Protein Cell, 2025, **16**, 815–821
https://doi.org/10.1093/procel/pwaf015
Advance access publication 18 February 2025
Letter

Protein & Cell

Letter

Step-wise organization of genomic nuclear speckle-associated domains during mammalian embryonic development

Mengyao Kang[1,4,7,‡], Tongzhen Zhang[1,4,‡], Chao Ning[4,5,‡], Yibing Bao[2,3], Zhenbo Liu[1], Lei Gao[1], Linghui Luan[1,4], Chao Wang[2,3], Jiang Liu[1,4,6,*], Yuwen Ke[2,3,*]

[1]Key Laboratory of Epigenetic Regulation and Intervention, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China
[2]Frontiers Science Center for Molecular Design Breeding (MOE), China Agricultural University, Beijing 100193, China
[3]State Key Laboratory of Animal Biotech Breeding, China Agricultural University, Beijing 100193, China
[4]College of Life Sciences, University of the Chinese Academy of Sciences, Beijing 100049, China
[5]National Laboratory of Biomacro-molecules, CAS Center for Excellence in Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China
[6]School of Future Technology, University of Chinese Academy of Sciences, Beijing 100049, China
[7]Laboratory of Bone Tissue Engineering, Beijing Laboratory of Biomedical Materials, National Center for Orthopaedics, Beijing Research Institute of Traumatology and Orthopaedics, Beijing Jishuitan Hospital, Capital Medical University, Beijing 100035, China

[‡]These authors contributed equally.
[*]Correspondence: liujiang@ibp.ac.cn (J. Liu), keyw@cau.edu.cn (Y. Ke)

Dear Editor,

Nuclear speckles, hubs for gene expression and storage sites for RNA processing factors in typical mammalian nuclei (Chen and Belmont, 2019), concentrate numerous functional components (Gordon et al., 2021). Recent evidence suggests that genomic nuclear speckle-associated domains (SPADs) are transcriptionally active, whereas those nuclear lamina-associated domain (LADs) are inactive (Payne et al., 2021; Quinodoz et al., 2018). Disruptions in chromatin organization lead to aberrant gene expression and altered cell fate decisions (Cuartero et al., 2023; Wang et al., 2022). Notably, LAD formation during mouse embryogenesis has been shown to occur *de novo* through interactions with the nuclear lamina (Borsos et al., 2019). However, the dynamics of chromatin organization at nuclear speckles and their role in regulating gene activation during embryogenesis remain unexplored.

To investigate genome-nuclear speckle interactions in mouse embryos, we developed a low-input CUT&Tag method using an SC35 antibody targeting the nuclear speckle core protein Srrm2. This approach generated high-quality SPADs profiles across developmental stages (Figs. S1 and S2; Table S1). In PN3 zygotes, SPADs were barely detectable. Weak and broad SPADs signals emerged in PN5 zygotes, and more defined smaller SPADs became increasingly prevalent from late 2-cell stage onward (Fig. 1A). SPADs strength significantly increased from late 2-cell to 8-cell stages (Fig. 1B). SPADs signal intensity and strength in embryonic stem cells (ESCs) were comparable to those at blastocyst stage but markedly higher in the prefrontal cortex (PFC) (Fig. 1A and 1B). SPADs genomic coverage gradually expanded during development, consistently exhibiting a mean size exceeding 1 Mb (Fig. S3A). Before 8-cell stage, SPADs underwent significant reorganization, marked by transitions between SPADs and adjacent nonSPADs regions. These transitions could be categorized into seven distinct types (Figs. 1A, S3B, S3C; Table S1). Overall, SPADs undergo profound reorganization following fertilization.

In mice, major zygotic genome activation (ZGA) occurs at late 2-cell stage. Notably, SPADs established at PN5 and early 2-cell stages formed larger domains (Fig. 1A). We classified SPADs formed before ZGA as primary SPADs (pSPADs) and those formed after ZGA as secondary SPADs (sSPADs). The proportion of pSPADs remained above 60% at all stages with minimal decreases throughout development (Fig. 1C). Compared to sSPADs, pSPADs

Accepted 24 January 2025.

