bioRxiv preprint doi: https://doi.org/10.1101/747279; this version posted September 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Single-cell transcriptomic reveals temporal dynamics of critical regulators of germ cell fate during mouse sex determination

Authors: Chloé Mayère[1,2], Yasmine Neirijnck[1], Pauline Sararols[1], Isabelle Stévant[1,2], Françoise Kühne[1], Anne Amandine Chassot[3], Marie-Christine Chaboissier[3], Emmanouil T. Dermitzakis[1,2], Serge Nef[1,2,*].

Affiliations:

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

[2]iGE3, Institute of Genetics and Genomics of Geneva, University of Geneva, 1211 Geneva, Switzerland;

[3]Université Côte d'Azur, CNRS, Inserm, iBV, France;

Lead Contact:

[*]Corresponding Author: Serge.Nef@unige.ch

bioRxiv preprint doi: https://doi.org/10.1101/747279; this version posted September 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Summary

Despite the importance of germ cell differentiation for sexual reproduction, gene networks underlying their fate remain unclear. Here, we describe a comprehensive characterization of gene expression dynamics during sex determination based on single-cell RNA sequencing on 14,750 XX and XY mouse germ cells between embryonic days 10.5 and 16.5. By computational gene regulation inference analysis, we identified sex-specific, sequential waves of master regulator genes during germ cells differentiation and unveiled that the meiotic initiator *Stra8* is regulated by positive and negative master regulators acting in an antagonistic fashion. Consistent with the importance of the somatic environment, we found that rare adrenal germ cells exhibit delayed meiosis and display altered expression of master genes controlling the female and male genetic programs. Our study provides a molecular roadmap of germ cell sex determination at single-cell resolution that will serve as a valuable resource for future studies of gonad development, function and disease.

Keywords:

Single-cell RNA-Sequencing (scRNA-seq), sex determination, ovary, testis, gonocytes, oocytes, prospermatogonia, meiosis, gene regulatory network, regulon

bioRxiv preprint doi: https://doi.org/10.1101/747279; this version posted September 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Introduction

In mice, primordial germ cells (PGCs) arise in the posterior proximal epiblast around embryonic day (E) 6.25. PGCs rapidly proliferate and colonize the gonads at around E10.5 (Saitou and Yamaji, 2012; Tam and Snow, 1981). The germ cell fate depends on sex-specific somatic cues provided by the ovarian and testicular environments rather than the chromosomal sex of the germ cells themselves (Byskov and Saxen, 1976; Evans et al., 1977; McLaren, 1983).

In fetal ovaries, germ cells enter meiosis asynchronously in a wave from anterior to posterior over about 2 days, between E12.5 and E14.5 (Bullejos and Koopman, 2004; Menke et al., 2003; Yao et al., 2003). This entry into meiosis is considered a hallmark of commitment to oogenesis. It is triggered by the expression of the pre-meiotic marker *Stra8* and the meiosis-associated gene *Rec8* and, at the same time, by the downregulation of pluripotency markers such as *Oct4* (*Pou5f1*), *Sox2* and *Nanog* (Baltus et al., 2006; Bowles et al., 2006; Bullejos and Koopman, 2004; Koubova et al., 2014; Koubova et al., 2006; Menke et al., 2003; Yao et al., 2003).

In contrast, germ cells in fetal testes differentiate into prospermatogonia through a number of distinct, potentially interrelated events that occur asynchronously over a period of several days, but this does not involve entry into meiosis (for a review see (Kocer et al., 2009) and (Spiller and Bowles, 2019)). Germ cells adopting the male fate up-regulate cell-cycle inhibitors such as *Cdkn2b* (Western et al., 2008) and are mitotically arrested from E12.5 onward (McLaren, 1984). They transiently activate the NODAL/CRIPTO signaling pathway (Souquet et al., 2012; Spiller et al., 2012a; Spiller et al., 2012b) and down-regulate pluripotency genes such as *Nanog*, *Sox2* and *Pou5f1* (Western et al.,

bioRxiv preprint doi: https://doi.org/10.1101/747279; this version posted September 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

2010). From E13.5 onward, they begin to express male-specific genes including Nanos homolog 2 (*Nanos2*) (Suzuki and Saga, 2008), *Dnmt3l* (La Salle et al., 2004) and *Piwil4* (Aravin et al., 2008) , which ensure normal male gametogenesis by regulating spermatogonial stem cell properties.

Although the cellular origins of oogonia and spermatogonia are well documented, numerous open questions related to the molecular mechanisms underlying their differentiation and cell fate remain. For instance, the transcriptional programs mediating germ cell sex determination are incompletely understood, and the essential genes and transcriptional regulators orchestrating such specifications remain poorly defined. In developing ovaries, the factors regulating *Stra8* expression are still questioned, and in developing testes it is unclear how the different events mediating the commitment and differentiation of germ cells toward spermatogenesis are initiated and coordinated. Finally, our understanding of how the somatic environment, whether gonadal or extra-gonadal, directs the transcriptional cascade mediating entry into meiosis and the commitment to oogenesis is still unclear.

To date, most transcriptional analyses relevant for mouse or human germ cell sex determination have been conducted using traditional methods such as microarrays or bulk RNA-seq on either whole gonads or isolated germ cell populations at few selected time points (Gkountela et al., 2015; Guo et al., 2015; Houmard et al., 2009; Irie et al., 2015; Jameson et al., 2012; Lesch et al., 2013; Molyneaux et al., 2004; Rolland et al., 2008; Rolland et al., 2011; Small et al., 2005; Soh et al., 2015; Tang et al., 2015). These studies, although informative, provided only an average transcriptional summary, masking the inherent variability of individual cells and lineage types and thereby limiting their

bioRxiv preprint doi: https://doi.org/10.1101/747279; this version posted September 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

capacity to reveal the precise dynamics of gene expression during germ cell sex determination.

To determine the sequence of transcriptional events that is associated with germ cell commitment and differentiation toward oogenesis and spermatogenesis, we performed time-series single-cell RNA sequencing (scRNA-seq) on developing gonads. We recovered 14,750 germ cells from XX and XY gonads across five developmental time points from E10.5 to E16.5, encompassing the entire developmental process of gonadal sex determination and differentiation. We reconstructed the developmental trajectories of male and female germ cells, characterized the associated genetic programs, and predicted gene regulatory networks that regulate germ cell commitment and differentiation. In particular, we found (i) sex-specific, sequential waves of master regulator genes during germ cells differentiation, (ii) the meiotic initiator *Stra8* is regulated by positive and negative master regulators acting in an antagonistic fashion, (iii) mRNA transcription and mRNA splicing are often disconnected, either temporally or in a sex-specific manner, and (iv) ectopic XY adrenal germ cells enter into meiosis with delay together with significant alterations in ovarian-specific genes and upregulation of testis-specific genes.

Results

A single-cell transcriptional atlas of germ cells sex determination and differentiation

To generate a gene expression atlas of germ cells sex determination, we used droplet-based 3’ end scRNA-seq (10x Genomics Chromium) of XX and XY gonads from mouse embryos at five developmental stages (E10.5, E11.5, E12.5, E13.5, and E16.5) (**Fig. 1A**

bioRxiv preprint doi: https://doi.org/10.1101/747279; this version posted September 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

and B). The selected time points cover the entire process of gonadal sex determination and span the emergence and differentiation of the major testicular and ovarian lineages including the gonocytes. For each of the 10 conditions, we generated two independent replicates from different pregnancies and sequenced an average of 10,000 cells. The transcriptomes of the individual cells were sequenced at the depth of ~150,000 reads/cell. Using ten well-established germ cell markers, namely *Ddx4*, *Dazl*, *Mael*, *Dppa3*, *Sycp3*, *Pecam*, *Pou5f1*, *Stra8*, *Dmrt1*, and *Dnmt1*, we identified 14,750 germ cells among a total of 92,267 cells (**Fig. S1A-C**) (see also **Supplementary information**). Among germ cells, 8,248 were XX (55.9%) and 6,502 were XY (44.1%). It included 70 cells from E10.5, 953 cells from E11.5, 4,365 cells from E12.5, 6,593 cells from E13.5 and 2,769 cells from E16.5. As expected, the median count of UMIs and median number of detected genes were higher in germ cells than in somatic cells with 22,584 versus 14,650 and 5,536 and 4,502, respectively (**Fig. S1D**) (Soumillon et al., 2013).

Cell lineage reconstruction identifies the dynamics of gene expression during XX and XY germ cell sex determination

UMAP projection of the 14,750 germ cells (**STAR Methods**) revealed that at early stages (E10.5 and E11.5), the transcriptomes of XX and XY cells globally overlapped, while cells from later stages formed two sex-specific branches (**Fig. 1C** and **D**). We then analyzed how the transcriptomes of XX and XY cells progress during the process of sex determination. To order cells along a pseudotime, we used ordinal regression modeling (Telley et al., 2018; Teo et al., 2010) using prior knowledge about the developmental stage of each cell capture (see **Fig. 1E**, **Fig. S1E**, **STAR Methods** and **Supplementary**

bioRxiv preprint doi: https://doi.org/10.1101/747279; this version posted September 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

information). We then represented the smoothed expression of the 3,013 top weighted genes (**Table S1**) in two distinct ordinal regression models trained on XY and XX cells with a double heatmap in which the center represents the earliest cells (pseudotime 0, E10.5 cells) and the extremities represent the lineage endpoints (pseudotime 100, E16.5) of XX germ cells (left) and XY germ cells (right), respectively (see **Supplementary information**). The heatmap revealed that XX and XY germ cells diverged as early as E12.5, exhibiting dynamic and sex-specific differentiation programs mediated by thousands of genes (**Fig. S2**). In addition, using 67 well-known genes involved in mouse germ cell pluripotency, sexual development and differentiation (Hill et al., 2018; Spiller and Bowles, 2019), we confirmed that our single-cell data accurately recapitulated male and female germ cell specification and was consistent with available literature (**Fig. 1F**).

Reconstructing Gene Regulatory Networks mediating germ cell sex determination

The identity and transcriptional state of a cell emerges from an underlying gene regulatory network (GRN or regulome) in which the activities of a small number of key transcription factors and co-factors regulate each other and their downstream target genes (Aibar et al., 2017). To comprehensively infer the GRNs at play during XX and XY germ cell sex determination, we applied the pySCENIC pipeline (Aibar et al., 2017) to our single-cell data. In brief, SCENIC links cis-regulatory sequence information together with scRNA-seq data. It first catalogs coexpression modules between transcription factors and candidate target genes and then identifies modules for which the regulator's binding motif is significantly enriched across target genes; it then creates regulons with only direct target genes. Finally, the activity of each regulon is attributed to each cell, using

bioRxiv preprint doi: https://doi.org/10.1101/747279; this version posted September 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

the AUCell method (Aibar et al., 2017). For germ cells, we identified 837 regulons (512 positive and 325 negative regulons) containing 13,381 genes (**Table S2**). These genes represented 62% of the total number of genes (21,553) detected in germ cells, indicating that our GRN analysis was comprehensive, covering the majority of the germ cell transcriptome. The size of each regulon varied from 2 to 2,625 genes, with a median size of 19 genes. To compare how XX and XY germ cells acquire their sex-specific identity, we selected the 394 positive regulons with AUCell-determined activity in more than 1,000 cells (activity score >0) and classified them according to their expression pattern into 30 profiles or modules (M1 to M30) along pseudotime (**Fig. 2**). We represented the smoothed regulon expression level of XX and XY germ cells with a double heatmap, in which the center represents the starting point of the lineage (pseudotime 0, E10.5) and the extremities represent the lineage endpoints (pseudotime 100, E16.5) of the XX germ cells (left) and the XY germ cells (right), respectively (**Fig. 2**). Strikingly, the expression patterns revealed numerous transient, sex-specific regulon profiles, mostly at late developmental stages (late E13.5 and E16.5). Initially, two modules common to both XX and XY gonocytes (M20 and M17) were observed at early developmental stages, which were superseded sequentially by a handful of transient and overlapping sex-specific regulon modules (M10, M16 and M29 in XX germ cells, and M12, M5 and M14 in XY germ cells). By late E13.5 and E16.5, we noted numerous oogonia-specific (M1, M30, M21, M28, M6, and M18) and spermatogonia-specific modules (M11, M24, M9, M22, M4, M3, M2 and M23).

