

Unique Patterns of H3K4me3 and H3K27me3 in 2-Cell-like Embryonic Stem Cells

Yanping Zhang,^{1,4} Yixin Huang,^{1,4} Yu Dong,¹ Xiaoyu Liu,³ Xiaochen Kou,¹ Yanhong Zhao,¹ Anqi Zhao,¹ Jiatong Sun,² Zhongqu Su,² Zongyu Li,¹ Huan Zhang,¹ Yunwei Li,¹ Shuyuan Cao,² Junhao Wei,¹ Jiqing Yin,¹ Lan Kang,³ Yawei Gao,³ Jiayu Chen,¹ Yixuan Wang,³ Chong Li,¹ Rui Gao,¹ Hong Wang,¹ Shaorong Gao,^{1,5,*} and Rongrong Le^{1,*}

¹Clinical and Translational Research Center of Shanghai First Maternity & Infant Hospital, Frontier Science Center for Stem Cell Research, School of Life Sciences and Technology, Tongji University, Shanghai 200092, China

²College of Animal Science and Technology, Shandong Key Laboratory of Animal Bioengineering and Disease Prevention, Shandong Agricultural University, Taian, Shandong 271018, China

³Institute for Regenerative Medicine, Shanghai East Hospital, Shanghai Key Laboratory of Signaling and Disease Research, School of Life Sciences and Technology, Tongji University, Shanghai 200092, China

⁴These authors contributed equally

⁵Lead contact

*Correspondence: gaoshaorong@tongji.edu.cn (S.G.), lerongrong@tongji.edu.cn (R.L.)

<https://doi.org/10.1016/j.stemcr.2021.01.020>

SUMMARY

A small subgroup of embryonic stem cells (ESCs) exhibit molecular features similar to those of two-cell embryos (2C). However, it remains elusive whether 2C-like cells and 2C embryos share similar epigenetic features. Here, we map the genome-wide profiles of histone H3K4me3 and H3K27me3 in 2C-like cells. We found that the majority of genes in 2C-like cells inherit their histone status from ESCs. Among the genes showing a switch in their histone methylation status during 2C-like transitions, only a small number acquire 2C-embryo epigenetic signatures. In contrast, broad H3K4me3 domains display extensive loss in 2C-like cells. Most of the differentially expressed genes display decreased H3K4me3 and H3K27me3 levels in 2C-like cells, whereas *de novo* H3K4me3 deposition is closely linked with the expression levels of upregulated 2C-specific genes. Taken together, our study reveals the unique epigenetic profiles of 2C-like cells, facilitating the further exploration of totipotency in the future.

INTRODUCTION

Following fertilization, extensive epigenetic reprogramming takes place and plays an essential role in the restoration of totipotency (Xu and Xie, 2018). Zygotic genome activation (ZGA) is closely linked with the epigenetic reprogramming and the acquisition of totipotency (Lu and Zhang, 2015). In mice, ZGA mainly takes place during 2-cell-embryo stage (Bouniol et al., 1995; Schultz, 1993). The disruption of the tight control of ZGA genes causes embryo developmental failure and disease pathogenesis (Falco et al., 2007; Gabriels et al., 1999; Geng et al., 2012; Guo et al., 2019). Although considerable progress has been made recently, there are still a large number of unsolved questions, and the scarcity of early embryos greatly impeded the study of the molecular basis of totipotency.

Recently, it has been reported that a rare group of ESCs are capable of transiently entering a state resembling two-cell embryos (2Cs) (Macfarlan et al., 2012; Zalzman et al., 2010). Genes and repeats that are restricted to 2C stage, such as *Dux*, *Zscan4* gene clusters, and MERVL, are highly upregulated in these so-called 2C-like cells (Macfarlan et al., 2012). In addition to showing similar transcriptome features, 2C-like cells exhibit chromatin architecture

resembling that of the 2C stage (Eckersley-Maslin et al., 2016; Ishiuchi et al., 2015). In addition to their unique molecular signature, 2C-like cells show an expanded ability to contribute to both embryonic and extraembryonic lineages (Macfarlan et al., 2012). The identification of 2C-like cells provides a new opportunity for understanding the mechanism of ZGA and the molecular basis of totipotency. The underlying principles of totipotency, cell plasticity, and lineage segregation are beginning to emerge.

To further explore the epigenetic signatures underlying the totipotency, we mapped the genome-wide profiles of H3K4me3 and H3K27me3 in 2C-like cells and conducted detailed comparisons among 2Cs, 2C-like cells, and ESCs. In general, 2C-like cells were found to retain the histone methylation status at gene promoters observed in ESCs, which differs from the situation in 2C embryos. Only a small fraction of genes acquire histone methylation marks resembling those of 2C embryos. In contrast to the preservation of histone methylation locations, extensive loss of broad H3K4me3 domains is observed in 2C-like cells. In addition, the patterns of H3K27me3 and H3K4me3 level alterations among upregulated genes are diverse in 2C-like cells. Notably, *de novo* H3K4me3 deposition is observed in the promoter regions of highly upregulated 2C genes, including *Zscan4* and *Dux*.



RESULTS

Genome-wide Profiling of H3K4me3 and H3K27me3 in 2C-like ESCs

We constructed a double reporter cell line containing a stable transgene of MERVL-tdTomato (Macfarlan et al., 2012) and a pZscan4c-EGFP fluorescent reporter (Zalzman et al., 2010) (Figure 1A). The correlation between the reporters and MERVL/Zscan4 expression was confirmed by quantitative RT-PCR (Figure 1B) and fluorescence immunostaining (Figure S1A). We first functionally characterized MERVL⁻/Zscan4⁺ cells (hereafter Zscan4⁺ cells) and MERVL⁺/Zscan4⁺ cells, which are the main 2C-like subpopulations in ESCs (Eckersley-Maslin et al., 2016). We sorted Zscan4⁺ cells and MERVL⁺/Zscan4⁺ cells by fluorescence-activated cell sorting and injected these two cell populations into eight-cell embryos respectively. The Zscan4⁺ cells contributed exclusively to the inner cell mass (ICM) of all chimeric embryos (Figures S1B and S1C). By contrast, the MERVL⁺/Zscan4⁺ cells can also contribute to the trophectoderm (in six out of 10 chimeric blastocysts) in addition to the ICM (in eight out of 10 chimeric blastocysts) (Figures S1B and S1C). We also investigated the differentiation potential of Zscan4⁺ cells and MERVL⁺/Zscan4⁺ cells into trophoblast *in vitro*. We cultured MERVL⁻/Zscan4⁻ cells, Zscan4⁺ cells, and MERVL⁺/Zscan4⁺ in TS medium for 3–5 days. Under this condition, key TSC genes were dramatically increased in differentiated MERVL⁺/Zscan4⁺ cells but not in MERVL⁻/Zscan4⁻ cells and Zscan4⁺ cells (Figure S1D). This result was consistent with previous studies that have shown that MERVL⁺/Zscan4⁺ cells show higher competency to contribute to chimera mice than Zscan4⁺ cells, although these two cell populations share similar transcriptome features (Amano et al., 2013; Eckersley-Maslin et al., 2016; Macfarlan et al., 2012).

