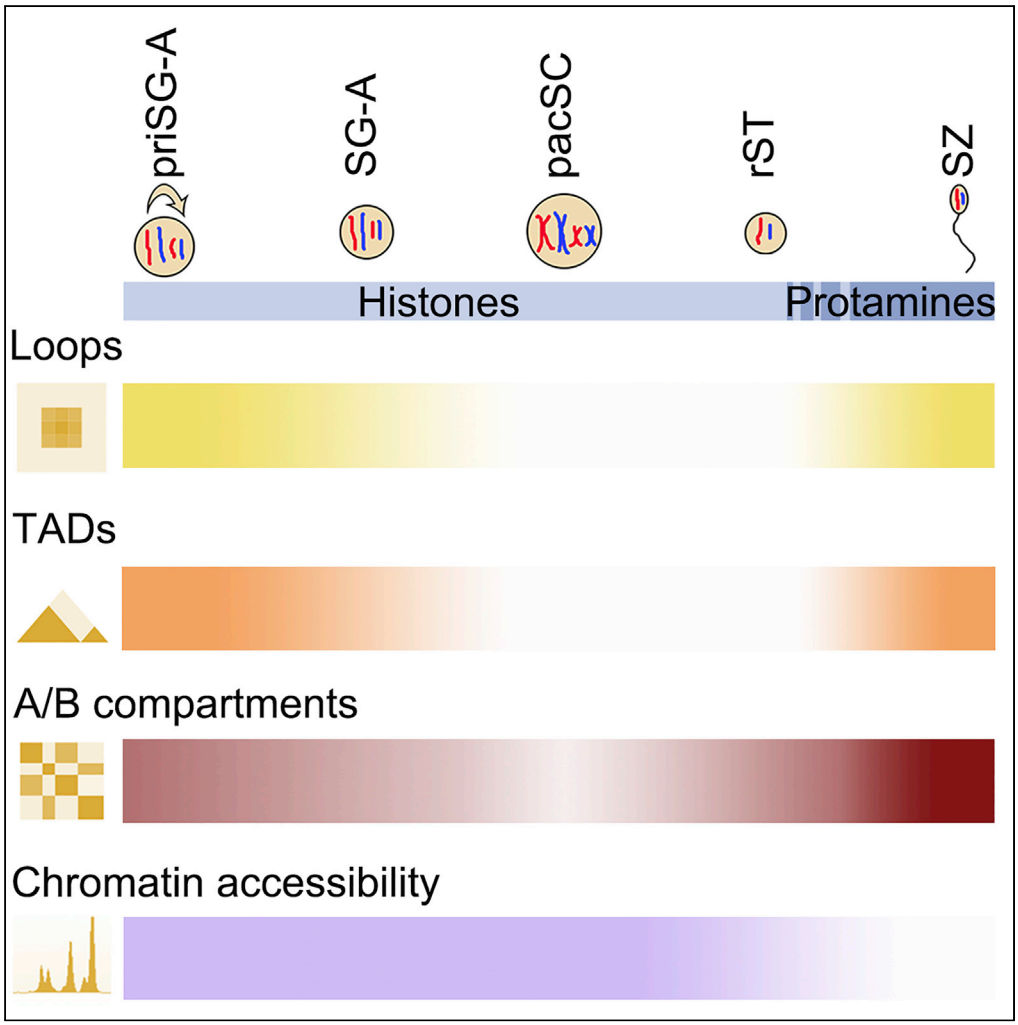
iScience

CellPress

Article

Reorganized 3D Genome Structures Support Transcriptional Regulation in Mouse Spermatogenesis



Zhengyu Luo, Xiaorong Wang, Hong Jiang, ..., Michael G. Rosenfeld, Fei Sun, Xiaoyuan Song

sunfei@ntu.edu.cn (F.S.)
songxy5@ustc.edu.cn (X.S.)

HIGHLIGHTS

Chromatin loops are reorganized during mouse spermatogenesis, being absent in pacSC

CTCF and cohesin remain bound to pacSC chromatin while TADs and loops are lost

Chromatin accessibility per se is not involved in the loss of TADs or loops in pacSC

A/B compartments switching is related to meiosis-specific mRNAs and piRNAs expression

Luo et al., iScience *23*, 101034
April 24, 2020
https://doi.org/10.1016/j.isci.2020.101034

iScience

CellPress

Article

Reorganized 3D Genome Structures Support Transcriptional Regulation in Mouse Spermatogenesis

Zhengyu Luo,[1,11] Xiaorong Wang,[2,11] Hong Jiang,[1,11] Ruoyu Wang,[1,3,4,11] Jian Chen,[5,6] Yusheng Chen,[6,7] Qianlan Xu,[1] Jun Cao,[1] Xiaowen Gong,[8] Ji Wu,[8] Yungui Yang,[6,7] Wenbo Li,[3,4] Chunsheng Han,[5,6] C. Yan Cheng,[9] Michael G. Rosenfeld,[10] Fei Sun,[2,*] and Xiaoyuan Song[1,12,*]

SUMMARY

Three-dimensional chromatin structures undergo dynamic reorganization during mammalian spermatogenesis; however, their impacts on gene regulation remain unclear. Here, we focused on understanding the structure-function regulation of meiotic chromosomes by Hi-C and other omics techniques in mouse spermatogenesis across five stages. Beyond confirming recent reports regarding changes in compartmentalization and reorganization of topologically associating domains (TADs), we further demonstrated that chromatin loops are present prior to and after, but not at, the pachytene stage. By integrating Hi-C and RNA-seq data, we showed that the switching of A/B compartments between spermatogenic stages is tightly associated with meiosis-specific mRNAs and piRNAs expression. Moreover, our ATAC-seq data indicated that chromatin accessibility *per se* is not responsible for the TAD and loop diminishment at pachytene. Additionally, our ChIP-seq data demonstrated that CTCF and cohesin remain bound at TAD boundary regions throughout meiosis, suggesting that dynamic reorganization of TADs does not require CTCF and cohesin clearance.

INTRODUCTION

The process of chromatin organization in three-dimensions (3D) has been largely mysterious since the first recognition of chromatin structures (Dounce et al., 1972). Hi-C technology (high-throughput genome-wide chromatin conformation capture, Lieberman-Aiden et al., 2009) is developed from chromosome conformation capture (3C) (Dekker et al., 2002), which enables the global characterization of the 3D chromatin architectures. Results from Hi-C (original version and its derived versions such as *in situ* Hi-C) revealed that chromatin architecture comprises a hierarchy of structures in mammals, with chromatin A/B compartments at multi-megabase scale, topologically associating domains (TADs) at hundreds of kilobases scale, and chromatin loops at kilobases to hundreds of kilobases scale (Dixon et al., 2012; Lieberman-Aiden et al., 2009; Rao et al., 2014). Previous studies have indicated that TADs (Flavahan et al., 2016; Hnisz et al., 2016; Lupiáñez et al., 2015) and/or enhancer-promoter loops impact gene expression and cellular physiology (Bonev et al., 2017; Isoda et al., 2017; Weintraub et al., 2017) and transcription elongation also affects the 3D genome organization in specific regions (Heinz et al., 2018). However, the elimination of TADs and chromatin loops by rapid degradation of the cohesin complex only modestly affects gene transcription programs (Rao et al., 2017), and inhibiting transcription with α-Amanitin does not substantially affect the establishment of 3D chromatin organization during early development (Du et al., 2017; Ke et al., 2017). Thus, much remains to be learned about how differentially arranged 3D chromatin architectures control key biological processes.

Mammalian spermatogenesis is known to involve extensive chromatin re-organization, of which analyses using imaging techniques have revealed these dynamic changes (Hao et al., 2019; Sassone-Corsi, 2002). During spermatogenesis, a small number of primitive type A spermatogonia (priSG-A) differentiate into more developed spermatogonia including type A spermatogonia (SG-A) and type B spermatogonia (SG-B) in sequence via mitosis (de Rooij, 2001). Subsequently, meiosis is initiated and marked by double-strand break (DSB) formation and is followed by full synapses of homologous chromosomes in pachytene spermatocytes (pacSC) (Tong and Lin, 2018). After meiosis I, spermatocytes rapidly divide in the ensuing meiosis II stage to form haploid round spermatids (rST) and spermatozoa (SZ) via spermiogenesis. Recent

Continued

[1]Hefei National Laboratory for Physical Sciences at the Microscale, CAS Key Laboratory of Brain Function and Disease, School of Life Sciences, Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei, Anhui 230026, China

[2]Institute of Reproductive Medicine, School of Medicine, Nantong University, Nantong, Jiangsu 226000, China

[3]Department of Biochemistry and Molecular Biology, McGovern Medical School, University of Texas Health Science Center, Houston, TX, USA

[4]Graduate School of Biomedical Sciences, University of Texas MD Anderson Cancer Center and UTHealth, Houston, TX, USA

[5]State Key Laboratory of Stem Cell and Reproductive Biology, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China

[6]University of Chinese Academy of Sciences, Beijing 100049, China

[7]CAS Key Laboratory of Genomic and Precision Medicine, Collaborative Innovation Center of Genetics and Development, CAS Center for Excellence in Molecular Cell Science, College of Future Technology, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing 100101, China

[8]Bio-X Institutes, Shanghai Jiao Tong University, Shanghai 200240, China

efforts using omics technologies revealed dynamic changes in transcriptomes, DNA methylomes, chromatin accessibility, etc., during spermatogenesis (Gan et al., 2013; Hammoud et al., 2014; Helsel et al., 2017; Hermann et al., 2018; Law et al., 2019; Maezawa et al., 2018; Rathke et al., 2014; Sohni et al., 2019). These findings have motivated researches into understanding the detailed molecular events in spermatogenesis, such as the 3D chromatin architectural dynamics and its effects on transcription regulation.

Because spermatogenesis represents a fundamental biological process that involves highly orchestrated rearrangement of chromosomes, characterization of 3D chromatin architectures during meiosis and spermatogenesis has been a highly active research area. For example, 3D genome structures of mammalian spermatogenesis have recently been studied by four groups at different sub-stages in mouse and rhesus monkey (Alavattam et al., 2019; Patel et al., 2019; Vara et al., 2019; Wang et al., 2019). One common insight from these reports is that TADs are depleted and compartment integrity is relatively weak during the pachytene stage. At the compartment level, however, a recent study has identified compartments that are more refined in resolution than the conventional A and B compartments and these "refined compartments" alternate between transcribing and non-transcribing regions during spermatogenesis of rhesus monkeys (Wang et al., 2019). In addition, cohesin occupancy in active compartments was shown to affect expression in pacSC and rST (Vara et al., 2019). Despite these findings, the relationship between the 3D genome and chromatin accessibility and their contribution to meiosis-specific gene expression programs (e.g., piRNAs) in spermatogenesis remain unclear.

Here, by using Hi-C and a series of other complementary technologies including assay for transposase-accessible chromatin using sequencing (ATAC-seq), chromatin immunoprecipitation sequencing (ChIP-seq), and RNA sequencing (RNA-seq), we demonstrated the following advances provided by this work that were not already described in any of the earlier publications. First, chromatin loops are present prior to and after, but not at, the pachytene stage. Second, chromatin accessibility *per se* is not involved in the diminishment of TADs or chromatin loops at the pacSC stage. Third, although the spermatozoa (SZ) has similar 3D chromatin structures as the primary spermatogonia (priSG-A) at the level of TADs and compartments, these two stages of cells have distinct chromatin loops. Finally, by tracing several factors (meiosis, DSBs, piRNAs, etc.) during the reprogramming of 3D chromatin architectures, our study elucidated the influence of 3D architecture on these factors related to spermatogenesis and suggests that reprogrammed chromatin compartments and loops in spermatogenesis underline differential transcriptional regulations that support and maintain spermatogenesis and prepare the sperms for the subsequent possible embryogenesis.

RESULTS

Higher-Order Chromatin Structures Are Reorganized during Mouse Spermatogenesis

To investigate the relationship between 3D genome structures and transcriptional regulation during mammalian spermatogenesis, we isolated spermatogenic cells from five different stages of mouse spermatogenesis (priSG-A, SG-A, pacSC, rST, and SZ) (Figure 1A) using the unit gravity sedimentation procedure (STA-PUT method) (Bellvé et al., 1977; Bryant et al., 2013; Gan et al., 2013; Hur et al., 2016; Korhonen et al., 2015; Liu et al., 2015; Luense et al., 2016; Wang et al., 2019) except for the SZ (see Transparent Methods). After validating the purities of these isolated cells (Figure S1A) by two methods, morphological characterization (Figure 1A) and immunofluorescence staining with cell-specific markers (Figure S1B), we performed *in situ* Hi-C with these different spermatogenic cells, respectively. The correlation analyses of the Hi-C data supported the high quality of our datasets among the biological replicates (Table S1 and Figure S2A), and we randomly sampled the same amount of Hi-C valid pairs (172,252,595) from spermatogenic cells of each stage to do the following analysis.

The chromatin interaction heatmaps revealed dramatic architectural reorganization throughout mouse spermatogenesis, as represented by chromosome 1 (Figure 1B), which confirmed recent reports using sub-stages (Vara et al., 2019; Wang et al., 2019). The A/B compartment patterns on autosomes were largely unaltered during mouse spermatogenesis (Figure 1C), whereas the strength, shown by saddle plots (Imakaev et al., 2012) (Figure S2B) and compartment strength (A-A and B-B compartments interaction strength relative to A-B compartments interaction strength) (Figure S2C) varied throughout the process. The strength of compartmentalization decreased from priSG-A to SG-A and dropped to the lowest level in pacSC before again increasing in rST and finally reaching its highest detected level in SZ. We also

[9]The Mary M. Wohlford Laboratory for Male Contraceptive Research, Center for Biomedical Research, Population Council, New York, USA

[10]Howard Hughes Medical Institute, School and Department of Medicine, University of California, San Diego School of Medicine, La Jolla, CA 92093-0651, USA

[11]These authors contributed equally

[12]Lead Contact

*Correspondence: sunfei@ntu.edu.cn (F.S.), songxy5@ustc.edu.cn (X.S.)

https://doi.org/10.1016/j.isci.2020.101034

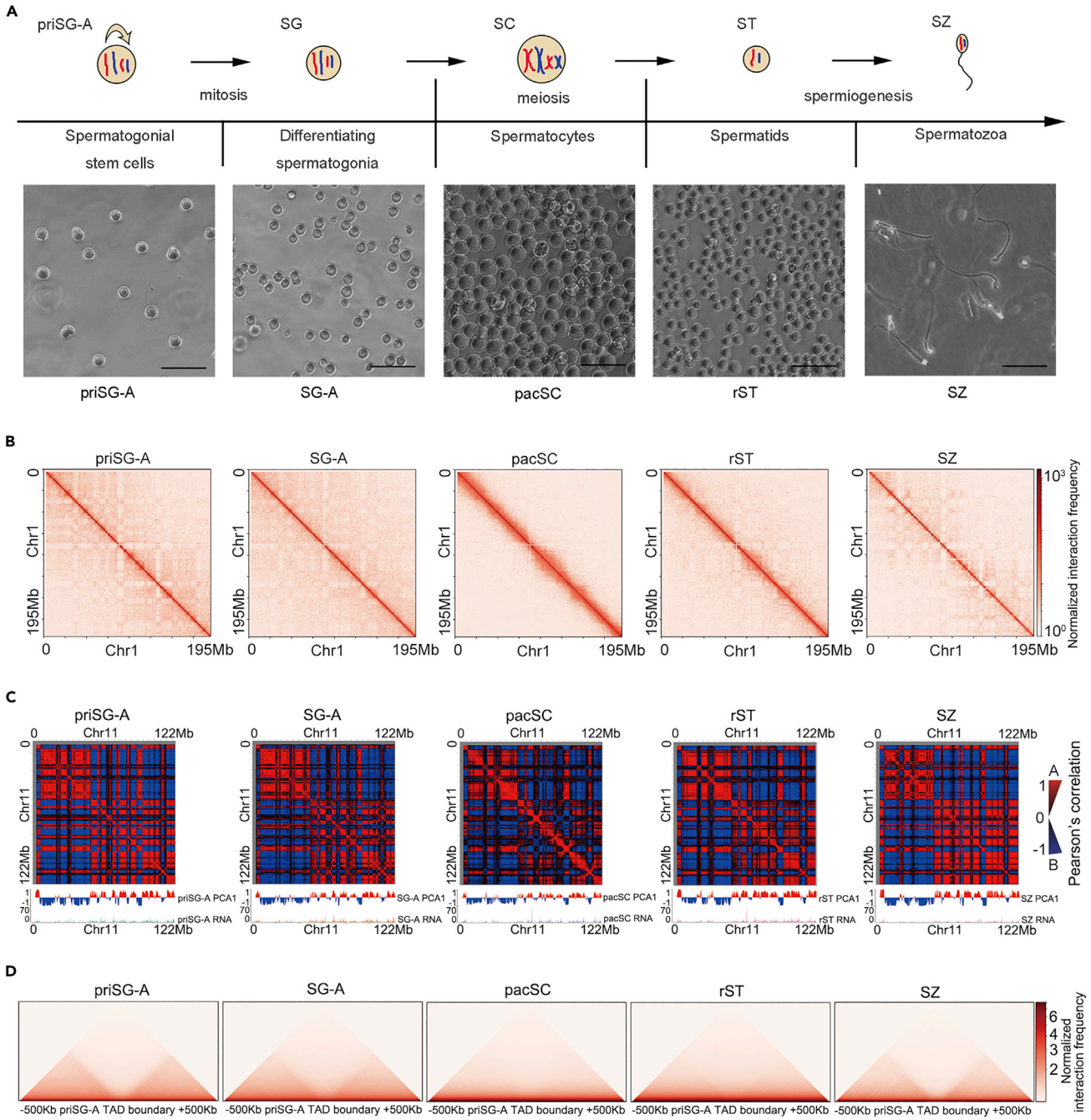


Figure 1. Reprogramming of Compartments and Topologically Associating Domains (TADs) throughout Mouse Spermatogenesis

(A) The schematic of mouse spermatogenesis (top panel). Phase-contrast microscopic images show the morphological characterization of five isolated spermatogenic cell types, including primitive type A spermatogonia (priSG-A), type A spermatogonia (SG-A), pachytene spermatocytes (pacSC), round spermatids (rST), and spermatozoa (SZ). Scale bar, 50 μm.

