1   **Integrated analyses of the single-cell ATAC-seq and RNA-seq reveal the epigenetic landscape of**

2   **human ovarian aging**

3   **Short title: Profiling human ovarian aging by single cell seq analyses**

4   Yunzhao Xu[1,2*], Jinling Chen[3*], Shuting Gu[1,2*], Yuanlin Liu[1], Huihua Ni[1], Yan Huang[1], Sainan
5   Gao[1], Nan Sheng[1], Xiaojing Zhang[4], Xin He[5], Songlin Zhou[6#], Wenliang Ge[7#]

6

7   Affiliation:

8   [1]Department of Obstetrics and Gynecology, Affiliated Hospital of Nantong University, Nantong
9   226001, Jiangsu, China
10   [2]Department of Obstetrics and Gynecology, School of Medicine, Nantong University, Nantong
11   226001, Jiangsu, China
12   [3]Department of Pathogen Biology, School of Medicine, Nantong University, Nantong, 226001,
13   Jiangsu, China
14   [4]Department of Clinical Biobank, Affiliated Hospital of Nantong University, Nantong 226001,
15   Jiangsu, China
16   [5]Department of Pathology, Affiliated Hospital of Nantong University, Nantong 226001, Jiangsu,
17   China
18   [6]Key Laboratory of Neuroregeneration of Jiangsu and Ministry of Education, NMPA Key
19   Laboratory for Research and Evaluation of Tissue Engineering Technology Products, Jiangsu
20   Clinical Medicine Center of Tissue Engineering and Nerve Injury Repair, Co-Innovation Center
21   of Neuroregeneration, Nantong University, Nantong, Jiangsu Province, China.

22   [7]Department of Pediatric Surgery, Affiliated Hospital of Nantong University, Nantong 226001,
23   Jiangsu, China
24
25   *These authors contributed equally to this study.
26
27   #Correspondence to: Email: songlin.zhou@ntu.edu.cn (SZ); gewl@ntu.edu.cn (WG)

28

33   Author contributions: Conceived and designed the experiments: Y.X., J.C., S.G., S.Z., and W.G.

34   Performed the experiments: Y.X., J.C. and S.G. Analyzed the data: Y.X., J.C. and S.G. Contributed

35   reagents/materials/analysis tools: Y.L., H.N., Y.H., S.G., N.S., X.Z.and X.H. Wrote the manuscript:

36   Y.X., S.Z., and W.G. All authors have read and approved the final manuscript.

37   **Abstract**

38   Studying the molecular mechanisms of ovarian aging is crucial for understanding the age-related fertility

39   issues in females. Recently, a single-cell transcriptomic roadmap of ovarian aging based on non-human

40   primates revealed the molecular signatures of the oocytes at different developmental stages. Herein, we

41   present the first epigenetic landscape of human ovarian aging, through an integrated analysis of the single-

42   cell assay for transposase-accessible chromatin using sequencing (scATAC-seq) and single-cell RNA-seq.

43   We depicted the transcriptional profiles and chromatin accessibility of the ovarian tissues isolated from old

44   (n=4) and young (n=2) donors. The unsupervised clustering of data revealed seven distinct cell populations

45   in the ovarian tissues and six subtypes of oocytes, which could be distinguished by age difference. Further

46   analysis of the scATAC-seq data from the young and old oocytes revealed that the interaction between the

47   Notch signaling pathway and AP-1 family transcription factors may crucially determine oocyte aging.

48   Finally, a machine-learning algorithm was applied to calculate the optimal model based on the single-cell

49   dataset for predicting oocyte aging, which exhibited excellent accuracy with a cross-validated area under

50   the receiver operating characteristics score of 0.99. In summary, this study provides a comprehensive

51   understanding of human ovarian aging at both the transcriptomic and epigenetic levels, based on an

52   integrated analysis of large-scale single-cell datasets. We believe our results will shed light on the discovery

53   of potential therapeutic targets or diagnostic markers for age-related ovarian disorders.

Itroduction

55   The ovary is the chief female reproductive organ that not only stores and provides oocytes but also supplies

56   essential sex hormones, fundamentally determining the reproductive lifespan and maintaining the endocrine

57   homeostasis in females[1]. However, women are born with a finite number of follicles[2]. As the number of

58   follicles decreases with age, female fertility declines and eventually leads to menopause as a physiological

59   consequence of ovarian aging[3]. Also, the specific reason for the cessation of ovulation in females after 35

60   years of age remains unknown[4]. Ovarian aging is also closely related to a wide spectrum of diseases,

61   including ovarian cancer, breast cancer, cardiovascular disease, and type 2 diabetes mellitus[5,6]. This

62   necessitates the elucidation of the mechanism of ovarian aging and identification of the possible target to

63   delay the process.

64   Ovary possesses a complex structure with a heterogeneous population of cells at progressive developmental

65   stages[7]. The follicle is the main functional unit of the ovary[8]. Each follicle consists of a partially

66   differentiated oocyte, enclosed within increasing layers of somatic granulosa cells and theca-interstitial

67   cells that support the oocyte during maturation[1]. Other components of the ovary include the stromal cells,

68   smooth muscle cells, endothelial cells, and various immune cells[9,10]. Therefore, unraveling the cellular

69   landscape of the ovary is indispensable for a better understanding of cell-type-specific dynamics during

70   ovarian aging.

71   Advances in single-cell RNA sequencing (scRNA-seq) technology enable the cellular landscape to be

72   mapped at the single-cell level and the cell trajectory to be inferred for examining the dynamic changes in

73   the cell populations[11]. Previous studies performed the scRNA-seq on the ovaries of young and old

74   cynomolgus monkeys[7]. Researchers have depicted the transcriptional profile of four oocyte subtypes and

75   demonstrated that oxidative damage and subsequent responses are essential for the aging-related changes

76   in early-stage oocytes and granulosa cells[7]. Another study focused on the growth and regression of the

77   follicular populations in human adult ovaries[10]. Various cell types have been identified, and a significant

78　association was found between the complement system and follicular remodeling[10]. However, the

79　epigenetic landscape of ovarian aging in humans is unclear.

80　In addition to the transcriptional regulation, epigenetic reprogramming also takes place in the postnatal

81　oocyte growth and ovarian aging. Previous studies reported that persistence of methylation at CpG islands

82　affected the expression of genes related to the preimplantation development in oocytes[12,13]. Nevertheless,

83　the chromatin landscape and its dynamics during the ovarian aging are still unexplored. Single cell assay

84　for transposase-accessible chromatin using sequencing (scATAC-seq) is an exquisite tool for sketching the

85　chromatin accessibility profile at single-cell resolution[14]. Combined with the scRNA-seq, the

86　comprehensive analysis through scATAC-seq is promising for unveiling the epigenetic mechanism of

87　ovarian aging.

