FOXL2 interaction with different binding partners regulates the dynamics of granulosa cell differentiation across ovarian development

Roberta Migale[a,1], Michelle Neumann[a*], Richard Mitter[b*], Mahmoud-Reza Rafiee[c], Sophie Wood[d], Jessica Olsen[d] & Robin Lovell-Badge[a,1]

Affiliations

[a]Laboratory of Stem Cell Biology and Developmental Genetics, The Francis Crick Institute, London NW1 1AT, UK

[b]Bioinformatics core, The Francis Crick Institute, London NW1 1AT, UK

[c]RNA Networks Laboratory, The Francis Crick Institute, London NW1 1AT, UK

[d]Genetic Modification Service, The Francis Crick Institute, London NW1 1AT, UK

*contributed equally

[1]Correspondence to: roberta.migale@crick.ac.uk & robin.lovell-badge@crick.ac.uk

Abstract

FOXL2 is a transcription factor essential for female fertility, expressed in somatic cells of the ovary, notably granulosa cells. In the mouse, *Foxl2* deletion leads to partial sex reversal postnatally, with mutants developing dysgenic ovaries devoid of oocytes. However, deletion of the gene in 8-week-old females leads to granulosa to Sertoli cell transdifferentiation and gonadal sex reversal. We hypothesise that different outcomes of *Foxl2* deletion in embryonic versus adult ovary may derive from a different role played by FOXL2 across ovarian development. Therefore, in this study, we take a multi-omics approach to characterise the dynamics of gene expression and chromatin accessibility changes in purified murine granulosa cells across key developmental stages (E14.5, 1 and 8 weeks). We coupled these analyses with genome wide identification of FOXL2 target genes and on-chromatin interacting partners by ChIP-SICAP to reconstruct the gene regulatory networks underpinned by this essential transcription factor. In the embryonic ovary, FOXL2 interacts with factors important for early stages of gonadal development, such as GATA4 and WT1, whilst postnatally it interacts with factors regulating primordial follicle activation, such as NR5A2, and with factors regulating steroidogenesis including AR and ESR2. Integration of chromatin landscape dynamics with gene expression changes and FOXL2 binding sites analysis revealed that its critical role in ovarian cell fate maintenance goes beyond repression of the Sertoli-specific gene *Sox9*. Our chromatome analysis revealed also that FOXL2 interacts with several proteins involved in chromatin remodelling, DNA repair, splicing and gene repression. In summary, in this study we identified target genes dynamically regulated by FOXL2 across ovarian development including known and newly identified FOXL2 targets with a role in embryonic ovarian development and folliculogenesis, as well as cofactors that point towards additional roles played by FOXL2 besides transcriptional regulation. This work constitutes a comprehensive resource for exploration of the molecular mechanisms of ovarian development and causes of female infertility.

53 **INTRODUCTION**
54
55 The development of primary reproductive organs, ovaries in the female and testes in the male,
56 starts from a pair of bipotential structures called genital ridges. These develop from E9.5 in the
57 mouse and from the 5th week of gestation in humans [1]. Bipotential genital ridges contain
58 undifferentiated cell types poised to develop into one of two sexes, depending on the genetic
59 switch activated [2]. The time and robustness of this cell fate decision is critical for the healthy
60 development of the gonads, successful reproduction and ultimately for the survival of a species.
61 The presence or absence of the male sex determination gene *Sry*, usually located on the Y
62 chromosome in mammals, determines which type of gonad it will become [3]. Its expression,
63 starting at E10.75 and peaking at E11.5 in the mouse [4, 5], is sufficient to induce male
64 differentiation in an XX female context[3].
65 This is also the case for its downstream target *Sox9*, which is able to induce male development
66 in the absence of *Sry* [6]. Ovarian development was considered as a default pathway activated
67 in the absence of *Sry* [7]. Recent work has challenged this idea showing that female sex
68 determination is not a default pathway, but an active process involving the constant repression
69 of the genetic program of the opposite sex [8-10]. A critical player in ovarian development and
70 function is *Foxl2*, a gene encoding for the forkhead box L2 (FOXL2) protein, a transcription
71 factor involved in regulating the differentiation and maintenance of granulosa cells in the ovary
72 [11, 12]. Mutations in the FOXL2 gene in humans have been linked to premature ovarian
73 insufficiency (POI), a condition in which the ovaries stop functioning before the age of 40,
74 leading to infertility [13-15].
75 Null mutant (*Foxl2*$^{-/-}$) and conditional mutant (R26CreERT2;*Foxl2*$^{fl/fl}$) murine models aimed
76 at studying the role of this factor in ovarian development provided intriguing results.
77 *Foxl2*$^{-/-}$ ovaries displayed morphological anomalies only from 1 week postpartum, involving
78 the lack of fragmentation of sex cords into individual follicles, failure to activate the
79 primordial-to-primary transition of granulosa cells, thus resulting in impaired folliculogenesis
80 and infertility. At the molecular level, mutant ovaries were characterised by downregulation of
81 somatic sex markers of folliculogenesis and upregulation of the pro-testis factor *Sox9* as well
82 as other male gonadal-specific markers including *Fgf9*, *Dhh*, *Dmrt1* and *Gata4* [16, 17].
83 These studies suggest that FOXL2 is not critical for initial ovarian specification because ovaries
84 will differentiate correctly up to when folliculogenesis should have started, around one week
85 postnatally [17]. In contrast, conditional deletion of *Foxl2* from adult ovaries resulted into a
86 more dramatic phenotype, involving transdifferentiation of the supporting cells of the ovary,
87 granulosa cells, into Sertoli cells, their opposite sex counterpart [9]. Crucially, the ectopic
88 upregulation of SOX9 in this model occurred within three days following *Foxl2* deletion,
89 indeed as soon as FOXL2 protein was lost, much earlier compared to the *Foxl2* null mutant
90 model. This study demonstrated the plasticity of the adult ovary and the crucial role played by
91 FOXL2 in this process. More recently, a similar role in ovarian sex maintenance was
92 discovered for TRIM28, which we show to be recruited on chromatin together with FOXL2
93 and to regulate the sex-specific SUMOylation profile of ovarian-specific genes through its E3-
94 SUMO ligase activity [8].

Collectively these data support a differential role played by FOXL2 depending on the developmental stage of the ovary, as well as supporting the presence of additional factors cooperating with FOXL2 in ensuring correct ovarian development.

We hypothesised that this differential function of FOXL2 during development may be dependent on the availability of key interacting partners. Therefore, in this study, we combined FOXL2 ChIP-SICAP with RNA-Seq and ATAC-Seq on purified *Foxl2*^{EGFP} expressing cells to address the differences in the role played by FOXL2 at embryonic versus adult stages and to reconstruct the gene regulatory network underpinned by this transcription factor fundamental to female fertility. We chose three timepoints for the analysis: E14.5 for embryonic stages, 1 week postnatally, because this is equivalent to prepubertal age and corresponds to the timepoint at which SOX9 protein is first detected in *Foxl2*$^{-/-}$ ovaries, and 8 weeks when it is imperative for ovary maintenance. We identified: i) the direct targets of FOXL2 at different stages of ovarian development, ii) on-chromatin protein interactors, and iii) the complement of gene expression changes and cis-regulatory element dynamics occurring throughout ovarian development.

Our data support a more prominent role of FOX2 in postnatal stages compared to early timepoints and provide a rich resource of novel protein interactors colocalising with FOXL2 on nearby genomic locations that could be used to generate new hypotheses on the molecular mechanisms of ovarian development, POI and other causes of female infertility.

RESULTS

ChIP-SICAP reveals differential genomic binding of FOXL2 throughout ovarian development.

To profile the dynamic changes in FOXL2 DNA binding and to identify its on-chromatin protein interactors, we performed a modified version of ChIP-SICAP (Methods & [18, 19]) on whole mouse ovaries at three timepoints: E14.5, one week postnatally (1W) and at 8 weeks (8W) (Fig.1A). ChIP-SICAP is a dual approach which allowed us to simultaneously identify genome-wide FOXL2 binding sites, via ChIP-Sequencing, and proteins colocalising with FOXL2 on chromatin, via mass spectrometry.

E14.5 was chosen because this embryonic stage is characterised by high FOXL2 expression in the bipotential pre-granulosa cells of the ovarian medulla [20]; 1W coincides with the activation of primordial follicles which start to develop into primary follicles, a process where FOXL2 plays a central role by regulating the transition of the granulosa cells surrounding the oocytes from squamous to cuboidal, and by inhibiting the growth of oocytes [12]; at 8W FOXL2 is widely expressed in the follicles where it regulates folliculogenesis [21] and maintains ovarian cell fate [9].

Principal-component analysis (PCA) of ChIP-Seq peaks highlighted the differences between FOXL2 binding sites across the timecourse. PC1 (54% of variance, Figure 1B) showed a clear separation of the E14.5 from both 1W an 8W. PC2 (27% of variance) further separated the 1W and 8W samples. Embryonic and postnatal FOXL2 binding sites differed in their genomic location. Annotation of peaks showed that most identified at E14.5 (50%) were mapped to promoters, whilst about 30% of peaks identified at 1W and 8W were located within promoters

(Figure S1A). Indeed, most of the peaks from these two timepoints localised to either the gene body or to distal intergenic regions. Motif enrichment analysis of consensus peaks across the timecourse confirmed enrichment of FOXL2 and other forkhead factors motifs (Figure S1B). Furthermore, we found enrichment of steroid hormone receptors amongst the top 10 predicted motifs. These included estrogen receptors and androgen receptor factors known to regulate hormone production as well as granulosa cell proliferation throughout ovarian development [22, 23], as well as NR5A1, an important transcriptional regulator of genes involved in the hypothalamic-pituitary-gonadal steroidogenic axis [24] and FOXL2 binding partner [25], as well as it being essential for gonadal development in both sexes [26].

Figure 1

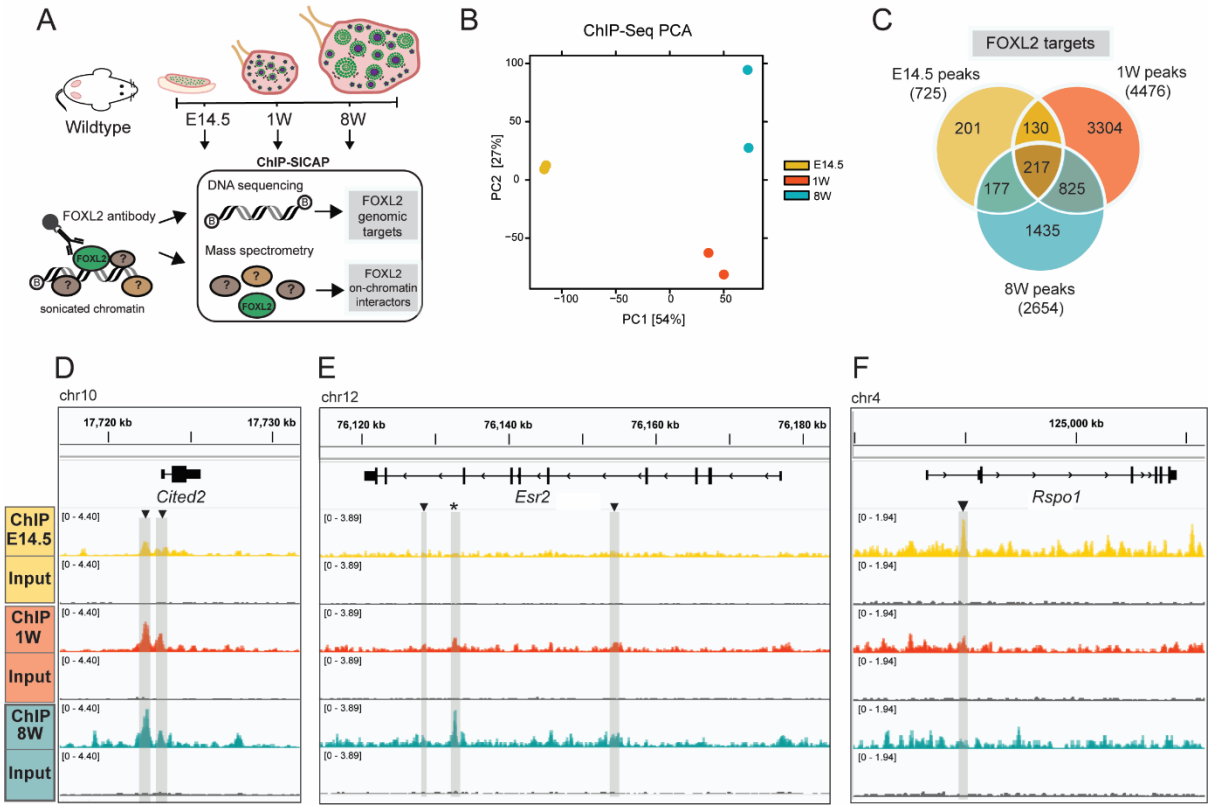


Figure 1: Timecourse ChIP-SICAP reveals differential FOXL2 binding across ovarian development.

(A) Experimental design. FOXL2 ChIP-SICAP was performed on mouse ovaries collected at three timepoints (E14.5, 1W and 8W, n=2). A FOXL2-specific antibody was used to capture genomics fragments bound by FOXL2 as well as proteins colocalising with FOXL2 on chromatin. DNA sequencing was used to identify FOXL2 target genes and mass-spectrometry was used to find its on-chromatin interacting partners. (B) PCA of FOXL2 ChIP-Seq consensus peaks identified with the nf-core MACS analysis (n=2). (C) Venn diagram depicting the number of peaks identified at each timepoint and their overlap across time. Only peaks present in both replicates (n=2) over the input were considered. (D-F) Integrative Genome Viewer (IGV) tracks of representative FOXL2-bound regions. Main tracks show normalised read-depth coverage (RPM). Grey bars and black arrow highlight significant peaks over input. (D) *Cited2*, (E) *Esr2* [27], and (F) *Rspo1*. Star denotes proven functional *Esr2* enhancer (see [27]).

163 In total, we detected 725 peaks at E14.5, 4476 at 1W and 2654 at 8W (Figure 1C, Table S1).
164 Amongst these, we identified known FOXL2 targets bound throughout the timecourse, such as
165 *Cited2* (Figure 1D), as well as targets dynamically bound including *Esr2* (Figure 1E) and a
166 novel target gene *Rspo1* (Figure 1F) encoding a positive effector of β-catenin signalling that is
167 required for the development of functional ovaries [28].
168 Differential binding analysis allowed the identification of FOXL2-binding events that were
169 statistically enriched in one timepoint or the other (Figure S1C-E & Table S1). We detected
170 the greatest number of differential binding (2503 differential peaks) when comparing E14.5
171 with 8W (Figure S1D), followed by the comparison between E14.5 and 1W (Figure S1C, 2098
172 differential peaks) and lastly by the comparison between 1W and 8W (Figure S1E, 634
173 differential peaks). Comparison of our FOXL2 ChIP-Seq dataset with a microarray analysis of
174 differentially expressed genes in mouse ovaries lacking *Foxl2* compared to wildtype controls
175 (*Foxl2*$^{-/-}$, Ortiz *et al.* [29], Table S1) allowed us to identify 770 genes more likely to be activated
176 by FOXL2 (because downregulated upon loss of *Foxl2* as well as bound by FOXL2, Figure
177 S2A). This group included, besides *Cited2* and *Esr2*, the cell-cycle repressor gene *Cdkn1b*
178 (Figure S2B), a known FOXL2 target that regulates granulosa cell proliferation [30], *Smad7*
179 which encodes for a regulator of primordial follicle activation together with SMAD3 [31]
180 (Figure S2C), and the follistatin gene (*Fst*) which encodes a secreted inhibitor of the TGFβ
181 pathway, essential for ovarian development [32] (Figure S2D).
182 Comparison of the FOXL2 targets identified, with genes upregulated in the absence of FOXL2
183 (Figure S2E) highlighted a total of 376 genes more likely to be repressed by FOXL2 (because
184 they are upregulated when *Foxl2* is deleted and are FOXL2 targets as assessed by ChIP-Seq).
185 This group included genes encoding for signalling molecules crucial for female sexual
186 development such as *Rspo1* [33] (Figure 1F), and *Wnt4* [34] (Figure S2F), as well as genes
187 important for testis development such as *Gadd45g* [35] (Figure S2G), and *Inhbb* (Figure S2H)
188 [36]. Lastly, we evaluated FOXL2 binding in the genomic region containing *Sox9*, a gene
189 critical for testis development [6, 37] (Figure S2I). We did not detect significant binding of
190 FOXL2 to the *Sox9* promoter at any timepoint. We then scanned the 1 Mb gene desert upstream
191 of *Sox9*, which was previously found to contain cis-regulatory elements essential for *Sox9*
192 expression in Sertoli cells [38, 39]. As previously shown [8, 27], neither TESCO nor enhancer
193 13, were significantly bound by FOXL2 at any point during ovarian development. Instead, we
194 observed a dynamic binding of FOXL2 in other known *Sox9* enhancers. For example, FOXL2
195 was bound to Enh33 at 1W and to Enh32 at E14.5. Significant binding events of FOXL2 to the
196 female-specific *Sox9* Enh8 [38], *Sox9* Enh2 and *Sox9* Enh6 were detected at 8W. The
197 remaining 5 FOXL2 binding events, 4 of which were detected at 8W and 1 at 1W, were
198 distributed in other regions, not previously identified as Sertoli-specific *Sox9* enhancers, and
199 therefore may constitute granulosa-specific enhancers of *Sox9* used by FOXL2 to repress it.
200 In total, by scanning the whole gene desert 1Mb upstream of *Sox9*, we detected one significant
201 peak at E14.5, two at 1W and eight at 8W (Figure S2J).
202 We then clustered the ChIP-Seq peaks into sets that showed similar patterns of variation in
203 FOXL2 occupancy across the timecourse (Figure 2A) and performed Gene Ontology (GO)
204 analysis on each cluster (Figure 2B, Table S1) to investigate which biological processes
205 FOXL2 is regulating across ovarian development.
206