(B) Heatmaps showing the normalized Hi-C interaction frequencies (ratio of observed interaction frequencies to expected interaction frequencies [observed/expected]) (100-kb bins, chromosome 1) (pooled data from two biological replicates).

(C) Heatmaps showing Pearson's correlation for Hi-C interactions (50 kb bins, chromosome 1), which captured genomic A/B compartmentalization patterns in each tested spermatogenic stage. PCA1 (the first eigenvalues) is shown underneath the heatmap and normalized transcriptomic coverage is shown at the bottom.

(D) Heatmaps showing averaged normalized Hi-C interaction frequencies (ratio of observed interaction frequencies to expected interaction frequencies [observed/expected]) of each tested spermatogenic stage around the TAD boundaries (which were defined based on identifications in priSG-A samples).

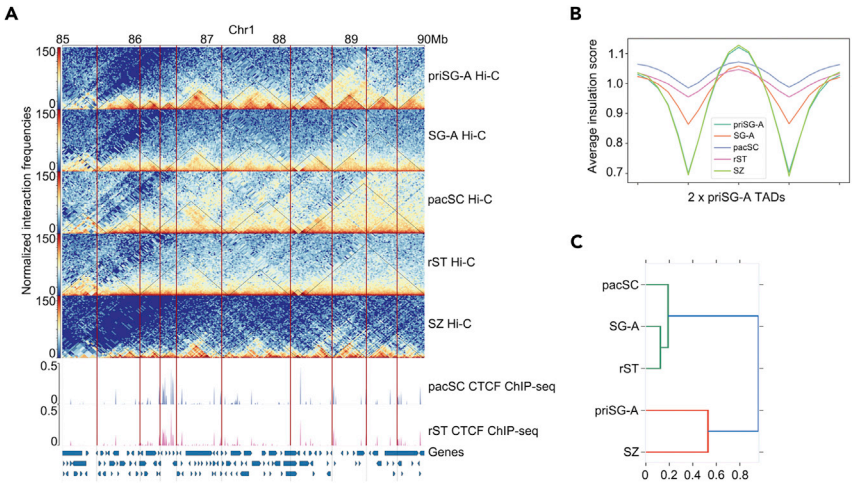


Figure 2. Topologically Associating Domains (TADs) Are Reorganized during Mouse Spermatogenesis

(A) Chromatin observed/expected interaction frequencies views (25 kb bin) of cells at different tested spermatogenic stages. Black triangular lines marked TADs of each stage and red vertical lines indicated boundary loci at the priSG-A stage. CTCF ChIP-seq coverage at the pacSC and rST stages and gencode genes were depicted at the bottom.

(B) The average insulation scores (ISs) of cells at different spermatogenic stages at TADs (defined in the priSG-A stage) and nearby regions (±0.5 TADs length).

(C) Dendrogram showing the hierarchical clustering of IS values calculated with window size of 100 kb from interaction matrices (25 kb bin) of cells at different spermatogenic stages.

observed that the distinct TAD structures in priSG-A and SG-A largely disappeared in pacSC and rST. Interestingly, they were largely restored in the SZ stage (Figures 1D and 2A). This trend of TAD reorganization was most clearly manifested by global analysis of averaged observed/expected interaction frequencies centered on these boundaries (Figure 1D) and was also very apparent as seen on a random representative genomic region such as Chromosome 1, 85–90 Mb region (Figure 2A). We further adopted an "insulation score (IS)" approach to calculate the strength of TAD boundaries, which analyzes the relative chromatin interaction frequency across a boundary (Crane et al., 2015). We found that the IS values decreased gradually from priSG-A to pacSC, then increased from rST to SZ again, and the IS value of SZ almost exactly matched that of priSG-A (Figure 2B).

priSG-A and SZ Share Similar 3D Chromatin Interactions and Folding Patterns

With these IS values, we performed hierarchical clustering and found that TAD structures that once existed in priSG-A, and disappeared in pacSC, had reappeared in SZ, making these two stages (priSG-A and SZ) clustered together (Figure 2C). These data led us to examine the 3D genome structures at the different spermatogenic stages in more detail. We found that the global 3D chromatin interactions were similar between priSG-A, SG-A, and SZ (Figures 3A–3C), whereas pacSC showed stronger short-range chromosome interactions (<5 Mb) but weaker long-range chromosome interactions (>5 Mb) as compared with SG-A (Figure 3A). Specifically, the contact frequency *P(s)* with genomic distances demonstrated that 3D genome structures were reorganized during spermatogenesis (Figure 3C), showing that priSG-A, SG-A, and SZ had similar chromatin folding patterns and that pacSC represented an extreme case, consistent with the above IS hierarchical clustering data.

Recent studies showed that mitotic chromatin exhibits a folding pattern distinct from that of interphase chromatin (Gibcus et al., 2018; Naumova et al., 2013). Because both meiotic and mitotic cells undergo chromosome condensation, we asked what is the difference between meiosis (pacSC) and mitosis in terms of 3D chromatin structures. The *P(s)* curves showed that at a scale less than 1 Mb, mitotic and meiotic chromatin shared similar folding patterns with a $P(s) \sim s$ slope at about −0.5 (Figure 3D). However, they displayed a striking difference at longer distances, where the meiotic *P(s)* curve had a steep drop at ~8 Mb, whereas the mitotic curve dropped at ~30 Mb (Figure 3D). This result suggested that meiotic pacSC chromatin is condensed via distinctive mechanisms as compared with that of mitotic chromatin.

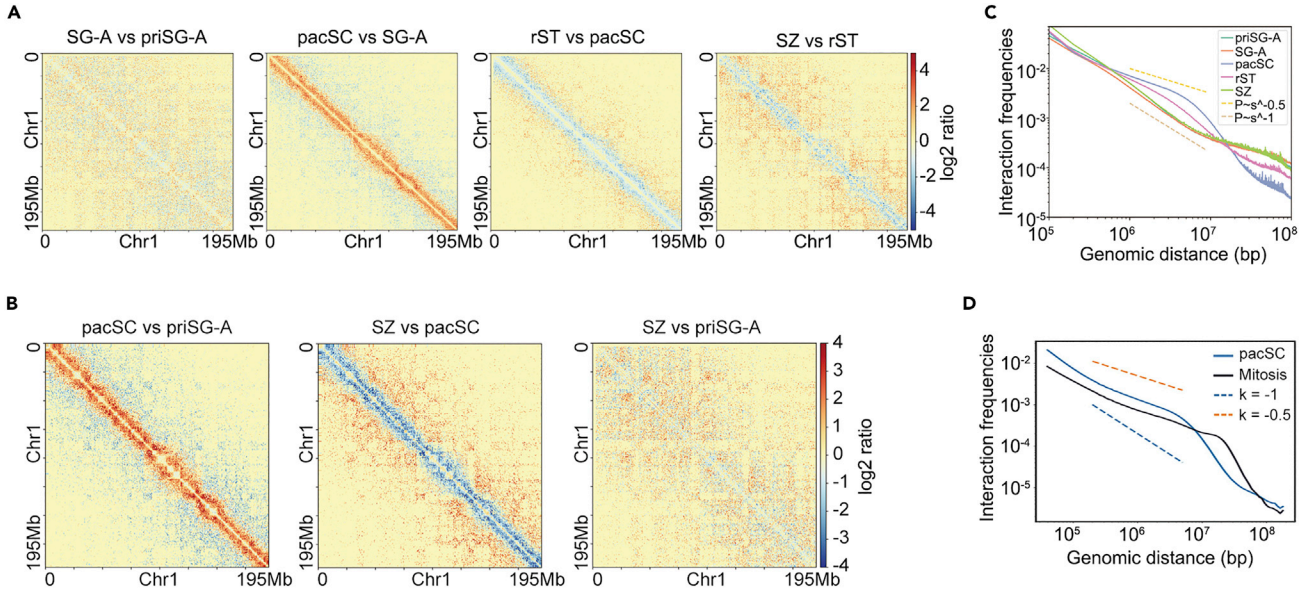


Figure 3. Global 3D Architecture Reprograms throughout Mouse Spermatogenesis

(A) Heatmaps of log_2 ratio comparisons of the normalized Hi-C interaction frequencies (observed/expected) (100 kb bins, chromosome 1) at different tested spermatogenic stages.

(B) Heatmaps of log_2 ratio comparisons of the observed/expected interaction frequencies (100 kb bins, chromosome 1) for pacSC versus priSG-A, SZ versus pacSC, SZ versus priSG-A cells, respectively.

(C) The *P(s)* curves (relationship between interaction probability and genomic distance) for each tested spermatogenic stage.

(D) The *P(s)* curves (relationship between interaction probability and genomic distance) for pacSC and mitosis, and slopes (k) −1 (blue) and −0.5 (orange) were shown by the dotted lines.

Active Compartments for Genes and piRNA Clusters with Meiosis-Related Functions Are Enriched in pacSC

We further examined A/B compartmentalization in these spermatogenic stages in detail. The A compartments are defined as regions with a positive first eigenvector value when eigenvector decomposition is applied on observed/expected intra-chromatin interaction matrices (50 kb bin size in this study). These computed A compartments correspond to the active, euchromatic nuclear compartments. By contrast, the B compartments largely correspond to those inactive, heterochromatic genomic regions (Lieberman-Aiden et al., 2009). We investigated if there is A/B compartment switching between priSG-A and pacSC. We found A compartments accounted for 43.2% in priSG-A and 13.9% of them changed to B compartments in pacSC. Of the other 56.8% B compartments in priSG-A, 15.3% changed to A compartments in pacSC (Figure 4A). Notably, the number of genome regions in the A (active) compartments was higher in pacSC, suggesting that pacSC chromatin is in a more transcriptionally active state.

We then identified the genes harbored by compartments that showed switching behaviors and examined their expression using RNA-seq data at the corresponding priSG-A and pacSC stages (Lin et al., 2016). Although there were genes that switched from A to B or B to A compartments with both up- and down-regulation of their expression, we found that genes that were originally located in compartment B regions in priSG-A but switched to compartment A regions in pacSC, and at same time increased their expression, were the most abundant (Figure 4B, 1,037 genes). This result implies that compartment switching from B to A is correlated with up-regulation of gene expression for their proper biological functions. Indeed, a number of these genes function in cilium formation and DNA DSBs repair (Figure S3A), which are critical for normal sperm physiology and function (Figure S3B). One example for both compartments switching and increased gene expression in pachytene cells is the DAZ family protein, Boll (Figure 4C), which is required for the meiotic G2/M transition and germ cell development (VanGompel and Xu, 2010).

PIWI proteins and PIWI-interacting RNAs (piRNAs) direct the silencing of target nucleic acids in animal germline and soma (Ozata et al., 2019). In early pachytene, piRNAs are produced from large genomic loci referred to as piRNA clusters; these RNAs function in controlling transposons and have been shown to contribute to paternal

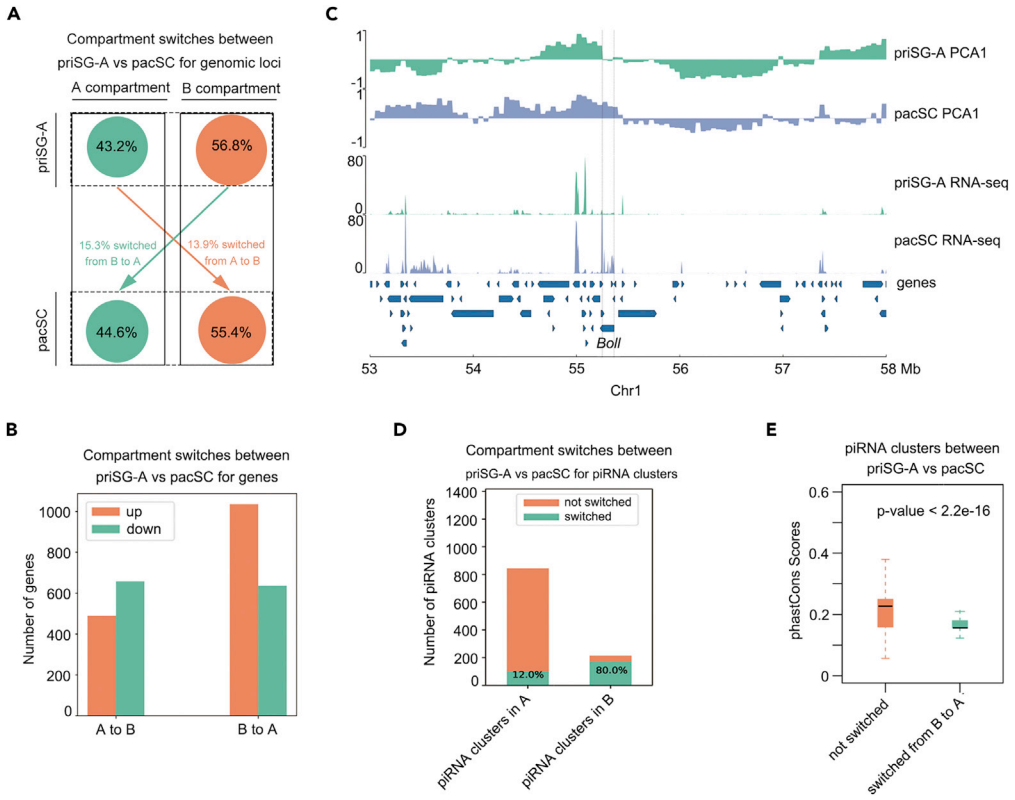


Figure 4. Spermatogenesis-Related Genes and piRNA Clusters Are Located in Active Compartments Emerging in pacSC Cells

(A) Schematic of the proportion of genome regions that switched their compartment states (A to B or B to A) comparing between priSG-A and pacSC.

(B) Bar chart showing the numbers of differentially expressed genes located in compartments transiting between the A and B states.

(C) PCA1 (the first eigenvalues) and RNA-seq coverage track of chromosome 1, 53–58 Mb showed a locus where a given gene, *Boll*, exhibited differential expression at two different stages. PCA1 was calculated via eigenvector decomposition on the observed/expected intra-chromatin interaction matrices. Normalized values for the extent of transcriptomic coverage in priSG-A and pacSC were shown in the middle. Gencode genes were shown at the bottom.

(D) Bar chart showing the number of piRNA clusters positioned in A or in B compartments in cells at the priSG-A stage, and the number of piRNA clusters that switched to the other compartment state (A to B, or vice versa) in cells at the pacSC stage.

(E) Averaged ucsc phastCons scores on piRNA clusters stable within the A compartments at both the priSG-A and pacSC stages or on piRNA clusters that switched from the B to A compartment.

imprinting (Aravin et al., 2007; Watanabe et al., 2011) and are required for the subsequent meiosis process (Fu and Wang, 2014). In addition to identifying genes in the compartments switched regions, we also analyzed the expression of piRNA clusters that locate to compartments showing stage-specific switching (piRNA clusters were identified from a published database, Rosenkranz, 2016). We found that 79.9% of all piRNA clusters had compartment A status in priSG-A, of which 88% remained in A compartments in pacSC; in contrast, 80% of the piRNA clusters with compartment B status in priSG-A switched to compartment A status in pacSC, correlating with their increased expression changes (Figure 4D).

