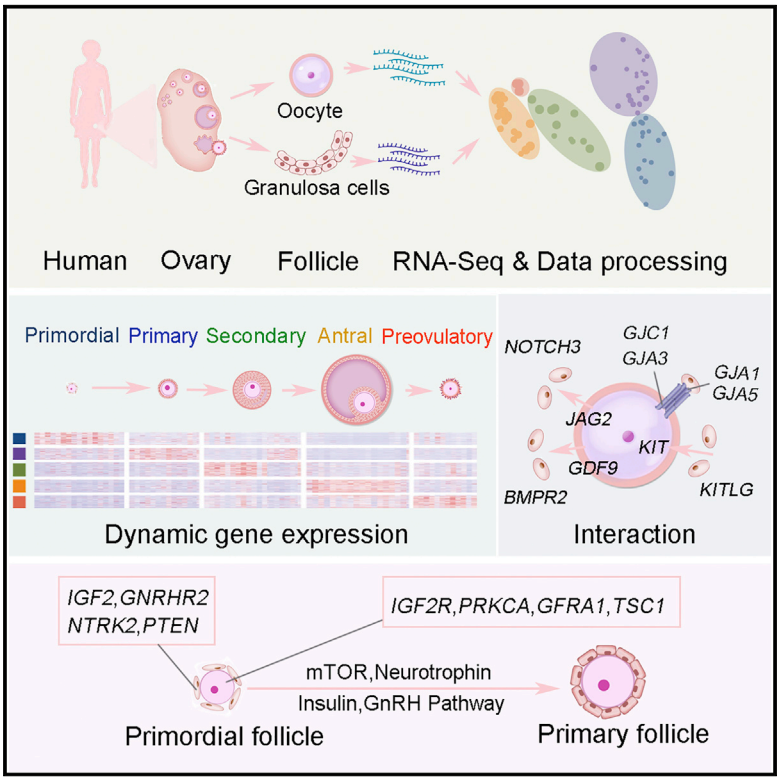
Resource

Molecular Cell

Transcriptome Landscape of Human Folliculogenesis Reveals Oocyte and Granulosa Cell Interactions

Graphical Abstract



Authors

Yaoyao Zhang, Zhiqiang Yan, Qingyuan Qin, ..., Jie Qiao, Jie Yan, Liying Yan

Correspondence

jie.qiao@263.net (J.Q.), yanjiebjmu@bjmu.edu.cn (J.Y.), yanliyingkind@aliyun.com (L.Y.)

In Brief

Zhang et al. investigated the transcriptomic profiles from both the oocyte and somatic follicular compartments during folliculogenesis using single-cell RNA-seq analysis. They provided a transcriptomic landscape of folliculogenesis as well as key insights into the transcriptional regulation in oocytes and granulosa cells during the stepwise folliculogenesis.

Highlights

- RNA-seq of oocytes and granulosa cells mapped transcriptome and signature genes
- KEGG/GSEA analysis uncovered pathways involved in primordial follicle activation
- Oocyte-granulosa cell interactions exhibit stage- and species-specific patterns
- RNA-seq analysis identified candidate secretory biomarkers of ovarian reserve

Data Resources

GSE107746

Zhang et al., 2018, Molecular Cell 72, 1021–1034
December 20, 2018
https://doi.org/10.1016/j.molcel.2018.10.029

CellPress

Molecular Cell
Resource

Transcriptome Landscape of Human Folliculogenesis Reveals Oocyte and Granulosa Cell Interactions

Yaoyao Zhang,[1,2,3,9] Zhiqiang Yan,[1,5,8,9] Qingyuan Qin,[1,3,4,9] Vicki Nisenblat,[1,4] Hsun-Ming Chang,[1,4] Yang Yu,[1,2,3] Tianren Wang,[1,2,3] Cuiling Lu,[1,2,3] Ming Yang,[1,2,3] Shuo Yang,[1,2] Ying Yao,[5] Xiaohui Zhu,[1,3,4] Xi Xia,[1,2,3] Yujiao Dang,[1,3,4] Yixin Ren,[1,3,4] Peng Yuan,[1,3,4] Rong Li,[1,2,3] Ping Liu,[1,2,3] Hongyan Guo,[5] Jinsong Han,[5] Haojie He,[5] Kun Zhang,[5] Yiting Wang,[5] Yu Wu,[5] Meng Li,[5] Jie Qiao,[1,2,3,5,6,7,10,*] Jie Yan,[1,2,3,*] and Liying Yan[1,2,3,*]

[1]Center for Reproductive Medicine, Department of Obstetrics and Gynecology, Peking University Third Hospital, No. 49 North HuaYuan Road, HaiDian District, Beijing 100191, China
[2]National Clinical Center for Obstetrics and Gynecology, Beijing 100191, China
[3]Key Laboratory of Assisted Reproduction, Ministry of Education, Beijing 100191, China
[4]Beijing Key Laboratory of Reproductive Endocrinology and Assisted Reproduction, Beijing 100191, China
[5]Peking-Tsinghua Center for Life Sciences, Peking University, Beijing 100871, China
[6]Department of Obstetrics and Gynecology, Peking University Third Hospital, No. 49 North HuaYuan Road, HaiDian District, Beijing 100191, China
[7]Beijing Advanced Innovation Center for Genomics (ICG), Peking University, Beijing 100871, China
[8]Academy for Advanced Interdisciplinary Studies, Peking University, Beijing 100871, China
[9]These authors contributed equally
[10]Lead Contact
*Correspondence: jie.qiao@263.net (J.Q.), yanjiebjmu@bjmu.edu.cn (J.Y.), yanliyingkind@aliyun.com (L.Y.)
https://doi.org/10.1016/j.molcel.2018.10.029

SUMMARY

The dynamic transcriptional regulation and interactions of human germlines and surrounding somatic cells during folliculogenesis remain unknown. Using RNA sequencing (RNA-seq) analysis of human oocytes and corresponding granulosa cells (GCs) spanning five follicular stages, we revealed unique features in transcriptional machinery, transcription factor networks, and reciprocal interactions in human oocytes and GCs that displayed developmental-stage-specific expression patterns. Notably, we identified specific gene signatures of two cell types in particular developmental stage that may reflect developmental competency and ovarian reserve. Additionally, we uncovered key pathways that may concert germline-somatic interactions and drive the transition of primordial-to-primary follicle, which represents follicle activation. Thus, our work provides key insights into the crucial features of the transcriptional regulation in the stepwise folliculogenesis and offers important clues for improving follicle recruitment *in vivo* and restoring fully competent oocytes *in vitro*.

INTRODUCTION

Human folliculogenesis is a remarkably complex, well-orchestrated process that relies on synchronization between oocyte maturation and proliferation of the neighboring granulosa cells (GCs). Follicle growth and oocyte maturation are associated with dynamic transcriptional regulation in both oocyte and GC compartments of the follicle (Sánchez and Smitz, 2012). Despite the impressive body of data produced in recent years on oocyte biology, many questions regarding the key developmental events, the activation of oocyte maturation, and the transition of primordial-to-primary follicle during folliculogenesis in humans remain unanswered. Oocyte-GC bidirectional communications via signal transduction or direct cell-to-cell contact provide the molecular and structural basis for effective oocyte-GC crosstalk, which is required for adequate follicular development. It is poorly understood, however, what signal initiates (Li and Albertini, 2013) the communication between human follicle compartments and how this interaction is regulated and maintained.

To date, most information on follicular transcriptome is obtained from studies in animal models due to the restricted availability of human follicles for research. However, human follicles have a unique developmental pattern and transcriptional regulation network that is different from mouse, and the expression patterns of germline genes are reported to display considerable interspecies variations (Knight and Glister, 2006). Characterizing human oocyte and GC transcriptome during different stages of development is the key to understanding the molecular interactions that coordinate human oocyte maturation and follicle growth, which in turn is expected to provide remarkable opportunities for developing novel diagnostic and therapeutic approaches for improving fertility. This work aimed at analyzing the gene expression dynamics throughout folliculogenesis by exploring the transcriptomes of the human oocyte and GCs at five key stages of follicular development *in vivo*, from primordial to preovulatory stage, which could offer a solid reference dataset for understanding transcriptional regulation of folliculogenesis.

RESULTS

Global Transcriptome Profiling of Human Oocytes and GCs

With ethics approval, we obtained a total of 83 oocytes and 92 GC samples (each GC sample comprised of randomly selected 10 cells per follicle, owing to low abundancy of RNA in these cells) to systemically investigate the human follicular development and the regulatory relationships between the oocyte and its neighboring follicle cells (Figures 1A, S1A, and S1B). We performed RNA sequencing (RNA-seq) of these individual cells. For the subsequent analyses, we downloaded the data of metaphase II (MII) oocytes from our previous research (Li et al., 2013) and merged with the newly generated data. After quality control and filtration of the RNA-seq data, 80 oocytes and 71 GC samples were retained for the analysis (Figure S1C). A down sampling analysis revealed an adequate sequencing depth for gene expression detection and a subsequent reliable analysis (Figure S1D).

We identified an average of 9,881 genes (fragments per kilobase per million mapped reads [FPKM] > 1) per oocyte and 7,660 genes per GC sample (Figure S1E). In total, there were 21,976 expressed genes with 20,100 expressed in oocytes and 19,537 expressed in GCs. There were 333 genes expressed in all the samples (oocytes and GCs) and were considered as ubiquitously expressed genes (FPKM > 1 in all samples). By utilizing the GENCODE database, we analyzed the expression of long non-coding RNA (lncRNA) at all stages of folliculogenesis. There is an average of 2,752 lncRNAs (FPKM > 0.1) in each oocyte and 1,881 lncRNAs (FPKM > 0.1) in each GC sample (Table S1). Both oocytes and GCs from secondary follicles expressed most abundant amount of lncRNAs on average among all follicular stages.

To explore the variance in the mRNA dataset, the obtained RNA-seq normalized data (FPKM, log) were subjected to principal component analysis (PCA) by using an unsupervised approach. The first principal component (PC1) was dominated by differences between oocytes and GCs, and the second principal component (PC2) separated the samples according to their developmental stages. There was a clear separation between 80 oocytes and 71 GCs with distinction between the developmental stages (Figure 1B). Interestingly, we observed distinct stage-specific clustering of oocytes, whereas the clustering of GCs showed some overlap between the stages of follicular development. This suggested that oocytes and GCs have distinct gene expression profiles with a difference in the dynamics of transcription between the two cell types.

Next, we analyzed the expression patterns of the known cell-type-specific markers for oocyte (*ZP2*, *DDX4*, *SYCP3*, *SOX30*, *ZAR1*, *DAZL*, *YBX2*, and *LHX8*; Dean, 2002; Meczekalski, 2009; Zheng and Dean, 2007; Figure 1C) and GC (*CYP11A1*, *STAR*, *INHBA*, and *AMH*; Hatzirodos et al., 2014; Figure 1D), which confirmed the validity of the PCA classification. Then, we analyzed the gene expression patterns of the oocyte and GC clusters and selected the most variable genes that contributed to PC1, which could be used to distinguish these two cell types. Notably, the most variable genes that contributed to PC1 in oocytes were *NOL4*, *CTCTF*, *ITIH2*, *SNAP91*, and *NAALAD2*, with *NOL4* and *CTCTF* as the two top variable genes (Figure 1E). On the other hand, the most variable genes that contributed to PC1 in GCs were *ZEB2*, *CD44*, *HSPG2*, *KDSR*, and *SRRM3*, with *ZEB2* and *CD44* as the two top variable genes (Figure 1F). These most variable genes derived from oocyte and GCs could be potentially used as the candidate cell-type-specific markers.