We also selected 303 negative regulons with AUCell-determined activity in more than 1000 cells (activity score >0) and classified them according to their expression pattern

bioRxiv preprint doi: https://doi.org/10.1101/747279; this version posted September 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

into 10 modules (M1 to M10) (**Fig. S3**). Negative regulons all displayed activity in opposition to their repressing master regulator (data not shown) and most of them showed a pattern with high expression at early times points, and a progressive downregulation/repression over time (M6, M1, M7 and M2 containing 264 regulons). Among them, regulons of M1 and M7 could be distinguished by the onset of their sex-specific repression: regulons from modules 1 and 7 started to be repressed at E16.5 and E13.5 in XY germ cells, respectively, while in XX germ cells this repression was observed at E13.5 and E16.5, respectively. We also detected 4 modules formed by 34 regulons with sequential transient expression (M3, M4, M8 and M10). Among these, target genes belonging to regulons of module 3 were transiently repressed specifically in XY germ cells, in accordance with the male-specific expression of their 12 master regulators (data not shown).

Overall, we identified 697 regulons whose activities were grouped under 40 modules based on their sex-specific and temporal expression. The transient, sequential, and often sex-specific profiles likely represent a sequential/hierarchical organization of regulatory modules required for oogonia and spermatogonia differentiation.

Sex-specific, sequential waves of cell cycle gene expression during germ cells differentiation

Following PGC colonization of the gonad around E10.5, XX and XY gonocytes undergo rapid mitotic proliferation while maintaining pluripotency (McLaren, 2003; Surani et al., 2007). How germ cells exit the rapid proliferative phase and enter into mitotic arrest in testes, or meiosis in ovaries, remains poorly understood. Our scRNA-seq analysis

bioRxiv preprint doi: https://doi.org/10.1101/747279; this version posted September 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

allowed us to comprehensively evaluate the expression of multiple key genes involved in the pluripotency and proliferation of germ cells during their sex-specific differentiation. We observed dynamic regulation and a strong sexual dimorphism among mitotic genes between XX and XY germ cells (**Fig. 3** and **Fig. S4**). As expected, we observed in both XX and XY gonocytes around E11.5 a downregulation of transcription factors mediating pluripotency such as *Nanog*, *Pou5f1* (*Oct4*) as well as other pluripotency-associated genes, including *Dppa2* and *Dppa4* (Maldonado-Saldivia et al., 2007; Pesce and Scholer, 2001; Surani et al., 2007; Western et al., 2005). The profile was similar for genes regulating G1-S phase, including the genes encoding Cyclin A2, B1, D3 (*Ccna2*, *Ccnb1*, *Ccnd3*) as well as E2F transcription factor genes *E2f1*, *E2f2*, *E2f3*, *E2f7*, and *E2f8*. In XY germ cells, consistent with mitotic arrest in G0 between E12.5 and E14.5, we observed an upregulation of cell-cycle inhibitors essential in the control of G1/G0 arrest, including *Cdkn1b* ($p27^{Kip1}$) and *Cdkn2b* ($p15^{INK4b}$) around E13.5, followed by *Cdkn1c* ($p57^{Kip2}$), *Cdkn2a* ($p16^{INK4a}$) and *Cdkn2c* ($p18^{INK4c}$) at E16.5. These results confirmed and extended previous studies (Western et al., 2008).

In XX germ cells, we observed sequential waves of sex-specific upregulation of cell-cycle genes including genes encoding Cyclin H, B2 and G1 (*Ccnh*, *Ccnb2*, *Ccng1*) and the cell cycle inhibitor *Cdkn1a* ($p21^{Cip1}$) around E12.5, cyclins C, E1, E2, D1 (*Ccnc*, *Ccne1*, *Ccnde2*, *Ccnd1)*, E2F transcription factor genes *E2f1*, *E2f2*, and the cell cycle inhibitor genes *Cdkn2a* ($p16^{INK4a}$), *Cdkn2c* ($p18^{INK4c}$), *Cdkn2d* ($p19^{INK4d}$) around E13.5, as well as genes encoding Cyclins B3, C, O, J (*Ccnb3*, *Ccnc*, *Ccno*, *Ccnj)*, and E2F transcription factor genes *E2f4*, *E2f5* at E16.5. Overall, our scRNA-seq analysis

bioRxiv preprint doi: https://doi.org/10.1101/747279; this version posted September 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

accurately revealed the complex and sex-specific expression of cell cycle regulators when germ cells exit the proliferative phase and either arrest mitotically or enter meiosis.

Predicting regulatory factors promoting meiosis and *Stra8* expression

Numerous studies have attempted to identify factors regulating the expression of *Stra8* (Stimulated by Retinoic Acid 8 gene), which triggers the DNA duplication step that precedes meiosis, thus engaging the meiotic program in germ cells (Baltus et al., 2006). Retinoic acid (RA) has been proposed as a meiosis initiating substance (Bowles et al., 2006; Koubova et al., 2006) as it induces *Stra8* mRNA accumulation in RA-treated P19 pluripotent cell lines (Oulad-Abdelghani et al., 1996), but recent lines of evidence indicated that RA signaling is actually dispensable for entry into meiosis, but instead stabilizes *Stra8* expression (Kumar et al., 2011; Vernet et al., 2019). To acquire a better understanding of the signals instructing oogonia to transition from mitosis to meiosis, we took advantage of our GRN analysis and investigated the positive and negative regulons predicted to control the expression of *Stra8* and *Rec8*, a gene that encodes a component of the cohesin complex accumulating during the meiotic S-phase. Consistent with the literature, *Stra8* and *Rec8* were transiently upregulated in XX germ cells between E12.5 and E16.5, coinciding with the entry into meiosis (**Fig. 2B** and **4B**). GRN analysis revealed that both *Stra8* and *Rec8* expressions were predicted to be regulated by a combination of positive and negative master regulators (**Fig. 4A, B** and **Fig. S5**). In particular, *Stra8* was predicted to be negatively regulated by the mitotic cohesin RAD21 and the Y-box binding protein YBX1. Both factors were downregulated specifically in developing XX germ cells, while expression was maintained in XY germ cells (**Fig. 4B**).

bioRxiv preprint doi: https://doi.org/10.1101/747279; this version posted September 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

In parallel, the histone demethylase KDM5A (Jumonji/JARID1) and the transcription factor PBX3, both preferentially expressed in XX over XY germ cells (**Fig. 4B**), were predicted to be key positive regulators of *Stra8* (Ge et al., 2018). KDM5A is also a positive regulator of *Ythdc2*, a gene encoding an RNA helicase that acts as critical regulator of the transition from mitosis to meiosis in both male and female germlines (Bailey et al., 2017; Gonczy et al., 1997). *Rec8* expression was predicted to also be negatively regulated by RAD21 and YBX1, but positively regulated, among others, by KDM5A and MSX1 and MSX2, two nuclear receptors known to promote meiosis initiation by maintaining or enhancing *Stra8* expression (Le Bouffant et al., 2011). EZH2, a member of the polycomb repressive complex 2 (PRC2) (Margueron and Reinberg, 2011), also appeared as a positive regulator of *Rec8* expression that could act in concert with KDM5A to allow control of chromatin opening and the correct timing of the mitosis-to-meiosis transition. Overall, applying single-cell regulatory network inference to germ cell sex determination leads to the prediction of new critical regulators of meiosis as well as *Stra8* and *Rec8* expression.

Variable rates of mRNA splicing in male and female germ cells

mRNA splicing represents another powerful mechanism to modulate gene expression and is known to contribute to the fine-tuning of cell differentiation programs (Kalsotra and Cooper, 2011). To investigate splicing dynamics, we applied RNA velocity analysis to the developing mouse germ cells to estimate the rates of nascent (unspliced) and mature (spliced) transcripts during germ cell differentiation and evaluated whether there were gene- or sex-specific differences in transcriptome kinetics. Velocity analysis considers

bioRxiv preprint doi: https://doi.org/10.1101/747279; this version posted September 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

both spliced and unspliced mRNA counts to predict developmental trajectories and speed of cell state transitions (La Manno et al., 2018). Consistent with our previous analyses, we observed strong directional flow toward the most differentiated germ cells, both in the XX and XY branches (**Fig. S6**). We then ordered the cells according to our ordinal regression model and examined the temporal progression of RNA biogenesis of XX and XY germ cells during the process of sex determination. As expected, unspliced mRNAs consistently preceded spliced mRNAs (**Fig. 5**). We also observed variation in transcriptomic kinetics. For example, among the meiotic-related genes, *Rec8* and *Stra8* exhibited fast kinetics with little differences between unspliced and spliced mRNAs whereas genes such as *Hormad2*, *Msh5*, *Tex11*, and *Spo11* exhibited increasing delays in spliced transcripts (**Fig. 5A**).

We also investigated the rates of mature and immature transcripts of some of the genes previously described as regulators of *Stra8* and *Rec8* expression (**Fig. 5B**) or as being involved in mitosis (**Fig. 5C**). In several cases, we observed a significant disconnection between transcription (i.e. the presence of unspliced mRNA) and the presence of mature (spliced) mRNAs. For example, this was the case for the aforementioned RNA helicase gene *Ythdc2*. *Ythdc2* exhibited increasing levels of unspliced transcripts in both XX and XY cells (dotted lines in **Fig. 5B**, top panel) along pseudotime, but spliced mRNAs were present only in XX germ cells (solid lines in **Fig. 5B**, top panel), suggesting male-specific intronic retention or degradation of *Ythdc2* mRNAs. The Cyclin gene *Ccnb3* also revealed sex-specific differences in gene splicing (**Fig. 5C**, fourth panel). While a transient sequential increase of unspliced and spliced mRNAs was observed in XX germ cells around E16.5, only unspliced mRNAs were observed in XY germ cells at late

bioRxiv preprint doi: https://doi.org/10.1101/747279; this version posted September 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

E16.5. A similar pattern was observed for Cyclin *Ccnd3* (**Fig. 5C**, bottom panel). Overall, we observed gene-specific or sex-specific putative intronic retention in numerous genes, emphasizing once more the importance of post-transcriptional regulation in male and female germ cell lineages.

Ectopic adrenal germ cells also enters into meiosis, but numerous major transcriptional regulators of oocyte differentiation are absent or downregulated

While the majority of PGCs migrate toward the gonadal ridges, a small fraction of germ cells are lost along the way and end up in extragonadal organs such as the nearby adrenal glands and mesonephroi (Heeren et al., 2016; Upadhyay and Zamboni, 1982; Zamboni and Upadhyay, 1983). These adrenal germ cells, irrespective of their genetic sex, have been reported to undergo meiosis, differentiate into oocytes and display morphological characteristics identical to those of young oocytes in primordial follicles before disappearing around 3 weeks of age (Upadhyay and Zamboni, 1982; Zamboni and Upadhyay, 1983). To evaluate how an extragonadal somatic environment affects germ cell fate, we investigated at the transcriptional level how ectopic adrenal germ cells enter into meiosis and commit toward the female fate.