To explore the difference between piRNA clusters that had compartment-A status in both priSG-A and pacSC cells (not switched) and that switched compartments from B to A between the priSG-A and pacSC stages, we examined the conservation status of these two categories of piRNA clusters using phastCons score, which measures the probability that each nucleotide belongs to a conserved element (Siepel et al., 2005). We averaged phastCons scores of every nucleotide in each piRNA cluster and compared the average scores of not switched and switched piRNA clusters between the priSG-A and pacSC stages.

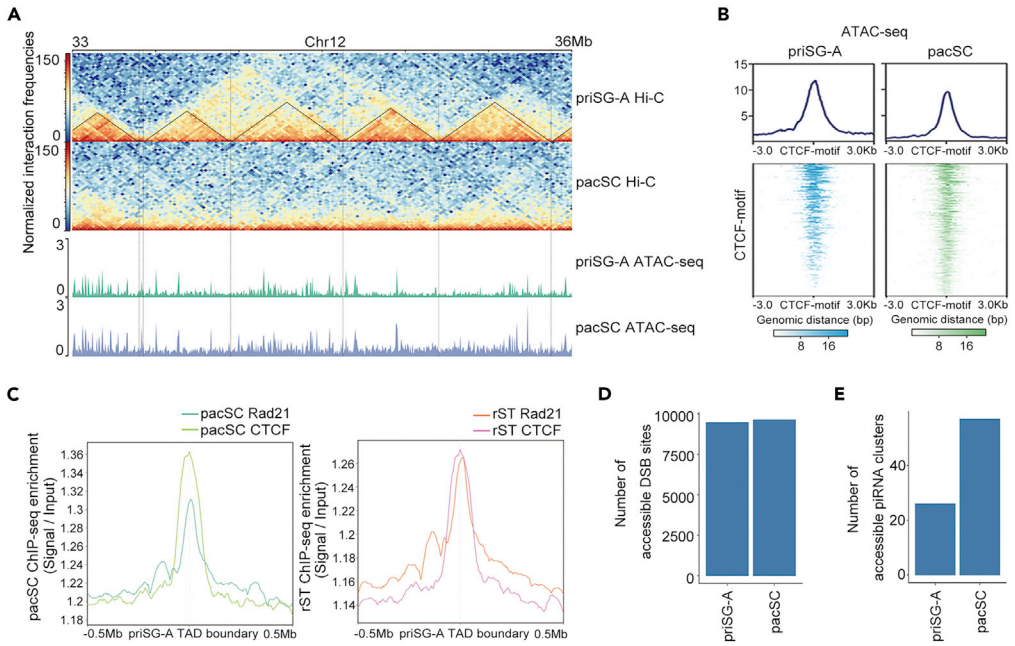


Figure 5. Genome Features of Accessible Regions at the pacSC Stage

(A) Chromatin observed/expected interaction frequencies views (25 kb bin) of cells at the priSG-A and pacSC stages, respectively. Vertical dashed lines indicated TAD boundary loci in priSG-A samples. Normalized ATAC-seq coverage in pacSC and rST samples were depicted at the bottom.

(B) Heatmaps showing the ATAC-seq signals enrichment around CTCF motif regions (±3kb), average ATAC-seq signals around all CTCF motif regions were plotted at the top.

(C) Profile plot showing the average CTCF and Rad21 signals (CTCF or Rad21 RPKM divided by input RPKM) in pacSC and rST samples around the TAD boundaries (±0.5Mb) defined in priSG-A samples.

(D) The number of accessible DSB sites in priSG-A and pacSC samples.

(E) The number of accessible piRNA clusters in priSG-A and pacSC samples.

The former (not switched) was more conserved than the switched between these two stages in a phylogenetic analysis of 60 vertebrate species (Figure 4E). This result indicated that the piRNA clusters with compartment A status in both priSG-A and pacSC cells exhibit signs of selection pressure and are conserved among different species, suggesting functional constraints on them and thus their potential functions in these species. These data suggested that the functional activity of piRNAs might be strongly impacted by their locations in compartment A or B in the 3D genome.

Chromatin Accessibility and CTCF/Cohesin Binding Are Retained throughout Mouse Spermatogenesis

Since TADs dramatically changed throughout spermatogenesis and disappeared in pacSC cells, we asked whether the 3D architecture of pacSC cells correlates with other genomic/epigenomic features. We conducted ATAC-seq to measure chromatin accessibility (Buenrostro et al., 2015) in four spermatogenic cell types (priSG-A, SG-A, pacSC, and rST) (Figure S4A). We first found that the accessible chromatin regions were mostly located at gene promoters or enhancers and the intensity of the accessible chromatin signal decreased as spermatogenesis proceeds (Figures S4B and S4C). When we compared the chromatin accessibility in priSG-A and pacSC, we found minimal differences between them (Figure 5A). This similarity of chromatin accessibility was rather surprising because the 3D genome architecture in these two cell types was largely different (Figure 5A). These data suggested that 3D genome reorganization might not be functionally linked to chromatin opening, at least at the level of TADs.

Chromatin accessibility is closely related to transcriptional activity. We thus examined transcriptional activity using ChIP-seq with antibodies against the elongating form of RNA Polymerase II (Pol II S2P). The enrichment of Pol II S2P on gene transcriptional start sites (TSSs) showed that these genes were actively

transcribing in both pacSC and rST (Figure S5A). However, the chromatin accessibility showed distinct patterns in pacSC and rST samples (Figure S5B). The chromatin accessibility in these two cell types was also very different from patterns in the other two cell types: priSG-A and SG-A (Figure S5B). Specifically, between pacSC and priSG-A, we identified 6,580 (out of 15,815) differentially accessible chromatin regions, whereas between pacSC and rST, there were 1,288 (of 15,815) differentially accessible regions. The chromatin regions displaying altered accessibilities between priSG-A and pacSC were analyzed for GO terms, and we found that they enriched for genes related to meiosis, and especially to the key events in meiosis I, e.g., meiotic recombination or synapsis (Figure S5C). A similar comparison of differentially accessible chromatin regions between pacSC and rST also revealed enrichment of meiotic genes, but notably, there was no longer any enrichment for gene groups specific to meiotic recombination or synapsis (Figure S5D). These results provided a molecular explanation of the biological knowledge that pacSC are at the prophase of the first meiotic division and that rST are produced after meiosis completion. Moreover, these data highlighted the utility of assessing chromatin accessibility states for inferring the functional status of transcriptional activity.

We then further investigated whether the differential chromatin accessibility may provide a molecular basis to the observed disappearance of TADs at the pacSC stage. As CTCF is enriched at TAD boundaries and the depletion of CTCF results in loss of TADs (Dixon et al., 2012; Nora et al., 2017), we checked if the chromatin accessibility of CTCF motifs may be affected at the TAD boundaries. Intriguingly, the chromatin accessibilities at CTCF motifs appeared highly similar between priSG-A and pacSC (Figure 5B), indicating that CTCF could bind to chromatin, which does not form conspicuous TADs (i.e., at the pacSC and rST stages). We experimentally examined CTCF binding in pacSC and rST cells via CTCF ChIP-seq. We found that, at both the pacSC stage and the rST stage, CTCF was still enriched at the TAD boundaries that were defined from the priSG-A stage (Figure 5C). In addition to CTCF, we also checked the binding of cohesin in pacSC and rST by Rad21 ChIP-seq. The result showed that in both pacSC and rST, Rad21 was also enriched at the TAD boundaries that were defined from the priSG-A stage (Figure 5C). These results suggested that the TAD disappearance in the meiotic stages of spermatogenesis could not be explained by the loss of CTCF or cohesin binding. Our data also renewed the understanding of CTCF and cohesin function in 3D genome structure, i.e., although CTCF and cohesin may be the molecules required for TAD maintenance in mammals (Nora et al., 2017; Rao et al., 2017), our data suggest that CTCF or cohesin binding at TAD boundaries is insufficient to maintain TADs at the pacSC stage.

During mammalian spermiogenesis, a highly condensed chromatin structure of sperm nuclei is in part formed by histone-protamine exchange and 1%–8% of histones in mouse are retained in sperm chromatin (Jung et al., 2019). Sperm histones were found to be mainly retained in distal intergenic regions, as shown by ChIP-seq using nucleoplasmin (NPM)-treated sperm (Yamaguchi et al., 2018). However, whether TAD boundaries play a role in retaining these histones in the sperm is unknown. We analyzed published protamine 1 targeted ChIP-seq data from sperms (Yoshida et al., 2018) and averaged the ChIP-seq signals around TAD boundaries in the sperm. Although protamine 1 was shown to occupy TSS regions, there was no enrichment of protamine 1 at the TAD boundaries (Figure S5E). We also analyzed the published histone H3 ChIP-seq data (Yamaguchi et al., 2018) and found no enriched localization of histone H3 around TAD boundaries, either (Figure S5F). These results implied that the TAD structure in the sperm may not be related to the histone or protamine constitution. Another question we asked was whether sperm-retained histones play substantial roles in the next generation. We found that, although histone H3 was depleted from the TSS regions, 171 genes had at least one exon in histone H3 retaining regions, and 16 of these 171 genes had been previously reported to be expressed in the early embryonic stage (Theiler stage 1–5, Table S2) (Smith et al., 2019). Further work examining the distribution of different modifications of retained histones and their potential roles in subsequent embryogenesis will provide more accurate answers to the function of retained histones.

Meiotic DSB Sites Are Pre-opened at the priSG-A Stage and piRNA Clusters Are Uniquely Open at the pacSC Stage

DNA double-strand break (DSB) is essential for the recombination in meiosis, and it begins at the leptotene stage and ends at the zygotene stage (Baudat et al., 2013; Smagulova et al., 2011). We examined the accessibility of chromatin regions surrounding previously reported DSB hotspots midpoints by analyzing our ATAC-seq data. We first identified meiotic DSB hotspots midpoints by the published Dmc1 ChIP-seq data obtained from DSB-positive cells in Hop2 knockout mice, whose meiotic DSBs are not repaired and

meiotic progression is blocked at the pachytene-like stage (Smagulova et al., 2011). We then compared the chromatin accessibility around these meiotic DSB hotspots midpoints at the priSG-A and pacSC stages. Remarkably, this analysis revealed that the numbers of the accessible DSB hotspots midpoints were largely unchanged between these two stages (Figure 5D). These results raised an intriguing possibility that meiotic DSB hotspots might be pre-opened (i.e., "primed") prior to their starting point, perhaps at a time even earlier than the priSG-A stage, and remained open after their repairing at the pacSC stage.

Since we observed that there were more piRNA clusters located in active compartments at the pacSC stage than at the priSG-A stage (Figure 4D), we further compared the numbers of accessible pacSC piRNA clusters between priSG-A and pacSC. A larger number of accessible piRNA clusters were found in pacSC but not in priSG-A (Figure 5E). These data were consistent with the above observation that more piRNA clusters were located in active compartments at the pacSC stage compared with that at the priSG-A stage, and it was also in line with the fact that pacSC-specific piRNAs are transcribed at the pacSC stage. These data together supported a refined coordinated functional contribution of DSBs and piRNA clusters to the meiosis process of murine spermatogenesis.

Reorganized Chromatin Loops in Sperm Contribute to Early Embryo Development

To examine the finer-scale architecture of the chromatin, we identified chromatin loops from the Hi-C data in each spermatogenic stage using cLoops (Cao et al., 2020). We observed that these loop interactions were reprogrammed throughout the spermatogenesis process (Figures 6A and 6B), similar to the dynamics at TAD or compartment strength levels. The number of loops was significantly decreased at the pacSC and rST stages, followed by an increase to a maximum value in SZ (Figure 6C). We aggregated the interaction values around loops defined in priSG-A and found that the strength of loops, as shown by the APA values (Rao et al., 2014), showed the same trend as the loop numbers (Figure S6A).

To test the functional roles of the most abundant loops formed in SZ, we used ChIP-seq data of histone marks and RNA Pol II binding in testis from previous reports (Shen et al., 2012) to annotate potential enhancers. This analysis revealed that 1,572 genes were putatively involved in the formation of promoter-enhancer loops at the SZ stage (Figure 6C). Of these 1,572 genes, only a small percentage (317, or 20%) were also anchored in loops in priSG-A samples (Figure S6B), suggesting that, although chromatin architecture in SZ samples restored the pattern in priSG-A at the larger scale (of compartments and TADs), at the smaller scale, the reestablished loops in SZ samples were distinct from those in the priSG-A samples. Genes that were involved in loops in SZ might tend to be transcriptionally active in subsequent embryonic stages, and previous studies have demonstrated the expression of a majority (1,174, or 75%) of these 1,572 genes in embryos, with 217 genes specifically detected in the developmental window of Theiler stages 1–5, although many of them repeatedly appeared in different stages of these early stages (Smith et al., 2019) (Figure 6D). Specifically, 63 genes could be detected from Theiler stage 2 and 26 genes could be detected from Theiler stage 1 until Theiler stage 4 (Figure S6C). These 217 genes were found to be involved in biological processes like regulation of gene expression and cell differentiation (Figure S6D), and their reported knockout phenotypes in mammals include embryonic growth abnormality, arrest, and lethality (Figure 6E). As examples, Actr3Gt$^{(A009F03)VauA}$/Actr3Gt$^{(A009F03)VauA}$ mice showed embryonic growth arrest (Vauti et al., 2007) and Adar^{tm2Knk}/Adar^{+} mice showed abnormal embryonic erythropoiesis and abnormal liver development (Wang et al., 2000). These results implied that the reorganized loops we detected in SZ might facilitate the expression of genes required for cell differentiation during the subsequent embryogenesis.

DISCUSSION

Recent studies have shown the dynamic changes in TADs and chromatin compartmentalization during mammalian spermatogenesis (Alavattam et al., 2019; Patel et al., 2019; Vara et al., 2019; Wang et al., 2019). Here, we further explored dynamic higher-order chromatin structures and their functional impacts on transcriptional regulation during mouse spermatogenesis combining Hi-C, ATAC-seq, ChIP-seq, and RNA-seq data. Our results complement recent reports by demonstrating that chromatin loops are dynamically reorganized during spermatogenesis in a pattern similar to that of TADs: they are present at stages prior to pacSC, are "erased" in pacSC, and get re-established after it. We also showed that the switching of A/B compartments between spermatogenic stages is tightly associated with meiosis-specific mRNAs and piRNAs expression. For mechanism, we showed that chromatin accessibility and the binding of CTCF and cohesin at TAD boundaries are not responsible for the TAD and loop diminishment at the pachytene stage.

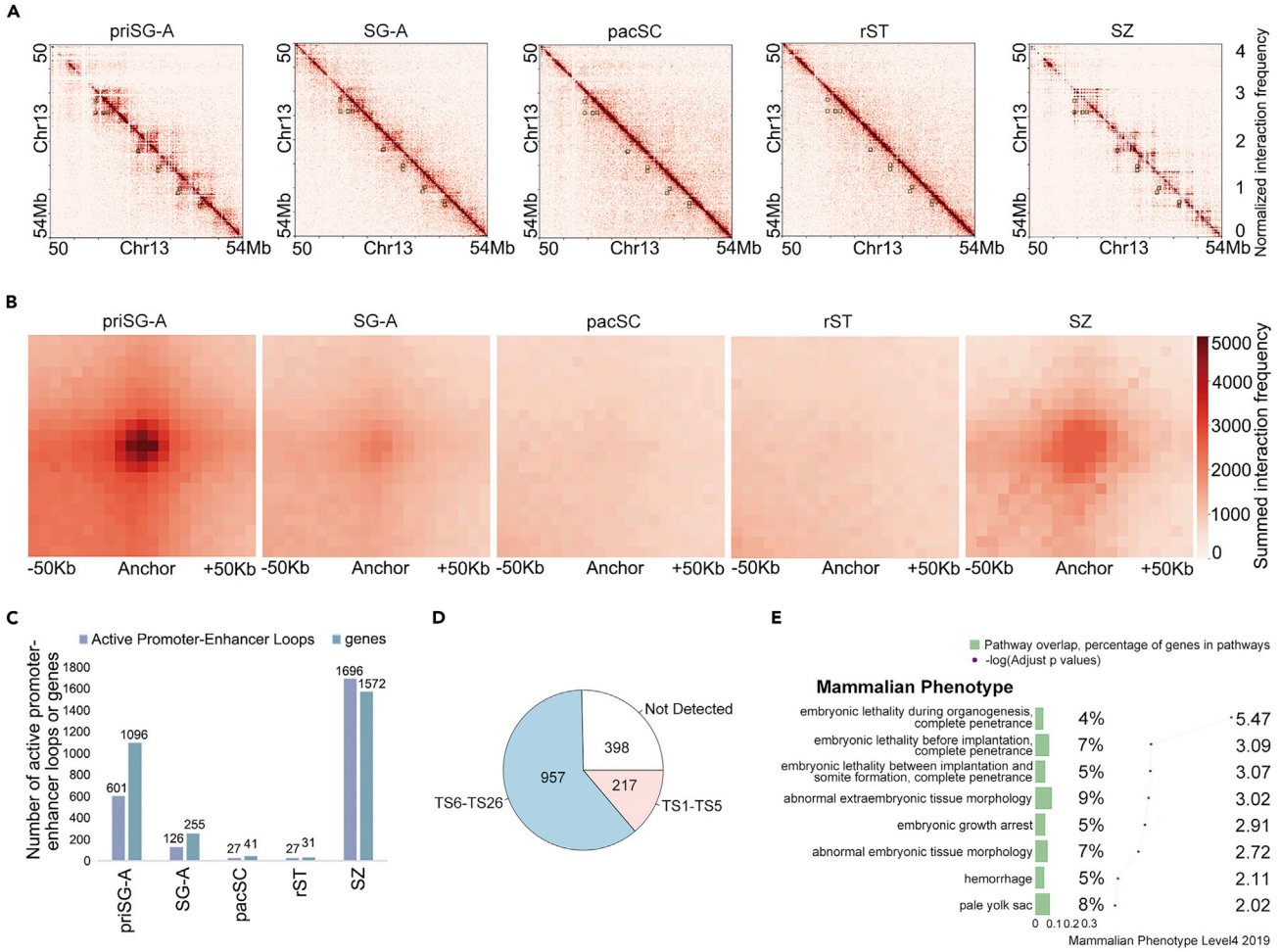


Figure 6. Reorganization of Chromatin Loops during Mouse Spermatogenesis and Their Implicated Roles in Early Embryo

(A) Heatmaps showing the observed/expected Hi-C interaction frequencies (5 kb bin, chromosome 13, 50–54 Mb) of cells at different tested spermatogenic stages. Black boxes showed the loop loci that were defined in priSG-A samples.