Gene Expression Dynamics and Transcriptional Characteristics of Oocytes during Folliculogenesis

To explore the gene expression dynamics in oocytes during folliculogenesis, we performed PCA within the cohort of oocytes and observed five distinct subpopulations, which corresponded to the morphological classification of follicular development (Figure 2A). As shown in Figure S1F, the stage of follicular development is a much more powerful discriminator than the sampling entity (inter-subject variability), indicating that transcriptome profiles of oocytes reflect the physiological status of maturation rather than the individual genetic background. We then characterized the differentially expressed genes (DEGs) between the stages of follicular development in oocytes (false-positive discovery rate [FDR] < 0.05; fold change [FC] of log2 transformed FPKM > 1.5; Figure 2B). Considering the volume of oocytes is expanding dramatically with a substantial increase in mRNA content during follicular development, the DEGs may reflect changes in transcript relative abundance, and only some of them might correspond to the changes in transcriptional status.

To further validate the information obtained from the DEG expression profiles, we performed immunohistochemistry (IHC) of the ovarian tissues from the RNA-seq samples to examine the expression patterns of *RBM24*, *GPD1*, *NTF4*, and *LCP2*. In agreement with the results from the RNA-seq analysis, the IHC results showed that *RBM24* was expressed only at the primordial and primary stages. *GPD1* was primarily expressed at the secondary and antral stages. *NTF4* and *LCP2* were specifically expressed at antral stage (Figure 2C). To investigate the biological functions involved in folliculogenesis, we performed Gene Ontology (GO) analysis of DEGs in oocytes. The analysis revealed several significantly enriched biological processes in association with each stage of follicular development ($p < 0.05$; Figure 2D).

We also investigated a number of oocyte markers, which exhibited stable expression levels in oocytes, such as *DDX4* and *ZP3* during follicular development (Figure S2A). On the other hand, some oocyte markers, such as *ZP1*, *GDF9*, and *H1FOO*, were progressively upregulated with oocyte maturation (Figure S2A). Notably, the expression levels of *ZP2*, *ZP3*, and *ZP4* remained stable across all stages (Figure S2A). In contrast to the high expression levels of *NANOG* and *POU5F1* in primordial germ cells (PGCs) (Guo et al., 2015; Li et al., 2017), the pluripotency markers *NANOG* and *POU5F1* were under-expressed in most stages in oocytes (Figure S2A).

Next, we investigated the expression patterns of "meiosis-related genes" in oocytes. The sets of 195 genes involved in meiosis I and 13 genes involved in meiosis II in human were gathered from the GO Consortium. Among these meiosis-related genes, 44 genes were differentially expressed across different stages of oocyte maturation. Most of these genes demonstrated

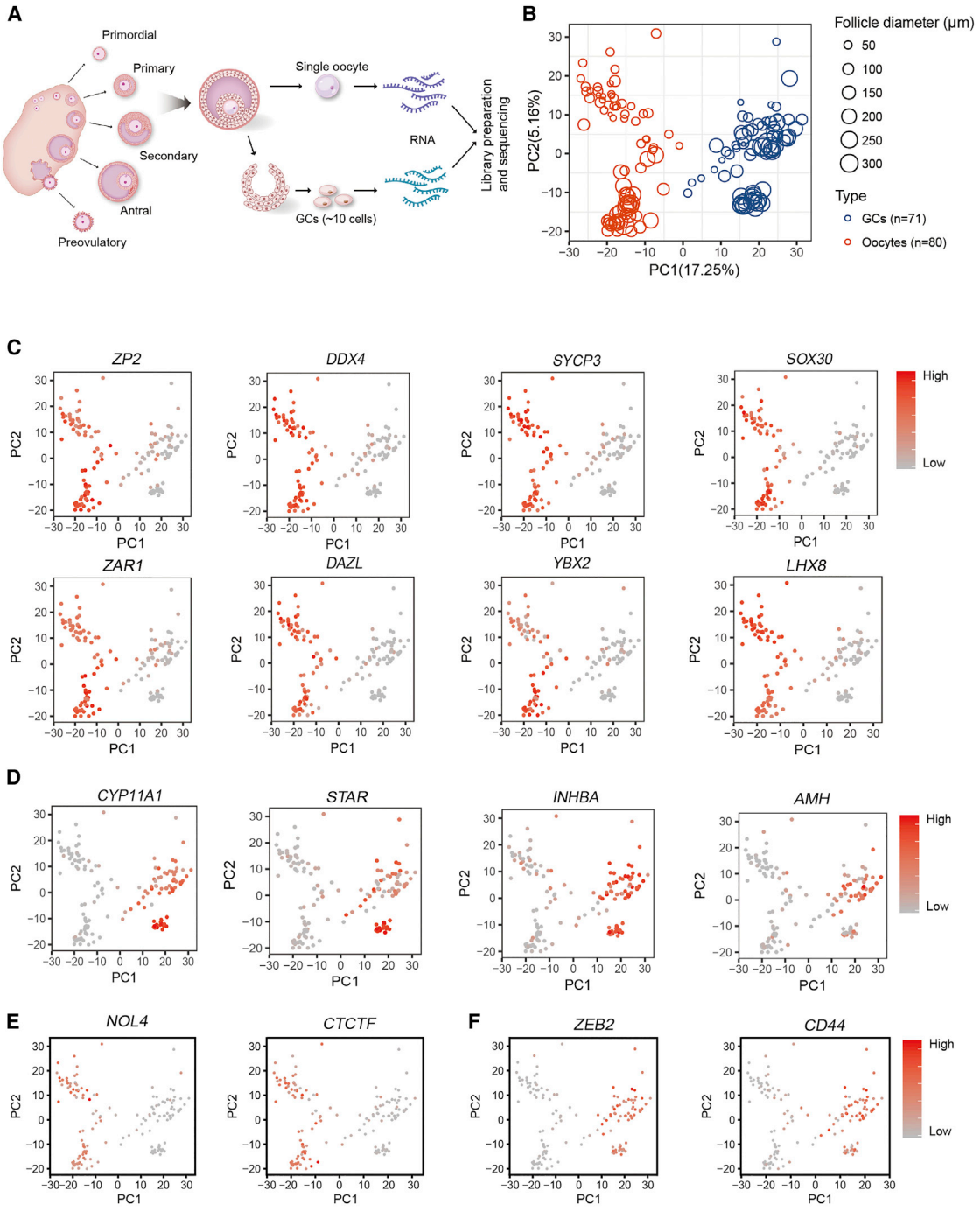


Figure 1. The Global Transcriptome Patterns of Human Oocytes and GCs

(A) Schematic illustration of the study workflow. antral, antral follicles; preovulatory, preovulatory follicles; primary, primary follicles; primordial, primordial follicles; secondary, secondary follicles.

(B) Principal component analysis (PCA) of the transcriptome of RNA-seq data from all oocytes and GCs included in this study. The PC1 separates two follicular compartments (oocytes versus GCs). The PC2 separates the samples according to their follicular stages. Red color indicates oocytes, and blue color indicates GCs. The continuous sizes of the points represent follicles with different sizes.

(C) Expression patterns of oocyte marker genes exhibited on PCA plots; a gradient of gray to red indicates the low to high gene expression level.

(D) Expression patterns of GC marker genes exhibited on PCA plots.

(E) Expression of the candidate cell-type-specific markers of oocytes on PCA plots.

(F) Expression of the candidate cell-type-specific markers of GCs on PCA plots.

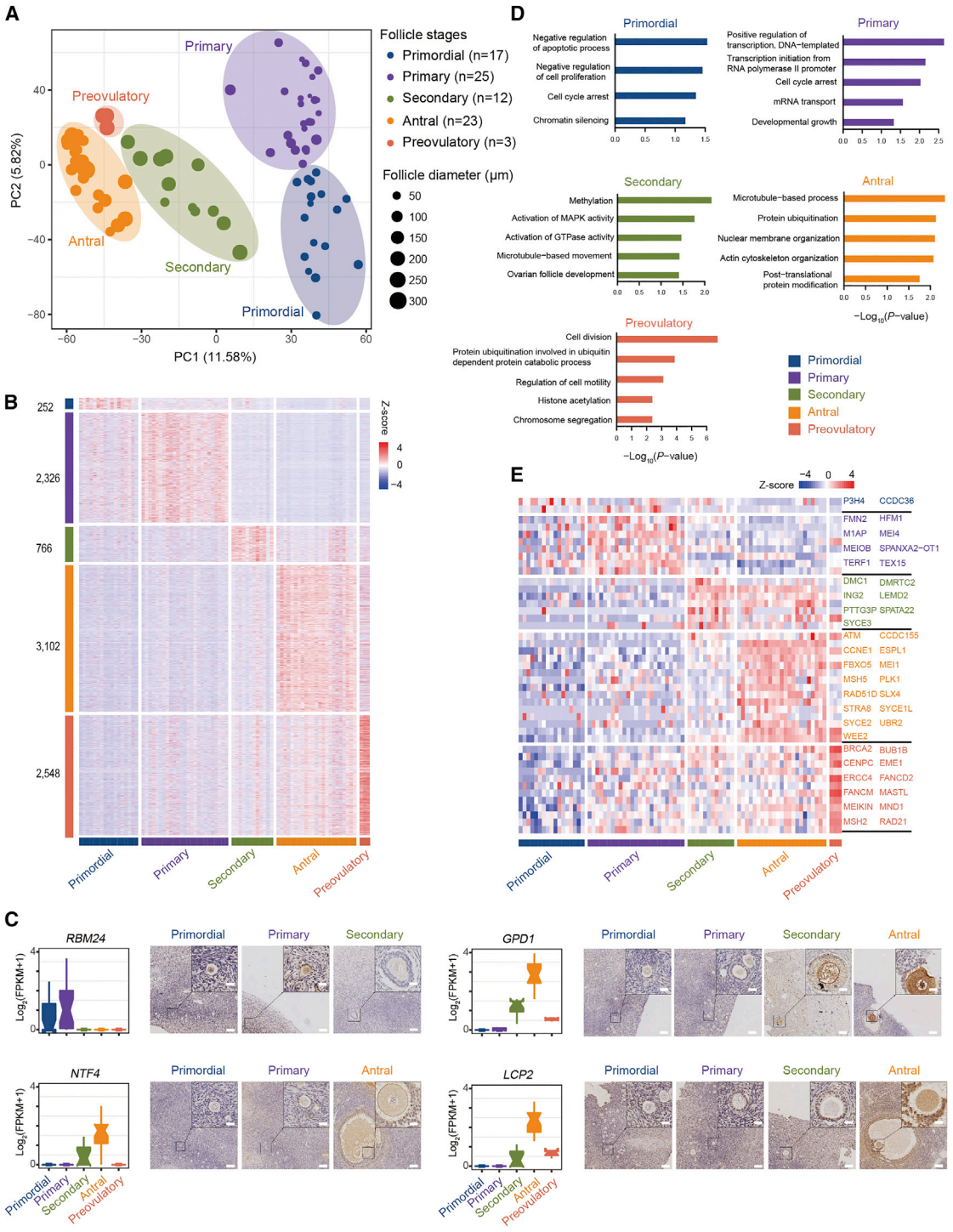


Figure 2. Gene Expression Dynamics and Transcriptional Characteristics of Oocytes

(A) PCA of the transcriptome of RNA-seq data from 80 oocytes collected from follicles at different stages of development. Oocytes are clustered into five subpopulations corresponding to the morphological stages. Different colors and sizes of the points represent follicles of different stages and sizes, respectively. See also Figure S1F for the results of PCA that portrays sampling entities within the oocytes.

(B) Heatmap of all the differentially expressed genes (DEGs) in oocytes at five stages of folliculogenesis. The numbers of identified DEGs are indicated on the y axis, and the stages of follicular development are presented along with the x axis. The color key from blue to red indicates the relative gene expression level from low to high, respectively.

(legend continued on next page)

a shift toward upregulation as follicular development proceeds, with the highest expression at both antral and preovulatory stages (Figure 2E). In addition, we analyzed the expression of 38 genes involved in chiasma formation (Youds and Boulton, 2011) during oocyte maturation. The results showed that some of these genes (including *SYCP2*, *SYCP3*, *MSH2*, and *RAD50*) showed high expression levels during all stages, and some genes (such as *FANCD2*, *MRE11A*, and *FANCM*) exhibited a gradual increase during oocyte development. Our data suggest that majority of these genes involved in chiasma formation are expressed during human folliculogenesis (Figure S3A).

