo. *Nat. Rev. Genet.* **10**, 467–477 (2009). doi: [10.1038/nrg2564](https://doi.org/10.1038/nrg2564); PMID: 19536196
- C. Chazaud, Y. Yamanaka, T. Pawson, J. Rossant, Early lineage segregation between epiblast and primitive endoderm in mouse blastocysts through the Grb2-MAPK pathway. *Dev. Cell* **10**, 615–624 (2006). doi: [10.1016/j.devcel.2006.02.020](https://doi.org/10.1016/j.devcel.2006.02.020); PMID: 16678776
- A. Smith, Formative pluripotency: The executive phase in a developmental continuum. *Development* **144**, 365–373 (2017). doi: [10.1242/dev.142679](https://doi.org/10.1242/dev.142679); PMID: 28143843
- K. Hayashi, H. Ohta, K. Kurimoto, S. Aramaki, M. Saitou, Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. *Cell* **146**, 519–532 (2011). doi: [10.1016/j.cell.2011.06.052](https://doi.org/10.1016/j.cell.2011.06.052); PMID: 21820164
- J. Nichols, A. Smith, Naive and primed pluripotent states. *Cell Stem Cell* **4**, 487–492 (2009). doi: [10.1016/j.stem.2009.05.015](https://doi.org/10.1016/j.stem.2009.05.015); PMID: 19497275
- J. Rossant, Genetic Control of Early Cell Lineages in the Mammalian Embryo. *Annu. Rev. Genet.* **52**, 185–201 (2018). doi: [10.1146/annurev-genet-120116-024544](https://doi.org/10.1146/annurev-genet-120116-024544); PMID: 30183407
- F. Spitz, E. E. Furlong, Transcription factors: From enhancer binding to developmental control. *Nat. Rev. Genet.* **13**, 613–626 (2012). doi: [10.1038/nrg3207](https://doi.org/10.1038/nrg3207); PMID: 22868264
- R. A. Young, Control of the embryonic stem cell state. *Cell* **144**, 940–954 (2011). doi: [10.1016/j.cell.2011.01.032](https://doi.org/10.1016/j.cell.2011.01.032); PMID: 21414485
- E. Wicklow *et al.*, HIPPO pathway members restrict SOX2 to the inner cell mass where it promotes ICM fates in the mouse blastocyst. *PLoS Genet.* **10**, e1004618 (2014). doi: [10.1371/journal.pgen.1004618](https://doi.org/10.1371/journal.pgen.1004618); PMID: 25340657
- A. A. Avilion *et al.*, Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev.* **17**, 126–140 (2003). doi: [10.1101/gad.224503](https://doi.org/10.1101/gad.224503); PMID: 12514105
- X. Chen *et al.*, Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* **133**, 1106–1117 (2008). doi: [10.1016/j.cell.2008.04.043](https://doi.org/10.1016/j.cell.2008.04.043); PMID: 18555785
- R. Blassberg *et al.*, Sox2 levels regulate the chromatin occupancy of WNT mediators in epiblast progenitors responsible for vertebrate body formation. *Nat. Cell Biol.* **24**, 633–644 (2022). doi: [10.1038/s41556-022-00910-2](https://doi.org/10.1038/s41556-022-00910-2); PMID: 35550614
- S. J. Hainer, S. Bošković, K. N. McCannell, O. J. Rando, T. G. Fazio, Profiling of Pluripotency Factors in Single Cells and Early Embryos. *Cell* **177**, 1319–1329.e11 (2019). doi: [10.1016/j.cell.2019.03.014](https://doi.org/10.1016/j.cell.2019.03.014); PMID: 30955888
- J. J. Thompson *et al.*, Extensive co-binding and rapid redistribution of NANOG and GATA6 during emergence of divergent lineages. *Nat. Commun.* **13**, 4257 (2022). doi: [10.1038/s41467-022-31938-5](https://doi.org/10.1038/s41467-022-31938-5); PMID: 35871075