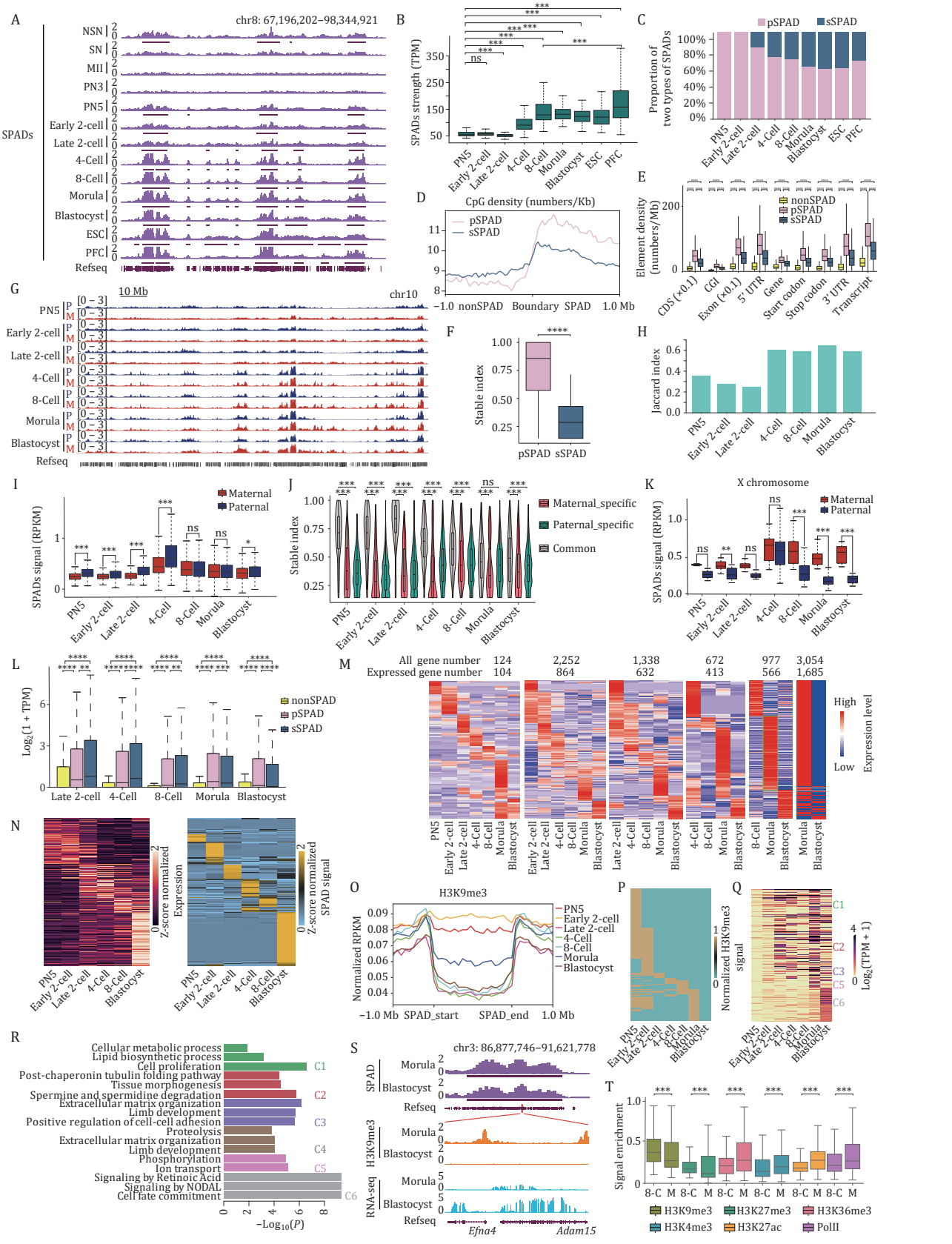

Figure 1. Dynamic SPADs in embryonic development. (A) Snapshot of speckle-associated domains (SPADs) in mouse GV-NSN oocytes (NSN), mouse GV-SN oocytes (SN), mouse mature oocytes (MII), mouse embryos, mouse embryonic stem cell (ESC) and mouse prefrontal cortex (PFC), pooled biological replicates; $n = 2$. Below bars represent SPADs. (B) The SPAD signal strength comparison for different stages (see Methods). The middle lines in the boxes represent medians. Box hinges indicate the 25th and 75th percentiles,

exhibited broader domains, higher CpG density, lower AT content, greater enrichment for genomic elements, higher stable index and higher conservation score (Figs. 1D–F and S3D–F). These findings indicate that SPADs formed before and after ZGA differ significantly in their characteristics.