88　**Results**

89　**Single-cell sequencing of over 49,000 cells from the ovarian tissues of six patients**

90　To study the transcriptomic heterogeneity and compare the old and young ovarian tissues based on the cell

91　types, the tissues were collected from six patients (**Materials and Methods, Supplementary Figure 1**)

92　and the single-cell sequencing was performed using the 10x Genomics platform. After the quality control

93　and batch integration, 49,651 cells were clustered into ten initial clusters (**Figure 1A**). The marker genes

94　from the literature were used to annotate these clusters, including CDKN1C for the stromal cells, TAGLN

95　and SOD3 for the oocytes, SPARC and TIMP1 for the smooth muscle cells, INHA for the granulosa cells,

96　CDH5 and VWF for the endothelial cells, CD14 and CD74 for the macrophages, IL32 and CCL5 for the

97　natural killer (NK) cells (**Figure 1B**). The clusters were then integrated and renamed based on the marker

98　genes and clustered cells into seven cell types (**Figure 1C**). The top markers of each cell type were

99　calculated and showed clear patterns for each cell type (**Figure 1D**). To further validate the annotations of

100　the clusters, the functional enrichment analyses were performed for the significantly highly expressed genes

101　for each cluster (p-value < 0.05, fold-change > 1.2) (**Supplementary Figure 2**). As expected, the enriched

102 Gene Ontology (GO) terms revealed critical functions of the corresponding cells. For example, the top

103 enriched GO terms for the macrophages and NK cells were mainly immune-related pathways, such as

104 "regulation of immune response," "immune response-activating signal transduction" and "activation of

105 immune response."

106 We further explored the changes in the composition of each cell type between the old and young ovarian

107 samples (**Figure 1E**). It is worth mentioning that the number of macrophages and NK cells was significantly

108 decreased in the old patient samples compared to the young patient samples, while the stromal cells

109 increased in the old patient samples. Aging has been reported to impair many aspects of macrophage

110 functions[15-17] and has also been identified to have critical impacts on the macrophage-involved immune

111 response.

112 **Six subtypes of oocytes show distinct gene expression signatures at the sequential development stage**

113 We conducted an unsupervised analysis of oocyte gene expression profiles to explore the dynamic changes

114 in the oocyte transcriptomic profiles and identified six subtypes of oocytes (C0 to C5). Uniform Manifold

115 Approximation and Projection (UMAP) revealed that the six subtypes were distributed along the UMAP1

116 dimension (**Figure 2A**). To investigate the dynamic changes in gene expression during oocyte development,

117 we calculated a subset of feature genes that exhibited high subtype-to-subtype variations (**Figure 2B**).

118 Notably, we found that most of the C1–C2 oocytes were from young tissues, while the C3–C5 oocytes were

119 from the old. We subsequently examined the gene expression across the six subtypes and identified the top

120 10 highly expressed genes for each subtype (**Figure 2D**). To validate the functional difference of each

121 subtype, we conducted the GO analysis for the significantly highly expressed genes for each cell type ($p<$

122 0.05) (**Figure 2D**). Consistent with the previous reports and the above canonical gene expression patterns,

123 the GO analysis revealed the biological processes that were enriched for each stage of follicular

124 development. Notably, the "protein targeting to endoplasmic reticulum" and "ribosome biogenesis" were

125 enriched for subtype C0 with high expression of ribosomal protein-related genes (representative genes

126 *RPS8, RPS9*, and *RPS14*). We also identified that "oxidative phosphorylation" (representative genes

127   *ATP5ME* and *COX5B*) and "ATP synthesis coupled electron transport" (representative genes *NDUFA4* and

128   *COX6C*) were enriched for the subtype C4. The variety of enriched GO terms revealed the functional

129   heterogeneity of these subtypes during folliculogenesis. As shown in **Figure 2B**, we identified a subset of

130   feature genes that exhibited high subtype-to-subtype variations. We then examined the expression trends

131   of these feature genes and identified the genes that gradually increased or decreased with oocyte

132   development. It is worth mentioning that we found that two critical oocyte-related genes, *TIMP1* and *TIMP2*,

133   had significant expression trends in correlation with oocyte development (**Figure 2E**). Previous studies

134   have reported that TIMP1 and TIMP2 are associated with oocyte maturity and quality and are necessary

135   for normal follicular development [18,19]. In our results, TIMP1 and TIMP2 displayed a gradually decreasing

136   expression along with oocyte development.

137   **The scATAC-seq profiling of human ovaries identified different cell types and gene expression**

138   **signatures**

139   To profile the genome-wide chromatin accessibility of the cells from the old and young ovarian tissues, a

140   scATAC-seq approach was applied to investigate the differences between the old and young ovarian

141   samples. We utilized "ArchR," a state-of-the-art scATAC-seq analysis R package, for quality control,

142   doublet removal, cell clustering, and annotation, peak calling, differential chromatin accessibility analysis,

143   and footprinting analysis were performed[20]. Cell barcodes with more than 4,000 total unique fragments and

144   a transcription start site enrichment score of more than 8 were selected for downstream analysis

145   (**Supplementary Figures3A and B**). After quality control and removal of doublets, a total of 38,312 single

146   cells were obtained from four old and two young patients' samples (**Supplementary Figures3C, 4, and 5**).

147   For the downstream cell clustering and visualization, the graph-based Louvain algorithm and UMAP were

148   implemented in the 'Seurat' R package. Based on chromatin accessibility profiling, 38,312 cells were

149   represented by a total of 16 clusters (**Figure 3A**). To accurately annotate the clusters, the aforementioned

150   scRNA-seq annotation information was integrated into our scATAC-seq analysis. Finally, seven distinct

151   cell clusters representing the cell types were discovered by scRNA-seq data, including endothelial cells,

152   granulosa cells, macrophages, NK cells, oocytes, smooth muscle cells, and stromal cells (**Figure 3B**).

153   By using ArchR's default peak caller MACS2 with recommended settings, a total of 258,273 peaks were

154   identified in the old and young ovarian samples. After identification of the robust peak sets for each cell

155   type, we further predicted the type of transcription factors (TFs) that may mediate the binding events

156   creating accessible chromatin sites. The motif enrichment results revealed distinct sets of TFs for cell types,

157   indicating the critical functions of these TFs in regulating the cells (**Figure 3C**). The footprinting analysis

158   also revealed the specific upregulated binding of the TFs for each cell type (**Figure 3D**). For example, the

159   RUNX1 TF has been reported to be associated with the NK cell clonal expansion and memory formation.

160   In addition, STAT6 is a key TF in macrophage M2 polarization[21,22]. Our results were supported by the

161   results of previous studies and reveal the novel functional TFs for each cell type during ovarian aging.

The integrated scRNA-seq and scATAC-seq analyses demonstrate the functional roles of the AP-1

163   **TF family and Notch signaling pathway in ovarian aging**

164   To investigate the molecular mechanisms of oocyte aging, the chromatin accessibility of oocyte cells from

165   old and young patients was further compared. The MACS2 with recommended settings was applied to

166   identify the peaks for the old and young oocyte cells and predict the TFs that mediate the binding events

167   creating accessible chromatin sites (**Figure 4A**). The motif enrichment results revealed distinct sets of TFs

168   in the old and young oocytes, which implied the distinct functional roles of TFs in regulating oocyte aging

169   (**Figure 4B**). Notably, we found that multiple critical AP-1 TF family members, such as FOS, JUNB,

170   FOSL1, and FOSL2, were highly enriched in the old oocyte cells compared to the young oocytes. We

171   further performed the functional enrichment analysis for the differentially enriched TFs between the old

172   and young oocytes. It is worth mentioning that several critical aging pathways, including "AP-1 TF

173   network," "tumor necrosis factor signaling pathway" and "Aging" terms, were enriched for significantly

174   upregulated TFs in the old oocyte cells (**Figure 4C**). The aging term was enriched with several critical

175   upregulated genes, including *FOS, CREB1, JUN, TP63*, and *PAX2*. For the enriched pathways for

