

1 **The regulatory landscapes of human ovarian ageing**

2 Chen Jin^{1,9}, Xizhe Wang^{1,9}, Adam D. Hudgins¹, Amir Gamliel², Mingzhuo
3 Pei¹, Seungsoo Kim¹, Daniela Contreras¹, Jan Hoeijmakers^{3,4,5}, Judith
4 Campisi^{6,7}, Rogerio Lobo¹, Zev Williams¹, Michael G. Rosenfeld² and
5 Yousin Suh^{1,8*}

6 ¹ Department of Obstetrics and Gynecology, Columbia University Medical
7 Center, New York, NY

8 ² Howard Hughes Medical Institute, Department and School of Medicine,
9 University of California, San Diego, La Jolla, CA, USA

10 ³ Department Molecular Genetics, Erasmus University Medical Center
11 Rotterdam, Rotterdam, The Netherlands

12 ⁴ Princess Máxima Center for Pediatric Oncology, Onco Institute,
13 Utrecht, The Netherlands

14 ⁵ Institute for Genome Stability in Ageing and Disease, Cologne
15 Excellence Cluster for Cellular Stress Responses in Aging-Associated
16 Diseases (CECAD), University Hospital of Cologne, Cologne, Germany

17 ⁶ Buck Institute for Research on Aging, Novato, CA, USA

18 ⁷ Lawrence Berkeley National Laboratory, Berkeley, CA, USA

19 ⁸ Department of Genetics and Development, Columbia University Medical
20 Center, New York, NY

21 ⁹ These authors contributed equally: Chen Jin, Xizhe Wang

22

23 *Correspondence: ys3214@cumc.columbia.edu (Y.S.)

24

Summary paragraph

The ovary is the first organ to age in the human body, affecting both fertility and overall health in women¹⁻⁸. However, the biological mechanisms underlying human ovarian ageing remain poorly understood. Here we performed single-nuclei multi-omics analysis of young and reproductively aged human ovaries to understand the molecular and cellular basis of ovarian ageing in humans. Our analysis reveals coordinated changes in transcriptomic output and chromatin accessibility across cell types during ageing, including elevated mTOR and MAPK signaling, decreased activity of the oxidative phosphorylation and DNA damage repair pathways, and an increased signature of cellular senescence. By constructing cell type-specific regulatory networks, we uncover enhanced activity of the transcription factor CEBPD across cell types in the aged ovary, with a corresponding significant loss of activity of most cell identity-associated transcription factors. Moreover, by performing integrative analyses of our single-nuclei multi-omics data with common genetic variants associated with age at natural menopause (ANM) from genome-wide association studies, we demonstrate a global impact of functional variants on changes in gene regulatory networks across ovarian cell types. Finally, we nominate about a dozen of functional non-coding variants, their target genes and cell types and regulatory mechanisms that underlie genetic association with ANM. This work provides a comprehensive multimodal landscape of human ovarian ageing and mechanistic insights into inherited variation of ANM.

Main

The ovary is the primary female reproductive organ, and the first tissue to undergo profound age-associated loss of function in humans, characterized

by a progressive decline in follicle number and oocyte quality¹. The rate of follicular depletion increases throughout reproductive life, but begins a more accelerated decline around age 37⁹. This results in a higher risk of both infertility, and aneuploidy and congenital disabilities in offspring². There is also overwhelming evidence that female reproductive ageing influences lifespan and diverse health outcomes^{3-8,10,11}. Consequently, an in-depth understanding of ovarian ageing can benefit not only reproduction but also longevity and health in women. However, thus far very little is known about basic biological mechanisms that underlie human ovarian ageing.

Menopause is the time marking the cessation of menstrual cycling and production of fertile oocytes, and age at natural menopause (ANM) has profound implications for health and disease risk in women³⁻⁸. Family and twin studies have demonstrated a strong relationship between genetics and ANM¹²⁻¹⁵, suggesting up to a ~6-fold increase in risk of early menopause for a woman with a family history of early menopause^{12,15}. Identification of the genes contributing to ANM will provide mechanistic insights into the biological processes underlying ovarian ageing. Genome-wide association studies (GWAS) have identified hundreds of genetic loci associated with ANM¹⁶. However, the great majority (~94%) of the risk variants reside in non-coding regions of the genome, making it difficult to assign their functional role in ovarian ageing

Many recent studies show that functional non-coding GWAS variants are significantly enriched in cell type-specific transcriptional regulatory elements such as enhancers¹⁷⁻²⁵. Enhancers have emerged as major points of integration of intra- and extracellular signals associated with development, homeostasis, and disease, resulting in context-specific transcriptional outputs²⁶. Cell-specific enhancer activation is driven by

combinatorial actions of lineage-determining and signal-dependent transcription factors (TFs)²⁷. Genetic variation affecting enhancer selection and function is considered a major determinant of differences in cell-specific gene expression and disease risk between individuals²⁷. Therefore, the identification of functional ANM-associated non-coding regulatory variants, as well as the target genes and cell types through which they confer their effects on ANM, is a powerful way to understand the biological processes underlying ovarian ageing. However, we currently lack an atlas of the transcriptional regulatory elements that are active in every cell type in the ovary during ageing.

In this study, we systematically characterize human ovarian ageing by performing single-nuclei multi-omics analysis and by superimposing these data with ANM-associated GWAS risk variants. Through these efforts, we identify the functional transcriptional regulatory elements, and functional non-coding variants and their target genes associated with ANM, across all cell types in the ovary.

Single nucleus multi-omics profiling ageing

We performed single-nuclei RNA-seq (snRNA-seq) and single-nuclei assay for transposase-accessible chromatin using sequencing (snATAC-seq) on the same flash-frozen human ovarian tissues, which were from young (n=4; ages 23-29 years) and reproductively old (n=4; ages 49-54 years) autopsy samples of sudden death with normal ovarian histology (Fig. 1a and Supplementary Table 1). After stringent quality control, we retained 42,568 nuclei for snRNA-seq and 41,550 nuclei for snATAC-seq (Methods). Seurat-based unsupervised clustering²⁸ and Harmony-based²⁹ batch correction on snRNA-seq revealed eight distinct clusters (Methods,

Fig. 1b, and Supplementary Fig. 1a). All major somatic cell types in the ovary, including stromal cells (SC), endothelial cells (blood vessel endothelial cells (BEC) and lymphatic endothelial cells (LEC)), granulosa cells (GC), smooth muscle cells (SMC), immune cells (IC), epithelial cells (EpiC), and theca cells (TC) were identified based on well-defined cell type-specific markers (Figs. 1b, c). For snATAC-seq, Signac-based unsupervised clustering³⁰ and Harmony-based batch correction revealed seven distinct clusters (Methods, Fig. 1d, and Supplementary Fig. 1b). To annotate the clusters, we used canonical correlation analysis (CCA) and mutual nearest neighbors (MNNs)³¹ to transfer the cell type labels from snRNA-seq to snATAC-seq (Methods). Consistently, all major cell types were also present in snATAC-seq (Fig. 1d and Supplementary Fig. 1c). Additionally, we confirmed cell type identities by examining chromatin accessibility at the promoter regions of known markers and calculating a gene activity score that quantified chromatin accessibility within the gene body and promoter regions (Fig. 1e and Supplementary Fig. 1d). Furthermore, we identified cell type-specific differentially expressed genes (DEGs) and differentially accessible chromatin regions (DARs) for each cell type (Supplementary Table 2 and 3), and found that the cell types can be well-distinguished by those DEGs and DARs (Supplementary Figs. 1e,f).

Altered cell type composition with age

To investigate the dynamic changes in cell type composition during human ovarian ageing, we compared the cell type proportions of aged and young ovaries in the snRNA-seq data. We found significant changes in the proportions of several cell types during ageing (Fig. 2a). For example, the abundance of granulosa and theca cells, two critical components of ovarian

follicles, were significantly decreased in aged compared to young ovaries (Fig. 2a), in line with the well-known phenomenon of decreasing follicle number with increasing age³². In addition, blood vessel and lymphatic endothelial cells, the cell layers lining the blood and lymph vessels, respectively, also markedly decreased in proportion in aged ovaries (Fig. 2a), consistent with the observed negative correlation between ovarian vascularity and age³³. Interestingly, epithelial cells were the only cell type that increased in proportion in aged ovaries (Fig. 2a), potentially reflecting the lifetime of ovulation-induced rupture and repair experienced by aged ovaries³⁴. Consistently, aged epithelial cells exhibited significantly elevated expression of cell cycle-associated genes, while aged granulosa cells, theca cells, and endothelial cells exhibited increased expression of apoptosis-associated genes and/or decreased expression of cell cycle-associated genes compared to young counterparts in the ovary (Supplementary Fig. 2a). In agreement with the snRNA-seq results, we observed almost identical age-related changes in the cellular composition estimated from the snATAC-seq data (Fig. 2b). These results indicated that ageing significantly remodels the cellular architecture of the human ovary.

Coordinated changes in ageing hallmarks

To investigate the dynamic changes in gene expression during human ovarian ageing, we identified ageing-associated DEGs for each cell type (Methods). In total, we identified 3,341 ageing-associated DEGs (Supplementary Table 4), the number of which ranged from a few hundred to several thousand, depending on cell type (Supplementary Fig. 2b). Specifically, granulosa cells have the largest number of ageing-associated DEGs (n=2,255) (Supplementary Fig. 2b), suggesting that granulosa cells are more vulnerable to ageing than other cell types in the human ovary.

Interestingly, we found that most ageing-associated DEGs are shared among cell types (Supplementary Fig. 2c) and show congruent changes in expression (Fig. 2c). Among the “common DEGs” which were significantly differentially expressed in at least four cell types (Supplementary Fig. 2d), 218 genes were up-regulated and 182 were down-regulated in the aged ovary (Supplementary Fig. 2d). Common DEGs include those reported in the GenAge database as human ageing-related genes, such as *RICTOR*, *IGF1R*, *MAP3K5*, and *APOE* (Supplementary Figs. 2e,f). Gene ontology (GO) analysis³⁵ indicated that common DEGs were enriched in the “hallmarks of ageing”³⁶, including pathways involved in nutrient sensing signaling, cellular senescence, proteostasis, cellular communication, and mitochondrial function (Supplementary Fig. 2g). We also found cell-type-specific ageing-associated DEGs (Supplementary Fig. 2h), such as those enriched in cell type-relevant functions, including vasculogenesis for blood vessel endothelial cells, follicle development for granulosa cells, and smooth muscle contraction for smooth muscle cells (Supplementary Fig. 2i).