We further explored parental differences in SPADs. The paternal SPADs were distinct from maternal SPADs at PN5 and 2-cell stages (Figs. 1G, 1H, S4A and S4B). After 4-cell stage, paternal and maternal SPADs became highly correlated (Figs. 1G, 1H, S4A and S4B). Consistently, paternal SPADs showed stronger signals at early stages (Fig. 1I). Nonetheless, distinctions persisted in certain imprinted regions, such as the non-canonical imprinted gene *Jade1* (Inoue et al., 2017) (Fig. S4C). At PN5 and 2-cell stages, most SPADs were allelic-specific. From the 4-cell stage onward, common SPADs accounted for over 60% of the total (Fig. S5A). Common SPADs were longer, had higher CpG density, and exhibited greater stability than allelic-specific SPADs (Figs. 1J, S5B and S5C). Notably, the majority of common SPADs were classified as pSPADs (Fig. S5D). SPADs organization also exhibited distinct patterns on the X chromosome. SPADs signals on the maternal X chromosome gradually increased, while those on the paternal X peaked at 4-cell stage before declining (Figs. 1K and S5E). This trend aligns with the paternal X inactivation observed during mouse embryogenesis (Lee and Bartolomei, 2013). Overall, parental SPADs differ in early stages and converge in late stage during embryogenesis.

We then investigated the relationship between SPADs establishment and gene expression. The genes in SPADs exhibit higher expression levels compared to those in nonSPADs (Fig. 1L). Notably, minor and major ZGA genes, housekeeping genes, and numerous transposable elements were significantly enriched in SPADs, particularly in pSPADs (Fig. S6A and S6B). These findings suggest that pSPADs create a favorable environment for early embryonic gene expression.

Genes in newly formed SPADs of each stage exhibited sequential expression patterns and were associated with stage-specific signaling pathways (Figs. 1M and S6C). Most stage-specific expression genes were localized within stage-specific SPADs (Fig. 1N). These data suggest that SPADs are pre-conFig.d and primed for gene activation. Critical genes involved in triggering ZGA and lineage segregation, such as *Klf17*, *Nfya*, *Yy1*, and *Obox* genes including *Obox1*, *Obox2*, *Obox3*, and *Obox5* (Kravchenko and Tachibana, 2025), showed significant enrichment of SPADs binding sites, particularly within pSPADs (Fig. S6D and S6E). These results highlight that SPADs