To investigate the mechanism underlying the different competency of Zscan4⁺ and MERVL⁺/Zscan4⁺ cells, we performed chromatin immunoprecipitation sequencing (ChIP-seq) to generate the genome-wide profiles of H3K4me3 and H3K27me3 in the MERVL⁻/Zscan4⁻, Zscan4⁺, and MERVL⁺/Zscan4⁺ cell populations. We used MACS2 to call H3K4me3 and H3K27me3 peaks (Table S1). The fraction of the genome covered by H3K4me3 peaks was significantly reduced in 2C-like cells compared with that in ESCs, reaching a level similar to that in 2Cs (Figure 1C). In contrast, the proportion of the genome covered by H3K27me3 peaks was slightly increased in 2C-like cells (Figure 1C). Next, we compared peak lengths. The average H3K4me3 and H3K27me3 peak lengths in ESCs were longer than those in 2C embryos (Figure 1D). An extensive reduction of H3K4me3 peak length was observed in 2C-like cells (Figure 1D). H3K4me3 peak length in MERVL⁺/Zscan4⁺ cells was reduced to a level similar to that in 2C em-

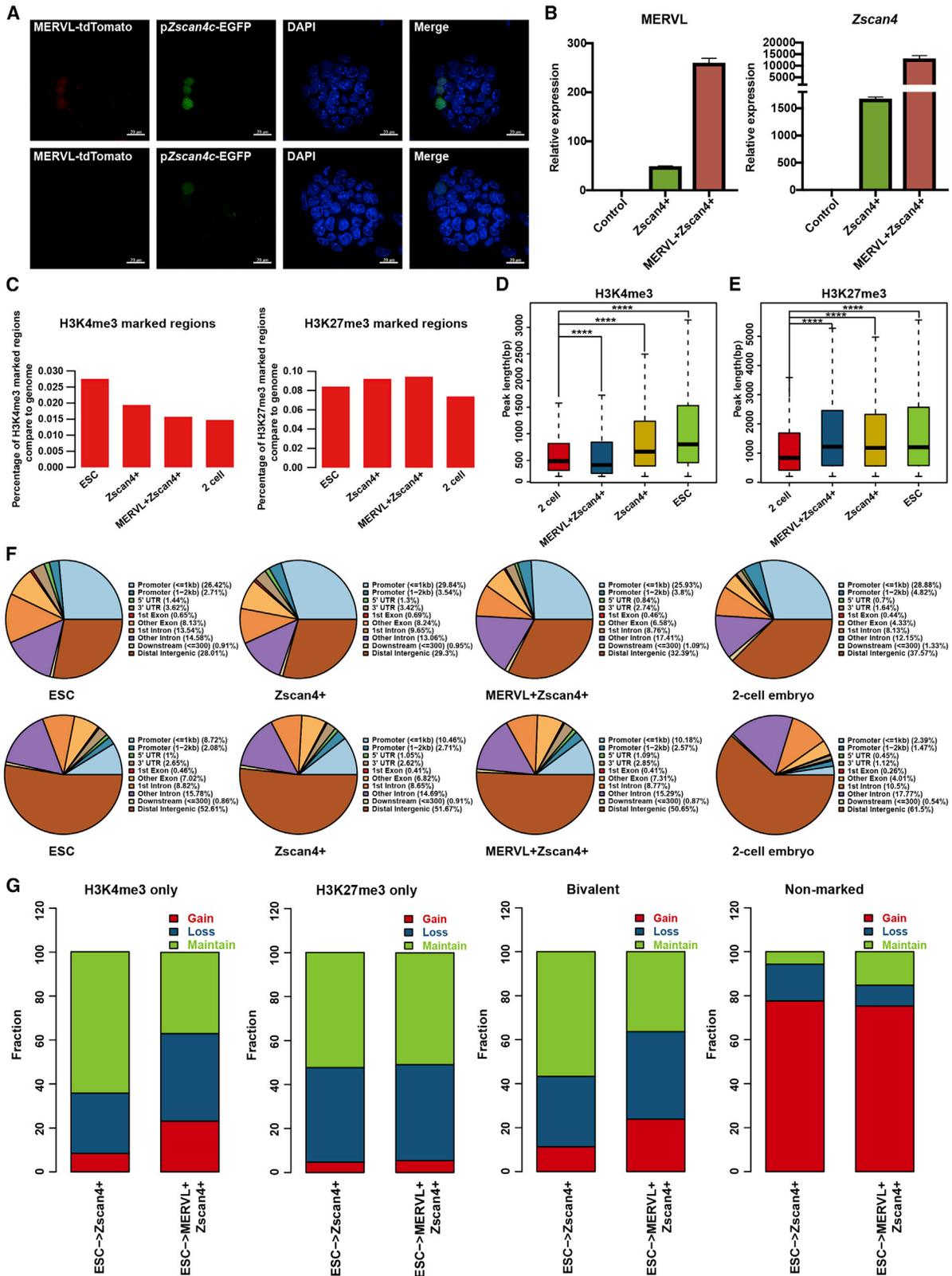
bryos, and Zscan4⁺ cells displayed an intermediate level (Figure 1D). In contrast, H3K27me3 peak lengths in 2C-like cells remained similar to those in ESCs (Figure 1E).

We next analyzed genome-wide enrichment of H3K4me3 and H3K27me3 in 2C-like cells. Peaks were assigned to genome features, including proximal promoters, 5' and 3' UTRs, exons, introns, and distal intergenic introns. ESCs, 2C-like cells, and 2C embryos shared similar H3K4me3 distribution patterns with preferential enrichment in promoters and distal intergenic regions (Figure 1F). The extent of enrichment in promoter regions was increased in 2C-like cells, displaying an intermediate level between those in ESCs and 2C embryos (Figure 1F). ESCs and 2C-like cells shared similar H3K27me3 enrichment patterns, with the greatest number of H3K27me3 peaks residing in distal intergenic regions, followed by promoter regions (more than 10%) (Figure 1F). In contrast, less than 4% of H3K27me3 peaks were found in promoter regions in 2C embryos, in line with our previous findings (Figure 1F) (Liu et al., 2016). Next, we assessed the distribution of H3K4me3 and H3K27me3 peaks with a TSS centered plot. Compared with ESCs, the proportion of H3K4me3 peaks near transcription start sites (within ±1 kb regions) was significantly increased in 2C-like cells, resembling that in 2C embryos (Figures S1E and S1F). H3K27me3 enrichment profiles showed diverse patterns in 2C-like cells. In 2C embryos, H3K27me3 peaks tended to be enriched farther upstream or downstream of the TSS (Figure S1G and S1H). Zscan4⁺ cells displayed a bimodal H3K27me3 distribution around the TSS, similar to the patterns in ESCs (Figure S1H). In contrast, H3K27me3 was enriched in the immediate vicinity of the TSS in MERVL⁺/Zscan4⁺ cells, unlike the patterns in 2C embryos and ESCs (Figure S1H).

Subsequently, the genome was classified into four chromatin states: H3K4me3-only, H3K27me3-only, bivalent, and non-marked regions. In Zscan4⁺ cells, nearly 40% of H3K4me3-only, H3K27me3-only, bivalent regions were retained and less than 10% of regions newly acquired histone methylation marks (Figure 1G). MERVL⁺/Zscan4⁺ cells exhibited a similar pattern of alteration of H3K27me3-only regions while the fractions of newly formed H3K4me3-only and bivalent regions were much higher than that in Zscan4⁺ cells (Figure 1G). Non-marked regions displayed more dynamic behavior, as nearly 80% of these regions were newly acquired during the 2C-like transition (Figure 1G).

