Origin, specification and differentiation of a rare supporting-like lineage in the developing mouse gonad

Short Title (50 characters): Description of a rare gonadal cell lineage

Authors: Chloé Mayère^{1,2,†}, Violaine Regard^{1,†}, Aitana Perea-Gomez³, Corey Bunce⁴, Yasmine Neirijnck^{1,3}, Cyril Djari¹, Pauline Sararols¹, Richard Reeves⁵, Simon Greenaway⁵, Michelle Simon⁵, Pam Siggers⁵, Diana Condrea⁶, Françoise Kühne¹, Ivana Gantar⁷, Furong Tang³, Isabelle Stévant¹, Laura Batti⁷, Norbert B. Ghyselinck⁶, Dagmar Wilhelm⁸, Andy Greenfield⁵, Blanche Capel⁴, Marie-Christine Chaboissier³, Serge Nef^{1,2,*}.

Affiliations:

¹Department of Genetic Medicine and Development, University of Geneva, 1211 Geneva, Switzerland.

²iGE3, Institute of Genetics and Genomics of Geneva, University of Geneva, Switzerland.

³Université Côte d'Azur, Inserm, CNRS, Institut de Biologie Valrose (iBV), 06108 Nice, France.

⁴Department of Cell Biology, Duke University Medical Center, Durham, NC 27710, USA;

⁵Mammalian Genetics Unit, Medical Research Council Harwell Institute, Oxfordshire, OX11 0RD, United Kingdom.

⁶Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), Département de Génétique Fonctionnelle et Cancer, Centre National de la Recherche Scientifique (CNRS UMR7104), Institut National de la Santé et de la Recherche Médicale (INSERM U1258), Université de Strasbourg (UNISTRA), 1 rue Laurent Fries, BP1014267404 ILLKIRCH CEDEX, France.

⁷Wyss Center for Bio and Neuroengineering, Geneva, Switzerland. ⁸Department of Anatomy and Physiology, University of Melbourne, Parkville, VIC 3010, Australia.

[†]These authors share first authorship

*Corresponding Author: Email: Serge.Nef@unige.ch

5

Abstract:

Gonadal sex determination represents a unique model for studying cell fate decisions. However, a complete understanding of the different cell lineages forming the developing testis and ovary remains elusive. Here, we investigated the origin, specification and subsequent sex-specific differentiation of a previously uncharacterized population of supporting-like cells (SLC) in the developing mouse gonads. The SLC lineage is closely related to the coelomic epithelium and specified as early as E10.5, making it the first somatic lineage to be specified in the bipotential gonad. SLC progenitors are localized within the genital ridge at the interface with the mesonephros and initially co-express *Wnt4* and *Sox9*. SLCs become sexually dimorphic around E12.5, progressively acquire a Sertoli- or granulosa-like identity and contribute to the formation of the rete testis and rete ovarii. Finally, we found that *WNT4* is a crucial regulator of the SLC lineage and is required for the formation of the rete testis.

Teaser: Description of an uncharacterized multipotent gonadal cell lineage involved in testis and ovary development

Keywords: Single-cell RNA-sequencing (scRNA-seq), gonadal sex determination, ovary, testis, rete testis, rete ovarii, supporting-like cells (SLC), Sertoli cells, granulosa cells, lineage tracing, PAX8

6

8

56 Introduction

57 Gonadal sex determination is the process by which the bipotential gonad commits to
 58 either the testicular or ovarian fate. In the mouse, the gonadal primordium is established
 59 at embryonic day 10 (E10), and develops from the thickening of the coelomic epithelium
 60 (CE) on the ventromedial surface of the mesonephros (for a review see (1)). At this point,
 61 CE cells have already acquired a molecular signature unique to the gonads, including the
 62 GATA binding protein 4 (*Gata4*), Wilms tumor 1 (*Wt1*), and nuclear receptor
 63 steroidogenic factor 1 (*Nr5a1* or *Sf1*) (2-4). The genital ridge grows through the
 64 proliferation of CE cells that either delaminate and invade the underlying mesenchyme as
 65 a result of basement membrane disintegration or as a result of an oriented division
 66 pattern, so that daughter cells end up within the developing genital ridge (2). Both *in vitro*
 67 and *in vivo* lineage tracing experiments and single cell RNA-seq analyses have revealed
 68 that the ingressing CE cells constitute the most important source of gonadal somatic cells
 69 in both sexes, and contribute to both the supporting and the steroidogenic cell lineages (5-
 70 8). Initially, these XX and XY CE-derived *Nr5a1*⁺ multipotent progenitor cells do not
 71 exhibit obvious sexual dimorphism. By E11.5, a subset of these progenitors adopt a
 72 supporting cell fate, before initiating robust sex-dependent genetic programs. This leads
 73 to the differentiation of this CE-derived supporting lineage into Sertoli and granulosa
 74 cells in the male and female, respectively (7-9). In parallel, another subset of these cells
 75 maintains its multipotent progenitor state and, starting around E12.5 in XY and E13.5 in
 76 XX gonads, acquires a steroidogenic progenitor identity at the origin of Leydig cells and
 77 peritubular myoid cells in the developing testis (10-18).

78 A complete characterization of the different cell lineages forming the bipotential gonad is
 79 essential for our understanding of the process of sex determination. In particular, it
 80 remains unclear whether CE-derived progenitor cells contribute to other uncharacterized
 81 populations of the developing testis and ovary. Recent investigations leveraging
 82 transcriptomics, immunofluorescence and 3D reconstruction revealed that rete testis

9

11

83 precursors possess a dual molecular signature reflecting both a Sertoli (*Sox9*, *Gata4*, *Wt1*,
84 *Nr5a1*) and a rete precursor (*Pax8*, *Krt8* and *Cdh1*) identity, suggesting a common origin
85 with the CE-derived supporting lineage(19-22). The rete testis forms an essential bridging
86 system connecting the seminiferous tubules with efferent ducts. In the adult testis, it
87 forms an anastomosing network of epithelium-lined tubules located at testicular
88 mediastinum (20). It is connected with, and receives the luminal contents of, the
89 seminiferous tubules. These fluids move from the rete into the efferent ducts and from
90 there into the epididymis and further into the vas deferens. The rete ovarii is the female
91 homolog of the rete testis and is characterized by a set of anastomosed tubules located in
92 the hilus of the ovary that may also extend through the ovarian medulla (23, 24). The
93 physiological role of the rete ovarii remains unclear, but it has been proposed to be
94 necessary for the initiation of meiosis and a source of granulosa cells during
95 embryogenesis (25-27). Although rete testis and rete ovarii are essential structures
96 connecting the gonads to the reproductive tract, some important gaps remain in the
97 characterization of rete progenitors, their embryonic origin, transcriptomic signature,
98 specification and sex-specific differentiation.

99 In the present study, we built a transcriptomic atlas of the entire cell population of the
100 developing female and male gonads during the process of mouse sex determination. We
101 identified a previously uncharacterized population of supporting-like cells (SLC) and
102 shed light on the origin and the developmental trajectory of this important cell lineage
103 during the process of testis and ovary development using a combination of single cell
104 RNA sequencing (scRNA-seq) analyses, immunofluorescence, gene expression analysis,
105 *in vivo* cell lineage tracing and loss-of-function experiments.

106

12

Results

Single-cell transcriptional atlas of gonadal sex determination in mice

We generated a single-cell transcriptomic atlas of mouse XX and XY gonads covering the entire process of sex determination and differentiation, from the emergence of the genital ridges at E10.5, to the late fetal gonads at E16.5, using droplet-based 3' end scRNA-seq. We collected and sequenced the gonadal cells at five different developmental stages (E10.5, E11.5, E12.5, E13.5, and E16.5) for a total of 94,705 cells after quality controls (**Figure 1A, B and Material & Methods**). We found that XX and XY cells tend to overlap at E10.5 and E11.5, as expected for time points prior to gonadal sex determination (**Figure 1C and D**). At later stages, cells diverge and display sex-specific gene expression patterns, reflecting the emergence of differentiated cell populations. Gonadal cells were further clustered based on transcriptional similarity using the Leiden algorithm (28) (**Figure 1E**). The clustering identified 43 transcriptionally distinct clusters covering XX and XY cells at different stages of gonadal differentiation. Using enriched marker gene expression, we were able to assign a cell type to each of the 43 clusters (**Figure 1E, F and Figure S1**). The relative abundance of each cell type was then determined for both the genital ridges and mesonephros at E10.5 and E11.5 and for the testes and ovaries at E12.5, E13.5 and E16.5 (**Figure 1G and Table S1**). All major cell populations of the gonad have been identified such as germ cells, Sertoli and granulosa cells, endothelial cells and fetal Leydig cells. It is interesting to note the high proportion of interstitial progenitors in the developing testis and ovary (25% and 58% respectively at E16.5) and the comparatively low abundance of germ cells in the testis (6%) compared to the ovary (34%) at E16.5 as a consequence of mitotic arrest of XY germ cells.

Identification of a previously uncharacterized population of supporting-like cells

Our single-cell transcriptomic analysis also identified all major cell populations derived from the *Nr5a1* lineage, such as supporting progenitors, fetal Leydig, Sertoli, and

17

135 granulosa cells, but also two clusters containing a previously uncharacterized population
 136 of supporting cells, which we refer to as supporting-like cells (SLC). These cells are
 137 present in XX and XY gonads from E11.5 to E16.5 (**Figure 1F**). In total, we identified
 138 1,352 unique SLCs forming an early group consisting of 460 cells at E11.5 (hereafter
 139 referred to as early SLCs) and 892 cells of later stages (E12.5-E16.5, referred to as late
 140 SLCs) both strongly expressing the transcription factor *Pax8* (**Figure 1H**). Although they
 141 are part of the same cell population (see below), SLCs group into two separate clusters
 142 that reflect two distinct differentiation stages. We identified SLCs in both developing
 143 ovaries and testes, with 289 XY and 171 XX early SLCs and 312 XY and 580 XX late
 144 SLCs. We observed that the relative abundance of early and late SLCs decreases over
 145 time compared to other gonadal cell populations, from 2.7% at E11.5 to 0.1-0.3% at
 146 E16.5 (**Figure 1G** and **Table S1**). This is probably due to the significant increase in the
 147 number of cells from other cell types during testicular and ovarian development, rather
 148 than a decrease in the total number of SLCs. Spearman's correlation analysis and
 149 hierarchical clustering revealed that the transcriptome of the early and late SLC
 150 populations share similarities with those of other cell types of the supporting lineage,
 151 such as granulosa, pre-supporting and Sertoli cells (**Figure 1I**). SLCs express CE-derived
 152 somatic cell markers such as *Nr5a1*, *Wt1*, and *Gata4*, but also *Pax8*, a marker of cells of
 153 the mesonephric tubules, the rete testis and rete ovarii (21, 29). In contrast, SLCs do not
 154 express *Pax2*, a well-known marker of mesonephric tubules (30) (**Figure 1J**). Overall,
 155 the expression profile of SLCs suggests that they are closely related to the CE-derived
 156 supporting cells and potentially involved in the formation of the rete ovarii in females,
 157 and the rete testis and its connection to efferent ducts in males.