Figure 2

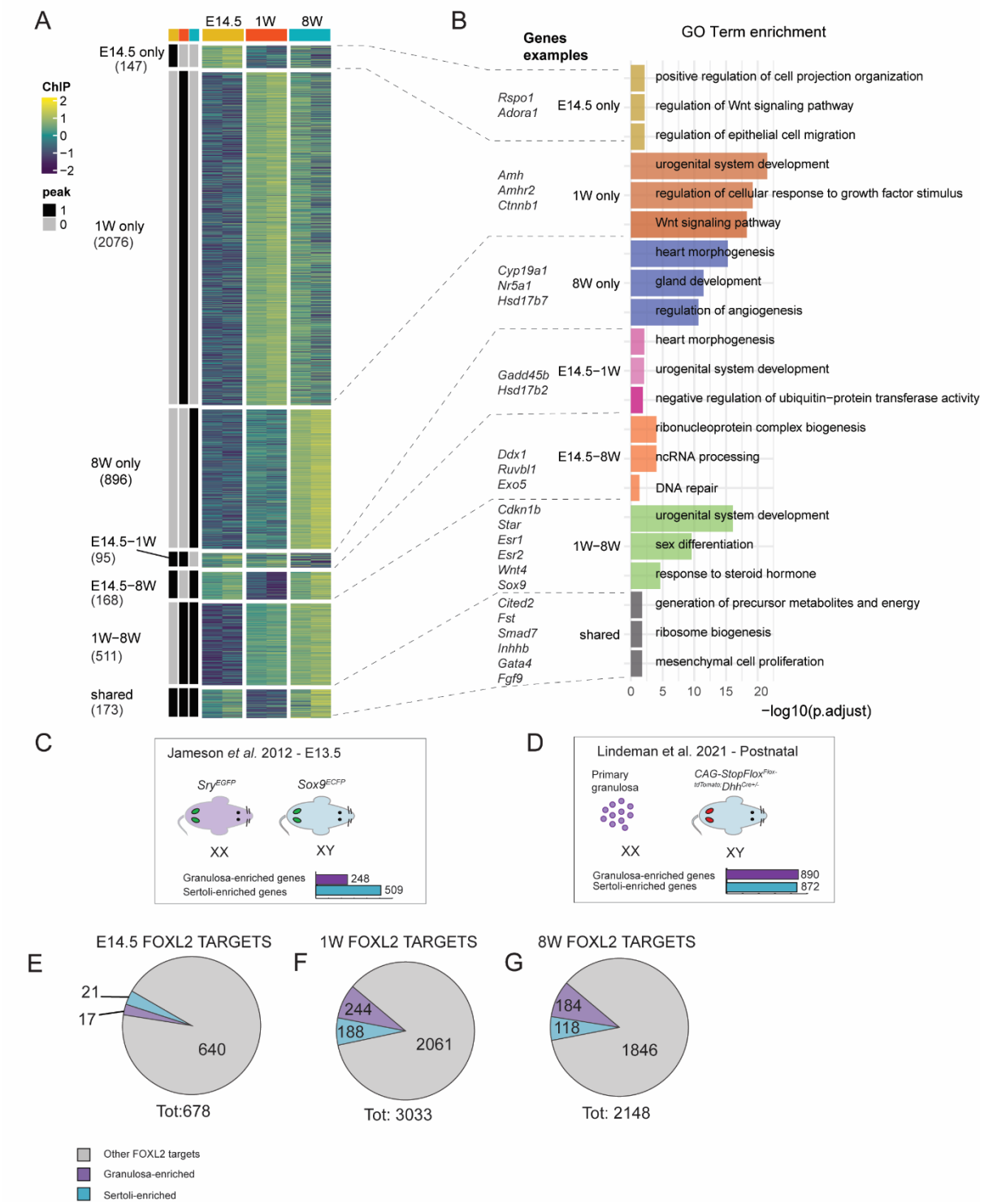


Figure 2: The contribution of FOXL2 to the regulation of granulosa and Sertoli-enriched genes is greater at postnatal stages.

(A) ChIP peak occupancy scores (gene-level). The heatmap has been split based on ChIP peak presence/absence across the conditions. Rows are scaled by z-score to aid visualisation. Black/grey bars on the left indicate presence/absence of a peak in a specific cluster. Number of genes in each cluster is shown below the cluster name. (B) Bar chart showing representative GO biological processes enriched in each dynamic cluster of FOXL2 binding events. See also

Table S1. (C-D) Overview of datasets used for the comparison. Bar charts indicate number of genes enriched in either granulosa or Sertoli cells in the embryonic stage (E13.5, C) and postnatal stage (P7-P29, D). Microarray gene expression analysis performed at E13.5 by Jameson *et al*. [40], RNA-Seq analysis by Lindeman *et al.* [37] of primary granulosa cells from mice between 23-29 days of age compared to Sertoli cells isolated from *CAG-Stop*$^{flox\text{-}tdTomato}$*;DhhCre* P7 XY pups [41]. Significant genes enriched in Sertoli cells are coloured blue, granulosa-enriched genes are coloured purple. (E) Pie chart depicting the proportion of FOXL2 target genes, as identified at E14.5, that are classified as either granulosa- or Sertoli enriched at E13.5 [40]. Grey denotes other FOXL2 target genes identified by our ChIP-SICAP and not overlapping with the lists of granulosa/Sertoli-enriched genes as found in the other two studies. (F) Proportion of FOXL2 target genes, as identified at 1W and (G) 8W, that are classified as either granulosa- or Sertoli and enriched postnatally.

Known FOXL2 target genes were identified amongst those shared across the timecourse and enriched for cell proliferation and ribosome biogenesis pathways, including the granulosa-enriched genes *Cited2*, *Smad7*, *Fst*, *Thbs1*, and the Sertoli-specific genes *Inhbb* and *Fgf9*. *Rspo1*, which encodes for a secreted signalling molecule that regulates WNT4 expression and is essential for female sex determination [42], was a target uniquely regulated by FOXL2 at E14.5, together with other genes including *Adora1*, *Pdgfc*, *Bcl9* and *Tert*, involved in cell migration and proliferation, cell projection organisation and Wnt signalling. Amongst the targets significantly bound by FOXL2 at 1W, we found genes involved in urogenital system development including *Amh*, *Amhr2*, *Fgfr2*, *Ctnnb1* and *Igf2r*. Of those genes uniquely bound by FOXL2 at 8W, we found an enrichment in those implicated in gland development and steroidogenesis pathways. These include known FOXL2 targets such as *Cyp19a1*, the gene encoding for aromatase [43], *Nr5a1*, known to physically interact with FOXL2 [25], and *Hsd17b7*, encoding for a steroidogenic enzyme involved in the estradiol synthesis. Whilst the timepoints E14.5-1W shared several targets involved in urogenital development, the group of genes bound by FOXL2 between 1W and 8W featured genes encoding for members of the steroidogenesis pathway, including the known FOXL2 targets *Esr2*, for which we detected binding in the functional enhancer "Peak 3", located in one intronic region as previously demonstrated [27] (Figure 1E), the human steroidogenic acute regulatory *Star* gene *[44]* and *Cyp17a1*. Additionally, sex differentiation genes were stably bound by FOXL2 between 1W and 8W, and these include granulosa-specific genes such as *Esr1* [45], *Irx3* [46] and *Wnt4* [34], as well as Sertoli-specific genes, such as *Sox9* [37], *Inha* [47], *Fndc3a* [48] and bivalent factors such as *Wt1* [49]. Additionally, we found binding within the cell-cycle repressor gene *Cdkn1b*, a known FOXL2 target gene that regulates granulosa cell proliferation (Figure 2 and Figure S2B) [30]. Lastly, targets shared between E14.5 and 8W were enriched for DNA repair pathways, a recently discovered function of FOXL2 [50], and included genes such as *Ddx1*, *Ruvbl1*, and *Exo5*.

Next, we assessed the relative contribution of FOXL2 to the regulation of granulosa *versus* Sertoli-enriched genes at different developmental stages. To obtain genes enriched in granulosa or Sertoli cells relevant for the embryonic and the postnatal stages of our timecourse, we used datasets from two studies (Table S2): one by Jameson et *al.* reporting microarray analysis of

259 sorted XX (*Sry*^{EGFP}) and XY (*Sox9*^{ECFP}) supporting cells collected at E13.5 [40], and another
260 by Lindeman *et al.* reporting RNA-Seq analysis of primary granulosa cells from mice between
261 23-29 days of age compared to Sertoli cells isolated from *CAG-Stop*$^{flox\text{-}tdTomato}$*;DhhCre* P7 XY
262 pups [41], which constitute a reference for our postnatal timepoints.
263 We calculated the overlap between FOXL2 targets characterising each stage of ovarian
264 development, and the granulosa-, and Sertoli-enriched genes identified by comparing XX and
265 XY supporting cells transcriptomes at either E13.5, for the comparison with E14.5 FOXL2
266 targets (Figure 2C), or postnatally (Figure 2D), for the comparison with 1W and 8W FOXL2
267 target genes. We observed that during the embryonic stage E14.5 only 6% of FOXL2 targets
268 corresponded to either granulosa or Sertoli-enriched genes, whilst this percentage increased at
269 later timepoints, as 15% of FOXL2 target genes identified at 1W or 8W consisted of either
270 granulosa-enriched genes or Sertoli-enriched genes. These comparisons, together with the
271 greater number of targets identified at 1 and 8W compared to E14.5 in our timecourse analysis,
272 support a greater role played by FOXL2 in controlling supporting cell-specific genes
273 postnatally.
274
275
276 Identification of FOXL2 on-chromatin protein interactors
277
278 We next used ChIP-SICAP to identify on-chromatin protein interactors (chromatome) of
279 FOXL2 by applying mass-spectrometry on chromatin samples isolated at different stages of
280 ovarian development (Figure 1A). Proteins co-localising with FOXL2 on the same chromatin
281 fragments were pulled down using a specific antibody against FOXL2 followed by DNA
282 biotinylation to separate chromatin-bound proteins (see Methods). The two-step purification
283 strategy improves the purity of the samples and reduces the artificial interactions that are
284 generated during the immunoprecipitation procedure. We identified a total of 112 proteins at
285 E14.5, 430 at 1W, and 300 at 8W (>4-fold change FOXL2/no antibody control, n=2, adj-
286 *pvalue*<0.1, Fig. 3A and Table S3), thus reconstructing the FOXL2 chromatome across key
287 stages of ovarian development. Comparison of the relative protein intensities showed that the
288 most highly abundant proteins were known nuclear proteins, indicating a successful enrichment
289 for nuclear factors over potential cytoplasmic contaminants (Figure 3B). GO enrichment
290 analysis of FOXL2 interactors showed enrichment of proteins involved in mRNA splicing,
291 transcription, DNA damage and repair, and cell division (Figure 3C and Figure S3). FOXL2
292 was amongst the top 20 mostly enriched proteins identified (Figure 3D-F), together with
293 previously identified FOXL2 nuclear interactors namely GATAD2B, SUPT16H, ATRX,
294 SMARCE1, XRCC6 [51], and TRIM28 [8]. Amongst the FOXL2 protein partners present
295 across the timecourse we found members of the WNT signalling pathway including CTNN1B
296 (β-catenin), TCF4, and TCF21 (Figure S3, "transcription factors" and "transcriptional
297 regulators"), cell cycle regulators such as HMGA2, heterochromatin-interacting proteins such
298 as CBX3 and members of the splicing system including the DNA helicase DHX9, SRSF1, and
299 DDX17.
300

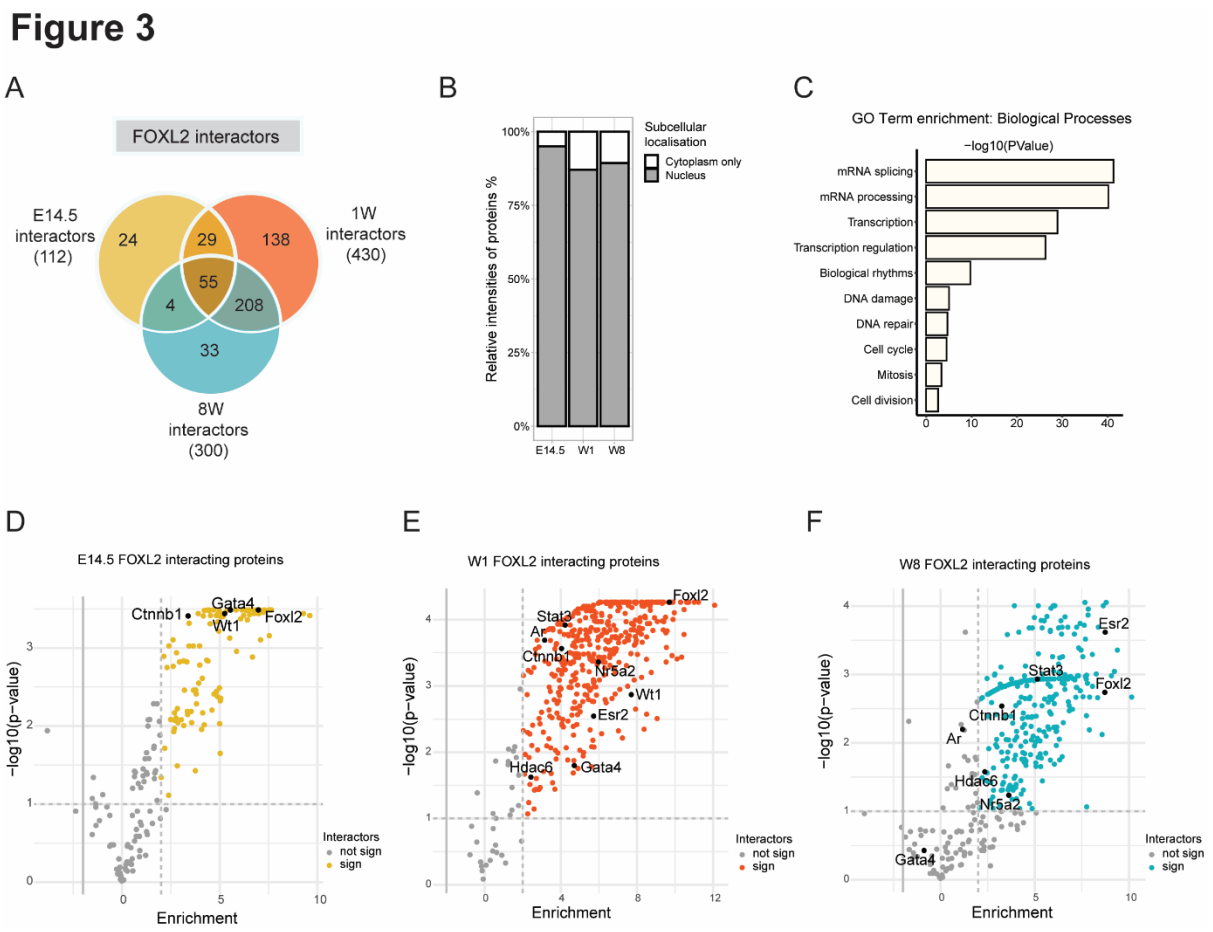


Figure 3: Identification of FOXL2 on-chromatin interactors by ChIP-SICAP

(A) Venn diagram representing the FOXL2 interactors identified at each timepoint by FOXL2 ChIP-SICAP. Total number of proteins identified at each stage is reported between round brackets. (B) Bar chart showing the relative intensities calculated using iBAQ (intensity Based Absolute Quantification) of the proteins identified in each timepoint and classified as either nuclear or cytoplasmatic only (proteins shuttling between nucleus and cytoplasm are classified as nuclear). (C) GO Biological processes enrichment analysis of the FOXL2 interactors. (D-F) Volcano plot showing enrichment (mean fold enrichment >4 and adj. p value <0.1, Benjamini-Hochberg adj. p values, FOXL2 immunoprecipitate *vs* no antibody control, n=2/timepoint) of proteins co-localising with FOLX2 on chromatin at E14.5 (D), 1W (E), and at 8W (F). Dotted lines separate significant proteins, also highlighted in yellow, orange and blue, depending on the timepoint. Black dots: selected FOXL2 interactors.