176   upregulated TFs in young oocyte cells, the Notch signaling pathway was notably one of the most

177   significantly enriched terms (**Figure 4D**). AP-1 was reported to negatively regulate the Notch signaling

178   pathway [23], which is consistent with our findings, indicating an interaction between these two pathways. In

179   our results, the AP-1 TF family activity was significantly upregulated in the old oocyte cells, whereas the

180   Notch signaling pathway was downregulated. The Notch signaling pathway has been widely reported to be

181   associated with aging and plays important role in age-related diseases[24,25]. Our results support the negative

182   regulation between AP-1 and the Notch signaling pathway and suggest the potential functional roles of the

183   AP-1 TF family in aging. To further validate the upregulation of the AP-1 family in the old oocyte cells,

184   we conducted the footprinting analysis for several critical AP-1 family members and integrated their

185   expression in the scRNA-seq data. We found that FOSL2 and ATF3 showed profoundly upregulated

186   patterns in the old oocyte cells in both the footprinting and scRNA-seq results (**Figure 4E; Supplementary**

187   **Figure 6**). FOSL2 and ATF3 worked orchestrally in the AP-1 complex. In order to validate this crucial

188   finding, we performed immunostaining assays on both young and old ovarian tissues. Interestingly, the

189   positive signal of AP-1 complex (c-fos and c-jun) can only be found in old oocyte cells, but not in the young

190   oocyte cells (**Figure 4F**). Therefore, together with the experimental validation results, our integrated

191   analysis of the scRNA-seq and scATAC-seq indicated consistent patterns of upregulation of AP-1 signal in

192   oocyte aging, thus indicating the potential role of AP-1 regulation in ovarian aging.

193   **Machine-learning-based prediction of ovarian aging**

194   It is widely established that bulk RNA-seq data have been used to investigate the biological processes of

195   aging. Recently, expression profiles at the single-cell level have been applied to human skin aging and

196   mouse aging in multiple tissues [26,27]. We found that the gene expression of oocyte cells was significantly

197   different between the young and old ovaries in our scRNA-seq dataset. Next, we applied our scRNA-seq

198   data to predict ovarian aging. A statistical model was then built and optimized to describe the relationship

199   between gene expression and ovarian aging. The mutual information scores were used to rank the genes.

200   The cross-validated AUROC scores for models with different numbers of top-ranked genes are shown in

201   **Figure 5A**. We included the 10 top-ranked genes in the predictive model, as including more genes did not

202   further improve the performance of the model. The performance of the optimized model (see **Materials**

203   **and Methods**) was characterized using Receiver Operating Characteristic (ROC) analysis (**Figure 5B**).

204   The model with the selected 10-gene signature exhibited excellent accuracy, with a cross-validated AUROC

205   score of 0.99. The other metrics of the model were also commensurate with the AUROC score (**Figure 5C**).

206   The gene signature included those of *MT-ATP6, MT-CO3, MT-CYB, IGFBP7, VIM, RPL3, RPL27A,*

207   *HIST1H4C, HLA-A*, and *CD63* (**Figure. 5D**). We also performed the Mann–Whitney–Wilcoxon test with

208   Bonferroni correction and found that all of these genes were significantly differentially expressed between

209   the two groups with *p*-values smaller than 10^{-4}. These findings suggest excellent predictability of ovarian

210   aging using the 10-gene signature expression of the oocyte cells.

211   **Profiling the cell-cell communication in young ovarian samples**

212   To further investigate the mechanisms of ovarian aging, we examined the cross talk of various cell types.

213   As mentioned earlier, we found that the Notch signaling pathway was significantly regulated in young

214   oocyte cells and it also played critical roles in cell communications. Therefore, we used CellPhoneDB[28] to

215   profile the communication among cell types in the scRNA-seq data of young patients. The interactions

216   among smooth muscle cells, endothelial cells and granulosa cells are the strongest, while others are

217   relatively weaker **(Figure 6A, B)**. Furthermore, we examined the interactions between oocyte and other

218   cell types. Notably, several Notch family receptors were significantly enriched, which include Notch2,

219   Notch3 and Notch 4 (**Figure 6C**). These pairs of interactions could play important roles in ovarian aging

220   for oocytes to interacting with other cell types.

221   **Discussion**

222   The ovary is characterized by highly heterogeneous populations at progressive developmental stages, and

223   its cellular landscape is of indispensable importance to clarify the mechanism of ovarian aging. In this study,

224   we integrated the scRNA-seq and scATAC-seq analyses to depict the transcriptional profiles and chromatin

225   accessibility of old and young ovarian tissues. Unsupervised clustering of the scRNA-seq data revealed

226   seven distinct cell populations. Annotation results were integrated with and confirmed by the scATAC-seq

227   data. Marker genes, enriched pathways, and featured motifs for each cluster were identified. Notably, six

228   subtypes of oocytes were recognized, representing different stages of oocyte development. Finally, a

229   machine-learning algorithm was applied to calculate the optimal model for predicting ovarian aging.

230   The difficulty in procuring human ovary samples has long been an obstacle in revealing the cellular

231   landscape of ovarian aging. Combining scRNA-seq and scATAC-seq data, we presented the first integrative

232   study of ovarian aging at single-cell resolution. After multiple clustering and annotation, seven cell clusters

233   were identified, including endothelial cells, granulosa cells, macrophages, NK cells, oocytes, smooth

234   muscle cells, and stromal cells. This is consistent with the results of previous studies. Wang et al. performed

235   the scRNA-seq analysis on the ovary samples collected from four young and four aged cynomolgus

236   monkeys[7]. The same clustering results were found in their analysis, whereas the marker genes were slightly

237   different[7]. Fan et al. applied scRNA-seq to human adult ovaries and depicted the molecular signature of

238   growing and regressing cell populations[10]. Somatic cell types, including endothelial cells, granulosa cells,

239   smooth muscle cells, and stromal cells, were also found in their study [10]. In conclusion, the major cell types

240   in human ovary samples were unveiled in this study.

241   The reclustering and cell trajectory analysis showed six subtypes of oocytes with distinct gene expression

242   signatures during the sequential development stage. The differentially expressed genes were examined

243   across the six subtypes for expression trends and oocyte development. Interestingly, TIMP1 and TIMP2

244   were negatively correlated with folliculogenesis. The TIMP gene family encodes proteins that inhibit matrix

245   metalloproteinases (MMPs), thus preventing degradation of the extracellular matrix[29]. Previous studies

246   have reported that TIMP1 and TIMP2 are involved in follicular development. Robinson et al. demonstrated

247   that TIMP family members are localized in the oocyte cytoplasm and play an important role in the

248   remodeling of the extracellular matrix during human gonadal development[30]. By investigating the gene

249   expression levels of TIMP1 and TIMP2 in the granulosa and cumulus cells, Luddi et al. demonstrated that

250   MMP and TIMP expression is involved in the regulation of oocyte maturation[31]. In this study, we showed

251   that TIMP1 and TIMP2 were both significantly downregulated in the C4 and C5 oocytes (**Figure 2C and**

252   **2E**), which were mostly from old patients. This result indicates that the expression levels of TIMP1 and

253   TIMP2 may determine the oocyte development stage as well as the aging process. Future studies should

254   focus on deciphering the mechanistic roles of TIMP1 and TIMP2 in oocyte aging.