whiskers indicate the hinge ± 1.5 × interquartile range. (C) Proportion of primary SPADs (pSPADs) and secondary SPADs (sSPADs) at different stages. (D) Average CpG density around the boundaries (± 1.0 Mb) of pSPADs and sSPADs. (E) Boxplots of different genomic elements in pSPADs, sSPADs and nonSPAD regions. The middle lines in the boxes represent medians. Box hinges indicate the 25th and 75th percentiles, whiskers indicate the hinge ± 1.5 × interquartile range. (F) The stability analysis of pSPADs and sSPADs by defined stable index (see Methods). The middle lines in the boxes represent medians. Box hinges indicate the 25th and 75th percentiles, whiskers indicate the hinge ± 1.5 × interquartile range. (G) Snapshot of SPADs in parental genomes from PN5 to blastocyst stage. (H) Measure of the similarity between paternal and maternal SPAD regions in embryos by Jaccard index (see Methods). (I) Box plots show SPAD signal (RPKM) differences between paternal and maternal genomes. The middle lines in the boxes represent medians. Box hinges indicate the 25th and 75th percentiles, whiskers indicate the hinge ± 1.5 × interquartile range. (J) The stability analysis of maternal_specific, paternal_specific and common SPAD regions by defined stable index (see Methods). Violin plots show median (black bars), quartiles (upper and lower borders in the white box), and density estimation (the fatter the image, the more concentrated the data) for each distribution. (K) Box plots show X chromosome SPAD signal (RPKM) differences between paternal and maternal genomes. The middle lines in the boxes represent medians. Box hinges indicate the 25th and 75th percentiles, whiskers indicate the hinge ± 1.5 × interquartile range. (L) Expression values as $\log_2(1 + \text{TPM})$ of genes locating at pSPADs, sSPADs and nonSPADs in embryos. (M) The expression level of genes on the newly formed SPADs of each stage at current stage or later stage. Top: all gene number means the total genes locating on the newly formed SPADs; expressed gene number means the total genes which expression at current stage or later stage. Bottom: Heat maps show relative genes expression at the current stage or later stage. The color bar identifies high expression level (red) and low expression level (blue). (N) Heat maps show the expression of stage-specific genes (left heatmap) and the SPADs enrichment signals in stage-specific genes vicinity (right heatmap). The color bar identifies high expression level (white) and low expression level (black), and high SPAD signal (yellow) and low SPAD signal (light blue). (O) Line plots show average H3K9me3 signals at SPADs in embryos. (P) Heatmap shows the H3K9me3 signals on genes at each stage. Only the genes locating in SPADs and covering by H3K9me3 are included in this analysis. The color bar identifies high H3K9me3 signal (yellow) and low H3K9me3 signal (light blue). (Q) Heatmap shows the expression levels of genes corresponding in (P). The color bar identifies high expression level (dark red) and low expression level (light yellow). (R) Gene ontology analysis for the genes (C1–C6) corresponding in (Q). $P < 0.05$, *P*-values adjusted for multiple comparisons. (S) Example shows that SPADs signals (purple tracks), SPADs (purple bars) and H3K9me3 signals (orange track) around *Efna4* and *Adam15*, and their expression level (light blue track) at morula and blastocyst. (T) Bar plots show the dynamics of epigenetic features signals on the 293 genes. The 293 genes locate at 8-cell new formed SPADs, highly expressing at morula, showing in (M) (Green box line). The middle lines in the boxes represent medians. Box hinges indicate the 25th and 75th percentiles, whiskers indicate the hinge ± 1.5 × interquartile range. 8-C, 8-cell; M, Morula. In (B, E, F, I, J, K, L, T),$P \le 0.0001$; *** $P \le 0.001$; ** $P \le 0.01$; * $P \le 0.05$; ns, not statistically significant; *P*-values are calculated by two-sided Wilcoxon rank-sum test.

