**Spring 2020 – Systems Biology of Reproduction**
**Discussion Outline – Gametogenesis / Stem Cells / Cloning**
**Michael K. Skinner – Biol 475/575**
**CUE 418, 10:35-11:50 am, Tuesday & Thursday**
**March 26, 2020**
**Week 11**

# Gametogenesis / Stem Cells / Cloning

**Primary Papers:**

1. Ball, et al. (2016) BMC Genomics 17:628
2. Valli, et al. (2014) Fert & Steril 101:3-13
3. Guo, et al. (2020) Cell Stem Cell 26:262-276

## Discussion

Student 25: Reference 1 above

- What regulatory factor was identified?
- How was spermatogonial stem cell renewal identified?
- What insights into spermatogonial stem cells was provided?

Student 26: Reference 2 above

- What are the technologies reviewed?
- What are the limitations of organ cultures?
- What clinical applications for germ cell transplantation exist?

Student 1: Reference 3 above

- What developmental stage and cell types were investigated?
- What technology was used, and advantages compared to other technologies?
- What conclusions were made about the onset of testis function and gametogenesis?

Ball et al. BMC Genomics (2016) 17:628
DOI 10.1186/s12864-016-2865-1

BMC Genomics

RESEARCH ARTICLE Open Access

# Regulatory complexity revealed by integrated cytological and RNA-seq analyses of meiotic substages in mouse spermatocytes

Robyn L. Ball[1†], Yasuhiro Fujiwara[1†], Fengyun Sun[1], Jianjun Hu[1], Matthew A. Hibbs[1,2], Mary Ann Handel[1*] and Gregory W. Carter[1*]

## Abstract

**Background:** The continuous and non-synchronous nature of postnatal male germ-cell development has impeded stage-specific resolution of molecular events of mammalian meiotic prophase in the testis. Here the juvenile onset of spermatogenesis in mice is analyzed by combining cytological and transcriptomic data in a novel computational analysis that allows decomposition of the transcriptional programs of spermatogonia and meiotic prophase substages.

**Results:** Germ cells from testes of individual mice were obtained at two-day intervals from 8 to 18 days post-partum (dpp), prepared as surface-spread chromatin and immunolabeled for meiotic stage-specific protein markers (STRA8, SYCP3, phosphorylated H2AFX, and HISTH1T). Eight stages were discriminated cytologically by combinatorial antibody labeling, and RNA-seq was performed on the same samples. Independent principal component analyses of cytological and transcriptomic data yielded similar patterns for both data types, providing strong evidence for substage-specific gene expression signatures. A novel permutation-based maximum covariance analysis (PMCA) was developed to map co-expressed transcripts to one or more of the eight meiotic prophase substages, thereby linking distinct molecular programs to cytologically defined cell states. Expression of meiosis-specific genes is not substage-limited, suggesting regulation of substage transitions at other levels.

**Conclusions:** This integrated analysis provides a general method for resolving complex cell populations. Here it revealed not only features of meiotic substage-specific gene expression, but also a network of substage-specific transcription factors and relationships to potential target genes.

**Keywords:** Meiosis, Spermatogenesis, Maximum covariance analysis, Mouse, Transcriptome, RNA-seq

## Background

Spermatogenesis is a complex developmental process with a unique cell division, meiosis, as a major defining event. The entire process includes maintenance of a small population of spermatogonial stem-cells, mitotic divisions of differentiating spermatogonia, meiotic prophase and ensuing divisions of spermatocytes, and post-meiotic differentiation of spermatids, by a process known as spermiogenesis. In mammalian testes, spermatogenesis occurs within seminiferous tubules, where all germ cells associate with one kind of somatic cell, the Sertoli cell, which provides the appropriate niche and microenvironment for the spermatogenic process. The adult testis is characterized by presence of all of the cells types in the spermatogenic lineage, with waves of differentiation throughout the testis propelled by retinoic acid signaling [1]. In mice, the first wave of spermatogenesis is initiated by spermatogonia shortly after birth, producing a sequential and orderly appearance of each of the more differentiated stages at regular intervals though the first four weeks of life. Although semi-synchronous with respect to the advancing wave of the most differentiated cells, the juvenile onset of

* Correspondence: maryann.handel@jax.org; greg.carter@jax.org
†Equal contributors
[1]The Jackson Laboratory, Bar Harbor, ME, USA
Full list of author information is available at the end of the article

spermatogenesis also includes regular and asynchronous initiation of subsequent waves of spermatogenic differentiation. This asynchronous and continuous process of spermatogenesis has made it difficult to achieve molecular characterization of specific cell types in the lineage. This has been particularly the case with respect to analysis of the defining process of gametogenesis, meiosis, which occurs in spermatocytes. The complex events of meiosis I prophase include recombination, homologous chromosome pairing, and synapsis, taking place as spermatocytes progress through the leptotene, zygotene, pachytene and diplotene substages. These events culminate in the first meiotic division, a reductive division in which homologous chromosomes are separated, producing secondary spermatocytes that rapidly undergo the second, equational, meiotic division to produce haploid round spermatids. Because of the genetic importance of meiotic recombination for the production of chromosomally normal gametes and offspring, there has been great interest in elucidating the molecular hallmarks and their underlying transcriptional signatures that define the meiotic spermatocyte substages.

Toward the goal of achieving a molecular understanding of spermatogenesis, considerable effort has been devoted to separation of specific germ-cell differentiation substages from the histologically complex seminiferous epithelium. A widely used approach, commonly known as the "STA-PUT" sedimentation process [2–4], involves enzymatic dissociation of germ and somatic cells and enrichment for specific stages by cell-size-based sedimentation at unit gravity on a bovine albumin gradient. Reasonably good enrichment of the most uniquely sized cells (large pachytene spermatocytes and small round spermatids) can be obtained from testes of adult mice. However, many cells of interest (spermatogonia, early meiotic prophase spermatocytes) are not retrieved from adults because of their anatomical position bounded by Sertoli cell tight junctions. While these early cell types can be retrieved from juvenile testes, the amount of biological material required is daunting and enrichment is not robust. Isolation of specific cell types by fluorescence-activated cell sorting, FACS [5–8] is promising and becoming more widely applied; however sample sizes are small. Finally, rather than cell separation, total testis germ-cell populations can be collected in a developmental continuum during the first two to three weeks of juvenile development of mouse testes, to take advantage of the first wave of spermatogenesis. In this way, molecular features have been defined with respect to the sequential appearance of more advanced cell stages of spermatogenesis. This approach is useful primarily for correlating the appearance of a molecular entity to the developmental appearance of a specific cell type. However, the degree of resolution has been suboptimal, because gene or protein expression has not been related to absolute frequencies of cell stages, a challenge we tackle in this study.

Together, these methods for enriching or inferring spermatogenic (and meiotic) substages have contributed to studies over the past decade on the developmental transcriptome of mammalian spermatogenesis, as recently reviewed [9]. Most such studies have taken advantage of microarrays querying known coding sequences [6, 10–12], and thus sequence-biased, but also provided interesting views of alternative splicing and other features of the spermatogenic transcriptome [13], and identified previously unknown potential targets for contraception and fertility [14]. More recently, methods for unbiased high-throughput deep sequencing of the transcriptome by RNA-seq have been employed [5, 15–19]. These studies have revealed unappreciated regulation of piRNAs [17], and global whole-genome views of the male germ-cell transcriptome [5, 18, 19]. However, the challenge for all transcriptomic analyses, particularly RNA-seq approaches, has been to computationally deconstruct the entire testis or germ-cell transcriptome into substage-specific transcriptomes. This is an important goal, given both the abundance and complexity of RNA species in the testis, particularly with respect to the imputed contribution of both coding and non-coding RNA from spermatocytes and spermatids [19]. In one example of such a computational approach, frequencies of specific germ-cell stages throughout the first wave of spermatogenesis previously published in another analysis [4] were used to estimate cell type-specific expression patterns in a separate data set [18]. While coming closer to the goal of substage-specific transcriptomes, this study relied on a low sample size and on integrating non-contemporaneous data. Other approaches [5, 6], employed cell sorting by FACS and subsequent validation of purity by meiotic markers. While these studies have yielded important insights into global gene expression switches, they rely on FACS, not always available or practical for small samples, such as from infertile mutant models.

Here we have tackled the problem of deconvolving transcriptomes of complex germ-cell populations into stage-specific transcriptomes by computationally integrating highly detailed and combinatorial cytological staging of the same cell samples as subjected to RNA-seq analysis. Deep and accurate cell stage phenotyping was conducted using antibodies to STRA8, SYCP3, phosphorylated H2AFX, and HISTH1T, all exhibiting meiotic substage specificity of expression and/or localization. By collecting highly enriched germ-cell populations from testes at two-day intervals through the first wave of spermatogenesis, and by multiple sampling of individual mice (30 total) for both cytological composition and RNA composition by high-throughput sequencing, we developed an unusually deep data set that portrays variation in cell substage

composition and transcript abundance. To decompose these data into substage-specific transcriptome patterns in an unbiased way, we chose a nonparametric solution with minimal assumptions about data structure, thus developing a novel permutation-based maximum covariance analysis. This method enabled us to assign co-expressed transcripts to one or more meiotic substages, thereby linking distinct molecular programs to cytologically defined cell states. Moreover, to better understand the regulation of each germ-cell substage transcriptome, we integrated transcription factor information to identify some of the key molecular regulators driving these cell stages. This approach provides a model for deconvolving transcriptomes of complex cell populations with well-defined cytological attributes. Together, these data provide an unprecedented view of the complexity of meiotic transcription programs and their coordinate regulation.

## Results

### Experimental design

Germ cells were obtained from individual mice at two day intervals from 8 dpp to 18 dpp ($N = 5$ biological replicates at each age; $N = 30$ samples total) and divided into two aliquots, one for cytology and one for RNA-seq (Fig. 1a). For the C57BL/6J (B6) strain used in this study, this developmental window captures the initial differentiation of spermatogonia into early leptonema, through the stages of pachynema and diplonema. For each sample,

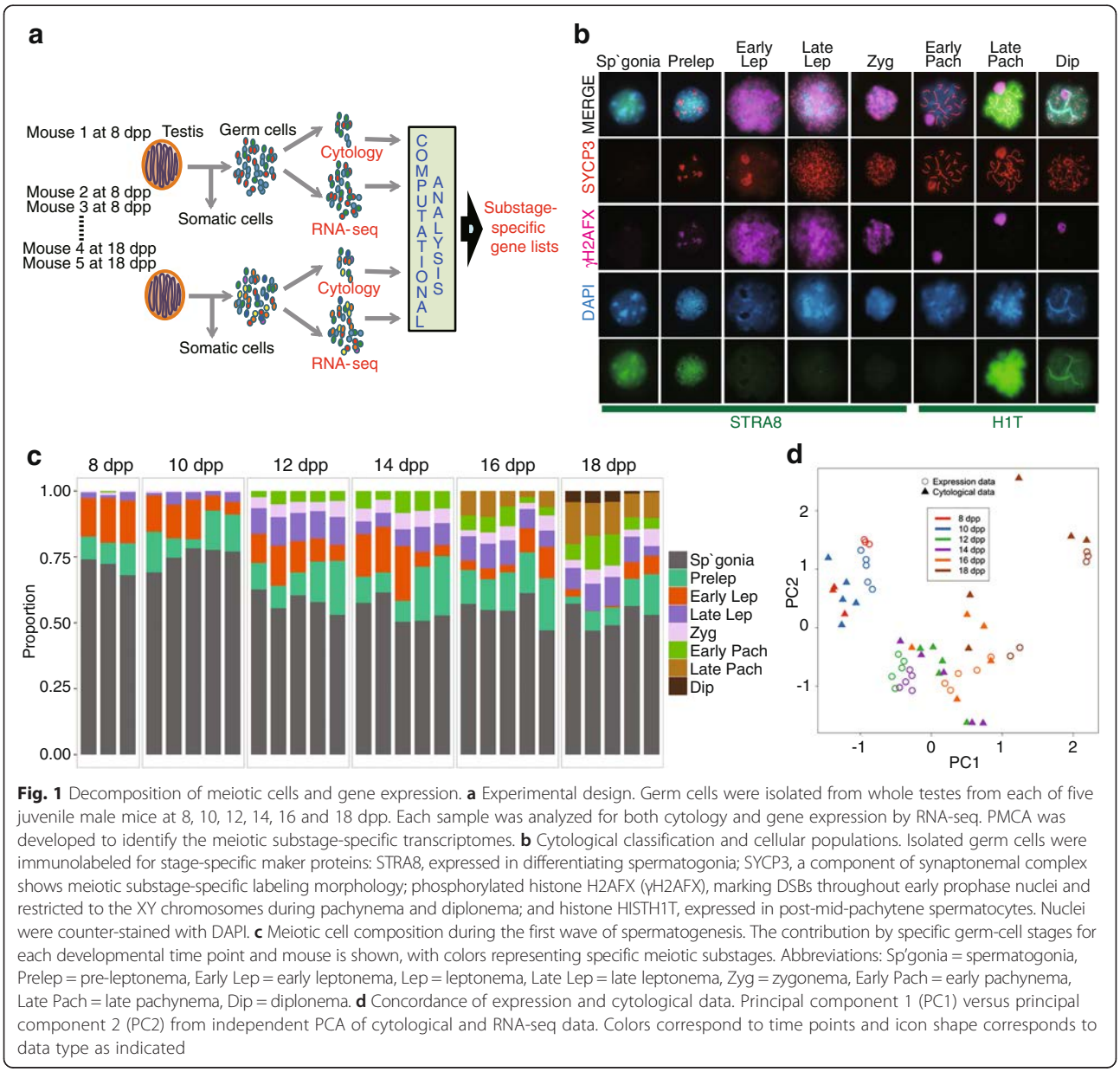



**Fig. 1** Decomposition of meiotic cells and gene expression. **a** Experimental design. Germ cells were isolated from whole testes from each of five juvenile male mice at 8, 10, 12, 14, 16 and 18 dpp. Each sample was analyzed for both cytology and gene expression by RNA-seq. PMCA was developed to identify the meiotic substage-specific transcriptomes. **b** Cytological classification and cellular populations. Isolated germ cells were immunolabeled for stage-specific maker proteins: STRA8, expressed in differentiating spermatogonia; SYCP3, a component of synaptonemal complex shows meiotic substage-specific labeling morphology; phosphorylated histone H2AFX (γH2AFX), marking DSBs throughout early prophase nuclei and restricted to the XY chromosomes during pachynema and diplonema; and histone HISTH1T, expressed in post-mid-pachytene spermatocytes. Nuclei were counter-stained with DAPI. **c** Meiotic cell composition during the first wave of spermatogenesis. The contribution by specific germ-cell stages for each developmental time point and mouse is shown, with colors representing specific meiotic substages. Abbreviations: Sp'gonia = spermatogonia, Prelep = pre-leptonema, Early Lep = early leptonema, Lep = leptonema, Late Lep = late leptonema, Zyg = zygonema, Early Pach = early pachynema, Late Pach = late pachynema, Dip = diplonema. **d** Concordance of expression and cytological data. Principal component 1 (PC1) versus principal component 2 (PC2) from independent PCA of cytological and RNA-seq data. Colors correspond to time points and icon shape corresponds to data type as indicated

we determined the relative numbers of cells of each substage by cytological criteria ($N$ = about 400 germ cells per mouse, for a total of 11,990 germ cells scored cytologically by criteria described below); by scoring germ cells from each individual mouse, we captured variability between mice. The purity of germ cells in each sample, computed as germ-cell count divided by total DAPI-stained cell count, was greater than 90 % for all samples. Upon inspection of both cytological and RNA-seq data, two germ-cell samples collected at 8 dpp revealed poor concordance with age-matched samples, likely due to insufficient cell numbers for successful RNA-seq library preparation (Additional file 1: Table S1); thus these samples were omitted from all subsequent analyses (final $N$ = 28).

### Cytological deconvolution of the meiotic substage composition of germ cells during the first wave of spermatogenesis

Germ-cell substage frequencies in each cytological preparation were determined by combinatorial immunolabeling with antibodies to meiotic substage marker proteins, scoring 400 germ cells per mouse. Marker proteins assessed were STRA8, a meiosis-initiating factor [20] present in differentiated spermatogonia and some leptotene spermatocytes; SYCP3, a synaptonemal complex (SC) protein present in the chromosomal axes of leptotene and zygotene spermatocytes, in the lateral elements of the mature SC in pachytene spermatocytes, and disassembling in diplotene spermatocytes [21, 22]; the phosphorylated form of histone H2AFX (PH2AFX, also known as γH2AX), which localizes to chromatin surrounding DNA double-strand breaks (DSBs) in characteristic patterns that discriminate early meiotic prophase from the pachytene and diplotene stages [22, 23]; and the testis germ cell-specific histone H1 variant HIST1H1T (herein referred to by its common designation of H1T), which is a marker of the mid-to-late pachytene stage [22]. The combinatorial labeling patterns for each marker protein allowed categorization of germ cells of each sample into eight substages, achieving a higher degree of meiotic substage discrimination than previous transcriptome analyses (Table 1 and Fig. 1b).

Together, these data provide a comprehensive picture of postnatal spermatogenic progress through meiotic prophase of the first wave of differentiating cells for comparison to data previously obtained by histological analyses [4]. At every time point in the current data, greater than 50 % of the retrieved germ cells were spermatogonia (Fig. 1c), reflecting continually initiating separate waves of spermatogenesis. Representation of these cells decreased over time (Fig. 1c), likely due to initiation of additional waves of meiotic cells, increasing numbers of post-spermatogonial spermatocytes, and establishment of cell junctions that impede cell retrieval. Throughout this period, due in part to initiation of subsequent waves of spermatogenesis, the average contribution to the total germ-cell population by preleptotene and leptotene spermatocytes remained relatively constant (Fig. 1c), thus reducing the power to identify transcripts specific to just these substages. The first appreciable numbers of H1T-negative early pachytene spermatocytes were observed at 12 dpp, and H1T-positive late-pachytene cells were abundant by 16-18 dpp, consistent with previously published data [24]. By 18 dpp, less than 10 % of the cells had progressed to the diplotene stage, and only about 0.1 % to metaphase. Overall substage frequencies were similar at 8 and 10 dpp, and also at 12 and 14 dpp (Fig. 1c). For computational analyses (below), we combined substages that exhibited similar frequency patterns across the developmental time span, specifically late leptotene and zygotene substages, and late pachytene and diplotene substages.

### High concordance of gene expression and cytological data

To assess the feasibility of associating gene expression with cytologically-defined substages, we performed independent principal component analyses (PCA) on each data set. The high concordance between the cytological frequency of cell substages and RNA expression data (Fig. 1d) suggested that changes in gene expression in pooled germ cells can be explained by variation in cytological proportions of the

**Table 1** Immunolabeling criteria for cell types and meiotic substages

| Substages | SYCP3 | γH2AX | STRA8 | H1T |
|---|---|---|---|---|
| Spermatogonia[a] | None | Negative | Positive | Negative |
| Preleptonema | Patches | Negative | Positive | Negative |
| Early Leptonema | Patches | Positive | Positive | Negative |
| Late Leptonema | Fine foci | Positive | Weakly positive | Negative |
| Zygonema | Partially paired | Patches | Negative | Negative |
| Early Pachynema | Paired | Restricted to XY body | Negative | Negative |
| Late Pachynema | Paired | Restricted to XY body | Negative | Positive |
| Diplonema | Partially dissociated | Restricted to XY body | Negative | Positive |

[a]Spermatogonia are a separate mitotic stage, not a meiotic substage

differentiation states during spermatogenesis. Although cells in different substages may contribute different amounts of RNA to each sample, this analysis demonstrates an overall quantitative agreement between the two data types. This supported the validity of combining these two datasets to identify meiotic substage-specific transcriptional programs.

### Meiotic substage-specific gene expression derived by covariance analysis

To identify signatures for meiotic substage-specific gene expression, we developed a novel permutation-based maximum covariance analysis (PMCA), which maps groups of co-expressed genes to combinatorial cytological marker-based staging based antibodies to STRA8, SYCP3, phosphorylated H2AFX, and HISTH1T. This statistical approach demonstrates *concordance* of distinct cellular programs to each meiotic substage based on the similarity of "cyto-pattern" and "gene-expression pattern" across samples. In brief, the concordant patterns are derived from the most preeminent features of covariance between cytological and transcriptome data across all samples (Methods). In this RNA-seq dataset, a total of 15,025 Ensembl genes (http://www.ensembl.org) and 5267 NONCODE lncRNA genes [25] were detected (at a Transcripts per Million (TPM) ≥ 3) in at least one sample of the isolated germ cells. We also detected piRNA precursors transcripts that had previously been identified [17] with appropriate substages, although they are not the focus in the current study. Through PMCA, we identified 1235 spermatogonia genes and 6052 meiosis substage-specific genes. Expression of many transcripts could not be assigned to distinct substages and instead, were shared across several consecutive substages; 131 genes were shared among late leptotene, zygotene, and early pachytene stages while 106 genes were shared among early pachytene, late pachytene, and diplotene stages (Fig. 2; Additional file 2: Figure S1; Table 2; Additional file 1: Table S2). Notably, increased numbers of expressed genes were observed at 16 dpp when cells first reach the late pachytene stage (Table 2; Additional file 1: Table S2).

In addition to identifying genes concordant with specific substages, PMCA also identified genes that were *negatively concordant* with specific substages. The negatively concordant genes have patterns of expression that are opposite to the cytological patterns; thus negatively concordant defines gene expression that is lower when the cytological proportion of a specific substage is higher and vice versa (Fig. 2; Additional file 2: Figure S2; Table 2;

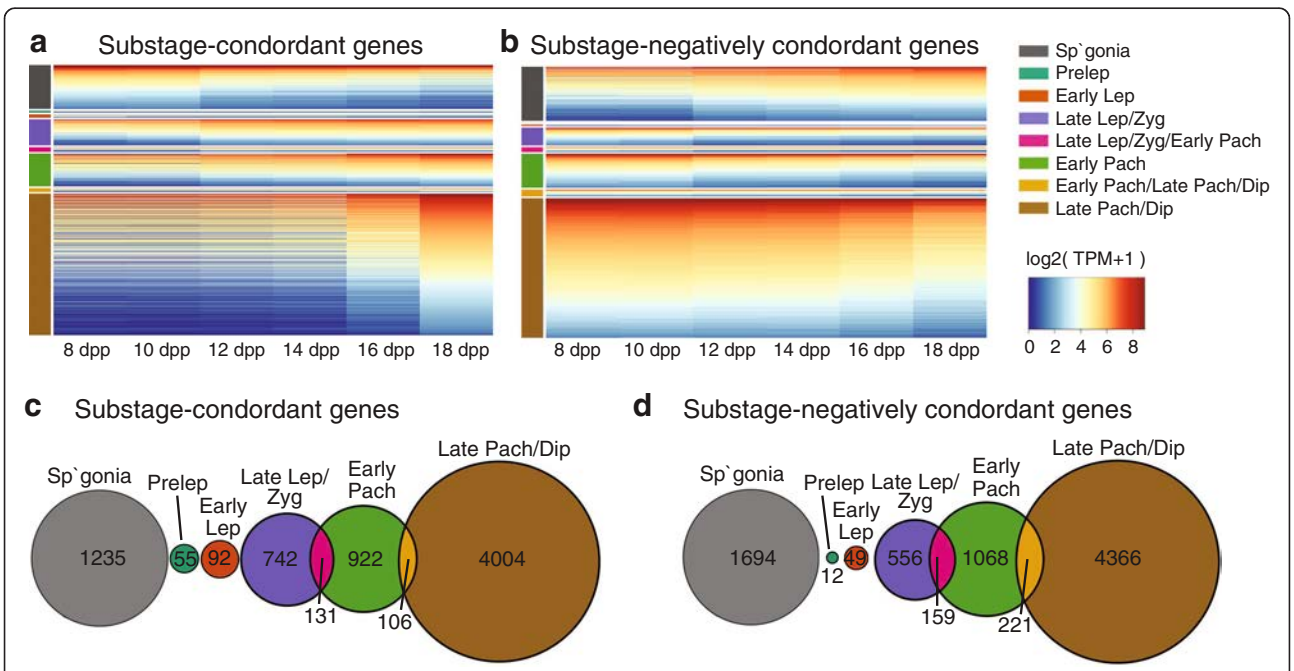



**Fig. 2** Expression of substage- concordant and negatively concordant genes. **a** Heat maps show gene expression of substage-concordant genes at each time point. **b** Heat maps show gene expression of substage-negatively concordant genes at each time point. Color bars on the left side of the heat maps represent meiotic substages (Abbreviations: Sp'gonia = spermatogonia, Pre Lep = pre-leptonema, Early Lep = early leptonema, Lep = leptonema, Late Lep = late leptonema, Zyg = zygonema, Early Pach = early pachynema, Late Pach = late pachynema, Dip = diplonema). Gene expression is shown as average log2(TPM + 1) over time point replicates. **c** Overlap among substage-concordant gene lists, illustrating the number of transcripts in common between each pair of lists. **d** Overlap among substage-negatively concordant gene lists, illustrating the number of transcripts in common between each pair of lists

**Table 2** Substage-concordant and negatively concordant genes

| Substage | Number of genes concordant | Number of genes negatively concordant |
|---|---|---|
| Spermatogonia[a] | 1235 | 1694 |
| Preleptonema | 55 | 12 |
| Early Leptonema | 92 | 49 |
| Late Leptonema/Zygonema | 742 | 556 |
| Late Leptonema/Zygonema/ Early Pachynema | 131 | 159 |
| Early Pachynema | 922 | 1068 |
| Early Pachynema/Late Pachynema/Diplonema | 106 | 221 |
| Late Pachynema/Diplonema | 4004 | 4366 |

[a]Spermatogonia are a separate mitotic stage, not a meiotic substage

Additional file 1: Table S2). Genes detected as negatively concordant with one stage are often detected as concordant with another substage. For example, genes negatively concordant with spermatogonia are genes expressed across the meiotic cell substages, but not enriched in spermatogonia. Not surprisingly, the 1694 genes negatively concordant with spermatogonia were enriched for known meiotic genes. Indeed, among all genes expressed, many of those with known meiotic function and/or Gene Ontology annotation for meiotic function were not concordant with any single meiotic substage, but were expressed across meiotic substages.

### Validation of substage-specific gene expression patterns

We bolstered and validated these computationally derived results in three distinctly different ways. First, gene expression analyses by qRT-PCR supported the RNA-seq expression data. Second, the validity of the PMCA procedure was queried by analysis of genes expressed in a highly enriched cell population of mid- to late-pachytene spermatocytes retrieved from adult testes by unit gravity sedimentation. Finally, we used our data to confirm a known pattern of male germ-cell specific gene expression, X-chromosome silencing during meiotic prophase.

We compared transcript expression patterns detected by RNA-seq to independent qRT-PCR assays. We tested representative genes among those increasing, decreasing, or showing little change in expression across the sequential time points, as well as genes known to be highly expressed in pachynema. Germ cells of 10 and 16 dpp males were subjected to qRT-PCR to determine expression of each tested gene relative to a reference housekeeping gene (*Actb*); data are shown in Additional file 2: Figure S3. The qRT-PCR assays reflected RNA-seq findings. For example, in 16 dpp samples we found elevated expression of pachytene-enriched genes and of genes shown by RNA-seq to increase in expression throughout the juvenile period tested (Additional file 2: Figure S3). Overall, patterns of gene expression are mostly concordant between the two different quantification methods and between sample sets, although, as expected, RNA-seq provides finer resolution and higher information content.

We further validated the PMCA-derived meiotic substage transcriptomes by comparison to highly enriched adult pachytene spermatocytes obtained by sedimentation at unit gravity [2–4, 26], which sorts cells based on size, not cytology or DNA content ($N = 4$ samples, each from germ cells pooled from testes of 6 mice at 9 weeks of age). The purity of late pachytene/diplotene spermatocytes in this preparation was 90 % based on cytological immunostaining with antibodies to stage-specific proteins phosphorylated H2AFX, SYCP3, and HIST1H1T. RNA isolated from the size-enriched pachytene spermatocytes was subjected to RNA-seq, allowing us to compare genes expressed in the pachytene spermatocytes to the meiotic substage-specific gene lists derived by PMCA. Cross-referencing from this dataset, there was a highly significant enrichment of pachynema-expressed genes among the gene lists from later meiotic substages (hypergeometric tests, all $p < 2.86 \times 10^{-18}$), but there was no enrichment for pachytene genes in the gene lists for spermatogonia or the early substages of meiosis (Table 3). In fact, 99 % of the genes in the late pachytene/diplotene list were also found in the enriched adult pachytene spermatocyte samples.

Finally, we queried the meiotic substage-specific patterns of gene expression determined by PMCA by confirming meiotic sex-chromosome inactivation (MSCI), a well-known feature of meiotic gene expression. In spermatocytes, most of the axial length of X and Y chromosomes, which are non-homologous, is not paired or synapsed. By the onset of the pachytene stage of meiotic prophase, the unpaired regions of the sex chromosomes become transcriptionally inactivated [11, 27, 28]. We

**Table 3** Representation of enriched pachytene spermatocyte genes in meiotic substages

| Substage | Number represented/ total concordant | *p*-value |
|---|---|---|
| Spermatogonia[a] | 614/1235 | 1 |
| Preleptonema | 17/55 | 1 |
| Early Leptonema | 28/92 | 1 |
| Late Leptonema/Zygonema | 550/742 | $1.2 \times 10^{-09}$ |
| Late Leptonema/Zygonema/ Early Pachynema | 97/131 | $9.0 \times 10^{-03}$ |
| Early Pachynema | 809/922 | $2.0 \times 10^{-62}$ |
| Early Pachynema/Late Pachynema/Diplonema | 102/106 | $1.3 \times 10^{-15}$ |
| Late Pachynema/Diplonema | 3955/4004 | $<1.0 \times 10^{-62}$ |

[a]Spermatogonia are a separate mitotic stage, not a meiotic substage

assessed expression of X-linked genes from our RNA-seq data; overall, 788 X-linked genes (and 24 Y-linked genes) were detected in the substage-associated gene lists. Of these, only 190 genes were detected in the highly enriched adult pachytene spermatocytes, providing evidence of robust MSCI. These 190 transcripts, which do not show any specific regional localization on the X chromosome, might reflect ongoing transcription and escape from MSCI, but, perhaps more likely, may represent stable transcripts persisting after inactivation. In the developmental transcriptome, the strongest X-linked gene signals are those represented in the negatively concordant gene lists (Table 4), that is, genes with expression pattern opposite to the cytologically determined frequencies of each substage. This strong signal is particularly apparent for genes in the late-pachytene negatively concordant list, which exhibit relatively constant expression from 8 to 14 dpp but then sharply decrease in expression at 16 and 18 dpp; this is evidence for pachytene substage MSCI. In addition to MSCI being revealed by the negatively concordant gene lists, there is also diminished representation of X-linked genes in gene lists concordant with later stages of meiotic prophase (Table 4). Overall, 730 X-linked genes appear to be down-regulated by pachynema to late pachynema, or even earlier (Fig. 3a) and only 58 X-linked genes with early expression are detected at later stages (Fig. 3b).

