Epigenetic regulation of mouse preimplantation embryo development
Xudong Fu1,2,3,6, Chunxia Zhang1,2,3,6 and Yi Zhang1,2,3,4,5

After fertilization, mouse embryos go through preimplantation development to give rise to blastocyst. Two key molecular events, zygotic genome activation (ZGA) and the first cell lineage specification, are essential for the process. Recent advances in low-input epigenomics profiling techniques allow the analysis of these events at a molecular level, which revealed a critical role of epigenetic and chromatin reprogramming in ZGA and the first cell lineage specification. Additionally, the establishment of an in vitro embryonic stem cell (ESC) to two-cell embryo-like conversion system have also contributed to the molecular understanding of preimplantation development. In this review, we summarize recent advances in epigenetic regulation of mouse preimplantation development, point out the remaining questions, and propose strategies to tackle these questions.

Addresses
1 Howard Hughes Medical Institute, Boston Children’s Hospital, Boston, MA 02115, USA
2 Program in Cellular and Molecular Medicine, Boston Children’s Hospital, Boston, MA 02115, USA
3 Division of Hematology/Oncology, Department of Pediatrics, Boston Children’s Hospital, Boston, MA 02115, USA
4 Department of Genetics, Harvard Medical School, Boston, MA 02115, USA
5 Harvard Stem Cell Institute, Boston, MA 02115, USA

Corresponding author: Zhang, Yi (y.zhang@genetics.med.harvard.edu)

Epigenetic dynamics during mouse preimplantation development
Since the sperm and egg genomes are organized very differently, it is anticipated that they go through dramatic changes during preimplantation development to become epigenetically equalized, with the exception of imprinted loci. However, the limited availability of mammalian oocytes and embryos impeded global analyses of this dynamic process. Now, this challenge has been largely resolved with the development of low-input techniques for epigenomic profiling.

Immunostaining and whole-genome bisulfite sequencing (WGBS) analysis revealed global DNA demethylation soon after fertilization (Figure 1). While both paternal and maternal genomes undergo demethylation, the paternal genome demethylation is at a faster rate. This demethylation process is driven by a combination of Tet3-mediated 5-methylcytosine (5mC) oxidation and DNA-replication-dependent dilution [4–6]. However, Tet3-dependent DNA demethylation is not required for ZGA and preimplantation development [5,7].

In addition to DNA methylation, histone modifications are also reprogrammed during mouse preimplantation development (Figure 1). The most notable change on histone modifications is the trimethylation of histone 3 lysine 4 (H3K4me3), a mark of active transcription when located at gene promoter. However, broad non-canonical H3K4me3 domains have also been identified in transcriptionally silent oocytes, which may function as a repressive mark to prevent transcription before ZGA [8]. The broad H3K4me3 domains are replaced by the
Overview of the mRNA level, DNA methylation, and chromatin landscapes during mouse preimplantation development.

After fertilization, maternal mRNA (orange line) are rapidly degraded while two waves of ZGA (both minor and major ZGA, green line) take place. Paternal genome 5-methylcytosine (5mC, blue line) is rapidly decreased in zygotes, while maternal 5mC level (orange line) is gradually decreased during preimplantation development. The broad H3K4me3 (green) domain in oocytes is removed at two-cell stage and replaced by the promoter H3K4me3 peaks. H3K27ac (orange) level is increased at two-cell stage. The H3K27me3 (purple) at promoter region is rapidly erased after fertilization, while the H3K27me3 at distal region is gradually decreased from two-cell stage. The H3K9me3 (blue) at promoter region is gradually decreased during preimplantation development. Maternal H3K36me3 (pink) across the entire gene body is attenuated after fertilization and lost at eight-cell stage, zygotic H3K36me3 starts to establish from two-cell stage and becomes stronger at eight-cell stage. Chromatin accessibility (yellow) is increased from two-cell stage and further enhanced during preimplantation development. Higher order chromatin structure, indicated by compartments and topologically associating domains (TADs), is disordered in MII oocytes and zygotes (light brown triangles with dashed lines) and established from late two-cell stage (brown triangles with solid lines). Polycomb-associating domains (PADs, red triangles with solid lines) are present transiently during two-cell to eight-cell stages.
promoter-associated canonical H3K4me3 peaks at the late two-cell stage [8,9]. Interestingly, removal of the broad H3K4me3 domains appears to be essential for ZGA as knockdown of the H3K4me3 demethylases Kdm5a/b results in preimplantation developmental arrest with defects in activation of a subset of ZGA genes [9]. The establishment of promoter H3K4me3 peaks coincides with the appearance of histone 3 lysine 27 acetylation (H3K27ac) [9], another mark for gene activation, indicating that these active marks may function together to mediate ZGA. It has been proposed that genomic elements close to the broad H3K4me3 that harbor H3K27ac may serve as putative active cis-regulatory elements (cREs) for ZGA [9]. Removal of the H3K4me3 broad domains after fertilization may allow binding of transcription factors (TFs) at cREs to deposit active H3K27ac marks, resulting in activation of the ZGA genes [9]. Whether this is indeed the case awaits to be shown.