There were relatively high expression levels of *DNMT1*, *DNMT3A*, and *DNMT3B* in the oocytes of all stages with mounting abundance (Figure S2A), indicating that the degree of DNA methylation is continually increasing with oocyte maturation. In contrast to our previous studies in PGCs (Guo et al., 2015), our data showed that ten-eleven translocation (TET) family genes were under-expressed in oocytes at all stages of follicular development (Figure S2A). Taken together, the overexpression of DNMT genes in association with the loss of TET activity suggests an active methylation transition in human oocyte during folliculogenesis.

To explore the global expression profiles of maternal-effect genes, we integrated the oocyte transcriptome data derived from this study with the transcriptome of human preimplantation embryos that had been identified in our previous work (Li et al., 2013). We searched for the genes that are expressed in the MII oocytes and zygotes, but not in 8-cell embryos, as these genes are carried over from the oocyte to early embryo and then are degraded following zygotic genome activation (ZGA). Overall, we identified 1,785 putative maternal-effect genes. The unsupervised cluster analysis identified four clusters of genes according to the expression patterns (Figure S3B; Table S2). The genes in cluster 1 (n = 47) and cluster 2 (n = 588) were under-expressed in the primordial stage and were upregulated during advanced stages of follicular maturation, with a peak expression level in the MII oocytes. However, the expression patterns of the gene in cluster 3 (n = 987) and cluster 4 (n = 163) showed that these transcripts could accumulate before primordial follicle assembly. The GO analysis revealed that genes that are related to biological functions associated with the meiosis process were over-represented in cluster 2. While in cluster 3 and 4, genes were enriched in ubiquitous functions (Figure S3C).

Gene Expression Dynamics and Transcriptional Profiles of GCs during Folliculogenesis

In GCs, the PCA of all samples showed clustering into five groups according to the stages of follicular development. As shown in Figure 3A, an overlap observed between the clusters of primary and secondary follicles indicated a less pronounced difference between the transcriptome levels of these two stages. We then examined the DEGs (FDR < 0.05; FC of log2 transformed FPKM > 1.5) in GCs (Figure 3B). Interestingly, some of the DEGs have been previously reported to be associated with oocyte competence (Huang and Wells, 2010; Parks et al., 2016; Table S3). Consistent with the results of RNA-seq, the IHC results showed a similar expression pattern of *CDCA3*, *BNIPL*, and *TST* on the ovarian tissues obtained from the RNA-seq sample set (Figure 3C). GO analysis of annotated genes revealed that these DEGs were involved in several crucial functions during folliculogenesis (Figure 3D).

The expression levels of 243 cell-cycle genes involved in the G1/M and G2/S phases (Tirosh et al., 2016) in GCs suggested that GCs from each follicular stage exhibited distinct expression patterns, which is similar with previous findings in animal studies (Douville and Sirard, 2014; Girard et al., 2015). These genes are relatively under-expressed with a low proliferative activity at the primordial stage. Then they were upregulated from primary follicle and reached a high expression level in antral follicle. In preovulatory follicles, most of these genes were downregulated in GCs (Figure 3E).

Next, we focused on genes that have been previously reported to be associated with steroidogenesis (Figure S2B). Several genes that encode steroidogenic enzymes, including *CYP11A1*, *CYP19A1*, *HSD17B1*, and *HSD3B2*, were upregulated in the antral follicles, with a peak expression level in the preovulatory follicles. The expression patterns of these genes indicate a progressive increase in steroidogenic activity in the mature follicles, which culminates before ovulation. *NR5A1*, a key regulator of the steroidogenic enzyme-encoding genes that has been previously described in sheep GCs (Bonnet et al., 2013), exhibited an upregulation with the progression of follicular growth. *AR*, *ESR* (*ESR1* and *ESR2*), and *PGR*, encoding for androgen, estrogen, and progesterone hormone receptors, respectively, were expressed in GCs from primary follicles, with the highest expression level in antral follicles. Consistent with previous animal studies (Cheng et al., 2002; Hu et al., 2004; Robker et al., 2000; Rosenfeld et al., 2001), these findings suggest that androgen, estrogen, and progesterone receptors are involved in the control of human follicular development.

Stage-Specific Signature Genes Identified in Oocytes and GCs

Notably, we identified a subset of stage-specific genes for human oocytes and GCs that were exclusively pertained to the individual stage of folliculogenesis. These stage-specific genes were referred as the "signature genes." Conjointly, 382 signature genes were identified, comprising the genes expressed in the oocytes from each follicular stage (Figures S4A and S4B; Table S4). The distinct stage-specific expression patterns of *NTF4* and *LCP2* in oocytes were further confirmed by the IHC staining (Figure 2C), showing that these genes were preferentially expressed in the oocyte of antral stage. Of note, *NTF4* was proposed to facilitate follicular development by inducing *FSH* receptor (*Fshr*) in mouse (Kerr et al., 2009). Consistent with the expression pattern of *NTF4*, *FSHR* gene expression

(C) Immunohistochemistry staining of selected DEGs. The boxplot demonstrating gene expression level is presented on the left of each corresponding immunohistochemistry panel. The scale bars represent 100 μm in low-magnification view (×100) and 25 μm in high-magnification view (×400).
(D) Significantly enriched GO terms (biological processes) of DEGs in the oocytes at five stages of folliculogenesis.
(E) Heatmap of genes involved in meiosis that are differentially expressed in the oocytes at five stages of folliculogenesis.

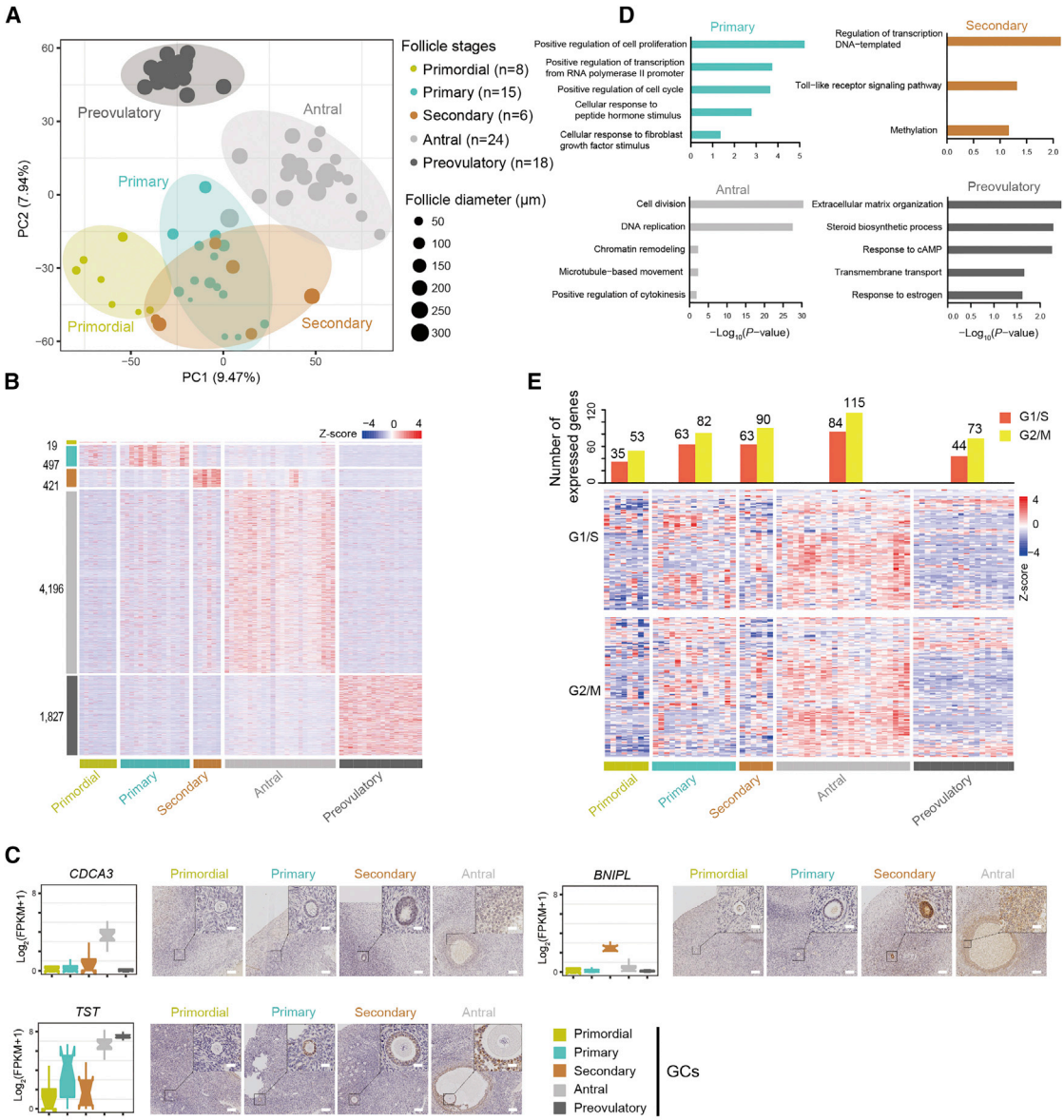


Figure 3. Dynamic Gene Expression Patterns and Transcriptional Features in GCs

(A) PCA of the transcriptome of RNA-seq data from 71 GCs collected from follicles at different stages of development. GCs are clustered into five subpopulations. Different colors and sizes of the points represent follicles of different stages and sizes, respectively.

(B) Heatmap of all the DEGs in GCs at five stages of folliculogenesis. The numbers of identified DEGs are indicated on the y axis, and the stages of follicular development are presented along with the x axis. The color key from blue to red indicates the relative gene expression level from low to high, respectively.

(C) Immunohistochemistry staining of selected DEGs. The boxplot demonstrating gene expression level is presented on the left of each corresponding immunohistochemistry panel. The scale bars represent 100 μm in low-magnification view (×100) and 25 μm in high-magnification view (×400).

(D) Significantly enriched GO terms (biological processes) of DEGs in GCs at five stages of folliculogenesis.

(E) Heatmap of cell-cycle-related genes that are expressed in GCs at five stages of folliculogenesis. The number of expressed cell-cycle-related genes in each stage was showed at the top.

was pertained to the GCs of antral follicles (Figure S4D). Our data suggest that *NTF4* might upregulate expression of *FSHR* in human GCs. Furthermore, we also identified a subset of stage-specific signature genes for GCs during folliculogenesis (Figures S4C and S4D; Table S4). Among these GC signature genes, *CDCA3* and *BNIPL* were validated by the IHC staining (Figure 3C). Interestingly, *FSHR*, previously observed in somatic cells from early preantral to antral follicles in sheep (Tisdall et al., 1995), was preferentially expressed in the GCs of antral follicles in this study (Figure S4D). The full list of the signature genes for both oocytes and GCs is presented in Table S4. Notably, these identified signature genes can be potentially used as

candidate cell-specific markers for each follicle stage. Most importantly, the subsets of 156 signature genes of oocytes in preovulatory stage and 175 signature genes of GCs in preovulatory stage could be served as markers of oocyte maturation and developmental competency.

Secretory Protein-Coding Genes throughout Folliculogenesis for Predicting Ovarian Reserve

We searched for the secretory protein-coding genes that were exclusively pertained to either oocytes or GCs and were expressed either at early or at more advanced stages of follicular development (as described in the STAR Methods). In total, we identified 61 protein-coding genes in oocytes and 16 genes in GCs that can be clustered into five groups (Figure S5A). Cluster 1 comprised 16 oocyte genes with specific expression in the oocytes of primordial and primary follicles, suggesting that their expression level reflect the resting follicle pool. Cluster 2 included 18 oocyte genes with specific expression in the oocytes obtained from secondary, antral, and preovulatory follicles, suggesting that their expression level indicates the growing follicle numbers. Cluster 3 consisted of 27 oocyte genes that were represented across the folliculogenesis. Cluster 4 (n = 4) and cluster 5 (n = 12) included GC-derived genes with specific expression at preovulatory and secondary-antral stages, respectively. The expression patterns of the genes from each cluster are exemplified in Figure S5B. These findings may provide subsets of candidate molecular biomarkers that predict the ovarian reserve. The subsets of genes in cluster 1 and cluster 2 may indicate resting and growing follicle numbers, respectively. Further evaluation of their expression levels in the biological fluids may contribute to the development of novel biomarkers for ovarian reserve in clinical application.

