Inaugural dissertation

for
obtaining the doctoral degree
of the
Combined Faculty of Mathematics, Engineering and Natural Sciences
of the
Ruprecht - Karls - University
Heidelberg

Presented by
M.Sc. Noe Mbengue
Born in Stuttgart, Germany
Oral examination June 30th 2025

Transcriptome evolution of the mammalian germline

Referees: Prof. Dr. Henrik Kaessmann
Prof. Dr. Duncan Odom

Summary

Sexual reproduction plays a central part in life and evolution, as it allows the transmission of genetic information across generations. As the offspring is generated from gametes of both parents, sexual reproduction is a source for diversity within a population. The gametes are formed through gametogenesis of the germ line, in a process which can generally be divided into three phases: 1) Mitotic expansion of the progenitor pool. 2) Meiotic reductive divisions. 3) Post-meiotic changes. For males this takes place in the testis and generates sperm, while in females this process takes place in the ovary and generates oocytes. From an evolutionary perspective, especially the testis is an interesting tissue, as it has been shown to evolve rapidly both morphologically as well as molecularly. Yet, so far it was not fully dissected, which cells are the driver of this observation. In my thesis work, I address this question in adult testis on the basis of single cell level transcriptional data from 11 species covering the three main mammalian lineages (monotremes, marsupials, and eutherians) and a bird. Among the mammalian species are seven primates, including all extant great apes, except for Orangutan. To also further our understanding of the female germline evolution, I also analyzed a dataset of single cell level transcriptional data of developing ovaries for three glires and a bird.

In these analyses, I show how we dissected the testicular cell type transcriptomes. We found that the fast divergence of the testis is especially driven by the post-meiotic cell types. Furthermore, we showed that this fast divergence is driven by both a relaxation of negative selection, as well as an increase of positive selection. In addition, we found support for the notion of the “out of the testis” hypothesis, according to which the testis plays an important role in the origin of new genes.

Through assessing the expression dynamics of individual genes, we found shared and lineage specific aspects of the transcriptome. In this analysis we identified a core set of genes, which maintained their expression across mammalian evolution and is likely connected to central testicular function and furthermore connected to maintenance of fertility. In the analysis of lineage specific expression patterns, we also found genes with relevance for fertility. Tracing the functional impact of individual genes, we also found conserved aspects of Sertoli cell to germ cell communication.

Specifically focusing on the analysis of sex chromosome expression, we found an accumulation of testis specific X-linked genes in spermatogonia. Furthermore, we dissected the

transcriptomal differences between the X- and Y-carrying spermatids. In addition to this, we assessed the silencing of sex chromosomes upon meiotic entry and showed, for the first time transcriptomal evidence for meiotic sex chromosome inactivation (MSCI) in platypus.

In the analysis of the developmental ovary dataset I generated, I identified the basic somatic cell types. In the assessment of the germ cells I found, that I could not clearly distinguish multiple subtypes of germ cells in the mammalian data. In the bird data I was able to distinguish germ cell subtypes.

Analyzing the cell type transcriptomes in this dataset, I found that germ cells diverge faster than somatic cells.

In the dedicated analysis of the bird germ line transcriptomes in regards to sex chromosome expression, I found support for the presence of MSCI in birds.

Taken together my work helps to explain the observed diversity in testicular phenotypes and explain their molecular sources. For the ovary, my data provides an exciting starting step for further exploration.