In conclusion, evidence from both MSCI and the transcriptome of highly enriched adult pachytene spermatocytes validates the PMCA-derived substage- transcriptome signatures derived by PMCA. Moreover, as discussed below, we find good concordance among our stage-specific gene lists and functions with those derived by other studies.

### Enrichment analysis of meiotic substage gene expression

The PMCA-derived meiosis substage-specific gene lists, coupled with Gene Ontology (GO) analysis [29], can provide insight into meiotic and spermatogenic function. Genes assigned to each substage were ranked by their similarity score, a measure of how closely the gene's expression pattern follows the substage's cytological pattern. Using ranked lists, we performed a ranked GO analysis in GOrilla [30]; these ranked lists provide approximations of meiotic substage gene ontology. Significantly, many meiosis-specific genes are not prominently represented among the meiosis substage-specific gene lists derived by PMCA. However, GO terms for genes negatively concordant with spermatogonia were predominantly meiotic terms (16 out of 31 GO terms for process, all $p < 9.97 \times 10^{-4}$), confirming that meiotic genes tend to be expressed across multiple meiotic substages and not unique to a particular meiotic substage (Additional file 1: Table S3). Thus meiotic substage transitions are probably not acutely regulated at the level of transcription of many of the known meiotic genes, although among late leptonema/zygonema-associated GO terms, 8 out of 10 GO "biological process" terms were associated with meiosis regulation (all $p < 8.6 \times 10^{-4}$), and similar GO terms were also identified in the early pachynema gene list (13 out of 33 GO terms for process, all $p < 5.56 \times 10^{-4}$) (Additional file 1: Table S4). We extracted 447 genes for *M. musculus* that are associated with any GO terms containing "meiosis" or "meiotic" (www.mousemine.org), and of these genes we considered the 404 genes which are expressed in our developmental time series (those genes not expressed in our time series might include female meiosis-related genes). We found that 229 (57 %) genes were not concordant with any specific substage, while 26 (6 %) were concordant with late leptonema and zygonema, 43 (11 %) were concordant with early pachynema, and 92 (23 %) were concordant with late pachynema and diplonema. Of meiotic terms that were negatively concordant with substages, 67 (17 %) were negatively concordant with spermatogonia and 97 (24 %) were negatively concordant with late pachynema and diplonema (Additional file 1: Table S5). GO terms for genes expressed in late meiotic prophase were significantly enriched for spermatogenesis, spermiogenesis and fertilization, reflecting transcription of mRNAs to be stored for later translation during the haploid phase

**Table 4** Substage-specific X-linked gene analysis

| Substage | X-linked Concordant | Y-linked Concordant | X-linked Negatively Concordant | Y-linked Negatively Concordant |
|---|---|---|---|---|
| Spermatogonia[a] | 28 | 0 | 55 | 2 |
| Preleptonema | 2 | 0 | 0 | 0 |
| Early Leptonema | 2 | 0 | 1 | 0 |
| Late Leptonema/Zygonema | 34 | 2 | 6 | 0 |
| Late Leptonema/Zygonema/Early Pachynema | 3 | 0 | 1 | 0 |
| Early Pachynema | 21 | 1 | 41 | 0 |
| Early Pachynema/Late Pachynema/Diplonema | 0 | 0 | 13 | 0 |
| Late Pachynema/Diplonema | 14 | 3 | 329 | 9 |

[a]Spermatogonia are a separate mitotic stage, not a meiotic substage

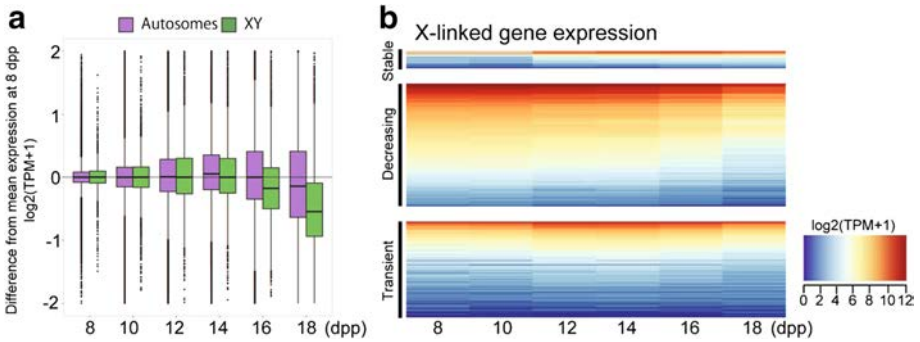



**Fig. 3** X-linked gene expression during the first wave of spermatogenesis. **a** Boxplots of the difference from mean gene expression at 8 dpp are shown for each time point, with genes on the autosomes in purple and those on the XY chromosomes in green. Gene expression is log2(TPM + 1). **b** Heatmap of X-linked gene expression at each time point. Genes are clustered based on the pattern of expression: an increase of expression at 12 dpp and then stable throughout (Stable), steadily decreasing over time (Decreasing), or low expression from 8-10 dpp and high expression from 12-14 dpp followed by low expression at 16-18 dpp (Transient). Gene expression is shown as average log2(TPM + 1) over time-point replicates

of spermatogenesis [31]. For example, early pachynema gene lists are enriched for GO terms associated with spermiogenesis (5 out of 33 GO terms for process, all $p < 7.39 \times 10^{-4}$), and the majority of late meiotic prophase- or late-pachynema/diplonema-associated genes were associated with spermiogenesis or fertilization-related GO terms (11 out of 16 GO terms for process, all $p < 7.11 \times 10^{-4}$) (Additional file 1: Table S4). Although genes concordant with early leptonema had transcription-related GO terms in 11 out of 32 GO terms for process (all $p < 1.00 \times 10^{-10}$), we found that many genes negatively concordant with leptotene through zygotene substages have GO terms for transcription and related processes (Additional file 1: Table S3), possibly reflecting the cytologically diminished incorporation of RNA precursors during the earliest meiotic prophase substages [32]. Also among the genes lists for these early stages we found extracellular membrane- or molecular transport-related GO terms prominently represented (Additional file 1: Table S4), which may related to the fact that these cells transit through the Sertoli cell junctions that create the blood-testis barrier [33].

Additionally we performed feature enrichment analysis with the hypergeometric test on the substage-specific gene lists. Interestingly, protein-coding genes are enriched in gene lists concordant with the late-leptotene/zygotene stages ($p = 4.52 \times 10^{-18}$) and the early pachytene stage ($p = 1.19 \times 10^{-8}$). Moreover, they are enriched in the set of genes that is negatively-concordant with late-pachytene and diplotene substages ($p = 5.94 \times 10^{-167}$). However, the genes concordant with late-pachytene and diplotene substages are not depleted in protein-coding gene, which comprise the expected majority of this gene set. This suggests that while certain protein-coding genes are abundant in the Late Pach/Dip substage, large numbers of other protein-coding genes are downregulated in this stage, perhaps reflecting cessation of mRNA transcription in preparation for the meiotic division stage, or, alternatively, post-transcriptional regulation by PIWI-interacting RNAs (piRNAs). Recent studies show that piRNAs play important roles in genome stability by suppressing harmful transposons as well as by regulating mRNAs [34], and future analyses could integrate piRNA expression with these data.

### Transcription factor analysis

Because our results point to a large and diverse meiotic germ-cell transcriptome, also noted by others [19], we inferred the underlying regulatory networks accounting for these patterns, using the iRegulon bioinformatic approach to identify transcription factors (TFs) potentially regulating substage-specific genes (Methods). TFs were identified for each substage-specific gene list with high normalized enrichment scores (NES ≥ 4), corresponding to an estimated false discovery rate of less than 0.01 [35]. We then determined which TF genes were unique to, or shared among, the substages. As can be seen in Additional file 1: Table S6 and Additional file 2: Figure S4, mRNA transcripts for some TFs are substage-specific. For example, *Pou5f1* transcript is specific to the preleptotene stage while *Zbtb33* transcript is specific to the pachytene/diplotene substages. Other TF transcripts, similar to meiosis-specific genes in general, are shared across substages but may be specific to early or late meiosis (Additional file 2: Figure S4); *Zfp143* transcript is shared across late meiosis substages while *Tpbl1* transcript is common to the preleptotene (in the "Jazf1 + 3" cluster) and late leptotene/zygotene/early pachytene (in the "Mybl* + 9" cluster) substages. We also found that some TFs, for example MYBl1, also have target genes that are negatively concordant with substages (Additional file 2: Figure S5; Additional file 1: Table S7). We used the MGI bioinformatics Batch Query tool to determine that 160 of the 181 TFs in this analysis were

previously found to be highly expressed specifically and/or significantly in the testis or male germ cells and 39 are annotated to male-reproduction-related phenotypes (Additional file 1: Table S6).

To gain a deeper insight into the regulatory patterns of meiosis, we selected the highest scoring TF genes for further analysis, based on iRegulon's NES score which corresponds to a low false discovery rate. Of these highest scoring TFs, half of them are annotated to meiosis-related functions: ETV4, E2F1, GATA2, RFX4, and ZFP143. We included four other well-known meiosis TFs: RARG, MYBL1, ETV5, and TBPL1. We compared the mRNA expression pattern of these TF genes with the expression patterns of their target genes in order to identify putative regulatory relationships, an analysis that is based on the assumption that a relevant TF protein appears more or less contemporaneously with its transcript (i.e., no translational delay). In cases where the expression pattern of the TF gene had the opposite expression pattern of its target genes, we infer that the TF acts as a repressor on these targets; but in cases where the expression patterns of the TF gene and its target genes were concordant, we postulate that the TF enhances expression of its targets. For TFs with genes that did not change over our time course, such as the known candidate NRF1, we could not infer a relationship to target genes and excluded them from further analysis.

Following these assumptions, as illustrated in Figs. 4 and 5, we suggest that OVOL2 and YY1 act as repressors of target genes in early leptonema and ETV5, ETV6, and ZFP143 act as repressors on target genes after initiation of meiosis, beginning with the late leptotene and zygotene substages. Evidence suggests that ZBTB33, GATA2, and ETV4 also act as repressors on gene targets in pachynema (Fig. 4 and Additional file 2: Figure S6); while RFX4 appears to be an enhancer of target genes in late pachytene and diplotene substages (Additional file 2: Figure S6), but may repress its target genes in preleptonema, although this relationship is less clear. Interestingly, these underlying assumptions based on relative expression levels suggest that MYBL1, TBPL1, and E2F1 act as both enhancers and repressors of their target genes. MYBL1 and TBPL1 appear to repress target genes in early meiosis while enhancing target genes in late meiosis, while E2F1 has the opposite pattern and may repress target genes in late meiosis while enhancing target genes in early leptonema (Additional file 2: Figure S6). Similar regulatory switching has been shown to result from changes in protein co-factors [36] and post-translational modifications [37]. Moreover, genes for some of the TFs described above are also themselves target genes associated with specific substages. By considering these relationships, we inferred candidate regulatory interactions among TFs; for example, ZFP143 and ETV5 suppress *Mybl1* gene expression, while MYBL1 enhances expression of *Rfx4*, *Ovol2*, *Yy1*, and *Tbpl1* genes (Fig. 5).

Overall, 41 % of the genes associated with specific meiotic substages (2483 of 6052 genes) are predicted targets of one or more of these eight meiosis-related TFs. This noteworthy observation suggests that this concise regulatory network can account for a substantial portion of the meiotic program of transcription.

## Discussion

Here we identify a male meiotic germ-cell transcriptome using a novel analysis based on a dense dataset of cytological substage-specific signatures. The high concordance between RNA-seq expression data and cytological proportions of isolated germ cells across all samples we analyzed allowed us to develop a novel PMCA to identify the substage-specific transcriptomes for meiotic prophase. This computational method does not require the use of FACS or sedimentation sorting of cells and can be applied to other complex cell populations for which there are well-defined cytological criteria.

### Cytological deconvolution of RNA-seq data from the first wave of spermatogenesis

In mammalian males, spermatogenesis is continuous and asynchronous, ensuring daily capacity to deliver sperm backed up by a testis comprised of abundant numbers of germ cells at all stages of spermatogenesis. While advantageous to reproductively active males, this biological imperative has frustrated attempts to isolate stage-purified spermatogenic cells. Methods for stratification of cell samples by either sedimentation [3, 4] or flow cytometry [7, 8] enrich specific spermatogenesis cell stages, while the juvenile first wave of spermatogenesis (used in this study) provides a leading edge of differentiating cells against a background of less differentiated cells. In the laboratory mouse, meiosis is initiated at about 8 dpp in a subset of spermatogonia by retinoic acid stimulation of STRA8 [1, 20]; however, initiation of subsequent waves of meiosis make the cellular population increasingly heterogeneous (Fig. 1c), resulting in a complex histology [4]. The cell-stage frequencies presented here (Fig. 1c) are based on a powerful combinatorial application of antibodies recognizing well-characterized and highly stage-specific marker proteins. We find a continuous and relatively constant background of the earliest spermatogenic cells, namely, spermatogonia and the preleptonema and leptonema (Fig. 1c), and, not surprisingly, these "background" spermatogenic cells are the ones most refractory for computational assignment of a substage-specific transcriptome. On the other hand, pachytene spermatocytes are not present at the earliest juvenile stages, and appear in a discrete time period; consequently, it was possible to assign a robust gene list to this stage.

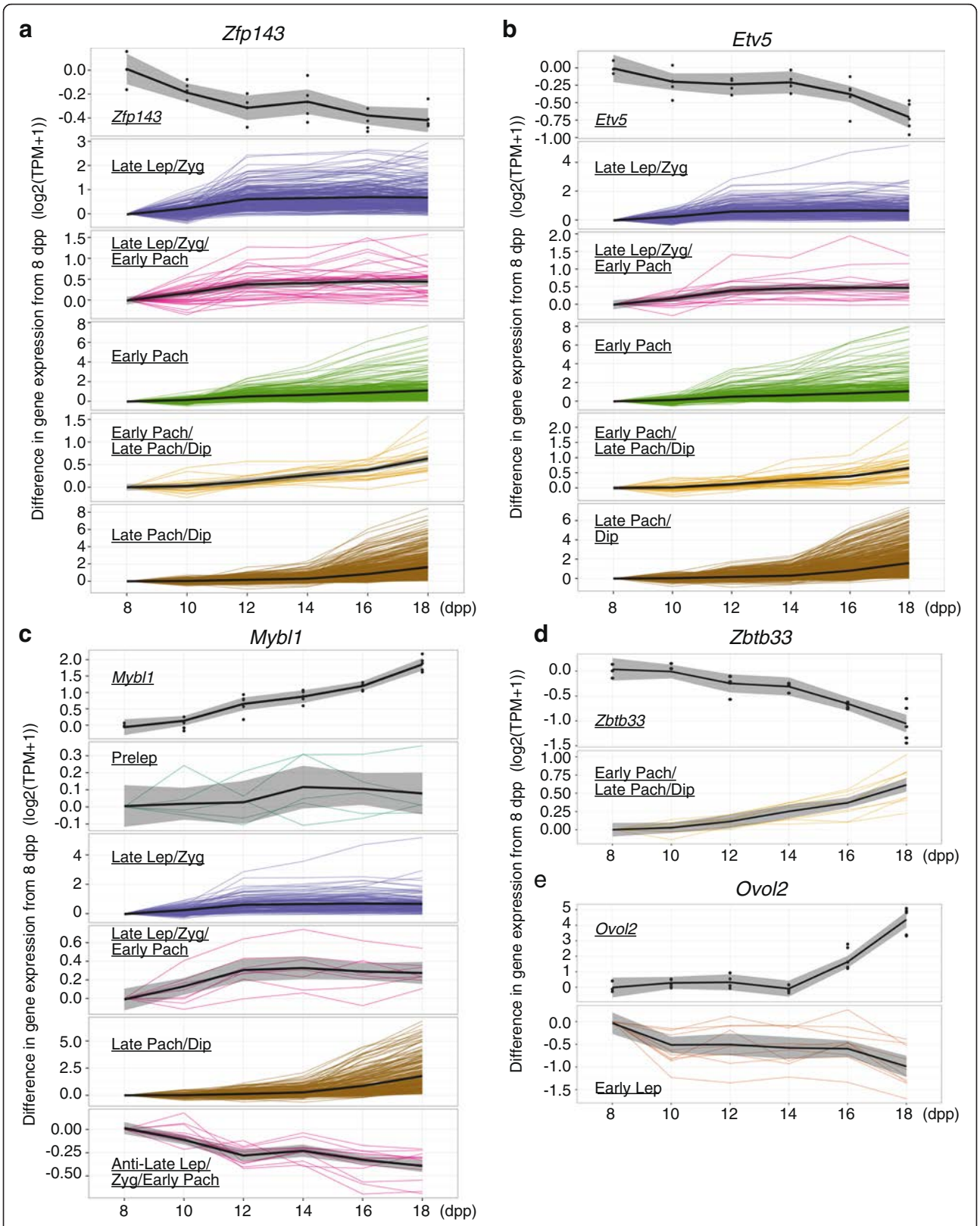



**Fig. 4** Expression of TFs and target genes for each substage. The difference from mean gene expression at 8 dpp of the TF and its target genes at each time point. To illustrate the overall pattern, a smoothed line was fit to the substage-specific gene expression. Color of lines represents the substage of target genes. Gene expression is average log2(TPM + 1), averaged over time point replicates. Expression is plotted for (**a**) *Zpf143*, (**b**) *Etv5*, (**c**) *Mybl1*, and (**d**) *Zbtb33*

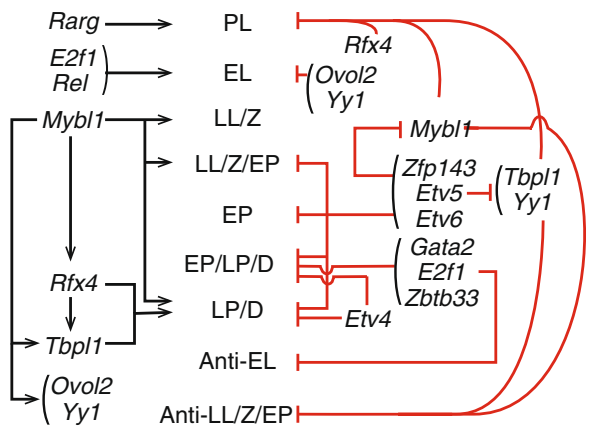



**Fig. 5** Regulation of substage-concordant gene expression during meiotic prophase. Analyses suggest that substage-concordant gene expression is regulated both positively and negatively by a network of TFs. TFs are shown by their gene names, and the meiotic substages of predicted TF target genes are shown in the center. Black arrows represent a TF acting as an enhancer on its target genes and red bars represent a TF acting as a suppressor on its target genes. (Abbreviations: PL = pre-leptonema, EL = early leptonema, LL/Z = late leptonema and zygonema, LL/Z/EP = late leptonema, zygonema, and early pachynema, EP = early pachynema, EP/LP/D = early and late pachynema and diplonema), LP/D = late pachynema and diplonema, Anti-EL = genes negatively concordant with early leptonema, Anti-LL/Z/EP = genes negatively concordant with late leptonema, zygonema, and early pachynema)

Gene assignments to specific substages were guided by the covariance of transcriptomic data and the cytological data. Assignment of expressed transcripts to cellular subpopulations in heterogeneous samples has long been a computational challenge [18, 38–40]. In this study, we exploited the advantage of paired RNA-seq and cytological data to develop the novel PMCA approach that isolates changes of gene expression specific to each meiotic substage. Maximum covariance analysis (MCA) was first developed in the meteorological sciences [41] and was popularized in the climatological sciences in the 1990s [42, 43]. More recently, an MCA approach has been used in a bioinformatics context to clarify relationships between gene and protein expression [44]. While MCA is often an effective tool for detecting common signals in two sets of variables, it can be limited by a tendency to fit spurious patterns when faced with increased sampling variation [45, 46]. Current methods [47] employ a parametric smoothing model using principal component regression, which requires a normality assumption or rely on GO analysis [44]. We developed a novel PMCA method that not only overcomes the spurious pattern identification liability but that does so without the need for any parametric assumptions or reliance on GO analysis. Our PMCA approach is broadly applicable to multi-dimensional data derived from a common set of samples.

### Substage specificity of meiotic gene expression and regulation

Historically, meiotic prophase substages are characterized by the morphology of chromatin, and correlated genetic mechanisms have been revealed in the past decades [48]. It is known that there is widespread transcription of protein-coding genes in the testis [19], and indeed, we found that among germ cells alone, more than 15,000 genes are transcribed, suggesting that a significant portion of testis transcriptional complexity is due to germ cells. Using PMCA, we successfully identified genes with expression patterns that matched the "cyto-patterns" of stage-specific cell frequencies determined by antibody labeling (Fig. 1c). Some genes were shared among substages; for example, late leptotene and zygotene shared 131 genes with early pachytene and early pachytene shared 106 genes with late pachytene and diplotene. Due to the strong temporal signal of contribution of the late pachytene germ cells, which do not appear until 16 dpp and greatly increase in representation by 18 dpp, we identified over 4000 late pachytene genes. However, since the proportion of preleptotene and early leptotene cells did not vary greatly over time, we were unable to deconvolve stage-specific transcriptomes for the earliest meiotic substages as confidently as for the later substages.

We compared these substage-specific gene lists to those developed in other recent analyses of the germ-cell testis transcriptome. The striking increase in number of genes unique to the pachytene stage was also observed by Soumillon, et al. [19], where, in fact, greater numbers of genes were considered expressed than in this study. However, we have used a more stringent cutoff (TPM ≥ 3) to eliminate from the data transcripts expressed at low levels. When we relaxed this requirement and compared the number of genes expressed with a cutoff of TPM > 0, we found the number of expressed transcripts to be similar to that reported by Soumillion, et al. [19]. We also compared the substage-specific gene lists to the relevant clusters identified by Margolin, et al. [18]. In their study, gene lists were temporally clustered from single samples taken at 6, 10, 12, 14, 16, 18, 20, and 38 dpp. While our study was based on multiple replicates at each time point as well as fine-grained cytological analysis, our PMCA gene lists substantially overlapped with their derived gene list clusters (Additional file 1: Table S8). In all, 4013 of the 4077 genes in clusters identified by Margolin, et al. [18] were also expressed in our dataset.

We also compared our results with those derived from flow-cytometry based methods of sorting by DNA content [5, 6]. For both of these studies, we found overall concordance in genes assigned to meiotic substages (Additional file 1: Table S9). These similarities were especially pronounced for late pachytene and diplotene substages due to the large numbers of genes in these

sets. However, we also observed a broader trend of alignment between early and late substages. Furthermore, our use of multiple markers to cytologically characterize cells paired with our PMCA analysis allowed for more precise substage assignment than possible by cell sorting. For instance, our spermatogonia and pre-leptotene sets were the top two strongest overlaps with the "secondary spermatocyte" (2C) cell fraction in da Cruz, et al. [5] (hypergeometric test $p = 4.5 \times 10^{-4}$ and $p = 1 \times 10^{-3}$, respectively), but split these assignments to provide improved substage resolution. Similarly, our joint late-leptotene and zygotene genes and early pachytene genes had the two greatest overlaps with the combined "leptotene and zygotene" and "pachytene spermatocytes" (LZ + PS) genes (hypergeometric test $p = 8.7 \times 10^{-7}$ and $p = 1.4 \times 10^{-3}$, respectively), and further partition these transcripts into more precise substages. Analysis of Fallahi et al. [6] data was limited by different experimental platforms (microarrays versus RNA-seq), different assay timing (adult versus juvenile mice), and generally low numbers of uniquely assigned genes, but also revealed significant overlaps in late substages (Additional file 1: Table S9). By assessing functional concordance by similarity in GO annotations, we determined that our early leptotene and Fallahi et al. [6] pre-leptotene sets share a common enrichment in RNA transcription genes. Finally, the most prominent divergences between our data and these previous results were in the spermatogonia fraction. We note that cell sorting techniques are susceptible to inclusion of somatic-cells in this fraction, whereas our spermatogonia genes were significantly negatively-concordant with many standard meiosis genes (Additional file 1: Table S3). In sum, these comparisons suggest that our transcript sets encompass and substantially expand these previous findings.

In this study, we characterized each meiotic substage using the list of substage-specific genes from PMCA analysis. Some meiosis genes were associated to particular substages. For example, *Spata22* [49] and *Msh4/5* [50, 51] are both highly associated with late leptotene/zygotene substages, and both are required for recombination. However, many canonical meiosis genes were found throughout all meiotic prophase substages; these genes include *Rad51*, *Rec8*, and *Syce2*. This may well reflect a lack of acute transcriptional regulation for these important transcripts. Rather than transcription at the precise meiotic substage of use, quality and efficiency of meiosis may be ensured by having transcripts present and available for translation throughout meiotic prophase, with substage transitions regulated post-transcriptionally and/or post-translationally. Further, although we confirmed that spermatogonia negatively concordant genes are enriched with meiotic genes, we found several meiotic genes in the gene list concordant with spermatogonia, including *Fign*, which was reported to be required for meiotic recombination [52], suggesting that some of the meiotic program is set up prior to meiotic entry.

Not surprisingly, this developmental transcriptome analysis also revealed that many genes expressed during meiosis do not have known functions directly contributing to meiosis. These may instead be part of a parallel program of spermatogenic gene expression. For example, many of the genes expressed in early leptonema are associated with transcription or RNA metabolism, as well as with cellular processes such as cell-cell interactions, which are of considerable importance in the biology of the seminiferous tubule. PMCA also revealed that late pachytene/diplotene-expressed genes are significantly associated with GO terms associated with spermiogenic processes, correlating well with findings of other studies [5, 6]. During post-meiotic stages of nuclear condensation, transcription is globally repressed [28, 53], underlying the biological rationale for prior transcription to support the active protein synthesis during spermiogenesis. In addition, the late pachytene/diplotene-associated gene lists are significantly enriched with non-protein-coding RNAs, including lincRNAs (long intergenic non-coding RNA) and piRNA precursors.

### Meiotic regulatory networks

Analysis of predicted TF targets [35] generated a regulatory network that potentially governs the meiotic and spermatogenic programs of gene expression identified by this study (Fig. 5). The protein ZFP143 was central in our network, and is required for embryo development in zebrafish [54] and human [55, 56]. The human ortholog, ZNF143, is ubiquitously expressed [57], and binds to the promoter region of target genes where it is required for formation of chromatin loops by interacting with POLII and RAD21 [58]. Because RAD21 is differentially expressed after the pachytene stage of meiosis [59], it is possible that ZFP143 is crucial for spermatogenic processes by regulating the transcription of genes during pachytene/diplotene substage, an idea that should be tested experimentally since our analysis indicates that ZFP143 may target genes strongly associated with genes in these substages. Reinforcing the validity of our approach, our regulatory network also involved MYBL1 (also known as A-MYB). Our results support previous findings that a subset of late pachynema/diplonema-expressed genes, many involved with the spermatogenic and spermiogenic processes, are associated with MYBL1 [60, 61]. MYBL1 regulates transcription of cell cycle-related genes. Germ cells with a homozygous mutation in the *Mybl1* gene exhibit defects in meiotic chromosome synapsis [60]. Moreover, our results suggest that MYBL1 also is in a network associated with the transcription of genes encoding proteins required for piRNA biogenesis, including *Tdrd1, Tdrd6* and *Tdrd7*, consistent with previous reports [62]. Going beyond

our core regulatory network, we also identified several other TFs, whose target genes include the Tdrd family and Piwi genes, potentially involved in piRNA processing: ATF3, ELF1, ELK3, ELK4, FLI1, NFYA, NFYB, NRF1, PBX3, RFX2, RFX7, SP1, YY1, ZBTB33, ZFP143 and ZFP42. This analysis not only revealed a core network for transcriptional regulation of meiotic progression (Fig. 5 and Additional file 2: Figures S4 and S5) but also suggested that a significant proportion of the genes expressed in the meiotic transcriptome may be controlled by a concise entourage of transcription factors.

## Conclusions

This study has untangled in part the complexity and parallel process of spermatogenesis and meiosis by focusing on associating gene expression with highly specific cytological signatures defining meiotic prophase substages. This unique and powerful approach to deconvolving transcriptomes of complex cell populations is applicable for discovery of transcriptional signals in other such complex cell populations or heterogeneous tissues.

## Methods

### Experimental design and mice

All C57BL/6J mice used for this study were obtained from The Jackson Laboratory (Bar Harbor, USA). All animal procedures were in accordance with the National Institute of Health guide and U.S. Department of Agriculture standards for animal care and use and were approved by the Animal Care and Use Committee at The Jackson Laboratory (Protocol #05004). Mice were euthanized at 8, 10, 12, 14, 16 and 18 days post partum (dpp) to follow the leading edge of meiotic progression during the first wave of spermatogenesis. For each time point, 5 biological replicates were sampled and germ cells were isolated from the pooled two testes of each mouse. A portion of germ cell sample was used for cytological analysis and the rest of cells were used for RNA-seq analysis. Both cytological and RNA-seq analyses were performed on all 30 samples (Fig. 1a).

### Cytological methods

#### *Isolation of mixed testicular germ cells*

The procedure was as previously described with some modifications [26]. Briefly, seminiferous tubules were transferred into 20 ml DMEM (Gibco, Life Technologies, Carlsbad, CA, USA) containing 0.5 mg of Liberase TL Research Grade (05401020001, Roche, Basel, Switzerland) and incubated for 20 min at 32 °C. To remove interstitial cells, tubules were washed three times with the same media. In the final wash media, the tubules were pipetted several times to form fragments, which were digested with 0.5 mg of Liberase and 10 μg of DNase in 20 ml DMEM for 13 min at 32 °C in shaking water bath. The isolated cells were further digested by pipetting for 3 min, and germ cells were isolated by filtering through Nitex mesh (53-70 μm pore size). The crude germ cells were washed three times by centrifugation for 7 min at 500 g using 10 ml of the media containing 10 μg of DNase. The cells were resuspended in 1 ml of ice-cold PBS, and cell concentration determined. 1.25 x $10^5$ cells (about 10 % of total) were used for cytological scoring, with the remainder used for RNA-seq (see below).

#### *Isolation of enriched populations of adult pachytene germ cells*

Each enriched population (4 biological replicates) of pachytene/diplotene spermatocytes was obtained from the testes of six 9-week-old mice by sedimentation at unit gravity [3]. Mixed germ-cell suspensions were prepared as described above, and after the three 0.5 % BSA/KRB washes, cells were separated by cellular sedimentation at unit gravity in a 2–4 % BSA gradient generated over 2.5 h in a STA-PUT apparatus (ProScience Inc., GlassShop, Toronto, ON, Canada). Following the sedimentation process, 10 ml fractions were collected and examined using light microscopy and differential interference contrast Nomarski optics. Cells were identified based on morphological criteria and size [3]. Fractions containing pachytene/diplotene spermatocytes of average purity ~90 % were pooled. For every cell separation, and for each population of cells collected, an aliquot of cells was snap-frozen for subsequent RNA extraction as described below.