255   Chromatin accessibility analysis and robust peak identification revealed that the AP-1 TF family members,

256   such as FOS, JUNB, FOSL1, and FOSL2 were highly enriched in the old oocyte cells compared to the

257   young oocyte cells, consistent with the functional enrichment analysis and foot-printing analysis. Previous

258   studies have reported that the AP-1 TFs participate in the early stages of murine follicle growth [32]. Moreover,

259   the AP-1 TF family has been reported to be involved in the aging of various cells. Shi et al. observed

260   decreased AP-1 transcriptional activity in senescent human keratinocytes[33]. Medicherla et al. demonstrated

261   that oxidative stress may contribute to impaired AP-1 binding activity, leading to adrenal aging in rats[34].

262   ATF3, an AP-1 TF, was found to be critical for the remodeling chromatin accessibility, thus promoting

263   senescence of the human umbilical vein endothelial cells[35]. Herein, we demonstrated the crucial role of AP-

264   1 genes in ovarian aging. Additionally, the Notch signaling pathway was found to be negatively regulated

265   by the AP-1 complex [36]. Chen et al. demonstrated that treatment with a Notch signaling pathway inhibitor

266   led to a significant decrease in oocytes in the primordial follicles of newborn mice[37], which implied that

267   downregulation of the Notch signaling pathway is correlated with the dysfunction of the ovary. These

268   studies indicated that the AP-1 TF family and its interaction with the Notch signaling pathway might

269   regulate follicle formation and growth, thus contributing to ovarian aging. Further functional studies are

270   needed to clarify their potential roles in the human ovary.

271   Finally, we built and optimized a predictive model based on the signature genes that were differentially

272   expressed between the young and old ovaries. These genes include *MT-ATP6, MT-CO3, MT-CYB, IGFBP7,*

273   *VIM, RPL3, RPL27A, HIST1H4C, HLA-A*, and *CD63*. The model presented excellent accuracy in cross-

274   validation, suggesting its predictive and diagnostic significance for ovarian aging. Taken together, these

275   observations provide novel insights into ovarian aging in humans. More detailed studies are essential to

276   evaluate the function of the TIMP gene family and AP-1 TF family in oocyte development and maturation.

277   Moreover, a predictive model based on machine learning promises to assist in the prediction of ovarian

278   aging. Signature genes identified in this study may serve as new biomarkers and therapeutic targets for the

279   diagnosis and treatment of age-related ovarian disorders.

280   **Material and methods**

281   **Sample information**

282   The experiments performed in this study were approved by the Ethics Committee of Affiliated Hospital of

283   Nantong University. With informed consent, fresh ovarian tissues used for this study were from 6 donors

284   who all underwent surgical treatment in Department of Obstetrics and Gynecology, Nantong University

285   Affiliated Hospital between 2019-2020. All cases were excluded the use of hormones and malignant tumors.

286   4 cases menopausal women represent the elderly group (mean age 55.5 years, 50-62 years), who underwent

287   whole uterus and double attachment resection because of uterine leiomyoma (n=3) and benign ovarian mass

288   (n = 1). And the other 2 cases with normal fertility aged 31 and 34 years old who both underwent ovarian

289   cyst removal surgery, we trimmed the irregular residual ovarian tissue to neat, and took a very small piece

290   of normal ovarian tissue at the edge of the residual tissue during the operation. All the normal ovarian tissue

291   were confirmed without histopathological abnormality by at least two independent pathologists.

292   **Single nuclei preparation and ATAC-library construction**

293   Single   nuclei   ATAC-seq   was   performed   following   10x   Genomics

294   NextGEM_SingleCell_ATAC_ReagentKits_v1.1 protocol. Homogenization was applied to tissues for

295   breaking the tissues and obtaining the purified nuclei. Nuclei concentration was determined with

296   hemocytometer under microscope and diluted in nuclei buffer (10x Genomics). The transposition

297   experiments were then set up with thermocycler; The GEMs were then produced and cleaned up with the

298   dynabeads, followed by performing the PCR reactions to add sample index. After quality control (QC) and

299   quantification of the products, it was used for following library construction under the standard protocol

300   (10x Genomics). The library was then sequenced using Illumina NovaSeq platform.

301   **Single cell preparation and library construction (for scRNA-seq)**

302   Separated ovarian tissues were dissociated to single-cell resuspension for the single-cell transcriptome

303   sequencing. The tissues were rinsed in cold DPBS for three times after dissection. Then the tissues were

304   carefully transferred to digestion buffer (2mg/ml collagenase I, 1ml; 2mg/ml collagenase II, 1ml; 0.9U

305   dispase, 0.5 ml; trypsin, 0.5 ml; DMEM, 1.5ml. Pre-heat to 37 degrees). The tissues were cut to appropriate

306   smaller pieces in the digestion buffer and then gently shook in a 37-degree metal heater-shaker for 15

307   minutes. After the treatment, they were taken 10 µl from the resuspension and observed with a

308   hemocytometer under microscope. If the cells were not sufficiently digested, it would be treated with

309   digestion buffer for another 5 minutes. The cell suspension was then filtered by a 70-µm mesh filter,

310   followed by washing the filter with another 5ml DMEM. The cells were concentrated by centrifugation

311   (500g, 5 minutes at 4 degrees), and then the supernatant was removed. The cells were resuspended with 50

312   to 100 µl DMEM with 10% phosphate-buffered saline (PBS). The cells were then observed and counted

313   with hemocytometer. The cell concentration was adjusted to 700 to 1200 cells per µl before loading to 10x

314   Genomics Chromium machine.

315   The GEMs were produced and collected to perform reverse transcription for barcoding, and the first strand

316   cDNA was purified with magnetic beads. After QC and quantification of the cDNA, it was then used for

317   library construction under standard protocol (10X Genomics). The library was sequenced using Illumina

318   NovaSeq platform.

319   **Single cell RNA-seq analysis**

320   Seurat[38] was used to read the Cellranger outputs (expression matrices) and used as the platform for

321   downstream analysis. For quality control of each dataset, lowly-detected genes and cells with limited

322   number of genes were discarded from the downstream analysis, in order to avoid the analysis driven by

323   noise or low-quality cells. The gene filter was set to detection in at least 0.5% of all cells and the cell filter

324   was set to a minimum number of 1000 of total expressed genes per cell. After the QC step, for each

325   expression matrix, the expression value of each gene was normalized and transformed using default

326   NormalizeData function in Seurat. Using the variation stabilizing transformation method ('vst' from Seurat),

327　the top 2000 variable genes were selected in each matrix and were used as input for the

328　'FindIntegrationAnchors' function of Seurat. The expression matrices were then integrated with the

329　'IntegrateData' function. The integrated data were dimension reduced with principal component analysis

330　(PCA; top20 dimensions). In the PCA space, nearest neighbors were defined among cells with KNN method

331　('FindNeighbors', top 15 PCs were selected) and cells were then grouped with Louvain algorithm

332　('FindClusters' in Seurat, resolution equal to 0.5). UMAP dimension reduction was then performed using

333　'RunUMAP' function for visualization purpose. CellPhoneDB was used to profile the communication

334　among cell types in the scRNA-seq data of young and old patients[28].

Detection of differentially expressed genes (DEGs).