Dynamic Histone Methylation Changes at Promoters During 2C-like Transitions

Nearly half of the altered H3K4me3 and H3K27me3 peaks were located in gene regions (Figure S1I). Thus, we focused on H3K4me3 and H3K27me3 in promoter regions. We performed principal component analysis (PCA) of promoter



(legend on next page)



regions. The PCA plot of H3K4me3 enriched promoters showed that 2C-like cells and 2Cs were clustered together (Figure 2A). In contrast, the H3K27me3 enriched promoters of 2C-like cells were clearly separated from those of ESCs and 2Cs (Figure 2A).

Next, we categorized genes into four groups based on the profiles of the two histone modifications: H3K4me3-only, H3K27me3-only, bivalent, and non-marked. The histone methylation status (H3K4me3-only, H3K27me3-only, bivalent, and non-marked) of every gene in the 2Cs and three-cell populations are provided in Table S2. A large fraction of H3K4me3-only genes (8,565 genes) were shared between 2C embryos and ESCs (Figures 2B and S2A). We found 6,841 genes showing different H3K4me3-only patterns (5,839 genes in 2C embryos and 1,002 genes in ESCs, respectively) (Figure 2B). We defined these genes as 2C-specific H3K4me3 genes or ESC-specific H3K4me3 genes (Figure 2B). A marked difference in H3K27me3-only genes was observed between ESCs and 2C embryos (Figure 2B). Only 664 H3K27me3-only genes were shared between 2C embryos and ESCs (Figure 2B). There were fewer H3K27me3-only genes in 2C embryos, as previously reported (Liu et al., 2016), and 1,863 genes showing H3K27me3-only patterns were ESC specific (Figure 2B). To address whether 2C-like cells acquire chromatin features similar to those of 2C embryos, we examined the histone methylation switch during the 2C-like transition. A total of 279 genes acquired promoter H3K4me3 in *MERVL*⁺/*Zscan4*⁺ cells and we found that 50 genes (18%) gained 2C-specific H3K4me3 (Figure 2C and Table S2). Similar observations were made in *Zscan4*⁺ cells. In *Zscan4*⁺ cells, 118 genes newly formed H3K4me3 enriched domains in promoters, among which 33 genes (18%) gained 2C-specific H3K4me3 (Figure 2C and Table S2). The erasure of promoter H3K4me3 was extensive during the 2C-like transition. A total of 524 genes lost promoter H3K4me3, including 348 genes carrying ESC-specific H3K4me3 marks during the transition from ESCs to *MERVL*⁺/*Zscan4*⁺ cells (Figure 2C and Table S2). In *Zscan4*⁺ cells, 209 genes lost promoter H3K4me3 (Figure 2C). Among these genes, 170 belonged to the ESC-specific

H3K4me3 group (Figure 2C and Table S2). H3K27me3 marked domains in promoters exhibited massive loss during 2C-like transitions. We identified 937 and 873 genes showing the loss of promoter H3K27me3 in *MERVL*⁺/*Zscan4*⁺ cells and *Zscan4*⁺ cells, respectively, most of which were ESC-specific-H3K27me3-only genes (734 genes in *MERVL*⁺/*Zscan4*⁺ cells and 684 genes in *Zscan4*⁺ cells) (Figure 2D and Table S2). Newly formed H3K27me3 domains in promoters were rare in 2C-like ESCs. Only 91 and 78 genes gained H3K27me3 domains in *MERVL*⁺/*Zscan4*⁺ cells and 684 genes in *Zscan4*⁺ cells (Figure 2D and Table S2). Since only a small fraction of 2C-specific genes are activated during 2C-like transitions, we next investigated the changes in the epigenetic fate of 2C-specific genes. As shown in Figure S2B, a majority of these genes retained histone marks in 2C-like ESCs. Next, we investigated chromatin status of 2C-specific genes during 2C-like transitions. In 2C embryos, 2C-specific genes largely belonged to the H3K4me3-only and non-marked groups (Figure S2B). The fraction of bivalent genes was dramatically increased in 2C-like cells and ESCs (Figure S2B). A small fraction of genes underwent histone mark switching during 2C-like transitions, which exhibited patterns distinct from those in 2C embryos (Figure S2B).

H3K4me3 and H3K27me3 are associated with gene activation and repression, respectively (Liu et al., 2016). Next, we investigated the correlation between histone methylation changes and 2C-specific gene activation in 2C-like cells. We reported above that a small fraction of genes gained 2C-specific promoter H3K4me3 domains in 2C-like cells. However, these two subgroups showed different expression patterns in early embryogenesis, especially at 2C stage. In early embryogenesis, this subgroup in *MERVL*⁺/*Zscan4*⁺ cells was highly upregulated in 2C embryos, while the subgroup in *Zscan4*⁺ cells displayed low expression levels in 2C embryos (Figures 2E and S2C). Notably, the average expression levels of these two subgroups were not significantly increased in both *MERVL*⁺/*Zscan4*⁺ cells and *Zscan4*⁺ cells, indicating a lack of other epigenetic regulation mechanisms or transcriptional

Figure 1. Genome-wide Profiling of H3K4me3 and H3K27me3 in 2C-like ESCs

- (A) Fluorescence of a double reporter cell line containing the stable *MERVL*-tdTomato transgene construct and the *pZscan4c*-EGFP reporter. *Zscan4* expression was visualized with the GFP-Emerald reporter gene (green), and *MERVL* expression was visualized with tdTomato. Cell nuclei were visualized with DAPI. Scale bars, 20 μ m.
- (B) qRT-PCR showing *MERVL* and *Zscan4* expression in 2C-like cells. Shown is mean \pm SD; n = 2 biological replicates.
- (C) Bar graphs showing the fraction of the genome covered by H3K4me3 and H3K27me3 in ESCs, 2C-like cells, and 2C embryos.
- (D and E) Box plots showing the average H3K4me3 peak lengths (D) and H3K27me3 peak lengths (E) in ESCs, 2C-like cells, and 2C embryos.
- (F) Pie charts showing the genome-wide distributions of H3K4me3 (top) and H3K27me3 (bottom) peaks in ESCs, 2C-like cells, and 2C embryos.
- (G) Bar graphs showing the fractions of the gains, losses and maintenance of H3K4me3-only, H3K27me3-only, bivalent, and unmarked regions in 2C-like cells. Significance was analyzed by using Student's t test (****p < 0.0001). See also Figure S1.