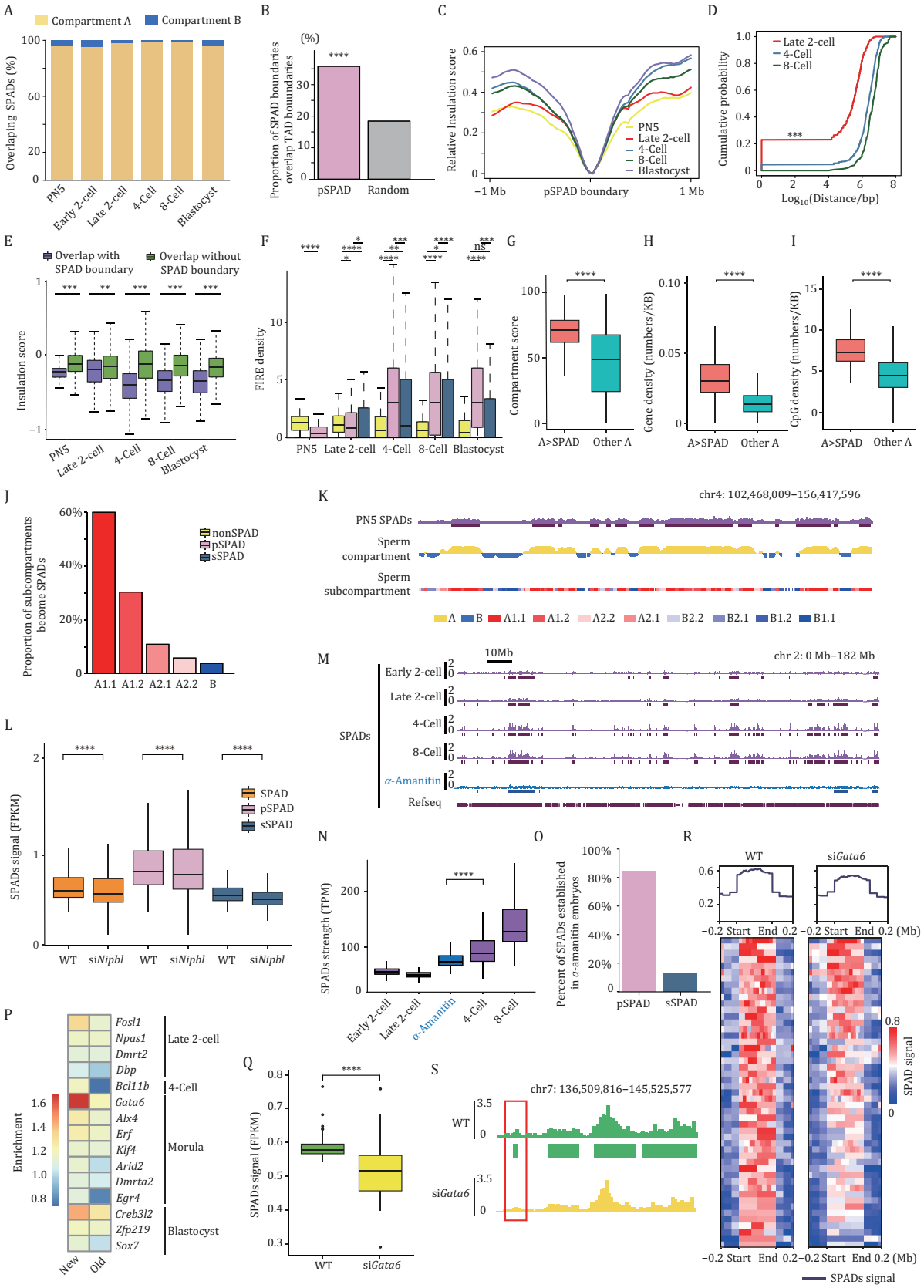

Figure 2. Key factors contribute for SPADs formation in embryos. (A) Bar plots show that the proportion of SPADs overlapping A or B compartments in embryos. (B) Bar plots show that the proportion of pSPAD boundaries overlapping blastocyst topologically associated domains (TAD) boundaries. **** $P \leq 0.0001$; *P*-value is calculated by chi-square test. (C) Line plots show relative insulation scores of PN5 SPAD boundaries at each stage. The score reflects the aggregate of interactions occurring across each genomic interval,

establishment is intricately linked to stage-specific gene activation, fulfilling developmental requirements.

We next explored the correlation between SPADs and epigenetic features in mouse embryos. In general, SPADs were negatively correlated with repressive epigenetic features such as DNA methylation, LADs signals, H2AK119ub1 (H2Aub), and H3K27me3, but they were positively correlated with active epigenetic features such as DNase I hypersensitive sites (DHSs), higher replication timing scores, RNA polymerase II (pol II), H3K36me3, H3K4me3, and H3K27ac (Figs. S7, S8, and S9). These correlations were more pronounced in pSPADs than sSPADs at blastocyst stage. Although, the SPADs are largely inversely related to repressive mark H3K9me3 (Fig. 1O), we observed some H3K9me3 signals within SPADs, and these signals were sequentially removed during development (Fig. 1P), accompanied by the associated genes activation (Fig. 1Q). These associated genes played critical roles in stage-specific biological processes (Fig. 1R), such as *Efna4* and *Adam15* (Fig. 1S). We further analyzed 293 genes (Table S2) located in SPADs newly formed at 8-cell stage (green box in Fig. 1M). These genes showed a decrease in repressive mark enrichment and an increase in active mark enrichment from 8-cell to morula stage (Fig. 1T), consistent with their higher expression at morula. These findings suggest that SPADs interact with other epigenetic features to finely regulate gene expression.

