CELL BIOLOGY

Mga safeguards embryonic stem cells from acquiring extraembryonic endoderm fates

Jinzhong Qin*, Congcong Wang†, Yaru Zhu†, Ting Su†, Lixia Dong, Yikai Huang, Kunying Hao

Polycomb group (PcG) proteins form multiprotein complexes that affect stem cell identity and fate decisions by still largely unexplored mechanisms. Here, by performing a CRISPR-based loss-of-function screen in embryonic stem cells (ESCs), we identify PcG gene *Mga* involved in the repression of endodermal transcription factor *Gata6*. We report that deletion of *Mga* results in peri-implantation embryonic lethality in mice. We further demonstrate that *Mga*-null ESCs exhibit impaired self-renewal and spontaneous differentiation to primitive endoderm (PE). Our data support a model in which Mga might serve as a scaffold for PRC1.6 assembly and guide this multimeric complex to specific genomic targets including genes that encode endodermal factors Gata4, Gata6, and Sox17. Our findings uncover an unexpected function of Mga in ESCs, where it functions as a gatekeeper to prevent ESCs from entering into the PE lineage by directly repressing expression of a set of endoderm differentiation master genes.

INTRODUCTION

Mouse preimplantation development is a process of progressive cell fate specification. At the time of implantation in the developing embryo, three distinct cell lineages are present in the implanting blastocyst: the extraembryonic trophectoderm (TE), the primitive endoderm (PE, also called the hypoblast), and the embryonic epiblast (EPI). These three cell types are specified by two consecutive rounds of binary cell fate decisions. During the first cell fate decision, TE, the precursor lineage of the placenta, is segregated from inner cell mass (ICM) in a transition stage between morula and blastocyst [around E3.5 (embryonic day 3.5) in mouse]. The pluripotent ICM will be further separated in the second lineage segregation into the extraembryonic PE that predominantly forms the extraembryonic endoderm (ExEn) layers of the visceral and parietal yolk sacs, and the pluripotent EPI that will later give rise to the entire fetus. Although trophoblast stem (*1*) cells, embryonic stem (ES) (*2*, *3*) cells, and extraembryonic endoderm stem (XEN) (*4*) cells have been established from the three tissue lineages present in the mouse blastocyst, the mechanism (s) initiating these cell fate decisions remains largely unknown (*5*). The second lineage decision has been proposed to be mediated by reciprocal repression between Gata6 and Nanog, the earliest markers of the PE and EPI lineages, respectively (*6*). *Gata6*-deficient embryos fail to specify the PE in the blastocyst (*7*), while forced expression of *Gata6* in ES cells (ESCs) triggers their differentiation to PE cells (*8*). Furthermore, *Nanog*-null ICM fails to form a proper EPI and is prone to differentiate into parietal endoderm-like cells (*6*). A recent replication of these experiments, however, suggest that ICMs deficient in Nanog are only capable of TE differentiation or else undergo apoptosis (*9*). Therefore, more work is needed to decipher the molecular mechanisms governing the specification of pluripotent EPI and PE lineages within the ICM of the early blastocyst.

Polycomb group (PcG) proteins are a group of evolutionarily conserved proteins that are essential for embryonic development and stem cell maintenance and are often deregulated in various cancers (*10*, *11*). They broadly assemble into two multimeric, chromatin-associated protein complexes in mammals, known as PRC2 (polycomb repressive complex 2) and PRC1, and are primarily responsible for the epigenetic repression of target genes. Ezh1/2 are the catalytically active component of PRC2 and, together with Suz12 and Eed, establish and maintain the di- and trimethylation of the N-terminal tail of histone H3 at Lys^{27} (H3K27me2 and H3K27me3) (*11*). Compared to PRC2, the composition and patterns of assembly of PRC1 are far more diverse. Six major mammalian PRC1 complexes (named PRC1.1 to PRC1.6) have been identified, and each contains one of six distinct, mutually exclusive members of the Pcgf protein family (Pcgf1 to Pcgf6), one of two RING finger E3 ubiquitin ligases (namely, Ring1b and Ring1a), and a unique set of ancillary partners with diverse functional properties (*12*). Canonical PRC1 is constituted of Ring1a/b, Phc (Phc1/2/3), Cbx (Cbx2/4/6/7/8), and either Pcgf2 or Pcgf4 referred to as PRC1.2 or PRC1.4, respectively. Canonical PRC1 complexes mediate cross-talk with PRC2 through recognition of PRC2-mediated H3K27me3 by a chromodomain in the Cbx proteins. By contrast, noncanonical PRC1s (PRC1.1, PRC1.3, PRC1.5, and PRC1.6) contain Rybp or Yaf2 instead of a Cbx component and one of the four Pcgf proteins (Pcgf1, Pcgf3, Pcgf5, and Pcgf6) and therefore are targeted to chromatin independently of H3K27me3 (*12*, *13*).

Mga (Max's giant associated protein), the largest and poorly understood member of the Max-interacting transcription factor network (*14*), is a dual-specificity transcription factor containing a T-domain DNA binding motif in its N-terminal half and a second DNA binding domain, known as the basic helix-loop-helix zipper (bHLHZip) domain, in its C-terminal half. Biochemical studies showed that Mga forms heterodimers with Max presumably through its bHLHZip domain and binds the DNA sequence CACGTG (E-box) present in the promoter (*14*). The T-domain of Mga is closely related to that of the Brachyury/Tbx proteins, a family known to play important roles in the early development of vertebrates (*15*). It has recently been reported that *Mga* is crucial for survival of mouse pluripotent ICM cells during peri-implantation development and for the growth of ESCs in culture (*16*). Moreover, somatic *Mga* loss-of-function mutations have recently been identified in a variety of tumor or cancers including chronic lymphocytic leukemia and lung cancer (*17*, *18*), suggesting that Mga might act as a tumor suppressor. Mga was initially purified together with several well-known epigenetic factors such as Pcgf6, Ring1a/b, Yaf2, L3mbtl2, Cbx3, G9a (EHMT2), and Glp (EHMT1), as well as sequence-specific DNA binding proteins

State Key Laboratory of Pharmaceutical Biotechnology and MOE Key Laboratory of Model Animals for Disease Study, Model Animal Research Center, School of Medicine, Nanjing University, Nanjing, China.
†These authors contributed equally to this work.
*Corresponding author. Email: qinjz@nju.edu.cn

such as E2F6, Max, and TFDP1, an assembly annotated as E2F6.com (*19*). Subsequently, most components of this complex were found to be associated with L3mbtl2 in ESCs (*20*), and this complex was classified into a variant PRC1 subgroup referred to as PRC1.6 in somatic cells (*12*). In line with this, loss-of-function studies in ESCs highlight essential roles for the corresponding PRC1.6 subunits in pluripotency, proliferation, and differentiation. Most substantial, as mentioned previously, Mga deficiency leads to the massive death of pluripotent ICM cells in vivo and in vitro (*16*). Also, ESCs depleted of Max and L3mbtl2 as well as Pcgf6 display impaired proliferation and differentiation (*20*–*22*) but less severe as *Mga*-null ESCs (*16*). Their functional significance is also evident from in vivo studies. Mice lacking unique subunits of PRC1.6, Max, Mga, and L3mbtl2 exhibit early embryonic lethality (*16*, *20*, *23*). Furthermore, loss of *Pcgf6* in mice results in partially penetrant embryonic lethality (*24*, *25*). While these studies clearly suggest the critical roles for PRC1.6 subunits in the early cell fate decisions, precisely how this occurs has yet to be elucidated.

Here, we perform a CRISPR-based loss-of-function screen of the PcG superfamily to identify PcG genes involved in the repression of *Gata6* gene expression in mouse ESCs, and the *Mga* gene is identified as a strong candidate. We show that disruption of *Mga* results in peri-implantation embryonic lethality. Using *Mga* conditional knockout ESC lines, we show further that Mga is essential for ESC self-renewal and pluripotency. *Mga*-null ESCs lose their characteristic morphological appearance and undergo precocious commitment toward the PE lineage even in conditions that otherwise favor their pluripotency. Mechanistically, Mga directly represses the expression of endodermal genes such as G*ata4*, *Gata6*, and *Sox17*, and *Mga*-null phenotypes can be partially rescued in ESCs lacking *Gata4* or *Gata6*. Together, our data provide strong evidence for a critical and unique role of Mga in safeguarding ESC identity by preventing induction of ExEn fates.

RESULTS

Mga is the most prominent PcG family member in the transcriptional repression of Gata6 gene in ESCs

It has been proposed that Oct4, Sox2, and Nanog form the core of a transcription factor network that supports ESC pluripotency and self-renewal while simultaneously repressing genes that promote differentiation (*26*). Although the functions of this network in sustaining pluripotency is well established (*6*, *27*–*29*), its role in ExEn differentiation is more poorly understood. To study the role of this network in ExEn differentiation, we attempted to generate *Oct4*, *Sox2*, or *Nanog* knockout ESCs (Fig. 1A). A *Sox2*- or *Oct4*-null ESC line could not be established, confirming the role of Oct4 and Sox2 in maintaining the self-renewal and pluripotency capacities of ESCs as previously reported (*27*, *28*). In contrast, *Nanog*-null ES cell lines can readily be propagated indefinitely in an undifferentiated state (Fig. 1, A and B). The $Nanog^{\Delta/\Delta}$ ESCs were morphologically normal and continued to express ESC-specific markers except Nanog itself (Fig. 1B) but had an extended population doubling time (~15.5 hours) relative to their wild-type counterparts (~12.5 hours). Western blot failed to detect the protein expression of the PE markers (Fig. 1C). Consistently, when injected subcutaneously into immunodeficient mice, $Nanog^{\Delta/\Delta}$ ESCs produced teratomas containing tissues representative of the three germ layers (Fig. 1D). These findings indicate that *Nanog*-null ESCs do not undergo precocious differentiation to PE.

Given the importance of PcG in repression of developmental genes (*30*), we explore the role of PcG in repressing endoderm fate in ESCs. To this end, we first screened for factors in the PcG superfamily (Fig. 1E and fig. S1), whose deficiency resulted in up-regulation of *Gata6*, the master transcription regulator central to ExEn formation. We found that CRISPR-mediated knockout of *Mga* was coupled with a marked induction of *Gata6* expression (Fig. 1F), as evidenced by reverse transcription quantitative polymerase chain reaction (RT-qPCR). In addition to *Mga*, the screen has identified other members of the PcG superfamily (namely, *Eed*, *Ring1b*, *Pcgf6*, *L3mbtl2*, and *Max*), although their inactivation was accompanied by a moderate induction of *Gata6*. Four gene hits (*Mga*, *Pcgf6*, *L3mbtl2*, and *Max*) from our CRISPR screen are unique members of the PRC1.6 complex. Thus, we decided to further investigate Mga's potential role as suppressor of endoderm differentiation in ESCs.

Mga deficiency results in peri-implantation embryonic lethality

We generated a conditional *Mga* knockout construct in which both exons 16 and 17 were flanked by loxP sites. Excision of these exons by Cre recombinase is predicted to generate an in-frame termination codon, enabling conditional *Mga* inactivation (fig. S2A). Heterozygous *Mga* mice were viable and fertile and appeared phenotypically normal as observed over 2 years. Heterozygous *Mga* mice were mated, and genotypes of newborn offspring were determined by PCR analysis of biopsied tail samples (fig. S2, B and C). Of 108 newborn animals genotyped from 22 independent litters, 38 were $Mga^{+/+}$, 70 were $Mga^{+/\Delta}$, and none were $Mga^{\Delta/\Delta}$ (fig. S2D), indicating a Mendelian segregation ratio typical for a prenatal lethality (1:2:0). To pinpoint the time of death in utero, embryos obtained from heterozygous intercrosses were dissected and genotyped at various gestational times. As shown in fig. S2D, embryos at different days of gestation that could be genotyped were either wild type or heterozygous for the *Mga* allele. Notably, the resorbed embryos in the empty decidua had very little embryonic material for genotyping and therefore were assumed to be $Mga^{\Delta/\Delta}$ on the basis of the Mendelian distribution. PCR genotyping of the blastocysts (E3.5) directly from heterozygous intercrosses by flushing uteri revealed that $Mga^{\Delta/\Delta}$ blastocysts were present at nearly the expected Mendelian ratio (fig. S2D). Furthermore, $Mga^{\Delta/\Delta}$ blastocysts appeared morphologically identical to the wild-type control. Our ability to detect $Mga^{\Delta/\Delta}$ blastocysts isolated in utero suggests that $Mga^{\Delta/\Delta}$ blastocysts are capable of implantation but die shortly thereafter. This is consistent with previous findings that $Mga^{\Delta/\Delta}$ embryos initially formed both EPI and PE but lost the EPI by apoptosis (*16*). As shown in fig. S2E, in contrast to the normal morphogenic events seen for their littermate controls at E7.5, the $Mga^{\Delta/\Delta}$ embryos were significantly reduced in size and developmentally retarded. These mutant embryos failed to form the amniotic, exocoelomic, and ectoplacental cavities and only consisted of an abnormal, unstructured mass of irregular cells. These results indicate that disruption of *Mga* in mice results in early embryonic lethality at the peri-implantation stage.