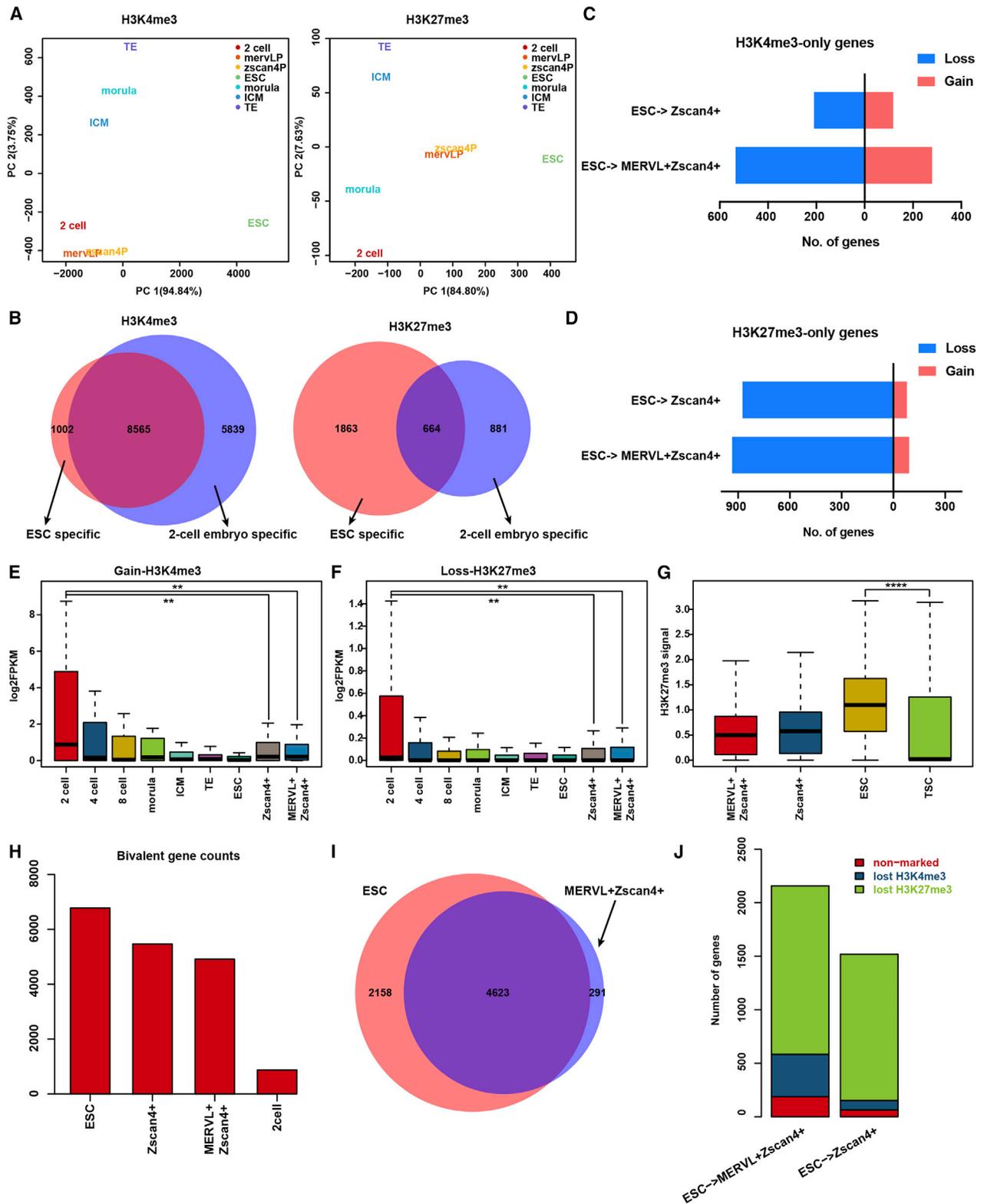


Figure 2. Dynamic Histone Methylation Changes at Promoters During 2C-like Transitions

(A) PCA plots showing the H3K4me3 and H3K27me3 signals of gene promoters in preimplantation stage and indicated cells. Zscan4P: *Zscan4*⁺ cells; mervLP: *MERVL*⁺/*Zscan4*⁺ cells.

(legend continued on next page)



regulators (Figures 2E and S2C). We also investigated the expression pattern of genes showing the loss of promoter H3K27me3 in 2C-like ESCs. Strikingly, these genes exhibited low expression levels during early embryogenesis and 2C-like transitions (Figures 2F and S2D). We and others have previously reported that H3K27me3 is absent from the promoter in early cleavage embryos and re-established in the transition from morula to ICM/TE stage (Liu et al., 2016; Zheng et al., 2016). Given that 2C-like cells acquire an expanded ability to contribute to both embryonic and ExEm lineages (Macfarlan et al., 2012), the genes showing the loss of promoter H3K27me3 during 2C-like transitions may be associated with the first cell lineage segregation. Indeed, the H3K27me3 levels of this subgroup were markedly lower in TSCs than in ESCs (Figures 2G and S2E).

In summary, neither the acquisitions of 2C-specific H3K4me3-only domains nor the loss of H3K27me3-only domains in promoter regions is associated with 2C-specific gene activation during 2C-like transitions.

Bivalent Genes Are Preserved in 2C-like ESCs

In ESCs, bivalent domains are preferentially enriched in key developmental regulators (Bernstein et al., 2006). The majority of bivalent domains are established during the transition from morula to ICM/TE stages and ESC derivation (Liu et al., 2016). The number of bivalent genes in 2C embryos was much lower than that in ESCs (Figures 2H and S2F and Table S2). During the 2C-like transition, the total number of bivalent genes was reduced (Figure 2H). The number of bivalent genes in *Zscan4*⁺ cells was reduced to a lesser extent than that in MERVL⁺/*Zscan4*⁺ cells (Figure 2H).

Newly established bivalent domains were rare in the 2C-like transition, and more than 70% of the bivalent domains in 2C-like cells were inherited from ESCs (Figures 2I and S2G). A total of 1,519 bivalent domains and 2,158 bivalent domains were lost in *Zscan4*⁺ and MERVL⁺/*Zscan4*⁺ cells respectively, compared with ESCs (Figures 2I and S2G). Bivalency loss during the 2C-like transition mostly depended on the removal of H3K27me3 (Figure 2J). We also noticed

that approximate 20% of bivalent genes lost H3K4me3 compared with less than 5% of bivalent genes in MERVL⁺/*Zscan4*⁺ cells, indicating that distinct epigenetic reprogramming occurs in *Zscan4*⁺ and MERVL⁺/*Zscan4*⁺ cells (Figure 2J).

Highly Upregulated 2C-Embryo-Specific Genes Are Correlated with Increased H3K4me3 Levels

To correlate histone methylation patterns with gene expression, we studied differentially expressed genes using publicly available RNA sequencing (RNA-seq) data from MERVL⁺/*Zscan4*⁺ and *Zscan4*⁺ cells (Eckersley-Maslin et al., 2016). By comparing the transcriptome of ESCs with that of 2C-like cells (MERVL⁺/*Zscan4*⁺, *Zscan4*⁺), we identified 1,100 upregulated genes and 1,300 downregulated genes in MERVL⁺/*Zscan4*⁺ cells (fold change > 1.5, adjusted $p < 0.005$) (Figure 3A). In *Zscan4*⁺ cells, we identified 615 upregulated genes and 900 downregulated genes (Figure S3A). A total of 570 upregulated genes were shared between MERVL⁺/*Zscan4*⁺ and *Zscan4*⁺ cells (Table S3). The majority of 2C-specific transcripts (171 genes) such as *Zscan4* gene clusters, *Zfp352* were shared between MERVL⁺/*Zscan4*⁺ and *Zscan4*⁺ cells (Table S3).

Next, we sought to study whether the changes in H3K4me3 and H3K27me3 features (localization, levels) were correlated with changes in gene expression in 2C-like ESCs. We generated an epigenetic fate map of genes that were upregulated in 2C-like cells (Figure S3B). These genes exhibited different histone marks between ESCs and 2C embryos (Figure S3B). During the 2C-like transitions, a large fraction of upregulated genes retained their histone status (Figure S3B). However, *Dux* and *Zscan4c*, which have been previously shown to play essential roles in initiating 2C-like transcriptome, acquired H3K4me3 marks in 2C-like cells (Figure 3B). These results indicated that *de novo* H3K4me3 deposition may be associated with the activation of 2C-like transcriptome.