Importantly, we identified a series of FOXL2-interacting proteins with well-recognised roles in ovarian development and primordial follicle activation in human and mice, being differently present across the timecourse, as follows (Figure 3).

At E14.5 and 1W, factors essential for gonadal development such as GATA4 [52] and WT1 [53] were identified (Figure 3D). These were not detected at 8W (Figure 3F). Our analysis revealed the presence of NR5A2, a newly identified regulator of primordial follicle activation [54], co-localising with FOXL2 at 1W and 8W (Figure 3E-F). Additionally, we detected two

323 additional known regulators of the primordial-to-primary follicle transition, namely the
324 deacetylase HDAC6 [55] and the transcription factor STAT3 [56].
325 We found androgen receptor (AR), a factor that contributes greatly to normal follicle
326 development [22], enriched at 1W, but not later at 8W. Importantly, we identified estrogen
327 receptor 2 (ESR2) amongst the FOXL2 interactors at both postnatal timepoints, in agreement
328 with the previously postulated interplay between these two factors in orchestrating functions
329 key to ovarian physiology such as steroidogenesis [27] and cell fate maintenance [9].
330 We then compared our *in vivo* FOXL2 chromatome with a FOXL2 co-immunoprecipitation,
331 whole-cell proteomics dataset performed on the granulosa-like cell line AT29C [51] to assess
332 the level of overlap between our chromatin-specific dataset and one that would also include
333 cytoplasmic interactors. We found that 111 out of the 490 proteins we identified throughout
334 the timecourse were also found by this study (22%). Of these shared FOXL2 interacting
335 partners, 31 were found in our E14.5 timepoint, 95 at 1W and 72 at 8W (Figure S3A, Table
336 S3). These shared proteins included the known FOXL2 interactor involved in sex maintenance
337 TRIM28 [8], chromatin remodelling enzymes such as ATRX, which is involved in sexual
338 differentiation [57], regulators of the WNT signalling pathway such CXXC5, and DNA repair
339 proteins such as XRCC6, previously identified in a yeast-two-hybrid screen and co-
340 immunoprecipitation study [58].
341 In summary, these results represent the first description of the chromatin interactome of the
342 female-specific transcription factor FOXL2. We showed that this factor, essential for female
343 fertility, differentially interacts with proteins playing a role in a range of gonadal-essential
344 functions, from gonadogenesis, through primordial follicle activation, to steroidogenesis.
345
346
347 Generation of a *Foxl2*^{EGFP} mouse line to isolate granulosa cells.
348
349 We next sought to analyse the transcriptomes and chromatin landscapes of granulosa cells
350 expressing FOXL2, to ultimately integrate these datasets with genomic binding locations of
351 this factor (ChIP-Seq), thus reconstructing the gene regulatory networks at play at different
352 developmental stages of the ovary. We therefore used CRISPR-Cas9 genome editing to
353 generate a *Foxl2*^{EGFP} mouse line, allowing the purification of cells expressing FOXL2 (Figure
354 4A, D and Figure S4, see also Methods). *Foxl2*$^{EGFP+/-}$ ovaries imaged by fluorescence
355 microscopy showed a bright GFP signal specifically in the ovarian tissue at all stages examined
356 (E14.5, 1W, and 8W, Figure 4B). As expected, given the known absence of FOXL2 expression
357 in the testis, no GFP signal was detected in *Foxl2*$^{EGFP+/-}$ males, similarly to *Foxl2*$^{EGFP-/-}$
358 wildtype testes and *Foxl2*$^{EGFP-/-}$ XX wildtype ovaries (Figure 4B and S5A). Colocalization of
359 GFP and endogenous FOXL2 protein in *Foxl2*$^{EGF+/-}$ ovaries and absence of this in *Foxl2*$^{EGFP-/-}$
360 ovaries and *Foxl2*$^{EGFP+/-}$ testes was confirmed by immunofluorescence (Figure 4C). The line
361 was maintained as heterozygotes. Fertility performance of *Foxl2*$^{EGFP+/-}$ females was monitored
362 over 12 weeks and found to be normal compared to control wildtype females (C57BL/6J stock
363 mice, Figure S5B).
364 We dissected ovaries from *Foxl2*$^{EGFP+/-}$ mice at the following timepoints: E14.5, 1W and 8W
365 (Figure 4D). We used fluorescence-activated cell sorting (FACS) to isolate *Foxl2*^{EGFP} positive
366 cells (Figure S6).

Figure 4

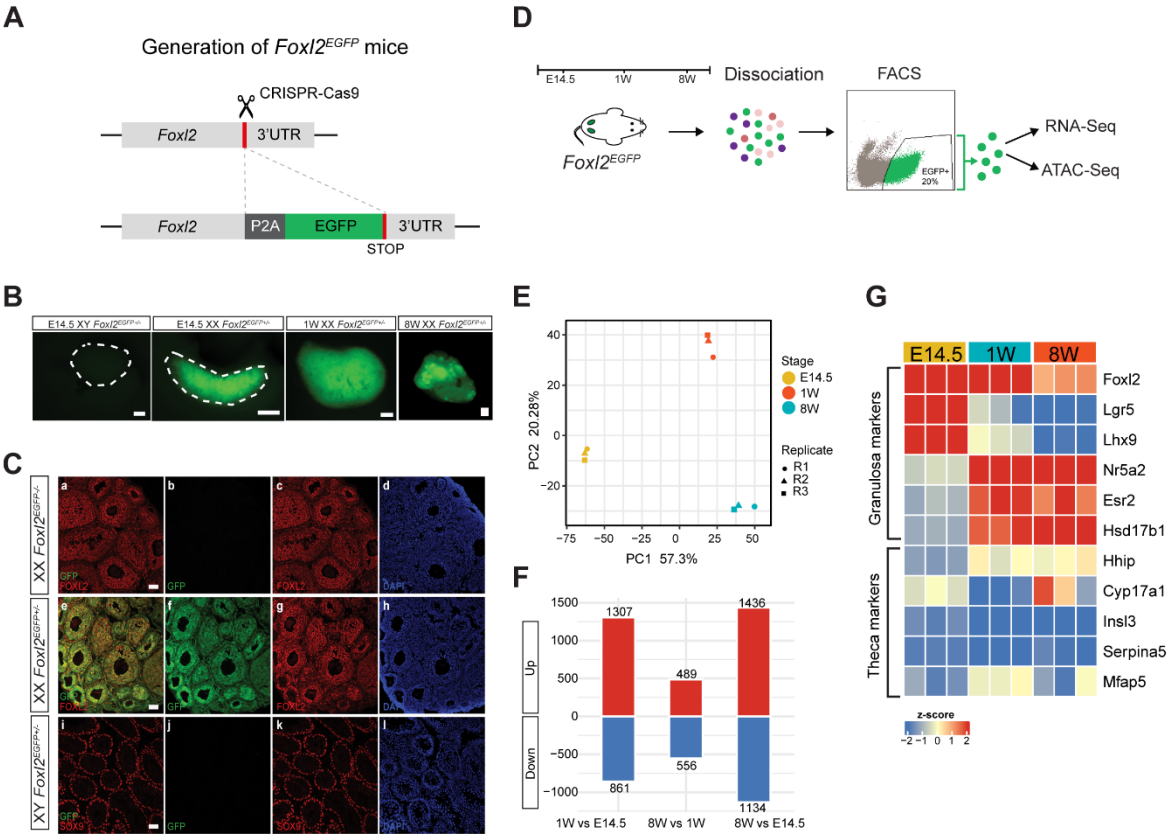


Figure 4: Generation of a *Foxl2*^{EGFP} mouse line to study granulosa cells across ovarian development

(A) Overview of the CRISPR/Cas9 mediated EGFP knock-in at the *Foxl2* locus. (B) Representative images of EGFP fluorescence in freshly collected gonads. Scale bars: 200 µm. Dotted lines outline the embryonic testis (left) and ovary (right). A minimum of three biological replicates were analysed. (C) Immunofluorescence staining of EGFP and endogenous FOXL2 in adult XX *Foxl2*$^{EGFP-/-}$ ovaries (a-d), XX *Foxl2* $^{EGFP+/-}$ ovaries (e-h), and of EGFP and endogenous SOX9 in XY *Foxl2* $^{EGFP+/-}$ testes. Scale bars: 100 µm. (D) Overview of sample collection and downstream analyses performed. (E) PCA plot of RNA-Seq analysis of cells isolated from E14.5, 1W and 8W gonads (50k cells/replicate, n=3 biological replicates/timepoint). (F) Barplot summary illustrating the number of differentially expressed genes (DEG) identified in each pairwise comparison (DESeq2's Wald test, *pvalue*<0.01, fold-change of >2) (G) Heatmap of expression changes in genes validated as theca or granulosa cell markers.

A total of 50k and 100k of *Foxl2*^{EGFP} positive cells were used to perform bulk RNA-Seq and ATAC-Seq respectively, using three biological replicates for each timepoint (See Methods). PCA of RNA-Seq data (Figure 4E) based on the variance stabilised abundance estimates showed good separation between the timepoints indicating major transcriptional changes characterising granulosa *Foxl2*^{EGFP} positive cells throughout time. Most of the variance in the dataset was evident along the first principal component (57.3% variance), which separated

390 E14.5 samples from both W1 and W8. Poisson similarity metric indicated more similarity
391 between the transcriptome of postnatal samples (1W and 8W) compared to E14.5 (Figure S7A).
392 To examine the global transcriptome dynamics, we identify and selected genes significantly
393 changing across the timecourse in at least one pairwise comparison (Table S4 and Figure S7B).
394 Differential gene expression analysis comparing 8W and E14.5 revealed the greatest changes
395 with 1436 genes upregulated and 1134 downregulated; comparing 1W with E14.5 revealed
396 upregulation of 1307 genes and downregulation of 861 whilst the least number of changes were
397 detected between 8W and 1W, with 489 genes upregulated and 556 genes downregulated
398 (Figure 4F). Granulosa cells and the steroidogenic cell lineage including theca cells both
399 express FOXL2 [59]. We therefore checked expression of proven markers [60] expressed
400 exclusively in either theca or granulosa cells (Figure 4 G). This analysis indicated enrichment
401 of granulosa cells over theca cells in cells isolated using the *Foxl2*^{EGFP} transgene.
402 We then used k-means clustering to analyse the dynamics of gene expression changes across
403 time (Figure S7B-C) and identified seven main patterns of gene expression. Clusters 7 and 6
404 consisted of genes highly expressed at E14.5 including markers of pre-granulosa cells such as
405 *Rspo1*, *Lgr5* and *Wnt4*, important for the initial proliferation of embryonic gonads [28]. GO
406 Biological processes analysis (GO BP, Figure S7D) of these clusters revealed enrichment of
407 biological processes such as Wnt signalling (*Wnt9a*, *Axin2*, *Fgfr2*, *Rspo1*) and urogenital
408 system development pathways (*Cdkn1b*, *Wt1*, *Wnt4*, *Cdkn1c*, *Rara*, *Emx2*, *Yap1*). Cluster 5
409 included genes moderately expressed at E14.5 and upregulated specifically at 1W and included
410 genes involved in kidney and urogenital system development (*Bmp7*, *Gli3*, *Smad3*, *Hey1*),
411 genetic imprinting (*Zfp57*, *Airn*, *Xist*), and extracellular matrix organisation (*Col18a1*,
412 *Adamts9*, *Eln*). Cluster 1 consisted of genes uniquely up regulated at 1W and participating in
413 extracellular matrix organisation, cell cycle phase transition, and DNA replication (*Pola2*,
414 *Prim1*, *Rad51*, *Cdc7*, *Dna2*). Cluster 4 grouped genes activated as the ovary develops from 1W
415 and remain upregulated in adulthood at 8W. GO analysis showed enrichment in nuclear
416 division, cell cycle (*Ube2c*, *Fbxo5*, *Top2a*), lipid transport, and synapses organisation (*Sema3e*,
417 *Tanc2*, *Ctnna2*). Cluster 3 included genes specifically upregulated at 8W and enriched in
418 steroid metabolic processes (*Star*, *Hsd17b1*, *Cyp19a1*) and angiogenesis (*Sema5a*, *Cd36*, *Jak1*,
419 *Gadd45a*). Lastly, Cluster 2 included genes expressed at the two divergent timepoints of E14.5
420 and 8W and characterised by high representation of lipid metabolism (*Apoe*, *Cyp17a1*, *Runx1*),
421 purine ribonucleotide biosynthesis (*Impdh1*, *Stat3*, *Bcl2l1*) and small molecule catabolic
422 processes (*Aldh1l1*, *Fah*, *Bcat2*). These data show the timely activation of genes important for
423 embryonic and postnatal ovarian development and function in our system.
424
425
426 <u>An integrative approach to reconstruct gene regulatory networks controlled by FOXL2.</u>
427
428 Having identified the gene expression changes characterising *Foxl2*$^{EGFP+/-}$ granulosa cells, we
429 then sought to reconstruct the gene regulatory networks (GRN) underpinned by this factor
430 (Figure 5 and Table S5).
431 We took the total number of FOXL2-bound genes identified by ChIP-Seq at any given
432 timepoint (Figure 1) and intersected it with the entire set of differentially expressed genes
433 (DEG) identified by RNA-Seq (Figure 4). Genes that changed in their expression value at any

point between E14.5, 1W and 8W (3333 genes) and that were assigned to a FOXL2 ChIP peak in at least one timepoint (4066 genes) were considered as "dynamic/FOXL2-bound" genes. The rest of the genes bound by FOXL2 but not changing their expression values were considered as "stable" (2966 in total) and divided between "stably repressed" and "stably activated", based on a variance stabilised (VST) cut-off of 8.3 (Figure S8 A-B). This threshold was chosen based on the expression values of known granulosa-enriched genes (*Foxl2*, *Esr2*, *Rspo1*), versus genes residing on the Y chromosome therefore not expected to be expressed at all in XX ovaries (*Sry*, *Usp9y* and *Uty*), as well as on the expression values of Sertoli-enriched genes (*Dmrt1*, *Sox9* and *Fgf9*) (Figure S8C).

Figure 5

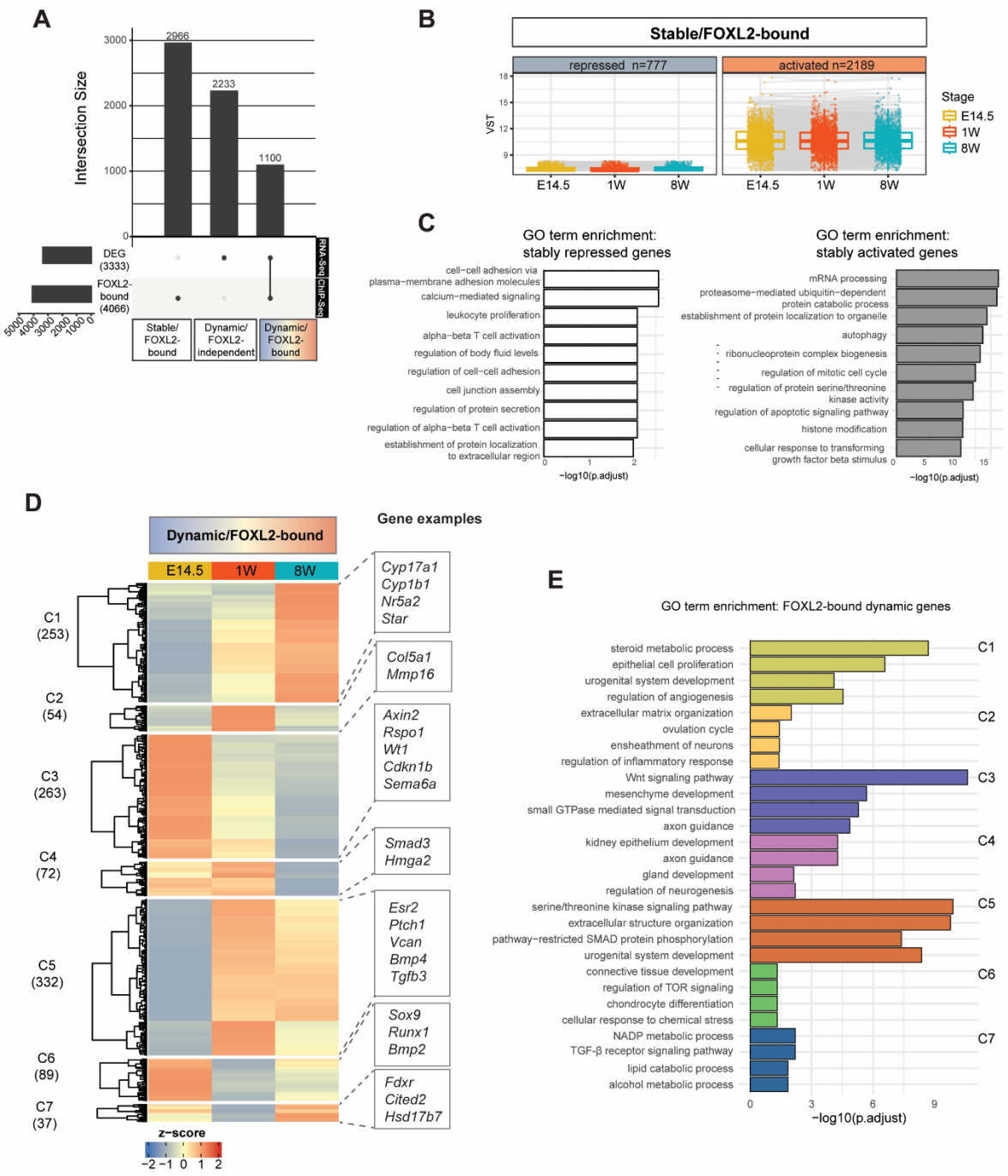


Figure 5: Integrative approach combining RNA-Seq and ChIP-Seq to refine the gene regulatory networks controlled by FOXL2.