To gain insight into the dynamic changes in biological pathways during human ovarian ageing, we used Gene Set Variation Analysis (GSVA)³⁷ to estimate the pathway activity score for individual cells and compare the pathway activity between young and aged ovaries in each cell type (Methods). We found that half (92/186) of KEGG pathways were significantly up- or down-regulated in at least six cell types in aged ovaries (Supplementary Fig. 3a). Genes involved in these pathways showed congruent changes in expression direction across cell types (Supplementary Fig. 3a). Notably, expression of genes involved in the nutrient-sensing signaling pathways, including the mTOR, insulin, and MAPK pathways, increased in the aged ovaries across cell types, while

those involved in oxidative phosphorylation and base excision repair decreased (Fig. 2d and Supplementary Figs. 3b-e). To validate the age-related changes in gene expression, we performed *in situ* hybridization assays and confirmed the increased expression of the mTOR signaling gene *RICTOR*, and decreased expression of the oxidative phosphorylation gene *MT-ATP6* in aged ovaries *in vivo* (Figs. 2e-h). Together, these results indicate that the human ovary undergoes coordinated transcriptomic changes during ageing, resulting in profound alterations to processes central to the biology of ageing.

Cellular senescence in the human ovary

Senescent cell burden increases with age in various tissues in the context of physiological ageing and ageing-related disease³⁸⁻⁴¹. To test if cellular senescence increases during human ovarian ageing, we examined the expression of the widely used senescence markers, *CDKN1A* (p21) and *CDKN2A* (p16), in the human ovary. On average, very few (~0.43%) ovarian cells expressed *CDKN2A*, while a considerable proportion of young ovarian cells expressed *CDKN1A* (~9.49%) and a significantly higher proportion of *CDKN1A*⁺ cells (~15.56%) was observed in aged ovaries (Supplementary Figs. 4a,b). We then calculated the proportion of *CDKN1A*⁺ cells for each cell type in young and aged ovaries. We found a significant increase in the proportion of *CDKN1A*⁺ cells with age in stromal, granulosa, theca, blood vessel endothelial, and smooth muscle cells (Fig. 3a). Using *in situ* hybridization, we found a ~3-fold increase in both the proportion of cells expressing *CDKN1A*, and in the average expression of *CDKN1A*, in aged compared to young ovaries (Figs. 3b,c). In addition, a subset of senescence-associated secretory phenotype (SASP) genes were up-regulated in *CDKN1A*⁺ cells in the human ovary (Fig. 3d). To gain

insight into the transcriptional signatures of *CDKN1A*⁺ cells, we identified the DEGs in *CDKN1A*⁺ cells compared to *CDKN1A*⁻ cells from both young and aged stromal cells (Supplementary Fig. 4c). GO analysis indicated that genes involved in “response to oxygen level” and “HIF-1 signaling pathway” are up-regulated in *CDKN1A*⁺ stromal cells (Supplementary Fig. 4c). We then compared the transcriptomes of *CDKN1A*⁺ cells between young and aged stromal cells and found that genes involved in the HIF-1 pathway as well as those in the nutrient-sensing signaling were up-regulated in aged- compared to young *CDKN1A*⁺ stromal cells (Supplementary Fig. 4c). To test whether the up-regulation of the HIF-1 pathway is a senescence signature in the ovary, we computed HIF-1 pathway scores based on the HIF-1 pathway-related DEGs that were up-regulated in *CDKN1A*⁺ cells (Supplementary Table 5), including the key HIF-1 target genes that regulate NAD⁺ metabolism (*NAMPT*)⁴², cellular respiration (*PDK1*)⁴³, and apoptosis (*DDIT4*)⁴⁴ in response to hypoxia (Supplementary Fig. 4d). We found that the HIF-1 pathway was significantly enhanced in *CDKN1A*⁺ cells and further elevated during ageing in most cell types (Fig. 3e). Given the reduced vasculature of aged ovaries³³, our results suggested that the hypoxic environment might be a critical factor in driving cellular senescence in human ovaries, through upregulation of HIF-1 signaling.

Ageing alters cellular communication

Altered intercellular communication is a hallmark of ageing³⁶. To explore potential age-related alterations to the ovarian cellular communication network, we used CellChat⁴⁵, which models the probability of the cell-cell interaction network based on gene expression and prior knowledge of ligand-receptor interactions. To gain insight into the intercellular

communication between ovarian somatic cells and oocytes, we integrated our snRNA-seq with publicly available human oocyte single-cell transcriptomes from young and reproductively aged ovaries⁴⁶. We found that ageing slightly reduced both the total number and overall strength of intercellular interactions (Supplementary Fig. 5a,b). Strikingly, aged granulosa and theca cells received far fewer signals from all other cell types, and the same trend was observed for blood/lymphatic endothelial cells and immune cells (Fig. 3f). Of particular relevance to fertility, the number of signals received by oocytes from granulosa and theca cells profoundly decreased with age (Fig. 3f), consistent with evidence that the functions of granulosa and theca cells in supporting oocytes fail during ageing⁴⁷. In contrast, the interaction number and strength from all cell types to epithelial cells increased (Fig. 3f and Supplementary Fig. 5c). We further identified the incoming and outgoing signaling pathways that exhibited a significant difference in communication probability between young and aged ovaries for each cell type (Fig. 3g). In total, we identified 46 pathways with significant differential communication probability with age (Fig. 3g). Notably, the COLLAGEN and FN1 (Fibronectin) pathways, core components of extracellular matrix (ECM) biology, exhibited a significantly higher communication probability in most cell types in the young ovary, suggesting an essential role of the ECM in maintaining ovary function. Interestingly, epithelial cells were the only cell type that exhibited a significantly higher communication probability of COLLAGEN and FN1 signaling in the aged ovary (Fig. 3g). COLLAGEN and FN1 signaling is known to promote the proliferation of epithelial cells^{48,49}, suggesting that activation of these pathways may explain the increased proportion of epithelial cells we observed in the aged ovary (Figs. 2a,b). In contrast, the JAM (Junction adhesion molecule), PARs (Protease-activated receptors), and NCAM (Neural cell adhesion molecule) pathways, which mediate cell-

cell adhesion processes, exhibited a significantly higher communication probability in most cell types in the aged ovary (Fig. 3g). The PDGF pathway, which has known roles in fibroinflammatory processes, was significantly enriched in all the cell types in the aged ovary (Fig. 3g), consistent with the finding of elevated fibroinflammatory cytokines in aged human ovarian follicular fluid⁵⁰. In contrast to the age-related loss of interactions between granulosa and theca cells and oocytes (Fig. 3f), several known signaling pathways that are critical to maintaining the function of oocytes and granulosa cells were specifically enriched in young oocytes and granulosa cells, including the FSH and GDF pathways that maintain follicle growth and function⁵¹ (Fig. 3g). Our results demonstrate that human ovarian ageing is characterized by significant changes in cellular communications among oocytes and somatic cell types, potentially contributing to an age-related loss of follicular function, tissue fibrosis, and epithelial hyperplasia.

Ageing alters cell identity TF networks

Master transcription factors (TFs) largely determine cell identity⁵². As loss of cell identity with age has been implicated in age-related tissue dysfunction, we next identified cell identity-associated TFs, and investigated for age-related changes in their activity (Methods) in the snATAC-seq data. As expected, individual cell types can be distinguished by predicted motif activity (Fig. 4a and Supplementary Table 6). The motifs of folliculogenesis-related TFs, mainly AP-1 and RUNX transcription factors⁵³, were predominantly enriched in granulosa cells (Fig. 4b and Supplementary Fig. 6a). Steroidogenesis-related TFs^{54,55} were mainly enriched in granulosa cells and theca cells (Supplementary Fig. 6b). The TF footprinting analysis highlighted cell type-specific enrichment of

those TFs in granulosa cells and theca cells (Supplementary Fig. 6c). ETS TFs are central regulators of endothelial and immune cells^{56,57}, both of which originate from the hemogenic endothelium during embryogenesis⁵⁸. Consistently, a family of ETS TFs were enriched in endothelial and immune cells (Fig. 4b and Supplementary Fig. 6d). Additionally, we identified the TFs enriched in epithelial cells and stromal cells, respectively (Fig. 4b and Supplementary Fig. 6e).

To reveal the TFs that govern human ovarian ageing, we compared the predicted motif activity between young and aged cells in each cell type (Supplementary Table 7). Surprisingly, CCAAT/enhancer binding proteins (C/EBPs) motif activities significantly increased in most cell types except immune cells and epithelial cells (Fig. 4c). Among the members of this TF family, *CEBPD* was highly expressed across cell types, and its age-related changes in expression were in line with the changes in motif activity during ovarian ageing (Fig. 4d). In addition, most cell identity-associated TFs exhibited significantly decreased motif activity, while epithelial cell identity-associated TFs exhibited significantly enhanced motif activity during ovarian ageing (Fig. 4c). We further calculated cell identity scores in each cell type by examining the expression level of the top 100 cell type-specific genes (Methods). We found that cells in young ovaries exhibited high expression of their corresponding cell type-specific genes and minimal expression of other cell type-specific genes (Supplementary Fig. 6f). Surprisingly, granulosa cells, immune cells, and theca cells in aged ovaries expressed deficient levels of their corresponding cell type-specific genes, and instead expressed high levels of stromal cell-specific genes (Supplementary Fig. 6f). These results suggest a prevalent loss of cell identity in aged ovaries.

We next sought to build cell type-specific TF regulatory networks for

human ovarian ageing, and first constructed the cis co-accessibility networks (CCANs) in each cell type using Cicero⁵⁹. Next, we defined the putative enhancers and promoters by overlaying the CCAN peaks with human ovary tissue enhancer and promoter annotations from the ENCODE database⁶⁰. Finally, we defined a gene as a CCAN-linked gene if one of the CCAN peaks lies in its putative promoter (Supplementary Fig. 7a). In this way, we identified varying numbers of CCANs and CCAN-linked genes for each cell type (Supplementary Fig. 7b). Most cell type-associated DEGs and ageing-associated DEGs significantly overlapped with CCAN-linked genes in all cell types (Supplementary Fig. 7c), suggesting that CCANs play essential roles in determining cell identity and the regulation of ovarian ageing. For each cell type, we built the ageing-associated TF regulatory networks governed by the top TFs (n=3~4) that change with age, as defined by the predicted motif activity. Those ageing-associated DEGs whose promoters or putative enhancers contained both accessible peaks and motifs of the top ageing-associated TFs within the peaks were defined as candidate targets of the selected TFs (Supplementary Fig. 7a). We generated the ovarian ageing-associated TF regulatory network for each cell type and found a critical role for CEBPD in human ovarian ageing (Figs. 4e-k). We found that CEBPD target genes are enriched in processes of known importance to the basic biology of ageing, including mTOR signaling, MAPK signaling, and cellular senescence, in multiple cell types (Supplementary Fig. 7d).

