Genome-wide analysis of H3K4me3 and H3K27me3 modifications throughout the mouse urogenital ridge at E11.5.

Yisheng Yang and Megan J Wilson*.

Developmental Biology and Genomics Laboratory, Department of Anatomy, School of Biomedical Sciences, University of Otago, P.O. Box 56, Dunedin 9054, New Zealand

*Corresponding author

Email: meganj.wilson@otago.ac.nz. Ph. +64 3 470 4695, Fax: +64 479 7254

Abstract

In mammals, the adrenal gland, testis and ovary arise from a common progenitor tissue known as the urogenital ridge (UGR). This small population of cells will adopt a number of different cell fates following sex determination, including forming the precursors of somatic cells (such as Sertoli and granulosa cells) and steroidogenic cells. In addition, this tissues also contains the Wolffian and Müllerian ducts that later form components of the reproductive tracts. A potential mechanism to maintain developmental plasticity of the UGR until gonad formation is through the epigenetic modification of histone proteins.

In order to provide a resource for future studies, we used chromatin immunoprecipitation followed by high throughput sequencing (ChIP-seq) for two histone modifications, H3K4me3 and H3K27me3, in the E11.5 mouse UGR. These marks are both known to reflect the active, repressive or a poised chromatin state. We found that enrichment for each histone mark reflected transcriptional activity in precursor cells of the developing gonad. From the analysis of potential enhancer/regulator peak regions for DNA binding motifs, we identified several candidate transcription factors that may contribute to gonadal cell lineage specification. We additionally identified signaling pathway genes that are targeted by both chromatin modifications. Together, these datasets provide a useful resource for investigating gene regulatory networks functioning during UGR development at E11.5.

Keywords: epigenetics, histone, gonad, urogenital ridge, cell fate

Introduction

The mammalian urogenital ridge (UGR) is a unique developmental structure that has the ability to adopt two quite distinct fates during late embryonic development. Each bi-potential UGR consists of the progenitor cells required to form either a testis or ovary: the supporting cells (Sertoli or granulosa cells), the germ cells (GCs), the interstitial cells (including steroidogenic cells) and endothelial cells (Capel, 2017). Cell lineage RNA analysis (Jameson et al., 2012) has suggested that the gonadal progenitor cells are biased towards one of two cell fates: germ cells are considered male-biased, whereas supporting cells are female-biased. The supporting cells adopt a sex-specific fate first, and this process is completed by embryonic day 12.5 (E12.5) in mice. The germ and interstitial cells follow, starting at E12.5 (Jameson et al., 2012). How cell fate plasticity is maintained within the developing UGR prior to and during early sex determination is unknown.

Sex determination in mammals is a critical step in specification of gonadal cell fate and depends on the chromosomal makeup of the embryo. The presence of a Y chromosome (XY genotype) shifts the bi-potential UGR towards a testicular fate, through the initial expression of the testis determining gene, *Sex determining region Y* (*Sry*). *Sry* in turn activates *Sry-box 9* (*Sox9*) gene expression required for Sertoli cell differentiation and the activation of signaling pathways for the maturation of testis-specific cell types (Sekido and Lovell-Badge, 2008; Wilhelm et al., 2007; Wilson et al., 2005). SOX9 up-regulates *Anti-Müllerian hormone* (*AMH*) to promote the regression of the Müllerian ducts, whereas the Wolffian ducts develop into male-specific structures under the influence of testosterone, produced by the Leydig cells (Munsterberg and Lovell-Badge, 1991; Wilhelm et al., 2006).

In the absence of a Y chromosome (XX genotype), the female developmental pathway is initiated. Unlike male development, the morphology of the gonad does not change drastically in the immediate period after initiation of the female pathway (Mork et al., 2012). Pre-granulosa cell specification occurs at ~E12.5 accompanied by expression of *Forkhead box L2* (*Foxl2*) (Schmidt et al., 2004). *Sox9* gene expression is repressed in the XX gonad by ovary-specific Wnt4/β-catenin/Rspo1 signaling pathway that promotes the expression of *Foxl2* and the differentiation of the Müllerian ducts into female-specific structures (Chassot et al., 2008; Pannetier et al., 2016; Tanaka and Nishinakamura, 2014). The primordial GCs arise near the yolk sac and migrate into the UGR, arriving at ~E10.5 (Molyneaux et al., 2001). GCs located in the developing ovary enter meiosis due to the presence of retinoic acid (RA) produced from the neighboring mesonephros. In contrast, male germ cells are prevented from beginning meiosis, as RA is degraded in the developing testis by the CYP26B1 enzyme (Bowles et al., 2006; Spiller et al., 2017).

Due to the bi-potential nature of the UGR, in some disorders of sex development (DSD) an ovotestis may result, a mix of both male and female associated cell types located within a single adult gonad. This disorders are often a result of mutations in the *SRY* or *SOX9* genes (Vilain, 2011). Conditional knockout of *Foxl2* in the adult ovary results in the formation of Sertoli- and Leydig-like cells, suggesting a role for *Foxl2* in maintaining the ovarian phenotype in the adult as well as the embryo (Uhlenhaut et al., 2009). Likewise, inactivation of *DMRT1* in adult Sertoli cells increases *Foxl2* gene expression, producing granulosa and theca-like cells, along with the production of estrogen (Matson et al., 2011). Together, these studies demonstrate the incredible plasticity of sex-development and the need to continually maintain one pathway (either male or female) over the other throughout the adult lifetime.

One possible mechanism to regulate and maintain gonadal cell fate is through the use of specific histone modifications. Global genome-wide changes to histone modifications could also be involved in initiating differentiation of precursor cells within the gonad and mesonephros. Epigenetic modifications play key roles in cell lineage specification during development by manipulating chromatin structure and altering gene expression (Atlasi and Stunnenberg, 2017). Chromatin is found in two states: euchromatin, an open, unwound formation that allows for an active state of transcription, and heterochromatin, a closed, tightly wound formation that represses transcription (Parker et al., 2004). Post-translational modifications to histone proteins largely occur to amino acids located at the N-terminal tail, influencing DNA accessibility to transcriptional regulators of gene expression. Trimethylation of lysine N-terminal residues can induce either an active or repressed chromatin configuration depending on location. Histone 3 lysine 4 trimethylation (H3K4me3) is a hallmark of actively transcribed genes and is commonly associated with transcription start sites (TSS) and promoter regions, whereas histone 3 lysine 27 trimethylation (H3K27me3) is strongly associated with inactive promoter regions and repressed gene transcription (Bannister and Kouzarides, 2011). Both active and repressive marks can be present at the same genomic regions, indicating a poised state common in undifferentiated cells such as embryonic stem cells (Bernstein et al., 2006; Voigt et al., 2013).

There is limited data regarding the epigenetic profile of gonadal somatic and germ cells *in vivo*. Using cell sorting, researchers have isolated the primordial GCs (PGCs) using Oct4-GFP mouse line (Ng et al., 2013; Sachs et al., 2013) to study epigenetic modifications using chromatin immunoprecipitation followed by sequencing (ChIP-seq). PGCs tend to have high levels of the repressive H3K27me3 mark and are transcriptionally silent (Ng et al., 2013;

Sachs et al., 2013). Bivalent regions, areas with both active and repressive histone modifications, are enriched within PGCs and tend to be located near developmental genes (Sachs et al., 2013). Recently, DNaseI-seq and ChIP-seq for H3K27ac was used to identify active enhancer elements in Sertoli cells at E15.5 (Maatouk et al., 2017).

To obtain a more global view of histone modifications in all cells of the developing bi-potential gonad, we used ChIP-seq to study the distribution of H3K27me3 and H3K4me3 in E11.5 gonads. Profiling of histone mortifications not only gives information about the transcriptional activity of nearby genes, but can aid in the identification of tissue-specific enhancer elements. Through ChIP-seq, we discovered different histone signal profiles that reflect gene expression levels in the E11.5 gonad. Using this data, we identified overrepresented DNA binding motifs to determine candidate transcription factors, and then extracted their cell-type expression values during gonad development using publically available data (Jameson et al., 2012). This data set provides a resource for future studies of early gonadal development.

Results

Immunostaining of UGR at E11.5

Sectioned E11.5 embryos were stained for H3K4me3 and H3K27me3 histone marks (Fig. 1). While all cells were positive for some level of these histone modifications, there were many cells located within the gonad that stained much more strongly than nearby cells (Fig. 1). These strongly-staining cells were located along the coelomic epithelium and scattered

146 throughout the gonad (Fig. 1A). While some of these cells are likely to be primordial GCs,
147 many of these cells are also somatic cells. For instance, cells that line the coelomic epithelium
148 are precursors of supporting cell lineages such as Sertoli cells (Karl and Capel, 1998). Staining
149 for each histone modification was found in the nucleus, often co-localized at nuclear foci (Fig.
150 1B).
151
152 *Genome-wide identification of H3K27me3 and H3K4me3*
153 Urogenital ridge tissue was dissected from each E11.5 stage embryo and pooled (two
154 biological replicates, five litters per replicate,), prior to cross-linking briefly with
155 formaldehyde. Extracted chromatin was fragmented by MNase digestion, followed by
156 immunoprecipitation using antibodies for either H3K27me3 or H3K4me3. Two biological
157 replicates (from separate ChIP experiments) were sequenced on an Illumina platform to
158 detect each histone mark (our approach is summarized in Fig. 2A). Sequencing reads were
159 aligned to the mouse reference genome (mm9, File 1) and correlation analysis revealed good
160 reproducibility between each biological replicate (>93% Pearson correlation, Fig. S1). For
161 global signal distribution analysis, both replicates were pooled.
162
163 We used MACS broad peak caller (bdgbroadcall) to call peak regions, genomic regions with
164 statistically enriched signal compared to the input controls. This identified 32,182 peak
165 regions for H3K4me3 (with an average peak length of 1809 bp) and 13,556 regions for
166 H3K27me3 (with an average length of 1310 bp). Peak location, with respect to gene body
167 features, was analyzed with HOMER (Heinz et al., 2010) and CEAS (Shin et al., 2009) for each
168 histone mark. Of the MACS broadcall regions, 19476 (H3K4me3) and 7527 (H3K27me3)
169 peaks were located near known RefSeq genes. H3K4me3 was more enriched at the 5' end of
170 genes, near the transcriptional start site (TSS) (Fig. 2B and Fig. S2) including CpG islands (Fig

2C). The repressive histone mark, H3K27me3 was more generally enriched for within intergenic regions and introns, and a smaller proportion of peaks were located within promoter regions (Fig. 2B and 2C, Fig. S2). This data is consistent with previous studies, showing that H3K4me3 is found at active promoter regions and enhancers (Santos-Rosa et al., 2002), whereas the H3K27me3 mark is largely distributed across the bodies of genes, with some limited enrichment near the TSS, often at bivalently marked genes (Young et al., 2011).

The distribution of histone marks at five key sex-determining genes was examined further (Fig 3. and Fig. S3). Active histone modification enrichment near promoter regions, and within either an intron or nearby intergenic regions, can indicate the presence of an enhancer region (Atlasi and Stunnenberg, 2017). *Sox9*, *Wt1* and *Sf1* genes are all expressed in the developing UGR at E11.5 and, following sex-determination, play important roles in testicular development (Hammes et al., 2001; Kent et al., 1996; Luo et al., 1994). *Lhx9* and *Cbx2* are required for formation of the mouse gonad primordium (Birk et al., 2000; Katoh-Fukui et al., 2005). The *Sf1* (also known as *Nr5a1*) gene showed enrichment for both histone marks, peak regions were located within the 4th intron and near the promoter region of the gene (Fig. 3A). H3K4me3 signal was strongly enriched for around the promoter region and first exons for *Wt1* and *Wt1os* genes (Fig. 3B). *Wt1os* encodes a long non-coding RNA, co-expressed with *Wt1* in many tissues (Dallosso et al., 2007). A narrow peak region for H3K27me3 signal, overlapping with H3K3me4 signal, was also identified in the 2nd intron (Fig. 3B). An additional H3K4me3 peak region is located ~3 kb, 3' of the *Wt1* gene (Fig. 3B); this may represent an active enhancer element.