Mga is required for maintenance of the pluripotent state in ESCs

In an attempt to understand the cellular basis of the embryonic lethality observed in *Mga* mutant mice, we established conditional $Mga^{F/F}$ ESCs by flanking exons 16 and 17 on both alleles with loxP sites (Fig. 2A). $Mga^{F/F}$ ESCs exhibited compact and dome-shaped

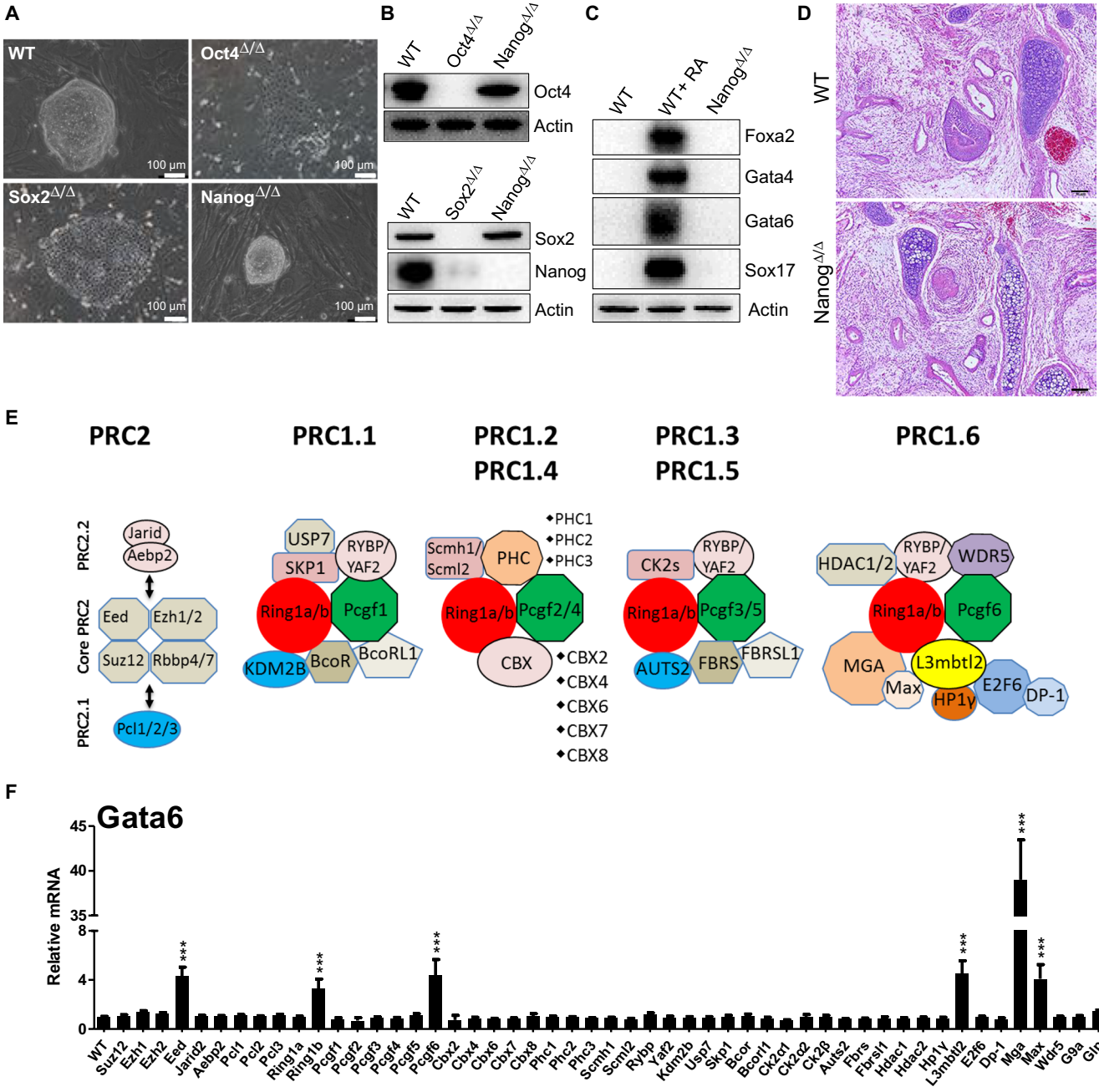

Fig. 1. Effect of knockout of PcG superfamily genes on the expression of *Gata6* gene in ESCs. (**A**) Representative phase-contrast images of ESC lines of the indicated genotypes grown on a feeder layer of mouse embryo fibroblasts (MEFs). WT, wild type. (**B**) Western blot showing the changes in the levels of Oct4, Sox2, and Nanog in ESCs of indicated genotypes. β-Actin was used as a loading control. (**C**) Western blot analysis of endodermal markers in ESCs of indicated genotypes. As a positive control, ESCs were treated with 5 mM all-trans retinoic acid (RA) for 14 days. β-Actin served as a loading control. (**D**) Hematoxylin and eosin (H&E) staining of representative teratomas from ESCs of the indicated genotypes showing derivatives of all three embryonic germ layers. Scale bars, 50 μm. Shown is a representative of four injected mice. (**E**) Schematic view of the two main mammalian PcG complexes, PRC1 and PRC2. (**F**) Effect of CRISPR-mediated knockout of PcG superfamily genes on the expression of *Gata6* gene in ESCs. Each value was normalized to actin expression, and the expression level of Gata6 in wild-type ESCs was arbitrarily set to 1. Data are shown as the means ± SD for triplicate analysis. ***$P < 0.001$.

colony morphology typical of undifferentiated ESC colonies and were morphologically indistinguishable from wild-type ESCs cultured under the same conditions (Fig. 2B). These cells were infected with a cre-expressing lentivirus and plated onto mitotically inactive mouse embryo fibroblast (MEF) feeders for colony formation. As shown in Fig. 2B, the *Mga* gene deficiency ($Mga^{\Delta/\Delta}$) ESCs yielded extremely small colonies of dispersed cells. Cell survival defects were partially rescued by exogenous putrescine, the end-product of ornithine decarboxylase, consistent with the importance of Mga in the regulation of polyamine homeostasis in ESCs (*16*). The proliferative

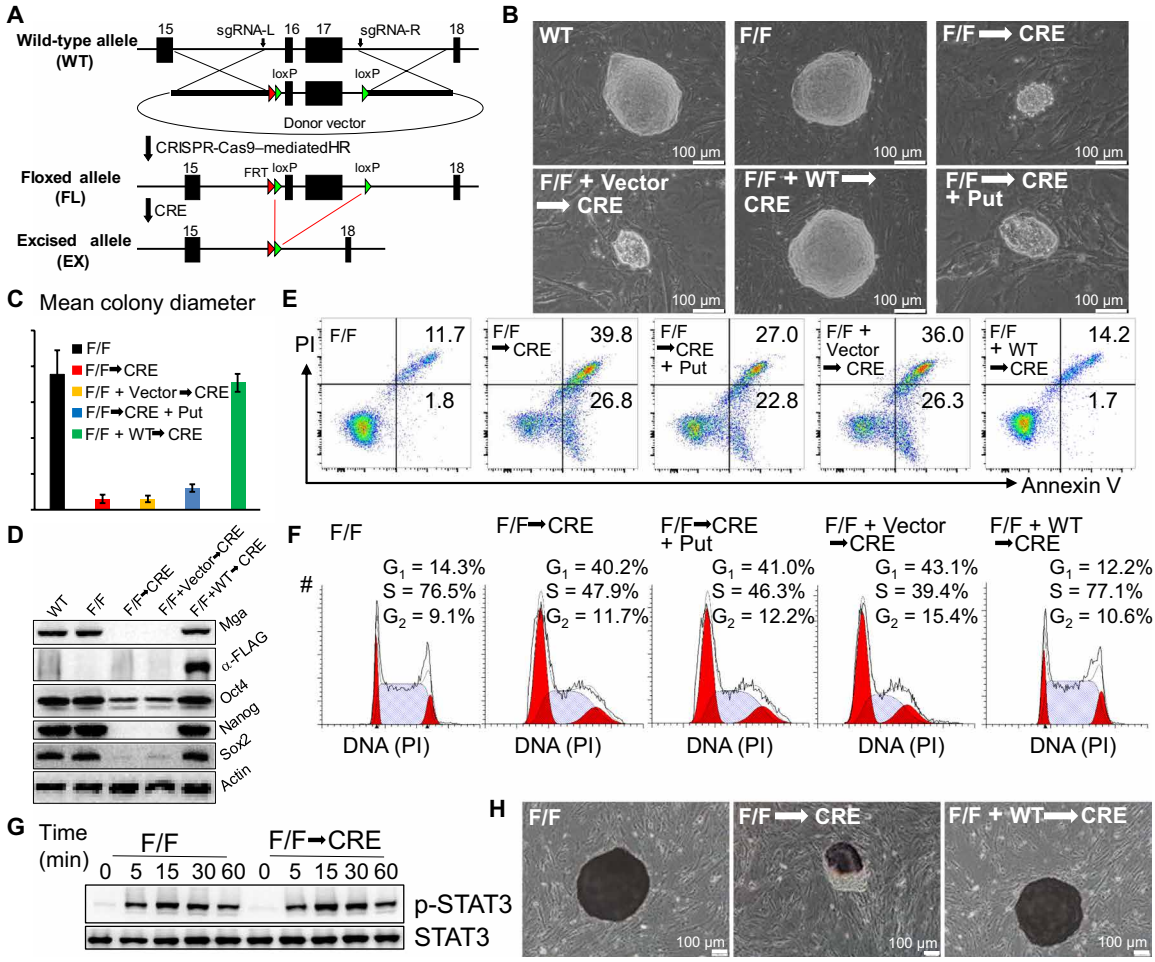

Fig. 2. Mga plays pivotal roles in ESC self-renewal and maintenance. (**A**) Schematic overview of the strategy to generate a *Mga* conditional allele using the Cre/loxP system. HR, homologous recombination; sgRNA, single-guide RNA. (**B**) Representative phase-contrast images of ESC lines in the presence or absence of CRE expression cultured on a feeder layer of MEFs. Shown is the colony size 7 days after seeding single-cell suspensions onto inactivated MEF feeder layers in the presence of leukemia inhibitory factor (LIF). Put, putrescine. (**C**) Bar graphs show the mean largest diameter of 100 random ESC colonies of the indicated genotypes. (**D**) Western blot demonstrating changes in the levels of Mga, Oct4, Sox2, and Nanog in ESCs of indicated genotypes. β-Actin was used as a loading control. (**E**) Flow cytometric analysis of apoptosis with annexin V/propidium iodine (PI) double staining. The percentages of annexin V–positive (apoptotic) cells are within the two right quadrants. (**F**) Representative cell cycle profiles of propidium iodide–stained ESCs of indicated genotypes obtained by flow cytometric analysis. (**G**) $Mga^{\Delta/\Delta}$ ESCs showed a normal response to LIF, as determined by phospho-STAT3 (signal transducer and activator of transcription 3) levels. (**H**) Alkaline phosphatase staining and morphological appearance of ESC colonies of indicated genotypes on feeders. Scale bars, 100 μm.

and cell-cell contact phenotypes observed in the knockout cells were specific to *Mga* deficiency, since they were rescued by the introduction of a transgenic copy of full-length wild-type *Mga* using lentivirus-mediated transduction (Fig. 2, B to D).

The impaired growth of $Mga^{\Delta/\Delta}$ ESCs might be due to altered cell cycle profile and/or increased apoptosis. To address these possibilities, we compared cell cycle progression and survival in $Mga^{\Delta/\Delta}$ and $Mga^{F/F}$ ESCs. The deletion of the *Mga* gene in ESCs resulted in massive cell death by apoptosis, as assayed by staining with fluorescein-conjugated annexin V (Fig. 2E). In addition, $Mga^{\Delta/\Delta}$ cells exhibited a lengthening of the G_1 phase and a reduction of cells in S phase (Fig. 2F). These effects associated with deletion of *Mga* gene were rescued by exogenous expression of wild-type *Mga* (Fig. 2, E and F). It appears that the impaired growth in $Mga^{\Delta/\Delta}$ cells reflect a combinatory effect of altered cell cycle and survival capabilities. Furthermore, the addition of putrescine in *Mga*-null ESCs modestly restored their viability but had no effect on cell cycle alterations (Fig. 2, E and F). In addition, $Mga^{\Delta/\Delta}$ ESCs retained the ability to activate signal transducer and activator of transcription 3 in response to leukemia inhibitory factor (LIF) (Fig. 2G). Of particular note, upon plating of cells onto a feeder layer, $Mga^{\Delta/\Delta}$ cells from passage 2 underwent spontaneous differentiation, as determined by flattened and spreading morphology and by loss of alkaline phosphatase staining (Fig. 2H). Consistently, we observed that the expression levels of the pluripotency-associated transcription factors Nanog and Sox2 in $Mga^{\Delta/\Delta}$ cells were almost undetectable (Fig. 2D). Similarly, Oct4 expression was also reduced upon *Mga* deletion, but to a lesser extent. Overall, these results indicate that Mga has an essential role in maintaining pluripotency and controlling cell fate decision in ESCs.