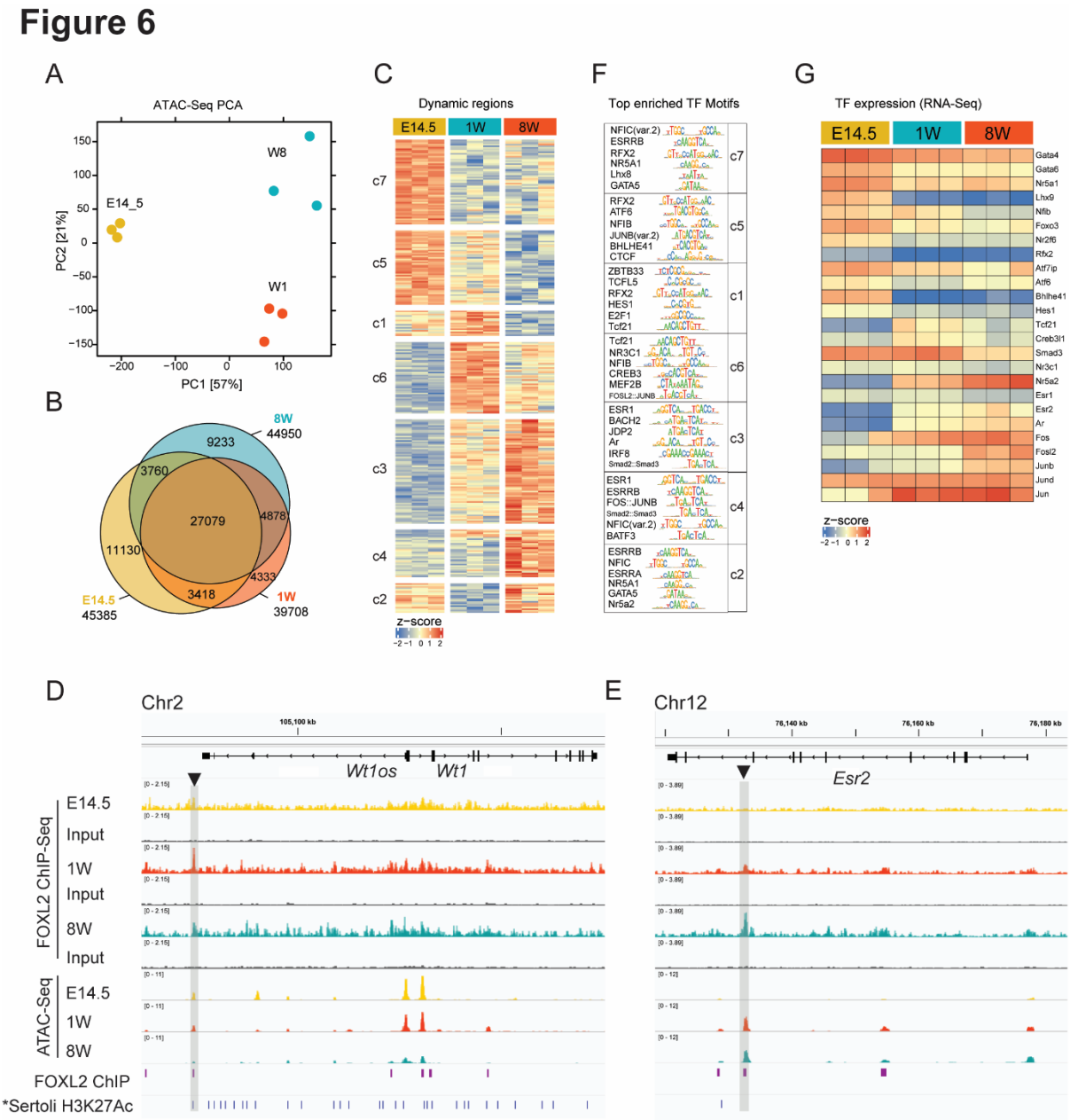
(A) Upset plot depicting the number of differentially expressed genes (DEG) identified by RNA-Seq, FOXL2 gene targets identified by ChIP-SICAP and the overlap between the two datasets. (B) Plot of variance stabilised (VST) gene expression values as detected by RNA-Seq on *Foxl2* $^{EGFP+/-}$ sorted ovarian cells (n=3). A threshold of VTS=8.3 was used to differentiate stably repressed from stably activated genes. (C) GO Biological processes enrichment analysis of the two groups of genes, stably repressed (left), and stably activated (right). (D) Hierarchical clustering of the 1100 genes bound by FOXL2 and differentially expressed across the timecourse. Gene examples from each cluster are shown on the right. (E) GO enrichment analysis of the clusters. Four representative pathways are depicted for each cluster.

Using this approach, we identified a total of 777 genes bound by FOXL2, but which are not expressed at any point during the timecourse and are therefore assumed to be repressed by FOXL2 (Figure 5B and Figure S8B). These were enriched in cell-cell adhesion, immune, and protein secretion pathways (Figure 5C). These included also Sertoli-enriched genes such as *Spz1* [61], *Cdh8*, *Rims1*, *Fgf9* [62], *Hsd17b3*, *Hopx* and *Tuba3* [41]. In contrast, a significant higher number of genes were bound by FOXL2 and were expressed throughout the timecourse but did not significantly change in levels across this (2189 genes). These were enriched in mRNA processing, ubiquitin-dependent catabolic processes, cell cycle and histone modification pathways (Figure 5C). In contrast, the "dynamic/FOXL2-bound" group of genes contained a total of 1100 genes (Figure 5D). We used unsupervised hierarchical clustering to analyse their pattern of expression throughout the timecourse and identified 7 main clusters (Figure 5D). GO enrichment analysis of these clusters showed that FOXL2 regulates a plethora of dynamically expressed genes enriched for functions critical for different stages of ovarian development. For example, we found enrichment in steroid metabolic and angiogenic processes, typical of folliculogenesis, in cluster 1 (C1), which included genes upregulated specifically at 8W. Of these, 53% were also a direct target of FOXL2 specifically at 8W (Figure 5D-E, Table S5). On the contrary, if considering C3, C4 and C6, clusters included genes upregulated at E14.5 compared to 8W, only 12% of the genes were significant targets of FOXL2 at E14.5. This confirmed the preponderant role of FOXL2 in regulating significantly changing genes at postnatal rather than during embryonic development. The high enrichment of Wnt signalling pathways found in C3, which contained a total of 263 genes upregulated at E14.5 only 36 of which were bound by FOXL2 at this timepoint (including 8 Wnt signalling pathway effectors such as *Rspo1*, *Zfp703* and *Bcl9*), was consistent with the established role of this pathway in ovarian development [34].

<u>*Foxl2*^{EGFP} positive cells have a dynamic chromatin landscape across ovarian development</u>

To explore the dynamics of chromatin remodelling underlying the gene expression changes throughout granulosa cells development, we performed ATAC-Seq on *Foxl2*^{EGFP} positive cells purified with FACS (Figure 6) at three timepoints (E14.5, 1W and 8W). PCA of genome-wide

488 chromatin accessibility on a total of 63831 consensus peaks (n=3) showed that most of the
489 differences in chromatin landscape (57% variance) were found between the embryonic
490 timepoint E14.5 and the late timepoints 1W and 8W (Figure 6A), in agreement with the
491 transcriptional changes shown in Figure 4.
492



493

494 **Figure 6: Granulosa cells have a dynamic chromatin landscape throughout development.**

495 (A) PCA of ATAC-Seq performed on *Foxl2* $^{EGFP+/-}$ sorted ovarian cells (n=3) at E14.5, 1W and
496 8W. (B) Venn diagram reporting the number of significant peaks identified at each stage. (C)
497 Hierarchical clustering of chromatin landscape remodelling. (D-E) IGV snapshots of
498 representative regions. *Sertoli H3K27Ac ChIP-Seq was re-analysed to match mm10
499 coordinates from Maatouk e*t al* [63]. (F) MonaLisa analysis of TF motifs enrichment analysis
500 within the open chromatin regions as detected by ATAC-Seq. Top 6 motifs for each cluster are

501 reported (Table S6). (G) Heatmap displaying gene expression changes of representative TF, as
502 assessed by RNA-Seq (Table S4).

503

504

505 The second largest difference (21% variance) was observed between 1W and 8W, when
506 additional remodelling occurs in granulosa cells as folliculogenesis starts. Poisson metric also
507 confirmed a higher similarity between 1W and 8W samples, whilst E14.5 timepoint was
508 clustering separately (Figure S9A).
509 In total, we identified 45385 accessible regions at E14.5, 39708 at 1W and 44950 at 8W (Figure
510 6B, Table S6). Of these, 27079 were accessible throughout the timecourse, while 11130 were
511 unique to E14.5, 4333 unique to 1W, and 9233 unique to 8W. The genomic annotation of open
512 chromatin regions was similar across timepoints, with 32-35% mapping to promoters, 34-38%
513 to the gene body, 3% downstream of a gene, and 25-28 % mapping to distal intergenic regions
514 (Figure S9B). K-means clustering of consensus peaks revealed a total of 7 patterns representing
515 similar dynamic opening and closing of chromatin regions across ovarian development (Figure
516 6C). To assess the potential contribution of FOXL2 in binding to these open chromatin regions,
517 we specifically looked for enrichment values for the FOXL2 canonical motif (JASPAR id:
518 MA1607.1) in each ATAC-Seq cluster (Figure S9C-D). Motif analysis of these dynamic
519 regions revealed a greater enrichment for the canonical FOXL2 motif in clusters 6, 3, 4, 2 and
520 the least in cluster 7 (fold enrichment 1.4-1.2, adjPvalue>4). Amongst these, clusters 6 and 3,
521 which included peaks accessible preferentially at 1W and 8W and closed at E14.5, were the
522 two most significantly enriched with FOXL2 motifs (Figure S9C-D). In contrast, amongst
523 regions accessible at E14.5 (c1, c5 and c7) only c7 contained significant enrichment of FOXL2
524 binding sites, albeit with an enrichment score much lower than c6 and c3, containing elements
525 opened at 1 and 8W.
526 Differential accessibility analysis was then performed using DESeq2 [64] (Figure S10, Table
527 S6). This analysis revealed a total of 40165 differentially accessible regions (DARs) throughout
528 the timecourse, 32611 of which were changing between E14.5 and 8W, 24026 were changing
529 between E14.5 and 1W, and 16054 were changing between 1W and 8W (Figure S10A-D).
530 These chromatin landscape changes were consistent with the transcriptomic data, which
531 showed that most of the differences were detected comparing E14.5 and 8W samples (Figure
532 S7A).
533 Our profiling of chromatin accessibility dynamics contained novel potential cis-regulatory
534 regions, for example one within *Rspo1* gene, differentially bound by FOXL2 according to our
535 ChIP-Seq analysis (Figure 1F), and showing a similar change in its accessibility profile, as
536 indicated by ATAC-Seq (Figure S10E). This region was highly conserved and contained a
537 candidate cis-regulatory element sequence as predicted by the ENCODE Registry [65].
538 We also detected FOXL2 binding by ChIP-Seq at a known *Wt1* enhancer located 50 kb
539 upstream of its promoter [63] (Figure 6D). This binding event correlated with closing of this
540 region as ovarian development progressed, with no opening detected at 8W. In contrast, *Esr2*
541 was bound by FOXL2 at its known functional enhancer, located in the 8th intron of the gene,
542 only at postnatal stages (Figure 6E) [27]. This coincided with an opening of this region at 1W

543 and 8W. This enhancer was not marked by H3K27Ac (active enhancer mark) in Sertoli cells,
544 therefore can be considered as a granulosa-specific cis-regulatory region of *Esr2*.
545 K-means clustering of the changes identified seven clusters representative of the main
546 chromatin remodelling patterns (Figure 6C). Cluster 7 (c7) included peaks accessible at E14.5,
547 followed by c5, which included peaks significantly opened at E14.5 and starting to become
548 accessible at 1W. c1 included peaks closing down at E14.5 and opening up at 1W, whilst
549 remaining closed at 8W. c6 contained the majority of peaks accessible at 1W, which are also
550 starting to open up at 8W. c3 includes peaks progressively closing at 1W and opening at 8W,
551 while c4 included peaks that have gained accessibility at 8W. Lastly, c2 contained a group of
552 2713 peaks opened at the embryonic stage, E14.5, and in adulthood, at 8W, but not at 1W
553 (Figure 6C). Amongst the DARs we identified regions located nearby granulosa cell markers
554 such as *Fst* [66] and *Runx1* [10] which were highly conserved (Figure S10F). Additionally, we
555 identified open elements nearby markers of Sertoli cells, including structural markers such as
556 *Cldn11* and *Itga6* (Figure S11A-B), hormonal markers such as *Amh* and *Hsd17b1* (Figure
557 S11C-D), and genes encoding for transcription factors such as *Fgf9* and *Nr5a1* (Figure S11E-
558 F). These were all bound by FOXL2 as shown by our ChIP-SICAP.
559 To investigate what transcription factors may bind to these dynamic regions that could
560 potentially drive the observed gene expression changes or that may direct changes in chromatin
561 organisation, we performed a transcription factor (TF) motif enrichment analysis of the DAR
562 dynamics (Figure 6F-G, Table S5). We selected only the significant DARs and used monaLisa
563 [67] to identify TF binding motifs enriched in accessible peaks. Regions that were more
564 accessible at E14.5, but lost accessibility at later stages (namely c7 and c5) were enriched with
565 nuclear factor family motifs (e.g. NFIC) and GATA motifs (GATA5, GATA3). NR5A1 motifs
566 were highly represented, as well as motifs from the RFX (RFX2, RFX5). Steroid receptor
567 motifs (ESR1 and ESRRB) were also enriched. Cluster 5 was also particularly enriched in
568 motifs typical of the transcriptional repressor CTCF. Regions accessible at 1W (c1 and c6)
569 were enriched in motifs for factors involved in Wnt signalling (TCFL5, TCF21) and Notch
570 signalling (HES1) and for ZBTB33 motifs, a transcriptional repressor also known as Kaiso,
571 which negatively regulates the Wnt signalling pathway [68]. Amongst these, TCF21, ATF1
572 were also found to localise with FOXL2 on chromatin, as we found by FOXL2 ChIP-SICAP
573 (Table S3 & Table S5). Clusters containing regions opened at 8W were highly enriched with
574 motifs for members of the steroid hormone receptor family, including ESRRB (containing a
575 half-site for the binding of ESR2) and AR, factors known to play critical roles in
576 folliculogenesis and ovulation [22, 69], which were also found in our proteomics analysis
577 (Figure 3). ESRRB motifs were also enriched at E14.5 although no expression of *Esr2* mRNA
578 was found at this timepoint by RNA-Seq (Figure 6G). Additionally, we noticed a remarkable
579 enrichment in motifs of members of the Jun and Fos family of transcription factors. Lastly, in
580 c2 we found enrichment for the TF NR5A2, known to regulate primordial follicle activation
581 [54], which was also detected with FOXL2 ChIP-SICAP at 1W and 8W timepoints (Figure 3).
582 RNA-Seq analysis revealed that some of these factors also show an expression pattern
583 matching the binding motif predictions (Figure 6G). For example, GATA4 and LHX9
584 expression was increased at E14.5, whilst factors predicted to bind at postnatal stages including
585 NR5A2, AR and ESR2, were upregulated from 1 week onwards.
586

587 **DISCUSSION**

588 Previous work has highlighted the importance of FOXL2 in many aspects of female sexual
589 development[21] as well as in granulosa cell-related pathologies[70]. It is a sex-determining
590 gene in goats [71], mutations of FOXL2 in humans lead to primary ovarian insufficiency [13],
591 and it was shown to cooperate with other factors to establish ovarian sexual identity in the
592 mouse embryonic gonad [72, 73]. Studies from our laboratory and that of Mathias Treier also
593 demonstrated that FOXL2 becomes the sole gatekeeper of ovarian cell fate maintenance in
594 adulthood [9].