Time-series 3' single-cell RNA sequencing of developing mouse adrenal cells from E12.5, E13.5, E16.5, and E18.5 XY embryos identified 312 adrenal germ cells based on the expression of the classical germ cell markers *Dazl*, *Ddx4* and *Mael* (see **STAR methods**). Overall, we captured 187 adrenal germ cells at E12.5, 92 cells at E13.5, 18 cells at E16.5, and 15 cells at E18.5. The relatively low number of germ cells at later stages may reflect the smaller proportion of germ cells in the growing adrenal glands.

bioRxiv preprint doi: https://doi.org/10.1101/747279; this version posted September 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

UMAP representation of these 312 XY adrenal germ cells combined with 14,718 gonadal germ cells revealed that the transcriptome of XY adrenal germ cells partially overlapped with the transcriptome of XX ovarian germ cells, suggesting that their transcriptomes are similar and that XY adrenal germ cells enter into meiosis and differentiate into oocytes in synchrony with gonadal oocytes (**Fig. 6A**). However, a more refined analysis investigating the expression of selected key marker genes mediating germ cell specification revealed a more complex picture. First, meiosis-related genes, including *Stra8*, *Sycp1*, *Sync3*, *Spo11*, *Ccdc155*, *Dmc1*, *Mei1*, *Mei4*, *Meioc*, *Hormad1*, *Hormad2*, *Msh5*, *Tex11*, *Prdm9* and *Smc1b* displayed similar profiles and expression levels between XY adrenal germ cells and XX ovarian germ cells (**Fig. 6B** and **Fig. S7**). Only a slight temporal delay was observed in adrenal germ cells. One notable exception was *Rec8*, whose expression was blunted in adrenal compared to ovarian germ cells (**Fig. S7**). These results confirmed published data that meiosis is not significantly affected in ectopic adrenal germ cells (Upadhyay and Zamboni, 1982; Zamboni and Upadhyay, 1983). While several genes and master regulators of oogonia differentiation were unaffected (e.g. cyclins *Ccnb3*, *Ccne2*; *Pparg*; the histone demethylases *Kdm5a* and *Phf8*; *Brca2*) (**Fig. S7**), we found numerous key female genes exhibiting either downregulation or a testis-like profile. It included, for example, genes involved in the WNT-β-catenin pathway (*Axin2*, *Lef1*, and *Sp5*), the transcription factor genes *Msx1* and *Msx2*, the cell cycle gene *E2f1*, and the oocyte-specific basic helix-loop-helix transcription factor gene *Figla* (**Fig. 6B** and **Fig. S7**). Finally, we found that most genes and master regulators involved in male germ cell fate were not upregulated in adrenal germ cells with few notable exceptions including the male fate marker *Nanos2*, the NODAL target genes

bioRxiv preprint doi: https://doi.org/10.1101/747279; this version posted September 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Lefty1 and *Lefty2*, the retinoic receptor gene *Rara*, the male germ cell regulator gene *Ctcfl* and the spermatogonial stem cell self-renewal gene *Lhx1* (**Fig. 6B** and **Fig. S7**). Overall, these results indicated that ectopic adrenal germ cells enter into meiosis at around the same time as ovarian germ cells, but numerous genes and master regulators related to both the female and male genetic programs were misregulated.

Discussion

This study represents the first large-scale scRNA-seq analysis of germ cells throughout the gonadal sex determination and differentiation process. The large number of individual germ cells profiled, both XX and XY, allowed us to reconstruct a continuous representation of changes in gene expression occurring during the process of gonadal differentiation, including the mitosis-to-meiosis switch in germ cells in developing ovaries, and spermatogonial commitment and differentiation in fetal testes. This represents a major advance beyond previous work and has broad implications for studies in germ cell development, sex determination and the etiology of human germ cell diseases.

How do germ cells commit to and acquire sex-specific fates during differentiation of the gonad into a testis or an ovary? Our experimental design based on scRNA sequencing to profile five developmental stages encompassing gonadal sex differentiation is well suited to tracing gene regulatory programs in which specific combinations of transcription factors drive sex- and cell type-specific transcriptomes in germ cells. Through computational analyses, we have comprehensively constructed the GRNs for XX and XY

bioRxiv preprint doi: https://doi.org/10.1101/747279; this version posted September 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

germ cells during the process of sex differentiation. We found that the gene regulatory circuitry mediating germ cell sex determination is composed of 512 positive regulons that can be grouped into 30 modules, each of them exhibiting transient, sequential and often sex-specific profiles. The fact that regulons are grouped into modules that display transient sex-specific profiles suggests a sequential organization of regulatory modules that work together to fine-tune the interrelated cellular events that lead to XX and XY germ cell differentiation. The master regulator genes present in each specific module may not be controlling a single cellular event, but instead a combination of overlapping sex-specific developmental processes including mitotic arrest, suppression of pluripotency genes, prospermatogonia commitment and *de novo* methylation for XY germ cells, as well as entry into meiosis, and suppression of pluripotency genes for XX germ cells.

This regulome analysis also provides an opportunity to identify new critical master regulators of germ cell sex determination. While various master regulators have already been implicated in playing a key role in pluripotency and germ cell sex-specific commitment and differentiation, such as *Dazl*, *Pou5f1*, *Dmc1*, *Rec8*, *Stra8*, *Nodal*, *Nanos2*, and *Dnmt3l*, our analysis predicted more than 800 positive and negative regulons (**Fig. 2, Fig. S3** and **Table S2**), including a large set of new potentially critical regulators of germ cell commitment and differentiation, for example KDM5A, KDM5B, NR3C1 and PHF8. While most of these remain predictions, they provide an important framework and guide for future experimental investigation.

Identifying the molecules controlling the fundamental decision of germ cells to exit the cell cycle and enter meiosis represents a major challenge for the reproductive medicine community. An example of the usefulness of such a regulome is provided by the

bioRxiv preprint doi: https://doi.org/10.1101/747279; this version posted September 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

prediction of new factors positively and negatively regulating *Stra8* expression. STRA8 is the only gatekeeper described to date that engages the meiotic program in developing female germ cells (Baltus et al., 2006). The onset of *Stra8* expression in germ cells of the developing ovary and its lack of expression in germ cells of the developing testis led to a search for the presence of the female "meiosis-initiating substance" (MIS) or male "meiosis-preventing substance" (MPS) (Kocer et al., 2009; McLaren and Southee, 1997). While RA has emerged as a potential MIS (Bowles et al., 2006; Koubova et al., 2006), recent reports revealed that female germ cells enter meiosis normally even in the absence of RA signaling (Kumar et al., 2011; Vernet et al., 2019) . Our regulome analysis revealed that the expression of both *Stra8* and *Rec8* is regulated by a combination of positive and negative master regulators. Among them, two factors, RAD21 and YBX1, are predicted to act as negative regulators of both *Stra8* and *Rec8* expression. *Rad21* and *Ybx1* genes are initially expressed in both developing ovary and testis but are specifically downregulated in developing ovaries at the time of entry into meiosis. Among the positive regulators, we found the transcription factor PBX3 and KDM5A, a histone demethylase, transiently expressed in XX germ cells around E13.5 when *Stra8* is upregulated. KDM5A has recently been shown to regulate temperature-dependent sex determination in red-eared slider turtle by promoting the transcription of the male sex-determining gene *Dmrt1* (Ge et al., 2018). Interestingly, KDM5A also acts as a negative regulator for the expression of *Rad21* and *Ybx1*, suggesting mutual antagonism between the male RAD21/YBX1 factors and KDM5A/PBX3 female factors in regulating *Stra8* expression and the entry into meiosis.

bioRxiv preprint doi: https://doi.org/10.1101/747279; this version posted September 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Another relevant master regulator is the nuclear receptor MSX2, a member of the Msh homeobox gene family composed of three members (*Msx1*, *2* and *3*). MSX1 and MSX2 function cooperatively to control the regulation of primordial germ cell migration (Sun et al., 2016) and later in XX germ cells to promote meiosis initiation by maintaining or enhancing Stra8 expression (Le Bouffant et al., 2011). We found that Msx2 regulon activity as well as *Msx1* and *Msx2* gene expression are present in module 16, consistent with their dual function (**Fig. 2** and **Fig. 4B**). *Msx2* is expressed in both XX and XY PGCs by E10.5 and E11.5 and then is rapidly downregulated in XY germ cells. In contrast, *Msx1* is specifically upregulated in XX germ cells from E11.5 onward with a peak around E13.5 (**Fig. 4B**). We found 66 predicted target genes regulated positively by the master regulator MSX2. Among these genes, we found for example *Msx1*, *Rec8*, *Dlx3*, *Lef1*, *Sp5*, *Axin2*, *E2f1*, *Notch1* (**Table S2**).

As alluded above, two histone demethylases KDM5A (also called Jarid1A, RBP2) and KDM5B (also called Jmjd3, Jarid1B or RBP2-H1) have been identified as master regulators specifically expressed in E13.5 XX germ cells (**Fig. 2** and **Fig. 4B**). KDM5A and KDM5B are histone demethylases that specifically demethylate H3K4me2/me3 and H3K27me3, respectively. Both histone demethylases are involved in epigenetic regulation of transcription and are essential for embryonic development (Christensen et al., 2007; Dahl et al., 2016; Klose et al., 2007). We found that both *Kdm5a* and *Kdm5b* genes are transiently expressed in XX germ cells around E13.5 in modules 13 and 14, respectively (**Fig. 2** and **Fig. 4B**). In male rat, KDM5A has been shown to be also expressed in quiescent gonocytes, mitotic gonocytes and spermatogonia at 6 dpp, and in spermatocytes at 12, 15 and 18 dpp (Nishio et al., 2014). We identified more than 2625

bioRxiv preprint doi: https://doi.org/10.1101/747279; this version posted September 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

and 1116 predicted positive target genes for KDM5A and KDM5B in germ cells including *Ctnnb1* (beta catenin 1), DEAD-Box Helicases *Ddx4* and *Ddx6*, the Retinoic Acid Receptor alpha (*Rara*) and the Lysine Demethylase *Kdm2a*, *Kdm2b* and *Kdm3a*, the RNA binding protein *Dazl* and the Mitogen-Activated Protein Kinase Kinase Kinase 4 *Map3k4* (**Table S2**). The biological functions of KDM5A and KDM5B, particularly in the context of XX germ cell meiosis regulation and oocyte differentiation remains poorly characterized.

Splicing kinetics and intron retention as post-transcriptional regulation in developing germ cells

Cellular RNAs are regulated at multiple stages, including transcription, RNA maturation, and degradation. To study posttranscriptional regulation, and more precisely RNA processing, during the process of germ cell sex determination we evaluated the abundance of nascent (unspliced) and mature (spliced) mRNAs in the 14,750 XX and XY germ cells. We exploited the fact that 23% of reads contained unspliced intronic sequences when performing single-cell RNA-seq based on the 10x Genomics Chromium protocols (La Manno et al., 2018). Although these splicing events are usually located at the very 3' end of mRNA transcripts due to the oligo-dT-based protocol, the RNA velocity analysis approach can be used to directly estimate unspliced and spliced mRNAs present in single cells (La Manno et al., 2018).