Loss of *Mga* unleashes precocious differentiation of ESCs

To determine why *Mga*-null ESCs fail to maintain pluripotency, RNA sequencing (RNA-seq) was used to screen for genes immediately

down- or up-regulated after reduction of Mga levels. For comparison, we also performed RNA-seq on ESCs lacking *Rybp*, *Pcgf6*, *L3mbtl2*, or *Max*, other components of the PRC1.6 complex. Consistent with the known function of Mga as a transcriptional repressor (*14*), after 72 hours of exposure to CRE recombinase in $Mga^{F/F}$ ESCs, RNA-seq differential gene expression analysis identified many more genes as up-regulated (1739) than down-regulated (1199) (Fig. 3A and table S1), indicating that Mga acts predominantly as a transcriptional repressor in ESCs. Gene ontology analysis of these genes identified categories associated with cellular and developmental differentiation, including pattern specification process, cell fate commitment, organ morphogenesis, and germ cell specification (Fig. 3B and table S1). Comparison of the set of deregulated genes with published genome-wide occupancy of Mga in ESCs (*31*) revealed that stronger binding was observed in the transcription start sites of up-regulated genes than down-regulated ones in $Mga^{\Delta/\Delta}$ cells. A large fraction of the Mga-bound genes were transcriptionally inactive (Fig. 3C). As expected, a large number of the genes deregulated in $Mga^{\Delta/\Delta}$ were also observed in ESCs individually deficient for *Max* (table S2), *L3mbtl2* (table S3), *Pcgf6* (table S4), or *Rybp* (*32*) and vice versa (Fig. 3D). Consistent with our cell cycle analysis (Fig. 2F), we found 76 genes related to proliferation to be differentially expressed in $Mga^{\Delta/\Delta}$ ESCs (Fig. 3E). Furthermore, among the up-regulated transcripts, we found that specific sets of genes during germ cell specification previously described to be direct targets of Rybp (*33*), L3mbt2 (*20*), Pcgf6 (*21*, *24*), and Max (*22*) were also derepressed in ESCs lacking *Mga* (Fig. 3F), consistent with them residing within the same complex. However, in contrast to the modest to moderate effect on expression of the target genes in *Rybp*-, *Pcgf6*-, *L3mbtl2*-, or *Max*-deficient cells, removal of *Mga* affected a large number of genes (Fig. 3E), most of which were involved in the differentiation process. Examination of up-regulated genes also showed a notable overrepresentation of endoderm transcripts, including master regulators of endoderm formation (such as Gata4, Gata6, and Sox17) (Fig. 3G), genes involved in cell adhesion (such as lamb2, Lrp2, Col5a2, and Col12a1), and a large number of genes associated with the function of PE and its derivatives, visceral and parietal endoderm (such as Sdc4, Sparc, Hnf1b, Pax9, and Lhx1). In contrast, expression levels of key markers of ectoderm (Nestin, Pax6, Sox3, and Sox11) or mesoderm (Flk1, Brachyury [T], Hand1, Eomes, and Msx2) either decreased or remained the same in mutant cells (table S1). Together, these results suggest that Mga normally functions to maintain ESCs in an undifferentiated state by repressing differentiation toward endoderm lineages.

Mga-null ESCs display profound differentiation defects

To evaluate whether spontaneous lineage specific differentiation occurred during normal passage (serum/LIF) of ESCs upon *Mga* deletion (fig. S3, A and B), we examined the expression levels of developmentally regulated genes by RT-qPCR (fig. S3B, day 0). Marked increases in transcript levels for PE markers (Foxa2, Gata4, Gata6, and Sox17) were observed but not for the early mesoderm markers, namely, T, Emoes, and Flk1, and the primitive ectoderm markers FGF5, Nestin, and Sox11. These results further indicate that when cultured under conditions that favor ESC maintenance, $Mga^{\Delta/\Delta}$ cells undergo differentiation to PE.

To further explore the in vitro differentiation potential of $Mga^{\Delta/\Delta}$ ESCs, these cells were tested for their ability to aggregate in suspension to form embryoid bodies (EBs). To this end, ESCs at early passage (p2) were induced to differentiate by withdrawal of LIF and growing in suspension culture. As shown in fig. S3A, $Mga^{\Delta/\Delta}$ ESCs in early passage (P2) did form EBs, but they were reduced in size and less compact and more irregular in shape than those from control ESCs during differentiation. RT-qPCR analysis of these EBs that were kept in differentiation medium over 12 days of EB formation revealed that these cells were biased toward endoderm formation upon differentiation. Specifically, $Mga^{\Delta/\Delta}$ EBs showed further elevated expression of endoderm markers Gata4, Gata6, Foxa2, and Sox17 during differentiation compared with the $Mga^{F/F}$ EBs in which these markers only increased expression after day 3 (fig. S3B). Either no or only a modest up-regulation of the mesoderm markers and ectoderm marker was observed during $Mga^{\Delta/\Delta}$ EB formation (fig. S3B). $Mga^{F/F}$ cells, however, did up-regulate these other lineage markers during EB differentiation, suggesting that loss of *Mga* promotes differentiation of ESCs toward endoderm at the expense of mesoderm and ectoderm lineages. In contrast to differentiation related gene expression, pluripotency genes, namely, *Oct4*, *Nanog*, and *Sox2*, were sharply depleted during the differentiation of $Mga^{\Delta/\Delta}$ cells, while $Mga^{F/F}$ cells exhibited similar trends, but of a lesser magnitude (fig. S3B). Of particular note, $Mga^{\Delta/\Delta}$ ESCs after passage 4, however, lost their capability to aggregate into EBs. Similarly, the *Mga*-null cells failed to generate detectable teratomas upon subcutaneous injection into immune-deficient mice (fig. S3C). Collectively, our findings demonstrate that *Mga* disruption exclusively specifies ESCs fate into EXEn lineages.

Inactivation of *Mga* leads to accelerated and more uniform differentiation into the endoderm fate

To evaluate transcriptional mediators responsible for PE specification upon *Mga* deficiency, we first assessed the temporal expression of the key endoderm transcription factors *Gata4/6*, *Foxa2*, and *Sox17* in relation to the onset of endoderm specification. Real-time RT-qPCR analysis revealed that *Foxa2* was first detected as early as 2 days after Cre transfection and plateaued around day 9 (Fig. 4A). ESCs also increased levels of *Gata4* and *Gata6* expression, both of which peaked at 12 days after Cre transfection. Maximal up-regulation of *Sox17* was noted at passage 2, presumably as a consequence of increased *Foxa2* and/or *Gata4/6* expression. In parallel with increased endodermal marker expression, mRNA levels of core pluripotency factors *Oct4*, *Sox2*, and *Nanog* were initially slightly elevated (up to 6 days) but declined thereafter following *Mga* deletion in ESCs (Fig. 4A), indicating a progression toward differentiation. Western blot analysis over the differentiation period confirmed these dynamics (fig. S4A).

After extended culture of $Mga^{\Delta/\Delta}$ ESC clones under ESC self-renewal LIF/serum conditions, which are known to indefinitely maintain ESC self-renewal and pluripotency in culture, $Mga^{\Delta/\Delta}$ ESC colonies gradually changed from a compact dome-shaped to a flattened shape as cells migrated away from the center of the colony by passage 3 (Fig. 4B). By passage 3, a homogeneous population of cells with the dispersed, refractile, and stellate morphology characteristic of XEN was established (Fig. 4B) (*4*). Consistent with spontaneous differentiation toward the ExEn lineage, the loss of a tightly packed colony morphology was accompanied by loss of alkaline phosphatase staining.

The endoderm differentiation of $Mga^{\Delta/\Delta}$ ESCs was also monitored by immunofluorescence microscopy to detect the expression profile of endoderm markers alongside pluripotency factors. In $Mga^{F/F}$ ESCs cells, Nanog, Oct4, and Sox2 expression was high, whereas cells positive for endoderm markers were rarely observed (Fig. 4C and fig. S4B), indicating the undifferentiated state of the

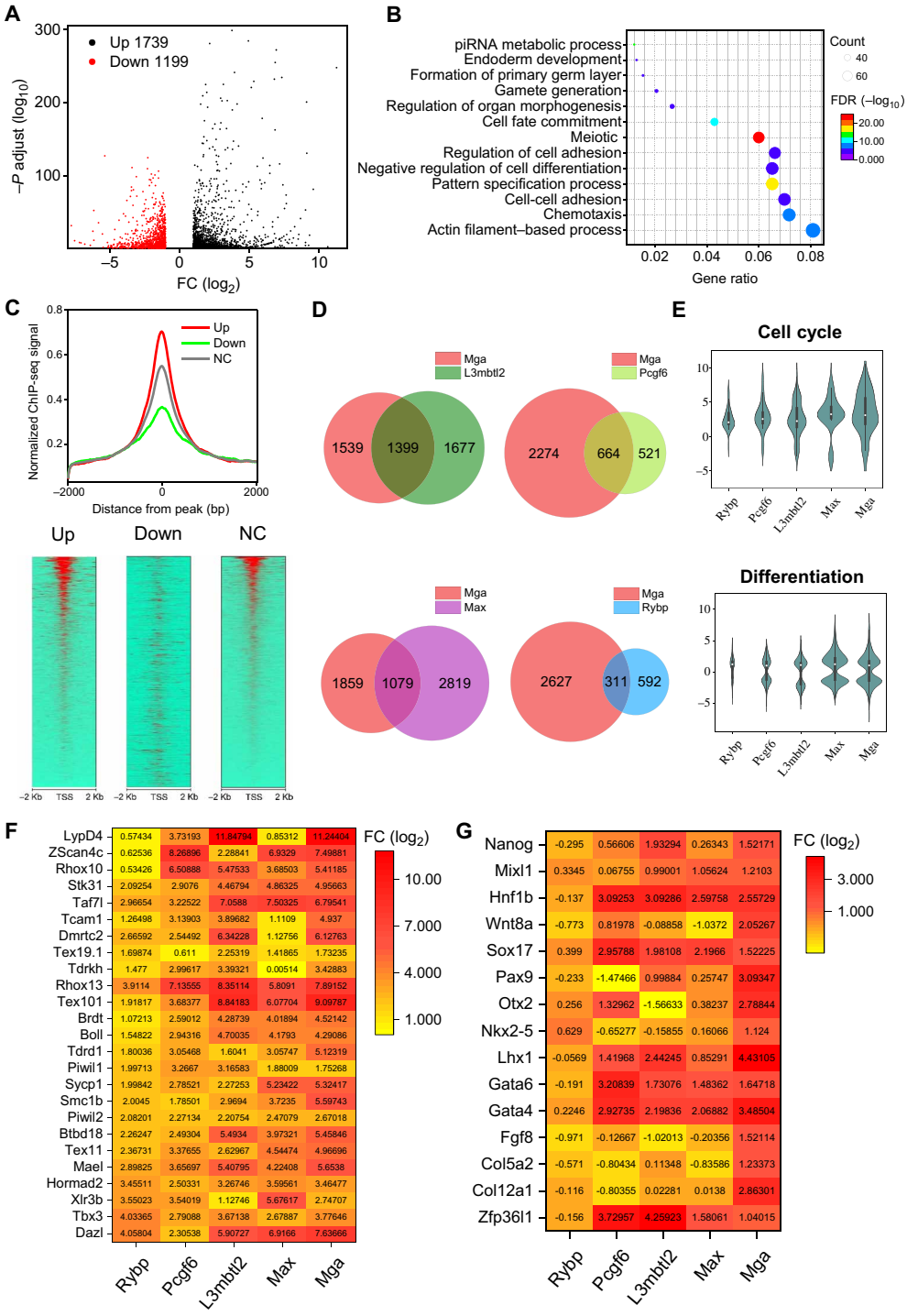

Fig. 3. Deletion of *Mga* triggers precocious differentiation of ESCs. (**A**) Volcano plots of $-\log_{10}$ (*P* value) against $\log_2$ fold change representing the differences in gene expression between $Mga^{F/F}$ and $Mga^{\Delta/\Delta}$ ESCs. FC, fold change. (**B**) Gene ontology analysis of genes up-regulated and down-regulated in $Mga^{\Delta/\Delta}$ ESCs. FDR, false discovery rate. (**C**) Average Mga binding profile for up-regulated (Up) and down-regulated (Down) genes and genes showing no change in transcript levels (NC) following ablation of the Mga gene (top) and heatmap representation of this binding (bottom), with red indicating enrichment and green denoting lack of enrichment in chromatin immunoprecipitation sequencing (ChIP-seq) sample versus input control. Notably, of 1739 genes that were up-regulated in $Mga^{\Delta/\Delta}$, 856 genes were bound by Mga on promoters or gene bodies, while 687 of 1199 down-regulated genes were Mga-bound. Published ChIP-seq data for Mga were obtained from National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) (accession number GSM1041373). bp, base pair; TSS, transcription start site. (**D**) Venn diagram showing significant overlapping between the gene lists between Mga and Rybp, Pcgf6, L3mbtl2, or Max in ESCs. RNA for RNA-seq analysis was obtained from $L3mbtl2^{F/F}$, $Pcgf6^{F/F}$, or $Max^{F/F}$ cells transfected with CRE or control vector for 72 hours in three independent experiments. Published RNA-seq data for Rybp were obtained from a published study (*32*). (**E**) A violin plot comparing $\log_2$ fold changes of genes involved in cell cycle (top) or differentiation (bottom) in ESCs deficient for *Rybp*, *Pcgf6*, *L3mbtl2*, *Max*, or *Mga*. (**F** and **G**) Heatmap representation of expression levels of genes involved in germ cell (F) or endoderm (G) specification in ESCs upon *Rybp*, *Pcgf6*, *L3mbtl2*, *Max*, or *Mga* ablation.