Sox9 is a key male sex-determination gene and target of *Sry* (Kent et al., 1996; Sekido and Lovell-Badge, 2008). The H3K4me3 modification was strongly enriched around the *Sox9* gene

196 (Fig. 3C), reflecting the fact that this gene is expressed in both sexes at E11.5 (Kent et al.,
197 1996). There was an additional peak for H3K27me3, located 8.4 kb upstream of the *Sox9*
198 promoter. This genomic region is ~3.5 kbp away from the TESCO (testis-specific enhancer of
199 *Sox9* core element) enhancer, thought to restrict *Sox9* expression to the testis (Bernard et al.,
200 2012; Gonen et al., 2017; Sekido and Lovell-Badge, 2008).
201
202 *Lhx9* and *Cbx2* (also known as *M33*) genes are required for formation of the UGR and are
203 expressed at E11.5 (Birk et al., 2000; Katoh-Fukui et al., 2005). H3K4me3 signal was high
204 around the TSS for both genes (Fig. S3), indicating active promoter regions. There was also a
205 peak for *2310009B15Rik,* a protein-coding gene of unknown function located near the *Lhx9*
206 gene, suggesting that this gene is also expressed at E11.5. Signal for the repressive histone
207 mark H3K27me3 was also found near the Lhx9 TSS and throughout the gene body (Fig. S3A),
208 indicating that in some gonadal cells *Lhx9* is repressed. Previous spatial expression studies
209 have shown that *Lhx9* mRNA is strongest in the cells of the coelomic epithelium, and absent
210 from some of the cells of the inner mesenchyme and mesonephros (Birk et al., 2000). In
211 contrast, H3K27me3 ChIP-seq signal was low around the *Cbx2* gene, with no peak regions
212 being identified by MACS (Fig. S3B), indicating little repression by H3K27me3 modification in
213 the UGR cell population. *Cbx2* exhibits broader cell expression, being found in all cells of the
214 gonad and the neighboring mesonesphros (Katoh-Fukui et al., 2012).
215
216 *Levels of H3K4me3 and H3K27me3 modifications reflect gene expression levels in the UGR.*
217
218 To identify genes with similar histone mark profiles, heatmaps with K-means clustering were
219 generated using deepTools (Ramirez et al., 2014) to visualize the region around the TSS for
220 RefSeq gene annotations and separate them into clusters with similar signal profiles (Fig. 4A

and 5A, File 2). Gene ontology (GO) analysis was carried out for clustered genes using PANTHER to identify overrepresented biological annotations (Fig. 4B and 4C, Fig. 5B and 5C, Files 3 and 4). The expression levels of genes grouped within each signal profile cluster was also extracted from Affymetrix expression data previously generated for each cell lineage present in the E11.5 gonad by Jameson *et al*. (Jameson et al., 2012). This enabled us to determine the relationship between signal clusters and gene expression levels (Fig. 6).

Three clusters were generated for H3K4me3 signal distribution. The first cluster had high levels of H3K4me3 located near the gene body and TSS (Fig. 4A). Cluster 1 genes function in core cellular biological processes including general essential cellular processes such as RNA processing and metabolism (Fig. 4B). The second cluster (cluster 2) genomic regions have moderate levels of H3K4me3 signal (Fig. 4A), and are enriched for genes involved largely in sensory reception and receptor signaling genes (Fig. 4C). There are no significant differences in gene expression levels between cluster 1 and 2 genes (Fig. 6B). However, cluster 3 genes are expressed at significantly lower levels compared to cluster 1 and 2 genes (Fig. 5, two-way ANOVA, $P < 0.0001$). Genomic regions grouped into the third cluster have little low/no H3K4me3 signal (Fig. 4A). Out of 13,043 genes (18,043 RefSeq transcripts) associated within cluster 3 regions, only 3,392 genes are noted as expressed in the E11.5 gonad Affymetrix dataset (Jameson et al., 2012) (26% of genes, File 3). These genes function in a variety of different biological processes (Fig. S4A). In comparison, of the 6166 genes in cluster 1 (corresponding to 7005 RefSeq transcripts), 5056 of these are expressed in the E11.5 gonad (~82% of genes, File 3). This supports previous research showing that H3 genome-wide enrichment is linked to gene expression (Guenther et al., 2007; Mikkelsen et al., 2007).

H3K27me3 signal distribution can be divided into 5 clusters with different profiles. Cluster 1 histone mark covers over 20 kbp of DNA including the gene TSS (Fig. 5A), and is likely to represent areas of repressed chromatin. GO annotation revealed that these genome regions are enriched for genes that function in essential early developmental processes such as germ layer development and axis specification (Fig. 5B). Of these 219 cluster 1 genes, only 58 had detectable expression in single cell profiling (File 4), with expression levels much lower than genes found in other H3K27me3 signal profile clusters (Fig. 5A). The signal profile for cluster 2 genes is strongest around a +/- 5 kb region near the TSS. GO analysis revealed that many cluster 2 genes function in cell-cell signaling, differentiation and transcriptional regulation (Fig. 5C). This signal cluster is associated with 1559 genes (1750 RefSeq transcripts) and 33% of these genes (518 genes) are expressed in the gonad at E11.5 (File 4). As the H3K27me3 mark is associated with gene repression, we would expect expression of genes with high H3K27me3 signal, such as those genes associated with clusters 1 and 2, to be low or not detectable.

Cluster 3 had reduced levels of signal near the TSS compared to nearby regions (Fig. 5A, see enlarged view in Fig. S4B). These genes function in essential biological processes such as metabolism and sub-cellular organization (Fig 5C). Of 3844 genes with this signal profile, 56% (3138 genes) are expressed in the gonad at E11.5 (File 4). H3K27me3 signal profile for cluster 4 is confined to narrow region near the TSS (Fig. 5A). Cluster 4 associated genes are involved in the intracellular signaling including localization and transport of proteins (Fig. S5). Of the 5294 genes associated with this cluster, 2,878 are expressed in the early gonad (File 4). Cluster 5 gene regions had no enrichment for H3K27me3 signal, and this was the largest group of genes (17219 genes corresponding to 22163 RefSeq transcripts). Approximately 54% of these genes are expressed at E11.5 (9,447 genes, File 4). When comparing normalized

gene expression levels between grouped genes, cluster 3-5 genes are expressed at levels significantly higher than cluster 1 and 2 genes (Fig. 6C, $P < 0.0001$).

Genomic regions common to both data sets

To identify regions of enrichment common to both histone marks, peak regions generated by MACSbroadpeakcall for each mark were intersected based on at least 1bp overlap, to produce a list of 12,296 regions (Fig. 7A). Examples of signal distribution for two genes of interest, *Fgf9* and *Esr1,* is shown in Fig. 7C and D, with genome regions common to both marks indicated. These common peak regions were used with GREAT (McLean et al., 2010) to identify genes nearby each genomic region (File 6). These genes may lie in regions of the genome that is being actively transcribed, or either repressed or poised for future activation due to the presence of H3K27me3 in addition to H3K4me3.

The resulting gene list (6523 genes) was used for GO functional annotation in PANTHER (Mi et al., 2010) (Fig. 7B and File 6). PANTHER pathway analysis showed that Wnt, Fgf and Tgfβ signaling pathways were significantly overrepresented (Fig. 7B), pathways that have previously been shown to be essential for early gonad development (Colvin et al., 2001; Gustin et al., 2016; Josso and di Clemente, 1999; Kim et al., 2006). In particular, Wnt signaling pathway associated genes (123 genes) were particularly enriched within this list (Fig. 7B). Other pathways of interest included angiogenesis, important for formation of the gonadal blood supply and patterning of the testicular cords during development (Combes et al., 2009; Coveney et al., 2008). Integrin signaling is also required for correct development of the gonad and their primordial germ colonization (Anderson et al., 1999; Messina et al., 2011).

Genes linked to pathways with unknown roles in gonad development were also overrepresented. Gonadotrophin-releasing hormone (GnRH) signaling has been best studied in the adult mouse, where GnRH neuronal signaling peptide loss of function leads to hypogonadism in adult mice. GnRH neuronal signaling does not appear to be functional until after E16.5 (Wen et al., 2010). The *GnRH receptor* (*GnrhR)* gene is expressed in the rat embryonic testis Leydig cells, adult ovary and breast tissues (Ishaq et al., 2013; Schang et al., 2012) but has not been reported to have significant expression in the developing embryo gonad and was included in the array RNA expression data (Jameson et al., 2012). However, the gene for the *Gnrh1* hormone is expressed in developing chick gonad (Carre et al., 2011) and mouse gonad (Fig. S6, (Jameson et al., 2012)).

Several genes linked to glutamate-receptor signaling pathways were also overrepresented on the resulting gene list. These genes are expressed in many non-neural tissues including the adult testis (Julio-Pieper et al., 2011; Marciniak et al., 2016), the developing gonad (Jameson et al., 2012) and are believed to contribute to organ homeostasis. Cholecystokinin receptor (CCKR) pathway components are also overrepresented. While cholecystokinin (a peptide hormone) itself is not expressed in the gonad or has any known function in gonad development, many CCKR downstream pathway genes such as β-catenin, *Jun* and c-Myc are known developmental regulators and are expressed in the developing gonad (Jameson et al., 2012). This may be why this pathway is significantly overrepresented. The CCKR pathway has also been linked to sex dimorphic responses in the brain (Xu et al., 2012).

DNA sequences associated with H3K4me3 enrichment

To identify candidate factors that might bind to enhancers associated with each histone mark, we used HOMER motif analysis. Narrow peak regions (~1kb) were first identified, as these

320 are more likely to represent a regulatory element rather than global repression/activation of
321 an area of the genome (Fig. 9A). This was followed by HOMER motif enrichment for these
322 narrow peak regions. The HOMER software produces a list of known motifs for transcription
323 factors significantly overrepresented when compared to a set of randomly selected
324 background sequences. This list of enriched motifs was further filtered to those sites whose
325 associated transcription factors are expressed in the developing gonad, by comparison to the
326 available Affymetrix array data (Jameson et al., 2012).
327
328 We identified 9332 narrow (1kb) regions with H3K4me3 enrichment (Fig. 8A and File 6). A
329 total of 67 motifs were enriched within these peak sequences, of which 33 of the known
330 motifs correspond to transcription factors that are expressed during gonad development (File
331 6). Seven of these factors are expressed in a cell-type specific manner (Fig. 8B-H).
332
333 Sixty one percent of H3K4me3 regions had a Myb (c-Myb) DNA binding motif (q-value =
334 0.007). Myb is a helix-turn-helix transcription factor with roles in cell cycle regulation (Nakata
335 et al., 2007), hematopoiesis (Mukouyama et al., 1999) and oocyte meiotic maturation(Zheng
336 et al., 2012). *Myb* gene expression was higher in GCs at E11.5 compared to other gonadal cell
337 types (Fig. 8B).
338
339 Nanog is a well-characterized transcription factor involved in maintaining pluripotency in
340 embryonic stem cells, proliferation and nuclear reprogramming in primordial germ cells
341 (Theunissen et al., 2011; Yamaguchi et al., 2009). Over 80% of H3K4me3peak regions (q-
342 value <0.0001) had potential Nanog binding motifs (File 6). Nanog was expressed at much
343 higher levels in GCs compared to other cell types found in the E11.5 gonad (Fig. 8C).
344

345 *Nuclear factor 1* (*NF1*) gene encodes a ubiquitous transcription factor with roles in chromatin
346 remodeling (Gonen et al., 2017) and has been found to function with Sf-1 in adrenal
347 steroidogenesis (Aigueperse et al., 2001). Sixty percent of H3K4me3 peak sequences had a
348 potential NF1 binding site (q-value = 0.02). *Nf1* mRNA is expressed at high levels in all
349 gonadal cell types with the exception of the primordial GCs at E11.5 (Fig. 8D). Expression in
350 male and female GCs eventually increases following sex determination (E12.5) to similar
351 levels of other gonadal cell types by E13.5 (Fig. 8D).
352
353 Nine members of the Ets family of transcription factors were identified as binding to motifs
354 located in H3K4me3 narrow peak regions. This protein family has roles in cell lineage
355 specification including blood cell differentiation (Ciau-Uitz et al., 2013; Remy and Baltzinger,
356 2000). The ETS factor, *Elf1* (*E74-like factor 1)* has an essential role in vascular development
357 (Gaspar et al., 2002; Huang et al., 2006). Around a third (34%) of H3K4me3 narrow peak
358 regions contained *Elf1* binding motifs (q-value <0.0001). *Elf1* mRNA expression levels were
359 lower in germ cells at E11.5, when compared to somatic and interstitial/stroma cells (Fig. 8F,
360 P<0.0001). mRNA for another ETS factor, *Friend leukemia integration 1 (Fli1),* was expressed
361 at higher levels in endothelial cells (P<0.0001) compared to other cell lineages (Fig. 9E).
362
363 *Interferon regulatory factor 2* (*Irf2*) was expressed at much higher levels in the endothelial
364 cell types, compared to both the somatic and germ cells (Fig. 8G, P<0.001 and P<0.0001
365 respectively). A small subset of sequences had a Irf2 DNA binding motif (5%, q-value <0.0001,
366 File 6). *Elf-1* and *Irf2* have both been found to induce stem cells to form blood cells in culture
367 (Yamamizu et al., 2013).
368

The DNA binding motif for TEA domain family member 4 (Tead4, also known as TEF3) was enriched for 30% of sequences (q-value 0.01, File 6). Expression profiling indicates the *Tead4* gene is expressed at higher levels in endothelial cells compared to other cell types present in the gonad (Fig. 8H). Previously, Tead4 was found to directly regulate *Sf1* expression in reporter assays (Sakai et al., 2008), and has been linked to VEGF-mediated angiogenesis (Liu et al., 2011)

DNA sequences associated with H3K27me3 marks

For H3K27me3, 4,847 regions were used in motif analysis (File 6) and identified enrichment for 15 known motifs for transcriptions factors. Of these, 6 transcription factors are expressed in the E11.5 gonad (Fig. 9B-E).

TG-interacting factor 1 (*Tgif1*) gene encodes for a homeobox protein that acts as a repressor of TGF-β and RA signaling (Bartholin et al., 2006; Wotton et al., 1999), and can repress embryonic stem cell factors, binding directly to Oct4 (Lee et al., 2015). The majority of H3K27me3 peak regions (88%) had a predicted Tgif1 DNA binding motif (q-value = 0.025). At E11.5, *Tgif1* mRNA is expressed at high levels in the undifferentiated somatic, endothelial and interstitial/stroma tissue, compared to GCs (Fig. 8B, P<0.001). By E13.5, *Tgfif1* mRNA is expressed at higher levels in male GCs, compared to female GCs E13.5 (Fig. 9B, P<0.0001).