Transcription Factor Regulatory Networks in the Oocytes and GCs

To investigate the master regulators and construct the transcriptional regulatory network along the steps of human folliculogenesis, we utilized the ARACNe method to analyze all 1,469 known transcription factors (TFs) from Animal Transcription Factor Database (TFDB v2.0; Margolin et al., 2006; Zhang et al., 2012). In the oocytes, the expressions of *GTF2I*, *CSDE1*, *SOHLH2*, *SMARCE1*, *TUB*, *HBP1*, *SOX30*, and *HIF1A* were upregulated in primary follicles, indicating that these TFs may play a critical role in the transition from the primordial to the primary stage (Figure 4). *KLF2*, *YBX2*, *FOXO6*, *SOX13*, *ETV5*, *TEAD2*, and *OTX2* were overexpressed in the oocytes from secondary follicles compared to those from primary follicles, implying that they are likely the regulators of the primary-to-secondary stage transition. *PINX1*, *PBX1*, *MTF1*, *SOX15*, *UBTF*, *SOX13*, and *POU2F1* had higher expression levels in the oocytes of antral follicles compared to those of secondary follicles, indicating possible regulatory roles in the cytoplasmic and nuclear maturation of oocyte at the antral stage. *ATF2* and *EOMES* were abundantly expressed in the MII oocytes of preovulatory follicles, indicating their potential roles in the unique transcription networks. Of note, we found that *SOX13* and *SOX15*, members of the SOX family of TFs, were upregulated in the oocytes from secondary to antral stage, and *SOX30* was abundantly expressed in the oocytes during the transition from primordial to primary stage. These findings may suggest potential roles of *SOX30* (primordial follicle activation) as well as *SOX13* and *SOX15* (antral formation) in regulating transcription networks during follicular development.

In the GC-expressed TFs, we speculated that *CREB1*, *NFKB1*, *MEF2A*, *PIAS1*, *FOSL2*, *KLF13*, and *PRDM4* might potentially regulate the activation of primordial to primary follicle, whereas *CRX*, *HES2*, *ZNF554*, *ZKSCAN3*, *LMX1B*, *FOXK1*, and *SIX4* might be involved in the transition from primary to secondary follicle. Additionally, *MBD1*, *FIZ1*, *GABPA*, *TGIF2*, *PIAS3*, *E4F1*, and *IRF3* most likely drive the initiation of the transcription network in antral follicles. Furthermore, *MEIS3*, *PRDM15*, and *VTN* that were expressed in GCs of preovulatory follicles could be the key regulators of cumulus cells progression in preovulatory follicles (Figure S6). Further evaluation of these TFs, the upstream signaling pathways, ligands, receptors, and downstream targets in oocytes and GCs will provide insight into the transcriptional control of human folliculogenesis.

Key Pathways and Interactions between the Oocyte and GC Compartments throughout Folliculogenesis

The primordial-to-primary transition is a key step in follicle activation. However, the molecular mechanism and signaling pathways that drive this transition are unclear. Therefore, the gene set enrichment analysis (GSEA) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis were applied to perform the pairwise comparisons of the follicular stages associated with the primordial-to-primary transition (Figure 5; Table S5). Notably, the functional pathways that were significantly overrepresented both in oocytes and GCs ($p < 0.05$) included insulin, gonadotropin-releasing hormone (GnRH), neurotrophin, and mammalian target of rapamycin (mTOR)-phosphatidylinositol 3-kinase (PI3K) signaling pathways. For the insulin signaling pathway, *IGF2R* and *IGF2* were significantly upregulated in the oocyte and GCs of primary stage, respectively. For the GnRH signaling pathway, *PRKCA* and *GNRHR2* were upregulated in the oocyte and GCs of primary stage, respectively. For the neurotrophin signaling pathway, *GFRA1* and *NTRK2* were upregulated in the oocyte and GCs of primary stage, respectively. For the mTOR-PI3K signaling pathway, *TSC1* and *PTEN* were upregulated in the oocyte and GCs of primary stage, respectively. In addition, the JAK-STAT signaling pathway was overrepresented in GCs of the primary stage and the *JAK1* of JAK-STAT pathway was upregulated. These intra-follicular signaling pathways most likely mediate the transition from primordial to primary follicular stage, which concert interaction of the two follicular compartments.

To investigate the interactions between the oocyte and GCs throughout folliculogenesis, we analyzed the expression of the components (ligands, receptors, and target genes) of key cell signaling pathways, including NOTCH, transforming growth factor β (TGF-β), KITLG-KIT signaling pathway, and gap junctions. The results showed that ligands *DLL3* and *JAG2* of NOTCH signaling were predominantly expressed in the oocytes before the preovulatory stage. Their receptors *NOTCH2* and *NOTCH3*

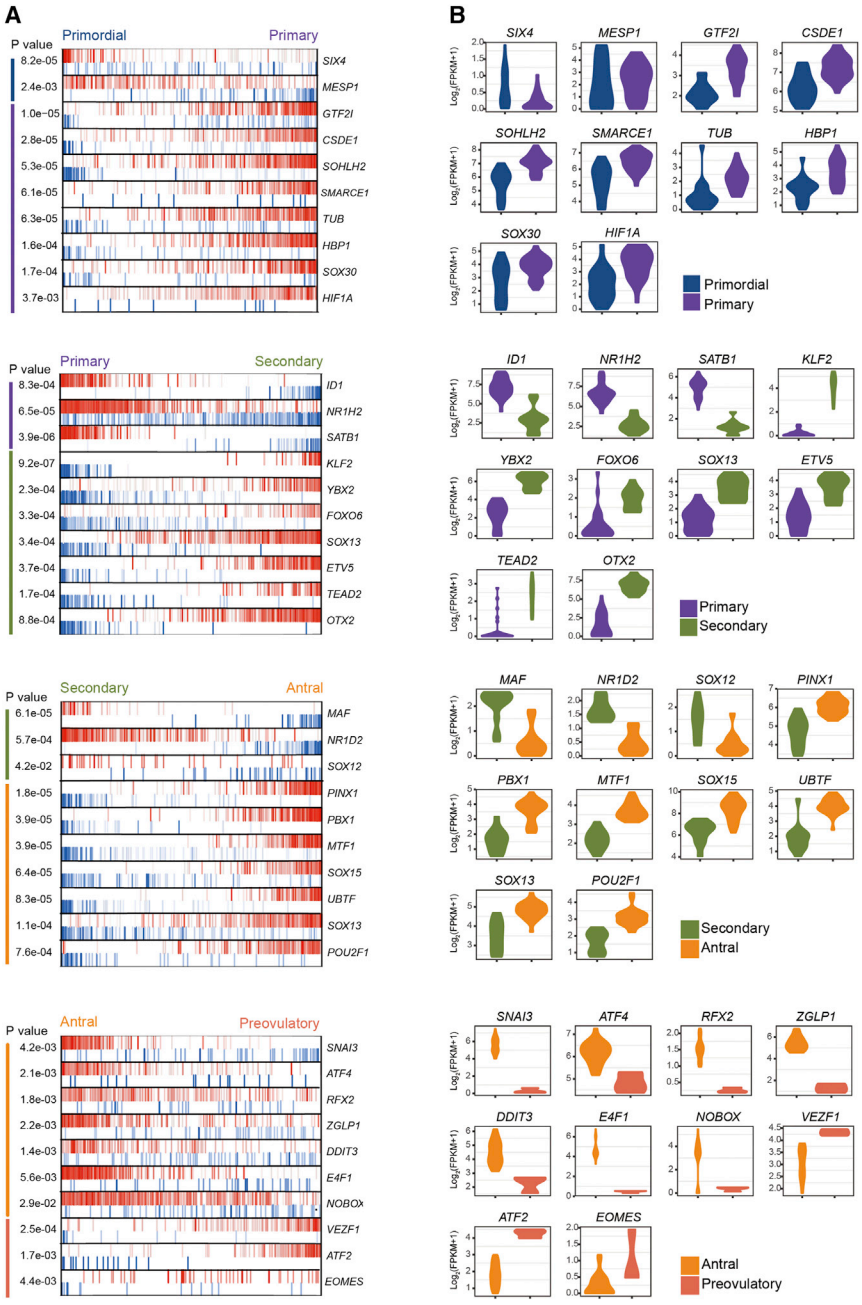


Figure 4. Inferred Key Transcriptional Factors in Oocytes at Each Stage-to-Stage Transition of Folliculogenesis

(A) MARINa plots of targets for each candidate master regulator. Red vertical bar represents the activated targets; blue vertical bar represents the repressed targets. On the x axis, genes are rank sorted according to the significance of differential expression between the two developmental stages. The candidate master regulators are displayed on the right. The corresponding p values of these master regulators indicating the significance of enrichment are displayed on the left. See also Figure S6 for the analogous information in GCs.

(B) Violin plots show the relative expression levels (log2 [FPKM+1]) of each master regulator in two consecutive stages. See also Figure S6 for the analogous information in GCs.

and the downstream target gene *HES1* were highly expressed in GCs (Figure 6A). These findings have highlighted the role of NOTCH signaling in the control of oocyte-mediated GC proliferation and differentiation.

Next, we analyzed the expression of the key components of the TGF-β signaling pathway in the oocytes and GCs. Our data showed that the expression levels of *GDF9* in oocytes were consistently high at all stages of follicular development, whereas *BMP15* was highly expressed only in the oocytes of antral and preovulatory follicles (Figure 6B). This observation points out that *GDF9* may maintain its function throughout the folliculogenesis, whereas *BMP15* mostly exerts its cellular action at more advanced stages. In contrast, studies in animal have shown that the activation of the primordial follicles is mediated by *BMP15* (Kim, 2012), suggesting that two oocyte-derived TGF-β superfamily members may play different roles in the development of follicles across the species. Interestingly, BMP type II receptor gene *BMPR2* and its target *ID3* were expressed in both oocytes and GCs. The presence of the gene and its target in two different cell types, herein, allows us to assume existence of both autocrine and paracrine mechanisms underlying the BMP signaling. Consistent with the previous findings in animal models (Thomas and Vanderhyden, 2006), we identified that *KITLG* and its receptor *KIT*, previously implicated in paracrine signaling in folliculogenesis, were expressed in GCs and oocytes, respectively (Figure S2B). Together, our findings provide evidence that human folliculogenesis is coordinated by both autocrine and paracrine signaling pathways that could be initiated by either the oocytes or GCs.