Zusammenfassung

Sexuelle Fortpflanzung spielt eine zentrale Rolle im Leben und in der Evolution, da sie die Übertragung von genetischer Information über Generationen hinweg ermöglicht. Da Nachwuchs aus den Gameten beider Eltern entsteht, ist sexuelle Reproduktion eine Diversitätsquelle innerhalb einer Population. Gameten werden durch Gametogenese der Keimbahn gebildet. Dieser Vorgang kann in drei Phasen untergliedert werden: 1) Mitotische Expansion des Progenitorpools. 2) Reduktive meiotische Zellteilungen. 3) Postmeiotische Veränderungen. In Männern findet dieser Vorgang in den Hoden statt und generiert Spermazellen. In Frauen findet dieser Vorgang in den Eierstöcken statt und generiert Eizellen. Aus der Perspektive der Evolution ist der Hoden ein besonders interessantes Gewebe, da gezeigt wurde, dass dieser sowohl morphologisch als auch molekular, sehr rasch evolviert. Aber, bisher war noch nicht vollständig geklärt, welche Zellen diese schnelle Evolution vorantrieben. In meiner Dissertation spreche ich diese Fragen an. Grundlage hierfür sind Daten mit Einzelzell Auflösung von 11 Spezies, welche die Hauptzweige der Säugetiere (Kloakentiere, Beuteltiere und Eutheria) und eine Vogelart. Unter den Säugerspezies finden sich Sieben Primatenarten, einschließlich aller rezenten Menschenaffen, außer Orangutan. Um auch das Verständnis der Evolution der weiblichen Keimbahn voranzutreiben, habe ich auch einen Transckriptionsdatensatz mit Einzelzell Auflösung von Eierstöcken in der Entwicklung dreier Nager und eines Vogels analysiert.

In diesen Analysen zeige ich, wie wir das Transkriptom testikulärer Zelltypen untersucht haben. Wir haben herausgefunden, dass die rasche Hodenevolution insbesondere durch postmeiotische Zelltypen vorangetrieben wird. Weiterhin haben wir gezeigt, dass diese rasche Evolution sowohl durch eine Entspannung negative Selektion, als auch durch einen Anstieg positiver Selektion getrieben wird. Zusätzlich haben weitere Hinweise für die „out oft he testis“-Hypothese gefunden, nach welcher der Hoden eine wichtige Rolle im Ursprung neuer Gene spielt.

Durch die Untersuchung der Expressionsdynamik einzelner Gene haben wir sowohl geteilte, als auch abstammungslinienspezifische Aspekte des Transkriptoms gefunden. Bei dieser Analyse haben wir einen Kernsatz an Genen identifiziert, welcher über die Säugetierevolution hinweg, sein Expressionsmuster beibehalten hat und wahrscheinlich mit zentralen Hodenfunktionen und weiterhin mit dem Aufrechterhalt der Fortpflanzungsfähigkeit

verbunden werden kann. In der Analyse abstammungslinienspezifischer Muster haben wir auch Gene gefunden, welche für die Fortpflanzung relevant sind. Bei der Untersuchung der funktionellen Auswirkungen einzelner Gene haben wir auch konservierte Aspekte der Sertolizell zu Keimzell Kommunikation gefunden.

Durch fokusierte Analyse der Sexchromosomexpression haben wir eine Anhäufung testisspezifischer X-chromosomaler Gene in Spermatogonien gefunden. Weiterhin haben wir die transkriptionellen Unterschiede zwischen X- und Y-tragenden Spermatiden untersucht. Zusätzlich haben wir die Inaktivierung meiotischer Sexchromosomen (MSCI) untersucht und zum ersten Mal transkriptionelle Belege für diese in Schnabeltieren gezeigt.

In der Analyse des sich entwickelnden Eierstockdatensatz, welchen ich generiert habe, habe ich die basalen somatischen Zelltypen identifiziert. In der Analyse der Keimzellen konnte ich in den Säugerdaten Keimzellsubtypen nicht klar differenzieren. In den Vogeldaten konnte ich Keimzellsubtypen unterscheiden.

Durch die Analyse der Zelltyptranskriptome dieses Datensatzes habe ich herausgefunden, dass Keimzellen schneller divergieren als somatische Zellen.

In der gezielten Analyse der Vogelkeimzelltranskriptome, mit dem Fokus auf Expression der Sexchromosomen, habe ich Hinweise auf das Vorhandensein von MSCI in Vögeln gefunden.

In the dedicated analysis of the bird germ line transcriptomes in regards to sex chromosome expression, I found support for the presence of MSCI in birds.

