The identification of a gene expression signature of primordial follicle activation in mouse pregranulosa cells

Emily R Frost[1,2,3*] Güneş Taylor[3*#], Stefan Boeing[4,5], Christophe Galichet[3], Mark A Baker[1,2], Jessie M Sutherland[1,2+#], Robin Lovell-Badge[3+#]

1 Priority Research Centre for Reproductive Science, School of Biomedical Science & Pharmacy, University of Newcastle, Ring Road, Callaghan, New South Wales, Australia, 2308

2 Hunter Medical Research Institute, New Lambton Heights, New South Wales, Australia, 2305

3 Stem Cell Biology and Developmental Genetics Lab, The Francis Crick Institute, London, United Kingdom, NW1 1AT

4 Bioinformatics and Biostatistics Facility, The Francis Crick Institute, London, United Kingdom

5 Scientific Computing – Digital Development Team, The Francis Crick Institute, London, United Kingdom

* These authors contributed equally and should be regarded as co-first authors

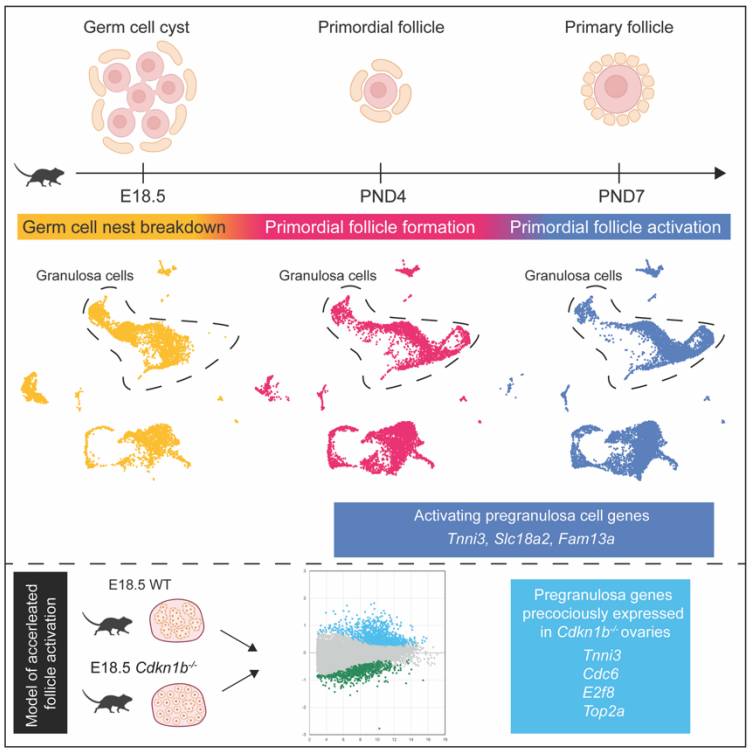
+ These authors contributed equally and should be regarded as co-last authors

Corresponding authors: Dr Güneş Taylor, The Francis Crick Institute, gunes.taylor@crick.ac.uk

Dr Jessie Sutherland, University of Newcastle, jessie.sutherland@newcastle.edu.au

Prof Robin Lovell-Badge, The Francis Crick Institute, robin.lovell-badge@crick.ac.uk

Graphical abstract



Abstract

Female fertility is dictated by the number of oocyte-containing primordial follicles within the ovary. These follicles are established early in development and their selective activation represents the definitive first step towards oocyte maturation and ovulation. The intrinsic molecular mechanisms that regulate ovarian follicle activation are largely uncharacterised. In this study, we report a single-cell RNA sequencing (scRNAseq) dataset to examine mouse embryonic and neonatal ovaries across the developmental window of primordial follicle formation and activation. We identified five distinct clusters of granulosa cells and bioinformatic analysis revealed a population of pregranulosa cells that appeared to be undergoing follicle activation. This cluster was uniquely differentiated by the expression of *Tnni3*, *Slc18a2*, *Fam13a* and *Klf2*, all of which are novel to follicle activation. Finally, we examined the transcriptome of a mouse model where all primordial follicles enter folliculogenesis simultaneously (*Cdkn1b*$^{-/-}$) and showed that the signature genes of activating pregranulosa cells was observed precociously in *Cdkn1b*$^{-/-}$ ovaries, further demonstrating that p27^{kip1} acts as an important repressor of follicle activation. Combined, this data provides insight into how primordial follicle activation is regulated in mammals and could be utilised to identify pathways to target to improve fertility preservation and infertility conditions in the future.

37 **Introduction**

38 Mammalian ovary development is a tightly regulated process that is crucial for establishing the
39 ovarian reserve of oocytes. The ovarian reserve must be able to sustain fertility over several
40 months in mice and several decades in humans (Monniaux et al., 2014). The ovarian reserve is
41 defined by the number of oocytes present in the ovary in structures called primordial follicles.
42 Primordial follicles consist of an oocyte with a layer of surrounding pregranulosa cells and are
43 formed during gestation (Kezele, Ague, Nilsson, & Skinner, 2005; Pepling & Spradling, 2001;
44 Tingen, Kim, & Woodruff, 2009). To create the ovarian reserve, both mitosis and meiosis occur
45 in co-ordinated events in the pregranulosa cells and the oocytes respectively. Following the
46 determination of ovarian fate in the bipotential gonad, pregranulosa cells and oogonia develop
47 in close contact, throughout germ cell nest formation and breakdown, resulting in primordial
48 follicle formation and activation (Lei & Spradling, 2013; Pepling & Spradling, 1998; Suzuki,
49 Kanai-Azuma, & Kanai, 2015; Wear, McPike, & Watanabe, 2016).

50

51 In the mouse ovary, it is well-established that two specific populations of primordial follicles
52 exist, classified by their spatial distribution and origin, namely the medullary and the cortical
53 primordial follicles (Mork et al., 2012; Niu & Spradling, 2020; W. Zheng, Zhang, Gorre, et al.,
54 2014; W. Zheng, Zhang, & Liu, 2014). Two precise populations of pregranulosa cells
55 contribute to the medullary and cortical classes of primordial follicles. Both lineages of
56 pregranulosa cells derive from the ovarian surface epithelium but at different times (Niu &
57 Spradling, 2020; W. Zheng, Zhang, Gorre, et al., 2014). The first pregranulosa cell population
58 that develops express FOXL2 and $p27^{kip1}$ and contribute to medullary primordial follicles,
59 termed bipotential pregranulosa cells (BPGs) (Niu & Spradling, 2020; J.-J. Wang et al., 2020).
60 The second population, termed epithelial pregranulosa cells (EPGs) ingress from the
61 epithelium later. They express LGR5 and contribute to cortical primordial follicles, expressing

62 FOXL2 and p27^{kip1} postnatally (Niu & Spradling, 2020; J.-J. Wang et al., 2020). The formation
63 of these two classes of primordial follicles are asynchronous, with medullary primordial
64 follicles formed first (J.-J. Wang et al., 2020). Aside from the distinct differences in
65 developmental timing, it is not yet clear whether other characteristics vary between the two
66 pregranulosa classes. Recent advances have identified molecular pathways in both the oocytes
67 and pregranulosa cells that regulate the formation of primordial follicles, including the
68 NOTCH, KITL, TGFß and PI3K pathways (Fukuda, Muraoka, Kato, & Saga, 2021; Niu &
69 Spradling, 2020; Rosario et al., 2019; J.-J. Wang et al., 2020).
70
71 In contrast to follicle formation, primordial follicle activation is more challenging to
72 interrogate, due to the continuous selection of follicles for activation throughout prepubertal
73 and adult life. However, an initial wave of primordial follicle activation occurs postnatally in
74 mouse, with the activation of medullary follicles entering into folliculogenesis (Cordeiro et al.,
75 2015; Mork et al., 2012). This wave presents an opportunity to understand the factors and
76 pathways that govern the activation of primordial follicles into folliculogenesis. Changes in
77 cell cycle status, granulosa cell morphology (from squamous to cuboidal) and oocyte size are
78 key indicators of follicle activation and the upregulation of essential factors including PI3K,
79 KITL and BMP4 are indicative of entry into the developing pool of follicles (Adhikari & Liu,
80 2009; Hardy et al., 2018; Hutt, McLaughlin, & Holland, 2006; John, Gallardo, Shirley, &
81 Castrillon, 2008; McLaughlin & McIver, 2009). Protocols have recently been developed for
82 the dissociation of the neonatal mouse ovary, generating single-cell suspensions for
83 sequencing, that can be used during this initial wave of activation (Y.-Y. Chen et al., 2022;
84 Frost et al., 2021a). To date, most studies have utilised genetically modified mouse models and
85 microarray experiments to explore the expression of gene-driven signals in stimulating or

86 inhibiting primordial follicle activation (reviewed in (Y. Chen et al., 2020; McLaughlin &
87 McIver, 2009; Sobinoff, Sutherland, & McLaughlin, 2013)).
88
89 In this study, we use single-cell RNA sequencing (scRNA-seq) to examine the transcriptional
90 profile of mouse embryonic and neonatal ovaries at three timepoints: embryonic day (E) 18.5,
91 postnatal day (PND) 4 and PND7. In total we sequenced 24,810 cells from the three
92 developmental stages and distinguished five clusters of granulosa cells. We identified a
93 transcriptomic signature of pregranulosa cells during primordial follicle activation in granulosa
94 cell cluster 1 (Gc_1), by comparing with the activated granulosa cells present in granulosa cell
95 cluster 4 (Gc_4). Moreover, analysis of a genetically modified mouse line with an accelerated
96 follicle activation phenotype (*Cdkn1b*$^{-/-}$) revealed that this signature was observed precociously
97 in the mutant mice as early as E18.5. Combined, this data reveals pregranulosa cell-specific
98 pathways that control primordial follicle activation in mice. Further interrogation of these
99 pathways in the infertile mouse model shows disruptions in the follicle activation process,
100 uncovering key insights into how primordial follicle activation establishes fertility.
101
102 **Results**
103 ***Characterisation of mouse ovarian transcriptome before and during primordial follicle***
104 ***activation***
105 To investigate the transcriptional signature of granulosa cells both prior to and during
106 primordial follicle activation, we performed single-cell RNA sequencing of the mouse ovary.
107 We focused on three critical timepoints representative of ovary development before, during
108 and after follicle activation, namely: embryonic day 18.5 (E18.5), when germ cell nest
109 breakdown is initiated to form primordial follicles; post-natal day 4 (PND4), when the initial
110 wave of primordial follicle activation is occurring; and PND7, when populations of activated

111 follicles (termed primary follicles) are present within the ovary (Figure 1A). A previously
112 described and validated dissociation protocol using the enzyme papain was utilised to achieve
113 a single cell suspension of the ovarian tissue at each stage (Frost et al., 2021b). In total, our
114 dataset contained 24,810 cells, with 8,210 cells from one E18.5 ovary, 7,587 cells from one
115 PND4 ovary and 9,013 cells from one PND7 ovary (Table S1). Initially, the data from all three
116 timepoints was integrated and uniform manifold approximation and projection (UMAP)
117 analysis was used to assess cell contribution to the clustering based on the developmental stage
118 (Figure 1B). Across the integrated timepoints, 18 distinct clusters were found, and the six major
119 cell types present in the ovary across these developmental stages were annotated based on their
120 expression of established cell type specific genetic markers (granulosa cells, mesenchymal
121 cells, oocytes, epithelial cells, immune cells and blood cells) (Figure 1C and Figure S1). Across
122 all three timepoints, most cells mapped to either the mesenchymal (43%) or granulosa cell
123 (39%) clusters. The number of clusters per cell type were as follows: five granulosa cell
124 clusters, six mesenchymal clusters, one oocyte cluster, two epithelial clusters, one immune
125 cluster, two blood cell clusters and one cluster unable to be identified (labelled 'other').
126 Granulosa cells were identified by the expression of *Fst*, *Amhr2*, *Kitl*, *Runx1*, *Foxl2, Hmgcs2*
127 and *Lgr5* and clusters were labelled as Gc_1 to Gc_5 (Bagheri-Fam et al., 2020; Jones &
128 Pepling, 2013; Meinsohn et al., 2021; Nicol et al., 2019; Ottolenghi et al., 2005; Rastetter et
129 al., 2014; Z. Wang et al., 2015). Mesenchymal cells expressed *Nr2f2* and *Col1a1* and clusters
130 were named Mc_1 to Mc_6 (Rastetter et al., 2014). Epithelial cells were characterised by the
131 expression of *Upk3b* and *Krt19*, with epithelial clusters labelled as Ep_1 and Ep_2 (Gaytan,
132 Morales, Roa, & Tena-Sempere, 2018; Niu & Spradling, 2020). Oocytes were distinguished
133 by *Ddx4*, *Dazl* and *Sycp3* expression and contained within cluster Oo_1 (Cooke, Lee, Kerr, &
134 Ruggiu, 1996; Cordeiro et al., 2015; Reynolds, Collier, Bingham, Gray, & Cooke, 2007;
135 Rosario et al., 2019; Yang et al., 2020). Blood cells were characterised by expression of *Cldn5*

136 and *Cx3cr1* and clusters were labelled as Bl_1 and Bl_2 (Landsman et al., 2009; Ma et al.,
137 2017), while immune cells expressed *Cd53*, contained in cluster Im_1 (Greenberg et al., 2020;
138 J.-J. Wang et al., 2020). Taken together, these results validated the successful isolation of
139 mouse ovarian cells during late embryonic and early postnatal development and illustrated
140 distinct cell-specific transcriptomic profiles essential for further interrogation.

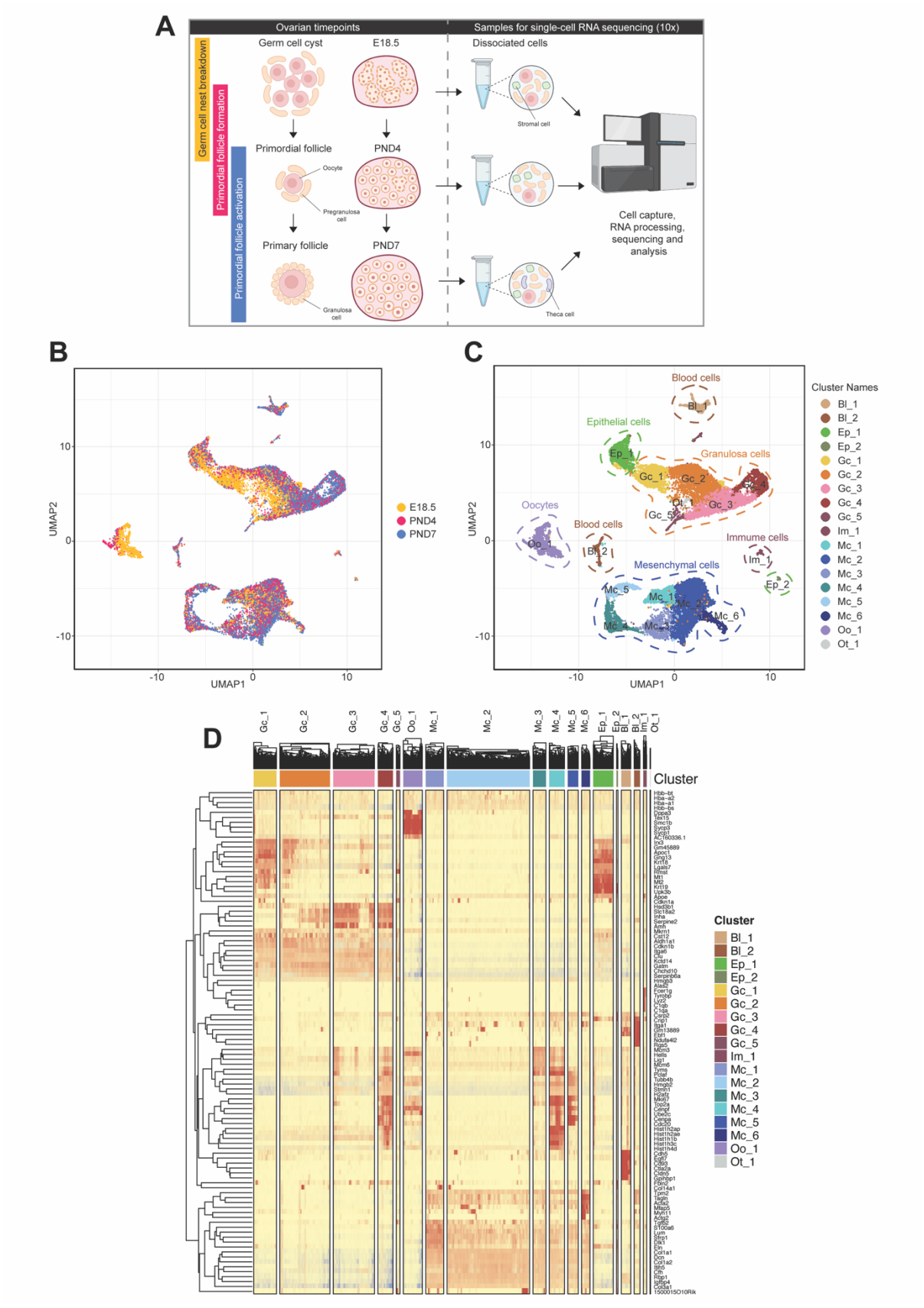


Figure 1. Transcriptional profiling of whole ovaries before, during and after primordial follicle activation in the mouse