To understand the characteristics of gap junctions in human follicles, we analyzed the connexin-encoding genes that were previously reported in the mammalian ovary (Gershon et al., 2008). Four of these connexin genes exhibited a compartment-specific pattern across the follicle stages (Figure 6C). We found that *GJA3* (Cx46) was preferentially expressed in the oocytes of primordial and primary follicles and *GJC1* (Cx45) was highly expressed in the oocytes from the secondary to pre-ovulatory follicles. *GJA1* (Cx43) was mainly expressed in the GCs throughout all the follicular stages, and *GJA5* (Cx40) was pertained to the GCs of antral and preovulatory follicles (Figure 6C). The IHC

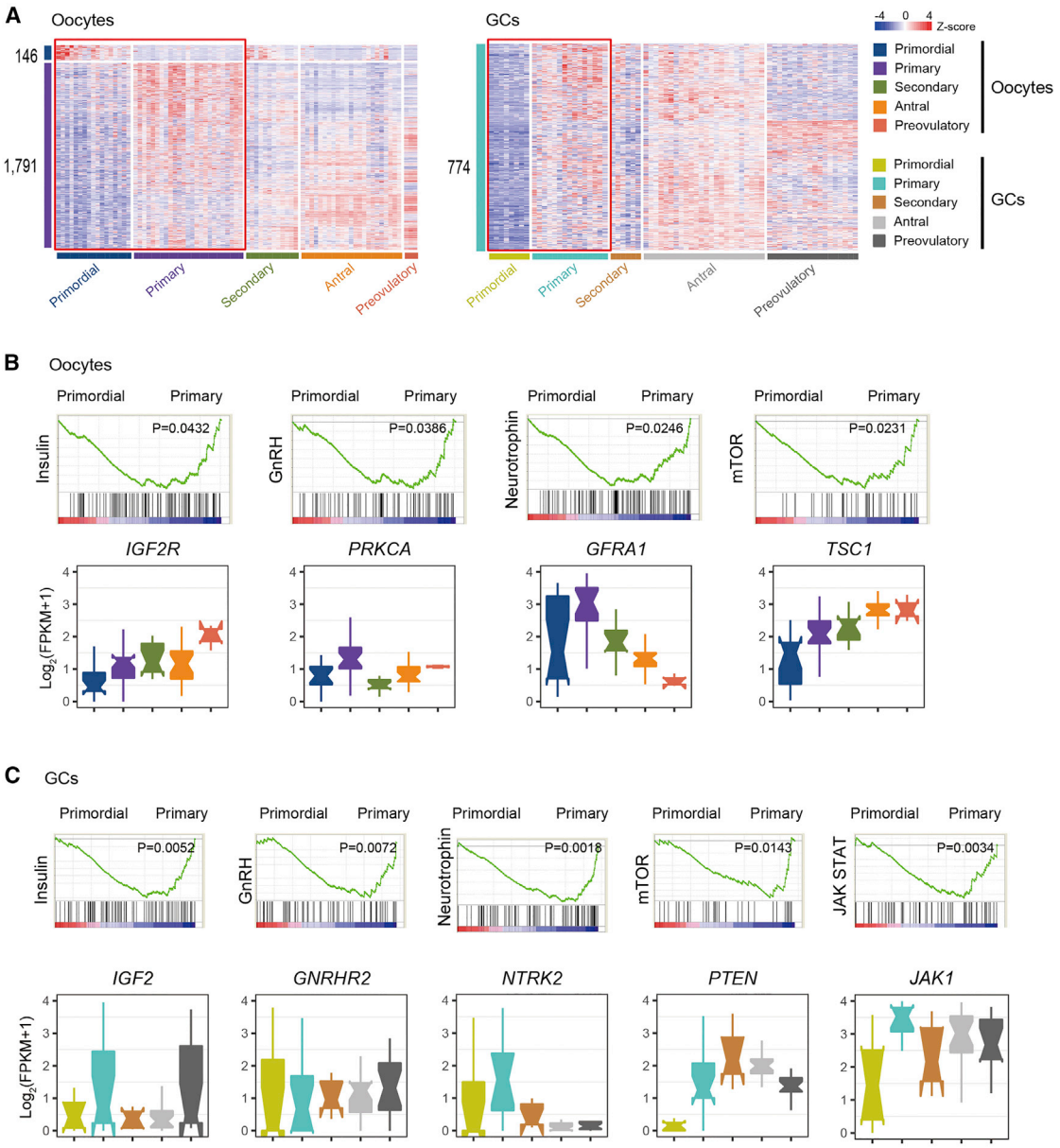


Figure 5. Signaling Pathways Enriched in Follicle Recruitment by GSEA/KEGG Analysis

(A) Heatmaps of DEGs between primordial and primary stage in oocytes and GCs. The numbers of identified DEGs are indicated on the y axis, and the stages of follicular development are presented along with the x axis. The color key from blue to red indicates the relative gene expression level from low to high, respectively.

(B) GSEA enrichment plots of KEGG signaling pathways and boxplots of key component genes in oocytes between the primordial and primary stage.

(C) GSEA enrichment plots of KEGG signaling pathways and boxplots of key component genes in GCs between the primordial and primary stage.

staining showed that the expression of the connexin proteins were largely consistent with the RNA-seq results (Figure 6C).

To explore the oocyte-GC interaction more precisely, we have analyzed the matched oocytes and GC samples (Figure S7) to illustrate the matched oocyte-GCs from the same follicle in each stage. In total, there are 20 pairs of oocytes and GCs from primordial to antral stage. As shown in Figure S7, the expression patterns of those ligand and receptors in the matched samples are concordant with the expression patterns of all samples (both matched and unmatched samples).

Comparison of Human and Mouse Oocyte Transcriptomic Profiles during Folliculogenesis

To better understand the interspecies complexity of oocyte transcriptome in human and mouse, we compared our findings in human oocytes with the reported RNA-seq datasets in mouse

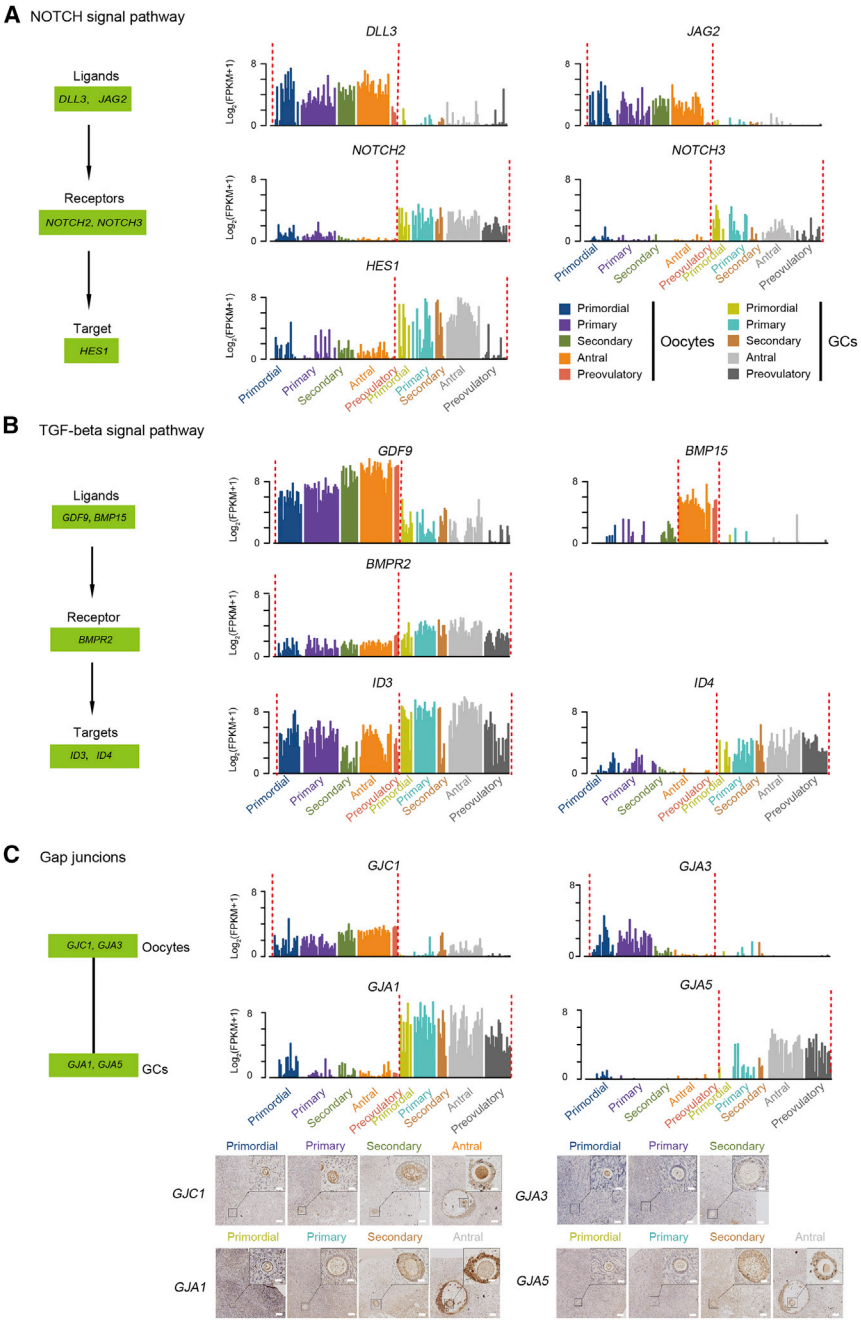


Figure 6. Signaling Pathways and Gap Junction Involved in Oocyte-GC Crosstalk throughout Folliculogenesis

(A) NOTCH signaling pathway is involved in the oocyte-GC crosstalk in folliculogenesis. The relative expression levels (log2 [FPKM+1]) of the specific ligands, receptors, and target genes are shown. The diagrams at the left show the relationship among these genes.

(B) The TGF-β signaling pathway is involved in the oocyte-GC crosstalk.

(C) The gap junction is involved in the oocyte-GC crosstalk. The relative expression levels (log2 [FPKM+1]) of the connexin genes are shown. The diagrams at the left show the putative relationship among these genes. The immunohistochemistry staining of the four genes were shown. The scale bars represent 100 μm in low-magnification view (×100) and 25 μm in high-magnification view (×400).

(Gahurova et al., 2017). Our single-cell oocyte data were merged by each stage in order to compare with the bulk RNA-seq data of mouse oocytes. The standardized analytical approach to the merged raw data revealed a total of 13,973 genes expressed in human (our data) and 11,585 genes expressed in mouse oocytes (Figure 7A). We focused on analyzing 16,175 one-to-one homologous genes shared by humans and mouse from our data and Gahurova's study. The homologous genes relationships were obtained from Mouse Genome Informatics (MGI) database (Blake et al., 2017). Among these homologous genes, there were 8,838 genes overlapped between human and mouse oocyte samples. Among these overlapped genes, we found that there were 5,923 house-keeping genes, 1,557 consistently expressed genes, and 1,358 DEGs between different stages of folliculogenesis (Figure 7A). The GO analysis of the co-expressed DEGs revealed that these genes exhibited ubiquitous functions, and none were specific for oocyte development (Figure 7B), suggesting that distinct molecular mechanisms underlying the modulation of folliculogenesis exist in humans. Moreover, several homologous DEGs that have been implicated in oocyte development, such as *POU5F1*, *GJC1*, *TEAD2*, and *BMP15*, showed different expression patterns in human and mouse oocytes (Figure 7C). Because the number of human oocytes in each stage is still far less than the mouse oocytes from Gahurova's study, some of the non-overlap could be due to technical or bioinformatics differences between these two approaches.

