Three-dimensional Genome Structure Reveals Distinct Chromatin Signatures of Mouse Female Germline Stem Cells During Development

Geng G. Tian [1,9], Xinyan Zhao [2,9], Wenhai Xie [3], Xiaoyong Li [1], Changliang Hou [1], Yinjuan Wang[1], Lijuan Wang [3], Xiaodong Zhao [4], Hua Li [5], Jing Li [6, *], and Ji Wu [1, 2, 7, 8, *]

[1]Renji Hospital, Key Laboratory for the Genetics of Developmental & Neuropsychiatric Disorders (Ministry of Education), Bio-X Institutes, School of Medicine, Shanghai Jiao Tong University, Shanghai 200240, China

[2] Key Laboratory of Fertility Preservation and Maintenance of Ministry of Education, Ningxia Medical University, Yinchuan 750004, China

[3] School of Life Sciences, Shandong University of Technology, No.266 Xincun Road, Zibo, Shandong 255000, China.

[4]Shanghai Center for Systems Biomedicine, Shanghai Jiao Tong University, Shanghai 200240, China

[5]State Key laboratory for Oncogenes and Bio-ID Center, School of Biomedical Engineering, Shanghai Jiao Tong University, 800 Dongchuan Road, Shanghai 200240, China

[6]Department of Bioinformatics and Biostatistics, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, 800 Dongchuan Road, Shanghai 200240, China

[7] Shanghai Key Laboratory of Reproductive Medicine, Shanghai 200025, China

[8] Lead Contact

[9] These authors contributed equally

[*] Correspondence: jiwu@sjtu.edu.cn (W. J), jing.li@sjtu.edu.cn

SUMMARY

The three-dimensional configuration of the genome ensures cell-type-specific gene expression profiles by placing genes and regulatory elements in close spatial proximity. Here, we revealed the distinct features of the chromatin architecture in female germline stem cells (FGSCs) by *in situ* high-throughput chromosome conformation analysis. We also showed that the X chromosome structures were similar in spermatogonial stem cells and FGSCs. Using integrative analysis of the three-dimensional chromatin structure, we observed that the TADs were attenuated in germinal vesicle oocytes and disappeared in metaphase II oocytes during FGSC development. Finally, we identified conserved compartments belonging to the paternal/maternal genomes during early embryonic development, which were related to imprinted genes. These results will provide a valuable resource for studying and further our understanding of the fundamental characteristics of oogenesis and early embryo development.

INTRODUCTION

The chromatin architecture of germline stem cells (GSCs) carries the information necessary for the cells to exert their unique functions, and is thus an essential factor in

the transmission of the genome from generation to generation. GSCs can renew themselves and differentiate into gametes, including sperm and metaphase II (MII) oocytes [1-3]. During this process, spermatogonial stem cells (SSCs) differentiate into sperm by packaging the chromatin into a highly condensed configuration. Recent identified female GSCs (FGSCs) in postnatal ovaries were shown to differentiate into MII oocytes after transplantation into the ovaries of infertile mice [4-10], thus reshaping the idea that female mammals lose the ability to produce oocytes at birth [11,12]. Unlike other stem cells, GSCs can undergo meiosis to produce haploid gametes with chromatin remodeling. It is therefore necessary to characterize the chromatin structure of GSCs during their development to further our understanding of GSC biology.

High-throughput chromosome conformation (Hi-C) is a powerful technology for studying the genome-wide architecture, allowing the high-order chromatin structure to be displayed and revealing the chromatin organization in the nucleus [13]. The spatial organization of chromatin, as the structural and functional basis of the genome, can affect DNA localization, with important roles in gene transcription, the prevention of DNA damage, and ensuring DNA duplication and other biological processes [14,15]. Previous studies reported that the chromatin architecture changed dynamically during spermatogenesis, with dissolved and reappeared topologically associated domains (TADs) and A/B compartments [16,17]. However, the signature of the chromatin architecture during the development of FGSCs is unknown. Furthermore, the paternal and maternal chromatin architectures have been shown differences during early embryonic development[18,19], but the respective changes in high-order allelic genome

structure in early embryonic development remain to be explored.

In this study, we used *in situ* Hi-C technology to compared the chromatin organization of FGSCs with pluripotent stem cells (induced pluripotent stem cells, iPSCs), adult stem cells (ASCs) including SSCs and neural stem cells (NSCs), and somatic cells (mouse SIM embryonic fibroblast cells, STOs) to explore the chromosome structure character of FGSCs. Together with RNA sequencing (RNA-Seq) and chromatin immunoprecipitation sequencing (ChIP-Seq), we identified the distinct features of chromatin organization in FGSCs at three major levels: A/B compartments, TADs, and chromatin loops, and demonstrated that FGSCs were most similar to other ASCs, and largely different from iPSCs and STOs. We also identified similarities in X chromosomes between SSCs and FGSCs by principal component analysis of the X chromosome. Further analysis of Hi-C data during female germline cell development showed that TADs were attenuated but still present in germinal vesicle (GV) oocytes, disappeared in MII oocytes. Finally, we found conserved compartment regions in the paternal/maternal genomes in early embryonic development, related to imprinted genes. Together , these findings revealed the unique chromatin signature of FGSCs, and presented a whole landscape of high-order genome structure during development of female germline cells and early embryos.

RESULTS

Biological Characterization of FGSCs and Other ASCs

FGSCs were isolated and cultured from the ovaries of Ddx4-Cre;mT/mG neonatal

mice, as described previously [20]. After culture for at least 18 passages, the cells exhibited a characteristic morphology similar to that previously described for FGSCs [2,20]. The expression of female germline marker genes was determined by reverse-transcription polymerase chain reaction (RT-PCR). FGSCs after long-term culture expressed *Oct4*, *Fragilis*, *Mvh* (mouse vasa homologue, expressed exclusively in germ cells), *Stella*, *Gfrα1*, and *Dazl* genes. Furthermore, immunofluorescence analysis revealed that these cells also expressed MVH, confirming their identity as FGSCs (Figure 1A).

We isolated and cultured SSCs from the testes of 6-day-old DBA/2×CAG-EGFP F_1 mice, as described previously [21]. Long-term cultured SSCs (>20 passages) were assessed by RT-PCR and were shown to express male germline marker genes (*Etv5*, *Oct4*, *Plzf* (promyelocytic leukaemia zinc finger), *Gfrα1*, and *Mvh*). These results were confirmed by immunofluorescence and most of the cultured cells were also positive for PLZF expression, confirming their identity as SSCs (Figure 1B).

Isolated primary NSCs could self-proliferate and be cultured for 5–8 passages in NSC proliferation medium. Morphologically, cultured NSCs were spindle-shaped with a high nucleus-to-cytoplasm ratio, as reported previously [22]. The cultured NSCs were positive for several NSC-specific markers including *Nestin*, *Sox2*, *Pax6*, *Olig2*, and *Ascl1*, as determined by RT-PCR. Immunocytochemical staining confirmed that most of the cultured NSCs were positive for NESTIN and SOX2, as typical markers specific for NSCs. Following the removal of epidermal growth factor and basic fibroblast growth factor from the medium, NSCs spontaneously differentiated into

neurons and astrocytes, characterized by prominent dendrites with long axons and by extensive cytoplasm with thick processes, respectively. The differentiation potential of cultured NSCs was confirmed by immunochemical staining of the neural- and astrocyte-specific markers TUJ-1 (β3 Tubulin) and GFAP (glial fibrillary acidic protein). These results confirmed the identity of the cultured cells as NSCs (Figure 1C). The morphology of STOs is shown in Figure 1D.

Global Chromosome Organization Map in FGSCs

To study the signature of the chromatin architecture in FGSCs, we performed *in situ* Hi-C [23] with two biological replicates of FGSCs and other cells (SSCs, NSCs, iPSCs and STOs) generating approximately 400 million reads for each replicate. After filtering artificial reads and normalization, we obtained a total of over 2 billion valid Hi-C reads, including an average of 1 billion long-range (>20 kb) intra-chromosomal *cis* contacts and 400 million inter-chromosomal *trans* contacts (Table S1). We confirmed the high reproducibility of the Hi-C data (Figure S1) and combined the two biological replicates into a single set of merged Hi-C data per cell type, to reach a maximum resolution of 20 kb.

Analysis of the Hi-C data showed that the high-order chromatin organization of the whole genome in FGSCs differed from that in the other cells (Figure S2). An overview of the intra-chromosomal contact heat maps revealed that FGSCs displayed a distinct chromatin organization (Figure 2A). We further examined the characteristics of the chromatin organization by analyzing the patterns of compartment status and

TADs in the autosomes across cells, avoiding sex chromosome effects. The compartment status was classified as active (A) or inactive (B) (Table S2). FGSCs were more similar to SSCs and NSCs in terms of A/B compartments, compared with iPSCs and STOs (Figure 2B), suggesting that FGSCs were ASCs. The patterns of TADs and directional indexes (DI) were almost the same for these cells (Figure 2B). We counted the numbers of compartments and TADs in the cells and showed that FGSCs had the lowest number of TADs, and the number of compartments was similar to that in SSCs (Figure 2C and Table S3). We also calculated the average intra-chromosomal contact probability of cells and found that the chromatin interaction frequency decreased monotonically from 10^5 to 10^8 bp for FGSCs and the other cells (Figure 2D). The contact probability curves were similar in five cell types from 10^5 to $10^{6.5}$ bp, but changes occurred at long-distance genome, as reported previously [18].

