Specific transcriptomic signatures and dual regulation of steroidogenesis between fetal and adult mouse Leydig cells.

Pauline Sararols[1*], Isabelle Stévant[1*], Yasmine Neirijnck[1], Diane Rebourcet[2], Annalucia Darbey[2], Michael K. Curley[3], Françoise Kühne[1], Emmanouil Dermitzakis[1,4], Lee B. Smith[2,3], Serge Nef[1,#]

Affiliations:

[1] Department of Genetic Medicine and Development, Faculty of Medicine, University of Geneva, CH-1211 Geneva, Switzerland;

[2] College of Engineering, Science and Environment, University of Newcastle, Callaghan, New South Wales 2308, Australia.

[3] Medical Research Council Centre for Reproductive Health, The Queen's Medical Research Institute, University of Edinburgh, Edinburgh EH16 4TJ, United Kingdom.

[4] Institute of Genetics and Genomics of Geneva (iGE3), Faculty of Medicine, Geneva, CH-1211 Geneva, Switzerland

[*] These authors share first authorship

[#]Correspondence to: Serge.Nef@unige.ch

Keywords: fetal Leydig cell, adult Leydig cell, androgen, testis, RNA sequencing, single cell RNA sequencing

Disclosure statement: The authors declare no conflicts of interest.

Funding sources: This work was supported by the Swiss National Science Foundation (Grant 31003A_152636 to S.N.); and the Département de l'Instruction Publique of the State of Geneva (to S.N.). Department of Health | National Health and Medical Research Council (NHMRC), Grant/Award Number: APP1158344 (to L.S. & D.R.).

Abstract

Leydig cells (LC) are the main testicular androgen-producing cells. In eutherian mammals, two types of LCs emerge successively during testicular development, fetal Leydig cells (FLCs) and adult Leydig cells (ALCs). Both display significant differences in androgen production and regulation. Using bulk RNA sequencing, we compared the transcriptomes of both LC populations to characterise their specific transcriptional and functional features. Despite similar transcriptomic profiles, a quarter of the genes show significant variations in expression between FLCs and ALCs. Non-transcriptional events, such as alternative splicing was also observed, including a high rate of intron retention in FLCs compared to ALCs. The use of single-cell RNA sequencing data also allowed the identification of nine FLC-specific genes and 50 ALC-specific genes. Expression of the corticotropin-releasing hormone 1 (*Crhr1*) receptor and the ACTH receptor melanocortin type 2 receptor (*Mc2r*) specifically in FLCs suggests a dual regulation of steroidogenesis. The androstenedione synthesis by FLCs is stimulated by luteinizing hormone (LH), CRH and ACTH whereas the testosterone synthesis by ALCs is dependent exclusively on LH. Overall, our study provides a useful database to explore LC development and function.

Introduction

Leydig cells (LC) are the main steroidogenic cells of the testes. They synthesise androgens that are essential for both masculinisation of the organism and spermatogenesis. In mice, two populations of Leydig cells arise sequentially, one during embryonic development referred as the fetal Leydig cells (FLCs) and the other postnatally referred as the adult Leydig cells (ALCs) (Baker et al., 1999; Habert et al., 2001; O'Shaughnessy et al., 2002; Haider, 2004; Chen et al., 2009). The mouse FLCs appear in the interstitial compartment of the testis shortly after sex determination at embryonic day (E)12,5. The FLC population expands considerably during fetal testis development through the recruitment and differentiation of Leydig progenitor cells rather than by mitotic division of differentiated FLCs (Byskov, 1986; Kerr et al., 1988; Migrenne et al., 2001; Brennan et al., 2003; Barsoum and Yao, 2010; Ademi et al., 2020). The maximum number of FLCs is reached around birth and regresses over the first two weeks of postnatal life (Kerr et al., 1988; Shima, 2019). The ALCs appear around one week after birth and increase in number during puberty. They arise from LC progenitors located in the testicular interstitium (Davidoff et al., 2004; Barsoum et al., 2013; Shima et al., 2013; Kilcoyne et al., 2014; Ademi et al., 2020). Two recent studies showed that both fetal and adult Leydig cells derive from a common pool of progenitor cells originating from the gonadal surface epithelium and mesonephric mesenchymal cells present from fetal life (Ademi et al., 2020; Shen et al., 2020). Evidence also shows that a subset of FLCs dedifferentiate at fetal stages to serve as potential ALC stem cells (Shima et al., 2018).

The rodent FLCs and ALCs have distinct morphological and functional differences. The FLCs display a high proportion of lipid droplets, while mostly absent in the ALCs (Huhtaniemi and Pelliniemi, 1992; Shima, 2019). Unlike ALCs, the FLCs are not capable of fully synthesising testosterone on their own. They express all the enzymes necessary for androgen synthesis except HSD17B3, which converts androstenedione to testosterone. The conversion of androstenedione produced by the FLCs is achieved by the adjacent fetal Sertoli cells that express HSD17B3 (O'Shaughnessy et al., 2000; Shima et al., 2013). Another notable difference between fetal and adult LCs is their regulation by the pituitary gonadotropins. Although the luteinizing hormone (LH) receptor is expressed from E16.5 in FLCs and later in ALCs (O'Shaughnessy et al., 1998)), LH signalling is dispensable for FLCs development, but prove to be essential for ALCs development and testosterone production. Neonatal mouse mutants for LH/CG receptors display testes indistinguishable from control mice. In contrast, testes from adult mutants for LH/CG receptors are

reduced in size, with fewer and hypoplastic ALCs, and show impaired testosterone production (Lei et al., 2001; Zhang et al., 2001; O'Shaughnessy and Fowler, 2011; Teerds and Huhtaniemi, 2015). FLC function is normal in the absence of endogenous circulating gonadotropins (O'Shaughnessy et al., 1998) but markedly reduced in late gestation in *T/ebp/Nkx2.1* null mice lacking a pituitary gland (Pakarinen et al., 2002). This suggests that additional hypothalamo/pituitary hormones, other than LH, may be required for FLC function and androgen production. Interestingly, two additional hormones have been reported to stimulate testosterone production in fetal testis. Adrenocorticotropic hormone (ACTH) has been reported to stimulate *in vitro* testosterone production in fetal and neonatal testes (O'Shaughnessy et al., 2003). In parallel, corticotropin-releasing hormone (CRH) has been reported to stimulate steroidogenesis by direct activation of FLCs in fetal rat and mouse testes *ex vivo* and in MA-10 mouse Leydig cells (McDowell et al., 2012), but not in primary ALCs (Huang et al., 1995; McDowell et al., 2012).

While testosterone synthesis is subjected to intensive studies, our knowledge of FLCs and ALCs origins, development, and in particular similarities and differences is still incomplete. Multiple transcriptomic studies including either mouse fetal or adult Leydig populations have been performed on whole gonads or purified cell populations in a given context (Nef et al., 2005; Beverdam and Koopman, 2006; Jameson et al., 2012; Munger et al., 2013; McClelland et al., 2015; Inoue et al., 2016; Miyabayashi et al., 2017). To date, no comprehensive comparison of FLCs and ALCs has been yet performed, and the identification of discriminant transcriptional signatures would be useful in distinguishing the two LC populations. In the present study, we employed a combination of bulk and single-cell RNA sequencing (RNA-seq and scRNA-seq) analyses to compare the transcriptome of FLCs and ALCs. While deep RNA-seq on purified Leydig cell populations allows the exploration of the transcriptomic landscape of FLCs and ALCs, including low expressed genes and alternative splicing; the high resolution of scRNA-seq allows us to bypass contamination issues inherent to cell population purification methods, and to identify specific marker genes that discriminate FLCs and ALCs amongst the other testicular cells. Our results provide a comprehensive view of the FLCs and ALCs transcriptional similarities and differences, unveiling important variations in terms of gene expression level and alternative splicing between the two populations of LCs. Furthermore, our analyses uncovered FLC- and ALC-specific markers that represent useful tools to study these two steroidogenic populations. Finally, amongst the FLCs specific markers we found the corticotropin-releasing hormone receptor 1 (*Crhr1*), suggesting

FLCs androgen synthesis is influenced by CRH.

Material & Methods

Mouse strains

Embryos were collected from timed pregnant female CD-1 outbred mice (Charles River) and heterozygous Tg(*Nr5a1-GFP*) transgenic male mice (Stallings, 2002). The mating plug observed the next morning is designated as E0.5 were used in this study. Animals were housed and cared according to the ethical guidelines of the Direction Générale de la Santé of the Canton de Genève (GE/57/18 30080).

Fetal and adult Leydig cell purification

The purification of fetal Leydig cells was carried out in seven independent experiments based on a previously described experimental protocol (Nef et al., 2005; Pitetti et al., 2013; Stévant et al., 2018). The resulting cells were pooled to achieve the amount of RNA required for the preparation of the RNA sequence library (**Supplementary Table S6**). In short, adult CD-1 females were time-mated with heterozygous Tg(*Nr5a1-GFP*) transgenic male and checked for the presence of vaginal plugs the next morning (E0.5). On the relevant days of gestation (E18.5), females were sacrificed by CO2 inhalation and the embryos collected in PBS. The sex and the presence of the *Nr5a1-GFP* transgene in the embryos were assessed under a fluorescent binocular microscope. Testes were isolated and incubated 20 minutes with trypsin-EDTA 0.05%, mechanically dissociated with gentle pipetting, and filtered through a 70µm cell strainer to obtain single cell suspension.

Adult Leydig cells purification has been performed in four independent experiments (**Supplementary Table S6**). Hundred-day-old (P100) heterozygous *Tg(Nr5a1-GFP*) transgenic male mice were used for this experiment. Mice were sacrificed with Esconarkon injection and ~1 mL of blood was collected by intracardiac puncture for serum extraction. Tunica albuginea of the testes were delicately removed and testes were incubated in DMEM supplemented with collagenase (1 mg/mL C0130; Sigma-Aldrich, St. Louis, MO), hyaluronidase (2 mg/mL H3506; Sigma-Aldrich), and DnaseI (0.8 mg/mL dN25; Sigma-Aldrich) at 37°C for 20 minutes with gentle agitation. After two rounds of seminiferous tubules sedimentation, the supernatants enriched in

interstitial cells were collected and incubated 10 minutes with Trypsine-EDTA 0.05%. Cells were centrifuged and filtered through a 70µm cell strainer to obtain single cell suspension.
Nr5a1-GFP$^+$ cells from E18.5 and P100 testes were then sorted by fluorescent-active cell sorting (BD FACS ARIA II), excluding cell doublets, and the dead cells with Draq7TM dye staining. Cells were collected directly into RLT buffer from Qiagen RNeasy Mini kit for RNA extraction.

Bulk RNA-sequencing library preparation and sequencing

RNA was extracted from Nr5a1-GFP$^+$ cells with the RNeasy Mini kit (Qiagen) to obtain a minimum of 260 ng of total RNA. The composition of the different samples is detailed in **Supplementary Table S6**. Sequencing libraries were prepared from 150 ng of DNA with the TruSeq Stranded Total RNA Library Prep Gold (Ribo-Zero) and sequenced on an Illumina HiSeq 2500 (50 bp, paired end, ~35 million reads expected) at the Genomics Platform of the University of Geneva.

RNA-sequencing analysis

Reads were demultiplexed with Casava (v1.8.2), mapped with GemTools (v1.7.1) (Marco-Sola et al., 2012) and read counts and RPKM gene expression quantifications were calculated with an in-house pipeline based on Gencode annotation GRCm38 (v4). Globally, over 88% of the reads mapped to exonic regions. Pre-ranked gene set enrichment analysis was performed using GSEA (v4.1.0) with the mean RPKM expression as ranking, and using the genes with a mean RPKM>1 (Subramanian et al., 2005). Spearman correlation and principal component analysis (PCA) using the R base stats package, revealed a very high correlation for both biological triplicates of ALCs and FLCs (Spearman correlation score >98%, see **Supplementary Figure 1C** and **D**). Similarly, the correlation between the two conditions, ALCs and FLCs, is also very high (Spearman correlation score >88%). We used the R package DESeq2 (v1.24.0) (Love et al., 2014) for the differential expression analysis. Genes with fewer than 10 reads were not taken into account. GO terms enrichment analysis was computed on the selected genes enriched in FLCs and in ALCs with the R packages ClusterProfiler (v3.12.0) (Yu et al., 2012).
Differential splicing analysis was performed using rMATS (v3.2.5) (Shen et al., 2012, 2014) with the bam files as input. Splicing event with an FDR<=0.05 were considered as significant.

Single-cell RNA-sequencing library preparation and sequencing

Single-cell library of E16.5 testis from Neirijnck et al. (Neirijnck et al., 2019) (**GSE123119**) was prepared with the Chromium Single Cell 3' Library v2 kit and sequenced targeting 5,000 cells with an Illumina Hiseq 4000 (100 bp, paired-end, 100,000 reads per cell expected) with Macrogen (http://foreign.macrogen.com/eng/).

Adult testis from adult mice (C57BL/6J mice) single-cell RNA sequencing data come from Ernst et al. (Ernst et al., 2019) (**E-MTAB-6946**). We used the libraries do15983, do15984, do17622, do17623, do17815, do17816, do18197, do18198, do18199 of mice older than 60 days old from Ernst et al. The libraries were prepared with the Single Cell 3' Library v2 kit and sequenced with an Illumina HiSeq 2500.

Data processing with the Cell Ranger package, cell selection and in-house quality controls

Computations were performed at the Vital-IT Centre for high-performance computing of the SIB (Swiss Institute of Bioinformatics) (http://www.vital-it.ch). Demultiplexing, alignment, barcode filtering and UMI counting were performed with the Cell Ranger v2.1 pipeline (10x Genomics). Reference genome has been modified to include the eGFP transgene with the mkref function. Data were mapped to the mouse reference genome GRCm38.p5 in which the *eGFP* and the bovine GH 3′ splice/polyadenylation signals (NM_180996.1) (Stallings, 2002) sequences have been added, and annotated with Gencode vM15. Only protein coding genes and long non-coding RNAs were retained for further analysis.

To set a threshold between cell containing barcodes and empty ones, we computed the knee point and the inflection point of the ranked barcode distribution plot. Then, we detected the local minimum between these points on the density curve (density base R function, bw=500, n=4,096) of the UMI number per barcode using the quantmod R package (v0.4-16). This local minimum was used as a threshold for cell selection.

Single-cell RNA-sequencing analysis

E16.5 mouse RNA-sequencing data

The single-cell RNA sequencing analysis of the E16.5 testis (Neirijnck et al., 2019) were performed using the Seurat software package (v2.3.4). From the raw matrix obtained with Cell Ranger version 2.0 (10X Genomics), we filtered cells based on the UMI count per cell (>2300 UMI) and on the

198 percentage of mitochondrial genes (>0.05% of mitochondrial genes) resulting in 3,781 cells. Then,
199 we reduced the size of the dataset using Principal Component Analysis (PCA) on the genes
200 expressed in more than 50 cells (9,576 genes) and calculated the UMAP representation (Becht et
201 al., 2019) and finally grouped the cells with Louvain algorithm (Waltman and van Eck, 2013)
202 (resolution=1) using the 15th first PCs. We classified the clusters according to the marker genes of
203 the different cell types in the literature, including the genes *Nr5a1, Star, Cyp11a1, Insl3, Hsd17b3,*
204 *Amh, Sox9, Arx, Nr2f2, Pecam1, Cdh5, Esam, Pou5f1, Ddx4* (Castrillon et al., 2000; Wakayama et
205 al., 2003; Okamura et al., 2008; Yu et al., 2009; Buaas et al., 2012; Wen et al., 2016; Kumar et al.,
206 2017; Stévant et al., 2018; Ernst et al., 2019; Mucenski et al., 2019) and identified cluster 7 as
207 Leydig cells.
208
209 *Adult mouse single-Cell RNA-Sequencing data*
210 We selected the libraries do15983, do15984, do17622, do17623, do17815, do17816, do18197,
211 do18198, do18199 of mice older than 60 days old from Ernst et al. (Ernst et al., 2019). These
212 libraries were analysed with Cell Ranger v1.3.1 software using the default threshold to obtain high-
213 quality cells with large numbers of UMIs. We filtered out cells with <500 UMI and we excluded
214 cells with more than 5% of reads mapping to the mitochondrial genome. We selected only protein-
215 coding genes. Then, we inferred cell labels with the annotation furnished by Ernst et al., and
216 removed cells labelled as “Outliers”. So, 24,672 cells were used. We confirmed the cell type
217 classification of Ernst et al. with genes known from the literature (*Cyp11a1, Star, Insl3, Cyp17a1,*
218 *and Fabp3*) (Stévant et al., 2018; Ernst et al., 2019).
219
220 *Merge of Fetal and Adult single-cell RNA Sequencing data*
221 In order to compare the Leydig cells present at fetal and adult stages, we merged the single-cell
222 RNA sequencing datasets with the MergeSeurat function (Seurat, v2.3.4). We normalised and
223 computed the Principal Component Analysis (PCA) using the genes expressed in more than 50
224 cells. The 10 first components of this PCA were used to compute the corrected neighbour graph
225 with BBKNN (balanced batch KNN, v1.3.8) (Polański et al., 2019) and then, the umap
226 representation. We made use of the previous cell annotation to distinguish fetal and adult
227 populations and we ensured the cell identity using marker gene expression. This list includes the
228 genes *Nr5a1, Star, Cyp11a1, Insl3, Hsd17b3, Amh, Sox9, Arx, Nr2f2, Pecam1, Cdh5, Esam,*

Pou5f1, *Ddx4*, *Dmrt1*, *Piwil1*, *Pex21*, and *Tnp1*. We performed a differential expression analysis (Mann-Whitney Wilcoxon test) between FLCs and ALCs clusters using Seurat FindMarkers function (only positive markers, min.pct=0.25, thresh.use=0.25). We identified genes showing a high expression in the one population of Leydig cells and a low expression in the other population (adj. p-valuej<0.01, avg logFC>0.5, pct.1>0.5, pct.2<0.25), and overexpressed as well in the same population in the DESeq2 analysis.