#### *Chromatin spread preparation*

Germ-cell suspensions prepared as described above were fixed in 2 % PFA containing 0.03 % SDS and mixed with an equal volume of 0.04 % Photo-Flo (Kodak, Rochester, NY, USA). The cell suspension was applied to wells of 12-well Shandon slides, and incubated in a humid chamber for 1 h at RT. After fixation, the cells were briefly air-dried, and subjected to further fixation: 2 % PFA with SDS for 3 min and 2 % PFA without SDS for 3 min. The slides were then washed three times with 0.04 % Photo-Flo. Air-dried slides were stored at -20 °C for further use.

#### *Immunostaining of spread chromatin*

Spread chromatin preparations were incubated with 10 % ADB blocking solution (ADB: PBS containing 2 % BSA and 0.05 % Triton-X 100) for 10 min, the same blocking solution with SDS for 10 min, and the blocking solution without SDS. Immunolabeling was performed with rat polyclonal anti-SYCP3 (1:1000 dilution, Handel lab), mouse monoclonal anti-phosphorylated histone H2AX (1:200 dilution; 05-636, Millipore, Billerica, MA, USA), rabbit polyclonal anti-STRA8 (1:1000 dilution; ab49405, Abcam, Cambridge, England) and guinea pig polyclonal

anti-H1t antibodies (1:500 dilution; Handel lab). Subsequently, secondary antibodies conjugated with Alexa 647, 594 or 488 (Molecular Probes, Invitrogen, Carlsbad, CA, USA) were used at 1:500 dilution. Nuclei were stained with DAPI (0.5 μg/mL) for 10 min, and the slides were mounted with glycerol. Images were observed using a Zeiss AxioImager.Z2 epifluorescence microscope equipped with a Zeiss AxioCam MRm CCD camera (Carl Zeiss, Jena, Germany).

## RNA methods

### *Isolation of RNA and sequencing library preparation*

Isolated germ-cell samples were resuspended in QIAzol Lysis Reagent according to the manufacturer's instructions, and total RNAs were purified from homogenized cells using Qiagen RNeasy Mini Kit (74104). The quality of the isolated RNA was assessed using an Agilent 2100 Bioanalyzer instrument (Agilent Technologies, Santa Clara, CA, USA) and RNA 6000 Nano LabChip assay (5067-1511, Agilent Technologies).

The mRNA sequencing libraries were prepared using the Illumina TruSeq methodology. mRNAs were purified from total RNA using biotin tagged poly dT oligonucleotides and streptavidin coated magnetic beads. The mRNAs were then fragmented and double stranded cDNA was generated by random priming. The library was then analyzed for quality using an Agilent 2100 Bioanalyzer instrument (Agilent Technologies) and DNA 1000 LabChip assay.

### *RNA sequencing*

Short 100 bp paired-end reads were generated and sequenced using an Illumina® HiSeq (Illumina, San Diego, CA, USA). Sequenced reads were filtered to keep reads for which 70 % of the base pair quality score was > 20, and the 3′ end was trimmed if the base pair quality score was < 20. Two technical replicates for each paired-end were run in different lanes and then merged.

### *RNA extraction and quantitative real-time quantitative RT-PCR*

For real-time RT-PCR, total RNA was isolated from isolated germ cells or enriched germ cells (see above) using the RNeasy Mini Kit (Qiagen, Hilden, Germany), and 1 μg RNA was reverse transcribed using QuantiTect Reverse Transcription Kit (Qiagen) according to the manufacture's instruction. The real-time RT-PCR was performed by the Applied Biosystems 7500 Real-Time PCR system (Foster City, CA. USA) using the QuntiTect SYBR Green RT-PCR kit (Qiagen). Transcript levels were normalized to the levels of *Actb* by the standard curve method [63], and are presented as the mean normalized expression in 10 μg RNA. Data are represented as mean ± estimated standard deviation. Gene-specific primers are listed in Additional file 1: Table S10.

## Computational methods

### *Alignment and expression*

All RNA-seq samples were aligned with Bowtie 1.0.0 [64] and expression levels were estimated by RSEM 1.2.8 [65]. A Bowtie index was prepared for alignment to a combined (mm10) transcriptome of Ensembl Genome Reference Consortium, build 38, release 75 [66], NONCODE v4 lncRNA [25], and piRNA precursor transcripts [17]. The 214 piRNA precursors were obtained from Dr. Christian Roy. Both the NONCODE lncRNA and piRNA precursors were converted to mm10 coordinates using liftOver [67]. For this study, we used log2(TPM + 1) as the expression level, where TPM is transcripts per million, defined by RSEM [65]. A gene was deemed expressed if TPM ≥ 3 for at least one of the 28 samples.

### *Principal component analysis*

To test the concordance of the expression and cytological data, independent Principal Component Analyses (PCA) were performed on each dataset using prcomp(x, scale = TRUE) in the rgl package [68] in R [69]. To produce Fig. 1d, the first and second principal components for both datasets were scaled to have the same range.

### *Permutation-based Maximum Covariance Analysis (PMCA)*

We developed a novel permutation-based maximum covariance analysis (PMCA) that not only overcomes the spurious pattern identification liability of traditional maximum covariance analysis but also does not require any parametric assumptions about the data. Instead, we implemented a permutation procedure that assures the patterns are valid within a given false positive rate (FPR). To begin, let

$$X_{6\times 28}\ \&\ Y_{20368\times 28} \tag{0.1}$$

be the cytological and gene expression data, respectively. Then mean center each row, where $\tilde{X}, \tilde{Y}$ are mean centered (Eq. 1.2.)

$$\tilde{x}_{is} = x_{is} - \frac{1}{28}\sum_{s=1}^{28} x_{is},\ \tilde{y}_{ks} = y_{ks} - \frac{1}{28}\sum_{s=1}^{28} y_{ks} \tag{0.2}$$

$$i = 1, 2, \ldots, 6,\ s = 1, 2, \ldots, 28,\ k = 1, 2, \ldots, 20368$$

We compute the covariance matrix and the SVD of the covariance matrix in Eq. 0.3.

$$C_{6\times 20368} = \frac{1}{28}\tilde{X}\tilde{Y}^T = U\Sigma V^T \tag{0.3}$$

Because we are interested in mapping the genes of $Y$ onto substages of $X$ we consider $P_{x(6\times 28)} = U^T\tilde{X}$, the

principal components of the covariance matrix that correspond to the substages, and calculate the homogeneous and heterogeneous regressions:

$$Z_{x(6\times6)} = XP^T{}_x,\ Z_{y(20368\times6)} = YP_x^T \tag{0.4}$$

To allow for direct comparison, each row of $Z_x, Z_y$ is divided by its respective root mean square, Eq. 0.5.

$$Z_{x(ij)} = z_{x(ij)} \Big/ \sqrt{\sum_{j=1}^{6} z_{x(ij)}^2 / 5} \tag{0.5}$$

$$Z_{y(kj)} = z_{y(kj)} \Big/ \sqrt{\sum_{j=1}^{6} z_{y(kj)}^2 / 5}$$

For each substage cyto-pattern (the row of $Z_x$ corresponding to the substage), we find the genes that have a similar gene expression pattern across the columns of $Z_y$. Through computation of the SVD, we lose a degree of freedom so the maximum number of patterns is 5; there are 6 substages. For each substage cyto-pattern, we call a gene pattern similar if it is within a certain window of the substage's cyto-pattern.

To determine the optimal window width, we devised a permutation method that iterates through varying window widths and chooses the optimal width based on the estimated false positive rate (FPR). We stipulated that the estimated FPR be less that 0.05 by the third component ($j = 3$). Because the gene lists get finer, and more specific, as we progress through the components, the estimated FPR for cyto-patterns 4 and 5 are less that the specified 0.05. For the substage-specific gene lists described in this paper, the estimated FPR < 0.005. A detailed explanation of the PMCA optimal width selection is provided in Additional file 3.

### Bioinformatic analysis

#### *Gene Ontology (GO) analysis*

GO enrichment analysis was performed using GOrilla, Gene Ontology enRIchment anaLysis and visuaLizAtion tool [30] with ranked gene lists. Gene lists were ranked for each substage based on a score that measures how closely each gene pattern follows the substage cytopattern. GO terms was established by the GO Consortium [70] and used to group genes according to their biological or molecular functions. A total of 13,363 of the 15,025 genes expressed in our time series have at least one annotation in GO.

#### *Transcription factor analyses*

TFs for substage-concordant and negatively concordant genes and the TF target genes were identified using iRegulon, Version 1.3 [35] in Cytoscape, version 3.1.0 [71] with the substage-concordant and negatively concordant gene lists.

## Additional files

**Additional file 1:** This file contains a mini-website of Tables S1-S10. The same mini-website is available at http://carterdev.jax.org/dtx/a2/index.html. Interactive expression plots. **Table S1.** Number and proportion of isolated cells for cytological analysis and RNA-seq. **Table S2.** Gene lists and gene expression values for each of the substages, separated by worksheets. An additional worksheet provides results for X-linked genes. Abbreviations for each substage are as follows: Sp'gonia for spermatogonia, PL for preleptotene, EL for early leptotene, LL + Z for late leptotene and zygotene, LL + Z + EP for those genes found in late leptotene, zygotene, and early pachytene, EP for early pachytene, EP + LP + D for early pachytene, late pachytene, and diplotene, and LP + D for late pachytene and diplotene. Sheet names prefixed with "Anti-"correspond to genes that are negatively concordant with the substage, (e.g., Anti-PL corresponds to genes that are negatively concordant with preleptotene. Genes are ranked by how closely the expression pattern follows the substage pattern (rank = 1 is best). Gene ids and Gene names for piRNA precursors and NONCODE lncRNAs have prefix "pi-" and "NON", respectively. Expression is given in log2(TPM + 1). **Table S3.** GO analysis for substage-negatively concordant genes. Abbreviations for each substage are as follows: Sp'gonia for spermatogonia, EL for early leptotene, LL-Z for late leptotene and zygotene, EP for early pachytene, and LP-D for late pachytene and diplotene. **Table S4.** GO analysis for substage-concordant genes. Abbreviations for each substage are as follows: Sp'gonia for spermatogonia, EL for early leptotene, LL-Z for late leptotene and zygotene, EP for early pachytene, and LP-D for late pachytene and diplotene. **Table S5.** Analysis of genes annotated to meiosis. **Table S6.** TFs for substage-concordant genes. The column names are explained as follows: "NES" is short for Normalized Enrichment Score in iRegulon, "ngenes" is the number of substage-concordant genes associated with the TF, "name" used in the arc-diagram (* indicates a family of genes, e.g., Rfx* for Rfx1, Rfx2, Rfx3, Rfx4). Bolded TFs are highly expressed in the testis and those highlighted in red are associated with male reproduction. **Table S7.** TFs for substage-negatively concordant genes. The column names are explained as follows: "NES" is short for Normalized Enrichment Score in iRegulon, "ngenes" is the number of substage negatively-concordant genes associated with the TF, "name" used in the arc-diagram (* indicates a family of genes, e.g., Rfx* for Rfx1, Rfx2, Rfx3, Rfx4). **Table S8.** Comparison of substage-concordant gene lists with derived clusters in Margolin, et al. [14]. **Table S9.** Comparison of substage-concordant gene lists with sets derived from FACS-sorted cells [5, 6]. **Table S10.** Primers for qRT-PCR. (ZIP 6513 kb)

**Additional file 2:** This file contains a mini-website of Figures S1-S6 and a link to interactive expression plots. The same mini-website is available at http://carterdev.jax.org/dtx/a1/index.html. **Figure S1.** Expression of substage-concordant genes. (A-H) Heat maps show gene expression of substage-concordant genes at each time point. Gene expression is shown as average log2(TPM + 1) of time point replicates. **Figure S2.** Expression of substage-negatively concordant genes. (A-H) Heat maps show gene expression of substage-negatively concordant genes at each time point. Gene expression is shown as average log2(TPM + 1) of time point replicates. **Figure S3.** Validation of RNA-seq expression pattern by qRT-PCR. Box plots (top) show gene expression detected by RNA-seq at 10 dpp and 16 dpp. Bar graphs (bottom) show gene expression detected by qRT-PCR using isolated germ cells from pooled testes from 3 mice at 10 dpp and 16 dpp. Gene expression for RNA-seq is shown as log2(TPM + 1). Gene expression for qRT-PCR is shown as fold increase relative to *Actb* expression. **Figure S4.** TF arc diagram illustrating the most significant TFs across substages for the concordant gene lists. Node color represents the substage. Node size is related to significance (larger = more significant but all are highly significant with a iRegulon NES > = 4). Width of arc corresponds to how many TFs are shared between nodes (wider = more shared TFs). A node is connected to another node if a node's TFs are a subset of the other node's TFs. Color of the arc (degree) is related to how many connections the node has. Degree equals the number of connections. The name of each node is an abbreviation, where an asterisk indicates multiple members of the TF family are included in the node (e.g., Rfx* for

Rfx1, Rfx2, Rfx3, Rfx4) and when a node name is followed by + integer, as in "Mef2* + 1", it indicates that there are TFs in the Mef2 family plus one other TF. All TFs associated with each node are listed in Additional file 1: Table S6. The number in parenthesis next to the node name indicates the number of substage-concordant genes associated with the TF cluster. **Figure S5.** TF arc diagram illustrating the most significant TFs across substages for the negatively concordant gene lists. Node color represents the substage. Node size is related to significance (larger = more significant but all are highly significant with a iRegulon NES > = 4). Width of arc corresponds to how many TFs are shared between nodes (wider = more shared TFs). A node is connected to another node if a node's TFs are a subset of the other node's TFs. Color of the arc (degree) is related to how many connections the node has. Degree equals the number of connections. The name of each node is an abbreviation, where an asterisk indicates multiple members of the TF family are included in the node (e.g., Rfx* for Rfx1, Rfx2, Rfx3, Rfx4) and when a node name is followed by + integer, as in "Mef2* + 1", it indicates that there are TFs in the Mef2 family plus one other TF. All TFs associated with each node are listed in Additional file 1: Table S7. The number in parenthesis next to the node name indicatea the number of substage-negatively-concordant genes associated with the TF cluster. **Figure S6.** Expression of TF and its target genes for each substage. The difference of mean gene expression at 8 dpp of the TF and its target genes at each time point. To illustrate the overall pattern, a smoothed line was fit to the substage-specific gene expression. Color of lines represents the substage of target genes. Gene expression is shown as average log2(TPM + 1) across time point replicates. (ZIP 13261 kb)

**Additional file 3:** Details of permutation-based maximum covariance analysis (PMCA). (PDF 150 kb)

#### Acknowledgements
We are indebted to members of the Handel and Carter laboratories for discussion of this work in progress, to Sabrina Petri for animal care and technical assistance, and to Drs. G. Howell and S. Munger for critical comments on the manuscript. This work was supported by a P01 grant from the NIH (GM99640) and by fellowships to YF from the Japan Society for the Promotion of Science (JSPS), the Strategic Young Researcher Oversea Visits Program for Accelerating Brain Research. Research reported in this publication was also partially supported by the National Cancer Institute under award number P30 CA034196; the content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

#### Availability of data and materials
Transcriptome data are available in Gene Expression Omnibus under accession number GSE72833. All other data are available as Additional Files as detailed below.

#### Authors' contributions
The project was designed by GWC, MAH and MH with YF and RLB; the research was conducted by YF and RLB with contributions from FS, JH and MH. YF, RLB, GWC and MAH analyzed the data and wrote the manuscript. All authors read and approved the final manuscript.

#### Competing interests
The authors declare that they have no competing interests.

#### Author details
[1]The Jackson Laboratory, Bar Harbor, ME, USA. [2]Department of Computer Science, Trinity University, San Antonio, TX, USA.

Received: 11 September 2015 Accepted: 28 June 2016
Published online: 12 August 2016

#### References
1. Hogarth CA, Arnold S, Kent T, Mitchell D, Isoherranen N, Griswold MD. Processive pulses of retinoic acid propel asynchronous and continuous murine sperm production. Biol Reprod. 2015;92(2):37.
2. Romrell LJ, Bellve AR, Fawcett DW. Separation of mouse spermatogenic cells by sedimentation velocity. A morphological characterization. Dev Biol. 1976;49(1):119–31.
3. Bellve AR. Purification, culture, and fractionation of spermatogenic cells. Methods Enzymol. 1993;225:84–113.
4. Bellve AR, Cavicchia JC, Millette CF, O'Brien DA, Bhatnagar YM, Dym M. Spermatogenic cells of the prepuberal mouse. Isolation and morphological characterization. J Cell Biol. 1977;74(1):68–85.
5. da Cruz I, Rodriguez-Casuriaga R, Santinaque FF, Farias J, Curti G, Capoano CA, Folle GA, Benavente R, Sotelo-Silveira JR, Geisinger A. Transcriptome analysis of highly purified mouse spermatogenic cell populations: gene expression signatures switch from meiotic-to postmeiotic-related processes at pachytene stage. BMC Genomics. 2016;17(1):294.
6. Fallahi M, Getun IV, Wu ZK, Bois PRJ. A global expression switch marks pachytene initiation during mouse male meiosis. Genes. 2010;1(3):469–83.
7. Gaysinskaya V, Soh IY, van der Heijden GW, Bortvin A. Optimized flow cytometry isolation of murine spermatocytes. Cytometry A. 2014;85(6):556–65.
8. Getun IV, Torres B, Bois PR. Flow cytometry purification of mouse meiotic cells. J Vis Exp. 2011;(50):2602.
9. Fujiwara Y. Differential gene expression revealed by transcriptomic analyses of male germ cells. J Anim Genet. 2014;42(2):91–9.
10. Chalmel F, Rolland AD, Niederhauser-Wiederkehr C, Chung SS, Demougin P, Gattiker A, Moore J, Patard JJ, Wolgemuth DJ, Jegou B, et al. The conserved transcriptome in human and rodent male gametogenesis. Proc Natl Acad Sci U S A. 2007;104(20):8346–51.
11. Namekawa SH, Park PJ, Zhang LF, Shima JE, McCarrey JR, Griswold MD, Lee JT. Postmeiotic sex chromatin in the male germline of mice. Curr Biol. 2006; 16(7):660–7.
12. Shima JE, McLean DJ, McCarrey JR, Griswold MD. The murine testicular transcriptome: characterizing gene expression in the testis during the progression of spermatogenesis. Biol Reprod. 2004;71(1):319–30.
13. Kleene KC. A possible meiotic function of the peculiar patterns of gene expression in mammalian spermatogenic cells. Mech Dev. 2001;106(1-2):3–23.
14. Schultz N, Hamra FK, Garbers DL. A multitude of genes expressed solely in meiotic or postmeiotic spermatogenic cells offers a myriad of contraceptive targets. Proc Natl Acad Sci U S A. 2003;100(21):12201–6.
15. Chalmel F, Lardenois A, Evrard B, Rolland AD, Sallou O, Dumargne MC, Coiffec I, Collin O, Primig M, Jegou B. High-resolution profiling of novel transcribed regions during rat spermatogenesis. Biol Reprod. 2014;91(1):5.
16. Laiho A, Kotaja N, Gyenesei A, Sironen A. Transcriptome profiling of the murine testis during the first wave of spermatogenesis. PLoS One. 2013;8(4): e61558.
17. Li XZ, Roy CK, Dong X, Bolcun-Filas E, Wang J, Han BW, Xu J, Moore MJ, Schimenti JC, Weng Z, et al. An ancient transcription factor initiates the burst of piRNA production during early meiosis in mouse testes. Mol Cell. 2013;50(1):67–81.
18. Margolin G, Khil PP, Kim J, Bellani MA, Camerini-Otero RD. Integrated transcriptome analysis of mouse spermatogenesis. BMC Genomics. 2014;15:39.
19. Soumillon M, Necsulea A, Weier M, Brawand D, Zhang X, Gu H, Barthes P, Kokkinaki M, Nef S, Gnirke A, et al. Cellular source and mechanisms of high transcriptome complexity in the mammalian testis. Cell Rep. 2013; 3(6):2179–90.
20. Anderson EL, Baltus AE, Roepers-Gajadien HL, Hassold TJ, de Rooij DG, van Pelt AM, Page DC. Stra8 and its inducer, retinoic acid, regulate meiotic initiation in both spermatogenesis and oogenesis in mice. Proc Natl Acad Sci U S A. 2008;105(39):14976–80.
21. Yuan L, Liu JG, Zhao J, Brundell E, Daneholt B, Hoog C. The murine SCP3 gene is required for synaptonemal complex assembly, chromosome synapsis, and male fertility. Mol Cell. 2000;5(1):73–83.
22. Inselman A, Eaker S, Handel MA. Temporal expression of cell cycle-related proteins during spermatogenesis: establishing a timeline for onset of the meiotic divisions. Cytogenet Genome Res. 2003;103(3-4):277–84.
23. Rogakou EP, Pilch DR, Orr AH, Ivanova VS, Bonner WM. DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. J Biol Chem. 1998;273(10):5858–68.
24. Cobb J, Cargile B, Handel MA. Acquisition of competence to condense metaphase I chromosomes during spermatogenesis. Dev Biol. 1999;205(1):49–64.
25. Xie C, Yuan J, Li H, Li M, Zhao G, Bu D, Zhu W, Wu W, Chen R, Zhao Y. NONCODEv4: exploring the world of long non-coding RNA genes. Nucleic Acids Res. 2014;42(Database issue):D98–103.
26. La Salle S, Sun F, Handel MA. Isolation and short-term culture of mouse spermatocytes for analysis of meiosis. Methods Mol Biol. 2009;558:279–97.
27. Handel MA. The XY body: a specialized meiotic chromatin domain. Exp Cell Res. 2004;296(1):57–63.

28. Turner JM. Meiotic sex chromosome inactivation. Development. 2007; 134(10):1823–31.
29. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet. 2000;25(1):25–9.
30. Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics. 2009;10:48.
31. Iguchi N, Tobias JW, Hecht NB. Expression profiling reveals meiotic male germ cell mRNAs that are translationally up- and down-regulated. Proc Natl Acad Sci U S A. 2006;103(20):7712–7.
32. Monesi V. Ribonucleic acid synthesis during mitosis and meiosis in the mouse testis. J Cell Biol. 1964;22:521–32.
33. Smith BE, Braun RE. Germ cell migration across Sertoli cell tight junctions. Science. 2012;338(6108):798–802.
34. Watanabe T, Lin H. Posttranscriptional regulation of gene expression by Piwi proteins and piRNAs. Mol Cell. 2014;56(1):18–27.
35. Janky R, Verfaillie A, Imrichova H, Van de Sande B, Standaert L, Christiaens V, Hulselmans G, Herten K, Naval Sanchez M, Potier D, et al. iRegulon: from a gene list to a gene regulatory network using large motif and track collections. PLoS Comput Biol. 2014;10(7):e1003731.
36. Smith JJ, Ramsey SA, Marelli M, Marzolf B, Hwang D, Saleem RA, Rachubinski RA, Aitchison JD. Transcriptional responses to fatty acid are coordinated by combinatorial control. Mol Syst Biol. 2007;3:115.
37. Bruckner S, Kohler T, Braus GH, Heise B, Bolte M, Mosch HU. Differential regulation of Tec1 by Fus3 and Kss1 confers signaling specificity in yeast development. Curr Genet. 2004;46(6):331–42.
38. Lahdesmaki H, Shmulevich L, Dunmire V, Yli-Harja O, Zhang W. In silico microdissection of microarray data from heterogeneous cell populations. BMC Bioinformatics. 2005;6:54.
39. Lu P, Nakorchevskiy A, Marcotte EM. Expression deconvolution: a reinterpretation of DNA microarray data reveals dynamic changes in cell populations. Proc Natl Acad Sci U S A. 2003;100(18):10370–5.
40. Shen-Orr SS, Tibshirani R, Khatri P, Bodian DL, Staedtler F, Perry NM, Hastie T, Sarwal MM, Davis MM, Butte AJ. Cell type-specific gene expression differences in complex tissues. Nat Methods. 2010;7(4):287–9.
41. Prohaska J. A technique for analyzing the linear relationships between two meteorological fields. Mon Weather Rev. 1976;104(11):1345–53.
42. Bretherton CS, Smith C, Wallace JM. An intercomparison of methods for finding coupled patterns in climate data. J Clim. 1992;5(6):541–60.
43. Wallace JM, Smith C, Bretherton CS. Singular value decomposition of wintertime sea-surface-temperature and 500 mb height anomalies. J Clim. 1992;5(6):561–76.
44. Tan CS, Salim A, Ploner A, Lehtio J, Chia KS, Pawitan Y. Correlating gene and protein expression data using Correlated Factor Analysis. BMC Bioinformatics. 2009;10:272.
45. Cherry S. Singular value decomposition and canonical correlation analysis. J Clim. 1996;9(9):2003–9.
46. Cherry S. Some comments on singular value decomposition analysis. J Clim. 1997;10(7):1759–61.
47. Salim A, Pawitan Y, Bond K. Modelling association between two irregularly observed spatiotemporal processes by using maximum covariance analysis. Appl Statist. 2005;54(3):555–73.
48. Handel MA, Schimenti JC. Genetics of mammalian meiosis: regulation, dynamics and impact on fertility. Nat Rev Genet. 2010;11(2):124–36.
49. La Salle S, Palmer K, O'Brien M, Schimenti JC, Eppig J, Handel MA. Spata22, a novel vertebrate-specific gene, is required for meiotic progress in mouse germ cells. Biol Reprod. 2012;86(2):45.
50. Edelmann W, Cohen PE, Kneitz B, Winand N, Lia M, Heyer J, Kolodner R, Pollard JW, Kucherlapati R. Mammalian MutS homologue 5 is required for chromosome pairing in meiosis. Nat Genet. 1999;21(1):123–7.
51. Kneitz B, Cohen PE, Avdievich E, Zhu LY, Kane MF, Hou H, Kolodner RD, Kucherlapati R, Pollard JW, Edelmann W. MutS homolog 4 localization to meiotic chromosomes is required for chromosome pairing during meiosis in male and female mice. Genes Dev. 2000;14(9):1085–97.
52. Yuan J, Chen J. FIGNL1-containing protein complex is required for efficient homologous recombination repair. Proc Natl Acad Sci U S A. 2013;110(26): 10640–5.
53. Lalancette C, Miller D, Li Y, Krawetz SA. Paternal contributions: new functional insights for spermatozoal RNA. J Cell Biochem. 2008;104(5): 1570–9.
54. Halbig KM, Lekven AC, Kunkel GR. The transcriptional activator ZNF143 is essential for normal development in zebrafish. BMC Mol Biol. 2012;13:3.
55. Chen X, Fang F, Liou YC, Ng HH. Zfp143 regulates Nanog through modulation of Oct4 binding. Stem Cells. 2008;26(11):2759–67.
56. Chia NY, Chan YS, Feng B, Lu X, Orlov YL, Moreau D, Kumar P, Yang L, Jiang J, Lau MS, et al. A genome-wide RNAi screen reveals determinants of human embryonic stem cell identity. Nature. 2010;468(7321):316–20.
57. Myslinski E, Krol A, Carbon P. ZNF76 and ZNF143 are two human homologs of the transcriptional activator Staf. J Biol Chem. 1998;273(34):21998–2006.
58. Bailey SD, Zhang X, Desai K, Aid M, Corradin O, Cowper-Sal Lari R, Akhtar-Zaidi B, Scacheri PC, Haibe-Kains B, Lupien M. ZNF143 provides sequence specificity to secure chromatin interactions at gene promoters. Nat Commun. 2015;2:6186.
59. Ishiguro K, Kim J, Fujiyama-Nakamura S, Kato S, Watanabe Y. A new meiosis-specific cohesin complex implicated in the cohesin code for homologous pairing. EMBO Rep. 2011;12(3):267–75.
60. Bolcun-Filas E, Bannister LA, Barash A, Schimenti KJ, Hartford SA, Eppig JJ, Handel MA, Shen L, Schimenti JC. A-MYB (MYBL1) transcription factor is a master regulator of male meiosis. Development. 2011;138(15):3319–30.
61. Toscani A, Mettus RV, Coupland R, Simpkins H, Litvin J, Orth J, Hatton KS, Reddy EP. Arrest of spermatogenesis and defective breast development in mice lacking A-myb. Nature. 1997;386(6626):713–7.
62. Chuma S, Nakano T. piRNA and spermatogenesis in mice. Philos Trans R Soc Lond B Biol Sci. 2013;368(1609):20110338.
63. Bustin SA. Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. J Mol Endocrinol. 2000; 25(2):169–93.
64. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 2009;10(3):R25.
65. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics. 2011;12:323.
66. Flicek P, Amode MR, Barrell D, Beal K, Billis K, Brent S, Carvalho-Silva D, Clapham P, Coates G, Fitzgerald S, et al. Ensembl 2014. Nucleic Acids Res. 2014;42(Database issue):D749–755.
67. Fujita PA, Rhead B, Zweig AS, Hinrichs AS, Karolchik D, Cline MS, Goldman M, Barber GP, Clawson H, Coelho A, et al. The UCSC Genome Browser database: update 2011. Nucleic Acids Res. 2011;39(Database issue):D876–882.
68. Adler D, Murdoch D, Nenadic O, Urbanek S, Chen M, Gebhardt A, Bolker B, Csardi G, Strzelecki A, Senger A. rgl: 3D visualization device system (OpenGL). 2014.
69. R Core Team. R: A Language and Environment for Statistical Computing. In. Vienna, Austria: R Foundation for Statistical Computing; 2014.
70. Harris MA, Clark J, Ireland A, Lomax J, Ashburner M, Foulger R, Eilbeck K, Lewis S, Marshall B, Mungall C, et al. The Gene Ontology (GO) database and informatics resource. Nucleic Acids Res. 2004;32(Database issue):D258–261.
71. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13(11): 2498–504.

# Germline stem cells: toward the regeneration of spermatogenesis

Hanna Valli, B.S.,[a,b] Bart T. Phillips, Ph.D.,[a,b] Gunapala Shetty, Ph.D.,[c] James A. Byrne, Ph.D.,[d,e] Amander T. Clark, Ph.D.,[e] Marvin L. Meistrich, Ph.D.,[c] and Kyle E. Orwig, Ph.D.[a,b]

[a] Department of Obstetrics, Gynecology and Reproductive Sciences, Molecular Genetics and Developmental Biology Graduate Program, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania; [b] Magee-Womens Research Institute, Pittsburgh, Pennsylvania; [c] Department of Experimental Radiation Oncology, The University of Texas MD Anderson Cancer Center, Houston, Texas; [d] Department of Molecular and Medical Pharmacology, Center for Health Sciences, Los Angeles, California; and [e] Eli and Edythe Broad Center of Regenerative Medicine and Stem Cell Research, University of California, Los Angeles, California

Improved therapies for cancer and other conditions have resulted in a growing population of long-term survivors. Infertility is an unfortunate side effect of some cancer therapies that impacts the quality of life of survivors who are in their reproductive or prereproductive years. Some of these patients have the opportunity to preserve their fertility using standard technologies that include sperm, egg, or embryo banking, followed by IVF and/or ET. However, these options are not available to all patients, especially the prepubertal patients who are not yet producing mature gametes. For these patients, there are several stem cell technologies in the research pipeline that may give rise to new fertility options and allow infertile patients to have their own biological children. We will review the role of stem cells in normal spermatogenesis as well as experimental stem cell–based techniques that may have potential to generate or regenerate spermatogenesis and sperm. We will present these technologies in the context of the fertility preservation paradigm, but we anticipate that they will have broad implications for the assisted reproduction field. (Fertil Steril® 2014;101:3–13. ©2014 by American Society for Reproductive Medicine.)