(B) Heatmaps showing the summed obs/exp Hi-C interaction frequencies within the vicinity (±50 kb) of the loop centers defined in priSG-A samples.

(C) The number of active promoter-enhancer loops at different spermatogenic stages. Loops were defined by cLoops. Enhancers were marked by H3K4me1 or H3K27ac outside promoter regions in testis, and active promoters in testis were indicated by enrichment of H3K4me3 or Pol II binding signals. The number of genes whose promoters were overlapped with loop anchors was shown.

(D) Pie plot showing the number of genes located in enhancer-promoter loops in SZ samples according to whether their expression was detected in early embryonic stages (Theiler stages 1–5) or in Theiler stages 6–26.

(E) GO-based mammalian phenotype enrichment data based on 217 genes positioned at enhancer-promoter loops expressed in early embryonic stages (Theiler stages 1–5), and a bar plot showing the proportion of these genes pertaining to the annotation of each phenotype over the number of all genes annotated in each phenotype. Line plot showing the significance (-log (adjusted p values)) of each ontology.

It was reported that A/B compartment switching is correlated with the expression of specific genes during spermatogenesis (Vara et al., 2019). We also found some meiosis-related genes switch A/B compartment states and change expression between priSG-A and pacSC. In addition, we focused on piRNA clusters and demonstrated that most piRNA clusters are located in active A compartments and their chromatin states are quite stable, yet a majority of the piRNA clusters that are located in inactive B compartments in priSG-A are switched to active A compartments in pacSC. This is in line with the fact that pacSC cells are known to transcribe a characteristic set of pacSC-specific piRNAs (Aravin et al., 2006; Girard et al., 2006; Ozata et al., 2019); thus, our findings indicate that the chromatin compartment switch likely affects the activity of piRNAs.

Our results showed that, although the 3D genome organization changes greatly at the pacSC stage, chromatin at regions containing meiotic genes are still accessible, indicating that chromatin accessibility *per se*

is not functionally associated with TADs or chromatin loops at the pacSC stage. In addition, enrichment of meiotic DSB sites in active A compartments has been reported during the pachytene stage (Patel et al., 2019). Chromatin accessibility was suggested to correlate with the occurrence of meiotic DSBs by some authors (Patel et al., 2019), but others have shown that accessible chromatin regions in pacSC have limited overlap with DSB sites (Maezawa et al., 2018). In the present study, we showed that many meiotic DSB sites are in accessible chromatin regions, and our data unexpectedly revealed that these sites are already "primed" in the priSG-A stage, which may play a functional role in generating DSB sites.

Chromatin tracing by super-resolution imaging at the single-cell level revealed TAD-like structures and showed that TAD boundaries tend to be located near CTCF and cohesin binding sites (Bintu et al., 2018). These TAD-like structures were still found in individual cells after removal of cohesin from chromatin, although the signal for the TAD-like structures was buried when the cells were analyzed as a population (Bintu et al., 2018). Upon specific chemical degradation of cohesin, the boundaries of the TAD-like structures in single cells appeared to be randomly distributed, suggesting that cohesin might not be required for TAD maintenance but may rather function to establish developmentally appropriate TAD boundaries (Bintu et al., 2018). Moreover, cohesin-mediated transcription in genomic regions out of the chromosomal axes in primary spermatocytes can provide an environment conducive to both gene expression and the formation of DSBs (Vara et al., 2019). We speculate that the population-averaged signal for TADs in the CTCF-depleted cells of the previous report (Nora et al., 2017) might also mask the occupancy of TAD structures in individual cells. However, this is not the case for meiotic chromatin: our ChIP-seq data showed that specific binding of CTCF and cohesin (at TAD boundaries that were defined at the priSG-A stage) is retained at the pacSC stage when TADs disappear, suggesting that there are additional factors that regulate the formation or maintenance of TAD structures, at least during meiosis. Multiple methodological strategies that combine 3D genome detection and immunoprecipitation or chromatin isolation and mass spectrometry and experiments better control for variables could be employed to identify candidates of such factors. In addition, loss of TADs and CTCF binding was reported to occur in mitotic chromatin (Gibcus et al., 2018; Naumova et al., 2013; Oomen et al., 2019), which, when taken together with our findings, strongly implies that TADs maintenance may be controlled via distinct mechanisms in mitosis versus meiosis.

In addition to lost TADs and chromatin loops, the global chromatin organization between mitotic and meiotic chromatin is also similar when compared with that in interphase, as revealed by our *P(s)* curve analysis. Nevertheless, meiotic chromatin exhibits A/B compartments (Figure 1C), whereas mitotic chromatin does not (Gibcus et al., 2018). Thus, a comparison of factors modulating 3D genome organization at the compartmentalization level in meiotic and mitotic chromatin may discover unknown factors involved in higher-order chromatin organization. In addition to CTCF binding, transcriptional activity represents another major difference between meiosis and mitosis and is known to be affected by higher-order chromatin organization. Although both meiotic and mitotic chromatin are compacted, meiotic chromatin remains transcriptionally active (Figure S5A), whereas mitotic chromatin is largely transcriptionally inert (Gottesfeld and Forbes, 1997; Wang et al., 2019). We propose that the distinct chromatin interaction patterns may be one of the causes underlying this difference.

Previous studies have reported that TADs and chromatin loops are associated with gene expression (Bonev and Cavalli, 2016; Flavahan et al., 2016; Hnisz et al., 2016; Lupiáñez et al., 2015), whereas there are also some findings that show that loss of TADs and chromatin loops by rapid degradation of cohesin or CTCF have only modest effects on gene expression (Nora et al., 2017; Rao et al., 2017). Thus, the impacts of chromatin organization at TAD and chromatin loop scales on gene expression are still not very clear. A recent study by Micro-C that explores the chromatin organization from single nucleosomes to whole genome-scale found that there are microTADs and gene-level folding within TADs that form in a gene-dependent manner and are highly associated with chromatin accessibility and transcriptionally active chromatin (Hsieh et al., 2020). Here, our results together with several recent reports show that TADs and chromatin loops are almost lost at the pacSC stage during mouse spermatogenesis (Patel et al., 2019; Vara et al., 2019; Wang et al., 2019), whereas we also show that CTCF/cohesin binding on chromatin, chromatin accessibility at promoter/enhancer regions, and Pol II binding remain at the pacSC stage. Thus, although TADs and chromatin loops are largely lost at the pacSC stage, the microTADs and gene-level folding, which cannot be detected by *in situ* Hi-C, might remain on pacSC chromatin and contribute to transcription regulation and gene expression at the pacSC stage. Our chromatin accessibility and Pol II S2P ChIP-seq data support this possibility,

and future studies with nucleosome-resolution Micro-C and single-cell methods will further deepen our understanding of the relationship between chromatin organization and transcription regulation during mouse spermatogenesis. Altogether, beyond confirming several major conclusions about TAD reorganization and the dynamics of A/B compartments strength, our study also reveals mechanistic insights about the functional relationships between 3D genome organization, CTCF or cohesin binding, and TAD maintenance and demonstrates the roles of A/B compartments switching in controlling meiotic gene transcription programs as well as potential function of chromatin loops in SZ in preparing for subsequent embryogenesis.

Limitations of the Study

In this study, we explored the 3D genome organization during mouse spermatogenesis with isolated spermatogenic cells from five different stages. We showed that chromatin loops and TADs are reorganized during mouse spermatogenesis, which almost disappear at the pacSC stage, whereas the transcriptional activity, CTCF binding on chromatin, and chromatin accessibility of promoters and enhancers are largely maintained at the pacSC stage. Thus, it is unclear how the transcriptional regulation is going on at the pacSC stage when the pacSC chromatin has no chromatin loops and TADs. A recent study about transcription-linked mammalian 3D chromatin folding with Micro-C uncovered a finer-scale genome organization that is associated with transcription and chromatin accessibility more tightly, compared with chromatin loops and TADs (Hsieh et al., 2020). Future studies that focus on the finer-scale genome organization at the pacSC stage might reveal the relationship between 3D genome structures and transcriptional regulation, as whether the finer-scale genome organization remains at the pacSC stage while chromatin loops and TADs are almost lost is still unclear.

In addition, we and others only explored the 3D genome organization during mouse and rhesus monkey spermatogenesis with limited isolated cell types (Alavattam et al., 2019; Patel et al., 2019; Vara et al., 2019; Wang et al., 2019), which restrict us to get a more comprehensive understanding about the 3D genome organization during mouse spermatogenesis. Future studies with single-cell Hi-C technologies and related methods might give us a more comprehensive view about the 3D genome organization during mouse spermatogenesis and uncover the underlying mechanisms about the reorganized chromatin structures.

METHODS

All methods can be found in the accompanying Transparent Methods supplemental file.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.isci.2020.101034.

ACKNOWLEDGMENTS

We sincerely thank Dr. Mingxi Liu (Nanjing Medical University), Dr. Bo Wen (Fudan University), and Dr. Yong Zhang (Institute of Zoology, Chinese Academy of Sciences) for their helpful advices. We greatly appreciate Dr. Shouhong Guang (University of Science and Technology of China, USTC) for insightful editing of the manuscript. We gratefully acknowledge English language assistance from Prof. Zachary J. Smith (USTC). This work was supported by grants from the National Key Research and Development Program of China (No. 2018YFC1003500 to F.S.), National Key Scientific Program of China (No. 2016YFA0100502 and 2015CB943002 to X.S.), the National Natural Science Foundation of China (Grant No. 31671490 to X.S., 81671510 to F.S., and 81901528 to X.W.), and the Cancer Prevention and Research Institute of Texas (CPRIT) Scholar Award (RR160083 to W.L.).

AUTHOR CONTRIBUTIONS

Z.L. and X.W. performed experiments with assistance from J.C. (under the guidance of C.H.), X.G. (under the guidance of J.W.) and Q.X.; H.J. and R.W. performed data analysis with assistance from Y.C. (under the guidance of Y.Y.) and J.C.; X.S. and F.S. designed the project; X.S., Z.L., H.J., R.W., X.W., C.H., W.L., C.Y.C., and M.G.R. wrote and reviewed of manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: November 20, 2019
Revised: December 30, 2019
Accepted: March 30, 2020
Published: April 24, 2020

REFERENCES

Alavattam, K.G., Maezawa, S., Sakashita, A., Khoury, H., Barski, A., Kaplan, N., and Namekawa, S.H. (2019). Attenuated chromatin compartmentalization in meiosis and its maturation in sperm development. Nat. Struct. Mol. Biol. *26*, 175–184.

Aravin, A., Gaidatzis, D., Pfeffer, S., Lagos-Quintana, M., Landgraf, P., Iovino, N., Morris, P., Brownstein, M.J., Kuramochi-Miyagawa, S., Nakano, T., et al. (2006). A novel class of small RNAs bind to MILI protein in mouse testes. Nature *442*, 203–207.

Aravin, A.A., Sachidanandam, R., Girard, A., Fejes-Toth, K., and Hannon, G.J. (2007). Developmentally regulated piRNA clusters implicate MILI in transposon control. Science *316*, 744–747.

Baudat, F., Imai, Y., and de Massy, B. (2013). Meiotic recombination in mammals: localization and regulation. Nat. Rev. Genet. *14*, 794.

Bellvé, A.R., Cavicchia, J., Millette, C.F., O'brien, D.A., Bhatnagar, Y., and Dym, M. (1977). Spermatogenic cells of the prepuberal mouse: isolation and morphological characterization. J. Cel. Biol. *74*, 68–85.

Bintu, B., Mateo, L.J., Su, J.-H., Sinnott-Armstrong, N.A., Parker, M., Kinrot, S., Yamaya, K., Boettiger, A.N., and Zhuang, X. (2018). Super-resolution chromatin tracing reveals domains and cooperative interactions in single cells. Science *362*, eaau1783.

Bonev, B., and Cavalli, G. (2016). Organization and function of the 3D genome. Nat. Rev. Genet. *17*, 661.

Bonev, B., Mendelson Cohen, N., Szabo, Q., Fritsch, L., Papadopoulos, G.L., Lubling, Y., Xu, X., Lv, X., Hugnot, J.-P., Tanay, A., et al. (2017). Multiscale 3D genome rewiring during mouse neural development. Cell *171*, 557–572.e24.

Bryant, J.M., Meyer-Ficca, M.L., Dang, V.M., Berger, S.L., and Meyer, R.G. (2013). Separation of spermatogenic cell types using STA-PUT velocity sedimentation. J. Vis. Exp. *80*, e50648.

Buenrostro, J.D., Wu, B., Chang, H.Y., and Greenleaf, W.J. (2015). ATAC-seq: a method for assaying chromatin accessibility genome-wide. Curr. Protoc. Mol. Biol. *109* (1), 21.29.1–21.29.9.

Cao, Y., Chen, Z., Chen, X., Ai, D., Chen, G., McDermott, J., Huang, Y., Guo, X., and Han, J.J. (2020). Accurate loop calling for 3D genomic data with cLoops. Bioinformatics *36*, 666–675.

Crane, E., Bian, Q., McCord, R.P., Lajoie, B.R., Wheeler, B.S., Ralston, E.J., Uzawa, S., Dekker, J., and Meyer, B.J. (2015). Condensin-driven remodelling of X chromosome topology during dosage compensation. Nature *523*, 240–U299.

de Rooij, D.G. (2001). Proliferation and differentiation of spermatogonial stem cells. Reproduction *121*, 347–354.

Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing chromosome conformation. Science *295*, 1306–1311.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376–380.

Dounce, A.L., Chanda, S.K., Ickowicz, R., Volkman, D., Palermiti, M., and Turk, R. (1972). The structure of eukaryote chromatin. Acta Endocrinol. Suppl. *168*, 86–111.

Du, Z., Zheng, H., Huang, B., Ma, R., Wu, J., Zhang, X., He, J., Xiang, Y., Wang, Q., Li, Y., et al. (2017). Allelic reprogramming of 3D chromatin architecture during early mammalian development. Nature *547*, 232–235.

Flavahan, W.A., Drier, Y., Liau, B.B., Gillespie, S.M., Venteicher, A.S., Stemmer-Rachamimov, A.O., Suvà, M.L., and Bernstein, B.E. (2016). Insulator dysfunction and oncogene activation in IDH mutant gliomas. Nature *529*, 110–114.

Fu, Q., and Wang, P.J. (2014). Mammalian piRNAs. Spermatogenesis *4*, e27889.

Gan, H., Wen, L., Liao, S., Lin, X., Ma, T., Liu, J., Song, C.-x., Wang, M., He, C., and Han, C. (2013). Dynamics of 5-hydroxymethylcytosine during mouse spermatogenesis. Nat. Commun. *4*, 1995.

Gibcus, J.H., Samejima, K., Goloborodko, A., Samejima, I., Naumova, N., Nuebler, J., Kanemaki, M.T., Xie, L., Paulson, J.R., Earnshaw, W.C., et al. (2018). A pathway for mitotic chromosome formation. Science *359*, eaao6135.

