

## METHODS AND RESOURCES

## Single-cell transcriptome landscape of ovarian cells during primordial follicle assembly in mice

Jun-Jie Wang<sup>1</sup>, Wei Ge<sup>1</sup>, Qiu-Yue Zhai<sup>1</sup>, Jing-Cai Liu<sup>1</sup>, Xiao-Wen Sun<sup>1</sup>, Wen-Xiang Liu<sup>1</sup>, Lan Li<sup>1</sup>, Chu-Zhao Lei<sup>2</sup>, Paul W. Dyce<sup>3</sup>, Massimo De Felici<sup>4</sup>, Wei Shen<sup>1\*</sup>

**1** College of Life Sciences, Institute of Reproductive Sciences, Qingdao Agricultural University, Qingdao, China, **2** Key Laboratory of Animal Genetics, Breeding and Reproduction of Shaanxi Province, College of Animal Science and Technology, Northwest A&F University, Yangling, China, **3** Department of Animal Sciences, Auburn University, Auburn, Alabama, United States of America, **4** Department of Biomedicine and Prevention, University of Rome Tor Vergata, Rome, Italy

\* wshen@qau.edu.cn, shenwei427@163.com



## OPEN ACCESS

**Citation:** Wang J-J, Ge W, Zhai Q-Y, Liu J-C, Sun X-W, Liu W-X, et al. (2020) Single-cell transcriptome landscape of ovarian cells during primordial follicle assembly in mice. *PLoS Biol* 18(12): e3001025. <https://doi.org/10.1371/journal.pbio.3001025>

**Academic Editor:** Polina V. Lishko, University of California, Berkeley, UNITED STATES

**Received:** October 22, 2019

**Accepted:** December 3, 2020

**Published:** December 22, 2020

**Peer Review History:** PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pbio.3001025>

**Copyright:** © 2020 Wang et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** The accession number of ovarian single-cell RNA sequencing raw data reported in this paper is NCBI GEO: GSE134339. The R code using for data processing

## Abstract

Primordial follicle assembly in the mouse occurs during perinatal ages and largely determines the ovarian reserve that will be available to support the reproductive life span. The development of primordial follicles is controlled by a complex network of interactions between oocytes and ovarian somatic cells that remain poorly understood. In the present research, using single-cell RNA sequencing performed over a time series on murine ovaries, coupled with several bioinformatics analyses, the complete dynamic genetic programs of germ and granulosa cells from E16.5 to postnatal day (PD) 3 were reported. Along with confirming the previously reported expression of genes by germ cells and granulosa cells, our analyses identified 5 distinct cell clusters associated with germ cells and 6 with granulosa cells. Consequently, several new genes expressed at significant levels at each investigated stage were assigned. By building single-cell pseudotemporal trajectories, 3 states and 1 branch point of fate transition for the germ cells were revealed, as well as for the granulosa cells. Moreover, Gene Ontology (GO) term enrichment enabled identification of the biological process most represented in germ cells and granulosa cells or common to both cell types at each specific stage, and the interactions of germ cells and granulosa cells basing on known and novel pathway were presented. Finally, by using single-cell regulatory network inference and clustering (SCENIC) algorithm, we were able to establish a network of regulons that can be postulated as likely candidates for sustaining germ cell-specific transcription programs throughout the period of investigation. Above all, this study provides the whole transcriptome landscape of ovarian cells and unearths new insights during primordial follicle assembly in mice.

is publicly available at GitHub at [https://github.com/WangLab401/2020scRNA\\_murine\\_ovaries](https://github.com/WangLab401/2020scRNA_murine_ovaries).

**Funding:** This work was supported by National Key Research and Development Program of China (2018YFC1003400), National Nature Science Foundation (31671554 and 31970788) and Taishan Scholar Construction Foundation of Shandong Province (ts20190946). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

**Abbreviations:** AHR, aryl hydrocarbon receptor; BMPs, bone morphogenetic protein; BPGs, bipotential pre-granulosa cells; D-BPG, differentiating bipotential pre-granulosa cell; D-EPG, differentiating epithelial pre-granulosa cell; E, embryonic day; EIF4A1, eukaryotic translation initiation factor 4A1; EPGs, epithelial pre-granulosa cells; FIGLA, folliculogenesis specific basic helix-loop-helix transcription factor; FOXL2, forkhead box L2; FOXO3, Forkhead box O3; FSH, follicle-stimulating hormone; G3BP2, GTPase-activating protein-binding protein 2; GDF9, growth differentiation factor 9; GO, Gene Ontology; GRN, global research network; HLH, helix-loop-helix; HSPB11, heat shock protein family B member 11; IRX, iroquois homeobox; KEGG, Kyoto Encyclopedia of Genes and Genomes; KITL, KIT proto-oncogene, receptor tyrosine kinase ligand; LATS, large tumor suppressor kinase; LGR5, leucine-rich repeat-containing G-protein coupled receptor 5; LHX8, LIM homeobox 8; MVH, mouse vasa homologue; NOBOX, NOBOX oogenesis homeobox; OOE, oocyte expressed protein; OR, ovarian reserve; PC, principal component; PD, postnatal day; PF, primordial follicle; PGC, primordial germ cell; PI, propidium iodide; POF, premature ovarian failure; PVDF, polyvinylidene fluoride; RA, retinoic acid; RSP01, R-spondin-1; SCENIC, single-cell regulatory network inference and clustering; SCMC, subcortical maternal complex; scRNA-seq, single-cell RNA sequencing; SMAD, homologues of *Drosophila* protein, mothers against decapentaplegic and *Caenorhabditis elegans* protein sma; SOHLH, spermatogenesis and oogenesis specific basic helix-loop-helix; TAF4b, TATA-box binding protein associated factor 4b; TAZ, tafazzin family protein; TF, transcriptional factor; TGF-beta, transforming growth factor-beta; t-SNE, t-distributed stochastic neighbor embedding; UMAP, uniform manifold approximation and projection; WNT4, Wnt family member 4; Yap1, yes-associated protein 1.

## Introduction

Gametogenesis is a finely regulated and complex process beginning from germline specification that gives rise to primordial germ cells (PGCs) and ending with either mature oocytes or sperms. In female mammals, primordial follicles (PFs) are the first functional unit of reproduction, each comprising a single oocyte surrounded by supporting somatic cells termed granulosa cells. Within 1 week after birth, the PFs constitute the ovarian reserve (OR), responsible for continued folliculogenesis and oocyte maturation throughout the adult life [1,2]. The PF population has long been believed to be nonrenewable, although this notion has recently been challenged [3,4]. Putting the latter discussion aside, it is evident that during the reproductive life span, the number of PFs progressively diminishes due to atresia as well as recruitment, maturation, and ovulation. In primates, the depletion of the reserve is the major determining driver of menopause in adult females and of various conditions of infertility in younger individuals grouped as having premature ovarian failure (POF) [5].

In the mouse, germline specification occurs before gastrulation in the epiblast at around embryonic day (E) 6.25, and this is primarily in response to bone morphogenetic protein (BMP) signals [6,7]. Around E7.0 to E8.0, PGCs are identifiable in the extraembryonic mesoderm at the angle between the allantois and the yolk sac [8]. From here, PGCs begin to migrate, and most of them move into the embryo proper to initiate colonization by E10.5, between E11.5 and E12.5, and the gonadal ridges form from the coelomic epithelium [9,10]. After active proliferation, PGCs/oogonia cease dividing at E13.5 and enter meiosis with the resulting formation of oocytes [11]. The oocytes are closely associated in clusters, termed germ cell nests, in which they are connected through cytoplasmic bridges formed because of incomplete cytokinesis during mitosis [12]. Oocyte loss and nest breakdown begin after birth in the cortical region of the ovary, but in the medullar region, these processes begin as early as E17.5 [13]. Nest breakdown begins after oocytes undergo the first meiotic block at the diplotene stage of prophase I. The failure of mouse oocytes to reach the diplotene stage impairs this process and consequently PF assembly [14,15]; however, it is still understood the 2 events of meiotic progression and PF formation were independent [16]. As reported above, germline nest breakdown is associated with a massive loss of oocytes that, in the mouse, has been reported to occur following programmed cell death mainly in the form of apoptosis and autophagy [15,17,18]. The surviving single oocytes become enveloped by granulosa cells from the surrounding nest, thereby producing the PFs. Synchronized development of oocytes and granulosa cells, as well as their interactive communication are required for efficient nest breakdown and the subsequent PF assembly.

Both in mouse and human, mutations of *Wnt4* and *Rspo1* genes induce failure of pre-granulosa cell differentiation, impair nest breakdown, and finally result in POF [19]. Recent studies in the mouse identified 2 waves of granulosa cell differentiation that contribute to 2 discrete populations of follicles [20,21]. The precursors of both populations arise from the ovarian surface epithelium. The first population, identifiable by the early expression of the transcription factor FOXL2, encircles the germ cell nests and eventually form the earliest PFs that develop within the ovarian medulla and undergo rapid maturation and atresia. The second population is composed of cells expressing leucine-rich repeat-containing G-protein coupled receptor 5 (LGR5), a likely receptor of Wnt family member 4 (WNT4)/R-spondin 1 (RSP01), that subsequently express forkhead box L2 (FOXL2) and down-regulate LGR5 and will form the OR of PFs in the cortex [22,23]. Other somatic cell transcription factors, such as iroquois homeobox (IRX)-3 and IRX-5, and the folliculogenesis specific basic helix-loop-helix transcription factor (FIGLA), spermatogenesis and oogenesis specific basic helix-loop-helix (SOHLH)-1, and SOHLH-2, LIM homeobox 8 (LHX8), NOBOX oogenesis homeobox

(NOBOX), TATA-box binding protein associated factor 4b (TAF4b), and aryl hydrocarbon receptor (AHR) also appear to play a role in nest breakdown and PF assembly both in mice and humans [24,25]. Several signaling exchanges between oocytes and pre-granulosa cells are also implicated in such processes, such as Notch, KIT proto-oncogene, receptor tyrosine kinase ligand (KITL)/KIT system, neurotrophins, transforming growth factor-beta (TGF-beta) and its family members growth differentiation factor 9 (GDF9), BMP15, Activin A, inhibin, follistatin, and anti-Mullerian hormone, as well estrogens, progesterone, and finally follicle-stimulating hormone (FSH) [1,15].

Despite these results and previous transcriptome analyses performed in rat [26,27] and mouse fetal and early postnatal ovaries [28,29], the precise molecular mechanisms underlying oocyte survival/death, granulosa cell differentiation, and the crosstalk among them during PF assembly are still incomplete. Moreover, all the latter studies were performed on the entire ovary or follicles so that the contribution of each cell type to the transcriptomes was not distinguishable. The recent advent of single-cell RNA sequencing (scRNA-seq) technologies makes it possible to identify specific cell subpopulations and the genetic programs regulating their differentiation in reproductive organs [10,30].

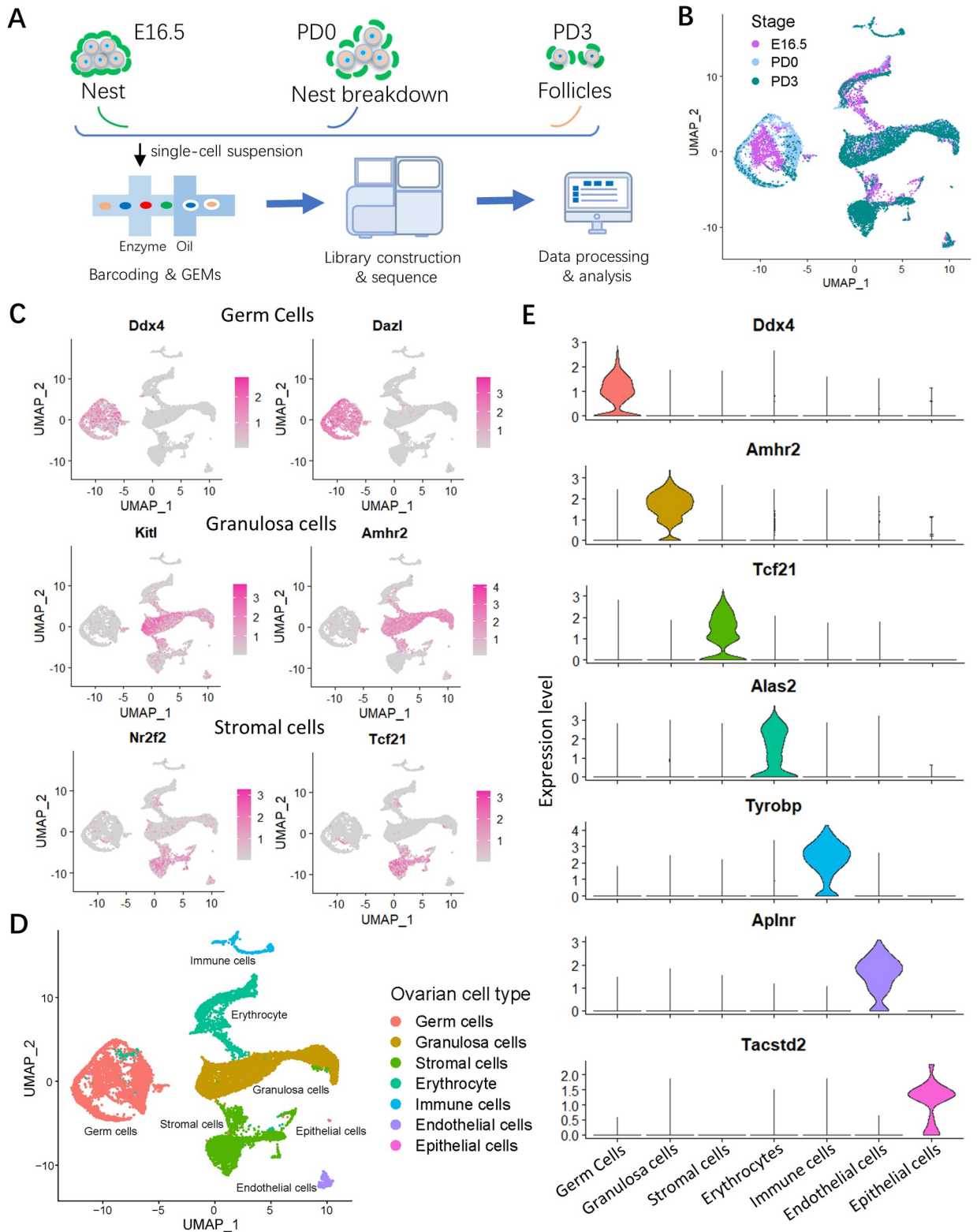
In the present paper, to achieve the mentioned goals, we focused on PF assembly in the mammalian ovary by performing scRNA-seq analyses on late fetal and early postnatal murine ovaries. The results and bioinformatics information reported here will certainly be useful for elucidating the molecular mechanisms underlying PF assembly and for selecting candidate regulatory factors for further investigation.

## Results

### scRNA-seq identified several types and subpopulations of germ cells and somatic cells in late fetal and early postnatal ovaries

To trace the distinct cell lineages and characterize the gene expression dynamics of ovarian cells during the crucial period of ovarian development lasting from late fetal and early postnatal age, ovaries were collected from E16.5, postnatal day 0 (PD0) and 3 (PD3) mice and subjected to immunohistochemical staining and scRNA-seq analyses. As expected, results from stained sections showed that the germ cell nests gradually broke down, while the percentage of germ cells enclosed in PFs increased from  $9.95 \pm 1.38\%$  at E16.5 to  $46.02 \pm 1.46\%$  at PD3 (S1A and S1B Fig).