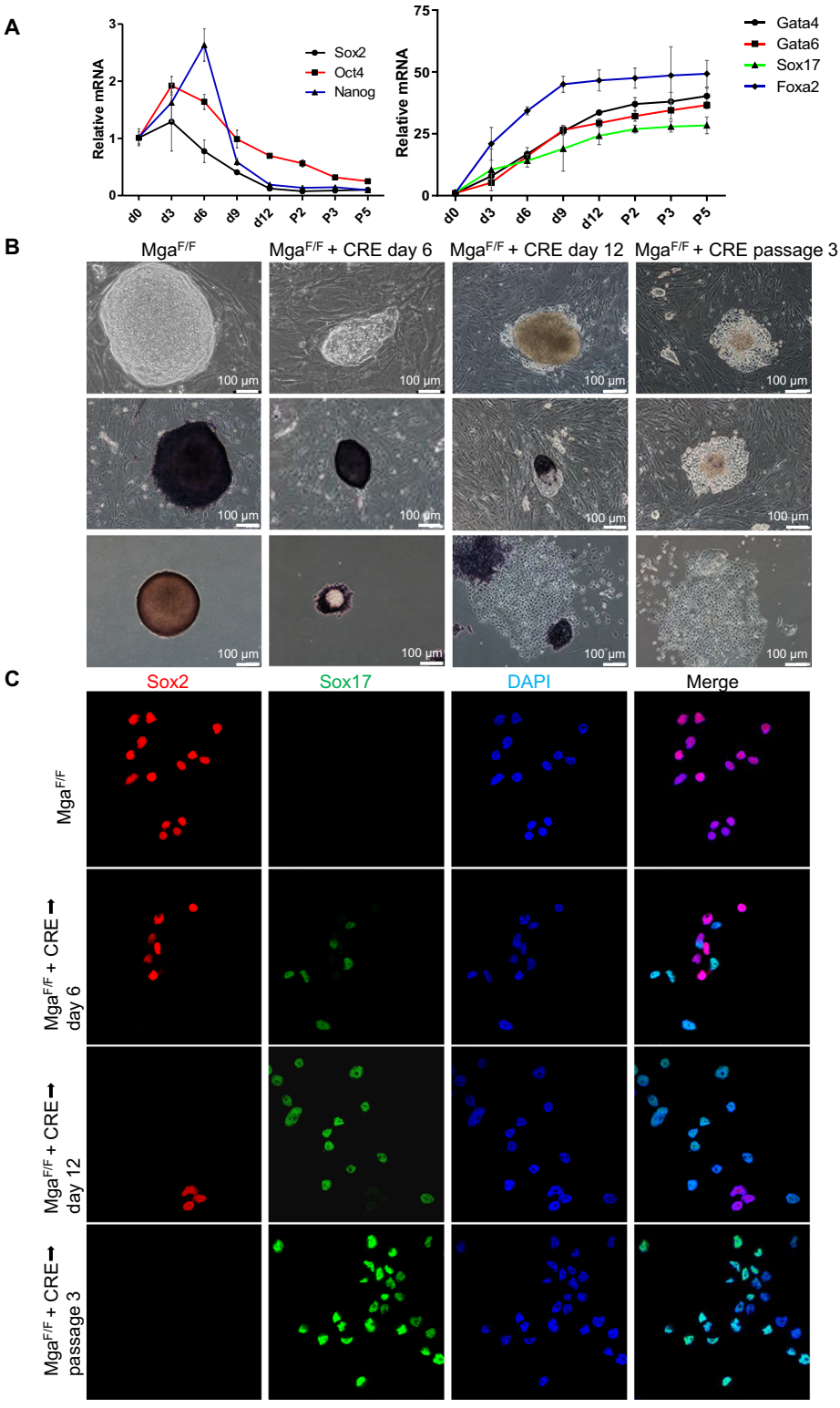

Fig. 4. *Mga* deficiency results in dynamic changes in cell morphology and gene expression characteristic of Xen cells. (**A**) RT-qPCR analysis for selected pluripotency (left) and endoderm (right) transcripts in *Mga*$^{F/F}$ ESCs after lenti-Cre infection. Relative expression is reflected as fold difference over uninfected ESCs normalized to β-actin. Data are expressed as means ± SEM of three biological replicates. d0, day 0; P2, passage 2. (**B**) Top: Phase-contrast images of *Mga*$^{F/F}$ in the presence or absence of CRE expression cultured on a feeder layer of MEFs. Middle and bottom: Images of *Mga*$^{F/F}$ in the presence or absence of CRE expression cultured on a feeder layer of MEFs (middle) or gelatin (bottom) after alkaline phosphatase staining. Scale bar, 100 μm. (**C**) Immunofluorescence analysis for Sox2 (red), Sox17 (green), or 4′,6-diamidino-2-phenylindole (DAPI) (blue) in *Mga*$^{F/F}$ or *Mga*$^{F/F}$ mES cells following lenti-Cre infection. Pictures were taken at ×63 magnification using confocal microscopy. Merge, merged images.

cells. As illustrated in Fig. 4C, Sox17 was readily detected in a manner mutually exclusive to Sox2 signal upon *Mga* depletion, further pointing to a close association between *Mga* ablation and the acquisition of PE identity. Immunofluorescence analysis also confirmed that $Mga^{\Delta/\Delta}$ cells by passage 3 ubiquitously expressed known XEN marker genes, including Gata4, Gata6, and Foxa2 proteins (fig. S4B). Moreover, $Mga^{\Delta/\Delta}$ cells after passage 3 no longer expressed Sox2 and Nanog proteins, suggesting that they have been reprogrammed to XEN cells. Although Sox2 and Nanog were mostly undetectable in $Mga^{\Delta/\Delta}$ cells after passage 3, *Oct4* mRNA and protein were present in $Mga^{\Delta/\Delta}$ cells, but at a reduced level of expression compared to the $Mga^{F/F}$ (Fig. 4A and fig. S4B). The persistence of low *Oct4* expression suggests that $Mga^{\Delta/\Delta}$ cells are more reminiscent of the primitive XEN cells (*34*). Hierarchical clustering demonstrated that genes differentially expressed in $Mga^{\Delta/\Delta}$ and in embryo-derived XEN cells (*35*) only partially overlapped (fig. S5A), suggesting that $Mga^{\Delta/\Delta}$ cells at early passages may represent an intermediate state of XEN conversion (Fig. 3G). Further chimera formation assays are ongoing to test the in vivo developmental potential of $Mga^{\Delta/\Delta}$ cells.

Functional synergy between the T-box and bHLHZip domains in Mga-mediated pluripotency

To test the importance of two known domains of Mga under physiological conditions, we have used a knock-in approach to generate mutant alleles of the *Mga* locus and ask whether they could support normal pluripotency in vivo (fig. S6, A and B). To this end, targeting vectors were constructed, in which the codons for amino acids 79 to 261 were deleted from exon 2 (named $Mga^{\Delta T\text{-box}}$) and the codons for amino acids 2376 to 2444 were deleted from exon 20 (named $Mga^{\Delta bHLHZip}$) (Fig. 5A and fig. S6, C to H). The final targeting vectors were introduced into ESCs via CRISPR-Cas9–facilitated homologous recombination. These *Mga* knock-in ESCs express mutant proteins at levels comparable to wild-type cells (fig. S6, E and H). We found that the $Mga^{\Delta T\text{-box}}$ and $Mga^{\Delta bHLHZip}$ ESCs displayed moderately reduced proliferative potential as measured by colony formation assays (Fig. 5, A and B); however, disruption of both domains led to markedly impaired proliferation of ESCs due to changes in cell cycle progression and apoptosis (Fig. 5, A and B, and fig. S6, I and J). Although these mutant cells displayed decreased ability for self-renewal, we did not observe a loss of ESC colony features and progression toward an endoderm-like phenotype over time in culture (Fig. 5A). Furthermore, to determine the roles of the T-box and bHLHZip domains in the activation of endodermal genes, we examined the mRNA and protein expression levels of the *Foxa2*, *Gata4*, *Gata6*, and *Sox17* genes in wild-type and mutant cells. Unexpectedly, double deletion of T-box and bHLHZip domains led to moderate elevation of these genes but less severe as $Mga^{\Delta/\Delta}$ ESCs (Fig. 5, C and D). These effects require cooperative activities of T-box and bHLHZip, since none of these endodermal markers exhibited significant changes in ESCs with a single deletion. Consistent with this, we found that the protein levels of Oct4, Sox2, and Nanog were maintained in ESCs containing single or combined deletions of the T-box and bHLHZip domains (Fig. 5D). $Mga^{\Delta T\text{-box}}$ still bound *Mga* target promoters as efficiently as wild-type Mga, whereas binding of the $Mga^{\Delta bHLHZip}$ to target promoters was moderately affected (Fig. 5, E and F). However, the combined loss of both DNA binding activities strongly reduced Mga and PRC1.6 recruitment to its target sites but were less severe as that seen in $Mga^{\Delta/\Delta}$ ESCs.

Consistent with these results, $Mga^{\Delta T\text{-box}}$ and $Mga^{\Delta bHLHZip}$ ESCs were shown to be pluripotent in a teratoma assay forming tumors with tissue representative of all three germ layers (Fig. 5, G and H). However, the size of $Mga^{\Delta T\text{-box}}$ and $Mga^{\Delta bHLHZip}$ teratomas were moderately reduced, whereas teratomas with double domain deletion were even more reduced in size in comparison to single deletion. Unexpectedly, despite the significant growth disadvantage of $Mga^{\Delta T\text{-box};\Delta bHLHZip}$ ESCs, they were able to form all three germ layers in teratoma assays with no obvious contribution bias toward any germ layer compared to single deletion or wild-type cells (Fig. 5H). These results suggest that both T-box and bHLHZip domains have redundant functions and independently contribute to the maintenance of ESC state. These results also indicated that disruption T-box or bHLHZip domains were not sufficient to impair Mga-mediated functions but suggested that Mga has broader structural functions responsible for ensuring proper ESC pluripotency programs.

Mga acts as a scaffold to stabilize the PRC1.6 complex

Biochemical studies have shown that Mga, Max, L3mbtl2, and Pcgf6 are unique core components of the PRC1.6 complex (Fig. 6A) (*16*, *20*–*22*, *24*). Their targeted mutations in ESCs displayed moderate or severe defects in self-renewal and differentiation (*16*, *20*–*22*), whereas knockout of other subunits displayed a mild negative effect on ESC growth (Fig. 6B), supporting that they work as core components in vivo. In addition, although Max, L3mbtl2, and Pcgf6 have been reported to cooperatively mediate the chromatin recruitment of the PRC1.6 complex (*36*), the function of Mga as a unique component of the PRC1.6 complex has not been fully elucidated. Domain deletion analysis has revealed DNA binding–independent activities of Mga in PRC1.6 complex (Fig. 5). Mga is a large protein with multiple domains involved in the direct interaction with major components of the PRC1.6 complex (*14*, *19*), which enables Mga to function as a scaffold protein within the complex. To test this possibility, we analyzed the effect of *Mga* deficiency on the protein level of the remaining components of the complex. As shown in Fig. 6C, *Mga* deletion led to a substantial reduction of the L3mbtl2, Max, and Pcgf6 proteins, whereas single knockout of other components did not appear to appreciably affect the protein levels of the remaining components of the complex. However, this reduction at the L3mbtl2, Max, and Pcgf6 protein levels was not accompanied by that of their mRNA levels (Fig. 6D). Therefore, the drop in the protein levels of L3mbtl2, Max, and Pcgf6 in $Mga^{\Delta/\Delta}$ ESCs cannot be accounted for by a change in their mRNA levels, suggesting that their protein levels are regulated posttranscriptionally. These results strongly support the notion that Mga serves as an integral component of the PRC1.6 complex and its loss results in the destabilization of the other subunits in ESCs. *Gata6* mRNA was detectable in ESCs deficient for *Pcgf6*, *L3mbtl2*, *Max*, or *Ring1b* by reverse transcription PCR (RT-PCR) (Fig. 1F) but not by Western blot analysis (Fig. 6C), indicating very weak expression of this gene. However, the cytological changes reminiscent of Xen cells and strong up-regulation of genes associated with endoderm lineages were observed in ESCs deficient for *Mga*, but not other components in the complex (Fig. 6, B and C), thereby suggesting that among the PRC1.6 family members, only Mga protects ESCs from spontaneous differentiation to ExEn fates. In further support of the notion that Mga functions as a scaffold, immunoprecipitation analysis for Ring1b in $Mga^{\Delta/\Delta}$ ESCs revealed that *Mga* ablation did not affect the association of Ring1b