We found a large variation in splicing kinetics. When analyzing, for example, meiosis-related genes we identified genes with rapid splicing (e.g. *Stra8*, *Rec8*) as well as genes with variable delays in splicing (e.g. *Hormad2*, *Sycp1*, *Spo11*, *Tex11*). This suggested that

bioRxiv preprint doi: https://doi.org/10.1101/747279; this version posted September 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

splicing retention is another layer of post-transcriptional regulation in developing germ cells, ensuring precise temporal expression of meiotic and differentiation genes. In addition, we identified genes with sex-specific intronic retention/degradation patterns. One example is the RNA helicase YTHDC2, a critical regulator of the transition from mitosis to meiosis in both male and female germline (Bailey et al., 2017; Gonczy et al., 1997). YTHDC2, through post-transcriptional control of RNA, both down-regulates the mitotic program and facilitates the proper expression of meiotic and differentiation genes. *Ythdc2* mutant male and female mice are infertile and mutant germ cells show defects soon after the mitosis to meiosis transition (Bailey et al., 2017). Consistent with its role in ensuring a transition from mitosis to meiosis, *Ythdc2* expression profile displayed a female-specific upregulation between E13.5-E16.5. However, RNA velocity analysis revealed the presence of unspliced *Ythdc2* mRNAs in both XX and XY germ cells, suggesting that in XY germ cells these unspliced mRNAs remain immature or are degraded in a sex-specific manner. Two other examples are provided by the Cyclins *Ccnb3* and *Ccnd3*, which both displayed increasing levels of unspliced mRNAs in XY germ cells without the presence of mature mRNAs. Intron retention has been shown to be a prominent feature of the meiotic transcriptome of mouse spermatocytes (Naro et al., 2017). It can either favor accumulation, storage, and timely usage of specific transcripts during highly organized cell differentiation programs or cause transcript instability at specific developmental transitions (Edwards et al., 2016; Naro et al., 2017; Pimentel et al., 2016; Wong et al., 2013; Yap et al., 2012). The temporal or sex-specific variation in intronic retention appears to be surprisingly frequent during the process of germ cell sex determination and may ensure proper and timely expression of selected transcripts.

bioRxiv preprint doi: https://doi.org/10.1101/747279; this version posted September 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Comparing adrenal and gonadal germ cells development provides a tool to investigate the importance of the somatic environment in the process of oogenesis

By comparing the transcriptome of adrenal and gonadal germ cells, we have been able to investigate how germ cells respond to three different somatic environments: the adrenal, ovarian and testicular environments. It allowed us to also investigate whether the gene regulatory circuitries mediating germ cell sex determination, composed of 837 regulons, are interconnected or act independently. The dynamic expression pattern of key marker genes of meiosis is strikingly similar in both adrenal and ovarian germ cells with a slight delay in adrenal germ cells. It suggests that the initiation and maintenance of meiosis proceeds relatively normally in adrenal germ cells. However, we also observed a significant alteration in the expression of some key female master regulator genes as well as upregulation of male master regulator genes, indicating that the somatic environment in the adrenal gland cannot completely support a female fate for these gonocytes. In particular, we observed a lack of upregulation of genes involved in the canonical WNT/β-catenin signaling pathway (*Axin2*, *Lef1*, and *Sp5*), suggesting that germ cells in this environment are unable to respond to WNT signals, or to express their receptors. This may also explain the slight delay in adrenal meiosis progression (Chassot et al., 2011; Chassot et al., 2008; Naillat et al., 2010). Other master regulator genes such as the transcription factor genes *Msx1* and *Msx2*, *Cdx2* and the oocyte-specific basic helix-loop-helix transcription factor gene *Figla* also displayed significant alteration. Interestingly, the absence of *Msx1* and *Msx2* expression may explain why *Rec8* expression, but not other meiotic genes such as *Stra8*, is blunted in adrenal germ cells (**Fig. S7**). Based on

bioRxiv preprint doi: https://doi.org/10.1101/747279; this version posted September 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

our GRN analysis, we found that both MSX1 and MSX2 are strong positive regulators of *Rec8* expression (**Fig. 4A**). Concerning male-specific master regulators, the large majority of them are not expressed in XY adrenal germ cells but with few notable exceptions including *Nanos2*, *Lefty1*, *Lefty2*, *Rara*, *Ctcfl*, *Lhx1*. Overall, the adrenal environment does not provide all the necessary signal(s) required to commit germ cells to oogenesis, resulting in adrenal germ cells characterized by an altered identity and delayed meiosis.

Compiling single cell transcriptomes from mouse germ cells at five developmental stages during the process of sex determination, both in gonadal and extragonadal tissues, allowed us to provide a comprehensive insight into the sex-specific genetic programs and gene regulatory networks that regulate germ cell commitment toward a male or female fate. As such, we have created a valuable and rich resource for the broad biology community that should provide insights into how this fundamental decision impacts the etiology of infertility and human gonadal germ cell tumours, two of the main clinical consequences of defects in germ cell sex determination.

Acknowledgments:

This work was supported by grants from the Swiss National Science Foundation (grants 31003A_173070 and 51PHI0-141994) and by the Département de l'Instruction Publique of the State of Geneva (to S.N.). We thank Luciana Romano and Deborah Penet for the sequencing, Cécile Gameiro and Gregory Schneiter (Flow Cytometry Facility, University of Geneva), the team of the Animal Facility (Faculty of Medicine, University of Geneva),

bioRxiv preprint doi: https://doi.org/10.1101/747279; this version posted September 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Julien Prados (Basic Neuroscience, University of Geneva) for his help for pseudotime computation and Valentin Durand Graphic Design for help with artwork. We thank also Andy Greenfield, the members of the Nef and Dermitzakis laboratories for helpful discussion and critical reading of the manuscript.

Author Contributions

Conceptualization, S.N.; data generation and investigation, C.M., Y.N., P.S., A.A.C., I.S. and F.K.; Formal Analysis and Data Curation, C.M and P.S.; Writing – Original Draft, S.N., C.M., and M.-C.C.; Funding Acquisition, S.N., and E.T.D.; Resources, S.N., and E.T.D.; Supervision, S.N., and E.T.D.

Declaration of Interests

The authors declare no competing interests

bioRxiv preprint doi: https://doi.org/10.1101/747279; this version posted September 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

STAR Methods:

KEY RESOURCES TABLE

bioRxiv preprint doi: https://doi.org/10.1101/747279; this version posted September 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, Peptides, and Recombinant Proteins		
Trypsin-EDTA (0.05%), phenol red	Thermo Fisher Scientific	25300054
DPBS, no calcium, no magnesium	Thermo Fisher Scientific	14190144
Fetal bovine serum	Thermo Fisher Scientific	26140087
Draq7™ #B25595	Beckman Coulter	B25595
Critical Commercial Assays		
Papain dissociation system	Worthington	LK003150
Chromium™ i7 Multiplex Kit	10x Genomics	120262
Chromium™ Single Cell 3' Library & Gel Bead Kit v2	10x Genomics	120237
Chromium™ Single Cell A Chip Kit	10x Genomics	1000009
Qubit dsDNA High Sensitivity	Life Technologies	Q32854
Agilent High Sensitivity DNA Kit Reagents	Agilent	5067-4626
Deposited Data		
Raw data, normalized counts	GEO	GSE136220
Scripts for analysis	GitHub	Available soon
Experimental Models: Organisms/Strains		
Mus musculus: CD-1	Charles River	Strain code 022
Mus musculus: CD-1-Tg(*Nr5a1* GFP)	(Stallings, 2002)	N/A
Oligonucleotides		
Primers for sex genotyping	(McFarlane et al., 2013)	N/A
Software and Algorithms		
CellRanger (version 2.3)	10X	www.10xgenomics.com
Python (version 3.6)	Python	Pyth
R (version version 3.6.1)	R-Project	https://www.R-project.org/
Seurat (Version 2.3.0)	Stuart et al, 2018	
heatmaps (version 1.8.0)	Bioconductor	https://bioconductor.org/packages/release/bioc/html/heatmaps.html
pheatmap (version 1.0.12)	Bioconductor	https://github.com/raivokolde/pheatmap
Bmrm (version 4.1)	(Teo et al., 2010)	
ggplot2 (version 3.2.0)	(Wickham, 2009)	https://ggplot2.tidyverse.org
scanpy (version 1.4.4)	(Wolf et al, 2018)	https://github.com/theislab/scanpy
scvelo (version 0.1.19)		https://github.com/theislab/scvelo
matplotlib (version 3.0.3)		
pyscenic (version 0.9.14)	(Aibar et al., 2017)	
bbknn (version 1.3.5)	(Park et al, 2018)	https://github.com/Teichlab/bbknn
velocyto (version 0.17.8)	(La Manno et al 2018)	http://velocyto.org/

bioRxiv preprint doi: https://doi.org/10.1101/747279; this version posted September 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Contact for Reagent and Resource Sharing

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Serge Nef (serge.nef@unige.ch).

Experimental Model Details

Transgenic Mice

All animal work was conducted according to the ethical guidelines of the Direction Générale de la Santé of the Canton de Genève (experimentation ID GE/57/18). *Tg(Nr5a1-GFP)* mouse strain was described previously (Stallings et al., 2002) and has been maintained on a CD1 genetic background.

Method Details

Mouse urogenital ridges, testes, ovaries and adrenal glands collection

CD-1 female mice were bred with heterozygous *Tg(Nr5a1-GFP)* transgenic male mice. Adult females were time-mated and checked for the presence of vaginal plugs the next morning (E0.5). E10.5 (8±2 tail somites (ts)), E11.5 (19±4 ts), E12.5, E13.5, E16.5 and E18.5 embryos were collected and the presence of the *Nr5a1*-GFP transgene was assessed under UV light. Sexing of E10.5 and E11.5 embryos was performed by PCR with a modified protocol from (McFarlane et al., 2013). Urogenital ridges from each sex, XY adrenal glands, testes or ovaries were pooled for tissue dissociation.

Single cell suspension and library preparations

bioRxiv preprint doi: https://doi.org/10.1101/747279; this version posted September 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Urogenital ridges and adrenal glands were enzymatically dissociated at 37ºC for 20 and 40 minutes, respectively, using the Papain dissociation system (Worthington #LK003150). Cells were resuspended in DMEM 2%FBS, filtered through a 70 µm cell strainer and stained with the dead cell marker Draq7™ (Beckman Coulter, #B25595). Viable single cells were collected on a BD FACS Aria II by excluding debris (side scatter vs. forward scatter), dead cells (side scatter vs. Draq7 staining), and doublets (height vs. width). Testes and ovaries (from E12.5 to E16.5) were enzymatically dissociated at 37ºC during 15 minutes in Trypsin-EDTA 0.05% (Gibco #25300054), resuspended in DMEM 2%FBS and filtered through a 70 µm cell strainer. After counting, 3000 to 7000 single cells were loaded on a 10x Chromium instrument (10x Genomics). Single-cell RNA-Seq libraries were prepared using the Chromium Single Cell 3′ v2 Reagent Kit (10x Genomics) according to manufacturer's protocol. Each condition (organ, sex and developmental stage) was performed in two biological independent replicates.

Sequencing

Library quantification was performed using the Qubit fluorometric assay with dsDNA HS Assay Kit (Invitrogen). Library quality assessment was performed using a Bioanalyzer Agilent 2100 with a High Sensitivity DNA chip (Agilent Genomics). Libraries were diluted, pooled and sequenced on an Illumina HiSeq4000 using paired-end 26 + 98 bp as the sequencing mode. Libraries were sequenced at a targeted depth of 100 000 to 150 000 total reads per cell. Sequencing was performed at the Health 2030 Genome Center, Geneva.

bioRxiv preprint doi: https://doi.org/10.1101/747279; this version posted September 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Bioinformatic Analysis

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

Computations were performed at the Vital-IT Center 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). Algorithms and versions used are listed in the key resources table. Data were mapped to the mouse reference genome GRCm38.p5 in which the eGFP (NC_011521.1) combined with the bovine GH 3′-splice/polyadenylation signals (Stallings et al., 2002) (NM_180996.1) sequences have been added.