- S. Masui *et al.*, Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. *Nat. Cell Biol.* **9**, 625–635 (2007). doi: [10.1038/ncb1589](https://doi.org/10.1038/ncb1589); PMID: 17515932
- K. S. Zaret, J. S. Carroll, Pioneer transcription factors: Establishing competence for gene expression. *Genes Dev.* **25**, 2227–2241 (2011). doi: [10.1101/gad.176826.11](https://doi.org/10.1101/gad.176826.11); PMID: 22056668
- A. Soufi *et al.*, Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming. *Cell* **161**, 555–568 (2015). doi: [10.1016/j.cell.2015.03.017](https://doi.org/10.1016/j.cell.2015.03.017); PMID: 25892221
- T. Boroviak, R. Loos, P. Bertone, A. Smith, J. Nichols, The ability of inner-cell-mass cells to self-renew as embryonic stem cells is acquired following epiblast specification. *Nat. Cell Biol.* **16**, 513–528 (2014). doi: [10.1038/ncb2965](https://doi.org/10.1038/ncb2965); PMID: 24859004
- S. Okumura-Nakanishi, M. Saito, H. Niwa, F. Ishikawa, Oct-3/4 and Sox2 regulate Oct-3/4 gene in embryonic stem cells. *J. Biol. Chem.* **280**, 5307–5317 (2005). doi: [10.1074/jbc.M410015200](https://doi.org/10.1074/jbc.M410015200); PMID: 15557334
- J. Wu *et al.*, The landscape of accessible chromatin in mammalian preimplantation embryos. *Nature* **534**, 652–657 (2016). doi: [10.1038/nature18606](https://doi.org/10.1038/nature18606); PMID: 27309802
- B. Nabet *et al.*, The dTAG system for immediate and target-specific protein degradation. *Nat. Chem. Biol.* **14**, 431–441 (2018). doi: [10.1038/s41589-018-0021-8](https://doi.org/10.1038/s41589-018-0021-8); PMID: 29581585
- Y. Xiang *et al.*, Epigenomic analysis of gastrulation identifies a unique chromatin state for primed pluripotency. *Nat. Genet.* **52**, 95–105 (2020). doi: [10.1038/s41588-019-0545-1](https://doi.org/10.1038/s41588-019-0545-1); PMID: 31844322
- B. Plusa, A. Pliliszek, S. Frankenberg, J. Artus, A. K. Hadjantonakis, Distinct sequential cell behaviours direct primitive endoderm formation in the mouse blastocyst. *Development* **135**, 3081–3091 (2008). doi: [10.1242/dev.021519](https://doi.org/10.1242/dev.021519); PMID: 18725515
- Y. Ohnishi *et al.*, Cell-to-cell expression variability followed by signal reinforcement progressively segregates early mouse lineages. *Nat. Cell Biol.* **16**, 27–37 (2014). doi: [10.1038/ncb2881](https://doi.org/10.1038/ncb2881); PMID: 24292013
- G. G. Stirparo *et al.*, OCT4 induces embryonic pluripotency via STAT3 signaling and metabolic mechanisms. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2008890118 (2021). doi: [10.1073/pnas.2008890118](https://doi.org/10.1073/pnas.2008890118); PMID: 33452132
- F. Lu *et al.*, Establishing Chromatin Regulatory Landscape during Mouse Preimplantation Development. *Cell* **165**, 1375–1388 (2016). doi: [10.1016/j.cell.2016.05.050](https://doi.org/10.1016/j.cell.2016.05.050); PMID: 27259149