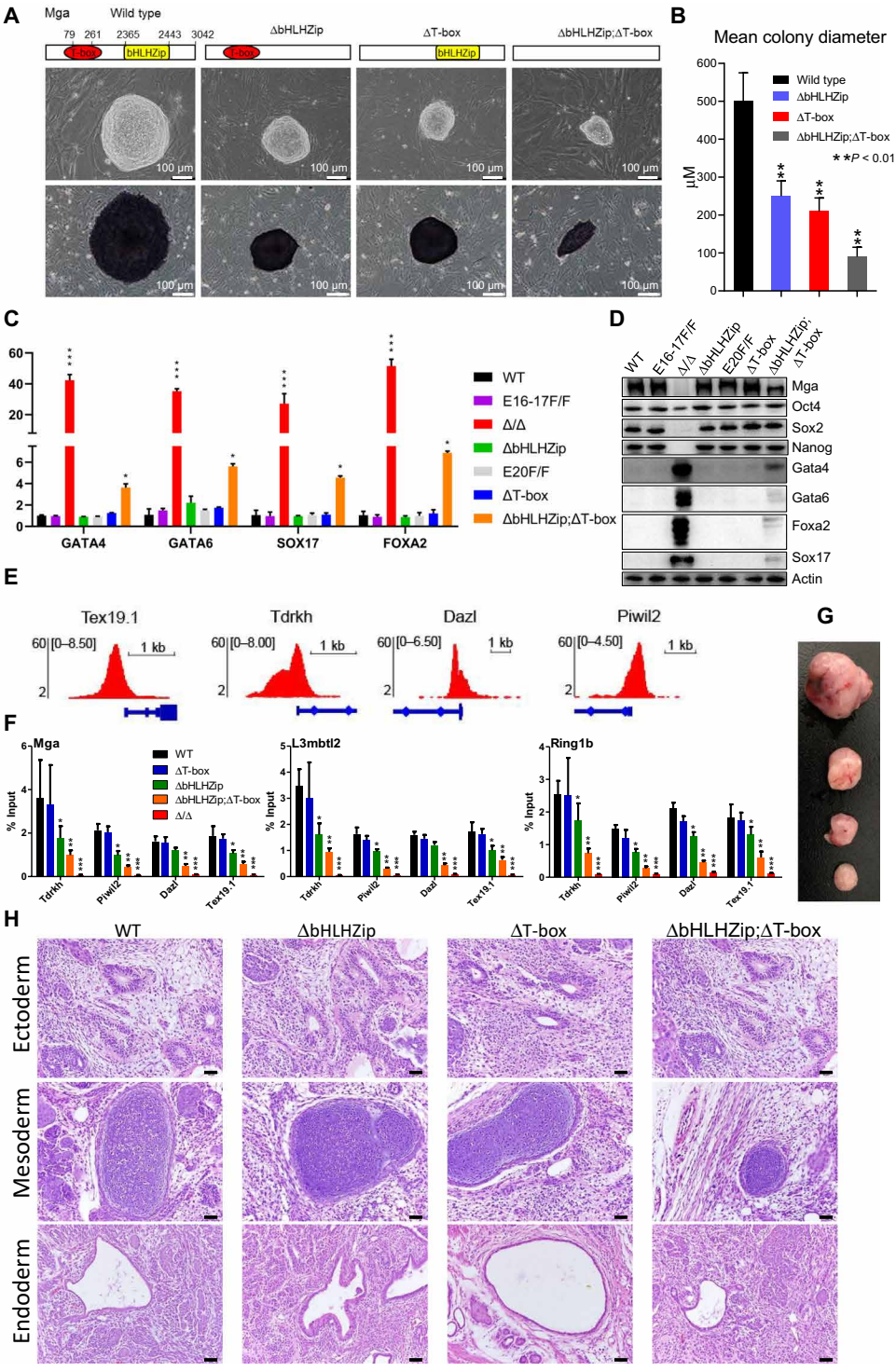

Fig. 5. The key functions of Mga lie beyond its DNA binding abilities. (**A**) The domain structures of the wild-type and mutant proteins are shown schematically at the top. Middle: Phase-contrast images of ESC colonies on MEF feeders. Bottom: Representative images of alkaline phosphatase staining of wild-type and mutant ESC colonies. (**B**) Bar graphs depict the mean largest diameter of 100 random ESC colonies of the indicated genotypes. (**C**) ChIP-qPCR analysis of selected endodermal markers in ESCs of indicated genotypes. (**D**) Western blot analysis of Mga, pluripotency factors, and endodermal markers in ESCs of indicated genotypes. β-Actin was used as a loading control. (**E**) Genomic snapshots of Mga ChIP-seq profiles at selected germline gene loci in wild-type ESCs. Published ChIP-seq data for Mga were obtained from NCBI GEO (accession number GSM1041373). (**F**) ChIP-qPCR analysis of Mga, L3mbtl2, and Ring1b binding to the promoters of a panel of selected germ cell–related target genes in wild-type, $Mga^{\Delta T\text{-box}}$, $Mga^{\Delta bHLHZip}$, $Mga^{\Delta t\text{-box};\Delta bHLHZip}$, and $Mga^{\Delta/\Delta}$ ESCs. The error bars denote SD; $n = 3$. (**G**) Teratoma formation in immunodeficiency mice by wild-type, $Mga^{\Delta bHLHZip}$, $Mga^{\Delta T\text{-box}}$, and $Mga^{\Delta bHLHZip;\Delta T\text{-box}}$ ESCs (from top to bottom). (**H**) Representative images of each embryonic germ layer from H&E staining of teratomas generated from ESCs of the indicated genotypes. Shown is a representative of four injected mice. Scale bars, 50 μm. Photo credit: (G) Jinzhong Qin, Nanjing University.

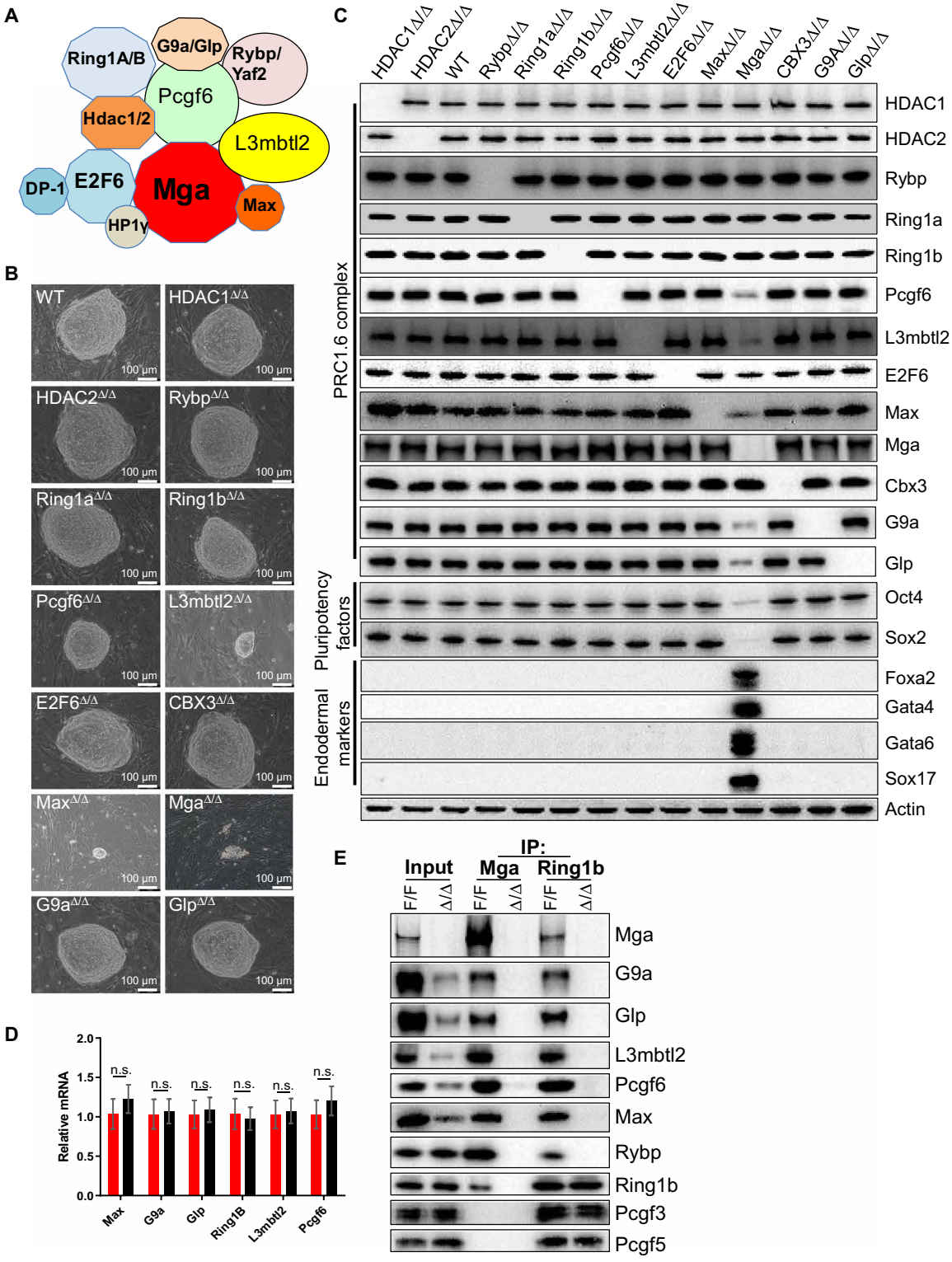

Fig. 6. Mga posttranscriptionally stabilizes major components of the PRC1.6 complex. (**A**) Schematic representation of the composition of PRC1.6 complex. (**B**) Representative phase-contrast images of colonies resulting from feeder-cultured ESCs of the indicated genotypes. (**C**) Western blot demonstrating changes in the global levels of PRC1.6 proteins, pluripotency factors, and endodermal markers in ESCs of indicated genotypes. β-Actin was used as a loading control. (**D**) Gene expression analysis of *Max*, *G9a*, *Glp*, *Ring1b*, *L3mbtl2*, and *Pcgf6* using RT-qPCR in wild-type and $Mga^{\Delta/\Delta}$ ESCs. n.s., not significant. (**E**) Western blot analysis using the indicated antibodies in Mga or Ring1b immunoprecipitations from nuclear extracts prepared from $Mga^{F/F}$ or $Mga^{\Delta/\Delta}$ ESCs. Input is shown as loading control.

with the components of other PRC1 subcomplexes Pcgf3 and Pcgf5 but that deletion of *Mga* disrupted the PRC1.6 complex (Fig. 6E). All of these results suggest that Mga functions as a scaffold protein, instead of a simple PRC1.6 recruiter, to stabilize the major components of the complex.

Mga blocks XEN lineage formation by repressing the expression of *Gata4/6* and *Sox17*

Previous studies have, in fact, shown that *Sox17*, *Gata6*, or *Gata4* overexpression is sufficient to reprogram ESCs into XEN-like cells (*8*, *37*, *38*). As described above, *Gata4/6* and *Sox17* transcripts were significantly up-regulated in $Mga^{\Delta/\Delta}$ ESCs (Figs. 3G, 4A, and 6C). Upon further examination of Mga-bound targets identified by chromatin immunoprecipitation sequencing (ChIP-seq) in ESCs, *Gata4/6*, *Sox17*, and *Foxa2* were also identified as potential targets (Fig. 7A). The binding of Mga was confirmed in an independent biological experiment by ChIP-qPCR on the promoters of these endoderm regulators (Fig. 7B). Notably, other PRC1.6 components including Pcgf6, Max, L3mbtl2, Ring1b, G9a as well as resulting H2AK119ub1 and the H3K9me2 were present at these target sites in $Mga^{F/F}$ ESCs but were displaced or, in the case of the H2AK119ub1 or H3K9me2, markedly reduced in $Mga^{\Delta/\Delta}$ ESCs, whereas deletion of Mga yielded no effect on global H2AK119ub or H3K9me2 levels (fig. S5B). These results implied that Mga targets to endodermal genes exclusively in the context of the PRC1.6 complex. These observations also suggest that Mga can suppress XEN lineage formation through regulation of *Gata4/6* and *Sox17*.

To address the question whether these endodermal transcription factors are responsible for the differentiation of pluripotent stem cells to PE in $Mga^{\Delta/\Delta}$ ESCs, we generated single or combined knockout of *Sox17*, *Foxa2*, and *Gata4/6* ESCs harboring floxed alleles of *Mga* ($Mga^{F/F}$). Notably, these mutant ESCs did not show any overt phenotype and formed compact colonies morphologically indistinguishable from $Mga^{F/F}$ (Fig. 7C). Next, all these mutants were transfected with CRE to eliminate the floxed *Mga* alleles. As shown in Fig. 7 (D and E), loss of *Gata4* or *Gata6* in $Mga^{\Delta/\Delta}$ ESCs resulted in a subtle, but noticeable induction in expression of the endodermal genes as evidenced by Western blot and immunofluorescence analysis, and these mutant cells had a morphology very similar to that observed in $Mga^{F/F}$ ESCs (Fig. 7C). These phenotypes were also shared with $Mga^{\Delta/\Delta}$;$Sox17^{\Delta/\Delta}$, indicating a rescue of the endoderm lineage differentiation in the absence of *Sox17*, *Gata4*, or *Gata6* (Fig. 7C). Astonishingly, $Mga^{\Delta/\Delta}$;$Gata4/6^{\Delta/\Delta}$ ESCs formed colonies that morphologically resembled the control undifferentiated ESCs. Accordingly, combined knockout of *Gata4* and *Gata6* completely eliminated the up-regulation of the endoderm markers Sox17 and Foxa2 in $Mga^{\Delta/\Delta}$ ESCs, consistent with retention of pluripotency markers Oct4, Nanog, and Sox2 (Fig. 7, D and E), suggesting that there is functional redundancy between *Gata4* and *Gata6* in endoderm lineage differentiation. In contrast, deletion of *Foxa2* in $Mga^{\Delta/\Delta}$ ESCs resulted in a morphological change similar to that observed in $Mga^{\Delta/\Delta}$ ESCs (Fig. 7C). These morphological changes were accompanied by the normal patterns of Gata4/6 and Sox17 up-regulation and reduced to undetectable levels of pluripotency markers Nanog and Sox2, evaluated by Western blot and immunofluorescence staining (Fig. 7, D and E), indicating that *Foxa2* has a marginal impact on Mga-mediated pluripotency maintenance of ESCs. These findings place Mga directly upstream of Gata4/6 and Sox17, where it negatively regulates the differentiation of ESCs into the endoderm lineage by repressing these master regulators of endoderm differentiation.