144 (A) A diagram of the mouse ovary during late embryonic and early postnatal development, highlighting
145 the key events of germ cell nest breakdown, primordial follicle formation and primordial follicle
146 activation, and the timepoints chosen for this study. The diagram also shows the workflow of sample
147 preparation for single-cell RNA sequencing. UMAP Dimensionality Reduction Plots showing
148 individual cells coloured by timepoint (B) and by cell type (C). Clusters were named by cell type as
149 follows: granulosa cells (Gc_1, Gc_2, Gc_3, Gc_4, and Gc_5), mesenchymal cells (Mc_1, Mc_2,
150 Mc_3, Mc_4, Mc_5 and Mc_6), epithelial cells (Ep_1 and Ep_2), oocytes (Oo_1), blood cells (Bl_1
151 and Bl_2), immune cells (Im_1) and other cells (Ot_1). (D) Heat map showing the most distinct marker
152 genes in each cluster, used to identify the clusters in the UMAP. For this analysis, 5,000 cells were
153 randomly selected from the 24,810 cells in this experiment.
154

155 ***Activated granulosa cells are transcriptionally distinct postnatally***

156 Next, we analysed the transcriptomic signature of the ovarian cells at each developmental
157 stage, by generating individual UMAP analyses by timepoint, to examine the gene expression
158 of cells before and during primordial follicle activation (Figure 2A). Most of the clusters that
159 were identified were present in all three developmental stages (Figure 2A). However, changes
160 in cell contribution were observed in the oocyte (Oo_1) and granulosa cell cluster 1 and 4
161 (Gc_1 and Gc_4) clusters across developmental time. The number of cells contributing to the
162 Oo_1 and Gc_1 clusters decreased over time (Figure 2A and Figure S2A and S2B). In contrast,
163 of the five pregranulosa cell clusters, Gc_4 was absent at E18.5 but appeared at PND4 and was
164 maintained at PND7 (Figure 2A and Figure S2A and S2B). Cluster Gc_4 expressed granulosa
165 cell marker genes including *Foxl2, Amhr2* and *Hmgcs2* (Figure 2B and 2C). However, unlike
166 the other granulosa cell clusters Gc_1, Gc_2 and Gc_3, the Gc_4 cluster also expressed high
167 levels of genes associated with proliferation, including *Mki67* and *Pcna* (Figure 2D and 2E)
168 (Miller et al., 2018; Xu et al., 2011). Using gene ontology analysis, the top biological processes
169 from the genes upregulated in Gc_4 are related to proliferation and include cell cycle
170 (GO:0007049), mitotic nuclear division (GO:0007067), cell division (GO:0051301),
171 chromosome segregation (GO:0007059) and mitotic cytokinesis (GO:0000281) (Figure 2F).

172 Combined, the Gc_4 cluster gene expression signature is indicative of differentiated granulosa
173 cells typical of primary follicles, which have transitioned into cuboidal granulosa cells and
174 have entered the cell cycle and are proliferating to generate multiple layers of granulosa cells
175 surrounding the oocyte. Given the appearance of a population of activated granulosa cells in
176 the postnatal ovary, we then examined the transcriptomic profile of the remaining pregranulosa
177 cell populations to characterise the key gene expression differences between clusters and
178 identify a population of pregranulosa cells in a primed state for activation.

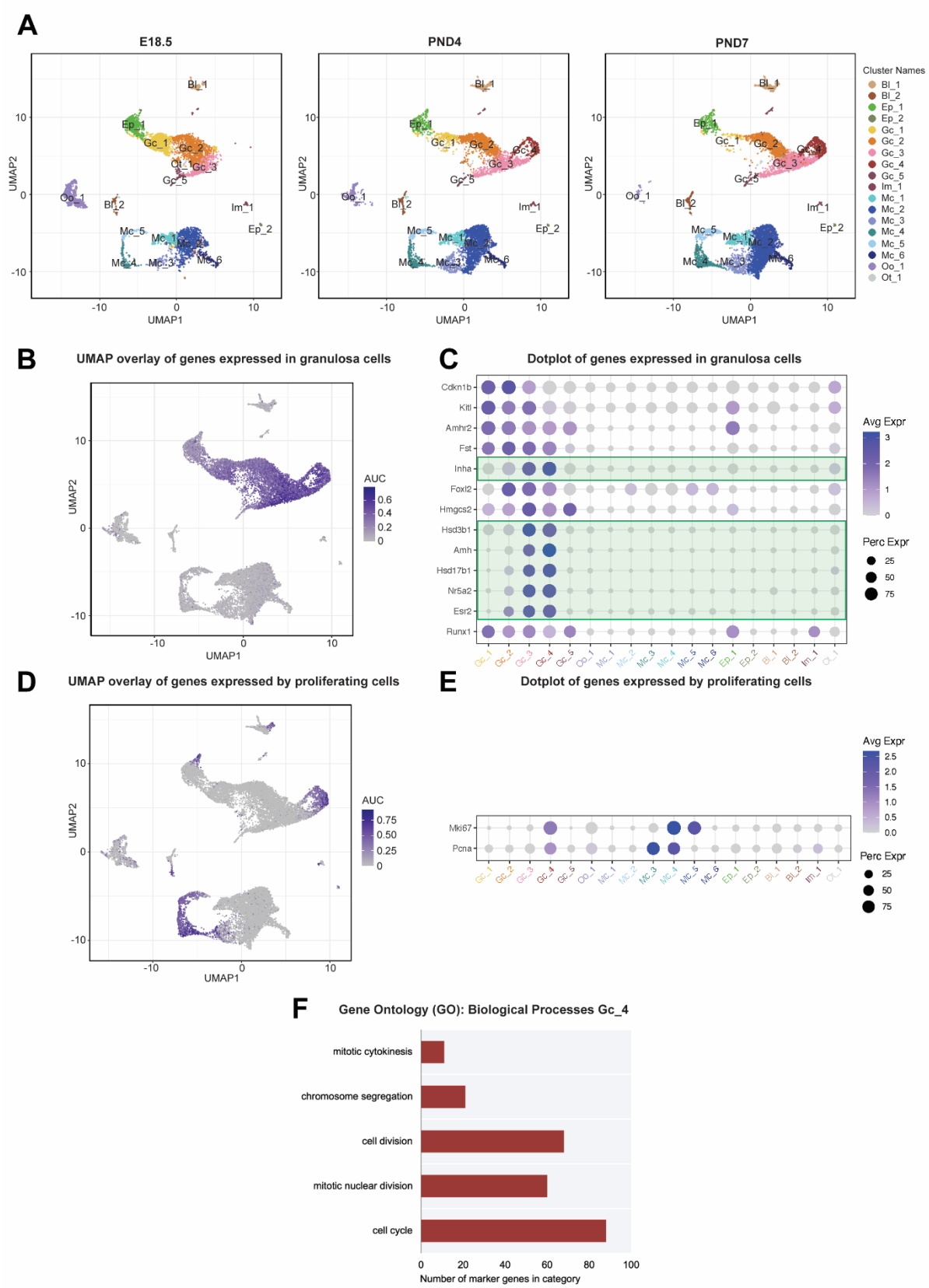


Figure 2. A distinct population of highly proliferative granulosa cells appear during primordial follicle activation in the mouse ovary

182 (A) UMAP Dimensionality Reduction Plots showing individual cells from each individual timepoint
183 coloured by cell type. (B) UMAP Dimensionality Reduction Plots showing cells from all three samples,
184 overlaid with their expression of the following genes associated with granulosa cells: *Amh, Amhr2,*
185 *Cdkn1b, Esr2, Foxl2, Fst, Hmgcs2, Hsd17b1, Hsd3b1, Inha, Kitl, Nr5a2,* and *Runx1.* Dots are coloured
186 according to their level of gene expression, with high levels of expression shown in purple. Expression
187 was determined by area under the curve (AUC). (C) Dotplot showing the expression of the above genes
188 associated with granulosa cells, with high levels of expression shown in purple and the percentage of
189 cells in each cluster expressing the gene indicated by the size of the dot. (D) UMAP Dimensionality
190 Reduction Plots showing cells from all three timepoints, overlaid with their expression of the following
191 proliferation markers: *mKi67* and *Pcna.* Dots are coloured according to their level of gene expression,
192 with high levels of expression shown in purple. Expression was determined by area under the curve
193 (AUC). (E) Dotplot showing the expression of the above genes associated with proliferation, with high
194 levels of expression shown in purple and the percentage of cells in each cluster expressing the gene
195 indicated by the size of the dot. (F) Gene ontology (GO) information of the biological process categories
196 with highest count of genes from the genes upregulated in granulosa cell cluster 4 (Gc_4). The list of
197 genes upregulated in Gc_4 was identified by differential gene expression analysis and annotated via
198 GO analysis in DAVID (v6.8).
199

200 ***Epithelial and bipotential derived pregranulosa cells exist in distinct clusters***

201 We next asked if we could characterise the remaining granulosa cell clusters by their
202 differences in gene expression. As cluster Gc_5 consisted of fewer cells with variable
203 expression of granulosa cell marker genes, we focused on the four largest pregranulosa cell
204 clusters (Gc_1 through to Gc_4). We compared the genes expressed in our granulosa cell
205 clusters with the published data from (Niu & Spradling, 2020) to similarly classify our
206 heterogenous granulosa cell populations. In granulosa cell cluster 1 (Gc_1), differential gene
207 expression analysis showed that *Gng13*, *Lgas7* and *Apoc1*, among other genes, had higher
208 levels of average expression than in all other clusters (Figure 3A and 3B, S2 Table). The gene
209 list generated from the differential gene expression analysis, along with the proximity of the
210 cluster to an epithelial cell cluster (Ep_1), suggested that this cluster is likely to be a population
211 of EPGs. This was also strengthened by the observation that most of the cells in Gc_1 were
212 from the E18.5 timepoint, as EPGs do not ingress from the ovarian epithelium to contribute to

213 follicles postnatally (Figure S2B) (Niu & Spradling, 2020). To confirm the identification of the
214 Gc_1 cluster, we used a selection of genes described as marker genes of EPGs and examined
215 their expression in all 18 clusters (Niu & Spradling, 2020) (Figure 3C). This supported the
216 differential gene expression analysis, with specific expression of EPG marker genes in
217 Gc_1.We used this same approach to identify granulosa cell cluster 2 (Gc_2) as a population
218 of pregranulosa cells derived from the bipotential *Nr5a1*+ somatic cell lineage (BPGs). This
219 cluster showed more similarity in gene expression to the other granulosa cell clusters (Figure
220 3D and 3E). However, differential gene expression analysis showed that *Serpinb6a* and
221 *Cdkn1b*, among other genes, had higher levels of average expression in Gc_2 than in all other
222 clusters (Figure 3E). This combined with the position of the cluster between the EPGs and the
223 mitotically active granulosa cells highlighted the Gc_2 cluster as a likely BPG population. To
224 verify, we compared our expression with the selection of marker genes for BPGs and found
225 high levels of *Foxl2* and *Aard*, which are first expressed in BPGs (Niu & Spradling, 2020)
226 (Figure 3F). Additionally, Gc_2 cells express low levels of *Hsd3b1* and *Hmgcs2*, markers of
227 pregranulosa cells later in development (Figure 3F). This supported our data indicating Gc_2
228 cells are BPGs.

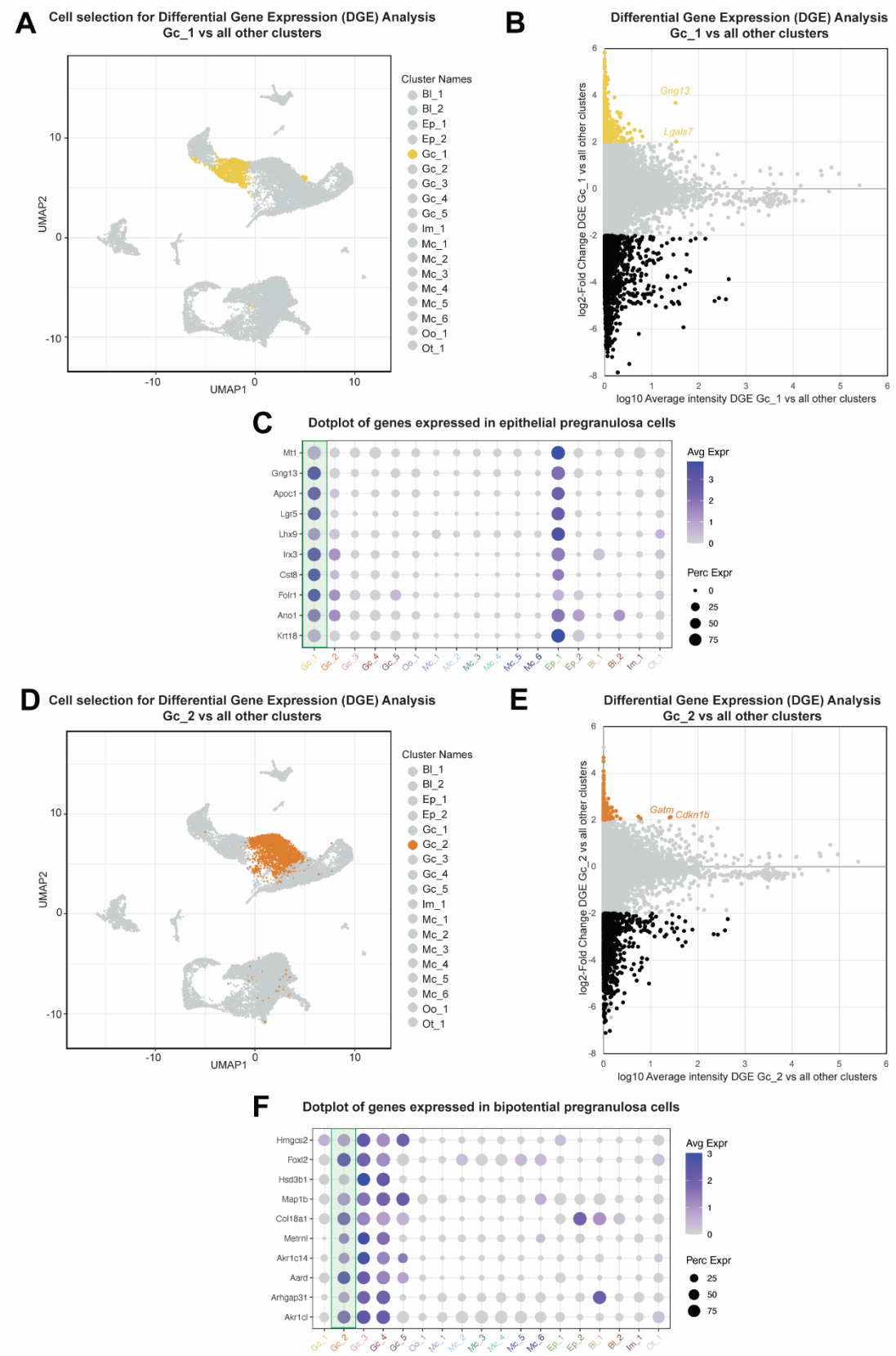


Figure 3. Clear origins of two populations of pregranulosa cells based on their gene expression signature

232 (A) UMAP Dimensionality Reduction Plots showing cells from all three samples, overlaid with the
233 cells selected for differential gene expression (DGE) analysis, Gc_1 cells are shown in yellow and all
234 other clusters are shown in grey. DGE analysis performed using the method detailed by Ahlmann-Eltze
235 (Ahlmann-Eltze & Huber, 2020). (B) MA plot showing the differential expression of gene transcripts
236 between Gc_1 cells and cells from all other clusters; genes upregulated in Gc_1 cells are shown in
237 yellow and genes upregulated in all other cells shown in black. Grey dots indicate genes below the $\log_2$
238 fold change ± 2; $p \leq 0.05$ cutoff. (C) Dotplot showing the expression of genes associated with epithelial
239 pregranulosa cells, with high levels of expression shown in purple and the percentage of cells in each
240 cluster expressing the gene indicated by the size of the dot. Gene list obtained from (Niu & Spradling,
241 2020). (D) UMAP Dimensionality Reduction Plots showing cells from all three samples, overlaid with
242 the cells selected for differential gene expression (DGE) analysis, Gc_2 cells are shown in orange and
243 all other clusters are shown in grey. DGE analysis performed using the method detailed by Ahlmann-
244 Eltze (Ahlmann-Eltze & Huber, 2020). (E) MA plot showing the differential expression of gene
245 transcripts between Gc_2 cells and cells from all other clusters; genes upregulated in Gc_2 cells shown
246 in orange and genes upregulated in all other cells shown in black. Grey dots indicate genes below the
247 $\log_2$ fold change ± 2; $p \leq 0.05$ cutoff. (F) Dotplot showing the expression of genes associated with
248 bipotential pregranulosa cells, with high levels of expression shown in purple and the percentage of
249 cells in each cluster expressing the gene indicated by the size of the dot. Gene list obtained from (Niu
250 & Spradling, 2020).
251