Ovarian cell populations were then prepared and barcoded following single cell capture, and RNA sequencing was performed (Fig 1A). Following filtering out of low-quality cells (including empty droplets, dying cells, and potential doublets or multiplets), the remained number of cells were 5,742 in E16.5; 5,255 in PD0; and 5,871 in PD3, and the mean number of genes expressed in each sample was 2,258, 2,844 and 2,768, respectively (S1C Fig). According to uniform manifold approximation and projection (UMAP), the cell populations at 3 developmental stages were clustered (Fig 1B), and 7 cell types were identified, and according to their distinct transcriptomic signatures (S1 Table), they could be subdivided as follows: germ cells, granulosa cells, stromal cells, erythrocytes, immune cells, endothelial cells and epithelial cells (Fig 1C and 1D and S1D Fig). The expression of representative genes known to be specific for each cell type was reported in Fig 1E. In S1E Fig, the percent changes within the ovarian populations of the 7 identified cell types, during the analyzed developmental period, were reported. Following, the differentiated fate was examined mainly focusing on 2 major participants for PF formation, germ cells, and granulosa cells.



**Fig 1. scRNA-seq identified ovarian cell types during PF formation.** (A) Schematic diagram of the scRNA-seq analysis procedure. Ovaries were isolated and disaggregated into single cell suspensions; cells were barcoded and used for library construction onto 10x Genomics platform and the data produced after sequencing analyzed by dedicated software. (B) UMAP plot of ovarian cells colored by the selected developmental time points (that is Stage). E16.5 for medium orchid, PD0 for light blue and PD3 for dark cyan. (C) Feature plots of specific marker genes of

different cell types, including germ cells, granulosa cells, and stromal cells. The detailed marker genes of each cell type are shown in [S1 Table](#). (D) UMAP plot of 7 ovarian cell types. (E) Vlnplots of specific marker gene in each ovarian cell type. The sequencing data was deposited available in GSE134339, and this figure can be produced using scripts at [https://github.com/WangLab401/2020scRNA\\_murine\\_ovaries](https://github.com/WangLab401/2020scRNA_murine_ovaries). scRNA-seq, single cell RNA-sequencing; UMAP, uniform manifold approximation projection; E16.5, embryonic day 16.5; PD0, postnatal day 0; PD3, postnatal day 3.

<https://doi.org/10.1371/journal.pbio.3001025.g001>

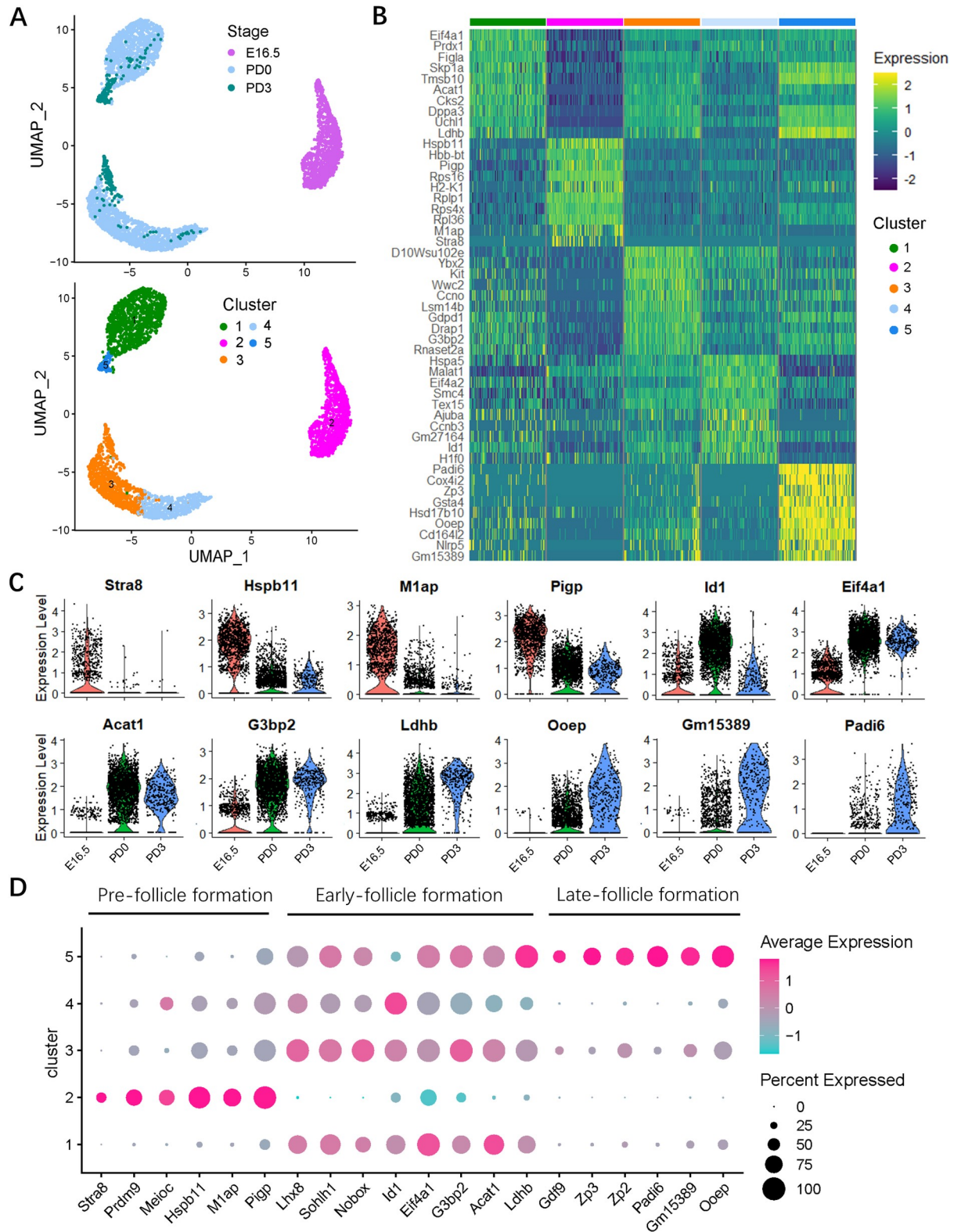
## Genetic dynamics of female germ cells from the nest stage to the primordial follicle formation

To dissect the heterogeneity of germ cells during development in the period of investigation through UMAP analysis, the germ cell populations were identified as 5 clusters: cluster 2 cells were mostly present at E16.5; clusters 1, 3, and 4 were mostly present at PD0; and cluster 5, at PD3 ([Fig 2A](#)). To obtain dynamic pattern of gene signature, the germ cell cluster-specific marker genes were identified ([S2 Table](#)). The top 10 expressed genes of each cluster were shown in a heatmap in [Fig 2B](#). According to these analyses, a genetic dynamic model, including pre-, early- and late-follicle formation stages during germ cell development was drawn ([Fig 2C and 2D](#)). Genes known to be expressed throughout female germ cell development at the beginning of meiosis are *Stra8*, *Prdm9*, and *Meioc* [31–34]; during PF formation are *Figla*, *Lhx8*, *Sohlh1*, *Nobox*, and *Ybx2* [4, 15, 24]; and for early oocyte growth are *Gdf9*, *Zp2*, and *Zp3* [35–38] ([Fig 2B and 2D](#)). Importantly, several new genes coupled with development stages of germ cells at the pre- (i.e., *M1ap*, *Hspb11*, and *Pigp*), early (i.e., *Eif4a1*, *G3bp2*, *Id1*, *Acat1*, and *Ldhb*), and late (i.e., *Ooep*, *Gm15389*, and *Padi6*) follicle formation stages were identified ([Fig 2D](#)). Moreover, these genes are also consistent with developmental time points ([S2A Fig](#)). In addition, the percentage changes of cells at these stages (pre-, early- and late-follicle formations) from E16.5 to PD3 were shown in [S2B Fig](#).

To improve the reliability of the dataset, the expressions of heat shock protein family B member 11 (HSBP11), GTPase-activating protein-binding protein 2 (G3BP2), and oocyte expressed protein (OOEP) representing the markers of pre-, early- and late-follicle formation stages, respectively, were confirmed by immunohistochemistry ([Fig 3A–3C](#)). Importantly, the same expression trend of these proteins was also detected with western blot, as well as eukaryotic translation initiation factor 4A1 (EIF4A1) for early follicle formation stage ([Fig 3D](#)).

## Fate transition of oocytes along pseudotime trajectories

To dissect the fate determination of germ cells throughout the investigated period, they were ordered along pseudotime trajectories according to the gene expression reported above. Three states (a jargon in pseudotime analysis; “state” is assigned to mark the segment of the trajectory tree in Monocle) were obtained, in which the majority of cells at the pre-follicle formation stage belonged to state 1, termed as “germ cells in nest”, while those at the early and late follicle formation stages were in states 2 and 3; in particular, cells of state 3 were “germ cells within follicle” ([Fig 4A–4C](#)). Meanwhile, the expressions of representative genes (marker genes of 3 follicle formation stages identified above) along with pseudotime trajectories were consistent with our inference ([S2C Fig](#)). In such pseudotime trajectories, 3 branches implied 3 differentiated stages of germ cells; moreover, 5 gene sets with distinct patterns were identified ([S3 Table](#)), which were likely involved in the commitment of germ cells within nests into follicles; clusters 1, 2, 3, 4, and 5 including 722; 1,819; 359; 1,077; and 1,904 genes, respectively ([Fig 4D and 4E](#)). The heatmap showed the dynamics of gene expression in the germ cells at the nest and follicle stages: Set 1 includes genes with steady high expression in follicular oocytes; set 2 genes showed an anabatic high expression in follicular oocytes; these 2 gene sets were both slightly or scarcely expressed at germ cells in nests (or state 1); set 3 seemed to function mainly at pre-



**Fig 2. Molecular characterization of the germ cell subsets.** (A) Cluster analysis of germ cells with UMAP plots based on developmental timeline (upper) and transcriptional patterns (below). (B) Heatmap of top 10 marker genes of the germ cell clusters. Top 50 marker genes in each cluster are shown in [S2 Table](#). (C) Violin plots of the expression level of representative genes along with developmental timeline. (D) Dot plot of general and stage-specific (pre-, early- and late-follicle formations) germ cell marker genes expressed in cell clusters. Dot size represents the

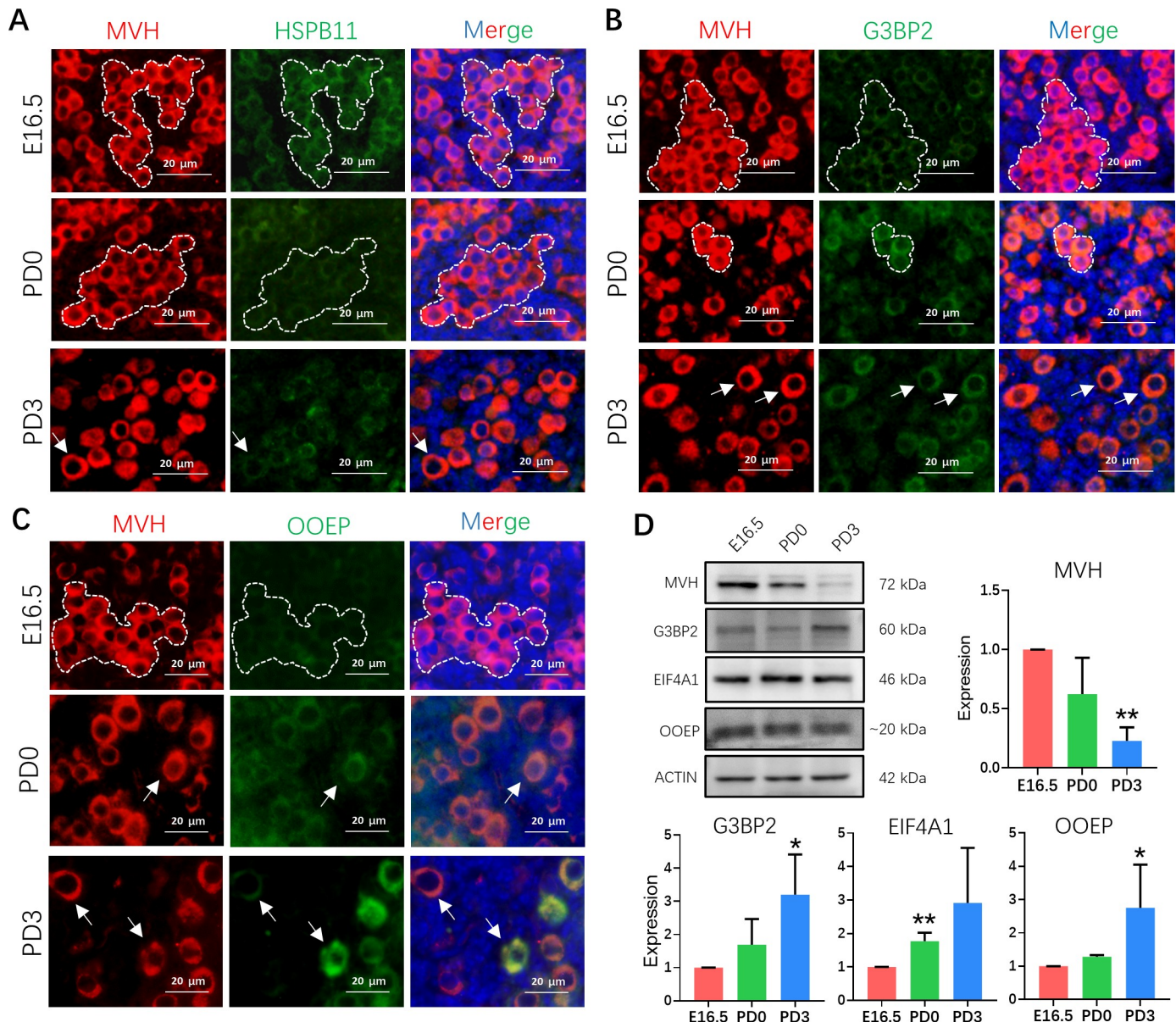
cell percentage expressed in each cluster; dot color marks the average expression level of each gene. The sequencing data was deposited available in GSE134339, and this figure can be produced using scripts at [https://github.com/WangLab401/2020scRNA\\_murine\\_ovaries](https://github.com/WangLab401/2020scRNA_murine_ovaries). UMAP, uniform manifold approximation projection.

<https://doi.org/10.1371/journal.pbio.3001025.g002>

and early follicle formation stages and was significantly decreased for germ cells within follicles; sets 4 and 5 comprised genes specifically up-regulated in germ cells in nests. The expression kinetics of each gene set are shown in Fig 4E. Enriched Gene Ontology (GO) terms of the most expressed 100 genes of these gene sets are shown in Fig 4F (S4 Table). For example, gene set 1 enriched “circadian regulation of gene expression, piRNA metabolic process, female gonad development, and cell division”; gene set 2 contained genes involved in “histone deacetylation, and cell cycle phase transition”; gene set 3 enriched genes in “chromatin organization, mRNA processing, sister chromatid cohesion, and cellular response to DNA damage stimulus”; and gene set 4 enriched genes of “regulation of chromosome organization, gene silencing, mRNA metabolic process, and in utero embryonic development”, as well as “chromatin organization”; importantly, gene set 5 were most related to “female gamete generation, DNA repair, regulation of RNA splicing, and regulation of gene expression, epigenetic.” Meanwhile, the representative genes in each set are displayed with pseudotime trajectories in Fig 4G. Specifically, representative genes enriched in GO terms of “histone deacetylation” (*Hdac2* and *Morf4l1*) and “regulation of gene expression and epigenetic” (*Morc2a* and *Smarca4*) were highlighted (S2D and S2E Fig). To further characterize the germ cell fate transition from nests to follicles, gene sets 1 and 2 that were dramatically increased in follicular germ cells were merged and analyzed with Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway, and the results showed abundant transcripts for “RNA transport, spliceosome, oxidative phosphorylation, nucleotide excision repair, and citrate cycle (TCA cycle)” (S2F Fig). Furthermore, KEGG analysis of genes in sets 4 and 5 that were actively transcribed in germ cells of the nest stage enriched “ribosome, regulation of actin cytoskeleton, focal adhesion, and adherens junction,” and several pathways, including “Ras, Rap1, mTOR, Wnt, insulin, and AMPK signaling” (S2G Fig and S5 Table).