DISCUSSION

Specification of the EPI and PE lineages represents the early cell fate changes that occur in the developing mammalian embryo. However, the lineage-specific transcription factors driving these cell fate decisions remain poorly defined. It has previously been reported that *Nanog*-null ICMs fail to generate EPIs and only produce parietal endoderm-like cells (*6*). To determine whether Nanog prevents endoderm differentiation, we analyzed $Nanog^{\Delta/\Delta}$ ESCs for expression of the PE markers *Gata4*, *Gata6*, *Sox17*, and *Foxa2* (Fig. 1C). Unexpectedly, $Nanog^{\Delta/\Delta}$ ESCs remained undifferentiated and morphologically indistinguishable from wild type and did not exhibit precocious PE marker gene expression (Fig. 1, C and D). This observation, paralleling the findings reported by others (*9*, *29*, *39*), raised the possibility of a Nanog requirement for PE formation. The PRC1 and PRC2 have emerged as key regulators of cell fate decisions during early development, although the precise mechanisms responsible for their cell fate allocation during blastocyst lineage formation remain elusive (*30*). In this study, our functional screening of the PcG superfamily uncovered Mga as a strong suppressor of the expression of endodermal gene *Gata6* in ESCs (Fig. 1F). Consistently, we show that the loss of *Mga* in ESCs leads to their spontaneous differentiation to XEN cells. The morphological change was additionally accompanied by the extensive loss of expression of core transcription factors, *Oct4*, *Sox2*, and *Nanog*. We have further demonstrated the importance of Mga function in maintenance of pluripotency by directly repressing the expression of *Gata4/6* and *Sox17*. In addition, the conversion of ESCs to XEN cells was completely rescued in *Mga/Gata4/6* triple-mutant ESCs. These findings thus demonstrate Mga as a crucial upstream repressor of master lineage-defining genes in the ExEn. Our findings in *Mga*-null ESCs may reflect a physiological role for Mga during the in vivo segregation of the PE lineage from ICM. *Mga*-null ESCs and $Mga^{+/\Delta}$ mice may serve as useful tools for studying the molecular mechanisms governing the cell fate decision between PE and EPI in the mouse embryo. Given the importance of the Mga-mediated transcriptional program in ESCs, our data on Mga targets are highly informative and helpful for future identification and characterization of additional genes involved in cell fate decisions and the establishment of different lineages in the early embryonic development. How the interaction between Mga and Sox17 or Gata4/6 is integrated with signaling cues to direct early embryonic cell-fate decisions is currently under investigation by using $Mga^{+/\Delta}$ mice.

Mga exists within a large Max network of transcription factors composed of a group of bHLHZip proteins that form heterodimers with the small bHLHZip partner Max. This network includes Myc paralogs (Myc, Mycn, and Mycl1), their obligate heterodimer Max, and the Max-binding antagonists of Myc (Mxd1 to 4, Mnt, and Mga) (*40*). Proteins of this network have been implicated in multiple aspects of cell function, including proliferation, differentiation, and apoptosis. Most of bHLHZip proteins form homodimers or heterodimers through their bHLHZip domains to bind the palindromic CACGTG E-box motif within regulatory regions of target genes (*40*). In addition to the bHLHZip domain, Mga harbors a second DNA binding domain, the T-box. On the basis of in vitro studies (*14*), Mga has been previously proposed to function as a dual-specificity transcription factor, regulating the expression of both the T-box

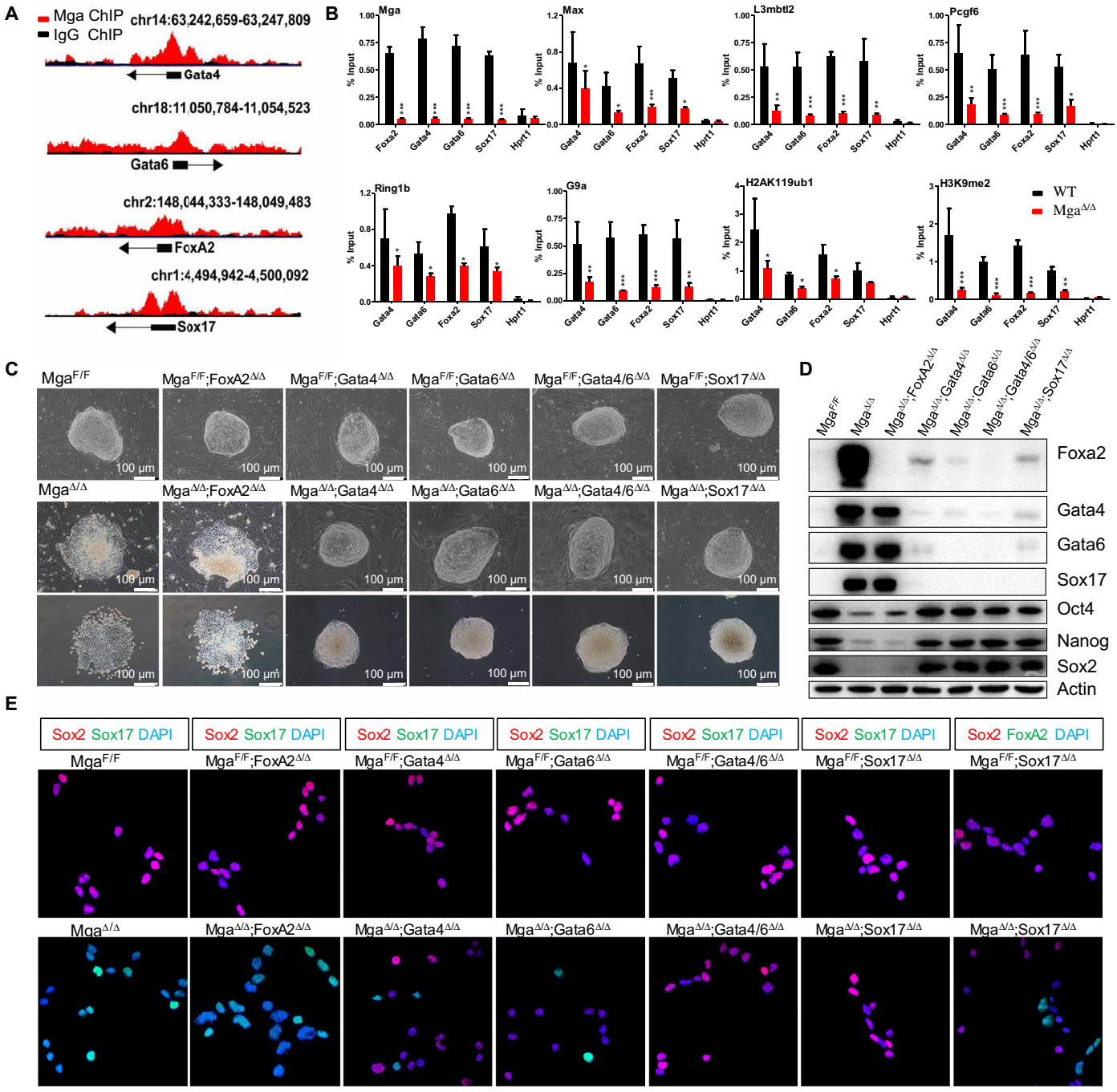

Fig. 7. Mga sustains pluripotency by imposing a PE differentiation blockade directly involving the master regulator *Gata4/6* and *Sox17*. (**A**) Genomic snapshots of Mga ChIP-seq profiles at selected target gene loci, performed in wild-type ESCs. Published ChIP-seq data for Mga were obtained from NCBI GEO (accession number GSM1041373). IgG, immunoglobulin G. (**B**) ChIP-qPCR analysis of Mga binding to selected promoters. Error bars represent SEM of three independent biological replicates. (**C**) Phase-contrast images of ESCs of indicated genotypes cultured on a feeder layer of MEFs (top and middle) or gelatin (bottom). Scale bar, 100 μm. (**D**) Western blot analysis of endodermal markers and pluripotency factors in ESCs of indicated genotypes. β-Actin was used as a loading control. (**E**) Immunofluorescence analysis for Sox17 or Foxa2 (green) and Sox2 (red) with DAPI (blue) merge in ESCs of indicated genotypes. Images were acquired using a Zeiss 880 laser confocal microscope at ×63 magnification.

family and the Max-network genes. Mga-Max heterodimers were often found as part of a large multimeric PRC that also includes E2f6, DP1, Hp1γ, Pcgf6, L3mbtl2, Ring1a/b, G9a, Glp, and Rybp/Yaf2, implicating Mga with functions related to development and differentiation (*12*, *19*, *20*). Consistently, this study, together with previous work (*16*), shows that disruption of *Mga* in mice results in peri-implantation embryonic lethality. This correlates with severely compromised proliferation and abnormal differentiation of $Mga^{\Delta/\Delta}$ ESCs. Most significantly, after loss of *Mga*, ESCs fail to retain pluripotency and preferentially differentiate toward PE.

Aspects of the phenotype after *Mga* ablation resemble findings after disruption of other PRC1.6 components. For example, embryos lacking *Ring1b* (*41*), *Max* (*23*), *L3mbtl2* (*20*), and *Rybp* (*42*) exhibit early embryonic lethality, whereas loss of *Pcgf6* in mice results in partially penetrant embryonic lethality (*24*, *25*). Likewise, significantly increased expression of meiosis- and germline-specific genes (Fig. 3F) has been observed in ESCs lacking *Ring1b* (*43*), *Pcgf6* (*21*), *L3mbtl2* (*20*), *Max* (*22*), and *G9a/Glp* (*22*). Like *Mga*-null ESCs, ESCs deficient in *Ring1b* (*43*), *L3mbtl2* (*20*), or *Max* (*22*, *44*) fail to repress a subset of endoderm genes, as determined by RT-qPCR. Last, like $Mga^{\Delta/\Delta}$ EBs (fig. S3), EBs lacking *L3mbtl2* (*20*), *Pcgf6* (*21*), and *Ring1b* (*43*) display skewed differentiation toward endoderm. Despite these similarities, none of the phenotypes associated with loss of other PRC1.6 components phenocopy the findings with *Mga* knockout. Apparently, the most unique feature is Mga's strong negative impact on the conversion of ESCs to XEN cells. None of the other PRC1.6 complex subunits are associated with an appreciable XEN conversion upon gene disruption.

Mechanistically, the most severe phenotype associated with *Mga* ablation may be the consequence of the displacement of PRC1.6 from chromatin upon loss of *Mga* expression. Consistent with this notion, ChIP-seq experiments demonstrated that Mga colocalized with other components in the PRC1.6 complex (i.e., Pcgf6, L3mbtl2, and E2F6) on its target promoters (*31*). Moreover, genome-wide binding of Pcgf6, L3mbtl2, and E2F6 is completely abolished in *Mga*-depleted cells (*31*), indicating that Mga is absolutely crucial for genomic targeting of the entire PRC1.6. To determine whether the two DNA binding domains of Mga account for genomic occupancy of PRC1.6, we generated *Mga* mutants deficient for DNA binding by deleting the T-box domain (MGA-ΔT) or bHLHZip domain (ΔbHLHZip). Unexpectedly, in $Mga^{\Delta T\text{-box}}$ or $Mga^{\Delta bHLHZip}$ ESCs, the genomic binding of these Mga mutants and their interacting partners was only moderately affected. Consistently, $Mga^{\Delta T\text{-box}}$ or $Mga^{\Delta bHLHZip}$ ESCs retain the capacity for differentiation but exhibit a mild defect in proliferation. Our observations are supported by the study in human embryonic kidney–293 cells, showing that the T-box and bHLHZip domains are essential for Mga chromatin occupancy at a subset of loci, whereas at other sites, they make little or no contribution (*31*). These results strongly indicate that sequence-specific DNA binding via Mga is not sufficient to ensure PRC1.6 recruitment but suggest the requirement of the combinatorial activities of PRC1.6 components in mediating the proper genomic localization of PRC1.6.