Signal transducer and activator of transcription 4 (*STAT4*) motif was enriched in ~43% (q-value = 0.03) of H3K27me3 narrow peak sequences (File 6). The *STAT4* gene is expressed at higher levels in both male and female germ cells at E11.5, in comparison to other cell types (Fig. 9C, P<0.0001). STAT4 has multiple roles in development, including in endothelial cell

393 proliferation in zebrafish (Meng et al., 2017), and is expressed in the germ cells of the adult
394 mouse testis (Herrada and Wolgemuth, 1997)
395
396 The *Nuclear factor of activated T-cells (NfatC1)* gene has roles in lymphatic and valve
397 endothelial development (Kulkarni et al., 2009; Wu et al., 2011). *NfatC1* was expressed at
398 higher levels endothelial cells compared to other cell types in the E11.5 gonad (Fig. 9D,
399 P<0.0001). Nine percent of H3K27me3-associated sequences (q=0.0009) were predicted to
400 contain NfactC1 DNA binding motifs (File 6).
401
402 The *Forkhead box protein O1* (*FoxO1*) gene has a role in regulating human and mouse ESC
403 pluripotency through regulation of *SOX2* and *OCT4* gene expression (Zhang et al., 2011), and
404 also functions in the adult mouse male germline and ovarian granulosa cells (Liu et al., 2013).
405 *Foxo1* mRNA levels were higher in GCs and endothelial cells at E11.5 compared to other
406 gonadal cell types (Fig. 9E, P<0.0001). The FoxO1 binding motif was present in 70% of peak
407 sequences (q-value = 0.0001).
408
409 **Discussion**
410
411 We successfully performed ChIP-seq analysis for H3K4me3 and H3K27me3, histone marks
412 that represent two different transcriptional states, on E11.5 UGRs. The resulting signal
413 distribution profiles reflected transcriptional activity in the E11.5 gonad and may indicate
414 areas of the genome either in an active or repressed state, including potential enhancer
415 regions.
416

Immunofluorescence staining of the UGR revealed that many gonadal cells had higher levels of these histone modifications compared to nearby cells. While some of these cells are likely to be PGCs, others are likely to represent precursors of other cell lineages required for gonad development. Future studies will address how these marks are first established and what happens to these marks over the course of gonadal development. This study used a time-point at the beginning of sex determination and prior to hormone production, and we do not address epigenetic dimorphism between the sexes at E11.5. Therefore, we cannot discount the possibly that sexually dimorphic epigenetic patterns exist prior to sex determination (earlier than E11), and perhaps are based on the karyotype of the embryo. Histone modifications are enzymatically added and removed, therefore the enzymes involved are important regulators of cell differentiation and fate (Butler et al., 2012). While the X and Y-chromosomes both have lysine demethylase genes, known as *Jarid1c* and *Jarid1d* respectively, it is unknown if they have a role in early gonad development. However, in terms of sex-determination, the histone demethylase enzyme, JmjC domain-containing protein (Jmjd1a) was recently shown to be required for correct expression of the *Sry* gene and male sex determination in mice (Kuroki et al., 2017). It will also be important to consider when and how any sex-specific histone modifications are first established during cell-fate specification in the developing gonad (Butler et al., 2012). (Kuroki et al., 2017).

Epigenetic modifications are essential for cell fate and lineage specification. DNA binding motifs located within narrow peak regions included transcription factors that were themselves expressed in a cell-specific manor. This identified several new candidate transcription factors that may have a role in the undifferentiated gonad. Tgif1 is a known repressor of retinoic acid signaling (Bartholin et al., 2006) and is expressed at significantly higher levels in male GCs by E13.5 (Fig. 9B). Repression of RA signaling prevents male GC

entry into meiosis (Bowles et al., 2006). Expression of another transcription factor, cMyb was also highest in E11.5 GCs (Fig. 8B). This gene was previously proposed to regulate *KIT* expression (Mithraprabhu and Loveland, 2009), KIT is required for germ cell migration and maturation (Fleischman, 1993).

Another transcription factor gene of interest was *Nfatc1*. Gonad *Nfatc1* gene expression is much higher in endothelial cells (Fig 9D), suggesting it has a role in vascular cell fate in the gonad. Previously, it was shown that *Sox9* inactivation in the mouse heart resulted in ectopic expression of *Nfatc1* in heart mesenchymal cells, and in wild-type animals *Nfatc1* is expressed only in the endothelial cells of the heart (Akiyama et al., 2004). This suggests a negative relationship between *Sox9* and *Nfatc1*, perhaps in restricting cell lineage specific-gene expression.

Several transcription factors predicted to bind to other DNA binding motifs enriched within our dataset were not expressed in the developing gonad based on available array data (File 6). However, these factors may have roles later in development, or in the post-natal gonad. *Recombination signal binding protein for immunoglobulin kappa J region* (*Rbpj*) gene encodes for a Notch pathway regulator, and conditional knockout results in reduced testis size. *Rbpj* is required for the testis stem cell niche (Garcia et al., 2014). Therefore, while *Rbpj* gene expression is detected in the early embryonic gonad, it also affects male fertility later in life.

We determined a set of genomic regions that were enriched with both histone marks in chromatin extracted from E11.5 gonads. Different cell lineages have distinct epigenetic profiles (Hawkins et al., 2010). From this data, we cannot be sure that a given genome region

466 is actually marked by both modifications in the same cell, as we used a mixed population of
467 cells, and we thus are observing the average signal profile over multiple cell types. Therefore,
468 these regions cannot be termed bivalent regions based on this analysis. Future studies, using
469 cell-sorting methods, such as those used by Jameson *et al.,* 2012, and the development of new
470 ChIP methods requiring less material may yield enough sequencing data to examine the
471 epigenetic landscape of individual cell populations within the UGR.
472
473 Wnt signaling pathway genes were significantly overrepresented as being epigenetically
474 marked with both modifications. This pathway has roles in formation of the UGR for both
475 sexes, but at later stages of gonad development, it functions in largely in ovarian development
476 only (Boyer et al., 2010; Ottolenghi et al., 2007). Wnt4 and Rpo1 are essential for proliferation
477 of coelomic epithelium in both XY and XX gonads (Chassot et al., 2012). The absence of *Wnt4*
478 in female mice results in masculinization with the development of Wolffian ducts and absence
479 of Müllerian ducts (Bernard and Harley, 2007; Vainio et al., 1999). Further studies have
480 shown that canonical *Wnt4* expression persists in the developing ovary to prevent testicular
481 formation by repressing the SOX9-Fgf9 positive-feedback loop (Kim et al., 2006). *Wnt5a* is
482 required for the development of posterior Müllerian duct structures (the cervix and vagina),
483 primordial germ cell migration and testicular development (Chawengsaksophak et al., 2012;
484 Mericskay et al., 2004).
485
486 Overall, this study provides first insight into the epigenetic landscape of the entire UGR at
487 E11.5 and provides an important dataset for further analysis of regulatory elements active
488 during early gonad development
489

Methods

Animals

The University of Otago Animal Ethics Committee granted ethical approval for this study (approvals ET17/12 (November 2012) and ET25/15 (May 2015)). Embryonic development was timed from the presence of a copulation plug, beginning at 0.5 days post coitum (dpc) to account for the period between fertilization and time of plug observation.

Chromatin immunoprecipitation

Gonad primordium and associated mesonephros tissue was dissected at E11.5 and pooled in 1 mL 1x PBS on ice and fixed in 25 µL 37% formaldehyde to crosslink together the DNA and associated proteins while rocking for 10 min. The crosslinking reaction was stopped with glycine (final concentration of 1 mg/mL). Five litters of crosslinked UGR tissue were pooled together, pelleted by centrifugation and resuspended in 500 µL membrane extraction buffer (20 mM TrisCl pH 8.0, 85mM KCl, 0.5% NP-40) plus 5 µL of proteinase inhibitor (20 µg/µL).

Homogenized tissue was placed on ice for 10 min and centrifuged at 9,000g for 3 min. Pellets were resuspended in 500 µL digestion buffer (50 mM TrisCl pH 8.0, 5mM $CaCl_2$, containing 0.5 µL DTT (100 mM) and 1 µL MNase (300 U/µL)) and incubated at 37^{0}C for 30 min, with mixing every 5 min to digest the chromatin. After the incubation, 50 µL of stop solution (200mM EGTA (pH 8.0)) was added and the sample was placed on ice for 5 min. Samples were centrifuged at 9,000g for 3 min and the supernatant removed. To extract the digested chromatin, nuclear extraction buffer (1% SDS, 10 mM EDTA pH 8.0, 50 mM TrisCl pH 8.0) was added to the pellet to give a final volume of 250 µL and sample left on ice for 15 min, with vortexing every 5 min for 15 s. Samples were centrifuged at 9,000g for 3 min, with the supernatant now containing the desired fragmented chromatin. 10 µl was retained to serve as the "input" control sample for sequencing.

To bind the antibodies to the magnetic beads, 25 µL of Dynabeads (Life Technologies) were washed 3 times in block solution (0.5% BSA in PBS). 10 µg of primary antibody (either anti-H3K27me3 or anti-H3K4me3) in 1 mL block solution was added to the Dynabeads. Mock tubes were set up without primary antibodies, containing 1 mL of block solution, 25 µL Dynabeads and preimmune serum. Beads were rotated at 4^0C overnight, washed 3 times in block solution to remove unbound antibody and resuspended in 1.25 mL block solution. Equal volumes of the chromatin-containing supernatant was added to the primary antibody or mock tubes. Microfuge tubes containing Dynabeads were left to rotate at 4^0C overnight. Antibodies used for ChIP and immunofluorescence were anti-H3K4me3 (ab8580) and anti-H3K27me3 (ab6002), both from Abcam.

Following immunoprecipitation, the beads were washed five times with wash buffer 1 (50 mM HEPES-KOH (pH 7.5), 500mM LiCl, 1 mM EDTA, 1.0% NP-40 and 0.7% sodium-deoxycholate), followed by two washes with wash buffer 2 (TE with 50 mM NaCl). Bound nucleoprotein complexes were extracted with elution buffer (50 MM TrisCl pH 8.0, 1% SDS) at 65°C for 10 min. Protein-DNA cross-linking was reversed for input and ChIP fragments and proteins digested by Proteinase-K (1 µg) 65°C for 90 min, followed by phenol-chloroform extraction. DNA samples were used to construct Thruplex DNA libraries and sequenced on a Illuminia HiSeq platform generating 50 bp single end reads. Sequencing reads were aligned to the mouse genome (mm9) prior to peak calling.

ChIP-seq data analysis

538 Sequences returned were assessed for quality by Phred(Q) score. High quality reads were
539 aligned to the mouse genome (mm9) using Map with BWA (Galaxy Version 0.3.1, with default
540 settings). Poorly mapped reads (MAPQ score <20) were removed by filtering with
541 samtool_filter2 (1.1.1). To examine correlation between replicates, a correlation matrix was
542 generated using multiBamSummary (2.4.1) with a bin size of 10000. The correlation
543 heatmaps were plotted using plotCorrelation (2.4.1), using Pearson correlation.
544
545 To generate BigWig files for signal visualization and bedgraph files for MACS2 analysis, ChIP-
546 seq output (plus antibody) aligned reads were normalized to the Input signal using
547 bamCompare (2.5.0) using the following settings: 50 bin, readCount set to scale samples to
548 the smallest dataset and ratio to compare to input sample.
549
550 To annotate signal with respect to gene features, HOMER and CEAS software packages were
551 used (Heinz et al., 2010)(Heinz et al., 2010; Shin et al., 2009). HOMER was used to annotate
552 genes and gene regions (annotatepeaks function, tag count distribution across the gene body).
553 CEAS (1.0.0) analysis was carried out using default settings (bin size = 50) with previously
554 generated bigwig files and MACSbroadcall peak regions for each histone mark.
555
556 The Deeptools (Ramirez et al., 2014) package was firstly used to generate a matrix
557 (computematrix (2.5.0)) to calculate the signal around the gene reference point (start of the
558 region, TSS as the reference-point with a bin size of 50). PlotProfile was used to plot the signal
559 over gene regions. Heatmaps were created using plotHeatmap (2.5.0,
560 averageTypeSummaryPlot=mean, with kmeans clustering).
561

562 MACS2 (Feng et al., 2012) bdgbroadcall package (2.1.1) using the normalized (to input)
563 bedgraph file to call variable size peak regions, with the following settings: cutoff for peaks set
564 as 5, minimum length of peak to 200, maximum gap between significant peaks was set to tag
565 size (50) and maximum linking between significant peaks of 1500.
566
567 ***RNA expression data analysis.***
568 Robust Multi-array Average (RMA) normalized values for each gene and cell lineage were
569 obtained from Jameson et al., supplementary data files (Jameson et al., 2012). This study used
570 mouse lines with cell-specific markers and fluorescence-activated cell sorting (FACS) to
571 isolate difference cell populations. Markers used were: *Sry*-EGFP/*Sox9*-ECFP supporting cells,
572 *Mafb-EGFP* to isolate interstitial (XY) and stromal cells (XX), germ cells were isolated using
573 *Oct4*-EGFP and endothelial cells with *Flk1*-mCherry. Gene expression were analyzed using a
574 Affymetrix Genechip 1.0 ST array (Jameson et al., 2012) and reported as normalized RMA
575 values.
576
577 ***Motif identification using HOMER***
578 HOMER (Heinz et al., 2010) was used to create tag libraries for each ChIP-seq data set.
579 Findpeaks (part of the HOMER package) was used to find genomic regions of ~1kb signal
580 enrichment compared to surround regions, with following parameters: -minDist 2500 -
581 tagThreshold 25 -L 0. The coordinates of the regions identified for histone marks are listed in
582 additional data File 6. To find motifs within HOMER 1kb ChIP-seq regions, we used findMotifs
583 (HOMER) using default parameters (-size given).
584
585 ***PANTHER GO and Pathway analysis***

The RefSeq transcript list for each signal profile cluster was used for PANTHER GO-Slim-Biological processes with Fisher's Exact with FDR multiple test correction. The returned GO annotations and their adjusted p-values were visualized using Revigo and Treemap (Supek et al., 2011).