### Mapping oocyte-specific regulon networks by SCENIC

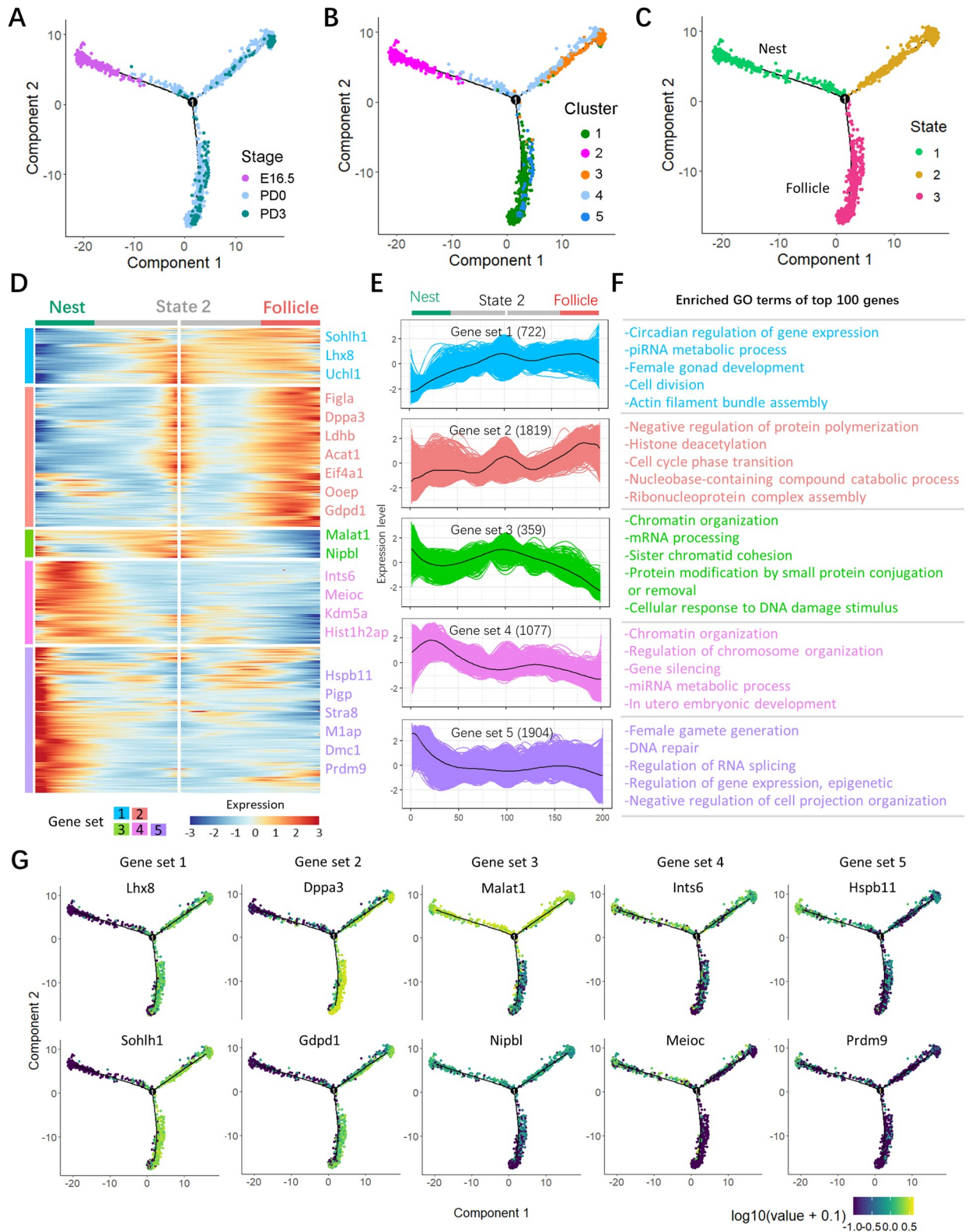
We further investigated the regulons (i.e., transcriptional factors (TFs) and their target genes) activity of germ cell-specific TFs using SCENIC, an algorithm developed to deduce global research networks (GRNs) and the cellular status for scRNA data [39]. Based on 201 regulons activity with 9,208 filtered genes with default filter parameters, the selected cells were clustered with cell state (right, pseudotime analysis label) and developmental stages (left) in t-distributed stochastic neighbor embedding (t-SNE) method calculated by SCENIC (S3A Fig), and the regulons density was also mapped to t-SNE plots (S3B Fig). Accordingly, regulon activity was binarized and matched with germ cell state and developmental stages (left panel); in addition, some representative regulons and their motifs (right panel) were listed (Fig 5A). As shown, a series of TFs displayed a more dynamic pattern. For example, *Nr3c1*, *Brca1*, *Kdm5a*, and *Kdm5b* were active mainly at the germ cell nest stage (Fig 5B and S3C Fig), and *Klf7* and *Gabpa* seemed to be mostly activated at the nest stage, but gradually turned off as follicular assembly occurred (Fig 5C). Similarly, *Smarca1*, *Hes1*, *Sox15*, and *Stat3* were mainly associated with states 2 and 3, corresponding to the period of PF assembly (Fig 5D and S3D Fig). Furthermore, a portion of TFs appeared to serve as a transient switch, such as *Foxo1* had a limited role in germ cells within nests and *Foxo3* functioned during a period of the follicular stage (Fig 5E). Moreover, several TFs, such as *Jun* and *Klf2*, were significantly represented during all oocyte stages, which supports a role for these TFs as likely candidates for sustaining a germ cell-specific transcription program throughout the investigated period (S3E Fig).



**Fig 3. Marker gene detection of germ cells at pre-, early- and late-follicle formation stages.** (A) Representative images of the HSPB11 staining of ovarian sections from E16.5 embryos, and PD0 and PD3 pups. Germ cells were marked by MVH (red); HSPB11 was stained with green; and nucleus was counterstained with Hoechst 33342 (blue). Dotted white line label the germline nest and white arrow mark the germ cell in follicle. Scale bar: 20  $\mu$ m. (B) Representative images of the G3BP2 of ovarian sections from E16.5 embryos, and PD0 and PD3 pups. Germ cells were marked by MVH (red); G3BP2 was stained with green; and nucleus was counterstained with Hoechst 33342 (blue). Dotted white line label the germline nest and white arrow mark the germ cell in follicle. Scale bar: 20  $\mu$ m. (C) Representative images of the OOEP staining of ovary sections from E16.5 embryos, and PD0 and PD3 pups. Germ cells were marked by MVH (red), OOEP was stained with green and nucleus was counterstained with Hoechst 33342 (blue). Dotted white line label the germline nest and white arrow mark the germ cell in follicle. Scale bar: 20  $\mu$ m. (D) Detection of G3BP2, EIF4A1, OOEP, ACTIN, and MVH proteins of E16.5 embryos, and PD0 and PD3 pups by western blot. ACTIN (only for MVH expression), and MVH was used as a loading control. Data were represented with mean  $\pm$  SD ( $n = 3$  for independent repeats). The relative expression level was calculated between PD0 or PD3 and E16.5. Unpaired  $t$  tests are performed. Statistical significance is shown as \* $P < 0.05$ ; \*\* $P < 0.01$ . The raw data image can be found in [S1 Data](#). The raw data used for quantification of D can be found in [S2 Data](#). EIF4A1, eukaryotic translation initiation factor 4A1; E16.5, embryonic day 16.5; G3BP2, GTPase-activating protein-binding protein 2; HSPB11, heat shock protein family B member 11; MVH, mouse vasa homologue; OOEP, oocyte expressed protein; PD0, postnatal day 0; PD3, postnatal day 3.

<https://doi.org/10.1371/journal.pbio.3001025.g003>





**Fig 4. Transition of germ cells throughout distinct states from nest to follicle stages.** (A) Single-cell trajectories of germ cell as a function of developmental timeline. (B) Single-cell trajectories of germ cells along with Seurat cluster. (C) Single-cell trajectories of the 3 germ cell states through the pseudotime. (D) Heatmap representing the expression dynamics of 5 gene sets with increased expression or reduced at the nest and follicle stage. The gene sets were selected with  $qval < 1 \times 10^{-4}$ . Detailed information in gene sets is shown in S3 Table. (E) Expression profiles of

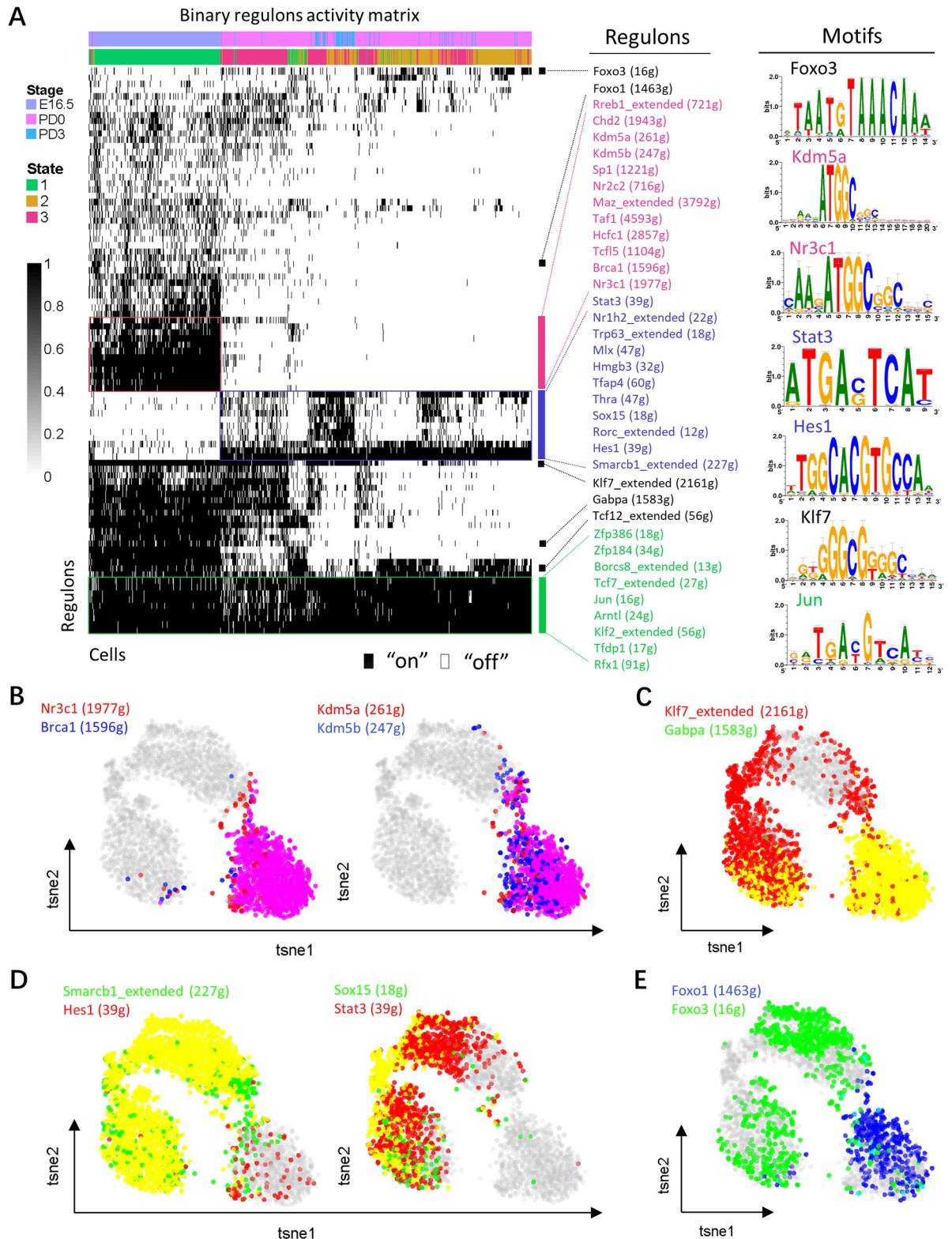
the differentially expressed genes in the transition from nest to follicle in each set. (F) GO terms of top 100 differentially expressed genes in each set. More GO terms are shown in [S4 Table](#). (G) Expression of representative genes in each set along single-cell pseudotime trajectories. The sequencing data was deposited available in GSE134339, and this figure can be produced using scripts at [https://github.com/WangLab401/2020scRNA\\_murine\\_ovaries](https://github.com/WangLab401/2020scRNA_murine_ovaries). GO, Gene Ontology.

<https://doi.org/10.1371/journal.pbio.3001025.g004>

## Gene expression signatures of the granulosa cell lineage

When the transcriptome of granulosa cells was plotted along with the developmental stages, 6 cell clusters were identified ([Fig 6A](#)). As noted, novel marker genes within each cluster were identified ([S6 Table](#)), whose expressions varied according to the developmental stages ([S4A and S4B Fig](#)). For example, *Wnt4* expression was most activated at E16.5, which may imply the differentiation of ovarian pre-granulosa cell and preparation of the second wave of follicle formation [[19](#), [40](#)]. Another, which showed a dramatic increase postnatally was *Aldh1a1*, a stem cell marker [[41](#), [42](#)], reported to be essential for the biosynthesis of retinoic acid (RA, a key regulatory factor of meiosis) [[43](#)]. At PD3, *Serpine2* is a gene encoding a member of the serpin family of proteins, a group of proteins that inhibit serine proteases and are also expressed by granulosa cells in human adult ovary [[44](#)].