Next, we assumed that oocyte genes with a concordant expression pattern and a high degree of correlation ($r > 0.8$) are involved in conserved molecular mechanisms in human and mouse oocytes. Among co-expressed DEGs with significant correlated expression, four highlighted co-expressed genes with equivalent expression patterns in human and mouse oocytes (*HAS3*, *HTR5A*, *SNPH*, and *NRAP*) exhibited upregulation with progression of oocyte, suggesting a possibility for their homologous function during oocyte maturation (Figure 7D). In the consistently expressed genes, the homologous genes (*ZP3*, *RAC1*, and *PATL2*), annotated in association with "oocyte maturation defect" in OMIM database, showed stable expression across

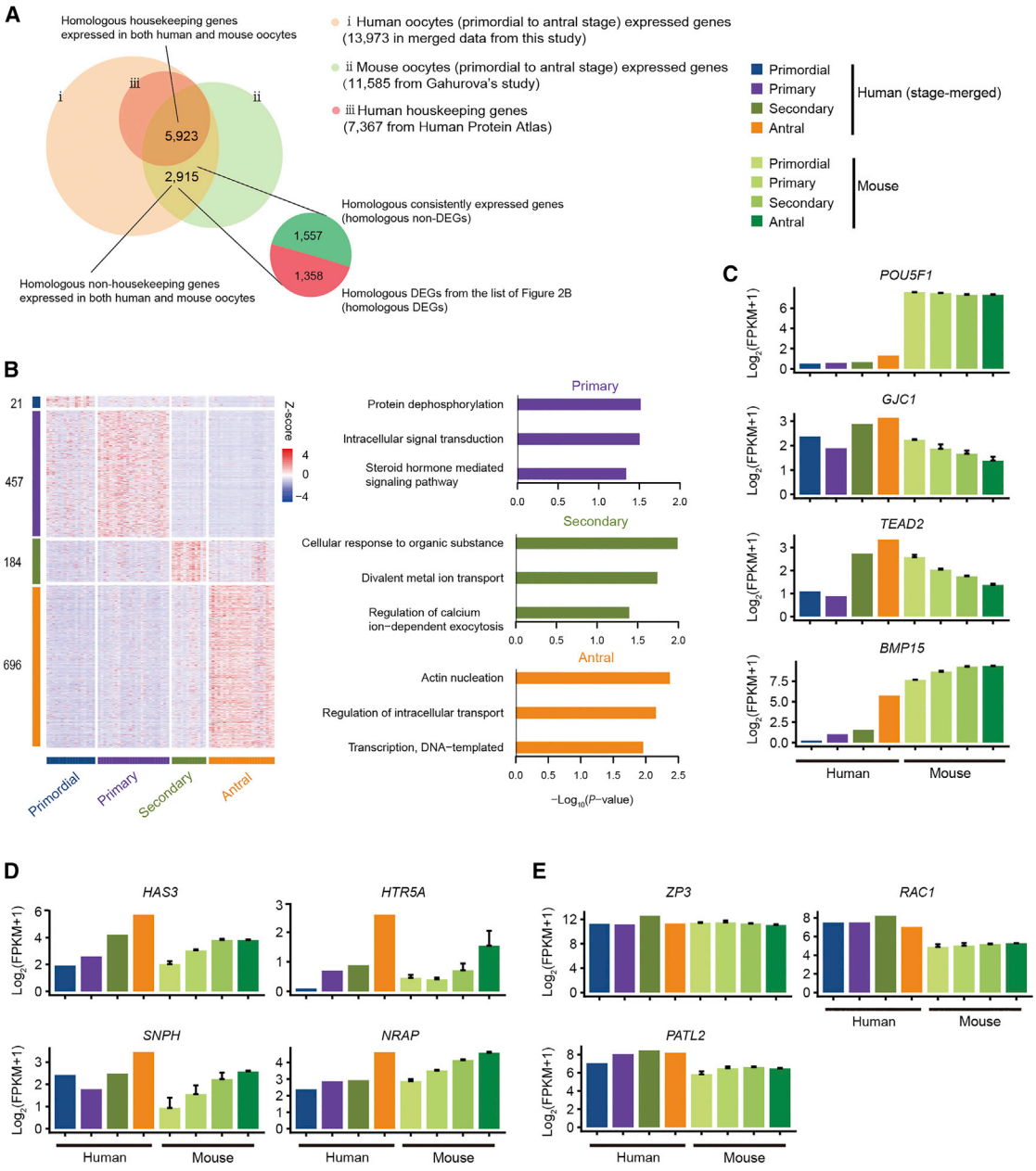


Figure 7. Comparison of Human and Mouse Oocyte Transcriptomic Profiles during Folliculogenesis

(A) Venn diagram of total expressed genes in human oocytes (peach) and mouse oocytes (light green) demonstrates overlap between the gene populations in two species and also overlap with the human housekeeping genes (deep orange). The single-cell oocyte data were merged by each stage in order to compare with the bulk RNA-seq data of mouse oocytes. The pie chart represents the non-housekeeping human-mouse oocyte-derived homologous genes (2,915), including homologous DEGs (red) and homologous consistently expressed genes (green).

(B) Heatmap of co-expressed homologous DEGs in oocytes at five stages of folliculogenesis. The numbers of identified DEGs are indicated on the y axis, and the stages of follicular development are presented along with the x axis. The color key from blue to red indicates the relative gene expression level from low to high, respectively. The enriched GO terms of human oocytes are shown on the right.

(C) Bar plots demonstrate a comparative analysis of the selected oocyte-derived homologous DEGs that show the distinct expression patterns in human (left) and mouse (right) oocytes during folliculogenesis. Different colors represent different stages of folliculogenesis.

(D) Bar plots demonstrate a comparative analysis of four oocyte-derived homologous DEGs that show a similar expression pattern in human (left) and mouse (right) oocytes during folliculogenesis.

(E) Bar plots demonstrate a comparative analysis of three homologous non-DEGs annotated as oocyte maturation defect genes (OMIM) in human (left) and mouse (right) oocytes during folliculogenesis.

Error bars indicate SE.

folliculogenesis in human and mouse oocytes (Figure 7E). Such shared gene expression pattern suggests the existence of evolutionarily conserved mechanisms involved in the mammalian folliculogenesis.

DISCUSSION

Our work revealed unique features in transcriptional machinery and reciprocal interactions between human oocytes and GCs as well as gene signatures in each follicular stage. These identified candidate compartment-specific and stage-specific genes in both oocytes and GCs may provide a valuable clue for future functional studies. These findings may have important implication for the development of essential genetic tools for cell-type-specific or stage-specific labeling and manipulations, which could be utilized in basic and translational research of the human ovary. In this study, we only evaluated the transcriptomic profiles of cumulus cells, located in proximity to oocytes and known to play the key roles in oocyte-GC communications. However, the mural GCs have distinct gene expression profiles that were not explored in this study (Grøndahl et al., 2012). The spatiotemporal transcriptomic analysis of mural GCs during follicular development requires further investigation.

We have proposed a set of candidate biomarkers that could be a valuable source of information to measure the reproductive potential and ovarian reserve. These identified subsets of secretory protein coding genes might be biomarkers addressing specific outcomes of interest. For example, the reproductive potential in general population is likely to be represented by the markers expressed at all stages of follicular development, and the age of menopause should be more accurately predicted by the markers that originate in the primordial follicle pool. It is more reasonable that the ovarian response to fertility treatments is predicted by using the biomarkers confined to the antral follicle, the main target of the gonadotropin treatments, and the subsequent pregnancy outcome is predicted by using the biomarkers secreted by the more advanced follicles. Further research aimed at elucidating the biological roles of the proposed candidates, in conjunction with comprehensive evaluation of their expression levels in biological fluids (e.g., serum and follicular fluid), will contribute to the development of novel ovarian reserve testing.

In this study, we uncovered signaling pathways that are coordinately and reciprocally regulated in human oocytes and their surrounding GCs by identifying the ligands and their receptors that are derived from the reciprocal compartments. This indicates that signaling pathways are activated by the ligands derived from the oocytes that act on the neighboring gonadal somatic cells and vice versa. For example, we found that the NOTCH signaling pathway was activated in GCs via oocyte-driven mechanisms, which is in line with our previous observation in the fetal ovary (Li et al., 2017). The KITLG-KIT pathway was initiated by *KITLG* expressed in GCs in paracrine effect, whereas the TGF-β pathway appeared to be activated in both autocrine and paracrine manners. Moreover, we identified five pathways that may govern the activation of primordial follicles, which potentially provides a treatment strategy for women with premature ovarian insufficiency.

We offer further insights into several species-specific gene expression patterns during folliculogenesis between human and mouse. Studies using mouse models are a mainstay of translational biomedical research and are widely implicated in the investigation of the ovarian transcriptome. However, it has been previously reported that the oocyte transcriptome is highly variable across mammals, and the human oocyte is likely to have greater complexity (Biase, 2017; Sylvestre et al., 2013). We evidenced strong transcriptional activity in both human and mouse oocytes, which dominated in humans and demonstrated considerable variability in oocyte transcriptome between human and mouse. We revealed gene expression dynamics during folliculogenesis in both species. We also showed some degree of similarity in gene expression between human and mouse oocytes, which highlights the potentially significant translational value of these conserved genes for future research. Based on the results of our analyses, we have proposed a cautious approach when mouse oocyte data are applied to the human domain and provided a dataset for further investigations of the molecular mechanisms associated with oocyte development through mouse models, which will improve the understanding of how well the oocyte transcriptome data translate from mouse to humans.

In summary, this is the first comprehensive investigation of transcriptomic profiles from both the oocyte and somatic follicular compartments in the adult human ovary. This work paves the way to understanding the molecular regulation of human folliculogenesis, which provided a valuable resource for developing targeted interventions aimed at improving follicle recruitment *in vivo* for premature ovarian failure (POF) patients and restoring fully competent oocytes *in vitro*.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Ovarian Histology Assessment
  - Immunohistochemistry
  - Isolation of Human Oocytes and GCs
  - Single-Cell cDNA Libraries Construction from Oocytes and GCs
  - RNA-Seq Data Processing
  - Principal Component Analysis (PCA)
  - Identification of Differentially Expressed Genes and Gene Ontology Analysis
  - Identifying Expression Patterns of Maternal-Effect Gene
  - Identifying Secretory Protein Coding Genes
  - Transcription Factor Network Construction
  - GSEA Analysis
  - Analysis of Conservation Between Human and Mouse
- DATA AND SOFTWARE AVAILABILITY
  - Data Resources

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures and seven tables and can be found with this article online at https://doi.org/10.1016/j.molcel.2018.10.029.

ACKNOWLEDGMENTS

We thank Prof. Fuchou Tang for technical support in single-cell RNA-seq experiments. This work was supported by the National Key Technology R&D Program of China (nos. 2017YFC1002002 and 2016YFC0900103) and National Science Foundation of China (nos. 31522034, 31571544, 81571386, 81730038, 81521002, and 31230047).

AUTHOR CONTRIBUTIONS

J.Q., J.Y., and L.Y. conceived the project. Y.Z., Q.Q., T.W., M.Y., S.Y., C.L., X.X., Y.D., and Y.R. performed the experiment. Y. Yao, P.Y., H.G., J.H., H.H., K.Z., Y. Wang, Y. Wu, and M.L. collected the samples. Z.Y., Y. Yu, and X.Z. conducted the bioinformatics analyses. J.Q., J.Y., L.Y., Y.Z., Z.Y., H.-M.C., V.N., R.L., and P.L. wrote the manuscript with help from all of the authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: May 6, 2018
Revised: August 23, 2018
Accepted: October 17, 2018
Published: November 21, 2018

REFERENCES

Alvarez, M.J., Shen, Y., Giorgi, F.M., Lachmann, A., Ding, B.B., Ye, B.H., and Califano, A. (2016). Functional characterization of somatic mutations in cancer using network-based inference of protein activity. Nat. Genet. *48*, 838–847.

Biase, F.H. (2017). Oocyte developmental competence: insights from cross-species differential gene expression and human oocyte-specific functional gene networks. OMICS *21*, 156–168.

Blake, J.A., Eppig, J.T., Kadin, J.A., Richardson, J.E., Smith, C.L., and Bult, C.J.; Mouse Genome Database Group (2017). Mouse Genome Database (MGD)-2017: community knowledge resource for the laboratory mouse. Nucleic Acids Res. *45*, D723–D729.

Bonnet, A., Cabau, C., Bouchez, O., Sarry, J., Marsaud, N., Foissac, S., Woloszyn, F., Mulsant, P., and Mandon-Pepin, B. (2013). An overview of gene expression dynamics during early ovarian folliculogenesis: specificity of follicular compartments and bi-directional dialog. BMC Genomics *14*, 904.

Braude, P., Bolton, V., and Moore, S. (1988). Human gene expression first occurs between the four- and eight-cell stages of preimplantation development. Nature *332*, 459–461.

Cheng, G., Weihua, Z., Mäkinen, S., Mäkelä, S., Saji, S., Warner, M., Gustafsson, J.A., and Hovatta, O. (2002). A role for the androgen receptor in follicular atresia of estrogen receptor beta knockout mouse ovary. Biol. Reprod. *66*, 77–84.

Dean, J. (2002). Oocyte-specific genes regulate follicle formation, fertility and early mouse development. J. Reprod. Immunol. *53*, 171–180.

Douville, G., and Sirard, M.A. (2014). Changes in granulosa cells gene expression associated with growth, plateau and atretic phases in medium bovine follicles. J. Ovarian Res. *7*, 50.

