DNA Research, 2019, 26(2), 105–117
doi: 10.1093/dnares/dsy042
Advance Access Publication Date: 27 December 2018
Full Paper

OXFORD

Full Paper

Molecular characteristics of early-stage female germ cells revealed by RNA sequencing of low-input cells and analysis of genome-wide DNA methylation

Binbin Ma[1†], Tin-Lap Lee[2,*†], Bian Hu[3,4†], Jing Li[5†], Xiaoyong Li[1], Xiaodong Zhao[6], Changliang Hou[1], Chen Zhang[1], Lin He[1], Xingxu Huang[3,4], Xuejin Chen[7*], Jing Li[5*], and Ji Wu[1,8*]

[1]Key Laboratory for the Genetics of Developmental & Neuropsychiatric Disorders (Ministry of Education), Bio-X Institutes, Shanghai Jiao Tong University, Shanghai 200240, China, [2]Reproduction, Development and Endocrinology Program, School of Biomedical Sciences, Faculty of Medicine, The Chinese University of Hong Kong, Shatin, Hong Kong, China, [3]MOE Key Laboratory of Model Animal for Disease Study, Model Animal Research Center of Nanjing University, Nanjing 210061, China, [4]School of Life Science and Technology, Shanghai Tech University, Shanghai 201210, China, [5]Department of Bioinformatics and Biostatistics, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, China, [6]Shanghai Center for Systems Biomedicine, Shanghai Jiao Tong University, Shanghai 200240, China, [7]Department of Laboratory Animal Sciences, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China, and [8]Key Laboratory of Fertility Preservation and Maintenance of Ministry of Education, Ningxia Medical University, Yinchuan 750004, China

*To whom correspondence should be addressed. Tel.+86 021 34207263. Email: jiwu@sjtu.edu.cn (J.W.); Tel. +86 021 34204348. Email: jing.li@sjtu.edu.cn (J.L.); Tel. +85239434436. Email: leetl@cuhk.edu.hk (T.-L.L.); Tel. +86 021-63846590. Email: chenxuejin@shsmu.edu.cn (X.C.)

†These authors contributed equally to this work.

Edited by Dr. Minoru Ko

Received 7 June 2018; Editorial decision 11 November 2018; Accepted 14 November 2018

Abstract

High-throughput stage-specific transcriptomics provides an unbiased approach for understanding the process of cell development. Here, we report transcriptome analysis of primordial germ cell, female germline stem cell (FGSC), germinal vesicle and mature oocyte by performing RNA sequencing of freshly isolated cells in mice. As expected, these stages and gene-expression profiles are consistent with developmental timing. Analysis of genome-wide DNA methylation during female germline development was used for confirmation. By pathway analysis and blocking experiments, we demonstrate PI3K-AKT pathway is critical for FGSC maintenance. We also identify functional modules with hub genes and lncRNAs, which represent candidates for regulating FGSC self-renewal and differentiation. Remarkably, we note alternative splicing patterns change dramatically during female germline development, with the highest occurring in FGSCs. These findings are invaluable resource for dissecting the molecular pathways and processes into oogenesis and will be wider applications for other types of stem cell research.

Key words: female germ cells, RNA sequencing, transcriptomics, DNA methylation, development

1. Introduction

In most multicellular organisms, including mammals, germ cells are the cells from which new organisms originate and thus serve to pass genetic information from one generation to the next. Primordial germ cells (PGCs) are the first germ cell population established during development.[1] PGCs from mice initially become identifiable as a cluster of approximately 40 cells at around embryonic day 7.25 (~E7.25).[2,3] They migrate and colonize the genital ridges at ~E10.5.[4–6] The PGCs from female embryos with the XX karyotype continue to proliferate until ~E13.5, reaching around 25,000 in number. Subsequently, these cells enter prophase I of meiotic division and then are arrested at the diplotene stage of prophase I of meiosis [germinal vesicle (GV) oocytes]. Upon hormonal stimulation, the GV oocytes complete the first meiotic division with concomitant extrusion of the first polar body [metaphase II (MII) oocytes].

However, recent studies have indicated that not all PGCs enter the differentiation pathway.[7–14] A study at our laboratory showed that FGSCs can be isolated from the ovaries of 5-day-old and adult mice using immunomagnetic sorting for the mouse Vasa homolog (*MVH*/*Ddx4*). After culture for more than 15 months or 6 months, respectively, the FGSCs still exhibited proliferative capacity and a normal karyotype. These cells can produce fertile offspring after transplantation into sterilized ovaries, followed by mating of the recipient with an adult male mouse.[14] The presence of FGSCs has also been demonstrated in the ovaries of post-natal rat, pig, and adult women.[9–11,15]

Our previous studies showed that FGSCs subjected to short-term culture shared similar morphological and molecular signatures to male germline stem cells.[16] Moreover, DNA methylation (5-cytosine methylation), a repressive epigenetic marker, has a key role in germ cell development.[17] Our recent study also showed that it is actively involved in specifying the development potential of FGSCs.[18] Recently, one independent group reported that FGSCs isolated from neonatal mice displayed the string-forming features and vigorous mitotic activity at 1-3 days post-partum.[19] However, the transcriptome or gene expression network or epigenetic signatures of freshly isolated FGSCs, and the network between the FGSCs and PGCs or the FGSCs and GV oocytes remain elusive. RNA sequencing (RNA-seq) technique for the analysis of was used, making it feasible to analyse the transcriptome and gene expression network of freshly isolated FGSC, and to decipher the spatiotemporal patterns of gene expression in PGC, FGSC, GV and MII oocyte. These studies are crucial to obtain a deeper understanding of the developmental processes of female germ cells and the molecular signatures of FGSC because the identity and behaviour of a cell are determined by its gene expression network.

2. Materials and methods

2.1. Animals

C57BL/6, pou5f1-GFP transgenic mice [CBA-Tg (pou5f1-EGFP) 2Mnn] (Jackson Laboratory) ×C57BL/6 F_1 hybrid mice were used. The Ddx4-Cre; mT/mG mice were produced as described previously[20] and the Stella-GFPcreERT2 strain was generated as described in Supplementary data.

2.2. Isolation and culture of FGSCs

FGSCs were isolated and purified from the ovaries of 3-5-day-old mice by the same procedures as used previously.[14] Then, the MVH^+ cells were plated to mitotically inactivated STO cell feeders. FGSC medium was changed every 2 days and cells were subcultured every 4-5 days. The culture system was maintained at 37°C and 5% CO_2, with saturated humidity.

2.3. Immunofluorescence and alkaline phosphatase staining of PGCs and FGSCs

Immunostaining was performed as previously described.[14] Cells were fixed with 4% PFA for 30 min, washed twice with 0.1% Tween-20 in PBS (PBST), incubated with 1% PBST for 10 min, blocked with 10% goat serum in PBS for 30 min, and then incubated with goat anti-rabbit MVH (1: 200; Abcam) or goat anti-mouse H3K27me3 (1: 150; Abcam) at 4°C overnight. After incubation, the cells were washed twice and then incubated with secondary TRITC-conjugated antibodies (1: 200; Invitrogen). Finally, the cells were incubated with DAPI and viewed under a fluorescence microscope. For the alkaline phosphatase staining, the samples were analysed using the Alkaline Phosphatase Detection Kit (Millipore), following the manufacturer's instructions.

2.4. RNA isolation and RT-PCR

Total RNA was isolated from the cultured FGSCs in accordance with the protocol supplied with TRIzol reagent (Invitrogen, Life Technologies). Then, 1 μg of RNA was reverse-transcribed with M-MLV reverse transcriptase. PCR analyses were performed with Taq DNA Polymerase. Primer sequences are listed in Supplementary Table 4.

2.5. RNA-seq of low-input cells

RNA-seq was performed with PGCs, FGSCs, GV and MII oocytes. The detailed information is included in Supplementary data.

2.6. Gene ontology analysis

GO analysis was applied to explore the functions of the DEGs identified in this study. It organizes genes into hierarchical categories and can uncover regulatory networks based on biological processes and molecular functions (http://www.geneontology.org) (The Gene Ontology Project, 2017). Specifically, two-sided Fisher's exact test was used to classify the GO category and the GO annotation list is greater than that expected by chance. The false discovery rate (FDR) was calculated to correct the *P*-value. We computed *P*-values for GOs enriched among the DEGs (recommended *P*-value <0.05).

2.7. MeDIP-seq

Ovaries from 30 female mice (Ddx4-Cre; mT/mG mice, 3-5 days) were collected. Isolation and fluorescence activated cell sorting (FACS) of FGSCs from mouse ovaries was performed as described previously.[21] The preparation of MeDIP and input DNA libraries were performed as previously described.[18] Briefly, genomic DNA was extracted and the residual RNA was removed by RNAse. Then Genomic DNA was randomly sheared by sonication to generate ~300 bp fragments. The fragmented chromatins were pre-cleared and then immunoprecipitated with Protein A + G Magnetic beads coupled with 5-methylcytosineantibody. After reverse crosslinking, MeDIP and input DNA fragments were end-repaired and A-tailed using the NEBNext End Repair/dA-Tailing Module (E7442, NEB) followed by adaptor ligation with the NEBNext Ultra Ligation Module (E7445, NEB). The DNA libraries were amplified and subjected to deep sequencing with an Illumina HiSeq 2000.

2.8. Analysis of DNA absolute methylation

DNA methylation data was downloaded from published papers.[18,22,23] DNA methylation levels detected by reduced representation bisulphite sequencing (RRBS) were calculated with Bismark tools. Meanwhile, FGSC methylation levels detected by MeDIP-seq were calculated to transfer the absolute methylation through QSEA package with default settings.