Zusammengefasst hilft mein Werk bei der Erklärung der beobachteten Diversität testikulärer Phänotypen und hilft bei der Erklärung der molekularen Quellen hierfür. Für den Eierstock liefern meine Daten einen aufregenden Einstiegspunkt für die weitere Erkundung.

Acknowledgments

There are many people whom I have to thank. Your support let me finally reach this point.

First, I want to say thank you to my supervisor Prof. Dr. Henrik Kaessmann. Thank you for giving me the opportunity to work with you and the many outstanding people you have gathered in the Kaessmann lab and I want to express my appreciation of your patient leadership style.

For his role in the testis project and as a colleague I also want to thank Dr. Florent Murat. Your work in the project made it what it is today. Together we not only had time for scientific discussions, but also for some fun, when the silly mood arrived in the office.

I want to also thank the many more colleagues who helped me during my PhD work. Dr. Mari Sepp, who guided me during my initial wet lab phase. Julia Schmid and Céline Schneider for the support in all the wet lab tasks. Nils Trost for the many impromptu troubleshooting session. Amir Fallahshahroudi for the help in chicken wrangling and input on ovaries.

Special thanks to my thesis advisory committee, Prof. Duncan Odom and Dr. Margarida Cardoso-Moreira who helped guide my thesis work with their support and critical comments. I also want to thank Dr. Ralph Grand Dr Anne Schlaitz for also being part of my examination committee. And thank you to all four members of my examination committee for enduring my administrative prowess.

Thank you to all the present and past members of the Kaessmann lab. Especially Francisca, Nils, Bassi, and Leti the people who started their path in the Kaessmann lab around the same time as me and with whom I by now have built a wealth of fond memories. Thank you to all the players in our D&D campaign, Evgeny, Matthias, Tetsuya, Daniel, and various guests.

I want to thank my mother for always being positive and encouraging. I also want to thank my uncle Werner and my siblings Issa, Mariama, and Djibril for the constant support you have given me throughout the years.

I am deeply thankful for my partner Elena. We have been side by side on our way to become a PhD since the early days of our Bachelors studies. Your presence is a warm embrace of love and support.

Thank you all.

Table of content

1 Introduction

In sexually reproducing organisms genetic information is transmitted to the next generation through specialized reproductive cells known as gametes. The gametes, eggs in females and sperm in males, are produced in the gonads and are formed in the germline[1]. The transfer of genetic information is achieved, when male and female gamete unite to form the zygote, which develops into an organism of the next generation. As such the germline plays a pivotal role in shaping the evolution of a species[2].

1.1 Gonads

The gonads - testes for males and ovaries for females - are different between the sexes[1]. However, the core functions of the tissue remain similar. The gonads produce gametes through gametogenesis and secrete steroid hormones[3,4]. As such the gonads are comprised both of germ cells and somatic cells. Additionally, the gonads share a common developmental origin, with differences emerging only after sex differentiation[1,5-8].

1.2 The somatic cells

1.2.1 Developmental origin of the bipotential gonad

The gonadal primordia are located on the ventral surface of the mesonephros, on both sides of the dorsal mesentery[6,8]. In mice, they arise from the coelomic epithelium (CE). The proliferating CE cells are marked by an expression signature of *GATA4*/*SF1*/*WT1*[6,9-14]. Prior to sex determination the bipotential gonadal primordia appear as long, narrow structures[12]. During the proliferation of the CE cells, their underlying basement membrane becomes fragmented, allowing introgression of CE cells into the mesenchyme of the mesonephros[6,15-17]. The epithelial-mesenchymal transition of the migrating CE cells is crucial for proper gonad formation, as the migrating CE cells contribute to the development of the primary sex cords[6,8,15,17,18]. At this stage, the gonad is still bipotential and transcriptionally does not yet show sexual dimorphism[8].

1.2.2 Sry driven male fate

In mammals, sex determination is dependent on the sex chromosomes. Individuals with an XY karyotype develop male gonads, the testes, while individuals with an XX karyotype develop female gonads, the ovaries. In therian mammals, the Y chromosome contains the *sex determining region of Y* (*SRY*) gene[19-21].