In addition to active histone marks, inactive marks such as trimethyl-histone 3 lysine 27 (H3K27me3) and lysine 9 (H3K9me3) also exhibit dynamic changes during preimplantation development. While promoter H3K27me3 and H3K9me3 are quickly erased or decreased after fertilization, the broad H3K27me3 domains in distal regions can be maintained until the blastocyst stage (Figure 1) [10–12]. Removal of H3K9me3 is essential for preimplantation development as failure to do so in somatic cell nuclear transfer (SCNT) embryos leads to developmental arrest due to ZGA defects [13]. Whether removal of H3K27me3 is required for preimplantation development still awaits to be shown. Yet, maternal H3K27me3-mediated non-canonical imprinting in preimplantation embryos is largely lost in the embryonic lineage after implantation, but maintained in the extraembryonic lineage at certain loci that include the maternal Xist and certain genes important for placenta development [14,15].

In oocytes, the establishment of the aforementioned DNA methylation, non-canonical H3K4me3, and H3K27me3 are largely modulated by another histone modification, trimethylation of histone 3 lysine 36 (H3K36me3) [16]. The depletion of maternal H3K36me3 in oocytes results in the disruption of the maternal epigenome, thus causing defects in ZGA and embryonic development [16]. After fertilization, maternal H3K36me3 is attenuated from the late two-cell stage and disappears at the eight-cell stage, while zygotic H3K36me3 is established gradually during preimplantation development (Figure 1) [16].

Local and high order chromatin dynamics during mouse preimplantation development

In eukaryotic cells, histones and DNA are organized in the form of chromatin. As mentioned earlier, the chromatin of sperm cells and eggs are organized differently. Upon fertilization, chromatin from both sperm cells and eggs undergo reprogramming to achieve similar chromatin accessibility so that gene expression on both parental alleles can be equalized. Using the assay for transposase-accessible chromatin sequencing (ATAC-seq) and low-input DNase I sequencing (liDNase-seq), the dynamics of chromatin accessibility during mouse preimplantation development have been profiled [17–19]. The analysis revealed a very rapid reprogramming process with the two parental genomes exhibit similar chromatin accessibility within 12 hours of fertilization and this reprogramming process is DNA replication-independent [19]. However, the protein factors that mediate the reprogramming process are largely unknown. During preimplantation development, chromatin accessibility gradually increases with zygotes and two-cell embryos exhibiting less chromatin accessibility. However, a rapid increase in chromatin accessibility takes place during and after ZGA and mainly occurs in the distal regions, which correlates with the usage of enhancer elements (Figure 1) [17,18]. Although the dynamics of chromatin accessibility in human preimplantation embryos are largely similar [20], species-specific differences have been observed, including the establishment of open chromatin at the OCT4 binding motifs at the time of ZGA [20], which does not occur in mouse until morula stage [18].

In addition to the local chromatin accessibility revealed by ATAC-seq and liDNase-seq, chromatin structural changes can also be assessed by their folding patterns, including the self-interacting topologically associating domains (TADs), and the non-DNA-structure-interacting domains, such as the lamina-associated domains (LADs) [21,22]. Dynamic changes in TADs and LADs have been reported during preimplantation development (Figure 1). Specifically, sperm cells show a structured organization, while the MII oocytes lack TADs and compartments [23,24]. After fertilization, TADs undergo reorganization, from weak TADs in zygotes and early two-cell embryos to newly established TADs at late two-cell embryos, and this reorganization is ZGA-independent [23,24]. A recent study has reported a new self-interacting compartmental domain, the Polycomb-associating domains (PADs), marked by H3K27me3 over large genomic blocks [25]. PADs are established during oocyte growth and disassembled upon meiotic resumption, but briefly reappeared in the maternal genome after fertilization. Polycomb repressive complex 1 (PRC1) and PRC2 are responsible for the establishment of PADs in oocytes and after fertilization, respectively. Another study also reported the existence of Polycomb-marked compartments in early embryos [26]. Besides TADs and PADs, LADs were reported to be established de novo after fertilization [27]. LADs formation, which may be H3K4me3-dependent, occurs after TAD consolidation and may help to prime the repressive compartments for certain genomic regions. Whether these higher-order
chromatin dynamics play any role in ZGA and preimplantation development, or they are the outcome of development cue, remains to be determined.