2.9. Pathway analysis

Pathway analysis was used to determine the significant pathway of DEGs. Pathway annotations of microarray genes were downloaded from WEGG (http://www. genome.jp/Wegg/). Fisher's exact test was used to determine the significant enrichment pathway. The resulting *P*-values were adjusted using the BH FDR algorithm. Pathway categories with FDR < 0.05 were reported. The enrichment provides a measure of the significance of the function: as the enrichment increases, the corresponding function is more specific, which helps us to identify more significant pathways in the experiment. The enrichment is calculated as follows: enrichment $= \left(\frac{n_g}{n_a}\right) / \left(\frac{N_g}{N_a}\right)$, where n_g is the number of DEGs within the particular pathway, n_α is the total number of genes within the same pathway, N_g is the number of DEGs that have at least one pathway annotation, and N_α is the number of genes that have at least one pathway annotation in the entire microarray.

2.10. Series clustering

We selected the genes differentially expressed among PGCs, FGSCs, GV and MII oocytes. In accordance with the different tendencies for RPKM change of genes at different stages, we identified a set of unique model expression tendencies. Using a strategy for clustering short time-series gene expression data, we defined some unique profiles. The expression model profiles are related to the actual or the expected number of genes assigned to each model profile. Significant profiles have a higher probability than expected by Fisher's exact test and multiple comparison tests.

2.11. Weighted gene co-expression network analysis

A signed weighted correlation network was constructed for any expressed gene (FPKM > 0.01) across the four developmental phases. The expression value was translated into a Z-score normalization value for the subsequent analysis. An adjacency matrix was constructed by raising the co-expression measure to the power $\beta = 14$, which was used to derive a pairwise distance matrix for selected genes. Based on the resulting adjacency matrix, the topological overlap was calculated. Genes with highly similar co-expression relationships were grouped together by performing average linkage hierarchical clustering on the topological overlap. In addition, the Dynamic Hybrid Tree Cut algorithm was used to cut the hierarchical clustering tree and define modules as branches from the tree cutting. The node centrality, defined as the within-cluster connectivity measures, was used to rank genes for "hubness" within each cluster. For visual analysis of the constructed networks, we exported the network into edge and node list files that Cytoscape can read with a threshold above 0.65 (some networks were too small to use 0.02). Then, we picked up the subnetwork using genes in GO terms that were related to the developmental process by using Cytoscape 3.1.0. We summarized the expression profile of each module by representing it as a module eigengene. Modules whose eigengenes were correlated at a level of $r > 0.25$ were merged.

2.12. miRNA-mRNA-lncRNA target network

We introduced the Miranda package to predict miRNA target on 3'UTR region of differentially expressed mRNA and the full-length sequence of differentially expressed lncRNA and miRNA sequence. Competing endogenous RNA (CeRNA) relations was constructed by a pair of lncRNA and mRNA affected by the same miRNA members. In this network, a circle represents mRNA, a diamond represents lncRNA, and a rectangle represents miRNA; a relationship is represented by an edge.

2.13. RNA extraction from low-input cells and XIST validation in FGSCs

Eight FGSCs were incubated in reverse transcription buffer supplemented with 0.1% NP-40 and RQ1 RNase-free DNase (Promega). Reverse transcription was carried out with random 6-mer primers and dNTP mix (Invitrogen). The mixture was incubated at 50°C for 1 h and then at 37°C for 15 min with RNase H (Invitrogen). The cDNA was amplified with the Multiple Annealing and Looping Based Amplification Cycles (MALBAC) kit. Then, the cDNAs were subjected to two rounds of PCR amplification to detect *XIST*, *TSIX*, and *MVH*.

2.14. Alternative splicing detection

We selected the Alternative Splicing Detector (ASD, http://www.novelbio.com/asd/ASD.html)[24] as a tool to detect the cases with differential alternative splicing based on a bam file after mapping and using the mouse genome sequence as a reference, according to *P-value* <0.05.

3. Results

3.1. Collection and biological characteristics of female germ cells

To perform RNA-seq analysis of female germ cells at different developmental stages, we collected PGCs, FGSCs, GV and MII oocytes from 12.5 days post-coitum (dpc), neonatal and adult ovaries, respectively (see Materials and methods, Fig. 1A, Supplementary Table S1). For PGCs and FGSCs, we used two-step enzymatic digestion and MVH-based immunomagnetic isolation and sorting or fluorescence-activated cell sorting (FACS) for analysis of DNA methylation in FGSCs (see Materials and methods). Most of the sorted cells were characterized by the round or ovoid shape with a large nucleus and small cytoplasm (Fig. 1A I, II). Furthermore, these cells were confirmed as germ cells by *MVH* expression (Fig. 1B I–III, Fig. 1C I, III). The sorted cells were also positive for OCT4 and alkaline phosphatase staining (Fig. 1B IV-IX, Fig. 1C II). After culturing for 2 days, the sorted FGSCs appeared as grape-like clusters (Fig. 1D I). In addition, FGSCs still expressed *MVH*,[25] *DAZL* [deleted in azoospermia-like][26] and *DPPA3* [developmental pluripotency-associated 3],[27] but didn't express *C-KIT* [also referred to as *SCFR*, stem cell factor receptor],[28] *SYCP3* [synaptonemal complex protein 3][29] and *ZP3* [zona pellucida glycoprotein 3][30] after long-term culture (Fig. 1D II-IV).

3.2. Transcriptional profiles of female germ cells at different developmental stages

Using the Illumina Hiseq 2500, we performed RNA-seq of 5-8 cells for PGCs, FGSCs, GV and MII oocytes and generated 173.1 million reads to detect the genes expressed in female germ cells at different

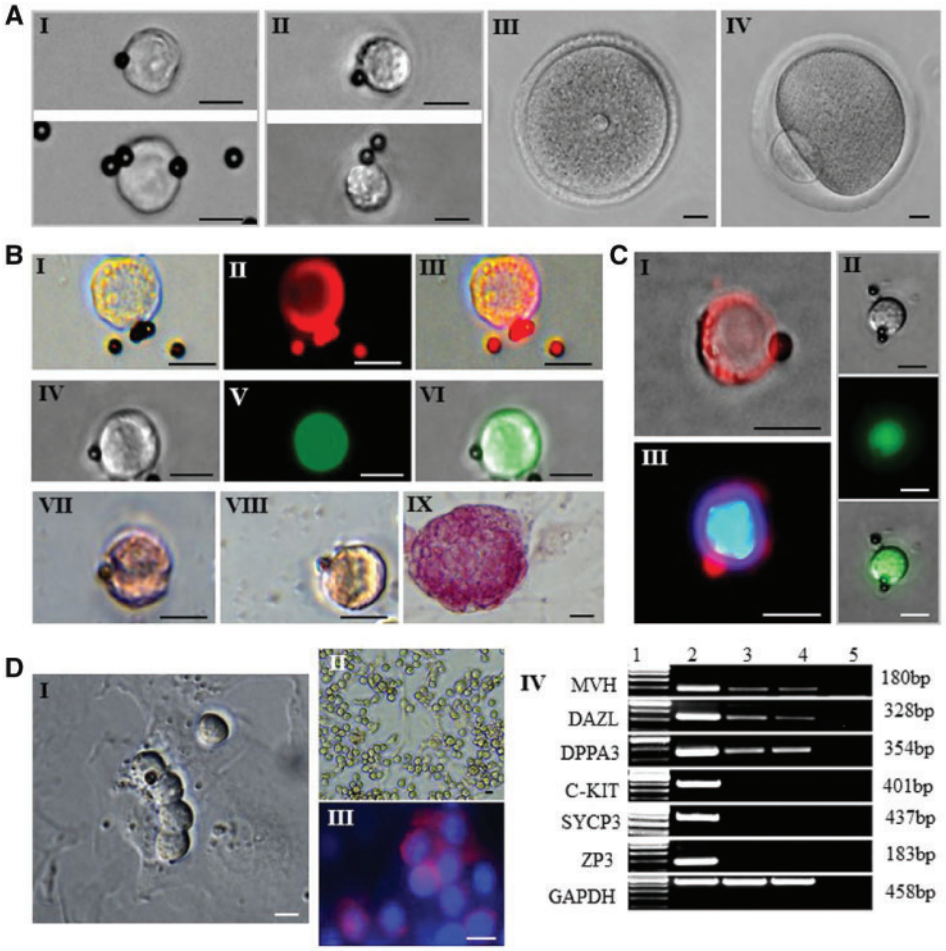

Figure 1. Morphology of freshly isolated mouse PGCs, FGSCs, GV and MII oocytes and their biological characteristics. **(A)** Representative imaging of freshly isolated PGC (I), FGSC (II), GV oocyte (III), and MII oocyte with the first polar body (IV). **(B)** Characterization of PGC: Immunofluorescence for MVH of PGC (I: bright field; II: MVH) and PGC from pou5f1-GFP transgenic mice (IV: bright field; V: field under UV light). Alkaline phosphatase staining of PGC (VII) and FGSC (VIII), with mESCs as positive control (IX). **(C)** Characterization of FGSC: Immunofluorescence for MVH of FGSC (I: merger of bright field and MVH; III: merger of MVH and DAPI), and FGSC from pou5f1-GFP transgenic mice (II) (top: bright field; middle: field under UV light). **(D)** Typical morphology of isolated FGSCs that had been cultured for 2 days, with a grape-like cluster (I); morphology of FGSC line (II); immunofluorescence of FGSC line for MVH and DAPI (III); Scale bars: 10 um. RT-PCR analysis of female germline-specific markers (IV). Lane 1: 100-bp DNA markers; lane 2: mouse ovary; lane 3: FGSC line A; lane 4: FGSC line B; lane 5: mock.