Gahurova, L., Tomizawa, S.I., Smallwood, S.A., Stewart-Morgan, K.R., Saadeh, H., Kim, J., Andrews, S.R., Chen, T., and Kelsey, G. (2017). Transcription and chromatin determinants of de novo DNA methylation timing in oocytes. Epigenetics Chromatin *10*, 25.

Gao, J.M., Yan, J., Li, R., Li, M., Yan, L.Y., Wang, T.R., Zhao, H.C., Zhao, Y., Yu, Y., and Qiao, J. (2013). Improvement in the quality of heterotopic allotransplanted mouse ovarian tissues with basic fibroblast growth factor and fibrin hydrogel. Hum. Reprod. *28*, 2784–2793.

Gershon, E., Plaks, V., and Dekel, N. (2008). Gap junctions in the ovary: expression, localization and function. Mol. Cell. Endocrinol. *282*, 18–25.

Girard, A., Dufort, I., Douville, G., and Sirard, M.A. (2015). Global gene expression in granulosa cells of growing, plateau and atretic dominant follicles in cattle. Reprod. Biol. Endocrinol. *13*, 17.

Gougeon, A. (1986). Dynamics of follicular growth in the human: a model from preliminary results. Hum. Reprod. *1*, 81–87.

Grøndahl, M.L., Andersen, C.Y., Bogstad, J., Borgbo, T., Boujida, V.H., and Borup, R. (2012). Specific genes are selectively expressed between cumulus and granulosa cells from individual human pre-ovulatory follicles. Mol. Hum. Reprod. *18*, 572–584.

Guo, F., Yan, L., Guo, H., Li, L., Hu, B., Zhao, Y., Yong, J., Hu, Y., Wang, X., Wei, Y., et al. (2015). The transcriptome and DNA methylome landscapes of human primordial germ cells. Cell *161*, 1437–1452.

Haddad-Tóvolli, R., Dragano, N.R.V., Ramalho, A.F.S., and Velloso, L.A. (2017). Development and function of the blood-brain barrier in the context of metabolic control. Front. Neurosci. *11*, 224.

Hatzirodos, N., Irving-Rodgers, H.F., Hummitzsch, K., Harland, M.L., Morris, S.E., and Rodgers, R.J. (2014). Transcriptome profiling of granulosa cells of bovine ovarian follicles during growth from small to large antral sizes. BMC Genomics *15*, 24.

Hu, Y.C., Wang, P.H., Yeh, S., Wang, R.S., Xie, C., Xu, Q., Zhou, X., Chao, H.T., Tsai, M.Y., and Chang, C. (2004). Subfertility and defective folliculogenesis in female mice lacking androgen receptor. Proc. Natl. Acad. Sci. USA *101*, 11209–11214.

Huang, Z., and Wells, D. (2010). The human oocyte and cumulus cells relationship: new insights from the cumulus cell transcriptome. Mol. Hum. Reprod. *16*, 715–725.

Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. *4*, 44–57.

Kerr, B., Garcia-Rudaz, C., Dorfman, M., Paredes, A., and Ojeda, S.R. (2009). NTRK1 and NTRK2 receptors facilitate follicle assembly and early follicular development in the mouse ovary. Reproduction *138*, 131–140.

Kim, J.Y. (2012). Control of ovarian primordial follicle activation. Clin. Exp. Reprod. Med. *39*, 10–14.

Knight, P.G., and Glister, C. (2006). TGF-beta superfamily members and ovarian follicle development. Reproduction *132*, 191–206.

Li, R., and Albertini, D.F. (2013). The road to maturation: somatic cell interaction and self-organization of the mammalian oocyte. Nat. Rev. Mol. Cell Biol. *14*, 141–152.

Li, R., Zhang, Q., Yang, D., Li, S., Lu, S., Wu, X., Wei, Z., Song, X., Wang, X., Fu, S., et al. (2013). Prevalence of polycystic ovary syndrome in women in China: a large community-based study. Hum. Reprod. *28*, 2562–2569.

Li, L., Dong, J., Yan, L., Yong, J., Liu, X., Hu, Y., Fan, X., Wu, X., Guo, H., Wang, X., et al. (2017). Single-cell RNA-seq analysis maps development of human germline cells and gonadal niche interactions. Cell Stem Cell *20*, 858–873.e4.

Macosko, E.Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A.R., Kamitaki, N., Martersteck, E.M., et al. (2015). Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. Cell *161*, 1202–1214.

Margolin, A.A., Nemenman, I., Basso, K., Wiggins, C., Stolovitzky, G., Dalla Favera, R., and Califano, A. (2006). ARACNE: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. BMC Bioinformatics *7* (*Suppl 1*), S7.

Meczekalski, B. (2009). Oocyte-specific genes: role in fertility and infertility. J. Endocrinol. Invest. *32*, 474–481.

Metsalu, T., and Vilo, J. (2015). ClustVis: a web tool for visualizing clustering of multivariate data using principal component analysis and heatmap. Nucleic Acids Res. *43* (W1), W566–W570.

Parks, J.C., Patton, A.L., McCallie, B.R., Griffin, D.K., Schoolcraft, W.B., and Katz-Jaffe, M.G. (2016). Corona cell RNA sequencing from individual oocytes revealed transcripts and pathways linked to euploid oocyte competence and live birth. Reprod. Biomed. Online *32*, 518–526.

Robker, R.L., Russell, D.L., Espey, L.L., Lydon, J.P., O'Malley, B.W., and Richards, J.S. (2000). Progesterone-regulated genes in the ovulation process: ADAMTS-1 and cathepsin L proteases. Proc. Natl. Acad. Sci. USA *97*, 4689–4694.

Rosenfeld, C.S., Roberts, R.M., and Lubahn, D.B. (2001). Estrogen receptor- and aromatase-deficient mice provide insight into the roles of estrogen within the ovary and uterus. Mol. Reprod. Dev. *59*, 336–346.

Sánchez, F., and Smitz, J. (2012). Molecular control of oogenesis. Biochim. Biophys. Acta *1822*, 1896–1912.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. USA *102*, 15545–15550.

Sylvestre, E.L., Robert, C., Pennetier, S., Labrecque, R., Gilbert, I., Dufort, I., Léveillé, M.C., and Sirard, M.A. (2013). Evolutionary conservation of the oocyte transcriptome among vertebrates and its implications for understanding human reproductive function. Mol. Hum. Reprod. *19*, 369–379.

Thomas, F.H., and Vanderhyden, B.C. (2006). Oocyte-granulosa cell interactions during mouse follicular development: regulation of kit ligand expression and its role in oocyte growth. Reprod. Biol. Endocrinol. *4*, 19.

Tirosh, I., Izar, B., Prakadan, S.M., Wadsworth, M.H., 2nd, Treacy, D., Trombetta, J.J., Rotem, A., Rodman, C., Lian, C., Murphy, G., et al. (2016). Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. Science *352*, 189–196.

Tisdall, D.J., Watanabe, K., Hudson, N.L., Smith, P., and McNatty, K.P. (1995). FSH receptor gene expression during ovarian follicle development in sheep. J. Mol. Endocrinol. *15*, 273–281.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-seq. Bioinformatics *25*, 1105–1111.

Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L. (2010). Transcript assembly and quantification by RNA-seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat. Biotechnol. *28*, 511–515.

Treutlein, B., Brownfield, D.G., Wu, A.R., Neff, N.F., Mantalas, G.L., Espinoza, F.H., Desai, T.J., Krasnow, M.A., and Quake, S.R. (2014). Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. Nature *509*, 371–375.

Uhlén, M., Fagerberg, L., Hallström, B.M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, Å., Kampf, C., Sjöstedt, E., Asplund, A., et al. (2015). Proteomics. Tissue-based map of the human proteome. Science *347*, 1260419.

Vanacker, J., Camboni, A., Dath, C., Van Langendonckt, A., Dolmans, M.M., Donnez, J., and Amorim, C.A. (2011). Enzymatic isolation of human primordial and primary ovarian follicles with Liberase DH: protocol for application in a clinical setting. Fertil. Steril. *96*, 379–383.e3S.

Wang, T.R., Yan, L.Y., Yan, J., Lu, C.L., Xia, X., Yin, T.L., Zhu, X.H., Gao, J.M., Ding, T., Hu, W.H., et al. (2014). Basic fibroblast growth factor promotes the development of human ovarian early follicles during growth in vitro. Hum. Reprod. *29*, 568–576.

Wang, T.R., Yan, J., Lu, C.L., Xia, X., Yin, T.L., Zhi, X., Zhu, X.H., Ding, T., Hu, W.H., Guo, H.Y., et al. (2016). Human single follicle growth in vitro from cryopreserved ovarian tissue after slow freezing or vitrification. Hum. Reprod. *31*, 763–773.

Xia, X., Yin, T., Yan, J., Yan, L., Jin, C., Lu, C., Wang, T., Zhu, X., Zhi, X., Wang, J., et al. (2015). Mesenchymal stem cells enhance angiogenesis and follicle survival in human cryopreserved ovarian cortex transplantation. Cell Transplant. *24*, 1999–2010.

Youds, J.L., and Boulton, S.J. (2011). The choice in meiosis - defining the factors that influence crossover or non-crossover formation. J. Cell Sci. *124*, 501–513.

Zhang, H.M., Chen, H., Liu, W., Liu, H., Gong, J., Wang, H., and Guo, A.Y. (2012). AnimalTFDB: a comprehensive animal transcription factor database. Nucleic Acids Res. *40*, D144–D149.

Zhang, H.M., Liu, T., Liu, C.J., Song, S., Zhang, X., Liu, W., Jai, H., Xue, Y., and Guo, A. (2015). AnimalTFDB 2.0: a resource for expression, prediction and functional study of animal transcription factors. Nucleic Acids Res. *43*, D76–D81.

Zheng, P., and Dean, J. (2007). Oocyte-specific genes affect folliculogenesis, fertilization, and early development. Semin. Reprod. Med. *25*, 243–251.

Zhou, F., Li, X., Wang, W., Zhu, P., Zhou, J., He, W., Ding, M., Xiong, F., Zheng, X., Li, Z., et al. (2016). Tracing haematopoietic stem cell formation at single-cell resolution. Nature *533*, 487–492.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit anti-RBM24 antibody	Abcam	Cat#ab94567; RRID: AB_10674832
Rabbit anti-GPD1(Glycerol 3 Phosphate Dehydrogenase) antibody	Abcam	Cat#ab153902; RRID: AB_153902
Rabbit anti-NTF4 antibody	Abcam	Cat#ab6205; RRID: AB_305371
Rabbit anti-LCP2 (SLP76) antibody	Abcam	Cat#ab196599; RRID: AB_196599
Rabbit anti-CDCA3 antibody	Abcam	Cat#ab167037; RRID: AB_167037
Rabbit anti-BNIPL antibody	CST	Cat#12396S; RRID: AB_2688036
Rabbit anti-TST antibody	Abcam	Cat#ab166625; RRID: AB_166625
Rabbit Anti-Connexin 43 / GJA1 antibody	Abcam	Cat#ab11370; RRID: AB_297976
Rabbit Anti-Connexin 40 / GJA5 antibody	Abcam	Cat#ab38580; RRID: AB_731705
Rabbit Anti-Connexin 45 / GJC1 antibody	Abcam	Cat#ab135474; RRID: AB_135474
Rabbit Anti-Connexin 46 / GJA3 antibody	GeneTex	Cat#GTX87220; RRID: AB_10722469
Biological Samples		
Human ovarian tissues	Donors from Peking University 3rd Hospital	N/A
Critical Commercial Assays		
Leibovitz's L-15 medium	Sigma-Aldrich	#L4386
Human serum albumin	LifeGlobal	#GHSA-125
αMEM medium	Sigma-Aldrich	#M0446
Liberase DH	Roche	#5401054001
DNase I	Sigma-Aldrich	#AMPD1
DPBS	Sigma-Aldrich	#D8537
Accutase	Sigma-Aldrich	#A6964
Kappa Hyper Prep Kit	Kappa Biosystems	#KK8504
Deposited Data		
RNA-Seq data	This paper	GEO: GSE107746
RNA-Seq data of human MII oocytes	Published manuscripts	GEO: GSE36552
Imaging dataset	This paper	https://doi.org/10.17632/63pf4hcn2k.1
Software and Algorithms		
TopHat (v2.0.12)	(Trapnell et al., 2009)	http://ccb.jhu.edu/software/tophat/downloads/
Cufflinks (v2.2.1)	(Trapnell et al., 2010)	http://cole-trapnell-lab.github.io/cufflinks/install/
GO (DAVID)	(Huang da et al., 2009)	https://david-d.ncifcrf.gov/
ARACNe software	(Margolin et al., 2006)	https://github.com/califano-lab/ARACNe-AP
Animal TFDB 2.0	(Zhang et al., 2012, 2015)	http://bioinfo.life.hust.edu.cn/AnimalTFDB#!/

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed and will be fulfilled by the Lead Contact, Jie Qiao (jie.qiao@263.net).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

The experiments performed in this study were approved by the Ethics Committee of Peking University Third Hospital. With oral and written informed consent, fresh ovarian tissues were obtained from 7 female donors who underwent ovariectomy for the following

indications: sex reassignment surgery (n = 1), fertility preservation for cervical cancer (n = 1), endometrial cancer (n = 2), benign ovarian mass (n = 2) and lymphoma (n = 1). All the donors were of reproductive age, ranging from 24 to 32 years, with a median age of 28 years (Table S6). All the participants had a regular menstrual cycle with no history of any autoimmune or genetic disease. These tissue donors involved patients who had not used hormonal therapy for more than 6 months before surgery, had no previous ovarian surgery and were not exposed to any cytotoxic agents or radiotherapy. All the ovary samples were without histopathological abnormality confirmed by the gynecological pathologists.