**Key Words:** Male fertility, male infertility, regenerative medicine, spermatogonial stem cells, stem cells

**Discuss:** You can discuss this article with its authors and with other ASRM members at http://fertstertforum.com/vallih-germline-stem-cells-spermatogenesis/

Use your smartphone to scan this QR code and connect to the discussion forum for this article now.*

* Download a free QR code scanner by searching for "QR scanner" in your smartphone's app store or app marketplace.

High-dose chemotherapy, whole-body radiation, or radiation to the gonads can cause permanent infertility (1). This is a significant human health concern because over 75,000 people under the age of 40 in the United States are diagnosed with cancer each year and most are cured (2). Thus, cancer patients can look beyond their diagnosis and treatment to quality of life after cancer. Parenthood is important to cancer survivors, and distress over infertility can have long-term psychological and relationship implications (3). Therefore, the American Society for Clinical Oncology (ASCO) (4) and the American Society for Reproductive Medicine Ethics Committee (5) recommend that the reproductive risks of gonadotoxic therapies and options for preserving fertility be discussed with patients before initiating treatment. While adoption and third-party reproduction provide alternative family-building options, the available data indicate that most cancer survivors prefer to have their own biological children (4).

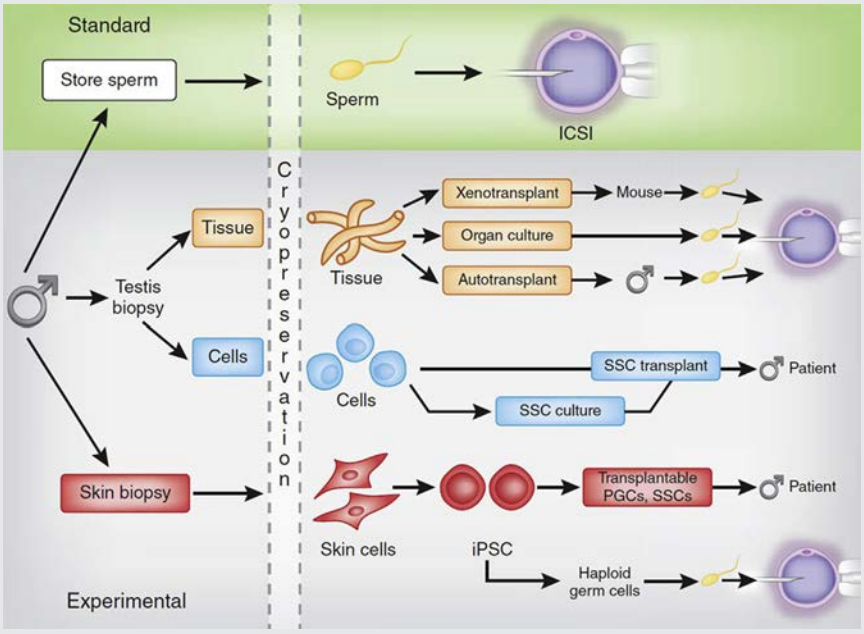
Postpubertal adolescent and adult males have the option to cryopreserve sperm before oncological treatment (Fig. 1, *top*). This is a simple and established method for preserving fertility potential and allows men to father their own genetic children. Nearly 17,000 men between the ages of 15 and 44 are diagnosed with cancer each year in the United States, and nearly 2,385 survivors will receive a treatment that puts them at high risk of azoospermia (2, 6). Unfortunately, only about 24% of men in this age range cryopreserve semen before their oncological treatment (7). Therefore, we calculate that each year in the United States, over 1,800 adult cancer survivors will be infertile with azoospermia and have limited options to have their

Received September 19, 2013; revised October 22, 2013; accepted October 28, 2013; published online December 5, 2013.
H.V. has nothing to disclose. B.T.P. has nothing to disclose. G.S. has nothing to disclose. J.A.B. has nothing to disclose. A.T.C. has nothing to disclose. M.L.M. has nothing to disclose. K.E.O. reports compensation for service on National Institutes of Health study section and honorarium as Keynote Speaker, American Association of Tissue Banks Annual Conference 2013.
Supported by National Institutes of Health Eunice Kennedy Shriver National Institute of Child Health and Human Development grant nos. HD055475 and HD061289; Magee-Womens Research Institute and Foundation; Richard King Mellon Foundation; and the United States–Israel Binational Science Foundation. (to K.E.O.).
Current address (B.T.P.): Laboratory of Structural Biology, National Institute of Environmental Health Sciences, National Institute of Health, Research Triangle Park, NC 27709.
Reprint requests: Kyle E. Orwig, Ph.D., Magee-Womens Research Institute, 204 Craft Avenue, Pittsburgh, PA 15213 (E-mail: orwigke@upmc.edu).

Fertility and Sterility® Vol. 101, No. 1, January 2014 0015-0282/$36.00

http://dx.doi.org/10.1016/j.fertnstert.2013.10.052

**FIGURE 1**




Standard and experimental options for preserving male fertility. *Top*, sperm obtained by ejaculation or surgical retrieval from the testes or epididymides are competent to fertilize oocytes using assisted reproductive techniques including IUI, IVF, or IVF ICSI that are standard in most fertility clinics. These options are not available to prepubertal boys who are not producing sperm or to adult azoospermic men. *Bottom*, testis tissue obtained via biopsy from prepubertal boys contains spermatogonial stem cells (SSCs) that can produce sperm in the context of the intact tissue by xenotransplant, organ culture, or autologous transplantation back into the individual (*orange boxes*). Sperm retrieved from cultured or transplanted tissue can be used for ICSI. Cells in suspension obtained from biopsied testicular tissue can be transplanted back into the endogenous seminiferous tubules of the patient (*blue boxes*). SSCs in the suspension can regenerate spermatogenesis and, in some cases, fertility. For infertile individuals who did not preserve germ cells before gonadotoxic therapy, iPSCs may be produced from their somatic cells (e.g., skin or blood) to differentiate into transplantable germ cells (primordial germ cells [PGCs] or SSCs) or haploid germ cells that can be used for ICSI (*red boxes*). Excerpted with permission from Clark AT, Phillips BT, Orwig KE, Nat Med 2000;1564–5.

*Valli. Stem cell therapies for male factor infertility. Fertil Steril 2014.*

own biological children because they did not save a semen sample. In some cases, sperm can be recovered surgically from small focal areas of spermatogenesis in the testes using the testicular sperm extraction method and can be used to fertilize oocytes by intracytoplasmic sperm injection (ICSI) (8).

There are no options to preserve the fertility of prepubertal boys, who are not yet making sperm. This is a significant problem because about 5,131 boys under the age of 15 in the United States are expected to develop cancer each year and 83% are expected to survive (2). A report from the Childhood Cancer Survivor Study indicates that the cytotoxic therapies for cancer reduce the number of young men subsequently able to have children by 44% (6, 9). Based on these statistics, we calculate that each year in the United States, 1,874 young male cancer patients will become sterile owing to their treatment. In addition to cancer survivors, over 500 patients under the age of 20 receive hematopoietic stem cell transplants each year in the United States for nonmalignant conditions (e.g., bone marrow failure, blood and immune deficiencies, autoimmune disorders) (10). Myeloablative conditioning therapy before bone marrow transplantation is associated with a high risk of infertility (4, 9, 11, 12). The ASCO report notes that "Impaired future fertility is difficult for children to understand, but potentially traumatic to them as adults" (4). The available data indicate that greater than 80% of parents consented to fertility preservation procedures on behalf of their children before initiation of gonadotoxic therapies (13, 14).

The summed incidence of chemotherapy or radiation-induced male factor infertility that cannot be treated with existing reproductive therapies is approximately 4,000 individuals each year in the United States. Therefore, responsible development of novel therapies to help these patients have biological children has a significant potential impact.

Promising results in animal models and human cell lines (Fig. 1, *bottom*) have generated enthusiasm that stem cells might be used or manipulated to preserve and/or restore the fertility of patients who are not producing sperm and have no other options to protect their future fertility before receiving gonadotoxic chemotherapy or radiation treatments (14–34). We will review the methods of spermatogonial stem cell (SSC) transplantation (Fig. 1, *blue boxes*), testicular tissue grafting, testicular tissue organ culture (Fig. 1, *orange boxes*), and induced pluripotent stem cell differentiation into gametes or transplantable male germ line stem cells (Fig. 1, *red boxes*). Table 1 summarizes published reports detailing the progress of each method. Enthusiasm for these experimental stem cell technologies is tempered by concerns about feasibility and safety, particularly for the vulnerable prepubertal

## TABLE 1

**Literature reporting progress in stem cell technology development.**

| Stem cell technologies | (q)Real time polymerase chain reaction | Histology | Immunocytochemistry | Immunohistochemistry | Flow/flow cytometry/ magnetic cell sorting | Xenotransplantation | Homologous transplantation | Autologous transplantation | Sperm[a] | Fertilization | Progeny |
|---|---|---|---|---|---|---|---|---|---|---|---|
| SSC transplant | | | | | | | | | | | |
| Rodents | | (30, 31, 48–51, 57, 58, 118, 119) | | (71, 120) | | (71, 118–120) | (30, 31, 48–51, 57, 58, 70, 121) | | (30, 31, 48, 50, 51, 57, 58, 71, 118–121) | (31, 48, 50, 51, 57, 58, 71, 120, 121) | (31, 48, 50, 51, 57, 58, 71, 120, 121) |
| Large animal[b] | | (54–56) | | (56, 64, 72) | | (64, 72) | (52–56) | (56) | (52–55) | (52, 55) | (52, 55) |
| Nonhuman primate | | (59, 60) | | (122–125) | | (122–125) | (29) | (29, 59–61) | (29, 59) | (29) | |
| Human | | | | (21, 38, 78, 89, 109, 115, 117) | | (21, 38, 78, 89, 109, 115, 117) | | (67) | | | |
| SSC culture | | | | | | | | | | | |
| Rodents | | (57, 58) | (70, 83) | (71) | (58, 70, 71, 121) | (71) | (57, 58, 70, 71, 121) | | (57, 58, 71, 83, 121) | (57, 58, 71, 121) | (57, 58, 71, 121) |
| Large animal | (73) | | (72–75) | (72) | | (72) | | | | | |
| Nonhuman primate | (76) | | (76) | | | | | | | | |
| Human | (21, 38, 79, 81) | | (77, 79, 81, 82, 126) | (21, 38, 78, 80) | (77, 80–82, 126) | (21, 38, 79) | | | | | |
| Testicular grafting | | | | | | | | | | | |
| Rodents | | (92, 93, 95) | | | | (95) | (92, 93, 95) | | (92, 93) | (92, 93) | (93) |
| Large animal | | (92) | | | | (92) | | | (92) | (92) | |
| Nonhuman primate | | (94–98) | | (97, 98) | | (94–96) | | (97, 98) | (94, 98) | (94) | |
| Human | | (99, 101–104) | | (100, 101, 103, 104) | | (99–104) | | | | | |
| Testicular organ culture | | | | | | | | | | | |
| Rodents | | (27) | | | | | (27) | | (27) | (27) | (27) |
| Large animal | | | | | | | | | | | |
| Nonhuman primate | | | | | | | | | | | |
| Human | | | | | | | | | | | |
| Pluripotent-derived male germ cells | | | | | | | | | | | |
| Rodents | | | (28) | | (28) | | (28) | | (28) | (28) | (28) |
| Large animal | | | | | | | | | | | |
| Nonhuman primate | (32, 33) | | (32, 33) | | | | | | | | |
| Human | (22–24, 26) | | (22–26) | | (22, 23, 25, 26) | | | | | | |

[a] Sperm and spermatogenesis.
[b] Farm animals and companion animals.

*Valli. Stem cell therapies for male factor infertility. Fertil Steril 2014.*

patient population (14, 35–37). Questions about feasibility stem from the early stage of technology development, uncertainty about optimal freezing conditions, concerns that a small testicular biopsy will contain few stem cells, and the lack of culture conditions to expand human germ line stem cells. Questions about safety are associated with the risks of surgery to obtain testicular tissue and the potential for malignant contamination in transplanted cells or tissue. Nonetheless, academic centers around the world are already cryopreserving testicular tissues for prepubertal boys and men in anticipation that stem cells in that tissue can be used safely and effectively to restore future fertility (13, 16, 21, 38–41).

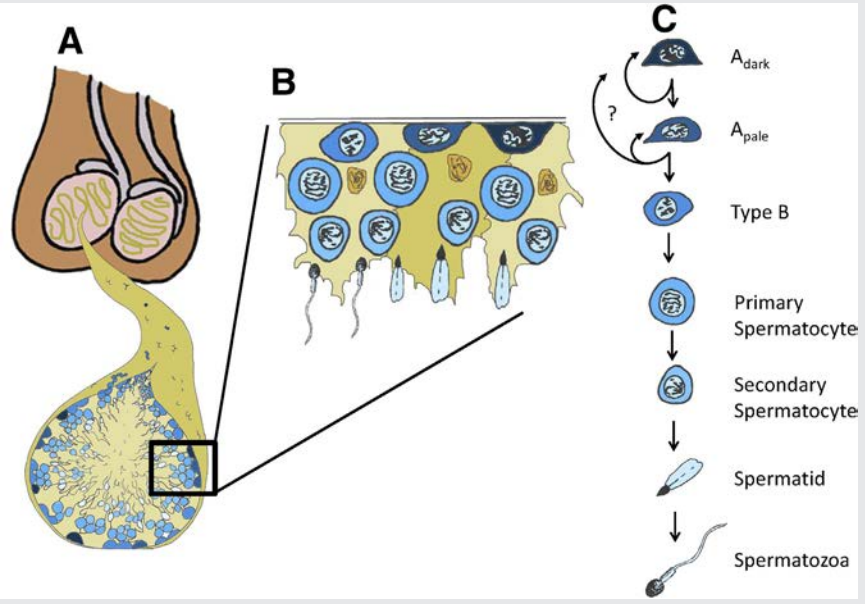
## SSCs AND SPERMATOGENESIS

SSCs are the adult tissue stem cells in the testes that balance self-renewing and differentiating divisions to maintain continuous sperm production throughout the postpubertal life of men. SSCs are rare cells located on the basement membrane of seminiferous tubules (Fig. 2A and B). These stem cells give rise to undifferentiated progenitor spermatogonia that undergo several transit-amplifying mitotic divisions, followed by two meiotic divisions and morphological differentiation to produce sperm. Spermatogenesis is an extraordinarily productive system that produces approximately 40 million sperm per gram of tissue per day in mice (42) and 5.5 million sperm per gram of tissue per day in men (43). The difference in sperm production in mice and men can be explained in part by differences in the number of transit-amplifying mitotic divisions that precede meiosis (44). In humans, the SSC pool is composed of $A_{dark}$ and $A_{pale}$ spermatogonia. $A_{dark}$ are relatively quiescent, while $A_{pale}$ are mitotically active and undergo one to two transit-amplifying divisions before giving rise to differentiating type B spermatogonia and then primary spermatocytes, which enter meiosis and migrate off the basement membrane of the seminiferous tubule (Fig. 2B and C). Owing to its highly proliferative nature, the spermatogenic system is an unintended target of chemo- and radiotherapies. The relative effects of various chemo- and radiotherapy regimens on male fertility are detailed in several previous reports (9, 11, 19, 45, 46). The extent that the fertility of a patient is affected depends on the dose, type, and frequency of treatment (1). There has been significant progress in minimizing the unwanted side effects of the cancer treatment by adjusting and modifying the therapeutic regimens without compromising the efficiency of oncological treatments (47). Nonetheless, some cancer treatments, like whole-body radiation, radiation to the gonads, and alkylating chemotherapy, can be particularly gonadotoxic and result in prolonged or permanent azoospermia. In the following sections, we will review experimental stem cell–based approaches that may have application for preserving and/or restoring male fertility.

## SSC TRANSPLANTATION

Brinster and colleagues pioneered the technique for SSC transplantation in mice in 1994, demonstrating that donor SSCs could engraft the seminiferous tubules of

**FIGURE 2**




Human SSCs and spermatogenesis. (A) Testes are composed of seminiferous tubules that start and end at the rete testis. (B) Cut-out of the basement membrane of the seminiferous epithelium. (B and C) The basement membrane of the seminiferous epithelium contains undifferentiated ($A_{dark}$ and $A_{pale}$) spermatogonia and differentiating type B spermatogonia. Type B spermatogonia give rise to primary spermatocytes that enter meiosis and migrate off the basement membrane. Subsequent meiotic divisions and morphogenesis give rise to secondary spermatocytes, spermatids, and the terminally differentiated spermatozoa that are released into the lumen of the seminiferous tubules.

*Valli. Stem cell therapies for male factor infertility. Fertil Steril 2014.*

chemotherapy-treated recipient mice and regenerate spermatogenesis and fertility leading to the production of viable progeny through normal breeding (31, 31). The SSC transplantation technique has become the experimental gold standard for quantifying stem cell activity and may have application for treating male factor infertility. Homologous species SSC transplantation has now been reported in mice, rats, pigs, goats, bulls, sheep, dogs, and monkeys, including the production of donor-derived progeny in mice, rats, goats, and sheep (29, 48–61). SSCs from donors of all ages, newborn to adult, can regenerate spermatogenesis (49, 62), and SSCs can be cryopreserved and retain spermatogenic function upon thawing and transplantation (29, 63, 64). We recently demonstrated that SSCs from the testes of prepubertal and adult rhesus macaques could be frozen, thawed, and transplanted to regenerate spermatogenesis and produce fertilization-competent sperm (29, 59). Progress in these previous studies suggests that it should be feasible for prepubertal boys or adult men to cryopreserve testicular tissue containing SSCs before treatment and have these cells reintroduced into their testes at a later date to regenerate spermatogenesis.

Radford and colleagues initially introduced the autologous SSC transplantation technique to the human clinic in 1999 (65). In Manchester, United Kingdom, testicular tissue from 12 male non-Hodgkin's lymphoma patients was cryopreserved as a cell suspension before the initiation of chemotherapy. At later dates, seven of the patients had the cells injected back into their testes (66, 67). To our knowledge, there have been no follow-up reports on the fertility status of those patients so the outcome of the experiment is unknown. Even if the men in that study fathered children, it would be difficult to demonstrate unequivocally (in the absence of a unique genetic marker) that those offspring resulted from sperm produced by transplanted stem cells rather than from surviving endogenous stem cells. There have been no subsequent reports of SSC transplantation in humans since the 1999 study. Nonetheless, this bold, pioneering study demonstrated that patients are willing to pursue experimental stem cell approaches to achieve fertility (65, 67). To date, published reports indicate that testicular tissue or cells have been cryopreserved for more than 150 prepubertal and adult male patients worldwide, and this is likely an underestimate of actual cases (13, 14, 21, 38–41, 65, 67).

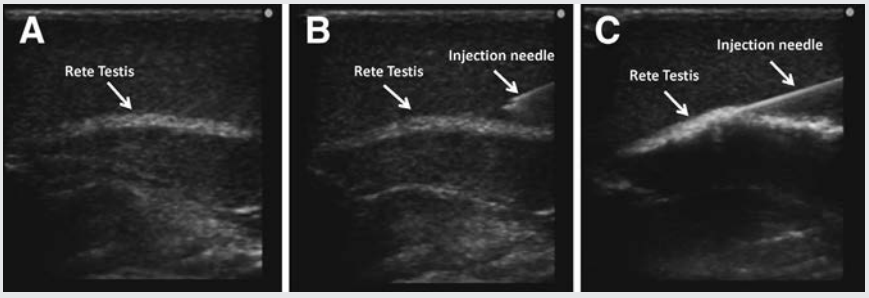
For SSC transplantation in rodents, the testes are typically accessed via a midventral abdominal incision. Testicular cells (including SSCs) are injected using a pulled glass capillary pipette inserted via the efferent ducts into the rete testis space, which can be visualized on the surface of the testis and is contiguous with all seminiferous tubules (68). Testis anatomy in larger animals, including nonhuman primates and humans, is different than that in rodents, with the rete testis being centrally located in the testes. Schlatt and colleagues (69) demonstrated that ultrasound can be used to visualize the rete testis and guide an injection needle into the rete testis space. Ultrasound-guided rete testis injection has now been used for SSC transplantation in several large animals species, including nonhuman primates (29, 52–56, 59–61). For this approach, an injection needle is simply inserted under ultrasound guidance through the scrotal skin and testicular parenchyma into the rete testis space (Fig. 3 depicts an example from nonhuman primates) (29).

Clinical translation of the SSC transplantation technique appears imminent considering successes in several large animal models (Table 1) and that many patients have already cryopreserved testicular tissue or cells. However, several questions and/or challenges remain. First, the small amount of tissue that can be obtained from testicular biopsies (especially from prepubertal boys) may contain a relatively small number of stem cells. Second, in the cancer survivor paradigm, it is essential to eliminate the risk of reintroducing malignant cells. The sections below describe approaches that may circumvent these two challenges. A third challenge for translating new stem cell technologies to the clinic is the limited availability of reagents and experimental tools for studying humans SSCs.

## SSC CULTURE

In rodents, SSCs can be greatly expanded in culture and maintain competence to generate spermatogenesis and restore fertility upon transplantation (57, 58, 70, 71). Several groups

**FIGURE 3**




Ultrasound-guided rete testis injections. For SSC transplantation into larger animals, including nonhuman primates, ultrasound is used to visualize the rete testis (echo-dense structure) (A). (B and C) The injection needle is inserted under ultrasound guidance through the scrotal skin into the rete testis space, which is continuous with the seminiferous tubules. (C) Positive pressure is applied so the cells are slowly injected into the rete testis space and seminiferous tubules.

*Valli. Stem cell therapies for male factor infertility. Fertil Steril 2014.*

have reported extending SSC culture to large animal species (72–76). This approach could theoretically be employed to amplify the small number of stem cells that can be obtained from a biopsy and also to demonstrate the absence of malignant contamination. A number of laboratories have reported culturing human SSCs (21, 38, 77–82), including three from the testes of prepubertal patients (38, 78, 80). These human tissue studies are promising but challenged by the inability to evaluate the full spermatogenic potential of cultured cells by homologous species transplantation into human testes. To date, each laboratory has used a different approach to culture human SSCs and different methods to assess outcomes. These published human SSC culture results need to be replicated in other laboratories and evaluated using robust markers of human spermatogonia and by xenotransplantation to nude mice (see section below on experimental methods to track human germ line stem cells).

In a recent study, Elhija and colleagues reported that testicular cells from 7-day-old immature mice could be expanded in a three-dimensional soft agar culture system and differentiated to produce postmeiotic germ cells, including morphologically normal sperm (83). Additional studies are needed to confirm that the resulting sperm are functionally competent to fertilize mouse oocytes. If a similar approach is successful in humans, this would eliminate the need to put cells back into the patient and thus the risk of reintroducing malignant cells into a cancer survivor.

## CELL-SORTING STRATEGIES TO ISOLATE SSCS AND REMOVE MALIGNANT CONTAMINATION

There are legitimate concerns about malignant contamination of preserved testicular tissue, especially from leukemia patients (84, 85). Several groups have demonstrated that it is feasible to isolate germ cells and remove malignant contamination from heterogeneous testicular cell suspensions using a fluorescence-activated cell sorter (FACS) and combinations of germ cell and/or cancer cell markers (86–90). Despite these encouraging results, caution is warranted because even low levels of contamination can lead to cancer and current assays may lack the sensitivity to detect very rare malignant cells (85, 91). Therefore, alternative methods for decontamination and screening as well as methods that do not require transplantation should be considered in some cases.

## TESTICULAR TISSUE GRAFTING

In contrast to SSC transplantation, which involves disaggregation of SSCs from their cognate niches, testicular tissue grafting and testicular tissue organ culture (next section) maintain the integrity of the stem cell/niche unit. Honaramooz and coworkers demonstrated that testicular tissues obtained from newborn mice, pigs, and goats could produce complete spermatogenesis when grafted under the skin of nude mice (92) and later reported the production of offspring from sperm obtained after ectopic (under the skin) grafting of mouse testicular tissue (93). Testis tissue from prepubertal rhesus macaques also produced complete spermatogenesis with fertilization-competent sperm after ectopic xenografting into nude mice (94). Survival and spermatogenesis from adult testicular tissue grafts have been less successful than immature grafts (95).

Cryopreservation is an essential component of the fertility preservation paradigm. Jahnukainen and coworkers evaluated several cryopreservation strategies with immature monkey tissues and found that freezing in 1.4 M DMSO provided good graft survival and spermatogenic induction up to the spermatocyte stage (96). However, haploid germ cells were not produced in any grafts from that study, fresh or frozen. Wistuba and colleagues performed autologous testicular grafting in two studies in marmoset monkeys (97, 98) and reported that complete spermatogenesis was obtained in orthotopic (in the scrotum) but not ectopic grafts. Frozen and thawed tissues were also grafted in that study, but these were only transplanted ectopically and did not produce spermatogenesis (98). Therefore, the question of whether frozen and thawed grafts can produce fertilization-competent haploid germ cells still needs to be addressed.

Several groups have reported xenografting of human testicular tissue to nude mice, and so far none have observed complete spermatogenesis or haploid gametes (99–104). In general, testicular tissues xenografted from prepubertal or adolescent boys survived long term (4–12 months) (100, 101, 103, 104), with a calculated 3.7% spermatogonial recovery rate after 6 months in one study (104) and three studies reporting spermatocytes as the most advanced stage of germ cell development between 6 and 12 months after transplantation (101, 103, 104). Some of those tissues were already postpubertal and contained meiotic or postmeiotic cells at the time of transplant, so it is difficult to exclude the possibility that the spermatocytes or occasional spermatids that were observed had persisted from the time of transplant (101, 104). Importantly, the study by Sato and colleagues (103) observed primary spermatocytes 1 year after xenografting of testicular tissue from a 3-month-old boy, which clearly did not contain spermatocytes before grafting. In contrast to prepubertal/adolescent tissues, xenografted adult testicular tissues regressed over time (99, 102). Although the human-to-mouse xenografting results are somewhat discouraging, the results of the monkey studies cited above suggest that autologous transplantation may be an option if suitable cryopreservation conditions are developed. Similar to SSC transplantation, autologous grafting will be problematic in cases where malignant contamination of the testicular tissue is suspected. Xenografting of human testicular tissue into animals could circumvent this problem but is associated with additional concerns about xenobiotics and has been unsuccessful to date.

## TESTICULAR TISSUE ORGAN CULTURE

Sato and colleagues (27, 105) demonstrated that haploid male germ cells from mice could be generated in testicular tissue organ culture, using a gas-liquid interface method that was originally devised to keep differentiated organs alive in vitro (106). Testis tissue was obtained from 2.5- to 3.5-day-old mice that contained only undifferentiated germ cells,

similar to prepubertal patients. Testicular tissue was minced into fragments (1–3 mm diameter) that were placed on an agarose gel that was half soaked in medium, such that the tissue was exposed to air and absorbed nutrients through the agarose. Haploid round spermatids or sperm were retrieved after 23 or 42 days in culture, respectively, and used to successfully fertilize mouse oocytes in vitro by round spermatid injection and ICSI. The resulting embryos were transferred to recipient females and generated live offspring with normal development to adulthood and normal fertility (27). The investigators also maintained frozen and thawed testicular tissue pieces in organ culture and were able to generate sperm, but the fertilization potential of that sperm was not tested. If these results in mice can be translated to humans, testicular organ culture would circumvent the need to put tissues or cells back into the patient and may be a safe option for patients with malignancies that contaminate the testes.

## GERM CELLS DERIVED FROM PLURIPOTENT STEM CELLS

For males who did not preserve sperm or SSCs before gonadotoxic treatment, generation of transplantable germ cells (primordial germ cells or SSCs) or haploid gametes from embryonic stem cells (ESCs) or induced pluripotent stem cells (iPSCs) has been explored. The iPSC technology is especially exciting because it would theoretically allow patients with no germ cells in their testes to father genetically related children. The derivation of germ cells from pluripotent cells will be covered in more detail in other articles in this issue. Here we will state briefly that several groups have reported the generation of germ cells from nonhuman primate ESCs (32, 33) and human ESCs and/or iPSCs (22–26), including evidence of rare haploid cells in some cases. These results are exciting and revolutionary but are challenged by the inability to test the function of the resulting germ cells in the human system. Therefore, safety and feasibility will need to be established in animal models. In 2011, Hayashi et al. showed that mouse ESCs and iPSCs can be induced to an epiblast-like cell that then gives rise to primordial germ cells when BMP4 was added to a culture medium (28). The resulting germ cells were transplanted into the testes of infertile mice and generated spermatogenesis with sperm that were used to fertilize oocytes using ICSI. The resulting embryos were transferred to recipient females and gave rise to viable offspring. These promising results need to be replicated in other laboratories, and more research is needed before this technique can be translated to the clinic, as some of the offspring developed malignant tumors around their neck area (28). For iPSC-based technologies, it will be important to understand the effects of reprogramming, culture, and differentiation on the genome and epigenome of in vitro–derived germ cells compared with endogenous germ cells (107).

## EXPERIMENTAL METHODS TO TRACK AND QUANTIFY HUMAN GERM LINE STEM CELLS

Studies on human cells and/or tissues are a valuable stepping stone toward clinical translation. However, these studies are challenged by the limited experimental tools for quantifying human spermatogonia and testing their function. In animal studies, the most compelling experiments are those that demonstrate the ability of a test cell population to regenerate spermatogenesis and produce fertilization-competent sperm, embryos, and offspring. This is an unrealistic expectation for human studies because fertilization of human eggs and production of human embryos are not universally legal or fundable. Based on progress in animal models, human-to-human SSC transplantation is likely to be feasible in the clinical setting, but this is not an option for the routine testing of spermatogenic potential in the laboratory setting. Therefore, experimental assessment of human germ line stem cell potential must rely on descriptive endpoints and xenotransplantation for functional testing.