252 ***Identification of the genetic signature of pre-granulosa cells in activating primordial***
253 ***follicles***

254 The identification of the Gc_1 and Gc_2 clusters as EPGs and BPGs respectively, led us to
255 hypothesise the genetic profile of Gc_3 as a potential population of activating granulosa cells.
256 Cells that contribute to Gc_3 originate from all three timepoints, but a large proportion of cells
257 arise from the PND7 timepoint (Figure S2B). Interestingly, cells in Gc_3 have similar gene
258 expression signatures to BPGs in Gc_2 (shown by the expression of *Foxl2* and *Hmgcs2*
259 amongst others), but also high average levels of expression of genes like *Kitl* and *Fst*, which
260 are well-established regulators of follicle activation (Figure 4A and 4B) (Hutt et al., 2006; Z.
261 Wang et al., 2015). Additionally, the cells in the Gc_3 cluster also express markers of
262 differentiated granulosa cells, such as *Hsd3b1*, *Amh*, *Esr2* and *Nr5a2*, albeit at lower average

263 levels of expression when compared to the differentiated proliferative granulosa cells in Gc_4
264 (Figure 2C and 4B). Differential gene expression analysis illustrated that *Hsd3b1, Fst* and *Inha*
265 among other genes, were on average more highly expressed in Gc_3 than in other cell clusters,
266 supporting the hypothesis that these cells are activating pregranulosa cells (Figure 2C and 4B
267 and S2 Table).
268
269 To determine if Gc_3 is an intermediate stage between quiescent BPG cells (Gc_2) and
270 differentiated proliferating granulosa cells (Gc_4), we performed two differential gene
271 expression (DGE) analyses between Gc_3 and Gc_2, and Gc_3 and Gc_4. The 3,302 cells that
272 contributed to Gc_3 and 1,803 cells that contributed to Gc_2 were selected and average gene
273 expression was analysed according to the methods detailed by Ahlmann-Eltze (Ahlmann-Eltze
274 & Huber, 2020) (Figure 4C). There was a total of 581 differentially expressed genes ($\log_2$ fold
275 change $\pm$ 2; $p \leq 0.05$) between cells from Gc_3 and Gc_2 (Figure 4D and S3 Table). 274 genes
276 were upregulated in Gc_3 cells and 307 genes were upregulated in Gc_2 cells (Figure 4D and
277 S3 Table). Gene ontology was used to identify the top biological pathways implicated in Gc_2
278 cells and Gc_3 respectively (Figure 4E and 4F). In the Gc_2 cluster, the top biological
279 processes based on 307 differentially expressed genes were ion transport (GO:0006811),
280 defence response to virus (GO:0051607), transmembrane transport (GO:0055085), immune
281 system process (GO:0002376) and immune response (GO:0006955) (Figure 4E). In contrast,
282 the top processes based on the 274 differentially expressed genes in the Gc_3 cluster were cell
283 cycle (GO:0007049), cellular response to DNA damage stimulus (GO:0006974), DNA
284 replication (GO:0006260), cell adhesion (GO:0007155) and DNA repair (GO:0006281)
285 (Figure 4F). These processes illustrate that cell cycle and the associated DNA replication
286 processes are occurring in the cells of Gc_3, which is a critical cell fate transition towards
287 differentiated granulosa cells.

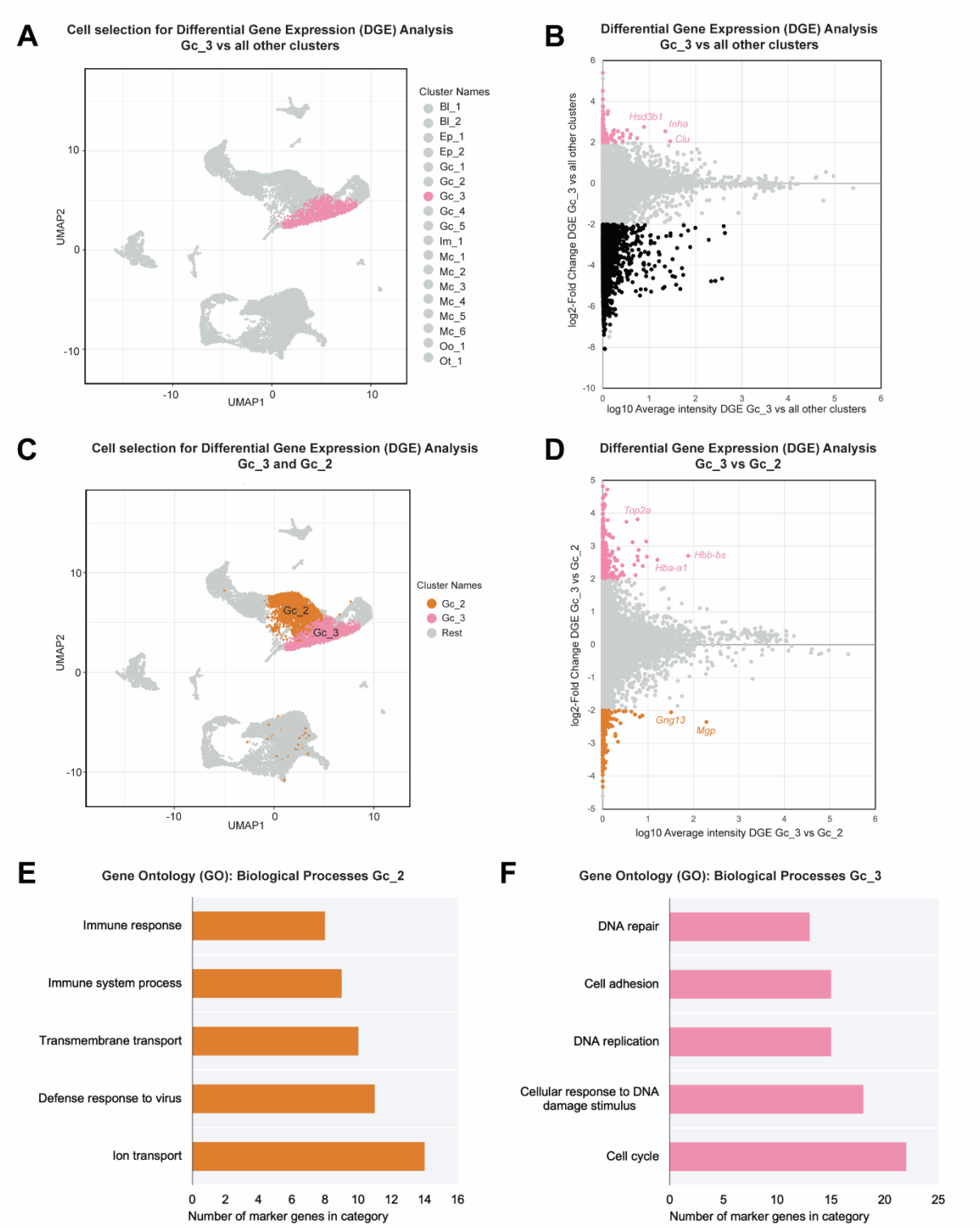


Figure 4. Comparison of gene expression between putative activating pregranulosa cells (Gc_3) and bipotential pregranulosa cells (Gc_2)

(A) UMAP Dimensionality Reduction Plots showing cells from all three samples, overlaid with the cells selected for differential gene expression (DGE) analysis, Gc_3 cells are shown in pink and all other clusters are shown in grey. DGE analysis performed using the method detailed by Ahlmann-Eltze

294 (Ahlmann-Eltze & Huber, 2020). (B) MA plot showing the differential expression of gene transcripts
295 between Gc_3 cells and cells from all other clusters; genes upregulated in Gc_3 cells are shown in pink
296 and genes upregulated in all other cells shown in black. Grey dots indicate genes below the $\log_2$ fold
297 change ± 2; $p \leq 0.05$ cutoff. (C) UMAP Dimensionality Reduction Plots showing cells from all three
298 samples, overlaid with the cells selected for differential gene expression (DGE) analysis, Gc_3 cells are
299 shown in pink and Gc_2 cells are shown in orange. DGE analysis performed using the method detailed
300 by Ahlmann-Eltze (Ahlmann-Eltze & Huber, 2020). (D) MA plot showing the differential expression
301 of gene transcripts between Gc_3 and Gc_2 cells; genes upregulated in Gc_3 cells shown in pink and
302 genes upregulated in Gc_2 cells shown in orange. Grey dots indicate genes below the $\log_2$ fold change
303 ± 2; $p \leq 0.05$ cutoff. (E) Gene ontology (GO) information of the biological process categories with
304 highest count of genes from the genes upregulated in Gc_2 shown in orange. (F) Gene ontology (GO)
305 information of the biological process categories with highest count of genes from the genes upregulated
306 in Gc_3 shown in pink. The lists of genes upregulated in Gc_2 and Gc_3 were identified by differential
307 gene expression analysis and annotated via GO analysis in DAVID (v6.8).
308

309 To characterise a gene expression signature of primordial follicle activation, we compared the
310 differentially expressed genes from the differential gene expression analysis performed in
311 Figure 4A (S2 Table) and the differentially expressed genes from the Gc_3 vs Gc_2
312 comparison performed in Figure 4C (S3 Table). Established pregranulosa marker genes of
313 primordial follicle activation including *Amh*, *Inha* and *Nr5a2* were observed when these gene
314 lists were compared (Figure S3). The expression of these genes validated our approach for
315 identifying a gene expression signature of activation, and we then looked for novel genes
316 potentially controlling activation. Genes including *Slc18a2*, *Tnni3* and *Fam13a* are suitable
317 candidates that share similar gene expression profiles to *Amh*, *Inha* and *Nr5a2*, and are
318 representative of cells entering activation (Figure 5).

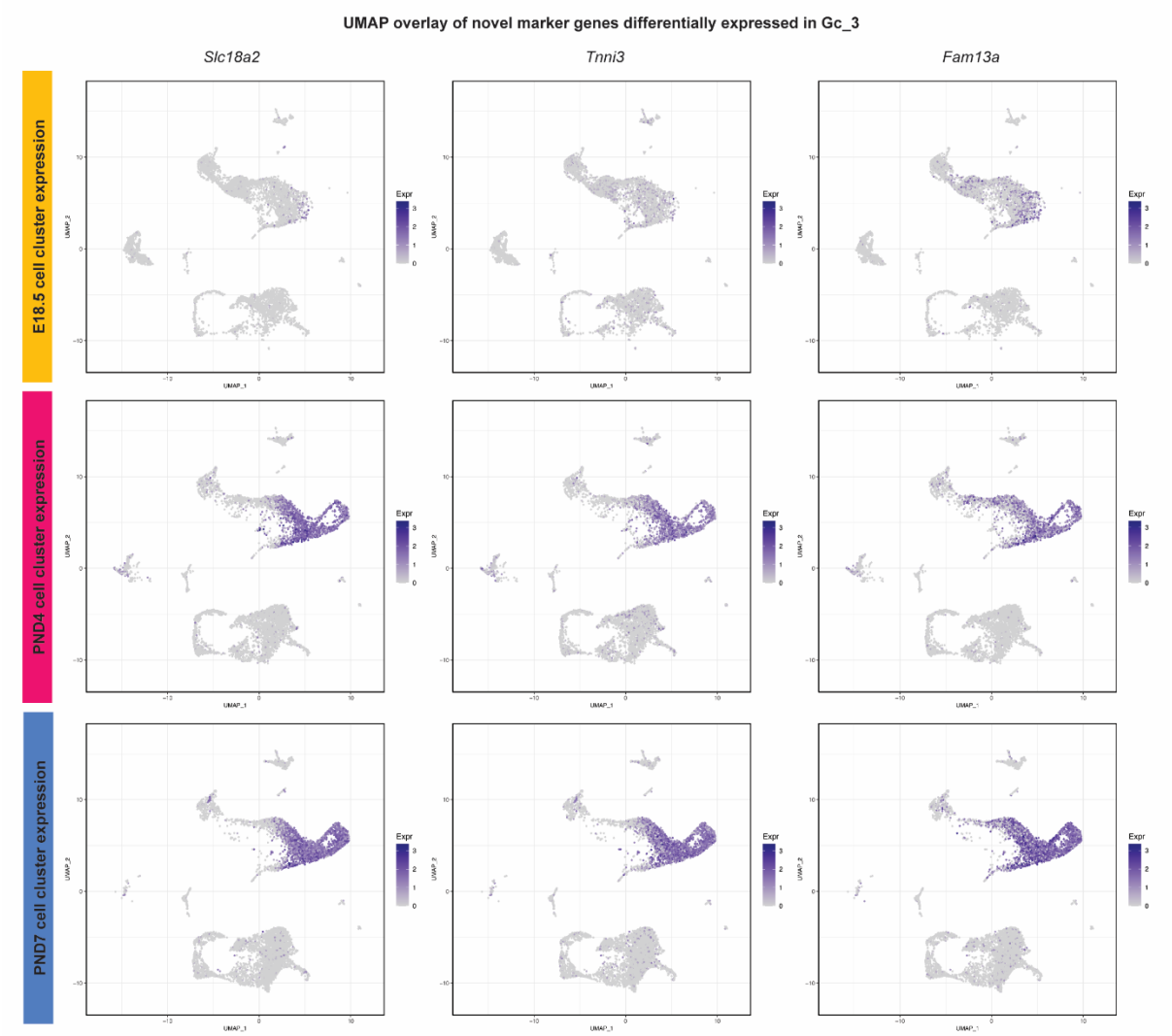


Figure 5. Expression of key genes upregulated during primordial follicle activation
The UMAP cluster maps show the expression of differentially expressed genes in Gc_3. UMAP cluster maps of novel pregranulosa marker genes (*Slc18a2*, *Tnni3* and *Fam13a*) identified in our differential gene expression analysis and DGE comparisons that are upregulated during primordial follicle activation. UMAP cluster maps are shown for each timepoint individually and only include cells from the timepoints in each panel (E18.5 = yellow, PND4 = pink and PND7 = blue). Dots are coloured according to their level of gene expression, with high levels of expression shown in purple.

We also performed DGE analysis between the activating pregranulosa cells (Gc_3) and the differentiated granulosa cells (Gc_4). The same method was utilised as previously, with 3,302 Gc_3 cells selected and 909 Gc_4 cells selected for gene expression analysis (Figure 6A). 667 differentially expressed genes ($\log_2$ fold change ± 2; $p \leq 0.05$) were identified between cells

332 from Gc_3 and Gc_4 (Figure 6B and S4 Table). 389 genes were upregulated in Gc_3 cells,
333 compared with 278 genes upregulated in Gc_4 cells (Figure 6B and S4 Table). The top
334 biological processes associated with the genes expressed in Gc_3 cells were cell differentiation
335 (GO:0030154), multicellular organism development (GO:0007275), spermatogenesis
336 (GO:0007283), meiotic cell cycle (GO:0051321) and metabolic process (GO:0008152) (Figure
337 6C). Gc_4 cells had distinct biological processes divergent from Gc_3 cells, namely cell cycle
338 (GO:0007049), mitotic nuclear division (GO:0007067), cell division (GO:0051301),
339 chromosome segregation (GO:0007059) and nucleosome assembly (GO:0006334) (Figure
340 6D). Taken together, these results provided a list of genes that are expressed in the pregranulosa
341 cells at the initiation of primordial follicle activation, prior to re-entry into the cell cycle.