**Regulators of ZGA in mouse**

ZGA is one of the key molecular events during preimplantation development yet the major regulators of this event still remains elusive. Recent studies using low-input genomic profiling approaches and the two-cell-embryo-like cell (2CLC) system, several candidate regulators for ZGA have been identified.

An integrative analysis of transcriptome and chromatin accessibility of zygotes and two-cell embryos revealed 

**Nfy** as one of the candidate TFs regulating ZGA because the Nfy-binding motif is the most enriched motifs in two-cell embryo activated gene promoters with newly gained chromatin accessibility [18]. The knockdown of maternal Nfy resulted in a decrease in chromatin accessibility and defects in the activation of a subset of ZGA genes [18]. A second candidate TF for ZGA is Dux as it is specifically expressed right before major ZGA and its enforced expression in embryonic stem cells (ESC) resulted in activation of a subset of ZGA genes [28–30]. However, Dux knockout mice (both zygotic KO and maternal/zygotic KO) only exhibit minor defects on ZGA and developmental potential, and homozygous knockout mice can survive to adulthood [31*,32,33], indicating that Dux does not play a major role in regulating ZGA. A third TF that might contribute to ZGA is Yap1 because maternal-Yap1 knockouts exhibit defects in the activation of a subset of ZGA genes [34]. However, the major defects of the knockout embryos are in cell compaction and TE lineage specification instead of two-cell or four-cell arrest [34].

In addition to TFs, chromatin remodeling is known to be essential for preparing the ZGA genes to be poised for activation. In this regard, it is interesting to note that maternal depletion of the chromatin remodeling factor Brg1 results in a two-cell arrest and causes defects in ZGA [35]. Another chromatin remodeler Snf2h has been shown to affect ZGA gene expression in an *in vitro* 2CLCs system [36]. However, whether it performs a similar function *in vivo* in developing embryos remains to be shown. As discussed above, ZGA is also affected by several histone modifications [8,9,16*].

In summary, although several candidate factors for mammalian ZGA have been proposed, their role in ZGA gene activation appears to be minor, suggesting major mammalian ZGA regulators remain to be identified. In contrast, master ZGA regulators in other organisms have been identified, including Zelda in *Drosophila* [37], Pou5f1, Sox2B1, and Nanog in zebrafish [38]. These factors are unlikely to be responsible for mammalian ZGA as no mammalian Zelda homology has been identified, and the pluripotent factors involved in zebrafish ZGA either do not express in mammalian embryos until after ZGA, or has no function in mice ZGA [39]. It is possible that the mammalian master ZGA regulators might be maternal proteins or newly translated proteins activated by fertilization triggered signaling. New approaches focusing on the maternal RNAs or proteins may reveal their identities.

**Epigenetic regulation of ICM and TE lineage specification**

After ZGA, embryos go through several cell divisions before the first cell lineage separation to generate TE and ICM cells. The TE cell lineage, with the specific expression of Cdx2 and Gata3, contributes to the placenta; while the ICM cells, marked by the expression of pluripotent factors such as Oct4, Nanog, and Sox2, differentiate into epiblast and primitive endoderm, which eventually give rise to all embryonic tissues and some extraembryonic membranes, respectively [40,41]. During the first cell lineage differentiation, the Hippo pathway controls the TE lineage specification [42–44,45], while epigenetic modifications are involved in the consolidation of the ICM cell lineage [46,47]. A recent study suggested that asymmetric distribution of Carm1-deposited dimethyl-arginine 26 of histone H3 (H3R26me2) in the blastomeres is observed as early as four-cell stage embryos [48]. Blastomeres with higher levels of Carm1 and H3R26me2 have a higher tendency to contribute to the ICM [48]. In addition, a higher level of H3R26me2 increases the expression of pluripotent genes and facilitates Sox2 binding to its targets for ICM specification [49,50]. Furthermore, overexpression of Carm1 can increase the frequency of asymmetric divisions and lead cells biased to take inner position of the embryo [51]. In addition to Carm1, another chromatin modifier Pdml14 is also asymmetrically expressed in four-cell embryos to modulate H3R26me2 level for ICM contribution [52].