Cell pseudotime trajectories revealed 3 granulosa cell states and 1 branch point ([Fig 6B](#)). During mouse ovary development, 2 waves of granulosa cell differentiation occur, characterized by partly distinct gene expression [[15](#)]; therefore, on the basis of the gene expression reported in [Fig 6C and 6D](#), we attributed state 1 cells to early pre-granulosa cells expressing *Foxl2* and its downstream gene *Cdkn1c* [[45](#), [46](#)]. At state 2, another differentiating pre-granulosa cells expressed *Lgr5*, *Gng13*, and *Krt19*, and at state 3, the pre-granulosa cells expressed *Hsd3b1*, *Aard*, and *Akr1c14*, as well as *Foxl2*. These results suggested that state 1 cells were early bipotential pre-granulosa cells (early BPGs); that state 2 cells were differentiating epithelial pre-granulosa cells (D-EPGs); and that state 3 cells were differentiating BPGs (D-BPGs) [[47](#)]. For further characterization of the granulosa cell lineage, the differentially expressed genes of the 3 states were divided in 4 clusters according to their distinct expression patterns, and a heatmap was generated ([Fig 6E](#), [S7 Table](#)). Gene sets 1 (460 genes) and 2 (722 genes) were assigned to state 3 (D-BPGs); set 3 (687 genes), to state 1 (early BPGs); and set 4 (1331 genes), to state 2 (D-EPGs). On this basis, *Cdkn1c* (set 3) expression decreased from early BPGs to D-BPGs; *Upk3b*, *Krt19*, and *Aldh1a2* (set 4) were high in D-EPGs and declined in D-BPGs; and *Aard*, *Hsd17b1*, *Inha*, and *Hsd3b1* (sets 1 and 2) were present at low levels in D-EPGs and high levels in D-BPGs ([Fig 6E](#)). Moreover, the expression dynamics of representative genes in each set are showed in [Fig 6F](#). Furthermore, GO analysis of the enriched top 100 genes in each set was performed ([Fig 6G](#)). Gene set 1 was characterized by the expression of genes associated with biological processes, such as “regulation of cellular response to growth factor stimulus,” “positive regulation of apoptotic process,” and “negative regulation of cell differentiation,” that were more highly expressed in D-BPGs and early BPGs than D-EPGs. Gene coding members of “c21-steroid hormone metabolic process” were among of set 2, confirming that c21-steroid hormones likely play a relevant role in PF formation. Whereas set 3 was composed of genes encoding proteins involved in “mitotic cell cycle process, muscle structure development, and epithelial cell proliferation,” which both highly expressed in early BPGs and D-EPGs, set 4 consisted of genes enriched in “epithelial cell proliferation, endothelial cell migration, and cell–cell adhesion”, which were more highly expressed in D-EPGs ([S8 Table](#)). In line with this, when sets 1 and 2 genes were analyzed by KEGG, they were highly enriched in “focal adhesion, oxytocin, and PPAR signaling” and cytochrome P450 related terms, as well as “ovarian steroidogenesis” in BPGs ([S4C Fig](#)). KEGG performed on sets 3 and 4 showed that the most represented terms of signaling pathways in EPGs were those of “PI3K-Akt, MAPK,



**Fig 5. Oocyte transcription factors involved in follicle assembly.** (A) Heatmap of regulon activity analyzed by SCENIC with default thresholds for binarization. The “regulon” refers to regulatory network of TFs and their target genes. “On” indicates active regulons; “Off” indicates inactive regulons. The top rows represent the transcriptome signature of selected germ cells, and they refer to the time point label and state character of Monocle (left panel). The representative TFs and part of their motifs involved in follicle assembly were listed in the

right panel. The numbers in parenthesis next to the regulon names indicate the number of genes enriched in regulons. (B) t-SNE projection of average binary regulon activity of *Nr3c1*, *Brca1*, *Kdm5a* and *Kdm5b* that were stage specific in germ cells in nests. (C) t-SNE projection of average binary regulon activity of *Klf7* and *Gabpa* that were most active in germ cells within nests. (D) t-SNE projection of average binary regulon activity of *Smad3*, *Hes1*, *Sox15*, and *Stat3* that were specific in follicular oocytes. (E) t-SNE projection of average binary regulon activity of *Foxo1* and *Foxo3* from the nest stage to follicle stage in germ cells. The sequencing data was deposited available in GSE134339, and this figure can be produced using scripts at [https://github.com/WangLab401/2020scRNA\\_murine\\_ovaries](https://github.com/WangLab401/2020scRNA_murine_ovaries). SCENIC, single-cell regulatory network inference and clustering; TF, transcriptional factor; t-SNE, t-distributed stochastic neighbor embedding.

<https://doi.org/10.1371/journal.pbio.3001025.g005>

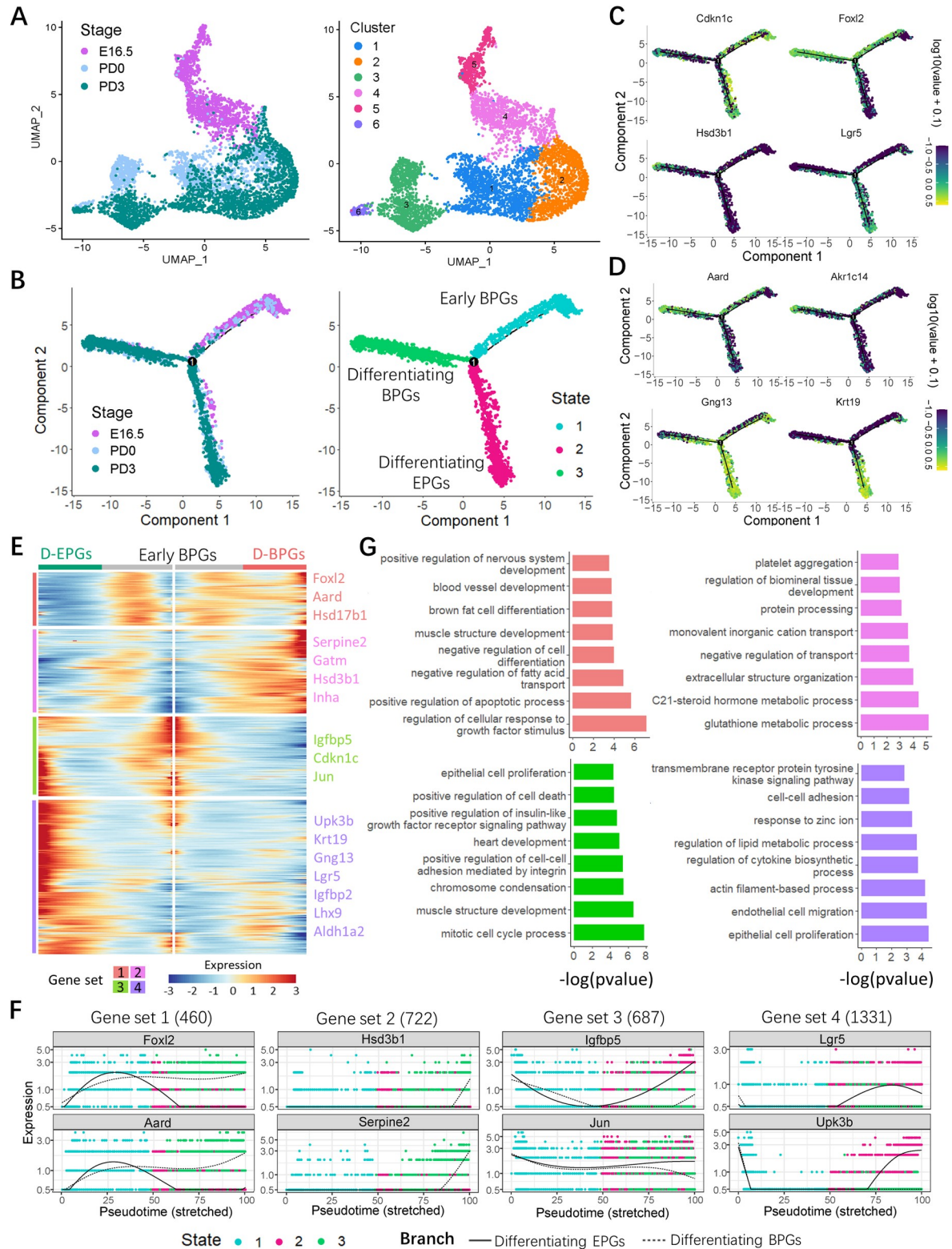
and Hippo” and to a lesser extent of “Rap1” (a Ras-related protein), while “ribosome, focal adhesions, tight junctions, and cell cycle” were among those more accounted for cellular functions (S4D Fig, S9 Table).

### Interactions between germ cells and granulosa cells during primordial follicle assembly

Since interactions between oocytes and granulosa cells are crucial during nest breakdown and PF assembly [35], and throughout the entire period of folliculogenesis, we compared the transcriptomes of germ cells and granulosa cells with each other according to the developmental stages, to verify the involvement of previously discovered players and identify new cell signaling pathways driving such interactions.

As seen in S5 Fig, and as previously reported, Vlnplots confirm that during the period of investigation, NOTCH signaling, members of the TGF-beta family, Kit/Kitl system and gap junction are important components of the oocyte–granulosa cell interaction [15,35]. In this regard, we found that genes encoding NOTCH ligands such as *Dll3* and *Jag1-2* were expressed by germ cells mainly at PD0 and/or PD3, respectively, while that for its receptor *Notch2* showed persistent expression in granulosa cells. The NOTCH-targeted gene *Hes1* showed transient expression in oocytes at PD0 and PD3 and resulted in ubiquitous expression in granulosa cells at all stages, whereas *Rbpj* was more highly expressed by germ cells at all stages and granulosa cells at PD0 and PD3 (S5A Fig). As expected, TGF-beta members, such as *Gdf9*, were mainly found in oocytes postnatally, while the gene encoding their receptors *Bmpr1a* and *Bmpr2* were expressed by oocytes in a paralleled manner and continually by granulosa cells. Genes encoding the transducers homologues of *Drosophila* protein, mothers against decapentaplegic and *Caenorhabditis elegans* protein sma (SMADs) were examined, and *Smad2* were expressed in both postnatal germ cells and granulosa cells. At the same time, there was no *Smad3* expression and decreasing levels of *Smad5* in germ cells, while *Smad3* and *Smad5* were found in granulosa cells throughout the investigated period. Finally, target *Id* genes were ubiquitously expressed at all stages both by germ cells and granulosa cells (S5B Fig). Again, as expected, *Kit* was similarly expressed by oocytes after birth, and *Kitl*, by granulosa cells at all stages (S5C Fig). Lastly, *Gjc1* and *Gja1* were found in oocytes and granulosa cells, respectively (S5D Fig), which encode gap junction proteins CX45 and CX43, supporting the participation of gap junction into the oocyte–granulosa cell interaction during the entire period of the study.

KEGG-enriched analyses were used to identify ongoing bioprocesses mainly in germ cells or granulosa cells or both (S10 Table). For example, “ribosome and RNA transport,” followed by “autophagy-animal, cell cycle, and oocyte meiosis,” were primarily found in germ cells; whereas “PI3K-AKT, MAPK, Rap1, and Hippo signaling pathways,” followed by “focal adhesion, tight junction, and cell cycle,” were mostly represented in granulosa cells (Fig 7A and 7B). Eight out of 25 processes found in both cell types referred to reproductive processes according to previous works [15,24] (S6A and S6B Fig), including “oocyte meiosis, cell cycle,



**Fig 6. Molecular characterization of granulosa cell subsets.** (A) Cluster analysis of granulosa cells with UMAP plots based on developmental timeline (left) and cell clusters (right). (B) Single-cell trajectories of granulosa cell subsets as a function of developmental timeline and cell States. (C) Expression of developmental marker genes along with pseudotime trajectories. (D) Expression of *Aard*, *Akr1c14*, *Gng13*, and *Krt19* along with pseudotime trajectories. (E) Heatmap of the gene expression programs of pre-granulosa cell with

different states. The gene sets were selected with  $qval < 1 \times 10^{-4}$ . Detail information in gene sets is shown in [S7 Table](#). Four gene sets represent increases or decreases in differentially expressed genes; part of the representative genes is listed in the right panel. “D-BPGs” refers to “Differentiating BPGs” and “D-EPGs” refers to “Differentiating EPGs”. (F) Expression of representative genes in each set with 2 branches of pre-granulosa cell. (G) Enriched GO terms of the top 100 genes of each set. More GO terms are shown in [S8 Table](#). The sequencing data was deposited available in GSE134339, and this figure can be produced using scripts at [https://github.com/WangLab401/2020scRNA\\_murine\\_ovaries](https://github.com/WangLab401/2020scRNA_murine_ovaries). BPGs, bipotential pre-granulosa cells; D-BPGs, differentiating BPGs; D-EPGs, differentiating EPGs; EPGs, epithelial pre-granulosa cells; GO, Gene Ontology; UMAP, uniform manifold approximation projection.

<https://doi.org/10.1371/journal.pbio.3001025.g006>

adherens junction, and tight junction,” and signaling pathways of “FoxO, Hippo, estrogen and p53.”

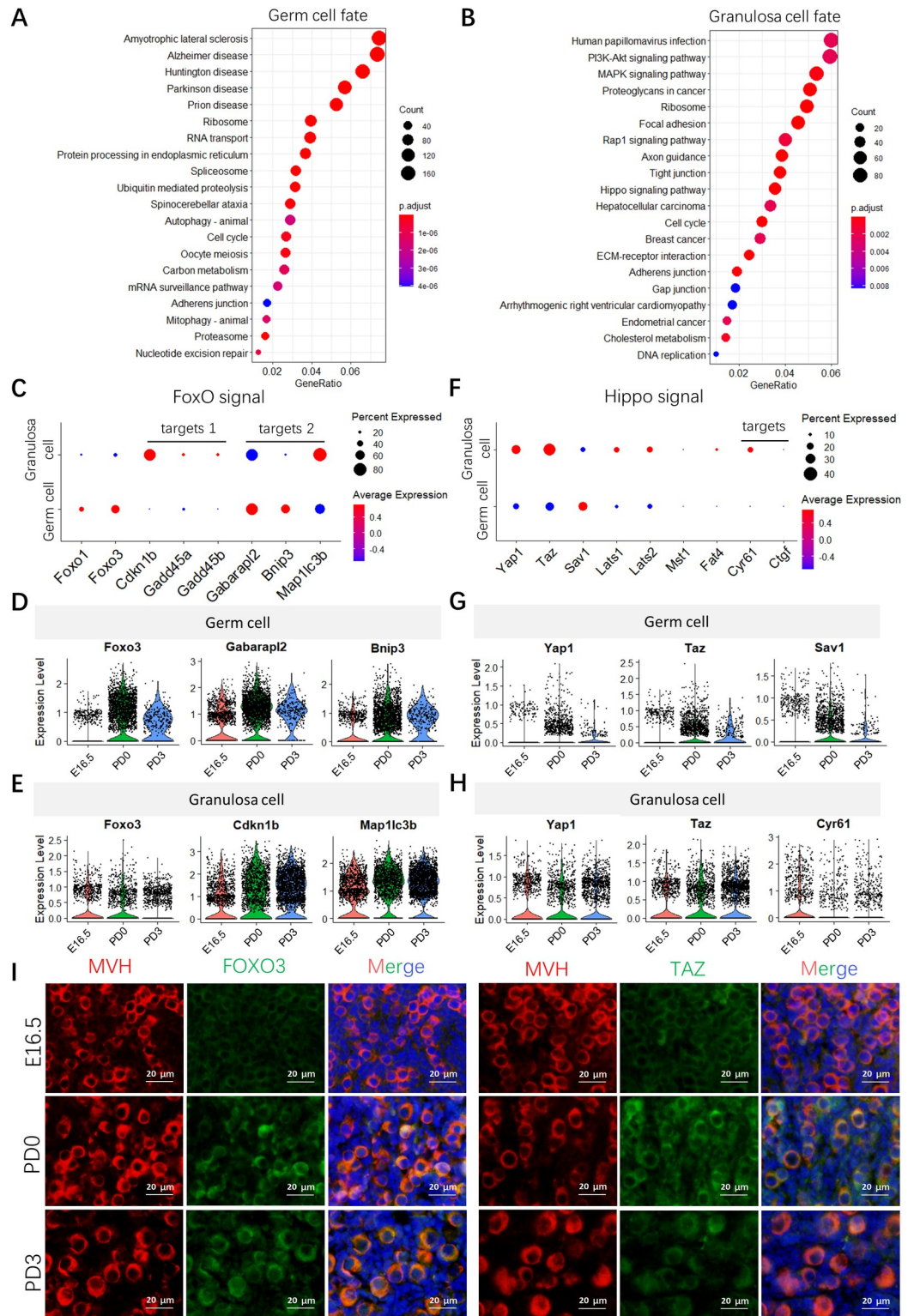
To characterize the crosstalk between the 2 cell types, germ cells and granulosa cell transcriptome clusters were combined. Concerning FoxO signaling, *Foxo1* and *Foxo3* were the most common transcripts mainly expressed postnatally in oocytes and to a lesser extent in granulosa cells ([Fig 7C–7E](#)). Among FoxO downstream genes involved in cell proliferation (target 1), *Cdkn1b* was expressed at a higher level in granulosa cells at PD3, while autophagy-related genes (target 2), such as *Gabarapl2*, *Map1lc3b*, and *Bnip3*, were expressed in both cell types throughout the examined stages although at different levels ([Fig 7C–7E](#)). As for Hippo signaling ([Fig 7F](#)), its key effectors, both *Yap1* and *Taz*, were more highly expressed in granulosa cells at all stages than in postnatal germ cells. For its downstream regulators, only *Sav1* had higher levels in germ cells, others, such as *Lats1* and *2*, were more highly expressed in granulosa cells postnatally; however, its target *Cyr61* is most expressed at E16.5 ([Fig 7G and 7H](#)). Moreover, the expressions of FOXO3 and TAZ in germ cells were confirmed at the protein level by immunohistochemical staining ([Fig 7I](#)), and they were mainly expressed at the postnatal stage, which verified the data analysis results. Moreover, the transcripts for tight junction were expressed at the postnatal stage in both granulosa cells and germ cells ([S6C and S6D Fig](#)), while those for adherens junction proteins appeared to follow constantly disparate patterns, and they were present in abundance throughout almost all the timeline ([S6E and S6F Fig](#)).