Genomic coordinates for regions that enriched within both MACS peak data sets were submitted Genomic Regions Enrichment of Annotations Tool (GREAT (McLean et al., 2010)) using default parameters and mouse genome (mm9) as background, to generate a list of genes nearby and overlapping with each peak region. This gene list was inputted into PANTHER for GO analysis (File 5). PANTHER pathway analysis P-values are adjusted for Bonferroni correction.

Immunofluorescence of paraffin sections

Sections of E11.5 embryos were de-paraffinized and rehydrated as described previously (Yang and Wilson, 2015). For antigen retrieval, slides were places in sodium citrate buffer (pH 6.0) and heated in a microwave for 30 min. Slides were left to cool at room temperature and then rinsed in dH_2O. Slides were washed three times with PBTx (PBS with 0.1% Triton X-100) and then blocked with 10% heat inactivated sheep serum in PBTx for 2 h at room temperature. Primary antibodies were diluted in blocking solution and incubated overnight at 4°C. For each secondary, a control staining (no primary antibody) was performed. Samples were washed with PBTx three times to remove unbound primary antibodies and then blocked for 20 min. Secondary antibodies were diluted in blocking solution and added to the slides before incubation at room temperature for 2 h. Samples were washed three times with PBTx before mounting. Staining was imaged on a Zeiss LSM 710 confocal microscope.

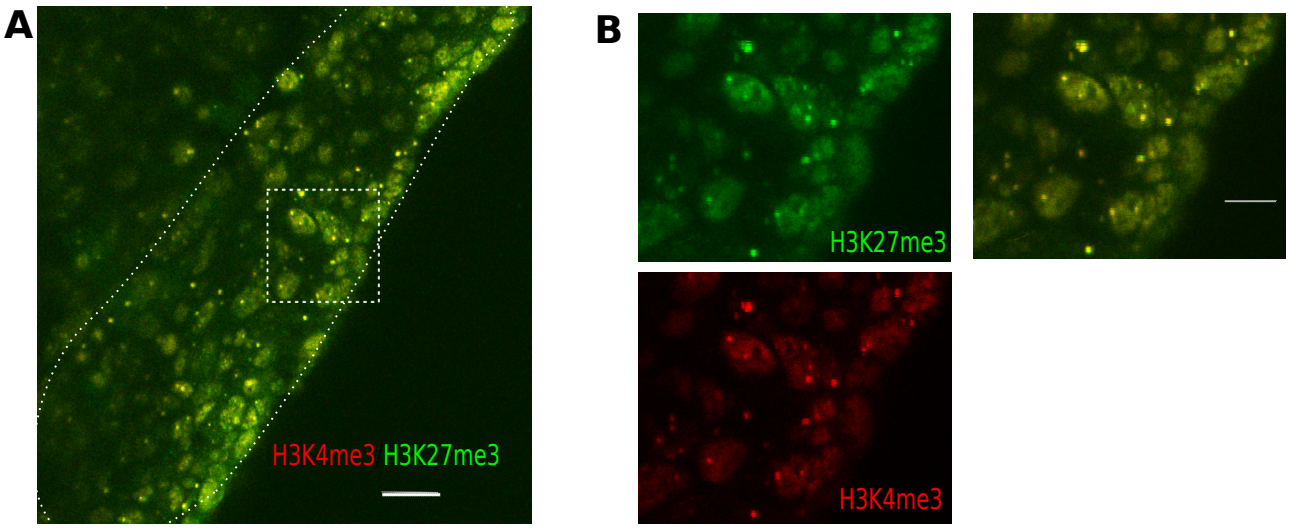
611 Secondary antibodies used were donkey anti-rabbit (Alexa Fluor 488, Life Technologies) and
612 anti-mouse (DyLight® 488, Abcam) to detect anti-H3K4me3 and anti-H3K27me3 respectively.
613 Antibody dilutions used were as follows: anti-H3K4me3 (1 in 50), anti-H3K27me3 (1 in 50),
614 donkey anti-rabbit (1 in 100), donkey anti-mouse (1 in 400).
615

616 **Data access**

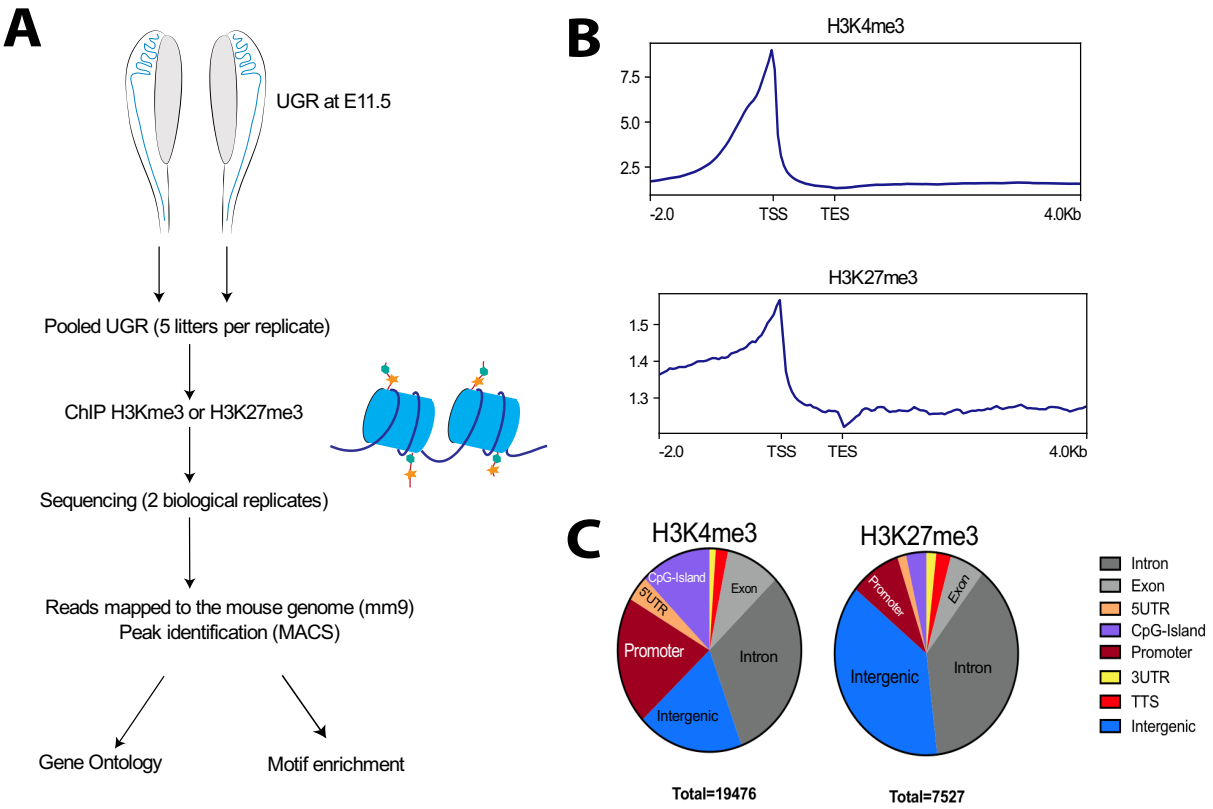
617 ChIP-seq data has been deposited in Gene Expression Omnibus (GEO) with the accession
618 number GSE109380.
619

FIGURE LEGENDS

Figure 1. Detection of histone marks H3K27me3 and H3K4me3 marks in the UGR. A. Double immunofluoresence staining for H3K27me3 (green) and H3K4me3 (red). The gonad is outlined with a dotted line. Scale bar = 20 μm. Dashed box indicates region shown in **B**.



628 **Figure 2 ChIP-seq for H3K4me3 and H4K27me3 histone marks. A**. Overview of the ChIP-
629 sequencing experiments. **B.** Signal profile with respect to the average gene body and 2 kb
630 upstream, 4kb downstream. **C**. ChIP-seq peak location with respect to genome annotations of
631 protein coding genes. Abbreviations: Transcriptional start site (TSS), transcription end site
632 (TES).