595 In this study, we describe the set of FOXL2 genomic targets and on-chromatin protein
596 interactors, as well as the gene expression changes and chromatin landscape dynamics
597 characterising granulosa cells throughout mouse ovarian development. We found that FOXL2
598 binds to a greater number of genes at postnatal stages (1W and 8W) compared to embryonic
599 (E14.5). The greatest number of targets overall was found at 1W, corresponding to the first
600 timepoint at which ectopic SOX9 expression is observed in *Foxl2* null mutants [17]. These data
601 support a more critical role for this factor during postnatal life compared to the embryo, at least
602 in mice. GO analysis of FOXL2 targets at E14.5 showed enrichment of Wnt signalling, in
603 keeping with the synergistic role played by these two pathways in reinforcing the female cell
604 fate commitment in the embryonic gonad [72]. The Wnt pathway has recently been shown to
605 be essential for pre-granulosa cell activation during primordial follicle activation [74]. This
606 pathway was also highly featured amongst FOXL2 target genes identified at 1W, supporting
607 the idea that a possible interplay between Wnt and FOXL2 around the time of primordial
608 follicle activation may happen through regulation of similar target genes. Targets bound at 8W,
609 once folliculogenesis is ongoing and mice have reached sexual maturity, are enriched for
610 steroidogenesis and DNA repair pathways. Notably, we found FOXL2 to directly bind to a
611 known ESR2 enhancer [27] specifically at 8W.

612 Indeed, in addition to binding sites for forkhead factors, FOXL2 peaks were highly enriched in
613 sites for steroid receptors including ESR1 and AR. This supports previous studies performed
614 in primary granulosa cell cultures which suggested a major role for FOXL2 on estradiol
615 signalling through its direct regulation of *Esr2* [27].
616 One of the main roles of FOXL2 is direct repression of SOX9 [9]. Our data show that this
617 repression may be time-specific, as we detected a greater number of FOXL2 peaks in the 'gene
618 desert' regulatory region 5' to *Sox9* at 8W compared to earlier timepoints. Although this does
619 not entirely exclude that the presence of a single peak at earlier timepoints may suffice and act
620 as repressor site for *Sox9*. In agreement with other studies, we did not find binding of FOXL2
621 to TESCO [8, 21, 75], nor to the *Sox9*-specific Enh13 [38, 76], but we detected peaks in other
622 regions, mostly not overlapping with Sertoli-specific *Sox9* enhancers [38]. The granulosa
623 specific *Sox9* Enh8 [38] appeared significantly bound by FOXL2 only at 8W. Some binding
624 seems to be present beforehand, however this was not considered significant because of our
625 stringent selection criteria for ChIP binding events (that is, a peak was considered only if
626 present above input in both replicates). FOXL2 appeared to bind to several Sertoli-enriched
627 genes, various with structural roles (*Cldn11*, *Itga6*) or involved in hormone signalling
628 (*Hsd17b1*, *Amh, Inhbb*). Comparison with published datasets of gene expression data obtained

629 from *Foxl2* null mutants ovaries [29], and from Sertoli and granulosa cells obtained at
630 embryonic and postnatal stages [41], also indicates a more prominent role of FOXL2
631 postnatally.
632 All together these data suggest that FOXL2 may repress testis development postnatally through
633 alternative *Sox9* enhancers, many of which are granulosa-specific, but also by directly
634 repressing genes important for Sertoli cell proliferation and function. The latter may as well be
635 the most important function of FOXL2 beyond repressing *Sox9*. This was suggested to be the
636 case in the conditional TRIM28 null mutant ovaries, where expression of Sertoli structural
637 genes precedes *Sox9* upregulation in an intermediate cell type along the female-to-male
638 transdifferentiation trajectory [8].
639 We then asked the question of which binding partners were co-localising on the same
640 chromatin regions with FOXL2 that could aid the multiple roles played by this factor along
641 ovarian development. Our ChIP-SICAP analysis represents the first *in vivo* analysis of
642 chromatin binding partners of FOXL2 and revealed that this factor interacts with proteins
643 crucial for each developmental milestone of ovarian development. For example, we found that
644 FOXL2 colocalises with the transcription coactivator of Wnt signalling Ctnnb1 (also known as
645 β-catenin) throughout the timecourse. It was previously shown that β-catenin stabilization and
646 activation in a normal XY testis leads to male-to-female sex-reversal [77] and is accompanied
647 by an increased expression of FOXL2 and reduced expression of SOX9 in the male gonad. The
648 co-localisation of β-catenin and FOXL2 on the same genomic locations observed in our study
649 may represent a mechanism to reinforce ovarian cell fate, suggesting that the functional
650 interplay between the two pathways may rely on their joint recruitment to DNA. This was also
651 supported by the identification of several other members of the pathway including TCF factors.
652 FOXL2 co-localised additionally with Wilms' tumour suppressor (WT1) at E14.5 and 1W.
653 WT1 is a transcriptional regulator essential for early development of the gonads [53] which is
654 found in two main splicing isoforms: +KTS, containing the amino acid trio KTS between the
655 last two zinc fingers, and -KTS, which lacks this. Importantly, only -KTS can act as a
656 transcriptional regulator because the addition of the +KTS portion abrogates WT1 ability to
657 bind to DNA, confining it to nuclear speckles [78, 79]. Given this knowledge and the fact that
658 our FOXL2 ChIP-SICAP experiments specifically identify only proteins bound to the
659 chromatin, which is typically absent from nuclear speckles [80], our data support an interplay
660 on the DNA of FOXL2 and the splicing isoform WT1 -KTS, which may aid the suggested role
661 of the latter in early ovarian development [81, 82]. We also identified interaction of FOXL2
662 with GATA4, which is essential for genital ridge formation in both sexes [52] and is known to
663 contribute to ovarian development and function [83, 84]. This is of interest because it was
664 previously shown that deficiency in GATA4-FOG2 interaction impaired ovarian development
665 through the alteration of the female gene expression program, including having a direct
666 negative effect on *Foxl2* expression [85]. Our data indicate that this novel FOXL2-GATA4
667 interaction may be at the base of GATA4's role in ovarian development.
668 FOXL2 is known to control the quiescence of primordial follicles by inhibiting the initiation
669 of the growth of oocytes. This is shown in *Foxl2* null mutants where the primordial follicles
670 form normally, but they do not develop into secondary follicles because pre-granulosa cells do
671 not undergo the necessary cuboidal transition [12]. FOXL2 was also shown to be differentially

expressed in subpopulations of pre-granulosa cells during primordial follicle formation being downregulated in primary compared to primordial granulosa cells [20, 56].

In our study we identified four interactors known to have a role in primordial follicle activation that had been presumed to be FOXL2-independent, namely NR5A2 [54], HDAC1 [86], HDAC6 [55] and CSNK1A1[87]. This indicates that FOXL2 may create a chromatin hub where factors regulating similar processes can interact. At 1 and 8W, we found FOXL2 interacting with ERS2, therefore confirming that the previously postulated interaction between FOXL2 and estrogen signalling is mediated through the physical interaction of ESR2 and FOXL2 at proximal genomic locations [27, 88]. In addition, we also detected AR at 1W in our proteomics analysis, which also confirms the motif analysis of our ChIP-SICAP data showing enrichment of AR motifs in the FOXL2 binding sites. Androgen receptor has a dual role in the two sexes, regulating granulosa cell proliferation in the ovary [22], while it is critical to ensure spermatogenesis is supported by Sertoli cells in the testis [89]. The co-localisation of AR and FOXL2 may hint to a role for AR in the transdifferentiation phenotype observed in the conditional *Foxl2* mutants [9]. Lastly, we also identified members of the DNA repair pathway, such as XRCC6, which was also previously implicated in a newly discovered role played by FOXL2 in DNA damage repair [50], alongside potential novel players such as PDS5B [90] and DHX9 [91].

As part of the studies reported here, we generated a mouse line to allow the purification of FOXL2-positive cells at all stages of ovarian development (*Foxl2*^{EGFP}). This valuable genetic tool allowed us to generate age-matched transcriptome and chromatin landscape datasets to integrate with our ChIP-SICAP data on the genomic binding sites of FOXL2 and be able to draw conclusions as to which genes may be repressed or activated by this factor. FOXL2 is also expressed in theca cells [59], therefore, we first assessed the expression values of known theca and granulosa cell markers [60]. We found an enrichment of granulosa over theca markers, confirming the suitability of this dataset for the exploration of granulosa-specific gene regulatory networks. Our bulk RNA-Seq analysis shows expression of the pathways expected to be activated at each developmental stage, notably an enrichment of WNT signalling and urogenital development at E14.5, ECM remodelling and cell cycle at 1W, steroidogenesis and angiogenesis at 8W.

By integrating the gene expression values with FOXL2 ChIP-SICAP we were able to identify a high-confidence set of genes likely to be repressed, activated, or dynamically regulated by FOXL2 across ovarian development. Our analysis showed that amongst all FOXL2 target genes, the majority was either stably expressed throughout the timecourse, or dynamic. A lower number of genes were bound by FOXL2 and stably repressed. Amongst the stably expressed genes, we found genes with a role in mRNA processing, apoptosis and cell cycle, function which have all been linked with FOXL2 in functional studies [92]. Analysis of those that are dynamically expressed and controlled by FOXL2 provided an overview of the range of functions critical for the ovary that this factor may regulate, from steroidogenesis and extracellular matrix organisation to lipid catabolic processes. Genes involved in axon guidance were also highlighted, which is of interest as the ovary is innervated, whereas the testis is not [93].

A smaller number of genes was stably repressed, and this group included genes encoding for Sertoli structural proteins suggesting that FOXL2 may inhibit the male pathway through the

regulation of genes critical for Sertoli cell function, beside regulating *Sox9* as previously suggested by Uhlenhaut *et al.* [9]. Our findings are supported by another adult sex reversal phenotype caused by gonadal deletion of TRIM28 in which authors found expression of structural Sertoli cell markers preceding that of *Sox9* [8]. This evidence was further corroborated by our ATAC-Seq analysis of the chromatin landscape of granulosa cells. Here, we found that in agreement with the transcriptomic analysis, extensive chromatin remodelling occurred across ovarian development. Importantly, we found that FOXL2 motifs were less enriched in chromatin regions opened exclusively at E14.5, whilst their presence increased significantly at 1 and 8W. This may suggest a more preponderant role of FOXL2 in regulating genes postnatally, which agrees with the greater number of FOXL2 binding sites found at 1 and 8W compared to E14.5 This analysis led to the identification of several potential enhancers of genes important for granulosa cell differentiation for example *Rspo1* and *Runx1*. Particularly in the latter, we found several FOXL2 binding sites coinciding with open chromatin regions, which is in agreement with the interplay between these two factors in regulating ovarian cell fate maintenance [10]. Finally, the integration of ATAC-Seq with FOXL2 ChIP-SICAP and comparison with ATAC-Seq datasets from sorted Sertoli cells [41, 94], also allowed us to observe that genes important for Sertoli cells (*Amh*, *Cldn11*, *Itga6*, *Fgf9*, *Hsd17b1* amongst others) contained regions opened in both Sertoli and granulosa cells and that these were bound by FOXL2 in granulosa cells. This supports the concept that the role of FOXL2 in repressing the male pathway may go beyond only repressing *Sox9* and may unfold additionally through the direct repression of genes encoding proteins critical for the structure and function of Sertoli cells.

In summary, our study represents the first description of the chromatin interactome of the ovarian-specific transcription factor FOXL2 *in vivo*. We showed that this factor, essential for female fertility, differentially interacts with proteins playing a role in a range of gonadal-essential functions, from gonadogenesis, through primordial follicle activation, to steroidogenesis and that this action is stage-specific. Further interrogation and mining of this dataset may provide novel insights into the FOXL2-centric mechanisms at play in the ovary during development and in the adult.

746 **MATERIALS AND METHODS**

747

748 **Mice and timed mating**

749 All animal experiments carried out were approved under the UK Animals (Scientific
750 Procedures) Act 1986 and under the project licences n. 80/2405 and PP8826065. *Foxl2*$^{EGFP+/-}$
751 mice were maintained on a C57Bl6 background. *Foxl2*$^{EGFP+/-}$ embryos were produced by
752 crossing heterozygotes with C57Bl6 wildtype animals and collected at E14.5 after timed
753 matings. Noon of the day of plug was considered as E0.5. Litters of 1 week of age were sexed
754 and females were culled to harvest the 1W timepoint. Fresh tissue was used for ATAC-Seq and
755 RNA-Seq, whilst ovaries for ChIP-SICAP were snap-frozen and stored at -80°C until the day
756 of the experiment.

757

758 **Generation of *Foxl2*^{EGFP} reporter strain**
759 The *Foxl2*^{EGFP} reporter strain was generated by the Genetic Modification Service at the Francis
760 Crick Institute using CRISPR-Cas9 assisted targeting to insert a P2A and EGFP sequence after
761 Exon 1 in frame of the *Foxl2* sequence, followed by a stop codon. The donor vector contained
762 a 783bp insert of P2A and eGFP, with 1kb homology arms either side, synthesised by
763 ThermoFisher. The guide sequence used was 5'-TCTCTGAGTGCCAACGCGCG-3' and was
764 cloned into the BbsI site of pX459 vector (Addgene #62988). The gene targeting was
765 performed by co-transfection of pX459 and the synthetic donor vector into B6N 6.0 embryonic
766 stem cells using Lipofectamine 2000. Clones were screened using an insertion PCR with
767 primers within the EGFP (Table 1), followed by a long-range PCR with primers across the
768 entire transgene (Table 1) and Sanger sequencing. Copy number was evaluated using digital
769 droplet PCR (ddPCR) (Table 1 and Figure S4) and one correctly targeted homozygous ESC
770 clone was microinjected into goGermline blastocysts [95], which were then transferred into
771 pseudo pregnant females. The resulting chimeras were crossed to C57BL/6J albino (B6(Cg)-
772 Tyr^{c-2J}/J) mice and offspring were validated for the correct integration of the transgene using
773 the above screening PCRs followed by Sanger sequencing confirmation of the resulting
774 amplicons. The ddPCR (Figure S4F) was performed for copy number evaluation of the F1
775 heterozygous mice and no evidence of random integrations of the transgene was found. Long
776 read sequencing (Oxford Nanopore Technologies) was performed on two 3kb amplicons
777 spanning the targeted insertion, generated by PCR and no aberrant on-target mutations or
778 deletions were found 1.6kb upstream and 1.8kb downstream of the insertion.
779

780 **Table 1. Primer Table**

Primer/Probe	Sequence
Insert F	aggacgacggcaactacaag (within EGFP sequence)
Insert R	gtccatgccgagagtgatcc (within EGFP sequence)
LR-F	gagaagagagtgagagccgc (from sequence upstream of 5' homology arm)

LR-R	gggtatctttgccggtgtga (from sequence downstream of 3' homology arm)
eGFP ddPCR F	agcagaagaacggcatca
eGFP ddPCR R	gggtgttctgctggtagtg
eGFP ddPCR Probe	caagatccgccacaacatcgagga
ONT-F1	tgtccattcgtgactgggac (from sequence upstream of 5' homology arm)
ONT-R1	gccgtcgtccttgaagaaga (within EGFP sequence)
ONT-F2	aggacgacggcaactacaag (within EGFP sequence)
ONT-R2	gatgaagtgaggggtgcaca (from sequence downstream of 3' homology arm)

781

782 **Immunofluorescence**

783 Ovaries and testes were fixed after dissection by immersion in 4% PFA at 4ºC.
784 Immunofluorescent stainings were performed on cryosections. The following primary
785 antibodies were used: rabbit anti-SOX9 (Millipore, cat.#ab5535), rat anti-GFP (Nacalai Tesque
786 04404-84), rabbit anti-FOXL2 (a generous gift from Dr Dagmar Wilhelm). Slides were then
787 incubated with Alexa Fluor secondary antibodies. Imaging of stained tissue sections was
788 performed on a Leica SPE confocal microscope.