Western blot analysis revealed that deletion of *Mga* in ESCs markedly reduced steady-state levels of L3mbtl2, Pcgf6, and Max. The reduced protein levels of these Mga interaction partners were likely due to impaired protein stability, as the mRNA levels of these PRC1.6 subunits were not reduced in *Mga*-null ESCs. These results suggest that Mga may function as a scaffold protein that interacts directly with other major components including L3mbtl2, Pcgf6, and Max. Notably, other core components of PRC1.6 do not likely function as scaffold. In support of this notion, the deficiency of *L3mbtl2* (*20*, *36*), *Pcgf6*, or *Max* (*36*) did not impair the assembly of the remaining components of the PRC1.6 complex. Furthermore, *L3mbtl2* (*20*), *Pcgf6* (*21*), or *Max* ablation did not reduce the protein levels of other components. Thus, all these results strongly suggest the critical role of Mga in the post-transcriptional stabilization of the major components of the PRC1.6 complex. The discovery that Mga plays a vital role as a scaffold to assemble the PRC1.6 complex provides a plausible explanation for why the loss of Mga alone leads to the most severe phenotype of *Mga*-null ESCs and mice. We have shown here that the T-box and bHLHZip motifs of Mag are not involved in the stabilization of components of the PRC1.6 complex. Our ongoing study is to define the domain in Mga responsible for posttranscriptional stabilization of the core components of the PRC1.6.complex. Considering that Mga also associated with G9a and Glp (Fig. 6E) and H3K9me2 at the promoter regions of endodermal genes markedly decreased upon Mga ablation (Fig. 7B), various epigenetic regulators associated with the complex could be involved in the repression of endodermal genes in ESCs. Future research should also focus on determining whether Mga-associated transcription factors and epigenetic factors function in concert to prevent ESCs from embarking on endodermal differentiation.

MATERIALS AND METHODS

Mice

The experimental animal facility has been accredited by the Association for Assessment and Accreditation of Laboratory Animal Care International, and the Institutional Animal Care and Use Committee of the Model Animal Research Center of Nanjing University approved all animal protocols used in this study. The strategy for generating *Mga* conditional targeting vector, in which exons 16 and 17 were flanked by loxP sites, was detailed previously (*20*). This targeting construct was linearized and transfected into C57BL/6 ESCs, followed by selection in G418 (500 μg/ml). ESCs were grown with MEFs on gelatinized tissue culture plates in Dulbecco's modified Eagle's medium (DMEM). Targeted ESC clones were identified by PCR analysis. Subsequently, the FRT-flanked neomycin resistance gene was selectively removed by transient expression of FLPe recombinase. Two such *Mga* floxed ESC clones were expanded and microinjected into C57BL/6 blastocysts to generate chimeras as determined by agouti coat color. Chimeric males were bred with C57BL/6 wild-type females for germline transmission. The transmitted floxed mice were intercrossed to β-actin–Cre recombinase mice to generate animals with *Mga*-null allele. Mice heterozygous for the null allele were then interbred to generate embryos lacking *Mga*.

Cell lines

Unless otherwise indicated, male mouse ESCs were grown on gelatin-coated tissue culture plates with mitomycin-treated feeder cells at 4°C and 5% CO_2, in DMEM, supplemented with 15% ESC-qualified fetal bovine serum, 2 mM L-glutamine, sodium pyruvate, nonessential amino acids, 0.5 mM β-mercaptoethanol, and 10 ng/ml of LIF and penicillin/streptomycin as described (*20*). Cells were regularly examined for the presence of mycoplasma contamination. Before harvesting these ESCs, the feeder cells were depleted by preplating trypsinized cells for 30 min at 4°C in maintenance medium on plates not coated with gelatin and discarding attached cells. To induce differentiation, ESCs were cultured in the absence of LIF and with 5 mM all-trans retinoic acid for up to 14 days with daily medium changes. ESC colonies were fixed in ice-cold methanol, stained with 0.1% crystal violet, washed, and left to air-dry. Colony number (100 colonies per sample) and colony size were quantified with ImageJ software.

RNA preparation and real-time qPCR

Cells were first rinsed in 1× phosphate-buffered saline (PBS) and total RNA was isolated using TRIzol reagent (Invitrogen) as described previously (*20*). A total of 2.5 μg of the corresponding RNA

yield was subjected to reverse transcription using HiScript II 1st Strand cDNA Synthesis Kit (Vazyme) and oligo(dT) as primer. Reverse transcription (RT) reactions carried out in the absence of reverse transcriptase were used as negative controls for genomic DNA contamination. All real-time assays were performed in triplicate on the StepOnePlus Real-Time PCR System and PowerUp SYBR Green Master Mix (Applied Biosystems), according to the manufacturer's protocol. Relative expression levels of each gene normalized to β-actin was calculated as $2^{-\Delta Ct}$, where $\Delta Ct = Ct_{\text{detected gene}} - Ct_{\beta\text{-actin}}$. Comparison between samples was performed using Student's *t* test. All primers are listed in table S5.

Genome engineering of ESCs via CRISPR–homology-directed repair

The guide RNA (gRNA) targeting the region of interest was designed by using the Zhang Lab Guide Design Resources (https://zlab.bio/guide-design-resources) tool and cloned into pSpCas9(BB)-2A-Puro (PX459) V2.0 (Addgene plasmid number 62988) by introducing synthesized oligo-DNA primers corresponding to the target sequence into Bbs I restriction sites. For the detailed gRNA sequences for each gene, see supplementary experimental procedures. To generate conditional knockout ESC clones, cells were cotransfected with 1-μg single-guide RNA vector and/or 1.5-μg DNA linearized donor plasmid using Lipofectamine 2000 according to the manufacturer's protocol. ESC clones that had stably integrated the targeting vector were identified by selection with G418 sulfate (400 μg/ml). The G418-resistant colonies were manually picked into 24-well plates and expanded for genomic DNA extraction and continued culture. Correctly targeted ESC clones were identified by genomic PCR analysis.

Mouse embryo study

Timed pregnancies were determined by daily examinations of the vaginas of females caged with adult male mice as described (*20*). The day of detection of a vaginal plug was designated E0.5 (day 0.5 of pregnancy). Embryos were dissected from the decidua and imaged followed by PCR genotyping with genomic DNA extracted from embryo. E6.5 and E7.5 decidua from $Mga^{+/\Delta}$ intercrosses were fixed in 4% paraformaldehyde, sectioned, and stained with hematoxylin and eosin (H&E). Mouse blastocysts were flushed from the uterus at E3.5 of gestation using M2 medium and then immediately placed in DMEM for culture.

Flow cytometry

Apoptosis assay

For apoptosis analysis, ESCs were harvested by trypsinization, washed twice with cold PBS, labeled with fluorescein isothiocyanate–Annexin V (BD Pharmingen) and propidium iodine (PI) in the dark at room temperature, and analyzed by flow cytometry as described (*20*).

Cell cycle analysis

ESCs were trypsinized, washed twice with cold PBS, fixed in ice-cold 70% ethanol at −20°C overnight. Subsequently, cells washed twice with PBS, incubated for 30 min at room temperature with ribonuclease A (RNase A) (200 μg/ml), and stained with PI solution (50 mg/ml in PBS) in the dark at 4°C for 30 min followed by analysis on a LSRFortessa flow cytometer equipped with Cell Quest software (BD Biosciences) as described (*20*).

EB formation and analysis

To induce the EB formation, modified hanging drop method from previously described protocol (*20*) was performed. Briefly, ES cells were harvested by trypsinization and resuspended in IMDM medium without LIF (nonessential amino acids, 2 mM L-glutamine, iron saturated holo-transferrin (200 μg/ml), sodium pyruvate, ascorbic (50 μg/ml), 450 μM monothioglycerol, 15% ES cell–grade serum, and penicillin/streptomycin). The cells were plated at a density of 600 cells per 30-μl drop in the lids of petri dishes for 3 days. Then, aggregates were transferred to uncoated petri dishes and incubated on a rotating shaker at 60 rpm in 5% CO_2 at 37°C. EBs were visualized or harvested at different time points. Total RNA was collected from EBs using TRIzol reagent (Invitrogen) according to the manufacturer's instructions and analyzed by RT-PCR. The levels of marker genes representative of the undifferentiated state and the three embryonic germ layers were examined to evaluate their in vitro differentiation potential. The mRNA of β-actin was used as an internal control. Primer details are given in table S5 and in a previous report (*21*).

Teratoma formation assay

Teratoma formation assay was performed as previously described (*20*). Briefly, ESCs were grown under differentiation conditions, harvested by trypsinization, washed in PBS, resuspended in PBS, and then injected subcutaneously into the nude mice. Teratomas were dissected, fixed, and subsequently processed by H&E staining.

Western blot and immunoprecipitation

Cell extraction and Western blots were performed as described (*20*). Briefly, to obtain whole-cell extracts, the harvested cell pellets were lysed on ice in radioimmunoprecipitation assay (RIPA) buffer [50 mM tris-HCl (pH 8.0), 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% SDS, 0.5% sodium deoxycholate, 2 mM Na_3VO_4, 1 mM dithiothreitol, 100 mM NaF, phenylmethylsulfonyl fluoride (10 mg/ml), and Protease Inhibitor Mix (Sigma-Aldrich)] for 30 min, and supernatant were mixed with SDS sample buffer and boiled for 5 min at 100°C and subjected to Western blot analysis. The blots were then incubated with an appropriate secondary antibody (horseradish peroxidase–conjugated anti-mouse, anti-rabbit, or anti-goat) and visualized using an enhanced chemiluminescence detection kit.

For immunoprecipitation assay, cells were washed in PBS and lysed in RIPA buffer without SDS. Cell lysates were clarified by centrifugation, and the supernatant was incubated with antibodies and 20 μl of slurry Protein A/G Sepharose beads (Amersham Biosciences) overnight at 4°C. The immunoprecipitates were then washed three times with lysis buffer with 500 mM NaCl, loaded on SDS–polyacrylamide gel electrophoresis, and probed with appropriate antibodies as indicated. For controls, either rabbit or mouse immunoglobulin G antibody was used. All of the antibodies used in this study are listed in table S6.

Chromatin immunoprecipitation

ChIP was performed as described previously with minor modifications (*20*, *36*). ESCs were cross-linked with 1% of formaldehyde at room temperature for 10 min while gently shaking, and the reaction was stopped by adding of 125 mM glycine. Cells were rinsed three times with cold PBS, resuspended in lysis buffer [50 mM tris (pH 8.0), 1% SDS, 10 mM EDTA, and protease inhibitors] and sonicated with the Bioruptor (Biogenide). Lysates were diluted 10 times with 1× ChIP dilution buffer [16.7 mM tris (pH 8.0), 165 mM NaCl, 1% Triton X-100, 0.01% SDS, 1.2 mM EDTA, and protease inhibitors] and then incubated overnight with 3- to 6-μg indicated antibodies. The next day, 40-μl prewashed protein A/G agarose beads were added and incubated at 4°C. After 6 hours, beads were washed in 1×

dilution buffer (once), low-salt buffer (twice) [20 mM tris (pH 8.0), 150 mM NaCl, 0.1% SDS,1 mM EDTA, 1% Triton X-100, 0.5% Na deoxycholate, and protease inhibitors], high-salt buffer (twice) [20 mM tris (pH 8.0), 500 mM NaCl, 0.1% SDS, 1 mM EDTA, 1% Triton X-100, 0.5% Na deoxycholate, and protease inhibitors], tris-EDTA buffer (twice) [10 mM tris (pH 8.0), 0.25 mM EDTA, and protease inhibitors]. Immune complexes were eluted by incubating the beads at 65°C twice in elution buffer (100 mM $NaHCO_3$ and 1% SDS). Eluates were mixed with RNase and 5 M NaCl (0.3 M final concentration) and then incubated overnight at 65°C to reverse cross-linking. The immunoprecipitated DNA was quantified by real-time qPCR using gene-specific primer pairs. The relative immunoprecipitate versus input ratio is presented as the percentage of input, which is calculated by $(2^{-\Delta Ct}) \times 100\%$, where $\Delta Ct = Ct_{(IP\ sample)} - Ct_{(Input)}$. Primers used for ChIP are listed in table S5.

Immunofluorescence

Cells cultured on gelatin-coated coverslips were fixed with 4% paraformaldehyde for 15 min at room temperature, permeabilized with 0.5% Triton X-100 in PBS, and subsequently incubated with blocking buffer (3% bovine serum albumin in PBS). Primary antibodies were diluted at 1:1000 in blocking solution, and samples were incubated overnight at 4°C rotating in a humidified atmosphere. Slides were washed with PBS followed by adding the appropriate secondary antibody conjugated with Alexa Flour 488 or 594 dyes (1:150 dilution) for 2 hours at dark room. For DNA staining, the slides were incubated in a 4′,6-diamidino-2-phenylindole solution in PBS for 5 min. Fluorescent images were acquired using a Zeiss 880 laser scanning confocal microscope.

RNA-seq analysis

To minimize sample variability caused by individual differences, we used three independent biological replicates to generate samples for RNA-seq assay. Normalization and differential expression analysis was carried out using the DESeq2 using default parameters. Genes with fold change >2 and false discovery rate <0.05 were considered as significantly differentially expressed.

Statistics

Results are expressed as means ± SEM from at least three independent experiments. Statistical analysis was obtained with GraphPad Prism 5 using two-tailed Student's *t* test for comparisons between two datasets, and a probability value (*P*) of less than 0.05, 0.01, or 0.001 was considered statistically significant and indicated by *, **, and ***, respectively ($*P < 0.05$, $**P < 0.01$, $***P < 0.001$).

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/7/4/eabe5689/DC1

View/request a protocol for this paper from *Bio-protocol*.