336　DEGs were first computed to test significantly highly expressed genes for each cluster. For each cluster,

337　only the genes expressed in more than 25% of that group were considered. All other cells were used as the

338　background. For statistical test, we used the 'MAST' method implemented in Seurat. DEGs were defined

339　as genes whose log fold-change were over 0.4 compared to the background, and with a q-value (FDR)

340　smaller than 0.05.

341　In each cell type, differentially expressed genes of the cells from tissues with primary tumors and those

342　with recurrent tumors were computed. The method is the same as above described except that log fold-

343　change threshold were set to 0.25.

Single-cell ATAC-seq analysis

345　Raw scATAC-seq data was first processed by the 'cellranger-atac' tool downloaded from 10x Genomics,

346　which mapped the sequenced reads to the hg38 genome reference and generated the fragment files. R

347　package 'archR' was then used for the downstream analysis with fragment files as input[20]. For quality

348　control, cells with less than 4000 unique fragments or TSS enrichment scores less than 8 were removed.

349　Doublets were also removed by using the default settings. Cells passing the filters were then clustered by

350　using the 'Seurat' method implemented in the 'archR' package with high resolution values (1.5 or 2).

351 UMAP with 30 neighbors was used for dimension reduction and visualization. For cell cluster annotation,

352 we first integrated the cell annotations generated by previous scRNA-seq data by using the

353 'addGeneIntegrationMatrix' function of 'archR' and then further curated the cell cluster annotations by

354 specific marker genes. For peak calling, we used the recommended 'MACS2' with default settings on

355 aggregated coverage profiles of identified cell populations. For the rest downstream analysis of motif

356 enrichment, footprinting analysis and trajectory plots, we used the functions provided by 'archR' packages

357 with the recommended settings.

358 **Statistical Analysis and machine learning based prediction model**

359 Mutual information scores were used to evaluate feature importance. We also considered permutation

360 importance score to select features, which were outperformed by mutual information in model accuracy

361 benchmarking. Each evaluation score was averaged over 100 repeated runs. Statistical models including

362 random forest classifier, gradient boost classifier, XGB classifier, Naïve Bayes were considered. We used

363 a genetic algorithm (TPOT) to select the best model and optimize its hyperparameters. The area under the

364 curve of the receiver operating characteristic curve (AUROC), was used to evaluate model performance.

365 We also calculated other metrics for comparison purposes. All models were evaluated using 5-fold cross

366 validation with stratified train-test splits that preserve the percentage of samples for the prediction target.

367 All metrics were averaged over 30 repeated runs.

368 **Hematoxylin-eosin staining and Immunofluorescence**

369 Histologic analysis was performed on ovarian tissues using hematoxylin and eosin staining (H&E). After

370 operation, ovarian tissues were fixed in 4% formaldehyde and processed for routine paraffin embedding

371 after fixation. We took 5-mm-thick sections for hematoxylin and eosin staining. Tissue sections were

372 deparaffinized and rehydrated in graded ethanol solutions. In order to repair the antigen, the sections were

373 immersed in antigen retrieval solution (P0088, Beyotime Biotechnology) and heated in 97 ℃ water bath

374 for 20 minutes. First of all, immunol staining blocking buffer (0.5% triton X-100, 5% Horse serum and

375 PBS) was used before staining, which can reduce non-specific staining. Immunostaining was performed

376 using the primary antibody against c-Fos (1:300, Abcam, ab208942, Cambridge, USA) and c-Jun (1:300,

377 Abcam, ab40766, Cambridge, USA) at 4℃ overnight. On completion of the incubation at 4℃, tissues were

378 washed three times with PBS. Then localization of c-Fos was monitored using Alexa Fluor 594 (Goat Anti-

379 Mouse, 1:500, Abcam, ab150116, Cambridge, USA) and Alexa Fluor 488 (Goat Anti-Rabbit, 1:500,

380 Abcam, ab150077, Cambridge, USA) for c-Jun. After 2 hours of co-cultivation at room temperature, tissues

381 were washed 3-5 times with PBS for 10 minutes each time. Finally, antifade mounting medium with DAPI

382 (P0131, Beyotime Biotechnology) was used to stain the nucleus. The tissues were observed under Zeiss

383 microscope and the images were processed by ZEN software.

387   **Figure and figure legends**

388   **Figure 1. Single-cell clustering and annotation.**

389   A. UMAP visualization of initial clustering results.

390   B. Marker genes reported by the literature (Materials and Methods) were used to define clusters. For gene

391   expression, the raw read counts were normalized and transformed.

392   C. UMAP visualization with cell-type annotation results.

393   D. Top 10 marker genes for each cell type visualized in a heatmap. For gene expression, the raw read

394   counts were normalized and transformed, and then scaled.

395   E. Proportions of each cell type in each patient sample.

396   **Figure 2. A single-cell cluster of oocyte cells.**

397   A. UMAP of six subclusters of the oocyte.

398   B. Subset of genes with high variation in subtype-to-subtype transition.

399   C. Proportion of six subclusters of oocyte in the young and old patient samples. Subcluster C0 = identity

400   0, C1 = identity 1 … C5 =identity 5, respectively.