From this selection, in order to select the genes with specific expression in one Leydig cell population and with a low expression in all other populations in the testis, we used an additional differential expression analysis between all cell types using Seurat FindAllMarkers function (only positive markers, p_val_adj<0.01 & avg_logFC>0.5 & pct.1>0.5 & pct.2<0.25) was used to compute marker genes for every cluster. The intersect of the two differential expression analysis with Seurat is used to get a list of the marker genes of FLCs and ALCs specifically.

RNAscope® analyses (*In situ* hybridization)

Adult (P100) and embryonic (E16.5) *Nr5a1-eGFP* samples from timed mated females were collected and fixed overnight in 4% paraformaldehyde, dehydrated and embedded in paraffin. Five µm thick sections were examined histologically *via* haematoxylin and eosin staining. We performed the RNAscope®® 2.5 HD DuplexAssay protocol following the recommendation of BioTechne. The *Star* probes (C2) to label Leydig cells, and the probes for the candidates *Crhr1* (C1), *Ren1* (C1), *Bhmt* (C1) and *Sult1e1* (C1) were tested. Slides were imaged using an Axioskop 2 plus confocal microscope and ZEN 2009 software (Carl Zeiss Ltd, Hertfordshire, UK). For reproducibility purpose, at least three different animals of each group were tested.

Immunostaining

Animals were bred and maintained in strict compliance with the Animals (Scientific Procedures) Act, 1986. All procedures were conducted in accordance with United Kingdom Home Office regulations under project licenses 60/4200 and 70/8804 held by Lee B. Smith.

Neonatal and adult tissues were fixed in Bouins for 6 hours, stored in ethanol 70% and embed in paraffin. Sections of 5 µm were dewaxed in xylene, rehydrated in graded ethanol solutions. For the double immunostaining, slides were antigen-retrieved in pressure cook with 0.01M citrate buffer (pH 6.0). To quench endogenous peroxidases activity, slides were incubated in 0.3% hydrogen

peroxide (v/v) in TBS for 30 min at room temperature (RT). The non-specific activity was blocked using the appropriate normal blocking serum for 30 min at RT followed by incubation overnight at 4°C with the first primary antibody diluted in blocking serum. After washing, slides were incubated for 30 min at RT with the appropriate secondary antibody conjugated to peroxidase and diluted 1/200 in blocking serum and left on the slides for 30 min at RT. Sections were then incubated with Tyramide Signal Amplification system ('TSA™', Perkin Elmer) diluted 1/50 for 10 min at RT according to the manufacturer's instructions. Slides were then stained with the second Primary antibody and washed as an incubated as described above with secondary and Tyramide. Sections were then counterstained in Sytox Green (Molecular Probes, life technologies, Paisley, UK) for 10 min at RT and mounted in PermaFluor mounting medium (Thermo Scientific, UK). Slides were scanned using an LSM 710 confocal microscope and ZEN 2009 software (Carl Zeiss Ltd, Hertfordshire, UK). The primary and adequate secondary antibodies used in this study are detailed in **Supplementary Table S7**. To assure the specificity of the stained tissue, sections incubated with no primary antibody were used as negative controls. For reproducibility purpose, at least three different animals of each group were tested.

Data Availability

The fetal and adult Leydig cell bulk RNA-seq data are available on GEO (NCBI) under accession number **GSE171746**.

Results

A global view of fetal and adult Leydig cell transcriptomes

To compare the transcriptomic signatures of FLCs and ALCs, we purified Leydig cells at embryonic day (E) 18.5 and postnatal day (P) P100. To proceed, we FAC-sorted the highest GFP positive cell population corresponding to Leydig cells from *Nr5a1-eGFP* testes (Stallings, 2002) (**Figure 1A; Supplementary Figure 1A and A'**). We then performed RNA sequencing of poly(A)$^+$ RNAs in biological triplicates for the two stages. In total, we identified 20,859 and 21,195 expressed genes in FLCs and ALCs respectively (with RPKM≥1), with 78.5% (±0.5) of them being protein coding. We evaluated the purity of our FLCs and ALCs samples by measuring the expression level of several genes specific to different testicular cells (**Supplementary Figure 1B**). Our samples display high expression levels of the Leydig cell marker genes *Cyp11a1, Insl3,*

Cyp17a1, Hsd3b1, and Star (Rebourcet et al., 2019). In contrast, expression of marker genes for Sertoli cells *(Sox9, Dhh, Amh*) (Bitgood et al., 1996; Liu et al., 2016; Rehman et al., 2017), interstitial progenitors cells *(Nr2f2, Arx, Tcf21*) , germ cells (*Dazl, Ddx4, Pou5f1*) (Castrillon et al., 2000; Okamura et al., 2008; Yu et al., 2009), and endothelial cells (*Tek, Pecam1, Esam*) (De Val and Black, 2009; Stévant et al., 2018), immune cells (*Cxcl2, Ptprc, Coro1a*), and peritubular myoid cells (*Acta2, Myh11, Cnn1*) were low, confirming the high degree of enrichment of our FLC and ALC samples (Chen et al., 2014; Rebourcet et al., 2014).

We show that the ALCs express in excess the *Insl3* gene with an average of 21,113 RPKM, followed by *Aldh1a1* (8,050 RPKM) and *Cyp17a1* (5,742 RPKM). Together the RNA abundance of three genes represents 9.4% of the whole transcriptome. The FLCs do not display such an extreme over-representation of the same genes, but the top three expressed genes are *mt-Co1* (7,501 RPKM), *Hsd3b1* (5,357 RPKM), and *Insl3* (3,947 RPKM), which represent 5.4% of the total transcriptome. To appreciate the biological functions enriched amongst the most expressed genes of FLCs and ALCs, we performed a pre-ranked gene set enrichment analysis of the transcriptomes, weighting genes by their level of expression. While in FLCs we observed a statistical enrichment of 345 GO terms (FDR<25%), including "glucocorticoid biosynthetic process", "C21 steroid hormone metabolic process" and "regulation of systemic arterial blood pressure by renin angiotensin" in the top terms, no statistical enrichment was observed for the ALCs. Although not statistically enriched, we could find in the top terms "glucocorticoid metabolic process", "C21 steroid hormone metabolic process", and "circadian sleep wake cycle" (of note, the P100 mice were euthanized in the morning, and we know testosterone synthesis is sensitive to the circadian rhythm (Chen et al., 2017)) (**Supplementary Table S2**). These results indicate that both fetal and adult Leydig cells are highly specialised cells dedicated to steroid production. For all subsequent analyses, we retained only the genes coding for proteins and long non-coding RNA (lncRNA).

Wide variations in gene expression levels were observed between FLCs and ALCs

We then thought to evaluate the extent to which the transcriptomes of the two Leydig cell populations are comparable and what genes and biological pathways are differentially expressed. Among the 21,083 protein coding and long-non-coding genes expressed in either FLCs or ALCs, 15,330 genes have no significant difference of expression (**Figure 1B** and **Supplementary Table S1**). Interestingly, a large proportion of genes (5,753 genes, *i.e.* 27% of the total) display significant

variations of expression levels between ALCs and FLCs (adjusted p-value<0.01). Of these genes, 2,357 are overexpressed in FLCs (FC>2) (**Figure 1B** and **C)**. The vast majority of genes overexpressed in FLCs have never been identified as such. This is particularly the case for the top 30 genes with the lowest adjusted p-value such as *Cdkn1c*, *Gpc3*, *Cbln1*, *Sfrp1*, *Myadm*, *Glis2*, *Peg3* and *Smoc2*. In addition, we found also *Thbs2 and Mc2r*, two genes known to be specifically expressed in FLCs (O'Shaughnessy et al., 2002, 2003) as well as 50 genes already reported to be expressed - although not specifically - in FLCs (Jameson et al., 2012; McDowell et al., 2012; McClelland et al., 2015; Inoue et al., 2016)). We grouped the genes differentially expressed into 15 clusters according to their expression profile using a hierarchical clustering (**Figure 1D**). Genes enriched in FLCs are grouped in clusters 3, 4, 7, 8, 9, and 13. Gene Ontology (GO) analysis of these clusters revealed an association with the development of the urogenital system (cluster 3), cell division and differentiation (cluster 4), various metabolic processes (clusters 7, 9 ,13), and response to reactive oxygen species (ROS) (cluster 13) (**Supplementary Table S2**). On the other side, 3,396 genes were found overexpressed in ALCs (FC<0.5) **(Figure 1B** and **C**). Again, the large majority of genes overexpressed in ALCs have never been identified as such. This is particularly the case for genes with the lowest adjusted p-value such as *Gstm2*, *Gstm1*, *Amy1, Csf1*, *Timp2*. As expected, we have found genes known to be specifically expressed in ALCs such as *Hsd3b6*, *Hsd17b3*, *Vcam1*, *Sult1e1* and *Hpgds* (O'Shaughnessy et al., 2002), as well as 750 genes already reported to be expressed - although not specifically - in ALCs (Sanz et al., 2013; O'Shaughnessy et al., 2014). Genes enriched in ALCs are grouped in clusters 2, 5, 6, 10, 11, 12, 14 and 15. This time, GO analysis of these clusters indicated a link with fertilisation (clusters 2, 5), regulation of cellular response (cluster 10), cell-substrate adhesion (cluster 11), regulation of ROS (cluster 14), and various metabolic processes (clusters 12, 14, 15) (**Supplementary Table S2**). The cluster 1 regroups genes with low expression in both LC populations that are involved in stress and immune response. This association with the immune system is consistent with the role of cytokines secreted by testicular macrophages in the regulation of Leydig cell functions (Hales, 2002). Overall, we observed significant variations in the level of expression of thousands of genes, with only a handful of genes exhibiting specific expression in one of the two Leydig cell populations.

Differential splicing between FLCs and ALCs

Alternative splicingis a ubiquitous regulatory mechanism that allows the generation of multiple

transcript isoforms from a single gene, thus expanding the complexity of the proteome. However, the extent of alternative splicing occurring in FLCs and ALCs and its functional relevance remain unclear. To investigate whether these two cell populations exhibit different alternative splicing profiles, we performed a multivariate analysis of transcript splicing. We found 1,971 splicing events that are statistically different between the two LC populations (FDR<0.05) (1,380 events in FLCs and 591 in ALCs) (**Figure 2A, Supplementary Table S2**). These splicing events occur in 1,437 genes, including 1,036 genes in FLCs, and 509 genes in ALCs (with 31 genes having an alternative splicing in both cell populations). We also examined if the genes detected as alternatively spliced correspond to differentially expressed genes. Of the 2,357 FLCs overexpressed genes, 86 of them show an alternative splicing. In the ALCs, 56 genes out of the 3,396 overexpressed genes show an alternative splicing. It appeared that the genes involved in steroidogenesis are not subject to alternative splicing in both FLCs and ALCs.

As shown in **Fig. 2A**, intron retention is the most represented type of alternative splicing in the FLCs, with 742 events found in FLCs but only 34 in ALCs. We investigated if the genes presenting the different type of alternative splicing in both populations are enriched in a particular biological function by performing a GO enrichment analysis (**Supplementary Table S3**). In FLCs, the genes presenting exon skipping are strongly enriched in RNA splicing functions (**Figure 2B**), while genes showing intron retention are involved in mRNA processing and chromatin rearrangements (**Figure 2C**). Regarding ALCs, genes showing exon skipping are involved in various processes such as cellular organization, cellular projection, or muscle system process (**Figure 2D**). No GO enrichment was found in the other types of alternative splicing due to the small number of genes.

Overall, we showed that intron retention is a landmark of the FLC transcriptome. It is known that alternative splicing is frequent during embryonic development, usually cell/organ specific and plays a role in gene expression regulation and protein diversity (Revil et al., 2010; Kalsotra and Cooper, 2011; Grabski et al., 2021).

Characterisation of mutually exclusive marker genes/signatures of FLCs and ALCs

A thorough identification of the genes specifically expressed in FLCs and ALCs has never been achieved. Although our comparative analysis identified many genes that are differentially expressed between FLCs and ALCs, there is no evidence that these are specific to FLCs or ALCs. Indeed, many of them may also be expressed in other testicular cell types. To identify FLC- and

ALC-marker genes, we took advantage of existing single-cell RNA sequencing data from E16.5 testes (3,781 cells) (Neirijnck et al., 2019) and adult testes (24,672 cells) (Ernst et al., 2019). We combined these fetal and adult datasets to evaluate the gene expression specificity. Using five well-established Leydig cell markers, namely *Hsd3b1, Star, Cyp11a1, Cyp17a1* and *Fabp3,* we identified 151 FLCs and 148 ALCs (**Figure 3)** (Ernst et al., 2019; Rebourcet et al., 2019). By comparing the sets of genes specifically enriched in ALCs and FLCs obtained with the single-cell RNA sequencing and the bulk RNA sequencing approaches described above, we generated a high-confidence selection of 62 genes enriched in FLCs showing a link with cell proliferation and differentiation, and hormone secretion (**Supplementary Table S3**) and 120 genes enriched in ALCs related to protein processing, spermatid development, and metabolic processes (**Supplementary Table S4**).

Identification of nine FLC-specific marker genes

Although we have identified 62 genes showing an expression in FLCs and low or no expression in ALCs, there is no evidence that these genes are specific to FLCs. Using single-cell transcriptomic data, we excluded the ones that were also expressed in other testicular cell populations. We have thus identified nine genes that were considered as specific testicular markers of FLCs (**Figure 4A-G** and **Table 1**). This includes genes *Crhr1* (McDowell et al., 2012; McClelland et al., 2015; Inoue et al., 2016) (**Figure 4C**), *Ren1* (Jameson et al., 2012; Inoue et al., 2016) (**Figure 4D**) and *Vsnl1* (Jameson et al., 2012; McClelland et al., 2015) (**Figure 4E**), whose expression in FLCs (but not their specificity) has already been demonstrated. In addition, we have identified six additional genes described for the first time as FLC specific markers, including *Cyp26b1* (**Figure 4F**), *Gsg1l, Pcsk6, Nppc, Cdon* and *Ppp2r5b*. Contrary to our expectations, *Mc2r* and *Thbs2* are not part of our selection. *Thbs2* is excluded by our filters because it does not show any enrichment in FLCs compared to ALCs, while *Mc2r* (**Figure 4G**) seems specific to FLCs but is not retained due to its low expression in single-cell transcriptomic data. The specific expression of FLC marker genes *Crhr1* and *Ren1* was validated by *in situ* hybridisation and compared with *Star* expression, a known marker of Leydig cells (**Figure 4H-K**). We found that both *Crhr1* and *Ren1* are co-expressed with *Star* in E18.5 testis but not in adult P100 testis, confirming that these genes are specifically expressed in FLCs and not in ALCs or any other testicular cells (**Figure 4H-K**).