Cellular targets of ANM genetic risk

The most comprehensive recent ANM GWAS identified 290 ANM-associated genetic risk loci¹⁶. Using gene expression data from several publicly-available datasets, the Ruth et al. study implicated hematopoietic

stem and progenitor cells as the major cellular targets of ANM-associated risk variants¹⁶. Since the Ruth et al. analysis did not include single-cell data from the human ovary, we next investigated whether any specific cell types in the human ovary were enriched for ANM-associated variants, by performing MAGMA^{61,62} analysis using our snRNA-seq dataset. For comparison, we also included GWAS data from 2 other ovarian phenotypes, ovarian epithelial cancer (OEC) and polycystic ovary syndrome (PCOS). We found that ANM-associated variants were significantly enriched in almost all cell types (Fig. 5a), indicating a systemic effect of ageing on the ovary. In contrast, OEC- and PCOS-associated variants were enriched in epithelial cells and granulosa cells, respectively (Fig. 5a). To investigate if any cell type-specific regions of chromatin accessibility were enriched for ANM-associated variants, we also performed cell type-specific linkage disequilibrium (LD) score regression⁶³ using our snATAC-seq dataset. Consistently, ANM-associated variants showed a significant enrichment in multiple cell types (Fig. 5b). Together with the coordinated changes in transcriptomes and chromatin accessibility we observed across cell types, the results from our analyses of ANM genetic signal suggest that all ovarian cell types contribute to ovarian ageing.

Functional ANM variants and genes nominated

To gain insights into how ANM-associated variants contribute to ovarian ageing, we performed post-GWAS analyses to identify functional regulatory variants and affected target genes. Using 290 ANM-associated GWAS lead variants from Ruth et al.¹⁶, we first compiled a comprehensive set of coinherited variants based on LD (R^2 value ≥ 0.8) calculated from phase 1 genotypes of individuals of European ancestry in the 1000

Genomes dataset (Methods). In total, we identified 5,555 ANM-associated variants (Supplementary Table 8). To identify the functional variants that may affect transcriptional regulatory activity in each cell type, we first overlapped ANM-associated variants with the putative enhancers and promoters we identified in each cell type. In this way, we identified 101 candidate functional variants (Supplementary Fig. 8 and Supplementary Table 9) and found that a substantial number of these variants were shared across several cell types (Fig. 5c). Next, we focused on the putative functional variants that were shared in at least four cell types, among which 6 variants occurred in DNA damage response (DDR)-related gene loci (Fig. 5d and Supplementary Figs. 9a-e). The DDR is the major pathway linked to ovarian ageing as detected by GWAS of ANM¹⁶ and our results highlight the functional role of ANM-associated variants on the regulation of DDR across major ovarian cell types. For example, the rs3741605 (T>C) allele, associated with delayed ANM (BETA>0; Supplementary Table 10), occurs in the putative promoter of the *HELB* gene, encoding DNA helicase B⁶⁴, that is active in most cell types in the ovary, including stromal, endothelial, theca, granulosa, and immune cells (Fig. 5d). Previously, multiple missense variants predicted to be deleterious have been identified in *HELB* that are associated with early ANM¹⁶. Remarkably, expression quantitative trait loci (eQTL) analysis from the GTEx database⁶⁵ indicated that the C allele of rs3741605 was significantly correlated with increased expression of *HELB* in the human ovary (Fig. 5f). This result suggests that the functional non-coding variant rs3741605 may contribute to delayed ANM by upregulating the expression of a critical DNA repair gene, thereby conferring improved genome maintenance.

To explore the potential mechanisms underlying the influence of variants on gene expression, we predicted the effect of candidate functional variants

on TF binding activity by applying gapped k-mer support vector machine-based methods (LS-GKM⁶⁶ and deltaSVM⁶⁷) (Methods). We found that the delayed ANM-associated C allele of rs3741605 could enhance TF binding activity (Fig. 5e), showing a high concordance of predicted beneficial allelic effect on the increased expression of *HELB* and delayed ANM through enhanced genome maintenance.

We also identified functional regulatory variants with predicted deleterious effects on ANM. The variant rs13263296 is located in the *DEPTOR* locus, a key gene involved in mTOR signaling (Fig. 5g and Supplementary Fig. 10a), and the T allele of rs13263296 is associated with early ANM (BETA<0; Supplementary Table 10) and occurs in the putative transcription start site (TSS)-proximal enhancer of the *DEPTOR* gene in most cell types (Fig. 5g). We observed significant up-regulation of *DEPTOR* expression in aged granulosa cells and epithelial cells (Supplementary Fig. 10b). Of note, rs13263296 (C>T) was not significantly associated with increased *DEPTOR* expression in the human ovary (Fig. 5i), perhaps due to the use of bulk tissue in the eQTL analysis. Mechanistically, the deltaSVM analysis suggested that rs13263296 (C>T) may affect *DEPTOR* expression by enhancing TF binding activity (Fig. 5h). In addition to these loci, we identified several functional regulatory variants located in oxidative phosphorylation and MAPK signaling-related gene loci (Supplementary Figs. 10c,d). Taken together, our integrated analyses revealed the global effects of ANM-associated non-coding variants on gene expression across ovarian cell types and nominated functional regulatory ANM risk variants that may dysregulate genes involved in pathways of relevance to the canonical hallmarks of ageing, such as mTOR and genome maintenance.

In summary, our single nuclei multi-omic analysis of young and

444 reproductively aged ovaries provides high-resolution characterization of
 445 the transcriptional regulatory landscape at the single-cell level, uncovering
 446 conserved mechanisms of ageing biology in action, such as the
 447 hyperactivity of mTOR observed across all ovarian somatic cell types with
 448 age. Our results raise the hope that geroprotectors targeting the basic
 449 biology of ageing, such as mTOR signaling, may be used to delay
 450 reproductive ageing in women. Furthermore, our integrative post-GWAS
 451 analyses of ANM provides biological insights into the role of inherited
 452 non-coding variants in ovarian ageing in humans, nominating new
 453 functional variants for follow-up interrogation. These findings expand our
 454 understanding of inherited variation in ANM and provide a roadmap for
 455 the functional dissection of the non-coding genetic variation influencing
 456 ANM, pointing towards the nomination of new therapeutic targets for
 457 reproductive health in women.

458 References

- 459 1 Broekmans, F., Soules, M. & Fauser, B. Ovarian aging: mechanisms and clinical
460 consequences. *Endocrine reviews* **30**, 465-493 (2009).
- 461 2 Herbert, M., Kalleas, D., Cooney, D., Lamb, M. & Lister, L. Meiosis and maternal aging:
462 insights from aneuploid oocytes and trisomy births. *Cold Spring Harbor perspectives*
463 *in biology* **7**, a017970 (2015).
- 464 3 Muka, T. *et al.* Association of age at onset of menopause and time since onset of
465 menopause with cardiovascular outcomes, intermediate vascular traits, and all-cause
466 mortality: a systematic review and meta-analysis. *JAMA cardiology* **1**, 767-776 (2016).
- 467 4 Rahman, I., Åkesson, A. & Wolk, A. Relationship between age at natural menopause
468 and risk of heart failure. *Menopause* **22**, 12-16 (2015).
- 469 5 Sullivan, S. D. *et al.* Effect of age of self-reported, non-surgical menopause on time to
470 first fracture and bone mineral density in the women's health initiative observational
471 study. *Menopause (New York, NY)* **22**, 1035 (2015).
- 472 6 Shen, L. *et al.* Association between earlier age at natural menopause and risk of
473 diabetes in middle-aged and older Chinese women: The Dongfeng-Tongji cohort
474 study. *Diabetes & metabolism* **43**, 345-350 (2017).
- 475 7 Green, J. *et al.* Reproductive factors and risk of oesophageal and gastric cancer in the
476 Million Women Study cohort. *British journal of cancer* **106**, 210-216 (2012).
- 477 8 Cancer, C. G. o. H. F. i. B. Menarche, menopause, and breast cancer risk: individual
478 participant meta-analysis, including 118 964 women with breast cancer from 117
479 epidemiological studies. *The lancet oncology* **13**, 1141-1151 (2012).
- 480 9 Faddy, M. & Gosden, R. Ovary and ovulation: a model conforming the decline in follicle
481 numbers to the age of menopause in women. *Human reproduction* **11**, 1484-1486
482 (1996).
- 483 10 Habermehl, T. L. *et al.* Extension of longevity and reduction of inflammation is ovarian-
484 dependent, but germ cell-independent in post-reproductive female mice.
485 *GeroScience* **41**, 25-38 (2019).
- 486 11 Cargill, S. L., Carey, J. R., Müller, H. G. & Anderson, G. Age of ovary determines
487 remaining life expectancy in old ovariectomized mice. *Aging cell* **2**, 185-190 (2003).
- 488 12 Cramer, D. W., Xu, H. & Harlow, B. L. Family history as a predictor of early menopause.
489 *Fertility and sterility* **64**, 740-745 (1995).
- 490 13 Torgerson, D. J., Thomas, R. E. & Reid, D. M. Mothers and daughters menopausal ages:
491 is there a link? *European Journal of Obstetrics & Gynecology and Reproductive Biology*
492 **74**, 63-66 (1997).
- 493 14 Murabito, J. M., Yang, Q., Fox, C., Wilson, P. W. & Cupples, L. A. Heritability of age at
494 natural menopause in the Framingham Heart Study. *The Journal of Clinical*
495 *Endocrinology & Metabolism* **90**, 3427-3430 (2005).
- 496 15 Morris, D. H., Jones, M. E., Schoemaker, M. J., Ashworth, A. & Swerdlow, A. J. Familial
497 concordance for age at natural menopause: results from the Breakthrough
498 Generations Study. *Menopause* **18**, 956-961 (2011).
- 499 16 Ruth, K. S. *et al.* Genetic insights into biological mechanisms governing human ovarian
500 ageing. *Nature* **596**, 393-397, doi:10.1038/s41586-021-03779-7 (2021).
- 501 17 Corradin, O. *et al.* Combinatorial effects of multiple enhancer variants in linkage
502 disequilibrium dictate levels of gene expression to confer susceptibility to common
503 traits. *Genome research* **24**, 1-13, doi:10.1101/gr.164079.113 (2014).
- 504 18 Trynka, G. *et al.* Chromatin marks identify critical cell types for fine mapping complex
505 trait variants. *Nature genetics* **45**, 124-130, doi:10.1038/ng.2504 (2013).
- 506 19 Hnisz, D. *et al.* Super-enhancers in the control of cell identity and disease. *Cell* **155**,
507 934-947, doi:10.1016/j.cell.2013.09.053 (2013).