FGSCs Exhibited Distinct Compartment Status

Systematic analysis of the compartment status in FGSCs showed that genes had higher expression levels in compartment A than in compartment B (Figure S3A), indicating that compartment status was correlated with gene expression. Combining with the ChIP-Seq data analysis, we observed that H3K27ac and H3K4me3 were highly correlated with compartment status (Figure S3B). The genome browser showed that H3K27ac was highly enriched in compartment A but not in compartment B (Figure S3C). By k-means clustering of the compartment status in FGSCs

compared with other cells, we obtained activate and inactive compartments of FGSCs (Figure 3A). Furthermore, switching compartments of FGSCs accounted for a high proportion (about 50%) of the total number in the genome compared with iPSCs and STOs, but a smaller proportion (about 30%–40%) compared with SSCs and NSCs (Figure S4A). These results suggested that FGSCs had a unique A/B compartment status that was more similar to other ASCs than to iPSCs or STOs. Additional RNA-Seq data revealed that the genes located in the switching compartment tended to be significantly differentially expressed compared with the stable compartments (Figure 3B). This was consistent with ChIP-Seq signal results, which showed dramatic differences of H3K27ac and H3K4me3 enrichment in the switching compartment compared with the stable compartment (Figure S4B). By calculating the covariation between gene expression and compartment status, we identified a subset of 1206 genes that were highly correlated with compartment status (Table S4). Gene Ontology (GO) analysis showed that these genes were particularly associated with stem cell population maintenance and cell proliferation (Figure S4C). Among these, *Coprs*, as an ASC marker located in the A compartment of FGSCs, SSCs, and NSCs, showed higher expression than in iPSCs in which it was located in the B compartment, consistent with a previous report [24] (Figure 3C). In addition, *Nanog* as a pluripotent stem cell marker, showed higher expression in iPSCs, being located in the A compartment in iPSCs while switching to the B compartment in other types of cells (Figure 3C). These findings suggested that FGSCs had a unique compartment status characteristic of ASCs, which could work together with histone modification to

177 regulate gene expression to determine their features.

178 We then identified the TADs in FGSCs using DI (Table S3). Well-defined TADs
179 were conserved in ~90% of the genome across cell types (Figure 3D). We classified
180 the TADs into five types: stable, merge, split, rearrangement, and unique. Most of the
181 TADs belonged to the stable type (Figure 3E), suggesting that TAD structure was
182 highly stable across all five types of cells, in accordance with a previous report [25].
183 However, the absolute DI showed that the strength of the TADs differed between
184 FGSCs and the other cells (Figure S5A), suggesting that, although TAD domains were
185 stable, FGSCs had a distinct frequency of intra-TAD interactions. We further explored
186 the relationship between TADs and gene expression, and observed that gene
187 expression was higher in TAD boundaries than in TADs in FGSCs and other cells
188 (Figure 3F), illustrating that genes were more activated in these TAD boundaries
189 (Figure S5B). We classified the boundaries into cell-type-specific and common
190 boundaries, and identified 417, 369, 286, 263, and 48 cell-type-specific boundaries
191 and 833 common boundaries, suggesting that most boundaries were stable across cell
192 types. Further study of the relationship between cell-type-specific boundaries and
193 gene expression indicated that gene expression changed significantly between specific
194 and common boundaries (Figure S5C). These results revealed that FGSCs had stable
195 TADs, but that gene expression was activated in the TAD boundary.

196

197 **Cell-type-specific Chromatin Loops Revealed FGSC Signature**

198 We systematically analyzed the chromatin loops and identified 4832, 1906, 7004,

3060, and 6951 chromatin loops in FGSCs, SSCs, NSCs, iPSCs and STOs, respectively. Using a Venn diagram, we observed that only a few (n=177) chromatin loops were shared across all cell types (cell-type shared loops) (Figure 4A), suggesting that most chromatin loops were cell-type-specific and that FGSCs had distinct chromatin loops. Furthermore, genes located in these cell-type-specific loops were highly enriched in cell-type-related GO categories (Figure 4B). For instance, *Sox17*, as a transcription factor involved in embryonic development [26], formed chromatin loops in iPSCs but gradually disappeared in FGSCs, SSCs, NSCs, and STOs, which was also most highly expressed in iPSCs (Figure S6A), suggesting that its expression could be regulated by *cis* regulation of the chromatin loops. Based on previous results, both A/B compartment status and TADs could affect gene expression [27,28]. We therefore investigated if the formation of cell-type-specific chromatin loops was related to compartment A/B status or to TADs, and found that 60% of cell-type-specific loops were commonly localized in the switching compartments (Figure S6B), while about 30% were localized in the stable TAD type (Figure S6C). This suggested that the chromatin loops relied more on compartment switching to regulate gene expression, and were not dependent on TAD type to exert their function.

We further examined the specific chromatin loops in FGSCs as GSCs by comparing them with the loops in SSCs using ChIP-Seq data. K-means clustering of ChIP-Seq signals identified four major classes on active promoter (H3K4me3) and enhancer (H3K27ac) sites (Figure 4C). Cluster 4 of active promoters and most of the active enhancers were obviously different between FGSCs and SSCs (Figure 4C).

Meanwhile, GO enrichment analysis further revealed that specific loops with different active promoters and enhancers were enriched in reproductive processes such as sex differentiation, sexual reproduction, and reproductive development (Fisher's exact test, $P < 0.05$; Figure S7). We divided these loops into two subsets: FGSCs lost and FGSCs gained (Figure 4D). *Dmrt1*, as a conserved transcriptional regulator in male germ cells required for the maintenance and replenishment of SSCs [29], was looped in SSCs but not in FGSCs. *Hif1an*, as a heterodimeric transcription factor related to interface with stem cell signaling pathways [30], gained loops in FGSCs and could be related to female germline cell development (Figure 4D). Overall, these findings indicated that FGSCs possessed cell-type-specific chromatin loops that provided the spatial space for histone modification to regulate gene expression.

X-Chromosome Structure Was Similar Between SSCs and FGSCs

One female X chromosomes is randomly inactivated during mammalian development to ensure matched dosages in males and females [31]. To dissect the high-order organization of the X chromosome in SSCs and FGSCs, we performed Pearson's correlation analysis of the Hi-C matrix in the X chromosome. The results showed that SSCs and FGSCs had a strong correlation (Figure 5A). Furthermore, upon extracting the eigenvectors of the whole chromosome interactions and using the first principal component (PC1) score to compare the structure of the X chromosome in SSCs and FGSCs, we found that the X chromosome was more similar (correlation = 0.87) than the autosomes (mean correlation = 0.21) between SSCs and FGSCs (Figure 5B and

Figure S8). By analyzing the PC1 score of the X chromosome in FGSCs, SSCs, and female embryonic stem cells (fESCs), considering that fESCs have two activated X chromosomes (Xa), we observed that FGSCs were more highly correlated with SSCs than with fESCs in the X chromosome (Figure 5C). This suggested that the X chromosome is similar between SSCs and FGSCs, probably due to one of the X chromosomes being inactivated (Xi) in FGSCs [32,33].

To identify the difference between the active X chromosome and the inactive one, we deconvoluted the respective Hi-C data of these chromosomes from FGSCs (Figure 5D). As expected, we visualized that the X chromosome was separated by a region containing the DXZ4 macrosatellite (which reportedly plays a crucial role in shaping Xi-chromosome structure) into two parts in FGSCs [34], as well as in deconvoluted Xi, whereas this was not the case in SSCs (Figure 5D). This demonstrated that one of the X chromosomes in FGSCs was inactivated. Deconvoluted Xi displayed that the long-range contacts were attenuated for intra-TADs and inter-TADs (Figure 5D), consistent with a previous report [32,33]. We next investigated the structure of the region containing Xist, which is a key factor for inactivation of the X chromosome. Notably, FGSCs were similar to SSCs in the Xist region, while in deconvolved Xi it was shown that the Xist region lost most long-range contacts and retained a sub-TAD-like structure (Figure 5E), in support of a previous report [32]. Taken together, our data suggested that one of the X chromosomes was inactivated in FGSCs, to maintain relative consistency with SSCs to balance the gene expression between males and females.

TADs Were Attenuated and then Disappeared During Female Germline Cell Development

FGSCs are derived from primordial germ cells and undergo meiosis into GV oocytes, and then to MII oocytes. To explore the dynamic changes of TADs during female germline cell development, we applied Hi-C to obtain the data of whole genome of chromatin architecture in mouse GV oocytes (Figure S9). After analysis of Hi-C data for MII oocytes, zygotes, two-cell embryos, and eight-cell embryos, we observed that TADs were attenuated in GV oocytes, disappeared in MII oocytes, and recovered in two-cell embryos (Figure 6A). A snapshot of TAD signals also showed that the TAD strength was weakened during female germline cell development and reprogramed at early embryonic development (Figure 6B), in contrast to changes in TADs during spermatogenesis [35].

We further examined the changes in TADs during female germline cell development by calculating the DI value, which reflected the degree of interactions of a given bin in upstream or downstream regions and was associated with calling TADs. Expectedly, the DI value was significantly reduced during female germline cell development and reestablished in early embryonic development (Figure 6C). Meanwhile, the insulation score showed that FGSCs had the strongest TAD boundaries, while these were decreased in GV oocytes and weakest in MII oocytes, and gradually increased in early embryonic development (Figure 6D). We further calculated the proportion of *cis*-short interactions (<2 Mb) and *cis*-long interactions

(>2 Mb) versus total *cis*-interactions. The relative proportions of *cis*-short interactions in FGSCs was similar to GV oocytes, lowest in MII oocytes, highest in zygotes, and then reduced during early embryonic development, with similar interactions in eight-cell embryos and FGSCs (Figure 6E). Overall, these results demonstrated that TADs were attenuated and then disappeared during female germ cell development, and reestablished during early embryonic development (Figure 6F), thus revealing the pattern of TADs throughout female germ cell and early embryonic development.