Identification of 50 ALC-specific marker genes

We used the same approach to identify testicular marker genes specific to ALCs. Among the 120 genes specifically enriched in ALCs (compared to FLCs), 50 genes were considered as specific markers of ALCs (**Figure 5A-G** and **Table 2**). Among these 50 genes, we found *Hsd17b3* and *Hsd3b6*, two known markers of ALCs confirming our analysis. The remaining 48 genes are described for the first time as ALC specific markers. The analysis of the GO terms highlighted cellular processes such as protein transformation (*Klk1b21, Klk1b24, Klk1b27, C1rl*), peptidase regulation (*Serpina3c, Serpina3g, Serpina3n, Serpina5*) as well as metabolic processes (*Hsd17b3, Sult1e1, Bhmt*). To validate further our analysis, the specific expression of ALC marker genes *Bhmt* and *Sult1e1* was confirmed by *in situ* hybridisation (**Figure 5H-K**). We found that *Bhmt* and *Sult1e1* are co-expressed with *Star* in adult P100 testis but not in fetal E18.5 testis confirming that these genes are specifically expressed in ALCs and not in FLCs or any other testicular cells.

Overall, our analysis combining both bulk RNA sequencing and single-cell RNA sequencing resulted in the identification of 9 and 50 specific markers for FLCs and ALCs, respectively, most of which are newly identified.

Discussion

The main purpose of our study was to characterize at transcriptomic level the similarities and differences between FLCs and ALCs, the major androgenic cells of the testis, using both bulk and single-cell RNA sequencing. Significant differences were observed both in terms of expression level, with 2,357 genes overexpressed in FLCs (11.2% of the total) and 3,396 genes overexpressed in ALCs (16.1%); and in terms of alternative splicing, with an over-representation of intron retention events in FLCs compared to ALCs. Our study also identified many specific markers for each Leydig cell populations, with 9 genes for FLCs and 50 genes for ALCs, most of them newly described.

Identification of FLC- and ALC-specific genes

The purity of the Leydig cell population is critical for the identification of Leydig cell markers using microarray and bulk RNA sequencing analyses. We have multiple indications that support the assertion that the 9 FLC-specific markers - and 50 ALC-specific markers - identified in this

study are robust and specific. First, our transcriptomic analysis combines two independent sources of data, namely those from the bulk RNA sequencing data of FLCs and ALCs, in which Leydig cells were sorted according to the level of GFP expression, but also a single-cell RNA sequencing of the testes of fetal and adult mice. In addition, the few markers already known, in particular *Hsd17b3* and *Hsd3b6* were also identified in the list of markers specific for ALCs. Finally, an independent validation by RNAscope® of *Ren1* and *Crhr1* as FLC-specific marker genes, and *Bhmt* and *Sult1e1* as ALC-specific marker genes, confirmed their specific expression.

Several genome-wide expression studies using microarray technology or bulk RNA sequencing have investigated the transcriptome of FLCs (Jameson et al., 2012; McDowell et al., 2012; McClelland et al., 2015). In these studies, they isolated and evaluated the transcriptome of several populations present in the fetal testis, such as germ cells, Sertoli cells, Leydig cells and interstitial cells. Differential analysis of expression among these cell populations led to the identification of 166 overexpressed genes in FLCs. However, since the analysed cell populations represent only a fraction of the cell types present in the fetal testis, this list of FLC overexpressed genes is overestimated. In contrast, our analysis combining bulk and single cell RNA-seq identified nine FLC-specific candidates of which three were already described in these previous studies. We found that the other genes initially described as enriched in FLCs are mostly non-specific, either expressed in ALCs or in other testicular cell types. Among the three genes specifically enriched in FLCs are the *Crhr1* (McDowell et al., 2012), *Vsnl1* (Jameson et al., 2012; McClelland et al., 2015) and *Ren1* (Jameson et al., 2012) genes. The *corticotropin releasing hormone receptor 1 (Crhr1)* is of particular interest as its ligand CRH (Corticotropin Releasing Hormone) is known to stimulate testosterone production in the fetal testes (McDowell et al., 2012) (see last paragraph of the discussion for more details and **Figure 6**). Moreover, our analysis revealed six new FLC-specific candidates including *Cyp26b1*, a gene coding for an enzyme degrading retinoic acid (RA), an active metabolite of retinol involved in meiosis regulation (Bowles et al., 2006; Koubova et al., 2006). It has recently been reported that in *Cyp26b1*−/− mutant mice, Leydig cell differentiation is impaired and steroidogenesis is decreased (Bowles et al., 2018).

Previous transcriptomics studies have used indirect methods to identify ALC-specific mRNA transcripts, such as the response of Leydig cells to hormones (Sanz et al., 2013), or cell ablation model using ethane dimethane sulphonate (EDS) to ablate LCs in adult male rats (O'Shaughnessy

et al., 2014). These studies have resulted in a combined list of over 2000 genes whose expression is enriched in ALCs. Although ingenious, these approaches due to their technical bias and limitations do not guarantee an ALC-specific expression. Here, we have identified 50 genes with ALC-specific expression (**Table 2**). Confirming our results, three genes known to be specific for ALCs, *Hsd3b6*, *Hsd17b3* and *Vcam1* are also present in our list (O'Shaughnessy et al., 2000; Shima et al., 2013; Wen et al., 2014). Among these 50 genes, six were not described in the previous studies, namely *Ces1d*, *Serpina3g*, *Rarres1*, *Bpifb5*, *Eppin*, and *Espn*. Hundreds of genes initially described as enriched in ALCs by previous studies were excluded because their expression was not specific. This is particularly the case for *Ptgds* and *Hsd11b1*, two genes often described as specific to ALCs (Baker and O'Shaughnessy, 2001; Wen et al., 2014). In this study, we validated by RNAscope® the ALC-specific expression of two genes: *Betaine-Homocysteine S-Methyltransferase* (*Bhmt*) and *Sulfotransferase Family 1E Member 1* (*Sult1e1*). We also proved the specific expression of BHMT at the protein level by immunohistochemistry (**Supplementary Figure S2**). BHMT plays a key role in regulating betaine concentration, that can be stored to control cellular osmolarity or metabolised to provide a methyl group for homocysteine methylation (Alirezaei et al., 2012). It has been shown that the testes are among the organs that contains the most betaine (Slow et al., 2009). SULT1E1, for its part, plays a protective role for Leydig cells and seminiferous tubules against oestrogen overstimulation by catalysing the sulfo-conjugation and inactivation of oestrogens (Song, 2007). It is also noteworthy to find several members of the SERPIN family among our ALC-specific genes (*Serpina3c*, *Serpina3g*, *Serpina3n*, *Serpina5*) as most of them have been found in Leydig cells and seem to be sensitive to hCG (human chorionic gonadotropin) (Odet et al., 2006).

Alternative splicing: high intron retention in FLCs

The main types of alternative splicing are alternative exon usage, alternative 5′ or 3′ splice sites, mutually exclusive exons, and intron retention. Intron retention is characterized by the inclusion of one or more introns in mature mRNA transcripts and has been previously considered to be an artefact of a dysfunctional spliceosome. It is known that alternative splicing, including intron retention, is frequent during embryonic development and contribute not only to the plasticity of the transcriptome but also the regulation of gene expression and protein diversity (Revil et al., 2010; Kalsotra and Cooper, 2011; Grabski et al., 2021). Here, we showed that intron retention is a

landmark of the FLC transcriptome. Messenger RNA displaying intron retention are generally restricted from exiting the nucleus. This was proposed as a mechanism to downregulate gene expression (Grabski et al., 2021). In FLCs, we also showed that genes presenting alternative exon skipping are involved in splicing regulation itself. The control expression levels and activities of RNA binding proteins (RBPs) that regulate RNA splicing is mediated by auto-regulatory feedbacks by directly influencing the splicing of their own mRNAs (Müller-McNicoll et al., 2019). In particular, the regulation of the splicing factors of the SR (Serine/arginine rich) family regulate their activity by modulating the inclusion of a cassette exon containing a premature termination codon to produce or nor a functional protein (Müller-McNicoll et al., 2019). Tight regulation of the splicing factors is necessary for post-transcriptional gene expression regulation. The intron retention events observed in FLCs might subsequently result from the auto-regulation feedback of the splicing factors. Post-transcriptional gene expression regulation through alternative splicing have been identified as key player in the differentiation of mesenchymal stem cells (Park et al., 2020). We can postulate that the regulation of the FLC differentiation might also be mediated by alternative splicing.

Differences in FLCs and ALCs transcriptomes affect steroidogenesis and its regulation

FLCs and ALCs display significant differences in both steroidogenic regulation and the type of androgen produced (androstenedione vs testosterone) (O'Shaughnessy et al., 2000, 2002; Shima et al., 2013). Our transcriptomic data confirmed the differences in androgen production, the expression of the *Hsd17b3* gene encoding the enzyme responsible for the conversion of androstenedione to testosterone is not expressed in FLCs but only in ALCs (**Table 2** and **Figure 6**), which explains why FLCs synthesise mainly androstenedione and ALCs are capable of producing testosterone (O'Shaughnessy et al., 2000; Rebourcet et al., 2020). Regarding the differences in steroidogenesis regulation, LH appears not to be essential for FLC function since androgen production and masculinization of the fetus occurs normally in LH/CG receptor knockout mice (Kendall et al., 1995; Lei et al., 2001; Zhang et al., 2001; Ma et al., 2004; O'Shaughnessy and Fowler, 2011; Teerds and Huhtaniemi, 2015). However in *T/ebp/Nkx2.1* null mice, which lack a pituitary gland, testicular androgen levels are markedly reduced in late gestation, suggesting that additional hypothalamo/pituitary hormones may be required for Leydig cell function and androgen production. Interestingly, our transcriptomic analysis revealed that the ACTH receptor,

melanocortin type 2 receptor (*Mc2r*), and the corticotropin releasing hormone receptor 1 *(Crhr1)* are both specifically expressed in FLCs and absent from ALCs (**Table 1** and **Figure 6**). ACTH has been reported to stimulate *in vitro* testosterone production in fetal but not in adult testes suggesting that FLCs, but not ALCs, are sensitive to ACTH stimulation(O'Shaughnessy et al., 2003). However, fetal testosterone levels were normal in *Proopiomelanocortin* (*POMC*)-deficient mice that lack circulating ACTH, indicating that ACTH, like LH, is not essential for FLC function. Corticotropin-releasing hormone (CRH) has been also reported to stimulate steroidogenesis by direct activation of FLCs in fetal rat and mouse testes *ex vivo* and in MA-10 mouse Leydig cells (McDowell et al., 2012). In contrast, CRH does not enhance steroidogenesis in primary ALCs (Huang et al., 1995; McDowell et al., 2012). Combined together, these results suggest a sequential regulation of steroidogenesis in LCs. In this model, androgen production by FLCs is stimulated by three potentially redundant hypothalamo/pituitary hormones, namely LH, CRH and ACTH. Fetal androgen production can occur in the absence of any of these hormones with the two other hormones able to maintain FLC steroidogenic activity. Conversely, activation of steroidogenesis in ALCs is LH-dependent and CRH- and ACTH-independent, since *Crhr1* and *Mc2r* are not expressed in these cells (**Figure 6**). Although this model of steroid regulation by Leydig cells needs to be confirmed by further studies, such a mechanism may have evolved to ensure the production of adequate levels of androgens during fetal development.

Acknowledgments

The authors thank Violaine Regard from the NEF laboratory, Cécile Gameiro and Gregory Schneiter from the Flow Cytometry Facility, the Genomics Platform of iGE3, and the Animal Facility of the Faculty of Medicine of the University of Geneva.

References

Ademi, H., Stévant, I., Rands, C., Conne, B., and Nef, S. (2020). Expression of Wnt5a defines the major progenitors of fetal and adult Leydig cells. *bioRxiv*, 2020.07.25.221069. doi:10.1101/2020.07.25.221069.

Alirezaei, M., Jelodar, G., and Ghayemi, Z. (2012). Antioxidant defense of betaine against oxidative stress induced by ethanol in the rat testes. *Int. J. Pept. Res. Ther.* 18, 239–247. doi:10.1007/s10989-012-9297-9.

Baker, P. J., and O'Shaughnessy, P. J. (2001). Expression of prostaglandin D synthetase during development in the mouse testis. *Reproduction* 122, 553–559. doi:10.1530/rep.0.1220553.

Baker, P. J., Sha, J. A., McBride, M. W., Peng, L., Payne, A. H., and O'Shaughnessy, P. J. (1999). Expression of 3beta-hydroxysteroid dehydrogenase type I and type VI isoforms in the mouse testis during development. *Eur. J. Biochem.* 260, 911–917. doi:10.1046/j.1432-1327.1999.00245.x.

Barsoum, I. B., Kaur, J., Ge, R. S., Cooke, P. S., and Yao, H. H.-C. H. C. (2013). Dynamic changes in fetal Leydig cell populations influence adult Leydig cell populations in mice. *FASEB J.* 27, 2657–2666. doi:10.1096/fj.12-225060.

Barsoum, I. B., and Yao, H. H.-C. (2010). Fetal Leydig Cells: Progenitor Cell Maintenance and Differentiation. *J. Androl.* 31, 11–15. doi:10.2164/jandrol.109.008318.

Becht, E., McInnes, L., Healy, J., Dutertre, C. A., Kwok, I. W. H., Ng, L. G., et al. (2019). Dimensionality reduction for visualizing single-cell data using UMAP. *Nat. Biotechnol.* 37, 38–47. doi:10.1038/nbt.4314.

Beverdam, A., and Koopman, P. (2006). Expression profiling of purified mouse gonadal somatic cells during the critical time window of sex determination reveals novel candidate genes for human sexual dysgenesis syndromes. *Hum. Mol. Genet.* 15, 417–431. doi:10.1093/hmg/ddi463.

Bitgood, M. J., Shen, L., and McMahon, A. P. (1996). Sertoli cell signaling by Desert hedgehog regulates the male germline. *Curr. Biol.* 6, 298–304. doi:10.1016/S0960-9822(02)00480-3.

Bowles, J., Feng, C. W., Ineson, J., Miles, K., Spiller, C. M., Harley, V. R., et al. (2018). Retinoic Acid Antagonizes Testis Development in Mice. *Cell Rep.* 24, 1330–1341. doi:10.1016/j.celrep.2018.06.111.

Bowles, J., Knight, D., Smith, C., Wilhelm, D., Richman, J., Mamiya, S., et al. (2006). Retinoid

Signaling Determines Germ Cell Fate in Mice. *Science (80-.).* 312, 596–600. doi:10.1126/science.1125691.

Brennan, J., Tilmann, C., and Capel, B. (2003). Pdgfr-alpha mediates testis cord organization and fetal Leydig cell development in the XY gonad. *Genes Dev.* 17, 800–810. doi:10.1101/gad.1052503.

Buaas, F. W., Gardiner, J. R., Clayton, S., Val, P., and Swain, A. (2012). In vivo evidence for the crucial role of SF1 in steroidproducing cells of the testis, ovary and adrenal gland. *Dev.* 139, 4561–4570. doi:10.1242/dev.087247.

Byskov, A. G. (1986). Differential of mammalian embryonic gonad. *Physiol. Rev.* 66, 71–117. doi:10.1152/physrev.1986.66.1.71.

Castrillon, D. H., Quade, B. J., Wang, T. Y., Quigley, C., and Crum, C. P. (2000). The human VASA gene is specifically expressed in the germ cell lineage. *Proc. Natl. Acad. Sci. U. S. A.* 97, 9585–90. doi:10.1073/pnas.160274797.

Chen, H., Gao, L., Xiong, Y., Yang, D., Li, C., Wang, A., et al. (2017). Circadian clock and steroidogenic-related gene expression profiles in mouse Leydig cells following dexamethasone stimulation. *Biochem. Biophys. Res. Commun.* 483, 294–300. doi:10.1016/j.bbrc.2016.12.149.

Chen, H., Ge, R.-S. S., and Zirkin, B. R. (2009). Leydig cells: From stem cells to aging. *Mol. Cell. Endocrinol.* 306, 9–16. doi:10.1016/j.mce.2009.01.023.