Cell-associated barcode selection and quality checks were performed with in-house tools. (see **Supplementary information** for details)

Gene expression normalization

UMI counts per gene and per cell were divided by the total UMI detected in the cell, multiplied by a scale factor of 10,000 and log transformed.

Germ cells selection

After barcode filtering based on the unique molecular identifiers (UMI) distribution, we obtained 92,267 cells. It included 14,904 cells from E10.5, 16,581 cells from E11.5, 19,551 cells from E12.5, 25,012 cells from E13.5, and 16,219 cells from E16.5. Among the 52,463 XY cells and the 39,804 XX cells, the median number of UMIs was 17,493 and 17,655 and the median number of detected genes was 4,802 and 4,658, respectively.

bioRxiv preprint doi: https://doi.org/10.1101/747279; this version posted September 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

To determine which one were germ cells, we selected all genes detected in more than 50 cells, and performed ICA on log normalized values. To assess for batch effect, we built a nearest neighbor graph using BBKNN function (BBKNN package). Clustering was performed using Scanpy Louvain method with resolution 1 and UMAP were generated using Scanpy UMAP method with default parameters. We selected clusters with a strong expression of 10 well-known germ cells markers (see **Supplementary table S3** and **Supplementary information** for details).

Pseudotime ordering of the cells

To order the cells along a pseudotime, we took advantage of the discrete prior knowledge we have about the embryonic day at which each cell were harvested and generated an ordinal regression model (adapted from (Teo et al., 2010)) to obtain a continuous pseudotime score reflecting the differentiation status of the cells (see **Supplementary information** for details).

Gene regulatory network generation

GRN analysis was generated using pyScenic package (see **Supplementary information** for details).

Regulons hierarchical clustering

Regulons were clustered using ward.d hierarchical clustering on the AUC matrix with Spearman correlation distance. Modulons were determined using cutree with k=10 for negative regulons and k=30 for positive regulons.

bioRxiv preprint doi: https://doi.org/10.1101/747279; this version posted September 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Heatmaps and expression curves

Heatmaps were generated using R (packages pheatmap and heatmaps). Expression curves were generated using ggplot2 (see **Supplementary information** for details).

Network visualization

Network views and layouts were generated with Cytoscape V3.7.1.

Velocity analysis

To generate spliced and unspliced counts data, the velocyto.py script from velocyto package was called on each bam file with aforementioned reference genome annotation (see **Supplementary information** for details).

Ectopic adrenal germ cells analysis

The analysis was performed using aforementioned steps: log normalization, ICA, neighbor graph, and clustering with the same parameters. Germ cell clusters were selected with the same 10 germ cell marker genes (see **Supplementary information** for details**)**. For pseudotime ordering of the cells, a model was trained on gonadal cells only and the pseudotime for adrenal germ cells was predicted from it.

bioRxiv preprint doi: https://doi.org/10.1101/747279; this version posted September 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Data and source code availability

Germ cells single-cell RNA-seq data is available on GEO (accession number GSE136220). Both adrenal and gonadal gene expression data are included in ReproGenomics Viewer (Darde et al., 2019; Darde et al., 2015).

bioRxiv preprint doi: https://doi.org/10.1101/747279; this version posted September 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

References

Aibar, S., Gonzalez-Blas, C.B., Moerman, T., Huynh-Thu, V.A., Imrichova, H., Hulselmans, G., Rambow, F., Marine, J.C., Geurts, P., Aerts, J., *et al.* (2017). SCENIC: single-cell regulatory network inference and clustering. Nat Methods *14*, 1083-1086.

Aravin, A.A., Sachidanandam, R., Bourc'his, D., Schaefer, C., Pezic, D., Toth, K.F., Bestor, T., and Hannon, G.J. (2008). A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. Mol Cell *31*, 785-799.

Bailey, A.S., Batista, P.J., Gold, R.S., Chen, Y.G., de Rooij, D.G., Chang, H.Y., and Fuller, M.T. (2017). The conserved RNA helicase YTHDC2 regulates the transition from proliferation to differentiation in the germline. Elife *6*.

Baltus, A.E., Menke, D.B., Hu, Y.C., Goodheart, M.L., Carpenter, A.E., de Rooij, D.G., and Page, D.C. (2006). In germ cells of mouse embryonic ovaries, the decision to enter meiosis precedes premeiotic DNA replication. Nat Genet *38*, 1430-1434.

Bowles, J., Knight, D., Smith, C., Wilhelm, D., Richman, J., Mamiya, S., Yashiro, K., Chawengsaksophak, K., Wilson, M.J., Rossant, J., *et al.* (2006). Retinoid signaling determines germ cell fate in mice. Science *312*, 596-600.

Bullejos, M., and Koopman, P. (2004). Germ cells enter meiosis in a rostro-caudal wave during development of the mouse ovary. Mol Reprod Dev *68*, 422-428.

Byskov, A.G., and Saxen, L. (1976). Induction of meiosis in fetal mouse testis in vitro. Dev Biol *52*, 193-200.

Chassot, A.A., Gregoire, E.P., Lavery, R., Taketo, M.M., de Rooij, D.G., Adams, I.R., and Chaboissier, M.C. (2011). RSPO1/beta-catenin signaling pathway regulates oogonia differentiation and entry into meiosis in the mouse fetal ovary. PLoS One *6*, e25641.

Chassot, A.A., Ranc, F., Gregoire, E.P., Roepers-Gajadien, H.L., Taketo, M.M., Camerino, G., de Rooij, D.G., Schedl, A., and Chaboissier, M.C. (2008). Activation of beta-catenin signaling by Rspo1 controls differentiation of the mammalian ovary. Hum Mol Genet *17*, 1264-1277.

Christensen, J., Agger, K., Cloos, P.A., Pasini, D., Rose, S., Sennels, L., Rappsilber, J., Hansen, K.H., Salcini, A.E., and Helin, K. (2007). RBP2 belongs to a family of demethylases, specific for tri-and dimethylated lysine 4 on histone 3. Cell *128*, 1063-1076.

Dahl, J.A., Jung, I., Aanes, H., Greggains, G.D., Manaf, A., Lerdrup, M., Li, G., Kuan, S., Li, B., Lee, A.Y., *et al.* (2016). Broad histone H3K4me3 domains in mouse oocytes modulate maternal-to-zygotic transition. Nature *537*, 548-552.

Darde, T.A., Lecluze, E., Lardenois, A., Stevant, I., Alary, N., Tuttelmann, F., Collin, O., Nef, S., Jegou, B., Rolland, A.D., *et al.* (2019). The ReproGenomics Viewer: a multi-omics and cross-species resource compatible with single-cell studies for the reproductive science community. Bioinformatics.

Darde, T.A., Sallou, O., Becker, E., Evrard, B., Monjeaud, C., Le Bras, Y., Jegou, B., Collin, O., Rolland, A.D., and Chalmel, F. (2015). The ReproGenomics Viewer: an integrative cross-species toolbox for the reproductive science community. Nucleic Acids Res *43*, W109-116.

Edwards, C.R., Ritchie, W., Wong, J.J., Schmitz, U., Middleton, R., An, X., Mohandas, N., Rasko, J.E., and Blobel, G.A. (2016). A dynamic intron retention program in the mammalian megakaryocyte and erythrocyte lineages. Blood *127*, e24-e34.

Evans, E.P., Ford, C.E., and Lyon, M.F. (1977). Direct evidence of the capacity of the XY germ cell in the mouse to become an oocyte. Nature *267*, 430-431.

Ge, C., Ye, J., Weber, C., Sun, W., Zhang, H., Zhou, Y., Cai, C., Qian, G., and Capel, B. (2018). The histone demethylase KDM6B regulates temperature-dependent sex determination in a turtle species. Science *360*, 645-648.

bioRxiv preprint doi: https://doi.org/10.1101/747279; this version posted September 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Gkountela, S., Zhang, K.X., Shafiq, T.A., Liao, W.W., Hargan-Calvopina, J., Chen, P.Y., and Clark, A.T. (2015). DNA Demethylation Dynamics in the Human Prenatal Germline. Cell *161*, 1425-1436.

Gonczy, P., Matunis, E., and DiNardo, S. (1997). bag-of-marbles and benign gonial cell neoplasm act in the germline to restrict proliferation during Drosophila spermatogenesis. Development *124*, 4361-4371.

Guo, F., Yan, L., Guo, H., Li, L., Hu, B., Zhao, Y., Yong, J., Hu, Y., Wang, X., Wei, Y., *et al.* (2015). The Transcriptome and DNA Methylome Landscapes of Human Primordial Germ Cells. Cell *161*, 1437-1452.

Heeren, A.M., He, N., de Souza, A.F., Goercharn-Ramlal, A., van Iperen, L., Roost, M.S., Gomes Fernandes, M.M., van der Westerlaken, L.A., and Chuva de Sousa Lopes, S.M. (2016). On the development of extragonadal and gonadal human germ cells. Biol Open *5*, 185-194.

Hill, P.W.S., Leitch, H.G., Requena, C.E., Sun, Z., Amouroux, R., Roman-Trufero, M., Borkowska, M., Terragni, J., Vaisvila, R., Linnett, S., *et al.* (2018). Epigenetic reprogramming enables the transition from primordial germ cell to gonocyte. Nature *555*, 392-396.

Houmard, B., Small, C., Yang, L., Naluai-Cecchini, T., Cheng, E., Hassold, T., and Griswold, M. (2009). Global gene expression in the human fetal testis and ovary. Biol Reprod *81*, 438-443.

Irie, N., Weinberger, L., Tang, W.W., Kobayashi, T., Viukov, S., Manor, Y.S., Dietmann, S., Hanna, J.H., and Surani, M.A. (2015). SOX17 is a critical specifier of human primordial germ cell fate. Cell *160*, 253-268.

Jameson, S.A., Natarajan, A., Cool, J., DeFalco, T., Maatouk, D.M., Mork, L., Munger, S.C., and Capel, B. (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.

Kalsotra, A., and Cooper, T.A. (2011). Functional consequences of developmentally regulated alternative splicing. Nat Rev Genet *12*, 715-729.

Klose, R.J., Yan, Q., Tothova, Z., Yamane, K., Erdjument-Bromage, H., Tempst, P., Gilliland, D.G., Zhang, Y., and Kaelin, W.G., Jr. (2007). The retinoblastoma binding protein RBP2 is an H3K4 demethylase. Cell *128*, 889-900.

Kocer, A., Reichmann, J., Best, D., and Adams, I.R. (2009). Germ cell sex determination in mammals. Mol Hum Reprod *15*, 205-213.

Koubova, J., Hu, Y.C., Bhattacharyya, T., Soh, Y.Q., Gill, M.E., Goodheart, M.L., Hogarth, C.A., Griswold, M.D., and Page, D.C. (2014). Retinoic acid activates two pathways required for meiosis in mice. PLoS Genet *10*, e1004541.

Koubova, J., Menke, D.B., Zhou, Q., Capel, B., Griswold, M.D., and Page, D.C. (2006). Retinoic acid regulates sex-specific timing of meiotic initiation in mice. Proc Natl Acad Sci U S A *103*, 2474-2479.

Kumar, S., Chatzi, C., Brade, T., Cunningham, T.J., Zhao, X., and Duester, G. (2011). Sex-specific timing of meiotic initiation is regulated by Cyp26b1 independent of retinoic acid signalling. Nat Commun *2*, 151.

La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K., Kastriti, M.E., Lonnerberg, P., Furlan, A., *et al.* (2018). RNA velocity of single cells. Nature *560*, 494-498.

La Salle, S., Mertineit, C., Taketo, T., Moens, P.B., Bestor, T.H., and Trasler, J.M. (2004). Windows for sex-specific methylation marked by DNA methyltransferase expression profiles in mouse germ cells. Dev Biol *268*, 403-415.