### Transcriptional signatures of ovarian somatic cells during primordial follicle assembly

Specially, the transcriptional signatures of ovarian somatic cell types were further explored. As shown in [S7A Fig](#), the stromal cell line was clustered, and 5 clusters were produced, which all expressed the specific marker genes of stromal cells ([S7B Fig](#)), while the novel marker genes in each cell cluster were identified ([S7C Fig](#)). Additionally, the cell clustering analyses of endothelial cells ([S7D Fig](#)) and immune cells ([S7E Fig](#)) were performed, and 3 and 4 clusters were generated, respectively. More notably, compared with transcriptome patterns of stromal and endothelial cells, whose transcriptome indicated the distinct patterns dependent on pre- and postnatal manners, the transcriptional signature of immune cells was distinguished from them both, and it seemed to be more irregular. However, the potential roles of ovarian somatic cells participating in PF assembly still need to be addressed.

### Discussion

In the present study, we used scRNA-seq analyses to identify the complete transcriptome programs of the different cell types present in mouse ovaries from the end of fetal age through the first few days after birth. As reported in the Introduction, in most mammalian species, this period is crucial for ovarian development and is characterized by several processes involving all ovarian cell populations. In this regard, our analyses were able to identify 7 distinct cell types, including germ cells, granulosa cells, stromal cells, epithelial cells, erythrocytes, immune



**Fig 7. Enrichment molecular pathways of oocytes and granulosa cell subsets during follicle assembly.** (A and B) Pathway enrichment of key genes regard to fate transitions of germ cells (A) (total of 5,881 genes at branch point) and granulosa cell (B) (total 3,200 genes at branch point). More pathways are shown in S10 Table. (C) Dot plots of FoxO signal and its target genes in germ cell and granulosa cell clusters. (D and E) Expression of representative genes of the FoxO signal in germ cells (D) and granulosa cells (E). (F) Dot plots of the Hippo signal in germ cells and granulosa cells. (G and H)

Expression of representative genes of the Hippo signal in germ cells (G) and granulosa cells (H). (I) Representative images of FOXO3 and TAZ immunohistochemistry staining of ovarian sections from E16.5 embryos and PD0 and PD3 pups. Germ cells were marked by MVH (red); FOXO3 and TAZ were stained with green; and nucleus was counterstained with Hoechst 33342 (blue). Scale bar: 20  $\mu$ m. The sequencing data was deposited available in GSE134339, and this figure can be produced using scripts at [https://github.com/WangLab401/2020scRNA\\_murine\\_ovaries](https://github.com/WangLab401/2020scRNA_murine_ovaries). E, embryonic day; FOXO3, Forkhead box O3; PD, postnatal day; TAZ, tafazzin family protein.

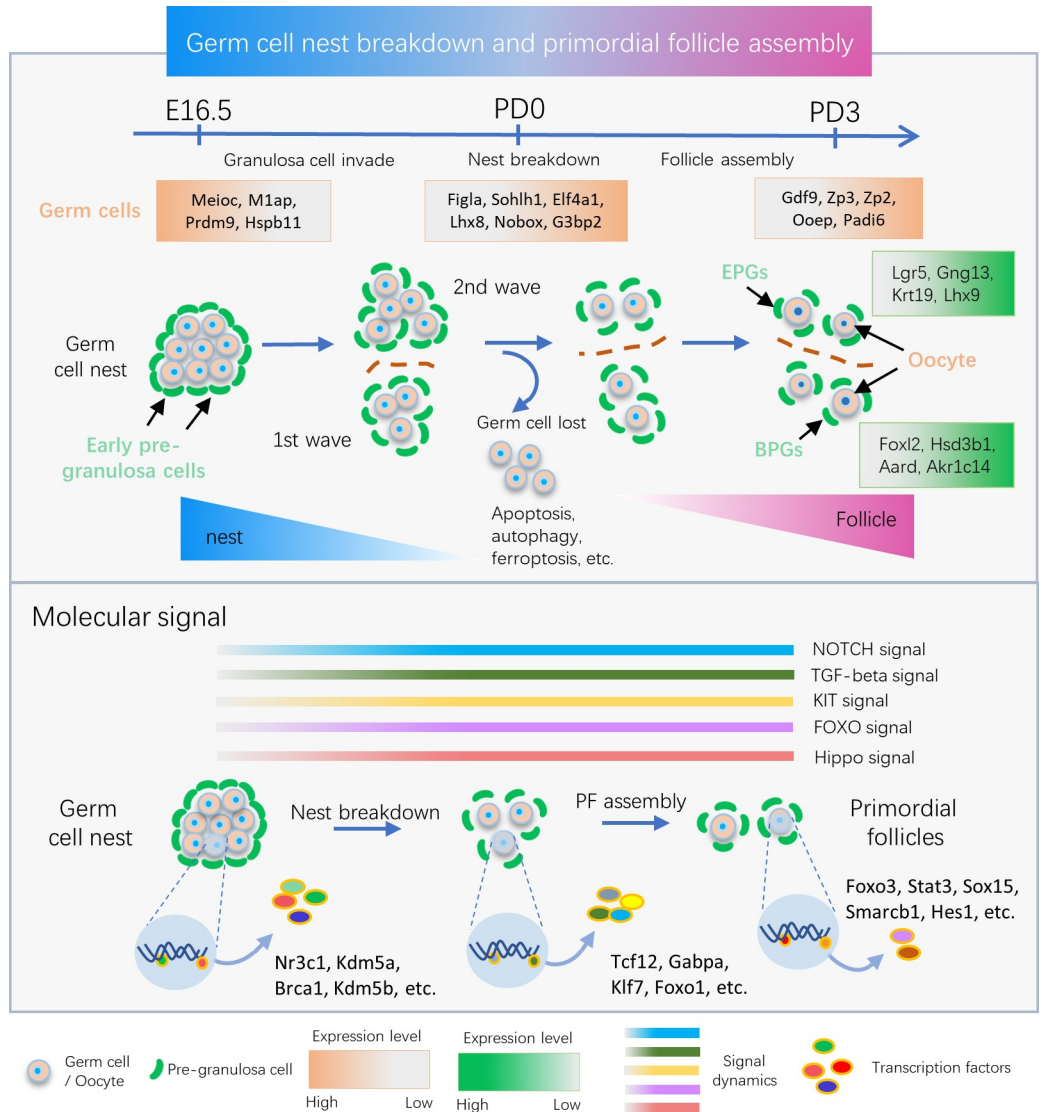
<https://doi.org/10.1371/journal.pbio.3001025.g007>

cells, and endothelial cells that underwent dynamic changes throughout the period of investigation (Fig 1B–1E). These changes likely reflect the dynamics of proliferation, differentiation, and death of the different cell populations that, with regards to germ cells, are characterized by nest breakdown associated with extensive degeneration. In fact, some ovarian somatic cells, such as stromal (S7A–S7C Fig), endothelial cells (S7D Fig), and immune cells (S7E Fig); however, due to the limited information available regarding other ovarian somatic compartments, the current study placed most attention on germ cells and granulosa cells.

Based on the expression of known cell lineage and developmental stage-specific genes, we generated 5 different cell clusters for germ cells. According to this analysis, a genetic dynamic model, including pre-, early- and late-follicle formation stages of germ cell development, and the changes of transcriptome signature in germ cell at these stages were delineated (Fig 2A–2D). Genes known to be related to meiotic processes, such as *Prdm9* [31], showed widespread high expression in cluster 2, as well as *Stra8*, encoding a protein crucial for the beginning of meiosis [48], only as expected since its expression was reported to be limited to the early stages of meiotic prophase 1 [49]. With the exception of *Figla* that was expressed at variable levels in all clusters, also other genes, such as *Lhx8*, *Nobox*, and *Sohlh1* (encoding transcription factors reported to be typical of follicular oocytes [15]), appeared restricted to a few clusters (1, 3, and 4), and assigned to the early follicle stage. Finally, *Gdf9* (encoding a growth factor of the TGF-beta family [50]), *Zp2*, and *Zp3* (encoding zona pellucida proteins [35,36]) showed an even more restricted expression that was limited mainly to cluster 5, corresponding to the late follicle formation stage. All these gene dynamics potentially play an important role in PF formation (upper panel in Fig 8).

Overall, the above results indicated the validity of our analyses and made us confident that we had identified several new genes expressed by germ cells at the 3 designated stages, whose importance and role can be the focus of future studies. For example, at the pre-follicle stage, *Mlap* and *Hspb11*, a protein required for meiosis I arrest during spermatogenesis [51] and a heat shock protein that inhibits cell death through stabilization of the mitochondrial membrane [52], respectively, are likely candidates for a role in the control of female germ cell meiosis and in their survival/apoptosis at the pre-follicle stage. At the early follicle stage, the expression of the following genes suggest an increase of both general and specific transcriptions in oocytes: *Eif4a1*, encoding an ATP-dependent RNA helicase, a subunit of the eIF4F complex involved in cap recognition and required for mRNA binding to ribosomes [53]; *G3bp2*, encoding an ATP- and magnesium-dependent helicase, a member of nuclear RNA-binding proteins [54], and *Id1*, encoding a helix-loop-helix (HLH) protein forming heterodimers with the basic HLH family of transcription factors that may play a role in cell growth, senescence, and differentiation [55]. At the same time, the expression of 2 genes: *Ooep* (which encodes an RNA binding protein of the subcortical maternal complex (SCMC) of the oocyte and likely participates oocyte homologous recombination [56,57]) and *Padi6* (which encodes a member of the peptidyl arginine deiminase family of enzymes that may play a role in cytoskeleton reorganization in the egg and in early embryo development [58–60]) both support findings that the oocytes begin to express these proteins at the late stage of PF assembly. It is important to note that the experimental approach used in the present paper identifies potential





**Fig 8. Schematic scheme of the main genetic and molecular signal dynamics detected in ovarian cells during follicle assembly.** The upper panel indicates the genetic programs and key biological events of germ cells and pre-granulosa cells during germline nest breakdown and follicle assembly, and the developmental time courses are indicated above; the orange box contains the marker genes and their expression dynamics in germ cell; the green ones comprise the marker genes and their expression dynamics in pre-granulosa cell. The dark orange dashed line separates ovarian cortex (above) from the medulla, labeled with 2 sources of pre-granulosa cells, i.e., BPG cells for the first wave in ovarian medulla and EPG cells in ovarian cortex for the second wave of follicle formation, respectively. In the bottom panel (upper part), the activation of 5 main interactive pathways between germ cells and granulosa cells during these processes is represented. In the lower part, the stage-specific TFs in oocytes from nest to follicle stage are reported. BPG, bipotential pre-granulosa; EPG, epithelial pre-granulosa; TF, transcriptional factor.

<https://doi.org/10.1371/journal.pbio.3001025.g008>

candidate genes involved in the investigated events (upper panel in Fig 8). Further analysis at an individual gene level is required to determine the specific function of a given gene product.

By ordering germ cells along pseudotime trajectories, we were able to reconstruct the gene dynamics involved in the transition of oocytes from the pre-follicle stage (mostly germ cells in nests) to the follicle stage (PF assembly) throughout 3 distinct differentiation states. The majority cells at pre-follicle stage belonged to state 1, most of which were germ cells in nests, while those at early- and late-follicle stages were in states 2 and 3 and were considered as

follicular oocytes. Five gene sets were identified including those genes likely involved in the commitment of germ cells present in nests into follicular oocytes. Among these, the most differentially expressed were *Sohlh1*, *Lhx8*, and *Uchl1* (set 1) and *Figla*, *Dppa3*, *Eif4a1*, and *Ooep* (set 2), with moderately and highly increased expression, respectively, in early and late follicular oocytes. *Malat1* and *Nipbl* (set 3) were highly expressed in germ cells in nests and early-follicle formation oocytes; *Ints6*, *Meioc*, *Kdm5a*, and *Hist1h2ap* (set 4) were dramatically up-regulated in germ cells in nest. Enriched GO applied to the 100 most highly expressed genes of these gene sets involved in specific biological process (Fig 4F). Of note, the GO terms of “histone deacetylation” and “regulation of gene expression and epigenetic” were enriched, and these results indicated the epigenetic modification of germ cell likely involved fate decisions of its own during follicle formation. Given the fact of germ cell loss, apoptosis has been the most commonly identified process; however, the present scRNA-seq platform does not excel in the determination of apoptotic cells, and dying cells are regarded as low quality and therefore needed to be discarded; another known mechanism of cell autophagy, which was enriched (Fig 7A); it is worth noting that ferroptosis, 1 form of cell death, was revealed in the current study (S10 Table), and its roles in oocyte loss deserves to be investigated further.

Moreover, using the SCENIC algorithm, we were able to establish a network of regulons, that can be postulated as likely candidates for sustaining germ cell specific TF programs throughout the period of investigation (lower panel in Fig 8). Through using this method, we revealed several TFs at distinct stages, such as *Nr3c1*, *Foxo1*, *Brcal*, *Kdm5a*, and *Kdm5b* at the germ cell nest stage; *Smrbc1*, *Hes1*, *Stat3*, and *Sox15* during the period of PF assembly; and *Foxo3* mainly limited postnatal period (Fig 5). It is worth noting that *Nr3c1* has been revealed as the meiotic specific TF in oocyte; furthermore, *Brcal*, *Kdm5a*, and *Kdm5b* have previously been detected in meiotic oocyte [61]. Meanwhile, recent studies have shown that FOXO3 plays a critical role in the maintenance of PF dormancy and *Foxo3* knockout lead to follicles overactivation [62–64], and the results indicate *Foxo3* is a suppressor of PF activation and may be required for their formation. In addition, SMARCB1 is essential for embryo development [65]; however, its role in PF has been unknown, which need further investigated.

Bioinformatics analyses of the granulosa cell transcriptome identified 6 clusters and 3 states (Fig 7B). As for the genes known to be expressed by the granulosa lineage during the period of investigation, it showed that some of these genes, such as *Foxl2* and *Lgr5*, varied considerably at different stages. Moreover, novel genes within each cluster were identified, whose expressions changed along with the developmental stages. Cell pseudotime trajectories revealed 3 granulosa cell states and 1 branch point (Fig 6B and 6C). Following the considerations reported in the results, state 1 was attributed to early pre-granulosa cells, and it branched at state 2 of pre-granulosa cells expressing *Lgr5* and *Gng13* and at state 3 of late stage pre-granulosa cells expressing *Foxl2* and *Hsd17b1* that are required for steroidogenesis [30]. Actually, the origin of granulosa cell is always indefinite, which was hypothesized to be derived from mesonephros or ovarian surface epithelial cells [66]. As recently reported [47], there are 2 granulosa cell differentiation pathways supporting follicle formation, which are marked by *Foxl2* and *Lgr5* expressions, and they show different transcriptional patterns and are involved in 2 waves of follicle formation. In our results, the 2 distinct cell populations above were indeed confirmed, and analogously, state 3 cells with high expression of *Foxl2*, *Aard*, and *Akr1c14* are suggested to be BPGs supporting follicle in the ovarian medulla; meanwhile, the EPGs are for oocytes destined for the ovarian cortex, as evidenced by the expression of *Lgr5*, *Gng13*, and *Krt19* at state 2 (Fig 6D). Thus, the consistent results in 2 studies further evidenced the notion that 2 kinds of pre-granulosa cells supporting follicle formation respectively at ovarian medulla and cortex in mice.