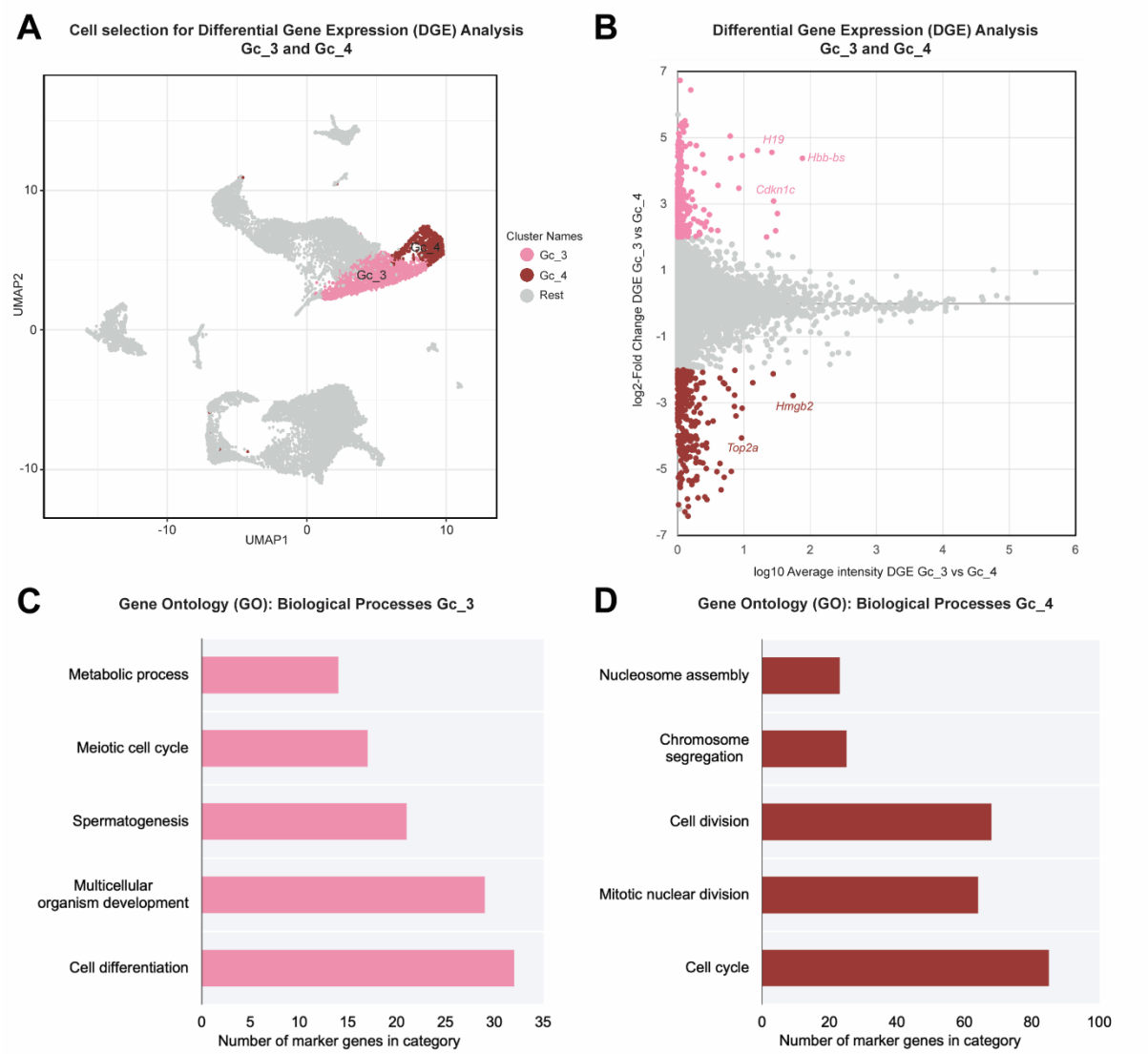


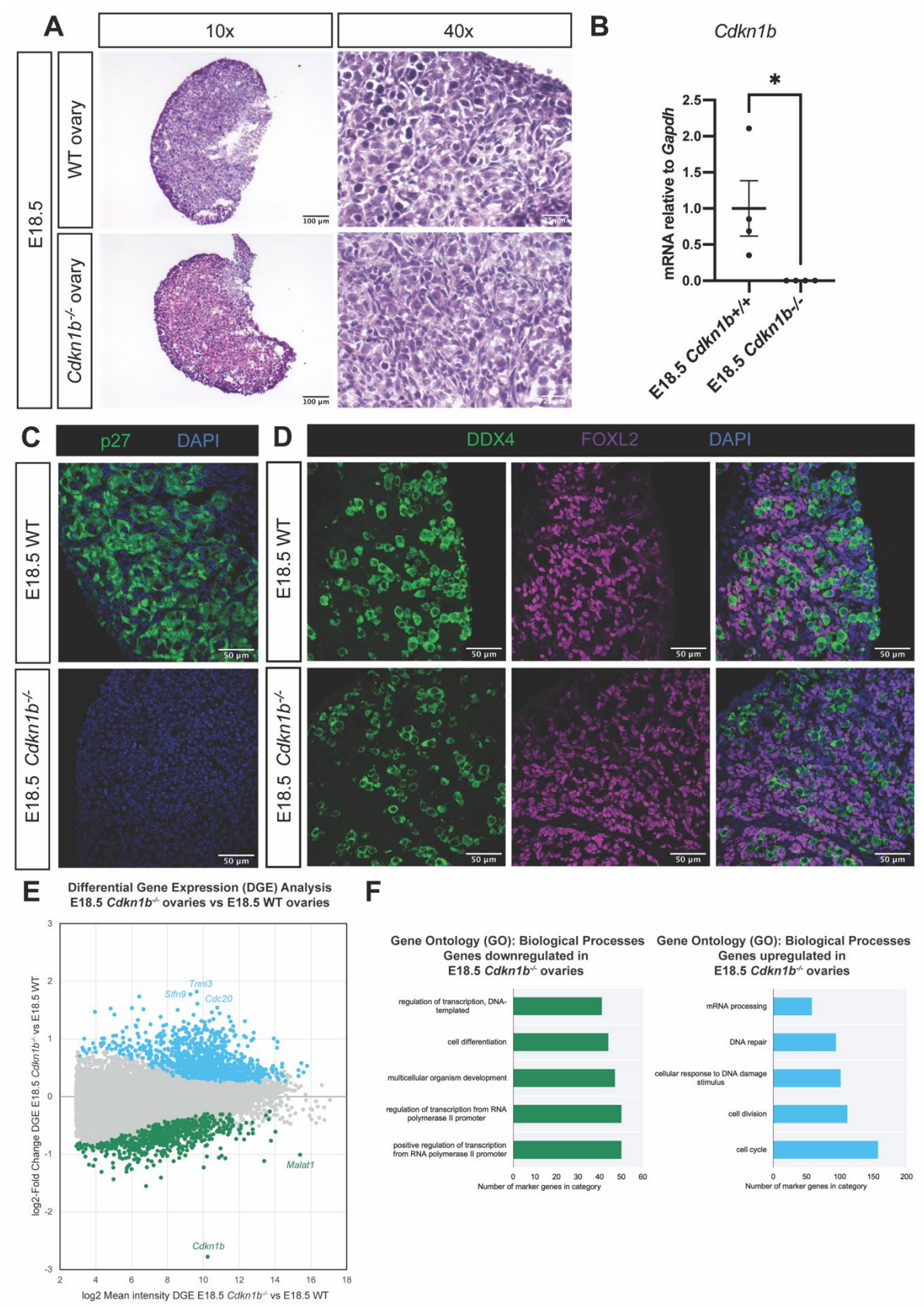
Figure 6. Comparison of gene expression between granulosa cell cluster 3 (Gc_3) and granulosa cell cluster 4 (Gc_4)

(A) UMAP Dimensionality Reduction Plots showing cells from all three samples, overlaid with the cells selected for differential gene expression (DGE) analysis, Gc_3 cells are shown in pink and Gc_4 cells are shown in maroon. DGE analysis performed using the method detailed by Ahlmann-Eltze (Ahlmann-Eltze & Huber, 2020). (B) MA plot showing the differential expression of gene transcripts between Gc_3 and Gc_4 cells; genes upregulated in Gc_3 cells shown in pink and genes upregulated in Gc_4 cells shown in maroon. Grey dots indicate genes below the $\log_2$ fold change ± 2; $p \leq 0.05$ cutoff. (C) Gene ontology (GO) information of the biological process categories with highest count of genes from the genes upregulated in Gc_3 shown in pink. (D) Gene ontology (GO) information of the biological process categories with highest count of genes from the genes upregulated in Gc_4 shown in maroon. The lists of genes upregulated in Gc_2 and Gc_3 were identified by differential gene expression analysis and annotated via GO analysis in DAVID (v6.8).

356 ***Examining the changes in gene expression signature in a model of accelerated primordial***
357 ***follicle activation***

358 Having identified putative genes expressed in the pregranulosa cells during primordial follicle
359 activation, we next looked to query these candidate genes in an established model of abnormal
360 follicle activation. Specifically, the genetically modified mouse line *Cdkn1b*$^{-/-}$ which lacks the
361 expression of the cell cycle inhibitor p27^{kip1} (gene name *Cdkn1b*). This mutant line shows an
362 uncontrolled activation of follicles which enter folliculogenesis perinatally, inducing infertility
363 prior to sexual maturation (Perez-Sanz et al., 2013; Rajareddy et al., 2007). No previous RNA
364 sequencing experiments have been performed on ovaries from *Cdkn1b*$^{-/-}$ mice. We examined
365 the ovaries of *Cdkn1b*$^{-/-}$ and littermate wildtype controls at E18.5. This timepoint was selected
366 as it was reported to be prior to primordial follicle activation occurring in both wild-type and
367 *Cdkn1b*$^{-/-}$ ovaries (Rajareddy et al., 2007). Additionally, this timepoint covers germ cell nest
368 breakdown and primordial follicle formation and overlaps with the first timepoint of the single-
369 cell sequencing experiment. As previously reported, no gross morphological differences in
370 histology were observed at E18.5 (Figure 7A). The absence of p27kip1 in *Cdkn1b*$^{-/-}$ ovaries
371 was confirmed by qRT-PCR (Figure 7B) and by immunofluorescence (Figure 7C). We
372 assessed that the follicle activation defect was not due to a failure of germ cells or pregranulosa
373 cells to specify correctly in *Cdkn1b*$^{-/-}$ mice, by staining for DDX4 (germ cell marker) and
374 FOXL2 (granulosa cell marker) at E18.5 (Figure 7D). Having established the *Cdkn1b*$^{-/-}$ mouse
375 line as an appropriate model to examine primordial follicle activation defects, we performed
376 bulk RNA-sequencing of WT and *Cdkn1b*$^{-/-}$ ovaries at E18.5. Surprisingly, at E18.5, 1,633
377 transcripts were already differentially expressed between WT and *Cdkn1b*$^{-/-}$ ovaries ($p < 0.05$).
378 Of these genes, 1,008 genes were upregulated in *Cdkn1b*$^{-/-}$ ovaries and 626 genes were
379 downregulated in *Cdkn1b*$^{-/-}$ ovaries (Figure 7E and S5 Table). Gene ontology was utilised to
380 identify the top biological pathways affected in *Cdkn1b*$^{-/-}$ ovaries (Figure 7F). The top

381 processes based on the 626 downregulated DEGs in *Cdkn1b*$^{-/-}$ ovaries were positive regulation
382 of transcription from RNA polymerase II promoter (GO:0045944), regulation of transcription
383 from RNA polymerase II promoter (GO:0006357), multicellular organism development
384 (GO:0007275), cell differentiation (GO:0030154) and regulation of transcription, DNA-
385 templated (GO:0006355) (Figure 7F). The top processes based on the 1,008 upregulated DEGs
386 in *Cdkn1b*$^{-/-}$ ovaries were cell cycle (GO:0007049), cell division (GO:0051301), cellular
387 response to DNA damage stimulus (GO:0006974), DNA repair (GO:0006281) and mRNA
388 processing (GO:0006397) (Figure 7F).



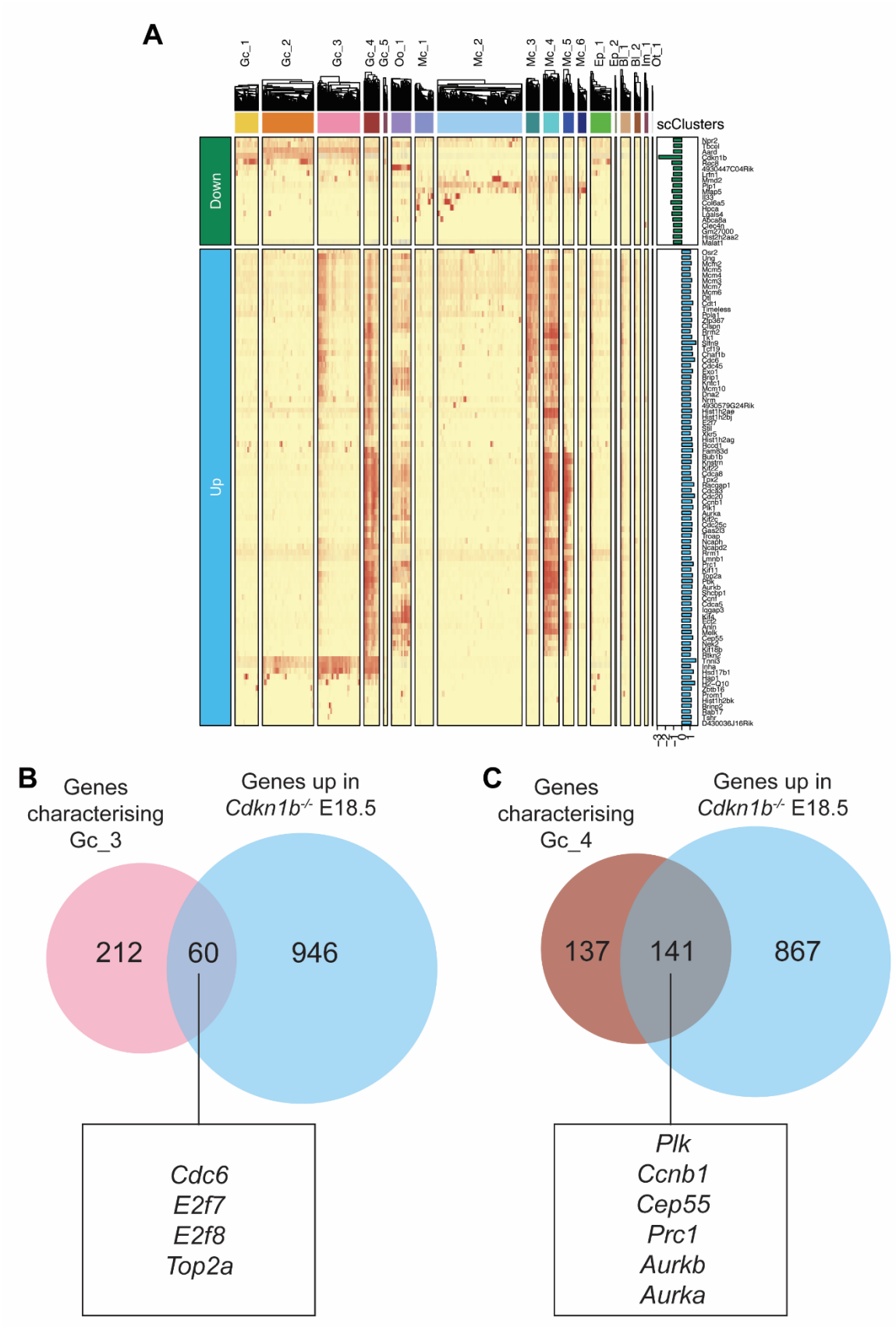
389

390 **Figure 7. Examination of ovarian development and gene expression in E18.5 *Cdkn1b*$^{-/-}$ mice.**

391 (A) Haematoxylin and Eosin (H&E) staining of E18.5 ovaries in both wild-type (WT) and *Cdkn1b*$^{-/-}$
392 mice. Images were taken at 10 x magnification and 40 x magnification, with scale bars of 100 µm and
393 25 µm respectively. (B) Relative mRNA expression of *Cdkn1b* in E18.5 wild-type (WT) and *Cdkn1b*$^{-/-}$
394 ovaries. The data are represented as mean ± SEM (* $p < 0.05$). (C) Cellular localisation of p27^{kip1} in
395 wild-type (WT) ovaries and *Cdkn1b*$^{-/-}$ ovaries at E18.5. Ovaries were stained for p27^{kip1} at E18.5 by
396 immunofluorescence. Images were taken at 20 x magnification, with scale bars of 50 µm. (D) Cellular
397 localisation of DDX4 and FOXL2 in E18.5 wild-type (WT) and *Cdkn1b*$^{-/-}$ ovaries. Ovaries were stained
398 with the oocyte marker DDX4 and the granulosa cell marker FOXL2 at E18.5 by immunofluorescence.
399 Images were taken at 20 x magnification, with scale bars of 50 µm. (E) MA plot indicates the differential
400 expression of gene transcripts between WT E18.5 dissociated ovarian cells and *Cdkn1b*$^{-/-}$ dissociated
401 ovarian cells (n = 4 replicates per group). Genes upregulated in *Cdkn1b*$^{-/-}$ ovaries are shown in blue and
402 genes downregulated in *Cdkn1b*$^{-/-}$ ovaries shown in green. Grey dots indicate genes below the $p \leq 0.05$
403 cutoff. (F) Gene ontology (GO) information of the biological process categories with top 5 highest
404 count of genes from the genes differentially regulated in *Cdkn1b*$^{-/-}$ ovaries. The list of genes was
405 identified by differential gene expression analysis ($p \leq 0.05$) and annotated via GO analysis in DAVID
406 (v6.8).
407

408 To confirm if the differentially expressed genes detected in *Cdkn1b*$^{-/-}$ ovaries at E18.5 were
409 altered due to changes in the activation status of granulosa cells, we overlaid the *Cdkn1b*$^{-/-}$ bulk
410 sequencing data on the single-cell sequencing data to identify genes that were dysregulated in
411 *Cdkn1b*$^{-/-}$ ovaries at E18.5 and are expressed in granulosa cell clusters in the wild-type context
412 (Figure 8A). This identified that the majority of differentially expressed genes in *Cdkn1b*$^{-/-}$
413 ovaries at E18.5 were expressed in granulosa cell clusters Gc_3 and Gc_4 and mesenchymal
414 clusters Mc_3, Mc_4 and Mc_5 in the wild-type ovary (Figure 8A). To better understand the
415 genes that are upregulated in granulosa cells of *Cdkn1b*$^{-/-}$ ovaries, we compared the 1,008
416 differentially expressed genes that were upregulated in *Cdkn1b*$^{-/-}$ ovaries with the 274 genes
417 upregulated in activating pregranulosa cells (Gc_3) (Figure 8B and S6 Table). 60 genes were
418 present in both groups, with most of these factors (70%) relating to cell cycle processes. This
419 suggested that the gene expression signature of activation that we see present in Gc_3 cells in
420 the single-cell experiment is similarly present in *Cdkn1b*$^{-/-}$ ovaries at E18.5 (Figure 8B and S6

421 Table). Moreover, this is developmentally earlier than primordial follicle activation in wild-
422 type ovaries, shown by the proportion of Gc_3 cells originating mostly from PND4 and PND7
423 (Figure S2). In summary, primordial follicle activation begins earlier in
424 *Cdkn1b*$^{-/-}$ ovaries, with the altered expression of genes more representative of activating
425 pregranulosa cells, rather than BPGs or EPGs (Figure 8B). To check if activated granulosa cell
426 genes could also be identified in E18.5 *Cdkn1b*$^{-/-}$ ovaries, we compared the 1,008 differentially
427 expressed genes that were upregulated in E18.5 *Cdkn1b*$^{-/-}$ ovaries with the 278 genes
428 upregulated in differentiated granulosa cells (Gc_4) (Figure 8C and S7 Table). Indeed, 141
429 genes were present in both groups, with several of these genes being known targets of p27^{kip1},
430 including *Ccnb1*, *Aurka* and *Aurkb*. Taken together, these findings show that the genes
431 associated with both activating and activated pregranulosa cells are precociously expressed in
432 *Cdkn1b*$^{-/-}$ ovaries at E18.5, despite no gross morphological changes in germ cell nest structures.