508 20 Maurano, M. T. *et al.* Systematic localization of common disease-associated variation
509 in regulatory DNA. *Science (New York, N.Y.)* **337**, 1190-1195,
510 doi:10.1126/science.1222794 (2012).

511 21 Altshuler, D., Daly, M. J. & Lander, E. S. Genetic mapping in human disease. *Science*
512 *(New York, N.Y.)* **322**, 881-888, doi:10.1126/science.1156409 (2008).

513 22 Musunuru, K. *et al.* From noncoding variant to phenotype via SORT1 at the 1p13
514 cholesterol locus. *Nature* **466**, 714-719, doi:10.1038/nature09266 (2010).

515 23 Stitzel, M. L. *et al.* Global epigenomic analysis of primary human pancreatic islets
516 provides insights into type 2 diabetes susceptibility loci. *Cell Metab* **12**, 443-455,
517 doi:10.1016/j.cmet.2010.09.012 (2010).

518 24 Harismendy, O. *et al.* 9p21 DNA variants associated with coronary artery disease
519 impair interferon- γ signalling response. *Nature* **470**, 264-268,
520 doi:10.1038/nature09753 (2011).

521 25 Lee, J. C. *et al.* Human SNP links differential outcomes in inflammatory and infectious
522 disease to a FOXO3-regulated pathway. *Cell* **155**, 57-69,
523 doi:10.1016/j.cell.2013.08.034 (2013).

524 26 Levine, M. Transcriptional enhancers in animal development and evolution. *Current*
525 *biology : CB* **20**, R754-763, doi:10.1016/j.cub.2010.06.070 (2010).

526 27 Heinz, S., Romanoski, C. E., Benner, C. & Glass, C. K. The selection and function of cell
527 type-specific enhancers. *Nat Rev Mol Cell Biol* **16**, 144-154, doi:10.1038/nrm3949
528 (2015).

529 28 Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* (2021).

530 29 Korsunsky, I. *et al.* Fast, sensitive and accurate integration of single-cell data with
531 Harmony. *Nature methods* **16**, 1289-1296 (2019).

532 30 Stuart, T., Srivastava, A., Madad, S., Lareau, C. A. & Satija, R. Single-cell chromatin
533 state analysis with Signac. *Nature methods*, 1-9 (2021).

534 31 Stuart, T. *et al.* Comprehensive integration of single-cell data. *Cell* **177**, 1888-1902.
535 e1821 (2019).

536 32 Hansen, K. R. *et al.* A new model of reproductive aging: the decline in ovarian non-
537 growing follicle number from birth to menopause. *Human reproduction* **23**, 699-708
538 (2008).

539 33 Costello, M. F. *et al.* Power doppler ultrasound assessment of the relationship
540 between age and ovarian perfollicular blood flow in women undergoing in vitro
541 fertilization treatment. *Journal of assisted reproduction and genetics* **23**, 359-365
542 (2006).

543 34 Fathalla, M. F. Incessant ovulation--a factor in ovarian neoplasia? *Lancet (London,*
544 *England)* **2**, 163, doi:10.1016/s0140-6736(71)92335-x (1971).

545 35 Zhou, Y. *et al.* Metascape provides a biologist-oriented resource for the analysis of
546 systems-level datasets. *Nature communications* **10**, 1-10 (2019).

547 36 López-Otín, C., Blasco, M. A., Partridge, L., Serrano, M. & Kroemer, G. The hallmarks
548 of aging. *Cell* **153**, 1194-1217 (2013).

549 37 Hänzelmann, S., Castelo, R. & Guinney, J. GSVA: gene set variation analysis for
550 microarray and RNA-seq data. *BMC bioinformatics* **14**, 1-15 (2013).

551 38 Wang, C. *et al.* DNA damage response and cellular senescence in tissues of aging mice.
552 *Aging cell* **8**, 311-323 (2009).

553 39 Krishnamurthy, J. *et al.* Ink4a/Arf expression is a biomarker of aging. *The Journal of*
554 *clinical investigation* **114**, 1299-1307 (2004).

555 40 Sis, B. *et al.* Accelerated expression of senescence associated cell cycle inhibitor
556 p16INK4A in kidneys with glomerular disease. *Kidney international* **71**, 218-226 (2007).

557 41 Chinta, S. J. *et al.* Cellular senescence is induced by the environmental neurotoxin
558 paraquat and contributes to neuropathology linked to Parkinson's disease. *Cell*

559 *reports* **22**, 930-940 (2018).

560 42 Bae, S.-K. *et al.* Hypoxic induction of human visfatin gene is directly mediated by
561 hypoxia-inducible factor-1. *FEBS letters* **580**, 4105-4113 (2006).

562 43 Kim, J.-w., Tchernyshyov, I., Semenza, G. L. & Dang, C. V. HIF-1-mediated expression
563 of pyruvate dehydrogenase kinase: a metabolic switch required for cellular adaptation
564 to hypoxia. *Cell metabolism* **3**, 177-185 (2006).

565 44 Shoshani, T. *et al.* Identification of a novel hypoxia-inducible factor 1-responsive gene,
566 RTP801, involved in apoptosis. *Molecular and cellular biology* **22**, 2283-2293 (2002).

567 45 Jin, S. *et al.* Inference and analysis of cell-cell communication using CellChat. *Nature*
568 *communications* **12**, 1-20 (2021).

569 46 Yuan, L. *et al.* Single - cell transcriptome analysis of human oocyte ageing. *Journal of*
570 *Cellular and Molecular Medicine* (2021).

571 47 Alberico, H. C. & Woods, D. C. Role of Granulosa Cells in the Aging Ovarian Landscape:
572 A Focus on Mitochondrial and Metabolic Function. *Front Physiol* **12**, 800739,
573 doi:10.3389/fphys.2021.800739 (2021).

574 48 Katsuno-Kambe, H., Teo, J. L., Ju, R. J., Hudson, J. & Stehbens, S. J. Collagen
575 polarization promotes epithelial elongation by stimulating locoregional cell
576 proliferation. *Elife* **10**, e67915 (2021).

577 49 Han, S. & Roman, J. Fibronectin induces cell proliferation and inhibits apoptosis in
578 human bronchial epithelial cells: pro-oncogenic effects mediated by PI3-kinase and
579 NF- κ B. *Oncogene* **25**, 4341-4349 (2006).

580 50 Machlin, J. H. *et al.* Fibroinflammatory Signatures Increase with Age in the Human
581 Ovary and Follicular Fluid. *International journal of molecular sciences* **22**, 4902 (2021).

582 51 Orisaka, M. *et al.* The role of pituitary gonadotropins and intraovarian regulators in
583 follicle development: A mini - review. *Reproductive Medicine and Biology* **20**, 169-175
584 (2021).

585 52 Whyte, W. A. *et al.* Master transcription factors and mediator establish super-
586 enhancers at key cell identity genes. *Cell* **153**, 307-319 (2013).

587 53 Pangas, S. A. & Rajkovic, A. Transcriptional regulation of early oogenesis: in search of
588 masters. *Human Reproduction Update* **12**, 65-76 (2006).

589 54 Parker, K. L. & Schimmer, B. P. Steroidogenic factor 1: a key determinant of endocrine
590 development and function. *Endocrine reviews* **18**, 361-377 (1997).

591 55 Felizola, S. J. *et al.* Estrogen-related receptor α in normal adrenal cortex and
592 adrenocortical tumors: involvement in development and oncogenesis. *Molecular and*
593 *cellular endocrinology* **365**, 207-211 (2013).

594 56 De Val, S. & Black, B. L. Transcriptional control of endothelial cell development.
595 *Developmental cell* **16**, 180-195 (2009).

596 57 Gallant, S. & Gilkeson, G. ETS transcription factors and regulation of immunity.
597 *Archivum immunologiae et therapias experimentalis* **54**, 149-163 (2006).

598 58 Hirschi, K. K. Hemogenic endothelium during development and beyond. *Blood* **119**,
599 4823-4827 (2012).

600 59 Pliner, H. A. *et al.* Cicero predicts cis-regulatory DNA interactions from single-cell
601 chromatin accessibility data. *Molecular cell* **71**, 858-871. e858 (2018).

602 60 Moore, J. E. *et al.* Expanded encyclopaedias of DNA elements in the human and mouse
603 genomes. *Nature* **583**, 699-710 (2020).

604 61 de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: generalized gene-
605 set analysis of GWAS data. *PLoS computational biology* **11**, e1004219 (2015).

606 62 Skene, N. G. *et al.* Genetic identification of brain cell types underlying schizophrenia.
607 *Nature genetics* **50**, 825-833 (2018).

608 63 Finucane, H. K. *et al.* Partitioning heritability by functional annotation using genome-
609 wide association summary statistics. *Nature genetics* **47**, 1228-1235 (2015).

610 64 Hazeslip, L., Zafar, M. K., Chauhan, M. Z. & Byrd, A. K. Genome Maintenance by DNA
611 Helicase B. *Genes* **11**, 578 (2020).
612 65 Consortium, G. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue
613 gene regulation in humans. *Science* **348**, 648-660 (2015).
614 66 Lee, D. LS-GKM: a new gkm-SVM for large-scale datasets. *Bioinformatics* **32**, 2196-
615 2198 (2016).
616 67 Lee, D. *et al.* A method to predict the impact of regulatory variants from DNA
617 sequence. *Nature genetics* **47**, 955-961 (2015).

618

619

620

621

622

623

624

625

626

627

628

629

630

631

632

633

Methods

Sample procurement

Fresh-frozen healthy human ovary samples were purchased from BioIVT (Baltimore, MD) and Cureline (Brisbane, CA). All samples were de-identified and the tissue source was anonymous to the researcher.

Nuclear dissociation and library preparation

For snATAC-seq, nuclei isolation was performed according to the 10× Genomics protocol CG000212 (Rev B) with ~100 mg frozen human ovary sample. Libraries were generated by using 10x Chromium Single Cell ATAC Reagent Kits (v1) and sequenced using the Illumina NextSeq 550 platform with 150-bp paired-end sequencing.

For snRNA-seq, the nuclei isolation was performed according to 10× Genomics protocol CG000393 (Rev A) with ~100 mg frozen human ovary sample. Libraries were generated using 10x Chromium Single Cell 3' Reagent Kits (v3) and sequenced using the Illumina NextSeq 550 platform with 150-bp paired-end sequencing.