Two recent studies extended the above observations by demonstrating that long non-coding RNAs (lncRNAs) are involved in Carm1-mediated H3R26me2 heterogeneity in four-cell embryos [53*,54*]. Wang et al. showed that a lncRNA, LineGET, is asymmetrically expressed during two-cell to four-cell stages [53*]. LineGET interacts with Carm1 and promotes its nuclear localization, leading to an increased H3R26me2 level. Importantly, injection of LineGET into one of the blastomeres of two-cell embryos can bias the injected blastomere toward ICM specification [53*]. Another lncRNA Neat1 is also required for Carm1-dependent H3R86me2 deposition and ICM specification [54*]. Neat1 is asymmetrically expressed between blastomeres in 4-cell embryo and recruits Carm1 to nuclear foci paraspeckles [54*]. Disruption of Neat1 results in reduced H3R26me2, increased Cdx2 expression, and biased TE specification [54*]. Whether LineGET and Neat1 coordinate

*Current Opinion in Genetics & Development* 2020, 64:13–20 www.sciencedirect.com
regulate Carm1 on lineage specification merits further investigation.

Taken together, the asymmetric distribution of H3R26me2 serves as one of the first signals for the first lineage specification. The heterogeneity of H3R26me2 is modulated by the asymmetric expression of Carm1, Prdm14, LincGET, and Neat1 in early blastomeres. What causes the initial biased expression of Carm1/Prdm14/LincGET/Neat1 in two-cell and four-cell embryos is currently unknown.

Understanding ZGA and totipotency to pluripotency transition using the 2CLC system

During preimplantation development, embryos go through ZGA and transit from totipotency to pluripotency [55]. Understanding the molecular mechanism of this process is impeded by the scarcity of mammalian early embryos. The observation that cultured mouse ESCs have a rare totipotent cell population, 2CLCs, that expresses two-cell-embryo-specific genes, provides an in vitro system for understanding this process [56].

The rare cell population in cultured ESCs gained their name 2CLCs because they share several key features of two-cell embryos, including expression of a group of two-cell-embryo-specific transcripts, loss expression of pluripotent genes, and have the capacity to contribute to both TE and ICM cell lineages [55,56,57]. Further studies show that the transition from pluripotent mESC to totipotent 2CLC establishes multiple two-cell-embryo-like molecular features, including chromatin decondensation [58], increased histone mobility [59], loss of chromocenters [58], decreased chromatin accessibility [29], and formation of MERVL-mediated insulating domain boundaries [60]. Since 2CLC has many features of the totipotent two-cell embryos, they provide a unique model for studying ZGA and totipotency to pluripotency transition.

As part of the ZGA genes, two-cell-embryo-specific genes are transiently activated in two-cell embryos and quickly shutdown in four-cell embryos [55]. Factors that regulate this transcriptional dynamic were largely unknown. Since the same group of genes, including MERVL repeats, are activated during ESC to 2CLC transition, this transition provides a unique system for understanding how the two-cell embryo-specific genes are activated during ZGA (Figure 2). Recent studies in the 2CLC system have revealed a transcriptional axis that includes Dppa2/4-mediated Dux activation, which in turn activates two-cell-embryo-specific genes [28–30,61–63]. Importantly, the loss function of either Dux or Dppa2/4 partly affected activation of some two-cell-embryo-specific genes during ZGA [62–65], indicating that Dppa2/4 and Dux may have some role in ZGA, although Dux is not essential for ZGA in mouse models [31*,32,33]. Whether Dppa2/4 has an essential role in ZGA remains to be determined. Additional factors, including Patz1, Nelfa, and Gata3, have also been shown to activate two-cell-embryo-specific genes in mESCs [36,61,66], their role in ZGA remains to be shown.

Relationship of 2CLCs with mouse preimplantation development.

2CLC are a rare two-cell-embryo-like cell population in cultured mESCs, which can serve as an in vitro model for studying preimplantation development. Factors, such as Dppa2/4 and Dux, that drive 2CLC formation may contribute to ZGA; while factors that impede 2CLC formation, such as LINE1, may drive totipotency to pluripotency transition in preimplantation embryos. 2CLCs spontaneously transit back to the pluripotent state and this reversal transition can serve as a model for understanding totipotency to pluripotency transition in embryos. For instance, the rapid clearance of Dux may contribute to both the reversal of the 2C-like state as well as the totipotency to pluripotency transition in preimplantation embryos. Dashed lines indicate the interactions merit further investigation.