433

434 **Figure 8. Precocious gene expression of differentiated granulosa cells in *Cdkn1b*$^{-/-}$ ovaries at E18.5**

435 (A) Heat map showing the expression of differentially expressed genes in *Cdkn1b*$^{-/-}$ ovaries ($p \leq 0.05$)
436 in the context of the cell clusters from the single-cell sequencing experiment. Genes highly expressed
437 in each cluster are shown in red and genes lowly expressed in each cluster are shown in yellow. Clusters
438 are labelled across the top of the heatmap and coloured according to the UMAP. The logFC values for
439 genes in the bulk *Cdkn1b*$^{-/-}$ experiment are shown on the right side of the UMAP, and clustered
440 according to their pattern of expression. Upregulated genes are shown in blue and downregulated genes
441 are shown in green. (B) Venn diagram comparing the 274 genes differentially expressed in Gc_3 cells
442 vs. Gc_2 cells (red circle) with the 1,008 genes upregulated in E18.5 *Cdkn1b*$^{-/-}$ ovaries vs. E18.5 WT
443 ovaries (blue circle). 60 genes upregulated in Gc_3 cells were also upregulated in *Cdkn1b*$^{-/-}$ ovaries. For
444 cell clusters, genes were identified as differentially expressed by single-cell RNA sequencing ($\log_2$ fold
445 change $\pm$ 2; $p \leq 0.05$) and for *Cdkn1b*$^{-/-}$ ovaries genes were identified by bulk RNA-sequencing ($p \leq$
446 0.05). (C) Venn diagram comparing the 278 genes differentially expressed in Gc_4 cells vs. Gc_3 cells
447 (orange circle) with the 1,008 genes upregulated in E18.5 *Cdkn1b*$^{-/-}$ ovaries vs. E18.5 WT ovaries (red
448 circle). 141 genes that are upregulated in Gc_4 cells are also upregulated in *Cdkn1b*$^{-/-}$ ovaries. For cell
449 clusters, genes were identified as differentially expressed by single-cell RNA sequencing ($\log_2$ fold
450 change $\pm$ 2; $p \leq 0.05$) and for *Cdkn1b*$^{-/-}$ ovaries genes were identified by bulk RNA-sequencing ($p \leq$
451 0.05).
452

453 **Discussion**

454 In this study, we have identified a putative primordial follicle activation gene expression
455 signature in the pregranulosa cells of the mouse ovary. By utilising single cell RNA-
456 sequencing, we were able to resolve five distinct clusters of granulosa cells in three timepoints:
457 E18.5, PND4 and PND7. One of these distinct granulosa cell clusters was identified as a
458 population of activating pregranulosa cells, enabling our examination of the pathways that
459 govern the first wave of primordial follicle activation. Genes which showed specific expression
460 in this cluster of pregranulosa cells were identified and hypothesised to interact to promote and
461 control follicle activation. To functionally interrogate this hypothesis, we utilised the
462 *Cdkn1b*$^{-/-}$ mouse line which has an accelerated primordial follicle activation phenotype in the
463 neonatal mouse ovary. Transcriptomic analysis of these ovaries revealed distinct gene
464 expression changes at E18.5, prior to the normal developmental window of primordial follicle
465 activation. These gene changes had not been observed in *Cdkn1b*$^{-/-}$ ovaries previously.

466 Differentially expressed genes upregulated in *Cdkn1b*$^{-/-}$ ovaries were also observed in the gene
467 expression signature of activating pregranulosa cells, illustrating a precocious primordial
468 follicle activation process in *Cdkn1b*$^{-/-}$ ovaries, regulated at least in part by p27^{kip1} maintaining
469 cell cycle arrest and preventing granulosa cell activation.
470
471 ***A transcriptomic dataset to explore primordial follicle activation in mice***
472 In this study, we have investigated the transcriptome of the embryonic and perinatal developing
473 mouse ovary. We successfully captured the dominant cell types in the ovary, namely the
474 oocytes, pregranulosa and mesenchymal cells at three developmental timepoints by analysing
475 24,810 cells. One advantage of this dataset is the ability to interrogate any ovarian cell type
476 and its contribution to primordial follicle activation. Additionally, this dataset complements
477 single-cell datasets generated to examine pre-granulosa cell specification (Niu & Spradling,
478 2020) and primordial follicle formation (J.-J. Wang et al., 2020). This study provides a valuable
479 resource for interrogating primordial follicle activation at the transcriptomic level either with
480 a cell-centric view or focusing on cell-cell interactions. Notably we observed a decrease in the
481 number of cells in the Oo_1 cluster, particularly at PND7 (Figure 2A). The dissociation
482 protocol used to prepare the ovarian samples for single cell RNA-sequencing in this study
483 enriches the suspension for somatic cells, rather than germ cells, via size selection, as oocytes
484 increase in size as activation occurs and primordial follicles transition to primary follicles
485 (Frost et al., 2021b). This may explain the decrease in oocyte numbers in the postnatal
486 timepoints (Zhang et al., 2014). While in the current study, our focus was on identifying
487 granulosa cell activators of follicle activation, opportunities exist to further analyse this dataset
488 by investigating receptor/ligand pathways and secreted factors that may be initiated in the
489 granulosa cells and then activated or passed to the oocytes to facilitate their growth.
490

491 ***Pregranulosa cells cluster separately depending on cellular origin and activation status***

492 We were able to identify five distinct clusters of granulosa cells in three timepoints: at E18.5,
493 PND4 and PND7, providing supporting evidence for heterogeneity in granulosa cell
494 populations (Cordeiro et al., 2015; Mork et al., 2012; Niu & Spradling, 2020). Activated
495 granulosa cells are the simplest granulosa cell type to identify, due to the extensive
496 characterisation of late-stage follicles and the function of granulosa cells to respond to follicle-
497 stimulating hormone and produce estrogen (Francois et al., 2017; O'Shaughnessy, McLelland,
498 & McBride, 1997). Activated granulosa cells also resume proliferation to support follicle
499 growth and ensure multiple layers of granulosa cells to form cumulus-oocyte complexes
500 necessary for ovulation (Cannon, Cherian-Shaw, Lovekamp-Swan, & Chaffin, 2007; Rebecca
501 L. Robker & JoAnne S. Richards, 1998; R. L. Robker & J. S. Richards, 1998). In our dataset,
502 GC_4 expressed *Amh*, *Nr5a1* and *Esr2*, along with higher levels of *Hsd3b1* and *mKi67*,
503 characteristic of activated granulosa cells (Figure 2C). The identification of activated granulosa
504 cells illustrates that our dataset captures not only primordial follicles, but also primary follicles,
505 which are the second stage of folliculogenesis and represent a follicle stage after primordial
506 follicle activation.

507

508 In addition to identifying activated granulosa cells, we also validated the presence of
509 populations of both BPG and EPG populations (Figure 3). BPGs express *Foxl2* from E12.5 and
510 create medullary primordial follicles during germ cell nest breakdown (Mork et al., 2012; Niu
511 & Spradling, 2020; J.-J. Wang et al., 2020). These cells appear primed to support activation
512 just after birth, based on the expression of genes including *Hmgcs2*, *Aard* and *Akr1cl*, which
513 are upregulated further in activated granulosa cells (Figure 3D). EPGs were more closely
514 clustered with epithelial cells, characterised by strong expression of *Lgr5*, *Gng13* and *Cst8*
515 (Figure 3B). In studies performed using an Lgr5-CreERT2 mouse line, Niu et al. established

516 that EPGs contribute to cortical primordial follicles, which support fertility throughout
517 adulthood (Niu & Spradling, 2020). In our study, the number of EPGs in GC_3 decreased
518 dramatically in the PND4 and PND7 timepoints (Figure S2B). It is unclear whether our
519 clustering strategy masks the subtle gene expression differences in EPGs when they upregulate
520 FOXL2 postnatally, or if EPGs become more transcriptionally similar to bipotential
521 pregranulosa cells (Pisarska, Barlow, & Kuo, 2011). It is of interest to consider whether the
522 differences in these two pregranulosa cell types persist into adulthood and whether
523 differentiated granulosa cells retain specific gene expression signatures relating to their origin.
524 Lineage tracing studies using *Lgr5*-GFP and/or *Foxl2*-GFP, combined with single-cell
525 sequencing experiments on later timepoints such as PND21 would shed light on the trajectory
526 of these cells and how similar or distinct their transcriptomes become.
527
528 ***A signature of gene activity regulating primordial follicle activation in pregranulosa cells***
529 In our analysis, one distinct granulosa cell cluster was observed to be a population of activating
530 pregranulosa cells, allowing us to map and examine the gene networks that govern the first
531 wave of primordial follicle activation. This represents one of the first single cell RNA-
532 sequencing dataset to examine the transcriptomic profile of granulosa cells during the first
533 wave of primordial follicle activation in the mouse ovary (Y.-Y. Chen et al., 2022). We have
534 identified a gene expression signature of primordial follicle activation in Gc_3 cells, using
535 novel bioinformatic differential gene expression approaches. This signature includes *Tnni3* and
536 *Slc18a2*. *Tnni3* has been identified in human granulosa cells and genetic variants of *Slc18a2*
537 are associated with polycystic ovary syndrome (Fan et al., 2019; Q. Li et al., 2016). Further
538 work to piece together where these genes function and how they cooperate with other
539 primordial follicle activation networks is essential. It is important to note that the mechanisms
540 controlling primordial follicle activation are poorly understood. Combining transcriptomic and

541 proteomic approaches, such as chromatin immunoprecipitation followed by selective isolation
542 of chromatin-associated proteins (ChIP-SICAP) will yield extensive detail on how these
543 networks translate to the protein level (Rafiee, Girardot, Sigismondo, & Krijgsveld, 2016).
544 Post-translational modifications can also be investigated using mass spectrometry, which has
545 not yet been performed on an early postnatal ovarian time-course to examine timepoint specific
546 modifications. Mechanisms of activation may not be only limited to gene expression
547 signatures, with mechanical stress and autophagy recently implicated in follicle activation
548 (Nagamatsu, Shimamoto, Hamazaki, Nishimura, & Hayashi, 2019; Zhihan et al., 2019). This
549 transcriptomic dataset provides the essential groundwork for future studies into primordial
550 follicle activation by providing cell-specific transcriptomic profiles and identifying a profile of
551 cells before and during follicle activation. Combining this dataset with previously published
552 single cell RNA-sequencing studies of the developing ovary could establish a time course of
553 gene expression across the processes of germ cell nest breakdown, primordial follicle
554 formation and activation (Niu & Spradling, 2020; Stevant et al., 2019; J.-J. Wang et al., 2020).
555 Exciting opportunities also exist to correlate this data with human ovary single cell RNA-
556 sequencing studies and that from other species to uncover how well the process is conserved.
557

558 ***Disruption of the genetic primordial follicle activation network in Cdkn1b$^{-/-}$ mice***

559 A valuable inclusion in this study was the functional assessment of our newly identified gene
560 expression signature against the sequencing profile of a mouse model of infertility. This
561 combination of single cell RNA-sequencing data with bulk RNA-sequencing supports and
562 strengthens our characterisation of the genes involved in primordial follicle activation. An
563 important confirmation was the observation that cell cycle control is not just a side effect of
564 activation, but cell cycle control genes are integrated in the network that activates primordial
565 follicles. Both in Gc_3 and Gc_4, cell cycle processes dominate the gene ontology analysis and

566 illustrate the essential role that cell cycle resumption plays in follicle activation. 36% of the
567 genes identified in activating pregranulosa cells (Gc_3) and differentiated granulosa cells
568 (Gc_4) were upregulated in *Cdkn1b*$^{-/-}$ ovaries. It remains to be elucidated if these changes are
569 primarily due to the loss of a cell cycle inhibitor or alternatively, a downstream effect of
570 increased activation. Two genes, *Cep55* and *Prc1,* are key factors in cytokinesis, which is
571 highly relevant in the ovary, because germ cells remain attached in nests through incomplete
572 cytokinesis until primordial follicle formation (Lei & Spradling, 2013; Rosario et al., 2019).
573 Higher transcript levels from these genes were observed in the *Cdkn1b*$^{-/-}$ ovary, which suggests
574 a developmentally early breakdown of germ cell nests, exacerbating the follicle activation
575 phenotype previously described (Rajareddy et al., 2007). Stringent experiments using inducible
576 *Cdkn1b* gene deletion or overexpression of alternate cell cycle inhibitors may reveal the
577 importance of cell cycle regulators in the developing ovary.
578
579 Questions remain as to what initial signal drives the precocious primordial follicle activation
580 in *Cdkn1b*$^{-/-}$ mutants. In this study, a non-cell autonomous effect of p27^{kip1} cannot be ruled out,
581 because p27^{kip1} functions in multiple tissues, including the pituitary (Moncho-Amor et al.,
582 2020). A conditional approach, such as using a *Cdkn1b*$^{fl/fl}$*;Nr5a1-Cre+* mouse line, would
583 provide a clear answer if primordial follicle activation is initiated in the pregranulosa cells and
584 dysregulated by the absence of p27^{kip1}. Regardless, it is clear from this body of work that further
585 interrogation is necessary to find the master regulators of follicle activation. The hierarchy of
586 the numerous factors that cause follicle activation are still being explored: as most previous
587 studies have used single gene-deletion models, which makes the construction of a gene
588 regulatory network challenging (Adhikari & Liu, 2009; Jagarlamudi, Reddy, Adhikari, & Liu,
589 2010; Uda et al., 2004). For p27^{kip1} in particular, potential upstream regulators that may initiate
590 primordial follicle activation by reducing p27^{kip1} expression include FOXL2, WNT4, SMAD2

591 and JAK2 (Granados-Aparici et al., 2019; Huang et al., 2018; Jakel, Weinl, & Hengst, 2011;
592 Maatouk, Mork, Chassot, Chaboissier, & Capel, 2013; Nicol et al., 2018). Additionally, given
593 that spatial and temporal control of activation is required to selectively activate follicles in
594 adulthood, epigenetic mechanisms and chromatin accessibility are two unexplored areas that
595 may provide the precise control necessary to activate multiple factors simultaneously. Future
596 experiments using ChIP-sequencing and ATAC-sequencing on pregranulosa cells during the
597 window of activation in the mouse will yield fascinating insights into how the genes
598 upregulated during primordial follicle activation are controlled.
599
600 This study describes a transcriptomic dataset to interrogate primordial follicle activation and
601 reveals key pathways that are controlling follicle activation in mice. This dataset provides a
602 platform for characterising a gene regulatory network that regulates primordial follicle
603 activation. The combination of single cell RNA-sequencing with a bulk RNA-sequencing
604 dataset in an infertility model reveals disruptions in the follicle activation process, but also
605 similar pathways that are consistently activated. This combination of transcriptomic analyses
606 adds a novel bioinformatic approach to this field to understand cell-specific changes in
607 infertility models. An increased understanding of the factors that act co-operatively to stimulate
608 or inhibit follicle activation will reveal key insights into how the ovarian reserve is established
609 and maintained into adulthood and how disruptions in this process lead to infertility conditions.
610