To further explore this hypothesis, we investigated changes in H3K27me3 and H3K4me3 levels during 2C-like transitions. We divided the upregulated genes into

(B) Venn diagrams showing the fraction of H3K4me3-only and H3K27me3-only genes shared between 2C embryos and ESCs.

(C) Bar graph showing the number of genes that gained or lost H3K4me3 in their promoters in 2C-like cells.

(D) Bar graph showing the number of genes that gained or lost H3K27me3 in their promoters in 2C-like cells.

(E) Box plots showing the expression levels of MERVL⁺/*Zscan4*⁺ cell genes that gained 2C-specific H3K4me3 domains in their promoters during 2C-like transitions.

(F) Box plots showing the expression levels of MERVL⁺/*Zscan4*⁺ cell genes that lost 2C-specific H3K27me3 domains during 2C-like transitions.

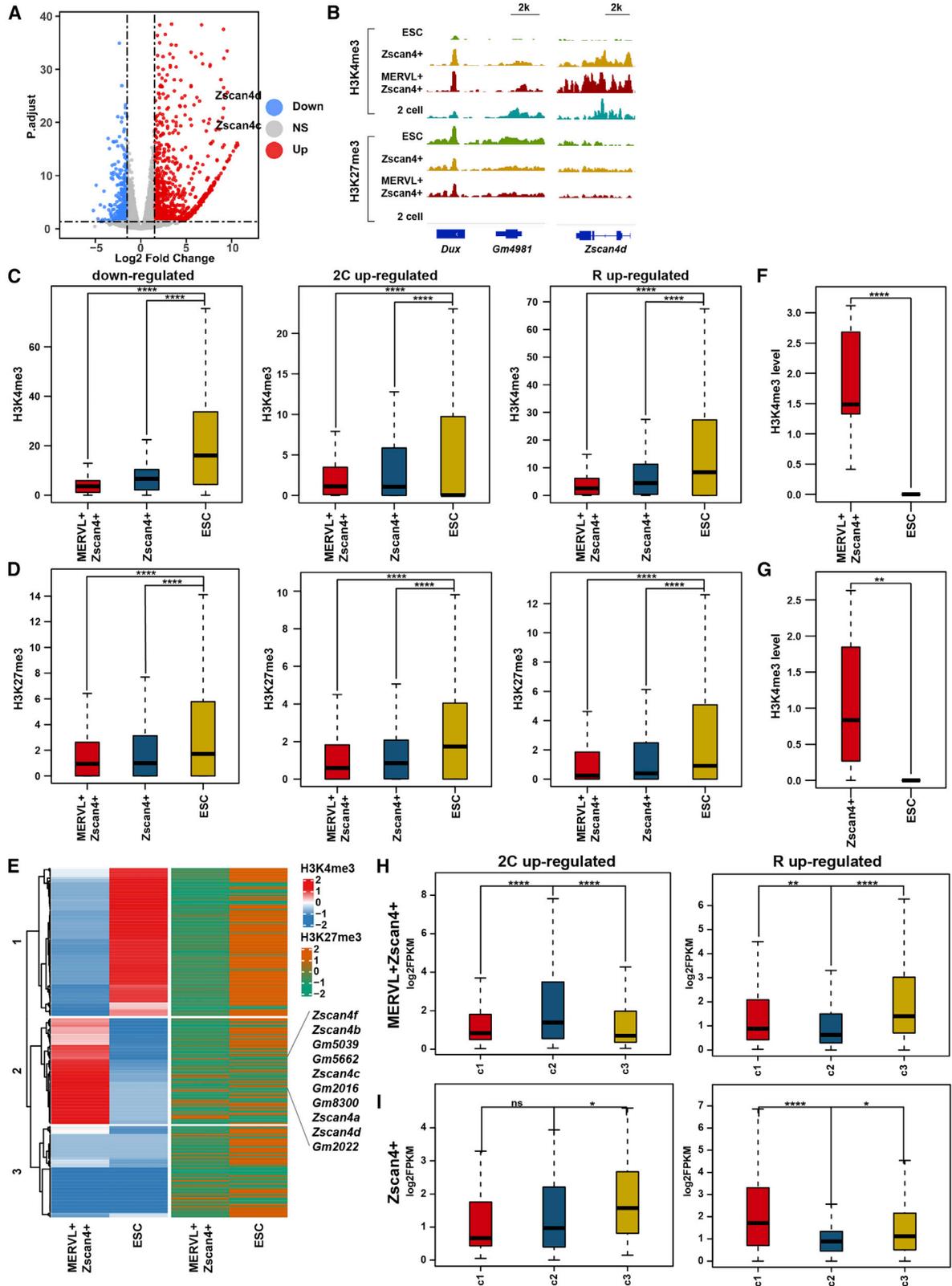
(G) Box plots showing the H3K27me3 signals in the promoter regions of MERVL⁺/*Zscan4*⁺ genes which display decreased H3K27me3 signals during 2C-like transitions.

(H) Bar graphs showing the number of bivalent genes in ESC, 2C-like cells, and 2Cs.

(I) Venn diagram showing the overlap of bivalent domains in MERVL⁺/*Zscan4*⁺ cells and ESCs.

(J) Bar graphs showing the number of genes that lost bivalent domains in 2C-like cells.

Significance was analyzed by using Student's t test. (** $p < 0.01$, **** $p < 0.0001$). See also Figure S2.



(legend on next page)



two subgroups: the 2C-specific genes (2C-up) and the remaining upregulated genes (R-up). The average H3K27me3 and H3K4me3 levels of downregulated and R-up genes were reduced in 2C-like cells (Figures 3C, 3D, S3C, and S3D). In contrast, 2C-up genes exhibited a more dramatic reduction in H3K27me3 levels and an increase in H3K4me3 levels, indicating unique epigenetic patterns of 2C-up genes (Figures 3C, 3D, S3C, and S3D). To further explore the link between H3K4me3 and 2C-genes activation in 2C-like cells, we identified 754 promoters in MERVL⁺/*Zscan4*⁺ cells and 1,194 promoters in *Zscan4*⁺ cells showing increased H3K4me3 signals compared with those in ESCs. Concomitantly, these regions exhibited decreased H3K7me3 signals in 2C-like ESCs. Interestingly, this group also exhibited markedly higher H3K4me3 levels in 2Cs, which were markedly decreased thereafter, especially at the transition from morula to ICM/TE stages (Figure S3E). In addition, this group displayed higher average expression levels in both 2C embryos and 2C-like cells (Figure S3F). Together, these data indicate that changes in H3K4me3 levels may be associated with 2C-specific genes activation in the 2C-like transitions. To investigate this hypothesis at individual-gene level, we performed cluster analysis of 2C-up genes and R-up genes based on the H3K4me3 and H3K27me3 level alterations during 2C-like transitions, which revealed three major gene groups in both gene sets (Figures 3E and S3G–S3I). H3K27me3 levels tended to reduce in the three groups while the change of H3K4me3 levels were diverse among the three groups (Figures 3E and S3G–S3I). The fractions of these three groups were different in 2C-up genes and R-up genes. The majority of R-up genes belong to the group showing reduced H3K4me3 levels (589 out of 933 genes in MERVL⁺/*Zscan4*⁺ cells, 356 out of 576 genes in *Zscan4*⁺ cells; Figures S3H–S3I, Table S3). In contrast, only 16% of R-up genes in MERVL⁺/*Zscan4*⁺ cells and 20% of R-up