Le Bouffant, R., Souquet, B., Duval, N., Duquenne, C., Herve, R., Frydman, N., Robert, B., Habert, R., and Livera, G. (2011). Msx1 and Msx2 promote meiosis initiation. Development *138*, 5393-5402.

Lesch, B.J., Dokshin, G.A., Young, R.A., McCarrey, J.R., and Page, D.C. (2013). A set of genes critical to development is epigenetically poised in mouse germ cells from fetal stages through completion of meiosis. Proc Natl Acad Sci U S A *110*, 16061-16066.

bioRxiv preprint doi: https://doi.org/10.1101/747279; this version posted September 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Maldonado-Saldivia, J., van den Bergen, J., Krouskos, M., Gilchrist, M., Lee, C., Li, R., Sinclair, A.H., Surani, M.A., and Western, P.S. (2007). Dppa2 and Dppa4 are closely linked SAP motif genes restricted to pluripotent cells and the germ line. Stem Cells *25*, 19-28.

Margueron, R., and Reinberg, D. (2011). The Polycomb complex PRC2 and its mark in life. Nature *469*, 343-349.

McFarlane, L., Truong, V., Palmer, J.S., and Wilhelm, D. (2013). Novel PCR assay for determining the genetic sex of mice. Sex Dev *7*, 207-211.

McLaren, A. (1983). Sex reversal in the mouse. Differentiation *23 Suppl*, S93-98.

McLaren, A. (1984). Meiosis and differentiation of mouse germ cells. Symp Soc Exp Biol *38*, 7-23.

McLaren, A. (2003). Primordial germ cells in the mouse. Dev Biol *262*, 1-15.

McLaren, A., and Southee, D. (1997). Entry of mouse embryonic germ cells into meiosis. Dev Biol *187*, 107-113.

Menke, D.B., Koubova, J., and Page, D.C. (2003). Sexual differentiation of germ cells in XX mouse gonads occurs in an anterior-to-posterior wave. Dev Biol *262*, 303-312.

Molyneaux, K.A., Wang, Y., Schaible, K., and Wylie, C. (2004). Transcriptional profiling identifies genes differentially expressed during and after migration in murine primordial germ cells. Gene Expr Patterns *4*, 167-181.

Naillat, F., Prunskaite-Hyyrylainen, R., Pietila, I., Sormunen, R., Jokela, T., Shan, J., and Vainio, S.J. (2010). Wnt4/5a signalling coordinates cell adhesion and entry into meiosis during presumptive ovarian follicle development. Hum Mol Genet *19*, 1539-1550.

Naro, C., Jolly, A., Di Persio, S., Bielli, P., Setterblad, N., Alberdi, A.J., Vicini, E., Geremia, R., De la Grange, P., and Sette, C. (2017). An Orchestrated Intron Retention Program in Meiosis Controls Timely Usage of Transcripts during Germ Cell Differentiation. Dev Cell *41*, 82-93 e84.

Nishio, H., Hayashi, Y., Moritoki, Y., Kamisawa, H., Mizuno, K., Kojima, Y., and Kohri, K. (2014). Distinctive changes in histone H3K4 modification mediated via Kdm5a expression in spermatogonial stem cells of cryptorchid testes. J Urol *191*, 1564-1572.

Oulad-Abdelghani, M., Bouillet, P., Decimo, D., Gansmuller, A., Heyberger, S., Dolle, P., Bronner, S., Lutz, Y., and Chambon, P. (1996). Characterization of a premeiotic germ cell-specific cytoplasmic protein encoded by Stra8, a novel retinoic acid-responsive gene. J Cell Biol *135*, 469-477.

Pesce, M., and Scholer, H.R. (2001). Oct-4: gatekeeper in the beginnings of mammalian development. Stem Cells *19*, 271-278.

Pimentel, H., Parra, M., Gee, S.L., Mohandas, N., Pachter, L., and Conboy, J.G. (2016). A dynamic intron retention program enriched in RNA processing genes regulates gene expression during terminal erythropoiesis. Nucleic Acids Res *44*, 838-851.

Rolland, A.D., Jegou, B., and Pineau, C. (2008). Testicular development and spermatogenesis: harvesting the postgenomics bounty. Adv Exp Med Biol *636*, 16-41.

Rolland, A.D., Lehmann, K.P., Johnson, K.J., Gaido, K.W., and Koopman, P. (2011). Uncovering gene regulatory networks during mouse fetal germ cell development. Biol Reprod *84*, 790-800.

Small, C.L., Shima, J.E., Uzumcu, M., Skinner, M.K., and Griswold, M.D. (2005). Profiling gene expression during the differentiation and development of the murine embryonic gonad. Biol Reprod *72*, 492-501.

Soh, Y.Q., Junker, J.P., Gill, M.E., Mueller, J.L., van Oudenaarden, A., and Page, D.C. (2015). A Gene Regulatory Program for Meiotic Prophase in the Fetal Ovary. PLoS Genet *11*, e1005531.

Soumillon, M., Necsulea, A., Weier, M., Brawand, D., Zhang, X., Gu, H., Barthes, P., Kokkinaki, M., Nef, S., Gnirke, A., *et al.* (2013). Cellular source and mechanisms of high transcriptome complexity in the mammalian testis. Cell Rep *3*, 2179-2190.

Souquet, B., Tourpin, S., Messiaen, S., Moison, D., Habert, R., and Livera, G. (2012). Nodal signaling regulates the entry into meiosis in fetal germ cells. Endocrinology *153*, 2466-2473.

bioRxiv preprint doi: https://doi.org/10.1101/747279; this version posted September 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Spiller, C., and Bowles, J. (2019). Sexually dimorphic germ cell identity in mammals. Curr Top Dev Biol *134*, 253-288.

Spiller, C.M., Bowles, J., and Koopman, P. (2012a). Regulation of germ cell meiosis in the fetal ovary. Int J Dev Biol *56*, 779-787.

Spiller, C.M., Feng, C.W., Jackson, A., Gillis, A.J., Rolland, A.D., Looijenga, L.H., Koopman, P., and Bowles, J. (2012b). Endogenous Nodal signaling regulates germ cell potency during mammalian testis development. Development *139*, 4123-4132.

Stallings, N.R., Hanley, N.A., Majdic, G., Zhao, L., Bakke, M., and Parker, K.L. (2002). Development of a transgenic green fluorescent protein lineage marker for steroidogenic factor 1. Mol Endocrinol *16*, 2360-2370.

Sun, J., Ting, M.C., Ishii, M., and Maxson, R. (2016). Msx1 and Msx2 function together in the regulation of primordial germ cell migration in the mouse. Dev Biol *417*, 11-24.

Surani, M.A., Hayashi, K., and Hajkova, P. (2007). Genetic and epigenetic regulators of pluripotency. Cell *128*, 747-762.

Suzuki, A., and Saga, Y. (2008). Nanos2 suppresses meiosis and promotes male germ cell differentiation. Genes Dev *22*, 430-435.

Tang, W.W., Dietmann, S., Irie, N., Leitch, H.G., Floros, V.I., Bradshaw, C.R., Hackett, J.A., Chinnery, P.F., and Surani, M.A. (2015). A Unique Gene Regulatory Network Resets the Human Germline Epigenome for Development. Cell *161*, 1453-1467.

Telley, L., Agirman, G., Prados, J., Fièvre, S., Oberst, P., Vitali, I., Nguyen, L., Dayer, A., and D., J. (2018). Single-cell transcriptional dynamics and origins of neuronal diversity in the developing mouse neocortex. BioRxiv.

Teo, C.H., Vishwanthan, S.V.N., Smola, A.J., and Le, Q.V. (2010). Bundle Methods for Regularized Risk Minimization. Journal of Machine Learning Research *11*, 311–336.

Upadhyay, S., and Zamboni, L. (1982). Ectopic germ cells: natural model for the study of germ cell sexual differentiation. Proc Natl Acad Sci U S A *79*, 6584-6588.

Vernet, N., Mark, M., Condrea, D., Féret, B., Klopfenstein, M., Alunni, V., Teletin, M., and Ghyselinck, N.B. (2019). Meiosis Initiates In The Fetal Ovary Of Mice Lacking All Retinoic Acid Receptor Isotypes. BioRxiv.

Western, P., Maldonado-Saldivia, J., van den Bergen, J., Hajkova, P., Saitou, M., Barton, S., and Surani, M.A. (2005). Analysis of Esg1 expression in pluripotent cells and the germline reveals similarities with Oct4 and Sox2 and differences between human pluripotent cell lines. Stem Cells *23*, 1436-1442.

Western, P.S., Miles, D.C., van den Bergen, J.A., Burton, M., and Sinclair, A.H. (2008). Dynamic regulation of mitotic arrest in fetal male germ cells. Stem Cells *26*, 339-347.

Western, P.S., van den Bergen, J.A., Miles, D.C., and Sinclair, A.H. (2010). Male fetal germ cell differentiation involves complex repression of the regulatory network controlling pluripotency. FASEB J *24*, 3026-3035.

Wong, J.J., Ritchie, W., Ebner, O.A., Selbach, M., Wong, J.W., Huang, Y., Gao, D., Pinello, N., Gonzalez, M., Baidya, K., *et al.* (2013). Orchestrated intron retention regulates normal granulocyte differentiation. Cell *154*, 583-595.

Yao, H.H., DiNapoli, L., and Capel, B. (2003). Meiotic germ cells antagonize mesonephric cell migration and testis cord formation in mouse gonads. Development *130*, 5895-5902.

Yap, K., Lim, Z.Q., Khandelia, P., Friedman, B., and Makeyev, E.V. (2012). Coordinated regulation of neuronal mRNA steady-state levels through developmentally controlled intron retention. Genes Dev *26*, 1209-1223.

Zamboni, L., and Upadhyay, S. (1983). Germ cell differentiation in mouse adrenal glands. J Exp Zool *228*, 173-193.

bioRxiv preprint doi: https://doi.org/10.1101/747279; this version posted September 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure legends

Figure 1. Generation of the germ cell sex determination atlas. (A) Schematic representation of developing testis and ovary highlighting the major events of male and female germ cell differentiation as well as the time points used in the study. (B) Illustration of the experimental workflow using the 10x Genomics Chromium platform. UMAP Projection of 14,750 germ cells colored by time (C), sex (D) and computed pseudotime going from 0 (E10.5 cells) to 100 (E16.5 cells) (E). (F) Heatmap of 67 well known genes involved in germ cell differentiation. Cells were ordered using a pseudotime score generated with ordinal regression modeling, expression was smoothed to reduce dropout effect and obtain a better visualization of expression tendencies (expression scale: log normalized counts normalized per gene). Cells with lowest score (E10.5) are in the center of the figure and those with highest scores (E16.5) are on the left side for XX cells and on the right side for XY cells. The relevant processes regulated by these genes (cell cycle, male sex determination, meiosis, pluripotency and others) are indicated on the left side of the heatmap.

Figure 2. Gene regulation network analysis reveals transient patterns of transcription activation during germ cell sex determination. (A) Activity (see STAR method) heatmap of the 394 regulons with positive association with their master regulator. Regulons were clustered in 30 modules (M1-M30) using hierarchical clustering with Spearman correlation distance. Cells were ordered according to their pseudotime score with lowest score (E10.5) in the center of the figure and highest score (E16.5) on the left side for XX cells and on the right side for XY cells. Boxes on the right display examples of master regulators of interest that are colored by dominant activity in XX

bioRxiv preprint doi: https://doi.org/10.1101/747279; this version posted September 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(pink) or XY (blue) cells. In brackets is the number of target genes for each master regulator. (B) Smoothed expression heatmap of well known marker genes involved in germ cells differentiation. Expression scale: log normalized counts normalized per gene.