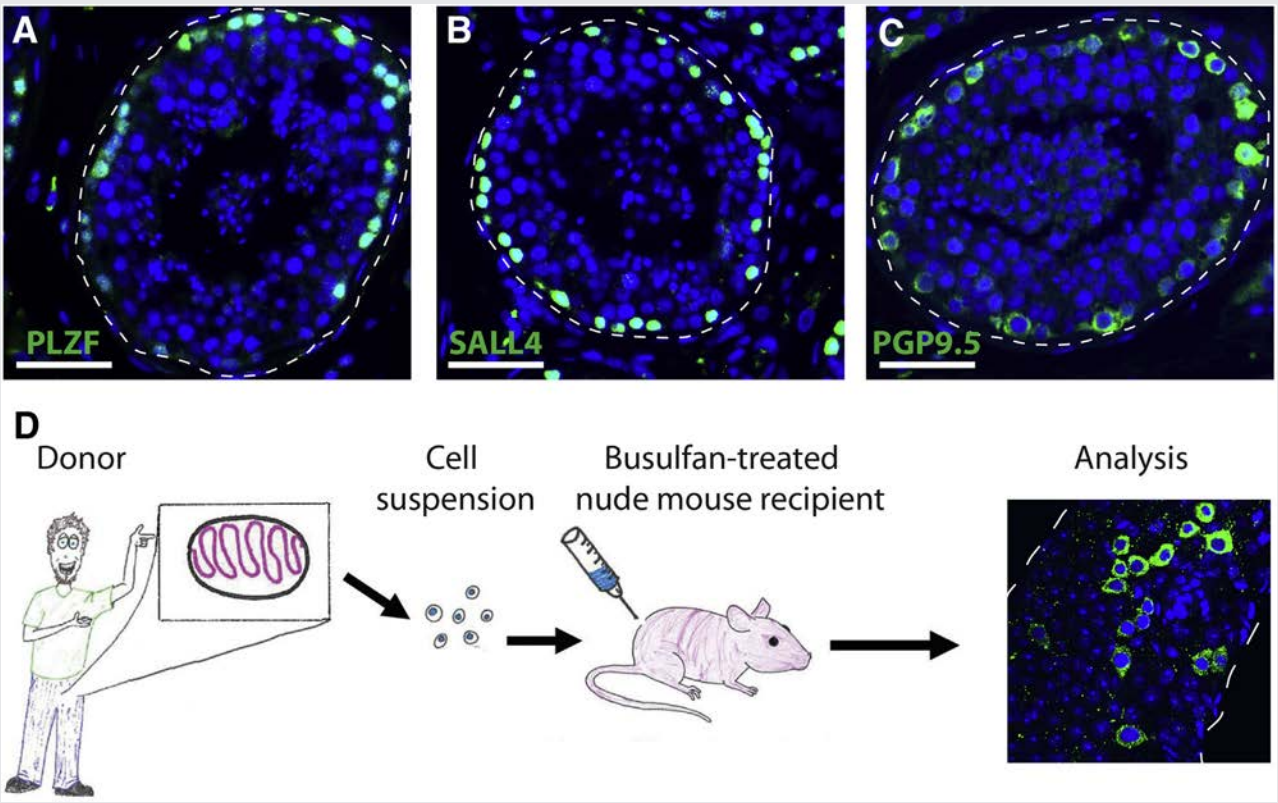
Here we propose that reliable markers of human undifferentiated spermatogonia are those with expression limited to germ cells located on the basement membrane of human seminiferous tubules and that meet one or more of the following criteria: 1) confirmed expression in Adark and/or Apale spermatogonia; 2) co-expression with other established markers of human undifferentiated spermatogonia (e.g., PLZF) or 3) that exhibit colonizing activity in the human to nude mouse xenotransplantation assay (see next paragraph and Figure 4D). Proteins that meet these criteria, based on personal experience and review of the literature, include PLZF, GFR$\alpha$1, GPR125, SALL4, LIN28, PGP9.5, UTF1, FGFR3, EXOSC10, DSG2, CBL, SSEA4, CD9, OCT2, and SSX (77, 78, 108–116). Examples of PLZF, SALL4, and PGP9.5 expression in the seminiferous tubules of adult human testes are shown in Figure 4A–4C (Valli and Orwig, unpublished data).

In rodents, SSC transplantation is the gold standard that allows investigators to quantify germ-line stem cells by observing their biological potential to produce and maintain spermatogenesis in infertile recipient animals. At present, human–to–nude mouse xenotransplantation is the best functional assay to test the SSC-like potential of a test cell population (21, 38, 78, 89, 109, 115, 117). This method does not recapitulate complete spermatogenesis from transplanted cells like mouse–to–mouse SSC transplantation, probably owing to the evolutionary distance between humans and mice. However, human–to–nude mouse xenotransplantation does assay the ability of transplanted cells to migrate to the basement membrane of seminiferous tubules, proliferate to produce characteristic colonies of spermatogonia, and persist long term as shown in Figure 4D (78, 89, 109, 115, 117).

## CONCLUSION

The assisted reproduction field is on the verge of a renaissance, which is fueled in part by exciting developments that merge new stem cell technologies with fertile outcomes. Thus, it is reasonable to expect in the next decade that the options to preserve and restore male fertility will expand from sperm freezing followed by IVF/ICSI to also include stem cell transplantation, tissue grafting, and/or culture to produce fertilization-competent gametes (Fig. 1 and Table 1). As with any rapidly developing field that has the potential to impact the clinic, it is essential to establish strict criteria to monitor progress and avoid sensationalism. Responsible technology

**FIGURE 4**




Experimental techniques to assay human spermatogonia. (A, B, and C) Expression of spermatogonia markers PLZF (A), SALL4 (B), and PGP9.5 (C) is limited to germ cells located on the basement membrane of human seminiferous tubules. Thus, they are reliable markers to screen test cell populations for human spermatogonia. DAPI (*blue*) stains all cell nuclei. Scale bar = 50 $\mu$m. (D) Human–to–nude mouse xenotransplantation assay. Human testicular tissue is made into a cell suspension and then transplanted into the testis of busulfan-treated infertile nude mice. Two months after the transplantation, the testes are recovered, the tunica is removed, and the seminiferous tubules are gently dispersed to make a whole mount. The tubules are then stained with antiprimate antibody (122) to recognize the colonies of human spermatogonia (*green*).

*Valli. Stem cell therapies for male factor infertility. Fertil Steril 2014.*

development should ideally include [1] studies in lower mammals (e.g., rodents and other domestic species) where it is possible to examine functional readouts such as regeneration of spermatogenesis, fertilization, embryo development, and generation of healthy offspring. [2] As technologies mature, preclinical studies in nonhuman primates that have anatomy and reproductive physiology similar to humans will demonstrate conservation of biological concepts and facilitate optimization of technical/surgical approaches (e.g., biopsy, cell processing, freezing, and transplantation). Nonhuman primate studies are also amenable to functional assessments, although it is recognized that such studies will be limited to institutions with space, knowledge, and technical expertise for research in primates and specifically primate assisted reproductive technology. [3] Direct investigations of human cells or tissues are particularly valuable on the road to clinical translation but are challenged by the limited availability of human samples, study-to-study variation in sample quantity and quality, and the limited functional assays to assess experimental outcomes. Despite these challenges, human tissue studies to optimize tissue and cell-processing procedures, cryopreservation methods, and cell/tissue culture conditions will be the most relevant to the clinics that are already preserving testicular tissue for patients. Although it is not popular in the current era, descriptive studies of human germ lineage development in situ are essential to guide experimental design and interpretation of results of human stem cell studies.

Clinics worldwide are preserving testicular tissue for patients who do not have sperm in anticipation that new male fertility technologies will be available in the future. Each technology described in this review has merits and limitations that are detailed in the sections above. Owing to uncertainty about which methods will ultimately be translated to the clinic, it seems reasonable to preserve intact testicular tissue fragments in a way that maximizes viability after freezing and thawing (an active focus of research in many laboratories) and will be amenable to tissue- or cell-based approaches in the future. Variability in patient circumstances and reproductive goals dictate the parallel development of multiple technologies for preserving and restoring male fertility.

## REFERENCES

1. Meistrich ML. Male gonadal toxicity. Pediatr Blood Cancer 2009;53:261–6.
2. Howlader N, Noone AM, Krapcho M, Neyman N, Aminou R, Waldron W, et al. SEER Cancer Stat Rev 1975–2008. Bethesda, MD: National Cancer Institute; 2010.

3. Schover LR. Patient attitudes toward fertility preservation. Pediatr Blood Cancer 2009;53:281–4.
4. Lee SJ, Schover LR, Partridge AH, Patrizio P, Wallace WH, Hagerty K, et al. American Society of Clinical Oncology recommendations on fertility preservation in cancer patients. J Clin Oncol 2006;24:2917–31.
5. Fertility preservation and reproduction in cancer patients. Fertil Steril 2005; 83:1622–8.
6. Meistrich ML, Vassilopoulou-Sellin R, Lipshultz LI. Adverse effects of treatment: gonadal dysfunction. In: DeVita VT, Hellman S, Rosenberg SA, editors. Principles and practice of oncology. 7th ed. Philadelphia: Lippincott Williams & Wilkins; 2004:2560–74.
7. Schover LR, Brey K, Lichtin A, Lipshultz LI, Jeha S. Knowledge and experience regarding cancer, infertility, and sperm banking in younger male survivors. J Clin Oncol 2002;20:1880–9.
8. Hsiao W, Stahl PJ, Osterberg EC, Nejat E, Palermo GD, Rosenwaks Z, et al. Successful treatment of postchemotherapy azoospermia with microsurgical testicular sperm extraction: the Weill Cornell Experience. J Clin Oncol 2011;29:1607–11.
9. Green DM, Kawashima T, Stovall M, Leisenring W, Sklar CA, Mertens AC, et al. Fertility of male survivors of childhood cancer: a report from the Childhood Cancer Survivor Study. J Clin Oncol 2010;28:332–9.
10. Pasquini MC, Wang Z. Current use and outcome of hematopoietic stem cell transplantation: CIBMTR Summary Slides. http://www.cibmtr.org; 2012.
11. Wallace WH, Anderson RA, Irvine DS. Fertility preservation for young patients with cancer: who is at risk and what can be offered? Lancet Oncol 2005;6:209–18.
12. Mitchell RT, Saunders PT, Sharpe RM, Kelnar CJ, Wallace WH. Male fertility and strategies for fertility preservation following childhood cancer treatment. Endocr Dev 2009;15:101–34.
13. Wyns C, Curaba M, Petit S, Vanabelle B, Laurent P, Wese JF, et al. Management of fertility preservation in prepubertal patients: 5 years' experience at the Catholic University of Louvain. Humanit Rep 2011;26:737–47.
14. Ginsberg JP. New advances in fertility preservation for pediatric cancer patients. Curr Opin Pediatr 2011;23:9–13.
15. Meistrich ML, Shetty G. Hormonal suppression for fertility preservation in males and females. Reproduction 2008;136:691–701.
16. Ginsberg JP, Carlson CA, Lin K, Hobbie WL, Wigo E, Wu X, et al. An experimental protocol for fertility preservation in prepubertal boys recently diagnosed with cancer: a report of acceptability and safety. Hum Reprod 2010; 25:37–41.
17. Ginsberg JP. Educational paper: the effect of cancer therapy on fertility, the assessment of fertility and fertility preservation options for pediatric patients. Eur J Pediatr 2011;170:703–8.
18. Orwig KE, Schlatt S. Cryopreservation and transplantation of spermatogonia and testicular tissue for preservation of male fertility. J Natl Cancer Inst Monographs 2005;34:51–6.
19. Levine J, Canada A, Stern CJ. Fertility preservation in adolescents and young adults with cancer. J Clin Oncol 2010;28:4831–41.
20. Schlatt S, Ehmcke J, Jahnukainen K. Testicular stem cells for fertility preservation: preclinical studies on male germ cell transplantation and testicular grafting. Pediatr Blood Cancer 2009;53:274–80.
21. Sadri-Ardekani H, Mizrak SC, van Daalen SK, Korver CM, Roepers-Gajadien HL, Koruji M, et al. Propagation of human spermatogonial stem cells in vitro. J Am Med Assoc 2009;302:2127–34.
22. Park TS, Galic Z, Conway AE, Lindgren A, van Handel BJ, Magnusson M, et al. Derivation of primordial germ cells from human embryonic and induced pluripotent stem cells is significantly improved by coculture with human fetal gonadal cells. Stem Cells 2009;27:783–95.
23. Easley CA, Phillips BT, McGuire MM, Barringer JM, Valli H, Hermann BP, et al. Direct differentiation of human pluripotent stem cells into haploid spermatogenic cells. Cell Rep 2012;2:440–6.
24. Kee K, Gonsalves JM, Clark AT, Pera RA. Bone morphogenetic proteins induce germ cell differentiation from human embryonic stem cells. Stem Cells Dev 2006;15:831–7.
25. Kee K, Angeles VT, Flores M, Nguyen HN, Reijo Pera RA. Human DAZL, DAZ and BOULE genes modulate primordial germ-cell and haploid gamete formation. Nature 2009;462:222–5.
26. Panula S, Medrano JV, Kee K, Bergstrom R, Nguyen HN, Byers B, et al. Human germ cell differentiation from fetal- and adult-derived induced pluripotent stem cells. Hum Mol Genet 2011;20:752–62.
27. Sato T, Katagiri K, Gohbara A, Inoue K, Ogonuki N, Ogura A, et al. In vitro production of functional sperm in cultured neonatal mouse testes. Nature 2011;471:504–7.
28. Hayashi K, Ohta H, Kurimoto K, Aramaki S, Saitou M. Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. Cell 2011;146:519–32.
29. Hermann BP, Sukhwani M, Winkler F, Pascarella JN, Peters KA, Sheng Y, et al. Spermatogonial stem cell transplantation into rhesus testes regenerates spermatogenesis producing functional sperm. Cell Stem Cell 2012;11:715–26.
30. Brinster RL, Zimmermann JW. Spermatogenesis following male germ-cell transplantation. Proc Natl Acad Sci U S A 1994;91:11298–302.
31. Brinster RL, Avarbock MR. Germline transmission of donor haplotype following spermatogonial transplantation. Proc Natl Acad Sci U S A 1994;91:11303–7.
32. Teramura T, Takehara T, Kawata N, Fujinami N, Mitani T, Takenoshita M, et al. Primate embryonic stem cells proceed to early gametogenesis in vitro. Cloning Stem Cells 2007;9:144–56.
33. Yamauchi K, Hasegawa K, Chuma S, Nakatsuji N, Suemori H. In vitro germ cell differentiation from cynomolgus monkey embryonic stem cells. PLoS ONE 2009;4:e5338.
34. Clark AT, Phillips BT, Orwig KE. Fruitful progress to fertility: male fertility in the test tube. Nat Med 2011;17:1564–5.
35. Holoch P, Wald M. Current options for preservation of fertility in the male. Fertil Steril 2011;96:286–90.
36. Wallace WH. Oncofertility and preservation of reproductive capacity in children and young adults. Cancer 2011;117:2301–10.
37. Lamar CA, DeCherney AH. Fertility preservation: state of the science and future research directions. Fertil Steril 2009;91:316–9.
38. Sadri-Ardekani H, Akhondi MA, van der Veen F, Repping S, van Pelt AM. In vitro propagation of human prepubertal spermatogonial stem cells. JAMA 2011;305:2416–8.
39. Keros V, Hultenby K, Borgstrom B, Fridstrom M, Jahnukainen K, Hovatta O. Methods of cryopreservation of testicular tissue with viable spermatogonia in pre-pubertal boys undergoing gonadotoxic cancer treatment. Hum Reprod 2007;22:1384–95.
40. Orwig KE, Shaw PH, Sanfilippo JS. Fertility preservation in Pittsburgh. http://www.mwrif.org/220.
41. Goossens E, Van Saen D, Tournaye H. Spermatogonial stem cell preservation and transplantation: from research to clinic. Hum Reprod 2013;28: 897–907.
42. Thayer KA, Ruhlen RL, Howdeshell KL, Buchanan DL, Cooke PS, Preziosi D, et al. Altered prostate growth and daily sperm production in male mice exposed prenatally to subclinical doses of 17a-ethinyl oestradiol. Hum Reprod 2001;16:988–96.
43. Sharpe RM. Regulation of spermatogenesis. In: Knobil E, Neill JD, editors. The physiology of reproduction. New York: Raven Press, Ltd.; 1994:1363–434.
44. Hermann BP, Sukhwani M, Hansel MC, Orwig KE. Spermatogonial stem cells in higher primates: are there differences from those in rodents? Reproduction 2010;139:479–93.
45. Meistrich M, Wilson G, Shetty G, Shuttlesworth G. Restoration of spermatogenesis after exposure to toxicants: genetic implications. In: Robaire B, Hales B, editors. Advances in male mediated developmental toxicity. Springer US; 2003:227–37.
46. Meistrich ML. Effects of chemotherapy and radiotherapy on spermatogenesis. Eur Urol 1993;23:136–41. discussion 42.
47. van den Berg H, Furstner F, van den Bos C, Behrendt H. Decreasing the number of MOPP courses reduces gonadal damage in survivors of childhood hodgkin disease. Pediatr Blood Cancer 2004;42:210–5.
48. Ogawa T, Dobrinski I, Avarbock MR, Brinster RL. Transplantation of male germ line stem cells restores fertility in infertile mice. Nat Med 2000;6: 29–34.
49. Shinohara T, Orwig KE, Avarbock MR, Brinster RL. Remodeling of the postnatal mouse testis is accompanied by dramatic changes in stem cell number and niche accessibility. Proc Natl Acad Sci U S A 2001;98:6186–91.

50. Nagano M, Brinster CJ, Orwig KE, Ryu BY, Avarbock MR, Brinster RL. Transgenic mice produced by retroviral transduction of male germ-line stem cells. Proc Natl Acad Sci U S A 2001;98:13090–5.
51. Brinster CJ, Ryu BY, Avarbock MR, Karagenc L, Brinster RL, Orwig KE. Restoration of fertility by germ cell transplantation requires effective recipient preparation. Biol Reprod 2003;69:412–20.
52. Honaramooz A, Behboodi E, Megee SO, Overton SA, Galantino-Homer H, Echelard Y, et al. Fertility and germline transmission of donor haplotype following germ cell transplantation in immunocompetent goats. Biol Reprod 2003;69:1260–4.
53. Mikkola M, Sironen A, Kopp C, Taponen J, Sukura A, Vilkki J, et al. Transplantation of normal boar testicular cells resulted in complete focal spermatogenesis in a boar affected by the immotile short-tail sperm defect. Reprod Domest Anim 2006;41:124–8.
54. Kim Y, Turner D, Nelson J, Dobrinski I, McEntee M, Travis AJ. Production of donor-derived sperm after spermatogonial stem cell transplantation in the dog. Reproduction 2008;136:823–31.
55. Herrid M, Olejnik J, Jackson M, Suchowerska N, Stockwell S, Davey R, et al. Irradiation enhances the efficiency of testicular germ cell transplantation in sheep. Biol Reprod 2009;81:898–905.
56. Izadyar F, Den Ouden K, Stout TA, Stout J, Coret J, Lankveld DP, et al. Autologous and homologous transplantation of bovine spermatogonial stem cells. Reproduction 2003;126:765–74.
57. Richardson TE, Chapman KM, Tenenhaus Dann C, Hammer RE, Hamra FK. Sterile testis complementation with spermatogonial lines restores fertility to DAZL-deficient rats and maximizes donor germline transmission. PLoS ONE 2009;4:e6308.
58. Kanatsu-Shinohara M, Ogonuki N, Inoue K, Miki H, Ogura A, Toyokuni S, et al. Long-term proliferation in culture and germline transmission of mouse male germline stem cells. Biol Reprod 2003;69:612–6.
59. Shetty G, Uthanmanthil RK, Zhou W, Shao SH, Weng CC, Tailor RC, et al. Hormone suppression with GnRH antagonist promotes spermatogenic recovery from transplanted spermatogonial stem cells in irradiated cynomolgus monkeys. Andrology 2013;1:886–98.
60. Schlatt S, Foppiani L, Rolf C, Weinbauer GF, Nieschlag E. Germ cell transplantation into X-irradiated monkey testes. Hum Reprod 2002;17:55–62.
61. Jahnukainen K, Ehmcke J, Quader MA, Saiful Huq M, Epperly MW, Hergenrother S, et al. Testicular recovery after irradiation differs in prepubertal and pubertal non-human primates, and can be enhanced by autologous germ cell transplantation. Hum Reprod 2011;26:1945–54.
62. Ryu BY, Orwig KE, Avarbock MR, Brinster RL. Stem cell and niche development in the postnatal rat testis. Dev Biol 2003;263:253–63.
63. Dobrinski I, Avarbock MR, Brinster RL. Transplantation of germ cells from rabbits and dogs into mouse testes. Biol Reprod 1999;61:1331–9.
64. Dobrinski I, Avarbock MR, Brinster RL. Germ cell transplantation from large domestic animals into mouse testes. Mol Reprod Dev 2000;57:270–9.
65. Radford J, Shalet S, Lieberman B. Fertility after treatment for cancer. Br Med J 1999;319:935–6.
66. Brook PF, Radford JA, Shalet SM, Joyce AD, Gosden RG. Isolation of germ cells from human testicular tissue for low temperature storage and autotransplantation. Fertil Steril 2001;75:269–74.
67. Radford J. Restoration of fertility after treatment for cancer. Horm Res 2003;59(Suppl 1):21–3.
68. Ogawa T, Arechaga JM, Avarbock MR, Brinster RL. Transplantation of testis germinal cells into mouse seminiferous tubules. Int J Dev Biol 1997;41:111–22.
69. Schlatt S, Rosiepen G, Weinbauer GF, Rolf C, Brook PF, Nieschlag E. Germ cell transfer into rat, bovine, monkey and human testes. Hum Reprod 1999;14:144–50.
70. Kubota H, Avarbock MR, Brinster RL. Growth factors essential for self-renewal and expansion of mouse spermatogonial stem cells. Proc Natl Acad Sci U S A 2004;101:16489–94.
71. Ryu BY, Kubota H, Avarbock MR, Brinster RL. Conservation of spermatogonial stem cell self-renewal signaling between mouse and rat. Proc Natl Acad Sci U S A 2005;102:14302–7.
72. Izadyar F, Den Ouden K, Creemers LB, Posthuma G, Parvinen M, De Rooij DG. Proliferation and differentiation of bovine type A spermatogonia during long-term culture. Biol Reprod 2003;68:272–81.
73. Kala S, Kaushik R, Singh KP, Kadam PH, Singh MK, Manik RS, et al. In vitro culture and morphological characterization of prepubertal buffalo (Bubalus bubalis) putative spermatogonial stem cell. J Assist Reprod Genet 2012;29:1335–42.
74. Luo J, Megee S, Rathi R, Dobrinski I. Protein gene product 9.5 is a spermatogonia-specific marker in the pig testis: application to enrichment and culture of porcine spermatogonia. Mol Reprod Dev 2006;73:1531–40.
75. Kuijk EW, Colenbrander B, Roelen BA. The effects of growth factors on in vitro-cultured porcine testicular cells. Reproduction 2009;138:721–31.
76. Eildermann K, Gromoll J, Behr R. Misleading and reliable markers to differentiate between primate testis-derived multipotent stromal cells and spermatogonia in culture. Hum Reprod 2012;27:1754–67.
77. He Z, Kokkinaki M, Jiang J, Dobrinski I, Dym M. Isolation, characterization, and culture of human spermatogonia. Biol Reprod 2010;82:363–72.
78. Wu X, Schmidt JA, Avarbock MR, Tobias JW, Carlson CA, Kolon TF, et al. Prepubertal human spermatogonia and mouse gonocytes share conserved gene expression of germline stem cell regulatory molecules. Proc Natl Acad Sci U S A 2009;106:21672–7.
79. Mirzapour T, Movahedin M, Tengku Ibrahim TA, Koruji M, Haron AW, Nowroozi MR, et al. Effects of basic fibroblast growth factor and leukaemia inhibitory factor on proliferation and short-term culture of human spermatogonial stem cells. Andrologia 2012;44:41–55.
80. Chen B, Wang YB, Zhang ZL, Xia WL, Wang HX, Xiang ZQ, et al. Xeno-free culture of human spermatogonial stem cells supported by human embryonic stem cell-derived fibroblast-like cells. Asian J Androl 2009;11:557–65.
81. Kokkinaki M, Djourabtchi A, Golestaneh N. Long-term culture of human SSEA-4 positive spermatogonial stem cells (SSCs). J Stem Cell Res Ther 2011:S2. 003.
82. Liu S, Tang Z, Xiong T, Tang W. Isolation and characterization of human spermatogonial stem cells. Reprod Biol Endocrinol 2011;9:141.
83. Abu Elhija M, Lunenfeld E, Schlatt S, Huleihel M. Differentiation of murine male germ cells to spermatozoa in a soft agar culture system. Asian J Androl 2011;14:285–93.
84. Kim TH, Hargreaves HK, Brynes RK, Hawkins HK, Lui VK, Woodard J, et al. Pretreatment testicular biopsy in childhood acute lymphocytic leukaemia. Lancet 1981;2:657–8.
85. Jahnukainen K, Hou M, Petersen C, Setchell B, Söder O. Intratesticular transplantation of testicular cells from leukemic rats causes transmission of leukemia. Cancer Res 2001;61:706–10.
86. Fujita K, Ohta H, Tsujimura A, Takao T, Miyagawa Y, Takada S, et al. Transplantation of spermatogonial stem cells isolated from leukemic mice restores fertility without inducing leukemia. J Clin Invest 2005;115:1855–61.
87. Fujita K, Tsujimura A, Miyagawa Y, Kiuchi H, Matsuoka Y, Takao T, et al. Isolation of germ cells from leukemia and lymphoma cells in a human in vitro model: potential clinical application for restoring human fertility after anticancer therapy. Cancer Res 2006;66:11166–71.
88. Hou M, Andersson M, Zheng C, Sundblad A, Soder O, Jahnukainen K. Decontamination of leukemic cells and enrichment of germ cells from testicular samples from rats with Roser's T-cell leukemia by flow cytometric sorting. Reproduction 2007;134:767–79.
89. Dovey SL, Valli H, Hermann BP, Sukhwani M, Donohue J, Castro CA, et al. Eliminating malignant contamination from therapeutic human spermatogonial stem cells. J Clin Invest 2013;123:1833–43.
90. Hermann BP, Sukhwani M, Salati J, Sheng Y, Chu T, Orwig KE. Separating spermatogonia from cancer cells in contaminated prepubertal primate testis cell suspensions. Hum Reprod 2011;26:3222–31.
91. Geens M, Van de Velde H, De Block G, Goossens E, Van Steirteghem A, Tournaye H. The efficiency of magnetic-activated cell sorting and fluorescence-activated cell sorting in the decontamination of testicular cell suspensions in cancer patients. Hum Reprod 2007;22:733–42.
92. Honaramooz A, Snedaker A, Boiani M, Scholer H, Dobrinski I, Schlatt S. Sperm from neonatal mammalian testes grafted in mice. Nature 2002;418:778–81.

93. Schlatt S, Honaramooz A, Boiani M, Scholer HR, Dobrinski I. Progeny from sperm obtained after ectopic grafting of neonatal mouse testes. Biol Reprod 2003;68:2331–5.
94. Honaramooz A, Li MW, Penedo MCT, Meyers S, Dobrinski I. Accelerated maturation of primate testis by xenografting into mice. Biol Reprod 2004;70:1500–3.
95. Schlatt S, Kim SS, Gosden R. Spermatogenesis and steroidogenesis in mouse, hamster and monkey testicular tissue after cryopreservation and heterotopic grafting to castrated hosts. Reproduction 2002;124:339–46.
96. Jahnukainen K, Ehmcke J, Hergenrother SD, Schlatt S. Effect of cold storage and cryopreservation of immature non-human primate testicular tissue on spermatogonial stem cell potential in xenografts. Hum Reprod 2007;22:1060–7.
97. Wistuba J, Luetjens CM, Wesselmann R, Nieschlag E, Simoni M, Schlatt S. Meiosis in autologous ectopic transplants of immature testicular tissue grafted to Callithrix jacchus. Biol Reprod 2006;74:706–13.
98. Luetjens CM, Stukenborg J-B, Nieschlag E, Simoni M, Wistuba J. Complete spermatogenesis in orthotopic but not in ectopic transplants of autologously grafted marmoset testicular tissue. Endocrinology 2008;149:1736–47.
99. Geens M, De Block G, Goossens E, Frederickx V, Van Steirteghem A, Tournaye H. Spermatogonial survival after grafting human testicular tissue to immunodeficient mice. Hum Reprod 2006;21:390–6.
100. Goossens E, Geens M, De Block G, Tournaye H. Spermatogonial survival in long-term human prepubertal xenografts. Fertil Steril 2008;90:2019–22.
101. Van Saen D, Goossens E, Bourgain C, Ferster A, Tournaye H. Meiotic activity in orthotopic xenografts derived from human postpubertal testicular tissue. Hum Reprod 2011;26:282–93.
102. Schlatt S, Honaramooz A, Ehmcke J, Goebell PJ, Rubben H, Dhir R, et al. Limited survival of adult human testicular tissue as ectopic xenograft. Hum Reprod 2006;21:384–9.
103. Sato Y, Nozawa S, Yoshiike M, Arai M, Sasaki C, Iwamoto T. Xenografting of testicular tissue from an infant human donor results in accelerated testicular maturation. Hum Reprod 2010;25:1113–22.
104. Wyns C, Van Langendonckt A, Wese FX, Donnez J, Curaba M. Long-term spermatogonial survival in cryopreserved and xenografted immature human testicular tissue. Hum Reprod 2008;23:2402–14.
105. Sato T, Katagiri K, Kubota Y, Ogawa T. In vitro sperm production from mouse spermatogonial stem cell lines using an organ culture method. Nat Protocols 2013;8:2098–104.
106. Trowell OA. The culture of mature organs in a synthetic medium. Exp Cell Res 1959;16:118–47.
107. Byrne JA. Nuclear reprogramming and the current challenges in advancing personlized pluripotent stem cell-based therapies. Gene Ther Reg 2012;07:1230002.
108. Grisanti L, Falciatori I, Grasso M, Dovere L, Fera S, Muciaccia B, et al. Identification of spermatogonial stem cell subsets by morphological analysis and prospective isolation. Stem Cells 2009;27:3043–52.
109. Zohni K, Zhang X, Tan SL, Chan P, Nagano M. CD9 is expressed on human male germ cells that have a long-term repopulation potential after transplantation into mouse testes. Biol Reprod 2012;87(2):27, 1-8.
110. Eildermann K, Aeckerle N, Debowski K, Godmann M, Christiansen H, Heistermann M, et al. Developmental expression of the pluripotency factor Sal-like protein 4 in the monkey, human and mouse testis: restriction to premeiotic germ cells. Cells Tissues Organs 2012;196:206–20.
111. Aeckerle N, Eildermann K, Drummer C, Ehmcke J, Schweyer S, Lerchl A, et al. The pluripotency factor LIN28 in monkey and human testes: a marker for spermatogonial stem cells? Mol Hum Reprod 2012;18:477–88.
112. von Kopylow K, Kirchhoff C, Jezek D, Schulze W, Feig C, Primig M, et al. Screening for biomarkers of spermatogonia within the human testis: a whole genome approach. Hum Reprod 2010;25:1104–12.
113. von Kopylow K, Staege H, Spiess AN, Schulze W, Will H, Primig M, et al. Differential marker protein expression specifies rarefaction zone-containing human Adark spermatogonia. Reproduction 2012;143:45–57.
114. von Kopylow K, Staege H, Schulze W, Will H, Kirchhoff C. Fibroblast growth factor receptor 3 is highly expressed in rarely dividing human type A spermatogonia. Histochem Cell Biol 2012;138:759–72.
115. Izadyar F, Wong J, Maki C, Pacchiarotti J, Ramos T, Howerton K, et al. Identification and characterization of repopulating spermatogonial stem cells from the adult human testis. Hum Reprod 2011;26:1296–306.
116. Lim J, Goriely A, Turner GD, Ewen KA, Jacobsen GK, Graem N, et al. OCT2, SSX and SAGE1 reveal the phenotypic heterogeneity of spermatocytic seminoma reflecting distinct subpopulations of spermatogonia. J Pathol 2011;224:473–83.
117. Nagano M, Patrizio P, Brinster RL. Long-term survival of human spermatogonial stem cells in mouse testes. Fertil Steril 2002;78:1225–33.
118. Clouthier DE, Avarbock MR, Maika SD, Hammer RE, Brinster RL. Rat spermatogenesis in mouse testis. Nature 1996;381:418–21.
119. Ogawa T, Dobrinski I, Avarbock MR, Brinster RL. Xenogeneic spermatogenesis following transplantation of hamster germ cells to mouse testes. Biol Reprod 1999;60:515–21.
120. Shinohara T, Kato M, Takehashi M, Lee J, Chuma S, Nakatsuji N, et al. Rats produced by interspecies spermatogonial transplantation in mice and in vitro microinsemination. Proc Natl Acad Sci U S A 2006;103:13624–8.
121. Kubota H, Avarbock MR, Brinster RL. Culture conditions and single growth factors affect fate determination of mouse spermatogonial stem cells. Biol Reprod 2004;71:722–31.
122. Hermann BP, Sukhwani M, Lin CC, Sheng Y, Tomko J, Rodriguez M, et al. Characterization, cryopreservation, and ablation of spermatogonial stem cells in adult rhesus macaques. Stem Cells 2007;25:2330–8.
123. Hermann BP, Sukhwani M, Simorangkir DR, Chu T, Plant TM, Orwig KE. Molecular dissection of the male germ cell lineage identifies putative spermatogonial stem cells in rhesus macaques. Hum Reprod 2009;24:1704–16.
124. Nagano M, McCarrey JR, Brinster RL. Primate spermatogonial stem cells colonize mouse testes. Biol Reprod 2001;64:1409–16.
125. Maki CB, Pacchiarotti J, Ramos T, Pascual M, Pham J, Kinjo J, et al. Phenotypic and molecular characterization of spermatogonial stem cells in adult primate testes. Hum Reprod 2009;24:1480–91.
126. He Z, Kokkinaki M, Jiang J, Zeng W, Dobrinski I, Dym M. Isolation of human male germ-line stem cells using enzymatic digestion and magnetic-activated cell sorting. Methods Mol Biol 2012;825:45–57.