Girard, A., Sachidanandam, R., Hannon, G.J., and Carmell, M.A. (2006). A germline-specific class of small RNAs binds mammalian Piwi proteins. Nature *442*, 199–202.

Gottesfeld, J.M., and Forbes, D.J. (1997). Mitotic repression of the transcriptional machinery. Trends Biochem. Sci. *22*, 197–202.

Hammoud, S.S., Low, D.H., Yi, C., Carrell, D.T., Guccione, E., and Cairns, B.R. (2014). Chromatin and transcription transitions of mammalian adult germline stem cells and spermatogenesis. Cell Stem Cell *15*, 239–253.

Hao, S.-L., Ni, F.-D., and Yang, W.-X. (2019). The dynamics and regulation of chromatin remodeling during spermiogenesis. Gene *706*, 201–210.

Heinz, S., Texari, L., Hayes, M.G.B., Urbanowski, M., Chang, M.W., Givarkes, N., Rialdi, A., White, K.M., Albrecht, R.A., Pache, L., et al. (2018). Transcription elongation can affect genome 3D structure. Cell *174*, 1522–1536.e22.

Helsel, A.R., Yang, Q.E., Oatley, M.J., Lord, T., Sablitzky, F., and Oatley, J.M. (2017). ID4 levels dictate the stem cell state in mouse spermatogonia. Development *144*, 624–634.

Hermann, B.P., Cheng, K., Singh, A., Roa-De La Cruz, L., Mutoji, K.N., Chen, I.C., Gildersleeve, H., Lehle, J.D., Mayo, M., Westernstroer, B., et al. (2018). The mammalian spermatogenesis single-cell transcriptome, from spermatogonial stem cells to spermatids. Cell Rep. *25*, 1650–1667.e8.

Hnisz, D., Day, D.S., and Young, R.A. (2016). Insulated neighborhoods: structural and functional units of mammalian gene control. Cell *167*, 1188–1200.

Hsieh, T.S., Cattoglio, C., Slobodyanyuk, E., Hansen, A.S., Rando, O.J., and Tjian, R. (2020). Resolving the 3D landscape of transcription-linked mammalian chromatin folding. Mol. Cell. https://doi.org/10.1016/j.molcel.2020.03.002.

Hur, S.K., Freschi, A., Ideraabdullah, F., Thorvaldsen, J.L., Luense, L.J., Weller, A.H., Berger, S.L., Cerrato, F., Riccio, A., and Bartolomei, M.S. (2016). Humanized H19/Igf2 locus reveals diverged imprinting mechanism between mouse and human and reflects Silver–Russell syndrome phenotypes. Proc. Natl. Acad. Sci. U S A *113*, 10938–10943.

Imakaev, M., Fudenberg, G., McCord, R.P., Naumova, N., Goloborodko, A., Lajoie, B.R., Dekker, J., and Mirny, L.A. (2012). Iterative correction of Hi-C data reveals hallmarks of chromosome organization. Nat. Methods *9*, 999–1003.

Isoda, T., Moore, A.J., He, Z., Chandra, V., Aida, M., Denholtz, M., Piet van Hamburg, J., Fisch, K.M., Chang, A.N., Fahl, S.P., et al. (2017). Non-coding transcription instructs chromatin folding and compartmentalization to dictate enhancer-promoter communication and T cell fate. Cell *171*, 103–119.e18.

Jung, Y.H., Kremsky, I., Gold, H.B., Rowley, M.J., Punyawai, K., Buonanotte, A., Lyu, X., Bixler, B.J., Chan, A.W.S., and Corces, V.G. (2019). Maintenance of CTCF- and transcription factor-mediated interactions from the gametes to the early mouse embryo. Mol. Cell *75*, 154–171.e5.

Ke, Y., Xu, Y., Chen, X., Feng, S., Liu, Z., Sun, Y., Yao, X., Li, F., Zhu, W., Gao, L., et al. (2017). 3D

chromatin structures of mature gametes and structural reprogramming during mammalian embryogenesis. Cell *170*, 367–381.e20.

Korhonen, H.M., Yadav, R.P., Da Ros, M., Chalmel, F., Zimmermann, C., Toppari, J., Nef, S., and Kotaja, N. (2015). DICER regulates the formation and maintenance of cell-cell junctions in the mouse seminiferous epithelium. Biol. Reprod. *93*, 139.

Law, N.C., Oatley, M.J., and Oatley, J.M. (2019). Developmental kinetics and transcriptome dynamics of stem cell specification in the spermatogenic lineage. Nat. Commun. *10*, 2787.

Lieberman-Aiden, E., Van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., and Dorschner, M.O. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. science *326*, 289–293.

Lin, X., Han, M., Cheng, L., Chen, J., Zhang, Z., Shen, T., Wang, M., Wen, B., Ni, T., and Han, C. (2016). Expression dynamics, relationships, and transcriptional regulations of diverse transcripts in mouse spermatogenic cells. RNA Biol. *13*, 1011–1024.

Liu, Y., Niu, M., Yao, C., Hai, Y., Yuan, Q., Liu, Y., Guo, Y., Li, Z., and He, Z. (2015). Fractionation of human spermatogenic cells using STA-PUT gravity sedimentation and their miRNA profiling. Sci. Rep. *5*, 8084.

Luense, L.J., Wang, X., Schon, S.B., Weller, A.H., Shiao, E.L., Bryant, J.M., Bartolomei, M.S., Coutifaris, C., Garcia, B.A., and Berger, S.L. (2016). Comprehensive analysis of histone post-translational modifications in mouse and human male germ cells. Epigenetics Chromatin *9*, 24.

Lupiáñez, D.G., Kraft, K., Heinrich, V., Krawitz, P., Brancati, F., Klopocki, E., Horn, D., Kayserili, H., Opitz, J.M., and Laxova, R. (2015). Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. Cell *161*, 1012–1025.

Maezawa, S., Yukawa, M., Alavattam, K.G., Barski, A., and Namekawa, S.H. (2018). Dynamic reorganization of open chromatin underlies diverse transcriptomes during spermatogenesis. Nucleic Acids Res. *46*, 593–608.

Naumova, N., Imakaev, M., Fudenberg, G., Zhan, Y., Lajoie, B.R., Mirny, L.A., and Dekker, J. (2013). Organization of the mitotic chromosome. Science *342*, 948–953.

Nora, E.P., Goloborodko, A., Valton, A.-L., Gibcus, J.H., Uebersohn, A., Abdennur, N., Dekker, J., Mirny, L.A., and Bruneau, B.G. (2017). Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization. Cell *169*, 930–944.e22.

Oomen, M.E., Hansen, A.S., Liu, Y., Darzacq, X., and Dekker, J. (2019). CTCF sites display cell cycle–dependent dynamics in factor binding and nucleosome positioning. Genome Res. *29*, 236–249.

Ozata, D.M., Gainetdinov, I., Zoch, A., O'Carroll, D., and Zamore, P.D. (2019). PIWI-interacting RNAs: small RNAs with big functions. Nat. Rev. Genet. *20*, 89–108.

Patel, L., Kang, R., Rosenberg, S.C., Qiu, Y., Raviram, R., Chee, S., Hu, R., Ren, B., Cole, F., and Corbett, K.D. (2019). Dynamic reorganization of the genome shapes the recombination landscape in meiotic prophase. Nat. Struct. Mol. Biol. *26*, 164–174.

Rao, S.S., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., et al. (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell *159*, 1665–1680.

Rao, S.S.P., Huang, S.C., Glenn St Hilaire, B., Engreitz, J.M., Perez, E.M., Kieffer-Kwon, K.R., Sanborn, A.L., Johnstone, S.E., Bascom, G.D., Bochkov, I.D., et al. (2017). Cohesin loss eliminates all loop domains. Cell *171*, 305–320.e24.

Rathke, C., Baarends, W.M., Awe, S., and Renkawitz-Pohl, R. (2014). Chromatin dynamics during spermiogenesis. Biochim. Biophys. Acta *1839*, 155–168.

Rosenkranz, D. (2016). piRNA cluster database: a web resource for piRNA producing loci. Nucleic Acids Res. *44* (D1), D223–D230.

Sassone-Corsi, P. (2002). Unique chromatin remodeling and transcriptional regulation in spermatogenesis. Science *296*, 2176–2178.

Shen, Y., Yue, F., McCleary, D.F., Ye, Z., Edsall, L., Kuan, S., Wagner, U., Dixon, J., Lee, L., Lobanenkov, V.V., et al. (2012). A map of the cis-regulatory sequences in the mouse genome. Nature *488*, 116–120.

Siepel, A., Bejerano, G., Pedersen, J.S., Hinrichs, A.S., Hou, M., Rosenbloom, K., Clawson, H., Spieth, J., Hillier, L.W., Richards, S., et al. (2005). Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome Res. *15*, 1034–1050.

Smagulova, F., Gregoretti, I.V., Brick, K., Khil, P., Camerini-Otero, R.D., and Petukhova, G.V. (2011). Genome-wide analysis reveals novel molecular features of mouse recombination hotspots. Nature *472*, 375–378.

Smith, C.M., Hayamizu, T.F., Finger, J.H., Bello, S.M., McCright, I.J., Xu, J., Baldarelli, R.M., Beal, J.S., Campbell, J., Corbani, L.E., et al. (2019). The mouse gene expression database (GXD): 2019 update. Nucleic Acids Res. *47*, D774–D779.

Sohni, A., Tan, K., Song, H.W., Burow, D., de Rooij, D.G., Laurent, L., Hsieh, T.C., Rabah, R., Hammoud, S.S., Vicini, E., et al. (2019). The neonatal and adult human testis defined at the single-cell level. Cell Rep. *26*, 1501–1517.e4.

Tong, M.H., and Lin, Z. (2018). m(6)A mRNA modification regulates mammalian spermatogenesis. Biochim. Biophys. Acta *1862*, 403–411.

VanGompel, M.J., and Xu, E.Y. (2010). A novel requirement in mammalian spermatid differentiation for the DAZ-family protein Boule. Hum. Mol. Genet. *19*, 2360–2369.

Vara, C., Paytuvi-Gallart, A., Cuartero, Y., Le Dily, F., Garcia, F., Salva-Castro, J., Gomez, H.L., Julia, E., Moutinho, C., Aiese Cigliano, R., et al. (2019). Three-dimensional genomic structure and cohesin occupancy correlate with transcriptional activity during spermatogenesis. Cell Rep. *28*, 352–367.e9.

Vauti, F., Prochnow, B.R., Freese, E., Ramasamy, S.K., Ruiz, P., and Arnold, H.H. (2007). Arp3 is required during preimplantation development of the mouse embryo. FEBS Lett. *581*, 5691–5697.

Wang, Q., Khillan, J., Gadue, P., and Nishikura, K. (2000). Requirement of the RNA editing deaminase ADAR1 gene for embryonic erythropoiesis. Science *290*, 1765–1768.

Wang, Y., Wang, H., Zhang, Y., Du, Z., Si, W., Fan, S., Qin, D., Wang, M., Duan, Y., Li, L., et al. (2019). Reprogramming of meiotic chromatin architecture during spermatogenesis. Mol. Cell *73*, 547–561.e6.

Watanabe, T., Tomizawa, S.-i., Mitsuya, K., Totoki, Y., Yamamoto, Y., Kuramochi-Miyagawa, S., Iida, N., Hoki, Y., Murphy, P.J., and Toyoda, A. (2011). Role for piRNAs and noncoding RNA in de novo DNA methylation of the imprinted mouse Rasgrf1 locus. Science *332*, 848–852.

Weintraub, A.S., Li, C.H., Zamudio, A.V., Sigova, A.A., Hannett, N.M., Day, D.S., Abraham, B.J., Cohen, M.A., Nabet, B., Buckley, D.L., et al. (2017). YY1 is a structural regulator of enhancer-promoter loops. Cell *171*, 1573–1588.e28.

Yamaguchi, K., Hada, M., Fukuda, Y., Inoue, E., Makino, Y., Katou, Y., Shirahige, K., and Okada, Y. (2018). Re-evaluating the localization of sperm-retained histones revealed the modification-dependent accumulation in specific genome regions. Cell Rep. *23*, 3920–3932.

Yoshida, K., Muratani, M., Araki, H., Miura, F., Suzuki, T., Dohmae, N., Katou, Y., Shirahige, K., Ito, T., and Ishii, S. (2018). Mapping of histone-binding sites in histone replacement-completed spermatozoa. Nat. Commun. *9*, 3885.

iScience, Volume *23*

Supplemental Information

Reorganized 3D Genome Structures Support Transcriptional Regulation in Mouse Spermatogenesis

Zhengyu Luo, Xiaorong Wang, Hong Jiang, Ruoyu Wang, Jian Chen, Yusheng Chen, Qianlan Xu, Jun Cao, Xiaowen Gong, Ji Wu, Yungui Yang, Wenbo Li, Chunsheng Han, C. Yan Cheng, Michael G. Rosenfeld, Fei Sun, and Xiaoyuan Song

Supplemental Figures

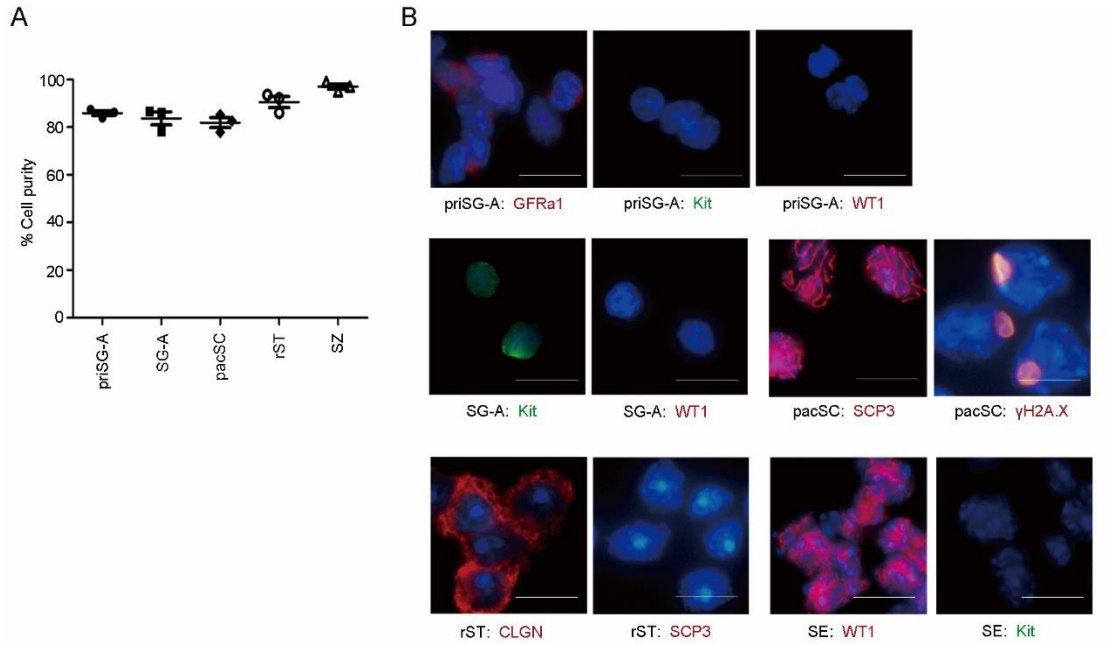


Figure S1. Purities of five isolated spermatogenic cell types. Related to Figure 1. **(A)** Cell purities according to the immunofluorescence data (shown in B). The total cell number for each sample was more than 200. Data are represented as mean +/- SEM. **(B)** Immunofluorescence revealed the expression of cell type-specific marker proteins in four isolated spermatogenic cell types and sertoli cell (SE). Specifically, GFRa1 for priSG-A, Kit for SG-A, SCP3 and γH2A.X for pacSC, CLGN for rST and WT1 for SE. Scale bar, 10 μm.