By integrating our Hi-C data (Ke et al., 2017) with SPADs profiles, we further investigated the relationship between SPADs and 3D chromatin structure. As expected, over 95.5% of SPADs overlapped with A compartments at each stage (Fig. 2A). When TADs are obscure, the pSPADs are already established at the PN5 stage. We found approximately 35% of pSPADs boundaries overlapped with blastocyst TAD boundaries (Fig. 2B). As TAD boundaries became clearer during development (Fig. S10A), insulation scores around pSPADs boundaries were lower than other regions and became more pronounced (Fig. 2C). The smallest distance between pSPADs boundaries and TAD boundaries were observed at the late 2-cell stage (Fig. 2D). Boundaries shared by SPADs and TADs exhibited significantly lower insulation scores compared to other boundaries (Figs. 2E and S10B). Collectively, these data suggest that pre-defined pSPADs might contribute to TAD structure establishment during embryogenesis. We also found that the proportion of frequently interacting regions (FIREs) overlapping with SPADs increased during development, with significant enrichment of FIREs within pSPADs after 4-cell stage (Figs. 2F, S10C and S10D), which indicates that SPADs play a role in FIREs organization. Together, the SPADs are closely associated with 3D chromatin structure during development.

Comparing the A/B compartments of sperm with zygotic SPADs, we observed that 97.53% of zygotic SPADs overlapped with sperm A compartments (Fig. S11A). SPADs also exhibited higher compartment scores, with constant SPADs consistently persisting as A compartments (Fig. S11B and S11C). Regions in sperm A compartments that formed SPADs in the zygote had higher compartment scores, gene density, and CpG density than regions without SPADs formation (Fig. 2G–I). Further subdivision of sperm A compartments into four sub-compartments (see Methods) revealed that the A1.1 sub-compartment had the largest overlap with PN5 SPADs (Fig. 2J and 2K). These findings suggest that sperm A compartments are closely associated with the formation of PN5 SPADs shortly after fertilization. In embryos with a certain extent knockdown of the chromatin

and minima of the insulation profile denote areas of high insulation are classified as TAD boundaries (see Methods). (D) Cumulative probability curves for the distance of TAD boundaries forming at late 2-cell, 4-cell and 8-cell to the closest SPAD boundaries.*** $P \le 0.001$; *P*-values is calculated by Kolmogorov-Smirnov (KS) test. (E) Box plots show the insulation scores of TAD boundaries which overlap with SPAD boundaries or do not overlap with SPAD in embryos. (F) Frequently interacting regions (FIREs) are regions of unusually high local contact frequency (see Methods). Box plots show the enrichment of FIREs in pSPADs, sSPADs, and nonSPADs in embryos. (G) Box plots show compartment score in two types of sperm A compartments. One type can become SPADs at PN5 stage (red) and the other can't (cyan). The compartment score is defined as in Methods. (H) Box plots show gene density in two types of sperm A compartments, relative in (G). (I) Box plots show CpG density in two types of sperm A compartments, relative in (G). (J) Proportion of sperm sub-compartments overlapping with the PN5 SPADs. (K) Example show relationship between sperm sub-compartment and PN5 SPADs. Purple track and purple bars, PN5 SPAD signals. Yellow and blue track, sperm A/B compartments. red and blue color bars, sperm sub-compartments. (L) Box plots show the pSPADs, sSPADs and nonSPADs signal (FPKM) in the wide type (WT) embryos and *Nipbl* knockdown (si*Nipbl*) embryos. (M) Snapshot of SPADs in mouse embryos and α-amanitin treated embryos. (N) SPAD strength variance of α-amanitin treated embryos compares with those of early embryos. (O) Bar plots show the proportion of pSPADs and sSPADs remained after α-amanitin treatment. (P) Heatmap shows motif enrichments of putative speckle targeting motif (STM) genes which highly express after ZGA in stage-specific newly formed SPADs. The color bar identifies high enrichment (red) and low enrichment (blue). (Q) Box plots show the SPAD signal (FPKM) in the WT embryos and *Gata6* knockdown (si*Gata6*) embryos. (R) Up, line plots show signals strength of SPADs in WT embryos and *Gata6* knockdown (si*Gata6*) embryos. Down, heat maps for SPADs signals, centered at SPADs. The color bar identifies high SPADs signals (red) and low SPADs signals (blue). (S) Example shows that SPADs signals in WT embryos (green tracks) and *Gata6* knockdown (si*Gata6*) embryos (yellow track). The morula stage newly formed SPADs are label by red rectangle. In (E, F, G, H, I, L, N, Q), The middle lines in the boxes represent medians. Box hinges indicate the 25th and 75th percentiles, whiskers indicate the hinge ± 1.5 × interquartile range. **** $P \le 0.0001$; *** $P \le 0.001$; ** $P \le 0.01$; * $P \le 0.05$; ns, not statistically significant; *P*-values are calculated by two-sided Wilcoxon rank-sum test.