- R. I. Sherwood *et al.*, Discovery of directional and nondirectional pioneer transcription factors by modeling DNase profile magnitude and shape. *Nat. Biotechnol.* **32**, 171–178 (2014). doi: [10.1038/nbt.2798](https://doi.org/10.1038/nbt.2798); PMID: 24441470
- P. Gu *et al.*, Orphan nuclear receptor LRH-1 is required to maintain Oct4 expression at the epiblast stage of embryonic development. *Mol. Cell Biol.* **25**, 3492–3505 (2005). doi: [10.1128/MCB.25.9.3492-3505.2005](https://doi.org/10.1128/MCB.25.9.3492-3505.2005); PMID: 15831456
- U. Werling, H. Schorle, Transcription factor gene *AP-2γ* essential for early murine development. *Mol. Cell Biol.* **22**, 3149–3156 (2002). doi: [10.1128/MCB.22.9.3149-3156.2002](https://doi.org/10.1128/MCB.22.9.3149-3156.2002); PMID: 11940672
- H. J. Auman *et al.*, Transcription factor AP-2γ is essential in the extra-embryonic lineages for early postimplantation development. *Development* **129**, 2733–2747 (2002). doi: [10.1242/dev.129.11.2733](https://doi.org/10.1242/dev.129.11.2733); PMID: 12015300
- Z. Cao *et al.*, Transcription factor AP-2γ induces early *Cdx2* expression and represses HIPPO signaling to specify the trophoblast lineage. *Development* **142**, 1606–1615 (2015). doi: [10.1242/dev.120238](https://doi.org/10.1242/dev.120238); PMID: 25858457
- M. Zhu *et al.*, Developmental clock and mechanism of de novo polarization of the mouse embryo. *Science* **370**, eabd2703 (2020). doi: [10.1126/science.abd2703](https://doi.org/10.1126/science.abd2703); PMID: 33303584
- J. Gassler *et al.*, Zygotic genome activation by the totipotency pioneer factor Nr5a2. *Science* **378**, 1305–1315 (2022). doi: [10.1126/science.abn7478](https://doi.org/10.1126/science.abn7478); PMID: 36423263
- N. Festuccia, N. Owens, A. Chervova, A. Dubois, P. Navarro, The combined action of Esrrb and Nr5a2 is essential for murine naive pluripotency. *Development* **148**, dev199604 (2021). doi: [10.1242/dev.199604](https://doi.org/10.1242/dev.199604); PMID: 34397088
- N. Festuccia *et al.*, Nr5a2 is essential for morula development. bioRxiv 2023.01.16.524255 [Preprint] (2023); <https://doi.org/10.1101/2023.01.16.524255>.
- F. Lai *et al.*, NR5A2 connects zygotic genome activation to the first lineage segregation in totipotent embryos. *Cell Res.* [10.1038/s41422-023-00887-z](https://doi.org/10.1038/s41422-023-00887-z) (2023). doi: [10.1038/s41422-023-00887-z](https://doi.org/10.1038/s41422-023-00887-z); PMID: 37935903
- T. Boroviak *et al.*, Lineage-Specific Profiling Delineates the Emergence and Progression of Naive Pluripotency in Mammalian Embryogenesis. *Dev. Cell* **35**, 366–382 (2015). doi: [10.1016/j.devcel.2015.10.011](https://doi.org/10.1016/j.devcel.2015.10.011); PMID: 26555056
- M. Kinoshita *et al.*, Capture of Mouse and Human Stem Cells with Features of Formative Pluripotency. *Cell Stem Cell* **28**, 453–471.e8 (2021). doi: [10.1016/j.stem.2020.11.005](https://doi.org/10.1016/j.stem.2020.11.005); PMID: 33271069