genes in *Zscan4*⁺ cells exhibited increased H3K4me3 level during 2C-like transitions (Figures S3H–S3I, Table S3). The proportions of the three groups were nearly the same in the 2C-up gene sets in both MERVL⁺/*Zscan4*⁺ cells and *Zscan4*⁺ cells (Figures 3E and S3G). Most of the top highly activated genes in 2C-like ESCs were 2C-specific genes. We noticed that many of these genes belonged to the group showing increased H3K4me3 levels, such as the *Zscan4* gene clusters, Eif1A-like cluster (*Gm2016*, *Gm2022*, *Gm8300*), implying a possible connection between increased H3K4me3 and expression levels of 2C-upregulated genes during the 2C-like transitions (Figure 3E and Table S3). Indeed, the top 50 highly upregulated 2C genes displayed significantly increased H3K4me3 levels in 2C-like cells compared with ESCs (Figures 3F–3G). To further validate this hypothesis, we compared the expression levels of different groups. As expected, 2C-up genes with increased H3K4me3 levels exhibited the highest expression levels in 2C-like cells (C2 in MERVL⁺/*Zscan4*⁺ cells, C3 in *Zscan4*⁺ cells), and the dominant group of R-up genes with decreased H3K4me3 (C3 in MERVL⁺/*Zscan4*⁺ cells, C1 in *Zscan4*⁺ cells) displayed the highest expression level (Figures 3E, 3H–3I, and S3G–S3I).

In summary, the histone methylation alteration pattern was distinct in upregulated 2C-embryo-specific genes during 2C-like transitions, implying that different types of epigenetic machinery may be involved during this process.

Extensive Loss of Broad H3K4me3 in 2C-like Cells

H3K4me3 breadth is a critical epigenetic signature and broad H3K4me3 domains are preferentially enriched in genes critical for cell identity (Benayoun et al., 2014; Chen et al., 2015). We have previously reported that early embryos possess far more broad H3K4me3 domains than later-stage embryos and ESCs (Liu et al., 2016). To investigate the dynamics of H3K4me3 length in gene promoter regions

Figure 3. Increased H3K4me3 Levels Correlate with 2C-Specific Gene Upregulation

- (A) Volcano plot showing differentially expressed genes between MERVL⁺/*Zscan4*⁺ cells and ESCs. Red dots represent upregulated genes and blue dots represent downregulated genes in MERVL⁺/*Zscan4*⁺ cells compared with ESCs.
- (B) Integrative Genomics viewer of H3K4me3 and H3K27me3 signals at *Dux* and *Zscan4d* loci in ESCs, 2C-like cells and 2C embryos.
- (C) Box plots showing the changes in H3K4me3 levels of different gene sets during the 2C-like transition. 2C upregulated indicates the upregulated 2C-embryo-specific genes and R-upregulated indicates the remaining upregulated genes in MERVL⁺/*Zscan4*⁺ cells compared with ESCs.
- (D) Box plots showing the changes in H3K27me3 levels of different gene sets during the 2C-like transition.
- (E) Heatmap generated from the cluster analysis of 2C-up genes in MERVL⁺/*Zscan4*⁺ cells based on H3K4me3 and H3K27me3 level alterations during 2C-like transitions.
- (F and G) Box plots showing the H3K4me3 signals in the promoters of the top 50 highly upregulated 2C genes in MERVL⁺/*Zscan4*⁺ cells (F) and *Zscan4*⁺ cells (G) compared with ESCs.
- (H and I) Box plots showing the expression levels of 2C-upregulated and R-upregulated gene subgroups. (H) Upregulated MERVL⁺/*Zscan4*⁺ cell subgroup genes; (I) upregulated *Zscan4*⁺ cell subgroup genes.
- Significance was analyzed by using Student's t test (*p < 0.05, **p < 0.01, ***p < 0.0001, ns, not significant). See also Figure S3.

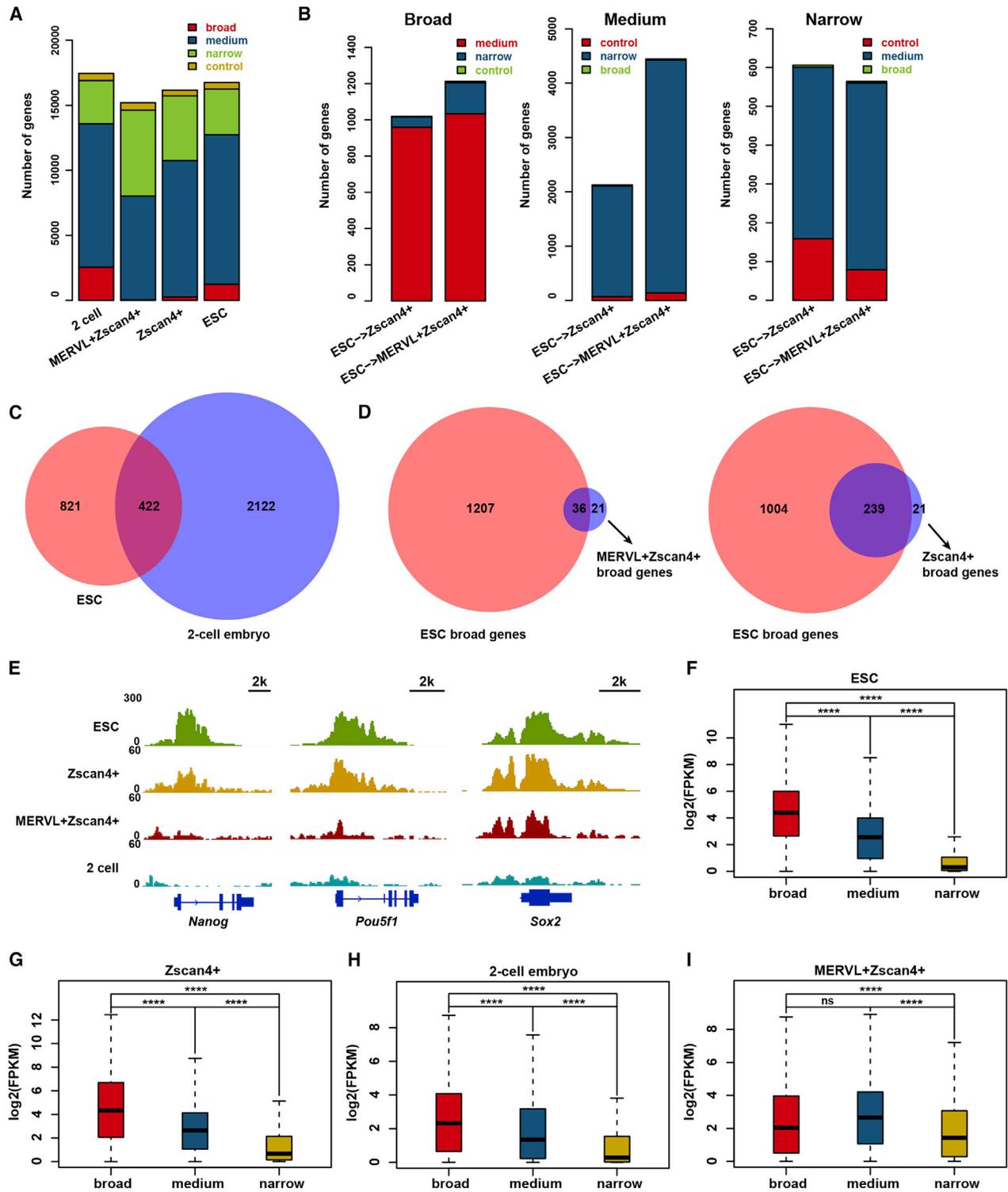


Figure 4. Extensive Loss of Broad H3K4me3 in 2C-like ESCs

(A) Bar graph showing numbers of genes with broad, medium, and narrow H3K4me3 domains in 2C embryos, 2C-like cells, and ESCs. (B) Bar graphs showing tendencies of the three types of H3K4me3 domains in promoters during the 2C-like transition. Each panel represents a specific type of H3K4me3 domain from left to right. Each bar represents the types and fractions of the promoters in the 2C-like cells. See supplemental experimental procedures for details.