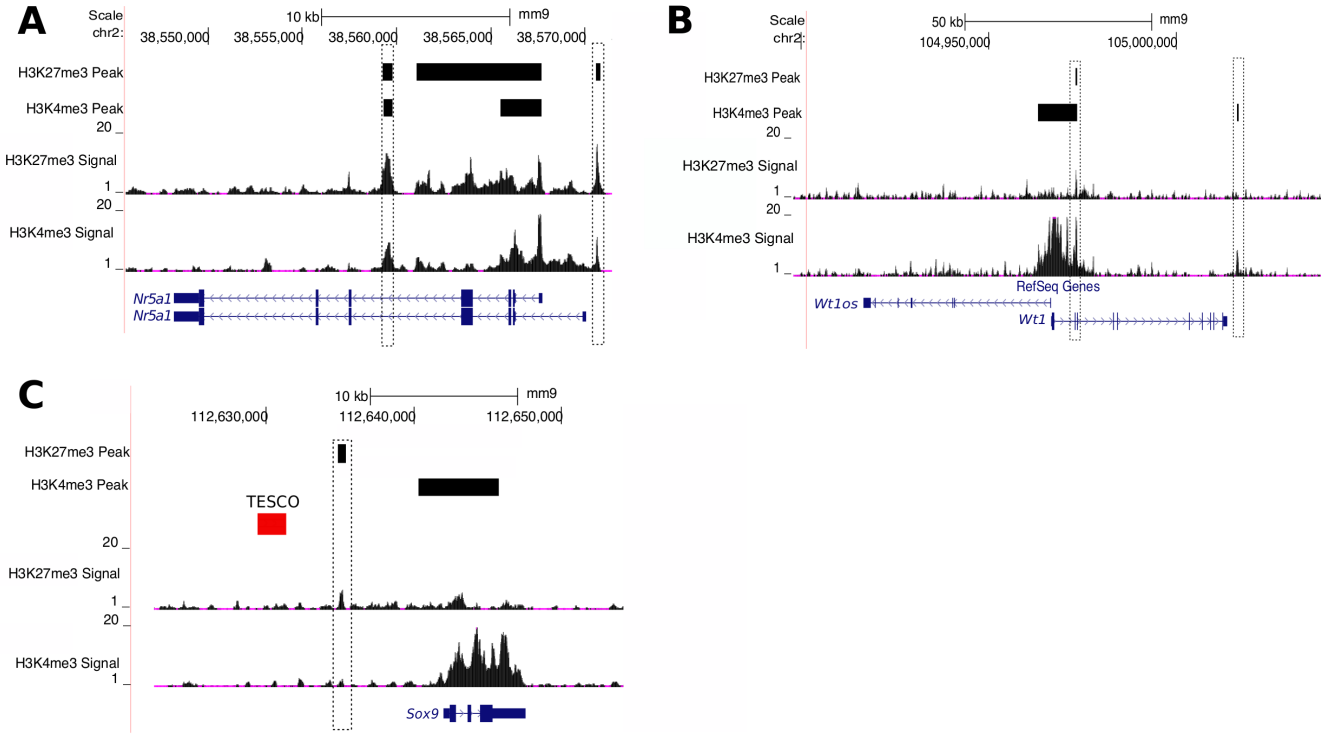


633

634

635

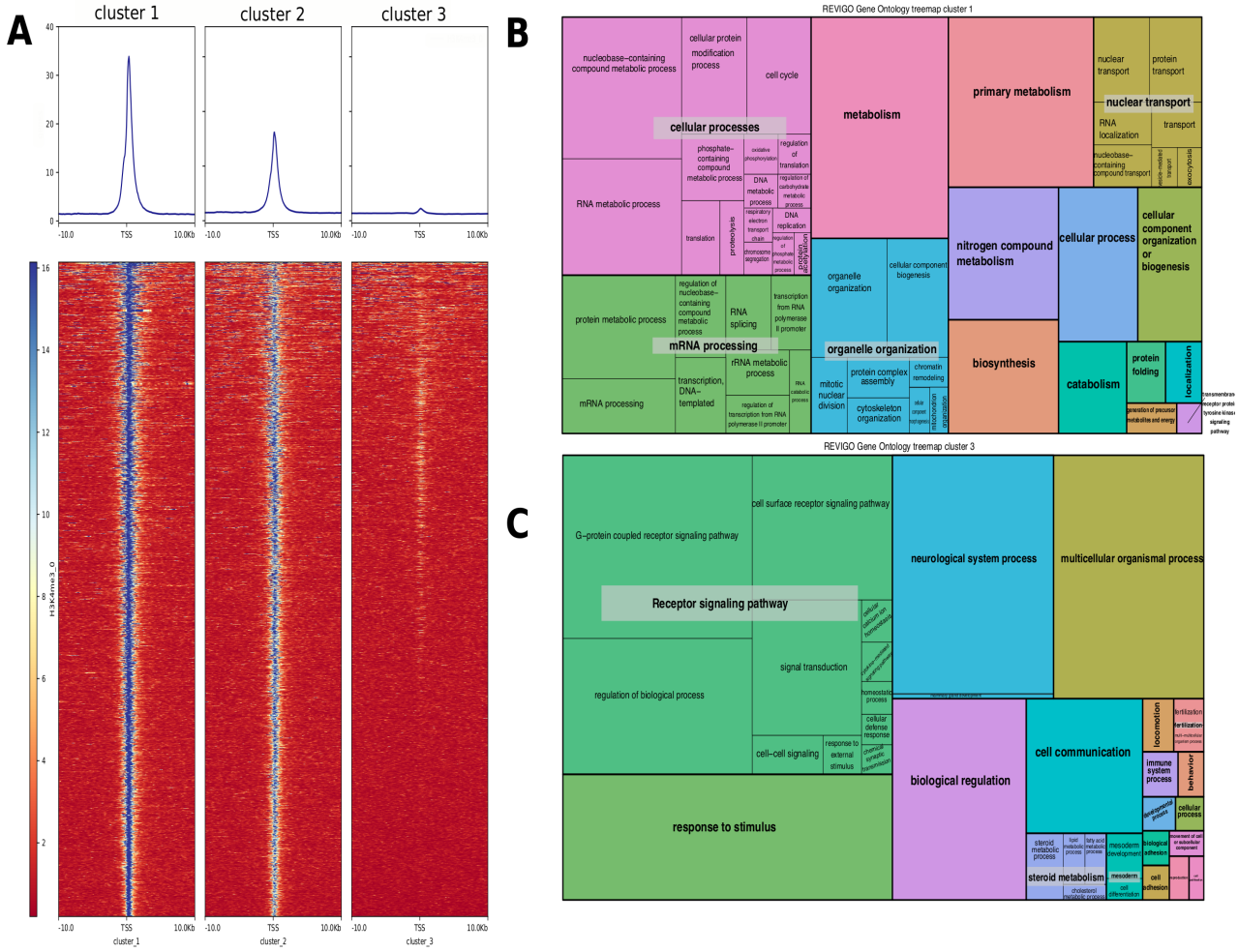
636 **Figure 3. Examples of signal distribution near three genes required for early gonad**
637 **formation and sex-determination. A**. *SF1/Nr5a1* gene **B.** *Wt1* gene. **C.** *Sox9* gene. The TESCO
638 enhancer (red bar) is also shown for the *Sox9* gene. Signal tracks for each histone mark are
639 shown above each gene model. Black bars indicate peak regions identified by MACS analysis.
640 Boxed region indicates enriched signal for both histone marks and therefore, the location of a
641 putative enhancer.
642



643
644

Figure 4 Heatmap profile for H3K4me3 and GO analysis for RefSeq genes cluster.

A. Heatmap of H3K4me3 signal density using K-means clustering. Average signal profile is shown above each heatmap. Read counts were considered around the TSS (+/- 10 kb). Blue indicates low signal and Red indicates an area of high signal. **B.** Treemap summarizing GO analysis for cluster 1 genes (PANTHER slim biological processes). **C.** Treemap visualizing cluster 2 GO analysis (PANTHER slim biological processes). Boxes with similar GO-terms are grouped together and are displayed with the same color. The box size represents the –log10 P-value of GO-term enrichment. Additional Treemap for cluster 3 is provided in supplementary Fig. S4.



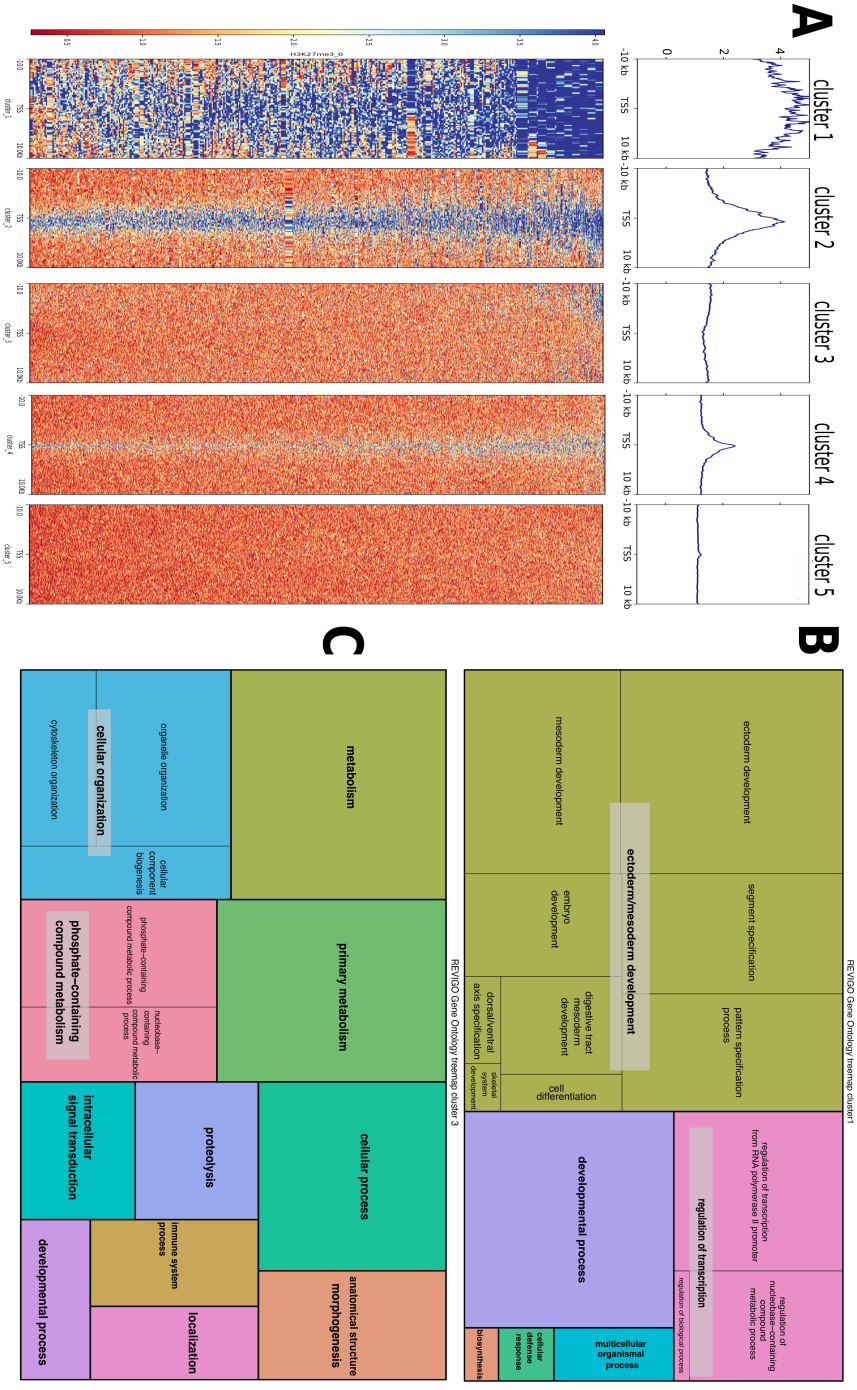
654

655 **Figure 5. Cluster and GO analysis for H3K27me3 whole genome distribution.**

656 **A.** Signal distribution heatmap and K-means clustering for H3K27me3. Read counts were
657 considered around the TSS (+/- 10 kb). Blue indicates low signal and Red indicates an area of
658 high signal. **B.** Treemap of GO pathways (PANTHER slim biological processes)
659 overrepresented for cluster 1. **C.** Treemap for cluster 3 enriched GO-terms (PANTHER slim
660 biological processes). GO treemaps for the other signal clusters are provided in Fig. S5.

661

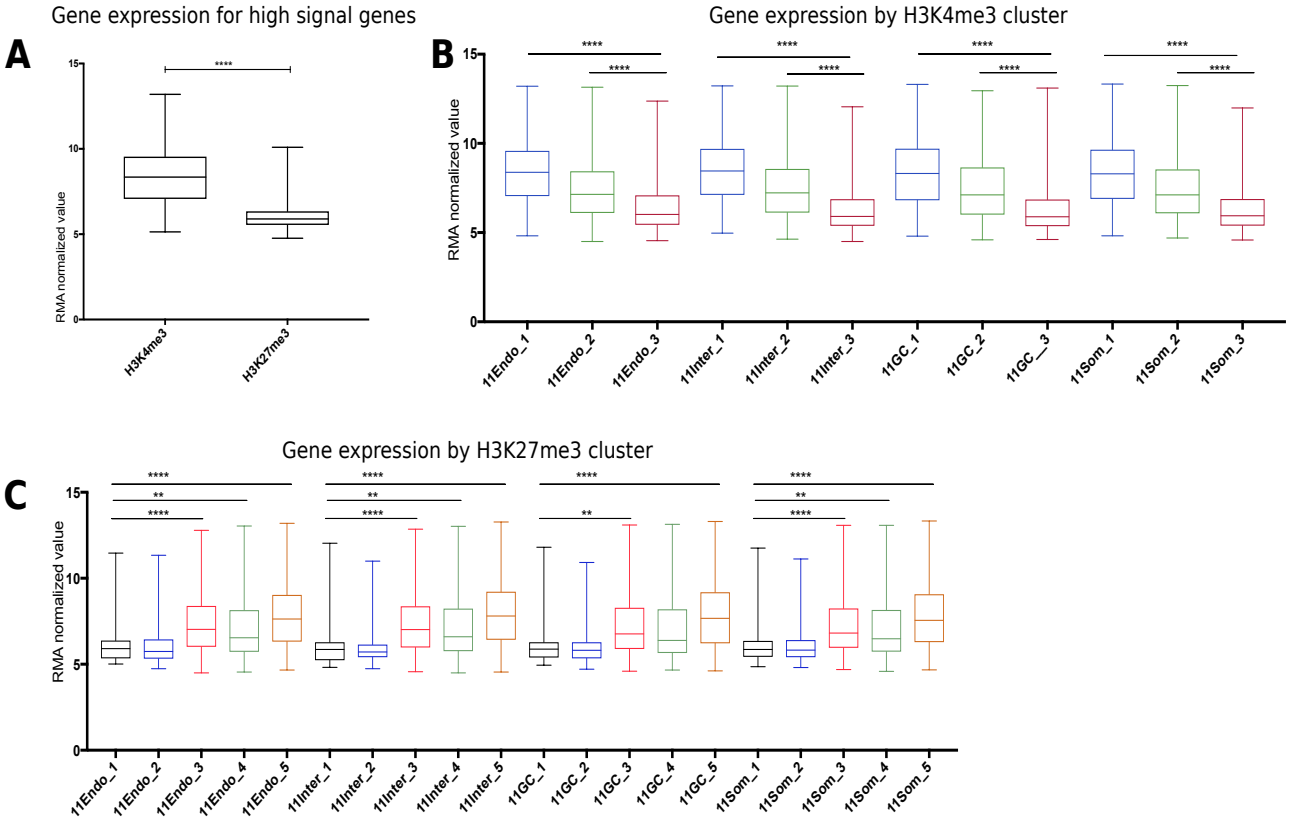
662



663
664

Figure 6. Gonadal expression of genes overlapping with distinct H3K4me3 and H3K27me3 signal profiles.

A. Average gene expression for all cell types, for H4K4me3 cluster 1 and H3K27me3 cluster 2 genes. These two clusters have high signal density for each histone mark around the gene body. **B and C**. Expression of genes associated with each signal profile cluster, per cell lineage. One-way ANOVA Tukey's multiple comparisons test adjusted P-value <0.001 (**), <0.0001 (****). Abbreviations: endothelial cells (Endo), germ cells (GC), interstitial/stroma (Inter), somatic cells (Som). Cluster number is indicated for each plot (eg 11Endo_1 is gene expression of cluster 1 genes in the endothelial cell lineage). Normalized gene expression data was taken from (Jameson et al., 2012).