611 **Materials and Methods**

612 **Key resources table**

Reagent type (species) or resource	Reagent or resource	Source	Identifier
Antibody	Rabbit monoclonal anti-p27^{kip1}	Cell Signalling Technology	Cat#3686S; RRID: AB_2077850
Antibody	Goat polyclonal anti-FOXL2	Abcam	Cat#ab5096; RRID: AB_304750
Antibody	Rat monoclonal anti-GCNA1	Abcam	Cat#ab82527; RRID: AB_1659152
Antibody	Rabbit polyclonal anti-DDX4	Abcam	Cat#ab13840; RRID: AB_443012
Antibody	Mouse monoclonal anti-DDX4	Abcam	Cat#ab27591; RRID: AB_11139638
Chemical	Paraformaldehyde	Thermofisher Scientific	Cat# 50-980-488
Chemical	ProLong™ Gold Antifade Mountant	Thermofisher Scientific	Cat# P36930
Chemical	Target Retrieval Solution, Citrate pH 6.1	Agilent	Cat#S169984-2
Chemical	Triton X-100	Sigma Aldrich	Cat# T8787
Commercial assay or kit	Chromium Next GEM Single Cell 3′ Kit v3.1	10x Genomics	Cat#1000268
Commercial assay or kit	NEBnext Low Input RNA Library Prep Kit	New England Biolabs	Cat# E6420L
Commercial assay or kit	RNeasy Plus Micro Kit	Qiagen	Cat#74034
Commercial assay or kit	Superscript III	Thermofisher Scientific	Cat#18080093
Genetic reagent (*Mus musculus*)	Mouse: Wildtype (WT; C57BL/6J)	The Jackson Laboratory	Cat#JAX000664; RRID:IMSR_JAX:000664
Genetic reagent (*Mus musculus*)	Mouse: *Cdkn1b*$^{-/-}$ (B6.129S4-Cdkn1b^{tm1Mlf}/J)	The Jackson Laboratory	Cat#JAX002781 RRID:IMSR_JAX:002781
Oligonucleotide	PCR forward primer for mouse *Cdkn1b* WT allele GATGGACGCCAGACAAGC	The Jackson Laboratory	N/A
Oligonucleotide	PCR reverse primer for mouse *Cdkn1b* WT allele CTCCTGCCATTCGTATCTGC	The Jackson Laboratory	N/A
Oligonucleotide	PCR forward primer for mouse *Cdkn1b* mutant allele CTTGGGTGGAGAGGCTATTC	The Jackson Laboratory	N/A
Oligonucleotide	PCR reverse primer for mouse *Cdkn1b* mutant allele AGGTGAGATGACAGGAGATC	The Jackson Laboratory	N/A
Oligonucleotide	qPCR forward primer for mouse *Cdkn1b* gene CAGAAGAAAATGTTTCAGACGGT	This paper	N/A
Oligonucleotide	qPCR reverse primer for mouse *Cdkn1b* gene TCTTAATTCGGAGCTGTTTACGT	This paper	N/A

Oligonucleotide	qPCR forward primer for mouse *Gapdh* gene TTCACCACCATGGAGAAGGC	(Nicol et al., 2018)	N/A
Oligonucleotide	qPCR reverse primer for mouse *Gapdh* gene GGCATGGACTGTGGTCATGA	(Nicol et al., 2018)	N/A
Software and algorithm	Adobe Illustrator CC	Adobe Systems Inc.	https://www.adobe.com/illustrator
Software and algorithm	DAVID Functional Annotation	(Dennis et al., 2003)	https://david.ncifcrf.gov
Software and algorithm	Fiji	(Schindelin et al., 2012)	https://fiji.sc/
Software and algorithm	Prism	GraphPad Inc.	https://www.graphpad.com/scientific-software/prism/
Software and algorithm	R version 3.6.0	R Project for Statistical Computing	https://scicrunch.org/resolver/SCR_001905 RRID:SCR_001905
Software and algorithm	Seurat package 3	R toolkit for single-cell genomics	https://satijalab.org/seurat/articles/install.html
Software and algorithm	Zen software	Zeiss	https://www.zeiss.com/microscopy/int/products/microscope-software/zen.html

613
614 ***Mouse strains, husbandry and genotyping***
615
616 All experiments carried out on mice were approved under the UK Animal (scientific
617 procedures) Act 1986 (Project license 70/8560). Animals were kept in individually ventilated
618 cages (IVC) with access to food and water *ad libitum*. Mouse husbandry, ear biopsies, breeding
619 and vaginal plug (VP) checks were performed by the Biological Research Facility team of the
620 Francis Crick Institute. The VP day was considered as 0.5 day from time of conception (E0.5)
621 and the day of birth considered P0. Adult female animals were euthanised by cervical
622 dislocation and ovaries were collected immediately following euthanasia. Embryos were
623 euthanised by decapitation and gonads were collected with the adjacent mesonephros. For
624 RNA-sequencing, the mesonephros was removed from the ovary using dissecting forceps.
625 C57BL/6J mice (RRID: IMSR_JAX:000664) were supplied by the Biological Research
626 Facility at the Francis Crick Institute and used for the single-cell sequencing experiments.
627 *Cdkn1b*^{tm1Mlf} (*Cdkn1b*$^{-/-}$) (RRID: IMSR_JAX:002781) (Fero et al., 1996) were maintained on
628 a mixed background and mated heterozygously. Experiments involving *Cdkn1b*$^{-/-}$ mice were

629 performed using littermate animals as controls. Genotyping of adult *Cdkn1b*$^{-/-}$ mice was
630 performed by Transnetyx®. Embryos from heterozygous matings were PCR genotyped using
631 primers for the WT allele and mutant allele as previously published (Fero et al., 1996).

632 ***RNA sequencing sample preparation***
633
634 ***Tissue dissociation and single-cell library preparation***
635
636 Female ovaries at E18.5, PND1 and PND7 were harvested and the ovarian bursa and fat were
637 removed, leaving only ovary tissue. Ovarian tissue was dissociated as previously described
638 (Frost et al., 2021b). Briefly, two ovaries from each timepoint were incubated separately in a
639 solution of papain (10108014001, Sigma-Aldrich, 1mg/ml in HBSS) for 15min at 37°C in
640 presence of 10µg/ml of both DNAse I (10104159001, Thermo Scientific) and Rock Inhibitor
641 (M1817, Abmole Bioscience) (Moncho-Amor et al., 2020). The papain solution containing the
642 cells was diluted with 1x HBSS containing 0.0025% BSA and 0.01% 1M Tris pH8.0. The cell
643 solution was then mechanically dissociated using a 10mL glass serological pipette, and passed
644 through a 20µM pre-separation filter (130-101-812, Miltenyi Biotec). Cells were counted using
645 the EVE Cell counter (NanoEnTek, Cambridge Bioscience), stained with trypan blue to
646 measure viability. Approximately 10,000 live cells were used to generate libraries using the
647 Chromium Single Cell 3' kit v3.1, according to the manufacturer's instructions (10x Genomics,
648 1000268) (G. X. Y. Zheng et al., 2017). The libraries were quantified using the TapeStation
649 (Agilent) to confirm quantity and purity, before sequencing on an Hiseq4000 (Illumina).
650 Sequencing outputs were processed using the 10x CellRanger (version.3.0.2) to generate
651 single-cell count data for each time point, using a mouse reference index provided by 10x
652 Genomics (refdata-cellranger-mm10v3.0). All subsequent analyses were performed in R
653 v.3.6.0 using the Seurat (v3) package (Stuart et al., 2019).

654 ***Tissue dissociation and bulk library preparation***

655 Ovaries from *Cdkn1b*$^{+/+}$ and *Cdkn1b*$^{-/-}$ mice were collected at E13.5, E18.5 and PND8 and the
656 ovarian bursa and fat were removed. A single ovary was dissociated as described above. Cells
657 were counted using the EVE Cell counter (NanoEnTek, Cambridge Bioscience) and stained
658 with trypan blue to measure viability. RNA was extracted from the dissociated cells using the
659 RNeasy Plus Micro kit (Qiagen) according to the manufacturer's protocol. RNA quality was
660 assessed the TapeStation (Agilent) to confirm quantity and purity. For E13.5 and E18.5 ovaries,
661 libraries were constructed using the NEBNext Low Input RNA Library Prep Kit (E6420L, New
662 England Biolabs) according to the manufacturer's instructions. For PND8 ovaries, libraries
663 were constructed using the polyA KAPA mRNA Hyper Prep Kit (Roche) according to the
664 manufacturer's instructions. All libraries were quantified using the TapeStation (Agilent) and
665 sequenced on an Hiseq400 (Illumina) to achieve an average of 25 million reads per sample.

666 ***scRNAseq Bioinformatics analysis***
667
668 Raw single-cell sequencing data from all three timepoints were processed using the 10x
669 CellRanger pipeline (10x Genomics). Count tables were loaded into R and analysed with the
670 Seurat package (Stuart et al., 2019). All cells with fewer than 1500 RNA features were
671 removed. Principal component analysis was performed on variable genes and 20 principal
672 components were used to generate the UMAP analysis. The UMAP clustering plot was
673 performed using the default method from the Seurat package. Differential Gene Expression
674 (DGE) analysis was performed as previously described (Ahlmann-Eltze & Huber, 2020).
675 DAVID Gene Ontology (GO) functional annotation tools were utilised to explore the biological
676 processes associated with the genes differentially expressed in specific clusters (Dennis et al.,
677 2003).

678 ***Bulk RNA sequencing analysis***
679
680 Twelve libraries were sequenced on an Illumina HiSeq 4000 machine. The 'Trim Galore!'
681 utility version 0.4.2 was used to remove sequencing adaptors and to trim individual reads with

682 the q-parameter set to 20 (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/).
683 The sequencing reads were then aligned to the mouse genome and transcriptome (Ensembl
684 GRCm38release-89) using RSEM version 1.3.0 (B. Li & Dewey, 2011) in conjunction with
685 the STAR aligner version 2.5.2 (Dobin et al., 2013). The sequencing quality of individual
686 samples was assessed using FASTQC version 0.11.5
687 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and RNA-SeQC version 1.1.8
688 (DeLuca et al., 2012). Differential gene expression was determined using the R-bioconductor
689 package DESeq2 version 1.14.1 (Development & Team, 2008; Love, Huber, & Anders, 2014).
690 Gene set enrichment analysis (GSEA) was conducted as described (Subramanian et al., 2005).

691 ***Immunofluorescence and in situ hybridisation***
692
693 Ovaries for immunofluorescence were fixed in chilled 4 % paraformaldehyde at 4 °C for 1 to
694 2 hrs. Following fixation, ovaries were washed in PBS 3 times for 1 hr each, before immersing
695 in 30 % sucrose in PBS overnight to cryoprotect the sample. The ovaries were then embedded
696 in OCT mounting medium, frozen on an ethanol/dry ice slurry and stored at -80 °C. Samples
697 were cryosectioned at 10 µm and sections were placed on Superfrost Plus glass slides and air
698 dried for 5 minutes, before storing at -80 °C. Before starting immunofluorescence, slides were
699 air dried for 5 minutes and then washed 3 times for 5 mins each in PBS. Antigen retrieval was
700 performed by immersing slides in 10 % target retrieval solution pH 6.1 (S169984-2, Agilent)
701 with distilled water. Antigen retrieval was done for 30 mins at 65 °C. Slides were washed 3
702 times in PBS for 5 mins and then incubated in blocking solution (10 % donkey serum in 0.1 %
703 triton X-100 PBS) for 2 hr at room temperature. Primary antibodies were added to the slide,
704 diluted in 10% blocking solution, and left overnight at 4 °C. Primary antibodies were used at
705 the following dilutions: rabbit anti-p27^{kip1} (D69C12, Cell Signalling) at 1:200, goat anti-
706 FOXL2 (ab5096, Abcam) at 1:100, goat anti-SOX9 (AF3075, R&D Systems) at 1:200, rat anti-
707 GCNA1 (ab82527, Abcam) at 1:200, rabbit anti-DDX4 (ab13840, Abcam) at 1:200 and mouse

708 anti-DDX4 (ab27591, Abcam) at 1:200. The next day, the slides were washed 3 times in 0.1 %
709 triton X-100 PBS for 5 minutes. Secondary antibodies were diluted in blocking solution (10 %
710 donkey serum in 0.1 % triton X-100 PBS) and slides were incubated with the appropriate
711 secondary antibodies for 1 hr at room temperature. The following secondary antibodies were
712 used at a 1:200 dilution: Donkey anti-rabbit 488 (Invitrogen, A21206), Donkey anti-goat 647
713 (Invitrogen, A21447), Donkey anti-mouse 647 (Invitrogen, A-31571) and Donkey anti-rat 647
714 (Abcam, ab150155). The slides were then washed in 0.1 % triton X-100 PBS for 5 minutes 3
715 times, then the slides were counter-stained with 4'-6-diamino-2-phenylindole (DAPI). Finally,
716 sections were washed once in distilled water, air dried and coverslip mounted with Prolong
717 Gold Antifade Mountant (P36930, Thermo Scientific) and Aqua-Poly/Mount (Polysciences,
718 Inc., Warrington, PA, USA). Images were taken using a Leica SPE confocal microscope and
719 processed using Fiji.

720 ***Histology staining***
721
722 Ovaries for hematoxylin and eosin (H&E) staining were fixed in chilled 4 % paraformaldehyde
723 at 4 °C for 1 to 2 hrs. Following fixation, ovaries were washed in PBS 3 times for 1 hr each,
724 before immersing in 30 % sucrose in PBS overnight to cryoprotect the sample. The ovaries
725 were then embedded in OCT mounting medium, frozen on an ethanol/dry ice slurry and stored
726 at -80 °C. Samples were cryosectioned at 10 µm and sections were placed on Superfrost Plus
727 glass slides and air dried for 5 minutes. H&E staining was performed by the Francis Crick
728 Institute Experimental Histopathology facility.

729 ***RNA extraction and qPCR***
730
731 Total mRNA was extracted from individual ovaries using the RNeasy Plus Micro kit (74034,
732 Qiagen) according to the manufacturer's protocol. Total mRNA was quantified using the
733 NanoDrop™ 2000 Spectrophotometer (Thermo Fisher Scientific). After quantification, 50 ng
734 of RNA was reverse transcribed into cDNA using the Superscript III cDNA synthesis kit

735 (Thermo Fisher Scientific) as per the manufacturer's instructions. Quantitative PCR was
736 performed on cDNA, with primer sequences listed in the key resources table. Reactions were
737 performed in triplicate using GoTaq® qPCR Master Mix (Promega) and 100 nM each of
738 forward and reverse primers and analysed on the LightCycler® 96 Instrument (Roche).
739 Relative mRNA levels were determined by calculating $2^{-\Delta\Delta Ct}$ values relative to the
740 housekeeping gene *Gapdh*. Data is shown as mean ± SEM from four biological replicates.

741 ***Statistical analyses***
742
743 Statistical analyses were performed using Prism v9.0 (GraphPad Software, USA). Student's
744 two-sided unpaired *t-tests* were used when comparing two groups of normally distributed
745 values. Mann Whitney tests were used when comparing two groups of non-parametric data.
746 For multiple comparisons, analysis of variance (ANOVA) was performed. All results are
747 represented as mean ± standard error of the mean (SEM) and standard significance levels were
748 used: *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001.

749 **Data availability**

750 The data supporting the findings of this study are available in the main text and the
751 supplementary materials. All sequencing data will be made publicly available after the
752 publication of this article.

753 **Acknowledgements**

754 We would like to thank members of the Lovell-Badge and Sutherland laboratories for their
755 comments on the manuscript. We thank the Advanced Sequencing Facility at the Francis Crick
756 Institute for their help in performing the sequencing studies in this manuscript. This work was
757 supported by the Francis Crick Institute, which receives its core funding from Cancer Research
758 UK (FC001107), the UK Medical Research Council (FC001107), and the Wellcome Trust
759 (FC001107). We acknowledge the generous financial assistance to R.L.B. by the UK
760 Biotechnology and Biological Sciences Research Council (BB/ N018680/1). We also

761 gratefully acknowledge the financial assistance to J.M.S. and M.A.B. by the Australian
762 National Health and Medical Research Council (G1600095) and the Hunter Medical Research
763 Institute (G1501433 and G1801335). E.R.F. is the recipient of an Australian Government
764 Research Training Program Scholarship.

765 **Author contributions**

766 E.R.F and G.T conceived the study and wrote the manuscript. J.M.S, C.G, M.A.B and R.L.B
767 provided expertise and feedback on the study design. All the authors critically revised the
768 manuscript. All experiments were performed by E.R.F except for the following: G.T and E.R.F
769 jointly prepared samples for single-cell RNA sequencing. S.B performed both RNA-seq and
770 single-cell RNA-seq analysis. J.M.S and R.L.B provided resources for this work.