stages, with a read length of 100 bp (Supplementary Table S2). For each high-throughput sequencing analysis, two biological replicates were performed. Because of the low-input samples, we first perform Fast-QC to control the data quality. We acquired the Q-score higher than 30 (error rate $< 0.1\%$) for each replicates, indicating the sequencing quality is very solid (Supplementary Fig. S1A). Finally, we detected on the average expression of 17,805 (47%) out of 37,980 Refseq genes. Nearly half of the known mouse genes were expressed in the female germ cells, especially in FGSCs, expressing 18,727 (49%) out of 37,980 Refseq genes. To explore whether these gene expression profiles correlated with the developmental stages, we used unsupervised hierarchical clustering to analyse the RNA-seq data. The results showed that these female germ cells were accurately in accordance with the developmental order, from PGCs to FGSCs and ends with GV and MII oocytes, as expected (Fig. 2A). The results of principal component analysis (PCA) revealed differences in expression patterns among developmental stages (Fig. 2B). PGCs, FGSCs, GV and MII oocytes were clustered separately in the PCA plots, suggesting they are four distinct stages. GV and MII oocytes were close together, indicating relatively similar transcriptomes between these two stages (Fig. 2B). Moreover, the correlation coefficient of each replicated sample show very good consistency (Supplementary Fig. S1B). The results from a pairwise comparison of gene expression for all female germ cell stages also demonstrated stage-specific gene expression patterns (Fig. 2C I-III). For instance, as many as 6,846 genes showed differential expression between FGSCs and GV oocytes (Fig. 2C II). Interestingly, the majority of differentially expressed genes (DEGs) could be clustered into four distinct stages reflecting stage-specific expression patterns and may thus play critical roles during female germline development (Fig. 2D).

3.3. Genome-wide DNA methylation patterns during female germ cell development

To carry out genome-wide DNA methylation analysis of female germ cell development, we firstly performed MeDIP-seq analysis of freshly isolated FGSCs as we described previously.[18] We generated 32.9 million sequencing reads; nearly 50% were uniquely mapped to the mouse

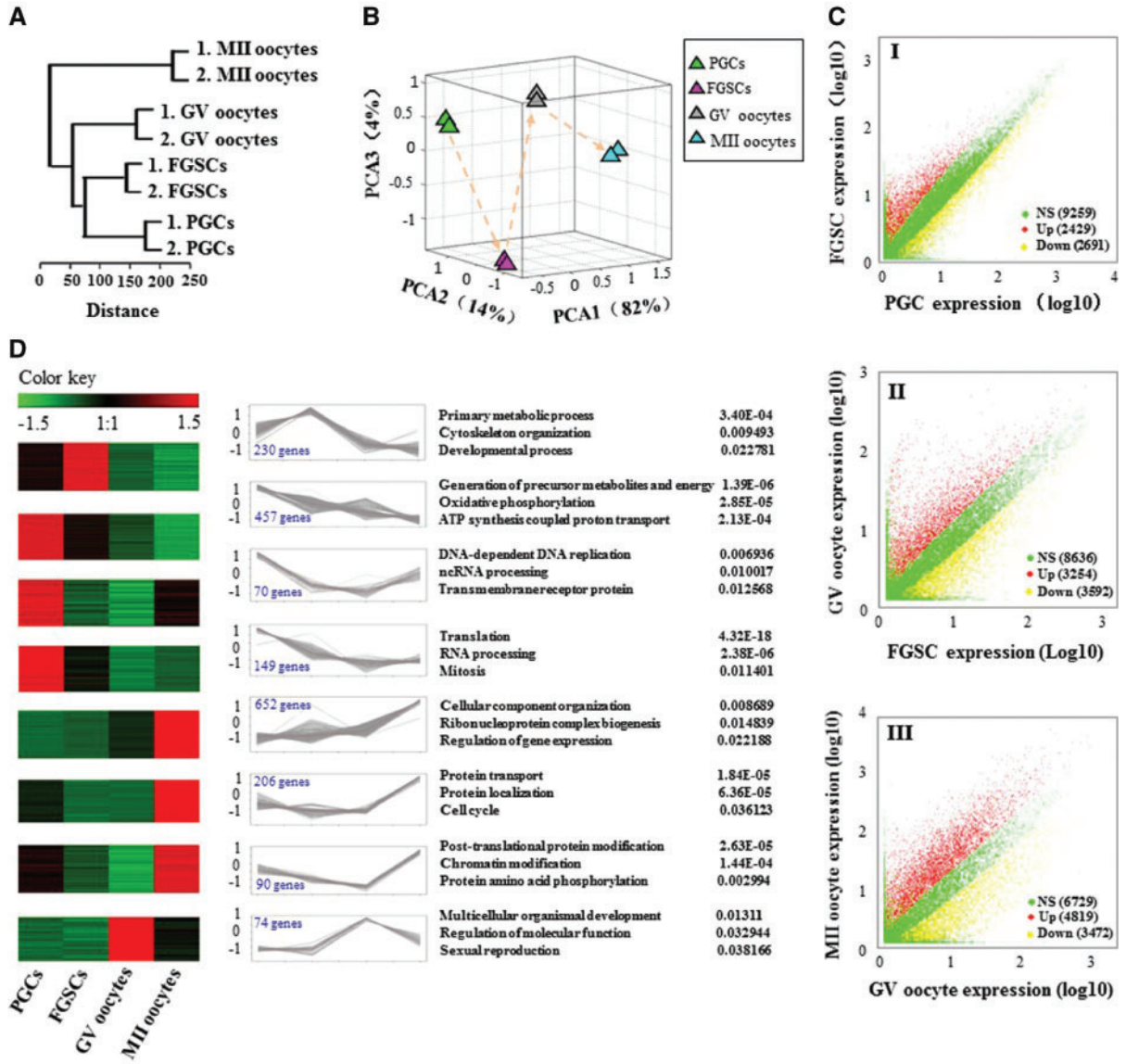

Figure 2. Global expression patterns of known RefSeq genes in the mouse female germline development. **(A)** Unsupervised clustering of the transcriptomes of PGCs, FGSCs, GV and MII oocytes. All RefSeq genes expressed with RPKM $\geq$ 0.1 were used for the analysis. **(B)** PCA of the transcriptomes of PGCs, FGSCs, GV and MII oocytes. Cells from the same developmental stage are shown with the same colour. The arrows indicate the developmental direction between consecutive stages. PCA1, PCA2 and PCA3 represent the top three dimensions of the genes showing differential expression from PGCs to MII oocytes, which account for 82%, 14% and 4% of the expressed RefSeq genes, respectively. **(C)** Scatter plots of RPKM between PGCs and FGSCs, FGSCs and GV oocytes, and GV and MII oocytes. Red, blue and yellow dots represent genes that were up-regulated, no significant change and down-regulated, respectively. The numbers in parentheses represent the numbers of genes. **(D)** Clusters of genes showing representative expression patterns in the mouse female germline. The top GO terms and corresponding enrichment *P*-values are shown on the right side.

reference genome. The sequencing reads were further processed to profile DNA methylation signal. Next, we analysed the latest data for genome-wide DNA methylation from PGCs,[22] GV oocytes and MII oocytes.[23] The result showed that PGC has relatively lower methylation level. Following the sex determination, FGSC also maintain lower methylation, while increases methylation with differentiation into GV and MII oocyte (Fig. 3A). This verified the results from RNA-seq data of female germ cells at different stages. To make it clear whether there is epigenetic pattern alteration between freshly isolated FGSCs and the cultured FGSCs,[18] we also performed the genome-wide DNA methylation analysis of these two populations. We observed that the raw sequencing datasets were fairly correlated (Fig. 3B). The fresh and cultured FGSCs showed very similar DNA methylation pattern, evidenced by both chromosome scale and *Stra8* locus (Fig. 3C). Next, we examined methylation status of CpG islands (CGIs) in the genome of fresh and cultured FGSCs. We found more than 95% methylated CGIs in cultured FGSCs were recaptured in fresh FGSCs (Fig. 3D). We performed functional annotation of shared methylated CGIs. Consistent with our previous finding,[18] we found the enriched items were mostly related to somatic development (Fig. 3E). Together, these observations suggest DNA methylation pattern of cultured FGSCs was almost identical to that of fresh FGSCs.

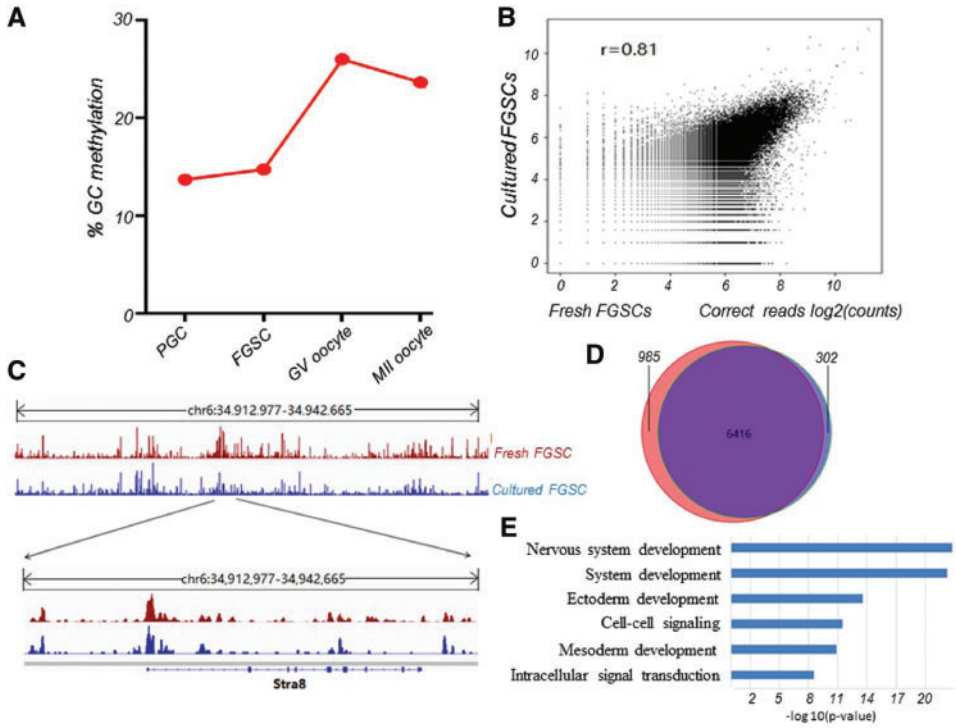

Figure 3. Genome-wide DNA methylation patterns during female germline development and methylome comparison of fresh and cultured FGSCs. **(A)** DNA methylation levels during female germline development. **(B)** Correlation analysis (log-transformed corrected read count) of the fresh FGSCs and cultured FGSCs MeDIP-ChIP data sets. **(C)** A snapshot of the IGV genome browser showing DNA methylation ($5^{m}C$) signal at the indicated region (top track) and Stra8 locus (bottom track). **(D)** Methylated CGIs identified in fresh FGSCs (the purple) and cultured FGSCs (the red) show high agreement. **(E)** Functional annotation of methylated genes shared by fresh FGSC and cultured FGSCs ($p < 0.05$).