Further, bioinformatics analyses of germ cell and granulosa cell transcriptomes paralleled in function with the developmental stages confirmed that NOTCH signaling, members of the TGF-beta family, Kit/Kitl system and gap junctions, are important components of the oocyte-granulosa cell interactions during the period of investigation (lower panel in Fig 8). In addition, it showed that ligands, receptors, and signaling mediators of these system were coherently expressed mainly in germ cells, pre-granulosa cells, or both, as detailed in the results. Moreover, basically in line with the finding discussed above, biological processes, such as “RNA transport, cell cycle, and meiosis,” were primarily found in germ cells, whereas “proteoglycan and cholesterol metabolism” were more commonly represented in granulosa cells and “ribosome biogenesis” in both cell types. Hippo pathways were mostly represented in granulosa cells, and FoxO signaling were present in both cell types although at different developmental stages. The components of Hippo signaling in granulosa cells have been reported in several mammalian species [67–69]. Previous reports suggest Hippo kinases large tumor suppressor kinase (LATS) 1 and 2 are required for ovarian granulosa cell fate maintenance [70] and knockout of yes-associated protein 1 (Yap1) in granulosa cells block follicle development [71]. Furthermore, disruption of the Hippo pathway has been reported to promote AKT-stimulated ovarian follicle growth [72]. Here, we reported its possible involvement in the early stages of granulosa cell development. As for FoxO, while a higher level of transcripts for *Foxo1* and *Foxo3* was present in postnatal oocytes, *Cdkn1b* was expressed at a higher level in granulosa cells at PD3, and autophagy-related genes (*Gabarapl2*, *Map1lc3b*, and *Bnip3*) were expressed in both cell types throughout the examined stages although at different levels. This suggests that FoxO signaling is not only important in controlling the cell cycle [73] and survival of the postnatal oocytes as reported in previous studies [74,75], but is involved in granulosa cell proliferation and autophagy.

Furthermore, despite the fact that the single-cell transcriptome technique is greatly expanding our insights into biological systems, especially regarding cellular heterogeneity, there are some unsolved issues that still need addressing within the field. Firstly, it is the current lack of standardization for data analysis [76]. The problem is triggered by the growing number of analysis methods, exploding dataset sizes, and technical aspects (such as a variety of programming languages for data analysis tools), all of which prevent computational biologists from developing a gold standard practices. Another is regard to the incorporate of biological replicates [77,78]. Although the biological replicates are suggested to eliminate batch effects, it seems the 1 replicate is still more popular in the single-cell dataset of recent publications [47,61]. In the present study, the application of 1 replicate may be considered as a potential limitation; however, the canonical knowledge in the field of PF formation, as well as the experimental verification, makes us confident for our results. Moreover, the cell proportions studied here, at least partially, do not reflect the true cell type compositions in ovarian tissue (such as germ cells), and this may be attributed to the methods of tissue handling and dissociation for single-cell sample preparation, which is also present in another publication of single cell transcriptome investigating the female ovary [79].

In summary, the present study highlighted the transcriptional dynamics taking place in oocytes and granulosa cells during follicular formation by scRNA-seq. Specifically, the development stage-specific marker genes of germ cells and granulosa cells are identified. Furthermore, the TF regulatory networks of germ cells and novel pathways involved in PF formation have been given preliminarily confirmation. Thus, the current study provides novel insights into follicle formation at the perinatal stage and offers a reliable resource for study of the molecular regulatory mechanism of follicular development in the mouse.

## Materials and methods

### Animals and preparation of cell suspensions

C57BL/6J strain mice were purchased for scRNA-seq and ovarian sections staining from Vital River Laboratory Animal Technology Co. (Beijing, China). CD-1 mice for western blot experiments only were purchased from Jinan Pengyue Experimental Animal Breeding Co. Ltd (Jinan, China). Females were mated with males in the afternoon, and the presence of a vaginal plug in the morning of the following day was considered as E0.5. Embryonic and postnatal ovaries were dissected from E16.5 embryos and pups at PD0 and PD3, respectively. The selected 3 time points for sequencing represent the key stage of a series of cellular events involving PF formation, i.e., E16.5 time point represents the germ cell preparation (see prerequisites cellular events for the establishment of PFs [15]), and PD0 is the initial time point of nest breakdown, when it comes to PD3, more PFs have formed.

To obtain single-cell populations, isolated ovaries were cut into small pieces and incubated in 0.25% trypsin (Hyclone, Beijing, China) and collagenase (2 mg/ml, Sigma-Aldrich, C5138, Shanghai, China) for 6 to 8 min at 37°C. Tissues were disaggregated with a pipette to generate single cells, and the solution was filtered through 40- $\mu$ m cell strainers (BD Falcon, 352340, United States of America) and washed 2 times with PBS containing 0.04% BSA. Cell viability was acceptable when after staining in 0.4% Trypan Blue; it was above 80%; and cell concentration (1,000 cells/ $\mu$ L) was checked to meet the requirement of sequencing, as well as single-cell rate (no connected cells observed during cell counting).

### Ethics statement

Animals were housed according to the national guideline, the mice were humanely killed by cervical dislocation, and other experimental operations on mice have been approved by Animal Care and Ethical Committee of Qingdao Agricultural University.

### Single-cell libraries and sequencing

With qualified single cell sample, cells in 0.04% BSA in PBS were loaded onto 10 $\times$  Chromium chip, and single-cell gel beads in emulsion were generated by using Single Cell 3' Library and Gel Bead Kit V2 (10 $\times$  Genomics Inc., 120237, Pleasanton, California, USA). Followed the manufacturer's instructions, single-cell RNA-seq libraries were constructed, and pair-end 150 bp sequencing was performed to produce high-quantity data on an Illumina HiSeq X Ten (Illumina, San Diego, California, USA). Further, the "mkfastq" module of Cell Ranger produced the FASTQ files with raw base call (BCL) files as input generated by Illumina sequencer alignment.

### Clustering analysis with Seurat and cell trajectory construction by Monocle

The Cell Ranger "count" pipeline (version 3.1.0) were applied with the produced FASTQ data to map the mouse reference genome (version mm10) and process feature-barcode matrices with default parameters, specially, to avoid potential bias during subsequent bioinformatics analysis, the captured cell of each sample in number is set to 6,000 with the parameter of "—force-cells." The data matrixes were then loaded in R (version 3.6) using the Seurat package (version 3.1.5) [80]. The Seurat object was created based on 2 filtering parameters of "min.cells = 5" and "low.thresholds = 200," and the exorbitant number of unique genes detected in each cell (i.e., "nFeature\_RNA") was adjusted in each sample to eliminate the empty drop and dying cell and also potential doublets/multiplets for subsequent analyses (S1C Fig). Moreover, the second check was performed in single-cell type to guarantee high-quality cells, including

the distribution of detected unique genes (cells with an over-high number of unique genes are viewed as potential doublets/multiplets) and abnormal gene expression in each cell type (for example, if the germ cell specifically expressed marker genes of other cell types, these cells would undergo abnormal gene expression and be discarded). Then, the multiple samples were performed an integrated analysis, and briefly, the “FindIntegrationAnchors” function was used to identify anchors (a strategy for identification of cell pairwise correspondences between single cells across datasets, termed “anchors,” which can be used to transform datasets into a shared space and construct the harmonized atlases at the tissue or organismal scale) of 3 Seurat objects, and the integration analysis was performed according to anchors among samples [81]. Followed by the normalizing and scaling steps, with the UMAP (a visualized method for cell clustering in high-dimensional transcriptomic data) technique [82], a series of command was executed to visualize cell clusters, such as “RunUMAP” function with proper combination of “resolution” and “dims.use”; “FindNeighbors” and “FindClusters” function to conduct cell clustering. To identify canonical cell cluster marker genes, “FindAllMarkers” function was used to identify conserved marker gene in clusters with default parameters.

Further, the subpopulation of germ line or granulosa cell line was imported into Monocle (version 2.10.1) [83], a novel unsupervised algorithm, which reordered the cells to maximize the transcriptional similarity by their progress through differentiation, and at the same time, it would distinguish genes activated early and later in differentiation, which contribute to dissecting of cell differentiation fate, also termed “pseudotime analysis.” With the gene count matrix as input, the new dataset for Monocle object was created, and functions of “reduceDimension” and “orderCells” were carried out to generate the cell trajectory based on pseudotime. Particularly, the ordering genes were differentially expressed genes between clusters in each cell type calculated by “differentialGeneTest” function in Monocle. In addition, the “BEAM” function was used to calculate the differentially expressed genes at a branch point in the trajectory, and a gene with “ $qval < 1 \times 10^{-4}$ ” showed with heatmap. Moreover, the root state (that is, a prebranch in the heatmap) was set and adjusted following consideration of the biological meanings of different cell branches.

### The regulon activity of transcription factors with SCENIC

The SCENIC algorithm had been developed to assess the regulatory network analysis regard to TFs and discover regulons (that is, TFs and their target genes) in individual cells. Following the standard pipeline, the gene expression matrix with gene names in rows and cells in columns was input to SCENIC (version 0.9.1) [39]. The genes were filtered with default parameter, and finally 9,208 genes were available in RcisTarget database, the mouse specific database (mm10) that is used as the default in SCENIC. The co-expressed genes for each TF were constructed with GENIE3 software, followed by Spearman’s correlation between the TF and the potential targets, and then the “runSCENIC” procedure assisted to generate the GRNs (also termed regulons). Finally, regulon activity was analyzed by AUCell (Area Under the Curve) software, where a default threshold was applied to binarize the specific regulons (“0” present “off” of TFs, and “1” refer to “on”). The parameters of t-SNE was set with 50 principal components (PCs) and 50 perplexity to visualize (after running a consistency test across several perplexity values and number of PCs); meanwhile, the cell states were mapped with specific regulons, and averages of binary regulon activity (Fig 5B–5E) and TF expressions (S3C–S3E Fig) were projected onto t-SNEs.

### GO and KEGG enrichment analysis of gene set

For a large gene set with more than 100 genes, the enrichment analysis of GO and KEGG is preformatted with ClusterProfiler [84], a R package in Bioconductor, was applied to detect the

gene-related biological process, and signaling pathways with threshold value of “pvalueCut-off = 0.05,” and top terms were displayed (Fig 7A and 7B and S2F, S2G, S4C and S4D Figs). The online tools Metascape is used to uncover the function GO terms of small gene sets [85] (Figs 4F and 6G).

### Immunohistochemistry

Immunohistochemistry was performed as previously described [86]. Ovarian samples were processed for paraffin inclusion after fixing with paraformaldehyde (Solarbio, P1110, China) and following a standard dehydration procedure. Sample sections of 5  $\mu\text{m}$  in thickness were blocked at room temperature for 30 min after gradient rehydration and antigen retrieval and incubated with the germ cell-specific mouse vasa homologue (MVH) antibody (Abcam, ab13840, USA) overnight at 4°C. After 3 washes in TBS, the sections were incubated with secondary antibodies of FITC-labeled goat anti-rabbit IgG (H + L) (Beyotime, A0562, China) for 30 min at 37°C. After a further 3 washes in TBS, slides were counterstained with propidium iodide (PI, Sangon Biotech, E607306, China) for 3 min and sealed. Photomicrographs were taken with a fluorescence microscope (Olympus, BX51, Japan). Double immunostaining was performed with MVH antibody from mouse (Abcam, ab27591) or rabbit (Abcam, ab13840) with others: HSPB11 (Proteintech, 15732-1-AP, USA), G3BP2 (ABclonal, A6026, USA), OOEP (Biorbyt, orb379068, USA), FOXO3 (Novus Biological, NBP2-16521, USA), and TAZ (ab84927, Abcam) antibodies, followed by secondary antibodies of CY3-labeled goat anti-mouse IgG (H+L) (Beyotime, A0521) and goat anti-rabbit IgG H&L (Alexa Fluor 488, Abcam, ab150077); Hoechst 33342 or PI were used for nucleus counterstaining. In addition, the germ cells with more than 2 cells interconnected were considered as nest germ cells in ovarian sections, whereas single germ cell without connection to each other were considered as follicle germ cells; the percentage of germ cells within nests or follicle was calculated.

### Western blot

The western blot was administrated as previous reports [86, 87]. Briefly, with the collection of 6 ovaries, proteins in each sample were extracted using RIPA lysis solution (Beyotime, P0013C). After SDS-PAGE, the proteins were separated and transferred on polyvinylidene fluoride membrane (PVDF, Millipore, ISEQ00010, USA). They were then blocked with TBST buffer (TBS with 1% Tween-20) containing 5% bovine serum albumin, and the membrane was subsequently incubated with ACTIN (Sangon Biotech, D110007) or EIF4A1 antibody (A5294, ABclonal) (other primary antibodies see above) overnight at 4°C. The next day, after 3 washes, the membranes were incubated with secondary antibodies of HRP-conjugated goat anti-rabbit (Beyotime, A0208) or anti-mouse IgG (Beyotime, A0216) at room temperature for 1.5 h. Finally, a BeyoECL Plus Kit (Beyotime, A0018) was applied for chemiluminescence, according to the manufacturer's instructions. The expression of target protein was quantified by using the AlphaView SA software (ProteinSimple, California, USA).

### Statistical analysis

Results data as mean  $\pm$  SD were obtained from 3 independent experiments. The statistical analysis was performed with GraphPad Prism software (version 5.0), and the significant difference was determined with 2-tailed student's unpaired *t* test. Significant and highly significant difference were set at  $P < 0.05$  and  $P < 0.01$ , respectively.