789

790 **Tissue collection and chromatin preparation for ChIP-SICAP**

791 Ovaries from E14.5, 1W and 8-weeks old C57BL/6 mice were dissected in PBS, snap-frozen,
792 and stored at −80 °C. A pool of 140 embryonic E14.5 gonads, 30 1W ovaries, and 5 8W ovaries
793 were used for each replicate. Chromatin was prepared using a modified version of the Active
794 Motif High Sensitivity Chromatin Prep Kit protocol. Frozen tissue was pulverised under liquid
795 nitrogen using mortar and pestle Bel-Art™ SP Scienceware™ Liquid Nitrogen-Cooled Mini
796 Mortar. Ovaries were pooled to reach a minimum weight of 10 mg allowing a chromatin yield
797 of at least 10 ug per replicate. Samples were fixed in 1 ml of Fixation Buffer (1.5 % methanol-
798 free formaldehyde in 1% PBS) for 15 minutes on a roller at room temperature. Fixation was
799 stopped using Stop buffer by Active Motif and incubating for 5 minutes. Chromatin preparation
800 continued as per Active Motif protocol except for the sonication steps which were performed
801 in a Bioruptor® Plus sonication device by Diagenode with the following settings: 40 cycles of
802 30"ON/30"OFF, power "high", constant temperature of 4ºC.

803

804 **ChIP-SICAP**

805 ChIP-SICAP experiments were performed in parallel using 2 biological replicates for each
806 timepoint. ChIP-SICAP [18] allows the identification of chromatin-bound proteins that
807 colocalise with a bait protein (FOXL2 in this study) on DNA. A total of 10 ug of sonicated
808 chromatin was used for each immunoprecipitation using 3µl of FOXL2 antibody kindly
809 donated by Dr Dagmar Wilhelm. The original ChIP-SICAP protocol [18, 19] was optimised to
810 work on frozen tissue samples as follows. Chromatin was prepared with Active Motif

811 chromatin preparation kit as described above. After overnight immunoprecipitation at 4ºC with
812 a FOXL2-specific antibody, immune-protein complexes were captured on protein G
813 Dynabeads. DNA was biotinylated by TdT in the presence of biotin-11-ddUTP and biotin-11-
814 ddCTP as well as by Klenow3'exo in the presence of Biotin-7-dATP and eluted. Protein-DNA
815 complexes were captured with protease-resistant streptavidin beads [96], and proteins were
816 digested overnight at 37°C with 300 ng of LysC. Streptavidin beads carrying DNA were
817 separated using a magnet to be processed for ChIP-seq library preparation. The supernatant,
818 containing the digested peptides was collected and digested further by 8 hours of incubation
819 with 200 ng trypsin. Digested peptides were cleaned using the stage-tipping technique. Briefly,
820 digested samples were acidified by the addition of 2ul of TFA 10%. For each sample, 50 µL of
821 80% acetonitrile/0.1% formic acid was aliquoted in a LCMS Certified Clear Glass 12 x 32mm
822 Screw Neck Total Recovery Vial (Waters). ZipTip with 0.6 µL C18 resin was pre-treated by
823 pipetting 100% acetonitrile twice by aspirating, discarding and repeating. The ZipTip was
824 equilibrated by pipetting 0.1% TFA, three times by aspirating, discarding and repeating.
825 Acidified samples were then pipetted 10 times with ZipTip to allow peptides to bind to the
826 polymer within the tip. Following tip wash in 0.1% TFA, peptides were eluted in 80%
827 acetonitrile/0.1% formic acid in the glass vial and the eluent dried in a speed vac. Peptides were
828 reconstituted in 8 µL of 2% DMSO/0.1% formic acid. Peptides were separated on a 50 cm,
829 75 µm I.D. Pepmap column over a 70min gradient to be injected into the mass spectrometer
830 (Orbitrap Fusion Lumos) according to the universal Thermo Scientific HCD-IT method. The
831 instrument ran in data-dependent acquisition mode with the most abundant peptides selected
832 for MS/MS by HCD fragmentation. The raw data were analysed using Proteome Discoverer
833 2.1 (Thermo Scientific). Sequest HT node was used to search the spectra using the UniProt
834 Mus musculus database. The search parameters were as follows: the precursor mass tolerance
835 was 10 ppm, and MS/MS tolerance was 0.05 Da. Variable modification included Methionine
836 oxidation and N-terminal acetylation. Fixed modification included Cystenine
837 carbamidomethylation. Trypsin and LysC were chosen as the enzymes, and maximum 2 missed
838 cleavages were allowed. Perculator node was used to eliminate false identifications
839 (FDR < 0.01). Protein ratios in Foxl2 assays (n = 2) over no-antibody assays (n = 2) were
840 calculated in Proteome Discoverer. The quantification values were exported to R studio to
841 analyse significantly enriched proteins using t test limma package to determine Bayesian
842 moderated t-test p-values and Benjamini-Hochberg (BH) adjusted p values. We considered
843 proteins with mean fold enrichment >4 and adj. p value <0.1 as enriched proteins.

844

845 **DNA purification for ChIP-SICAP-seq samples**

846 Protease-resistant streptavidin beads carrying DNA samples from the above ChIP-SICAP were
847 processed for DNA purification. SDS buffer and proteinase K (20mg/ml) were added to
848 resuspend the beads followed by incubation for 30 minutes at 55°C. The beads were then heated
849 at 80°C for two hours to reverse the cross-linking. A magnetic rack was used to separate the
850 beads from the DNA, now in the supernatant. Supernatant was collected and 50 µL of Ampure
851 beads used to purify the DNA. Purified input (no SICAP) and ChIP-SICAP DNA samples were

852 used to generate libraries. NEB Ultra II DNA kit was used. Samples were sequenced on the
853 NovaSeq 6000 system. Paired-end sequencing was performed (30 million reads).

854

855 **Ovary dissociation for fluorescence-activated cell sorting (FACS)**

856 Ovaries were rinse in 1x PBS, then dissociated using add 450ul 0.05% trypsin and 50ul of 2.5%
857 collagenase. Tissue was incubated at 37°C for 6 minutes (for E14.5, up 20 minutes for 1W and
858 8 W ovaries after mincing). Enzymatic reaction was quenched by adding 200ul PBS/3% FBS.
859 Supernatant was removed without disrupting the gonads, then tissue clumps were resuspended
860 in 300ul PBS/3% FBS and gently pipetted up and down until gonads were disaggregated. Cell
861 mixture was passed through a 30 µm filter cap into a sorting tube. An extra 200ul of PBS/3%
862 FBS was used to rinse the tube and filtered. Samples were sorted with a FACS Aria III to
863 exclude debris (side scatter versus forward scatter) and doublets (area versus width) and to
864 select only live cells by excluding cells positive for DAPI. Cells were collected directly into a
865 low-retention 1.5 ml tube with 200ul PBS/3% FBS and kept on ice.

866

867 **RNA extraction, library preparation and sequencing**

868 A total of 50000 GFP positive cells were sorted, and RNA was extracted using the RNeasy
869 Plus Micro Kit (cat. no. 74034) by Qiagen. RNA quality was assessed by Bioanalyzer and
870 samples with a RIN value>8 were processed for library preparation. A total of 10ng of RNA
871 was used to prepare libraries using the NEBNext® Low Input RNA Library Prep Kit (E6420L).
872 Samples were subsequently sequenced on the NovaSeq 6000 system, at a dept of 25 million
873 reads (PE100) per sample.

874

875 **ATAC-Seq, library preparation and sequencing**

876 GFP positive cells isolated by FACS from gonads at E14.5, 1W and 8W (three biological
877 replicates per timepoint) were used for ATAC-Seq library preparation using Active Motif
878 ATAC-Seq Kit (cat. number 53150). In brief, 100 000 GFP positive sorted cells were washed
879 in PBS, pelleted, and lysed in ATAC Lysis Buffer to isolate intact nuclei. Tagmentation of
880 nuclei was performed for 30 minutes in a thermomixer at 37°C set at 800 rpm. Tagmented
881 DNA was purified using DNA purification columns and library amplification was performed
882 using Illumina's Nextera adapters. A unique combination of i7/i5 primers was used to allow
883 for multiplexing and sequencing on the same flow cell. PCR amplification included 5 minutes
884 incubation at 72°C, followed by 30 sec of DNA denaturation at 98°C and 10 cycles of: 98°C
885 for 10 sec, 63°C for 30 sec and 72°C for 1 minute. PCR products were cleaned up using SPRI
886 beads and size distribution of PCR amplified libraries was assessed on Tapestation. Samples
887 were subsequently sequenced on the NovaSeq 6000 system.

888

889

890 **mRNA-Seq alignment and quantification**

891 Reads were processed using the publicly available nf-core [97] rnseq pipeline v3.3 with the
892 STAR [98]/RSEM [99] option against mouse genome assembly GRCm38 and Ensembl release
893 81 transcript annotations. Other options were left as default. Gene-level RSEM abundance
894 estimates were imported into R using the Bioconductor package tximport's tximport function
895 [100] for further analysis using DESeq2 [64]. Normalisation factors were calculated per gene
896 using the default DESeq2 function, correcting for library size and feature length. Variance-
897 stabilised abundance estimates were calculated using the VST function. PCA analysis was
898 performed on the variance stabilised values using all available genes.

899

900 **mRNA-Seq differential expression analysis and clustering**

901 Differential expression between treatment groups was assessed using a DESeq2's Wald test in
902 a pairwise manner, with false-discovery rate control based on an independent hypothesis
903 weighting (IHW) (Ignatiadis et al 2016). Shrunken log2 fold changes were calculated using the
904 lfcShrink function (type="ashr"). Significance was assessed based on a combined FDR<0.01,
905 a shrunken log2FC>2 and a baseMean (i.e. mean abundance across all samples) >2. A
906 combined list of significantly differentially expressed genes from all pairwise comparisons was
907 created and used to generate a heatmap showing how patterns of expression changed across
908 samples. Per-gene z-scores were calculated from the variance stabilised abundance estimates.
909 K-means clustering was applied (km=7) to split the data into slices of similarly changing genes.
910 Within each slice genes were clustered using a Euclidean distance metric and complete-linkage.

911

912 **Gene Ontology enrichment analysis**

913 Gene Ontology (GO) Biological Process (BP enrichment analysis of the individual gene lists
914 and the kmeans clusters was calculated relative to a genomic background via the
915 compareCluster function from the clusterProfiler Bioconductor package [101]
916 (fun="enrichGO", OrgDb="org.Mm.eg.db", ont="BP", pAdjustMethod="BH",
917 pvalueCutoff=0.05, qvalueCutoff=0.01, minGSSize=10, maxGSSize=500).

918

919 **ChIP-SICAP**

920

921 **ChIP-Seq read alignment, peak calling and peak annotation**

922 Reads were processed using the publicly available nf-core [97] chipseq pipeline v1.2.2 against
923 the mouse genome assembly GRCm38 (--narrow_peak, --single_end, --deseq2_vst,
924 min_reps_consensus 2, --macs_fdr 0.05), applying the Encode GRCm38 blacklist definitions
925 to filter out problematic regions. A peak was only considered to be genuine if present in both
926 biological replicates for each treatment group. A consensus peak set was then constructed by
927 combining the peak calls from the individual treatment groups into a set of non-overlapping

intervals. Consensus peaks were annotated using the Bioconductor package ChIPpeakAnno's annotatePeakInBatch function [102] against the protein-coding gene component of Ensembl release 81. Gene assignments were made based on peak-centre minimal proximity to a gene's TSS (output="nearestLocation",multiple=FALSE,maxgap=1L,PeakLocForDistance="middle",FeatureLocForDistance="TSS",select="first",bindingType ="fullRange"). The distribution of peaks mapping to genomic features was similarly assessed using ChIPpeakAnno's "genomicElementDistribution" function, defining promoter regions as +/-2kb around a TSS and downstream regions as the 2kb region downstream of genic boundaries.

DiffBind analysis

Consensus peak occupancy analysis was conducted using the Bioconductor package DiffBind [103]. Briefly, reads were counted from the nfcore chipseq pipeline alignment files (BAM) over a combined set of consensus peaks (intervals). Control (input) read counts were subtracted from the targeted samples at each interval:

dba.count(samples.dba,peaks=consensus.gr,summits=FALSE,filter=0,bSubControl=TRUE,minCount=0,bUseSummarizeOverlaps=TRUE,mapQCth=0).

The resulting DiffBind object was normalised using a background approach, which calculates scaled factors from read counts against a set of 15000bp genomic bins considered large enough not to show differential enrichment across samples:

dba.normalize(samples.dba,background=TRUE, method=DBA_DESEQ2).

Normalised peak counts for all peaks were subsequently used for Pearson correlation analysis and PCA analysis. Differential binding affinity analysis was conducted in a pairwise fashion between treatment conditions using DESeq2 [64]. A FDR of <0.05 was used to threshold significant changes in peak occupancy. Peak showing significant changes in occupancy were submitted for read-depth profiling in the region +/-1.5kb from their centres. Plots are presented at the merged replicate level, stratified based on gain/loss status (i.e. direction of fold change).

Clustering of ChIP-Seq peaks

The normalised peak occupancy estimates were used to cluster the consensus peaks into sets that showed similar patterns of variation in occupancy across conditions. Peak estimates were merged across replicates using a mean, a pseudo-count of 0.1 was added prior to a log2 transformation. Per-peak z-scores were calculated. K-means clustering was applied (km=7, 1000 repeats) to split the data into slices of similarly behaving peaks. Within each slice peaks were clustered using a Euclidean distance metric and complete-linkage.

965 **Motif enrichment analysis**

966 Analysis was performed using the "calcBinnedMotifEnrR" from the monaLisa package [67],
967 by scanning the sequence of consensus peaks +/-500bp from their centres against the PWMs
968 from JASPAR 2020 [104]. Significance of enrichment was assessed against a background of
969 equivalently sized sequences randomly sampled from the genome. Sequences were weighted
970 to correct for GC and k-mer composition differences between fore- and background sets.
971 Statistically significant enriched motifs within each cluster were determined using a one-tailed
972 Fisher's exact test with Benjamini-Hochberg adjusted p-values pval<0.0001 and subsequently
973 prioritised based on their log2 enrichment scores (descending).

974

975 **ATAC-Seq**

976 **Read alignment and peak calling**

977 Reads were processed using the publicly available nf-core [97] atacseq pipeline v1.2.1 against
978 the mouse genome assembly GRCm38 (--deseq2_vst --min_reps_consensus), applying the
979 Encode GRCm38 blacklist definitions to filter out problematic regions. A peak was only
980 considered to be genuine if present in at least 2 biological replicates for each treatment group.
981 A consensus peak set was then constructed by combining the peak calls from the individual
982 treatment groups into a set of non-overlapping intervals.

983

984 **Peak annotation**

985 Consensus peaks were annotated using the Bioconductor package ChIPpeakAnno's
986 annotatePeakInBatch function [102] against the protein-coding gene component of Ensembl
987 release 81. Gene assignments were made based on peak-centre minimal proximity to a gene's
988 TSS: (output="nearestLocation", multiple=FALSE, maxgap=-1L,
989 PeakLocForDistance="middle", FeatureLocForDistance="TSS", select="first",
990 bindingType="fullRange"). The distribution of peaks mapping to genomic features was
991 similarly assessed using ChIPpeakAnno's genomicElementDistribution function, defining
992 promoter regions as +/-2kb around a TSS and downstream regions as the 2kb region
993 downstream of genic boundaries.

994

995 **DiffBind analysis**

996 Consensus peak occupancy analysis was conducted using the Bioconductor package DiffBind
997 [103]. Briefly, reads were counted from the nfcore chipseq pipeline alignment files (BAM)
998 over a combined set of consensus peaks (intervals) dba.count
999 (samples.dba,peaks=consensus.gr,summits=FALSE,filter=0,bSubControl=FALSE,minCount
1000 =0,bUseSummarizeOverlaps=TRUE,mapQCth=0). The resulting DiffBind object was
1001 normalised DESeq2's RLE method [64] based on the peak counts
1002 dba.normalize(samples.dba,method=DBA_DESEQ2,normalize=DBA_NORM_RLE,library=

1003 DBA_LIBSIZE_PEAKREADS). Normalised peak counts for all peaks were subsequently used
1004 for Pearson correlation analysis and PCA analysis.

1005 Differential binding affinity analysis was conducted in a pairwise fashion between treatment
1006 conditions using DESeq2 [64]. A FDR of <0.05 was used to threshold significant changes in
1007 peak occupancy. Peak showing significant changes in occupancy were submitted for read-
1008 depth profiling in the region +/-1.5kb from their centres. Plots are presented at the merged
1009 replicate level, stratified based on gain/loss status (i.e. direction of fold change).

1010

1011 **Clustering of peaks**

1012 The normalised peak occupancy estimates were used to cluster the consensus peaks into sets
1013 that showed similar patterns of variation in occupancy across conditions. Peak estimates were
1014 merged across replicates using a mean, a pseudo-count of 0.1 was added prior to a log2
1015 transformation. Per-peak z-scores were calculated. K-means clustering was applied (km=7,
1016 1000 repeats) to split the data into slices of similarly behaving peaks. Within each slice peaks
1017 were clustered using a Euclidean distance metric and complete-linkage.

1018

1019 **Data integration**

1020 Data from the RNA, ATAC and ChIP experiments were integrated at the gene level. The
1021 ATAC and ChIP data were assigned to protein-coding genes based on minimal peak proximity
1022 to a TSS. Where this resulted in multiple peaks being assigned the same gene, the one closest
1023 to the TSS was given precedence.