Identification of Conserved Allelic Chromatin Structures

Previous findings have reported the chromatin structure of the maternal genome is different with the paternal genome during early embryonic development [18,19]. We aimed to investigate how the conserved structures in these respective genomes affected their genome organization. Early embryonic Hi-C data were analyzed with single nucleotide polymorphisms between two mouse strains to track the maternal and paternal genomes. The correlation of compartment status according to the PC1 score, showed that the paternal genome was clustered in pachytene spermatocytes (PACs), sperm, paternal zygotes, and eight-cell embryos, while the maternal genome was clustered within MII oocytes, maternal zygotes, and eight-cell embryos (Figure 7A), indicating that the allelic chromatin structure was conserved in early embryonic development. FGSCs and SSCs, as early-stage germ cells, were clustered together (Figure 7A). Interestingly, the paternal genome clustered with the maternal genome at the two-cell embryo stage, but was completely separate at the eight-cell stage,

suggesting that the two-cell stage plays an important role in allelic chromatin structure development, possibly because of the reestablishment of TADs at this stage. Furthermore, we observed some regions of compartments in allelic genome were conserved during the early embryonic development (Figure 7B). We then systematically identified the conserved A/B compartment regions by calculating Pearson's correlation for the whole genome, with a sliding window of 2 Mb. The paternal genome had more conserved regions than the maternal genome (Figure 7C), indicating that the paternal genome was more conserved. Comparing the conserved regions, some allelic-specific conserved regions were identified by Venn diagrams (Figure 7D). By enriching for imprinted genes with Fisher's exact test ($P < 0.05$), we found that the allelic-specific regions were significantly related to imprinted genes (Table S5) such as *Igf2r*, *Dlk1*, and *Dio3*, which were reported to highly express in maternal or paternal [36,37]. The results suggested that those imprinted genes were affected by the allelic-specific conserved region, which could regulate paternal or maternal development.

DISCUSSION

Stem cells, including pluripotent stem cells and ASCs, have important implications in basic biology and regenerative medicine. However, regenerative medicine requires stem cells to be transplanted safely, with a particular focus on avoiding the development of cancer. Although gene mutations have been reported to be responsible for many diseases, including cancer [38,39], recent studies have revealed that diseases

331 such as cancer can also be caused by disruption of chromatin organization. The
332 chromatin architecture thus plays crucial roles in preventing DNA damage, in gene
333 mutation, and in ensuring appropriate gene transcription, DNA duplication, and
334 developmental processes [40-42]. It is therefore essential that stem cells are characterized
335 or identified in terms of their chromatin organization before their use in a clinical
336 context. FGSCs not only have the abilities of self-renewal and differentiation, but are
337 also responsible for passing on genetic information to the next generation. The
338 stability of that genetic information is affected by the high-order genome organization
339 [43].

340 To identify the chromosome structure character of FGSCs, we compared FGSCs
341 with pluripotent stem cells (iPSCs), ASCs (SSCs and NSCs) and somatic cells (STOs)
342 by Hi-C technology. The results revealed that FGSCs had a distinct high-order
343 genome structure in terms of the A/B compartment status, chromatin loops, and TADs.
344 For further characterization, we identified FGSCs specific activated and repressed
345 compartment regions, and obtained partially genes highly related with the switch of
346 FGSCs compartment status. These genes were related to stem cell maintenance and
347 differentiation pathways, strongly supporting the role of FGSCs as stem cells, with
348 some shared characteristics with SSCs and NSCs, and belonging to ASCs. Moreover,
349 FGSC-specific loops, which included active promoters and enhancers, could be
350 significantly enriched in reproductive-related pathways. Among these, *hif1an*, which
351 is related to Notch signaling, could be a potential marker for female germline cells.
352 Our findings indicated that FGSCs belong to ASCs in the high-order organization, and

353 further confirmed the existence of FGSCs, consistent with the previous reports for
354 cellular and molecular characteristic [4-7,44,45].

355 Previous studies indicated that both FGSCs and SSCs have their own unique
356 epigenetic signatures [44,46]. We analyzed and compared the 3D genomic architectures
357 of FGSCs and SSCs, and revealed that GSCs had their own unique high-order
358 chromatin organization, especially in terms of A/B compartments and chromatin loops.
359 Although FGSCs have one more X chromosome than SSCs, the architecture of the X
360 chromosome in FGCSs was more similar to SSCs than to autosomes, suggesting that
361 FGSCs maintain a balance of gene expression with SSCs by inactivating one X
362 chromosome. Furthermore, analysis of the X-chromosome matrix of FGSCs indicated
363 that it was separated into two domains by a region containing *Dxz4*, which has been
364 reported to be an essential regulator of X chromosome inactivation [32]. Meanwhile,
365 deconvoluted Xi data showed that *Xist* was also located in a region showing moderate
366 interactions with a TAD-like structure, consistent with previous findings [32]. These
367 results indicate that these differences could reflect Xi, supporting inactivation of one
368 X chromosome in FGSCs [47].

369 Interestingly, recent studies reported that the chromatin architecture changed
370 dynamically during spermatogenesis, with dissolved and then reappeared TADs and
371 compartments [16,17]. On the contrary, TADs were attenuated and then disappeared in
372 oogenesis. During early embryonic development, TADs recovered and started to
373 appear in two-cell embryos, suggesting that this stage could play an essential role in
374 genome organization, including paternal/maternal chromatin reprogramming. We

therefore further analyzed paternal and maternal chromatin structures during early embryonic development, and found that the paternal genome clustered with the maternal genome at the two-cell stage, but was completely separate at the eight-cell stage, consistent with the above results.

In conclusion, we present a comprehensive overview of the chromatin organization of FGSCs to create a rich resource of high-resolution genome-wide maps. Our findings revealed that the chromatin architecture of FGSCs included unique features, especially in terms of compartment status and chromatin loops, which may contribute to their cell-type-specific gene regulation. These data will provide a valuable resource for future studies of the features of chromatin organization in mammalian stem cells, with important implications for their role in medical research and their potential and actual clinical applications.

EXPERIMENTAL PROCEDURES

Additional information and details regarding this work may be found in the Supplemental Experimental Procedures.

ACCESSION NUMBERS

The accession number for the expression and sequencing data reported in this paper is GEO: GSE126014 and GEO: GSE137771.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, eight figures and four tables and can be found with this article online.

AUTHOR CONTRIBUTIONS

T.G.G performed the Hi-C experiments, analyzed the data and wrote the manuscript. Z.X carried out ChIP-Seq experiments. X.W and W.L performed the NSC culture and identification. L.X undertook the FGSC culture and identification. W.Y performed the SSC culture and identification. L.H finished RNA-Seq experiment. H.C did the Hi-C experiment of GV stage oocyte. W.J and L.J supervised the experiment work and devised this study.

ACKNOWLEDGMENTS

We thank Dr. Kang's lab (Tongji University) for providing the iPSCs cell line. This work was supported by National Basic Research Program of China (2018YFC1003501, 2017YFA0504201), National Nature Science Foundation of China (81720108017, 81501316), the National Major Scientific Instruments and Equipment Development Project, National Nature Science Foundation of China (61827814).

REFERENCES

1 Brinster, R. L. & Zimmermann, J. W. Spermatogenesis following male germ-cell transplantation. Proc Natl Acad Sci U S A 91, 11298-11302 (1994).

2 Zou, K. et al. Production of offspring from a germline stem cell line derived from neonatal ovaries. Nature cell biology 11, 631-636, doi:10.1038/ncb1869 (2009).

3 Wu, C. et al. Tracing and Characterizing the Development of Transplanted Female Germline Stem Cells In Vivo. Molecular therapy : the journal of the American Society of Gene Therapy 25, 1408-1419, doi:10.1016/j.ymthe.2017.04.019 (2017).

4 Zou, K. et al. Production of offspring from a germline stem cell line derived from neonatal ovaries. Nature Cell Biology 11, 631-U424, doi:10.1038/ncb1869 (2009).

5 White, Y. A. R. et al. Oocyte formation by mitotically active germ cells purified from ovaries of reproductive-age women. Nat Med 18, 413-U176, doi:10.1038/nm.2669 (2012).

6 Ding, X. B. et al. Human GV oocytes generated by mitotically active germ cells obtained from follicular aspirates. Sci Rep-Uk 6, doi:ARTN 2821810.1038/srep28218 (2016).

7 Wu, C. Q. et al. Tracing and Characterizing the Development of Transplanted Female Germline Stem Cells In Vivo. Mol Ther 25, 1408-1419, doi:10.1016/j.ymthe.2017.04.019 (2017).

8 Zhou, L. et al. Production of fat-1 transgenic rats using a post-natal female germline stem cell line. Mol Hum Reprod 20, 271-281, doi:10.1093/molehr/gat081 (2014).

9 Lu, Y. Q., He, X. C. & Zheng, P. Decrease in expression of maternal effect gene Mater is associated with maternal ageing in mice. Mol Hum Reprod 22, 252-260, doi:10.1093/molehr/gaw001 (2016).

10 Liu, J. et al. Isolation and characterization of string-forming female germline stem cells from ovaries of neonatal mice. Journal of Biological Chemistry 292, 16003-16013, doi:10.1074/jbc.M117.799403 (2017).