158

XY SLCs share a common CE-derived origin but diverge from the Sertoli lineage before the time of sex determination

161 To characterize the specification and fate of SLCs in the developing testis, we selected all
 162 of the 11 XY cell clusters expressing the three marker genes present in CE-derived

18

20

gonadal cells (*Nr5a1*, *Gata4* and *Wt1*) between E10.5 and E16.5 (clusters #6, 7, 8, 15, 16, 18, 21, 22, 25, 27 and 39, see **Figure 1E**). We then used the PAGA algorithm (Plass et al., 2018; Wolf et al., 2019) to generate a consolidated lineage graph that included all selected cell types rooted to the E10.5 CE cluster (**Figure 2A**). The reconstruction of the CE-derived cell lineages in the developing XY gonad revealed three distinct developmental trajectories, with all cells emerging initially from the CE-derived progenitor cells at E10.5 and E11.5 (clusters #16 and #18, **Figure 2A-C**). Three different cell lineages emerge from these CE-derived progenitors: the gonadal surface epithelium (SE) that expresses the *Upk3b* marker gene (clusters #8 and #27, **Figure 2A-E**); the supporting cell lineage expressing sequentially *Sry*, *Sox9* and *Amh* during the process of Sertoli cell differentiation (clusters #15 then #25, #21, #22 and #39, **Figure 2A-D and F-H**); and the SLC lineage, expressing *Pax8* as well as *Sox9*, but not *Sry* (see clusters #7 and #6, **Figure 2A-D, G and I**). The *Pax8* gene appears also transiently expressed in 5% of XY pre-supporting cells at E11.5 (threshold defined as 0.5 log normalized counts see also **Figure 2I**) suggesting that its expression is not completely SLC-specific. Cell cycle analysis using scRNA-seq data (31) revealed that cells in the coelomic epithelium, surface epithelium, and Sertoli cells actively proliferate, whereas pre-supporting cells expressing *Sry* and SLCs appear to remain in a quiescent state (**Figure 2J**). The absence of proliferation in the SLC lineage was confirmed by whole-mount immunofluorescence of gonads pulsed with 5-Ethynyl-2'-deoxyuridine (EdU) incorporation. Absence of EdU labelling of PAX8⁺/GATA4⁺ cells confirms that SLCs arrest cycling during the process of sex determination at E11.5 and early testis development at E13.5 (**Figure S2**).

To investigate the differences between SLCs and the supporting lineage in the developing testis, we performed differential expression analysis (32). As the number of SLCs decreases with time, we grouped these cells into two groups to maintain sufficient statistical power for the analysis: early SLCs (cells at E11.5) and late SLCs (cells between E12.5 and E16.5). We observed a large number of differentially expressed genes (DEGs) both at E11.5 (560 genes) and at later stages (814 genes) (**Figure 2K**). Genes

21

23

191 overexpressed in XY SLCs were linked with gene ontology terms such as the “BMP
192 pathway”, “mesonephric tubule development”, and “regulation of TGF β stimulus”,
193 suggesting that although these cells do not originate from mesonephric tubules, they share
194 some common transcriptomic features. Among the DEGs overexpressed in XY SLCs, we
195 find SLC lineage markers such as *Pax8*, *Ncam1*, *Tbx2* and *Ennp2* but also genes of the
196 WNT/ β -catenin signaling pathway such as *Wnt4*, *Lef1*, *Nkd1*, and the transcription factor
197 *Nr0b1* (*Dax1*). Similarly, *Bmp4* and the *Id1*, *Id2* and *Id3* genes, targets of the BMP
198 pathway, are also overexpressed in XY SLCs consistent with the GO term “BMP
199 pathway”. As expected, genes overexpressed in the supporting cell lineage were
200 associated with terms associated with “male gonad development” (see **Data S1** for more
201 details on DEGs and GO biological processes enrichment analysis) and include classical
202 Sertoli cell markers such as *Amh*, *Ptgds*, *Tesc*, *md2*, *Ned9*, *Mro*, *Dhh*, *Cst8*.

203

204 Overall, our lineage tracing reconstruction is consistent with a coelomic origin of the
205 SLCs that subsequently adopt a developmental trajectory independent of the supporting
206 cell lineage before or around the time of sex determination (**Figure 2D**).

207

208 **XX SLCs are already specified by E11.5 and contribute to the pool of granulosa** 209 **cells**

210 To characterize the specification and fate of SLCs in the developing ovaries, we used the
211 PAGA algorithm to generate a consolidated lineage graph that includes all the nine XX
212 cell clusters between E10.5 and E16.5 expressing *Nr5a1*, *Gata4* and *Wt1* (i.e. clusters #5,
213 6, 7, 8, 11, 15, 16, 18 and 19, see **Figure 1D, F-H** and **Figure 2L**). Due to the existence
214 of distinct populations of granulosa cells, as well as cellular heterogeneity in the
215 precursors of the granulosa cell lineage (33-35), the resulting PAGA graph is more
216 complex than its testicular counterpart. Nevertheless, we observed three different cell
217 lineages originating from CE-derived progenitors (clusters #18 and 16, **Figure 2L-O**):
218 the gonadal surface epithelium (SE) cells that expresses the *Upk3b* marker gene (cluster

24

26

219 #8, **Figure 2L-O and P**); the supporting cell lineage sequentially expressing *Foxl2* and
 220 *Lgr5* during the process of granulosa cell differentiation (clusters #15, #11, #5 and #19,
 221 **Figure 2L-O, Q and R**); and the SLC lineage, expressing *Pax8* and *Enpp2* (clusters #7
 222 and #6, **Figure 2L-O, S and T**). As for their male counterpart, the XX SLCs are already
 223 specified by E11.5 (cluster #7). In respect of *Pax8* gene, it also appears transiently
 224 expressed in 5% of XX pre-supporting cells at E11.5 (threshold defined as 0.5 log
 225 normalized counts; see **Table S1** and **Figure 2P**) suggesting that it is not specific to SLCs
 226 in the developing gonad. Cell cycle analysis revealed that CE and SE cells are highly
 227 proliferative, whereas late SLCs are in a quiescent phase (**Figure 2U**). Interestingly,
 228 PAGA analysis showed strong connections between granulosa cells (clusters #11, #5 and
 229 #19) and cells from the SE (cluster #8) as well as late SLCs (cluster #6, **Figure 2L**).
 230 Similarly, in the UMAP representations, the SE cluster (#8) showed a continuous
 231 connection with *Lgr5*-expressing granulosa cells, while the late SLCs were connected to
 232 *Foxl2*-expressing granulosa cells (**Figure 2L-N**). It is therefore likely that granulosa cells
 233 are derived from multiple sources of progenitors, as recently described (36). Our
 234 transcriptomic data suggest that granulosa cells initially derived from pre-supporting cells
 235 around E12.5 (cluster #15), while at later stages, both SE cells (cluster #8) and late SLCs
 236 (cluster #6) may contribute to the granulosa cell pool.

237 Differential expression analysis between XX SLCs and the supporting lineage revealed
 238 that the set of DEGs is significantly smaller compared to XY SLCs and Sertoli cells, both
 239 at E11.5 (450 genes) and at later stages (305 genes) (**Figure 2V**). This confirmed that XX
 240 and XY SLCs are transcriptionally closer to granulosa cells than to Sertoli cells (compare
 241 **Figure 2K and V**). Genes overexpressed in XX SLCs were linked with terms such as
 242 “response to BMP”, “negative regulation of growth factor stimulus”, “mesonephric
 243 tubule development”. It includes SLC marker genes such as *Pax8*, *Ncam1*, *Tbx2* and
 244 *Enpp2* but also the pro-testis factor *Sox9* as well as genes of the WNT/ β -catenin signaling
 245 pathway such as *Wnt4*, *Lef1*, *Nkd1* and the BMP responding genes *Id1*, *Id2* and *Id3*. As
 246 expected genes overexpressed in the granulosa cell lineage were associated with terms

27

29

247 such as “female gonad development” and include genes such as *Lhx9*, *Fst*, *Irx3*, *Bmp2*,
248 *Lgr5* (see **Data S2** for more details on DEGs and GO biological processes enrichment
249 analysis).

250 Overall, our lineage tracing reconstruction suggests that XX SLCs adopt a developmental
251 trajectory independent of the supporting lineage before or around the time of sex
252 determination and may also contribute to the pool of granulosa cells (**Figure 2O**).

253

254 **SLCs are specified at E10.5 and become gradually sexually dimorphic from E12.5**

255 Having shown that both XY and XX SLCs adopt a developmental trajectory that is
256 independent from the supporting cell lineage, we next aimed to identify their
257 specification at E10.5. We next examined the emergence of sexual dimorphism in the
258 SLC lineage as the XY and XX SLCs progress through gonadal development. To this
259 end, we used the PAGA algorithm to generate a consolidated lineage graph that includes
260 the CE clusters at E10.5 (cluster #18) and E11.5 (cluster #16), pre-supporting cell cluster
261 at E11.5 (cluster #15), as well as early and late SLC clusters between E11.5 and E16.5
262 (i.e. clusters #6 and #7, see **Figure 1D, F-H**). Applying a Leiden clustering restricted to
263 CE clusters, we found that the CE-derived progenitors represent a heterogeneous
264 population of cells that can be classified into seven sub-clusters (thereafter named CE-0
265 to CE-6, **Figure 3A**). At E10.5, only three sub-clusters express the classical CE-derived
266 markers *Gata4* and *Nr5a1*, namely CE-1, CE-5 and CE-6. Interestingly, our PAGA graph
267 indicated that sub-clusters CE-1 and CE-5 are connected to sub-cluster CE-6, which is
268 connected to the early and late SLCs clusters (**Figure 3A**). While cells from sub-clusters
269 CE-1 and CE-5 expressed classical CE-derived markers such as *Gata4*, *Nr5a1*, they were,
270 however, *Pax8* negative, therefore representing the canonical CE-derived progenitors. In
271 contrast, cells from sub-cluster CE-6 expressed not only *Gata4* and *Nr5a1*, but also *Pax8*,
272 suggesting that they are in the process of transitioning toward the SLC lineage (see
273 **Figure 3B, C and D**). To better characterize the transcriptomic changes related to the
274 specification of the SLC lineage at E10.5, we then investigated which genes were

30

32

275 differentially expressed between the *Pax8*-expressing cells of the CE-6 sub-cluster and
 276 the other two sub-clusters which represent the E10.5 CE population (CE-1 and CE-5). For
 277 this purpose, we retained only genes with a mean log fold change (LogFC) > 0.25 and an
 278 FDR-adjusted P-value <0.05 (**Data S3**). Among the 112 genes differentially expressed,
 279 we found 59 genes downregulated in the CE-6 sub-cluster including the CE markers
 280 *Lhx9*, *Upk3b* but also genes such as *Gata4*, *Tcf21* and *Wt1*. In addition, we observed 53
 281 upregulated genes including marker genes of the SLC lineage such as *Pax8*, *Ncam1* and
 282 *Enpp2* as well as genes known to be involved in the process of gonadal determination
 283 such as *Nr0b1*, *Sox9*, *Wnt4* as well as *Igf1* and *Igf2*. These results indicate that the SLC
 284 lineage specification is initiated as early as E10.5 in a subset of CE-derived progenitors
 285 with the latter rapidly acquiring SLC transcriptomic markers such as *Pax8*, *Ncam1* and
 286 *Enpp2*.

287 Initially, early XX and XY SLCs at E10.5 and E11.5 remain sexually undifferentiated
 288 and do not exhibit sexual dimorphism. Sexual dimorphism in SLCs develops gradually
 289 from E12.5 onward with the upregulation of Sertoli (e.g. *Ptgds*, *Dhh*) and granulosa (e.g.
 290 *Foxl2*, *Fst*) cell-specific markers in late SLCs (**Figure 3E and F**). Moreover, XX and XY
 291 SLCs at E11.5 co-express the pro-ovarian gene *Wnt4* and the pro-testicular gene *Sox9* at
 292 similar levels (**Figure S3A**), suggesting that both pathways may be active in early SLCs.
 293 This observation is consistent with whole-mount in situ hybridization (WISH) revealing
 294 the presence of *Sox9* and *Wnt4* transcripts in the antero-dorsal part of both E13.5 testis
 295 and ovary where the rete testis and rete ovarii are forming (see arrows in **Figure S3B**).