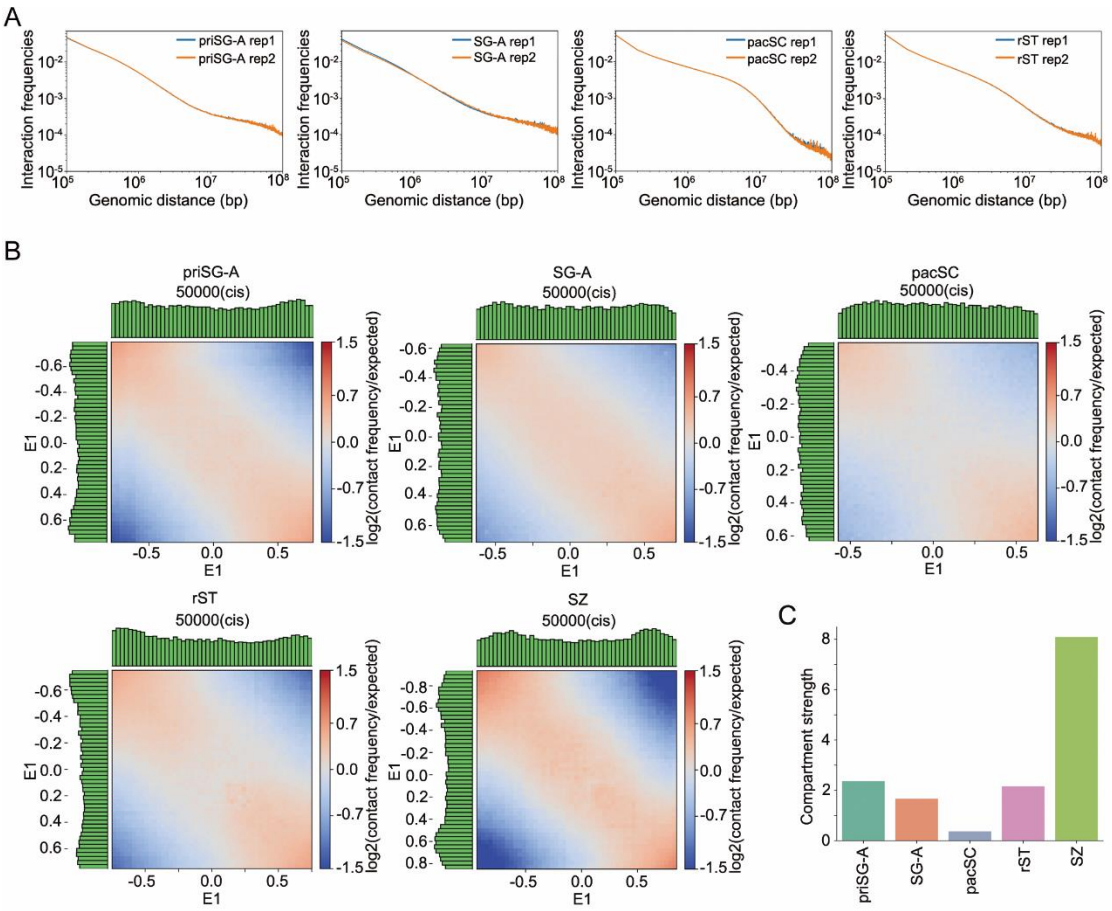


Figure S2. Dynamic compartmentalization during spermatogenesis. Related to Figure 1. **(A)** *P(s)* analysis of four stages' independent replicates of Hi-C data suggested a high correlation between biological replicates. **(B)** Compartmentalization saddle plot for each tested spermatogenic stage: average interactions between pairs of loci (50kb bin) arranged by their compartment signals (eigenvector values). Compartmentalization was weakened in pacSC and then enhanced in SZ. Histograms along the axes showed the distribution of eigenvector values. **(C)** Intrachromosomal interactions of each spermatogenic stage determined by the measurement of genomic compartment strength.

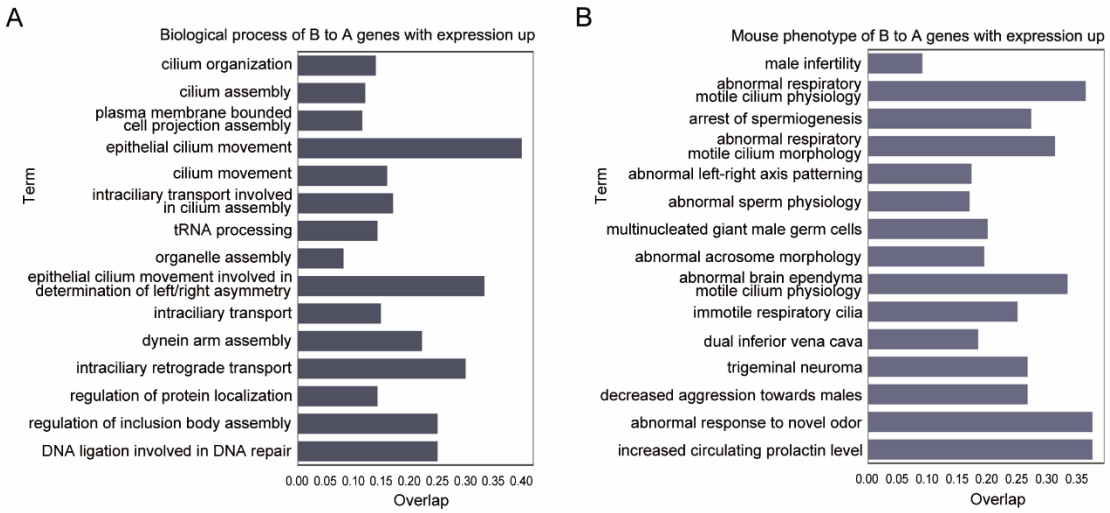


Figure S3. Potential functions of differential expressed genes with compartment changes comparing pacSC with priSG-A samples. Related to Figure 4. **(A)** Bar plot showing the significant biological process terms for those up-regulated genes with their genomic compartments switching from B to A when primitive type A spermatogonia (priSG-A) proceeded to meiosis I (pacSC). **(B)** Bar plot showing the significant mammalian phenotypes enriched from the same genes as in (A).

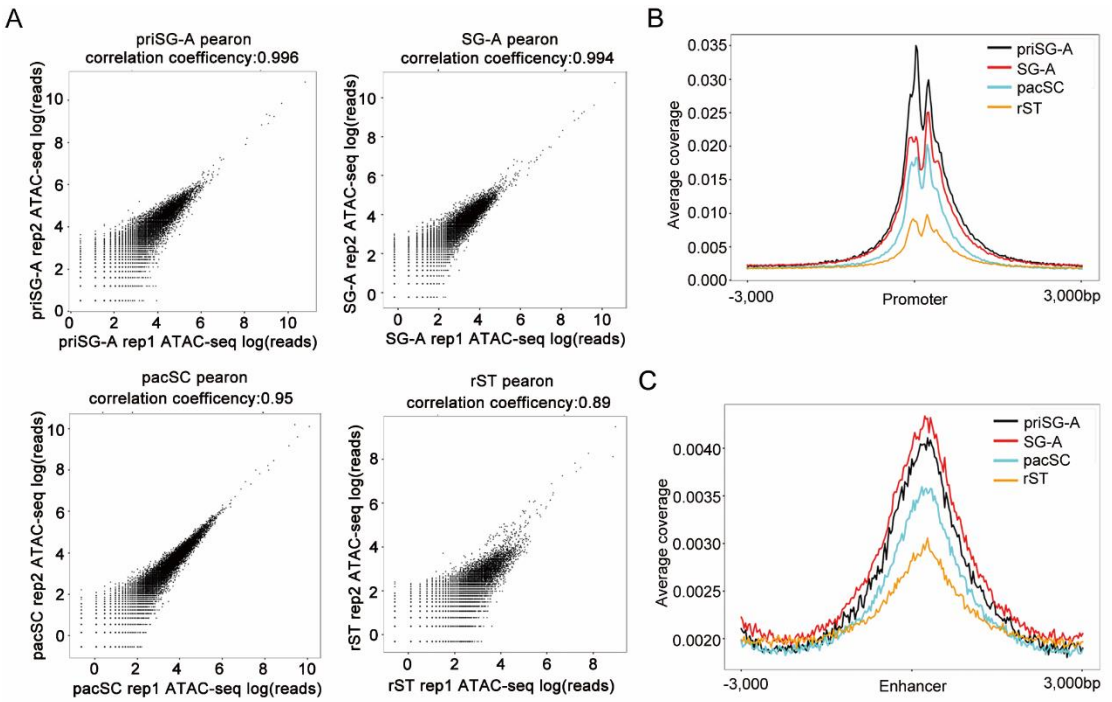


Figure S4. High correlation and the landscape of chromatin accessibility during spermatogenesis by ATAC-Seq. Related to Figure 5. **(A)** The scatter plot of two replicates of priSG-A, SG-A, pacSC and rST ATAC-Seq signals. **(B)** Metagene plot of ATAC-Seq signals at promoter regions (±3kb) of known genes showed that chromatin accessibility was gradually decreased during spermatogenesis. **(C)** Metagene plot of ATAC-Seq signals at enhancer regions (±3kb) showed that chromatin accessibility was gradually decreased during spermatogenesis.

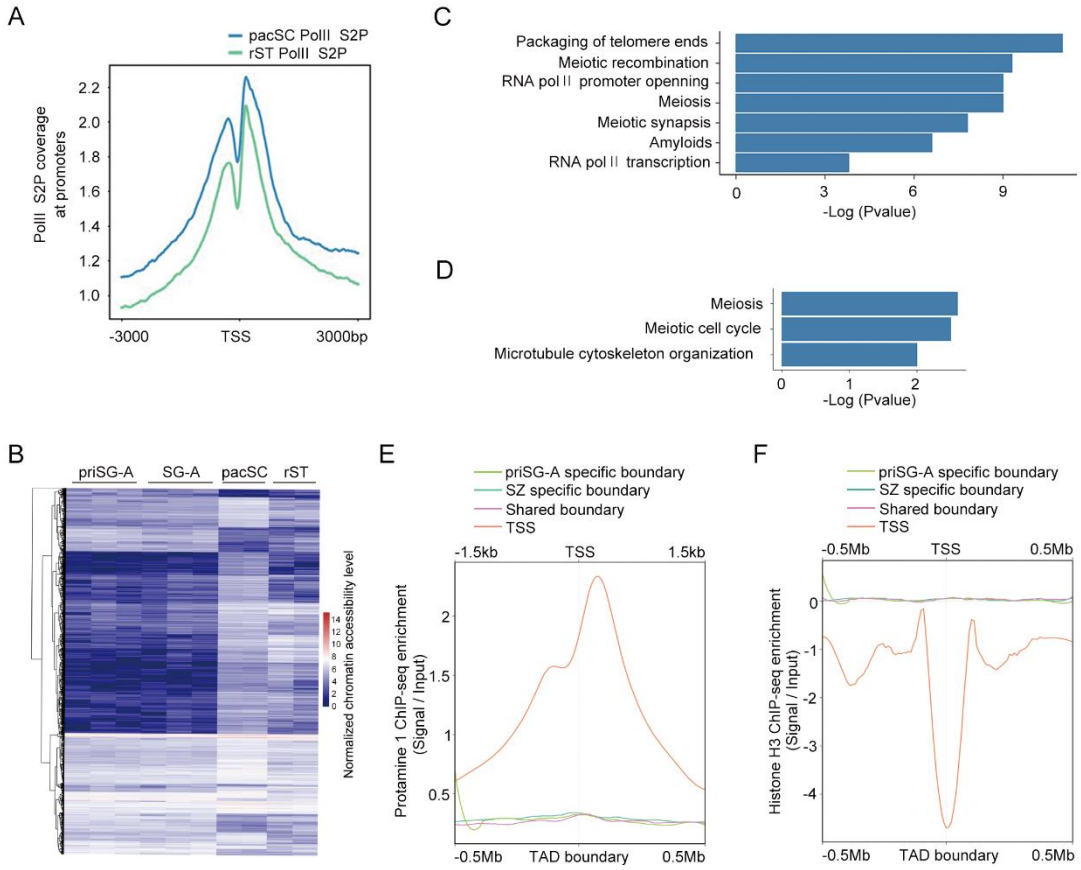


Figure S5. Chromatin accessibility during mouse spermatogenesis and histone retainment in the sperm. Related to Figure 5. **(A)** Profile plot showing the average PolI II signal (input RPKM was subtracted from Pol II RPKM) in pacSC and rST samples around the transcription start site (TSS) (±3kb) of all genes. **(B)** The quantification heatmaps of all chromatin accessible sites identified in four cell types of spermatogenesis. There were 2 or 3 biological replicates for each cell type. **(C)** The functional enrichment of the differentially accessible chromatin regions between priSG-A and pacSC. **(D)** The functional enrichment of the differentially accessible chromatin regions between pacSC and rST. **(E)** Profile plot showing the average protamine1 ChIP-Seq signals (input RPKM was subtracted from protamine 1 RPKM) in SZ around different kinds of TAD boundaries (±0.5Mb) (the priSG-A and SZ specific TAD boundaries and shared TAD boundaries, respectively). Orange line showing the average protamine1 ChIP-Seq signals in SZ around TSS (±1.5kb). **(F)** Profile plot showing the average histone H3 ChIP-Seq signals (input RPKM was subtracted from histone H3 RPKM) in SZ around different kinds of TAD boundaries (±0.5Mb) (the priSG-A and SZ specific TAD boundaries and shared TAD boundaries, respectively). Datasets of protamine 1 ChIP-Seq, histone H3 ChIP-Seq as well as their input control were obtained from (Yoshida et al., 2018) and (Yamaguchi et al., 2018).

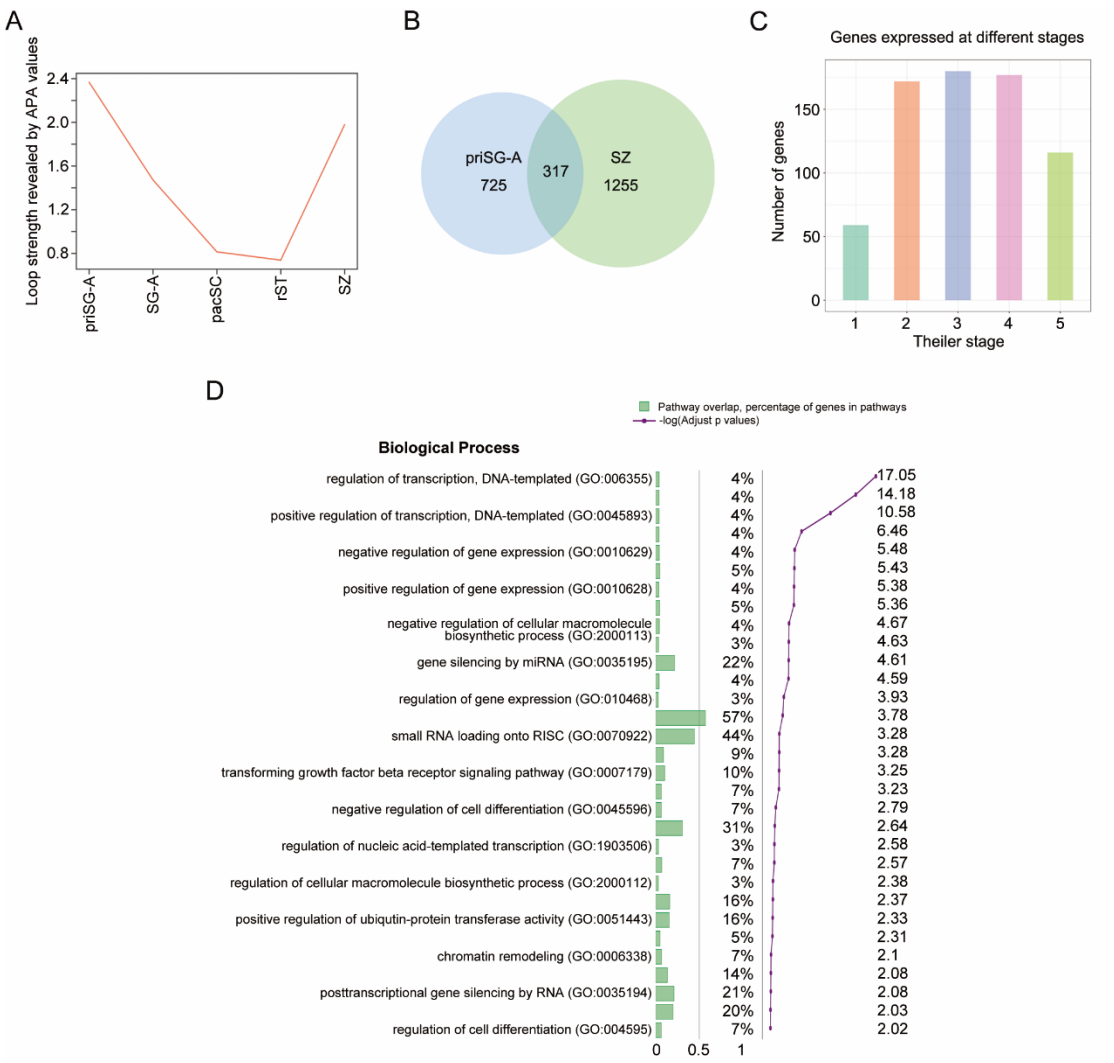


Figure S6. Loop strength during mouse spermatogenesis and the potential function of SZ loops related genes in early embryo. Related to Figure 6. **(A)** The APA Peak to Lower Left (P2LL, the ratio of the central pixel to the mean of the pixels in the lower-left corner, which represents the strength of chromatin loops) values quantitatively characterized the dynamics of chromatin loops strength across different spermatogenic stages. **(B)** There were 1,096 genes' promoters located in loop anchors in priSG-A and 1,572 genes' promoters located in loop anchors in SZ, there were 317 genes' promoters shared in both priSG-A and SZ. **(C)** Of the 1572 genes whose promoters located in loop anchors in SZ samples, the number of genes expressed in early embryo (Theiler Stage 1-5) was shown by bar plot. **(D)** GO-based biological process enrichment data based on genes positioned at enhancer-promoter loops expressed in early embryonic stages, and a bar plot showing the proportion of these genes pertaining to the annotation of each phenotype over the number of all genes annotated in each phenotype. Line plot showing the significance (-log (adjusted p-values) using Fisher exact test) of each ontology.