Processing of snRNA-seq data

Reads were aligned to a pre-mRNA GTF built on the GRCh38 genome using Cellranger (v3.1.0) to account for unspliced nuclear transcripts. The Cellranger function aggr was used to aggregate all snRNA-seq libraries without depth normalization to generate a gene by nucleus matrix. Nuclei with fewer than 200 genes, nuclei with more than 6000 genes, or nuclei with more than 15% of unique molecular identifiers stemming from mitochondrial genes were removed. In total, we obtained 42,568 nuclei for downstream analysis. Expression levels were normalized with the LogNormalize method in Seurat²⁸ (v4.0.4), and the top 2100 highly

variable genes (HVGs) were used for principal component analysis (PCA). To remove batch effects, the first 15 PCs were batch corrected using Harmony²⁹ (v0.1). Clustering was performed by constructing a K-nearest neighbor (KNN) graph with corrected PCs and applying the Louvain algorithm. Dimensional reduction was performed with Uniform Manifold Approximation and Projection (UMAP) and individual clusters were annotated based on expression of cell type-specific markers. Differentially expressed genes (DEGs) were identified with the Seurat FindMarkers function for genes detected in at least 25% of cells, using the MAST test and a log-fold-change threshold of 0.25. Bonferroni-adjusted p-values were used to determine significance at $P_{adj} < 0.05$. Gene ontology analysis was performed using Metascape³⁵. Cell cycle score, apoptosis score, HIF pathway score, and cell identity score were evaluated by AddModuleScore function in Seurat with corresponding gene lists, respectively.

Processing of snATAC-seq data

Reads were aligned to the GRCh38 genome using cellranger-atac (v2.0.0). Libraries were aggregated with cellranger-atac without depth normalization to generate a peak by nucleus matrix. Low-quality nuclei (peak region fragments < 200, peak region fragments > 10000, percentage of reads in peaks > 8, blacklist ratio < 0.01, TSS enrichment > 1.5 & nucleosome signal < 1.5) were removed using Signac³⁰ (v1.5.0). In total, we obtained 41,550 nuclei for downstream analysis. The peak by cell matrix was transformed using the term frequency-inverse document frequency (TF-IDF). Dimensional reduction was performed via singular value decomposition (SVD) of the TF-IDF matrix. The first 40 latent semantic indexing (LSI) components were batch corrected using Harmony (v0.1). Clustering was performed by constructing a K-nearest neighbor (KNN) graph with corrected LSI components and applying the Louvain

algorithm. Peak calling was performed with the CallPeaks function in MACS2⁶⁸ (v2.2.7.1) in each cluster. Gene activity was estimated by counting ATAC peaks within the gene body and 2 kb upstream of the TSS using protein-coding genes annotated in the Ensembl database. Canonical correlation analysis³¹ (CCA) was used to capture the shared feature correlation structure between snATAC-seq gene activity and snRNA-seq gene expression. Mutual nearest neighbors³¹ (MNNs) were then identified the pairs of corresponding cells that anchor the two datasets together. We assigned the cell types to the snATAC-seq clusters if the majority (>80%) of cells were aligned to the corresponding cell type. Differentially accessible chromatin regions (DARs) between cell types were assessed with the FindMarkers function for peaks detected in at least 5% of cells, using the MASTtest and a log-fold-change threshold of 0.25. Bonferroni-adjusted p-values were used to determine significance at $P_{adj} < 0.05$. The single-nuclei motif activity for a set of 452 human TFs from the JASPAR 2020⁶⁹ was computed by running chromVAR (v1.14.0) through the RunChromVAR function in Signac. Differential motif activity between young and old ovaries in each cell type was identified by using the FindMarker function for chromVAR motifs detected in at least 25% of cells, using the MAST test and a log-fold-change threshold of 0.50. Bonferroni-adjusted p-values were used to determine significance at $P_{adj} < 0.05$. To further analyze specific TFs of interest, we used the Footprint function in Signac to perform TF footprinting analysis.

Gene set variation analysis (GSVA)

Pathway analyses were performed on the 186 Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways in the MSigDB⁷⁰ database (v7.4.1). GSVA³⁷ (v1.40.1) was used to perform gene set variation analysis to estimate the pathway activity score for individual cells. To compare the

pathway activity scores between young and old ovaries in each cell type, we contrasted the activity scores using the `limma`⁷¹ package (v3.48.3). Bonferroni-adjusted p-values were used to determine significance at an $P_{\text{adj}} < 0.05$. T-values of pathways that exhibited significance in at least 6 cell types were visualized using heatmaps.

Generation of cis-coaccessibility networks with Cicero

We applied Cicero⁵⁹ (v1.3.4.11) to generate cis-accessibility networks (CCANs) for each cell type. Briefly, the Signac object for each cell type was converted to the CellDataSet format and then made into a Cicero object. The algorithm assigned the cells into many groups, each group comprised of 50 cells similarly positioned in clustering space. Graphical LASSO was used to calculate the correlations in adjusted accessibilities between all pairs of ATAC peaks within 500 kb. Finally, CCANs were identified through community detection.

Transcription factor regulatory network construction.

For a given TF, the ovarian ageing-associated DEGs whose promoters or putative enhancers contained both accessible peaks and motifs of the certain ageing-associated TFs within the peaks were defined as candidate targets of the selected TFs. We used this information to construct a directed TF regulatory network using the Gephi (v0.9.2).

Cell type enrichment analysis

For the snRNA-seq data, to estimate the association of gene-level GWAS trait association statistics with gene expression specificity in a given cell type, we used EWCE⁷² (v1.0.0) to calculate gene expression specificity in each cell type. Then, MAGMA.Celltyping⁶² (v1.0.0) was used to calculate the quantile groups for each cell type with the `prepare.quantile.groups` function. The GWAS variants were then annotated onto their neighbouring

genes (Genes were extended 10 kb upstream and 1.5 kb downstream).
Finally, MAGMA⁶¹ (v1.08) was used to test for a positive association (one-sided test) between the cell type specificity and the gene-level associations.
P-values were used to determine significance at $P < 0.05$.

For the snATAC-seq data, we used ldsc⁶³ (v1.0.1) to annotate each variant according to whether or not it overlapped ATAC peaks in each cell type for each GWAS summary statistics. We then estimated partitioned LD scores with the annotated files, HapMap SNPs, and PLINK data corresponding to 1000 genomes phase 3. The baseline model was downloaded from ldsc website. Finally, we used stratified LD score regression to assess the contribution of an annotation to each GWAS trait heritability. *P*-values were used to determine significance at $P < 0.05$.

All GWAS summary statistics for age at menopause^{16,73,74}, polycystic ovary syndrome (PCOS)^{75,76}, and ovarian epithelial cancer (OEC)⁷⁷⁻⁷⁹ were downloaded from GWAS Catalog (<https://www.ebi.ac.uk/gwas/>) or ReproGen (<https://www.reprogen.org/>), and re-formatted with MungeSumstats (v1.3.5) or munge_sumstats.py in the ldsc package.

The effect of variants on transcription factor binding activity

To predict the TFs binding activity score, we overlapped our ovary ATAC peaks with human ovary tissue enhancer and promoter annotations from the ENCODE database. In this way, we obtained 71,470 putative enhancers and promoter regions, which were used as the positive set. We generated the random length and GC-matched genome sequences as the negative set. We then used the gkmtrain function from LS-GKM (v0.1.1)⁶⁶, a new gkm-SVM software for large-scale datasets, to train the TFs binding model for human ovary with positive set, negative set, and “gkmrbf” kernel. For the variants of interest, we retrieved the ± 25 bp reference DNA sequence

around the variant. To generate the corresponding alternative DNA sequence, we replaced the 25th position with the effect allele. To compute deltaSVM scores, we generated all possible non-redundant k-mers of size 11 and scored each of them using the trained model. We then used deltaSVM to compute the deltaSVM scores with k-mer scores, reference sequences, and alternative sequences. For the GkmExplain scores, we used GkmExplain⁸⁰ on the reference sequences or alternative sequences of variants of interest. The GkmExplain scores were visualized using logomaker (v0.8) (<https://github.com/jbkinney/logomaker>).

Cellular communication

To build cell-cell interactions between somatic cells and oocytes in young and old human ovaries, we used CellChat⁴⁵ (v1.1.3) to infer the cell-cell interactions based on the expression of known ligand-receptor pairs in different cell types with a combination of our snRNA-seq datasets and publicly available human oocyte single-cell RNA-seq datasets⁴⁶ from reproductive young and old females. Briefly, we inferred the cell-cell interactions for young and old ovaries, respectively. Next, we used “rankNet” function in CellChat to identify the significant outgoing or incoming signaling enriched in young or old ovaries.

***In situ* hybridization assay**

Flash-frozen human ovary tissues were sectioned at 10 μ m. RNA in situ hybridization was performed using RNAscope Multiplex Fluorescent v2 kits (Advanced Cell Diagnostics) according to the manufacturer’s instructions, except fixed with 4% PFA 90 mins at RT and protease IV incubation was performed for 15 min. Probes used were MT-ATP6 (532961), RICTOR (544841), and CDKN1A (311401). Fluorophores used were Opal 690 (1:1500 dilution, Perkin Elmer). Images were taken on a

Leica TCS SP8 MP at $\times 40$ magnification. 4-10 regions per sample were analyzed using HALO (v3.2). The mRNA expression levels were evaluated according to the ACD scoring system (<https://acdbio.com/dataanalysisguide>) by counting number of dots per cell. The cells with at least 4 dots were recognized as *CDKN1A*⁺ cells.

Reporting Summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The snRNA-seq and snATAC-seq data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) under accession numbers: GSE202601. All other data supporting the findings of this study are available from the corresponding authors on reasonable request.

Code availability

The codes used to analyze the snRNA-seq and snATAC-seq data are available at <https://github.com/ChenJin2020/The-regulatory-landscapes-of-human-ovarian-ageing>.

Reference

- 68 Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome biology* **9**, 1-9 (2008).
- 69 Fornes, O. *et al.* JASPAR 2020: update of the open-access database of transcription factor binding profiles. *Nucleic acids research* **48**, D87-D92 (2020).
- 70 Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences* **102**, 15545-15550 (2005).
- 71 Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic acids research* **43**, e47-e47 (2015).
- 72 Skene, N. G. & Grant, S. G. Identification of vulnerable cell types in major brain disorders using single cell transcriptomes and expression weighted cell type enrichment. *Frontiers in neuroscience* **10**, 16 (2016).
- 73 Loh, P.-R., Kichaev, G., Gazal, S., Schoech, A. P. & Price, A. L. Mixed-model association for biobank-scale datasets. *Nature genetics* **50**, 906-908 (2018).