296 To better characterize how differentiating SLCs acquire their Sertoli or granulosa
 297 characteristics, we trained a LASSO model to determine a score based on the molecular
 298 signature of both cell types. We have identified a set of 167 and 74 genes that define the
 299 molecular signature of Sertoli and granulosa cells, respectively (see **Data S4**). According
 300 to this model, XY and XX progenitor cells of the CE exhibit low Sertoli and granulosa
 301 scores, consistent with their undifferentiated progenitor status (**Figure 3G-I**). In contrast,
 302 and as expected, pre-supporting cells, Sertoli cells and granulosa cells are characterized

33

35

303 by increased Sertoli and granulosa scores. Similarly, from E12.5 onward we observed a
304 progressive increase of the Sertoli score for XY SLCs, and a gradual increase of the
305 granulosa score in XX SLC cells. In-depth analysis indicated that late XY SLCs start to
306 gradually express Sertoli-specific genes such as *Dhh*, *Amh*, *Ptgds*, *Aard*, *Tesc*, *Vnn1*,
307 *Nedd9*, *Cbln4*, while maintaining the expression of genes enriched in SLCs such as *Pax8*
308 and *Ncam1* (**Figure S4**). Similarly, late XX SLCs gradually upregulated classical
309 granulosa markers, such as *Wnt4*, *Foxl2*, *Fst*, *Irx3*, and downregulated *Sox9* while
310 maintaining the expression of *Pax8* and *Ncam1*. Although the increase in expression of
311 Sertoli or granulosa markers in late XY and XX LSCs is significant, it does not reach the
312 expression levels present in the supporting cell line (**Figure S4B and C**). Overall, these
313 results demonstrate that SLC lineage specification occurs around E10.5 and that the
314 establishment of sexual dimorphism in the SLC lineage starts at E12.5 (**Figure 3B**).
315 Then, differentiating XY and XX SLCs gradually acquire a molecular signature similar to
316 Sertoli and granulosa cells respectively, reinforcing the possibility that SLCs have the
317 capacity to differentiate into either of these cell types.

318

319 **SLCs arise in both XY and XX genital ridges at E11.0 and become restricted to** 320 **either the rete testis or rete ovarii**

321 To confirm the results obtained using scRNA-seq and to characterize the spatiotemporal
322 localization of the SLC population during the process of gonadal development, we
323 analyzed the expression of the SLC marker PAX8 by whole-mount immunofluorescence
324 (IF) of wildtype XX and XY genital ridges from E10.5 to E12.5. Our analysis revealed
325 that a population of cells faintly expressing PAX8 and the gonadal somatic cell marker
326 GATA4 at around E11.0 are present throughout the XX and XY genital ridge at the
327 border between the gonad and the mesonephros (**Figure 4**, XY E11.0). These IF data
328 were consistent with the identification of a small cluster of CE-derived gonadal cells
329 expressing *Pax8* at E10.5 by scRNA-seq (**Figure 3A and D**). At early stages, the PAX8⁺
330 gonad/mesonephros border domain is adjacent to the mesonephric tubules, making it

36

38

331 difficult to distinguish SLC PAX8 from mesonephric PAX8. However, PAX8 is
 332 generally expressed at higher levels in the mesonephric tubules than in the
 333 PAX8⁺/GATA4⁺ SLC presumptive rete progenitors. At E11.0, we also observed rare
 334 PAX8⁺/GATA4⁺/SOX9⁺ cells in the anterior part of the XY gonads (arrowhead in **Figure**
 335 **4**). At E11.5, a portion of the PAX8⁺/GATA4⁺ population also expressed SOX9 in both
 336 XY and XX gonads. Similar to Sertoli and granulosa cells, PAX8⁺/GATA4⁺ cells along
 337 the border of the gonad were NR2F2⁻. IF analysis of laminin subunit Beta-1 (LAMB1)
 338 reveals that PAX8⁺/GATA4⁺ cells along the gonad/mesonephros border are associated
 339 with a basement membrane (**Figure 4**, XX E10.5 and XX E12.5). LAMB1 expression is
 340 continuous within the PAX8⁺/GATA4⁺ mesonephric tubules and the PAX8⁺/GATA4⁺
 341 coelomic epithelium from E10.5 to E12.5.
 342 Overall, the PAX8 expression analysis was consistent with our scRNA-seq analysis. IF
 343 further revealed that SLC progenitor cells can be found at the border to the mesonephros
 344 in XY and XX genital ridges as early as E11.0 and that they contribute to the rete testis
 345 and rete ovarii, respectively.

346

347 ***Pax8*⁺ progenitors contribute to both rete testis/ovarii and to the pool of Sertoli and** 348 **granulosa cells**

349 To determine the fate of SLC progenitors and their spatiotemporal localization in the
 350 developing testis and ovary, we first used light sheet fluorescence microscopy on cleared
 351 whole gonads to obtain a high-resolution tridimensional visualization of the urogenital
 352 system. The technique allowed us to examine the anatomical distribution of the targeted
 353 cells in the intact gonad, by preserving the three-dimensional information. The evaluation
 354 was carried out at E13.5 when the rete structure develops, at E16.5, and at birth (P0)
 355 when testis and ovary development is well underway and rete formation is completed
 356 (19, 37). To this end, we labelled *Pax8*-expressing cells and all their derivatives as well
 357 as *Nr5a1*-expressing cells using a transgenic line
 358 *Pax8:Cre;Rosa26:tdTomato;Nr5a1:GFP* (7, 38, 39). In the developing testis (**Figure**

39

41

359 **5A**), RFP⁺ cells were primarily clustered in the cranio-dorsal region of the gonad,
360 corresponding to the rete testis. In addition, a small set of RFP⁺ cells were also observed
361 scattered throughout the testis. Similarly, in the developing ovary (**Figure 5A**), RFP⁺
362 cells were present along the antero-dorsal region of the gonad near the mesonephros. The
363 clustering in the anterior part is not as pronounced as in the testis. We also observed some
364 RFP⁺ cells scattered within the developing ovary, although to a lesser extent than the
365 male counterpart.

366 Lineage tracing experiments using *Pax8:Cre;R26:tdTomato* transgenic animals at E13.5
367 and E16.5 coupled with IF for SOX9 in the testis and FOXL2 (granulosa cell marker) in
368 the ovary revealed that RFP⁺/SOX9⁺ and RFP⁺/FOXL2⁺ cells are predominantly localized
369 in the cranio-dorsal region of the gonad where the rete develops (**Figure S5**). We also
370 observed some Sertoli and granulosa cells expressing RFP, indicating that *Pax8*-
371 expressing cells contribute also to supporting cells.

372 To better characterize the ability of *Pax8*-expressing cells to differentiate into rete cells
373 and supporting cells in XX and XY embryos, testes and ovaries of
374 *Pax8:Cre;R26:tdTomato* transgenic animals were analyzed by IF for RFP and AMH
375 (Sertoli cell marker) or FOXL2 (granulosa cell marker) at P0. In the testis, RFP⁺ cells
376 were mainly detected in the rete but a significant number of positive cells were also
377 identified in testis cords (**Figure 5B**). In the rete testis, RFP⁺ cells are mostly AMH⁺
378 (**Figure 5B** and **inset 1**), while at the interface between the rete testis and the testis cords,
379 some RFP⁺ cells express AMH at a low level (**Figure 5B** and **inset 2**), suggesting that
380 they are Sertoli cells derived from *Pax8*⁺ SLC progenitors. Strikingly, we found that in
381 testis cords, Sertoli cells double positive for RFP and AMH are distributed in a gradient
382 (**Figure 5B** and **inset 3**), with the proportion of doubly positive cells being higher near
383 the rete testis. Double IF analysis revealed that 9.8% of all AMH⁺ Sertoli cells co-express
384 RFP (1,106 double AMH⁺/RFP⁺ cells amongst 11,271 AMH⁺/RFP⁻ and AMH⁺/RFP⁺
385 cells). Finally, to determine if *Pax8*-expressing cells give rise to other cell types in the
386 developing testis, we evaluated the co-expression of RFP⁺ with HSD3B, a marker of

42

44

387 Leydig cells, ACTA2 (also called α SMA), a marker of peritubular myoid cells and
 388 vascular smooth muscle cells, ARX, a specific marker of steroidogenic progenitors, and
 389 NR2F2 (also called COUP-TFII), a marker of interstitial progenitors. With rare
 390 exceptions, none of the RFP⁺ cells co-expressed these markers, consistent with the
 391 supporting-like transcriptomic profile of Pax8⁺ SLC progenitors (**Figure S6**).

392 A similar approach was used to assess the capacity of XX SLC progenitors to
 393 differentiate into rete ovarii cells and granulosa cells (**Figure 5B**). As expected, RFP⁺
 394 cells were detected mostly in the rete ovarii, but a gradient of positive cells was also
 395 identified within the ovary itself. We also found RFP and FOXL2 double-positive cells in
 396 a gradient, labeling granulosa cells from ovigerous cords or primordial follicles (**Figure**
 397 **5B** and **insets 4** and **5**, respectively). Double IF analysis revealed that 6.7% of all
 398 FOXL2⁺ granulosa cells co-express RFP (205 double FOXL2⁺/RFP⁺ cells amongst 3,055
 399 FOXL2⁺/RFP⁻ and FOXL2⁺/RFP⁺ cells). RFP⁺ cells co-expressing the granulosa cell
 400 marker FOXL2 are also found in the ovary at E13.5 and E16.5 (**Figure S5**). Overall,
 401 these results indicate that Pax8⁺ progenitor cells are capable not only of forming the rete
 402 system but also of giving rise to a significant fraction of the Sertoli and granulosa cell
 403 pool.

404

405 **The pro-ovary factor WNT4 is a crucial regulator of the SLC lineage and is**
 406 **required for the formation of rete testis**

407 While the signalling pathway(s) responsible for SLC differentiation and rete testis
 408 development remain unknown, our scRNA-seq data revealed that in the developing
 409 gonads, *Wnt4* is highly expressed in XX and XY SLCs around E11.5, before decreasing
 410 (**Figure 6A** and **Figure S3**). To assess whether WNT4 might act in an autocrine manner
 411 and activate the WNT/ β -catenin signaling pathway in both XX and XY SLCs, we first
 412 investigated the temporal expression of *Rspo1* and downstream targets of the WNT/ β -
 413 Catenin pathway in the developing SLC and supporting cell lineages. As expected,
 414 *Rspo1*, *Axin2*, *Sp5*, *Lef1* and *Nkd1* were highly expressed in granulosa cells but almost

45

47

absent in developing Sertoli cells (**Figure 6A**). By contrast, these genes are expressed in both XX and XY SLCs, indicative of an autocrine effect of WNT4 on SLCs. Initially, at E11.5, the expression levels of *Wnt4* and downstream targets of the WNT/ β -catenin pathway are generally similar in both XX and XY SLCs and then a sexual dimorphism appears at E12.5, with higher expression in XX SLCs compared to XY SLCs.