3.4. Pathway analysis of the DEGs in female developmental stages

To explore the pathways that are involved in biological processes during female germline development, we used a Bayesian model-based clustering method[31] (for details, see Materials and methods) to analyse the DEGs from PGCs to MII oocytes.[29] In accordance with different tendencies for RPKM change of genes under different situations, we identified 26 unique model expression tendencies (Fig. 4A). Using a strategy for clustering short time-series gene expression data, we defined 13 significant unique profiles. Significant profiles have a higher probability than expected by Fisher's exact test and multiple comparison tests. Then we enriched signalling pathway of the DEGs, and found that tendencies 1 and 10 showed gradual downward trends from PGCs to MII oocytes, which were thus deduced to be related to the self-renewal of FGSCs. In contrast, tendencies 8 and 24 exhibited increasing trends from FGSCs to GV oocytes, suggesting that these two groups are related to FGSC differentiation. Intriguingly, tendencies 1/10 shared a pathway, PI3K-AKT (Fig. 4B I), while tendencies 8/24 shared the TGF-β pathway (Fig. 4B II), suggesting these two pathways play vital roles in maintaining the self-renewal and regulating the differentiation of FGSCs, respectively.

To confirm the role of PI3K-AKT signalling pathway in FGSC self-renewal and survival, we added the AKT inhibitor IV to the medium of FGSCs. As expected, we observed that the proliferative capacity of FGSCs was significantly reduced (Fig. 4C I, II, IX). Furthermore, EdU incorporation assay showed that the proportion of EdU-positive cells was dramatically decreased with the AKT inhibitor IV (Fig. 4C III-VI, X). To observe cell apoptosis when the PI3K-AKT pathway was inhibited in cultured FGSCs, we stained cells with Annexin V and PI to assess apoptosis (Annexin V^{+}/PI^{-}) and cell death (Annexin V^{+}/PI^{+}) upon AKT inhibitor IV treatment for 2 days. This analysis revealed increased proportions of both apoptotic and dead cells in the AKT inhibitor IV-treated FGSCs (Fig. 4C VII, VIII, XI). qRT-PCR experiments that we carried out on the downstream target genes of AKT showed that RAF1, CCND1 and P53 were down-regulated, while *P21* and *BRCAL* were up-regulated, when AKT was inhibited, which was compatible with the present data set (Fig. 4C XII, XIII, Supplementary Fig. S2). For TGF-β pathway, a previous study identified BMP4 as a key regulator of PGC differentiation.[32] In addition, the output of *in vitro*-derived oocytes was significantly increased upon BMP4 supplementation because it could activate SMAD1/5/8 and key meiosis-initiating genes in FGSCs (or OSCs).[33] These results suggest that the TGF-β superfamily might regulate FGSC differentiation during mouse oogenesis. This finding matches the results from the pathway analysis, which further validates the RNA-seq data.

3.5. Gene expression dynamics during the development of female germ cells

To further clarify the molecular characteristics of early-staged female germ cells, especially FGSCs, we monitored the dynamic expression of germline and core stem-cell genes. We first analysed key genes associated with PGC specification and germline development. The expression of *FRAGILIS*, *PRDM14*, *MVH*, *OVOL2*, *PRMT5*, *PUM1*, *GJA1*, *REX1* and *NANOS3*, tended to decrease over the course of oogenesis, which is compatible with previous reports.[34–41] Conversely, the expression levels of *DAZL* and *STELLA* tended to increase gradually during these consecutive stages (Supplementary Fig. S3A), which

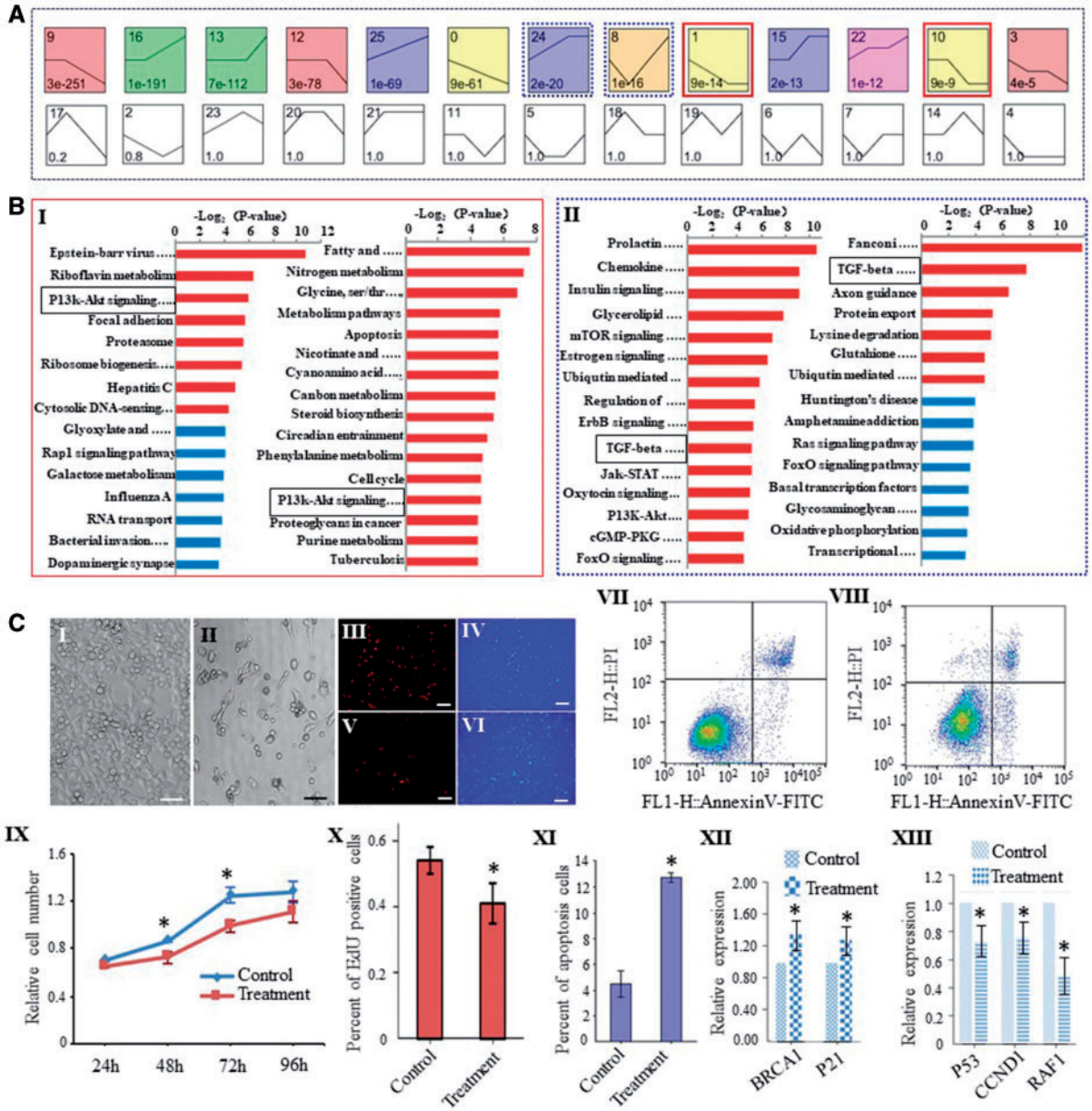

Figure 4. The regulation of biological process and pathway analysis during mouse female germline development. **(A)** Twenty-six unique tendencies were identified with different tendencies for RPKM change of the genes at different stages. Each box corresponds to one of the model temporal expression profiles. The number in the top left-hand corner of a profile box is the profile ID number, while the number at the bottom of a profile box is the *P*-value for the significance. The black curve shows a model expression pattern. Red line boxes indicate a downward trend from PGCs to MII oocytes. Blue line boxes indicate an increasing trend from FGSCs to GV oocytes. **(B)** The enriched pathway terms for representative tendencies related to self-renewal (I) and differentiation (II) during female germline development. Black line boxes indicate PI3K-AKT pathway and TGF-β pathway, respectively. **(C)** Validation of the function of the PI3K-AKT pathway in regulating self-renewal in FGSC lines. I, II, Morphology of control (I) and FGSCs treated with AKT inhibitor-IV (II). III-VI, Immunofluorescence of EdU in control and AKT inhibitor IV-treated FGSCs (III, V: Edu; IV, VI: DAPI). X: Proportion of EdU-positive cells. Scale bars: 50 μm. VII, VIII: FACS analysis showed the incidence of cell apoptosis with Annexin V and PI. VII: Control; VIII: FGSCs treated with the AKT inhibitor IV. XI: Proportion of apoptotic cells. IX: CCK8 assays were performed on FGSCs with or without AKT inhibitor-IV (treatment) after 24 h, 48 h, 72 h and 96 h. XII, XIII: qRT-PCR analysis confirmed the downstream target genes for AKT . $*P < 0.05$, Student's *t*-test, compared with control.

is probably related to their regulating functions in the female developmental process. To confirm these findings, we used qRT-PCR to verify the transcriptional change of *ZP1*, *ZP3*, *SYCP3*, *FKBP9*, *FHL1* and *REX1*, which are all consistent with the RNA-seq data analysis (Supplementary Fig. S3B). Meanwhile, we used Stella-GFPcreERT2 mouse strain (see Materials and methods) to track the protein activity of Stella. Interestingly, the dynamic pattern of the Stella protein coincided with its mRNA expression as determined by the qRT-PCR and RNA-seq analysis[42] (Supplementary Fig. S4), which in turn confirmed the validity of our RNA-seq data.