SRY is expressed only transiently around the time of male fate commitment[22-24]. Expression of SRY starts at the center of the gonad and then extends to both the anterior and posterior ends[6,22,23]. Expression of the high mobility group transcription factor SRY in turn leads to expression of *SRY-box transcription factor 9* (*SOX9)*, which is required for the development of the testis[6,8,25,26]. The activation of *SRY* and *SOX9* leads to an activation of male-specific genes and the downregulation of genes, which have been upregulated during commitment to the supporting cell fate[27]. Supporting cell precursors in the primary sex cords then rapidly commit to a Sertoli cell fate[28]. Through the aggregation of Sertoli cells and germ cells these cords develop into the seminiferous tubule[28].

The CE cells continue to proliferate, but the cells proliferating in this second phase exclusively differentiate into interstitial somatic cell types, not Sertoli cells[29]. Among the interstitial compartment are the steroidogenic progenitors, which will commit to become the steroidogenic Leydig cells[10,29].

By the time mammalian males reach reproductive age, the testis is encapsulated with a dense connective tissue layer[30]. Inside the capsule, the parenchyma of the testis is formed by the cord like seminiferous tubules[31,32]. Together, these tubules empty into the *rete testis*[33]. The tubules themselves are enveloped by contractile myoid cells, which form *the tunica propria*[34]. The basal membrane of the tubules is lined by an epithelium consisting of germ cells and the large mitotically inactive Sertoli cells[35]. The role of the Sertoli cells is to nurture the germ cells throughout testicular gametogenesis, which is called spermatogenesis[35]. Furthermore, the Sertoli cells form the blood-testis barrier, which separates the Sertoli cells into basal and adluminal compartments[36]. During spermatogenesis the germ cells traverse from the basal membrane into the lumen of the seminiferous tubule[37] (Figure 1).

The androgens in the testis are produced by the interstitial Leydig cells in response to Luteinizing Hormone[38]. The interstitial space also contains various other somatic cell types, including vasculature, neural, and immune cells[39].