To test the hypothesis that WNT4 is involved in regulating the SLC lineage and rete testis development, we performed double IF for PAX8 and AMH at E12.5 and E14.5 on testes that were either control or homozygous knockout for the *Wnt4* mutant allele (Jeays-Ward et al., 2004; Stark et al., 1994) (**Figure 6B**). As previously reported, we found a reduction in size and fewer testis cords in the *Wnt4* knockout testis (40). In addition, we observe a significant reduction in both PAX8⁺ and PAX8⁺/AMH⁺ cells (**Figure 6B and C**). The few remaining PAX8⁺ cells are mainly scattered at the base of the testis near the mesonephros but do not form a rete testis. Furthermore, it appears that the testis cords adjacent to the base of the testis and close to the mesonephros are also altered in shape and ill-defined, in particular at E14.5, as if the absence of PAX8⁺ cells and rete testis in *Wnt4* knockout testis prevents the optimal formation of testis cords. Overall, our results reveal that WNT4 is a crucial regulator of the canonical β -catenin signaling pathway in SLCs and is required for the formation of rete testis.

Our scRNA-seq data revealed also that the orphan nuclear receptor DAX1 (encoded by *Nr0b1*) is expressed at high levels in both XX and XY SLCs (**Figure S4**). The function of DAX1 in gonadal development, Sertoli cell differentiation and specification of SLC cells remain unclear and controversial (21, 41). DAX1 has been reported to play a role in testis morphogenesis as *Nr0b1*-deficient mice are infertile due to an obstruction of the rete testis and efferent ductules by dysregulated proliferation of Sertoli cells (42). In parallel, DAX1 has been identified as an all-trans retinoic acid (ATRA) effector gene. Several observations indicate that ATRA drives *Nr0b1* expression directly, antagonizing the testis determination pathway in somatic lineages (43-45). In particular, DAX1 has been reported to antagonize NR5A1 function and impede AMH production in Sertoli cells (46-

48

50

443 48). We decided to evaluate the hypothesis that abnormalities in ATRA signaling could
 444 lead to a defect in the SLC specification and formation of rete testis or rete ovarii. To this
 445 end, we investigated the presence of PAX8-positive cells and the structure of the rete
 446 testis/ovarii in the gonads of E14.5 embryos ubiquitously lacking all nuclear retinoic acid
 447 receptors (RARA, RARB and RARG isotypes) in all cell-types from embryonic day 11.5
 448 onwards (49). Our results revealed that RARs, and therefore endogenous ATRA, does not
 449 affect SLC numbers and are dispensable for rete testis and rete ovarii formation (see
 450 **Figure S7**). Nevertheless, it still remains possible that DAX1 may have a role in
 451 mediating SLC lineage specification and differentiation.

452

453 **Discussion**

454 Although the process of gonadal sex determination has been studied for several decades,
 455 many key aspects remain elusive. Our understanding of gonad development and
 456 differentiation will only be complete after full characterization of all its cell types. We are
 457 still missing a comprehensive understanding of the different cell lineages comprising this
 458 bipotential primordium over time and their relationships to each other. Similarly,
 459 molecular factors governing the specification of the many cell lineages that emerge from
 460 this process are still missing. The widespread adoption of advanced sequencing
 461 technologies such as scRNA-seq, has provided the field of developmental biology with an
 462 opportunity to discover previously unrecognized cell types, such as short-lived
 463 progenitors or rare cell lineages.

464

465 Here, we densely sampled gene expression at the single-cell level in developing mouse
 466 gonads during the critical period of sex determination. We describe the specification and
 467 differentiation of a rare, previously overlooked, gonadal cell lineage. We refer to these as
 468 supporting-like cells (SLCs) due to their transcriptomic similarities to the supporting cell
 469 lineage. Transcriptionally early SLCs are closely related to CE-derived progenitors and
 470 are specified at E10.5, prior to sex determination. These cells are localized in the XX and

51

53

471 XY genital ridge along the border to the mesonephros and express *Pax8* as well as CE-
 472 derived markers, such as *Nr5a1*, *Gata4* and *Wt1*. We found that SLCs are initially
 473 sexually undifferentiated and co-express both the pro-ovarian and pro-testis genes *Wnt4*
 474 and *Sox9*. From E12.5 onward sexual dimorphism appears with the gradual acquisition of
 475 a Sertoli-like and granulosa-like profiles. Finally, our lineage tracing experiments
 476 revealed that *Pax8*⁺ progenitors contribute primarily to the formation of the rete testis and
 477 ovarii and also to a significant fraction of the Sertoli and granulosa cell pool.

478

479 **Origin and specification of SLCs**

480 Our transcriptomic analyses reveal that SLCs are the first somatic cells of the gonad to be
 481 specified, as early as E10.5. By contrast, the commitment of CE-derived progenitors
 482 toward the supporting lineage is initiated just before the process of gonadal sex
 483 determination at E11.0 - E11.5 (7). We found that at E10.5, early somatic progenitor cells
 484 - known to express marker genes such as *Gata4*, *Nr5a1* and *Wt1* - are not a homogeneous
 485 population. Using scRNA-seq, we identified several CE-related sub-populations with
 486 close but distinct transcriptomic profiles, including one that expresses the SLC markers
 487 *Pax8*, *Ncam1* and *Enpp2*, indicating already a transition toward early and late SLCs. At
 488 this early stage (E10.5), the transcriptomes of *Pax8*⁻ and *Pax8*⁺ somatic progenitor sub-
 489 populations are extremely similar with only 112 differentially expressed genes including
 490 the SLC markers *Pax8*, *Ncam1* and *Enpp2* out of a total of about 15,000 genes (see **Data**
 491 **S3**). This single cell analysis strongly suggests that the SLC lineage originates and
 492 diverges transcriptionally as early as E10.5 from somatic progenitor cells expressing the
 493 CE marker genes *Gata4*, *Nr5a1* and *Wt1*. These findings are consistent with our IF
 494 results, revealing that at E11.0, a population of PAX8⁺ cells co-expressing GATA4 is
 495 localized at the base of the genital ridge in close contact with the mesonephric tubules
 496 (see **Figure 4**, **Figure S2**). While these PAX8⁺ cells are initially distributed all along the
 497 gonad at the border to the mesonephros, they are gradually restricted to the anterior part
 498 in close contact with the cranial mesonephric tubules. Our results are also in agreement

54

with the previous description of a population of cells expressing NR5A1 found in both sexes between the cranial mesonephric tubules and the undifferentiated gonad prior to sex determination (22). Finally, the transcriptomic similarities between SLCs and the supporting lineage, as well as the expression of Sertoli (*Dhh*, *Ptgds*, *Aard*, *Neddd9*, *Tesc*) and granulosa (*Foxl2*, *Irx3*, *Fst*) markers by rete testis and rete ovarii cells respectively, reinforce the hypothesis of a CE origin for the SLC lineage. Collectively, our data indicate that SLCs are derived from the CE and share a common origin with the supporting cell lineage.

***Pax8*⁺ progenitors contribute to both rete testis/ovarii and to the pool of Sertoli and granulosa cells**

By coupling lineage tracing analysis with immunofluorescence, we observed three distinct RFP⁺ cell types in the developing testis. The first type is RFP⁺/AMH⁻ and is localized in the center of the rete testis (**Figure 5G-K** and **inset 1**). These cells form the core network of interconnected tubules of the rete testis. The second population expresses RFP and low levels of AMH and is located at the interface between the rete testis and the testicular cords (**Figure 5G-K** and **inset 2**). These cells most certainly correspond to the transitional zone and/or the tubuli recti, both of which are composed exclusively of Sertoli cells (20, 50, 51). The third population is composed of Sertoli cells that are located within the testis cords and represent 10% of all Sertoli cells at birth (**Figure 5G-K** and **inset 3**). These results indicate that *Pax8*-expressing cells contribute not only to the rete testis but also significantly to the Sertoli cell pool, suggesting a potential unappreciated role for PAX8 in supporting progenitors. However, the remaining question is whether the RFP⁺ Sertoli cells are derived solely from the 5% pre-supporting progenitors that weakly express *Pax8* around E11.5 (see **Figure 2I**), or whether, remarkably, they are derived in part or exclusively from the SLC lineage. Our lineage

59

525 tracing experiment does not allow us to give a definitive answer to this question.
 526 Nevertheless, several indications suggest the possibility that SLC progenitors could be a
 527 second source of Sertoli cells. First, both the SLC and supporting lineages appear to
 528 derive from a common progenitor at around E10.5, which expresses CE-derived markers
 529 (see **Figure 3A, C and D**). Furthermore, the transcriptome of SLCs shows similarities
 530 with cells of the supporting lineage, both at early stages with pre-supporting cells and at
 531 later stages with granulosa and Sertoli cells (see **Figure 2I and J**). For example, from
 532 E12.5 onwards, SLCs acquire either a Sertoli- or granulosa-like identity with
 533 overexpression of many classical Sertoli or granulosa markers (**Figure S4**). Furthermore,
 534 our lineage tracing analysis of *Pax8*⁺ cells reveals a gradient of RFP⁺ Sertoli cells in the
 535 testis, with a higher density near the rete testis, suggesting that SLC progenitors may
 536 indeed contribute to these RFP⁺ Sertoli cells (**Figure 5A-B**). The ability of these potential
 537 SLC-derived Sertoli cells to colonize all regions of the testis reflects their early inclusion
 538 in the expanding testis cords following the rapid proliferation of fetal Sertoli cells.
 539 Finally, it also remains possible that at around E10.5 the specification of some early SLC
 540 progenitors is not yet well established redirecting a small fraction of these cells to a
 541 supporting fate instead. This could explain the transient presence of *Pax8* transcripts in
 542 5% pre-supporting progenitors around E11.5.

543 In the developing ovary, our results revealed that RFP⁺ cells have the ability to
 544 differentiate into either rete ovarii cells or granulosa cells. The RFP⁺/FOXL2⁻ cells
 545 forming the rete ovarii are concentrated mostly in the cranio-dorsal part of the ovary
 546 consistent with the IF and light sheet microscopy at E13.5 and E16.5 (**Figure 5A-B** and
 547 insets **4 & 5**). Regarding granulosa cells, our quantitative analyses reveal that about 6%
 548 of FOXL2-expressing granulosa cells are derived from *Pax8*⁺ progenitors. As for the
 549 testis, we observe a gradient of these cells, with RFP⁺/FOXL2⁺ double-positive cells
 550 mainly concentrated in the cranio-dorsal part of the ovary near the rete ovarii. In the
 551 cortical part, we also observe primordial follicles with some granulosa cells being RFP⁺.

552

60

62

553 **Factors regulating SLC specification, differentiation and rete formation**

554 Although our combined scRNA-seq analysis and lineage tracing approach has allowed us
 555 to better characterize the origins and developmental trajectories of the SLC progenitor
 556 cells that give rise to the rete system, some questions still remain. In particular, the
 557 factors, local signals and molecular pathways responsible for the specification of SLC
 558 progenitors remain unknown. We decided to focus our attention on the secreted protein
 559 WNT4 because of its high expression in the SLC lineage (**Figure S3**) and the essential
 560 role of the canonical WNT/ β -catenin signaling pathway for ovarian fate, particularly in
 561 the supporting cell lineage (52-56). We tested the possibility that WNT signaling also
 562 plays a role in the SLC lineage and the formation of the rete testis. Indeed, the high
 563 expression of *Wnt4* and its downstream target genes (*Lef1*, *Axin2*, *Sp5* and *Nkd1*) in the
 564 SLC lineage is consistent with an autocrine effect (**Figure 6A**). Direct experimental
 565 evidence found in the literature already suggested that canonical β -catenin signaling is
 566 activated in the rete ovarii/testis of the gonads at E12.5 (52). Using a reporter line
 567 carrying *LacZ* fused to *Axin2* (57), a target of β -catenin, Chassot and coworkers observed
 568 that at E12.5, a robust β -galactosidase staining was found at the interface between the XX
 569 and XY gonads and the mesonephros, a position corresponding to the rete testis and rete
 570 ovarii at this developmental stage.