11 Zuckerman, S. J. R. P. i. H. R. The number of oocytes in the mature ovary. 6, 63-109 (1951).

12 Borum, K. J. E. c. r. Oogenesis in the mouse: a study of the meiotic prophase. 24, 495-507 (1961).

13 Lieberman-Aiden, E. et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 326, 289-293, doi:10.1126/science.1181369 (2009).

14 Smallwood, A. & Ren, B. Genome organization and long-range regulation of gene expression by enhancers. Current opinion in cell biology 25, 387-394, doi:10.1016/j.ceb.2013.02.005 (2013).

15 Gorkin, D. U., Leung, D. & Ren, B. The 3D Genome in Transcriptional Regulation and Pluripotency. Cell Stem Cell 14, 762-775, doi:10.1016/j.stem.2014.05.017 (2014).

16 Wang, Y. et al. Reprogramming of Meiotic Chromatin Architecture during Spermatogenesis. Mol Cell 73, 547-561 e546, doi:10.1016/j.molcel.2018.11.019 (2019).

17 Vara, C. et al. Three-Dimensional Genomic Structure and Cohesin Occupancy Correlate with Transcriptional Activity during Spermatogenesis. Cell Reports 28, 352-+, doi:10.1016/j.celrep.2019.06.037 (2019).

18 Du, Z. et al. Allelic reprogramming of 3D chromatin architecture during early mammalian development. Nature 547, 232-235, doi:10.1038/nature23263 (2017).

19 Ke, Y. et al. 3D Chromatin Structures of Mature Gametes and Structural Reprogramming during Mammalian Embryogenesis. Cell 170, 367-381 e320, doi:10.1016/j.cell.2017.06.029 (2017).

20 Zhang, C. & Wu, J. Production of offspring from a germline stem cell line derived from prepubertal ovaries of germline reporter mice. Molecular human reproduction 22, 457-464, doi:10.1093/molehr/gaw030 (2016).

21 Gong, X., Xie, H., Li, X., Wu, J. & Lin, Y. Bisphenol A induced apoptosis and transcriptome differences of spermatogonial stem cells in vitro. Acta biochimica et biophysica Sinica 49, 780-791, doi:10.1093/abbs/gmx075 (2017).

22 Weixiang, G., Patzlaff, N. E., Jobe, E. M. & Xinyu, Z. Isolation of multipotent neural stem or progenitor cells from both the dentate gyrus and subventricular zone of a single adult mouse. Nature protocols 7, 2005 (2012).

23 Rao, S. S. P. et al. A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. Cell 159, 1665-1680, doi:10.1016/j.cell.2014.11.021 (2014).

485 24 Xie, W., Wang, H. & Wu, J. Similar morphological and molecular signatures
486 shared by female and male germline stem cells. Sci Rep 4, 5580,
487 doi:10.1038/srep05580 (2014).
488 25 Dixon, J. R. et al. Topological domains in mammalian genomes identified by
489 analysis of chromatin interactions. Nature 485, 376-380,
490 doi:10.1038/nature11082 (2012).
491 26 Kim, I., Saunders, T. L. & Morrison, S. J. Sox17 dependence distinguishes the
492 transcriptional regulation of fetal from adult hematopoietic stem cells. Cell
493 130, 470-483, doi:10.1016/j.cell.2007.06.011 (2007).
494 27 Dixon, J. R. et al. Chromatin architecture reorganization during stem cell
495 differentiation. Nature 518, 331-336, doi:10.1038/nature14222 (2015).
496 28 Bonev, B. et al. Multiscale 3D Genome Rewiring during Mouse Neural
497 Development. Cell 171, 557-572 e524, doi:10.1016/j.cell.2017.09.043 (2017).
498 29 Zhang, T., Oatley, J., Bardwell, V. J. & Zarkower, D. DMRT1 Is Required for
499 Mouse Spermatogonial Stem Cell Maintenance and Replenishment. Plos
500 Genet 12, doi:ARTN e100629310.1371/journal.pgen.1006293 (2016).
501 30 Simon, M. C. & Keith, B. The role of oxygen availability in embryonic
502 development and stem cell function. Nat Rev Mol Cell Bio 9, 285-296,
503 doi:10.1038/nrm2354 (2008).
504 31 Huynh, K. D. & Lee, J. T. X-chromosome inactivation: a hypothesis linking
505 ontogeny and phylogeny. Nature reviews. Genetics 6, 410-418,
506 doi:10.1038/nrg1604 (2005).

32 Giorgetti, L. et al. Structural organization of the inactive X chromosome in the mouse. Nature 535, 575-579, doi:10.1038/nature18589 (2016).

33 Nora, E. P. et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature 485, 381-385, doi:10.1038/nature11049 (2012).

34 Deng, X. et al. Bipartite structure of the inactive mouse X chromosome. Genome biology 16, 152, doi:10.1186/s13059-015-0728-8 (2015).

35 Wang, Y. et al. Reprogramming of Meiotic Chromatin Architecture during Spermatogenesis. Molecular Cell 73, 547-+, doi:10.1016/j.molcel.2018.11.019 (2019).

36 Da Rocha, S. T., Edwards, C. A., Ito, M., Ogata, T. & Ferguson-Smith, A. C. Genomic imprinting at the mammalian Dlk1-Dio3 domain. Trends in Genetics 24, 306-316, doi:10.1016/j.tig.2008.03.011 (2008).

37 Birger, Y., Shemer, R., Perk, J. & Razin, A. The imprinting box of the mouse Igf2r gene. Nature 397, 84-88, doi:Doi 10.1038/16291 (1999).

38 Vogelstein, B. et al. Cancer genome landscapes. Science 339, 1546-1558, doi:10.1126/science.1235122 (2013).

39 Cooper, D. N. et al. Genes, Mutations, and Human Inherited Disease at the Dawn of the Age of Personalized Genomics. Hum Mutat 31, 631-655, doi:10.1002/humu.21260 (2010).

40 Lupianez, D. G. et al. Disruptions of Topological Chromatin Domains Cause Pathogenic Rewiring of Gene-Enhancer Interactions. Cell 161, 1012-1025, doi:10.1016/j.cell.2015.04.004 (2015).

529 41 Hnisz, D. et al. Activation of proto-oncogenes by disruption of chromosome
530 neighborhoods. Science 351, 1454-1458, doi:10.1126/science.aad9024 (2016).
531 42 Franke, M. et al. Formation of new chromatin domains determines
532 pathogenicity of genomic duplications. Nature 538, 265-269,
533 doi:10.1038/nature19800 (2016).
534 43 Battulin, N. et al. Comparison of the three-dimensional organization of sperm
535 and fibroblast genomes using the Hi-C approach. Genome Biol 16, 77,
536 doi:10.1186/s13059-015-0642-0 (2015).
537 44 Zhang, X. L. et al. Integrative epigenomic analysis reveals unique epigenetic
538 signatures involved in unipotency of mouse female germline stem cells.
539 Genome Biol 17, 162, doi:10.1186/s13059-016-1023-z (2016).
540 45 Ma, B. et al. Molecular characteristics of early-stage female germ cells
541 revealed by RNA sequencing of low-input cells and analysis of genome-wide
542 DNA methylation. DNA Res 26, 105-117, doi:10.1093/dnares/dsy042 (2019).
543 46 Hammoud, S. S. et al. Chromatin and transcription transitions of mammalian
544 adult germline stem cells and spermatogenesis. Cell Stem Cell 15, 239-253,
545 doi:10.1016/j.stem.2014.04.006 (2014).
546 47 Li, X. et al. Dosage compensation in the process of inactivation/reactivation
547 during both germ cell development and early embryogenesis in mouse. Sci
548 Rep 7, 3729, doi:10.1038/s41598-017-03829-z (2017).

549

550

551 **Figure Legends**

552 **Figure 1. Morphology and biological characteristics of FGSCs and other ASCs**

553 (A) Morphology and biological characteristics of FGSCs. (I) Representative
554 morphology of cultured FGSCs (>18 passages) from Ddx4-Cre;mT/mG mice. (II)
555 Representative field under UV light for FGSCs with GFP expression. (III) Merged
556 images from I and II. Bar=25 μm. (IV) RT-PCR analysis of female germline markers.
557 Marker, 100-bp DNA markers. (V–VIII) Immunofluorescence of FGSCs for GFP (V),
558 MVH (VI), DAPI (VII), and merged (VIII). Bar=10 μm.

559 (B) Morphology and biological characteristics of SSCs. (I) Representative view of
560 cultured SSCs (>20 passages) from CAG-EGFP mice. (II) Representative image
561 under UV light for SSCs with GFP expression. (III) Merged images from I and II.
562 Bar=25 μm. (IV) RT-PCR analysis of male germline markers. Marker, 100-bp DNA
563 markers. (V–VIII) Immunofluorescence of SSCs for GFP (V), PLZF (VI), DAPI (VII),
564 and merged (VIII). Bar=50 μm.

565 (C) Morphology and biological characteristics of NSCs. (I) Representative
566 morphology of cultured NSCs. (II) Representative view of differentiated NSCs *in*
567 *vitro*. Bar=10 μm. (III) RT-PCR analysis of NSC markers. Marker, 100-bp DNA
568 markers. (V–VIII) Immunofluorescence of NSCs for Nestin (V), SOX2 (VI), DAPI
569 (VII), and merged (VIII). (IX–XII) Immunofluorescence of differentiated NCs for
570 TUJ-1 (IX), GFAP (X), DAPI (XI), and merged (XII). Bar=10 μm.

571 (D) Representative morphology of cultured STOs. Bar=10 μm.