829 74 Backman, J. D. *et al.* Exome sequencing and analysis of 454,787 UK Biobank
830 participants. *Nature* **599**, 628-634 (2021).
831 75 Day, F. *et al.* Large-scale genome-wide meta-analysis of polycystic ovary syndrome
832 suggests shared genetic architecture for different diagnosis criteria. *PLoS genetics* **14**,
833 e1007813 (2018).
834 76 Tyrmi, J. S. *et al.* Leveraging Northern European population history: novel low-
835 frequency variants for polycystic ovary syndrome. *Human Reproduction* **37**, 352-365
836 (2022).
837 77 Rashkin, S. R. *et al.* Pan-cancer study detects genetic risk variants and shared genetic
838 basis in two large cohorts. *Nature communications* **11**, 1-14 (2020).
839 78 Phelan, C. M. *et al.* Identification of 12 new susceptibility loci for different histotypes
840 of epithelial ovarian cancer. *Nature genetics* **49**, 680-691 (2017).
841 79 Jiang, L., Zheng, Z., Fang, H. & Yang, J. A generalized linear mixed model association
842 tool for biobank-scale data. *Nature genetics* **53**, 1616-1621 (2021).
843 80 Shrikumar, A., Prakash, E. & Kundaje, A. GkmExplain: fast and accurate interpretation
844 of nonlinear gapped k-mer SVMs. *Bioinformatics* **35**, i173-i182 (2019).

845

846

847

848

849

850

851

852

853

854

855

856

857

Acknowledgments

We thank Wilber Quispe (SingulOmics Corporation) for snRNA-seq and snATAC-seq libraries preparation and Columbia Genome Center for sequencing. This work was supported by NIH grants AG069750, DK127778, AG057433, AG061521, HL150521, AG055501, AG057341, AG057433, AG057706, AG057909, and AG17242 (Y.S), a grant GCRLE-1320 (Y. S.) from the Global Consortium for Reproductive Longevity and Equality at the Buck Institute, made possible by the Bia-Echo Foundation, and a grant from The Simons Foundation (Y.S.).

Author information

These authors contributed equally: Chen Jin, Xizhe Wang

Affiliations

Department of Obstetrics and Gynecology, Columbia University Medical Center, New York, NY

Chen Jin, Xizhe Wang, Adam Hudgins, Mingzhuo Pei, Seungsoo Kim, Daniela Contreras, Roger Lobo, Zev Williams & Yousin Suh

Howard Hughes Medical Institute, Department and School of Medicine, University of California, San Diego, La Jolla, CA, USA

Amir Gamliel, Michael Rosenfeld

Department Molecular Genetics, Erasmus University Medical Center Rotterdam, Rotterdam, The Netherlands

Jan Hoeijmakers

Princess Máxima Center for Pediatric Oncology, Oncode Institute, Utrecht,

882 The Netherlands

883 Jan Hoeijmakers

884 Institute for Genome Stability in Ageing and Disease, Cologne Excellence

885 Cluster for Cellular Stress Responses in Aging-Associated Diseases

886 (CECAD), University Hospital of Cologne, Cologne, Germany Jan

887 Hoeijmakers

888 Buck Institute for Research on Aging, Novato, CA, USA

889 Judith Campisi

890 Lawrence Berkeley National Laboratory, Berkeley, CA, USA

891 Judith Campisi

892 Department of Genetics and Development, Columbia University Medical

893 Center, New York, NY

894 Yousin Suh

895 **Contributions**

896 C.J. designed and performed the experiments, analyzed the data, and wrote

897 the manuscript; X.W. performed the *in situ* hybridization assay, and

898 analyzed the imaging data; A.H. provided the analytic advice and revised

899 the manuscript; A.G. generated part of snRNA-seq libraries; M.P. assisted

900 the snRNA-seq data analysis; S.K. identified the LD SNPs and provided

901 the analytic advises; D.C. assisted the *in situ* hybridization assay; J.H. and

902 J.C. revised the manuscript; R.L, Z.W. and M.R. provided conceptual

903 advice; Y.S. conceived and designed the research, analyzed and interpreted

904 the data, and wrote the manuscript.

905 **Corresponding author**

906 Correspondence to Yousin Suh

907 **Ethics declarations**

908 **Competing interests**

909 The authors declare no competing interests

910

911 **Figure Legends**

912 **Fig.1: Single-nuclei transcriptomic and chromatin accessibility**
 913 **profiling of the human ovary**

914 **a**, Schematic representation of experimental methodology. **b**, UMAP plots
 915 of human ovary snRNA-seq dataset. **c**, Dot plot representing relative
 916 mRNA expression of selected known markers for each cell type. Dot size
 917 indicates the proportion of cells in the cluster expressing a gene, the
 918 shading indicates the relative level of expression (low to high reflected as
 919 light to dark). **d**, UMAP plots of human ovary snATAC-seq dataset. **e**, Dot
 920 plot representing relative gene activity of selected known markers for each
 921 cell type. Dot size indicates the proportion of cells in the cluster expressing
 922 a gene, the shading indicates the relative level of expression (low to high
 923 reflected as light to dark).

924 **Fig.2: Ageing alters ovarian cellular composition and affects the**
 925 **transcriptional activity of pathways involved in the hallmarks of**
 926 **ageing across cell types**

927 **a**, Bar plots represent the proportion of each cell type in young and aged
 928 ovaries estimated from snRNA-seq data. (Permutation test; Asterisk (*))
 929 indicates $FDR < 0.05$ and $abs(log_2FD) > 1.5$; Methods). **b**, Bar plots
 930 represent the proportion of each cell type in young and aged ovaries

estimated from snATAC-seq data. (Permutation test; Asterisk (*) indicates FDR<0.05 and $\text{abs}(\log_2\text{FD}) > 1.5$). **c**, Heat map displaying $\log_2$ fold changes in gene expression (aged vs. young) of human ovarian ageing-associated DEGs in each cell type. **d**, Heat map showing selected up- and down-regulated pathways significantly altered in at least 6 cell types during human ovarian ageing. Asterisk (*) indicates a statistically significant difference ($P_{\text{adj}} < 0.05$). **e**, Representative *in situ* hybridization (RNAscope) images from fresh-frozen human ovary tissue for *RICTOR* staining. **f**, Quantification of *RICTOR* expression in human ovary (young versus old). $n = 2$; $*P < 0.05$. **g**, Representative *in situ* hybridization (RNAscope) images from fresh-frozen human ovary tissue for *MT-ATP6* staining. **h**, Quantification of *MT-ATP6* expression in human ovary (young versus old). $n = 2$; $**P < 0.01$.

Fig.3: Ageing increases signatures of cellular senescence and alters cellular communication in the ovary

a, Bar plots represent the proportion of *CDKN1A*⁺ for each cell type in young and aged ovaries. (Permutation test; Asterisk (*) indicates FDR<0.05 and $\text{abs}(\log_2\text{FD}) > 1.5$). **b**, Representative *in situ* hybridization (RNAscope) images from fresh-frozen human ovary tissue for *CDKN1A* (p21) staining. **c**, Quantification of *CDKN1A* expression and the proportion of *CDKN1A*⁺ cells in the human ovary (young versus old). $n = 2$; $*P < 0.05$. **d**, Heat map displaying $\log_2$ fold changes in gene expression (*CDKN1A*⁺ vs. *CDKN1A*⁻ cells) of selected SASP genes in each cell type. **e**, Violin plots showing the module score of HIF-1 pathway genes in *CDKN1A*⁺ cells and *CDKN1A*⁻ cells from each type. (Two-sided Wilcoxon test; NS: Not significant, $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, $****P < 0.0001$). **f**, Heat map of the differential number of interactions between cell types in young and aged ovaries. The top bar plots represent the sum of each column of values

displayed in the heatmap (incoming signaling). The right bar plots represent the sum of each row of values (outgoing signaling). **g**, Heat map showing the outgoing and incoming signaling pathways that were significantly enriched in young or aged ovaries for each cell type.

Fig.4: Cell type-specific TF regulatory networks implicate CEBPD as an important regulator of ageing-associated gene expression in the human ovary

a, Heat map showing the average chromVAR motif activity for each cell type. **b**, UMAP plots displaying the chromVAR motif activity of selected cell type-specific TFs. **c**, Heat map showing the TFs with significant changes in chromVAR motif activity in each cell type during ovarian ageing. **d**, Split violin plots showing the expression levels of *CEBPD* in each cell type from young and aged ovaries. (MAST test; $*P_{adj} < 0.05$). **e-j**, TF regulatory network plots showing the top regulators of ageing-associated DEGs in each cell type.

Fig.5: Integration of ANM GWAS, single-nuclei multi-omics, and machine-learning nominates causal variants and gene targets associated with human ovarian ageing

a, Heat map of enrichment significance of ovary-relevant trait GWAS variants in ovary cell type gene expression signatures. **b**, Heat map of enrichment significance of ovary-relevant trait GWAS variants in ovary cell type-specific chromatin accessibility. **c**, Upset plot showing the intersection size between sets of ANM-associated variants that overlap with transcriptional regulatory elements found in each cell type. The bar plot on the left shows the set size of variants for each cell type, and the bar plot on the top shows the number of overlapping variants shared by two or

more sets, or the number of unique variants in one set. **d**, Cis-regulatory architecture at the *HELB* gene in each cell type. The snATAC-seq tracks represent the aggregate signals of all cells from a given cell type. The co-accessible peaks inferred by Cicero for each cell type are shown. **e**, The gkm-SVM importance score for each base within the ± 25 -base pair (bp) region surrounding rs3741605. **f**, The eQTL effect of rs3741605 on *HELB* expression in human ovary tissue from the GTEx database. **g**, Cis-regulatory architecture at the *DEPTOR* gene in each cell type. The snATAC-seq tracks represent the aggregate signals of all cells from a given cell type. The co-accessible peaks inferred by Cicero for each cell type are shown. **h**, The gkm-SVM importance score for each base within the ± 25 -base pair (bp) region surrounding rs13263296. **i**, The eQTL effect of rs13263296 on *DEPTOR* expression in human ovary tissue from the GTEx database.