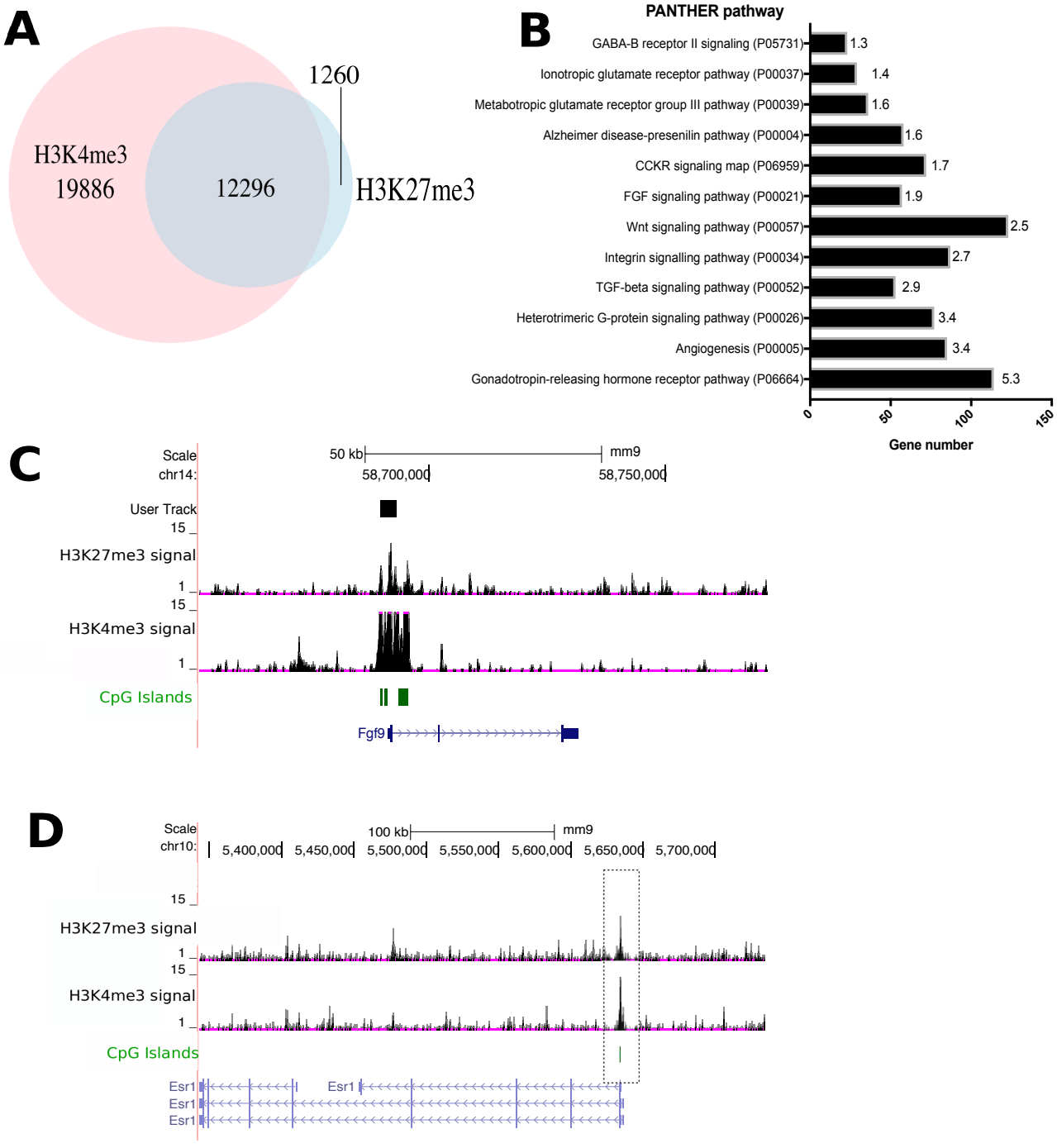
677
678

679 **Fig. 7. Regions overlap between H3K27me3 and H3K4me3.**

680 **A.** Identification of common H3K4me3 and H3K27me3 MACS peak regions. **B.** PANTHER
681 pathway classification for genes associated with H3K4me3-H3K27me3 intersection peak
682 regions. Example of shared peak regions for *Fgf9* (**C**) and *estrogen receptor gene* 1 (*Esr1,* **D**).
683 Regions that intersect between the two peak files are boxed.

684

685

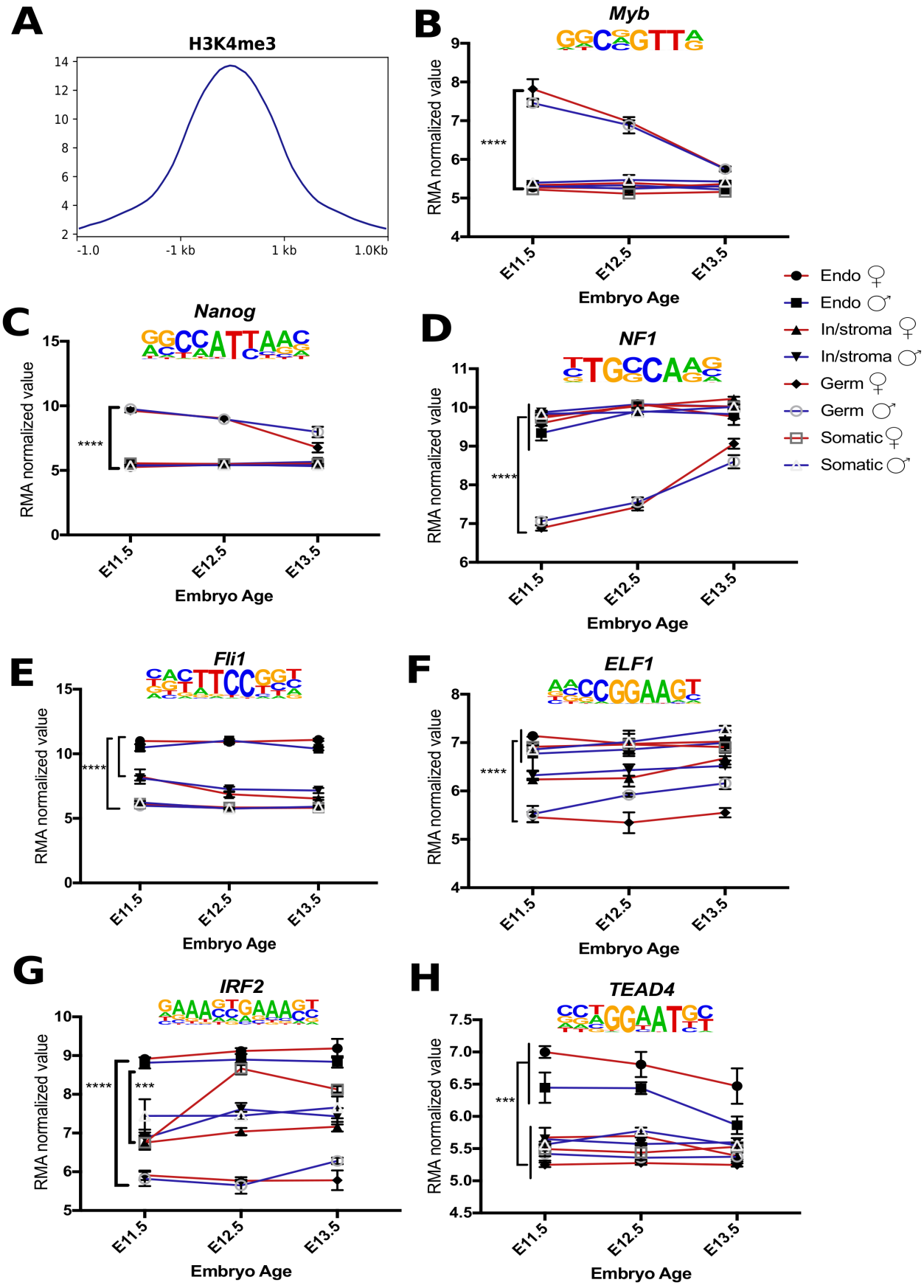


686
687

688 **Figure 8. DNA motifs enriched within H3K4me3 HOMER peak regions.**

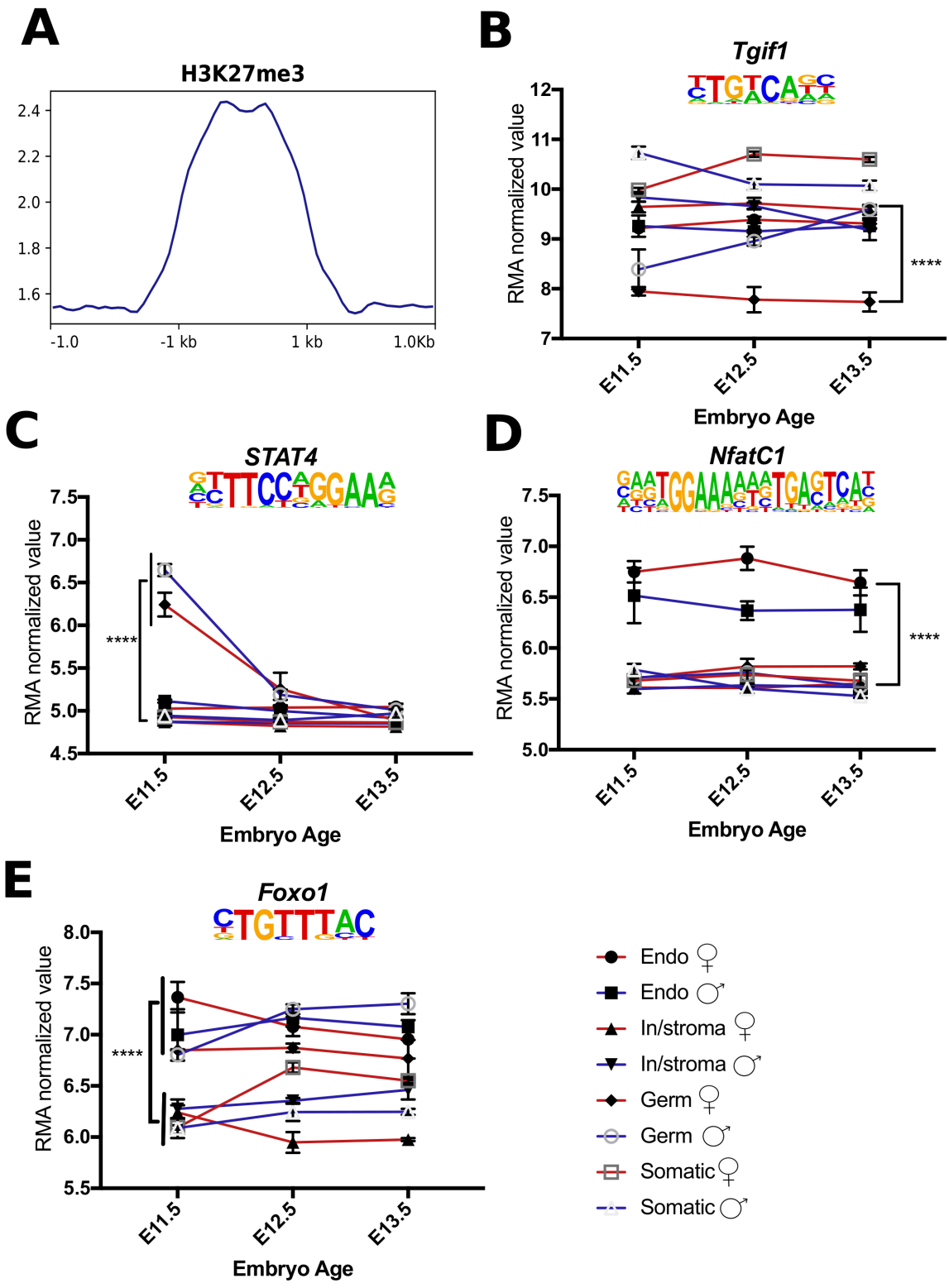
689 **(A**) Signal profile for H3K3me3 across the genome regions (3 kb), centered on the middle of
690 the peak interval. Enriched DNA binding motifs and their cell-lineage expression profiles for
691 germ, interstitial, supporting and endothelial cells at E11.5, E12.5 and E13.5 (Jameson et al.,
692 2012). Profiles are shown for *Myb* (**B**), *Nanog* (**C**), *Nf1* (**D**), *Fli1* (**E**), *Elf1* (**F**), *IRF2* (**G**) and
693 *TEAD4* (**H**) genes. Mean +/- standard error of the mean. Statistical analysis: two-way ANOVA
694 Tukey's multiple comparisons test, adjusted P-value <0.001 (**), <0.0001 (****).

695
696



697

698 **Figure 9**. **DNA motifs overrepresented within 1kb peak regions enriched with the**
699 **H3K27me3 modification. (A)** Signal profile for H3K27me3 across the genome regions (3 kb),
700 centered on the middle of the peak interval.
701 Enriched DNA binding motifs and their cell-lineage normalized expression profiles from E11.5
702 to E13.5 of gonad development (Jameson et al., 2012). Profiles are shown for *Tgif1* (**B**), *STAT4*
703 (**C**), *NfatC1* (**D**) and *Foxo1* (**E**) genes. Mean +/- standard error of the mean. Statistical analysis:
704 two-way ANOVA Tukey's multiple comparisons test, adjusted P-value <0.001 (**), <0.0001
705 (****).
706



707
708

Acknowledgements

Research described here was funded by a University of Otago Research Grant to Megan Wilson. Yisheng Yang was supported by a University of Otago PhD scholarship. We thank James Smith, Susie Szakats, Kathy Sircombe, Jeremy McCallum-Loudaec, Rebecca Clarke and Lyvianne Decouryte for feedback on manuscript drafts.

715
716

717 **Supplementary Data files**

718 **File 1. Sequencing data mapping statistics.**

719

720 **File 2. List of regions with similar signal distribution identified by**
721 **deepTools/plotHeatmap using k-means clustering analysis.** Output file giving region
722 coordinates, nearby feature/gene and cluster group. Sheet1: H3K27me3 signal clusters.
723 Sheet 2: H3K4me3 signal clusters.

724

725 **File 3. PANTHER Gene ontology analysis for H3K4me3 clusters.** Each sheet summarizes
726 the PANTHER analysis carried out for the cluster gene list. Each gene list includes genes
727 associated with each signal profile (cluster).

728

729 **File 4. PANTHER Gene ontology analysis for H3K27me3 clusters.** Each sheet summarizes
730 the PANTHER analysis carried out for the cluster gene list. Each gene list includes genes
731 associated with each signal profile (cluster).

732

733 **File 5. MACS broadcall peak regions.** Excel file with the peak regions identified for each
734 histone mark and a list of peak regions that are shared between both marks ("overlapping
735 regions"). Sheet 1: MACS broadcall peaks for H3K4me3. Sheet2: MACS broadcall peak regions
736 for H3K27me3. Sheet 3: Overlapping peak regions. Sheet 4: DAVID annotation analysis
737 (summarized in Fig. 7B). Sheet 5: Genes located near shared peak regions (identified using
738 GREAT).