572

573 **Figure 2. Overall chromosome structure in FGSCs**

574 (A) Contact matrices from chromosome 19 in FGSCs, SSCs, NSCs, iPSCs and STOs.
575 (B) First principal component (PC1) value, normalized Hi-C interaction heat maps,
576 and directional indexes (DIs) in FGSCs and other cells at 20-kb resolution. PC1 value
577 was used to indicate A/B compartment status, where a positive PC1 value represented
578 the A compartment (blue) and a negative one represented the B compartment
579 (yellow).
580 (C) Numbers of identified A/B compartments and TADs in FGSCs and other cells.
581 (D) The average contact probability across the genome decreased as a function of
582 genomic distance.
583
584 **Figure 3. FGSCs exhibited specific compartment status**
585 (A) K-means clustering (k=2) of PC1 values of the genome that change A/B
586 compartment status in FGSCs.
587 (B) Genes that changed compartment status (A to B or B to A) or that remained the
588 same (stable) compared with FGSCs (P value by Wilcoxon's test).
589 (C) IGV snapshots for *Coprs* and *Nanog* showed concordance between their
590 expression and PC1 values.
591 (D) Percentages of TADs and TAD boundaries in the genome.
592 (E) Numbers of TAD types, including stable, merged, split, unique, and rearrangement,
593 in stem cells.
594 (F) Comparison of gene expression between TADs and TAD boundaries.
595 Genome-wide average distribution of RNA-Seq, H3K27ac, H3K4me3, and

596 H3K27me3 reads around the domain boundaries in FGSCs.

597

598 **Figure 4. Cell-type-specific chromatin loops revealed FGSCs signature**

599 (A) Venn diagram showing that most chromatin loops were cell-type-specific in five

600 types of cells.

601 (B) GO enrichment of cell-type-specific chromatin loops.

602 (C) K-means clustering of H3K4me3 and H3K27ac ChIP-Seq data at promoter and

603 enhancer regions between FGSCs and SSCs.

604 (D) Chromatin loops gained or lost in FGSCs compared with SSCs.

605

606 **Figure 5. X-Chromosome structure conformation between SSCs and FGSCs**

607 (A) Pearson's correlation heat map showing similarity between FGSCs and SSCs.

608 (B) PC1 values in the X chromosome in FGSCs and SSCs.

609 (C) Spearman's correlation of PC1 values in the X chromosome in FGSCs, SSCs, and

610 fESCs.

611 (D) Deconvolution of Xi signal from Hi-C data obtained in FGSCs (Xa+Xi) by

612 subtracting the Xa contribution estimated from SSCs. A small region containing *Dxz4*

613 showed two domains in FGSCs and deconvoluted Xi, but not in SSCs (blue pixels

614 represent negative values, red pixels represent positive values).

615 (E) Normalized chromatin interaction maps around *Xist* at 20-kb resolution. Plots

616 show TAD signal (insulation score).

617

618 **Figure 6. TADs were attenuated and then disappeared during female germline**
619 **cell development**

620 (A) Normalized Hi-C interaction frequencies during female germline cell
621 development displayed as a heatmap.

622 (B) TAD signals at chromosome 17:33–37 Mb during female germline cell
623 development.

624 (C) Boxplot of absolute DI values during female germline cell development. P value
625 calculated by Kruskal-Wallis test.

626 (D) Average insulation scores (IS) of different stages in female germline cell
627 development at TADs (defined in FGSCs) and nearby regions.

628 (E) Relative proportions of *cis* interactions at different genome distances versus total
629 paired loci.

630 (F) Graphical model for dynamic changes in TADs during female germline cell
631 development.

632

633 **Figure 7. Identification of conserved allelic chromatin structures**

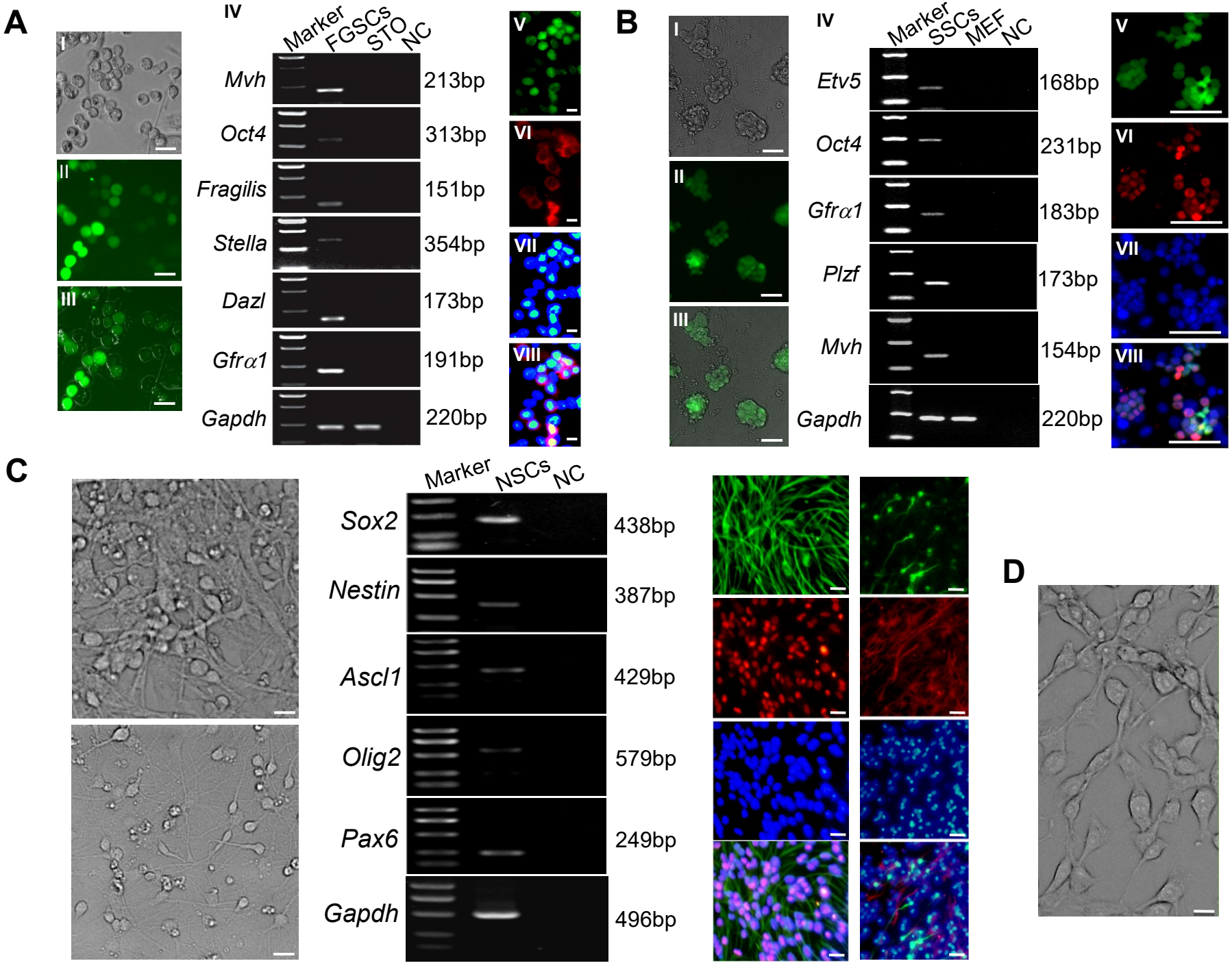
634 (A) Hierarchical clustering of PC1 values based on maternal (black) and paternal (red)
635 genome architectures. Reference shown in green.

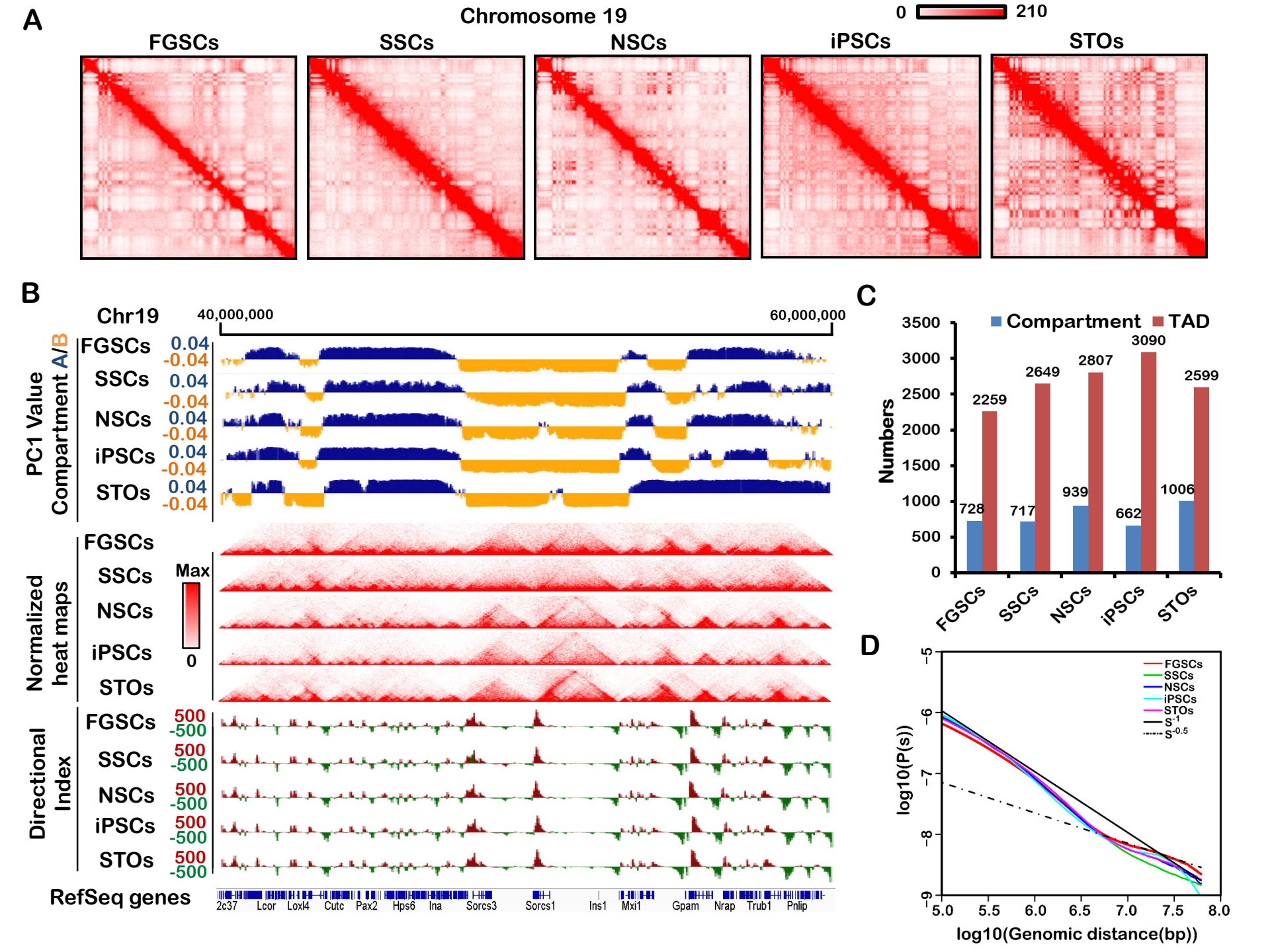
636 (B) Example of conserved compartment region at chromosome 19:34–60 Mb during
637 early embryonic development.