METHOD DETAILS

Ovarian Histology Assessment

Histological assessment was performed on all the ovarian tissue samples by using hematoxylin and eosin (HE) staining as described elsewhere (Xia et al., 2015). Briefly, ovarian tissues were fixed in 4% formaldehyde and processed for routine paraffin embedding after fixation. 5-μm-thick sections were prepared for hematoxylin and eosin staining. All the prepared tissue sections were reviewed by two independent pathologists and confirmed normal histology.

Immunohistochemistry

Using the ABC Staining System (Zhongshan Golden Bridge Biotechnology, Beijing, China), we performed immunohistochemistry staining of the ovarian tissue samples for the RNA-Seq analysis, as previously described (Gao et al., 2013). Briefly, the paraffin embedded ovarian tissue sections were deparaffinated and rehydrated in xylene and in decreasing graded ethanol. Next, the sections were demasking in citrate buffer (pH 6.0) for 20 min at 98°C and cooled down to room temperature, then were incubated in 0.3% H_2O_2 for 10 min to block endogenous peroxidase activity. The sections were treated overnight at 4°C with primary antibodies according to the manufacturer's instructions, negative controls were obtained by omission of primary antibodies and 1% PBS was used to wash the sections three times for 3 min each. Then the sections were incubated in reaction enhancer and secondary antibodies each for 20 min at room temperature and this procedure was repeated once. The sections were washed by 1%PBS three times for 3 min each. Then the DAB was added to each section by incubating for 2-5 min and hematoxylin was used for counterstaining for 2 min. Then the sections were dehydrated through increasing ethanol concentration, cleared in xylene, and sealed. Brown staining of the cytoplasm or nucleus of the cells was considered as positive.

Isolation of Human Oocytes and GCs

Human follicles were isolated from fresh ovarian tissues as described previously (Vanacker et al., 2011; Wang et al., 2014). Briefly, after removal of medulla tissues, the ovarian cortical pieces were placed in a tissue sectioner (McIlwain Tissue Chopper, The Mickle Laboratory, Guildford, UK) and cut into 0.5 × 0.5 × 1 mm pieces. Then the tissue pieces were enzymatically digested by a mixed digestion medium, which included αMEM (Sigma-Aldrich) media, 0.04 mg/ml Liberase DH (Dispase High; Roche Diagnostics GmbH, Mannheim, Germany), 10 IU/ml DNase I (Sigma-Aldrich), 100 IU/ml penicillin and 100 μg/ml streptomycin (Invitrogen) and incubated for 75 min on a shaker at 37°C (Thermo Fisher, Marietta, OH, USA). The incubation was terminated by double wash with DPBS (Sigma-Aldrich) supplemented with 10% HSA (LifeGlobal). For secondary follicles, we used liberase DH and DNase to digest the basement membrane while the theca cells and stromal cells are in the outer layers of follicle basement membrane. After the membrane was digested completely, the theca cells were removed together with the stromal cells. While for antral follicles, the oocyte-cumulus complex was mechanically isolated by aspiration with 29G needles without contamination of theca cells. After these steps, the oocyte-GC complex was obtained. Then we utilized the accutase to obtain the GCs in single cells. At last, we removed zona pellucida of oocytes by hydrolysis with 0.012 M hydrochloric acid to obtain naked oocyte without contamination of GCs (Figure S1A). The removal of surrounding GCs was confirmed under microscopy that oocyte was not attached by any GCs.

Each GC sample comprised randomly selected 10 cells because of low abundancy of RNA in these cells. GCs from the antral and preovulatory follicles included cumulus cells isolated from the cumulus-oocyte complex (COC) as following: 6 samples of cumulus cells from the antral follicles and 16 samples from the preovulatory follicles. Follicular stages were classified according to the criteria described by Gougeon (Gougeon, 1986). The diameters of follicles and oocytes were measured in a light microscope (Nikon).

Single-Cell cDNA Libraries Construction from Oocytes and GCs

The oocytes and GCs that were isolated from follicles were analyzed by RNA-Seq as previously described (Li et al., 2013; Wang et al., 2016). Briefly, the oocytes or 10 randomly selected GCs were transferred into the lysis buffer quickly using a mouth pipette. Then we performed reverse transcription on the cell lysate and terminal deoxynucleotidyl transferase was adopted to add a poly A tail to the 3′ end of the first-strand cDNAs. Next, we performed 20 cycles of PCR to amplify the single-cell cDNA library. qPCR analysis was conducted to check the quality of the cDNA libraries using two housekeeping genes, *GAPDH* and *RPS24*. The RNA-Seq libraries were constructed by a Kappa Hyper Prep Kit (Kappa Biosystems).

RNA-Seq Data Processing

The analysis of single-cell RNA-Seq data was carried out as previously described (Guo et al., 2015; Li et al., 2017; Zhou et al., 2016). Briefly, RNA-Seq raw reads with 10% low-quality bases, adapters and artificial sequences (including UP1, UP2, polyA sequences)

introduced during the experimental processes were trimmed by in-house scripts. Next, the trimmed clean reads were aligned to the UCSC human hg19 reference using the Tophat2 (v2.1.0) with default settings (Trapnell et al., 2009). Cufflinks (v2.2.1) was further used to quantify transcription levels of annotated genes (Trapnell et al., 2010). The GENECODE (https://www.gencodegenes.org) lncRNAs expressions were evaluated as FPKM values by Cufflinks.

Previously published data, including those from human MII oocytes (Li et al., 2013), human pre-implantation embryos (Li et al., 2013), and mouse oocytes (Gahurova et al., 2017), were downloaded from the GEO database, and the raw fastq reads were obtained and incorporated into our analysis. For all sequenced cells, we counted the number of genes detected in each cell. Cells with fewer than 2,400 genes or 500,000 mapped reads were filtered out. In total, 80 oocytes and 71 GCs at five developmental stages passed the filter standards. To ensure the accuracy of estimated gene expression levels, only genes with FPKM > 1 in at least one cell were analyzed (Treutlein et al., 2014). Expression levels of each gene were plus one then log2 transformed in the following analysis. The expression levels of each oocyte and GC are presented in Table S7.

Principal Component Analysis (PCA)

The Seurat method was applied to analyze the single-cell data (80 oocytes and 71 GCs) to observe the whole clustering profile (Macosko et al., 2015). Only highly variable genes (coefficient of variation > 0.5) were used as inputs for PCA. The marker genes in PCA plot were plotted by the FeaturePlot function in Seurat package. To complement the PCA clustering more accurately, we also clustered the oocytes and GCs separately using the FactoMineR package in R (Metsalu and Vilo, 2015).

Identification of Differentially Expressed Genes and Gene Ontology Analysis

The multiple t test was used to obtain the statistical significance of differentially expressed genes in each stage (one stage versus all other stages). Only the genes with significant p values and false discovery rate (FDR) less than 0.05 with a fold change of log2 transformed FPKM larger than 1.5 were considered as the differentially expressed genes. Gene ontology analysis of differentially expressed genes was performed using DAVID (Huang da et al., 2009).

Identifying Expression Patterns of Maternal-Effect Gene

Maternal-effect genes play a critical role in mammalian development before zygote genome activation (ZGA). It has been reported that ZGA mainly happened at 4-cell to 8-cell stages (Braude et al., 1988; Li et al., 2013).To explore the expression patterns of maternal-effect genes during folliculogenesis, we searched for genes that carried from oocytes to early embryo and then degraded before ZGA. Genes expressed (FPKM > 1) at MII oocytes and zygote stage but relatively low expressed (FPKM < 1) at 8-cell stage were taken as the candidates. The candidate genes were clustered by expression correlation and cut into four clusters with cutree function in R.

Identifying Secretory Protein Coding Genes

To identify candidate biomarkers for predicting ovarian reserve, we focused on the oocyte- or GC-derived secretory proteins. The secretory protein-encoding genes were downloaded from the Human Protein Atlas database (https://www.proteinatlas.org/). The GTEx v6 database (http://gtexportal.org) was utilized to obtain information on gene expression level in tissues of interest. We aimed to evaluate protein secretion genes perferentially expressed in ovary, and therefore we selected genes that were previously identified only in gonads and could be present in the brain, but not in any other tissues (Uhlén et al., 2015). We assumed that brain-blood barrier prevents secretion of the proteins into the systemic circulation and thus reduces the confounding effect on peripheral levels of these proteins. This assumption was based on the previously observed high levels of AMH in the brain, which supports the idea of the restrictive function of brain-blood barrier (Haddad-Tóvolli et al., 2017).

Transcription Factor Network Construction

TFs play key roles in regulating mammalian development. To find the driver factors and construct their regulatory network in every two consecutive stages, we used the ARACNe to perform the regulatory network analysis as previously described (Li et al., 2017). First, 1,469 human TFs in AnimalTFDB (Zhang et al., 2012) and gene expression matrix were taken as inputs for the ARACNe-AP software (Zhang et al., 2012). Then, viper package in R was used to visualize the TFs and their target genes in each consecutive stage (Alvarez et al., 2016). Regulators with p values less than 0.01 were inferred as driver factors in every two consecutive stages.

GSEA Analysis

To identify the significantly enriched pathways in the transition from primordial to primary stages, we used Gene Set Enrichment Analysis (GSEA, http://software.broadinstitute.org/gsea/index.jsp) and the KEGG pathway (Subramanian et al., 2005) to perform enrichment analysis (Subramanian et al., 2005). The gene sets that showed nominal p value less than 0.05 were chosen as enriched.

Analysis of Conservation Between Human and Mouse

Human and mouse homologous genes were downloaded from the Vertebrate Homology Database (http://www.informatics.jax.org/homology.shtml). Housekeeping genes were obtained from Human Protein Atlas. Human oocytes expressed genes were overlapped with mouse oocytes expressed homologous genes to find genes expressed in both human and mouse oocytes, and

the housekeeping genes were filtered out from the overlapped gene set. The retained genes were divided into homologous DEGs (genes in DEGs of human oocytes) and homologous non-DEGs. The genes involved in oocyte maturation defect were downloaded from the Online Mendelian Inheritance in Man (OMIM) database (https://www.omim.org/) and incorporated into analysis.

DATA AND SOFTWARE AVAILABILITY

Data Resources

The accession number for the RNA-Seq data reported in this paper is GEO: GSE107746.

Our IHC staining figures have been deposited in Mendeley datasets (https://doi.org/10.17632/63pf4hcn2k.1)