Figure 3. Sex-specific, sequential waves of cell cycle genes during germ cells differentiation. Smoothed expression heatmap of cell cycle genes. Cells were ordered according to their pseudotime. Cells with lowest score (E10.5) are in the center of the figure and those with the highest scores (E16.5) are on the left side for XX cells and on the right side for XY cells. Log normalized expression values where normalized per row.

Figure 4. Gene regulation network of the meiotic genes *Stra8, Rec8* and *Ythdc2*. (A) Gene regulation network of *Rec8, Stra8* and *Ythdc2* with connection to their regulators. Color of the edge represent positive (green) or negative (red) regulation. Edge width is proportional to association score of the target gene to the master regulator. Blue and red-fill colors indicate high expression in XY and XX germ cells, respectively (Log2 FC). (B) Expression profiles of selected genes involved in *Rec8, Stra8* and *Ythdc2* gene regulation. The solid line represents the smoothed expression curves of the gene in the cells ordered by pseudotime, and the fade band is the 95% confidence interval of the model.

Figure 5. Relative abundance of nascent (unspliced) and mature (spliced) transcripts reveals gene- or sex-specific differences in transcriptomic kinetics. Expression profile of spliced (solid line) and unspliced (dash line) forms of transcripts in XX (red) and XY (blue) germ cells across pseudotime for (A) meiosis-associated genes, (B) selected regulators of *Stra8, Rec8* and *Ythdc2* expression and (C) mitosis-associated

bioRxiv preprint doi: https://doi.org/10.1101/747279; this version posted September 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

genes. As unspliced transcripts are less detected, spliced expression levels were multiplied by the gamma factor as in La Manno, 2018. In (B) UMAP projection of *Ythdc2* transcripts in the cells colored by pooled abundance of spliced and unspliced transcripts. Expression scale: log normalized counts.

Figure 6. Altered identity and delayed meiosis in adrenal XY germ cells. (A) UMAP projections of 14,718 gonadal cells and 312 ectopic XY germ cells developing in the adrenal colored by organ (left) and time (right). Adrenal germ cells are represented with larger dots. (B) Expression profiles of selected genes involved in meiosis, oogonia and spermatogonia differentiation process. The solid line represents the smoothed expression curves of the gene in the cells ordered by pseudotime, and the fade band is the 95% confidence interval of the model.

Figure 1

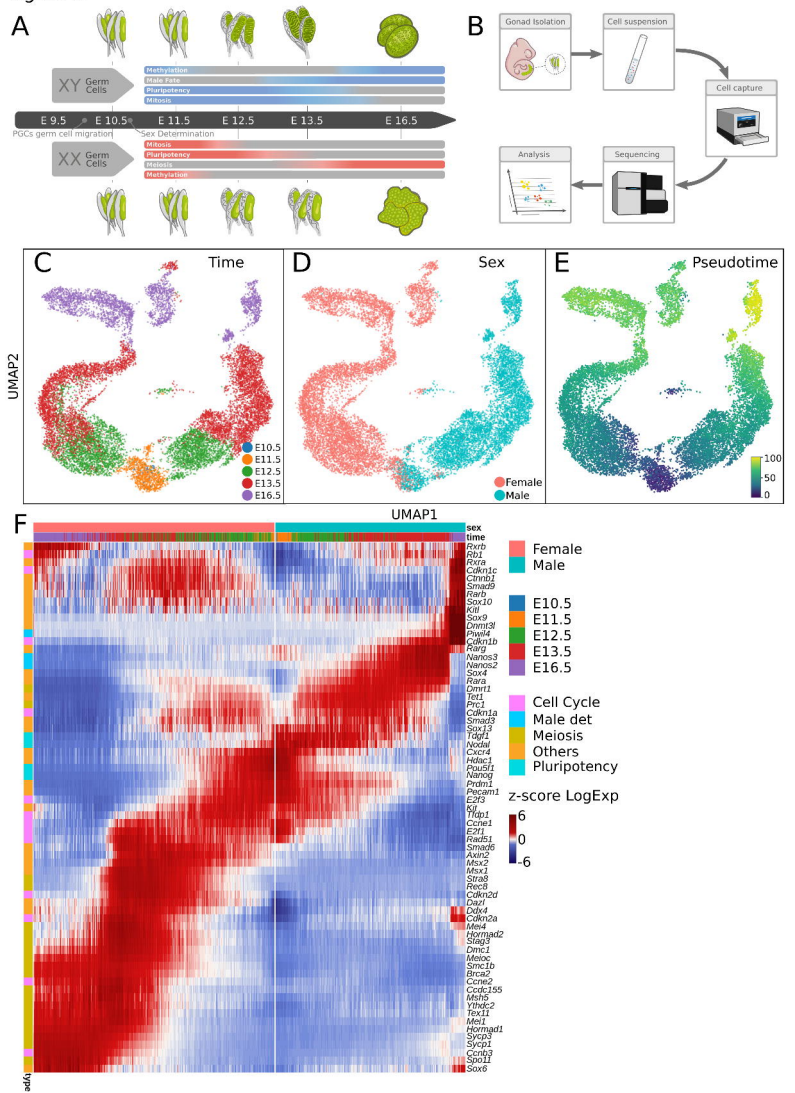

A
XY Germ Cells
Methylation
Male Pate
Pluripotency
Mitosis
E 9.5
E 10.5
E 11.5
E 12.5
E 13.5
E 16.5
PGCs germ cell migration
Sex Determination
XX Germ Cells
Mitosis
Pluripotency
Meiosis
Methylation
B
Gonad Isolation
Cell suspension
Cell capture
Sequencing
Analysis
C
Time
E10.5
E11.5
E12.5
E13.5
E16.5
D
Sex
Female
Male
E
Pseudotime
100
50
0
UMAP2
UMAP1
F
sex
time
Female
Male
E10.5
E11.5
E12.5
E13.5
E16.5
Cell Cycle
Male det
Meiosis
Others
Pluripotency
z-score LogExp
6
0
-6
Rxrb
Rb1
Rxra
Cdkn1c
Ctnnb1
Smad9
Rarb
Sox10
Kitl
Sox9
Dnmt3l
Piwil4
Cdkn1b
Rarg
Nanos3
Nanos2
Sox4
Rara
Dmrt1
Tet1
Prc1
Cdkn1a
Smad3
Sox13
Tdgf1
Nodal
Cxcr4
Hdac1
Pau5f1
Nanog
Prdm1
Pecam1
E2f3
Kif
Tfdp1
Ccne1
E2f1
Rad51
Smad6
Axin2
Msx2
Msx1
Stra8
Rec8
Cdkn2d
Daxl
Ddx4
Cdkn2a
Mei4
Hormad2
Stag3
Dmc1
Meioc
Smc1b
Brca2
Ccne2
Ccdc155
Msh5
Ythdc2
Tex11
Mei1
Hormad1
Sycp3
Sycp1
Ccnb3
Spo11
Sox6
type

Figure 2

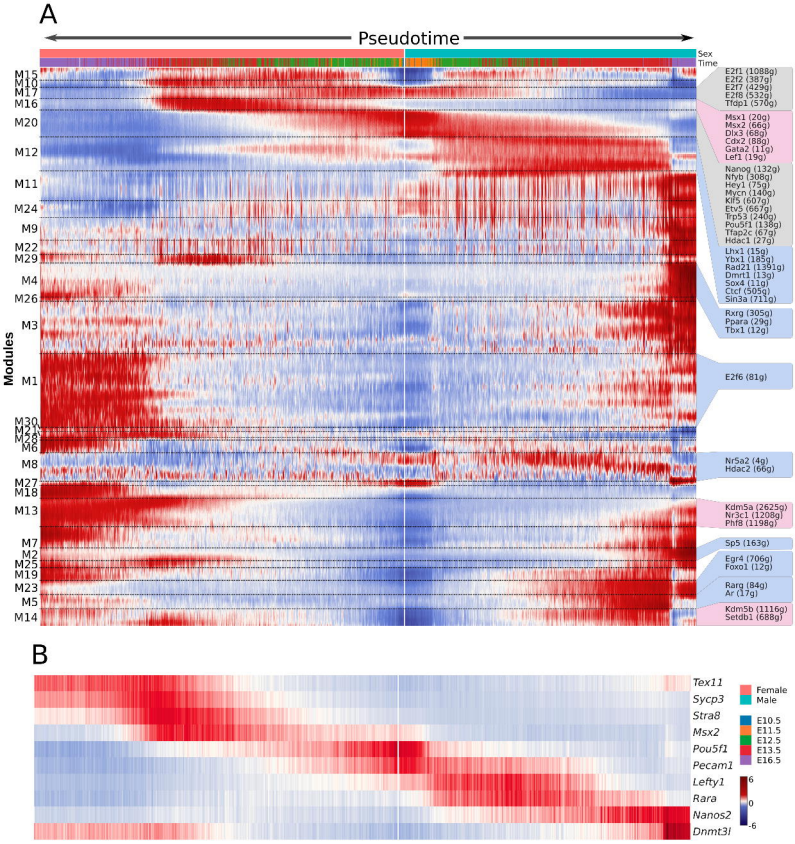

A
Pseudotime
Sex
Time
Modules
M15
M10
M17
M16
M20
M12
M11
M24
M9
M22
M29
M4
M26
M3
M1
M30
M21
M28
M6
M8
M27
M18
M13
M7
M2
M25
M19
M23
M5
M14
E2f1 (1088g)
E2f2 (387g)
E2f7 (429g)
E2f8 (532g)
Tfdp1 (570g)
Msx1 (20g)
Msx2 (66g)
Dlx3 (68g)
Cdx2 (88g)
Gata2 (11g)
Lef1 (19g)
Nanog (132g)
Nfyb (308g)
Hey1 (75g)
Mycn (140g)
Klf5 (607g)
Etv5 (667g)
Trp53 (240g)
Pou5f1 (138g)
Tfap2c (67g)
Hdac1 (27g)
Uhx1 (15g)
Ybx1 (185g)
Rad21 (1391g)
Dnmt1 (13g)
Sox4 (11g)
Ctcf (505g)
Sin3a (711g)
Rxrg (305g)
Ppara (29g)
Tbx1 (12g)
E2f6 (81g)
Nr5a2 (4g)
Hdac2 (66g)
Kdm5a (2625g)
Nr3c1 (1208g)
Phf8 (1198g)
Sp5 (163g)
Egr4 (706g)
Foxo1 (12g)
Rarg (84g)
Ar (17g)
Kdm5b (1116g)
Setdb1 (688g)
B
Tex11
Sycp3
Stra8
Msx2
Pou5f1
Pecam1
Lefty1
Rara
Nanos2
Dnmt3l
Female
Male
E10.5
E11.5
E12.5
E13.5
E16.5
6
0
-6