Supplemental Tables

Supplementary Table 1. Summary of the Hi-C data. Related to Figure 1, Figure 2, Figure 3, Figure4, Figure 5, and Figure 6.

	PriSG-A rep1	PriSG-A rep2	SG-A rep1	SG-A rep2	pacSC rep1	pacSC rep2	rST rep1	rST rep2
Total_pairs_processed	396,320,543	22,845,016	676,335,638	21,604,566	380,700,331	25,609,400	630,891,088	22,317,535
Unmapped_pairs	7,723,837	282,327	8,798,453	16,887,929	6,768,715	300,448	11,189,457	250,109
Low_qual_pairs	0	0		0	0	0		0
Unique_paired_alignments	243,850,710	15,259,577	469,710,800	3,187,617	225,862,507	16,429,889	395,619,288	14,924,284
Multiple_pairs_alignments	90,037,310	5,472,927	133,185,090	1,037,953	100,808,260	7,111,075	149,453,231	5,476,387
Pairs_with_singleton	54,708,686	1,830,185	64,641,295	491,067	47,260,849	1,767,988	74,629,112	1,666,755
Low_qual_singleton	0	0		0	0	0		0
Unique_singleton_alignments	0	0		0	0	0		0
Multiple_singleton_alignments	0	0		0	0	0		0
Reported_pairs	243,850,710	15,259,577	469,710,800	3,187,617	225,862,507	16,429,889	395,619,288	14,924,284
valid_interaction	254,161,487		457,604,091		237,259,853		403,391,384	
valid_interaction_rmdup	214,380,530		172,251,595		205,275,466		341,791,610	
trans_interaction	64,029,379		58,676,598		26,275,954		84,732,291	
cis_interaction	150,351,151		113,574,997		178,999,512		257,059,319	
cis_shortRange	16,732,869		16,069,826		25,302,081		40,767,712	
cis_longRange	133,618,282		97,505,171		153,697,431		216,291,607	
Unique_paired_alignments	254,161,487		457,604,091		237,259,853		403,391,384	

Supplementary Table 2. Genes with expression detected in Theiler Stage 1-5 from MGI-Mouse Gene Expression Database (Smith et al., 2019). PDG: protein coding gene. Related to Figure 6.

MGI Gene ID	Gene Symbol	Gene Name	Type	Chr	Genome Location-GRCm38	cM	Strand
MGI:87911	Acvr1	activin A receptor, type 1	PDG	2	58446438-58566828	33.05	-
MGI:2448562	Adnp2	ADNP homeobox 2	PDG	18	80126311-80151482	53.28	-
MGI:2442590	Ankrd35	ankyrin repeat domain 35	PDG	3	96670131-96691032	41.94	+
MGI:108405	Apbb2	amyloid beta (A4) precursor protein-binding, family B, member 2	PDG	5	66298703-66618784	34.43	-
MGI:108028	Atr	ataxia telangiectasia and Rad3 related	PDG	9	95857597-95951781	50.27	+
MGI:1298392	Bscl2	Berardinelli-Seip congenital lipodystrophy 2 (seipin)	PDG	19	8837467-8848683	5.76	+
MGI:88455	Col4a2	collagen, type IV, alpha 2	PDG	8	11312805-11449287	5.62	+
MGI:94865	Dbi	diazepam binding inhibitor	PDG	1	120113280-120121078	52.65	-
MGI:99892	Lama1	laminin, alpha 1	PDG	17	67697259-67822647	38.8	+
MGI:97275	Myod1	myogenic differentiation 1	PDG	7	46376474-46379092	30.03	+
MGI:2441856	Sf3b2	splicing factor 3b, subunit 2	PDG	19	5273932-5295455	4.29	-
MGI:108078	Sfrp2	secreted frizzled-related protein 2	PDG	3	83766321-83774316	37.37	+
MGI:103063	Stat1	signal transducer and activator of transcription 1	PDG	1	52119440-52161865	26.81	+
MGI:98729	Tgfbr2	transforming growth factor, beta receptor II	PDG	9	116087695-116175363	68.39	-
MGI:1341872	Tjp2	tight junction protein 2	PDG	19	24094505-24225030	19.17	-
MGI:1926031	Zc3hav1	zinc finger CCCH type, antiviral 1	PDG	6	38305286-38354603	17.72	-

Transparent Methods

Experimental Animals

C57BL/6 mice were housed in the Animal Center of the University of Science and Technology of China and were cultured under a 12-h light/dark cycle (lights off at 7 p.m.) at 23±2°C. All animal experiments were carried out in compliance with the guidelines of the institutional review board of the University of Science and Technology of China, which approved this study.

Purification of male germ cells during spermatogenesis

C57BL/6 mice were originally purchased from Vital River Laboratories in Beijing, China. Primitive type A spermatogonia (priSG-A), type A spermatogonia (SG-A), pachytene spermatocytes (pacSC), and round spermatids (rST) were isolated using a unit gravity sedimentation procedure (STA-PUT method) based on Bellvé and Gan´s description (Bellvé et al., 1977; Gan et al., 2013) with minor modifications. priSG-A were isolated from 35 male mice at 6-days postpartum (dpp); SG-A were isolated from 25 8-dpp male mice; pacSC and rST were isolated from 5 adult male mice. Briefly, the testes were removed and minced by scissors until a semiliquid state had been achieved, followed by incubation in 40 ml DMEM (Gibco, 11995-081) containing 1 mg/ml collagenase IV (Sigma, C5138-1G) and 1 unit/ml DNase I (Sigma, AMPD1-1KT) in a shaking water bath at 32 °C for 10 min. To stop the digestion, 20 ml of fresh DMEM containing 10% FBS was added to samples. The seminiferous tubules were collected by centrifugation at 500 g for 2 min. The pellet was washed once with DMEM and was resuspended in 40 ml DMEM containing 1 mg/ml Trypsin (Sigma, T1426-500MG) and 1 unit/ml DNase I (Sigma, AMPD1-1KT), and incubated in a shaking water bath at 32 °C for 10 min. 20 ml of fresh DMEM containing 10% FBS was added to stop the digestion. Cells were collected by centrifugation at 500 g for 2 min. The cell pellet was washed twice with DMEM and resuspended in 40 ml DMEM containing 0.5% BSA (Sangon, AD0023-100g). Then the cells were filtered through a 40 μm Nylon Cell Strainer (BD Falcon, 352340) and

separated by sedimentation velocity at unit gravity at 4°C, using 2-4% BSA gradient in DMEM. Only fractions of the expected cell type and purity (≥ 75%) were pooled together. The collected cells were then cultured in 10 ml DMEM containing 10% FBS in a 10 cm diameter tissue culture dish pre-coated with 0.1 mg/ml poly-D-lysine for 3 hours at 34 °C. Sertoli cells (SE) attached to the culture plates, and the germ cells in suspension were collected by centrifugation at 500 g for 5 min. Spermatozoa (SZ) were isolated from 5 adult male mice. The cauda epididymis was removed and cross-cut in a tissue culture dish containing PBS (Thermo, 21600-010) preheated at 37 °C, and SZ swam up to within about 10 min. The motile SZ were collected and washed twice with PBS, followed by filtration through a 40 μm Nylon Cell Strainer and collection via centrifugation at 500 g for 5 min. The purity of SZ was evaluated as ~95% based on morphological characterization. The purities of the other isolated cells were evaluated and identified by their morphological characterization and immunofluorescence staining with germ cell type-specific markers (GFRa1 for priSG-A, Kit for SG-A, SCP3 and γH2A.X for pacSC, CLGN for rST, and WT1 for SE).

Immunofluorescence

Immunofluorescence analysis was performed to identify the cell type and purity of the isolated cells. Since the isolation methods were routine for us with high purity and success (Gan et al., 2013), only part of the isolated cell types was used to perform Immunofluorescence. Specifically, the cell pellet was resuspended in 5 μl FBS and transferred to glass slides precoated with 1% gelatin. 10 min later, cells were fixed with 4% paraformaldehyde (Sigma, V900894-100G) for 30 min and washed three times with PBS, permeabilized with 0.5% Triton X-100 (Sigma, 93443-100ML)/PBS for 10 min and washed three times with PBS, and blocked with 3% BSA/PBS for 1 hour. The cells were then incubated with primary antibodies including rabbit anti-GFRa1 (Abcam, ab186855), rabbit anti-Kit (Abcam, ab5506), rabbit anti-WT1 (Abcam, ab89901),

rabbit anti-SCP3 (Abcam, ab15093), rabbit anti-γH2A.X (Cell Signaling Technology, 9718S), or rabbit anti-CLGN (Abcam, ab171971) at a 1 : 200 dilution and overnight at 4°C. After washing three times with 0.1% Triton X-100/PBS for 5 min, the cells were incubated with the secondary antibodies, including goat anti-rabbit IgG H&L (Alexa Fluor® 488) (Abcam, ab150077) or donkey anti-rabbit IgG H&L (Alexa Fluor® 555) (Abcam, ab150074) at a dilution with 1 : 200 for 1 hour at room temperature. After washing three times with 0.1% Triton X-100/PBS for 10 min, nuclei were stained with Hoechst 33342 (Sigma, 14533-100MG) and washed three times with PBS for 5 min. The fluorescent signals were examined using a fluorescence microscope. The purity of each sample was calculated as a percentage of positive signal cells among total cells (total cells > 200).

In situ Hi-C

In situ Hi-C was used to explore the higher order chromatin organization with isolated cells during mouse spermatogenesis (Rao et al., 2014). Briefly, about 5 x 10^5 to 5 x 10^6 cells were crosslinked using formaldehyde (a final 1% (v/v); Sigma, F8775-500ML), followed by quenching with a Glycine (final 125 mM; Sigma, G8898-1KG), mixed well and incubated for another 10 min. Cells were pelleted and washed once with ice-cold 1x PBS. After removing the supernatant, the cell pellets were stored at -80°C or directly used for the following Hi-C experiments. The cells were resuspended in 500 μl of ice-cold lysis buffer (10 mM Tris-HCl, pH8.0, 10 mM NaCl and 0.2% (v/v) Igepal CA-630 (Sigma, 18896-100 ML)) containing 1x proteinase inhibitor complex (PIC, Roche, 11 873 580 001), lysed on ice for at least 20 min. After lysis, cells were pelleted at 2500 g for 4 min at 4 °C then resuspended in 50 μl of 0.5% (w/v) SDS (Thermo, 24730020) and incubated at 62 °C for 10 min, followed by adding 25 μl 10% (v/v) Triton X-100 (Sigma, X100-100ML) and 145 μl of water, then incubated at 37 °C for 10 min with shaking (800rpm). 10 μl of MboI (NEB, R0147M) and 31 μl of 10x CutSmart buffer (NEB, B7204S) were added to the sample before an

overnight incubation at 37 °C with shaking for digestion. The next day, after incubating the tubes at 62 °C for 20 min, the total sample volume was brought to 1200 μl with final 1x NEB DNA ligase reaction buffer (NEB, B0202S) and 1x BSA (NEB, B9000S); 4 μl of T4 DNA ligase (NEB, M0202L) was added, and the tubes were incubated at room temperature for 4 hours with slow rotation. After Hi-C ligation, 120 μl of 10% (w/v) SDS and 20 μl of proteinase K (Thermo, EO0492) were added and incubated at 65 °C overnight. DNA was then purified using an ethanol precipitation method and sheared into 200-500 bp fragments using a sonicator (NingBoXinZhi, JY92-IIN). The DNA fragments were purified again using VAHTS DNA Clean Beads (Vazyme, N411-02) and eluted in 100 μl of water. 30 μl of washed Dynabeads M-280 (Thermo, 11205D) were resuspended in 100 μl of 2x Bind buffer and mixed with the purified DNA, and the tubes were incubated at room temperature for 30 min with slow rotation. After 3 washes, the samples were prepared for sequencing on beads. After end repairing, dATP tailing and adapter ligation, the DNA was washed 5 times with TWB buffer (5 mM Tris-HCl, pH7.5, 0.5 mM EDTA and 1 M NaCl) and resuspended in 50 μl of water. The DNA on the beads was then used as the template for PCR amplification with Phusion DNA polymerase (NEB, M0530) and purified with VAHTS DNA Clean Beads (Vazyme, N411-01) to select the DNA fragments between 200bp and 600bp. All of the Hi-C libraries were then sent to a commercial sequencing company (Novogene Co., LTD) and sequenced on the HiSeq X ten platform.

ATAC-Seq

ATAC-Seq was performed following a previously published method (Buenrostro et al., 2013). Briefly, 5×10^4 cells were spun at 500 g for 5 min at 4 °C, and the cell pellet was washed with 50 μl of cold 1× PBS and collected by centrifugation at 500 g for 5 min at 4 °C. Then the cell pellet was resuspended in 50 μl cold lysis buffer (10 mM Tris-HCl, pH 7.4, 10 mM NaCl, 3 mM MgCl2 and 0.1% (v/v) Igepal CA-630) and centrifuged at 500 g for 10 min at 4 °C. Discard the

supernatant and place the cell pellet on ice. Resuspend the cell pellet in the transposase reaction mix (25 µl 2× TD buffer, 2.5 µl transposase (Illumina, FC-121-1030) and 22.5 µl nuclease-free water). The transposition reaction was carried out for 30 min at 37 °C. Then, the samples were purified by using a Qiagen MinElute PCR Purification Kit (Qiagen, 28004) and amplified by PCR with NEBNext® High-Fidelity 2X PCR Master Mix (NEB, M0541L). After size selected by VAHTS DNA Clean Beads (Vazyme, N411-01), the samples were sequenced by HiSeq X Ten at Novogene.

Native ChIP-Seq

Native ChIP was performed following a previously published report (Gilfillan et al., 2012). Briefly, the cells were washed once with 1x PBS and then pelleted at 800 g for 5 min. The supernatant was discarded and the cells pellet was stored at -80 °C or directly used for the following ChIP experiments. For ChIP, about 2 x 10^5 cells were first resuspended in 20 μl of MNase working buffer (50 mM Tris-HCl pH8.0, 1 mM CaCl2, 0.2% Triton X-100 and final 1x proteinase inhibitor complex (PIC, Roche, 11 873 580 001) followed by the addition of 1 μl of 0.01U/μl MNase (for Histone) (Sigma, N3755-50UN) and incubation at 37°C for 2 min. 2.4 μl of 10x MNase stop buffer was then added and the tube was placed on ice. Subsequently, 23 μl of ice-cold 2x RIPA buffer (280 mM NaCl, 1.8% Triton X-100, 0.2% SDS, 0.2% Na-Deoxycholate, 5 mM EGTA, 1x PIC) and 155 μl of ice-cold 1x RIPA buffer (10 mM Tris pH 8.0, 1 mM EDTA, 140 mM NaCl, 1 % Triton X‐100, 0.1% SDS, 0.1 % Na‐Deoxycholate, 1x PIC) were added to the sample and mixed before centrifugation at 16000 g at 4 °C for 10 min, after which the supernatant was transferred to a new tube. Protein A/G beads (Thermo, 88803) that had been pre-washed with 1x RIPA buffer were then added (30 μl per IP reaction) followed by incubation at 4 °C for 1 hour with slow rotation. The beads were collected via a magnet and the supernatant was transferred to a new tube. 10% of the supernatant was then used as the input, to which 5 μl of proteinase K (Thermo, EO0492) was added and the volume

was brought up to 100 μl with TE buffer (10 mM Tris pH8.0, 5 mM EDTA pH 8.0), followed by incubating with shaking at 55 °C for 1 hour and storage at -20 °C. For IP, brought volume to 100 μl with RIPA buffer for each IP reaction, mixed with CTCF (Santa Cruz Biotechnology, sc-271514/sc-28198/sc-15914), Rad21 (Abcam, ab992) or RNA Pol II S2P (Abcam, ab5095) antibody and incubated overnight at 4 °C with slow rotation. The next day, 30 μl of protein A/G beads were added to each IP reaction and incubated for another 2 hours with slow rotation at 4 °C. The beads were then washed 5 times with RIPA buffer and washed once with LiCl wash buffer (250 mM LiCl, 10 mM Tris-HCl pH8.0, 1 mM EDTA, 0.5% NP-40, 0.5% sodium deoxycholate), resuspended beads in 100 μl of TE buffer with 5 μl of proteinase K and incubated for 1 hour with shaking at 55 °C. For IP samples containing IgG control and the Input sample, purified the DNA and eluted in 20 μl of water. The sequencing libraries were generated by Tn5 transposase and then PCR amplification and size selected by VAHTS DNA Clean Beads (Vazyme, N411-01). Libraries were sent for high throughput sequencing with the HiSeq X Ten at Novogene.