structure organization factor *Nipbl* (Gao et al., 2018) (Fig. S11D and S11E), sSPADs signals were significantly reduced, while pSPADs were not affected (Fig. 2L). This highlights the role of chromatin structural proteins in stage-specific SPADs formation.

To investigate the impact of transcription in SPADs establishment, we treated embryos with α-amanitin efficiently (Lallena, 1997) (Fig. S12A–C). SPADs signals and strength appeared weaker in α-amanitin treated embryos (Fig. 2M and 2N). Notably, most SPADs remaining in α-amanitin-treated embryos were classified as pSPADs (Fig. 2O). Additionally, following α-amanitin inhibition, virtually no maternal SPADs are formed, whereas a few paternal SPADs still emerge (Fig. S12D and S12E). Together, the data indicate that the establishment of sSPADs and maternal SPADs predominantly relies on ZGA.

We next explored the association between SPADs formation and the putative speckle targeting motif (STM) factors, which is essential for chromatin-nuclear speckle interactions (Yu et al., 2023). Focusing on STM genes with known DNA-binding motifs (Table S2), we found that most of them displayed stage-specific expression patterns (Fig. S12F). STM proteins coded from maternally deposited transcripts consistently exhibited binding motif enrichment in pSPADs (Fig. S12G), such as the speckled protein Sp110 (Fraschilla and Jeffrey, 2020) (Fig. S12H and S12I). These findings suggest that maternally deposited STM transcripts may participate in pSPADs formation before ZGA. The STM genes expressed after ZGA (Table S2) showed binding motif enrichment in stage newly formed SPADs, including *Gata6* (Schrode et al., 2014) (Fig. 2P). In morula embryos with efficient *Gata6* knockdown (Fig. S12J and S12K), signals of newly formed SPADs were significantly reduced (Fig. 2Q–S), while pre-existing SPADs were minimally affected. Collectively, these results indicate that a set of STM genes may participate in stage-specific SPADs formation.

In this study, we generated high-resolution SPADs maps of mouse embryos for the first time. The SPADs are re-established post-fertilization with distinct allelic features until 4-cell stage. The comprehensive analysis suggests dynamic SPADs formation play critical roles in the regulation of gene expression during embryonic development. SPADs were shown to pre-configure gene expression during embryonic development. Interestingly, the heterochromatin marker H3K9me3 within SPADs plays a precise role in tuning gene expression. We also confirmed the closely interplay between chromatin high-order structure and SPADs. Finally, we demonstrated that RNA transcription significantly influences SPADs formation, particularly stage-specific STM genes like the primitive endoderm transcription factor *Gata6* involve in stage-specific SPAD formation. This study provides novel insights into the complex epigenetic network regulating embryonic gene expression and highlights the transition of chromatin from a disordered state to an organized structure post-fertilization.

Supplementary data

Supplementary data is available at *Protein & Cell* online https://doi.org/10.1093/procel/pwaf015.

Footnotes

We are very grateful to Shuqin Jia for the preparation of experimental materials and research assistant services from the Key Laboratory of Epigenetic Regulation and Intervention, Institute of Biophysics, CAS. We thank Jinzhou Hu for collected samples and staining from Key Laboratory of Epigenetic Regulation and Intervention, Institute of Biophysics, CAS.