1000

1001

Fig.1

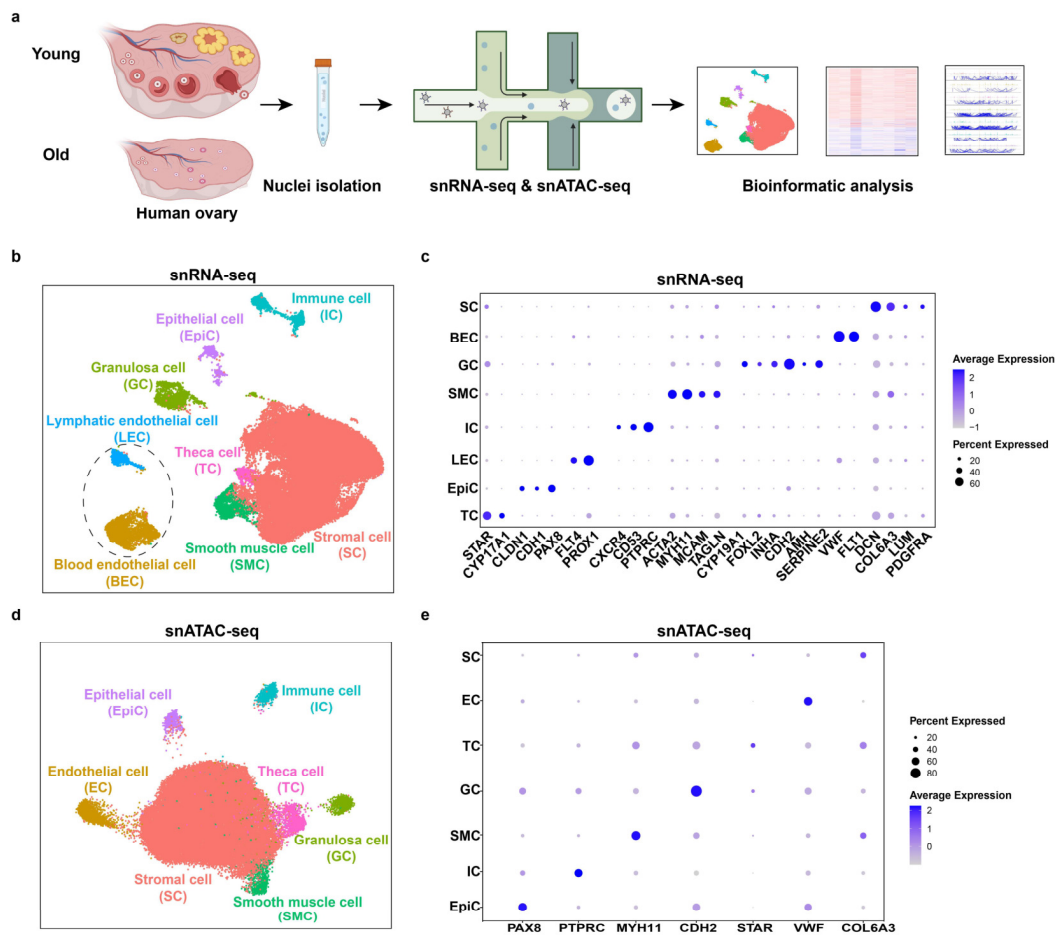


Fig.1: Single-nuclei transcriptomic and chromatin accessibility profiling of the human ovary

a, Schematic representation of experimental methodology. **b**, UMAP plots of human ovary snRNA-seq dataset. **c**, Dot plot representing relative mRNA expression of selected known markers for each cell type. Dot size indicates the proportion of cells in the cluster expressing a gene, the shading indicates the relative level of expression (low to high reflected as light to dark). **d**, UMAP plots of human ovary snATAC-seq dataset. **e**, Dot plot representing relative gene activity of selected known markers for each cell type. Dot size indicates the proportion of cells in the cluster expressing a gene, the shading indicates the relative level of expression (low to high reflected as light to dark).

Fig.2

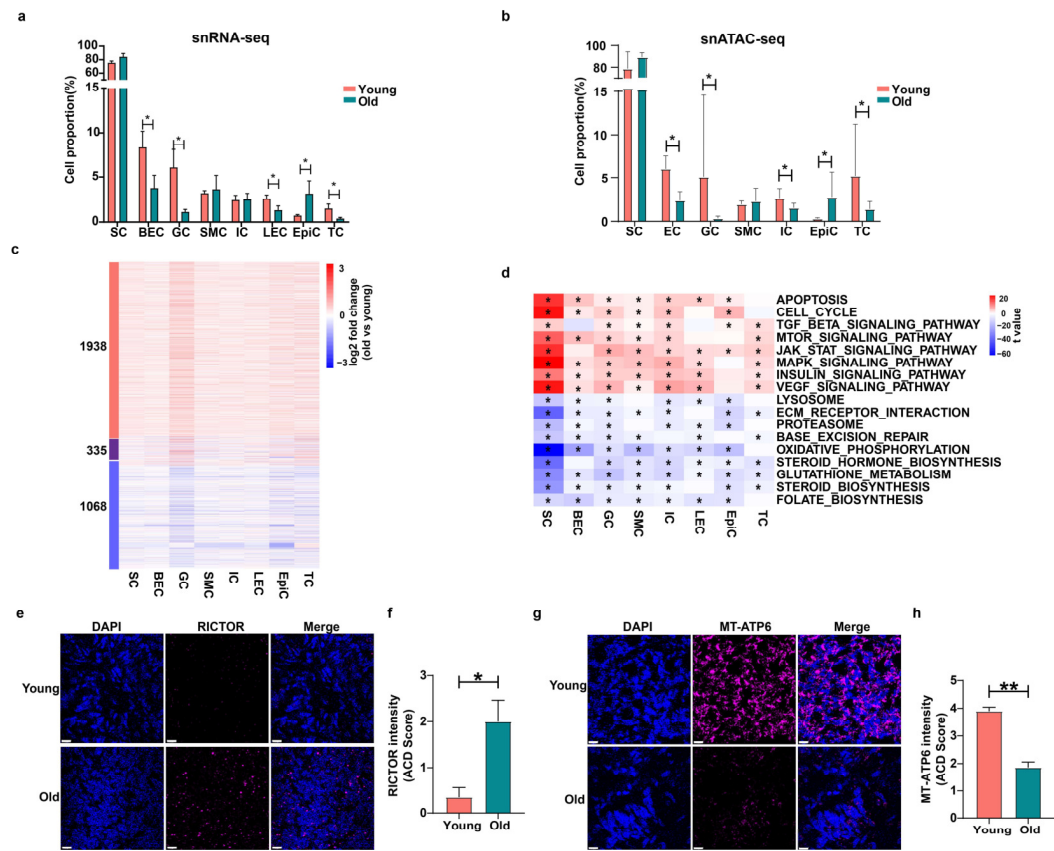


Fig.2: Ageing alters ovarian cellular composition and affects the transcriptional activity of pathways involved in the hallmarks of ageing across cell types

a, Bar plots represent the proportion of each cell type in young and aged ovaries estimated from snRNA-seq data. (Mean±SE; Permutation test; Asterisk (*) indicates $FDR < 0.05$ and $abs(log_2FD) > 1.5$; Methods). **b**, Bar plots represent the proportion of each cell type in young and aged ovaries estimated from snATAC-seq data. (Mean±SE; Permutation test; Asterisk (*) indicates $FDR < 0.05$ and $abs(log_2FD) > 1.5$). **c**, Heat map displaying log_2 fold changes in gene expression (aged vs. young) of human ovarian ageing-associated DEGs in each cell type. **d**, Heat map showing selected up- and down-regulated pathways significantly altered in at least 6 cell types during human ovarian ageing. Asterisk (*) indicates a statistically

significant difference ($P_{\text{adj}} < 0.05$). **e**, Representative *in situ* hybridization (RNAscope) images from fresh-frozen human ovary tissue for *RICTOR* staining. **f**, Quantification of *RICTOR* expression in human ovary (young versus old). Mean $\pm$ SE; n = 2; * $P < 0.05$. **g**, Representative *in situ* hybridization (RNAscope) images from fresh-frozen human ovary tissue for *MT-ATP6* staining. **h**, Quantification of *MT-ATP6* expression in human ovary (young versus old). Mean $\pm$ SE; n = 2; ** $P < 0.01$.

Fig.3

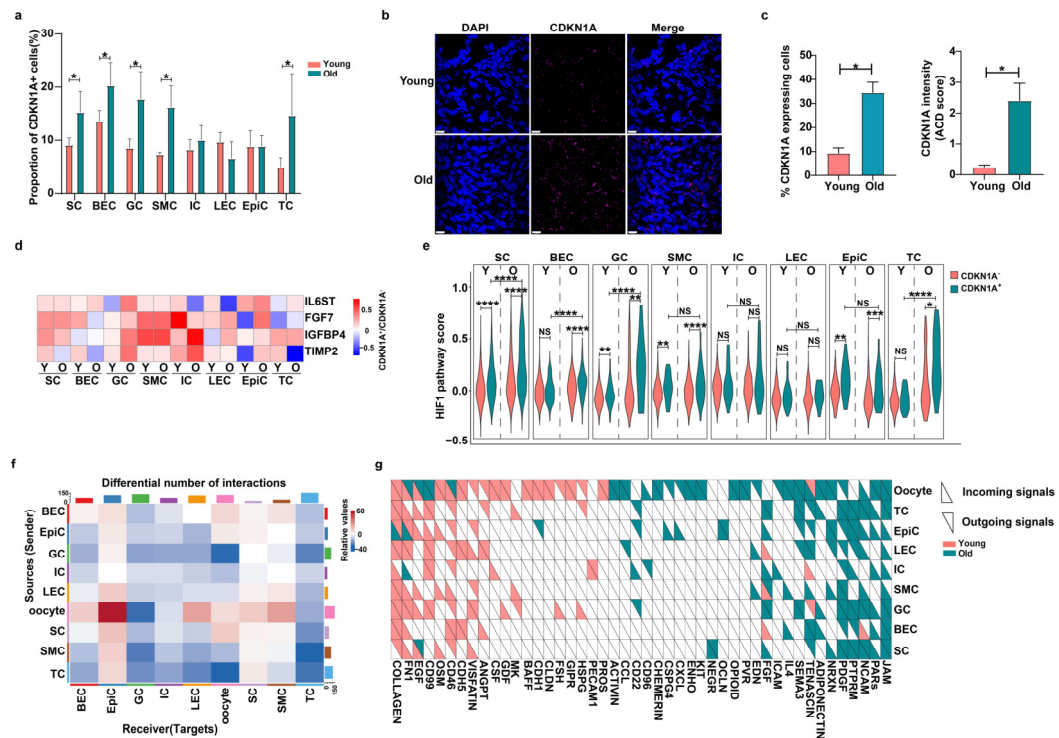


Fig.3: Aging increases signatures of cellular senescence and alters cellular communication in the ovary

a, Bar plots represent the proportion of $CDKN1A^+$ cells for each cell type in young and aged ovaries. (Mean $\pm$ SE; Permutation test; Asterisk (*) indicates FDR<0.05 and abs(log2FD)>1.5). **b**, Representative *in situ* hybridization (RNAscope) images from fresh-frozen human ovary tissue for $CDKN1A$ (p21) staining. **c**, Quantification of $CDKN1A$ expression and the proportion of $CDKN1A^+$ cells in the human ovary (young versus old). Mean $\pm$ SE; n = 2; * P < 0.05. **d**, Heat map displaying log2 fold changes in gene expression ($CDKN1A^+$ vs. $CDKN1A^-$ cells) of selected SASP genes in each cell type. **e**, Violin plots showing the module score of HIF-1 pathway genes in $CDKN1A^+$ cells and $CDKN1A^-$ cells from each type. (Two-sided Wilcoxon test; NS: Not significant, * P <0.05, ** P <0.01, *** P <0.001, **** P <0.0001). **f**, Heat map of the differential number of interactions between cell types in young and aged ovaries. The top bar plots

represent the sum of each column of values displayed in the heatmap (incoming signaling). The right bar plots represent the sum of each row of values (outgoing signaling). **g**, Heat map showing the outgoing and incoming signaling pathways significantly enriched in young or aged ovaries for each cell type.