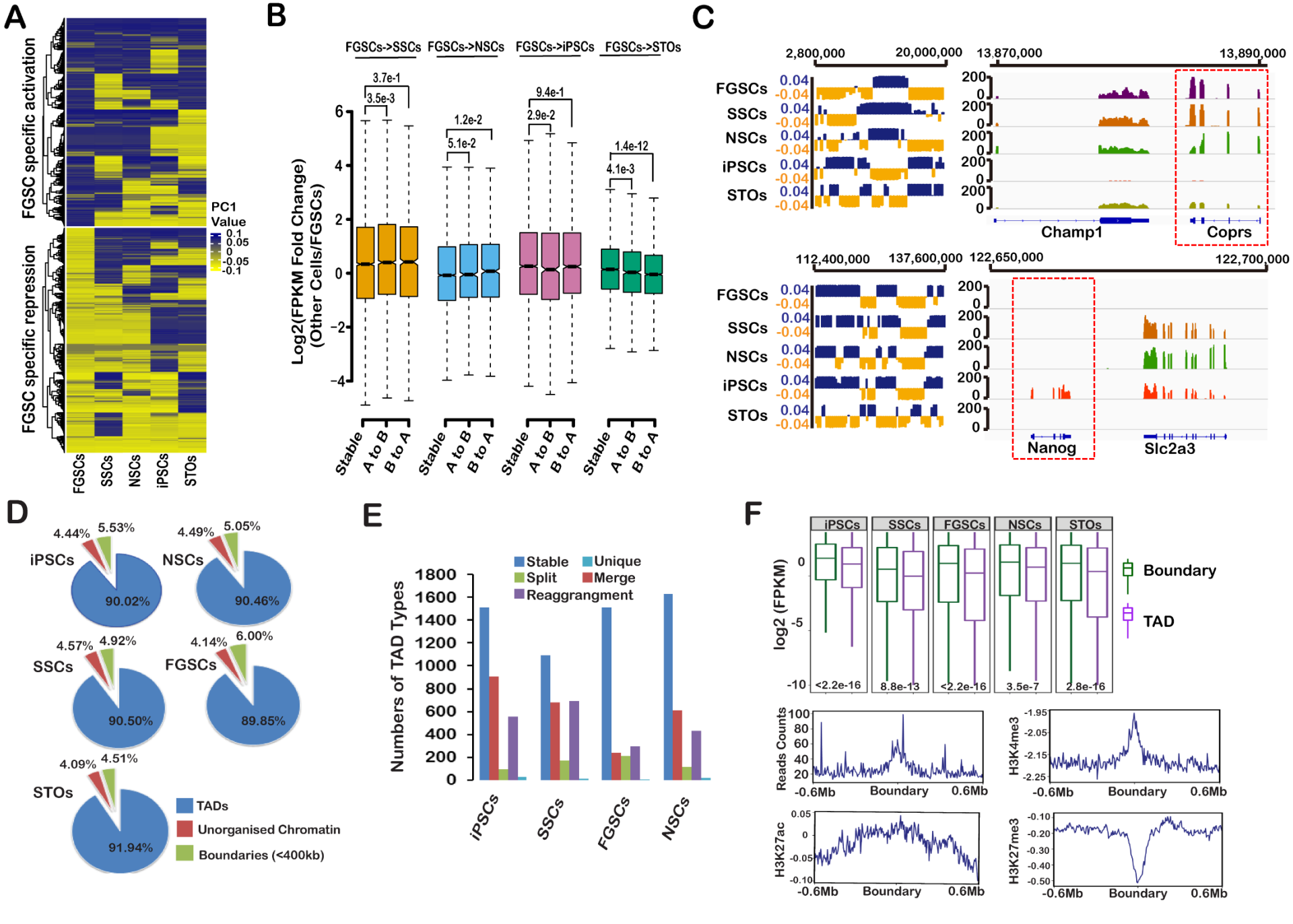
638 (C) Numbers of conserved compartment regions in paternal and maternal genomes.

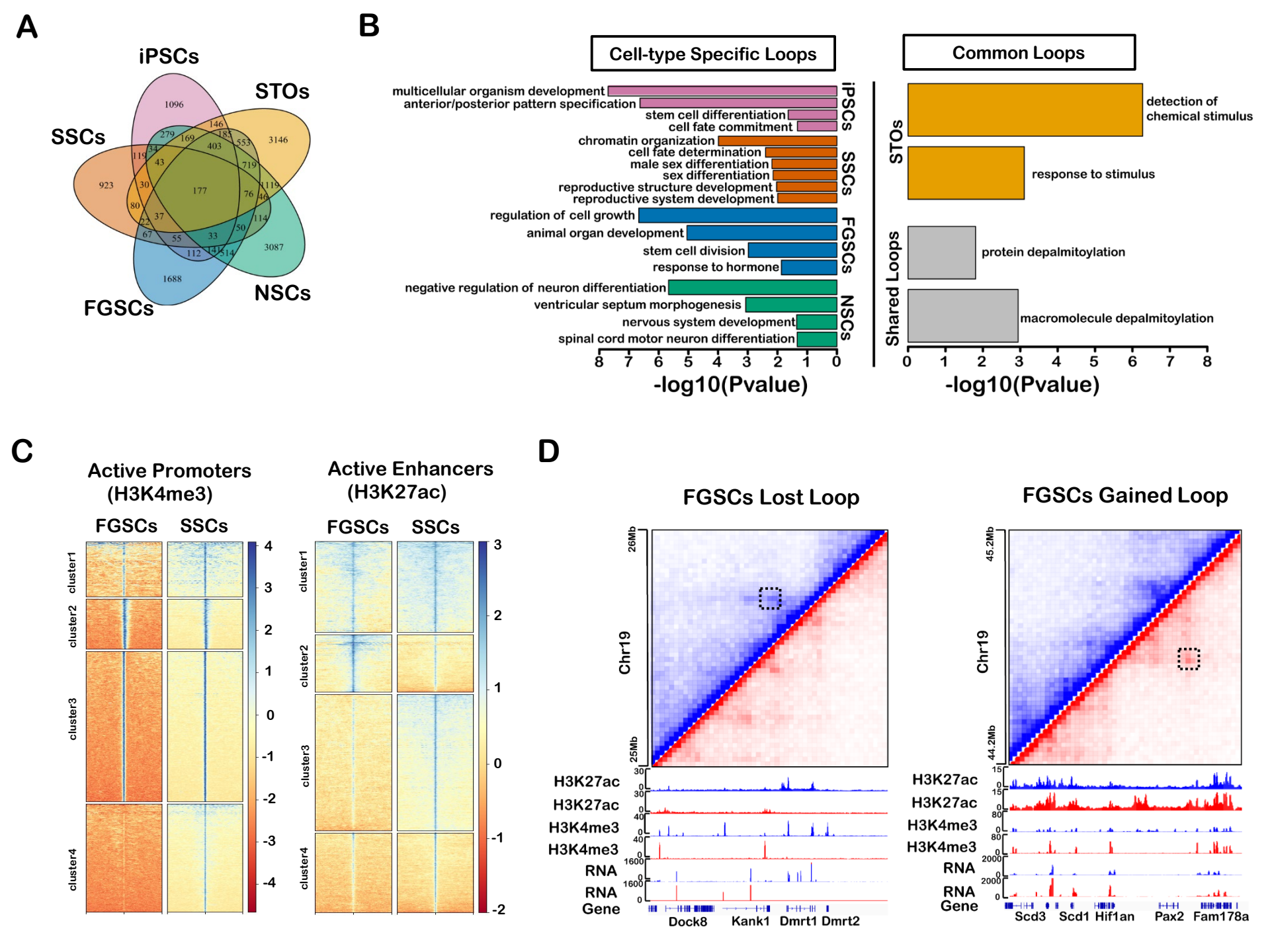
639 (D) Venn diagram showing overlap of conserved compartment regions between