CellPress

Cell Stem Cell
Resource

# The Dynamic Transcriptional Cell Atlas of Testis Development during Human Puberty

Jingtao Guo,[1,2] Xichen Nie,[1] Maria Giebler,[3] Hana Mlcochova,[3] Yueqi Wang,[4] Edward J. Grow,[1] DonorConnect,[5] Robin Kim,[6] Melissa Tharmalingam,[7,8] Gabriele Matilionyte,[7,8] Cecilia Lindskog,[9] Douglas T. Carrell,[2] Rod T. Mitchell,[7,8] Anne Goriely,[3,10] James M. Hotaling,[2,10] and Bradley R. Cairns[1,10,11,*]

[1]Howard Hughes Medical Institute, Department of Oncological Sciences and Huntsman Cancer Institute, University of Utah School of Medicine, Salt Lake City, UT 84112, USA
[2]The Andrology Laboratory, Department of Surgery (Andrology/Urology), Center for Reconstructive Urology and Men's Health, University of Utah Health Sciences Center, Salt Lake City, UT 84112, USA
[3]Radcliffe Department of Medicine, MRC Weatherall Institute of Molecular Medicine, University of Oxford, Oxford OX39DS, UK
[4]Department of Computer Science, Columbia University, New York, NY 10027, USA
[5]DonorConnect, Murray, UT 84107, USA
[6]Section of Transplantation, Department of Surgery, University of Utah School of Medicine, Salt Lake City, UT 84132, USA
[7]MRC Centre for Reproductive Health, The Queen's Medical Research Institute, The University of Edinburgh, Edinburgh EH16 4TJ, UK
[8]Royal Hospital for Children and Young People, Edinburgh EH91LF, UK
[9]Department of Immunology, Genetics and Pathology, Science for Life Laboratory, Uppsala University, Uppsala 751 85, Sweden
[10]Senior author
[11]Lead Contact
*Correspondence: brad.cairns@hci.utah.edu
https://doi.org/10.1016/j.stem.2019.12.005

## SUMMARY

**The human testis undergoes dramatic developmental and structural changes during puberty, including proliferation and maturation of somatic niche cells, and the onset of spermatogenesis. To characterize this understudied process, we profiled and analyzed single-cell transcriptomes of ~10,000 testicular cells from four boys spanning puberty and compared them to those of infants and adults. During puberty, undifferentiated spermatogonia sequentially expand and differentiate prior to the initiation of gametogenesis. Notably, we identify a common pre-pubertal progenitor for Leydig and myoid cells and delineate candidate factors controlling pubertal differentiation. Furthermore, pre-pubertal Sertoli cells exhibit two distinct transcriptional states differing in metabolic profiles before converging to an alternative single mature population during puberty. Roles for testosterone in Sertoli cell maturation, antimicrobial peptide secretion, and spermatogonial differentiation are further highlighted through single-cell analysis of testosterone-suppressed transfemale testes. Taken together, our transcriptional atlas of the developing human testis provides multiple insights into developmental changes and key factors accompanying male puberty.**

## INTRODUCTION

Human male puberty involves major changes in testis physiology, a large increase in testicular volume, and complex hormonal and molecular modulation, to accomplish both somatic cell proliferation/maturation and the initiation of spermatogenesis (Koskenniemi et al., 2017; Sharpe et al., 2003). Puberty is initiated by re-activation of the hypothalamo-pituitary-gonadal axis following a period of relative quiescence during childhood (Plant, 2015). This hormonal control requires hypothalamic gonadotrophin-releasing hormone (GnRH) stimulating the release of gonadotrophins, luteinizing hormone (LH), and follicle stimulating hormone (FSH). LH is responsible for stimulating Leydig cells to produce testosterone (T), while FSH stimulates Sertoli cells to support spermatogenesis (Herbison, 2016; Plant, 2015).

To date, our understanding of puberty derives mostly from physiological approaches in humans and primates complemented by extensive molecular and genetic approaches in rodents, which includes the use of sophisticated lineage-tracing and transgenic technologies. However, fundamental differences exist between human and rodent with respect to hormonal control of puberty and onset of spermatogenesis (Tharmalingam et al., 2018). In mice, type B spermatogonia begin to develop at day 8 (d8) resulting in a synchronous "first wave" of spermatogenesis (Bellvé et al., 1977), during which pro-spermatogonia undergo differentiation and synchronously generate both self-renewing and differentiating spermatogonia (Kluin and de Rooij, 1981; Yoshida et al., 2006). This results in the first round of murine sperm production at around d35 (Vergouwen et al., 1993). Subsequently, self-renewing spermatogonia, residing in the niche, commence differentiation and produce successive and continuous waves of gametogenesis from 2–3 weeks after birth (Kanatsu-Shinohara and Shinohara, 2013). However, humans lack the equivalent of this first wave of spermatogenesis and instead are believed to maintain spermatogonia in an undifferentiated (though largely uncharacterized) state prior to the initiation of puberty (Paniagua and Nistal, 1984).

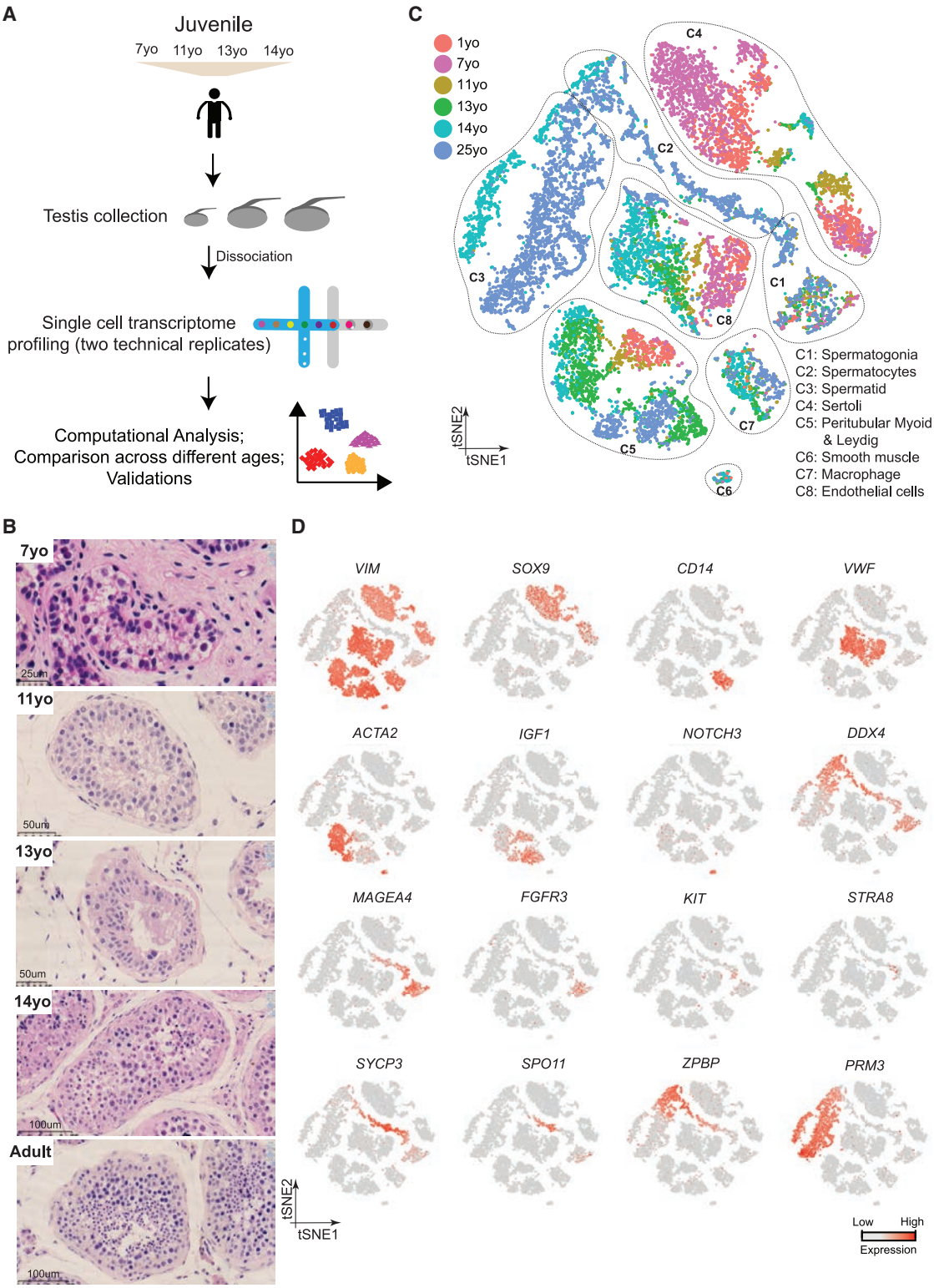



**Figure 1. Single-Cell Transcriptome Profiling of the Human Juvenile Testis**

(A) Schematic illustration of the experimental workflow.

(B) H&E (hematoxylin and eosin) staining of the different juvenile testes analyzed in this study illustrates the typical macroscopic testicular changes in morphology observed during puberty.

*(legend continued on next page)*

In humans, puberty typically begins between 10 and 13 years after birth. However, low levels of incomplete spermatogenesis is observed in portions of the testis in juveniles prior to puberty, though mature sperm are usually not present prior to age 10–11, and meiotic cells typically undergo apoptosis during peri-pubertal phases (Chemes, 2001) (note: we define "juvenile" as the developmental period after the infant stage but prior to full adult sexual maturity [typically ages from 1–14]). Morphologically, the pre-pubertal testis resembles that of the early postnatal mouse, with undifferentiated spermatogonia embedded in small "cords" of Sertoli cells, that lack a defined tubular lamina or lumen. During puberty, formation of the basal lamina and lumen creates structured seminiferous tubules (Paniagua and Nistal, 1984), which involves the coordinated development of myoid cells (contributing to the formation of the basal lamina), Leydig cells (responsible for hormonal signaling), and Sertoli cells with multiple functions within the tubule—and which together define the main niche components. However, our understanding of this crucial developmental period is limited, and much remains to be characterized at the molecular and genome-wide level, including the mechanisms that maintain human spermatogonial stem cells (SSCs) in a slowly self-renewing and undifferentiated state during childhood and the processes responsible for initiation and progression to spermatogenesis at puberty.

Single-cell RNA sequencing (scRNA-seq) approaches provide unique and powerful tools to examine human testis development, enabling the simultaneous examination of thousands of individual cells, faithfully representing the constituents of entire organs, without the need for prior sorting or enrichment procedures (Birnbaum, 2018; Guo et al., 2018; Hermann et al., 2018; Sohni et al., 2019; Wang et al., 2018). The lack of spatial information obtained from this approach can be alleviated by using orthogonal validation approaches such as co-localization of specific proteins in fixed tissues (Guo et al., 2018). Here, we profiled ~10,000 single-cell transcriptomes from whole-testes of four juvenile males (7, 11, 13, and 14 years old) and compared these data to our previously published infant (1 year old) and adult (25 years old) datasets (Guo et al., 2018). Our study uncovers dramatic changes in germline stem cells and heretofore unappreciated, complex modulations of the somatic niche revealing candidate factors and pathways that regulate somatic cell development during puberty. Certain changes in the niche, such as T production, are accompanied by changes in germ cell states—indicating communication between the germline and niche. To explore the role of T on germline and niche development and function, we analyzed the transcriptome of testes from T-suppressed adult transfemales (that retained spermatogonia), which revealed roles for T in promoting and/or supporting testis development during puberty. Taken together, our datasets and comparative analyses provide multiple insights into the development of SSCs and their niche during puberty in humans, while also providing a valuable data resource (https://humantestisatlas.shinyapps.io/humantestisatlas1/) for the research community.

## RESULTS

### Single-Cell Transcriptomes of Human Juvenile Testes

We collected whole-testes through a rapid autopsy program from 4 juvenile donors aged 7, 11, 13, and 14 years (Figure 1A), with Johnsen scores (Johnsen, 1970) at 3, 4, 4, and 7, respectively. Consistent with prior observations (Paniagua and Nistal, 1984), our histological examination revealed major changes in morphology and composition along this timeline (Figure 1B): in the 7-year-old sample, small testis cords lacking an apparent lamina or lumen were observed, forming structures reminiscent of those seen in fetal/early postnatal mouse testes. A tubular structure became progressively apparent in the 11-year-old sample, while a clear defined lamina and lumen were observed across the tubules of the 13- and 14-year-old samples (Figure 1B). To characterize the molecular features accompanying this developmental progression, we isolated single cells from these testicular tissues and performed scRNA-seq using the 10x Genomics platform. For each donor, two separate technical replicates were performed, resulting in eight datasets. From a total of ~10,000 cells, 7,675 passed standard quality-control (QC) dataset filters and were retained for downstream analysis (see STAR Methods for details). We obtained ~200K reads/cell, which enabled the analysis of ~1,500–2,000 genes/cell (Figure S1A). The sequencing saturation rate was ~85%, and technical replicates were highly similar ($r > 0.95$) (Figures S1A and S1B).

To describe pubertal development and to enable comparison to other time points, we combined the juvenile datasets with our previously published adult and infant datasets, yielding a total of 13,797 single cells (after QC filtering). We first performed t-distributed stochastic neighbor embedding (tSNE) analysis on the combined datasets using the Seurat package (Butler et al., 2018; Figure 1C), which yielded eight major clusters that were subsequently annotated using known gene markers (Figure 1D; Figures S1D and S1E). Clusters 1–3 are germ cells, with C1 representing spermatogonia (*UTF1*$^+$, *FGFR3*$^+$), C2 consisting of spermatocytes (*SYCP3*$^+$), and C3 identifying post-meiotic spermatids (*PRM3*$^+$). Clusters 4–8 correspond to a heterogenous mixture of Sertoli cells (C4; *SOX9*$^+$), Leydig and myoid-like cells (C5; *IGF1*$^+$ and/or *ACTA2*$^+$), smooth muscle (C6; *NOTCH3*$^+$), macrophages (C7: *CD14*$^+$), and endothelial cells (C8; *VWF*$^+$), respectively (see Figures S1D and S1E for additional markers).

### Distinct Phases of Spermatogonial Proliferation and Differentiation during Human Puberty

We examined changes in germ cell composition during puberty by analyzing the germ cells, following specific reclustering of initial clusters 1–3 (Figure 2A). Expression patterns of key

(C) tSNE and clustering analysis of single-cell transcriptome data from juvenile human testes (n = 7,680), combined with previous datasets of infant and adult scRNA-seq (Guo et al., 2018). Each dot represents a single cell and is colored according to its donor of origin. tSNE, t-distributed stochastic neighbor embedding. (D) Expression patterns of selected markers projected on the tSNE plot (Figure 1C). For each cell cluster, one cell marker is shown in the main figure accompanied by a gallery of further markers in Figures S1D and S1E. The two top rows show somatic/niche cell markers; representative germ cell markers are shown in the two bottom rows.

See also Figure S1 and Table S1.

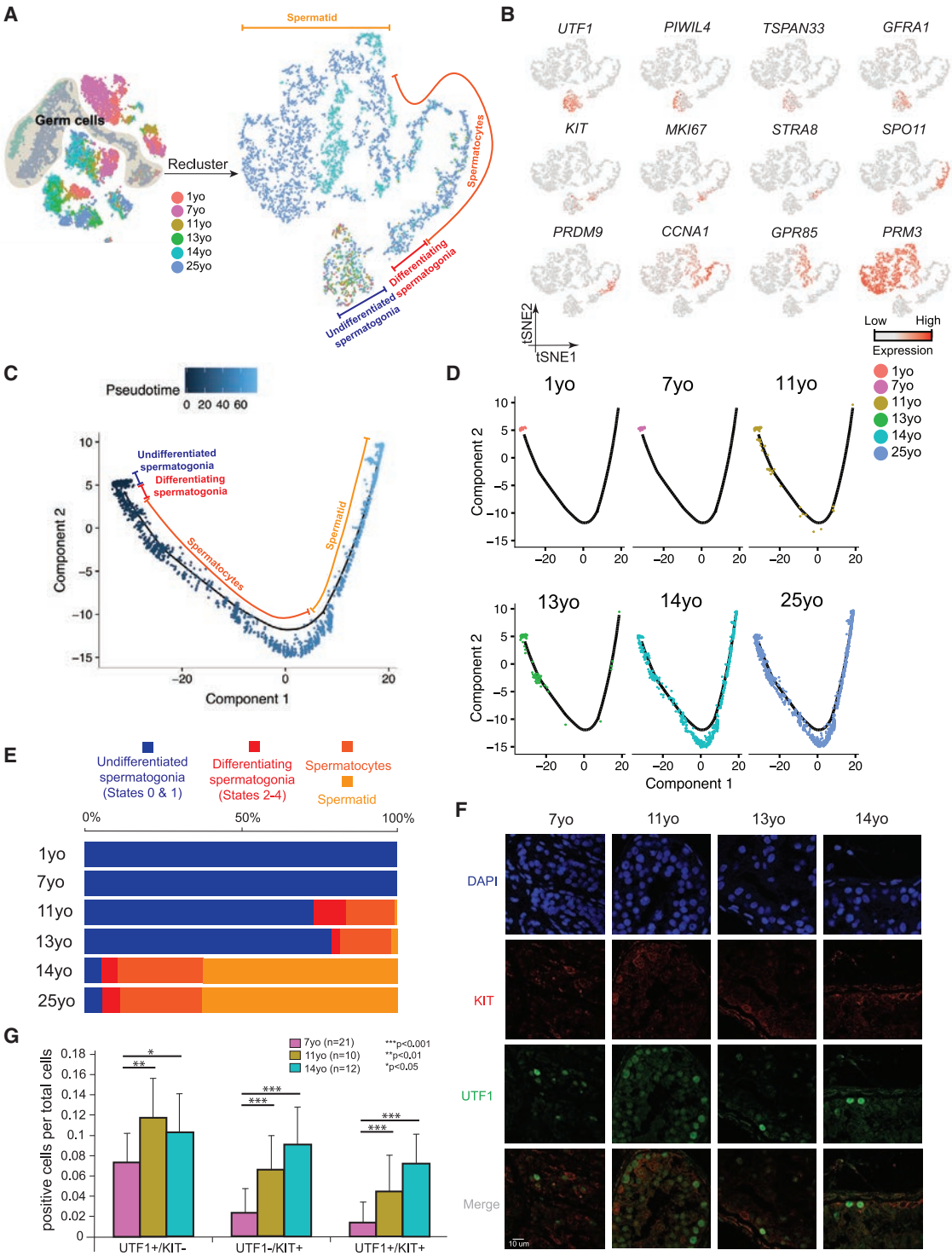



**Figure 2. Distinct Phases of Spermatogonial Proliferation and Differentiation during Human Puberty**

(A) Focused analysis (tSNE and clustering) of the germ cells (clusters C1, C2, and C3 from Figure 1C) reveals developmental progression of spermatogenesis during puberty. Cells are colored based on the ages/donors of origin.

(B) Expression patterns of known spermatogenic markers projected onto the tSNE plot from Figure 2A.

(C) Pseudotime trajectory (Monocle analysis) of the germ cells. Cells are colored based according to the predicted pseudotime.

(D) Deconvolution of the Monocle pseudotime plot according to ages/donors of origin.

(E) Relative proportion of the single cells at different spermatogenic stages in the samples analyzed.

*(legend continued on next page)*

markers and comparisons to prior work enable assignment of germ cells into four broad developmental stages: slowly self-renewing and undifferentiated spermatogonia (referred to as States 0–1 in [Guo et al., 2017, 2018]) marked by *UTF1*$^+$ and largely *MKI67*$^-$, differentiating spermatogonia (States 2–4) marked by *KIT*$^+$ and largely *MKI67*$^+$ (note: State 4 is preparing to enter meiosis and lacks *MKI67*), spermatocytes (from *STRA8*$^+$ to *GPR85*$^+$), and spermatids (*PRM3*$^+$) (Figures 2A and 2B). An orthogonal pseudotime analysis using the Monocle package (Qiu et al., 2017) further supports this pseudotime trajectory (Figures 2C and 2D; Figures S2A and S2B). Germ cells were then parsed out to examine their relative composition at different phases of puberty (Figures 2D and 2E). Interestingly, germ cells from both the 1- and 7-year-old samples consisted of only undifferentiated spermatogonia (corresponding to States 0–1) (Figures 2A, 2D, and 2E) that expressed *UTF1* and other early germline stem cell markers (Figure 2B; Figures S2A and S2B). In the 11-year-old sample, while a high proportion of cells were still States 0–1 spermatogonia, differentiating spermatogonia and meiotic cells also began to emerge (Figures 2D and 2E). Notably, in the infant and 7-year-old samples, spermatogonia were relatively rare (~3%–4% of total testicular cells), whereas in the 11-year-old and older samples, the relative proportion of spermatogonia increased considerably to represent ~10%–15% of total testicular cells (Figure S1C), consistent with a spermatogonial amplification and proliferative phase prior to a differentiation phase. The 13-year-old sample largely resembled the 11-year-old sample (likely reflecting the known age differences in puberty onset) though post-meiotic cells increased in proportion, indicating a more robust commitment to meiosis. Last, germ cell composition in the 14-year-old sample resembled the adult, indicating nearly full spermatogenesis (Figure 2E).

Next, we performed immunofluorescence (IF) to confirm our scRNA-seq findings (Figures 2F and 2G; Figures S2C and S2D). First, we observed UTF1$^+$ undifferentiated spermatogonia (State 0–1) at all ages analyzed. In contrast, proliferative and differentiating spermatogonia (States 2–3) display strong increases in multiple proliferative markers (e.g., cyclins, CDKs, *TOP2A*, MKI67, KIT [Figure S2E]) and MKI67$^+$ spermatogonia become apparent only from 11 years old onward, and significant levels of the meiotic marker SYCP3 were found only at/after 14 years old (Figure S2C). Co-staining for UTF1 and KIT confirmed the gradual and steady increase in KIT$^+$ spermatogonia during juvenile development (Figures 2F and 2G), indicative of progressive commitment to spermatogonial differentiation. Taken together, our transcriptomic and protein data combined with prior immunohistochemistry (IHC) studies (Paniagua and Nistal, 1984) point to three separate phases between infancy and puberty: (1) a phase where undifferentiated spermatogonia are quiescent or slowly self-renew, (2) a transient period of spermatogonial proliferation (mitosis) with limited and incomplete spermatogonial differentiation, occurring early in puberty, which gradually transitions to (3) the establishment of steady and balanced spermatogonial self-renewal and differentiation, alongside commitment to full gametogenesis.

## Gene Expression Dynamics during Sertoli Cell Maturation

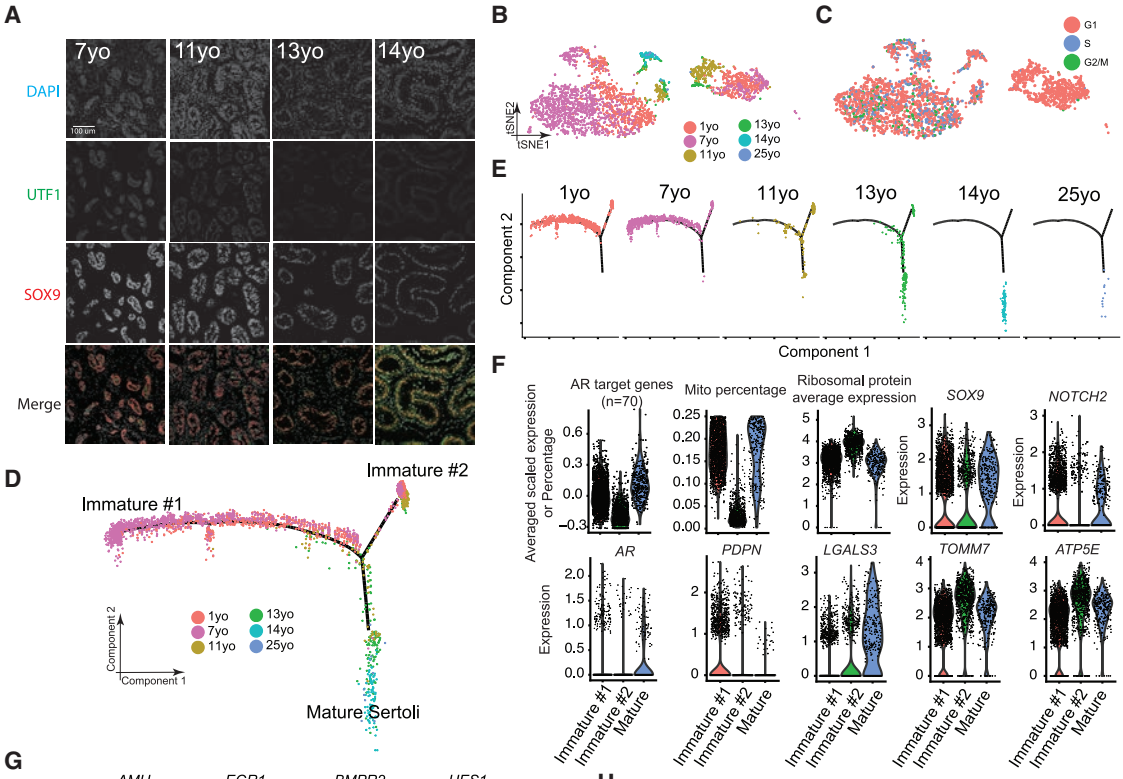
We next explored Sertoli cell development during puberty. As previewed above, in the pre-pubertal samples (7 years old), staining of UTF1 and SOX9 revealed cord-like structures consisting of intermingled spermatogonia and Sertoli cells (Figure 3A). On initiation of puberty (e.g., 11- and 13-year-old samples), Sertoli cells and spermatogonia align at the basement membrane of the tubules and a lumen begins to form—a pattern that developed further in the 14-year-old sample (Figure 3A). To characterize this morphological re-organization, we parsed out and examined the Sertoli cell populations (cluster 4 in Figure 1C). Here, we note that mature Sertoli cells are typically larger than the preferred size for capture in the 10x Genomics system, likely explaining the low proportion of Sertoli cells at older ages. Hence, although we obtained sufficient cell numbers to define Sertoli cell transcriptional signatures, this caveat prevents quantitation of changes in total Sertoli cell numbers during development. Our clustering analysis defined two large Sertoli clusters flanking two smaller clusters (Figures 3B and 3C). Notably, cell-cycle analysis (see STAR Methods) reveals that the large right cluster (Figures 3B and 3C) contains cells from all juvenile samples and displays a G1 phase-bias, while the left cluster is mainly composed of cells from the younger samples (1–7 years old) at different cell-cycle phases. Interestingly, Monocle pseudotime analysis also partitioned the 1- and 7-year-old Sertoli cells into two distinct states, termed Immature #1 and #2 (Figures 3D and 3E). Remarkably, these two states merge into one path that lead to mature Sertoli cells, which start emerging in samples from 11 years old onward (Figures 3D and 3E).

We observed multiple features of Immature #1 and mature Sertoli cells that distinguish them from Immature #2 cells, including high AR (Androgen Receptor) target gene expression (defined by a set of ~70 known AR-dependent genes) (Bolton et al., 2007), high mitochondrial transcription, and lower expression of ribosomal protein genes (Figure 3F). Gene Ontology (GO) analysis revealed enrichment for terms such as nucleus and transcription in Immature #1 (Figure S3H). IF co-staining for AR and SOX9 confirms this sequence of events, by showing gradual AR upregulation and co-localization in Sertoli cells as puberty proceeds (Figure 3H). Immature #2 cells have the converse properties of Immature #1 described above and are enriched in GO terms for translation and ribosome function (Figure S3H). They exhibit higher transcription of ribosomal protein genes, mitochondrial ATPases (e.g., *ATP5E*), and mitochondrial membrane proteins (e.g., *TOMM7*) and contain a higher fraction of cells in the G1 phase of the cell cycle (Figures 3F and S3E). Taken together, these two immature Sertoli cell states differ markedly in their metabolic/mitochondrial, translational, and cell-cycle properties (as indicated by transcriptional signatures).

(F) Protein co-immunofluorescence for two spermatogonial cell markers: UTF1 (States 0–1) and KIT (States 2–4) in the 4 analyzed samples. See Figure S2D for a wider field of view.

(G) Quantification of UTF1$^+$ and/or KIT$^+$ spermatogonia at different ages. The data shown are means ± SD of independent tubules. The p value was calculated via Student's t test.

See also Figure S2.