Chen, L. Y., Brown, P. R., Willis, W. B., and Eddy, E. M. (2014). Peritubular myoid cells participate in male mouse spermatogonial stem cell maintenance. *Endocrinology* 155, 4964–4974. doi:10.1210/en.2014-1406.

Davidoff, M. S., Middendorff, R., Enikolopov, G., Riethmacher, D., Holstein, A. F., and Müller, D. (2004). Progenitor cells of the testosterone-producing Leydig cells revealed. *J. Cell Biol.* 167, 935–44. doi:10.1083/jcb.200409107.

De Val, S., and Black, B. L. (2009). Transcriptional Control of Endothelial Cell Development. *Dev. Cell* 16, 180–195. doi:10.1016/j.devcel.2009.01.014.

Ernst, C., Eling, N., Martinez-Jimenez, C. P., Marioni, J. C., and Odom, D. T. (2019). Staged developmental mapping and X chromosome transcriptional dynamics during mouse spermatogenesis. *Nat. Commun.* 10, 1–20. doi:10.1038/s41467-019-09182-1.

Grabski, D. F., Broseus, L., Kumari, B., Rekosh, D., Hammarskjold, M., and Ritchie, W. (2021).

Intron retention and its impact on gene expression and protein diversity: A review and a practical guide. *WIREs RNA* 12, e1631. doi:10.1002/wrna.1631.

Habert, R., Lejeune, H., and Saez, J. M. (2001). Origin, differentiation and regulation of fetal and adult Leydig cells. *Mol. Cell. Endocrinol.* 179, 47–74. doi:10.1016/S0303-7207(01)00461-0.

Haider, S. G. (2004). Cell Biology of Leydig Cells in the Testis. *Int. Rev. Cytol.* 233, 181–241. doi:10.1016/S0074-7696(04)33005-6.

Huang, B. M., Stocco, D. M., Hutson, J. C., and Norman, R. L. (1995). Corticotropin-releasing hormone stimulates steroidogenesis in mouse Leydig cells. *Biol. Reprod.* 53, 620–626. doi:10.1095/biolreprod53.3.620.

Huhtaniemi, I., and Pelliniemi, L. J. (1992). Fetal Leydig Cells: Cellular Origin, Morphology, Life Span, and Special Functional Features. *Exp. Biol. Med.* 201, 125–140. doi:10.3181/00379727-201-43493.

Inoue, M., Shima, Y., Miyabayashi, K., Tokunaga, K., Sato, T., Baba, T., et al. (2016). Isolation and Characterization of Fetal Leydig Progenitor Cells of Male Mice. *Endocrinology* 157, 1222–1233. doi:10.1210/en.2015-1773.

Jameson, S. A., Natarajan, A., Cool, J., DeFalco, T., Maatouk, D. M., Mork, L., et al. (2012). Temporal transcriptional profiling of somatic and germ cells reveals biased lineage priming of sexual fate in the fetal mouse gonad. *PLoS Genet.* 8, e1002575. doi:10.1371/journal.pgen.1002575.

Kalsotra, A., and Cooper, T. A. (2011). Functional consequences of developmentally regulated alternative splicing. *Nat. Rev. Genet.* 12, 715–729. doi:10.1038/nrg3052.

Kendall, S. K., Samuelson, L. C., Saunders, T. L., Wood, R. I., and Camper, S. A. (1995). Targeted disruption of the pituitary glycoprotein hormone α-subunit produces hypogonadal and hypothyroid mice. *Genes Dev.* 9, 2007–2019. doi:10.1101/gad.9.16.2007.

Kerr, J. B. B., Risbridger, G. P. P., and Knell, C. M. M. (1988). Stimulation of interstitial cell growth after selective destruction of foetal Leydig cells in the testis of postnatal rats. *Cell Tissue Res.* 252, 89–98. doi:10.1007/BF00213829.

Kilcoyne, K. R., Smith, L. B., Atanassova, N., Macpherson, S., McKinnell, C., van den Driesche, S., et al. (2014). Fetal programming of adult Leydig cell function by androgenic effects on stem/progenitor cells. *Proc. Natl. Acad. Sci.* 111, E1924–E1932. doi:10.1073/pnas.1320735111.

656 Koubova, J., Menke, D. B., Zhou, Q., Cape, B., Griswold, M. D., Page, D. C., et al. (2006). Retinoic
657 acid regulates sex-specific timing of meiotic initiation in mice. *Proc. Natl. Acad. Sci. U. S. A.*
658 103, 2474–2479. doi:10.1073/pnas.0510813103.
659 Kumar, A., D'Souza, S. S., Moskvin, O. V., Toh, H., Wang, B., Zhang, J., et al. (2017).
660 Specification and Diversification of Pericytes and Smooth Muscle Cells from
661 Mesenchymoangioblasts. *Cell Rep.* 19, 1902–1916. doi:10.1016/j.celrep.2017.05.019.
662 Lei, Z. M., Mishra, S., Zou, W., Xu, B., Foltz, M., Li, X., et al. (2001). Targeted Disruption of
663 Luteinizing Hormone/Human Chorionic Gonadotropin Receptor Gene. *Mol. Endocrinol.* 15,
664 184–200. doi:10.1210/mend.15.1.0586.
665 Liu, C., Rodriguez, K., and Yao, H. H. C. H.-C. (2016). Mapping lineage progression of somatic
666 progenitor cells in the mouse fetal testis. *Development* 143, dev.135756.
667 doi:10.1242/dev.135756.
668 Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion
669 for RNA-seq data with DESeq2. *Genome Biol.* 15, 550. doi:10.1186/s13059-014-0550-8.
670 Ma, X., Dong, Y., Matzuk, M. M., and Kumar, T. R. (2004). Targeted disruption of luteinizing
671 hormone $β$-subunit leads to hypogonadism, defects in gonadal steroidogenesis, and
672 infertility. *Proc. Natl. Acad. Sci. U. S. A.* 101, 17294–17299. doi:10.1073/pnas.0404743101.
673 Marco-Sola, S., Sammeth, M., Guigó, R., and Ribeca, P. (2012). The GEM mapper: Fast, accurate
674 and versatile alignment by filtration. *Nat. Methods* 9, 1185–1188. doi:10.1038/nmeth.2221.
675 McClelland, K. S., Bell, K., Larney, C., Harley, V. R., Sinclair, A. H., Oshlack, A., et al. (2015).
676 Purification and Transcriptomic Analysis of Mouse Fetal Leydig Cells Reveals Candidate
677 Genes for Specification of Gonadal Steroidogenic Cells1. *Biol. Reprod.* 92, 145–145.
678 doi:10.1095/biolreprod.115.128918.
679 McDowell, E. N., Kisielewski, A. E., Pike, J. W., Franco, H. L., Yao, H. H.-C., and Johnson, K. J.
680 (2012). A Transcriptome-Wide Screen for mRNAs Enriched in Fetal Leydig Cells: CRHR1
681 Agonism Stimulates Rat and Mouse Fetal Testis Steroidogenesis. *PLoS One* 7, e47359.
682 doi:10.1371/journal.pone.0047359.
683 Migrenne, S., Pairault, C., Racine, C., Livera, G., Géloso, A., and Habert, R. (2001). Luteinizing
684 hormone-dependent activity and luteinizing hormone-independent differentiation of rat fetal
685 Leydig cells. *Mol. Cell. Endocrinol.* 172, 193–202. doi:10.1016/S0303-7207(00)00339-7.
686 Miyabayashi, K., Shima, Y., Inoue, M., Sato, T., Baba, T., Ohkawa, Y., et al. (2017). Alterations

687 in Fetal Leydig Cell Gene Expression during Fetal and Adult Development. *Sex. Dev.* 11, 53–
688 63. doi:10.1159/000453323.
689 Mucenski, M. L., Mahoney, R., Adam, M., Potter, A. S., and Potter, S. S. (2019). Single cell RNA-
690 seq study of wild type and Hox9,10,11 mutant developing uterus. *Sci. Rep.* 9, 4557.
691 doi:10.1038/s41598-019-40923-w.
692 Müller-McNicoll, M., Rossbach, O., Hui, J., and Medenbach, J. (2019). Auto-regulatory feedback
693 by RNA-binding proteins. *J. Mol. Cell Biol.* 11, 930–939. doi:10.1093/jmcb/mjz043.
694 Munger, S. C., Natarajan, A., Looger, L. L., Ohler, U., and Capel, B. (2013). Fine time course
695 expression analysis identifies cascades of activation and repression and maps a putative
696 regulator of mammalian sex determination. *PLoS Genet.* 9, e1003630.
697 doi:10.1371/journal.pgen.1003630.
698 Nef, S., Schaad, O., Stallings, N. R., Cederroth, C. R., Pitetti, J.-L. L., Schaer, G., et al. (2005).
699 Gene expression during sex determination reveals a robust female genetic program at the onset
700 of ovarian development. *Dev. Biol.* 287, 361–77. doi:10.1016/j.ydbio.2005.09.008.
701 Neirijnck, Y., Kühne, F., Mayère, C., Pavlova, E., Sararols, P., Foti, M., et al. (2019). Tumor
702 suppressor PTEN regulates negatively sertoli cell proliferation, testis size, and sperm
703 production in vivo. *Endocrinology* 160. doi:10.1210/en.2018-00892.
704 O'Shaughnessy, P. J., Baker, P. J., Heikkilä, M., Vainio, S., and McMahon, A. P. (2000).
705 Localization of 17$β$-Hydroxysteroid Dehydrogenase/17-Ketosteroid Reductase Isoform
706 Expression in the Developing Mouse Testis—Androstenedione Is the Major Androgen
707 Secreted by Fetal/Neonatal Leydig Cells*. *Endocrinology* 141, 2631–2637.
708 doi:10.1210/endo.141.7.7545.
709 O'Shaughnessy, P. J., Baker, P., Sohnius, U., Haavisto, A.-M. M., Charlton, H. M., and
710 Huhtaniemi, I. (1998). Fetal Development of Leydig Cell Activity in the Mouse Is
711 Independent of Pituitary Gonadotroph Function*. *Endocrinology* 139, 1141–1146.
712 doi:10.1210/endo.139.3.5788.
713 O'Shaughnessy, P. J., Fleming, L. M., Jackson, G., Hochgeschwender, U., Reed, P., and Baker, P.
714 J. (2003). Adrenocorticotropic Hormone Directly Stimulates Testosterone Production by the
715 Fetal and Neonatal Mouse Testis. *Endocrinology* 144, 3279–3284. doi:10.1210/en.2003-
716 0277.
717 O'Shaughnessy, P. J., and Fowler, P. A. (2011). Endocrinology of the mammalian fetal testis.

718 *Reproduction* 141, 37–46. doi:10.1530/REP-10-0365.
719 O'Shaughnessy, P. J., Monteiro, A., Fowler, P. A., and Morris, I. D. (2014). Identification of
720 Leydig cell-specific mRNA transcripts in the adult rat testis. *Reproduction* 147, 671–682.
721 doi:10.1530/REP-13-0603.
722 O'Shaughnessy, P. J., Willerton, L., and Baker, P. J. (2002). Changes in Leydig Cell Gene
723 Expression During Development in the Mouse1. *Biol. Reprod.* 66, 966–975.
724 doi:10.1095/biolreprod66.4.966.
725 Odet, F., Verot, A., and Le Magueresse-Battistoni, B. (2006). The Mouse Testis Is the Source of
726 Various Serine Proteases and Serine Proteinase Inhibitors (SERPINs): Serine Proteases and
727 SERPINs Identified in Leydig Cells Are under Gonadotropin Regulation. *Endocrinology* 147,
728 4374–4383. doi:10.1210/en.2006-0484.
729 Okamura, D., Tokitake, Y., Niwa, H., and Matsui, Y. (2008). Requirement of Oct3/4 function for
730 germ cell specification. *Dev. Biol.* 317, 576–584. doi:10.1016/J.YDBIO.2008.03.002.
731 Pakarinen, P., Kimura, S., El-Gehani, F., Pelliniemi, L. J., and Huhtaniemi, I. (2002). Pituitary
732 hormones are not required for sexual differentiation of male mice: Phenotype of the
733 T/ebp/Nkx2.1 null mutant mice. *Endocrinology* 143, 4477–4482. doi:10.1210/en.2002-
734 220052.
735 Park, J. W., Fu, S., Huang, B., and Xu, R.-H. (2020). Alternative splicing in mesenchymal stem
736 cell differentiation. *Stem Cells* 38, 1229–1240. doi:10.1002/stem.3248.
737 Pitetti, J.-L., Calvel, P., Romero, Y., Conne, B., Truong, V., Papaioannou, M. D., et al. (2013).
738 Insulin and IGF1 Receptors Are Essential for XX and XY Gonadal Differentiation and
739 Adrenal Development in Mice. *PLoS Genet.* 9, e1003160. doi:10.1371/journal.pgen.1003160.
740 Polański, K., Young, M. D., Miao, Z., Meyer, K. B., Teichmann, S. A., and Park, J.-E. (2019).
741 BBKNN: fast batch alignment of single cell transcriptomes. *Bioinformatics*.
742 doi:10.1093/bioinformatics/btz625.
743 Rebourcet, D., Mackay, R., Darbey, A., Curley, M. K., Jørgensen, A., Frederiksen, H., et al. (2020).
744 Ablation of the canonical testosterone production pathway via knockout of the steroidogenic
745 enzyme HSD17B3, reveals a novel mechanism of testicular testosterone production. *FASEB*
746 *J.* 34, 10373–10386. doi:10.1096/fj.202000361R.
747 Rebourcet, D., Monteiro, A., Cruickshanks, L., Jeffery, N., Smith, S., Milne, L., et al. (2019).
748 Relationship of transcriptional markers to Leydig cell number in the mouse testis. *PLoS One*

14, e0219524. doi:10.1371/journal.pone.0219524.

Rebourcet, D., O'Shaughnessy, P. J., Monteiro, A., Milne, L., Cruickshanks, L., Jeffrey, N., et al. (2014). Sertoli Cells Maintain Leydig Cell Number and Peritubular Myoid Cell Activity in the Adult Mouse Testis. *PLoS One* 9, e105687. doi:10.1371/journal.pone.0105687.

Rehman, Z. ur, Worku, T., Davis, J. S., Talpur, H. S., Bhattarai, D., Kadariya, I., et al. (2017). Role and mechanism of AMH in the regulation of Sertoli cells in mice. *J. Steroid Biochem. Mol. Biol.* 174, 133–140. doi:10.1016/j.jsbmb.2017.08.011.

Revil, T., Gaffney, D., Dias, C., Majewski, J., and Jerome-Majewska, L. A. (2010). Alternative splicing is frequent during early embryonic development in mouse. *BMC Genomics* 11, 399. doi:10.1186/1471-2164-11-399.

Sanz, E., Evanoff, R., Quintana, A., Evans, E., Miller, J. A., Ko, C., et al. (2013). RiboTag Analysis of Actively Translated mRNAs in Sertoli and Leydig Cells In Vivo. *PLoS One* 8, e66179. doi:10.1371/journal.pone.0066179.

Shen, S., Park, J. W., Huang, J., Dittmar, K. A., Lu, Z. X., Zhou, Q., et al. (2012). MATS: A Bayesian framework for flexible detection of differential alternative splicing from RNA-Seq data. *Nucleic Acids Res.* 40, e61–e61. doi:10.1093/nar/gkr1291.

Shen, S., Park, J. W., Lu, Z. X., Lin, L., Henry, M. D., Wu, Y. N., et al. (2014). rMATS: Robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. *Proc. Natl. Acad. Sci. U. S. A.* 111, E5593–E5601. doi:10.1073/pnas.1419161111.

Shen, Y., Larose, H., Shami, A. N., Moritz, L., Manske, G., Ma, Q., et al. (2020). Tcf21 + mesenchymal cells contribute to testis somatic cell development, homeostasis, and regeneration. *bioRxiv*, 2020.05.02.074518. doi:10.1101/2020.05.02.074518.

Shima, Y. (2019). Development of fetal and adult Leydig cells. *Reprod. Med. Biol.* 18, 323–330. doi:10.1002/rmb2.12287.