All authors reviewed and approved the paper. J.L. and Y.K. conceived the study. T.Z., M.K. and C.N. facilitated its designs. M.K. and Y.B. collected samples. M.K. and L.G. performed drugs treatment and staining. T.Z. performed CUT&Tag. T.Z. and C.N. performed the bioinformatics analyses. Z.L. and L.L. performed RNA-seq and Q-PCR. J.L., Y.K., T.Z., M.K., C.N., and C.W. interpreted the data. J.L., Y.K., T.Z., M.K., and C.N. wrote the paper with the assistance of other authors.

The authors declare no conflict of interest. All the authors agreed to participate in this paper and publish this manuscript.

This work was supported by grants from the Biological Breeding-National Science and Technology Major Project (2023ZD04075), the National Key Research and Development Program of China (2022YFC2703302), the National Natural Science Foundation of China (32170587), and the 2115 Talent Development Program of China Agricultural University.

This study was approved by the Institutional Animal Care and Use Committee (IACUC) of the Institute of Biophysics, CAS, Beijing, China. All institutional and national guidelines for the care were carefully followed.

All data supporting the results can be found in this paper and its Supplementary Materials. All data from this study have been deposited to GSA with the accession number CRA014840. Raw image files used in the Figs that support the findings of this study are available from the corresponding authors upon reasonable request. The codes employed for the analyses reported in this study are available at “github”.

References

Borsos M, Perricone SM, Schauer T *et al*. Genome-lamina interactions are established *de novo* in the early mouse embryo. *Nature* 2019;**569**:729–733.

Chen Y, Belmont AS. Genome organization around nuclear speckles. *Curr Opin Genet Dev* 2019;**55**:91–99.

Cuartero S, Stik G, Stadhouders R. Three-dimensional genome organization in immune cell fate and function. *Nat Rev Immunol* 2023;**23**:206–221.

Fraschilla I, Jeffrey KL. The Speckled Protein (SP) family: immunity's chromatin readers. *Trends Immunol* 2020;**41**:572–585.

Gao D, Zhu B, Cao X *et al*. Roles of NIPBL in maintenance of genome stability. *Seminars in Cell & Developmental Biology* 2018;**90**:181–186.

Gordon JM, Phizicky DV, Neugebauer KM. Nuclear mechanisms of gene expression control: pre-mRNA splicing as a life or death decision. *Curr Opin Genet Dev* 2021;**67**:67–76.

Inoue A, Jiang L, Lu F *et al*. Maternal H3K27me3 controls DNA methylation-independent imprinting. *Nature* 2017;**547**:419–424.

Ke Y, Xu Y, Chen X *et al*. 3D chromatin structures of mature gametes and structural reprogramming during mammalian embryogenesis. *Cell* 2017;**170**:367–381.e20.

Kravchenko P, Tachibana K. Rise and SINE: roles of transcription factors and retrotransposons in zygotic genome activation. *Nat Rev Mol Cell Biol* 2025;**26**:68–79.

Lallena MJ, Correas I. Transcription-dependent redistribution of nuclear protein 4.1 to SC35-enriched nuclear domains. *J Cell Sci* 1997;**110**:239–247.

Lee JT, Bartolomei MS. X-inactivation, imprinting, and long noncoding RNAs in health and disease. *Cell* 2013;**152**:1308–1323.

Payne AC, Chiang ZD, Reginato PL *et al*. *In situ* genome sequencing resolves DNA sequence and structure in intact biological samples. *Science* 2021;**371**:eaay3446.

Quinodoz SA, Ollikainen N, Tabak B *et al*. Higher-order inter-chromosomal hubs shape 3D genome organization in the nucleus. *Cell* 2018;**174**:744–757.e724.

Schrode N, Saiz N, Di Talia S *et al*. GATA6 levels modulate primitive endoderm cell fate choice and timing in the mouse blastocyst. *Dev Cell* 2014;**29**:454–467.

Wang Y, Elsherbiny A, Kessler L *et al*. Lamin A/C-dependent chromatin architecture safeguards naive pluripotency to prevent aberrant cardiovascular cell fate and function. *Nat Commun* 2022;**13**:6663.

Yu R, Roseman S, Siegenfeld AP *et al*. CTCF/cohesin organize the ground state of chromatin-nuclear speckle association. bioRxiv. 2023.