(legend continued on next page)



during 2C-like transitions, we categorized the H3K4me3 domains around TSSs into three groups: broad, exceeding 5 kb; medium, 1–5 kb; and narrow, less than 1 kb. The promoter regions without H3K4me3 were used as controls. The most extensive changes occurred at medium H3K4me3 domains with a large number of medium H3K4me3 domains switching to narrow H3K4me3 domains in 2C-like cells (Figures 4A and 4B). More than 1,000 genes with broad H3K4me3 domains shrank during 2C-like transitions and the dynamics of broad H3K4me3 were observed mainly between broad and medium H3K4me3 length groups (Figures 4A, 4B, and S4A and Table S4). We found that 2C embryos possessed far more broad H3K4me3 domains than ESCs (Figures 4A, 4C, S4A, and S4B and Table S4), consistent with our previous report (Liu et al., 2016). Gene Ontology (GO) analysis revealed that genes covered with broad H3K4me3 domains were enriched with genes involved in transcription, DNA binding, and RNA metabolic processes (Figure S4C). Globally, the number of broad H3K4me3 domains was dramatically decreased in 2C-like cells and a more extensive decrease was observed in *MERVL⁺/Zscan4⁺* cells (Figures 4D and 4E). The newly formed broad H3Kme3 domains were rare in 2C-like cells and only 10 genes in *MERVL⁺/Zscan4⁺* cells and six genes in *Zscan4⁺* cells acquired 2C-specific broad H3K4me3 domains (Figures 4D, S4D, and S4E).

We then compared transcriptional activity and H3K4me3 lengths among promoters. We observed that genes marked with broad H3K4me3 domains showed the highest expression levels in ESCs, *Zscan4⁺* cells, and 2C embryos (Figures 4F–4H). Strikingly, medium H3K4me3-domain-enriched genes displayed the highest expression levels in *MERVL⁺/Zscan4⁺* cells, possibly due to the extensive loss of broad H3K4me3 domains (Figure 4I). Additionally, we also observed that genes losing broad H3K4me3 domains retained high expression levels in the 2C-like cells (Figures S4F–S4G).

DISCUSSION

Our data indicate that, in promoter regions, the H3K4me3 and H3K27me3 profiles of 2C-like cells are distinct from those of 2C embryos, more closely resembling those of ESCs.

Unlike the global loss of DNA methylation patterns observed in 2C-like cells (Eckersley-Maslin et al., 2016),

the locations of H3K4me3 and H3K27me3 are highly preserved during 2C-like transitions (Figure S2E). Intriguingly, we detected *de novo* H3K4me3 deposition in some ZGA gene promoter regions that display high H3K4me3 levels in 2C embryos. However, most of these genes fail to be reactivated in 2C-like cells, implying lack of either additional epigenetic controls or transcription regulators.

In the present study, we also compared the epigenome profiles between *Zscan4⁺* cells and *MERVL⁺/Zscan4⁺* cells. Although the transcriptomes of these two cell populations were found to be highly similar, *MERVL⁺/Zscan4⁺* cells display higher competency to contribute to chimera mice, while *Zscan4⁺* cells show lower competency than both ESCs and *MERVL⁺/Zscan4⁺* cells (Amano et al., 2013; Macfarlan et al., 2012), which is also supported by our study (Figures S1B and S1C). *MERVL⁺/Zscan4⁺* cells display more extensive loss of epigenetic memory from ESCs and the acquisition of 2C-specific epigenetic features, whereas *Zscan4⁺* cells display an intermediate status between ESCs and *MERVL⁺/Zscan4⁺* cells. Our data are consistent with previous findings showing that *Zscan4⁺* cells are in an intermediate state during 2C-like transitions (Rodriguez-Terrones et al., 2018). Thus, the incomplete acquisition of 2C-like epigenetic and transcriptome features might both account for the low competency of *MERVL⁺/Zscan4⁺* cells to contribute to chimeric blastocysts. The investigation of the determinants of the differential competency of these two types of 2C-like cells in the future may deepen the notion of cell plasticity and expanded developmental potential.

In summary, our study provides a detail comparison of epigenetic profiles among 2C embryos, 2C-like cells, and ESCs. The epigenome profiles are distinct between 2C-like cells and 2C embryos. The exploration of epigenetic factors involved in further reprogramming the epigenome of 2C-like cells in the future may help us to understand the molecular mechanisms of ZGA and totipotency.

EXPERIMENTAL PROCEDURES

Ultra-Low-Input Native ChIP-Seq

For ultra-low-input native ChIP-seq (ULI-NChIP-seq), 500 cells were used per reaction, and two or three replicates were performed for each cell type. All isolated cells were washed three times in 0.5% BSA in PBS solution (Sigma) to avoid possible contamination. The

(C) Venn diagram displaying the overlap of genes with broad H3K4me3 domains in ESCs and 2C embryos.

(D) Venn diagram displaying the overlap of genes with broad H3K4me3 domains in ESCs and 2C-like cells.

(E) Integrative Genomics viewer of H3K4me3 domains in the *Nanog*, *Pou5f1*, and *Sox2* gene regions in ESCs, 2C-like cells, and 2C embryos.

(F–I) Box plots showing the expression levels of genes marked by broad, medium, and narrow H3K4me3 domains in ESCs (F), *Zscan4⁺* cells (G), 2-cell embryos (H), and *MERVL⁺/Zscan4⁺* cells (I).

Significance was analyzed by using Student's t test (ns, not significant, ****p < 0.0001). See also Figure S4.



cells were resuspended in nuclear isolation buffer (Sigma). The ULINChIP procedure was performed as previously described (Brind'Amour et al., 2015). To obtain ChIP DNA, Chromatin was incubated with 1 mg of H3K4me3 (Cell Signaling Technology, 97275) or H3K27me3 (Diagenode, C15410195) antibody-bead complexes (5 or 10 mL of 1:1 protein A:protein G Dynabeads, Life Technologies) overnight at 4°C. Paired-end 125-bp or 150-bp sequencing was performed on the HiSeq 2500 or X10 (Illumina) at the Berry Genomics and Cloudhealth Medical Group, respectively.

Statistical Analysis

Analysis was performed in GraphPad Prism version 5.0 (GraphPad Software, United States) using Student's *t* test. Data are presented as mean \pm SD, and *p* values <0.05 were considered statistically significant.

Data and Code Availability

The ChIP-seq datasets generated during this study are available under GEO accession number GEO: GSE164486. The accession number for the published H3K4me3 and H3K27me3 ChIP-seq data of early embryos is GEO: GSE73952 (Liu et al., 2016). The accession number for the published RNA-seq data of 2C-like cells is GEO: GSE75751 (Eckersley-Maslin et al., 2016).