## Supporting information

**S1 Fig. Information of single cell sequencing and cell types identification in ovary.** (A) Representative images of fetal ovaries at E16.5 and postnatal ovaries at PD0 and PD3. Germ cells were labeled with MVH (green) and nuclei counterstained with PI (red). Scale bar: 20  $\mu\text{m}$ . (B) Percentage of germ cells in nests or follicles at the indicated stages. Data were represented with mean  $\pm$  SD ( $n = 3$  for independent repeats). The relative level was calculated between PD0 or PD3 and E16.5. Unpaired  $t$  tests are performed. Statistical significance is shown as \*\*  $P < 0.01$ ; ns  $P > 0.05$ . The raw data used for quantification of B can be found in [S3 Data](#). (C) The sequenced detail information of 3 samples after CellRanger and Seurat workflow. (D) Feature plots of specific marker genes of erythrocytes, endothelial cells, immune cells, and epithelial cells. (E) Percentages of the 6 ovarian cell types at E16.5, PD0, and PD3. The sequencing data was deposited available in GSE134339, and this figure can be produced using scripts at [https://github.com/WangLab401/2020scRNA\\_murine\\_ovaries](https://github.com/WangLab401/2020scRNA_murine_ovaries). E16.5, embryonic day 16.5; PD0, postnatal day 0; PD3, postnatal day 3; MVH, mouse vasa homologue; PI, prodium iodide. (TIF)

**S2 Fig. Marker gene expression with developmental time course, cell proportion, and enrichment analysis in germ cells.** (A) Heatmap of top 10 marker genes of germ cell cluster with developmental timeline. Top 50 marker genes in each cluster are shown in [S2 Table](#). (B) Percentage of germ cells at pre-, early- and late-follicle formation stages. (C) The expressions of representative genes for 3 identified stages along with pseudotime trajectories. (D) Expression of representative genes (*Hdac2* and *Morf4l1*) in GO term of “histone deacetylation” along with pseudotime trajectories. (E) Expression of representative genes (*Morc2a* and *Smarca4*) in GO term of “regulation of gene expression and epigenetic” along with pseudotime trajectories. (F) Pathway enrichment of highly un-regulated genes of germ cells in follicle. (G) Pathway enrichment of most expressed genes of germ cells in nest. More pathways are shown in [S5 Table](#). GO, Gene Ontology. (TIF)

**S3 Fig. The t-SNE projection map of cell clustering, regulons density, and TF expressions.** (A) t-SNE based on 201 regulons with 50 PCs and 50 perplexity according to the developmental stages (left) and cell states (right). (B) t-SNE of regulons activity density. Regulon density means “the occurrence frequency of regulon in cells.” (C) TF expressions of *Nr3c1*, *Brca1*, *Kdm5a*, and *Kdm5b* in selected cells. (D) TF expressions of *Smarca1*, *Hes1*, *Stat3*, and *Sox15* in selected cells. (E) t-SNE projection of average binary regulon activity (left) of *Klf2* and *Jun* throughout the developmental stages and their expressions (right) in germ cells. TF, transcriptional factor; t-SNE, t-distributed stochastic neighbor embedding; PC, principal component. (TIF)

**S4 Fig. Marker gene expression with cell cluster and with developmental time points and pathway enrichment analysis of pre-granulosa cells.** (A) Heatmap of the top 5 marker genes in granulosa cell clusters. Top 50 marker genes in each cluster are shown in [S6 Table](#). (B) Vlnplots of the representative genes in granulosa cell clusters according to the developmental stages. (C) KEGG pathway enrichment of gene sets 1 and 2 that were related to BPGs fate. (D) KEGG pathway enrichment of gene sets 3 and 4 that have high expression in EPGs. More KEGG pathways are shown in [S9 Table](#). The sequencing data was deposited available in GSE134339, and this figure can be produced using scripts at [https://github.com/WangLab401/2020scRNA\\_murine\\_ovaries](https://github.com/WangLab401/2020scRNA_murine_ovaries). BPG, bipotential pre-granulosa; EPG, epithelial pre-granulosa; KEGG, Kyoto Encyclopedia of Genes and Genomes. (TIF)

**S5 Fig. Interaction of germ cell and pre-granulosa cell mediated by typical signal pathways.** (A) Vnlplots of the expression of NOTCH signal ligands, receptors, and targets in germ cells and granulosa cells. (B) Vnlplots of the expression of TGF-beta signal ligands, receptors, effectors, and targets in germ cells and granulosa cells. (C) Vnlplots of the expression of *Kit* and *Kitl* in germ cells and granulosa cells. (D) Vnlplots of the expression of connexin genes of gap junction in germ cells and granulosa cells. TGF-beta, transforming growth factor beta. (TIF)

**S6 Fig. Investigation of pathway signals between germ cell and granulosa cell.** (A) Venn diagram of the common and specific pathway between germ cells and granulosa cells. (B) Histogram of the most representative common pathway of germ cells and granulosa cells. (C and D) Dot plots (C) and Vnlplots (D) of tight junction related genes in germ cells and granulosa cells. (E and F) Dot plots (E) and Vnlplots (F) of adherens junction-related genes in germ cells and granulosa cells. UMAP, uniform manifold approximation projection. (TIF)

**S7 Fig. Clustering analyses of stromal cell, endothelial cell and immune cell in ovary.** (A) Cluster analysis of stromal cells with UMAP plots based on developmental timeline (upper) and cell clusters (below). (B) Feature plots of known marker genes of stromal cells. (C) Heatmap of top 5 marker genes of stromal cell clusters. (D) Cluster analysis of endothelial cells with UMAP plots based on developmental timeline (left) and cell clusters (right). (E) Cluster analysis of immune cells with UMAP plots based on developmental timeline (left) and cell clusters (right). The sequencing data was deposited available in GSE134339, this figure can be produced using scripts at [https://github.com/WangLab401/2020scrRNA\\_murine\\_ovaries](https://github.com/WangLab401/2020scrRNA_murine_ovaries). UMAP, uniform manifold approximation projection. (TIF)

**S1 Table. Marker genes of ovarian cell types during primordial follicle formation.** (XLSX)

**S2 Table. Top 50 marker genes of germ cell clusters.** (XLSX)

**S3 Table. Differentially expressed genes of germ cell trajectories.** (XLSX)

**S4 Table. Enriched GO term of top 100 genes in each gene set for germ cell trajectories.** GO, Gene Ontology. (XLSX)

**S5 Table. KEGG pathway of gene sets regard to germ cells in nest and follicle.** KEGG, Kyoto Encyclopedia of Genes and Genomes. (XLSX)

**S6 Table. Top 50 marker genes of granulosa cell clusters.** (XLSX)

**S7 Table. Differentially expressed genes of granulosa cells trajectories.** (XLSX)

**S8 Table. Enriched GO term of top 100 genes in each gene set for granulosa cell trajectories.** GO, Gene Ontology. (XLSX)



**S9 Table. KEGG pathway of gene sets regard to differentiation of granulosa cells.** KEGG, Kyoto Encyclopedia of Genes and Genomes.

(XLSX)

**S10 Table. KEGG pathways of key genes regard to germ cell and granulosa cell fates.**

KEGG, Kyoto Encyclopedia of Genes and Genomes.

(XLSX)

**S1 Data. Original western blot gel images of Fig 3D.**

(TIF)

**S2 Data. The raw data used for quantification of Fig 3D.**

(XLSX)

**S3 Data. The raw data used for quantification of S1B Fig.**

(XLSX)

## Author Contributions

**Conceptualization:** Wei Shen.

**Data curation:** Jun-Jie Wang, Wei Ge, Qiu-Yue Zhai, Xiao-Wen Sun, Wen-Xiang Liu, Lan Li.

**Formal analysis:** Jun-Jie Wang.

**Funding acquisition:** Wei Shen.

**Methodology:** Jun-Jie Wang, Wei Ge.

**Project administration:** Wei Shen.

**Resources:** Wei Shen.

**Supervision:** Wei Shen.

**Validation:** Jun-Jie Wang.

**Visualization:** Jun-Jie Wang.

**Writing – original draft:** Jun-Jie Wang, Chu-Zhao Lei, Paul W. Dyce, Massimo De Felici, Wei Shen.

**Writing – review & editing:** Jun-Jie Wang, Wei Ge, Qiu-Yue Zhai, Jing-Cai Liu, Xiao-Wen Sun, Wen-Xiang Liu, Lan Li, Chu-Zhao Lei, Paul W. Dyce, Massimo De Felici, Wei Shen.

## References

1. Tingen C, Kim A, Woodruff TK. The primordial pool of follicles and nest breakdown in mammalian ovaries. *Mol Hum Reprod*. 2009; 15 (12):795–803. <https://doi.org/10.1093/molehr/gap073> PMID: 19710243
2. Kerr JB, Myers M, Anderson RA. The dynamics of the primordial follicle reserve. *Reproduction*. 2013; 146 (6):R205–R15. <https://doi.org/10.1530/REP-13-0181> PMID: 23929903
3. Johnson J, Canning J, Kaneko T, Pru JK, Tilly JL. Germline stem cells and follicular renewal in the post-natal mammalian ovary. *Nature*. 2004; 428 (6979):145–50. <https://doi.org/10.1038/nature02316> PMID: 15014492
4. White YA, Woods DC, Takai Y, Ishihara O, Seki H, Tilly JL. Oocyte formation by mitotically active germ cells purified from ovaries of reproductive-age women. *Nat Med*. 2012; 18 (3):413–21. <https://doi.org/10.1038/nm.2669> PMID: 22366948

5. Skinner MK. Regulation of primordial follicle assembly and development. *Hum Reprod Update*. 2005; 11 (5):461–71. <https://doi.org/10.1093/humupd/dmi020> PMID: 16006439
6. Seki Y, Hayashi K, Itoh K, Mizugaki M, Saitou M, Matsui Y. Extensive and orderly reprogramming of genome-wide chromatin modifications associated with specification and early development of germ cells in mice. *Dev Biol*. 2005; 278 (2):440–58. <https://doi.org/10.1016/j.ydbio.2004.11.025> PMID: 15680362
7. Hayashi K, de Sousa Lopes SMC, Surani MA. Germ cell specification in mice. *Science*. 2007; 316 (5823):394–6. <https://doi.org/10.1126/science.1137545> PMID: 17446386
8. Lawson KA, Dunn NR, Roelen BA, Zeinstra LM, Davis AM, Wright CV, et al. Bmp4 is required for the generation of primordial germ cells in the mouse embryo. *Genes Dev*. 1999; 13 (4):424–36. <https://doi.org/10.1101/gad.13.4.424> PMID: 10049358
9. Nef S, Stévant I, Greenfield A. Characterizing the bipotential mammalian gonad. *Curr Top Dev Biol*. 2019; 134:167–94. <https://doi.org/10.1016/bs.ctdb.2019.01.002> PMID: 30999975
10. Stévant I, Nef S. Genetic Control of Gonadal Sex Determination and Development. *Trends Genet*. 2019; 35 (5):346–58. <https://doi.org/10.1016/j.tig.2019.02.004> PMID: 30902461
11. Chassot A-A, Gregoire EP, Lavery R, Taketo MM, de Rooij DG, Adams IR, et al. RSPO1/ $\beta$ -catenin signaling pathway regulates oogonia differentiation and entry into meiosis in the mouse fetal ovary. *PLoS One*. 2011; 6 (10):e25641. <https://doi.org/10.1371/journal.pone.0025641> PMID: 21991325
12. Pepling ME, Spradling AC. Female mouse germ cells form synchronously dividing cysts. *Development*. 1998; 125 (17):3323–8. PMID: 9693136
13. Pepling ME, Sundman EA, Patterson NL, Gephardt GW, Medico L, Wilson KI. Differences in oocyte development and estradiol sensitivity among mouse strains. *Reproduction*. 2010; 139 (2):349–57. <https://doi.org/10.1530/REP-09-0392> PMID: 19846484
14. Paredes A, Garcia-Rudaz C, Kerr B, Tapia V, Dissen GA, Costa ME, et al. Loss of synaptonemal complex protein-1, a synaptonemal complex protein, contributes to the initiation of follicular assembly in the developing rat ovary. *Endocrinology*. 2005; 146 (12):5267–77. <https://doi.org/10.1210/en.2005-0965> PMID: 16150897
15. Wang C, Zhou B, Xia G. Mechanisms controlling germline cyst breakdown and primordial follicle formation. *Cell Mol Life Sci*. 2017; 74 (14):2547–66. <https://doi.org/10.1007/s00018-017-2480-6> PMID: 28197668
16. Dutta S, Burks DM, Pepling ME. Arrest at the diplotene stage of meiotic prophase I is delayed by progesterone but is not required for primordial follicle formation in mice. *Reprod Biol Endocrinol*. 2016; 14 (1):82. <https://doi.org/10.1186/s12958-016-0218-1> PMID: 27919266
17. Klinger FG, Rossi V, De Felici M. Multifaceted programmed cell death in the mammalian fetal ovary. *Int J Dev Biol*. 2015; 59 (1-2-3):51–4. <https://doi.org/10.1387/ijdb.150063fk> PMID: 26374525
18. Pepling ME. From primordial germ cell to primordial follicle: mammalian female germ cell development. *Genesis*. 2006; 44 (12):622–32. <https://doi.org/10.1002/dvg.20258> PMID: 17146778
19. Chassot A-A, Bradford ST, Auguste A, Gregoire EP, Pailhoux E, De Rooij DG, et al. WNT4 and RSPO1 together are required for cell proliferation in the early mouse gonad. *Development*. 2012; 139 (23):4461–72. <https://doi.org/10.1242/dev.078972> PMID: 23095882
20. Zheng W, Zhang H, Gorre N, Risal S, Shen Y, Liu K. Two classes of ovarian primordial follicles exhibit distinct developmental dynamics and physiological functions. *Hum Mol Genet*. 2013; 23 (4):920–8. <https://doi.org/10.1093/hmg/ddt486> PMID: 24087793
21. Zheng W, Zhang H, Liu K. The two classes of primordial follicles in the mouse ovary: their development, physiological functions and implications for future research. *Mol Hum Reprod*. 2014; 20 (4):286–92. <https://doi.org/10.1093/molehr/gau007> PMID: 24448914
22. Rastetter RH, Bernard P, Palmer JS, Chassot A-A, Chen H, Western PS, et al. Marker genes identify three somatic cell types in the fetal mouse ovary. *Dev Biol*. 2014; 394 (2):242–52. <https://doi.org/10.1016/j.ydbio.2014.08.013> PMID: 25158167
23. Mork L, Maatouk DM, McMahon JA, Guo JJ, Zhang P, McMahon AP, et al. Temporal differences in granulosa cell specification in the ovary reflect distinct follicle fates in mice. *Biol Reprod*. 2012; 86 (2):37. <https://doi.org/10.1095/biolreprod.111.095208> PMID: 21976597
24. Pepling ME. Follicular assembly: mechanisms of action. *Reproduction*. 2012; 143 (2):139–49. <https://doi.org/10.1530/REP-11-0299> PMID: 22065859
25. Grive KJ, Freiman RN. The developmental origins of the mammalian ovarian reserve. *Development*. 2015; 142 (15):2554–63. <https://doi.org/10.1242/dev.125211> PMID: 26243868
26. Kezele PR, Ague JM, Nilsson E, Skinner MK. Alterations in the ovarian transcriptome during primordial follicle assembly and development. *Biol Reprod*. 2005; 72 (1):241–55. <https://doi.org/10.1095/biolreprod.104.032060> PMID: 15371273