571 Our results in *Wnt4* knockout mutant mice show a drastic reduction in the number of
 572 SLCs at E14.5, as well as disorganization of the rete testis and adjacent testicular cords
 573 (**Figure 6B**). This confirms that WNT4 and the canonical WNT/ β -catenin signaling
 574 pathway are required for rete testis formation. However, it is not clear whether the
 575 reduction in the number of PAX8⁺ cells and the alteration of the rete testis formation is
 576 due to a reduction in the total number of SLCs or to a change in their fate. The former is
 577 unlikely, as our transcriptomic data combined with results from EdU DNA synthesis
 578 monitoring assay clearly indicate that SLCs do not proliferate during the process of
 579 testicular determination. This absence of proliferation may explain why the rete system is
 580 a relatively small structure compared to the whole gonad (**Figures 2J and S2**). We

63

65

581 hypothesize instead that the ability of progenitors to maintain their SLC identity and form
582 the rete testis is dependent on the canonical WNT/ β -catenin pathway. Loss of the
583 WNT/ β -catenin pathway may therefore affect the fate of SLC progenitors. Since the rete
584 testis forms an essential bridging system connecting the testis cords to rete testis, an
585 indirect effect is the alteration of the morphology of the testicular cords in close
586 proximity to the defective rete structure, as observed in *Wnt4* knockout mutant testes. If
587 the WNT pathway is indeed key to the differentiation of SLC progenitors towards rete
588 testis fate, what is their fate in its absence or upon overexpression of *Wnt4* or the
589 canonical WNT pathway more generally? An increase in the WNT/ β -catenin pathway
590 could be tested by genetic studies promoting stabilization of β -catenin or ablation of
591 *Znrf3*, a negative regulator of ovarian development. Loss of *Znrf3* results in ectopic
592 canonical WNT signaling in XY gonads with reduced *Sox9* expression, resulting in
593 defects in testis determination, including gonadal sex reversal (58). We predict that
594 elevated WNT/ β -catenin signaling would promote SLC fate and increase the final number
595 of *Pax8*⁺ SLCs in the XY mutant ovary/ovotestis.

596

597 Overall, our results provide a single-cell transcriptional atlas of gonadal sex
598 determination in mice. This unique dataset allowed us to identify SLCs as a previously
599 uncharacterized cell population at the origin of the rete system. Understanding this
600 lineage is important for two reasons: first, proper formation of the rete is essential for
601 male fertility, as it provides a channel system allowing sperm export from the testis to the
602 efferent ductules. Second, from a fundamental point of view, the SLC lineage exhibits
603 remarkable characteristics, and represents an interesting model for the study of cell fate
604 decisions during mammalian organ development. We demonstrate that the SLC lineage is
605 the first somatic cell lineage to be specified, even before the initiation of the gonadal sex
606 determination process. We are unable to definitely demonstrate that the SLC lineage
607 possesses the capacity to differentiate into Sertoli-like or granulosa-like cells, a feature
608 that was previously thought to be unique to the supporting cell lineage. However, several

66

68

609 observations support this conclusion. The molecular mechanism involved in Sertoli-like
 610 differentiation is different from that of the supporting lineage, since it does not involve
 611 the well-known pro-testis gene *Sry*. It is now essential to identify the factors that control
 612 the specification of the SLC lineage as well as the molecular mechanisms controlling
 613 their differentiation into rete cells exhibiting transcriptomic profiles close to Sertoli or
 614 granulosa cells. In addition, a better characterization of the similarities and differences
 615 between the supporting and SLCs lineages during the process of testis/ovarian
 616 development is needed. With our results, we lay the foundation for further elucidation of
 617 mechanisms in mammalian sex determination, and improving our understanding of the
 618 genetic basis and physiology of human disorders/differences of sex development.

69

71

619 **Material & Methods**

620 *Animals*

621 Animals were housed and cared according to the ethical guidelines of the Service de la
622 Consommation et des Affaires Vétérinaires (SCAV) of the Canton de Genève
623 (experimentation ID GE214-19 and GE35) or the relevant institutional and European
624 animal welfare laws, guidelines and policies for the *Wnt4*^{KO} line.
625 *Pax8*^{tm1.1(cre)Mbu/J;Gt(ROSA)26Sor^{tm14(CAG-tdTomato)Hze}·Tg(Nr5a1/EGFP)1Klp} (abbreviated
626 *Pax8:cre;Rosa26:Tomato;Nr5a1:eGFP*) mouse strains were described previously (30, 39,
627 59). These mice have been maintained on a mixed 129/CD1 genetic background and were
628 genotyped between P6 and P11 from digit biopsies by PCR as described previously (39).
629 For *Pax8:cre* knock-in allele, primers used were 5'-TCTCCACTCCAACATGTCTGC-
630 3', 5'-AGCTGGCCCAAATGTTGCTGG-3' (30) and 5'-
631 GCCAGCAGCTATGAGGTTGA-3', giving a WT band of 601 bp and a specific Cre
632 amplicon of 673 bp. Control (*Sf1:Cre;Sox9*^{F/+}; *Wnt4*^{KO}+/+) and *Wnt4*^{KO} mutant
633 (*Sf1:Cre;Sox9*^{F/+}; *Wnt4:KO*^{-/-}) embryos were generated and genotyped as described in (60)
634 and the mouse line was kept on a mixed 129/C57Bl6/J background.

635

636 *Sample collection and single cell transcript profiling*

637 The day when a vaginal plug was designated as embryonic day E0.5. Embryos were
638 collected at E10.5 (8±2 caudal somites), E11.5 (19±4 ts), E12.5, E13.5, and E16.5 from
639 time-mated pregnant female CD-1 outbred mice (Charles River) and heterozygous
640 Tg(Nr5a1-GFP) transgenic male mice (39). The *Nr5a1-GFP* transgene was used to
641 facilitate identification of the position of the genital ridges during the dissection. Genetic
642 sexing was performed by PCR, as described previously (61). Urogenital ridges including
643 both the gonads and mesonephros (E10.5, E11.5) from both sexes, and testes or ovaries
644 (E12.5-E16.5), were isolated at each time point and enzymatically dissociated as
645 described previously (62). Approximately 5000 single cells were loaded on a 10x
646 Chromium instrument and scRNA-seq libraries were prepared using the 10X Genomics

72

74

647 Chromium Controller and Chromium Single Cell 3' Library & Gel Bead Kit V2, as per
648 the manufacturer's instructions. For each developmental stage and sex, two independent
649 biological replicates were processed. The prepared libraries were sequenced on an
650 Illumina HiSeq4000 using paired-end 26 + 98 + 8 bp as the sequencing mode. Libraries
651 were sequenced at a targeted depth of 100 000 to 150 000 total reads per cell. Sequencing
652 was performed at the Health 2030 Genome Center, Geneva.

653

654 *Data pre-processing with the Cell Ranger package, cell selection and quality controls*

655 Computations were performed at the Vital-IT Center for high-performance computing of
656 the SIB (Swiss Institute of Bioinformatics) (<http://www.vital-it.ch>). Demultiplexing,
657 alignment, barcode filtering and UMI counting were performed with the Cell Ranger v2.1
658 pipeline (10x Genomics). Data were mapped to the mouse reference genome GRCm38.p5
659 in which the eGFP (NC_011521.1) combined with the bovine GH
660 3'-splice/polyadenylation signals (39) (NM_180996.1) sequences have been added. It has
661 recently been established that the mouse *Sry* locus harbors a cryptic exon that is essential
662 for male sex determination (63). As this exon is located in a palindromic region, all reads
663 mapping to this region are also mapping to the positive strand and discarded as
664 multimappers. In order to include reads from exon 2 of the *Sry* gene in our counts
665 matrices, reads from the negative strand, mapped to location chrY(2653159, 2655636)
666 with number of alignments (NH) tag inferior to 3 were extracted and GN tag was set to
667 "Sry". The edited reads were saved in a specific bam and reads were deduplicated and
668 counted using umi-tools count function with --extract-umi-method=tag --per-cell --cell-
669 tag=CB --per-gene --gene-tag=GN --umi-tag=UB parameters.

670

671 To select barcodes associated to cells, we set a threshold as the local minima of a
672 standardized density curve of UMI counts located between the knee point and the
673 inflection point of the ranked barcodes distribution plot (DropletUtils package). When no
674 local minimum could be detected between the two points, the nearest local minimum was

75

77

675 used. Quality controls regarding abnormal mitochondrial or ribosomal content, UMI
676 number, detected gene numbers, unmapped reads and putative doublet identification
677 (Scrublet 0.2.1-0) were performed, but no data filtering was applied as no important
678 anomalies were detected. In total, for gonadal samples, we obtained 92,267 cells. It
679 included 13,060 cells from E9.0, 14,962 cells from E10.5, 16,670 cells from E11.5,
680 20,285 cells from E12.5, 25,794 cells from E13.5 and 16,994 cells from E16.5. In total,
681 for gonadal samples, we obtained 94,705 cells. It included 14,962 cells from E10.5,
682 16,670 cells from E11.5, 20,285 cells from E12.5, 25,794 cells from E13.5 and 16,994
683 cells from E16.5. In total, we obtained 53,962 XY cells and 40,743 XX cells.

684

685 *Gene expression normalization*

686 UMI counts per gene per cell were divided by the total UMI detected in the cell with
687 exclusion of very highly expressed genes (accounting to more than 5% of total counts in
688 cell), multiplied by a scale factor of 10,000 and log transformed.

689

690 *Dimensionality reduction, Batch correction, Clustering and cell-type annotation*

691

692 A first step of dimensionality reduction was performed using Independent Component
693 Analysis (icafast function, ica package version 1.0-2) on all genes expressed in more than
694 50 cells (100 components). Because ICA extracts non-Gaussian components from data
695 and creates non-orthogonal components (as in PCA) it allows a better discrimination of
696 small cell populations in heterogeneous dataset. We set the number of ICs to 100 as we
697 expected to have fewer than 100 different populations in our total dataset. Another
698 advantage of ICA is its good performance with nearly no gene filtering. We did not
699 perform any highly variable gene selection.

700 To minimize biological or technical variability on clustering, we created a neighbor graph
701 based on ICA components using the BBKNN python module (version 1.3.12) (64). The
702 correction was performed between biological replicates. BBKNN corrects the neighbor

78

80

703 graph by identifying the k-nearest neighbors in each batch instead of considering all cells
704 come from one unique batch. No correction of the expression levels was performed.
705 A first clustering was performed using the Scanpy Leiden method with resolution 1.3.
706 Clusters were annotated using known markers from described cell populations of
707 developing testis and ovary. To refine the clustering, a leiden clustering was applied to
708 the SLC cluster and granulosa cells to obtain early and late SLC. The parameters to
709 achieve this final clustering of the cells were resolution=0.37, restrict_to=('leiden', ["5",
710 "30"]).