401   C. Heatmap of six oocyte subtypes and their corresponding GO enrichment results.

402   D. Violin plots of TIMP1 and TIMP2.

403   **Figure 3. scATAC analysis of ovarian tissues.**

404   A. UMAP visualization of initial clustering results.

405   B. UMAP visualization of annotated clustering results.

406   C. Enriched transcription factor motifs for each cell type.

407   D. Footprinting plot of representative transcription factors for each cell type.

408   **Figure 4. scATAC analysis of the old and young oocyte cells.**

409   A. Heatmap of enriched motifs of the old and young oocytes.

410   B. Enriched transcription factors of the old and young oocytes.

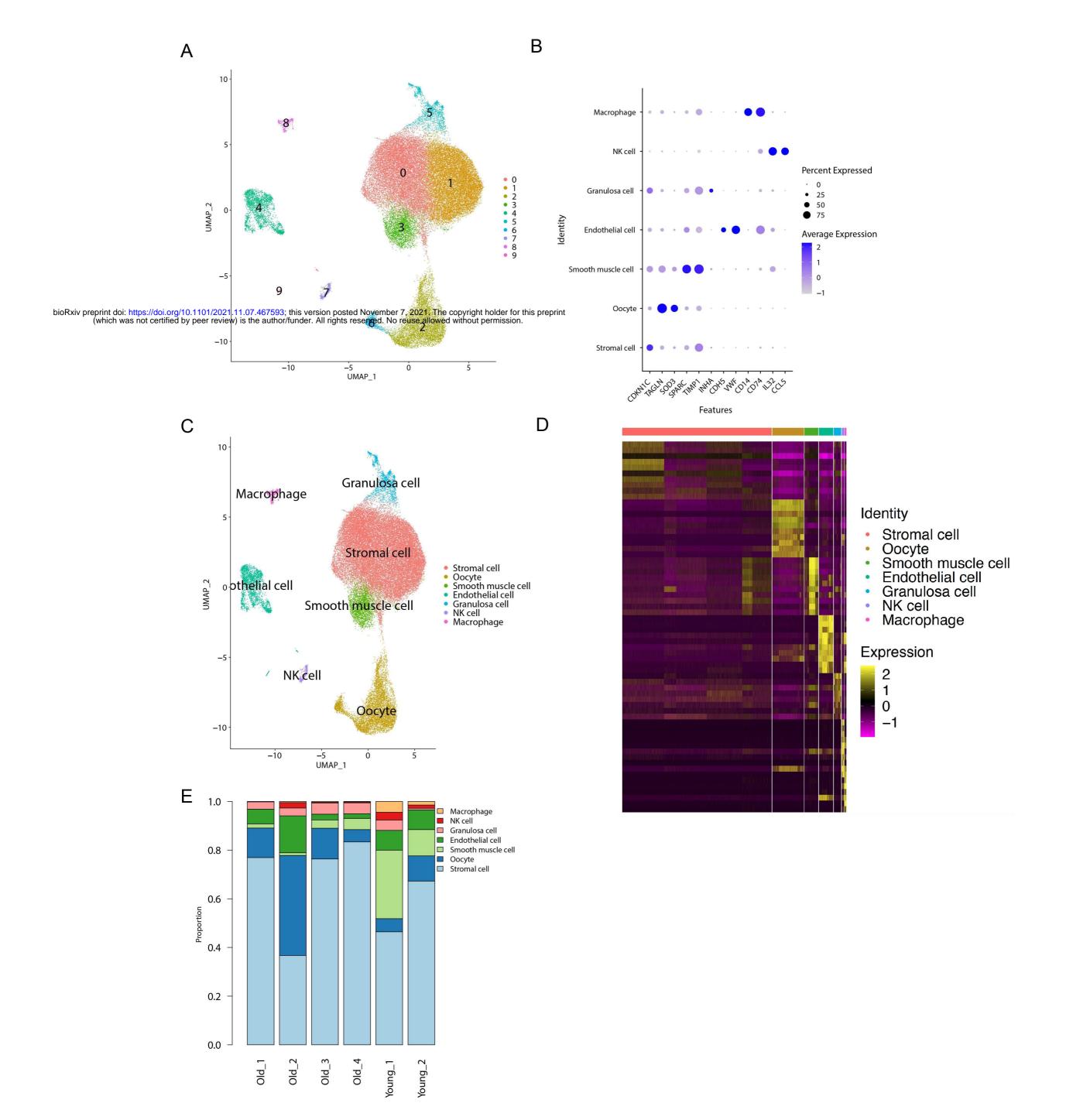411   C-D. Functional enrichment of transcription factor motifs identified for the old and young oocyte cells,

412   respectively.

413   E. Footprinting and violin plots of FOSL2 and ATF3.

414   **Figure 5. Old and young ovaries prediction based on a machine-learning algorithm**.

415   A. Cross-validated AUROC scores for different features included in the model.

416   B. ROC curve for the predictive model using the 10-gene signature.

417   C. Model evaluation metrics. Numbers outside (inside) the parentheses are the mean (standard deviation)

418   of the 100 validation sets in 30 repeated runs of cross-validation.

419   D. Expression of all genes in the signature between the old and young groups.

420   **Figure 6. Cell interactions in young ovarian scRNA-seq data.**

421   A. Heatmap of interaction counts cell-cell interaction in young ovarian samples.

422   B. Cell-cell interaction network in young ovarian samples. Colors and widths of edges represents number

423   of interaction pairs between cell types.

424   C. Bubble heatmap of Notch family receptors between oocyte and other cell types. Dot size represents -

425   log10 p-value, while color represents log2 mean expression value of the receptor-ligand pair.

426   Supplementary Figure 1: H&E staining for the collected tissue.

427   Supplementary Figure 2: Functional enrichment analysis of marker genes for each cell type.

428   Supplementary Figure 3. Quality control of scATAC-seq analysis.

429   A. scATAC-seq signals of the unique fragment number per cell for old and young groups.

430   B. scATAC-seq signals of old and young groups are highly enriched on TSS.

431   C. Fragment size distribution for old and young groups.

432   Supplementary Figure 4. Doublets removal by ArchR for four old patient samples.

433   Supplementary Figure 5. Doublets removal by ArchR for two young patient samples.

434   Supplementary Figure 6. Browser track showing the scATAC-seq signal of FOSL2 and ATF3 across old

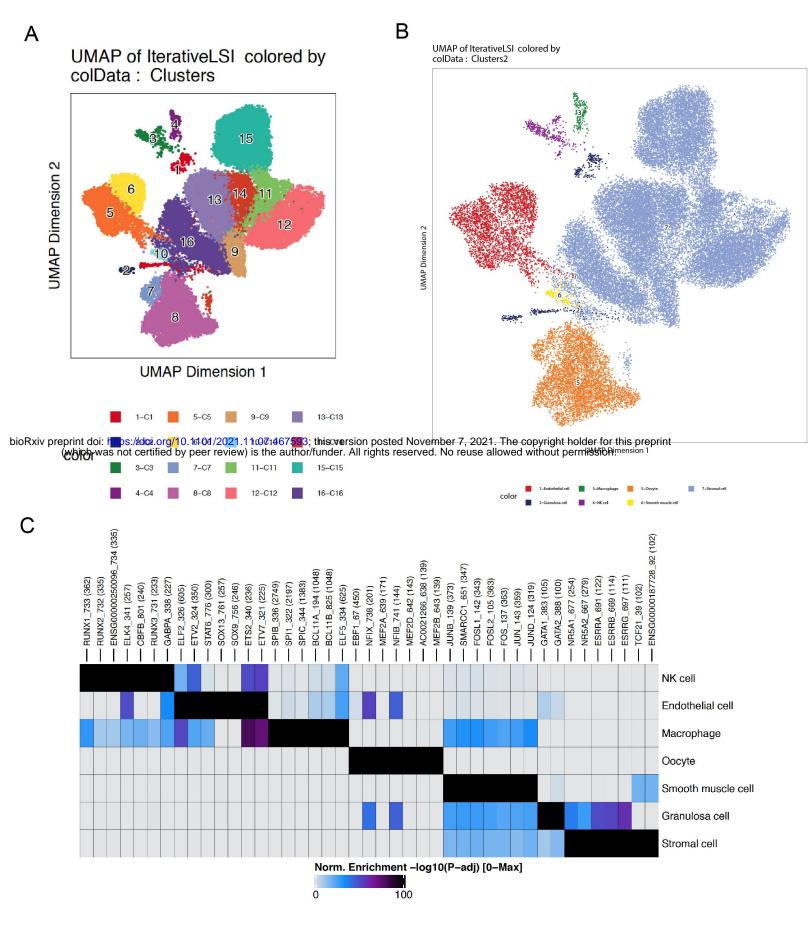435   and young patient samples.

436

References

1   Tilly, J. L. & Sinclair, D. A. Germline energetics, aging, and female infertility. *Cell Metab* **17**, 838-850, doi:10.1016/j.cmet.2013.05.007 (2013).

2   Perheentupa, A. & Huhtaniemi, I. Aging of the human ovary and testis. *Mol Cell Endocrinol* **299**, 2-13, doi:10.1016/j.mce.2008.11.004 (2009).

3   Pelosi, E., Simonsick, E., Forabosco, A., Garcia-Ortiz, J. E. & Schlessinger, D. Dynamics of the ovarian reserve and impact of genetic and epidemiological factors on age of menopause. *Biol Reprod* **92**, 130, doi:10.1095/biolreprod.114.127381 (2015).

4   Laisk, T. *et al.* Demographic and evolutionary trends in ovarian function and aging. *Hum Reprod Update* **25**, 34-50, doi:10.1093/humupd/dmy031 (2019).