1024 Identification of stably expressed and dynamically regulated genes: genes dynamically
1025 regulated by FOXL22 were defined as the intersect of i) genes differentially expressed in at
1026 least one of the pairwise comparisons from the RNA-seq analysis and ii) bound by Foxl2 based
1027 on the presence of a ChIP-seq peak in at least 1 timepoint. Stably expressed genes were defined
1028 as those bound by FOXL2 but now showing differential expression. Genes not bound by
1029 FOXL2 but showing differential expression were not considered to be directly regulated by
1030 FOXL2. Stably expressed genes were subsequently divided into "activated" and "repressed"
1031 subsets based on the variance stabilised (VST) normalised counts from the RNA-seq analysis.
1032 A VST score of >=8.3 in at least 1 sample was enough to quality a gene as “activated”, while
1033 all others were regarded as “repressed”.

1034

1035 **Cytoscape visualisation of FOXL2 protein interactors**

1036 GO enrichment analysis of FOXL2 interactors identified by ChIP-SICAP and fold changes
1037 over no-antibody control in Table S3 were loaded onto Cytoscape v.3.9.1 [105]. Proteins were
1038 clustered based on their function. The cytoscape app enhancedGraphics [106] was used to
1039 overlay bar charts depicting the fold changes over control of each protein in each timepoint.

1040

1041 **Data and materials availability**

1042 All data are available in the main text or in the supplementary materials. Codes are available
1043 on request.

1044

1045 **Aknowledgements**

1046 This work was supported by the Medical Research Council, U.K. (CC2116) and the Francis
1047 Crick Institute which receives its core funding from Cancer Research UK (CC2116), the UK
1048 Medical Research Council (CC2116), and the Wellcome Trust. RM holds a Postdoctoral
1049 Fellowship core-funded by the Francis Crick Institute (CC2116).

1050 We are grateful for the support from present and past members of the Lovell-Badge lab. We
1051 are particularly grateful to Dr Nitzan Gonen (Bar-Ilan University), to Dr Francis Poulat
1052 (University of Montpellier), and to Dr Marie-Christine Chaboissier (Institut de Biologie
1053 Valrose) for scientific discussions and support.

1054 We thank the Francis Crick institute platforms (GeMS, ASF, BABS, FLOW, Proteomics) for
1055 their excellent support and expertise. We thank Jacek Mor and all staff from the Biological
1056 Research facility (Francis Crick Institute) for taking excellent care of the animals. We thank
1057 Dr Dagmar Wilhelm for the generous gift of the anti-FOXL2 antibody.

1058

1059 **Author contributions:**

1060 Conceptualization: RM

1061 Visualization: RM

1062 Investigation: RM, MN

1063 Bioinformatic analyses: RM, RM, MRR

1064 Methodology: RM, MN, MRR, SW, JO

1065 Supervision: RM, RLB

1066 Funding acquisition: RLB

1067 Writing-original draft: RM

1068 Writing-review&editing: RM and all authors

1069

1070 **Competing interests**

1071 The authors declare no competing interests.

1072

1073

1074

1075

1076 **SUPPLEMENTARY FIGURE LEGENDS**

1077

1078 **Figure S1. FOXL2 genome-wide occupancy across ovarian development**, Related to Figure
1079 1.

1080 (A) Annotation of consensus peaks by ChIPpeakAnno against the default iGenomes Ensembl
1081 GTF file. (B) Enrichment analysis of transcription factor binding sites associated with
1082 consensus peaks significantly changing across the timecourse. Analysis performed using the
1083 calcBinnedMotifEnrR from the monaLisa package. Top 10 representative TF are shown (*p-*
1084 *value*<0.0001). (C-E) Heatmaps of differentially bound peaks in each pairwise comparison.

1085

1086 **Figure S2. Comparison of FOXL2 target genes with DEGs identified in a *Foxl2*$^{-/-}$ null**
1087 **model (Ortiz *et al.*)**, Related to Figure 1.

1088 (A) Venn diagram showing the overlap between FOXL2 target genes identified at any point
1089 throughout ovarian development (E14.5, 1W and 8W) and genes downregulated in *Foxl2*$^{-/-}$ null
1090 mutant ovaries compared to wildtype controls (Ortiz *et al.*[29] and Table S2). Genomic
1091 overviews of FOXL2 peaks in representative genes likely to be activated by FOXL2 including
1092 *Cdkn1b* (B), *Smad7* (C), and *Fst* (D). (E) Venn diagram showing overlap between FOXL2
1093 target genes and genes upregulated in in FOXL2^{KO} ovaries. Representative genes within this
1094 overlap are shown and include *Wnt4* (F), *Gadd45g* (G), *Inhbb* (H) and *Sox9* (I). (J) Bar chart
1095 displaying the total number of significant peaks identified in each timepoint in the 1Mb gene
1096 desert upstream of *Sox9*.

1097

1098 **Figure S3 Cytoscape network of FOXL2 interactome across ovarian development**,
1099 Related to Figure 3.

1100 (A) Proteins interacting with FOXL2 on-chromatin were clustered by GO Processes and
1101 visualised with Cytoscape [107]. Enrichment values over no-antibody control were depicted as
1102 barcharts (yellow=E14.5, orange=1W, blue=8W, n=2, fold change over no antibody control >4
1103 in at least one timepoint, adj-*pvalue*<0.1) and visualised within the network using
1104 enhancedGraphics Cytoscape plugin [106]. (B) Overlap of protein interactors found in our
1105 study compared to a cell line (AT29c) whole proteome study by Penrad-Mobayed *et al.* [51].

1106

1107 **Figure S4. Design of *Foxl2*^{EGFP} construct and genotyping strategy to ensure correct**
1108 **insertion of EGFP into *Foxl2* locus**, Related to Figure 4.

1109 **(**A) Design of gRNA (yellow box) and associated silent mutation (SM) in PAM sequence as
1110 well as 783bp P2A-EGFP construct (green boxes). (B) Location of primers for insert, junction,
1111 long-range and digital droplet PCR. (C) ES cell screening identified clone E11 (green arrows),
1112 which showed successful EGFP integration depicted by a single band in insert, junction, and
1113 long-range PCR. (D) Correct integration of EGFP into Foxl2 locus of clone E11 was confirmed
1114 via sanger sequencing. (E) Long-range PCR of chimeric offspring as well as clone E11
1115 confirmed heterozygous and homozygous EGFP insertion, respectively. No evidence of

1116 backbone integration was found. BB pos= backbone positive; BB nrg= backbone negative. (F)
1117 Copy number evaluation of chimeric offspring as well as ES cell clone E11 by ddPCR
1118 confirmed absence of random transgene integration.

1119

1120 **Figure S5. Characterisation of the *Foxl2*^{EGFP} mouse line**, Related to Figure 4.

1121 (A) Representative images of ovaries and testes dissected from adult mice (8 weeks) and
1122 imaged with a fluorescence microscope. Upper row: GFP channel, lower row: brightfield
1123 channel. Adult mouse ovaries wildtype (a-b) and heterozygous (c-d) as well as testes wildtype
1124 (e-f) and heterozygous (g-h) are shown. Scale bars=100 µM. (B) Fertility assessment of
1125 *Foxl2*$^{EGFP+/-}$ females compared to wildtype controls (C57BL/6J stock mice).

1126

1127 **Figure S6. Fluorescence-activated cell sorting of EGFP positive cells throughout ovarian**
1128 **development**, Related to Figure 4.

1129 Gating strategy used to isolate live cells (A), singlets (B), and positive for GFP. An example
1130 of FACS profile of a negative samples is shown in (C). (D-F) Dissociated cells were gated and
1131 purified by FACS at E14.5 (D), 1W (E), and 8W (F). A total of 50k cells at each timepoint was
1132 sorted for RNA-Seq and 100k for ATAC-Seq.

1133

1134 **Figure S7. Overview of RNA-Seq analysis of *Foxl2*$^{EGFP+/-}$ cells isolated throughout**
1135 **ovarian development**, Related to Figure 4 and Methods

1136 (A) Poisson dissimilarity metric to assess RNA-Seq sample similarity, n=3 biological
1137 replicates. (B) Boxplot illustrating the dynamics of gene expression changes granulosa
1138 *Foxl2*$^{EGFP+/-}$ cells. (C) K-means clustering of the gene expression dynamics and representative
1139 genes on the right-hand side. (D) Top terms from the GO enrichment test (Cluster Profiler)
1140 showing the processes associated with DEGs from each of the seven clusters.

1141

1142 **Figure S8. Integration of FOXL2 ChIP-Seq data with RNA-Seq of *Foxl2*$^{EGFP+/-}$ cells**,
1143 Related to Figure 5.

1144 (A) Plot of gene density and VST cut-off =8.3 used to distinguish genes stably repressed from
1145 those stably expressed. (B) Barplot depicting the number of genes either stably repressed or
1146 activated and bound by FOXL2. (C) Example of gene expression changes used to choose the
1147 VST cut-off. The male marker *Fgf9* was used as reference for the cut-off as known to not be t
1148 expressed in the ovary.

1149

1150 **Figure S9. ATAC-Seq analysis of chromatin accessibility of *Foxl2*^{EGFP} positive cells**
1151 **collected across ovarian development**, Related to Figure 6.

1152 (A) Poisson metric showing peak occupancy correlation between timepoints (E14.5, 1W and
1153 8W). (B) Genomic annotation of consensus peaks obtained with ChIPpeakAnno [102] against

1154 the default iGenomes Ensembl GTF file. (C) Boxplot of z-scores representing overall trends of
1155 chromatin opening dynamics over time within each cluster derived from the clustering of
1156 normalised abundancies of peaks. (D) Heatmap of enrichment scores for the FOXL2 canonical
1157 motif in each cluster, ranked on enrichment score and filtered by -log10pval. Cut-off: -
1158 log10pval > 4 (i.e. pval<0.0001).

1159

1160 **Figure S10. Overview of differential accessibility analysis of ATAC-Seq data**, Related to
1161 Figure 6.

1162 (A-C) Density plots of peaks differentially opened in each pairwise comparison. (D) Summary
1163 of shared and unique open chromatin regions across the timecourse. (E-F) IGV snapshots
1164 representative of several putative enhancers identified within *Rspo1*, *Fst*, *Runx1* and *Sox9*
1165 genes, as assessed by ATAC-Seq and FOXL2 ChIP-SICAP. UCSC Genome Brower
1166 conservation analysis in pop-up squares. Black tracks show E13.5 Sertoli ATAC-Seq data re-
1167 analysed from Garcia-Moreno *et al*. [94], and P7 Sertoli ATAC-Seq from Lindeman *et al.* [41].

1168

1169 **Figure S11. Overview of differential accessibility analysis of ATAC-Seq data, Sertoli-**
1170 **specific genes**, Related to Figure 6.

1171 IGV snapshots representative of several putative enhancers identified within Sertoli-enriched
1172 genes. (A) *Cldn11*, (B) *Itga6*, (C) *Amh*, (D) *Hsd17b1*, (E) *Nr5a1*, (F) *Fgf9*. Black tracks show
1173 E13.5 Sertoli ATAC-Seq data re-analysed from Garcia-Moreno *et al.*[94], and P7 Sertoli
1174 ATAC-Seq from Lindeman *et al.* [41].

1175

1176 **Sup. Table 1:** FOXL2 ChIP-SICAP analysis of genomic targets, GO of clusters and diffbind
1177 analysis, relative to Figure 1 and 2.

1178 **Sup. Table 2:** Comparison of FOXL2 ChIP-SICAP with other datasets from Jameson *et*
1179 *al.[40],* Lindeman *et al.[41]*, and Ortiz *et al.[29]*, relative to Figure 2.

1180 **Sup. Table 3:** FOXL2 ChIP-SICAP, proteomics dataset of FOXL2 interactors, relative to
1181 Figure 3.

1182 **Sup. Table 4:** RNA-Seq of *Foxl2*$^{EGFP+/-}$ cells isolated throughout ovarian development**,**
1183 relative to Figure 4.

1184 **Sup. Table 5:** Data integration of ChIP-Seq, ATAC and RNA-Seq and gene-level dynamics,
1185 relative to Figure 5.

1186 **Sup. Table 6:** ATAC-Seq normalised peak counts, diffbind analysis**,** clustering analysis.

1187 **References**

1188

1189 1. Bunce, C., J. McKey, and B. Capel, *Concerted morphogenesis of genital ridges and nephric ducts*
1190 *in the mouse captured through whole-embryo imaging.* Development, 2021. **148**(18).
1191 2. Migale, R., M. Neumann, and R. Lovell-Badge, *Long-Range Regulation of Key Sex*
1192 *Determination Genes.* Sex Dev, 2021. **15**(5-6): p. 360-380.
1193 3. Koopman, P., et al., *Male development of chromosomally female mice transgenic for Sry.*
1194 Nature, 1991. **351**(6322): p. 117-21.
1195 4. Hacker, A., et al., *Expression of Sry, the mouse sex determining gene.* Development, 1995.
1196 **121**(6): p. 1603-14.
1197 5. Bullejos, M. and P. Koopman, *Spatially dynamic expression of Sry in mouse genital ridges.* Dev
1198 Dyn, 2001. **221**(2): p. 201-5.
1199 6. Vidal, V.P., et al., *Sox9 induces testis development in XX transgenic mice.* Nat Genet, 2001.
1200 **28**(3): p. 216-7.
1201 7. Yao, H.H., *The pathway to femaleness: current knowledge on embryonic development of the*
1202 *ovary.* Mol Cell Endocrinol, 2005. **230**(1-2): p. 87-93.
1203 8. Rossitto, M., et al., *TRIM28-dependent SUMOylation protects the adult ovary from activation*
1204 *of the testicular pathway.* Nat Commun, 2022. **13**(1): p. 4412.
1205 9. Uhlenhaut, N.H., et al., *Somatic sex reprogramming of adult ovaries to testes by FOXL2*
1206 *ablation.* Cell, 2009. **139**(6): p. 1130-42.
1207 10. Nicol, B., et al., *RUNX1 maintains the identity of the fetal ovary through an interplay with*
1208 *FOXL2.* Nat Commun, 2019. **10**(1): p. 5116.
1209 11. Tucker, E.J., *The Genetics and Biology of FOXL2.* Sex Dev, 2022. **16**(2-3): p. 184-193.
1210 12. Schmidt, D., et al., *The murine winged-helix transcription factor Foxl2 is required for granulosa*
1211 *cell differentiation and ovary maintenance.* Development, 2004. **131**(4): p. 933-42.
1212 13. Crisponi, L., et al., *The putative forkhead transcription factor FOXL2 is mutated in*
1213 *blepharophimosis/ptosis/epicanthus inversus syndrome.* Nat Genet, 2001. **27**(2): p. 159-66.
1214 14. Harris, S.E., et al., *Identification of novel mutations in FOXL2 associated with premature*
1215 *ovarian failure.* Mol Hum Reprod, 2002. **8**(8): p. 729-33.
1216 15. De Baere, E., et al., *Spectrum of FOXL2 gene mutations in blepharophimosis-ptosis-epicanthus*
1217 *inversus (BPES) families demonstrates a genotype--phenotype correlation.* Hum Mol Genet,
1218 2001. **10**(15): p. 1591-600.
1219 16. Uda, M., et al., *Foxl2 disruption causes mouse ovarian failure by pervasive blockage of follicle*
1220 *development.* Hum Mol Genet, 2004. **13**(11): p. 1171-81.
1221 17. Ottolenghi, C., et al., *Foxl2 is required for commitment to ovary differentiation.* Hum Mol
1222 Genet, 2005. **14**(14): p. 2053-62.
1223 18. Rafiee, M.R., et al., *Expanding the Circuitry of Pluripotency by Selective Isolation of Chromatin-*
1224 *Associated Proteins.* Mol Cell, 2016. **64**(3): p. 624-635.
1225 19. Rafiee, M.R. and J. Krijgsveld, *Using ChIP-SICAP to Identify Proteins That Co-localize in*
1226 *Chromatin.* Methods Mol Biol, 2021. **2351**: p. 275-288.
1227 20. Niu, W. and A.C. Spradling, *Two distinct pathways of pregranulosa cell differentiation support*
1228 *follicle formation in the mouse ovary.* Proc Natl Acad Sci U S A, 2020. **117**(33): p. 20015-20026.
1229 21. Georges, A., et al., *FOXL2: a central transcription factor of the ovary.* J Mol Endocrinol, 2014.
1230 **52**(1): p. R17-33.
1231 22. Franks, S. and K. Hardy, *Androgen Action in the Ovary.* Front Endocrinol (Lausanne), 2018. **9**:
1232 p. 452.
1233 23. Drummond, A.E. and P.J. Fuller, *The importance of ERbeta signalling in the ovary.* J Endocrinol,
1234 2010. **205**(1): p. 15-23.
1235 24. Bashamboo, A. and K. McElreavey, *NR5A1/SF-1 and development and function of the ovary.*
1236 Ann Endocrinol (Paris), 2010. **71**(3): p. 177-82.