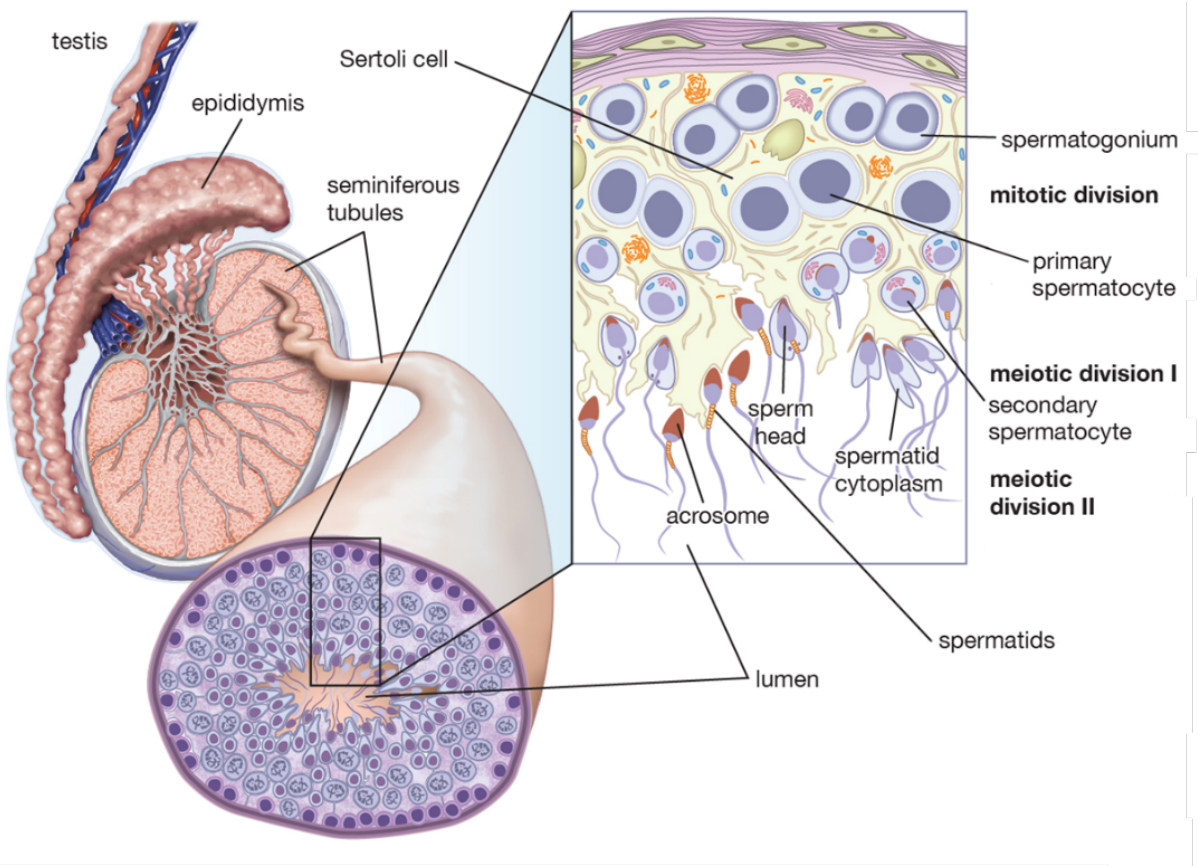


Figure 1: Testis structure. *The parenchyma of testis is formed by seminiferous tubules. While nurtured by Sertoli cells, germ cells progress through spermatogenesis within the seminiferous tubules. Adapted from Encyclopedia Brittanica 2013.*

1.2.3 Somatic cells committing to female fate

The common progenitors of supporting cells in male and female are marked by a signature of *R-spondin1*/*Wnt family member 4*/*catenin beta 1* (*RSPO1*/*WNT4*/*CTNNB1*) signaling[9]. In the developing female gonad, this signature is maintained and, by the time the gonad commits to the female fate, it promotes the expression of pro-ovarian genes while simultaneously antagonizing the expression of *SOX9* and *fibroblast growth factor 9* (*FGF9*), which are transcription factors driving masculinization[10,27,40,41]. As ovarian differentiation progresses, further genes with female specificity are upregulated. Among these is *forkhead box L2* (*FOXL2*), which in mice is needed for the maintenance of the granulosa cell fate after birth[42-44]. In mice, the knockout of *Foxl2* does not have a sex-reversal phenotype prenatally and at birth, but is required for maintaining the granulosa cell fate postnatally[45-47]. In contrast, knockout of *Foxl2* in goats leads to XX female to male sex-reversal[48,49]. This variability in role of *FOXL2* as a driver for female fate determination, together with the maintenance of the *RSPO1*/*WNT4*/*CTNNB1* signature, lead to the longstanding understanding that the female gonad fate is the default state in development, while the *SRY* driven male fate is an active

genetic program[50]. However, newer studies on developmental transcriptomes of ovaries challenge this notion, as they show active upregulation of female specific gene expression programs[8,10,41].

The *FOXL2* positive granulosa precursors will form the medullary granulosa cells and together with germ cells the primary sex cords[51,52]. Later, there is a secondary wave of CE cells differentiating into cortical granulosa cell precursors[53-56]. These precursors, marked by the expression of *leucine rich repeat containing G protein-coupled receptor 5* (*LGR5*), associate with germ cells to form the secondary sex cords[52]. The supporting cell progenitors together with germ cells, form germ cell nests. These nests eventually break down to give rise to the primordial follicles, which in turn mature to become follicles[57-59]. Each follicle contains one germ cell, initially surrounded by granulosa cells[55,56].

During the reproductive cycle, groups of primordial follicles progress through folliculogenesis[3,60]. Depending on species, one or multiple follicles complete folliculogenesis, while the other follicles of the group are removed through atresia[61 62]. As folliculogenesis advances, secondary follicles associate with steroidogenic theca cells[3]. These theca cells arise from CE derived stromal cells[3](Figure 2).