5   Perry, J. R., Murray, A., Day, F. R. & Ong, K. K. Molecular insights into the aetiology of female reproductive ageing. *Nat Rev Endocrinol* **11**, 725-734, doi:10.1038/nrendo.2015.167 (2015).

6   Broekmans, F. J., Soules, M. R. & Fauser, B. C. Ovarian aging: mechanisms and clinical consequences. *Endocr Rev* **30**, 465-493, doi:10.1210/er.2009-0006 (2009).

7   Wang, S. *et al.* Single-Cell Transcriptomic Atlas of Primate Ovarian Aging. *Cell* **180**, 585-600 e519, doi:10.1016/j.cell.2020.01.009 (2020).

8   Gougeon, A. Regulation of ovarian follicular development in primates: facts and hypotheses. *Endocr Rev* **17**, 121-155, doi:10.1210/edrv-17-2-121 (1996).

9   Zhang, A. W. *et al.* Interfaces of Malignant and Immunologic Clonal Dynamics in Ovarian Cancer. *Cell* **173**, 1755-1769 e1722, doi:10.1016/j.cell.2018.03.073 (2018).

10   Fan, X. *et al.* Single-cell reconstruction of follicular remodeling in the human adult ovary. *Nat Commun* **10**, 3164, doi:10.1038/s41467-019-11036-9 (2019).

11   Cao, J. *et al.* The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566**, 496-502, doi:10.1038/s41586-019-0969-x (2019).

12   Gu, C., Liu, S., Wu, Q., Zhang, L. & Guo, F. Integrative single-cell analysis of transcriptome, DNA methylome and chromatin accessibility in mouse oocytes. *Cell Res* **29**, 110-123, doi:10.1038/s41422-018-0125-4 (2019).

13   Hanna, C. W. *et al.* MLL2 conveys transcription-independent H3K4 trimethylation in oocytes. *Nat Struct Mol Biol* **25**, 73-82, doi:10.1038/s41594-017-0013-5 (2018).

14   Pott, S. & Lieb, J. D. Single-cell ATAC-seq: strength in numbers. *Genome Biol* **16**, 172, doi:10.1186/s13059-015-0737-7 (2015).

15   Shaw, A. C., Joshi, S., Greenwood, H., Panda, A. & Lord, J. M. Aging of the innate immune system. *Curr Opin Immunol* **22**, 507-513, doi:10.1016/j.coi.2010.05.003 (2010).

16   Plowden, J., Renshaw-Hoelscher, M., Engleman, C., Katz, J. & Sambhara, S. Innate immunity in aging: impact on macrophage function. *Aging Cell* **3**, 161-167 (2004).

17   Sebastián, C., Espia, M., Serra, M., Celada, A. & Lloberas, J. MacrophAging: a cellular and molecular review. *Immunobiology* **210**, 121-126 (2005).

18   McCaffery, F. H., Leask, R., Riley, S. C. & Telfer, E. E. Culture of bovine preantral follicles in a serum-free system: markers for assessment of growth and development. *Biology of reproduction* **63**, 267-273 (2000).

19   Ayvazova, N. P., Ilieva, L. O., Konova, E. I. & Atanasova, M. A. Levels of MMP2, TIMP1 and TIMP2 in Follicular Fluids in Women Undergoing in Vitro Fertilization and Their Relationship to Oocyte Quality. *Journal of Biomedical and Clinical Research* **12**, 100-107 (2019).

20   Granja, J. M. *et al.* ArchR is a scalable software package for integrative single-cell chromatin accessibility analysis. *Nat Genet*, doi:10.1038/s41588-021-00790-6 (2021).