We then analysed meiotic and oocyte specific genes. The results showed that early meiotic genes, *SYCP1*, *SYCP3* and *MOS*, were dramatically up-regulated during the meiotic progression from FGSCs to oocytes, supporting their critical roles in meiosis of female germ cells. Accordingly, during the differentiation process from FGSC to oocytes, we found clear up-regulation of oocyte marker genes, including *ZP1*, *ZP3*, *GDF9*, *OOG1*, *NLRP5*, *NOBOX*, *LHX8*, and *BMP15*, suggesting that they play regulatory roles in oocyte formation.

Furthermore, we analysed the core stemness genes (*DTYMK*, *ZMYM4*, *AW549877*, *NMEL*, *FKBP9*, *FHL1* and *CCS*) during these stages.[16] Intriguingly, these genes showed a global down-regulated trend, especially upon the transition from FGSCs to GV oocytes. The expression pattern of these genes is probably involved in early-stage female germ cell proliferation, especially cell fate specification of FGSCs.

3.6. Weighted gene co-expression network analysis for the female germline development

To investigate the co-expression relationships among female germ cell developmental stages with specific gene-regulatory modules, we performed weighted gene co-expression network analysis (WGCNA) in this study.[43,44] This unsupervised and unbiased analysis identified distinct co-expression modules corresponding to clusters of correlated transcripts (Fig. 5A). By applying WGCNA independently to our mouse RNA-seq data (RPKM > 0.1), we found that female germline development involves 25 co-expression modules (Fig. 5A). When we merged the modules that perform the same or similar functions in a particular stage, four composite modules were obtained. Interestingly, these modules showed stage-specific expression; that is, they comprised genes that tended to be specifically expressed in a single developmental stage. Gene Ontology (GO) enrichment showed that stage-specific modules shared stage-correlated biological processes, such as RNA processing, cell division and cell cycle (PGCs), mitotic cell cycle and nuclear division (FGSCs), gamete generation and sexual reproduction (GV oocytes), and meiotic chromosome segregation, and programmed cell death (MII oocytes) (Fig. 5B), indicating accurately step-wise developmental progression. We then examined the distribution of module expression during development, showing that the PGC modules undergo gradual degradation over the course of development, whereas the modules in the transitions from FGSCs to GV oocytes or from GV to MII oocytes showed trends of a sharp degradation and continuous activation (Fig. 5C).

We analysed the correlations between modules and developmental stages (Fig. 5D), demonstrating that mouse female germline development also involved stage-specific co-expression modules. Next, to identify the core stage-specific regulatory genes, we constructed hub gene networks and filtered a cluster of hub genes (*Ncapd2*, *INCENP*, *Rab5b*, *PSAP*, *GTPBP3* and *EIF4H*) from the FGSC-specific brown, middlelightblue and tan modules, using GO terms most abundantly represented in these modules (Fig. 5E). Collectively, we can speculate that FGSCs may be potentially associated with those hub genes and further specify the stem cell self-renewal through controlling their mitotic progression.

3.7. Dynamic expression of long non-coding RNAs

To trace the expression of long non-coding RNAs (lncRNAs) in female germline development, we mapped the Noncode v3.0 to the NCBI database, filtered out those less than 200 nt, and retained the ncRNAs from 200 nt to over 100 kb in length.[45] Finally, we obtained 7,859 known lncRNAs, including ncRNAs, pseudo-RNAs and miscRNAs. Among these, there were 632 significantly differentially expressed lncRNAs. After hierarchical clustering analysis, we found that, like protein-coding genes, these lncRNAs showed very distinct stage-specific expression patterns, suggesting that they have regulatory roles during female germline development (Fig. 6A). Subsequently, we used RNA-seq data that were directly or indirectly mappable to the mouse reference genome, but had no overlap with sequences in RefSeq, NCBI or Noncode v3.0 lncRNA databases, then carried out de novo transcript assembly using the reference-guided assembly software Cufflinks.[46] Finally, we obtained 376 novel lncRNAs. Interestingly, 13 novel lncRNAs were highly expressed at the FGSC stage (Supplementary Table S3).

Recently, miRNAs have been identified to be important regulators of development and homeostasis in mammalians. And lncRNAs were reported to regulate their corresponding mRNAs by acting as a decoy for miRNAs that bind to common sites in the 3′ UTR.[47] To obtain further insights into the crosstalk of lncRNA, mRNA and miRNA in female germline, we used three hub genes (*Ncapd2*, *CDKN1A* and *Rab5b*), two known germline marker genes (*MVH* and *C-KIT*) and one stemness-associated gene (*LHX1*) to construct the interaction networks. Intriguingly, we obtained functionally characterized lncRNAs, including *XIST* and *MALAT1* (Fig. 6B, C, Supplementary Figs. S5 and S6). Using single cells RT-PCR, we identified *XIST* expression at the FGSCs stage (Fig. 6D), which is consistent with the RNA-seq data. The immunofluorescent results also showed the expected frequencies of Xi-like enrichment for H3K27me3 in FGSCs (Fig. 6D), suggesting that one of X-chromosomes in FGSCs was inactive. Notably, we found ZFP783 could interact with *MVH* and *CDKN1A* in FGSCs through miR-6963-3p. Since *CDKN1A* has been confirmed to be important for maintenance of FGSCs (Fig. 4C XII), this reveals *ZFP783* could regulate FGSC germline specification and mitotic cell cycle through miRNA sponge effect.

3.8. Dynamic patterns of alternative splicing

According to the reads that were uniquely mapped to transcript-isoform-specific exons or exon-exon junctions, we analysed those genes with more than two known transcripts and found that there were 8,532 (52.8%) genes expressed in PGCs, 10,698 (57.12%) in FGSCs, 9,822 (53.3%) in GV oocytes, and 10,161 (53.76%) in MII oocytes (Fig. 7A). This data suggested that the dynamic patterns of alternative splicing (AS) pattern showed variations across female germline development, with the highest proportion occurring in FGSCs, indicating that this specific developmental stage features more active AS and that it potentially contributes to their ability of self-renewal and differentiation. The proportions of known genes with two or more transcripts at the FGSC stage are also shown in Fig. 7B.

To further investigate the FGSC-specific AS pattern, we obtained intersection elements of AS between PGCs vs FGSCs and FGSCs vs GV oocytes for subsequent analysis. With an adjusted *P*-value <0.05, we identified 698 AS events in FGSC stage, including 196 skipped exons, 116 retained intron, 83 alternative last exon, 80 alternative to 5′ splice site, 71 alternative to 3′ splice site, 53 alternative start exon, 28 mutually exclusive exon and 71 uncommon AS events (Fig. 7C). Surprisingly, we found that, besides cassette, retained intron was the second most prevalent type in FGSCs, which suggested that retained intron also greatly contributed to the molecular features of FGSCs. GO biological process of AS intersection enriched in