739

740 **File 6. Motif analysis results for narrow peak regions.** 1 kbp peak regions identified by
741 HOMER analysis for H3K27me3 and H3K4me3. Motifs Homer analysis for known DNA

742 binding motifs. Grouped into motifs of factors expressed in the early gonad (Jameson et al.,
743 2012) and those factors not present in array data by Jameson *et al*. 2012.
744
745
746 Supplementary figures are provided in a single pdf file.

References

Aigueperse, C., Val, P., Pacot, C., Darne, C., Lalli, E., Sassone-Corsi, P., Veyssiere, G., Jean, C., Martinez, A., 2001. SF-1 (steroidogenic factor-1), C/EBPbeta (CCAAT/enhancer binding protein), and ubiquitous transcription factors NF1 (nuclear factor 1) and Sp1 (selective promoter factor 1) are required for regulation of the mouse aldose reductase-like gene (AKR1B7) expression in adrenocortical cells. Mol Endocrinol 15, 93-111.

Akiyama, H., Chaboissier, M.C., Behringer, R.R., Rowitch, D.H., Schedl, A., Epstein, J.A., de Crombrugghe, B., 2004. Essential role of Sox9 in the pathway that controls formation of cardiac valves and septa. Proc Natl Acad Sci U S A 101, 6502-6507.

Anderson, R., Fassler, R., Georges-Labouesse, E., Hynes, R.O., Bader, B.L., Kreidberg, J.A., Schaible, K., Heasman, J., Wylie, C., 1999. Mouse primordial germ cells lacking beta1 integrins enter the germline but fail to migrate normally to the gonads. Development 126, 1655-1664.

Atlasi, Y., Stunnenberg, H.G., 2017. The interplay of epigenetic marks during stem cell differentiation and development. Nat Rev Genet 18, 643-658.

Bannister, A.J., Kouzarides, T., 2011. Regulation of chromatin by histone modifications. Cell Res 21, 381-395.

Bartholin, L., Powers, S.E., Melhuish, T.A., Lasse, S., Weinstein, M., Wotton, D., 2006. TGIF inhibits retinoid signaling. Mol Cell Biol 26, 990-1001.

Bernard, P., Harley, V.R., 2007. Wnt4 action in gonadal development and sex determination. Int J Biochem Cell Biol 39, 31-43.

Bernard, P., Ryan, J., Sim, H., Czech, D.P., Sinclair, A.H., Koopman, P., Harley, V.R., 2012. Wnt signaling in ovarian development inhibits Sf1 activation of Sox9 via the Tesco enhancer. Endocrinology 153, 901-912.

Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., Jaenisch, R., Wagschal, A., Feil, R., Schreiber, S.L., Lander, E.S., 2006. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell 125, 315-326.

Birk, O.S., Casiano, D.E., Wassif, C.A., Cogliati, T., Zhao, L., Zhao, Y., Grinberg, A., Huang, S., Kreidberg, J.A., Parker, K.L., Porter, F.D., Westphal, H., 2000. The LIM homeobox gene Lhx9 is essential for mouse gonad formation. Nature 403, 909-913.

Bowles, J., Knight, D., Smith, C., Wilhelm, D., Richman, J., Mamiya, S., Yashiro, K., Chawengsaksophak, K., Wilson, M.J., Rossant, J., Hamada, H., Koopman, P., 2006. Retinoid signaling determines germ cell fate in mice. Science 312, 596-600.

Boyer, A., Lapointe, E., Zheng, X., Cowan, R.G., Li, H., Quirk, S.M., DeMayo, F.J., Richards, J.S., Boerboom, D., 2010. WNT4 is required for normal ovarian follicle development and female fertility. Faseb J 24, 3010-3025.

Butler, J.S., Koutelou, E., Schibler, A.C., Dent, S.Y., 2012. Histone-modifying enzymes: regulators of developmental decisions and drivers of human disease. Epigenomics 4, 163-177.

Capel, B., 2017. Vertebrate sex determination: evolutionary plasticity of a fundamental switch. Nat Rev Genet 18, 675-689.

Carre, G.A., Couty, I., Hennequet-Antier, C., Govoroun, M.S., 2011. Gene expression profiling reveals new potential players of gonad differentiation in the chicken embryo. PLoS One 6, e23959.

Chassot, A.A., Bradford, S.T., Auguste, A., Gregoire, E.P., Pailhoux, E., de Rooij, D.G., Schedl, A., Chaboissier, M.C., 2012. WNT4 and RSPO1 together are required for cell proliferation in the early mouse gonad. Development 139, 4461-4472.

795 Chassot, A.A., Ranc, F., Gregoire, E.P., Roepers-Gajadien, H.L., Taketo, M.M., Camerino, G., de
796 Rooij, D.G., Schedl, A., Chaboissier, M.C., 2008. Activation of beta-catenin signaling by Rspo1
797 controls differentiation of the mammalian ovary. Hum Mol Genet 17, 1264-1277.
798 Chawengsaksophak, K., Svingen, T., Ng, E.T., Epp, T., Spiller, C.M., Clark, C., Cooper, H.,
799 Koopman, P., 2012. Loss of Wnt5a disrupts primordial germ cell migration and male sexual
800 development in mice. Biology of reproduction 86, 1-12.
801 Ciau-Uitz, A., Wang, L., Patient, R., Liu, F., 2013. ETS transcription factors in hematopoietic
802 stem cell development. Blood Cells Mol Dis 51, 248-255.
803 Colvin, J.S., Green, R.P., Schmahl, J., Capel, B., Ornitz, D.M., 2001. Male-to-female sex reversal in
804 mice lacking fibroblast growth factor 9. Cell 104, 875-889.
805 Combes, A.N., Wilhelm, D., Davidson, T., Dejana, E., Harley, V., Sinclair, A., Koopman, P., 2009.
806 Endothelial cell migration directs testis cord formation. Dev Biol 326, 112-120.
807 Coveney, D., Cool, J., Oliver, T., Capel, B., 2008. Four-dimensional analysis of vascularization
808 during primary development of an organ, the gonad. Proc Natl Acad Sci U S A 105, 7212-7217.
809 Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J.,
810 Lodato, M.A., Frampton, G.M., Sharp, P.A., Boyer, L.A., Young, R.A., Jaenisch, R., 2010. Histone
811 H3K27ac separates active from poised enhancers and predicts developmental state. Proc Natl
812 Acad Sci U S A 107, 21931-21936.
813 Dallosso, A.R., Hancock, A.L., Malik, S., Salpekar, A., King-Underwood, L., Pritchard-Jones, K.,
814 Peters, J., Moorwood, K., Ward, A., Malik, K.T., Brown, K.W., 2007. Alternately spliced WT1
815 antisense transcripts interact with WT1 sense RNA and show epigenetic and splicing defects
816 in cancer. RNA 13, 2287-2299.
817 Feng, J., Liu, T., Qin, B., Zhang, Y., Liu, X.S., 2012. Identifying ChIP-seq enrichment using MACS.
818 Nat Protoc 7, 1728-1740.
819 Fleischman, R.A., 1993. From white spots to stem cells: the role of the Kit receptor in
820 mammalian development. Trends in genetics : TIG 9, 285-290.
821 Garcia, T.X., Farmaha, J.K., Kow, S., Hofmann, M.C., 2014. RBPJ in mouse Sertoli cells is
822 required for proper regulation of the testis stem cell niche. Development 141, 4468-4478.
823 Gaspar, J., Thai, S., Voland, C., Dube, A., Libermann, T.A., Iruela-Arispe, M.L., Oettgen, P., 2002.
824 Opposing functions of the Ets factors NERF and ELF-1 during chicken blood vessel
825 development. Arterioscler Thromb Vasc Biol 22, 1106-1112.
826 Gonen, N., Quinn, A., O'Neill, H.C., Koopman, P., Lovell-Badge, R., 2017. Normal Levels of Sox9
827 Expression in the Developing Mouse Testis Depend on the TES/TESCO Enhancer, but This
828 Does Not Act Alone. PLoS Genet 13, e1006520.
829 Guenther, M.G., Levine, S.S., Boyer, L.A., Jaenisch, R., Young, R.A., 2007. A chromatin landmark
830 and transcription initiation at most promoters in human cells. Cell 130, 77-88.
831 Gustin, S.E., Stringer, J.M., Hogg, K., Sinclair, A.H., Western, P.S., 2016. FGF9, activin and
832 TGFbeta promote testicular characteristics in an XX gonad organ culture model. Reproduction
833 152, 529-543.
834 Hammes, A., Guo, J.K., Lutsch, G., Leheste, J.R., Landrock, D., Ziegler, U., Gubler, M.C., Schedl, A.,
835 2001. Two splice variants of the Wilms' tumor 1 gene have distinct functions during sex
836 determination and nephron formation. Cell 106, 319-329.
837 Hawkins, R.D., Hon, G.C., Lee, L.K., Ngo, Q., Lister, R., Pelizzola, M., Edsall, L.E., Kuan, S., Luu, Y.,
838 Klugman, S., Antosiewicz-Bourget, J., Ye, Z., Espinoza, C., Agarwahl, S., Shen, L., Ruotti, V., Wang,
839 W., Stewart, R., Thomson, J.A., Ecker, J.R., Ren, B., 2010. Distinct epigenomic landscapes of
840 pluripotent and lineage-committed human cells. Cell Stem Cell 6, 479-491.
841 Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H.,
842 Glass, C.K., 2010. Simple combinations of lineage-determining transcription factors prime cis-
843 regulatory elements required for macrophage and B cell identities. Mol Cell 38, 576-589.

844 Herrada, G., Wolgemuth, D.J., 1997. The mouse transcription factor Stat4 is expressed in
845 haploid male germ cells and is present in the perinuclear theca of spermatozoa. J Cell Sci 110
846 (Pt 14), 1543-1553.
847 Huang, X., Brown, C., Ni, W., Maynard, E., Rigby, A.C., Oettgen, P., 2006. Critical role for the Ets
848 transcription factor ELF-1 in the development of tumor angiogenesis. Blood 107, 3153-3160.
849 Ishaq, M., Schang, A.L., Magre, S., Laverriere, J.N., Guillou, A., Coudouel, N., Wargnier, R., Cohen-
850 Tannoudji, J., Counis, R., 2013. Rat Gnrhr promoter directs species-specific gene expression in
851 the pituitary and testes of transgenic mice. J Mol Endocrinol 50, 411-426.
852 Jameson, S.A., Natarajan, A., Cool, J., DeFalco, T., Maatouk, D.M., Mork, L., Munger, S.C., Capel, B.,
853 2012. Temporal transcriptional profiling of somatic and germ cells reveals biased lineage
854 priming of sexual fate in the fetal mouse gonad. PLoS Genet 8, e1002575.
855 Josso, N., di Clemente, N., 1999. TGF-beta Family Members and Gonadal Development. Trends
856 in endocrinology and metabolism: TEM 10, 216-222.
857 Julio-Pieper, M., Flor, P.J., Dinan, T.G., Cryan, J.F., 2011. Exciting times beyond the brain:
858 metabotropic glutamate receptors in peripheral and non-neural tissues. Pharmacol Rev 63,
859 35-58.
860 Karl, J., Capel, B., 1998. Sertoli cells of the mouse testis originate from the coelomic epithelium.
861 Dev Biol 203, 323-333.
862 Katoh-Fukui, Y., Miyabayashi, K., Komatsu, T., Owaki, A., Baba, T., Shima, Y., Kidokoro, T., Kanai,
863 Y., Schedl, A., Wilhelm, D., Koopman, P., Okuno, Y., Morohashi, K., 2012. Cbx2, a polycomb
864 group gene, is required for Sry gene expression in mice. Endocrinology 153, 913-924.
865 Katoh-Fukui, Y., Owaki, A., Toyama, Y., Kusaka, M., Shinohara, Y., Maekawa, M., Toshimori, K.,
866 Morohashi, K., 2005. Mouse Polycomb M33 is required for splenic vascular and adrenal gland
867 formation through regulating Ad4BP/SF1 expression. Blood 106, 1612-1620.
868 Kent, J., Wheatley, S.C., Andrews, J.E., Sinclair, A.H., Koopman, P., 1996. A male-specific role for
869 SOX9 in vertebrate sex determination. Development 122, 2813-2822.
870 Kim, Y., Kobayashi, A., Sekido, R., DiNapoli, L., Brennan, J., Chaboissier, M.C., Poulat, F.,
871 Behringer, R.R., Lovell-Badge, R., Capel, B., 2006. Fgf9 and Wnt4 act as antagonistic signals to
872 regulate mammalian sex determination. PLoS biology 4, e187.
873 Kulkarni, R.M., Greenberg, J.M., Akeson, A.L., 2009. NFATc1 regulates lymphatic endothelial
874 development. Mech Dev 126, 350-365.
875 Kuroki, S., Okashita, N., Baba, S., Maeda, R., Miyawaki, S., Yano, M., Yamaguchi, M., Kitano, S.,
876 Miyachi, H., Itoh, A., Yoshida, M., Tachibana, M., 2017. Rescuing the aberrant sex development
877 of H3K9 demethylase Jmjd1a-deficient mice by modulating H3K9 methylation balance. PLoS
878 Genet 13, e1007034.
879 Lee, B.K., Shen, W., Lee, J., Rhee, C., Chung, H., Kim, K.Y., Park, I.H., Kim, J., 2015. Tgif1
880 Counterbalances the Activity of Core Pluripotency Factors in Mouse Embryonic Stem Cells.
881 Cell Rep 13, 52-60.
882 Liu, X., Zhao, D., James, L., Li, J., Zeng, H., 2011. Requirement of the nuclear localization of
883 transcription enhancer factor 3 for proliferation, migration, tube formation, and angiogenesis
884 induced by vascular endothelial growth factor. Faseb J 25, 1188-1197.
885 Liu, Z., Castrillon, D.H., Zhou, W., Richards, J.S., 2013. FOXO1/3 depletion in granulosa cells
886 alters follicle growth, death and regulation of pituitary FSH. Mol Endocrinol 27, 238-252.
887 Luo, X., Ikeda, Y., Parker, K.L., 1994. A cell-specific nuclear receptor is essential for adrenal
888 and gonadal development and sexual differentiation. Cell 77, 481-490.
889 Maatouk, D.M., Natarajan, A., Shibata, Y., Song, L., Crawford, G.E., Ohler, U., Capel, B., 2017.
890 Genome-wide identification of regulatory elements in Sertoli cells. Development 144, 720-730.