- D. Acampora, L. G. Di Giovannantonio, A. Simeone, Otx2 is an intrinsic determinant of the embryonic stem cell state and is required for transition to a stable epiblast stem cell condition. *Development* **140**, 43–55 (2013). doi: [10.1242/dev.085290](https://doi.org/10.1242/dev.085290); PMID: 23154415
- C. Buecker *et al.*, Reorganization of enhancer patterns in transition from naive to primed pluripotency. *Cell Stem Cell* **14**, 838–853 (2014). doi: [10.1016/j.stem.2014.04.003](https://doi.org/10.1016/j.stem.2014.04.003); PMID: 24905168
- S. H. Yang *et al.*, ZIC3 Controls the Transition from Naive to Primed Pluripotency. *Cell Rep.* **27**, 3215–3227.e6 (2019). doi: [10.1016/j.celrep.2019.05.026](https://doi.org/10.1016/j.celrep.2019.05.026); PMID: 31189106
- F. Nakaki *et al.*, Induction of mouse germ-cell fate by transcription factors *in vitro*. *Nature* **501**, 222–226 (2013). doi: [10.1038/nature12417](https://doi.org/10.1038/nature12417); PMID: 23913270
- J. Ernst *et al.*, Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* **473**, 43–49 (2011). doi: [10.1038/nature09906](https://doi.org/10.1038/nature09906); PMID: 21441907
- M. Thomson *et al.*, Pluripotency factors in embryonic stem cells regulate differentiation into germ layers. *Cell* **145**, 875–889 (2011). doi: [10.1016/j.cell.2011.05.017](https://doi.org/10.1016/j.cell.2011.05.017); PMID: 21663792
- R. Argelaguet *et al.*, Multi-omics profiling of mouse gastrulation at single-cell resolution. *Nature* **576**, 487–491 (2019). doi: [10.1038/s41586-019-1825-8](https://doi.org/10.1038/s41586-019-1825-8); PMID: 31827285
- I. Muñoz-Sanjuán, A. H. Brivanlou, Neural induction, the default model and embryonic stem cells. *Nat. Rev. Neurosci.* **3**, 271–280 (2002). doi: [10.1038/nrn786](https://doi.org/10.1038/nrn786); PMID: 11967557
- P. Karagianni, P. Moulos, D. Schmidt, D. T. Odom, I. Talianidis, Bookmarking by Non-pioneer Transcription Factors during Liver Development Establishes Competence for Future Gene Activation. *Cell Rep.* **30**, 1319–1328.e6 (2020). doi: [10.1016/j.celrep.2020.01.006](https://doi.org/10.1016/j.celrep.2020.01.006); PMID: 32023452
- Y. I. Yeom *et al.*, Germ-line regulatory element of Oct-4 specific for the totipotent cycle of embryonic cells. *Development* **122**, 881–894 (1996). doi: [10.1242/dev.122.3.881](https://doi.org/10.1242/dev.122.3.881); PMID: 8631266
- J. Nichols *et al.*, Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* **95**, 379–391 (1998). doi: [10.1016/S0092-8674\(00\)81769-9](https://doi.org/10.1016/S0092-8674(00)81769-9); PMID: 9814708
- D. J. Rodda *et al.*, Transcriptional regulation of nanog by OCT4 and SOX2. *J. Biol. Chem.* **280**, 24731–24737 (2005). doi: [10.1074/jbc.M502573200](https://doi.org/10.1074/jbc.M502573200); PMID: 15860457
- J. L. Chew *et al.*, Reciprocal transcriptional regulation of *Pou5f1* and *Sox2* via the Oct4/Sox2 complex in embryonic stem cells. *Mol. Cell Biol.* **25**, 6031–6046 (2005). doi: [10.1128/MCB.25.14.6031-6046.2005](https://doi.org/10.1128/MCB.25.14.6031-6046.2005); PMID: 15988017