711

712 *Lineage reconstruction, UMAP visualization*

713 Lineage reconstruction was done using partition-based graph abstraction (PAGA) (65).
714 From the global k-nearest-neighbors graph and clustering, PAGA outputs a much simpler
715 graph with nodes representing clusters and edges showing confidence of connection
716 between clusters. PAGA and its visualization (Fruchterman-Reingold) were run prior to
717 UMAP to allow a better organization of clusters. For complete dataset global
718 visualization, UMAP was run with random_state=25, min_dist=0.2, init_pos ="paga"
719 parameters. (**Figure 1, Figure S1**).

720 In order to precisely reconstruct the SLC cell lineage, we created three subsets of data:
721 the first two groups contain XX or XY cells from clusters composed of cells from the
722 coelomic epithelium, surface epithelium, pre-supporting cell population, granulosa and
723 Sertoli cells as well as early and late SLC. The third subset is composed of XX and XY
724 cells from clusters composed of cells of the coelomic epithelium, as well as early and late
725 SLCs. Clusters with less than 50 cells were removed (a few XY cells clustering with
726 granulosa cells for instance). We observed that in some conditions PAGA outputs
727 stronger connections with small clusters. To address this issue, we randomly subsampled
728 big clusters to obtain balanced clusters and ran PAGA on this subsampled data. BBKNN,
729 PAGA and UMAP steps were re-ran independently on XX, XY and SLC datasets.

730

81

83

731 *Correlation analysis*

732 The similarity of the transcriptomes between the different cell populations was
733 determined using Spearman correlation analysis on the mean expression of the genes. We
734 then applied a hierarchical clustering on the correlation matrix using ward linkage (66).
735 The obtained dendrogram was reused to order cell types in figure 1J.

736

737 *Differential expression analysis*

738 Differential expression analysis was performed using the Seurat (version 3.2.2) wrapper
739 function for MAST (version 1.14.0). Genes were selected with a log-foldchange higher
740 than 0.25 and an adjusted *p*-value below 0.05. Gene ontology enrichment was performed
741 positive or negative genes obtained with differential expression analysis using
742 clusterProfiler package (version 3.16.1) and only on Gene Ontology biological processes
743 database. Background was set to genes detected in more than 50 cells in the complete
744 dataset and *p*-value cutoff was set to 0.01.

745

746 *Scoring of the cells*

747 To create a classification score for the different lineages of gonad specific cells, we
748 trained a Least Absolute Shrinkage and Selection Operator (LASSO) with cross-
749 validation classifier on a random subset of cells with a one versus all approach. The
750 LASSO approach uses a L1 regularization and allows a stringent selection of the genes
751 defining the cells of interest (<https://www.jstor.org/stable/2346178>). We used the
752 LassoCV function from scikit-learn python module. To avoid gene weighting bias due to
753 over-representation of some clusters, the cells were randomly subsampled keeping 300
754 cells per leiden clusters.

755

756 *Histological and Immunofluorescence analyses*

757 Embryonic or postnatal samples were collected at specific time points determined by the
758 day of vaginal plug (E0.5) and fixed overnight at 4°C in 4% paraformaldehyde. Samples

84

86

759 were embedded in paraffin and 5µm sections were prepared. Unmasking was performed
760 in TEG (10mM, pH 9) or citrate (10 mM, pH 6) buffer for 15 minutes in a pressure
761 cooker. Sections were incubated in blocking buffer (3% BSA, 0.1% Tween in PBS) for 2
762 hours at room temperature. Primary antibodies were incubated overnight at 4°C.
763 Secondary antibodies were incubated 1 hour at room temperature with DAPI (1:1000).
764 The following primary antibodies were used: sheep anti-SOX9-CT (1:200, a generous
765 gift from Prof. Francis Poulat, University of Montpellier, France), rabbit anti-FOXL2
766 (1:300, (67)), mouse anti-αSMA (1:500, a generous gift from Christine Chaponnier,
767 University of Geneva, Switzerland), rabbit anti-ARX (1:500, a generous gift from Prof.
768 Ken-ichirou Morohashi, Kyushu University, Japan), goat anti-AMH (C-20) (1:100, sc-
769 6886, SantaCruz Biotechnology, USA), mouse anti-AMH (1:20, MCA2246, Bio-Rad
770 Laboratories), rabbit anti-PAX8 (1:200, 10336-1-AP, Proteintech), mouse anti-RFP
771 (DsRed) (1:250, sc-390909, SantaCruz Biotechnology, USA), goat anti-RFP (1:50, 200-
772 101-379, Rockland Immunochemicals Inc., USA), rabbit anti-RFP (1:250, 600-401-379,
773 Rockland Immunochemicals Inc., USA), mouse anti-COUPTFII (1:200, PP-H7147-00,
774 Perseus Proteomics Inc., Japan), rabbit anti-HSD3B (1:200, KO607, TransGenic Inc,
775 Japan). Each immunostaining run included negative controls, with replacement of the
776 primary antibody by blocking solution. Fluorescent images were acquired using an Axio
777 Imager M2 or Z1 microscope (ZEISS, Germany) fitted with an AxioCam 702 mono
778 camera or MRm camera (ZEISS, Germany). Images were minimally processed for
779 global levels with ZEN (ZEISS, Germany). Whole gonad images were analyzed using
780 QuPath v0.2.0 (68). The percentages of RFP⁺/AMH⁺ or RFP⁺/FOXL2⁺ co-labeled cells
781 were determined from three independent animals. With respect to *Wnt4*^{KO} embryos,
782 paraffin sections covering the entire gonad were processed for immunostaining and the
783 number of PAX8⁺ cells or PAX8⁺ AMH⁺ cells was determined every 20µm using the cell
784 counter Plugin from Fiji. The graphs show the number of positive cells per 10 000µm² of
785 gonadal area measured in each section, the mean and s.e.m for control (n=3) and *Wnt4*^{KO}
786 mutant embryos (n=3).

87

89

787

788 *Whole-mount immunofluorescence*

789 For whole-mount immunofluorescence (IF), embryonic samples were fixed overnight at
790 4°C, rinsed three times for 20 minutes with PBS 0.1% Triton X-100 (PBST) and serially
791 dehydrated into 100% methanol for storage at -20°C. Immunofluorescence was
792 performed as follows: samples were rehydrated stepwise into PBS, washed three times
793 with PBST, rocked for 1 hour at room temperature in blocking solution (BS: 10% FBS,
794 3% BSA, 0.2% Triton X-100 in PBS), and incubated with primary antibodies in BS
795 overnight at 4°C. The following day, samples were washed three times for 30 minutes in
796 washing solution (WS: 1% FBS, 3% BSA, 0.2% Triton X 100 in PBS) and incubated
797 with secondary antibodies and DAPI (1:1000) in BS overnight at 4°C. The next day,
798 samples were washed three times in WS and mounted in DABCO (Sigma Aldrich, St
799 Louis, USA). Images were taken with an LSM710 confocal microscope using the
800 associated Zen software (Zeiss). The following primary antibodies were used: rabbit anti-
801 PAX8 (1:500, 10336-1-AP, Proteintech), mouse anti-GATA4 (1:500, Sc-25310, Santa
802 Cruz Biotechnology), goat anti-SOX9 (1:1000, AF3075, R&D Systems), rat anti-LAMB1
803 (1:500, RT-796-P, NeoMarkers), mouse anti-NR2F2 (1:250, PP-H7147-00, Perseus
804 Proteomics).

805

806 *EdU labeling*

807 For cell proliferation experiments, pregnant females were injected intraperitoneally with
808 25 mg/kg 5-ethynyl-2'-deoxyuridine (EdU; Lumiprobe) dissolved in PBS. Embryonic
809 samples were collected 1 hour after EdU injection and fixed overnight at 4°C. EdU
810 treated samples were imaged using the wholemount immunofluorescence protocol with
811 the addition of a click reaction step between rehydration and blocking. Samples were
812 incubated in click reaction solution (20 mg/mL ascorbic acid, 2mM cupric sulfate, 4 µM
813 sulfo-Cy3 azide dye (Lumiprobe) in PBS) for 1 hour rocking at room temperature.

814

90

92

815 *Sample clearing and light sheet fluorescence microscopy (LSFM)*

816 *Pax8:Cre;Rosa26:tdTomato;Nr5a1:eGFP* embryos at E11.5, E12.5, E13.5, E16.5 were
817 cleared using a modified passive CLARITY-based clearing protocol (69) while the
818 samples at embryonic stages E10.5 and dissected gonads at P0 went through a refractive-
819 index matching process. Briefly, samples at E11.5, E12.5, E13.5 and E16.5 embryonic
820 stage were incubated in 5 mL X-CLARITY™ Hydrogel Solution Kit (C1310X, Logos
821 Biosystems, South Korea) for 3 days at 4°C, allowing diffusion of the hydrogel solution
822 into the tissue. Polymerization of solution was carried in a Logos Polymerization system
823 (C20001, Logos Biosystems, South Korea) at 37°C for 3 hours. After two washes of 30
824 min in PBS, samples were immersed in a SDS based clearing solution and left at 37°C for
825 24h (E11.5), 48h (E12.5), 72h (E13.5) and 96h (E16.5). Once cleared, tissue was washed
826 twice in PBS-TritonX 0.1% and then placed in a Histodenz© based-refractive index-
827 matching solution (RI = 1.46). Early-stage embryos (E10.5) underwent a uniquely
828 refractive index-matching process, prior to 2 days of incubation in PBS-TritonX 2% at
829 room temperature (RT). Dissected gonads of embryos at P0 were embedded in low
830 melting agarose, placed in PBS-TritonX 2% for 2 days at RT and transferred in a
831 Histodenz© based-refractive index-matching solution (RI = 1.46) 2 days before imaging.
832 Light-sheet imaging was performed using a customized version of the Clarity Optimized
833 Light-sheet Microscope (COLM) (Tomer et al., 2014) at the Wyss Center Advanced
834 Light-sheet Imaging Center, Geneva. The samples were illuminated by one of the two
835 digitally scanned light sheets, using a 488 nm and 561 nm wavelength laser. Emitted
836 fluorescence was collected by 10X XLFLUOR4X N.A. 0.6 filtered (525/50 nm and
837 609/54 nm Semrock BrightLine HC) and imaged on an Orca-Flash 4.0 LT digital CMOS
838 camera at 4 fps, in rolling shutter mode. A self-adaptive positioning system of the light
839 sheets across z-stacks acquisition ensured optimal image quality over the whole thickness
840 of the sample. Z-stacks were acquired at 3 µm spacing with a zoom set at 10x resulting in
841 an in-plane pixel size of 0,59 µm (2048 x 2048 pixels). Images of multiple tiles were
842 reconstructed in 3D using the Grid Collection Stitching plugin tool in TeraStitcher (BMC

93

95

843 Bioinformatics, Italy). Volume reconstructions and 3D renderings were performed using
844 Imaris Bitplane software.

845

846 *Statistical analyses*

847 The number of PAX8⁺ cells or PAX8⁺AMH⁺ cells per unit of gonadal area in control and
848 Wnt4^{KO} mutant embryos was analyzed by a two-tailed t-test using Graph Pad Prism
849 version 9.