REFERENCES AND NOTES

1. S. Tanaka, T. Kunath, A. K. Hadjantonakis, A. Nagy, J. Rossant, Promotion of trophoblast stem cell proliferation by FGF4. *Science* **282**, 2072–2075 (1998).
2. M. J. Evans, M. H. Kaufman, Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**, 154–156 (1981).
3. G. R. Martin, Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc. Natl. Acad. Sci. U.S.A.* **78**, 7634–7638 (1981).
4. T. Kunath, D. Arnaud, G. D. Uy, I. Okamoto, C. Chureau, Y. Yamanaka, E. Heard, R. L. Gardner, P. Avner, J. Rossant, Imprinted X-inactivation in extra-embryonic endoderm cell lines from mouse blastocysts. *Development* **132**, 1649–1661 (2005).
5. Y. Yamanaka, A. Ralston, Early embryonic cell fate decisions in the mouse. *Adv. Exp. Med. Biol.* **695**, 1–13 (2010).
6. K. Mitsui, Y. Tokuzawa, H. Itoh, K. Segawa, M. Murakami, K. Takahashi, M. Maruyama, M. Maeda, S. Yamanaka, The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell* **113**, 631–642 (2003).
7. E. E. Morrisey, Z. Tang, K. Sigrist, M. M. Lu, F. Jiang, H. S. Ip, M. S. Parmacek, GATA6 regulates HNF4 and is required for differentiation of visceral endoderm in the mouse embryo. *Genes Dev.* **12**, 3579–3590 (1998).
8. J. Fujikura, E. Yamato, S. Yonemura, K. Hosoda, S. Masui, K. Nakao, J. Miyazaki Ji, H. Niwa, Differentiation of embryonic stem cells is induced by GATA factors. *Genes Dev.* **16**, 784–789 (2002).
9. J. Silva, J. Nichols, T. W. Theunissen, G. Guo, A. L. van Oosten, O. Barrandon, J. Wray, S. Yamanaka, I. Chambers, A. Smith, Nanog is the gateway to the pluripotent ground state. *Cell* **138**, 722–737 (2009).
10. J. A. Simon, R. E. Kingston, Mechanisms of polycomb gene silencing: Knowns and unknowns. *Nat. Rev. Mol. Cell Biol.* **10**, 697–708 (2009).
11. B. Schuettengruber, H. M. Bourbon, L. Di Croce, G. Cavalli, Genome regulation by polycomb and trithorax: 70 years and counting. *Cell* **171**, 34–57 (2017).
12. Z. Gao, J. Zhang, R. Bonasio, F. Strino, A. Sawai, F. Parisi, Y. Kluger, D. Reinberg, PCGF homologs, CBX proteins, and RYBP define functionally distinct PRC1 family complexes. *Mol. Cell* **45**, 344–356 (2012).
13. S. A. Turner, A. P. Bracken, A "complex" issue: Deciphering the role of variant PRC1 in ESCs. *Cell Stem Cell* **12**, 145–146 (2013).
14. P. J. Hurlin, E. Steingrímsson, N. G. Copeland, N. A. Jenkins, R. N. Eisenman, Mga, a dual-specificity transcription factor that interacts with Max and contains a T-domain DNA-binding motif. *EMBO J.* **18**, 7019–7028 (1999).
15. A. Kispert, B. Koschorz, B. G. Herrmann, The T protein encoded by Brachyury is a tissue-specific transcription factor. *EMBO J.* **14**, 4763–4772 (1995).
16. A. J. Washkowitz, C. Schall, K. Zhang, W. Wurst, T. Floss, J. Mager, V. E. Papaioannou, Mga is essential for the survival of pluripotent cells during peri-implantation development. *Development* **142**, 31–40 (2015).
17. L. De Paoli, M. Cerri, S. Monti, S. Rasi, V. Spina, A. Bruscaggin, M. Greco, C. Ciardullo, R. Fama, S. Cresta, R. Maffei, M. Ladetto, M. Martini, L. Laurenti, F. Forconi, R. Marasca, L. M. Larocca, F. Bertoni, G. Gaidano, D. Rossi, MGA, a suppressor ofMYC, is recurrently inactivated in high risk chronic lymphocytic leukemia. *Leuk. Lymphoma* **54**, 1087–1090 (2013).
18. Cancer Genome Atlas Research Network, Comprehensive molecular profiling of lung adenocarcinoma. *Nature* **511**, 543–550 (2014).
19. H. Ogawa, K.-I. Ishiguro, S. Gaubatz, D. M. Livingston, Y. Nakatani, A complex with chromatin modifiers that occupies E2F- and Myc-responsive genes in G0 cells. *Science* **296**, 1132–1136 (2002).
20. J. Qin, W. A. Whyte, E. Anderssen, E. Apostolou, H. H. Chen, S. Akbarian, R. T. Bronson, K. Hochedlinger, S. Ramaswamy, R. A. Young, H. Hock, The polycomb group protein L3mbtl2 assembles an atypical PRC1-family complex that is essential in pluripotent stem cells and early development. *Cell Stem Cell* **11**, 319–332 (2012).
21. W. Zhao, H. Tong, Y. Huang, Y. Yan, H. Teng, Y. Xia, Q. Jiang, J. Qin, Essential role for polycomb group protein pcgf6 in embryonic stem cell maintenance and a noncanonical polycomb repressive complex 1 (PRC1) integrity. *J. Biol. Chem.* **292**, 2773–2784 (2017).
22. I. Maeda, D. Okamura, Y. Tokitake, M. Ikeda, H. Kawaguchi, N. Mise, K. Abe, T. Noce, A. Okuda, Y. Matsui, Max is a repressor of germ cell-related gene expression in mouse embryonic stem cells. *Nat. Commun.* **4**, 1754 (2013).
23. H. Shen-Li, R. C. O'Hagan, H. Hou Jr., J. W. Horner II, H. W. Lee, R. A. DePinho, Essential role for Max in early embryonic growth and development. *Genes Dev.* **14**, 17–22 (2000).
24. M. Endoh, T. A. Endo, J. Shinga, K. Hayashi, A. Farcas, K. W. Ma, S. Ito, J. Sharif, T. Endoh, N. Onaga, M. Nakayama, T. Ishikura, O. Masui, B. M. Kessler, T. Suda, O. Ohara, A. Okuda, R. Klose, H. Koseki, PCGF6-PRC1 suppresses premature differentiation of mouse embryonic stem cells by regulating germ cell-related genes. *eLife* **6**, e21064 (2017).
25. M. Liu, Y. Zhu, F. Xing, S. Liu, Y. Xia, Q. Jiang, J. Qin, The polycomb group protein PCGF6 mediates germline gene silencing by recruiting histone-modifying proteins to target gene promoters. *J. Biol. Chem.* **295**, 9712–9724 (2020).
26. L. A. Boyer, T. I. Lee, M. F. Cole, S. E. Johnstone, S. S. Levine, J. P. Zucker, M. G. Guenther, R. M. Kumar, H. L. Murray, R. G. Jenner, D. K. Gifford, D. A. Melton, R. Jaenisch, R. A. Young, Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122**, 947–956 (2005).
27. J. Nichols, B. Zevnik, K. Anastassiadis, H. Niwa, D. Klewe-Nebenius, I. Chambers, H. Schöler, A. Smith, Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* **95**, 379–391 (1998).

28. A. A. Avilion, S. K. Nicolis, L. H. Pevny, L. Perez, N. Vivian, R. Lovell-Badge, Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev.* **17**, 126–140 (2003).
29. I. Chambers, D. Colby, M. Robertson, J. Nichols, S. Lee, S. Tweedie, A. Smith, Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell* **113**, 643–655 (2003).
30. L. A. Boyer, K. Plath, J. Zeitlinger, T. Brambrink, L. A. Medeiros, T. I. Lee, S. S. Levine, M. Wernig, A. Tajonar, M. K. Ray, G. W. Bell, A. P. Otte, M. Vidal, D. K. Gifford, R. A. Young, R. Jaenisch, Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* **441**, 349–353 (2006).
31. B. Stielow, F. Finkernagel, T. Stiewe, A. Nist, G. Suske, MGA, L3MBTL2 and E2F6 determine genomic binding of the non-canonical polycomb repressive complex PRC1.6. *PLOS Genet.* **14**, e1007193 (2018).
32. W. Zhao, M. Liu, H. Ji, Y. Zhu, C. Wang, Y. Huang, X. Ma, G. Xing, Y. Xia, Q. Jiang, J. Qin, The polycomb group protein Yaf2 regulates the pluripotency of embryonic stem cells in a phosphorylation-dependent manner. *J. Biol. Chem.* **293**, 12793–12804 (2018).
33. K. Hisada, C. Sanchez, T. A. Endo, M. Endoh, M. Roman-Trufero, J. Sharif, H. Koseki, M. Vidal, RYBP represses endogenous retroviruses and preimplantation- and germ line-specific genes in mouse embryonic stem cells. *Mol. Cell. Biol.* **32**, 1139–1149 (2012).
34. Y. Zhong, T. Choi, M. Kim, K. H. Jung, Y. G. Chai, B. Binas, Isolation of primitive mouse extraembryonic endoderm (pXEN) stem cell lines. *Stem Cell Res.* **30**, 100–112 (2018).
35. A. C. McDonald, S. Biechele, J. Rossant, W. L. Stanford, Sox17-mediated XEN cell conversion identifies dynamic networks controlling cell-fate decisions in embryo-derived stem cells. *Cell Rep.* **9**, 780–793 (2014).
36. Y. Huang, W. Zhao, C. Wang, Y. Zhu, M. Liu, H. Tong, Y. Xia, Q. Jiang, J. Qin, Combinatorial control of recruitment of a variant PRC1.6 complex in embryonic stem cells. *Cell Rep.* **22**, 3032–3043 (2018).
37. S. E. Wamaitha, I. del Valle, L. T. Cho, Y. Wei, N. M. Fogarty, P. Blakeley, R. I. Sherwood, H. Ji, K. K. Niakan, Gata6 potently initiates reprograming of pluripotent and differentiated cells to extraembryonic endoderm stem cells. *Genes Dev.* **29**, 1239–1255 (2015).
38. K. K. Niakan, H. Ji, R. Maehr, S. A. Vokes, K. T. Rodolfa, R. I. Sherwood, M. Yamaki, J. T. Dimos, A. E. Chen, D. A. Melton, A. P. McMahon, K. Eggan, Sox17 promotes differentiation in mouse embryonic stem cells by directly regulating extraembryonic gene expression and indirectly antagonizing self-renewal. *Genes Dev.* **24**, 312–326 (2010).
39. D. M. Messerschmidt, R. Kemler, Nanog is required for primitive endoderm formation through a non-cell autonomous mechanism. *Dev. Biol.* **344**, 129–137 (2010).
40. M. Conacci-Sorrell, L. McFerrin, R. N. Eisenman, An overview of MYC and its interactome. *Cold Spring Harb. Perspect. Med.* **4**, a014357 (2014).
41. J. W. Voncken, B. A. Roelen, M. Roefs, S. de Vries, E. Verhoeven, S. Marino, J. Deschamps, M. van Lohuizen, Rnf2 (Ring1b) deficiency causes gastrulation arrest and cell cycle inhibition. *Proc Natl Acad Sci U.S.A.* **100**, 2468–2473 (2003).
42. M. K. Pirity, J. Locker, N. Schreiber-Agus, Rybp/DEDAF is required for early postimplantation and for central nervous system development. *Mol. Cell. Biol.* **25**, 7193–7202 (2005).
43. M. Leeb, D. Pasini, M. Novatchkova, M. Jaritz, K. Helin, A. Wutz, Polycomb complexes act redundantly to repress genomic repeats and genes. *Genes Dev.* **24**, 265–276 (2010).
44. T. Hishida, Y. Nozaki, Y. Nakachi, Y. Mizuno, Y. Okazaki, M. Ema, S. Takahashi, M. Nishimoto, A. Okuda, Indefinite self-renewal of ESCs through Myc/Max transcriptional complex-independent mechanisms. *Cell Stem Cell* **9**, 37–49 (2011).

Acknowledgments: We are indebted to Z. Gan and L. Liu for helpful suggestions and for providing reagents. We thank members of our laboratory for helpful discussion. **Funding:** This work was supported by grants from the National Natural Science Foundation of China (31671532 and 31970810) to J.Q. **Author contributions:** Conceptualization: J.Q. Mouse models: J.Q. Methodology: J.Q., C.W., Y.Z., and K.H. Formal analysis: C.W., Y.Z., T.S., L.D., and J.Q. Mechanistic investigations: C.W., Y.Z., Y.H., and J.Q. Writing: J.Q. (original draft); J.Q., C.W., and Y.Z. (review and editing). Funding acquisition: J.Q. Supervision: J.Q. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** The accession numbers for the sequencing data reported in this paper are NCBI GEO: GSE143819 and GSE144141. All other data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.

Submitted 1 September 2020
Accepted 24 November 2020
Published 20 January 2021
10.1126/sciadv.abe5689

Citation: J. Qin, C. Wang, Y. Zhu, T. Su, L. Dong, Y. Huang, K. Hao, Mga safeguards embryonic stem cells from acquiring extraembryonic endoderm fates. *Sci. Adv.* **7**, eabe5689 (2021).