53. T. Frum *et al.*, Oct4 cell-autonomously promotes primitive endoderm development in the mouse blastocyst. *Dev. Cell* **25**, 610–622 (2013). doi: [10.1016/j.devcel.2013.05.004](https://doi.org/10.1016/j.devcel.2013.05.004); pmid: [23747191](https://pubmed.ncbi.nlm.nih.gov/23747191/)
54. G. C. Le Bin *et al.*, Oct4 is required for lineage priming in the developing inner cell mass of the mouse blastocyst. *Development* **141**, 1001–1010 (2014). doi: [10.1242/dev.096875](https://doi.org/10.1242/dev.096875); pmid: [24504341](https://pubmed.ncbi.nlm.nih.gov/24504341/)
55. K. Mitsui *et al.*, The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell* **113**, 631–642 (2003). doi: [10.1016/S0092-8674\(03\)00393-3](https://doi.org/10.1016/S0092-8674(03)00393-3); pmid: [12787504](https://pubmed.ncbi.nlm.nih.gov/12787504/)
56. R. Stadhouders, G. J. Filion, T. Graf, Transcription factors and 3D genome conformation in cell-fate decisions. *Nature* **569**, 345–354 (2019). doi: [10.1038/s41586-019-1182-7](https://doi.org/10.1038/s41586-019-1182-7); pmid: [31092938](https://pubmed.ncbi.nlm.nih.gov/31092938/)
57. D. Solter, B. B. Knowles, Immunosurgery of mouse blastocyst. *Proc. Natl. Acad. Sci. U.S.A.* **72**, 5099–5102 (1975). doi: [10.1073/pnas.72.12.5099](https://doi.org/10.1073/pnas.72.12.5099); pmid: [1108013](https://pubmed.ncbi.nlm.nih.gov/1108013/)
58. S. Picelli *et al.*, Full-length RNA-seq from single cells using Smart-seq2. *Nat. Protoc.* **9**, 171–181 (2014). doi: [10.1038/nprot.2014.006](https://doi.org/10.1038/nprot.2014.006); pmid: [24385147](https://pubmed.ncbi.nlm.nih.gov/24385147/)
59. S. M. Harrison, S. L. Dunwoodie, R. M. Arkell, H. Lehrach, R. S. Beddington, Isolation of novel tissue-specific genes from cDNA libraries representing the individual tissueconstituents of the gastrulating mouse embryo. *Development* **121**, 2479–2489 (1995). doi: [10.1242/dev.121.8.2479](https://doi.org/10.1242/dev.121.8.2479); pmid: [7671812](https://pubmed.ncbi.nlm.nih.gov/7671812/)
60. Y. Zhang *et al.*, Dynamic epigenomic landscapes during early lineage specification in mouse embryos. *Nat. Genet.* **50**, 96–105 (2018). doi: [10.1038/s41588-017-0003-x](https://doi.org/10.1038/s41588-017-0003-x); pmid: [29203909](https://pubmed.ncbi.nlm.nih.gov/29203909/)
61. P. J. Skene, J. G. Henikoff, S. Henikoff, Targeted *in situ* genome-wide profiling with high efficiency for low cell numbers. *Nat. Protoc.* **13**, 1006–1019 (2018). doi: [10.1038/nprot.2018.015](https://doi.org/10.1038/nprot.2018.015); pmid: [29651053](https://pubmed.ncbi.nlm.nih.gov/29651053/)
62. B. Zhang *et al.*, Allelic reprogramming of the histone modification H3K4me3 in early mammalian development. *Nature* **537**, 553–557 (2016). doi: [10.1038/nature19361](https://doi.org/10.1038/nature19361); pmid: [27626382](https://pubmed.ncbi.nlm.nih.gov/27626382/)
63. J. Wu *et al.*, Chromatin analysis in human early development reveals epigenetic transition during ZGA. *Nature* **557**, 256–260 (2018). doi: [10.1038/s41586-018-0080-8](https://doi.org/10.1038/s41586-018-0080-8); pmid: [29720659](https://pubmed.ncbi.nlm.nih.gov/29720659/)