Figure 2

![Diagram](https://example.com/diagram.png)
In addition to ZGA, the 2CLC system can also be used for understanding cell fate reprogramming between totipotent and pluripotent states. The barriers that prevent ESC to 2CLC transition in vitro might play an important role in preimplantation development to drive totipotency to pluripotency transition (Figure 2). One example is the retrotransposon LINE1 [67]. LINE1 has been shown to correlate with Kap1 and nucleolin to suppress Dux expression and the generation of 2CLCs in mESCs. Importantly, LINE1 also suppresses Dux in embryos and promotes totipotency to pluripotency transition [67]. Similarly, histone variant H3.3 and Zmym2 prevent 2CLC generation in mESCs and plays a role in the totipotency to pluripotency transition during preimplantation development [68–70]. Several other factors, including H3K4/H3K9 methylation, histone acetylation, Tet proteins, DNA methylation, Myc, the PRC1.6 and EP400-TIP60 complexes have also been shown to inhibit 2CLC generation in the mESC culture system [56,57*,71–74]. The role of these factors in totipotency to pluripotency transition in vivo merits further investigation.

It is known that the totipotent state is transient and that the 2CLCs spontaneously transit back into pluripotent mESCs. This reversal transition can also be used for understanding totipotency to pluripotency transition during preimplantation development (Figure 2). For instance, the transcripts of Dux and many two-cell-embryo-specific transcripts are quickly disappeared in four-cell embryos and this clearance appears to be crucial for totipotent to pluripotent state transition in embryos [63,65,67*,75]. The reversal of 2C-like state is accompanied by a similar clearance of Dux and two-cell-embryo-specific transcripts [57*,75], which can serve as a model for understanding how these transcripts are quickly removed.

Although 2CLC can be used for understanding some of the molecular events during preimplantation development, this system has its limitation [33,64,65]. For example, while Dux is necessary and sufficient in inducing 2C-like transition in mESCs, Dux depletion only has a minimum effect on ZGA and embryonic development [33,64,65]. This discrepancy is not too surprising as two-cell embryos still contain many maternal transcripts and proteins that are absent in the 2CLCs. Thus, caution should be taken when extrapolating observations in the 2CLC system to preimplantation embryos.

**Concluding remarks**

Recent advances in low-input epigenomic profiling techniques have greatly facilitated our understanding of the molecular mechanism of mouse preimplantation development, particularly on ZGA and the first cell lineage specification. In addition, the development of the 2CLC model system further facilitated dissecting molecular features of preimplantation embryos. Importantly, the majority of the new findings, including the dynamics of DNA methylation, histone modifications, and chromatin accessibility, are largely conserved between human and mouse, highlighting mouse as a useful model in understanding human preimplantation development [20,76–78]. Despite the great progress, many important questions still remain unaddressed, including the identity of the master transcription factors responsible for ZGA and the mechanisms of rapid clearance of the two-cell-embryo-specific transcripts. New techniques and approaches, such as low-input profiling of RNA modifications and polyA-tails, as well as an understanding of translational and post-translational regulations during the maternal-to-zygotic transition should shed light on these questions.

**Acknowledgements**

We thank Drs. Zhiyuan Chen and Wenhao Zhang for critical reading of the manuscript. This work was supported by HHMI and N.I.H. (R01HD092465). Y.Z. is an Investigator of the Howard Hughes Medical Institute. The authors declare no conflict of interest.

**References and recommended reading**

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest


The authors report that H3K36me3 is a central epigenetic regulator mediating the deposition of DNA methylation, H3K4me3, and H3K27me3 in oocytes, and plays an essential role during preimplantation development.


This paper is the first study reporting the formation of Polycomb-associating domains (PADs) during oocyte growth and their dynamics after fertilization, as well as how they are regulated by Polycomb repressive complexes.


This is the first study of the in vivo function of Dux in mice which demonstrate that Dux plays a minor non-essential role in ZGA, highlighting the difference between 2CLC and 2-cell embryos.


33. De Iaco A et al.: DUX is a non-essential synchronizer of zygotic genome activation. Development 2020, 147 dev177725.


This paper together with reference [54] reveal that lincRNAs are involved in the regulation of lineage specification by Cram1 and H3R26me2.


This paper together with reference [54] reveal that lincRNAs are involved in the regulation of lineage specification by Cram1 and H3R26me2.


This paper reveals the transcriptomic dynamics and candidate regulators for the transition between mESCs and 2CLCs, which provides a valuable resource for using 2CLC as a model to study preimplantation development.


62. De lao A et al.: DPPA2 and DPPA4 are necessary to establish a 2C-like state in mouse embryonic stem cells. EMBO Rep 2019, 20:e47382.


This paper reveals a novel function of LINE1 in suppressing Dux both in 2CLCs and early embryos, and suggests that the clearance of Dux may be critical for totipotency to pluripotency transition.