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at <https://doi.org/10.1016/j.stemcr.2021.01.020>.

AUTHOR CONTRIBUTIONS

R.L. and S.G. conceived and designed the experiments. Y.H. and Y.D. performed most of the experiments. Y.Z. performed the computational analysis. R.L., Y.H., and Y.Z. designed and performed the data analysis. Y.D., X.K., Y.Z., A.Z., J.S., Z.S., S.C., J.W., J.Y., L.K., Y.G., J.C., Y.W., C.L., R.G., and H.W. assisted with the experiments. R.L., and S.G. wrote the manuscript.

CONFLICTS OF INTERESTS

The authors have declared no competing interests.

ACKNOWLEDGMENTS

We are grateful to our colleagues in the laboratory for their assistance with the experiments and in the preparation of this manuscript. This work was primarily supported by the Ministry of Science and Technology of China (2016YFA0102200, 2017YFA0103302, 2016YFA0100400, 2017YFA0102602, 2018YFA0108900, 2017YFA0103300). This work was also supported by the National Natural Science Foundation of China (31970758, 31820103009, 31721003, 81630035, 31771646, 31501196, and 81525010, 32000562), the Shanghai municipal medical and health discipline construction projects (2017ZZ02015), the National Postdoctoral Program for Innovative Talents (BX201700307 and BX20170174), the China Postdoctoral Science Foundation (2017M621527), and the Fundamental Research Funds for the Central Universities (1515219049 and 22120200410).

Received: July 22, 2020

Revised: January 27, 2021

Accepted: January 27, 2021

Published: February 25, 2021

REFERENCES

- Amano, T., Hirata, T., Falco, G., Monti, M., Sharova, L.V., Amano, M., Sheer, S., Hoang, H.G., Piao, Y., Stagg, C.A., et al. (2013). Zscan4 restores the developmental potency of embryonic stem cells. *Nat. Commun.* 4, 1966.
- Benayoun, B.A., Pollina, E.A., Ucar, D., Mahmoudi, S., Karra, K., Wong, E.D., Devarajan, K., Daugherty, A.C., Kundaje, A.B., Mancini, E., et al. (2014). H3K4me3 breadth is linked to cell identity and transcriptional consistency. *Cell* 158, 673–688.
- Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., et al. (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 125, 315–326.
- Bouniol, C., Nguyen, E., and Debey, P. (1995). Endogenous transcription occurs at the 1-cell stage in the mouse embryo. *Exp. Cell Res.* 218, 57–62.
- Brind'Amour, J., Liu, S., Hudson, M., Chen, C., Karimi, M.M., and Lorincz, M.C. (2015). An ultra-low-input native ChIP-seq protocol for genome-wide profiling of rare cell populations. *Nat. Commun.* 6, 6033.
- Chen, K.F., Chen, Z., Wu, D.Y., Zhang, L.L., Lin, X.Q., Su, J.Z., Rodriguez, B., Xi, Y.X., Xia, Z., Chen, X., et al. (2015). Broad H3K4me3 is associated with increased transcription elongation and enhancer activity at tumor-suppressor genes. *Nat. Genet.* 47, 1149.
- Eckersley-Maslin, M.A., Svensson, V., Krueger, C., Stubbs, T.M., Giehr, P., Krueger, F., Miragaia, R.J., Kyriakopoulos, C., Berrens, R.V., Milagre, I., et al. (2016). MERVL/Zscan4 network activation results in transient genome-wide DNA demethylation of mESCs. *Cell Rep.* 17, 179–192.
- Falco, G., Lee, S.L., Stanghellini, I., Bassey, U.C., Hamatani, T., and Ko, M.S. (2007). Zscan4: a novel gene expressed exclusively in late 2-cell embryos and embryonic stem cells. *Dev. Biol.* 307, 539–550.
- Gabriels, J., Beckers, M.C., Ding, H., De Vriese, A., Plaisance, S., van der Maarel, S.M., Padberg, G.W., Frants, R.R., Hewitt, J.E., Collen, D., et al. (1999). Nucleotide sequence of the partially deleted D4Z4 locus in a patient with FSHD identifies a putative gene within each 3.3 kb element. *Gene* 236, 25–32.
- Geng, L.N., Yao, Z., Snider, L., Fong, A.P., Cech, J.N., Young, J.M., van der Maarel, S.M., Ruzzo, W.L., Gentleman, R.C., Tawil, R., et al. (2012). DUX4 activates germline genes, retroelements, and immune mediators: implications for facioscapulohumeral dystrophy. *Dev. Cell* 22, 38–51.
- Guo, M., Zhang, Y., Zhou, J., Bi, Y., Xu, J., Xu, C., Kou, X., Zhao, Y., Li, Y., Tu, Z., et al. (2019). Precise temporal regulation of Dux is important for embryo development. *Cell Res.* 29, 956–959.
- Ishuchi, T., Enriquez-Gasca, R., Mizutani, E., Boskovic, A., Ziegler-Birling, C., Rodriguez-Terrones, D., Wakayama, T., Vaquerizas, J.M., and Torres-Padilla, M.E. (2015). Early embryonic-like cells are



induced by downregulating replication-dependent chromatin assembly. *Nat. Struct. Mol. Biol.* *22*, 662–671.

Liu, X., Wang, C., Liu, W., Li, J., Li, C., Kou, X., Chen, J., Zhao, Y., Gao, H., Wang, H., et al. (2016). Distinct features of H3K4me3 and H3K27me3 chromatin domains in pre-implantation embryos. *Nature* *537*, 558–562.

Lu, F., and Zhang, Y. (2015). Cell totipotency: molecular features, induction, and maintenance. *Natl. Sci. Rev.* *2*, 217–225.

Macfarlan, T.S., Gifford, W.D., Driscoll, S., Lettieri, K., Rowe, H.M., Bonanomi, D., Firth, A., Singer, O., Trono, D., and Pfaff, S.L. (2012). Embryonic stem cell potency fluctuates with endogenous retrovirus activity. *Nature* *487*, 57–63.

Rodriguez-Terrones, D., Gaume, X., Ishiuchi, T., Weiss, A., Kopp, A., Kruse, K., Penning, A., Vaquerizas, J.M., Brino, L., and Torres-Pa-

dilla, M.E. (2018). A molecular roadmap for the emergence of early-embryonic-like cells in culture. *Nat. Genet.* *50*, 106–119.

Schultz, R.M. (1993). Regulation of zygotic gene activation in the mouse. *Bioessays* *15*, 531–538.

Xu, Q., and Xie, W. (2018). Epigenome in early mammalian development: inheritance, reprogramming and establishment. *Trends Cell Biol.* *28*, 237–253.

Zalzman, M., Falco, G., Sharova, L.V., Nishiyama, A., Thomas, M., Lee, S.L., Stagg, C.A., Hoang, H.G., Yang, H.T., Indig, F.E., et al. (2010). Zscan4 regulates telomere elongation and genomic stability in ES cells. *Nature* *464*, 858–863.

Zheng, H., Huang, B., Zhang, B., Xiang, Y., Du, Z., Xu, Q., Li, Y., Wang, Q., Ma, J., Peng, X., et al. (2016). Resetting epigenetic memory by reprogramming of histone modifications in mammals. *Mol. Cell* *63*, 1066–1079.