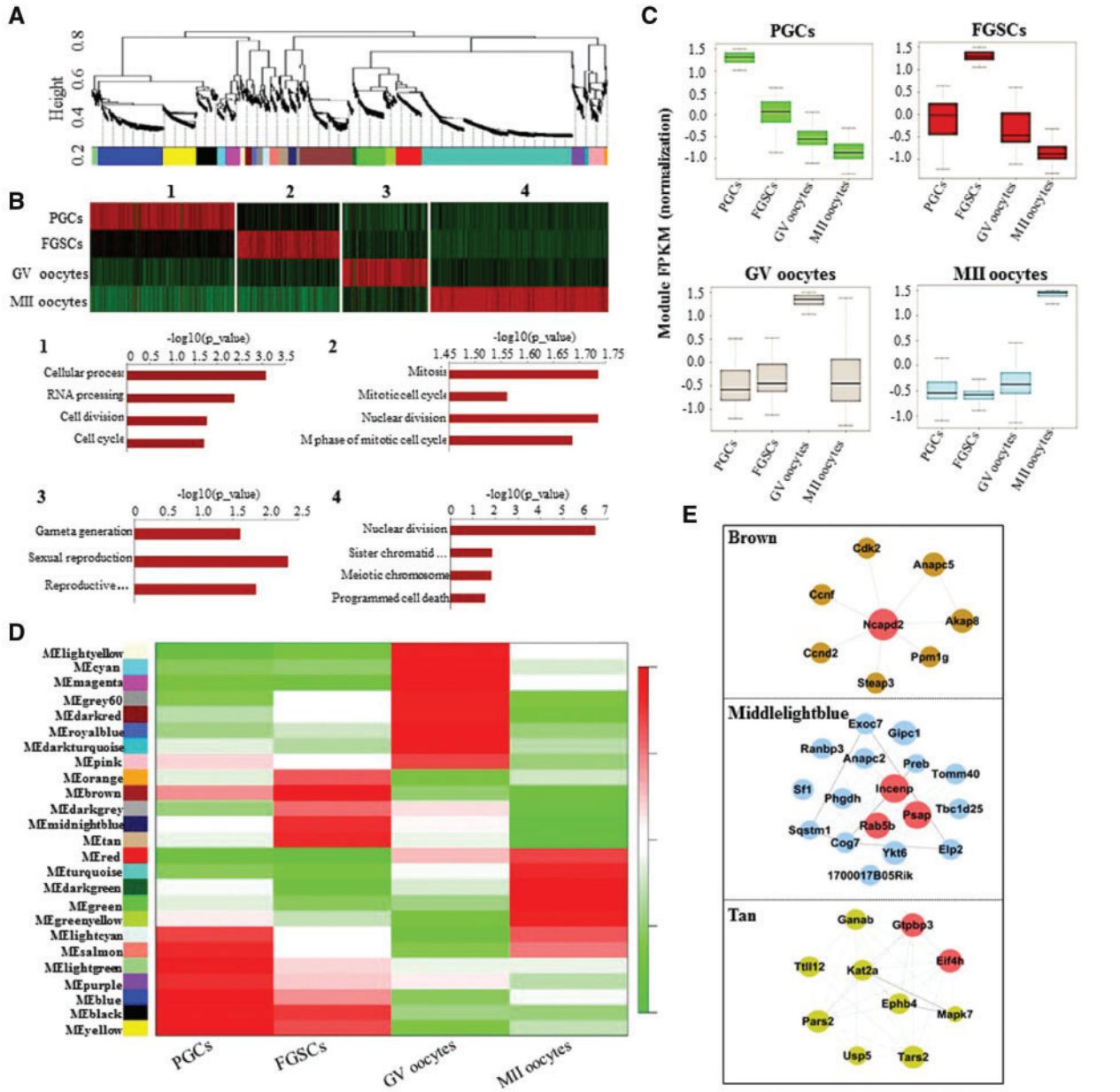

Figure 5. WGCNA revealed gene-network modules enriched in the female germline. **(A)** Hierarchical cluster tree shows co-expression modules identified using WGCNA. **(B)** GO enrichment and heatmap showing the stage-specific developmental events. Several modules were functionally aggregated and mapped to each stage with different biological processes. The most strongly related GO terms and their corresponding *P*-values are shown. **(C)** Boxplots showing the distribution of module expression (mean RPKM of all expressed genes in each module) for different cell types. **(D)** Heatmap reporting the correlations between each module and developmental stage. Colour legend indicates the level of correlation between gene co-expression and stage-specific expression. **(E)** Hub-gene networks of the FGSC-specific modules. Size of the dots represents hubness.

mitotic cell cycle and spindle organization. Intriguingly, we found G-protein coupled receptor pathway also participated in mitotic process (Fig. 7D). Based on previous study,[16] we presented core genes, *DTYMK*, *FHL1*, *CDK2*, *SYCP3* and *TBP* for their AS types (Fig. 7E). *DTYMK* and *FHL1* showed conversion process-specific AS occurrence from PGCs to FGSCs and from FGSCs to GV oocytes. Upon comparing the junction read counts, *DTYMK* and *FHL1* exhibited less exon exclusion at the FGSC stage. While *CDK2* and *TBP* showed a different trend, acquiring the highest proportion of exon exclusion in FGSCs. *SYCP3* showed a dramatically high exon-skipping ratio in GV oocytes compared with that in PGCs and FGSCs. Together with those results, dynamic AS patterns could affect the mitotic gene expressions and then contribute to the unique properties of FGSCs.

4. Discussion

To shed more light on the molecular genetics and signatures of early-stage female germ cells, we determined the transcriptome of freshly isolated germ cells using RNA-seq of low-input cells. The results showed that FGSCs were genetically (or evolutionarily) close to

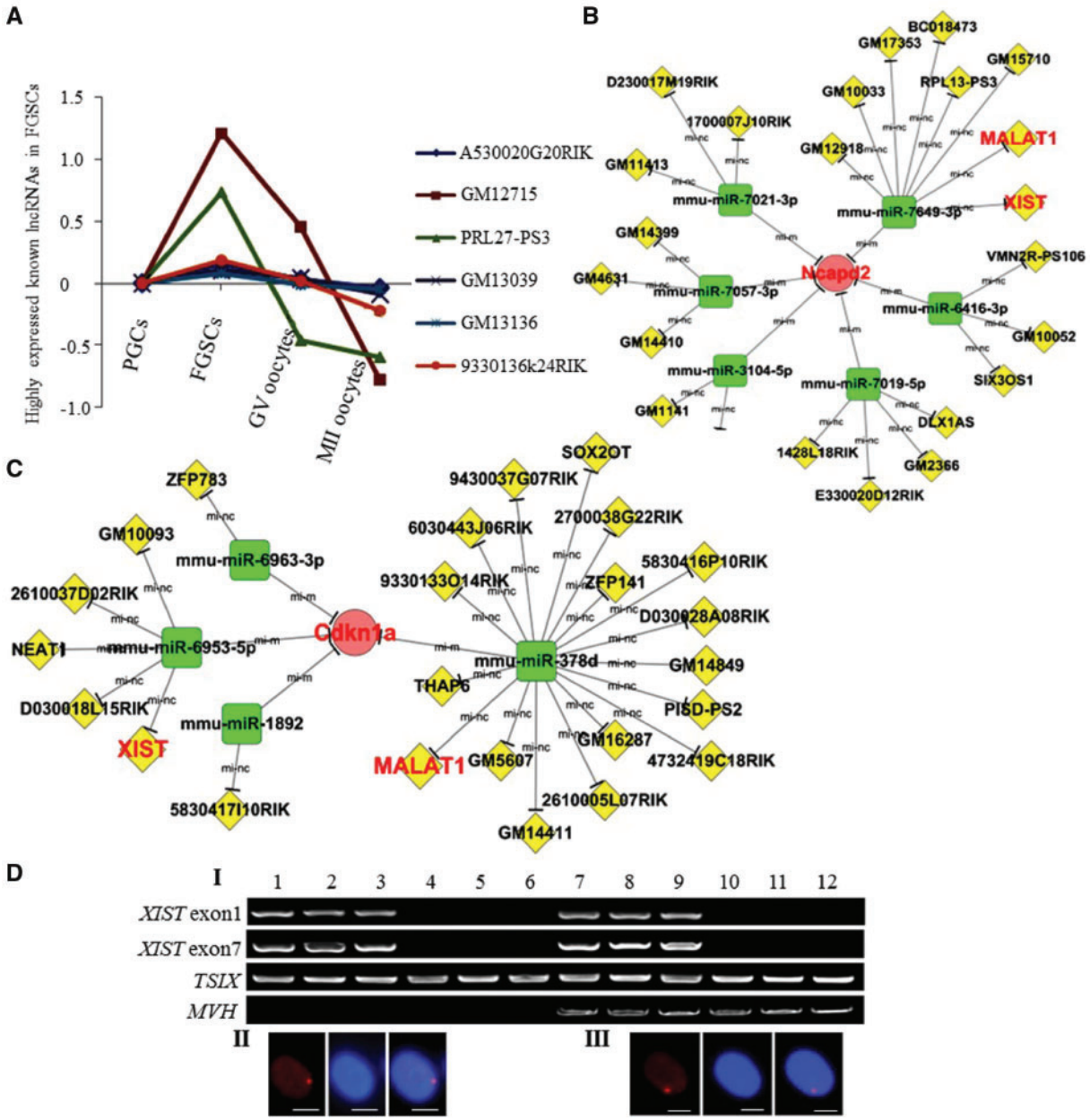

Figure 6. Dynamic expression of lncRNAs during female germline development. **(A)** Dynamic expression of FGSC-specific highly expressed known lncRNAs during female germline development. Through the series cluster and tendency analysis, those FGSC-specific highly expressed known lncRNAs are shown. **(B and C)** Interaction network [including the miRNA-mRNA predicted relationship (Energy<−20, Score>170, P<0.001) and the lncRNA-miRNA interaction (Energy<−20, Score>155, P<0.001)] with WGCNA-derived hub genes at the FGSC stage. Partial enlargement of the interaction network with *NCAPD2* (B) and *CDKN1A* **(C)** (WGCNA-derived hub genes at the FGSC stage) as representative examples. **(D)** XIST expression in FGSCs. I, single cells RT-PCR of *XIST*, *TSIX*, and *MVH*. Lanes 1-3, female tail fibroblasts (positive control); lanes 4-6, male tail fibroblasts (negative control); lanes 7-9, FGSCs; lanes 10-12, oocytes. II and III, FGSCs line showed the expected frequencies of the Xi-like enrichment for H3K27me3 (II). Female tail fibroblasts as a positive control (III). H3K27me3 (left) and DAPI (middle). Merge (right) of H3K27me3 and DAPI. Scale bars: 10 μm.