1237 25. Park, M., et al., *FOXL2 interacts with steroidogenic factor-1 (SF-1) and represses SF-1-induced*
1238 *CYP17 transcription in granulosa cells.* Mol Endocrinol, 2010. **24**(5): p. 1024-36.
1239 26. Shen, W.H., et al., *Nuclear receptor steroidogenic factor 1 regulates the mullerian inhibiting*
1240 *substance gene: a link to the sex determination cascade.* Cell, 1994. **77**(5): p. 651-61.
1241 27. Georges, A., et al., *The transcription factor FOXL2 mobilizes estrogen signaling to maintain the*
1242 *identity of ovarian granulosa cells.* Elife, 2014. **3**.
1243 28. Chassot, A.A., et al., *WNT4 and RSPO1 together are required for cell proliferation in the early*
1244 *mouse gonad.* Development, 2012. **139**(23): p. 4461-72.
1245 29. Garcia-Ortiz, J.E., et al., *Foxl2 functions in sex determination and histogenesis throughout*
1246 *mouse ovary development.* BMC Dev Biol, 2009. **9**: p. 36.
1247 30. Gustin, S.E., et al., *WNT/beta-catenin and p27/FOXL2 differentially regulate supporting cell*
1248 *proliferation in the developing ovary.* Dev Biol, 2016. **412**(2): p. 250-60.
1249 31. Granados-Aparici, S., et al., *SMAD3 directly regulates cell cycle genes to maintain arrest in*
1250 *granulosa cells of mouse primordial follicles.* Sci Rep, 2019. **9**(1): p. 6513.
1251 32. Yao, H.H., W. Whoriskey, and B. Capel, *Desert Hedgehog/Patched 1 signaling specifies fetal*
1252 *Leydig cell fate in testis organogenesis.* Genes Dev, 2002. **16**(11): p. 1433-40.
1253 33. Parma, P., et al., *R-spondin1 is essential in sex determination, skin differentiation and*
1254 *malignancy.* Nat Genet, 2006. **38**(11): p. 1304-9.
1255 34. Vainio, S., et al., *Female development in mammals is regulated by Wnt-4 signalling.* Nature,
1256 1999. **397**(6718): p. 405-9.
1257 35. Johnen, H., et al., *Gadd45g is essential for primary sex determination, male fertility and testis*
1258 *development.* PLoS One, 2013. **8**(3): p. e58751.
1259 36. Rodriguez, K.F., et al., *Somatic cell fate maintenance in mouse fetal testes via*
1260 *autocrine/paracrine action of AMH and activin B.* Nat Commun, 2022. **13**(1): p. 4130.
1261 37. Sekido, R., et al., *SOX9 is up-regulated by the transient expression of SRY specifically in Sertoli*
1262 *cell precursors.* Dev Biol, 2004. **274**(2): p. 271-9.
1263 38. Gonen, N., et al., *Sex reversal following deletion of a single distal enhancer of Sox9.* Science,
1264 2018. **360**(6396): p. 1469-1473.
1265 39. Sekido, R. and R. Lovell-Badge, *Sex determination involves synergistic action of SRY and SF1 on*
1266 *a specific Sox9 enhancer.* Nature, 2008. **453**(7197): p. 930-4.
1267 40. Jameson, S.A., et al., *Temporal transcriptional profiling of somatic and germ cells reveals*
1268 *biased lineage priming of sexual fate in the fetal mouse gonad.* PLoS Genet, 2012. **8**(3): p.
1269 e1002575.
1270 41. Lindeman, R.E., et al., *The conserved sex regulator DMRT1 recruits SOX9 in sexual cell fate*
1271 *reprogramming.* Nucleic Acids Res, 2021. **49**(11): p. 6144-6164.
1272 42. Chassot, A.A., et al., *Activation of beta-catenin signaling by Rspo1 controls differentiation of*
1273 *the mammalian ovary.* Hum Mol Genet, 2008. **17**(9): p. 1264-77.
1274 43. Wang, D.S., et al., *Foxl2 up-regulates aromatase gene transcription in a female-specific*
1275 *manner by binding to the promoter as well as interacting with ad4 binding*
1276 *protein/steroidogenic factor 1.* Mol Endocrinol, 2007. **21**(3): p. 712-25.
1277 44. Pisarska, M.D., et al., *Forkhead l2 is expressed in the ovary and represses the promoter activity*
1278 *of the steroidogenic acute regulatory gene.* Endocrinology, 2004. **145**(7): p. 3424-33.
1279 45. Dupont, S., et al., *Effect of single and compound knockouts of estrogen receptors alpha*
1280 *(ERalpha) and beta (ERbeta) on mouse reproductive phenotypes.* Development, 2000. **127**(19):
1281 p. 4277-91.
1282 46. Fu, A., et al., *IRX3 and IRX5 collaborate during ovary development and follicle formation to*
1283 *establish responsive granulosa cells in the adult mousedagger.* Biol Reprod, 2020. **103**(3): p.
1284 620-629.
1285 47. Cai, K., et al., *Action mechanism of inhibin alpha-subunit on the development of Sertoli cells*
1286 *and first wave of spermatogenesis in mice.* PLoS One, 2011. **6**(10): p. e25585.

48. Obholz, K.L., et al., *FNDC3A is required for adhesion between spermatids and Sertoli cells.* Dev Biol, 2006. **298**(2): p. 498-513.
49. Gao, F., et al., *The Wilms tumor gene, Wt1, is required for Sox9 expression and maintenance of tubular architecture in the developing testis.* Proc Natl Acad Sci U S A, 2006. **103**(32): p. 11987-92.
50. Jin, H., et al., *FOXL2 directs DNA double-strand break repair pathways by differentially interacting with Ku.* Nat Commun, 2020. **11**(1): p. 2010.
51. Penrad-Mobayed, M., et al., *Conventional and unconventional interactions of the transcription factor FOXL2 uncovered by a proteome-wide analysis.* FASEB J, 2020. **34**(1): p. 571-587.
52. Hu, Y.C., L.M. Okumura, and D.C. Page, *Gata4 is required for formation of the genital ridge in mice.* PLoS Genet, 2013. **9**(7): p. e1003629.
53. Kreidberg, J.A., et al., *WT-1 is required for early kidney development.* Cell, 1993. **74**(4): p. 679-91.
54. Meinsohn, M.C., et al., *A role for orphan nuclear receptor liver receptor homolog-1 (LRH-1, NR5A2) in primordial follicle activation.* Sci Rep, 2021. **11**(1): p. 1079.
55. Zhang, T., et al., *HDAC6 regulates primordial follicle activation through mTOR signaling pathway.* Cell Death Dis, 2021. **12**(6): p. 559.
56. Ernst, E.H., et al., *Granulosa cells from human primordial and primary follicles show differential global gene expression profiles.* Hum Reprod, 2018. **33**(4): p. 666-679.
57. Leon, N.Y. and V.R. Harley, *ATR-X syndrome: genetics, clinical spectrum, and management.* Hum Genet, 2021. **140**(12): p. 1625-1634.
58. L'Hote, D., et al., *Discovery of novel protein partners of the transcription factor FOXL2 provides insights into its physiopathological roles.* Hum Mol Genet, 2012. **21**(14): p. 3264-74.
59. Zhou, J., et al., *Dissecting the fate of Foxl2-expressing cells in fetal ovary using lineage tracing and single-cell transcriptomics.* Cell Discov, 2022. **8**(1): p. 139.
60. Morris, M.E., et al., *A single-cell atlas of the cycling murine ovary.* Elife, 2022. **11**.
61. Hsu, S.H., et al., *Spz1, a novel bHLH-Zip protein, is specifically expressed in testis.* Mech Dev, 2001. **100**(2): p. 177-87.
62. Stevant, I., et al., *Deciphering Cell Lineage Specification during Male Sex Determination with Single-Cell RNA Sequencing.* Cell Rep, 2018. **22**(6): p. 1589-1599.
63. Maatouk, D.M., et al., *Genome-wide identification of regulatory elements in Sertoli cells.* Development, 2017. **144**(4): p. 720-730.
64. Love, M.I., W. Huber, and S. Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.* Genome Biol, 2014. **15**(12): p. 550.
65. Consortium, E.P., et al., *Expanded encyclopaedias of DNA elements in the human and mouse genomes.* Nature, 2020. **583**(7818): p. 699-710.
66. Jorgez, C.J., et al., *Granulosa cell-specific inactivation of follistatin causes female fertility defects.* Mol Endocrinol, 2004. **18**(4): p. 953-67.
67. Machlab, D., et al., *monaLisa: an R/Bioconductor package for identifying regulatory motifs.* Bioinformatics, 2022. **38**(9): p. 2624-2625.
68. Kim, S.W., et al., *Non-canonical Wnt signals are modulated by the Kaiso transcriptional repressor and p120-catenin.* Nat Cell Biol, 2004. **6**(12): p. 1212-20.
69. Couse, J.F., et al., *Estrogen receptor-beta is critical to granulosa cell differentiation and the ovulatory response to gonadotropins.* Endocrinology, 2005. **146**(8): p. 3247-62.
70. Shah, S.P., et al., *Mutation of FOXL2 in granulosa-cell tumors of the ovary.* N Engl J Med, 2009. **360**(26): p. 2719-29.
71. Boulanger, L., et al., *FOXL2 is a female sex-determining gene in the goat.* Curr Biol, 2014. **24**(4): p. 404-8.
72. Ottolenghi, C., et al., *Loss of Wnt4 and Foxl2 leads to female-to-male sex reversal extending to germ cells.* Hum Mol Genet, 2007. **16**(23): p. 2795-804.

1337 73. Auguste, A., et al., *Loss of R-spondin1 and Foxl2 amplifies female-to-male sex reversal in XX*
1338 *mice.* Sex Dev, 2011. **5**(6): p. 304-17.
1339 74. Habara, O., et al., *WNT signaling in pre-granulosa cells is required for ovarian folliculogenesis*
1340 *and female fertility.* Development, 2021. **148**(9).
1341 75. Nicol, B., et al., *Genome-wide identification of FOXL2 binding and characterization of FOXL2*
1342 *feminizing action in the fetal gonads.* Hum Mol Genet, 2018. **27**(24): p. 4273-4287.
1343 76. Croft, B., et al., *Human sex reversal is caused by duplication or deletion of core enhancers*
1344 *upstream of SOX9.* Nat Commun, 2018. **9**(1): p. 5319.
1345 77. Maatouk, D.M., et al., *Stabilization of beta-catenin in XY gonads causes male-to-female sex-*
1346 *reversal.* Hum Mol Genet, 2008. **17**(19): p. 2949-55.
1347 78. Englert, C., et al., *Truncated WT1 mutants alter the subnuclear localization of the wild-type*
1348 *protein.* Proc Natl Acad Sci U S A, 1995. **92**(26): p. 11960-4.
1349 79. Larsson, S.H., et al., *Subnuclear localization of WT1 in splicing or transcription factor domains*
1350 *is regulated by alternative splicing.* Cell, 1995. **81**(3): p. 391-401.
1351 80. Spector, D.L. and A.I. Lamond, *Nuclear speckles.* Cold Spring Harb Perspect Biol, 2011. **3**(2).
1352 81. Barbaux, S., et al., *Donor splice-site mutations in WT1 are responsible for Frasier syndrome.*
1353 Nat Genet, 1997. **17**(4): p. 467-70.
1354 82. Hammes, A., et al., *Two splice variants of the Wilms' tumor 1 gene have distinct functions*
1355 *during sex determination and nephron formation.* Cell, 2001. **106**(3): p. 319-29.
1356 83. Heikinheimo, M., et al., *Expression and hormonal regulation of transcription factors GATA-4*
1357 *and GATA-6 in the mouse ovary.* Endocrinology, 1997. **138**(8): p. 3505-14.
1358 84. Efimenko, E., et al., *The transcription factor GATA4 is required for follicular development and*
1359 *normal ovarian function.* Dev Biol, 2013. **381**(1): p. 144-58.
1360 85. Manuylov, N.L., et al., *Ovarian development in mice requires the GATA4-FOG2 transcription*
1361 *complex.* Development, 2008. **135**(22): p. 3731-43.
1362 86. Ma, P., et al., *Compensatory functions of histone deacetylase 1 (HDAC1) and HDAC2 regulate*
1363 *transcription and apoptosis during mouse oocyte development.* Proc Natl Acad Sci U S A, 2012.
1364 **109**(8): p. E481-9.
1365 87. Zhang, D., et al., *Oocyte Casein kinase 1alpha deletion causes defects in primordial follicle*
1366 *formation and oocyte loss by impairing oocyte meiosis and enhancing autophagy in developing*
1367 *mouse ovary.* Cell Death Discov, 2022. **8**(1): p. 388.
1368 88. Herman, L., et al., *Genomic exploration of the targets of FOXL2 and ESR2 unveils their*
1369 *implication in cell migration, invasion, and adhesion.* FASEB J, 2021. **35**(4): p. e21355.
1370 89. De Gendt, K., et al., *A Sertoli cell-selective knockout of the androgen receptor causes*
1371 *spermatogenic arrest in meiosis.* Proc Natl Acad Sci U S A, 2004. **101**(5): p. 1327-32.
1372 90. Morales, C., et al., *PDS5 proteins are required for proper cohesin dynamics and participate in*
1373 *replication fork protection.* J Biol Chem, 2020. **295**(1): p. 146-157.
1374 91. Chakraborty, P. and K. Hiom, *DHX9-dependent recruitment of BRCA1 to RNA promotes DNA*
1375 *end resection in homologous recombination.* Nat Commun, 2021. **12**(1): p. 4126.
1376 92. Benayoun, B.A., et al., *Transcription factor FOXL2 protects granulosa cells from stress and*
1377 *delays cell cycle: role of its regulation by the SIRT1 deacetylase.* Hum Mol Genet, 2011. **20**(9):
1378 p. 1673-86.
1379 93. McKey, J., et al., *Neural crest-derived neurons invade the ovary but not the testis during mouse*
1380 *gonad development.* Proc Natl Acad Sci U S A, 2019. **116**(12): p. 5570-5575.
1381 94. Garcia-Moreno, S.A., et al., *Gonadal supporting cells acquire sex-specific chromatin landscapes*
1382 *during mammalian sex determination.* Dev Biol, 2019. **446**(2): p. 168-179.
1383 95. Koentgen, F., et al., *Exclusive transmission of the embryonic stem cell-derived genome through*
1384 *the mouse germline.* Genesis, 2016. **54**(6): p. 326-33.
1385 96. Rafiee, M.R., et al., *Protease-resistant streptavidin for interaction proteomics.* Mol Syst Biol,
1386 2020. **16**(5): p. e9370.

1387 97. Ewels, P.A., et al., *The nf-core framework for community-curated bioinformatics pipelines.* Nat
1388 Biotechnol, 2020. **38**(3): p. 276-278.
1389 98. Dobin, A., et al., *STAR: ultrafast universal RNA-seq aligner.* Bioinformatics, 2013. **29**(1): p. 15-
1390 21.
1391 99. Li, B. and C.N. Dewey, *RSEM: accurate transcript quantification from RNA-Seq data with or*
1392 *without a reference genome.* BMC Bioinformatics, 2011. **12**: p. 323.
1393 100. Soneson, C., M.I. Love, and M.D. Robinson, *Differential analyses for RNA-seq: transcript-level*
1394 *estimates improve gene-level inferences.* F1000Res, 2015. **4**: p. 1521.
1395 101. Wu, T., et al., *clusterProfiler 4.0: A universal enrichment tool for interpreting omics data.*
1396 Innovation (Camb), 2021. **2**(3): p. 100141.
1397 102. Zhu, L.J., et al., *ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip*
1398 *data.* BMC Bioinformatics, 2010. **11**: p. 237.
1399 103. Ross-Innes, C.S., et al., *Differential oestrogen receptor binding is associated with clinical*
1400 *outcome in breast cancer.* Nature, 2012. **481**(7381): p. 389-93.
1401 104. Fornes, O., et al., *JASPAR 2020: update of the open-access database of transcription factor*
1402 *binding profiles.* Nucleic Acids Res, 2020. **48**(D1): p. D87-D92.
1403 105. Shannon, P., et al., *Cytoscape: a software environment for integrated models of biomolecular*
1404 *interaction networks.* Genome Res, 2003. **13**(11): p. 2498-504.
1405 106. Morris, J.H., et al., *enhancedGraphics: a Cytoscape app for enhanced node graphics.* F1000Res,
1406 2014. **3**: p. 147.
1407 107. Cline, M.S., et al., *Integration of biological networks and gene expression data using*
1408 *Cytoscape.* Nat Protoc, 2007. **2**(10): p. 2366-82.

1409