#### ACKNOWLEDGMENTS

We are grateful to members of the Xie laboratory for discussions and comments during the SOX2 study and the preparation of the manuscript and to the Animal Center and Biocomputing Facility at Tsinghua University for their support. We are grateful to Q. Sun from the Institute of Zoology, Chinese Academic of Sciences, for providing us with the Stra8-cre mice. **Funding:** This work was funded by the National Key R&D Program of China (2021YFA1100102 to W.X.); the National Natural Science Foundation of China (31988101 and 31830047 to W.X.); the National Key R&D Program of China (2019YFA0508900 to W.X.); and the Tsinghua-Peking Center for Life Sciences (to W.X.) M.A.H., T.F., and A.R. have been supported by NIH R35 GM131759 and R01 HD108722. W.X. is a recipient of an HHMI International Research Scholar and a New Cornerstone Investigator. **Author contributions:** L.Li, F.L., and W.X. conceived and designed the project. L.Li performed the CUT&RUN and RNA-seq experiments. F.L. performed embryo experiments with help from X.L., Z.L., L.Liu, and Y.X. F.L. performed microinjection and immunostaining with the help of L.Liu. L.Li performed dTAG and EpiLC differentiation experiments. X.H. performed STAR ChIP-seq and constructed plasmids for reporter assay and rescue experiments.

B.L. and L.Li performed ATAC-seq experiments. L.Li performed the bioinformatics analysis with the help of B.L. and F.C. L.Li and W.X. prepared most figures and wrote the manuscript with help from all authors. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** The datasets generated and analyzed during this study are available in the Gene Expression Omnibus (GEO) database under accession no. GSE203194. Accession codes of the published data in GEO used in this study are as follows: RNA-seq and ATAC-seq of early embryos, GSE66390; RNA-seq and ATAC-seq of postimplantation epiblast and ectoderm, GSE125318; H3K27me3 of 2-cell, GSE76687; scRNA-seq of E4.5 epiblast and PrE, GSE159030; morula DNase-seq data, GSE92605; SOX2/NANOG ChIP of serum mESC, GSM2417143; H3K4me1 and OCT4/OTX2 ChIP of mESC and EpiLCs, GSM1355167; Hi-C of ESCs, GSE118911; Hi-C of EpiLCs, GSE183828; scRNA of Oct4 KO ICM, GSE159030; NR5A2 CUT&RUN and RNA-seq in the Nr5a2 knockdown embryos, GSE229740; and TFAP2C CUT&RUN and RNA-seq in the *Tfap2c* knockout mouse embryos, GSE203194. Accession code of the published data in ArrayExpress is as follows: ZIC3 ChIP-seq data in mESCs and EpiLCs, E-MTAB-7208. **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. <https://www.science.org/about/science-licenses-journal-article-reuse>

#### SUPPLEMENTARY MATERIALS

[science.org/doi/10.1126/science.ad5516](https://doi.org/10.1126/science.ad5516)

Materials and Methods

Figs. S1 to S14

References (64–77)

MDAR Reproducibility Checklist

Submitted 4 May 2023; accepted 9 November 2023  
10.1126/science.ad5516



## Multifaceted SOX2-chromatin interaction underpins pluripotency progression in early embryos

Lijia Li, Fangnong Lai, Xiaoyu Hu, Bofeng Liu, Xukun Lu, Zili Lin, Ling Liu, Yunlong Xiang, Tristan Frum, Michael A. Halbisen, Fengling Chen, Qiang Fan, Amy Ralston, and Wei Xie

*Science* **382** (6676), eadi5516. DOI: 10.1126/science.adi5516

### Editor's summary

Pluripotency transiently appears in early development and is controlled by pioneer transcription factors (TFs) such as OCT4 and SOX2. However, how these master TFs drive pluripotency progression in early embryos remains poorly understood. Li *et al.* investigated spatiotemporal chromatin occupancy of SOX2 in mouse early embryos across 4 days. In early blastocysts, SOX2 regulates the pluripotency program not just by opening global enhancers, but also by occupying preaccessible enhancers opened by early-stage-expressing TFs. SOX2 then redistributes and opens enhancers or primes enhancers for future activation when cells acquire naive and formative pluripotency. These data revealed flexible pioneer TF-chromatin interactions and a transitional “prepluripotency” state that connects totipotency and pluripotency. —Di Jiang

### View the article online

<https://www.science.org/doi/10.1126/science.adi5516>

### Permissions

<https://www.science.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of service](#)

---

*Science* (ISSN 1095-9203) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.

Copyright © 2023 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works