PGCs, which are the precursors of germline cells. However, FGSCs also exhibited some unique molecular properties, evidenced by PCA analysis and pairwise comparison of gene expression. These findings provide a comprehensive overview of the transcriptomic landscape of female germline stem cell maintenance, differentiation, and oocyte maturation. Intriguingly, this is consistent with the result from analysis of genome-wide DNA methylation of female germ cells. Moreover, DNA methylation pattern of cultured FGSCs were almost identical to that of fresh FGSCs, suggesting that cultured FGSCs still maintain characteristics of germline stem cells and possess the known molecular signatures of FGSC lines.[10,14,20]

Through the clustering short time-series gene expression analysis, we determined that PI3K-AKT pathway was associated with mitotic activity of FGSCs. With AKT inhibition, we found PI3K-AKT could affect the proliferation and apoptosis of FGSCs through regulating downstream target genes, including *BRCA1*, *P21*, *P53*, *CCND1* and

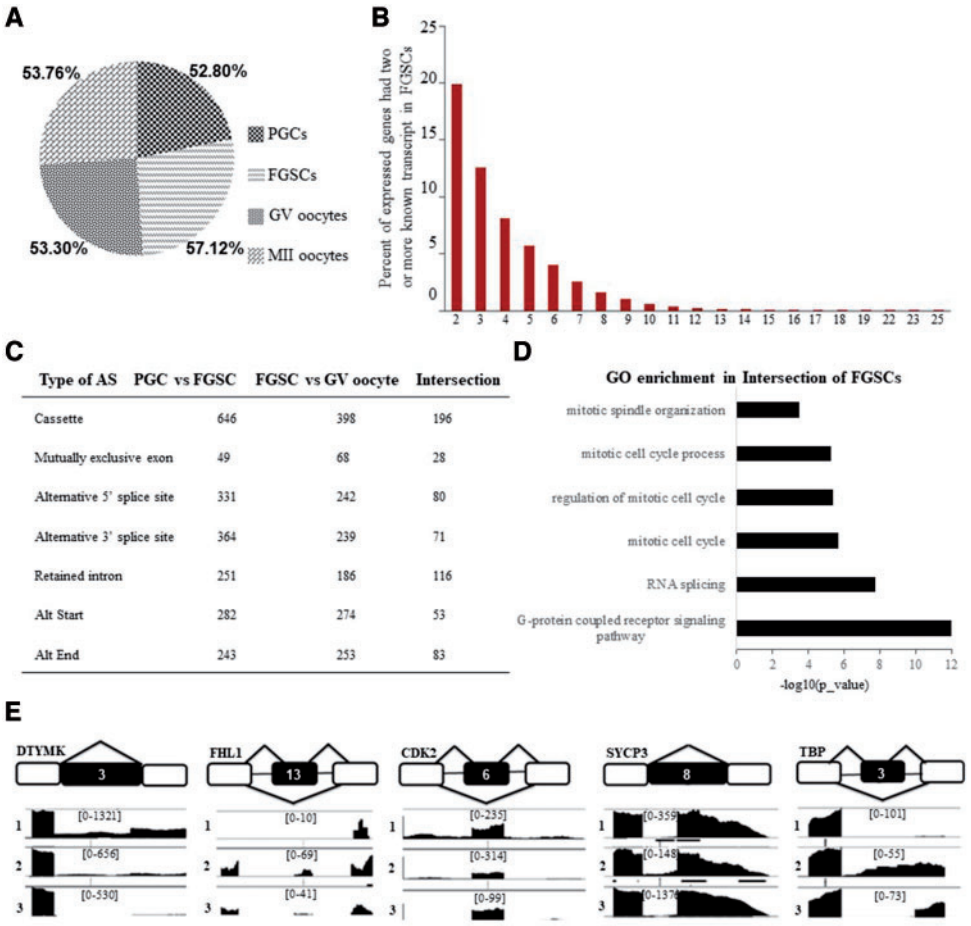

Figure 7. Dynamic patterns of alternative splicing during mouse female germline development. **(A)** Pie chart showing the proportions of known genes expressed with two or more transcript isoforms. **(B)** Proportions of the expressed genes having two or more known transcripts in FGSCs. **(C)** The detected AS types in female germline development. Orange box represents a splicing exon. **(D)** The statistical analysis of differential AS types between PGCs-FGSCs and between FGSCs-GV oocytes, and the intersection indicating the AS types at the FGSC-specific stage. **(E)** Schematic representation of representative AS types of genes related to FGSC self-renewal and differentiation (*DTYMK, FHL1, CDK2, SYCP3,* and *TBP*).

RAF1. Previously study showed that *P53* is strongly induced in the germline stem cells of dysgenic females, and is required for their maintenance,[48] which is also compatible with our results. These findings provide the new insight for future research on the mechanism of FGSCs self-renewal. It has been suggested before that some germline and stem-associated gene clusters are involved in female germline development and stem cell specification. The global expression of each gene cluster was compatible with the pattern reported in previous studies.[1,49,50] For instance, early germline specific genes and stem cell-related genes were mainly detected in PGCs and FGSCs, indicating their progenitor or stem cells identity in an undifferentiated state. By contrast, the meiotic or oocyte-specific genes were highly expressed in GV oocytes or MII oocytes. Our results thus clearly revealed the reliability of the RNA-seq data, which will be a stepping-stone towards exploring more innate cellular and molecular signatures of FGSCs.

By applying WGCNA analysis, we also found specific relevance between co-expression modules and female germline developmental stages. When focusing on the FGSC-specific modules, we identified six hub genes, *NCAPD2*, *RAB5B*, *INCENP*, *PSAP*, *GTPBP3* and *EIF4H*, by network construction. Interestingly, *NCAPD2* is a target of casein kinase 2, Aurora B and CDK1, which is related to mitotic pathways and cell cycle.[51] It contains a nuclear- and chromosome-targeting domain, which interacts with histone H1, H3 and PARP1 in mitotic cells.[52] Taking these together, we can deduce that *NCAPD2* may serve as a candidate regulator for the self-renewal of FGSCs through cell cycle regulatory subunits of condensing complex and associated chromatin modification.

It has been reported that lncRNAs can regulate mRNAs through the miRNA sponge effect. Therefore, we constructed interaction networks with WGCNA-derived hub genes and known genes expressed in FGSCs. As expected, we acquired several clusters of functionally regulating lncRNAs, such as those involved in epigenetic, transcriptional and translational regulation. Notably, XIST induces CpG island methylation by regulating its target genes, and leads to the random establishment of XCI in female germline.[53] Furthermore, we confirmed XIST expression and Xi-like enrichment of H3K27me3 in FGSCs. This supported the assertion that FGSCs maintained their internal signatures with one inactive X chromosome. Based on our results, we would like to propose that a small group of germline precursors still retain their stem cell identity through X chromosome inactivation with the interaction of mRNA, lncRNA and miRNA when most of PGCs enter the meiotic pathway.

AS of pre-mRNAs provides versatile regulatory mechanism that can affect genomic diversity and functional diversification of proteins.[54] During the female germline development, we found that FGSCs showed the highest frequency of occurrence of AS, which revealed that FGSCs acquired more slicing events than other developmental stages. When we got intersection elements of AS in FGSC stage, we were surprised to find that RI and cassette were both critical AS patterns at FGSC stage, which might make a critical contribution to its regulator genes and then affect the stem cell property. Moreover, differential molecular mechanisms allow PGCs to escape the meiotic pathway and continuously maintain their mitotic stem cell identity. Specifically, TBP could recruit PP2A and interact with condensin in order to transmit cell memory through mitosis to the daughter cells.[55] To the extent that TBP might be correlated with the transmission of stem cell memory in a manner other than transcriptional regulation, this might involve the delivery of epigenetic information to daughter cells in FGSCs.

Overall, this study systematically analysed the morphological and molecular features of PGCs, FGSCs, GV and MII oocytes. Using RNA-seq of known RefSeq genes, known lncRNAs, novel lncRNAs and dynamic patterns of AS, it should be possible to delineate many more developmental stage-specific and regulatory mechanisms behind female germline development and the derivation of functional oocytes from FGSCs. These findings will broaden future applications for other types of stem cell research.

Data Availability

RNA-seq data from this study have been submitted to the NCBI Gene Expression Omnibus (GEO; http://www.ncbi. nlm.nih.gov/geo/) under the accession number GSE75738.

Funding

This work was supported by National Natural Science Foundation of China (81720108017), National Basic Research Program of China (2017YFA0504201). We thank J.Z. and B.Z. of Novel Bioinformatics Company for technical help and helpful discussion with the RNA-seq data analysis. We also thank Dr. Xingguo Liu for providing us with the mESCs.

Conflict of interest

None declared.

Supplementary data

Supplementary data are available at *DNARES* online.