850

96

98

851 References

852

- 853 1. E. Rotgers, A. Jorgensen, H. H. Yao, At the Crossroads of Fate-Somatic Cell Lineage
854 Specification in the Fetal Gonad. *Endocr Rev* **39**, 739-759 (2018).
- 855 2. Y. C. Hu, L. M. Okumura, D. C. Page, Gata4 is required for formation of the genital
856 ridge in mice. *PLoS Genet* **9**, e1003629 (2013).
- 857 3. Y. Ikeda, W. H. Shen, H. A. Ingraham, K. L. Parker, Developmental expression of mouse
858 steroidogenic factor-1, an essential regulator of the steroid hydroxylases. *Mol Endocrinol*
859 **8**, 654-662 (1994).
- 860 4. J. A. Kreidberg *et al.*, WT-1 is required for early kidney development. *Cell* **74**, 679-691
861 (1993).
- 862 5. J. Karl, B. Capel, Sertoli cells of the mouse testis originate from the coelomic epithelium.
863 *Developmental Biology* **203**, 323-333 (1998).
- 864 6. J. Schmahl, E. M. Eicher, L. L. Washburn, B. Capel, Sry induces cell proliferation in the
865 mouse gonad. *Development* **127**, 65-73. (2000).
- 866 7. I. Stevant *et al.*, Dissecting Cell Lineage Specification and Sex Fate Determination in
867 Gonadal Somatic Cells Using Single-Cell Transcriptomics. *Cell Rep* **26**, 3272-3283
868 e3273 (2019).
- 869 8. I. Stevant *et al.*, Deciphering Cell Lineage Specification during Male Sex Determination
870 with Single-Cell RNA Sequencing. *Cell Rep* **22**, 1589-1599 (2018).
- 871 9. S. Nef, I. Stevant, A. Greenfield, Characterizing the bipotential mammalian gonad.
872 *Current topics in developmental biology* **134**, 167-194 (2019).
- 873 10. Y. T. Lin, L. Barske, T. DeFalco, B. Capel, Numb regulates somatic cell lineage
874 commitment during early gonadogenesis in mice. *Development* **144**, 1607-1618 (2017).
- 875 11. h. Ademi, I. Stévant, C. M. Rands, B. Conne, S. Nef, Expression of Wnt5a defines the
876 major progenitors of fetal and adult Leydig cells. *BioRxiv*, (2020).
- 877 12. I. B. Barsoum, H. H. Yao, Fetal Leydig cells: progenitor cell maintenance and
878 differentiation. *J Androl* **31**, 11-15 (2010).
- 879 13. R. K. Bhandari, E. N. Schinke, M. M. Haque, I. Sadler-Riggleman, M. K. Skinner, SRY
880 induced TCF21 genome-wide targets and cascade of bHLH factors during Sertoli cell
881 differentiation and male sex determination in rats. *Biol Reprod* **87**, 131 (2012).
- 882 14. J. Brennan, C. Tilmann, B. Capel, Pdgfr-alpha mediates testis cord organization and fetal
883 Leydig cell development in the XY gonad. *Genes Dev* **17**, 800-810 (2003).
- 884 15. S. Cui *et al.*, Disrupted gonadogenesis and male-to-female sex reversal in Pod1 knockout
885 mice. *Development* **131**, 4095-4105 (2004).
- 886 16. M. Inoue *et al.*, Isolation and Characterization of Fetal Leydig Progenitor Cells of Male
887 Mice. *Endocrinology* **157**, 1222-1233 (2016).
- 888 17. C. Liu, K. Rodriguez, H. H. Yao, Mapping lineage progression of somatic progenitor
889 cells in the mouse fetal testis. *Development* **143**, 3700-3710 (2016).
- 890 18. K. Miyabayashi *et al.*, Aristaless related homeobox gene, Arx, is implicated in mouse
891 fetal Leydig cell differentiation possibly through expressing in the progenitor cells. *PLoS*
892 *ONE* **8**, e68050 (2013).

99

100

101

893 19. A. Y. Kulibin, E. A. Malolina, Formation of the rete testis during mouse embryonic
894 development. *Dev Dyn* **249**, 1486-1499 (2020).
895 20. A. T. Major, M. A. Estermann, C. A. Smith, Anatomy, Endocrine Regulation and
896 Embryonic Development of the Rete Testis. *Endocrinology*, (2021).
897 21. E. A. Malolina, A. Y. Kulibin, The rete testis harbors Sertoli-like cells capable of
898 expressing DMRT1. *Reproduction (Cambridge, England)* **158**, 399-413 (2019).
899 22. T. Omotehara, X. Wu, M. Kuramasu, M. Itoh, Connection between seminiferous tubules
900 and epididymal duct is originally induced before sex differentiation in a sex-independent
901 manner. *Dev Dyn* **249**, 754-764 (2020).
902 23. J. McKey, D. N. Anbarci, C. Bunce, B. Capel, Integration of mouse ovary morphogenesis
903 with developmental dynamics of the
904 oviduct, ovarian ligaments, and rete ovarii. *bioRxiv preprint*, (2021).
905 24. J. G. Wenzel, S. Odend'hal, The mammalian rete ovarii: a literature review. *Cornell Vet*
906 **75**, 411-425 (1985).
907 25. A. G. Byskov, Does the rete ovarii act as a trigger for the onset of meiosis? *Nature* **252**,
908 396-397 (1974).
909 26. A. G. Byskov, N. E. Skakkebaek, G. Stafanger, H. Peters, Influence of ovarian surface
910 epithelium and rete ovarii on follicle formation. *J Anat* **123**, 77-86 (1977).
911 27. P. Smith, D. Wilhelm, R. J. Rodgers, Development of mammalian ovary. *J Endocrinol*
912 **221**, R145-161 (2014).
913 28. V. A. Traag, L. Waltman, N. J. van Eck, From Louvain to Leiden: guaranteeing well-
914 connected communities. *Sci Rep* **9**, 5233 (2019).
915 29. C. Bunce, J. McKey, B. Capel, Concerted morphogenesis of genital ridges and nephric
916 ducts in the mouse captured through whole-embryo imaging. *Development* **148**, (2021).
917 30. M. Bouchard, A. Souabni, M. Mandler, A. Neubuser, M. Busslinger, Nephric lineage
918 specification by Pax2 and Pax8. *Genes Dev* **16**, 2958-2970 (2002).
919 31. R. Satija, J. A. Farrell, D. Gennert, A. F. Schier, A. Regev, Spatial reconstruction of
920 single-cell gene expression data. *Nature biotechnology* **33**, 495-502 (2015).
921 32. G. Finak *et al.*, MAST: a flexible statistical framework for assessing transcriptional
922 changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome*
923 *Biol* **16**, 278 (2015).
924 33. L. Mork *et al.*, Temporal differences in granulosa cell specification in the ovary reflect
925 distinct follicle fates in mice. *Biol Reprod* **86**, 37 (2012).
926 34. R. H. Rastetter *et al.*, Marker genes identify three somatic cell types in the fetal mouse
927 ovary. *Dev Biol* **394**, 242-252 (2014).
928 35. W. Zheng *et al.*, Two classes of ovarian primordial follicles exhibit distinct
929 developmental dynamics and physiological functions. *Hum Mol Genet* **23**, 920-928
930 (2014).
931 36. W. Niu, A. C. Spradling, Two distinct pathways of pregranulosa cell differentiation
932 support follicle formation in the mouse ovary. *Proc Natl Acad Sci U S A* **117**, 20015-
933 20026 (2020).
934 37. H. Nakata, S. Iseki, Three-dimensional structure of efferent and epididymal ducts in
935 mice. *J Anat* **235**, 271-280 (2019).
936 38. M. Bouchard, A. Souabni, M. Busslinger, Tissue-specific expression of cre recombinase
937 from the Pax8 locus. *Genesis* **38**, 105-109 (2004).

102

103

104

39. N. R. Stallings *et al.*, Development of a transgenic green fluorescent protein lineage marker for steroidogenic factor 1. *Endocr Res* **28**, 497-504 (2002).
40. K. Jeays-Ward, M. Dandonneau, A. Swain, Wnt4 is required for proper male as well as female sexual development. *Dev Biol* **276**, 431-440 (2004).
41. F. J. Barrionuevo, M. Burgos, G. Scherer, R. Jimenez, Genes promoting and disturbing testis development. *Histol Histopathol* **27**, 1361-1383 (2012).
42. B. Jeffs *et al.*, Blockage of the rete testis and efferent ductules by ectopic Sertoli and Leydig cells causes infertility in Dax1-deficient male mice. *Endocrinology* **142**, 4486-4495 (2001).
43. J. Bowles *et al.*, Retinoic Acid Antagonizes Testis Development in Mice. *Cell Rep* **24**, 1330-1341 (2018).
44. S. Mahony *et al.*, Ligand-dependent dynamics of retinoic acid receptor binding during early neurogenesis. *Genome Biol* **12**, R2 (2011).
45. F. Nagl *et al.*, Retinoic acid-induced nNOS expression depends on a novel PI3K/Akt/DAX1 pathway in human TGW-nu-I neuroblastoma cells. *Am J Physiol Cell Physiol* **297**, C1146-1156 (2009).
46. M. Ito, R. Yu, J. L. Jameson, DAX-1 inhibits SF-1-mediated transactivation via a carboxy-terminal domain that is deleted in adrenal hypoplasia congenita. *Mol Cell Biol* **17**, 1476-1483 (1997).
47. M. W. Nachtigal *et al.*, Wilms' tumor 1 and Dax-1 modulate the orphan nuclear receptor SF-1 in sex-specific gene expression. *Cell* **93**, 445-454 (1998).
48. S. Y. Park *et al.*, Nuclear receptors Sf1 and Dax1 function cooperatively to mediate somatic cell differentiation during testis development. *Development* **132**, 2415-2423 (2005).
49. N. Vernet *et al.*, Meiosis occurs normally in the fetal ovary of mice lacking all retinoic acid receptors. *Sci Adv* **6**, eaaz1139 (2020).
50. R. A. Hess, L. Hermo, in *Encyclopedia of Reproduction*. (Elsevier, 2018), pp. 263-269.
51. E. C. Roosen-Runge, A. F. Holstein, The human rete testis. *Cell Tissue Res* **189**, 409-433 (1978).
52. A. A. Chassot *et al.*, Activation of beta-catenin signaling by Rspo1 controls differentiation of the mammalian ovary. *Hum Mol Genet* **17**, 1264-1277 (2008).
53. C. F. Liu, N. Bingham, K. Parker, H. H. Yao, Sex-specific roles of beta-catenin in mouse gonadal development. *Hum Mol Genet* **18**, 405-417 (2009).
54. D. M. Maatouk, L. Mork, A. A. Chassot, M. C. Chaboissier, B. Capel, Disruption of mitotic arrest precedes precocious differentiation and transdifferentiation of pregranulosa cells in the perinatal Wnt4 mutant ovary. *Dev Biol* **383**, 295-306 (2013).
55. B. Nicol, H. H. Yao, Gonadal Identity in the Absence of Pro-Testis Factor SOX9 and Pro-Ovary Factor Beta-Catenin in Mice. *Biol Reprod* **93**, 35 (2015).
56. S. Vainio, M. Heikkila, A. Kispert, N. Chin, A. P. McMahon, Female development in mammals is regulated by Wnt-4 signalling. *Nature* **397**, 405-409. (1999).
57. B. Lustig *et al.*, Negative feedback loop of Wnt signaling through upregulation of conductin/axin2 in colorectal and liver tumors. *Mol Cell Biol* **22**, 1184-1193 (2002).
58. A. Harris *et al.*, ZNRF3 functions in mammalian sex determination by inhibiting canonical WNT signaling. *Proc Natl Acad Sci U S A* **115**, 5474-5479 (2018).
59. L. Madisen *et al.*, A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nat Neurosci* **13**, 133-140 (2010).