Shima, Y., Miyabayashi, K., Haraguchi, S., Arakawa, T., Otake, H., Baba, T., et al. (2013). Contribution of Leydig and Sertoli cells to testosterone production in mouse fetal testes. *Mol. Endocrinol.* 27, 63–73. doi:10.1210/me.2012-1256.

Shima, Y., Miyabayashi, K., Sato, T., Suyama, M., Ohkawa, Y., Doi, M., et al. (2018). Fetal Leydig cells dedifferentiate and serve as adult Leydig stem cells. *Development* 145, dev169136. doi:10.1242/dev.169136.

Slow, S., Lever, M., Chambers, S. T., and George, P. M. (2009). Plasma dependent and

independent accumulation of betaine in male and female rat tissues. *Physiol. Res.* 58, 403–410.

Song, W.-C. (2007). "The Role of Estrogen Sulfotransferase in Leydig Cells," in *The Leydig Cell in Health and Disease* (Humana Press), 197–205. doi:10.1007/978-1-59745-453-7_14.

Stallings, N. R. (2002). Development of a Transgenic Green Fluorescent Protein Lineage Marker for Steroidogenic Factor 1. *Mol. Endocrinol.* 16, 2360–2370. doi:10.1210/me.2002-0003.

Stévant, I., Neirijnck, Y., Borel, C., Escoffier, J., Smith, L. B., Antonarakis, S. E., et al. (2018). Deciphering Cell Lineage Specification during Male Sex Determination with Single-Cell RNA Sequencing. *Cell Rep.* 22, 1589–1599. doi:10.1016/j.celrep.2018.01.043.

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* 102, 15545–15550. doi:10.1073/pnas.0506580102.

Teerds, K. J., and Huhtaniemi, I. T. (2015). Morphological and functional maturation of Leydig cells: From rodent models to primates. *Hum. Reprod. Update* 21, 310–328. doi:10.1093/humupd/dmv008.

Wakayama, T., Hamada, K., Yamamoto, M., Suda, T., and Iseki, S. (2003). The expression of platelet endothelial cell adhesion molecule-1 in mouse primordial germ cells during their migration and early gonadal formation. *Histochem. Cell Biol.* 119, 355–362. doi:10.1007/s00418-003-0528-1.

Waltman, L., and van Eck, N. J. (2013). A smart local moving algorithm for large-scale modularity-based community detection. *Eur. Phys. J. B* 86, 471. doi:10.1140/epjb/e2013-40829-0.

Wen, Q., Cheng, C. Y. Y., and Liu, Y.-X. X. (2016). Development, function and fate of fetal Leydig cells. *Semin. Cell Dev. Biol.* 59, 89–98. doi:10.1016/j.semcdb.2016.03.003.

Wen, Q., Zheng, Q.-S., Li, X.-X., Hu, Z.-Y., Gao, F., Cheng, C. Y., et al. (2014). Wt1 dictates the fate of fetal and adult Leydig cells during development in the mouse testis. *Am. J. Physiol. Endocrinol. Metab.* 307, E1131-43. doi:10.1152/ajpendo.00425.2014.

Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16, 284–7. doi:10.1089/omi.2011.0118.

Yu, Z., Ji, P., Cao, J., Zhu, S., Li, Y., Zheng, L., et al. (2009). Dazl Promotes Germ Cell Differentiation from Embryonic Stem Cells. *J. Mol. Cell Biol.* 1, 93–103.

811 doi:10.1093/jmcb/mjp026.

812 Zhang, F.-P., Poutanen, M., Wilbertz, J., and Huhtaniemi, I. (2001). Normal Prenatal but Arrested

813 Postnatal Sexual Development of Luteinizing Hormone Receptor Knockout (LuRKO) Mice.

814 *Mol. Endocrinol.* 15, 172–183. doi:10.1210/mend.15.1.0582.

815

816

817

Tables

Table 1. FLC specific genes

Gene	baseMean	log2FoldChange	Fold Change	p value	p adj
Gsg1l	6 453	5,109	34,511	2,69E-295	2,54E-292
Crhr1	3 887	10,022	1 039,829	4,24E-191	1,31E-188
Cyp26b1	23 010	6,389	83,795	1,96E-109	1,70E-107
Pcsk6	2 697	4,599	24,240	1,38E-56	3,84E-55
Ren1	24 384	5,091	34,091	1,11E-29	1,31E-28
Vsnl1	168	10,669	1 627,844	1,54E-17	1,02E-16
Ppp2r5b	8 531	1,116	2,168	1,18E-15	6,91E-15
Nppc	307	3,352	10,212	2,17E-14	1,17E-13
Cdon	28 637	2,341	5,067	1,88E-05	4,58E-05

Table 2. ALC specific genes

Gene	baseMean	log2FoldChange	Fold Change	p value	p adj
C2	18 052	-4,944	0,032	0	0
Gpt2	35 274	-5,798	0,018	0	0
Amy1	3 372	-6,032	0,015	0	0
Bhmt	20 022	-9,792	0,001	0	0
Hsd17b3	18 581	-9,798	0,001	1,74E-281	1,39E-278
Aqp2	2 982	-7,231	0,007	4,42E-245	2,15E-242
Prps2	4 552	-3,749	0,074	2,18E-209	8,13E-207
Ces1d	11 984	-5,741	0,019	2,83E-205	9,87E-203
Glb1l2	4 513	-7,003	0,008	2,53E-199	8,47E-197
Lgals3bp	20 133	-6,373	0,012	1,14E-198	3,73E-196
Serpina3c	3 338	-4,255	0,052	1,48E-188	4,40E-186
Gpd1	3 897	-6,797	0,009	1,39E-179	3,79E-177
Stxbp3a	4 244	-3,010	0,124	3,20E-162	6,75E-160
Klf15	2 929	-3,176	0,111	9,87E-156	1,69E-153
Figf	3 069	-4,574	0,042	7,10E-154	1,18E-151
Sdsl	1 122	-4,313	0,050	2,78E-150	4,37E-148
Acox3	8 398	-3,121	0,115	4,46E-131	5,19E-129
Snx10	2 540	-6,134	0,014	1,75E-126	1,90E-124
Il17rb	1 514	-3,448	0,092	3,59E-122	3,74E-120
Kcnk1	2 467	-3,389	0,095	3,32E-114	3,04E-112
Sult1e1	2 879	-10,672	0,001	1,02E-108	8,80E-107
Vcam1	46 596	-5,777	0,018	3,45E-107	2,87E-105
Slc7a11	4 286	-5,998	0,016	1,12E-104	8,95E-103

Serpina3n	2 169	-10,261	0,001	6,47E-87	3,75E-85
Hsd3b6	32 837	-11,130	0,000	8,96E-78	4,48E-76
Klk1b24	3 178	-11,822	0,000	1,09E-75	5,12E-74
Gm11744	1 030	-3,358	0,098	2,68E-74	1,21E-72
Hpgd	1 068	-4,030	0,061	2,07E-61	6,54E-60
Rhbg	2 881	-12,255	0,000	2,48E-60	7,62E-59
Serpina3g	462	-5,205	0,027	1,89E-58	5,51E-57
Lbp	3 087	-8,037	0,004	9,19E-54	2,39E-52
Adhfe1	1 626	-2,181	0,220	1,71E-49	3,93E-48
Spon1	8 151	-9,477	0,001	4,49E-47	9,60E-46
Paqr6	391	-5,686	0,019	6,54E-47	1,39E-45
Dcxr	1 124	-3,278	0,103	8,51E-46	1,77E-44
Nat8	725	-4,189	0,055	7,05E-42	1,28E-40
Klk1b21	5 524	-11,869	0,000	2,96E-40	5,07E-39
Rarres1	2 641	-8,946	0,002	2,68E-39	4,42E-38
Retsat	14 826	-3,468	0,090	9,21E-37	1,40E-35
Bpifb5	1 962	-4,132	0,057	1,77E-35	2,60E-34
Cml1	533	-2,338	0,198	2,78E-34	3,89E-33
Adck3	7 476	-2,174	0,222	3,28E-32	4,23E-31
Lrg1	1 040	-11,788	0,000	5,26E-30	6,29E-29
Cml5	687	-3,956	0,064	5,53E-29	6,33E-28
Eppin	472	-12,500	0,000	1,31E-25	1,33E-24
Klk1b27	3 990	-13,693	0,000	3,32E-25	3,31E-24
Espn	1 912	-1,994	0,251	5,15E-23	4,63E-22
C1rl	902	-2,248	0,211	7,72E-22	6,55E-21
Asgr1	79	-6,078	0,015	3,07E-20	2,39E-19
Serpina5	22 933	-1,330	0,398	0,004924	0,008635

823

Figure legends

Figure 1: Similarities and differences between FLCs and ALCs

(A) Sample collection, RNA-sequencing and analysis workflow. **(B)** Venn diagram showing the number of genes expressed at same levels in FLCs and ALCs (12,645 genes), enriched in FLCs (3,741 genes) and enriched in ALCs (4,657 genes). **(C)** Volcano plot displaying differential expressed genes between FLCs and ALCs. The x-axis corresponds to the log2(Fold Change) and the y-axis corresponds to the -log10(adjusted p-value). The salmon dots represent the significant up-regulated transcripts in FLCs (adj. pval<0.01); the blue dots represent the significant up-regulated transcripts in ALCs (adj. pval<0.01) and the grey dots represent the the not significantly differentially expressed transcripts (adj. pval>0.01). Top 30 genes according to the adjusted p-value are displaying on this volcano plot. **(D)** Heatmap displaying the normalized scaled expression of differentially expressed genes in the 6 samples. Genes are ordered with hierarchical clustering according to their expression pattern into 15 groups.

Figure 2: Alternative splicing

(A) Proportion of the differential alternative splicing events found in fetal and adult Leydig cells. **(B**, **C** & **D)** GO term enrichment of the biological functions in genes showing differential skipping exons in fetal and adult Leydig cells **(B, C)**, and intron retention in fetal Leydig cells (**D**).

Figure 3: Classification of cell populations in single-cell RNA sequencing data of fetal and adult mouse testis

(A) UMAP representation of single-cell transcriptomes colored by cell type, where each dot corresponds to one cell. Enlargement of the global UMAP on the area of interest which include both the fetal and adult Leydig, colored by cell populations **(A')** and by developmental stage **(C)**. The fetal cells are colored in pale green and the adult cells in dark green. **(D, E)** Enlargement on the UMAP representation colored according to the normalized expression of Leydig cells marker genes *Hsd3b1* **(D)** and *Cyp11a1* **(E)**. **(B)** Dotplot displaying the expression of selected marker genes of testis cell populations. The size of the dots is proportional to the fraction of cells in the population expressing the gene and the scaled gene expression level is indicated by the color scale. Leydig cells: *Nr5a1, Star, Cyp11a1*. Sertoli cells: *Hsd17b3, Amhg, Sox9*. Interstitial progenitors:

Arx, Lhx9, Nr2f2. Endothelial cells: *Pecam1, Cdh5, Esam*. Germ cells: *Pou5f1, Ddx4*. Spermatogonia: *Dmrt1*. Spermatocytes: *Piwil1*. Round spermatids: *Tex21*. Elongating spermatids: *Tnp1*.

Figure 4: Identification of genes labelling FLCs

(A, A') UMAP representation of single-cell transcriptomes highligting the two Leydig cell populations: fetal Leydig cells (FLCs) in pink and adult Leydig cells (ALCs) in grey pink. The other populations of the testis are colored in grey. **(C-G)** Enlargement on the UMAP representation colored according to the normalized expression of selected FLC-specific candidate marker genes, like *Crhr1* **(C)**, *Ren1* **(D)**, *Vsnl1* **(E)**, *Cyp26b1* **(F)**, and *Mc2r* **(G)**. **(B)** Dotplot displaying the scaled expression of the FLC-specfic candidate genes, and the top 15 non-specific candidates discriminating FLCs over ALCs. **(H-K)** *In situ* hybridization assay by RNAscope® labelling in red *Star*, a known marker of Leydig cells, and in green the FLC marker *Ren1* **(H, I)** and *Crhr1* **(J, K)** in embryonic mouse testis at E18.5 **(H, J)** and in adult mouse testis at P100 **(I, K)**. The colocalisation of the red and green probes indicate a co-expression of the two genes in the Leydig cells. The black scale bar corresponds to 100µm.

Figure 5: Identification of genes labelling ALCs

(A, A') UMAP representation of single-cell transcriptomes highligting the two Leydig cell populations: fetal Leydig cells (FLCs) in pink and adult Leydig cells (ALCs) in pale violet. The other populations of the testis are colored in grey. **(C-G)** Enlargement on the UMAP representation colored according to the normalized expression of selected ALC-specific candidate marker genes, like *Bhmt* **(C)**, *Sult1e1* **(D)**, *Lrg1* **(E)**, *Serpina3n* **(F)**, and *Klk1b27* **(G)**. **(B)** Dotplot displaying the scaled expression of the top 25 ALC-specific candidate genes. **(H-K)** *In situ* hybridization assay by RNAscope® labelling in red *Star*, a known marker of Leydig cells, and in green the ALC marker *Bhmt* **(H, I)** and *Sult1e1* **(J, K)** in adult mouse testis at P100 **(H, J)** and in embryonic mouse testis at E18.5 **(I, K)**. The colocalisation of the red and green probes indicate a co-expression of the two genes in the Leydig cells. The black scale bar corresponds to 100µm.

Figure 6: Dual and sequential model of steroidogenesis regulation

Drawing recapitulating the similarities and differences between FLCs (left panel) and ALCs (right

panel). The expression of the ACTH receptor (encoded by *Mc2r*) and the CRH receptor (encoded by *Crhr1*) in FLCs suggests that the regulation of androgen production is mediated by LH, CRH and CRH, while in the absence of CRH and ACTH receptors in ALCs the regulation is exclusively under the control of LH.

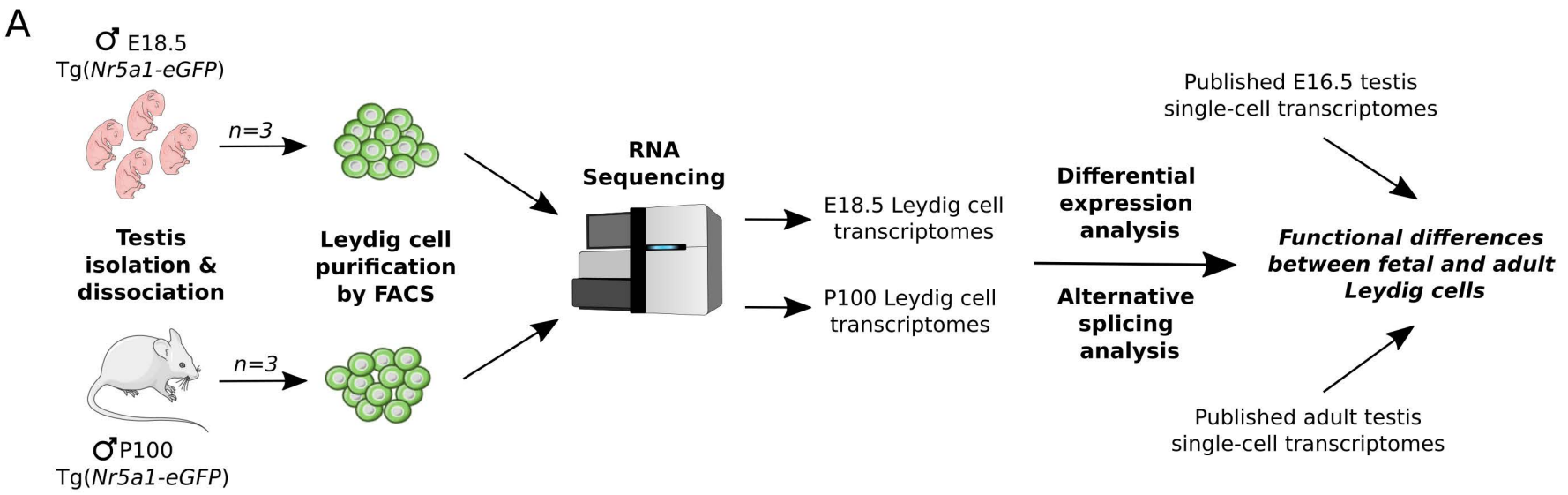
Supplementary Figure 1:

(A,A') Scatter plot showing the selection of Leydig cells (GFP+++) by FACS in E18.5 **(A)** and in P100 **(A')** testis cell suspension. The X axis corresponds to the GFP florescence level and the Y axis forward scatter area explaining the granularity of the events. **(B)** Heatmap showing the normalized scaled expression of selected marker genes in our three replicates at E18.5 (pale green) and in our three replicates at P100 (dark green). The color is representative of the expression level. The gene expression confirmed the purity of the samples in Leydig cells. *Cyp11a1, Star, Nr5a1, Hsd3b1, Insl3, Cyp17a1, Hsd17b3, eGFP*: Leydig cells. *Cdh5, Pecam1, Acta2, Rgs5*: Endothelial cells. *Pdgfra, Tcf21, Wnt5a, Nr2f2*: Interstitial progenitors. *Amh, Lhx9, Dhh, Sox9*: Sertoli cells. *Pou5f1, Mael, Dadx4, Dazl*: Germ cells. **(C)** Heatmap showing the Spearman correlation score between the six Leydig cells samples. The score is indicated by the color scale. **(D)** Principal Component Analysis (PCA) plot where each dot corresponds to a Leydig cells sample. The dots are colored according to the sample type. E18.5: pale green; P100: dark green.