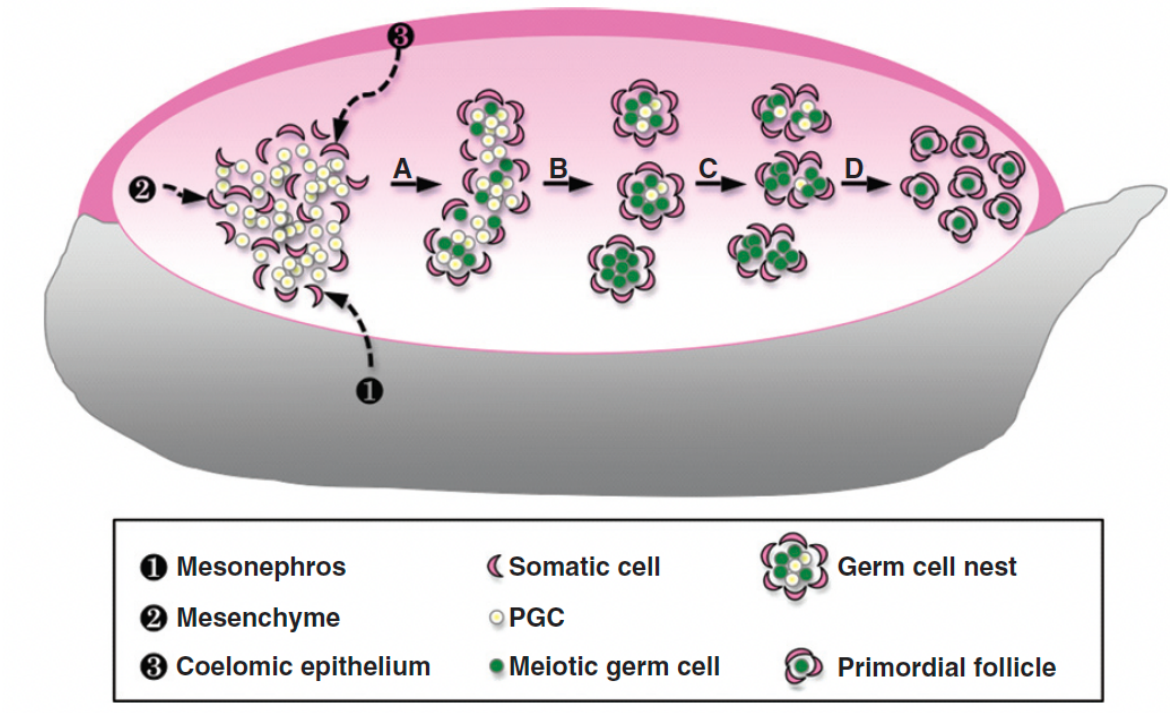


Figure 2: Formation of the mouse fetal ovary. *Somatic cells can originate from the mesonephros (1), mesenchyme (2), or the ceolomic epithelium (3). A-D They form germ cell nests, which break down around birth leading to follicle formation. Adapted from Liu et al. (2010)*

While supporting cells in male differentiate rapidly to form Sertoli cells in the well-structured sex cords, female supporting cells differentiate more slowly[8]. Furthermore, the process of folliculogenesis and the regionalization of cortex and medulla, vary among species[6,18,63].

1.3 The germline

Germ cells are the cells, which carry genetic information of the parent to the offspring, thus carrying this information across generations[64]. While germ cells differ between the sexes, they also share common features. In both males and females, germ cells have an extraembryonic origin as primordial germ cells (PGCs) before they colonize the gonads[65-67]. In the gonads they expand their number mitotically, before committing to meiosis[67,68].

The PGC precursors are proximal epiblast cells, which acquire the competence to become PGCs[65-67]. In mice, by embryonic day (E)7.25 approximately around 40 of those are further specified towards the PGC fate [67,69-71]. These colonize the gonadal primordia between E10 and E12.5, concurrently with the thickening of the primordia[67,68]. Through mitotic division, the number of germ cells increases significantly, reaching a total of around 25 000 germ cells by E13.5[57,72]. Early germ cells are marked by the expression of *deleted in azoospermia like* (*DAZL*) and a loss of pluripotency[73,74].