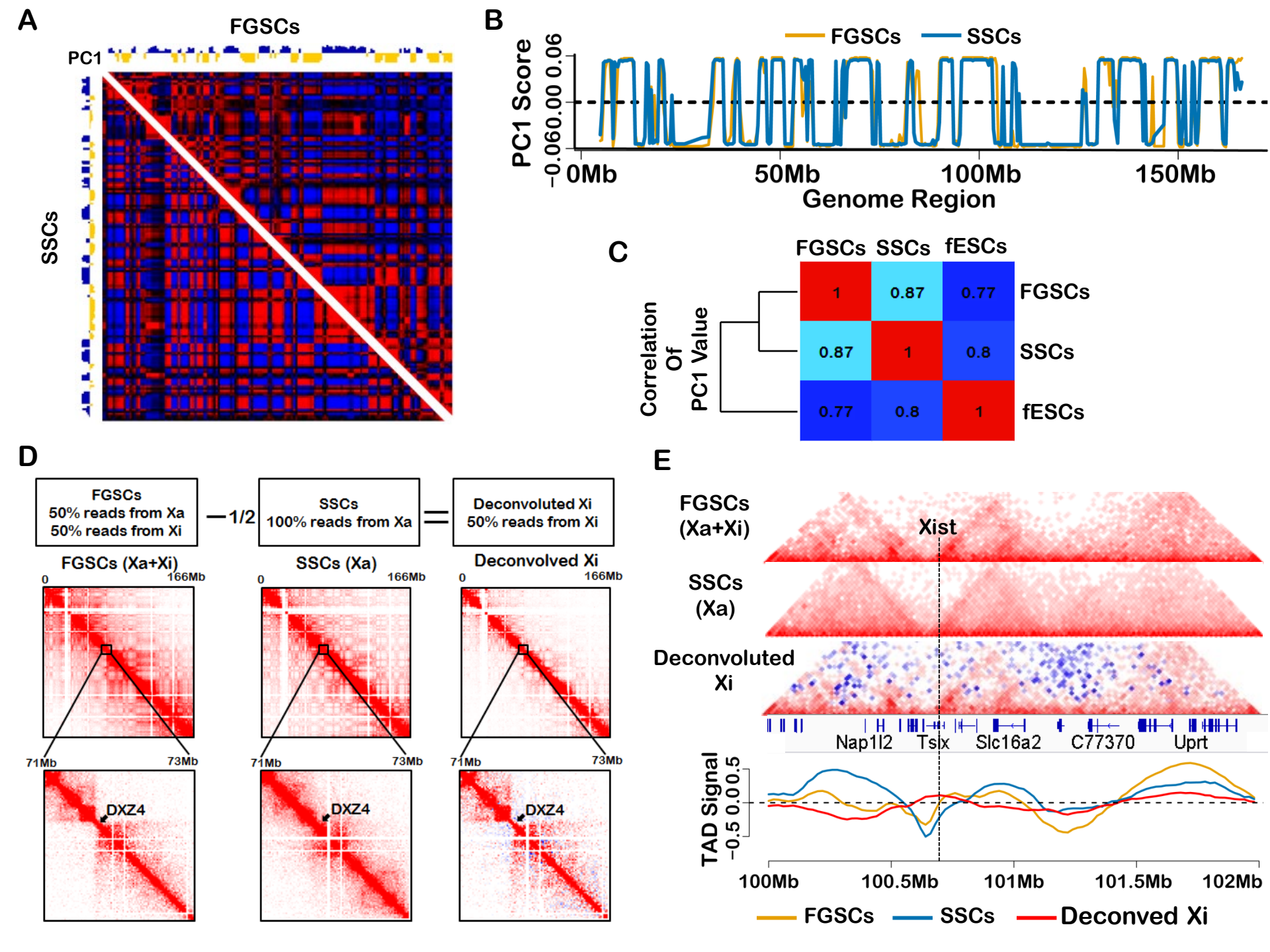
paternal and maternal genomes.

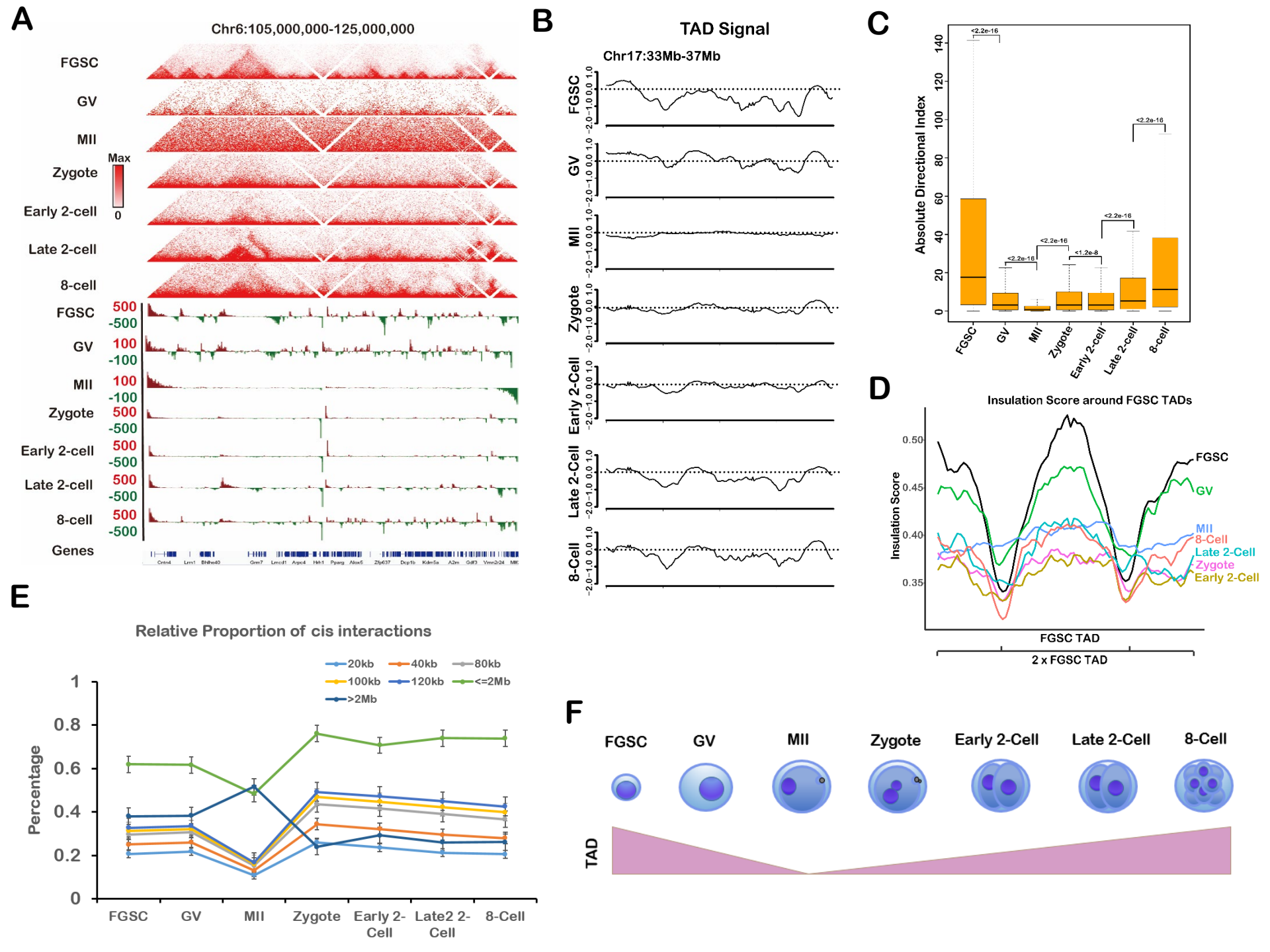

A
I
II
III
IV
Marker
FGSCs
STO
NC
Mvh
213bp
Oct4
313bp
Fragilis
151bp
Stella
354bp
Dazl
173bp
Gfrα1
191bp
Gapdh
220bp
V
VI
VII
VIII
B
I
II
III
IV
Marker
SSCs
MEF
NC
Etv5
168bp
Oct4
231bp
Gfrα1
183bp
Plzf
173bp
Mvh
154bp
Gapdh
220bp
V
VI
VII
VIII
C
Marker
NSCs
NC
Sox2
438bp
Nestin
387bp
Ascl1
429bp
Olig2
579bp
Pax6
249bp
Gapdh
496bp
D