Fig.4

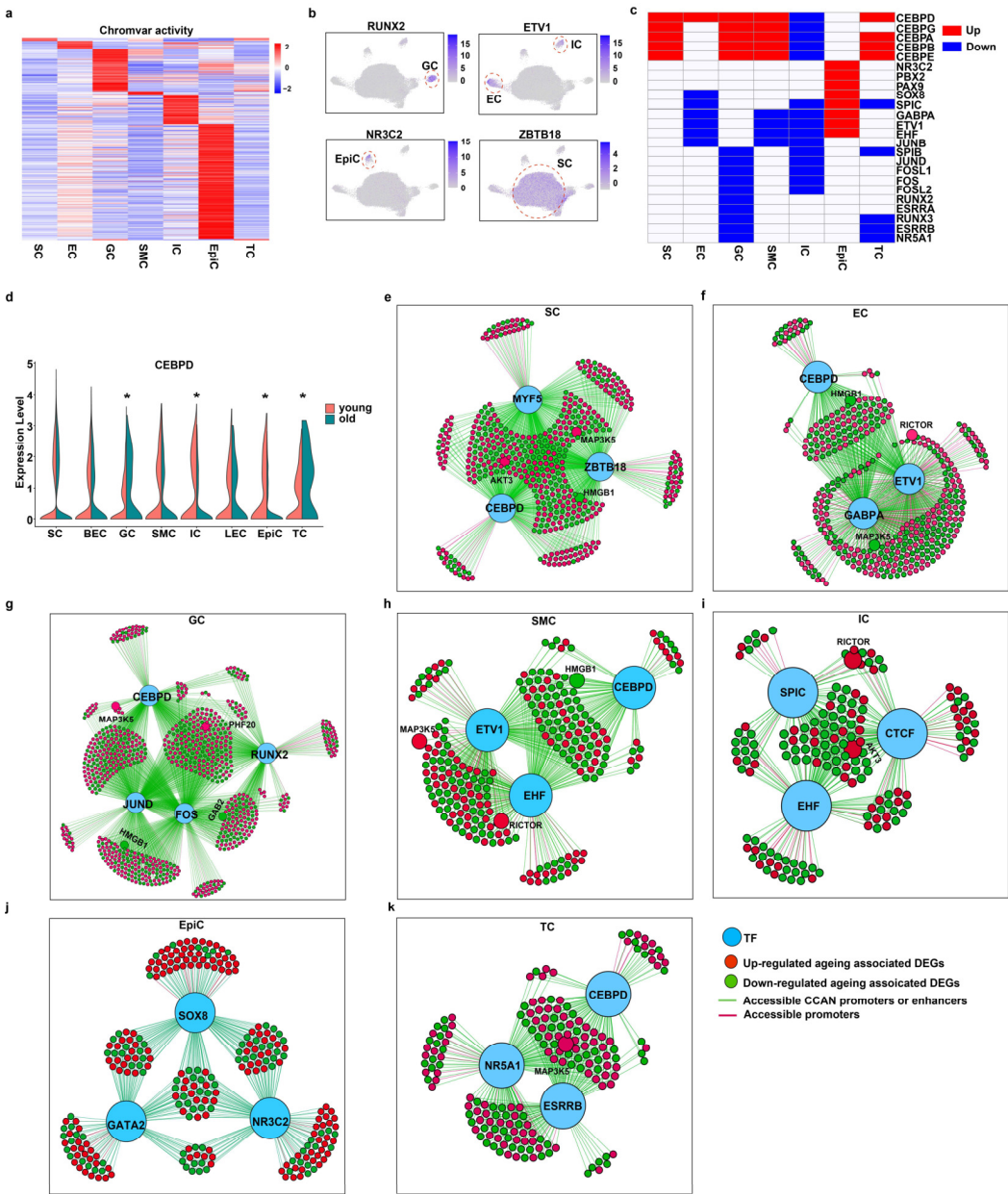


Fig.4: Cell type-specific TF regulatory networks implicate CEBPD as an important regulator of ageing-associated gene expression in the human ovary

a, Heat map showing the average chromVAR motif activity for each cell

type. **b**, UMAP plots displaying the chromVAR motif activity of selected cell type-specific TFs. **c**, Heat map showing the TFs with significant changes in chromVAR motif activity in each cell type during ovarian ageing. **d**, Split violin plots showing the expression levels of *CEBPD* in each cell type from young and aged ovaries. (MAST test; $*P_{\text{adj}} < 0.05$). **e-j**, TF regulatory network plots showing the top regulators of ageing-associated DEGs in each cell type.

Fig.5

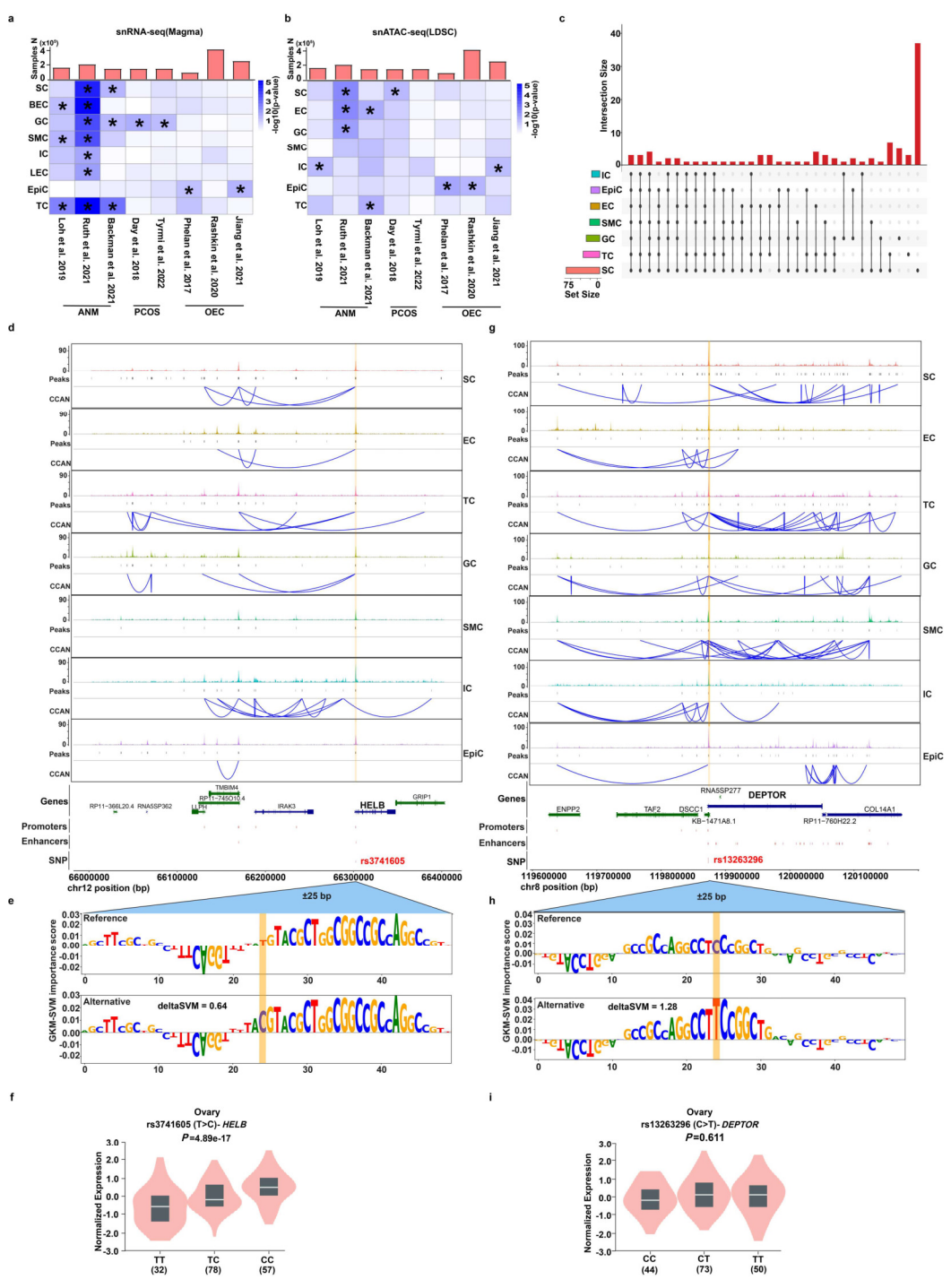


Fig.5: Integration of ANM GWAS, single-nuclei multi-omics, and machine-learning nominates causal variants and gene targets associated with human ovarian aging

a, Heat map of enrichment significance of ovary-relevant trait GWAS

variants in ovary cell type gene expression signatures. **b**, Heat map of enrichment significance of ovary-relevant trait GWAS variants in ovary cell type-specific chromatin accessibility. **c**, Upset plot showing the intersection size between sets of ANM-associated SNPs that overlap with transcriptional regulatory elements found in each cell type. The bar plot on the left shows the set size of variants for each cell type, and the bar plot on the top shows the number of overlapping SNPs shared by two or more sets, or the number of unique variants in one set. **d**, Cis-regulatory architecture at the *HELB* gene in each cell type. The snATAC-seq tracks represent the aggregate signals of all cells from a given cell type. The co-accessible peaks inferred by Cicero for each cell type are shown. **e**, The gkm-SVM importance score for each base within the ± 25 -base pair (bp) region surrounding rs3741605. **f**, The eQTL effect of rs3741605 on *HELB* expression in human ovary tissue from the GTEx database. **g**, Cis-regulatory architecture at the *DEPTOR* gene in each cell type. The snATAC-seq tracks represent the aggregate signals of all cells from a given cell type. The co-accessible peaks inferred by Cicero for each cell type are shown. **h**, The gkm-SVM importance score for each base within the ± 25 -base pair (bp) region surrounding rs13263296. **i**, The eQTL effect of rs13263296 on *DEPTOR* expression in human ovary tissue from the GTEx database.