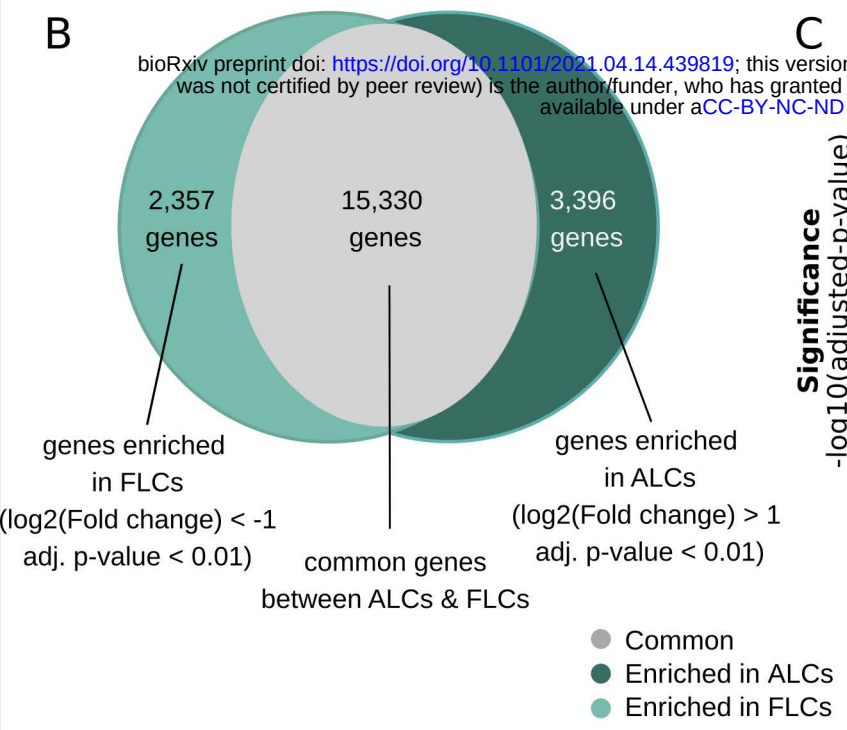
Supplementary Figure 2:

Immuno-histochemistry (IHC) staining on mouse fetal and adult testis. The DNA is colored in white, HSD3B labels Leydig cells in green and BHMT is colored in red. The white scale bar corresponds to 50µm.