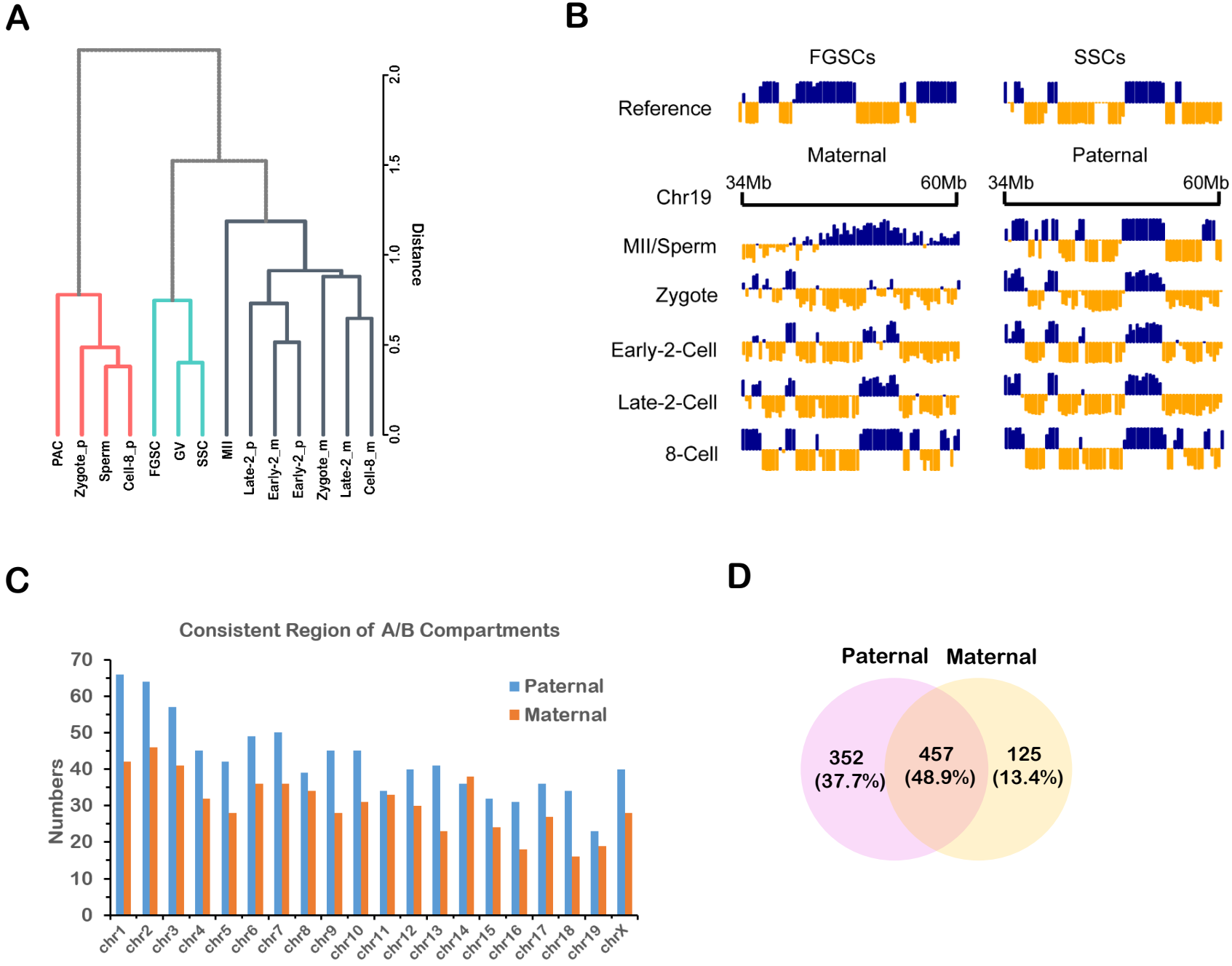

A
Chromosome 19
0
210
FGSCs
SSCs
NSCs
iPSCs
STOs
B
Chr19
40,000,000
60,000,000
PC1 Value
Compartment A/B
FGSCs
SSCs
NSCs
iPSCs
STOs
0.04
-0.04
Normalized heat maps
Max
0
Directional Index
500
-500
RefSeq genes
2c37
Lcor
Loxl4
Cutc
Pax2
Hps6
Ina
Sorcs3
Sorcs1
Ins1
Mxi1
Gpam
Nrap
Trub1
Pnlip
C
Compartment
TAD
Numbers
3500
3000
2500
2000
1500
1000
500
0
728
2259
717
2649
939
2807
662
3090
1006
2599
FGSCs
SSCs
NSCs
iPSCs
STOs
D
log10(P(s))
log10(Genomic distance(bp))
-5
-6
-7
-8
-9
5.0
5.5
6.0
6.5
7.0
7.5
8.0
FGSCs
SSCs
NSCs
iPSCs
STOs
S-1
S-0.5


A
FGSC specific activation
FGSC specific repression
PC1 Value
0.1
0.05
0
-0.05
-0.1
FGSCs
SSCs
NSCs
iPSCs
STOs
B
FGSCs->SSCs
FGSCs->NSCs
FGSCs->iPSCs
FGSCs->STOs
Log2(FPKM Fold Change) (Other Cells/FGSCs)
3.7e-1
3.5e-3
1.2e-2
5.1e-2
9.4e-1
2.9e-2
1.4e-12
4.1e-3
Stable
A to B
B to A
C
2,800,000
20,000,000
13,870,000
13,890,000
Champ1
Coprs
112,400,000
137,600,000
122,650,000
122,700,000
Nanog
Slc2a3
D
iPSCs
NSCs
SSCs
FGSCs
STOs
4.44%
5.53%
90.02%
4.49%
5.05%
90.46%
4.57%
4.92%
90.50%
4.14%
6.00%
89.85%
4.09%
4.51%
91.94%
TADs
Unorganised Chromatin
Boundaries (<400kb)
E
Stable
Unique
Split
Merge
Reaggrangment
Numbers of TAD Types
iPSCs
SSCs
FGSCs
NSCs
F
log2 (FPKM)
Boundary
TAD
<2.2e-16
8.8e-13
<2.2e-16
3.5e-7
2.8e-16
Reads Counts
H3K4me3
H3K27ac
H3K27me3
-0.6Mb
Boundary
0.6Mb


A
iPSCs
STOs
SSCs
FGSCs
NSCs
1096
146
279
169
185
553
3146
403
119
34
43
719
923
30
177
1119
76
46
80
37
114
22
67
55
33
50
112
141
514
3087
1688
B
Cell-type Specific Loops
Common Loops
multicellular organism development
anterior/posterior pattern specification
stem cell differentiation
cell fate commitment
chromatin organization
cell fate determination
male sex differentiation
sex differentiation
reproductive structure development
reproductive system development
regulation of cell growth
animal organ development
stem cell division
response to hormone
negative regulation of neuron differentiation
ventricular septum morphogenesis
nervous system development
spinal cord motor neuron differentiation
iPSCs
SSCs
FGSCs
NSCs
8 7 6 5 4 3 2 1 0
-log10(Pvalue)
STOs
Shared Loops
detection of chemical stimulus
response to stimulus
protein depalmitoylation
macromolecule depalmitoylation
0 1 2 3 4 5 6 7 8
-log10(Pvalue)
C
Active Promoters (H3K4me3)
Active Enhancers (H3K27ac)
FGSCs
SSCs
cluster1
cluster2
cluster3
cluster4
4 3 2 1 0 -1 -2 -3 -4
3 2 1 0 -1 -2
D
FGSCs Lost Loop
FGSCs Gained Loop
Chr19
26Mb
25Mb
45.2Mb
44.2Mb
H3K27ac
H3K4me3
RNA
Gene
Dock8
Kank1
Dmrt1
Dmrt2
Scd3
Scd1
Hif1an
Pax2
Fam178a


A
FGSCs
PC1
SSCs
B
FGSCs
SSCs
PC1 Score
−0.06 0.00 0.06
0Mb
50Mb
100Mb
150Mb
Genome Region
C
FGSCs SSCs fESCs
1 0.87 0.77 FGSCs
0.87 1 0.8 SSCs
0.77 0.8 1 fESCs
Correlation Of PC1 Value
D
FGSCs 50% reads from Xa 50% reads from Xi
1/2
SSCs 100% reads from Xa
Deconvoluted Xi 50% reads from Xi
FGSCs (Xa+Xi)
SSCs (Xa)
Deconvolved Xi
0
166Mb
71Mb
73Mb
DXZ4
E
FGSCs (Xa+Xi)
Xist
SSCs (Xa)
Deconvoluted Xi
Nap1l2
Tsix
Slc16a2
C77370
Uprt
TAD Signal
−0.5 0.0 0.5
100Mb
100.5Mb
101Mb
101.5Mb
102Mb
FGSCs
SSCs
Deconved Xi


A
Chr6:105,000,000-125,000,000
FGSC
GV
MII
Zygote
Early 2-cell
Late 2-cell
8-cell
Max
0
500
-500
100
-100
Genes
B
TAD Signal
Chr17:33Mb-37Mb
C
Absolute Directional Index
<2.2e-16
<1.2e-8
D
Insulation Score around FGSC TADs
Insulation Score
FGSC TAD
2 x FGSC TAD
E
Relative Proportion of cis interactions
20kb
40kb
80kb
100kb
120kb
<=2Mb
>2Mb
Percentage
FGSC
GV
MII
Zygote
Early 2-Cell
Late2 2-Cell
8-Cell
F
TAD


A
PAC
Zygote_p
Sperm
Cell-8_p
FGSC
GV
SSC
MII
Late-2_p
Early-2_m
Early-2_p
Zygote_m
Late-2_m
Cell-8_m
Distance
0.0
0.5
1.0
1.5
2.0
B
FGSCs
SSCs
Reference
Maternal
Paternal
Chr19
34Mb
60Mb
34Mb
60Mb
MII/Sperm
Zygote
Early-2-Cell
Late-2-Cell
8-Cell
C
Consistent Region of A/B Compartments
Paternal
Maternal
Numbers
0
10
20
30
40
50
60
70
chr1
chr2
chr3
chr4
chr5
chr6
chr7
chr8
chr9
chr10
chr11
chr12
chr13
chr14
chr15
chr16
chr17
chr18
chr19
chrX
D
Paternal
Maternal
352
(37.7%)
457
(48.9%)
125
(13.4%)