21   Chen, Z. *et al.* Threshold for neural tube defect risk by accumulated singleton loss-of-function variants. *Cell research* **28**, 1039-1041 (2018).
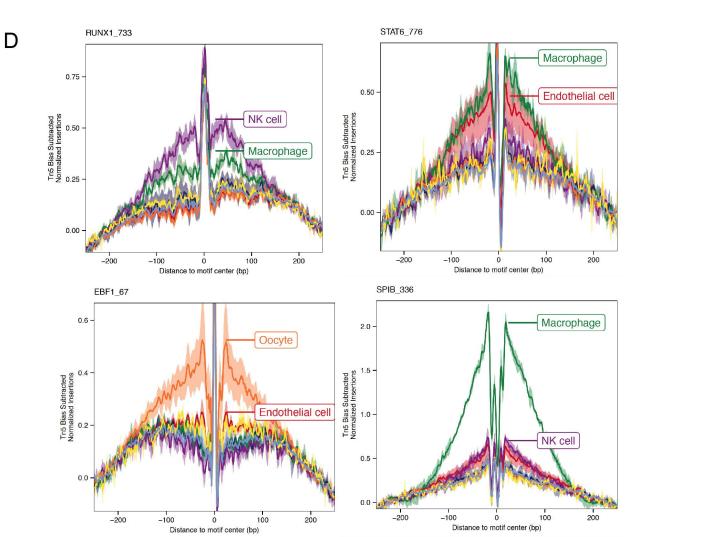
484 22   Yu, T. *et al.* Modulation of M2 macrophage polarization by the crosstalk between Stat6 and
485   Trim24. *Nature communications* **10**, 4353, doi:10.1038/s41467-019-12384-2 (2019).
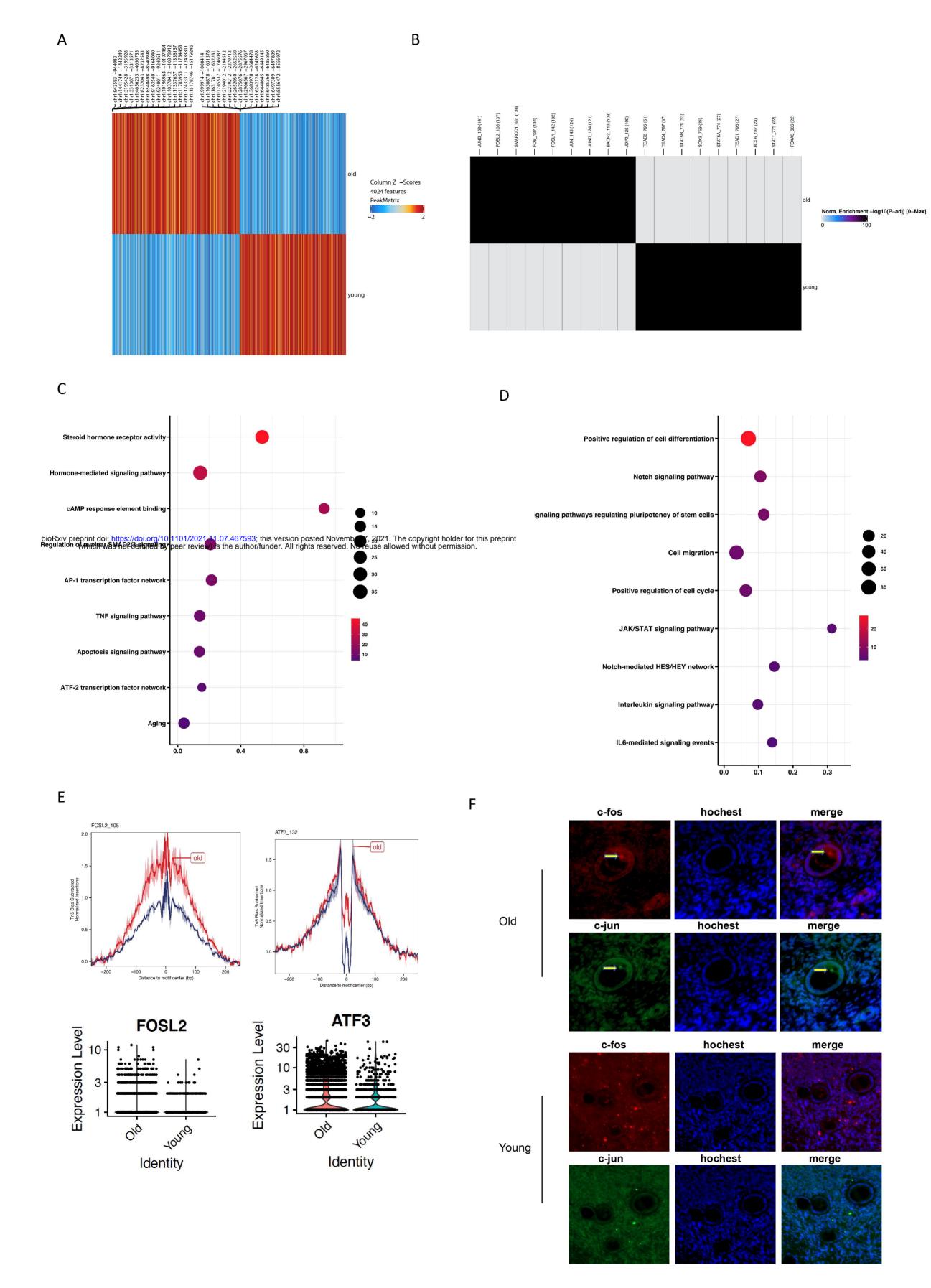486 23   Benhra, N. *et al.* AP-1 controls the trafficking of Notch and Sanpodo toward E-cadherin junctions
487   in sensory organ precursors. *Curr Biol* **21**, 87-95, doi:10.1016/j.cub.2010.12.010 (2011).
488 24   Balistreri, C. R., Madonna, R., Melino, G. & Caruso, C. The emerging role of Notch pathway in
489   ageing: Focus on the related mechanisms in age-related diseases. *Ageing Res Rev* **29**, 50-65,
490   doi:10.1016/j.arr.2016.06.004 (2016).
491 25   Polychronidou, E., Vlachakis, D., Vlamos, P., Baumann, M. & Kossida, S. Notch signaling and
492   ageing. *Adv Exp Med Biol* **822**, 25-36, doi:10.1007/978-3-319-08927-0_6 (2015).
493 26   Zou, Z. *et al.* A Single-Cell Transcriptomic Atlas of Human Skin Aging. *Dev Cell* **56**,
494   doi:10.1016/j.devcel.2020.11.002 (2021).
495 27   A single-cell transcriptomic atlas characterizes ageing tissues in the mouse. *Nature* **583**, 590-595,
496   doi:10.1038/s41586-020-2496-1 (2020).
497 28   Efremova, M., Vento-Tormo, M., Teichmann, S. A. & Vento-Tormo, R. CellPhoneDB: inferring
498   cell-cell communication from combined expression of multi-subunit ligand-receptor complexes.
499   *Nat Protoc* **15**, 1484-1506, doi:10.1038/s41596-020-0292-x (2020).
500 29   Berkholtz, C. B., Shea, L. D. & Woodruff, T. K. Extracellular matrix functions in follicle maturation.
501   *Semin Reprod Med* **24**, 262-269, doi:10.1055/s-2006-948555 (2006).
502 30   Robinson, L. L., Sznajder, N. A., Riley, S. C. & Anderson, R. A. Matrix metalloproteinases and
503   tissue inhibitors of metalloproteinases in human fetal testis and ovary. *Mol Hum Reprod* **7**, 641-
504   648, doi:10.1093/molehr/7.7.641 (2001).
505 31   Luddi, A. *et al.* Matrix metalloproteinases and their inhibitors in human cumulus and granulosa
506   cells as biomarkers for oocyte quality estimation. *Fertil Steril* **109**, 930-939 e933,
507   doi:10.1016/j.fertnstert.2018.01.030 (2018).
508 32   Oktay, K. H. & Oktay, M. H. Immunohistochemical analysis of tyrosine phosphorylation and AP-1
509   transcription factors c-Jun, Jun D, and Fos family during early ovarian follicle development in the
510   mouse. *Appl Immunohistochem Mol Morphol* **12**, 364-369, doi:10.1097/00129039-200412000-
511   00014 (2004).
512 33   Shi, B. & Isseroff, R. R. Epidermal growth factor (EGF)-mediated DNA-binding activity of AP-1 is
513   attenuated in senescent human epidermal keratinocytes. *Exp Dermatol* **14**, 519-527,
514   doi:10.1111/j.0906-6705.2005.00317.x (2005).
515 34   Medicherla, R., Leers-Sucheta, S., Luo, Y. & Azhar, S. Impaired activation of AP-1 and altered
516   expression of constituent proteins in rat adrenal during ageing. *Mech Ageing Dev* **122**, 1169-
517   1186, doi:10.1016/s0047-6374(01)00258-5 (2001).
518 35   Zhang, C. *et al.* ATF3 drives senescence by reconstructing accessible chromatin profiles. *Aging*
519   *Cell*, e13315, doi:10.1111/acel.13315 (2021).
520 36   Benhra, N. *et al.* AP-1 controls the trafficking of Notch and Sanpodo toward E-cadherin junctions
521   in sensory organ precursors. *Curr Biol* **21**, 87-95, doi:10.1016/j.cub.2010.12.010 (2011).
522 37   Chen, C. L. *et al.* Primordial follicle assembly was regulated by Notch signaling pathway in the
523   mice. *Mol Biol Rep* **41**, 1891-1899, doi:10.1007/s11033-014-3038-4 (2014).
524 38   Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic
525   data across different conditions, technologies, and species. *Nat Biotechnol* **36**, 411-420,
526   doi:10.1038/nbt.4096 (2018).

527

A
UMAP of IterativeLSI colored by
colData : Clusters

B
UMAP of IterativeLSI colored by
colData : Clusters2

color

1–C1	5–C5	9–C9	13–C13
3–C3	7–C7	11–C11	15–C15
4–C4	8–C8	12–C12	16–C16

color: 1–Endothelial cell, 2–Granulosa cell, 3–Macrophage, 4–NK cell, 5–Oocyte, 6–Smooth muscle cell, 7–Stromal cell

C

Norm. Enrichment −log10(P−adj) [0−Max]

D

A

B

C

Steroid hormone receptor activity

Hormone-mediated signaling pathway

cAMP response element binding

Regulation of nuclear SMAD2/3 signaling

AP-1 transcription factor network

TNF signaling pathway

Apoptosis signaling pathway

ATF-2 transcription factor network

Aging

D

Positive regulation of cell differentiation

Notch signaling pathway

Signaling pathways regulating pluripotency of stem cells

Cell migration

Positive regulation of cell cycle

JAK/STAT signaling pathway

Notch-mediated HES/HEY network

Interleukin signaling pathway

IL6-mediated signaling events

E

F

A

B

C

Metrics	Score
Precision	0.96 (0.003)
Recall	0.90 (0.016)
AUROC Score	0.99 (0.003)
F1 Score	0.93 (0.011)
Accuracy	0.97 (0.004)

D