Hi-C: alignment

Hi-C dataset from cells at the SZ stage was obtained from published work (Jung et al., 2017). Paired-end .fastq files of Hi-C libraries were mapped using HiC-Pro (v 2.8.1) (Servant et al., 2015). Matrix binning and balancing was performed with the cooler software package (0.8.6.post0). In brief, paired reads were iteratively mapped to the *Mus musculus* mm10 genome via bowtie2 (v 2.2.5) (Langmead and Salzberg, 2012); the unmapped reads contain the MboI ligation sites were then trimmed and aligned back to the mm10 genome again. After combing the two-round mapping results, uniquely aligned paired reads were kept and assigned to MboI restriction fragments. The paired reads with dangling ends and or predicted self-circle structures pairs were excluded from valid pairs. This alignment and filtering process resulted in sequencing reads described in (Table S1). We merged replicates and sampled all valid pairs of each sample

to the same 172,252,595 reads, which were further converted to Hi-C contact matrices binned into 100kb, 50kb, 25kb, and 5kb sizes respectively (Abdennur and Mirny, 2019). The binned contact matrices were further normalized by using an iterative correction method (Servant et al., 2015). The Hi-C matrices were further transferred into the .hic format to facilitate visualization with juicebox (Durand et al., 2016).

Hi-C: interaction frequency curves

The interaction frequency curves (*P(s)*) were calculated with normalized interaction matrices in 100kb resolution following methods used in (Naumova et al., 2013). We first divided distances into logarithmically spaced bins and counted the number of interactions at corresponding distances at each bin. To obtain the probability *P(s)*, we divided the number of interactions in each bin by the total number of possible region pairs.

Hi-C: A/B Compartment, TADs, and chromatin loops analysis

The compartment profiles of Hi-C matrices at 50kb resolution were identified using cooltools (https://github.com/mirnylab/cooltools) eigdecomp.py. To do this, we first calculate the intra-chromatin observed/expected interaction matrix. We used the interaction matrix at 500kb resolution and calculated the mean number of interactions at a given genomic distance, d, using a sliding window with linearly increasing size, and get the expected contact probability. Then the observed/expected matrix was used to generate a Pearson correlation matrix, which was subject to principal component analysis (PCA). The eigenvector of the first principal component defines the A/B compartment profile (Lieberman-Aiden et al., 2009). With convention, genomic compartments were assigned to ‘A’ (active/euchromatic compartments) and ‘B’ (inactive/heterochromatic compartments) primarily based on associated mRNA transcription. Compartment saddle analysis was performed by using cooltools (Gibcus et al., 2018; Schwarzer et al., 2017). Briefly, we used the observed/expected Hi-C

maps, obtained by dividing each diagonal of a matrix by its chromosome-wide average value from 50kb binned iteratively corrected interaction maps of cis interactions. In each observed/expected map, we rearranged the rows and the columns in the order of increasing eigenvector value. Finally, we aggregated the rows and the columns of the resulting matrix into 30 equally sized aggregated bins, thus obtaining a compartmentalization plot ("saddle plot"). The compartment strength was calculated from the deviation of the averaged interaction intra strong A (with highest 25% of pca1 values) and strong B (lowest 25% of pca1 values) loci from the interaction inter those string A and strong B loci: Mean (strong A-strong A interactions) + mean (strong B-strong B interactions) – 2 [mean (strong A-strong B interactions)].

The topologically associating domains (TADs) were identified following an insulation square analysis (Crane et al., 2015). Briefly, to calculate the 'insulation' score of each bin in the 25kb binned Hi-C data, we sided a sequential size of squares (from 100kb to 1000kb, 50kb as steps) along the diagonal of the interaction matrix for each chromosome, and aggregated all signals within the square. A 125kb window was used for calculation of the delta vector. Only boundaries for which more than half of the differentially sized squares had a 'boundary strength' value larger than 0.25 were retained. Insulation scores were plotted around all priSG-A TADs as well as their nearby regions (+/- 0,5 TAD length). Using all of the priSG-A TADs as control TADs, we then plotted the composite interaction by averaging interactions near TAD boundaries (range) for each of all the sampled stages of spermatogenesis.

The chromatin loops were identified using cLoops (Cao et al., 2018) with parameters -w -j -s -m 3. The APA (Aggregation Peak Analysis) was performed following the methods developed by (Rao et al., 2014). Briefly, to measure the enrichment of loops over the local background, the KR normalized contact frequency of pixels of loops, as well as the surrounding pixels up to 10 bins away in both x and y directions, i. e., 50kb*50kb local contact matrices, were collected and plotted. APA scores (P2LL, the ratio of the central pixel to the

mean of the pixels in the lower-left corner, representing the strength of chromatin loops) were determined by dividing the center pixel value by the mean value of the 25 (5*5) pixels in the lower right section of the APA plot.

We downloaded lists of active promoters (marked by H3K4me3 or PoI II binding signals) and enhancers (marked by H3K4me1 or H3K27ac) in the sperm (Shen, et al., 2012). We used their genomic loci to annotate our loop anchors in each stage of spermatogenesis. We then queried the genes with promoter-enhancer loops to MGI-Mouse Gene Expression Database (Smith et al., 2018) and found genes with detectable expression in Theiler Stage 1-5.

ATAC-Seq data analysis

All ATAC-Seq sequencing data were mapped to the mm10 reference genome using snap-aligner (v 1.0) (Zaharia et al., 2011). Samtools was then used to further sorted, index, and remove duplicates from the aligned reads (Li et al., 2009). The ATAC-Seq peaks of all samples were called using MACS2 (v 2.1.0) with the default parameters (Zhang et al., 2008). Pearson correlation coefficients were calculated based on the normalized reads number for merged peaks from all samples between different samples. The ATAC-Seq reads at identified peaks were quantified using bedtools (Quinlan and Hall, 2010) and normalized using the quantile norm function in R. The differential accessible chromatin regions were identified using DESeq2 (Love et al., 2014). The functional enrichment analysis of the differentially accessible chromatin regions was performed using GREAT (McLean et al., 2010). The motif enrichment analysis of differentially accessible chromatin regions was performed using the HOMER findMotif function (Heinz et al., 2010). The list of meiotic DNA double strand break sites was generated from previously published DMC1 ChIP-Seq (Smagulova et al., 2011). The list of piRNA cluster sites from 14.5dpp mouse was obtained from a piRNA cluster database (Rosenkranz, 2015). phastCons scores were downloaded from ucsc genome browser (http://hgdownload.cse.ucsc.edu/goldenPath/mm10/phastCons60way/

mm10.60way.phastCons.bw).

ChIP-Seq analysis

For our Pol II S2P ChIP-Seq data, CTCF and Rad21 ChIP-Seq data, and protamine 1 ChIP-Seq data from GSM2088400 and GSM2401441, histone H3 ChIP-Seq data from DRA006537, we trimmed their adapters using NGmerge (v 0.3) and trimmed right end low quality sequence from reads using BBDuk from BBMap (v 37.62). Reads were mapped to the mm10 reference genome using bowtie2 (v 2.2.5) (Langmead and Salzberg, 2012), and duplicated reads were removed using samblaster (v 0.1.24) (Faust and Hall, 2014). We subsampled 21445601 reads from our Pol II S2P ChIP-Seq and Rad21 ChIP-Seq align file (.bam) using sambamba (v 0.7.0) (Tarasov et al., 2015) and compared the same-depth treated bam files with Input bam files using bamCompare (v 3.3.0) with the parameters –normalizeUsing RPKM –operation ratio --binSize 30 --smoothLength 300 --scaleFactorsMethod None --extendReads 200. For public protamine1 and histone H3 ChIP-seq data, we subsampled reads from align file (.bam) using sambamba (v 0.7.0) and compared the same-depth treated bam files with Input bam files using bamCompare (v 3.3.0) with the parameters –normalizeUsing RPKM –operation subtract --binSize 30 --smoothLength 300 --scaleFactorsMethod None --extendReads 200. Public mESC CTCF and Rad21 ChIP-seq and Input .bw files were downloaded from GSE102997. The ChIP-Seq peaks of CTCF were called using MACS2 (v 2.1.0) with default parameters (Zhang et al., 2008). The CTCF signal profile near TADs boundary was calculated by averaging reads in each distance bin using computeMatrix and plotProfile functions of deepTools (3.3.0) (Ramírez et al., 2016). All of the visualizations of reads used pyGenomeTracks (3.0) (Ramírez et al., 2018).

RNA-Seq analysis

RNA-Seq data used in this study were obtained from published data (Lin et al.,

2016). Paired-end fastq files were quality confirmed by fastQC (Andrews S., 2010. FastQC: a quality control tool for high throughput sequence data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc) and aligned to the mm10 reference genome using hisat2 (Kim et al., 2015). Duplicated and multiple hit reads were excluded using samtools (Li et al., 2009). Annotation and counting of the aligned reads were performed using stringtie (Pertea et al., 2016). Deferentially expressed genes were identified via DESeq2 package in R (Love et al., 2014). Gene ontology of gene sets was performed using enrichR (Kuleshov et al., 2016).

DATA AND CODE AVAILABILITY

The raw and processed sequencing data generated in this study had been submitted to the NCBI Gene Expression Omnibus (GEO; http://www.ncbi.nlm.nih.gov/geo/) under accession number GSE147536.

Supplemental References

Abdennur, N., and Mirny, L.A. (2019). Cooler: scalable storage for Hi-C data and other genomically labeled arrays. Bioinformatics.

Bellvé, A.R., Cavicchia, J., Millette, C.F., O'brien, D.A., Bhatnagar, Y., and Dym, M. (1977). Spermatogenic cells of the prepuberal mouse: isolation and morphological characterization. The Journal of cell biology *74*, 68-85.

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nature Methods *10*, 1213-+.

Cao, Y., Chen, X., Ai, D., Chen, Z., Chen, G., McDermott, J., Huang, Y., and Han, J.-D.J. (2018). Accurate loop calling for 3D genomic data with cLoops. bioRxiv, 465849.

Crane, E., Bian, Q., McCord, R.P., Lajoie, B.R., Wheeler, B.S., Ralston, E.J., Uzawa, S., Dekker, J., and Meyer, B.J. (2015). Condensin-driven remodelling of X chromosome topology during dosage compensation. Nature *523*, 240-U299.

Durand, N.C., Robinson, J.T., Shamim, M.S., Machol, I., Mesirov, J.P., Lander, E.S., and Aiden, E.L. (2016). Juicebox provides a visualization system for Hi-C contact maps with unlimited zoom. Cell systems *3*, 99-101.

Faust, G.G., and Hall, I.M. (2014). SAMBLASTER: fast duplicate marking and structural variant read extraction. Bioinformatics *30*, 2503-2505.

Gan, H., Wen, L., Liao, S., Lin, X., Ma, T., Liu, J., Song, C.-x., Wang, M., He, C., and Han, C. (2013). Dynamics of 5-hydroxymethylcytosine during mouse spermatogenesis. Nature communications *4*.

Gibcus, J.H., Samejima, K., Goloborodko, A., Samejima, I., Naumova, N., Nuebler, J., Kanemaki, M.T., Xie, L., Paulson, J.R., Earnshaw, W.C*., et al.* (2018). A pathway for mitotic chromosome formation. Science.

Gilfillan, G.D., Hughes, T., Sheng, Y., Hjorthaug, H.S., Straub, T., Gervin, K., Harris, J.R., Undlien, D.E., and Lyle, R. (2012). Limitations and possibilities of low cell number ChIP-seq. BMC genomics *13*, 645.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Molecular cell *38*, 576-589.

Jung, Y.H., Sauria, M.E., Lyu, X., Cheema, M.S., Ausio, J., Taylor, J., and Corces, V.G. (2017). Chromatin states in mouse sperm correlate with embryonic and adult regulatory landscapes. Cell reports *18*, 1366-1382.

Kim, D., Langmead, B., and Salzberg, S.L. (2015). HISAT: a fast spliced aligner with low memory requirements. Nature Methods *12*, 357.

Kuleshov, M.V., Jones, M.R., Rouillard, A.D., Fernandez, N.F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S.L., Jagodnik, K.M., Lachmann, A*., et al.* (2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Research *44*, W90-W97.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nature methods *9*, 357.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The sequence alignment/map format and SAMtools. Bioinformatics *25*, 2078-2079.

Lieberman-Aiden, E., Van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., and Dorschner, M.O. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. science *326*, 289-293.

Lin, X., Han, M., Cheng, L., Chen, J., Zhang, Z., Shen, T., Wang, M., Wen, B., Ni, T., and Han, C. (2016). Expression dynamics, relationships, and transcriptional regulations of diverse transcripts in mouse spermatogenic cells. RNA Biology *13*, 1011-1024.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology *15*, 550.

McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. Nature biotechnology *28*, 495.

Pertea, M., Kim, D., Pertea, G.M., Leek, J.T., and Salzberg, S.L. (2016). Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. Nature Protocols *11*, 1650.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841-842.

Ramírez, F., Bhardwaj, V., Arrigoni, L., Lam, K.C., Grüning, B.A., Villaveces, J., Habermann, B., Akhtar, A., and Manke, T. (2018). High-resolution TADs reveal DNA sequences underlying genome organization in flies. Nature communications *9*, 189.

Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic acids research *44*, W160-W165.

Rao, S.S., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L.,

Machol, I., Omer, A.D., Lander, E.S., *et al.* (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell *159*, 1665-1680.
Rosenkranz, D. (2015). piRNA cluster database: a web resource for piRNA producing loci. Nucleic acids research *44*, D223-D230.
Schwarzer, W., Abdennur, N., Goloborodko, A., Pekowska, A., Fudenberg, G., Loe-Mie, Y., Fonseca, N.A., Huber, W., C, H.H., Mirny, L., *et al.* (2017). Two independent modes of chromatin organization revealed by cohesin removal. Nature *551*, 51-56.
Servant, N., Varoquaux, N., Lajoie, B.R., Viara, E., Chen, C.J., Vert, J.P., Heard, E., Dekker, J., and Barillot, E. (2015). HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. Genome biology *16*, 259.
Smagulova, F., Gregoretti, I.V., Brick, K., Khil, P., Camerini-Otero, R.D., and Petukhova, G.V. (2011). Genome-wide analysis reveals novel molecular features of mouse recombination hotspots. Nature *472*, 375-378.
Smith, C.M., Hayamizu, T.F., Finger, J.H., Bello, S.M., McCright, I.J., Xu, J., Baldarelli, R.M., Beal, J.S., Campbell, J., Corbani, L.E., *et al.* (2018). The mouse Gene Expression Database (GXD): 2019 update. Nucleic Acids Research *47*, D774-D779.
Smith, C.M., Hayamizu, T.F., Finger, J.H., Bello, S.M., McCright, I.J., Xu, J., Baldarelli, R.M., Beal, J.S., Campbell, J., Corbani, L.E., *et al.* (2019). The mouse Gene Expression Database (GXD): 2019 update. Nucleic Acids Res *47*, D774-D779.
Tarasov, A., Vilella, A.J., Cuppen, E., Nijman, I.J., and Prins, P. (2015). Sambamba: fast processing of NGS alignment formats. Bioinformatics *31*, 2032-2034.
Yamaguchi, K., Hada, M., Fukuda, Y., Inoue, E., Makino, Y., Katou, Y., Shirahige, K., and Okada, Y. (2018). Re-evaluating the Localization of Sperm-Retained Histones Revealed the Modification-Dependent Accumulation in Specific Genome Regions. Cell Rep *23*, 3920-3932.
Yoshida, K., Muratani, M., Araki, H., Miura, F., Suzuki, T., Dohmae, N., Katou, Y., Shirahige, K., Ito, T., and Ishii, S. (2018). Mapping of histone-binding sites in histone replacement-completed spermatozoa. Nat Commun *9*, 3885.
Zaharia, M., Bolosky, W.J., Curtis, K., Fox, A., Patterson, D., Shenker, S., Stoica, I., Karp, R.M., and Sittler, T. (2011). Faster and more accurate sequence alignment with SNAP. arXiv preprint arXiv:11115572.
Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., and Li, W. (2008). Model-based analysis of ChIP-Seq (MACS). Genome biology *9*, R137.