Similar to the divergence of female and male somatic cells the development of the fetal germ cells diverges between the sexes. One difference is at which developmental stage the germ cells commit to meiosis[7].

1.3.1 Meiosis

Germ cells committing to meiosis first progress prophase. The prophase of the first meiotic division is the longest phase of meiosis regardless of sex or species[37,75-78]. Based on the appearance of the chromosomes, this phase is further subdivided into five stages, leptotene, zygotene, pachytene, diplotene, and diakinesis[77,79,80]. During the leptotene the chromatin starts to condense, which makes it visible as thin threads[77,79,80]. During this stage the homologous chromosomes begin to align and the axial elements are assembled[81]. Furthermore, during this stage double strand breaks (DSBs) are induced through SPO11 initiator of meiotic double strand breaks (SPO11)[82-84]. These DSBs are needed as substrate for recombination[77,79,85]. The homologous recombination repair (HRR) machinery recognizes the breaks and initiates repairs[77,79,80,86].

During the zygotene, synapsis of the homologous chromosomes continues and the synaptonemal complex starts to form[81]. The DSBs are repaired by members of the HRR

machinery such as DNA meiotic recombinase 1 (DMC1)[87,88]. During this process at least one crossover has to be introduced per chromosome to physically tether the homologous chromosomes together[77].

During pachytene, synapsis of chromosomes through the synaptonemal complex is complete and remaining DSBs are resolved[77,79,80,86]. After pachytene, the cells transition to the diplotene stage in which the synaptonemal complex is disassembled and the homologous chromosomes are only connected through the chiasmata, where previously crossing over occurred (Figure 3)[77,79,80].

In diakinesis, the chromosomes condense further and the nuclear membrane starts to disintegrate[77,79,80]. The division at the end of meiosis I results in the segregation of the homologous chromosomes. During meiosis II the sister chromatids are segregated resulting in the haploid germ cells[77,79,80].

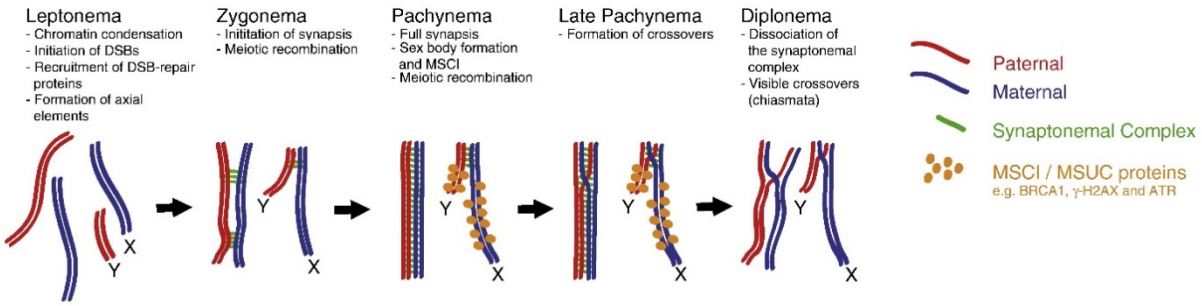


Figure 3: Schematic of prophase I stages. *The appearance of the chromosomes is used to separate prophase I into five distinct stages. DSBs: double strand breaks, MSCI: meiotic sex chromosome inactivation. Adapted from Jan et al., 2012.*

1.3.2 The female germline

In females, the cells formed by the mitotically active fetal germ cells are called oogonia. These cells express the retinoic acid (RA) receptor Stimulated by retinoic acid 8 (STRA8), which makes them responsive to retinoic acid produced in the mesonephros[89,90]. RA reception leads to a commitment of oogonia to meiosis in an anterior to posterior wave[91-93]. The dynamics of the wavelike commitment and thus degree to which asynchrony in the germ cells committing to meiosis can be observed, varies across species. For example, in humans, commitment is more asynchronous, occurring over several weeks during gestation (Figure 4)[94].