771 **Declaration of interests**

772 The authors declare no competing interests.

773 **References**

774 Adhikari, D., & Liu, K. (2009). Molecular mechanisms underlying the activation of mammalian
775 primordial follicles. *Endocr Rev, 30*(5), 438-464. doi:10.1210/er.2008-0048
776 Ahlmann-Eltze, C., & Huber, W. (2020). glmGamPoi: fitting Gamma-Poisson generalized linear
777 models on single cell count data. *Bioinformatics, 36*(24), 5701-5702.
778 doi:10.1093/bioinformatics/btaa1009
779 Bagheri-Fam, S., Chen, H., Wilson, S., Ayers, K., Hughes, J., Sloan-Bena, F., . . . Wilhelm, D.
780 (2020). The gene encoding the ketogenic enzyme HMGCS2 displays a unique expression
781 during gonad development in mice. *PLOS ONE, 15*(1), e0227411.
782 doi:10.1371/journal.pone.0227411
783 Cannon, J. D., Cherian-Shaw, M., Lovekamp-Swan, T., & Chaffin, C. L. (2007). Granulosa cell
784 expression of G1/S phase cyclins and cyclin-dependent kinases in PMSG-induced follicle
785 growth. *Mol Cell Endocrinol, 264*(1-2), 6-15. doi:10.1016/j.mce.2006.09.015
786 Chen, Y., Yang, W., Shi, X., Zhang, C., Song, G., & Huang, D. (2020). The Factors and Pathways
787 Regulating the Activation of Mammalian Primordial Follicles in vivo. *Front Cell Dev Biol,*
788 *8*(1018). doi:10.3389/fcell.2020.575706
789 Chen, Y.-Y., Russo, D. D., Drake, R. S., Duncan, F. E., Shalek, A. K., Goods, B. A., & Woodruff, T.
790 K. (2022). Single-cell transcriptomics of staged oocytes and somatic cells reveal novel
791 regulators of follicle activation. *Reproduction, 164*(2), 55-70. doi:10.1530/rep-22-0053
792 Cooke, H. J., Lee, M., Kerr, S., & Ruggiu, M. (1996). A murine homologue of the human DAZ gene
793 is autosomal and expressed only in male and female gonads. *Hum Mol Genet, 5*(4), 513-516.
794 doi:10.1093/hmg/5.4.513
795 Cordeiro, M. H., Kim, S. Y., Ebbert, K., Duncan, F. E., Ramalho-Santos, J., & Woodruff, T. K.
796 (2015). Geography of follicle formation in the embryonic mouse ovary impacts activation
797 pattern during the first wave of folliculogenesis. *Biology of Reproduction, 93*(4), 88.
798 doi:10.1095/biolreprod.115.131227

DeLuca, D. S., Levin, J. Z., Sivachenko, A., Fennell, T., Nazaire, M.-D., Williams, C., . . . Getz, G. (2012). RNA-SeQC: RNA-seq metrics for quality control and process optimization. *Bioinformatics, 28*(11), 1530-1532. doi:10.1093/bioinformatics/bts196

Dennis, G., Sherman, B. T., Hosack, D. A., Yang, J., Gao, W., Lane, H. C., & Lempicki, R. A. (2003). DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biology, 4*(5), P3. doi:10.1186/gb-2003-4-5-p3

Development, R. C., & Team. (2008). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., . . . Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics, 29*(1), 15-21. doi:10.1093/bioinformatics/bts635

Fan, X., Bialecka, M., Moustakas, I., Lam, E., Torrens-Juaneda, V., Borggreven, N. V., . . . Chuva de Sousa Lopes, S. M. (2019). Single-cell reconstruction of follicular remodeling in the human adult ovary. *Nat Commun, 10*(1), 3164. doi:10.1038/s41467-019-11036-9

Fero, M. L., Rivkin, M., Tasch, M., Porter, P., Carow, C. E., Firpo, E., . . . Roberts, J. M. (1996). A syndrome of multiorgan hyperplasia with features of gigantism, tumorigenesis, and female sterility in p27(Kip1)-deficient mice. *Cell, 85*(5), 733-744. Retrieved from https://ac.els-cdn.com/S0092867400812398/1-s2.0-S0092867400812398-main.pdf?_tid=81ed8a97-9c51-4fe7-82d6-725b355e7e1c&acdnat=1548794809_c846c788db4da96843436c6939aab1ba

Francois, C. M., Petit, F., Giton, F., Gougeon, A., Ravel, C., Magre, S., . . . Guigon, C. J. (2017). A novel action of follicle-stimulating hormone in the ovary promotes estradiol production without inducing excessive follicular growth before puberty. *Sci Rep, 7*, 46222. doi:10.1038/srep46222

Frost, E. R., Ford, E. A., Taylor, G., Boeing, S., Beckett, E. L., Roman, S. D., . . . Sutherland, J. M. (2021a). Two alternative methods for the retrieval of somatic cell populations from the mouse ovary. *Molecular Human Reproduction, 27*(6). doi:10.1093/molehr/gaab033

Frost, E. R., Ford, E. A., Taylor, G., Boeing, S., Beckett, E. L., Roman, S. D., . . . Sutherland, J. M. (2021b). Two alternative methods for the retrieval of somatic cell populations from the mouse ovary. *Molecular Human Reproduction*. doi:10.1093/molehr/gaab033

Fukuda, K., Muraoka, M., Kato, Y., & Saga, Y. (2021). Decoding the transcriptome of pre-granulosa cells during the formation of primordial follicles in the mouse. *Biology of Reproduction*. doi:10.1093/biolre/ioab065

Gaytan, F., Morales, C., Roa, J., & Tena-Sempere, M. (2018). Changes in keratin 8/18 expression in human granulosa cell lineage are associated to cell death/survival events: potential implications for the maintenance of the ovarian reserve. *Hum Reprod, 33*(4), 680-689. doi:10.1093/humrep/dey010

Granados-Aparici, S., Hardy, K., Franks, S., Sharum, I. B., Waite, S. L., & Fenwick, M. A. (2019). SMAD3 directly regulates cell cycle genes to maintain arrest in granulosa cells of mouse primordial follicles. *Sci Rep, 9*(1), 6513. doi:10.1038/s41598-019-42878-4

Greenberg, Z. J., Monlish, D. A., Bartnett, R. L., Yang, Y., Shen, G., Li, W., . . . Schuettpelz, L. G. (2020). The Tetraspanin CD53 Regulates Early B Cell Development by Promoting IL-7R Signaling. *The Journal of Immunology, 204*(1), 58-67. doi:10.4049/jimmunol.1900539

Hardy, K., Mora, J. M., Dunlop, C., Carzaniga, R., Franks, S., & Fenwick, M. A. (2018). Nuclear exclusion of SMAD2/3 in granulosa cells is associated with primordial follicle activation in the mouse ovary. *Journal of Cell Science, 131*(17). doi:10.1242/jcs.218123

Huang, K., Wang, Y., Zhang, T., He, M., Sun, G., Wen, J., . . . Wang, C. (2018). JAK signaling regulates germline cyst breakdown and primordial follicle formation in mice. *Biology Open, 7*(1), bio029470. doi:10.1242/bio.029470

Hutt, K. J., McLaughlin, E. A., & Holland, M. K. (2006). KIT/KIT Ligand in Mammalian Oogenesis and Folliculogenesis: Roles in Rabbit and Murine Ovarian Follicle Activation and Oocyte Growth1. *Biology of Reproduction, 75*(3), 421-433. doi:10.1095/biolreprod.106.051516

Jagarlamudi, K., Reddy, P., Adhikari, D., & Liu, K. (2010). Genetically modified mouse models for premature ovarian failure (POF). *Mol Cell Endocrinol, 315*(1-2), 1-10. doi:10.1016/j.mce.2009.07.016

Jakel, H., Weinl, C., & Hengst, L. (2011). Phosphorylation of p27Kip1 by JAK2 directly links cytokine receptor signaling to cell cycle control. *Oncogene, 30*(32), 3502-3512. doi:10.1038/onc.2011.68

John, G. B., Gallardo, T. D., Shirley, L. J., & Castrillon, D. H. (2008). Foxo3 is a PI3K-dependent molecular switch controlling the initiation of oocyte growth. *Dev Biol, 321*(1), 197-204. doi:10.1016/j.ydbio.2008.06.017

Jones, R. L., & Pepling, M. E. (2013). KIT signaling regulates primordial follicle formation in the neonatal mouse ovary. *Dev Biol, 382*(1), 186-197. doi:10.1016/j.ydbio.2013.06.030

Kezele, P. R., Ague, J. M., Nilsson, E., & Skinner, M. K. (2005). Alterations in the ovarian transcriptome during primordial follicle assembly and development. *Biology of Reproduction, 72*(1), 241-255. doi:10.1095/biolreprod.104.032060

Landsman, L., Bar-On, L., Zernecke, A., Kim, K.-W., Krauthgamer, R., Shagdarsuren, E., . . . Jung, S. (2009). CX3CR1 is required for monocyte homeostasis and atherogenesis by promoting cell survival. *Blood, 113*(4), 963-972. doi:10.1182/blood-2008-07-170787

Lei, L., & Spradling, A. C. (2013). Mouse primordial germ cells produce cysts that partially fragment prior to meiosis. *Development, 140*(10), 2075-2081. doi:10.1242/dev.093864

Li, B., & Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics, 12*(1), 323. doi:10.1186/1471-2105-12-323

Li, Q., Yan, Z., Kuang, Y., Zhou, X., Jin, L., He, L., . . . Wang, L. (2016). Genetic variations in the 3'-untranslated region of SLC18A2 are associated with serum FSH concentration in polycystic ovary syndrome patients and regulate gene expression in vitro. *Hum Reprod, 31*(9), 2150-2157. doi:10.1093/humrep/dew162

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology, 15*(12), 550. doi:10.1186/s13059-014-0550-8

Ma, S. C., Li, Q., Peng, J. Y., Zhouwen, J. L., Diao, J. F., Niu, J. X., . . . Jiang, W. G. (2017). Claudin-5 regulates blood-brain barrier permeability by modifying brain microvascular endothelial cell proliferation, migration, and adhesion to prevent lung cancer metastasis. *CNS Neurosci Ther, 23*(12), 947-960. doi:10.1111/cns.12764

Maatouk, D. M., Mork, L., Chassot, A. A., Chaboissier, M. C., & Capel, B. (2013). Disruption of mitotic arrest precedes precocious differentiation and transdifferentiation of pregranulosa cells in the perinatal Wnt4 mutant ovary. *Dev Biol, 383*(2), 295-306. doi:10.1016/j.ydbio.2013.08.026

McLaughlin, E. A., & McIver, S. C. (2009). Awakening the oocyte: controlling primordial follicle development. *Reproduction, 137*(1), 1-11. doi:10.1530/rep-08-0118

Meinsohn, M. C., Hughes, C. H. K., Estienne, A., Saatcioglu, H. D., Pépin, D., Duggavathi, R., & Murphy, B. D. (2021). A role for orphan nuclear receptor liver receptor homolog-1 (LRH-1, NR5A2) in primordial follicle activation. *Sci Rep, 11*(1), 1079. doi:10.1038/s41598-020-80178-4

Miller, I., Min, M., Yang, C., Tian, C., Gookin, S., Carter, D., & Spencer, S. L. (2018). Ki67 is a Graded Rather than a Binary Marker of Proliferation versus Quiescence. *Cell Rep, 24*(5), 1105-1112.e1105. doi:10.1016/j.celrep.2018.06.110

Moncho-Amor, V., Chakravarty, P., Galichet, C., Matheu, A., Lovell-Badge, R., & Rizzoti, K. (2020). SOX2 is required independently in both stem and differentiated cells for pituitary tumorigenesis in p27 null mice. *BioRxiv*, 2020.2006.2009.142638. doi:10.1101/2020.06.09.142638

Monniaux, D., Clement, F., Dalbies-Tran, R., Estienne, A., Fabre, S., Mansanet, C., & Monget, P. (2014). The ovarian reserve of primordial follicles and the dynamic reserve of antral growing follicles: what is the link? *Biology of Reproduction, 90*(4), 85. doi:10.1095/biolreprod.113.117077

Mork, L., Maatouk, D. M., McMahon, J. A., Guo, J. J., Zhang, P., McMahon, A. P., & Capel, B. (2012). Temporal differences in granulosa cell specification in the ovary reflect distinct follicle fates in mice. *Biology of Reproduction, 86*(2), 37. doi:10.1095/biolreprod.111.095208

Nagamatsu, G., Shimamoto, S., Hamazaki, N., Nishimura, Y., & Hayashi, K. (2019). Mechanical stress accompanied with nuclear rotation is involved in the dormant state of mouse oocytes. *Sci Adv, 5*(6), eaav9960. doi:10.1126/sciadv.aav9960

Nicol, B., Grimm, S. A., Chalmel, F., Lecluze, E., Pannetier, M., Pailhoux, E., . . . Yao, H. H. (2019). RUNX1 maintains the identity of the fetal ovary through an interplay with FOXL2. *Nat Commun, 10*(1), 5116. doi:10.1038/s41467-019-13060-1

Nicol, B., Grimm, S. A., Gruzdev, A., Scott, G. J., Ray, M. K., & Yao, H. H. (2018). Genome-wide identification of FOXL2 binding and characterization of FOXL2 feminizing action in the fetal gonads. *Hum Mol Genet, 27*(24), 4273-4287. doi:10.1093/hmg/ddy312

Niu, W., & Spradling, A. C. (2020). Two distinct pathways of pregranulosa cell differentiation support follicle formation in the mouse ovary. *Proceedings of the National Academy of Sciences, 117*(33), 20015-20026. doi:10.1073/pnas.2005570117

O'Shaughnessy, P. J., McLelland, D., & McBride, M. W. (1997). Regulation of luteinizing hormone-receptor and follicle-stimulating hormone-receptor messenger ribonucleic acid levels during development in the neonatal mouse ovary. *Biology of Reproduction, 57*(3), 602-608. doi:10.1095/biolreprod57.3.602

Ottolenghi, C., Omari, S., Garcia-Ortiz, J. E., Uda, M., Crisponi, L., Forabosco, A., . . . Schlessinger, D. (2005). Foxl2 is required for commitment to ovary differentiation. *Hum Mol Genet, 14*(14), 2053-2062. doi:10.1093/hmg/ddi210

Pepling, M. E., & Spradling, A. C. (1998). Female mouse germ cells form synchronously dividing cysts. *Development, 125*(17), 3323-3328.

Pepling, M. E., & Spradling, A. C. (2001). Mouse ovarian germ cell cysts undergo programmed breakdown to form primordial follicles. *Dev Biol, 234*(2), 339-351. doi:10.1006/dbio.2001.0269

Perez-Sanz, J., Arluzea, J., Matorras, R., Gonzalez-Santiago, N., Bilbao, J., Yeh, N., . . . de la Hoz, C. (2013). Increased number of multi-oocyte follicles (MOFs) in juvenile p27Kip1 mutant mice: potential role of granulosa cells. *Hum Reprod, 28*(4), 1023-1030. doi:10.1093/humrep/des436

Pisarska, M. D., Barlow, G., & Kuo, F. T. (2011). Minireview: roles of the forkhead transcription factor FOXL2 in granulosa cell biology and pathology. *Endocrinology, 152*(4), 1199-1208. doi:10.1210/en.2010-1041

Rafiee, M. R., Girardot, C., Sigismondo, G., & Krijgsveld, J. (2016). Expanding the Circuitry of Pluripotency by Selective Isolation of Chromatin-Associated Proteins. *Mol Cell, 64*(3), 624-635. doi:10.1016/j.molcel.2016.09.019

Rajareddy, S., Reddy, P., Du, C., Liu, L., Jagarlamudi, K., Tang, W., . . . Liu, K. (2007). p27kip1 (cyclin-dependent kinase inhibitor 1B) controls ovarian development by suppressing follicle endowment and activation and promoting follicle atresia in mice. *Mol Endocrinol, 21*(9), 2189-2202. doi:10.1210/me.2007-0172

Rastetter, R. H., Bernard, P., Palmer, J. S., Chassot, A. A., Chen, H., Western, P. S., . . . Wilhelm, D. (2014). Marker genes identify three somatic cell types in the fetal mouse ovary. *Dev Biol, 394*(2), 242-252. doi:10.1016/j.ydbio.2014.08.013

Reynolds, N., Collier, B., Bingham, V., Gray, N. K., & Cooke, H. J. (2007). Translation of the synaptonemal complex component Sycp3 is enhanced in vivo by the germ cell specific regulator Dazl. *Rna, 13*(7), 974-981. doi:10.1261/rna.465507

Robker, R. L., & Richards, J. S. (1998). Hormonal Control of the Cell Cycle in Ovarian Cells: Proliferation Versus Differentiation. *Biology of Reproduction, 59*(3), 476-482. doi:10.1095/biolreprod59.3.476

Robker, R. L., & Richards, J. S. (1998). Hormone-induced proliferation and differentiation of granulosa cells: a coordinated balance of the cell cycle regulators cyclin D2 and p27Kip1. *Mol Endocrinol, 12*(7), 924-940. doi:10.1210/mend.12.7.0138

Rosario, R., Crichton, J. H., Stewart, H. L., Childs, A. J., Adams, I. R., & Anderson, R. A. (2019). Dazl determines primordial follicle formation through the translational regulation of Tex14. *FASEB Journal, 33*(12), 14221-14233. doi:10.1096/fj.201901247R

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., . . . Cardona, A. (2012). Fiji: an open-source platform for biological-image analysis. *Nat Methods, 9*(7), 676-682. doi:10.1038/nmeth.2019

Sobinoff, A. P., Sutherland, J. M., & McLaughlin, E. A. (2013). Intracellular signalling during female gametogenesis. *Mol Hum Reprod, 19*(5), 265-278. doi:10.1093/molehr/gas065

Stevant, I., Kuhne, F., Greenfield, A., Chaboissier, M. C., Dermitzakis, E. T., & Nef, S. (2019). Dissecting Cell Lineage Specification and Sex Fate Determination in Gonadal Somatic Cells Using Single-Cell Transcriptomics. *Cell Rep, 26*(12), 3272-3283.e3273. doi:10.1016/j.celrep.2019.02.069