27. Nilsson E, Zhang B, Skinner MK. Gene bionetworks that regulate ovarian primordial follicle assembly. *BMC Genomics*. 2013; 14 (1):496. <https://doi.org/10.1186/1471-2164-14-496> PMID: 23875758
28. Bonnet A, Cabau C, Bouchez O, Sarry J, Marsaud N, Foissac S, et al. An overview of gene expression dynamics during early ovarian folliculogenesis: specificity of follicular compartments and bi-directional dialog. *BMC Genomics*. 2013; 14 (1):904. <https://doi.org/10.1186/1471-2164-14-904> PMID: 24350644
29. Tan X-M, Zhang X-L, Ji M-M, Yu K-L, Liu M-M, Tao Y-C, et al. Differentially expressed genes associated with primordial follicle formation and transformation to primary Follicles. *Reproductive and Developmental Medicine*. 2018; 2 (3):142.
30. Stévant I, Kühne F, Greenfield A, Chaboissier M-C, Dermitzakis ET, Nef S. Dissecting Cell Lineage Specification and Sex Fate Determination in Gonadal Somatic Cells Using Single-Cell Transcriptomics. *Cell Rep*. 2019; 26 (12):3272–83. e3. <https://doi.org/10.1016/j.celrep.2019.02.069> PMID: 30893600
31. Imai Y, Baudat F, Taillepierre M, Stanzione M, Toth A, de Massy B. The PRDM9 KRAB domain is required for meiosis and involved in protein interactions. *Chromosoma*. 2017; 126 (6):681–95. <https://doi.org/10.1007/s00412-017-0631-z> PMID: 28527011
32. Gu W, Gu W, Tekur S, Reinbold R, Eppig JJ, Choi Y-C. Mammalian male and female germ cells express a germ cell-specific Y-Box protein, MSY2. *Biol Reprod*. 1998; 59 (5):1266–74. <https://doi.org/10.1095/biolreprod59.5.1266> PMID: 9780336
33. Bolcun-Filas E, Schimenti JC. Genetics of meiosis and recombination in mice. *Int Rev Cell Mol Biol*. 2012; 179–227. <https://doi.org/10.1016/B978-0-12-394309-5.00005-5> PMID: 22878107
34. Takasaki N, Rankin T, Dean J. Normal gonadal development in mice lacking GPBOX, a homeobox protein expressed in germ cells at the onset of sexual dimorphism. *Mol Cell Biol*. 2001; 21 (23):8197–202. <https://doi.org/10.1128/MCB.21.23.8197-8202.2001> PMID: 11689708
35. Zhang Y, Yan Z, Qin Q, Nisenblat V, Chang H-M, Yu Y, et al. Transcriptome Landscape of Human Folliculogenesis Reveals Oocyte and Granulosa Cell Interactions. *Mol Cell*. 2018; 72 (6):1021–34. e4. <https://doi.org/10.1016/j.molcel.2018.10.029> PMID: 30472193
36. Dean J. Oocyte-specific genes regulate follicle formation, fertility and early mouse development. *J Reprod Immunol*. 2002; 53 (1–2):171–80. [https://doi.org/10.1016/S0165-0378\(01\)00100-0](https://doi.org/10.1016/S0165-0378(01)00100-0) PMID: 11730914
37. Rajkovic A, Pangas SA, Ballow D, Suzumori N, Matzuk MM. NOBOX deficiency disrupts early folliculogenesis and oocyte-specific gene expression. *Science*. 2004; 305 (5687):1157–9. <https://doi.org/10.1126/science.1099755> PMID: 15326356
38. Bayne RA, Kinnell HL, Coutts SM, He J, Childs AJ, Anderson RA. GDF9 is transiently expressed in oocytes before follicle formation in the human fetal ovary and is regulated by a novel NOBOX transcript. *PLoS One*. 2015; 10 (3):e0119819. <https://doi.org/10.1371/journal.pone.0119819> PMID: 25790371
39. Aibar S, González-Blas CB, Moerman T, Imrichova H, Hulselmans G, Rambow F, et al. SCENIC: single-cell regulatory network inference and clustering. *Nat Methods*. 2017; 14 (11):1083–6. <https://doi.org/10.1038/nmeth.4463> PMID: 28991892
40. Capel B. Ovarian epithelium regeneration by Lgr5+ cells. *Nat Cell Biol*. 2014; 16 (8):743–4. <https://doi.org/10.1038/ncb3020> PMID: 25082199
41. Hummitzsch K, Anderson RA, Wilhelm D, Wu J, Telfer EE, Russell DL, et al. Stem cells, progenitor cells, and lineage decisions in the ovary. *Endocr Rev*. 2015; 36 (1):65–91. <https://doi.org/10.1210/er.2014-1079> PMID: 25541635
42. Hummitzsch K, Hatzirodos N, Irving-Rodgers HF, Hartanti MD, Perry VE, Anderson RA, et al. Morphometric analyses and gene expression related to germ cells, gonadal ridge epithelial-like cells and granulosa cells during development of the bovine fetal ovary. *PLoS One*. 2019; 14 (3):e0214130. <https://doi.org/10.1371/journal.pone.0214130> PMID: 30901367
43. Chassot A-A, Le Rolle M, Jolivet G, Stévant I, Guignonis J-M, Da Silva F, et al. Retinoic acid synthesis by ALDH1A proteins is dispensable for meiosis initiation in the mouse fetal ovary. *Sci Adv*. 2020; 6(21): eaaz1261. <https://doi.org/10.1126/sciadv.aaz1261> PMID: 32494737
44. Fan X, Bialecka M, Moustakas I, Lam E, Torrens-Juaneda V, Borggreven N, et al. Single-cell reconstruction of follicular remodeling in the human adult ovary. *Nat Commun*. 2019; 10 (1):3164. <https://doi.org/10.1038/s41467-019-11036-9> PMID: 31320652
45. Rosario R, Araki H, Shelling AN. The transcriptional targets of mutant FOXL2 in granulosa cell tumours. *PLoS One*. 2012; 7 (9):e46270. <https://doi.org/10.1371/journal.pone.0046270> PMID: 23029457
46. Nicol B, Grimm SA, Gruzdev A, Scott GJ, Ray MK, Yao HH. Genome-wide identification of FOXL2 binding and characterization of FOXL2 feminizing action in the fetal gonads. *Hum Mol Genet*. 2018; 27 (24):4273–87. <https://doi.org/10.1093/hmg/ddy312> PMID: 30212841

47. Niu W, Spradling AC. Two distinct pathways of pregranulosa cell differentiation support follicle formation in the mouse ovary. *Proc Natl Acad Sci.* 2020; 117 (33):20015–26. <https://doi.org/10.1073/pnas.2005570117> PMID: 32759216
48. Feng C-W, Bowles J, Koopman P. Control of mammalian germ cell entry into meiosis. *Mol Cell Endocrinol.* 2014; 382 (1):488–97. <https://doi.org/10.1016/j.mce.2013.09.026> PMID: 24076097
49. Tedesco M, La Sala G, Barbagallo F, De Felici M, Farini D. STRA8 shuttles between nucleus and cytoplasm and displays transcriptional activity. *J Biol Chem.* 2009; 284 (51):35781–93. <https://doi.org/10.1074/jbc.M109.056481> PMID: 19805549
50. Wang J, Roy SK. Growth differentiation factor-9 and stem cell factor promote primordial follicle formation in the hamster: modulation by follicle-stimulating hormone. *Biol Reprod.* 2004; 70 (3):577–85. <https://doi.org/10.1095/biolreprod.103.023234> PMID: 14585807
51. Arango NA, Li L, Dabir D, Nicolau F, Pieretti-Vanmarcke R, Koehler C, et al. Meiosis I arrest abnormalities lead to severe oligozoospermia in meiosis 1 arresting protein (M1ap)-deficient mice. *Biol Reprod.* 2013; 88 (3):76. <https://doi.org/10.1095/biolreprod.111.098673> PMID: 23269666
52. Török Z, Pilbat A-M, Gombos I, Hocsák E, Sümegi B, Horváth I, et al. Evidence on cholesterol-controlled lipid raft interaction of the small heat shock protein Hspb11. *Cellular Trafficking of Cell Stress Proteins in Health and Disease*: Springer. 2012:75–85.
53. Kumar P, Hellen CU, Pestova TV. Toward the mechanism of eIF4F-mediated ribosomal attachment to mammalian capped mRNAs. *Genes Dev.* 2016; 30 (13):1573–88. <https://doi.org/10.1101/gad.282418.116> PMID: 27401559
54. Kobayashi T, Winslow S, Sunesson L, Hellman U, Larsson C. PKC $\alpha$  binds G3BP2 and regulates stress granule formation following cellular stress. *PLoS One.* 2012; 7 (4):e35820. <https://doi.org/10.1371/journal.pone.0035820> PMID: 22536444
55. Wang L-H, Baker NE. E proteins and ID proteins: helix-loop-helix partners in development and disease. *Dev Cell.* 2015; 35 (3):269–80. <https://doi.org/10.1016/j.devcel.2015.10.019> PMID: 26555048
56. He D-J, Wang L, Zhang Z-B, Guo K, Li J-Z, He X-C, et al. Maternal gene Ooep may participate in homologous recombination-mediated DNA double-strand break repair in mouse oocytes. *Zool Res.* 2018; 39 (6):387–95. <https://doi.org/10.24272/j.issn.2095-8137.2018.067> PMID: 29955025
57. Li L, Baibakov B, Dean J. A subcortical maternal complex essential for preimplantation mouse embryogenesis. *Dev Cell.* 2008; 15 (3):416–25. <https://doi.org/10.1016/j.devcel.2008.07.010> PMID: 18804437
58. Esposito G, Vitale A, Leijten F, Strik A, Koonen-Reemst A, Yurttas P, et al. Peptidylarginine deiminase (PAD) 6 is essential for oocyte cytoskeletal sheet formation and female fertility. *Mol Cell Endocrinol.* 2007; 273 (1–2):25–31. <https://doi.org/10.1016/j.mce.2007.05.005> PMID: 17587491
59. Yurttas P, Vitale AM, Fitzhenry RJ, Cohen-Gould L, Wu W, Gossen JA, et al. Role for PADI6 and the cytoplasmic lattices in ribosomal storage in oocytes and translational control in the early mouse embryo. *Development.* 2008; 135 (15):2627–36. <https://doi.org/10.1242/dev.016329> PMID: 18599511
60. Kan R, Yurttas P, Kim B, Jin M, Wo L, Lee B, et al. Regulation of mouse oocyte microtubule and organelle dynamics by PADI6 and the cytoplasmic lattices. *Dev Biol.* 2011; 350 (2):311–22. <https://doi.org/10.1016/j.ydbio.2010.11.033> PMID: 21147087
61. Ge W, Wang J-J, Zhang R-Q, Tan S-J, Zhang F-L, Liu W-X, et al. Dissecting the initiation of female meiosis in the mouse at single-cell resolution. *Cell Mol Life Sci.* 2020:1–19. <https://doi.org/10.1007/s00018-020-03533-8> PMID: 32367190
62. Shimamoto S, Nishimura Y, Nagamatsu G, Hamada N, Kita H, Hikabe O, et al. Hypoxia induces the dormant state in oocytes through expression of Foxo3. *Proc Natl Acad Sci.* 2019; 116 (25):12321–6. <https://doi.org/10.1073/pnas.1817223116> PMID: 31147464
63. Nagamatsu G, Shimamoto S, Hamazaki N, Nishimura Y, Hayashi K. Mechanical stress accompanied with nuclear rotation is involved in the dormant state of mouse oocytes. *Sci Adv.* 2019; 5(6):eaav9960. <https://doi.org/10.1126/sciadv.aav9960> PMID: 31249869
64. Castrillon DH, Miao L, Kollipara R, Horner JW, DePinho RA. Suppression of ovarian follicle activation in mice by the transcription factor Foxo3a. *Science.* 2003; 301 (5630):215–8. <https://doi.org/10.1126/science.1086336> PMID: 12855809
65. Conti M, Franciosi F. Acquisition of oocyte competence to develop as an embryo: integrated nuclear and cytoplasmic events. *Hum Reprod Update.* 2018; 24 (3):245–66. <https://doi.org/10.1093/humupd/dmx040> PMID: 29432538
66. Hummitzsch K, Irving-Rodgers HF, Schwartz J, Rodgers RJ. *Development of the mammalian ovary and follicles*. The ovary: Elsevier; 2019. p. 71–82.
67. Kawashima I, Kawamura K. Regulation of follicle growth through hormonal factors and mechanical cues mediated by hippo signaling pathway. *Syst Biol Reprod Med.* 2018; 64 (1):3–11. <https://doi.org/10.1080/19396368.2017.1411990> PMID: 29224376

68. Kawamura K, Cheng Y, Suzuki N, Deguchi M, Sato Y, Takae S, et al. Hippo signaling disruption and Akt stimulation of ovarian follicles for infertility treatment. *Proc Natl Acad Sci*. 2013; 110 (43):17474–9. <https://doi.org/10.1073/pnas.1312830110> PMID: 24082083
69. Plewes MR, Hou X, Zhang P, Liang A, Hua G, Wood JR, et al. Yes-associated protein 1 is required for proliferation and function of bovine granulosa cells in vitro. *Biol Reprod*. 2019; 101 (5):1001–17. <https://doi.org/10.1093/biolre/iox139> PMID: 31350850
70. Tsoi M, Morin M, Rico C, Johnson RL, Paquet M, Gévry N, et al. Lats1 and Lats2 are required for ovarian granulosa cell fate maintenance. *FASEB J*. 2019; 33 (10):10819–32. <https://doi.org/10.1096/fj.201900609R> PMID: 31268774
71. Lv X, He C, Huang C, Wang H, Hua G, Wang Z, et al. Timely expression and activation of YAP1 in granulosa cells is essential for ovarian follicle development. *FASEB J*. 2019; 33 (9):10049–64. <https://doi.org/10.1096/fj.201900179RRR> PMID: 31199671
72. Plewes M, Hou X, Zhang P, Wood J, Cupp A, Davis J. Yes-associated protein (YAP) is required in maintaining normal ovarian follicle development and function. *bioRxiv*. 2018:406256.
73. Furukawa-Hibi Y, Kobayashi Y, Chen C, Motoyama N. FOXO transcription factors in cell-cycle regulation and the response to oxidative stress. *Antioxid Redox Signal*. 2005; 7 (5–6):752–60. <https://doi.org/10.1089/ars.2005.7.752> PMID: 15890021
74. Sui X-X, Luo L-L, Xu J-J, Fu Y-C. Evidence that FOXO3a is involved in oocyte apoptosis in the neonatal rat ovary. *Biochem Cell Biol*. 2010; 88 (4):621–8. <https://doi.org/10.1139/O10-001> PMID: 20651833
75. Pelosi E, Omari S, Michel M, Ding J, Amano T, Forabosco A, et al. Constitutively active Foxo3 in oocytes preserves ovarian reserve in mice. *Nat Commun*. 2013; 4:1843. <https://doi.org/10.1038/ncomms2861> PMID: 23673628
76. Luecken MD, Theis FJ. Current best practices in single-cell RNA-seq analysis: a tutorial. *Mol Syst Biol*. 2019; 15 (6):e8746. <https://doi.org/10.15252/msb.20188746> PMID: 31217225
77. Hwang B, Lee JH, Bang D. Single-cell RNA sequencing technologies and bioinformatics pipelines. *Exp Mol Med*. 2018; 50 (8):1–14. <https://doi.org/10.1038/s12276-018-0071-8> PMID: 30089861
78. Jevitt A, Chatterjee D, Xie G, Wang X-F, Otwell T, Huang Y-C, et al. A single-cell atlas of adult *Drosophila* ovary identifies transcriptional programs and somatic cell lineage regulating oogenesis. *PLoS Biol*. 2020; 18 (4):e3000538. <https://doi.org/10.1371/journal.pbio.3000538> PMID: 32339165
79. Wagner M, Yoshihara M, Douagi I, Damdimopoulos A, Panula S, Petropoulos S, et al. Single-cell analysis of human ovarian cortex identifies distinct cell populations but no oogonial stem cells. *Nat Commun*. 2020; 11(1):1–15. <https://doi.org/10.1038/s41467-019-13993-7> PMID: 31911652
80. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol*. 2018; 36 (5):411–20. <https://doi.org/10.1038/nbt.4096> PMID: 29608179
81. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM III, et al. Comprehensive integration of single-cell data. *Cell*. 2019; 177 (7):1888–902. e21. <https://doi.org/10.1016/j.cell.2019.05.031> PMID: 31178118
82. Becht E, McInnes L, Healy J, Dutertre C-A, Kwok IW, Ng LG, et al. Dimensionality reduction for visualizing single-cell data using UMAP. *Nat Biotechnol*. 2019; 37 (1):38–44.
83. Qiu X, Mao Q, Tang Y, Wang L, Chawla R, Pliner HA, et al. Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods*. 2017; 14 (10):979–82. <https://doi.org/10.1038/nmeth.4402> PMID: 28825705
84. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics: a journal of integrative biology*. 2012; 16 (5):284–7. <https://doi.org/10.1089/omi.2011.0118> PMID: 22455463
85. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun*. 2019; 10 (1):1–10. <https://doi.org/10.1038/s41467-018-07882-8> PMID: 30602773
86. Wang J-J, Yu X-W, Wu R-Y, Sun X-F, Cheng S-F, Ge W, et al. Starvation during pregnancy impairs fetal oogenesis and folliculogenesis in offspring in the mouse. 2018; 9 (5):452. <https://doi.org/10.1038/s41419-018-0492-2> PMID: 29670080
87. Sun Z-Y, Zhang P, Wang J-J, Liu J-C, Li L, Shen W, et al. Melatonin alleviates meiotic defects in fetal mouse oocytes induced by Di (2-ethylhexyl) phthalate in vitro. *Aging (Albany NY)*. 2018; 10 (12):4175–87. <https://doi.org/10.18632/aging.101715> PMID: 30591620