105

106

107

- 984 60. F. Tang, N. Richardson, A. Albina, M. C. Chaboissier, A. Perea-Gomez, Mouse Gonad
985 Development in the Absence of the Pro-Ovary Factor WNT4 and the Pro-Testis Factor
986 SOX9. *Cells* **9**, (2020).
- 987 61. L. McFarlane, V. Truong, J. S. Palmer, D. Wilhelm, Novel PCR assay for determining
988 the genetic sex of mice. *Sex Dev* **7**, 207-211 (2013).
- 989 62. C. Mayere *et al.*, Single-cell transcriptomics reveal temporal dynamics of critical
990 regulators of germ cell fate during mouse sex determination. *FASEB J* **35**, e21452 (2021).
- 991 63. S. Miyawaki *et al.*, The mouse Sry locus harbors a cryptic exon that is essential for male
992 sex determination. *Science* **370**, 121-124 (2020).
- 993 64. K. Polanski *et al.*, BBKNN: fast batch alignment of single cell transcriptomes.
994 *Bioinformatics* **36**, 964-965 (2020).
- 995 65. F. A. Wolf *et al.*, PAGA: graph abstraction reconciles clustering with trajectory inference
996 through a topology preserving map of single cells. *Genome Biol* **20**, 59 (2019).
- 997 66. J. H. Ward, Hierarchical Grouping to Optimize an Objective Function. *Journal of the*
998 *American Statistical Association* **58**, 236-244 (1963).
- 999 67. J. C. Polanco, D. Wilhelm, T. L. Davidson, D. Knight, P. Koopman, Sox10 gain-of-
1000 function causes XX sex reversal in mice: implications for human 22q-linked disorders of
1001 sex development. *Hum Mol Genet* **19**, 506-516 (2010).
- 1002 68. P. Bankhead *et al.*, QuPath: Open source software for digital pathology image analysis.
1003 *Sci Rep* **7**, 16878 (2017).
- 1004 69. K. Chung *et al.*, Structural and molecular interrogation of intact biological systems.
1005 *Nature* **497**, 332-337 (2013).
- 1006

1007

108

Acknowledgments:

We thank Deborah Penet for the sequencing, Christelle Borel (GEDEV department, University of Geneva) for her advice and help with 10X technology, the team of the Animal Facility and the Histology Plateform (Faculty of Medicine, University of Geneva). We thank also the members of the Nef laboratory for helpful discussion and critical reading of the manuscript.

Funding

Swiss National Science Foundation grant 31003A_173070 (SN)
 Swiss National Science Foundation grant 310030_200316 (SN)
 Département de l'Instruction Publique of the State of Geneva (SN)
 Agence Nationale de la Recherche grant ANR-19-CE14-0022-01 SexDiff (MCC)
 UK Medical Research Council grant MC_U142684167 (AG)
 National Institutes of Health grant 1R01HD090050 (BC, CB)
 National Institutes of Health grant R37HD039963 (BC)
 NBG and CD were supported by grants from Agence Nationale de la Recherche (ANR-16-CE14-0017 and ANR-20-CE14-0022)

Author Contributions

SN, DW and AG contributed to conception and design of the study. IS, YN, and FK collected mouse samples and prepared sequencing libraries. CM, PS, RR, SG, MS performed the bioinformatics analysis. VR, APG, CB, PS, DC, PS, CD, IG, FT and FK carried out mouse work, cell lineage tracing, immunofluorescence, light sheet microscopy and other experimental analysis. CM, VR and SN wrote the first draft of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version. Funding and Resources, SN, AG, BC, MCC, LB, NBG and DW; Supervision, SN.

113

1036 **Competing Interests**

1037 The authors declare no competing interests.

1038

1039 **Data and materials availability**

1040 All data needed to evaluate the paper are present in the paper and/or the Supplementary

1041 Materials. Additional data related to this paper may be requested from the authors.

1042

Figure legends:

Figure 1. A single-cell atlas of gonadal development and sex differentiation. (A)

Schematic overview of the experimental procedure and dataset. A representative embryo at E11.5 is shown with (in inset) the region containing the gonads and mesonephros. (B) Barplots show the number of single cells RNA sequenced per stage and sex. (C-F) UMAP representation of the 94,705 cells colored by sex (B), stage (C), Leiden clustering (D) and annotation of the different cell clusters (E). (G) Proportion of cell types across sex and developmental stages. (H) UMAP representation of the expression level of *Pax8* gene. (I) Stacked violin plots showing expression (log-normalized counts) of major somatic cell markers of the gonad in annotated cell populations. Populations were ordered using hierarchical clustering based on correlation (Spearman) distance of expression levels between cell populations. (J) Heatmap showing correlation between transcriptomes of the different cell populations. Populations were ordered as in (I). Abbreviations: CE, coelomic epithelial cells; SE, surface epithelial cells; Pre-Sup., pre-supporting cells; Sertoli, Sertoli cells; Granulosa, granulosa cells; SLC, supporting-like cells; EIP, early interstitial progenitors; LIP, late interstitial progenitors; FLC, fetal Leydig cells; IM, invading mesonephric cells; PV, perivascular cells, Imm., immune cells; Meso., mesonephric mesenchymal cells; MT, mesonephric tubules; End., endothelial cells; AS, adrenosympathic cells; GC, germ cells; EC, erythrocytes.

Figure 2. Characterization of XY and XX supporting-like cell (SLC) populations.

The analysis was performed with both XY (A-K) and XX (L-V) gonadal cells. PAGA representation of the clusters of XY (A) and XX (L) cells expressing *Nr5a1*, *Gata4* and *Wt1*. Each node is a Leiden cluster and the links between clusters represent confidence of relation between two clusters. (B, C, M, N) UMAP projection of gonadal cell lineage at the origin of the supporting cells and SLCs, colored by cell annotation (B, M) and by developmental stage (C, M). (D) Simplified model of the cell lineage reconstruction in

the developing testis (**D**) and ovary (**O**). UMAP representations of XY (**E-I**) and XX (**P-T**) cells colored by expression levels (log-normalized counts) of selected genes. UMAP and barplot representation of the cell cycle phase of XY (**J**) and XX (**U**) cells. Barplot illustrates the proportion of cells in the different cell cycle phase for each population. (**K, V**) Barplots representing the number of differentially-expressed genes between SLCs and the supporting lineage either in XY (**K**) or XX (**V**) developing gonads at early (E11.5) and late (E12.5, E13.5 and E16.5) stages. Abbreviations: CE, coelomic epithelial cells; SE, surface epithelial cells; P-Sup., pre-supporting cells; Sertoli, Sertoli cells; Granulosa; *e*-SLC, early supporting-like cells; *l*-SLC, late supporting-like cells.

Figure 3. Sex-specific differentiation of XX and XY SLCs. (A) UMAP projection of SLC cell lineage colored by developmental stage, genetic sex and annotation with subclustering applied to the CE. The fourth panel shows the PAGA graph representation where each node is a Leiden cluster and the links between clusters represent confidence of relation between two clusters. (B) Schematic model of SLC lineage specification. (C-F) UMAP representation colored by expression levels (log-normalized counts) of CE-derived markers *Gata4* and *Nr5a1* (C), SLC markers *Pax8* and *Enpp2* (D), Sertoli cell markers *Ptgds* and *Dhh* (E), and granulosa cell markers *Fst* and *Foxl2* (F). (G) Scatterplots representing granulosa and Sertoli scores of individual cells using a LASSO model. A plot was generated for each cell type and individual cells were colored by developmental stage. (H) Pointplots showing granulosa and Sertoli cell scores in XX and XY SLC and supporting cells across time. (I) Barplots representing the number of significantly over-expressed genes found when comparing XX to XY SLCs at early (E11.5), late (E12.5, E13.5 and E16.5) or all stages.

Figure 4. A population of gonadal cells expressing both GATA4 and PAX8 is present from E10.5-E11.0 along the genital ridge next to the mesonephros. Whole-

mount immunofluorescence of XY and XX genital ridges from E10.5 to E12.5. Boxes in merged images indicate regions shown as isolated channels on right. The dotted line correspond to the border that separates the GATA4⁺ cells of the genital ridge (G) from the mesonephric cells (M). LAMB1 is expressed in the basement membrane throughout the CE and mesonephric tubules. SOX9 overlaps with a subset of PAX8⁺/GATA4⁺ cells in XY and XX tissues at E11.5. NR2F2 is a marker of gonad interstitial cells and mesonephric mesenchyme. Arrowhead indicates SOX9⁺/PAX8⁺ cell within the XY gonad at E11.0. Arrow indicates PAX8⁺/GATA4⁺ cells adjacent in the region of the presumptive rete that are not within mesonephric tubules. Scale bar = 50um.

Figure 5. Pax8⁺ progenitors contribute to both rete testis/ovarai and the pool of Sertoli and granulosa cells. (A) 3D reconstruction of *Pax8:Cre;Rosa26:tdTomato;Nr5a1:GFP* testis and ovary at E13.5, E16.5 and P0. Pax8⁺ cells are lineage-traced with RFP. GFP, expressed under the control of the *Nr5a1* promoter, is used to delineate gonadal cells. Note the presence of RFP⁺ cells close to the rete testis/ovarai but also throughout the gonads. White arrowhead indicates the Wolffian duct and empty arrowhead the Müllerian duct. (B) Double immunofluorescence for RFP/AMH and RFP/FOXL2 respectively in XY and XX *Pax8:Cre;Rosa26:tdTomato;Nr5a1:GFP* mice at P0. Note the presence of RFP⁺ cells increases near the rete testis and rete ovarii. Cells in the rete testis are exclusively RFP⁺ and do not express AMH (inset 1). At the junction between the rete testis and the testis cords, low AMH expression is observed in some RFP⁺ cells (inset 2). RFP⁺/AMH⁺ cells are also present in testis cords (inset 3). Similarly, RFP⁺/FOXL2⁺ cells are present in granulosa cells of the developing ovary (inset 4) including in some primordial follicles (inset 5). Red arrows indicate RFP⁺ cells, yellow arrows indicate RFP⁺/AMH⁺ or RFP⁺/FOXL2⁺ cells. DAPI was used as a nuclear counterstain. MT, mesonephric tubules; RT, rete testis. Scale bars 200 µm in (A) and 100 µm in (B).

Figure 6. *Wnt4* is required for rete testis formation. (A) Expression profiles of genes involved in WNT/ β -catenin pathway in SLCs and supporting cells. Data were extracted from scRNA-seq analysis of XX and XY embryos at E11.5, E12.5, E13.5 and E16.5. (B) Representative double immunofluorescence against PAX8 and AMH in control (*Sf1:Cre;Sox9^{F/+};Wnt4KO^{+/+}*) or *Wnt4^{KO}* (*Sf1:Cre;Sox9^{F/+};Wnt4:KO^{-/-}*) XY embryos at E12.5 and E14.5. Boxes indicate regions shown on the right magnifying the rete testis and adjacent testis cords. Note that the number of PAX8⁺ cells, the rete testis and the testis cords near the rete testis are reduced and disorganized in *Wnt4* mutant embryos both at E12.5 and E14.5. White and yellow arrowheads indicate PAX8⁺ and PAX8⁺/AMH⁺ cells, respectively. DAPI was used as a nuclear counterstain. Scale bar 100 μ m. (C) Quantification of PAX8⁺ and PAX8⁺/AMH⁺ cells per 10,000 μ m² of testis section of control and *Wnt4^{KO}* embryos at E14.5. Data are presented as a box-and-whisker plot to illustrate the heterogeneity of PAX8-positive cells according to the different sections. Each point represents the quantification of one testicular section. Three embryos per genotype were used. Student's t-test, two-sided unpaired (*p < 0.05).

Figure 1

Figure 2

138

139

140

Figure 5

1147

141

Figure 6