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W. M., 3rd, . . . Satija, R. (2019). Comprehensive Integration of Single-Cell Data. *Cell, 177*(7), 1888-1902.e1821. doi:10.1016/j.cell.2019.05.031

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., . . . Mesirov, J. P. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences, 102*(43), 15545-15550. doi:10.1073/pnas.0506580102

Suzuki, H., Kanai-Azuma, M., & Kanai, Y. (2015). From Sex Determination to Initial Folliculogenesis in Mammalian Ovaries: Morphogenetic Waves along the Anteroposterior and Dorsoventral Axes. *Sex Dev, 9*(4), 190-204. doi:10.1159/000440689

Tingen, C., Kim, A., & Woodruff, T. K. (2009). The primordial pool of follicles and nest breakdown in mammalian ovaries. *Mol Hum Reprod, 15*(12), 795-803. doi:10.1093/molehr/gap073

Uda, M., Ottolenghi, C., Crisponi, L., Garcia, J. E., Deiana, M., Kimber, W., . . . Pilia, G. (2004). Foxl2 disruption causes mouse ovarian failure by pervasive blockage of follicle development. *Hum Mol Genet, 13*(11), 1171-1181. doi:10.1093/hmg/ddh124

Wang, J.-J., Ge, W., Zhai, Q.-Y., Liu, J.-C., Sun, X.-W., Liu, W.-X., . . . Shen, W. (2020). Single-cell transcriptome landscape of ovarian cells during primordial follicle assembly in mice. *PLOS Biology, 18*(12), e3001025. doi:10.1371/journal.pbio.3001025

Wang, Z., Niu, W., Wang, Y., Teng, Z., Wen, J., Xia, G., & Wang, C. (2015). Follistatin288 Regulates Germ Cell Cyst Breakdown and Primordial Follicle Assembly in the Mouse Ovary. *PLOS ONE, 10*(6), e0129643. doi:10.1371/journal.pone.0129643

Wear, H. M., McPike, M. J., & Watanabe, K. H. (2016). From primordial germ cells to primordial follicles: a review and visual representation of early ovarian development in mice. *J Ovarian Res, 9*(1), 36. doi:10.1186/s13048-016-0246-7

Xu, B., Hua, J., Zhang, Y., Jiang, X., Zhang, H., Ma, T., . . . Shi, Q. (2011). Proliferating Cell Nuclear Antigen (PCNA) Regulates Primordial Follicle Assembly by Promoting Apoptosis of Oocytes in Fetal and Neonatal Mouse Ovaries. *PLOS ONE, 6*(1), e16046. doi:10.1371/journal.pone.0016046

Yang, C.-R., Rajkovic, G., Daldello, E. M., Luong, X. G., Chen, J., & Conti, M. (2020). The RNA-binding protein DAZL functions as repressor and activator of mRNA translation during oocyte maturation. *Nature Communications, 11*(1), 1399. doi:10.1038/s41467-020-15209-9

Zhang, H., Risal, S., Gorre, N., Busayavalasa, K., Li, X., Shen, Y., . . . Liu, K. (2014). Somatic cells initiate primordial follicle activation and govern the development of dormant oocytes in mice. *Curr Biol, 24*(21), 2501-2508. doi:10.1016/j.cub.2014.09.023

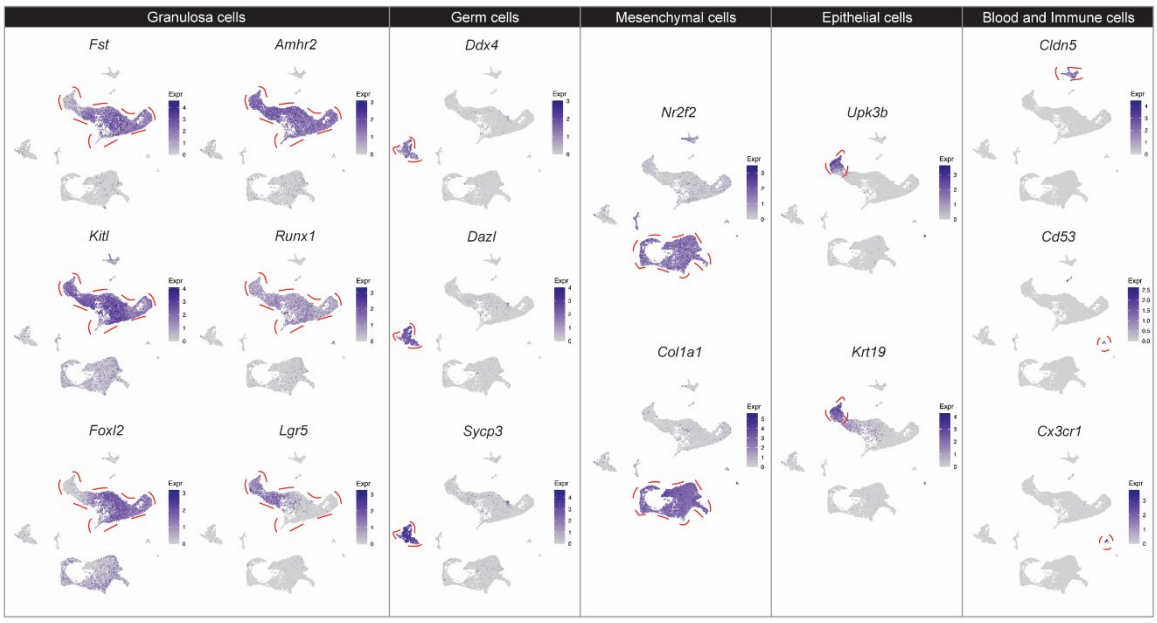
Zheng, G. X. Y., Terry, J. M., Belgrader, P., Ryvkin, P., Bent, Z. W., Wilson, R., . . . Bielas, J. H. (2017). Massively parallel digital transcriptional profiling of single cells. *Nature Communications, 8*(1), 14049. doi:10.1038/ncomms14049

Zheng, W., Zhang, H., Gorre, N., Risal, S., Shen, Y., & Liu, K. (2014). Two classes of ovarian primordial follicles exhibit distinct developmental dynamics and physiological functions. *Hum Mol Genet, 23*(4), 920-928. doi:10.1093/hmg/ddt486

Zheng, W., Zhang, H., & Liu, K. (2014). The two classes of primordial follicles in the mouse ovary: their development, physiological functions and implications for future research. *Mol Hum Reprod, 20*(4), 286-292. doi:10.1093/molehr/gau007

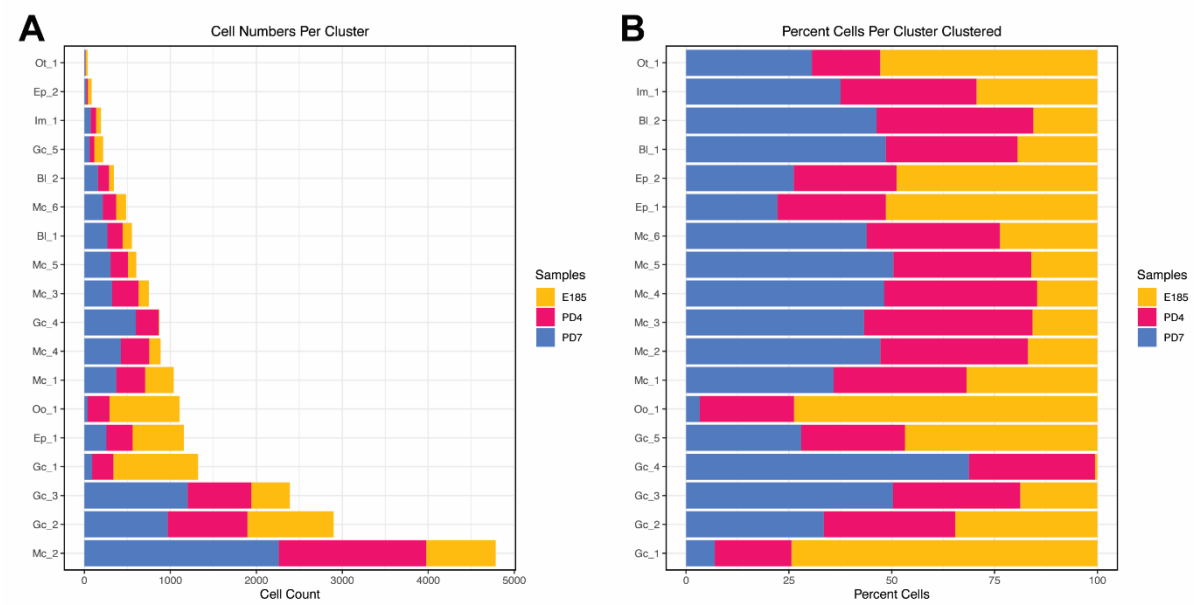
Zhihan, T., Xinyi, M., Qingying, L., Rufei, G., Yan, Z., Xuemei, C., . . . Junlin, H. (2019). Autophagy participates in cyst breakdown and primordial folliculogenesis by reducing reactive oxygen species levels in perinatal mouse ovaries. *J Cell Physiol, 234*(5), 6125-6135. doi:10.1002/jcp.27367

1016 **Supplemental figures**



1017
1018 **Supplemental Figure 1. Identification of cell types within the mouse ovary**

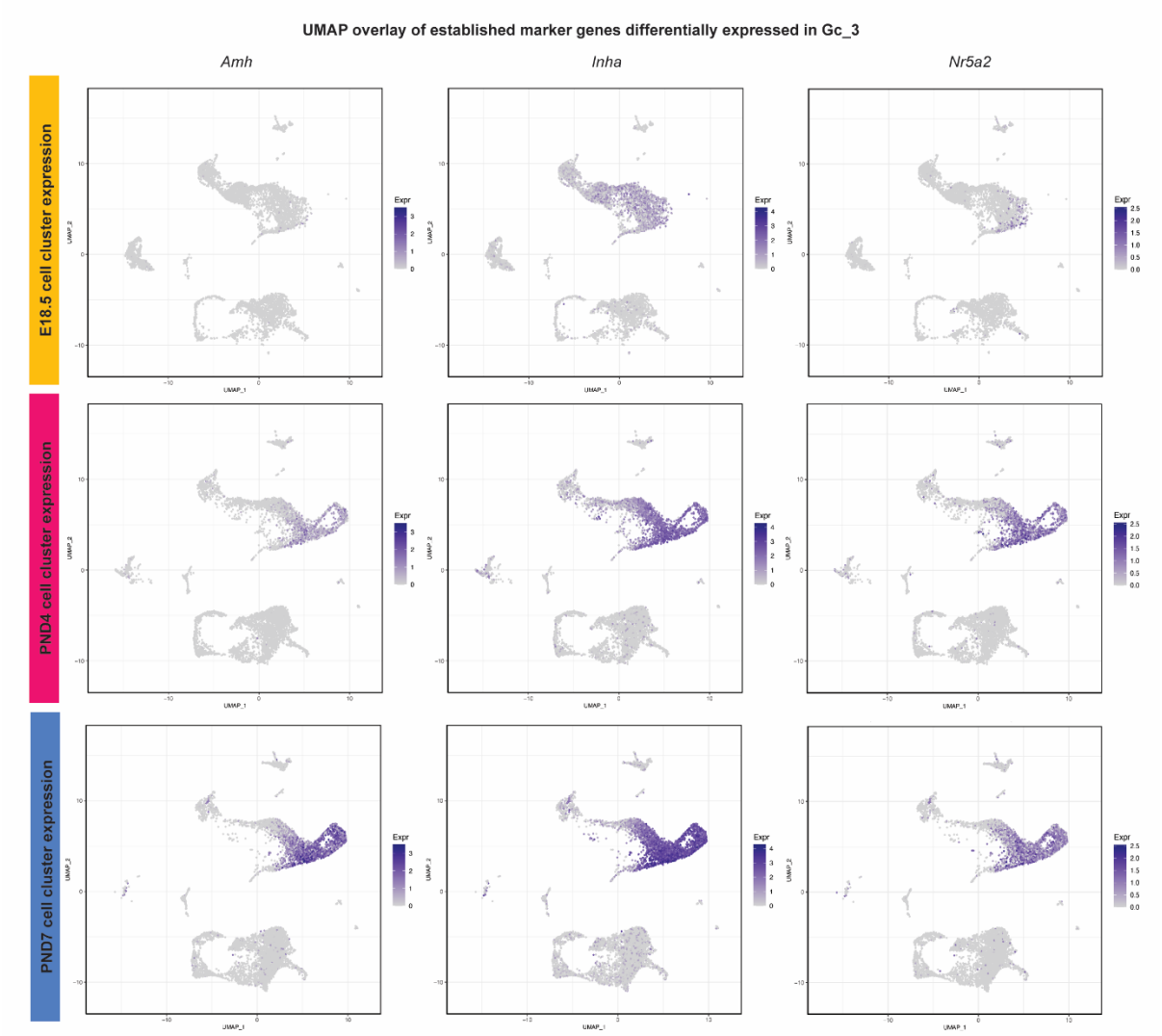
1019 The UMAP cluster maps show the expression of genes used to mark granulosa cells, mesenchymal
1020 cells, germ cells, epithelial cells and blood and immune cells. Dots are coloured according to their level
1021 of gene expression, with high levels of expression shown in purple. Red dashed lines show the clusters
1022 classified as each cell type.



1023
1024 **Supplemental Figure 2. Cell contributions to each cluster**

1025 (A) Barchart showing the cell count in each cluster and the number of cells allocated to each
1026 cluster from each timepoint (E18.5, PND4 and PND7). Cells from the E18.5 timepoint are
1027 shown in yellow, cells from the PND4 timepoint are shown in pink and cells from the PND7

timepoint are shown in blue. (B) Barchart showing the percentage of cells in each cluster that arise from each timepoint (E18.5, PND4 and PND7). Cells from the E18.5 timepoint are shown in yellow, cells from the PND4 timepoint are shown in pink and cells from the PND7 timepoint are shown in blue.



Supplemental Figure 3. Expression of key genes upregulated during primordial follicle activation The UMAP cluster maps show the expression of differentially expressed genes in Gc_3. UMAP cluster maps of established pregranulosa marker genes (*Amh*, *Inha* and *Nr5a2*) identified in our differential gene expression analysis and DGE comparisons that are upregulated during primordial follicle activation. UMAP cluster maps are shown for each timepoint individually and only include cells from the timepoints in each panel (E18.5 = yellow, PND4 = pink and PND7 = blue). Dots are coloured according to their level of gene expression, with high levels of expression shown in purple.

1044 **Supplemental Tables**
1045
1046 **Supplemental Table 1. Single cell RNA-seq quality control summary**
1047

Stage	Number of Cells	UMI/cell	Genes/cell	Number of reads	Fraction reads in cells	Reads mapped to genome
E18.5	8,210	12,011	3,698	516,012,231	90.3%	90.2%
PND4	7,587	10,267	3,413	468,968,268	91.4%	92.1%
PND7	9,013	9,850	2,982	435,022,244	88.8%	92.1%

1048
1049 **Supplementary Table 2: Gene list of differentially expressed transcripts in granulosa cell**
1050 **clusters compared to all other cell clusters from differential gene expression analysis,**
1051 **presented by highest log_2FC.**
1052
1053 **Supplementary Table 3: Gene list of differentially expressed transcripts Gc_3 vs. Gc_2**
1054 **(log_2 fold change ± 2; $p \leq 0.05$), presented by highest average differential gene expression.**
1055 **Positive values are upregulated in Gc_3 cells (cells coloured in pink) and negative values**
1056 **are upregulated in Gc_2 cells (cells coloured in orange).**
1057
1058 **Supplementary Table 4: Gene list of differentially expressed transcripts Gc_3 vs. Gc_4**
1059 **(log_2 fold change ± 2; $p \leq 0.05$), presented by highest average differential gene expression.**
1060 **Positive values are upregulated in Gc_3 cells (cells coloured in pink) and negative values**
1061 **are upregulated in Gc_4 cells (cells coloured in maroon).**
1062
1063 **Supplementary Table 5: Gene list of differentially expressed transcripts between**
1064 ***Cdkn1b*$^{-/-}$ and WT ovaries at E18.5 ($p \leq 0.05$), presented by highest log_2 fold change.**
1065 **Positive values are upregulated in *Cdkn1b*$^{-/-}$ ovaries (cells coloured in blue) and negative**
1066 **values are upregulated in WT ovaries (cells coloured in green).**
1067
1068 **Supplementary Table 6: Overlap of differentially expressed genes upregulated in Gc_3**
1069 **vs Gc_2 and genes upregulated in p27KO E18.5, presented by highest differential gene**
1070 **expression in Gc_3 in the single-cell RNA sequencing dataset. Data taken from**
1071 **Supplementary Tables 3 and 5.**
1072
1073 **Supplementary Table 7: Overlap of differentially expressed genes upregulated in Gc_4**
1074 **vs Gc_3 and genes upregulated in p27KO E18.5, presented by highest differential gene**
1075 **expression in Gc_4 in the single-cell RNA sequencing dataset. Data taken from**
1076 **Supplementary Tables 4 and 5.**
1077
1078