*(legend on next page)*

Analysis of Sertoli maturation via pseudotime and differential gene expression revealed dynamic expression of ~1,000 genes (Figure S3F) that differ between the immature and mature stages. As expected, genes associated with cytoskeleton, cell morphogenesis, and extracellular matrix were upregulated during Sertoli maturation (Griswold, 2018; Figure S3F). Genes encoding antimicrobial innate immunity peptides (including the defensin class of peptides) and immunity-related genes (e.g., *ISG15*) were also upregulated during puberty (Figures 3F, 3G, and S3G), supporting a role for Sertoli cells in protecting the testis from infection (Com et al., 2003), especially after sexual maturation. To confirm this, we performed IF/IHC staining and observed downregulation of the immature Sertoli marker AMH during puberty (Figures 3J and S3I), and upregulation of the immune peptides LGALS3 and S100A6 in Sertoli cells in the 14-year-old sample (Figures 3I and S3J). Interestingly, immunostaining of an Immature #1 marker, PDPN (Figure 3F), showed spatial separation in the 11-year-old testis (Figure S3D), suggesting that the maturation of Sertoli cells in different tubules is asynchronous and proceeds gradually. Taken together, our data describe complex developmental progression in Sertoli cell states late in puberty that involve morphogenesis, metabolism, and innate immunity.

## A Common Progenitor for Leydig and Peritubular Myoid Cells

Interestingly, the clustering analysis had initially placed Leydig and myoid cells together into a single large cluster (C5; Figure 1C), prompting further exploration. Reclustering of C5 revealed that cells from the 11-year-old and younger samples largely overlapped and shared the expression of known markers of fetal Leydig and myoid precursors in the mouse (e.g., *MAFB* and *NR4A1*; Figures 4A and 4B; DeFalco et al., 2011; Wen et al., 2016), as well as heterogeneous co-expression of markers typically associated with both mature Leydig and myoid cells (e.g., *DLK1* and *ACTA2*, respectively) (Guo et al., 2018). However, in samples from 13 years old to adult, we observed a clear separation of Leydig (marked by *DLK1* and *IGF1*) and myoid cells (marked by *ACTA2* and *MYH11*). In keeping, Monocle pseudotime analysis revealed that cells from pre-puberty (1–7 years old) tightly cluster together, and the trajectory bifurcates into two clusters as puberty proceeds (Figures 4C and 4D; Figures S4C and S4D). Thus, both modes of computational analysis suggest a common expression program defining a bipotential progenitor for Leydig and myoid cells prior to puberty.

We next identified lineage-specific genes and programs, yielding ~1,000 differentially expressed genes (Figure 4E). Early precursors expressed particular genes associated with transcription (e.g., *MAFB* and *NR4A1*); as cells progress toward the myoid lineage, cytoskeleton- and cell adhesion-related genes become expressed, consistent with myoid cells forming their distinctive peritubular structure (Maekawa et al., 1996). In contrast, cells progressing along the Leydig lineage express genes involved in secretion, consistent with the known steroidogenic function of these cells (Chen et al., 2009). IF stainings revealed that ACTA2 and MYH11 signals were low or absent in the tubule lamina in the 7- and 11-year-old samples, which display a layered ring structure around the tubules from 13 years old and onward (Figures 4F and 4G). Likewise, protein expression of CYP11A1, a marker for Leydig cell steroidogenesis, progressively appears as puberty proceeds (Figure 4H).

## Orthogonal Validation using Additional Juvenile Testis Samples

To further validate our results, we performed IHC on five additional infant-juvenile samples (1, 2, 4, 11, and 14 years old) (Figures S4E and S4F). In keeping with the results above, we observed progressive upregulation of AR expression during puberty, concomitant with gradual AMH downregulation as Sertoli cells mature (Figure S4E). Furthermore, in these independent samples the expression of CYP11A1 was evident only in the 14-year-old sample, indicating complete maturation of adult Leydig cells. Likewise, the myoid marker, ACTA2, displays clear laminar expression only later in puberty (14 years old, Figure S4F). These stainings provide additional confirmation for the generality of our genomics findings and validation of the 4 analyzed testis samples.

## Signaling Pathways Regulating Testis Development during Puberty

To explore germ cell-niche and niche-niche interactions during puberty, we examined signaling factor relationships (Figure S5). We observed dynamic and cell-type-specific expression of genes encoding ligands, inhibitors, receptors, and gene targets from multiple signaling pathways, including retinoic acid (RA), activin/inhibin, NOTCH, GDNF, FGF, and WNT (Figures S5B–S5H). This genome-wide view confirmed known relationships, previously documented in either humans or mice, while other observations provide useful cues regarding temporal and molecular relationships during puberty. For example, as *STRA8*

**Figure 3. Identification of Two Sertoli States in the Pre-pubertal Testis**

(A) Immunolocalization of germ cells and Sertoli cells in the analyzed samples (7–14 years old) illustrated by co- staining with UTF1 and SOX9.

(B) Focused analysis (tSNE and reclustering) of Sertoli cells (cluster C4 from Figure 1C), with cells colored by ages/donors.

(C) Focused analysis (tSNE and reclustering) of Sertoli cells, with cells colored according to their cell-cycle phase (G1/S/G2).

(D) Pseudotime trajectory (via Monocle) of Sertoli cells revealed two distinct early (immature) cellular states that progressively converge along the pseudotime.

(E) Deconvolution of the Monocle pseudotime plot according to ages/donors of origin.

(F) Selected key genes/programs showing differential expression in the distinct Sertoli states.

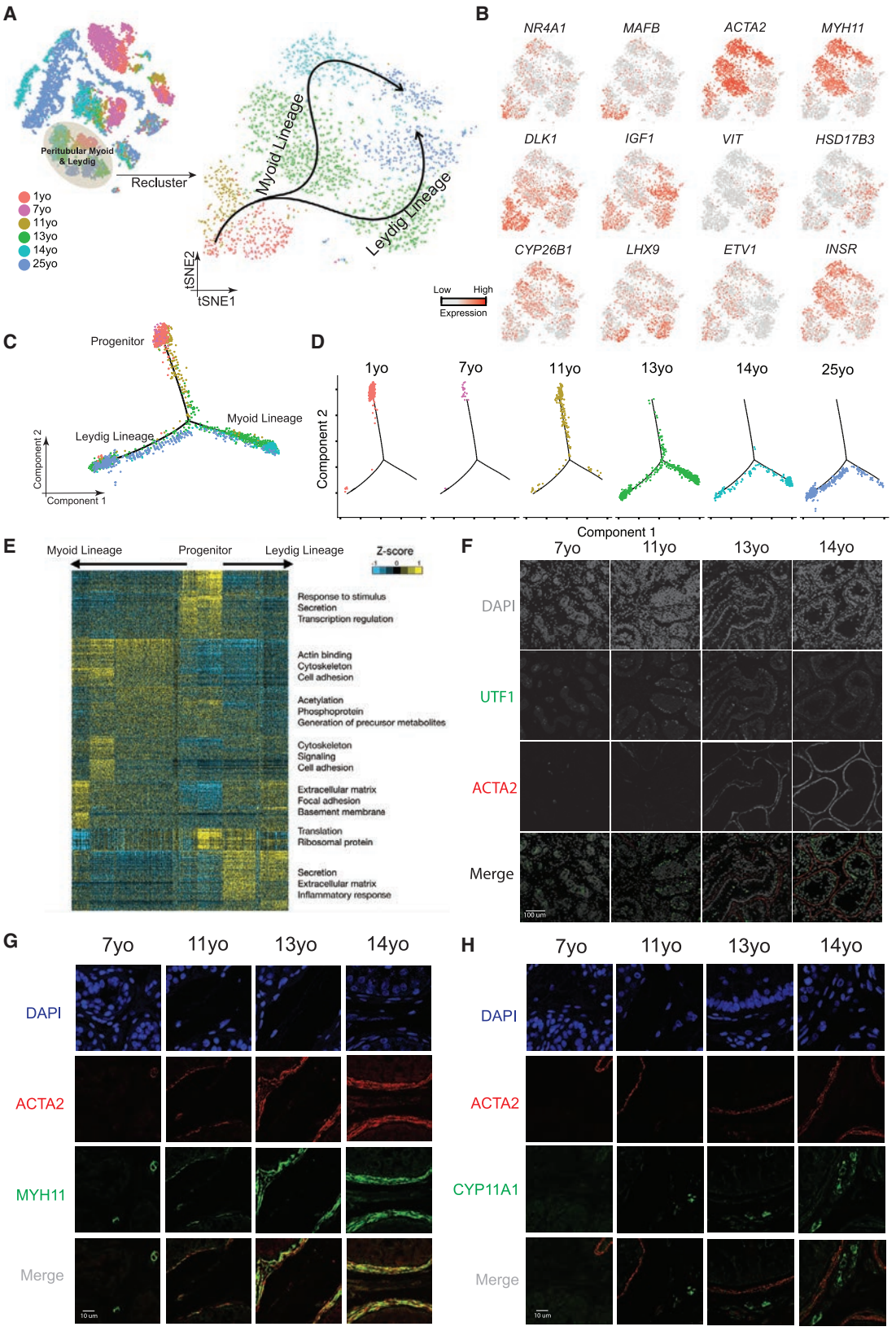
(G) Expression patterns of representative dynamic genes during Sertoli cell maturation, as predicted by pseudotime.

(H) Immunofluorescent co-staining for SOX9 and AR at different ages (7–14 years old) shows an increase in AR expression during juvenile development.

(I) Immunofluorescent co-staining for SOX9 and LGALS3 at different ages (7–14 years old) supports the progressive maturation of the Sertoli cells shown by pseudotime trajectory.

(J) Immunofluorescent co-staining for SOX9 and AMH at different ages (7–14 years old) supports the gradual maturation of Sertoli cells over time shown by pseudotime trajectory.

See also Figures S3–S5 and Table S2.




*(legend on next page)*

expression was absent from the 7- to 11-year-old samples (Figure S5B), RA is unlikely to be involved in the transition from undifferentiated to differentiating spermatogonia during human puberty, in contrast with findings in pre-pubertal mice (Velte et al., 2019). Furthermore, our data support Activin and BMP pathways as candidate key players during puberty: we observe that Sertoli cells downregulate *INHA* as they mature, while Leydig cells express high levels of *INHBA* after puberty, leading to increased Activin and lower Inhibin activity after puberty (Figure S5C). IHC staining confirmed that INHA protein is downregulated as puberty proceeds (Figure S5D). Interestingly, activin receptors (*ACVR1B*, *BMPR1B*, and *ACVR2B*) are expressed in spermatogonia, whereas key inhibitors of Activin signaling (*FST*, *BAMBI*, and *NOG*) are specifically expressed in the undifferentiated spermatogonia (States 0 and 1). Thus, the Activin pathway appears selectively inhibited in slowly self-renewing and undifferentiated spermatogonia but is active in proliferating and differentiating spermatogonia, warranting further exploration for functional roles during spermatogonial differentiation.

## Single-Cell Profiling of Testes from T-Suppressed Transfemales

Although previous data and our current analysis suggest a key role for AR and T in promoting testis development (both spermatogonia and niche cells) during puberty (Herbison, 2016; Koskenniemi et al., 2017), the molecular details involved are poorly characterized. Hence, to better define the functional role of T in the adult testis, we examined the expression profiles of testes from two adult transfemales who had been subjected to long-term reduction in T. One testis (Tf1) was from a T-suppressed 50-year-old transfemale, treated with spironolactone (a T antagonist) and estradiol for 19 months before testis removal during gender confirmation surgery. The other (Tf2) was obtained from a 26-year-old transfemale, treated with spironolactone and estradiol for 16 months before testis removal during gender confirmation surgery. In both cases, histological examination revealed major differences between the adult untreated and T-suppressed testes: macroscopically, the tubular architecture was disrupted and germ cell development was greatly impaired in the T-suppressed testes (Figure 5A). Here, scRNA-seq on isolated single cells from both T-suppressed testes was performed. For Tf1, our two technical replicates were highly similar (Figure S6A), yielding 2,161 single cells (after QC filtering). For Tf2, a single replicate yielding more than twice the cell numbers of Tf1 was analyzed. For both Tf1 and Tf2, tSNE and clustering analysis revealed the presence of the major testicular cell types as defined by clustering and annotation with key marker genes (Figures 5B and 5C). However, we noted that the relative proportion of cell types varied considerably in the two samples (Figure S6B). Thus, a larger study that examines and controls parameters (e.g., patient ages, treatment lengths, etc.) will be required to explain all differences. Nevertheless, our analyses reveal a set of consistent findings between the transfemale samples and replicates. For example, we observed a much lower proportion (in Tf1) or absence (in Tf2) of differentiating spermatogonia (e.g., *KIT*$^{+}$), spermatocytes or spermatids, whereas in both transfemale samples spermatogonia expressing undifferentiated markers (e.g., *UTF1*$^{+}$) were clearly present (Figures 5B–5E; Figure 6C; Figure S6B) and appeared relatively enriched in proportion when compared to untreated samples. In keeping with our scRNA-seq results, UTF1 protein expression was observed in both the T-suppressed and the untreated testes, while MKI67 and SYCP3 signals were both significantly reduced in the T-suppressed spermatogonia (Figure 5D). Interestingly, staining for the Sertoli cell marker SOX9 revealed a distorted localization pattern in the T-suppressed testes (Figure 5E), prompting further studies to better understand how T suppression affects Sertoli cells.

## T Promotes Germ Cell and Sertoli Cell Development *In Vivo*

To better understand the impact of T suppression on testicular cell development, we combined the scRNA-seq results from the two T-suppressed transfemales and those from all the 6 untreated males (from infancy through puberty into adulthood) and performed analysis (Figures 6A and 6B). Interestingly, spermatogonia derived from Tf1 and Tf2 cluster together with untreated spermatogonia (Figure 6B), suggesting that they do not transition to a new state but rather reside in state(s) present in untreated males. In keeping, examination of key markers that define spermatogonia states from the T-suppressed to the untreated testes revealed a higher proportion of undifferentiated spermatogonia (States 0 and 1) in the T-suppressed testes (Figure 6C). Further tSNE and clustering analysis of only spermatogonia demonstrated cells clustering based on "state" rather than "donor of origin" (Figures 6D and 6E). Results from two analyses further suggest that the undifferentiated spermatogonia from T-suppressed testes partition into State 0 and State 1 cells. First, we observed a clear separation of State 0 and State 1 cell

**Figure 4. A Common Progenitor for Leydig and Peritubular Myoid Cells**

(A) Focused analysis (tSNE and clustering) of the Leydig and peritubular myoid cells (cluster C5 from Figure 1C), with cells colored according to ages/donors of origin. The predicted developmental lineages are represented by the arrows.

(B) Expression patterns of key representative markers projected onto the tSNE plot from Figure 4A.

(C) Pseudotime trajectory (via Monocle) of the C5 cluster predicts a common early pre-pubertal progenitor state and two distinct developmental trajectories for the Leydig and myoid lineages during puberty.

(D) Deconvolution of the Monocle pseudotime plot according to the ages/donors of origin.

(E) K-means clustering of genes exhibiting differential expression (n = 1,005) along the Leydig and myoid cell lineages. Each row represents a gene, and each column represents a single cell, with columns/cells placed in pseudotime order defined in Figure 4C. Differential gene expression levels utilize a *Z* score as defined by the color key; associated GO terms (using DAVID v6.7) are given on the right of the corresponding gene clusters.

(F) Protein co-immunofluorescence for UTF1 and the peritubular myoid cell marker ACTA2 in the analyzed samples (7–14 years old) revealed that the myoid lineage (ACTA2$^{+}$) is progressively specified during puberty.

(G) Protein co-immunofluorescence for two known myoid cell markers, ACTA2 and MYH11, at different ages (7–14 years old).

(H) Immunofluorescent co-staining for ACTA2 and CYP11A1 (Leydig cell marker) shows the progressive expression of them during juvenile development.

See also Figures S4 and S5 and Table S3.

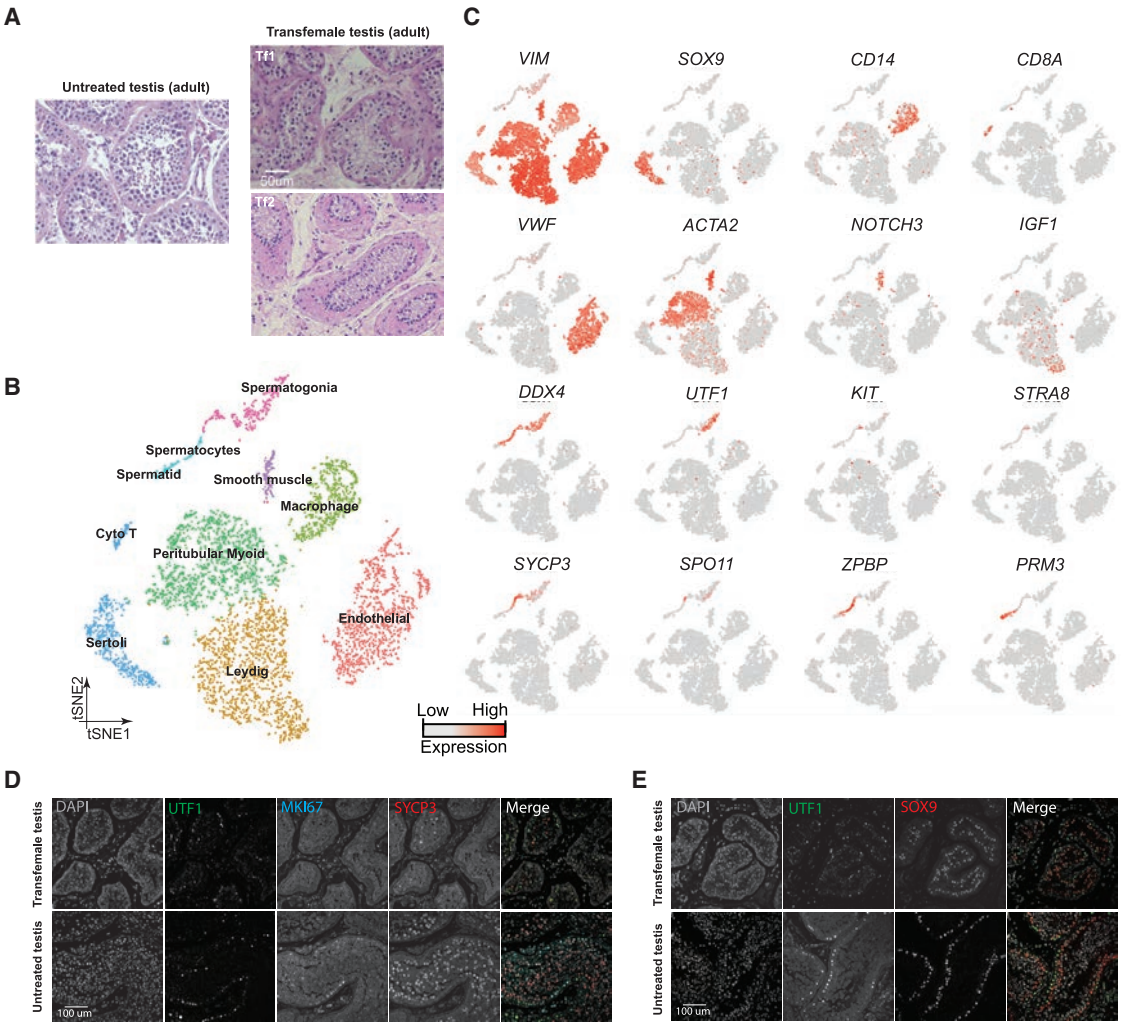



**Figure 5. Single-Cell Transcriptome Profiling of Testes from T-Suppressed Transfemales**

(A) H&E staining of the adult untreated (25 years old) and T-suppressed (Tf1 and Tf2) testicular sections.

(B) tSNE and clustering analysis of single-cell transcriptome data from two transfemale testes (n = 5,179).

(C) Expression patterns of selected markers projected on the tSNE plot. Top two rows are somatic/niche cell markers; bottom two rows are representative germ cell markers.

(D) Examination of germ cell compositions in T-suppressed (Tf1 as an example) and untreated (25 years old) testis by protein immunostaining of three germ cell markers.

(E) Immunolocalization of germ cells and Sertoli cells in T-suppressed (Tf1 as an example) and untreated (25 years old) testis by staining for UTF1 and SOX9.

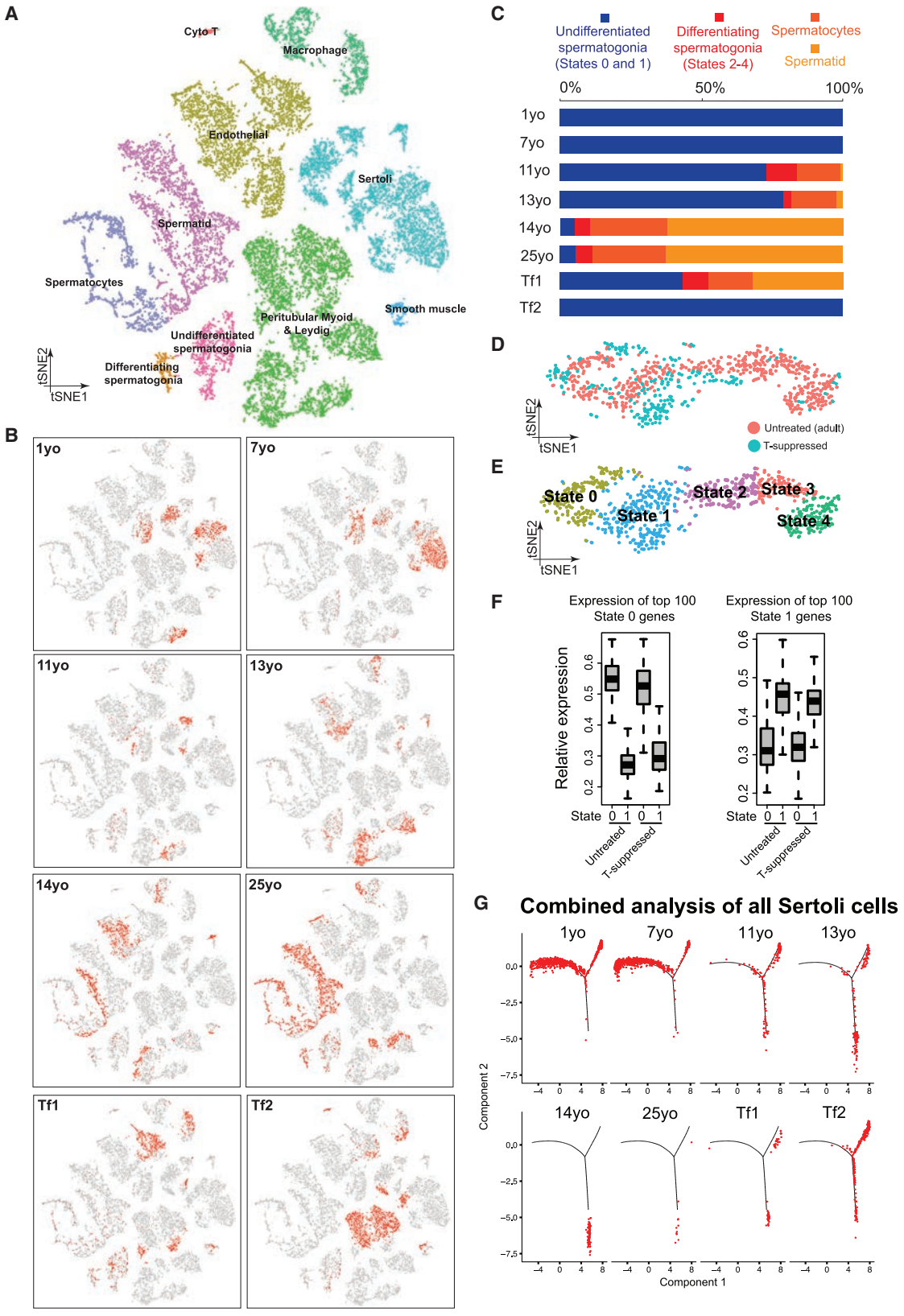
See also Figure S6 and Table S4.

clusters (Figures 6D and 6E); second, State 0 or 1 cells from either the untreated or T-suppressed testes expressed similar levels of State 0 or 1 specific markers (Figure 6F, based on a 100 gene classifier). Together, these results suggest that T has limited impact on transcriptional profiles of State 0–1 spermatogonia, though its suppression impairs progression to differentiation.

Additionally, we performed a combined analysis of Sertoli cells, and found that Sertoli cells from the two T-suppressed testes resembled more the 11- to 13-year-old samples than the 14-year-old or adult samples (Figure 6G). This is further supported by the upregulation of early Sertoli markers (including *AMH* and *HES1*), and downregulation of later Sertoli markers, including immune-related genes (*S100A6*, *ISG15*) (Figure S6E). Thus, T suppression in adults could lead to a partial reversal of Sertoli cell developmental state, mimicking aspects of the pubertal stage.

## DISCUSSION

Puberty in human males involves major changes in testis weight and physiology, including development of niche cells to establish the seminiferous tubule, spermatogonial differentiation, and modulation of hormonal signaling—culminating in the establishment of steady spermatogenesis (Koskenniemi et al., 2017; Sharpe et al., 2003). To better understand this complex process at the molecular and genome-wide level, we generated the first overview of the pubertal transcriptional cell atlas at single-cell




*(legend on next page)*

level, providing a foundational data resource coupled to multiple modes of analyses. To further facilitate data access and visualization, we have designed a website allowing browsing and query of expression patterns of genes of interest (https://humantestisatlas.shinyapps.io/humantestisatlas1/).

Regarding germ cell development, our prior dataset from adult human testes identified five cellular spermatogonial states (States 0–4) and provided evidence that States 0 and 1 correspond to the undifferentiated reserve germ cells (Guo et al., 2018). Notably, we observe a large increase (from 3% in the 7-year-old sample to ~13% in the 11-year-old sample [Figure S1C]) in the proportion of spermatogonia (relative to the total testicular cells profiled). A similar increase has been previously documented in IHC studies examining the relative numbers of type A spermatogonia (Berensztein et al., 2002; Paniagua and Nistal, 1984)—which together suggest a proliferative phase involving expansion of States 0–1 spermatogonia concomitant with seminiferous tubule reorganization and development. Of note, our scRNA-seq and protein validation do not support a simple molecular basis to account for the two histologically distinct population of $A_{dark}$ and $A_{pale}$ spermatogonia noted previously via histology (Clermont, 1966). Our results are consistent with recent work (Jan et al., 2017), which did not identify significant transcriptional differences within laser-capture microdissected $A_{dark}$ and $A_{pale}$ spermatogonia, suggesting that additional proteomic and functional work is needed to understand possible differences. Notably, tubule morphogenesis (with a defined basal lamina and lumen) coincided with the first emergence of differentiated spermatogonia (States 2–4), without completion of gametogenesis, as very few spermatids were observed. Thus, a phase of spermatogonial expansion (involving States 0-1) and initial differentiation (involving States 2–4) precedes the ability to complete spermatogenesis. Furthermore, our results from juveniles and the transfemale testes strongly suggest that States 0 and 1 are indeed reserve (undifferentiated, quiescent, or slow-cycling) spermatogonia that persist from infancy to adulthood even in the absence/suppression of T. For example, in the T-suppressed testes, while spermatogenesis is severely impaired, we observed a population of spermatogonia with the same transcriptional profile as the State 0–1 population identified in untreated adult testes (Guo et al., 2018). This likely accounts for the ability of spermatogonia in some transfemales to resume gametogenesis when T suppression is halted (Kohn et al., 2017; Schneider et al., 2017).

Interestingly, whereas AR knockout or T-suppressed rodents display a germ cell developmental block only at/after meiosis (Walker, 2011; Yeh et al., 2002), our data correlate T upregulation with the emergence of differentiating spermatogonia, as well as a role in spermiogenesis. Notably, in macaque monkeys, treatment with GnRH inhibitor for 2 weeks reduced T and impaired development beyond the undifferentiated/self-renewing ($A_{dark}$) spermatogonia but did not change the proportion of $A_{dark}$ spermatogonia (O'Donnell et al., 2001; Weinbauer et al., 1998)—suggesting that both in macaques and humans spermatogonial differentiation may require T. Whether this effect is mediated by Sertoli cells, which harbor the receptor for T (AR), and/or if T directly regulate spermatogonia through an AR-independent pathway, remains unknown. Finally, despite our observations of the role of T, our pathway analyses suggest that additional ligands (e.g., Activin, NOTCH, etc.) may control human spermatogonial differentiation.

Our work reinforces prior work that the physiology of the postnatal testis is markedly different in mice versus humans. Pre-pubertal males display small testis cords lacking an apparent lamina or lumen, structures resembling those in fetal/early postnatal mouse testes, which in peri-pubertal stages progressively acquire a more tubular structure. Regarding the location of spermatogonia in these cords and nascent tubules, consistent with prior work (Paniagua and Nistal, 1984), we observe spermatogonia at both the basement membrane and within the cord/tubule in pre-pubertal samples (Figure S2F), which in peri-pubertal and subsequent stages become fully aligned with the basement membrane. Therefore, our data and analysis—together with prior work by others—suggest that the development and physiology of human spermatogonia from 1 year old through pre-puberty is very different in humans versus mice.

Regarding somatic cell development, our study reveals several insights. First, both spermatogonia and niche cell transcriptional profiles showed very little change between the 1- and 7-year-old samples—and then changed dramatically. For example, we observed (by transcriptional criteria) two distinct populations of Sertoli cells at infant-juvenile stages that differ in metabolic, mitochondrial, and cell-cycle properties—but lack major changes in transcription and chromatin factors—indicating that these states may not represent different developmental states but rather alternative physiological states. Surprisingly, these states appear to converge into a single maturing Sertoli population during puberty. Here, the expression of PDPN in the 11-year-old testis in only a subset of tubules

**Figure 6. Testosterone Promotes Sertoli and Germ Cell Development *In Vivo***

(A) tSNE and clustering analysis of combined single-cell transcriptome data from T-suppressed (Tf1 and Tf2) and untreated (from infancy through puberty to adult) testis, with cells colored based on their identities. See Figure S6D for the markers used to assign the cell identities.

(B) Partitioning the combined tSNE analysis in Figure 6A based on the ages/donors of origin, with cells from each donor highlighted in red separately in different boxes.

(C) Relative proportion of cells at different spermatogenic stages in untreated testes (different ages) or T-suppressed testes.

(D and E) Comparison of spermatogonia from untreated adult (5 States as defined in Guo et al., 2018) and T-suppressed testes via tSNE analysis, with cells colored according to their T treatment (D) or spermatogonial states (E).

(F) Expression levels of spermatogonial State 0 genes (left) or State 1 genes (right) in untreated or T suppressed. Neither State 0, nor State 1 cells from untreated and T-suppressed testes showed statistically significant differences in gene expression via Wilcoxon test (p value = 0.528 [left] and 0.065 [right]).

(G) Comparison of Sertoli cells profile in all samples from T-suppressed (Tf1 and Tf2) and untreated (from infancy through puberty to adult) testes using focused pseudotime analysis (via Monocle).