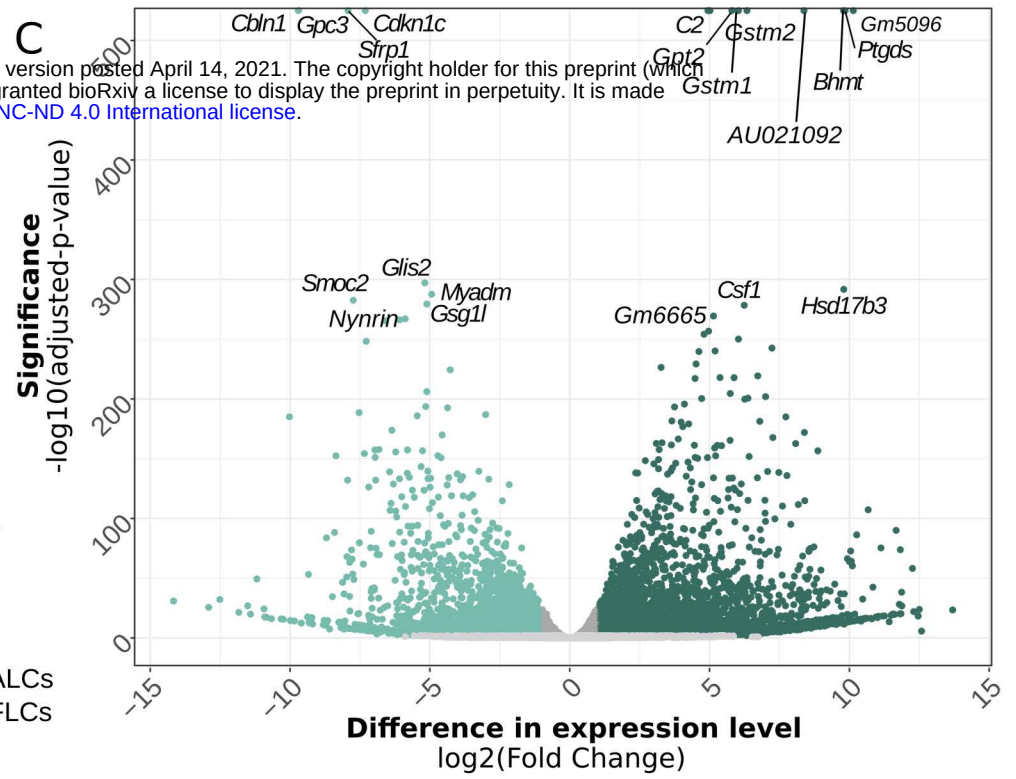
A



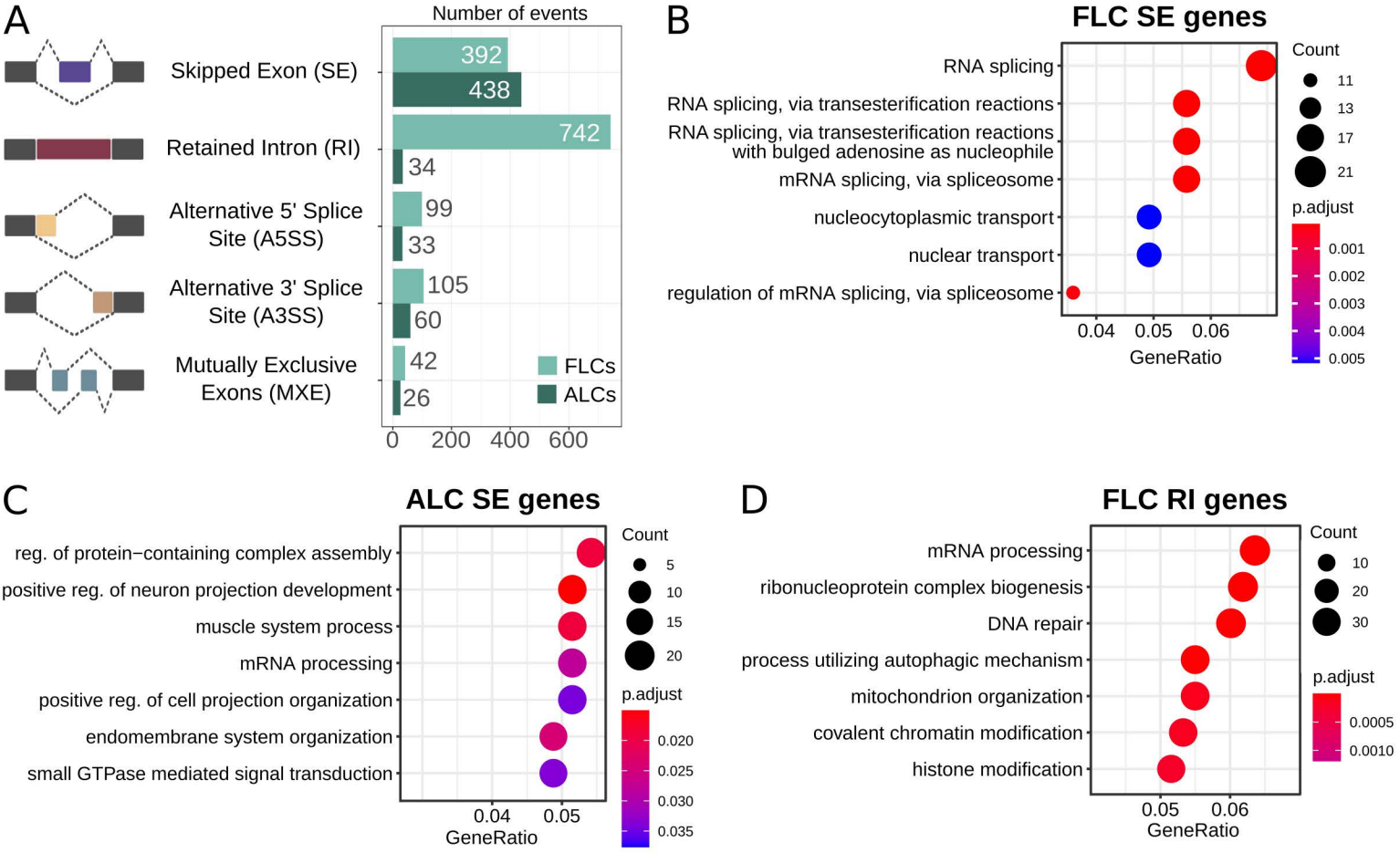
B

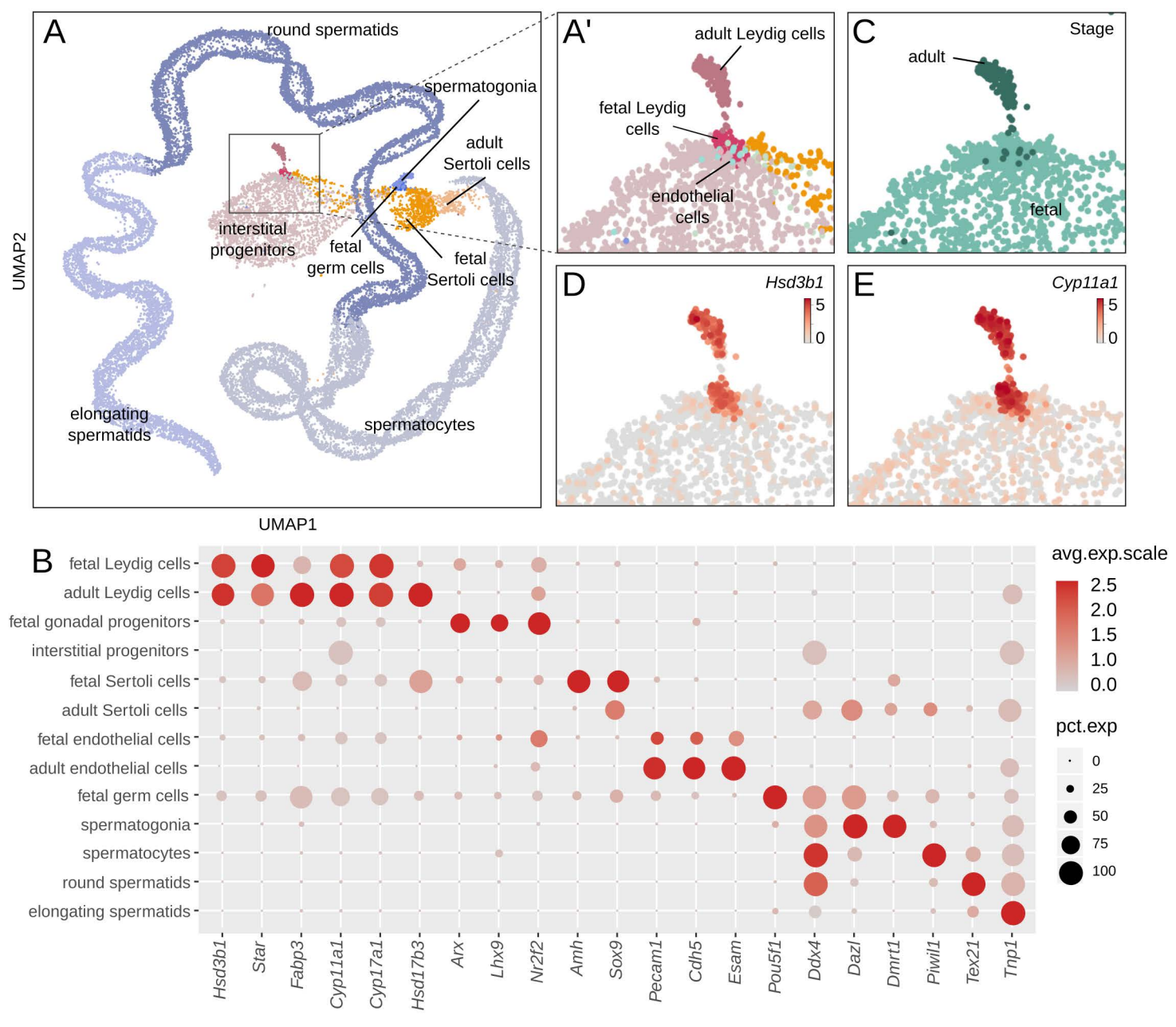


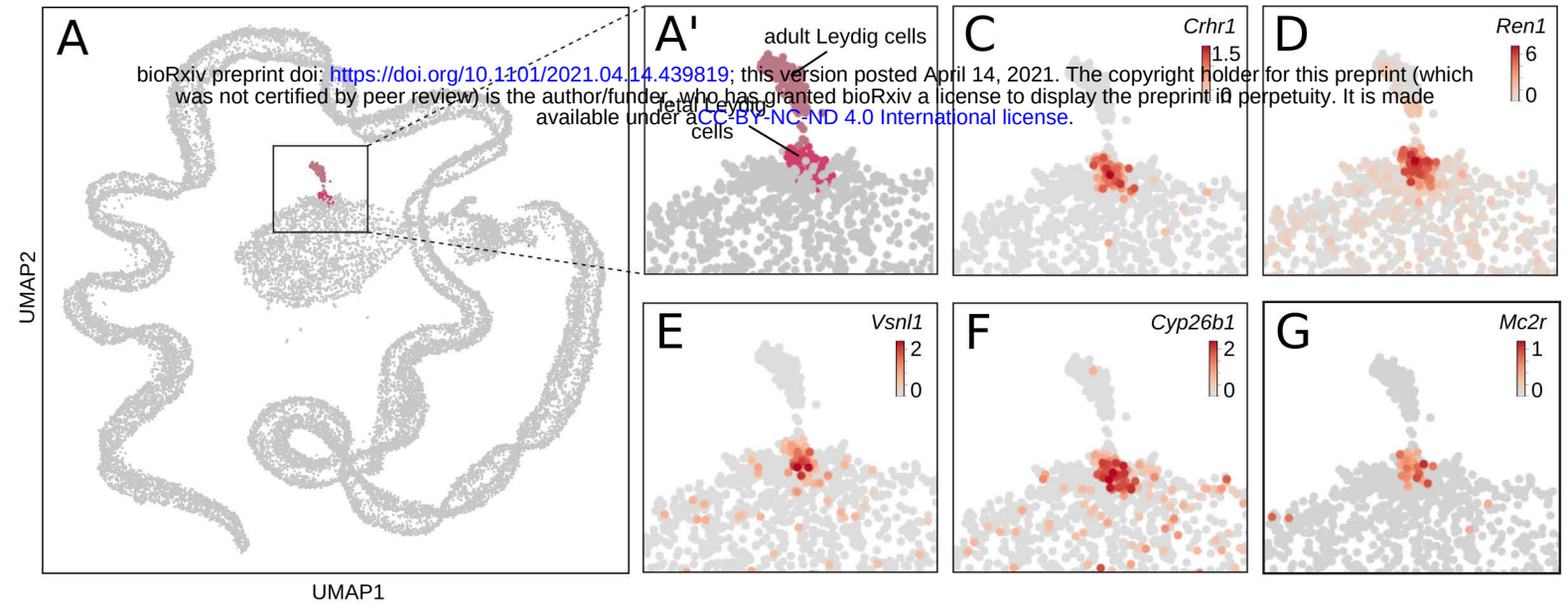
C

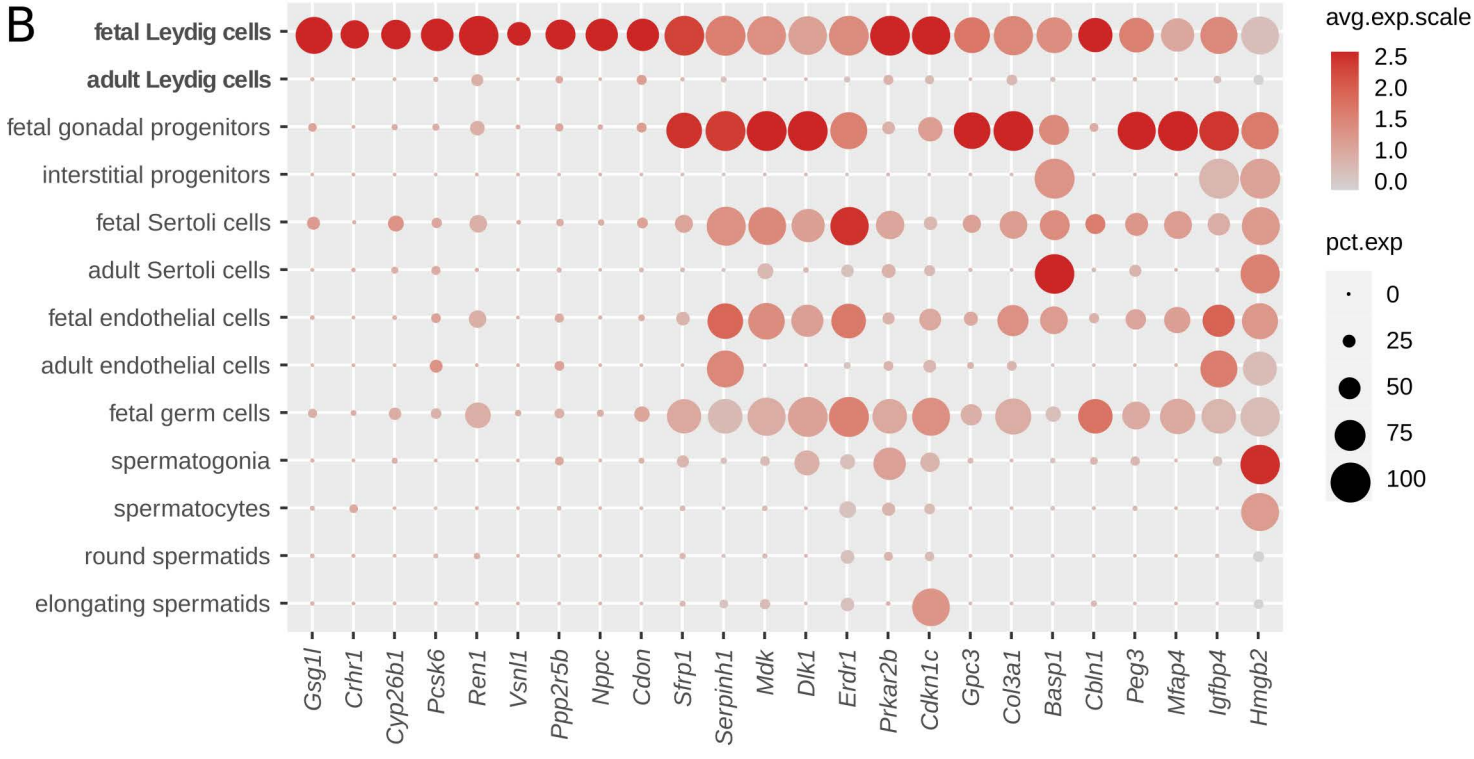


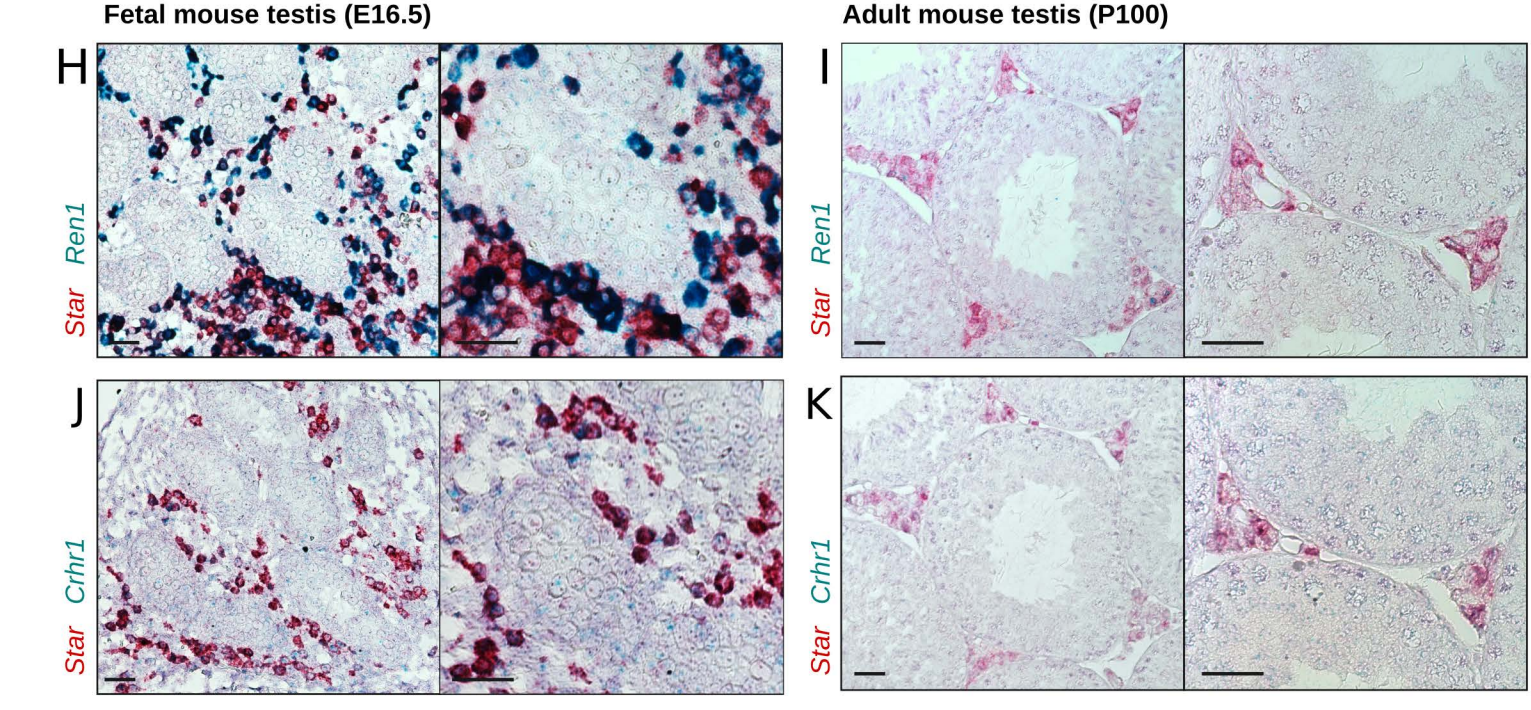
D

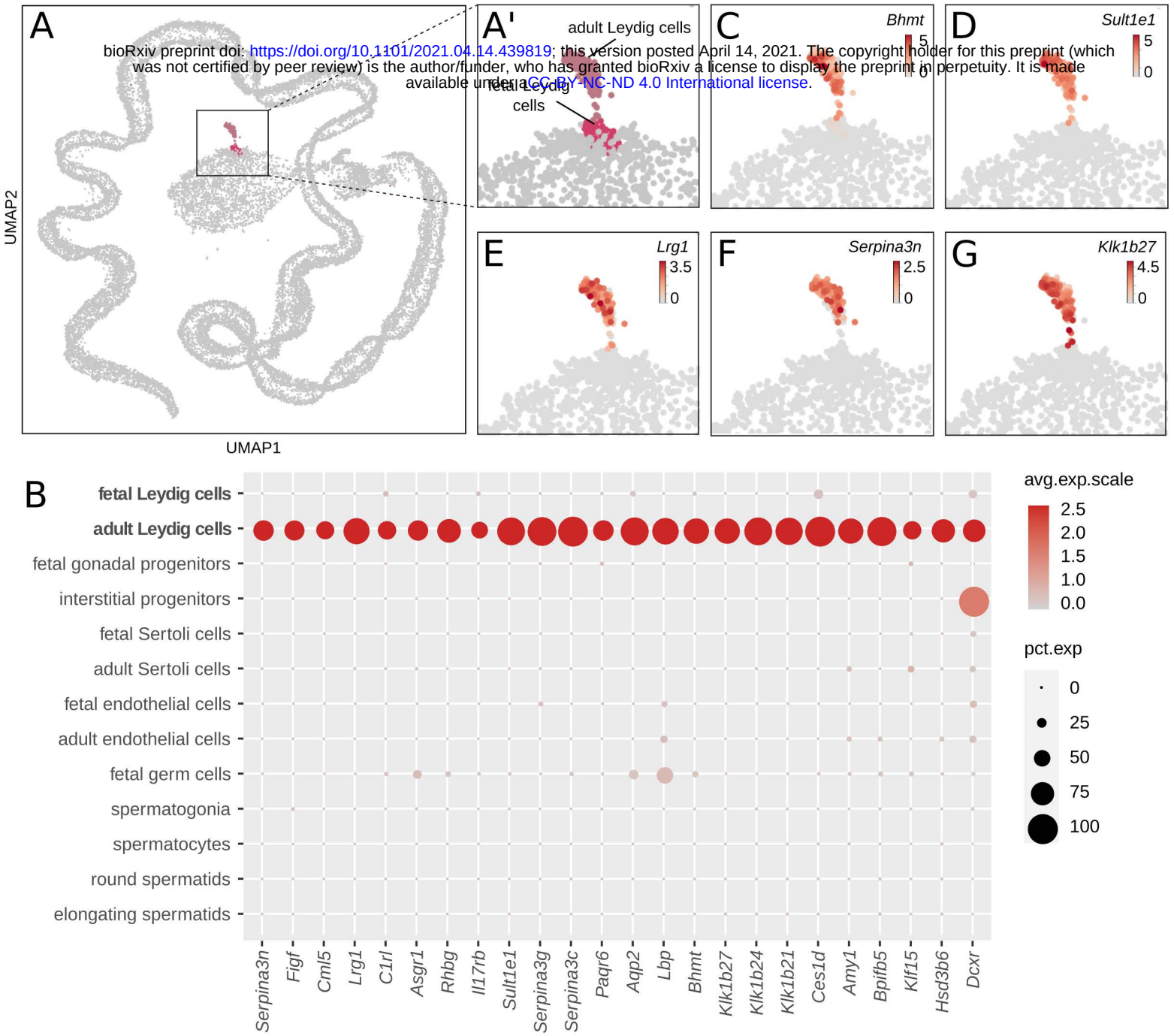

A
Number of events
Skipped Exon (SE)
392
438
Retained Intron (RI)
742
34
Alternative 5' Splice Site (A5SS)
99
33
Alternative 3' Splice Site (A3SS)
105
60
Mutually Exclusive Exons (MXE)
42
26
FLCs
ALCs
0
200
400
600
B
FLC SE genes
RNA splicing
RNA splicing, via transesterification reactions
RNA splicing, via transesterification reactions with bulged adenosine as nucleophile
mRNA splicing, via spliceosome
nucleocytoplasmic transport
nuclear transport
regulation of mRNA splicing, via spliceosome
0.04
0.05
0.06
GeneRatio
Count
11
13
17
21
p.adjust
0.001
0.002
0.003
0.004
0.005
C
ALC SE genes
reg. of protein−containing complex assembly
positive reg. of neuron projection development
muscle system process
mRNA processing
positive reg. of cell projection organization
endomembrane system organization
small GTPase mediated signal transduction
0.04
0.05
GeneRatio
Count
5
10
15
20
p.adjust
0.020
0.025
0.030
0.035
D
FLC RI genes
mRNA processing
ribonucleoprotein complex biogenesis
DNA repair
process utilizing autophagic mechanism
mitochondrion organization
covalent chromatin modification
histone modification
0.05
0.06
GeneRatio
Count
10
20
30
p.adjust
0.0005
0.0010

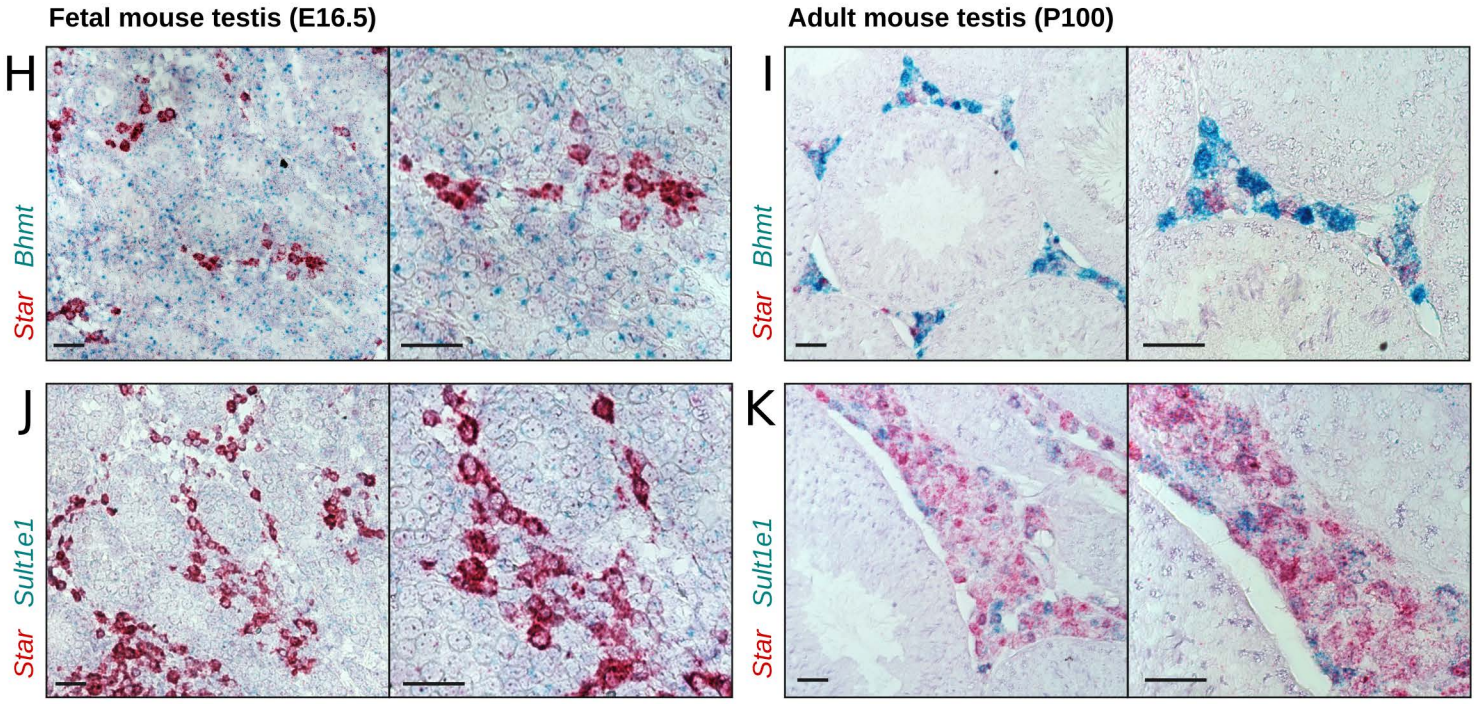

A
round spermatids
spermatogonia
adult Sertoli cells
interstital progenitors
fetal germ cells
fetal Sertoli cells
elongating spermatids
spermatocytes
UMAP2
UMAP1
A'
adult Leydig cells
fetal Leydig cells
endothelial cells
C
Stage
adult
fetal
D
Hsd3b1
5
0
E
Cyp11a1
5
0
B
fetal Leydig cells
adult Leydig cells
fetal gonadal progenitors
interstitial progenitors
fetal Sertoli cells
adult Sertoli cells
fetal endothelial cells
adult endothelial cells
fetal germ cells
spermatogonia
spermatocytes
round spermatids
elongating spermatids
Hsd3b1
Star
Fabp3
Cyp11a1
Cyp17a1
Hsd17b3
Arx
Lhx9
Nr2f2
Amh
Sox9
Pecam1
Cdh5
Esam
Pou5f1
Ddx4
Dazl
Dmrt1
Piwil1
Tex21
Tnp1
avg.exp.scale
2.5
2.0
1.5
1.0
0.0
pct.exp
0
25
50
75
100