References

1. Saitou, M. and Yamaji, M. 2012, Primordial germ cells in mice, Cold Spring Harb, *Perspect. Biol.*, **4**, 59–66.
2. Chiquoine, A.D. 1954, The identification, origin, and migration of the primordial germ cells in the mouse embryo, *Anat. Rec.* , **118**, 135–46.
3. Ginsburg, M., Snow, M.H. and McLaren, A. 1990, Primordial germ cells in the mouse embryo during gastrulation, *Development*, **110**, 521–8.
4. Molyneaux, K.A., Stallock, J., Schaible, K., et al. 2001, Time-lapse analysis of living mouse germ cell migration, *Dev. Biol.* , **240**, 488–98.
5. Seki, Y., Yamaji, M., Yabuta, Y., et al. 2007, Cellular dynamics associated with the genome-wide epigenetic reprogramming in migrating primordial germ cells in mice, *Development*, **134**, 2627–38.
6. Richardson, B.E. and Lehmann, R. 2010, Mechanisms guiding primordial germ cell migration: strategies from different organisms, *Nat. Rev. Mol. Cell Biol.*, **11**, 37–49.
7. Bukovsky, A., Gupta, S.K., Virant-Klun, I., et al. 2008, Study origin of germ cells and formation of new primary follicles in adult human and rat ovaries, *Germline Stem Cells. Humana Press*, **450**, 233–65.
8. De Felici, M. and Barrios, F. 2013, Seeking the origin of female germline stem cells in the mammalian ovary, *Reproduction*, **146**, R125–30.
9. White, Y.A.R., Woods, D.C., Takai, Y., et al. 2012, Oocyte formation by mitotically active germ cells purified from ovaries of reproductive-age women, *Nat. Med.* , **18**, 413–21.
10. Zhou, L., Wang, L., Kang, J.X., et al. 2014, Production of fat-1 transgenic rats using a post-natal female germline stem cell line, *Mol. Hum. Reprod.*, **20**, 271–81.
11. Ding, X., Liu, G., Xu, B., et al. 2016, Human GV oocytes generated by mitotically active germ cells obtained from follicular aspirates, *Sci. Rep.*, **6**, 28218.
12. Hummitzsch, K., Anderson, R.A., Wilhelm, D., et al. 2015, Stem cells, progenitor cells and lineage decisions in the ovary, *Endocr. Rev.* , **36**, 65–91.
13. Wu, C., Xu, B., Li, X., et al. 2017, Tracing and characterizing the development of transplanted female germline stem cells *in vivo*, *Mol. Ther.*, **25**, 1408–19.
14. Zou, K., Yuan, Z., Yang, Z., et al. 2009, Production of offspring from a germline stem cell line derived from neonatal ovaries, *Nat. Cell Biol.*, **11**, 631–6.
15. Bai, Y., Yu, M., Hu, Y., et al. 2013, Location and characterization of female germline stem cells (FGSCs) in juvenile porcine ovary, *Cell Prolif.*, **46**, 516–28.
16. Xie, W., Wang, H. and Wu, J. 2014, Similar morphological and molecular signatures shared by female and male germline stem cells, *Sci. Rep.*, **4**, 5580.
17. Sasaki, H. and Matsui, Y. 2008, Epigenetic events in mammalian germ-cell development: reprogramming and beyond, *Nat. Rev. Genet.*, **9**, 129–40.
18. Zhang, X.-L., Wu, J., Wang, J., et al. 2016, Integrative epigenomic analysis reveals unique epigenetic signatures involved in unipotency of mouse female germline stem cells, *Genome Biol.*, **17**, 162.
19. Liu, J., Shang, D., Xiao, Y., Zhong, P., Cheng, H. and Zhou, R. 2017, Isolation and characterization of string-forming female germline stem cells from ovaries of neonatal mice, *J. Biol. Chem.*, **292**, 16003–13.
20. Zhang, C. and Wu, J. 2016, Production of offspring from a germline stem cell line derived from prepubertal ovaries of germline reporter mice, *Mol. Hum. Reprod.*, **22**, 457–64.
21. Li, X., Ao, J. and Wu, J. 2017, Systematic identification and comparison of expressed profiles of lncRNAs and circRNAs with associated co-expression and ceRNA networks in mouse germline stem cells, *Oncotarget*, **8**, 26573–90.
22. Seisenberger, S., Andrews, S., Krueger, F., et al. 2012, The dynamics of genome-wide DNA methylation reprogramming in mouse primordial germ cells, *Mol. Cell.*, **48**, 849–62.
23. Smallwood, S.A., Tomizawa, S., Krueger, F., et al. 2011, Dynamic CpG island methylation landscape in oocytes and preimplantation embryos, *Nat. Genet.*, **43**, 811–4.
24. Zhou, X., Wu, W., Li, H., et al. 2014, Transcriptome analysis of alternative splicing events regulated by SRSF10 reveals position-dependent splicing modulation, *Nucleic Acids Res.*, **42**, 4019–30.
25. Toyooka, Y., Tsunekawa, N., Takahashi, Y., et al. 2000, Expression and intracellular localization of mouse Vasa-homologue protein during germ cell development, *Mech. Dev.*, **93**, 139–49.
26. Ruggiu, M., Speed, R., Taggart, M., et al. 1997, The mouse Dazla gene encodes a cytoplasmic protein essential for gametogenesis, *Nature*, **389**, 73–7.
27. West, J.A., Viswanathan, S.R., Yabuuchi, A., et al. 2009, A role for Lin28 in primordial germ-cell development and germ-cell malignancy, *Nature*, **460**, 909–13.
28. Driancourt, M.A., Reynaud, K., Cortvrindt, R., et al. 2000, Roles of KIT and KIT LIGAND in ovarian function, *Rev. Reprod.*, **5**, 143–52.

29. Yuan, L., Liu, J.G., Zhao, J., et al. 2000, The murine SCP3 gene is required for synaptonemal complex assembly, chromosome synapsis, and male fertility, *Mol. Cell*, **5**, 73–83.
30. Bleil, J.D. and Wassarman, P.M. 1980, Structure and function of the zona pellucida: identification and characterization of the proteins of the mouse oocyte's zona pellucida, *Dev. Biol.*, **76**, 185–202.
31. Ramoni, M.F., Sebastiani, P. and Kohane, I.S. 2002, Cluster analysis of gene expression dynamics, *Proc. Natl. Acad. Sci. USA.*, **99**, 9121–6.
32. Lawson, K.A., Dunn, N.R., Roelen, B.A.J., et al. 1999, Bmp4 is required for the generation of primordial germ cells in the mouse embryo, *Genes Dev.*, **13**, 424–36.
33. Park, E.-S., Woods, D.C. and Tilly, J.L. 2013, Bone morphogenetic protein 4 promotes mammalian oogonial stem cell differentiation via Smad1/5/8 signaling, *Fertil. Steril.*, **100**, 1468–75.
34. Ohinata, Y., Payer, B., O'Carroll, D., et al. 2005, Blimp1 is a critical determinant of the germ cell lineage in mice, *Nature*, **436**, 207–13.
35. Yabuta, Y., Kurimoto, K., Ohinata, Y., et al. 2006, Gene expression dynamics during germline specification in mice identified by quantitative single-cell gene expression profiling, *Biol. Reprod.*, **75**, 705–16.
36. Tsuda, M., Sasaoka, Y., Kiso, M., et al. 2003, Conserved role of nanos proteins in germ cell development, *Science*, **301**, 1239–41.
37. Mak, W., Fang, C., Holden, T., et al. 2016, An important role of pumilio 1 in regulating the development of the mammalian female germline, *Biol. Reprod.*, **94**, 134.
38. Juneja, S.C., Barr, K.J., Enders, G.C., et al. 1999, Defects in the germ line and gonads of mice lacking connexin431, *Biol. Reprod.*, **60**, 1263–70.
39. Francis, R.J.B. and Lo, C.W. 2006, Primordial germ cell deficiency in the connexin 43 knockout mouse arises from apoptosis associated with abnormal p53 activation, *Development*, **133**, 3451–60.
40. Berrens, R.V. and Reik, W. 2015, Prmt5: a guardian of the germline protects future generations, *Embo J.*, **34**, 689–90.
41. Hayashi, M., Shinozuka, Y., Shigenobu, S., et al. 2017, Conserved role of Ovo in germline development in mouse and Drosophila, *Sci. Rep.*, **7**, 40056.
42. Saitou, M. and Miyauchi, H. 2016, Gametogenesis from pluripotent stem cells, *Cell Stem Cell*, **18**, 721–35.
43. Langfelder, P. and Horvath, S. 2008, WGCNA: an R package for weighted correlation network analysis, *BMC Bioinformatics.*, **9**, 559.
44. Bin, Z. and Steve, H. 2005, A general framework for weighted gene co-expression network analysis, *Stat. Appl. Genet, Mol. Biol.*, **4**, 1–45.
45. Costa, V., Angelini, C., De Feis, I., et al. 2010, Uncovering the complexity of transcriptomes with RNA-Seq, *J. Biomed. Biotechnol.*, 20102010, 853916.
46. Trapnell, C., Williams, B.A., Pertea, G., et al. 2010, Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation, *Nat. Biotechnol.*, **28**, 511–5.
47. Ebert, M.S., Neilson, J.R. and Sharp, P.A. 2007, MicroRNA sponges: competitive inhibitors of small RNAs in mammalian cells, *Nat. Methods*, **4**, 721–6.
48. Tasnim, S. and Kelleher, E.S. 2018, p53 is required for female germline stem cell maintenance in P-element hybrid dysgenesis, *Dev. Biol.*, **434**, 215–20.
49. Hackett, J.A., Sengupta, R., Zylicz, J.J., et al. 2013, Germline DNA demethylation dynamics and imprint erasure through 5-Hydroxymethylcytosine, *Science*, **339**, 448–52.
50. Choi, Y., Yuan, D. and Rajkovic, A. 2008, Germ cell-specific transcriptional regulator Sohlh2 is essential for early mouse folliculogenesis and oocyte-specific gene expression, *Biol. Reprod.*, **79**, 1176–82.
51. Piazza, I., Haering, C.H. and Rutkowska, A. 2013, Condensin: crafting the chromosome landscape, *Chromosoma*, **122**, 175–90.
52. Kong, X., Stephens, J., Ball, A.R., et al. 2011, Condensin I recruitment to base damage-enriched DNA lesions is modulated by PARP1, *PLoS One*, **6**, e23548.
53. Clemson, C.M., Hall, L.L., Byron, M., et al. 2006, The X chromosome is organized into a gene-rich outer rim and an internal core containing silenced nongenic sequences, *Proc. Natl. Acad. Sci. USA.*, **103**, 7688–93.
54. Celotto, A.M. and Graveley, B.R. 2001, Alternative splicing of the Drosophila Dscam pre-mRNA is both temporally and spatially regulated, *Genetics*, **159**, 599–608.
55. Xing, H., Vanderford, N.L. and Sarge, K.D. 2008, The TBP-PP2A mitotic complex bookmarks genes by preventing condensin action, *Nat. Cell Biol.*, **10**, 1318–23.