See also Figure S6.

suggests that Sertoli maturation as an asynchronous and gradual process. Future studies should focus on investigating the physiological, metabolic, and functional differences between Sertoli states, as well as the roles of the immune peptides and their regulation by T. Finally, our data suggest that anti-bacterial/viral peptides may be important for protection from infection associated with sexual activity after puberty.

Importantly, we identified a common precursor for Leydig and myoid cells in humans that expresses (at low and heterogeneous levels) markers normally associated with both mature Leydig and myoid cells. We also identified certain factors known in mouse to reside in fetal somatic precursors (e.g., *MAFB*) but did not observe *NESTIN* (a known marker of Leydig precursors in the fetal mouse) during the human stages we examined (data not shown) (Kumar and DeFalco, 2018). These human Leydig/myoid precursors persist from infancy until peri-puberty and then diverge and mature as the seminiferous tubules develop. Thus, our results reveal major differences between mice and humans regarding the markers for somatic cell development and the timing of their appearance—with Leydig and myoid lineages being specified during fetal stages in mice (Svingen and Koopman, 2013) but only at pre-/peri-pubertal stages in humans. We note that T-suppressed Leydig and myoid cells displayed altered transcriptomes that were not simply developmentally regressed patterns, prompting further investigation on the effect of T on the maintenance of Leydig and myoid cell functions. Moreover, the developmental timeline differences between humans and mice may account for differences in the composition of peritubular myoid cells (multi-layered in humans versus mono-layered in mice)—differences that could be further explored in non-human primate models.

Regarding limitations, our current study profiled and analyzed four testis samples spanning human puberty. Given that age differences in puberty onset exist in humans, and that puberty is a progressive process, future studies with additional samples may reveal additional details regarding developmental processes and transitions. Likewise, given the differences observed between the two transfemales profiled here, additional studies are needed to understand the origins of the variation observed. Furthermore, our analyses and interpretations focus primarily on transcriptional data supported by initial protein validation studies, and post-transcriptional mechanisms exist that will complement and refine our understanding of the molecular strategies and the complex interplay between the different cellular components of the juvenile testis.

The human testis is one of the few organs that defines most of its cell types, physiology, and function after birth. Our work provides major advances in defining the strategy and timing of human testis development. During pre-pubertal stages, the testis maintains a pool of undifferentiated and largely quiescent germline stem cells within an immature niche. During early puberty, maturation of the niche occurs first, involving the differentiation of a common progenitor to form mature Leydig and peritubular myoid cells, which contribute to defining the lamina-luminal architecture of the seminiferous tubule to promote signaling pathways—in conjunction with hormonal induction from the hypothalamic-pituitary-gonadal axis. Likewise, Sertoli cell populations develop and change their localization, and their transcriptional, metabolic, and signaling profiles. Proliferation of undifferentiated spermatogonia then commences to progressively populate the elongating tubules, followed by spermatogonial differentiation and the generation of the first meiotic cells, culminating in the steady generation of spermatids and mature sperm. Our data further strengthen the notion that establishment of spermatogenesis requires a complex interplay between multiple signaling pathways. These results should provide the foundation for building hypothesis-driven research that can be explored in primate models and will support improvement of *in vitro* cultures of human seminiferous tubules and spermatogonial maturation, to ultimately help guide therapeutic options for male infertility.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- LEAD CONTACT AND MATERIALS AVAILABILITY
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Human Testicular Tissues
- METHOD DETAILS
  - Sample Storage by Cryopreservation
  - Sample Fixation for Immunostainings
  - Human Testis Sample Preparation for Single Cell RNA Sequencing
  - Single Cell RNA-seq Performance, Library Preparation and Sequencing
  - Processing of Single Cell RNA-seq Data
  - Immunostaining of Testicular Tissues
  - Immunohistochemistry staining
  - H&E staining
  - Cell counting
- QUANTIFICATION AND STATISTICAL ANALYSIS
- DATA AND CODE AVAILABILITY

## SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.stem.2019.12.005.

## ACKNOWLEDGMENTS

We are especially grateful to the donors and their families, who made this work possible. We thank Brian Dalley and Opal Allen (Genomics Core at Huntsman Cancer Institute) for sequencing expertise, Chris Conley and Tim Parnell (Bioinformatics Core at Huntsman Cancer Institute) for bioinformatics assistance, the Huntsman Cancer Institute Biorepository and Molecular Pathology Shared Resource for tissue sectioning, and DonorConnect staff for family consents and sample handling. Imaging was performed at the Fluorescence Microscopy Core Facility at the Health Sciences Cores at University of Utah, supported by an NCRR Shared Equipment Grant 1S10RR024761-01. Financial support was from Howard Hughes Medical Institute to B.R.C. and P30CA042014 from the National Cancer Institute to Huntsman Cancer Institute for core facilities. The Wellcome, UK (Senior Investigator Award 102731 to Prof. Andrew Wilkie) supports A.G. and H.M.; A.G. is supported by the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre (BRC), UK; M.G. is supported by the EU(ESF)-funded Research Exchange Network HAL-OX ZS/2016/08/80642 of the Martin Luther University Halle-Wittenberg, Germany; and the Wolfson Imaging Centre (Oxford) is supported by the WIMM Strategic Alliance (G0902418 and MC_UU_12025). C.L. is supported by the Knut and

Alice Wallenberg Foundation. R.T.M. is supported by a UKRI Future Leaders Fellowship (MR/S017151/1) and MRC Centre for Reproductive Health Centre Grant (MR/N022556/1).

## AUTHOR CONTRIBUTIONS

B.R.C. and J.G. conceived the project, designed the genomics and computational analysis, and advised on other sections; J.G. conducted genomics experiments and computational analyses with help from X.N.; Y.W. contributed to website construction with help from J.G.; sample acquisition was led by J.M.H. and R.K. with input from D.T.C., B.R.C., and J.G.; samples were transferred and processed by J.G., E.J.G., and X.N.; C.L. provided HPA antibodies and contributed to antibody validation; J.G., X.N., M.G., H.M., M.T., and G.M. performed IHC/IF experiments that were supervised and coordinated by A.G. and R.T.M.; the manuscript was written by J.G. and B.R.C. with input from R.T.M. and A.G. and agreement of all authors.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: August 9, 2019
Revised: November 3, 2019
Accepted: December 5, 2019
Published: January 9, 2020

## REFERENCES

Bellvé, A.R., Cavicchia, J.C., Millette, C.F., O'Brien, D.A., Bhatnagar, Y.M., and Dym, M. (1977). Spermatogenic cells of the prepuberal mouse. Isolation and morphological characterization. J. Cell Biol. *74*, 68–85.

Berensztein, E.B., Sciara, M.I., Rivarola, M.A., and Belgorosky, A. (2002). Apoptosis and proliferation of human testicular somatic and germ cells during prepuberty: high rate of testicular growth in newborns mediated by decreased apoptosis. J. Clin. Endocrinol. Metab. *87*, 5113–5118.

Birnbaum, K.D. (2018). Power in Numbers: Single-Cell RNA-Seq Strategies to Dissect Complex Tissues. Annu. Rev. Genet. *52*, 203–221.

Bolton, E.C., So, A.Y., Chaivorapol, C., Haqq, C.M., Li, H., and Yamamoto, K.R. (2007). Cell- and gene-specific regulation of primary target genes by the androgen receptor. Genes Dev. *21*, 2005–2017.

Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat. Biotechnol. *36*, 411–420.

Chemes, H.E. (2001). Infancy is not a quiescent period of testicular development. Int. J. Androl. *24*, 2–7.

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

Clermont, Y. (1966). Spermatogenesis in man. A study of the spermatogonial population. Fertil. Steril. *17*, 705–721.

Com, E., Bourgeon, F., Evrard, B., Ganz, T., Colleu, D., Jégou, B., and Pineau, C. (2003). Expression of antimicrobial defensins in the male reproductive tract of rats, mice, and humans. Biol. Reprod. *68*, 95–104.

DeFalco, T., Takahashi, S., and Capel, B. (2011). Two distinct origins for Leydig cell progenitors in the fetal testis. Dev. Biol. *352*, 14–26.

Griswold, M.D. (2018). 50 years of spermatogenesis: Sertoli cells and their interactions with germ cells. Biol. Reprod. *99*, 87–100.

Guo, J., Grow, E.J., Yi, C., Mlcochova, H., Maher, G.J., Lindskog, C., Murphy, P.J., Wike, C.L., Carrell, D.T., Goriely, A., et al. (2017). Chromatin and Single-Cell RNA-Seq Profiling Reveal Dynamic Signaling and Metabolic Transitions during Human Spermatogonial Stem Cell Development. Cell Stem Cell *21*, 533–546.e6.

Guo, J., Grow, E.J., Mlcochova, H., Maher, G.J., Lindskog, C., Nie, X., Guo, Y., Takei, Y., Yun, J., Cai, L., et al. (2018). The adult human testis transcriptional cell atlas. Cell Res. *28*, 1141–1157.

Herbison, A.E. (2016). Control of puberty onset and fertility by gonadotropin-releasing hormone neurons. Nat. Rev. Endocrinol. *12*, 452–466.

Hermann, B.P., Cheng, K., Singh, A., Roa-De La Cruz, L., Mutoji, K.N., Chen, I.-C., Gildersleeve, H., Lehle, J.D., Mayo, M., Westernströer, B., et al. (2018). The Mammalian Spermatogenesis Single-Cell Transcriptome, from Spermatogonial Stem Cells to Spermatids. Cell Rep. *25*, 1650–1667.

Huang, W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. *4*, 44–57.

Jan, S.Z., Vormer, T.L., Jongejan, A., Röling, M.D., Silber, S.J., de Rooij, D.G., Hamer, G., Repping, S., and van Pelt, A.M.M. (2017). Unraveling transcriptome dynamics in human spermatogenesis. Development *144*, 3659–3673.

Johnsen, S.G. (1970). Testicular biopsy score count–a method for registration of spermatogenesis in human testes: normal values and results in 335 hypogonadal males. Hormones *1*, 2–25.

Kanatsu-Shinohara, M., and Shinohara, T. (2013). Spermatogonial stem cell self-renewal and development. Annu. Rev. Cell Dev. Biol. *29*, 163–187.

Kluin, P.M., and de Rooij, D.G. (1981). A comparison between the morphology and cell kinetics of gonocytes and adult type undifferentiated spermatogonia in the mouse. Int. J. Androl. *4*, 475–493.

Kohn, T.P., Louis, M.R., Pickett, S.M., Lindgren, M.C., Kohn, J.R., Pastuszak, A.W., and Lipshultz, L.I. (2017). Age and duration of testosterone therapy predict time to return of sperm count after human chorionic gonadotropin therapy. Fertil. Steril. *107*, 351–357.e1.

Koskenniemi, J.J., Virtanen, H.E., and Toppari, J. (2017). Testicular growth and development in puberty. Curr. Opin. Endocrinol. Diabetes Obes. *24*, 215–224.

Kumar, D.L., and DeFalco, T. (2018). A perivascular niche for multipotent progenitors in the fetal testis. Nat. Commun. *9*, 4519.

Lindskog, C. (2016). The Human Protein Atlas—an important resource for basic and clinical research. Expert Rev. Proteomics *13*, 627–629.

Lun, A.T.L., McCarthy, D.J., and Marioni, J.C. (2016). A step-by-step workflow for low-level analysis of single-cell RNA-seq data with Bioconductor. F1000Res. *5*, 2122.

Maekawa, M., Kamimura, K., and Nagano, T. (1996). Peritubular myoid cells in the testis: their structure and function. Arch. Histol. Cytol. *59*, 1–13.

O'Donnell, L., Narula, A., Balourdos, G., Gu, Y.Q., Wreford, N.G., Robertson, D.M., Bremner, W.J., and McLachlan, R.I. (2001). Impairment of spermatogonial development and spermiation after testosterone-induced gonadotropin suppression in adult monkeys (Macaca fascicularis). J. Clin. Endocrinol. Metab. *86*, 1814–1822.

Paniagua, R., and Nistal, M. (1984). Morphological and histometric study of human spermatogonia from birth to the onset of puberty. J. Anat. *139*, 535–552.

Plant, T.M. (2015). Neuroendocrine control of the onset of puberty. Front. Neuroendocrinol. *38*, 73–88.

Qiu, X., Mao, Q., Tang, Y., Wang, L., Chawla, R., Pliner, H.A., and Trapnell, C. (2017). Reversed graph embedding resolves complex single-cell trajectories. Nat. Methods *14*, 979–982.

Schneider, F., Kliesch, S., Schlatt, S., and Neuhaus, N. (2017). Andrology of male-to-female transsexuals: influence of cross-sex hormone therapy on testicular function. Andrology *5*, 873–880.

Sharpe, R.M., McKinnell, C., Kivlin, C., and Fisher, J.S. (2003). Proliferation and functional maturation of Sertoli cells, and their relevance to disorders of testis function in adulthood. Reproduction *125*, 769–784.

Sohni, A., Tan, K., Song, H.-W., Burow, D., de Rooij, D.G., Laurent, L., Hsieh, T.-C., Rabah, R., Hammoud, S.S., Vicini, E., and Wilkinson, M.F. (2019). The Neonatal and Adult Human Testis Defined at the Single-Cell Level. Cell Rep. *26*, 1501–1517.e4.

Svingen, T., and Koopman, P. (2013). Building the mammalian testis: origins, differentiation, and assembly of the component cell populations. Genes Dev. *27*, 2409–2426.

Tharmalingam, M.D., Jorgensen, A., and Mitchell, R.T. (2018). Experimental models of testicular development and function using human tissue and cells. Mol. Cell. Endocrinol. *468*, 95–110.

Velte, E.K., Niedenberger, B.A., Serra, N.D., Singh, A., Roa-DeLaCruz, L., Hermann, B.P., and Geyer, C.B. (2019). Differential RA responsiveness directs

formation of functionally distinct spermatogonial populations at the initiation of spermatogenesis in the mouse. Development *146*, dev173088.

Vergouwen, R.P., Huiskamp, R., Bas, R.J., Roepers-Gajadien, H.L., Davids, J.A., and de Rooij, D.G. (1993). Postnatal development of testicular cell populations in mice. J. Reprod. Fertil. *99*, 479–485.

Walker, W.H. (2011). Testosterone signaling and the regulation of spermatogenesis. Spermatogenesis *1*, 116–120.

Wang, M., Liu, X., Chang, G., Chen, Y., An, G., Yan, L., Gao, S., Xu, Y., Cui, Y., Dong, J., et al. (2018). Single-Cell RNA Sequencing Analysis Reveals Sequential Cell Fate Transition during Human Spermatogenesis. Cell Stem Cell *23*, 599–614.e4.

Weinbauer, G.F., Schubert, J., Yeung, C.H., Rosiepen, G., and Nieschlag, E. (1998). Gonadotrophin-releasing hormone antagonist arrests premeiotic germ cell proliferation but does not inhibit meiosis in the male monkey: a quantitative analysis using 5-bromodeoxyuridine and dual parameter flow cytometry. J. Endocrinol. *156*, 23–34.

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

Yeh, S., Tsai, M.-Y., Xu, Q., Mu, X.-M., Lardy, H., Huang, K.-E., Lin, H., Yeh, S.-D., Altuwaijri, S., Zhou, X., et al. (2002). Generation and characterization of androgen receptor knockout (ARKO) mice: an in vivo model for the study of androgen functions in selective tissues. Proc. Natl. Acad. Sci. USA *99*, 13498–13503.

Yoshida, S., Sukeno, M., Nakagawa, T., Ohbo, K., Nagamatsu, G., Suda, T., and Nabeshima, Y. (2006). The first round of mouse spermatogenesis is a distinctive program that lacks the self-renewing spermatogonia stage. Development *133*, 1495–1505.

## STAR★METHODS

### KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| Antibodies | | |
| Mouse monoclonal anti-UTF1 Dilution: 1:100 - 1:1000 | Millipore | cat# Mab4337 (5G10.2), RRID:AB_827541 |
| Rat polyclonal anti-MKI67 Dilution: 1:100 | Abcam | cat# ab15580, RRID:AB_443209 |
| Rabbit monoclonal anti-MKI67 Dilution: 1:200 | Abcam | cat# ab16667, RRID:AB_302459 |
| Mouse polyclonal anti-ACTA2 Dilution: 1:400 | Abcam | cat# ab5694, RRID:AB_2223021 |
| Mouse monoclonal anti-ACTA2 Dilution: 1:300 | Atlas Antibodies | cat# cab013531, RRID:AB_2677383 |
| Mouse polyclonal anti-ACTA2 Dilution: 1:250 | Sigma Aldrich | cat# A2547, RRID:AB_476701 |
| Mouse monoclonal anti-AMH Dilution: 1:300 | Santa Cruz | cat# sc-166752 RRID:AB_2289536 |
| Mouse monoclonal anti-AR Dilution: 1:300 | Atlas Antibodies | cat# CAB000001, RRID:AB_2685535 |
| Rabbit monoclonal anti-AR Dilution: 1:250 | Abcam | cat# ab133273, RRID:AB_11156085 |
| Goat polyclonal anti-KIT Dilution: 1:25 | R&D systems | cat# AF332, RRID:AB_355302 |
| Rabbit polyclonal anti-INHA Dilution: 1:1000 | Atlas Antibodies | cat# HPA019141, RRID:AB_1851789 |
| Mouse monoclonal anti-LGALS3 Dilution: 1:500 | Santa Cruz | cat# sc-32790, RRID:AB_627657 |
| Rabbit polyclonal anti-MYH11 Dilution: 1:100 | Atlas Antibodies | cat# HPA015310, RRID:AB_1854261 |
| Rabbit polyclonal anti-S100A6 Dilution: 1:750 | Atlas Antibodies | cat# HPA007575, RRID:AB_1079859 |
| Rabbit polyclonal anti-SOX9 Dilution: 1:1000 | Millipore | cat# AB5535, RRID:AB_2239761 |
| Rabbit monoclonal anti-SOX9 Dilution: 1:300 | Abcam | cat# ab185966, RRID:AB_2728660 |
| Rabbit monoclonal anti-SYCP3 Dilution: 1:400 | Abcam | cat# ab15093, RRID:AB_301639 |
| Rabbit polyclonal anti-CYP11A1 Dilution: 1:250 | Atlas Antibodies | cat# HPA016436, RRID:AB_1847423 |
| Mouse polyclonal anti-PDPN Dilution: 1:500 | Abcam | cat# ab77854, RRID:AB_1566117 |
| Donkey anti-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 488 | Thermo Fisher Scientific | cat# A21202, RRID:AB_141607 |
| Donkey anti-Goat IgG (H+L) Highly Cross-Absorbed Secondary Antibody, Alexa Fluor 555 | Thermo Fisher Scientific | cat# A21432, RRID:AB_141788 |
| Donkey anti-Rabbit IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 488 | Thermo Fisher Scientific | cat# A31573, RRID:AB_141788 |
| Donkey anti-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 555 | Thermo Fisher Scientific | cat# A31570, RRID:AB_2536180 |
| Biological Samples | | |
| Human testis samples for donors | DonorConnect | NA |
| Human testis samples from transfemale | Universityf of Utah Andrology laboratory | NA |
| Human testis samples for immunohistochemistry | MRC Centre for Reproductive Health, The Queen's Medical Research Institute, The University of Edinburgh and Royal Hospital for Children and Young People | NA |
| Deposited Data | | |
| Single cell RNA-seq for pre- and peri-pubertal human testes | This paper | GEO: GSE134144 |
| Single cell RNA-seq for transfemale testes | This paper | GEO: GSE134144 |
| Software and Algorithms | | |
| Seurat (2.3.4) | Butler et al., 2018 | https://satijalab.org/seurat/ |
| Monocle (2.10.1) | Qiu et al., 2017 | http://cole-trapnell-lab.github.io/monocle-release/ |
| GO (David 6.7) | Huang et al., 2009 | https://david-d.ncifcrf.gov |
| Cell Ranger (2.2.0) | NA | https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/what-is-cell-ranger |

(*Continued on next page*)

*Continued*

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| Cluster 3.0 | NA | http://bonsai.hgc.jp/~mdehoon/software/cluster/software.htm |
| scran (1.6.5) | Lun et al., 2016 | https://bioconductor.org/packages/3.7/bioc/vignettes/scran/inst/doc/scran.html |
| Other | | |
| Single cell RNA-seq for infant and adult human testes | Guo et al., 2018 | GEO: GSE120508 |
| The Human Protein Atlas | Lindskog, 2016 | http://www.proteinatlas.org |

## LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for reagents should be directed to and will be fulfilled by the Lead Contact, Bradley R. Cairns (brad.cairns@hci.utah.edu). This study did not generate new unique reagents.

## EXPERIMENTAL MODEL AND SUBJECT DETAILS

### Human Testicular Tissues

Pre- and peri-pubertal human testicular samples were obtained from four healthy boys aged 7, 11, 13 and 14 years old, through the University of Utah Andrology laboratory and Connect Donor (formerly named Intermountain Donor Services). Those samples were removed from deceased individuals who consented to organ donation for transplantation and research.

Transfemale (testosterone-suppressed) samples were obtained through the University of Utah Andrology laboratory consented for research (IRB approved protocol #00075836).

Pre- and peri-pubertal tissues for orthogonal validation studies were obtained from males undergoing testicular tissue cryopreservation for fertility preservation with prior ethical approval from the South East Scotland Research Ethics Committee (Reference: 13SS/0145).

## METHOD DETAILS

### Sample Storage by Cryopreservation

Once collected, the pair of whole-testis was transported to the research laboratory on ice in saline or Hank's Balanced Salt Solution (HBSS; GIBCO cat # 24020117) and processed within 1 hour of removal by surgery. Around 90% of each testis was divided into smaller portions (~500 mg – 1g each) using scissors and directly transferred into cryovials (Corning cat # 403659) in DMEM medium (Life Technologies cat # 11995073) containing 10% DMSO (Sigma-Aldrich cat # D8779), 15% fetal bovine serum/FBS (GIBCO cat # 10082147) and cryopreserved in Mr. Frosty container (Thermo Fisher Scientific cat #5100-0001) at a controlled slow rate, and stored at −80°C for overnight. Cryovials were transferred to liquid nitrogen for long-term storage.

### Sample Fixation for Immunostainings

Around 10% of the remaining testis tissues were placed in 1X PBS containing 4% paraformaldehyde/PFA (Thermo Fisher Scientific cat # 28908) and incubated overnight at 4C with agitation on a rotor (60 rpm). Fixed samples were then washed with three times in cold PBS, and stored in PBS at 4C until immunostaining processing.

### Human Testis Sample Preparation for Single Cell RNA Sequencing

For each single-cell sequencing experiment, 1 or 2 cryovials were thawed quickly. Tissues were then washed twice with PBS, and subject to digestion as described previously (Guo et al., 2018). For pubertal samples, tissues were washed twice in 1 x PBS, and minced into small pieces for better digestion outcome. Tissues were then treated with trypsin/ethylenediaminetetraacetic acid (EDTA; Invitrogen cat # 25300054) for 20-25 min and collagenase type IV (Sigma Aldrich cat # C5138-500MG) at 37°C. For the transfemale samples, tissues were digested following the standard two step enzymatic isolation protocol as described in (Guo et al., 2018). The digestion was then stopped by adding 10% FBS (GIBCO cat # 10082147). Single testicular cells were obtained by filtering through 70 μm (Fisher Scientific cat # 08-771-2) and 40 μm (Fisher Scientific cat # 08-771-1) strainers. The cells were pelleted by centrifugation at 600 g for 15 min, and wash with PBS twice. Cell number was counted using hemocytometer, and the cells were then resuspended in PBS + 0.4% BSA (Thermo Fisher Scientific cat # AM2616) at a concentration of ~1,000 cells/uL ready for single-cell sequencing.

### Single Cell RNA-seq Performance, Library Preparation and Sequencing

Here, we aimed to capture ~4,000-5,000 cells. Briefly, cells were diluted following manufacturer's instructions, and 33.8 uL of total mixed buffer together with cells were loaded into 10x Chromium Controller using the Chromium Single Cell 3′ v2 reagents. The

sequencing libraries were prepared following the manufacturer's instructions, using 13 cycles of cDNA amplification, followed by an imput of ~100 ng of cDNA for library amplification using 12 cycles. The resulting libraries were then sequenced on a 2 X 100 cycle paired-end run on an Illumina HiSeq 2500 or Novaseq 6000 instruments.

### Processing of Single Cell RNA-seq Data

Raw data were demultiplexed using mkfastq application (Cell Ranger v2.2.0) to make Fastq files. Fastq files were then run with count application (Cell Ranger v2.2.0) using default settings, which performs alignment (using STAR aligner), filtering and UMI counting. The UMI count tables were used for further analysis.

### Immunostaining of Testicular Tissues

The immunofluorescence (IF) stainings were performed on 5μm formalin-fixed paraffin embedded (FFPE) or cryopreserved sections from portions of the collected testicular samples following deparaffinisation, rehydratation and heat-mediated antigen retrieval in 10mM sodium citrate buffer solution (pH 6). After treatment with Superblock (PBS) Blocking Buffer (Thermo Fisher Scientific, cat# 37515) for 30 mins, individual sections were incubated overnight at 4°C with a mix of diluted antibodies (for antibodies details and dilutions, see the Key Resources Table). Antigen detection was conducted using the appropriate combination of Alexa Fluor 488 and 555 secondary antibodies (all 1:500; Thermo Fisher Scientific, cat#A21202, cat#A21432, cat#A31570, cat#A21206, respectively) for 2 hr at room temperature in the dark. All primary/secondary antibodies were diluted in SignalBoost Immunoreaction Enhancer Kit (Calbiochem, cat#407207-1KIT). After three washes in PBS, sections were incubated with DAPI (4',6-Diamidine-2-phenylindole dihydrochloride) (Roche, cat#10236276001) to facilitate nuclear visualization (dilution 1:2000). Specificity of the antibody staining was confirmed using the same protocol but with omission of primary antibodies. Following multiple washes in PBS, slides were mounted using Vectashield mounting medium for fluorescence (Vector Laboratories, Inc., Burlingame, CA, cat#H-1000). Images were obtained under 25x objective (LD LCI PA 25x/0.8 DIC WD = 0.57 mm Imm Corr (UV)VIS-IR (Oil-Immersion) with a Zeiss LSM 780 Upright Multi-Photon Confocal Microscope and analyzed using ImageJ software (NIH, Bethesda, MD, USA).

### Immunohistochemistry staining

For validation using the four pubertal samples, immunohistochemistry on 5 μm sections from FFPE or cryopreserved testis samples was performed using the Rabbit specific HRP/DAB (ABC) Detection IHC Kit (Abcam, cat# 236466). Briefly, after deparaffinisation, rehydration and heat-mediated antigen retrieval in 10mM sodium citrate buffer solution (pH 6), sections were blocked with Protein block and incubated overnight at 4°C with a variety of different antibodies in 1x TBS. Subsequently, sections were incubated with HRP conjugate before chromogenic detection using DAB (3,3′Diaminobenzidine). Nuclei were counterstained with hematoxylin (Poly Scientific, cat#s212A), slides mounted in Aquatex (Merck, cat#108562) and images were acquired using the Hamamatsu NanoZoomer slide scanner and the NDP.view 2 software (Hamamatsu Photonics).

For validation in additional human testis tissue samples, immunohistochemistry was performed on Bouin's fixed, paraffin-embedded tissue sections of approximately 5μm thickness. The sections were deparaffinized with xylene and rehydrated in ethanol using a reducing concentration from 100% to 70% (20secs each). Between each step sections were washed in TBS. If required, heat-mediated antigen retrieval in 10mM citrate buffer, pH6 was performed in a pressure cooker. Sections were cooled in TBS and endogenous peroxidase activity was blocked in 3% (v/v) hydrogen peroxide (VWR International) in methanol (Fisher Chemicals) for 30mins. Sections were blocked in Normal Horse Serum (Diagnostics Scotland) diluted (1:5) in TBS with 5% Bovine Serum Albumin (Sigma Aldrich). Sections were incubated in primary antibody diluted in blocking serum overnight in a humidified chamber at 4°C. The following day, sections were incubated with the relevant Impress HRP secondary antibody at room temperature for 30mins. Visualization of staining was performed using DAB (1drop/ml) diluted in buffer. Sections were counterstained with Haematoxylin for 2.5mins, Scott's tap water for 30 s and dehydrated by gradual introduction to ethanol from 70% to 100% (20secs each) and Xylene (2x 5mins). Sections were mounted using Pertex (CellPath). Brighfield images were captured using the ZEISS Axio Scan.Z1 at 20x magnification.

### H&E staining

For hematoxylin and eosin (H&E) staining, deparaffinized and rehydrated 5 μm sections were incubated in hematoxylin for 3 min and rinsed with running tap water for 5 min. Afterward, the sections were dipped in acid alcohol (0,5% v/v hydrochloric acid in 70% ethanol), washed with distilled H2O, and incubated in eosin (Poly Scientific, cat#176) for 30 s. The sections were dehydrated before mounting with Histomount (National Diagnostics, cat# 12954910).

Johnsen scores are the standard quantitative/clinical histological grading system utilized by testis pathologists, which assesses the degree of spermatogenic maturation in individual tubular cross-sections, ranging between 1 to 10, with 1 representing complete azoospermia, and 10 representing full spermatogenesis(Johnsen, 1970).

### Cell counting

For quantification of UTF and/or KIT expressing cells, single-positive and double-positive cells were counted in cross-sections of seminiferous tubules. The number of positive cells per cross-section were normalized to the total number of cells located in the periphery of the seminiferous tubule. The bars represent means + SD of independent tubules. Data were analyzed using unpaired two-sided Student's t test. $p < 0.05$ was considered statistically significant.

## QUANTIFICATION AND STATISTICAL ANALYSIS

The Seurat program (https://satijalab.org/seurat/, R package, v.2.3.4) was used as a first analytical package. To start with, UMI count tables from both replicates from all four juvenile donors were loaded into R using Read10X function, and Seurat objects were built from each experiment. Each experiment was filtered and normalized with default settings. Specifically, cells were retained only if they contained > 500 expressed genes, and had < 25% reads mapped to mitochondrial genome. We first run t-SNE and clustering analysis on each replicate, which resulted in similar t-SNE map. Data matrices from different donors and replicates were then combined with the previously published infant and adult data (Guo et al., 2018). Next, cells were normalized to the total UMI read counts, as instructed in the tutorial (http://satijalab.org/seurat/). t-SNE and clustering analyses were performed on the combined data using the top 6,000 highly variable genes and 1-30 PCs, which showed the most significant p values.

Detailed pseudotime for different cell types were performed using the Monocle package (v2.10.1) following the default settings. After pseudotime coordinates/order were determined, gene clustering analysis was performed to establish the accuracy of pseudotime ordering. Here, cells (in columns) were ordered by their pseudotime, and genes (in rows) were clustered by k-means clustering using Cluster 3.0. Different k-mean numbers were performed to reach the optimal clustering number. Cell cycle analysis was performed using scran program (https://bioconductor.org/packages/3.7/bioc/vignettes/scran/inst/doc/scran.html, R Package; v1.6.5).

## DATA AND CODE AVAILABILITY

All software tools can be found online (see Key Resources Table). The accession number for all sequencing data reported in this paper is GEO: GSE134144.