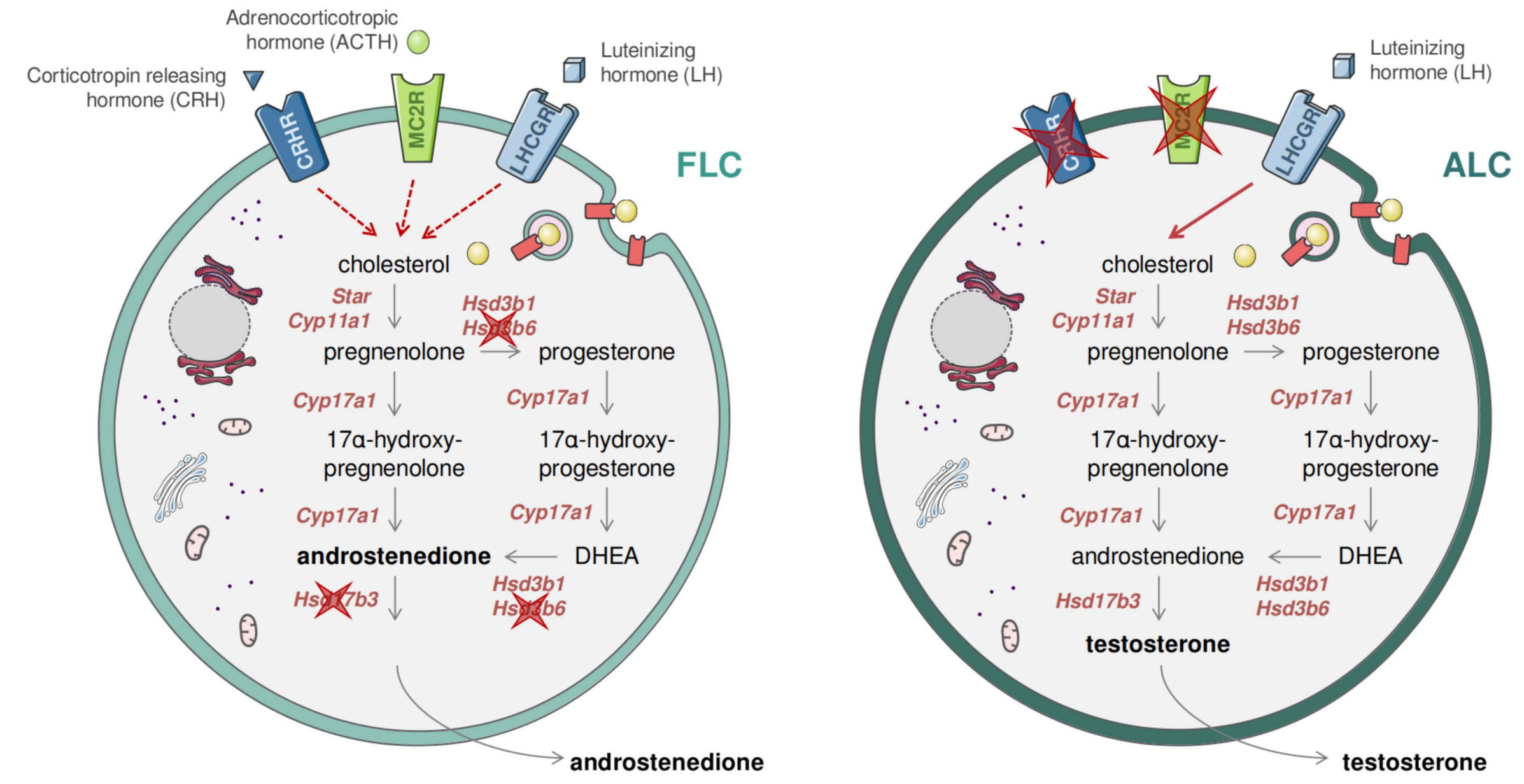


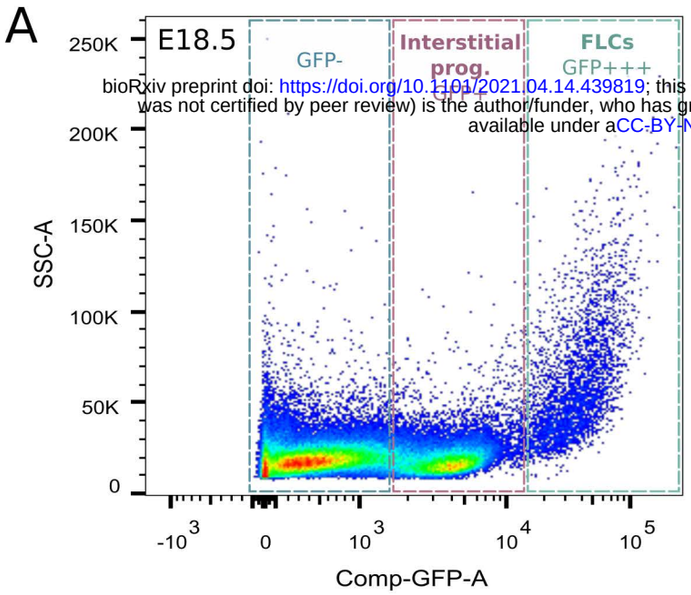


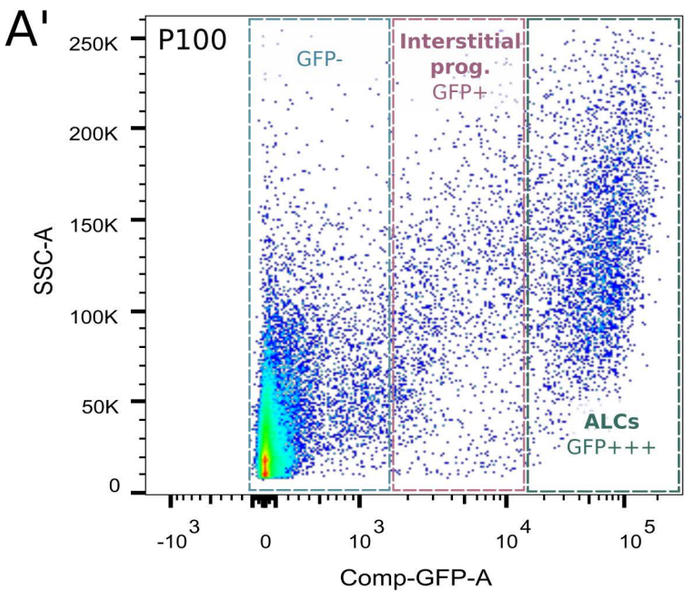


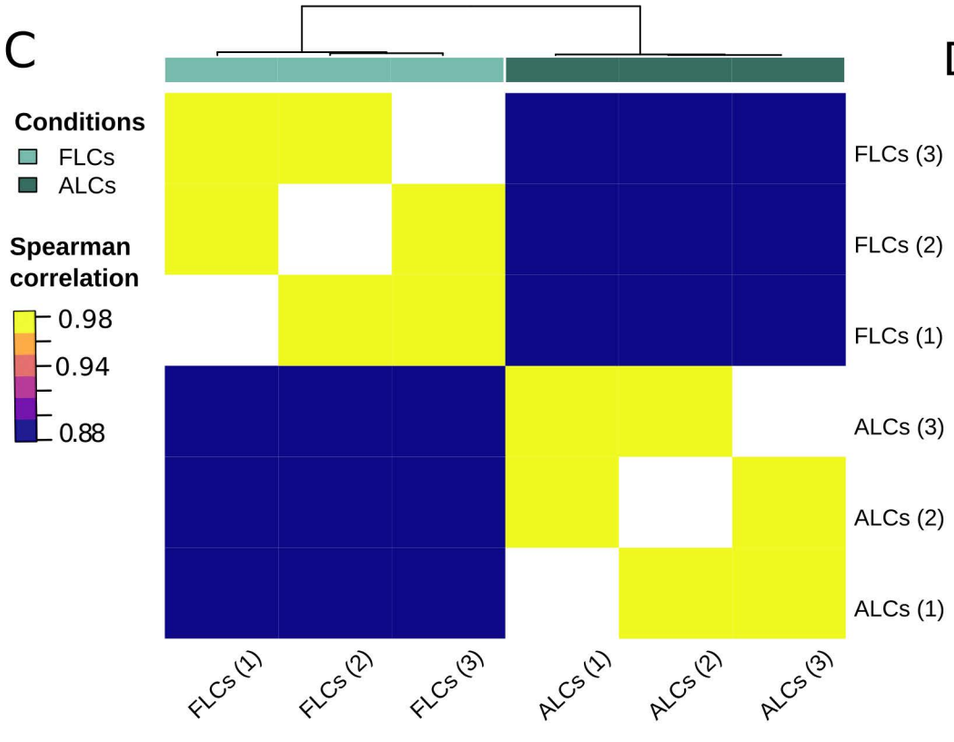


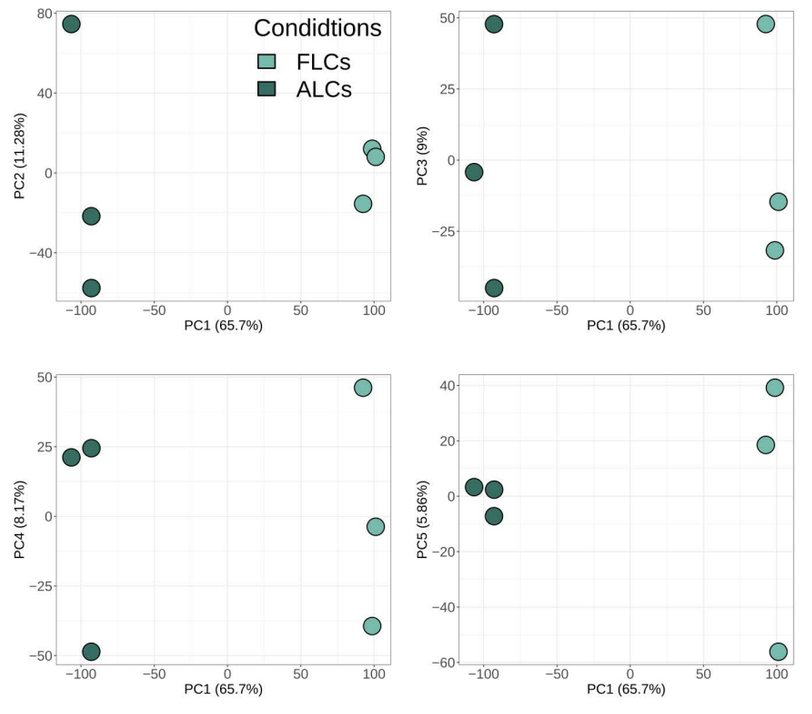


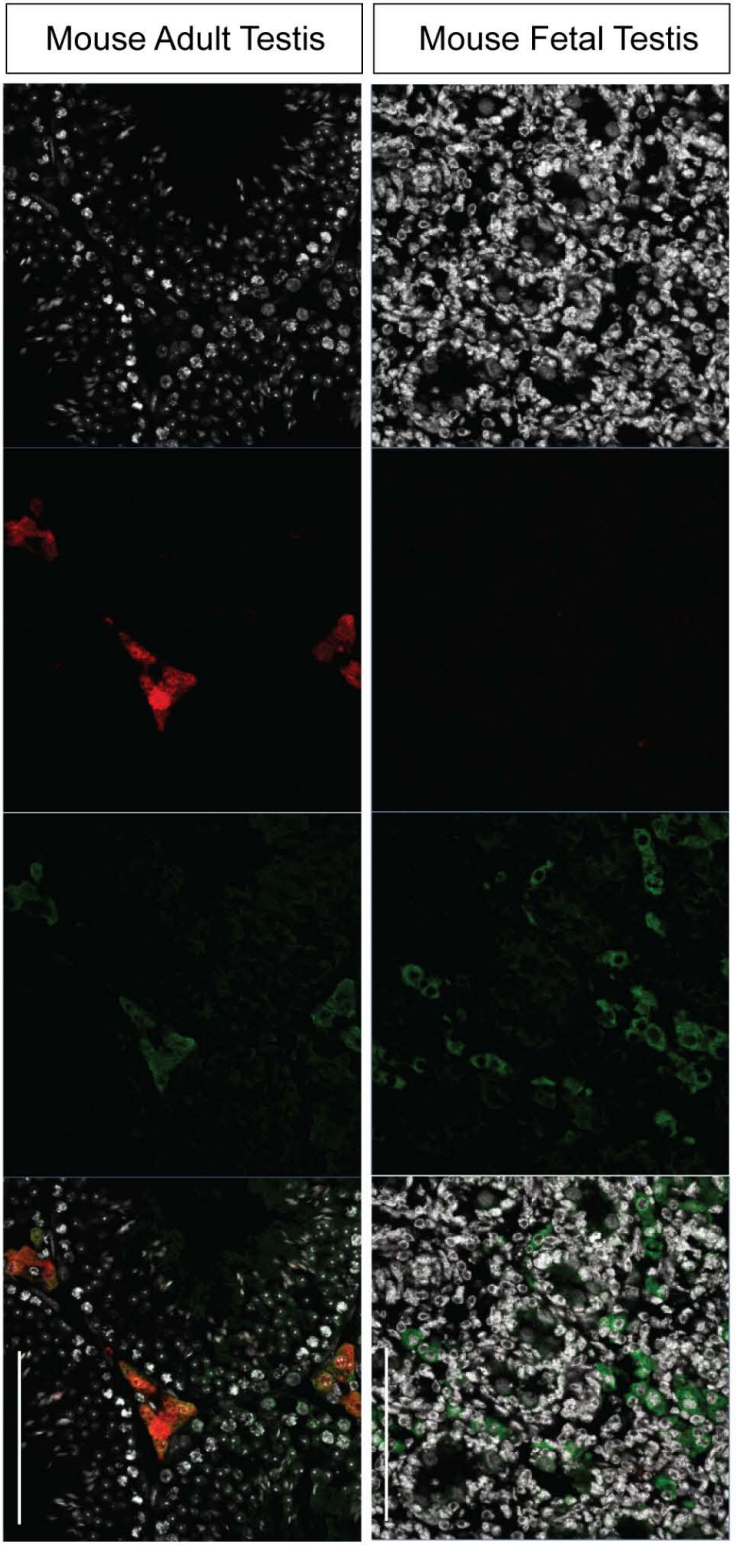

Adrenocorticotropic hormone (ACTH)
Corticotropin releasing hormone (CRH)
Luteinizing hormone (LH)
CRHR
MC2R
LHCGR
FLC
cholesterol
Star
Cyp11a1
pregnenolone
Hsd3b1
Hsd3b6
progesterone
Cyp17a1
Cyp17a1
17α-hydroxy-pregnenolone
17α-hydroxy-progesterone
Cyp17a1
Cyp17a1
androstenedione
DHEA
Hsd17b3
Hsd3b1
Hsd3b6
androstenedione
Luteinizing hormone (LH)
CRHR
MC2R
LHCGR
ALC
cholesterol
Star
Cyp11a1
pregnenolone
Hsd3b1
Hsd3b6
progesterone
Cyp17a1
Cyp17a1
17α-hydroxy-pregnenolone
17α-hydroxy-progesterone
Cyp17a1
Cyp17a1
androstenedione
DHEA
Hsd17b3
Hsd3b1
Hsd3b6
testosterone
testosterone










Mouse Adult Testis
Mouse Fetal Testis
HSD3B BHMT DNA