Marciniak, M., Chruscicka, B., Lech, T., Burnat, G., Pilc, A., 2016. Expression of group III metabotropic glutamate receptors in the reproductive system of male mice. Reproduction, fertility, and development 28, 369-374.
Matson, C.K., Murphy, M.W., Sarver, A.L., Griswold, M.D., Bardwell, V.J., Zarkower, D., 2011. DMRT1 prevents female reprogramming in the postnatal mammalian testis. Nature 476, 101-104.
McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., Bejerano, G., 2010. GREAT improves functional interpretation of cis-regulatory regions. Nat Biotechnol 28, 495-501.
Meng, Z.Z., Liu, W., Xia, Y., Yin, H.M., Zhang, C.Y., Su, D., Yan, L.F., Gu, A.H., Zhou, Y., 2017. The pro-inflammatory signalling regulator Stat4 promotes vasculogenesis of great vessels derived from endothelial precursors. Nat Commun 8, 14640.
Mericskay, M., Kitajewski, J., Sassoon, D., 2004. Wnt5a is required for proper epithelial-mesenchymal interactions in the uterus. Development 131, 2061-2072.
Messina, A., Ferraris, N., Wray, S., Cagnoni, G., Donohue, D.E., Casoni, F., Kramer, P.R., Derijck, A.A., Adolfs, Y., Fasolo, A., Pasterkamp, R.J., Giacobini, P., 2011. Dysregulation of Semaphorin7A/beta1-integrin signaling leads to defective GnRH-1 cell migration, abnormal gonadal development and altered fertility. Hum Mol Genet 20, 4759-4774.
Mi, H., Dong, Q., Muruganujan, A., Gaudet, P., Lewis, S., Thomas, P.D., 2010. PANTHER version 7: improved phylogenetic trees, orthologs and collaboration with the Gene Ontology Consortium. Nucleic Acids Res 38, D204-210.
Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.K., Koche, R.P., Lee, W., Mendenhall, E., O'Donovan, A., Presser, A., Russ, C., Xie, X., Meissner, A., Wernig, M., Jaenisch, R., Nusbaum, C., Lander, E.S., Bernstein, B.E., 2007. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature 448, 553-560.
Mithraprabhu, S., Loveland, K.L., 2009. Control of KIT signalling in male germ cells: what can we learn from other systems? Reproduction 138, 743-757.
Molyneaux, K.A., Stallock, J., Schaible, K., Wylie, C., 2001. Time-lapse analysis of living mouse germ cell migration. Dev Biol 240, 488-498.
Mork, L., Maatouk, D.M., McMahon, J.A., Guo, J.J., Zhang, P., McMahon, A.P., Capel, B., 2012. Temporal differences in granulosa cell specification in the ovary reflect distinct follicle fates in mice. Biology of reproduction 86, 37.
Mukouyama, Y., Chiba, N., Mucenski, M.L., Satake, M., Miyajima, A., Hara, T., Watanabe, T., 1999. Hematopoietic cells in cultures of the murine embryonic aorta-gonad-mesonephros region are induced by c-Myb. Current biology : CB 9, 833-836.
Munsterberg, A., Lovell-Badge, R., 1991. Expression of the mouse anti-mullerian hormone gene suggests a role in both male and female sexual differentiation. Development 113, 613-624.
Nakata, Y., Shetzline, S., Sakashita, C., Kalota, A., Rallapalli, R., Rudnick, S.I., Zhang, Y., Emerson, S.G., Gewirtz, A.M., 2007. c-Myb contributes to G2/M cell cycle transition in human hematopoietic cells by direct regulation of cyclin B1 expression. Mol Cell Biol 27, 2048-2058.
Ng, J.H., Kumar, V., Muratani, M., Kraus, P., Yeo, J.C., Yaw, L.P., Xue, K., Lufkin, T., Prabhakar, S., Ng, H.H., 2013. In vivo epigenomic profiling of germ cells reveals germ cell molecular signatures. Dev Cell 24, 324-333.
Ottolenghi, C., Pelosi, E., Tran, J., Colombino, M., Douglass, E., Nedorezov, T., Cao, A., Forabosco, A., Schlessinger, D., 2007. Loss of Wnt4 and Foxl2 leads to female-to-male sex reversal extending to germ cells. Hum Mol Genet 16, 2795-2804.

939 Pannetier, M., Chassot, A.A., Chaboissier, M.C., Pailhoux, E., 2016. Involvement of FOXL2 and
940 RSPO1 in Ovarian Determination, Development, and Maintenance in Mammals. Sex Dev 10,
941 167-184.
942 Parker, J., Klein, S.L., McClintock, M.K., Morison, W.L., Ye, X., Conti, C.J., Peterson, N., Nousari,
943 C.H., Tausk, F.A., 2004. Chronic stress accelerates ultraviolet-induced cutaneous
944 carcinogenesis. J Am Acad Dermatol 51, 919-922.
945 Ramirez, F., Dundar, F., Diehl, S., Gruning, B.A., Manke, T., 2014. deepTools: a flexible platform
946 for exploring deep-sequencing data. Nucleic Acids Res 42, W187-191.
947 Remy, P., Baltzinger, M., 2000. The Ets-transcription factor family in embryonic development:
948 lessons from the amphibian and bird. Oncogene 19, 6417-6431.
949 Sachs, M., Onodera, C., Blaschke, K., Ebata, K.T., Song, J.S., Ramalho-Santos, M., 2013. Bivalent
950 chromatin marks developmental regulatory genes in the mouse embryonic germline in vivo.
951 Cell Rep 3, 1777-1784.
952 Sakai, N., Terami, H., Suzuki, S., Haga, M., Nomoto, K., Tsuchida, N., Morohashi, K., Saito, N.,
953 Asada, M., Hashimoto, M., Harada, D., Asahara, H., Ishikawa, T., Shimada, F., Sakurada, K., 2008.
954 Identification of NR5A1 (SF-1/AD4BP) gene expression modulators by large-scale gain and
955 loss of function studies. J Endocrinol 198, 489-497.
956 Santos-Rosa, H., Schneider, R., Bannister, A.J., Sherriff, J., Bernstein, B.E., Emre, N.C., Schreiber,
957 S.L., Mellor, J., Kouzarides, T., 2002. Active genes are tri-methylated at K4 of histone H3.
958 Nature 419, 407-411.
959 Schang, A.L., Querat, B., Simon, V., Garrel, G., Bleux, C., Counis, R., Cohen-Tannoudji, J.,
960 Laverriere, J.N., 2012. Mechanisms underlying the tissue-specific and regulated activity of the
961 Gnrhr promoter in mammals. Front Endocrinol (Lausanne) 3, 162.
962 Schmidt, D., Ovitt, C.E., Anlag, K., Fehsenfeld, S., Gredsted, L., Treier, A.C., Treier, M., 2004. The
963 murine winged-helix transcription factor Foxl2 is required for granulosa cell differentiation
964 and ovary maintenance. Development 131, 933-942.
965 Sekido, R., Lovell-Badge, R., 2008. Sex determination involves synergistic action of SRY and
966 SF1 on a specific Sox9 enhancer. Nature 453, 930-934.
967 Shin, H., Liu, T., Manrai, A.K., Liu, X.S., 2009. CEAS: cis-regulatory element annotation system.
968 Bioinformatics 25, 2605-2606.
969 Spiller, C., Koopman, P., Bowles, J., 2017. Sex Determination in the Mammalian Germline. Annu
970 Rev Genet.
971 Supek, F., Bosnjak, M., Skunca, N., Smuc, T., 2011. REVIGO summarizes and visualizes long lists
972 of gene ontology terms. PLoS One 6, e21800.
973 Tanaka, S.S., Nishinakamura, R., 2014. Regulation of male sex determination: genital ridge
974 formation and Sry activation in mice. Cell Mol Life Sci 71, 4781-4802.
975 Theunissen, T.W., van Oosten, A.L., Castelo-Branco, G., Hall, J., Smith, A., Silva, J.C., 2011. Nanog
976 overcomes reprogramming barriers and induces pluripotency in minimal conditions. Current
977 biology : CB 21, 65-71.
978 Uhlenhaut, N.H., Jakob, S., Anlag, K., Eisenberger, T., Sekido, R., Kress, J., Treier, A.C., Klugmann,
979 C., Klasen, C., Holter, N.I., Riethmacher, D., Schutz, G., Cooney, A.J., Lovell-Badge, R., Treier, M.,
980 2009. Somatic sex reprogramming of adult ovaries to testes by FOXL2 ablation. Cell 139,
981 1130-1142.
982 Vainio, S., Heikkila, M., Kispert, A., Chin, N., McMahon, A.P., 1999. Female development in
983 mammals is regulated by Wnt-4 signalling. Nature 397, 405-409.
984 Vilain, E., 2011. The genetics of ovotesticular disorders of sex development. Adv Exp Med Biol
985 707, 105-106.
986 Voigt, P., Tee, W.W., Reinberg, D., 2013. A double take on bivalent promoters. Genes Dev 27,
987 1318-1338.

988 Wen, S., Ai, W., Alim, Z., Boehm, U., 2010. Embryonic gonadotropin-releasing hormone
989 signaling is necessary for maturation of the male reproductive axis. Proc Natl Acad Sci U S A
990 107, 16372-16377.
991 Wilhelm, D., Hiramatsu, R., Mizusaki, H., Widjaja, L., Combes, A.N., Kanai, Y., Koopman, P., 2007.
992 SOX9 regulates prostaglandin D synthase gene transcription in vivo to ensure testis
993 development. J Biol Chem 282, 10553-10560.
994 Wilhelm, D., Huang, E., Svingen, T., Stanfield, S., Dinnis, D., Koopman, P., 2006. Comparative
995 proteomic analysis to study molecular events during gonad development in mice. Genesis 44,
996 168-176.
997 Wilson, M.J., Jeyasuria, P., Parker, K.L., Koopman, P., 2005. The transcription factors
998 steroidogenic factor-1 and SOX9 regulate expression of Vanin-1 during mouse testis
999 development. J Biol Chem 280, 5917-5923.
1000 Wotton, D., Lo, R.S., Lee, S., Massague, J., 1999. A Smad transcriptional corepressor. Cell 97, 29-
1001 39.
1002 Wu, B., Wang, Y., Lui, W., Langworthy, M., Tompkins, K.L., Hatzopoulos, A.K., Baldwin, H.S.,
1003 Zhou, B., 2011. Nfatc1 coordinates valve endocardial cell lineage development required for
1004 heart valve formation. Circ Res 109, 183-192.
1005 Xu, X., Coats, J.K., Yang, C.F., Wang, A., Ahmed, O.M., Alvarado, M., Izumi, T., Shah, N.M., 2012.
1006 Modular genetic control of sexually dimorphic behaviors. Cell 148, 596-607.
1007 Yamaguchi, S., Kurimoto, K., Yabuta, Y., Sasaki, H., Nakatsuji, N., Saitou, M., Tada, T., 2009.
1008 Conditional knockdown of Nanog induces apoptotic cell death in mouse migrating primordial
1009 germ cells. Development 136, 4011-4020.
1010 Yamamizu, K., Piao, Y., Sharov, A.A., Zsiros, V., Yu, H., Nakazawa, K., Schlessinger, D., Ko, M.S.,
1011 2013. Identification of transcription factors for lineage-specific ESC differentiation. Stem Cell
1012 Reports 1, 545-559.
1013 Yang, Y., Wilson, M.J., 2015. Lhx9 gene expression during early limb development in mice
1014 requires the FGF signalling pathway. Gene Expr Patterns 19, 45-51.
1015 Young, M.D., Willson, T.A., Wakefield, M.J., Trounson, E., Hilton, D.J., Blewitt, M.E., Oshlack, A.,
1016 Majewski, I.J., 2011. ChIP-seq analysis reveals distinct H3K27me3 profiles that correlate with
1017 transcriptional activity. Nucleic Acids Res 39, 7415-7427.
1018 Zhang, X., Yalcin, S., Lee, D.F., Yeh, T.Y., Lee, S.M., Su, J., Mungamuri, S.K., Rimmele, P., Kennedy,
1019 M., Sellers, R., Landthaler, M., Tuschl, T., Chi, N.W., Lemischka, I., Keller, G., Ghaffari, S., 2011.
1020 FOXO1 is an essential regulator of pluripotency in human embryonic stem cells. Nat Cell Biol
1021 13, 1092-1099.
1022 Zheng, L.P., Huang, J., Zhang, D.L., Xu, L.Q., Li, F., Wu, L., Liu, Z.Y., Zheng, Y.H., 2012. c-erbB2 and
1023 c-myb induce mouse oocyte maturation involving activation of maturation promoting factor.
1024 DNA Cell Biol 31, 164-170.
1025