Figure 3

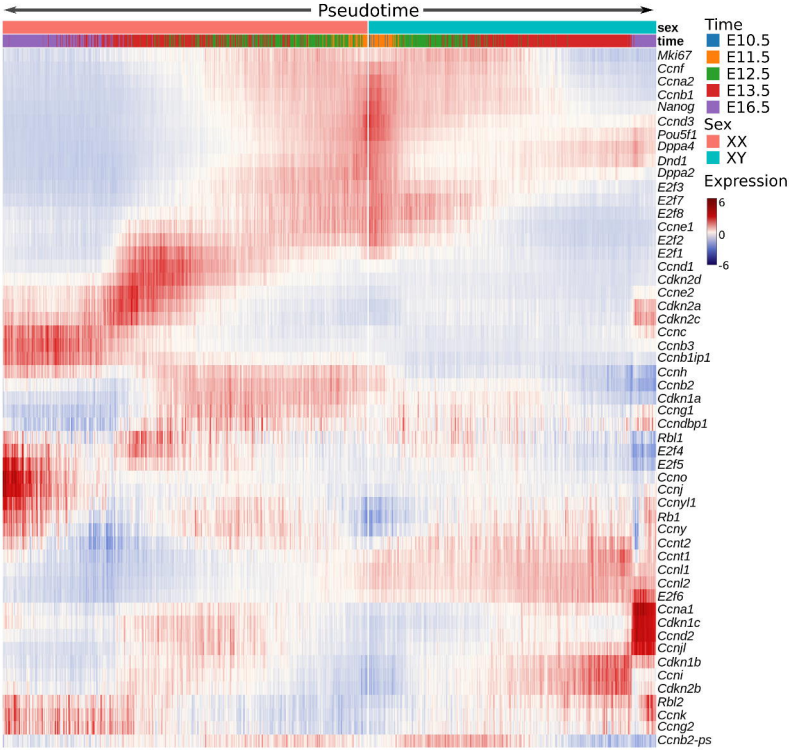

Pseudotime
sex
time
Mki67
Ccnf
Ccna2
Ccnb1
Nanog
Ccnd3
Pou5f1
Dppa4
Dnd1
Dppa2
E2f3
E2f7
E2f8
Ccne1
E2f2
E2f1
Ccnd1
Cdkn2d
Ccne2
Cdkn2a
Cdkn2c
Ccnc
Ccnb3
Ccnb1ip1
Ccnh
Ccnb2
Cdkn1a
Ccng1
Ccndbp1
Rbl1
E2f4
E2f5
Ccno
Ccnj
Ccnyl1
Rb1
Ccny
Ccnt2
Ccnt1
Ccnl1
Ccnl2
E2f6
Ccna1
Cdkn1c
Ccnd2
Ccnjl
Cdkn1b
Ccni
Cdkn2b
Rbl2
Ccnk
Ccng2
Ccnb2-ps
Time
E10.5
E11.5
E12.5
E13.5
E16.5
Sex
XX
XY
Expression
6
0
-6

Figure 4

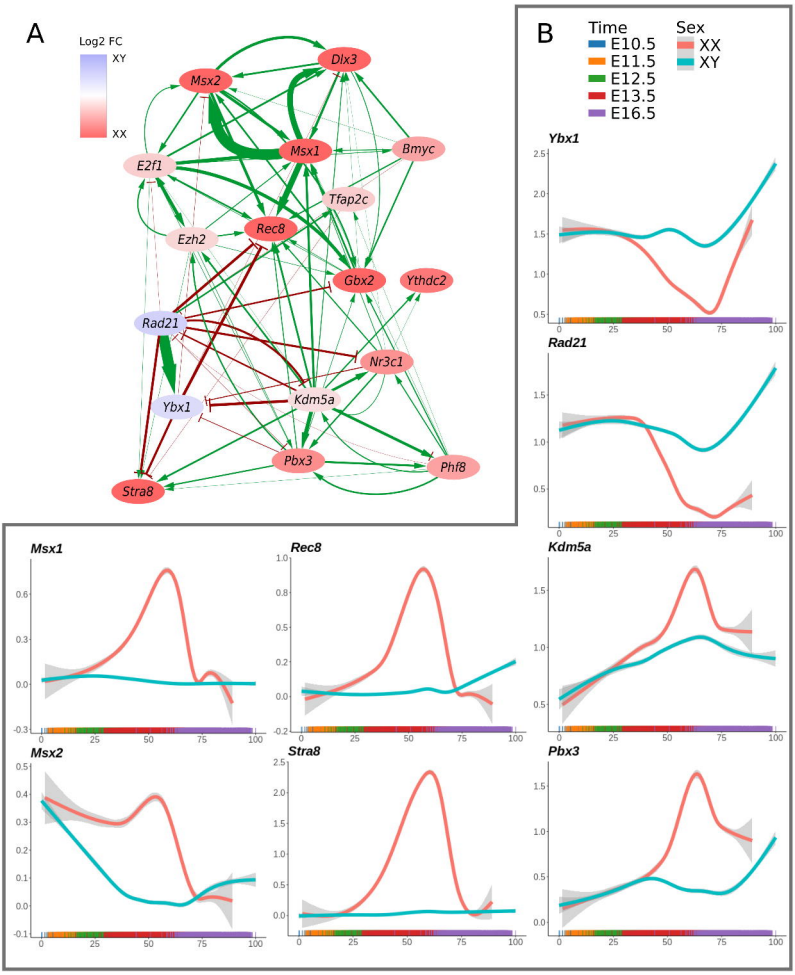

A
Log2 FC
XY
XX
Dlx3
Msx2
Msx1
Bmyc
E2f1
Tfap2c
Rec8
Ezh2
Gbx2
Ythdc2
Rad21
Nr3c1
Kdm5a
Ybx1
Pbx3
Phf8
Stra8
B
Time
E10.5
E11.5
E12.5
E13.5
E16.5
Sex
XX
XY
Ybx1
Rad21
Msx1
Rec8
Kdm5a
Msx2
Stra8
Pbx3

Figure 5

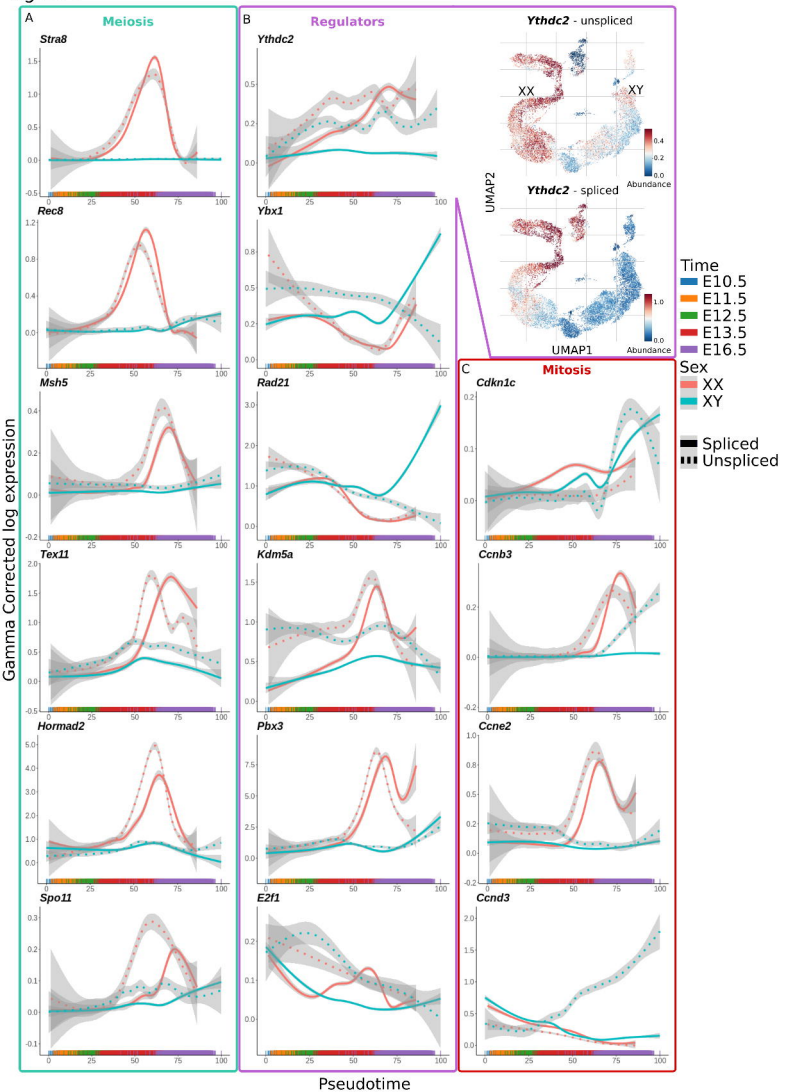
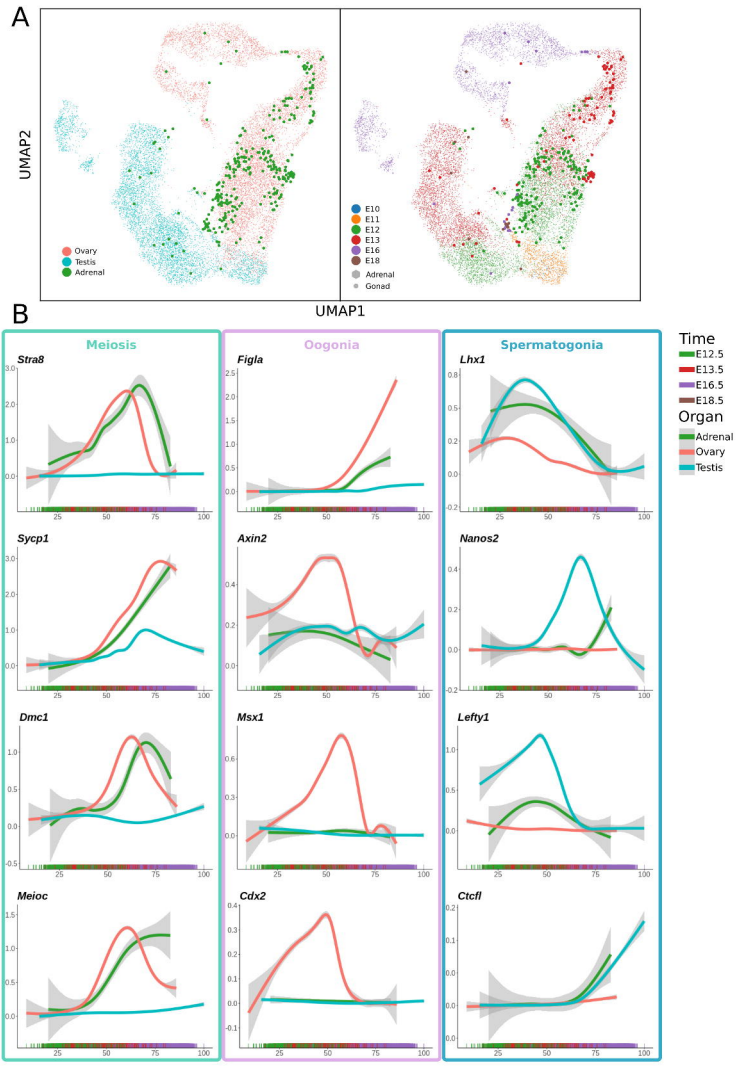

A
Meiosis
Stra8
Rec8
Msh5
Tex11
Hormad2
Spo11
B
Regulators
Ythdc2
Ybx1
Rad21
Kdm5a
Pbx3
E2f1
Ythdc2 - unspliced
XX
XY
Abundance
Ythdc2 - spliced
UMAP2
UMAP1
C
Mitosis
Cdkn1c
Ccnb3
Ccne2
Ccnd3
Time
E10.5
E11.5
E12.5
E13.5
E16.5
Sex
XX
XY
Spliced
Unspliced
Gamma Corrected log expression
Pseudotime

Figure 6


A
UMAP2
UMAP1
Ovary
Testis
Adrenal
E10
E11
E12
E13
E16
E18
Adrenal
Gonad
B
Meiosis
Oogonia
Spermatogonia
Stra8
Figla
Lhx1
Sycp1
Axin2
Nanos2
Dmc1
Msx1
Lefty1
Meioc
Cdx2
Ctcfl
Time
E12.5
E13.5
E16.5
E18.5
Organ
Adrenal
Ovary
Testis