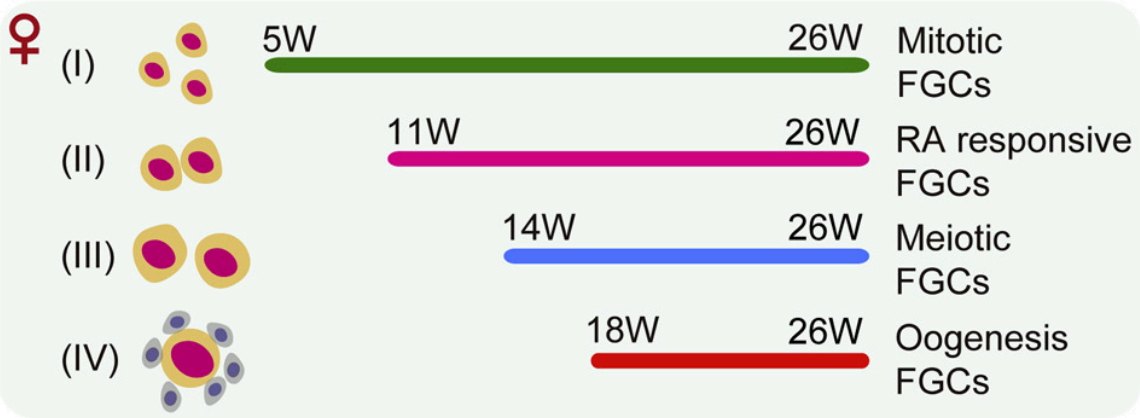


Figure 4: Human ovarian gametogenesis. *In the developing human ovary germ cell differentiation is initiated asynchronuously. Through this multiple germ cell stages are present at the same time. Adapted from Li et al. (2017)*

Prenatally, the oogonia committed to meiosis are called oocytes. However, prenatally they only progress until diplotene stage of prophase I. In this stage the oocytes become arrested in a stage called *dictyate*[78,95]. Together with the somatic supporting cells they form the primordial follicles as described earlier. The germ cells in these follicles form the ovarian reserve, which is not replenished by further mitotic divisions[96-99]. Interestingly, the number of germ cells is not only not increasing, but is actively reduced through apoptosis[96]. For example, a newborn human has around 1 million oocytes. This number is reduced to around 400,000 when puberty is reached[100].

After reaching reproductive age, clutches of follicles become activated per reproductive cycle. Central to this activation is the Phosphoinositide 3-kinase (PI3K) pathway[101-105]. Within the follicle the oocyte progresses to metaphase of meiosis II through an asymmetrical division, which results in the extrusion of the polar body[106,107]. At this stage, progression is halted again and meiosis II is only completed after fertilization by a spermatozoon[106,108].

1.3.3 The male germline

In males, the cells formed by the mitotically active fetal germ cells are called the pre-spermatogonia[7]. The supporting cells in the fetal male gonad prevent the germ cells from progressing towards meiosis, by degrading the RA in the gonad[89,109]. When reaching reproductive age, the germ cells become active again and new germ cells are produced through spermatogenesis[37,75,110].

Spermatogenesis can be divided into three main phases and the germ cells are named according to which phase of spermatogenesis they are in: (1) mitotic expansion of

spermatogonia; (2) meiotic divisions of spermatocytes; and (3) morphological differentiation of spermatids (also called spermiogenesis, Figure 1)[79,110]. This highly productive process can produce 40 million sperm per day per gram of testis in the mouse, while humans can produce 4 million sperm per day per gram of testis (Figure 5)[111]. In continuous breeders these high productive rates need to be maintained throughout the reproductive lifetime. In seasonal breeders, this productive rate has to be reestablished during the onset of the breeding seasons[112,113]. For both of these breeding strategies an active stem cell compartment among the germ cells is required.

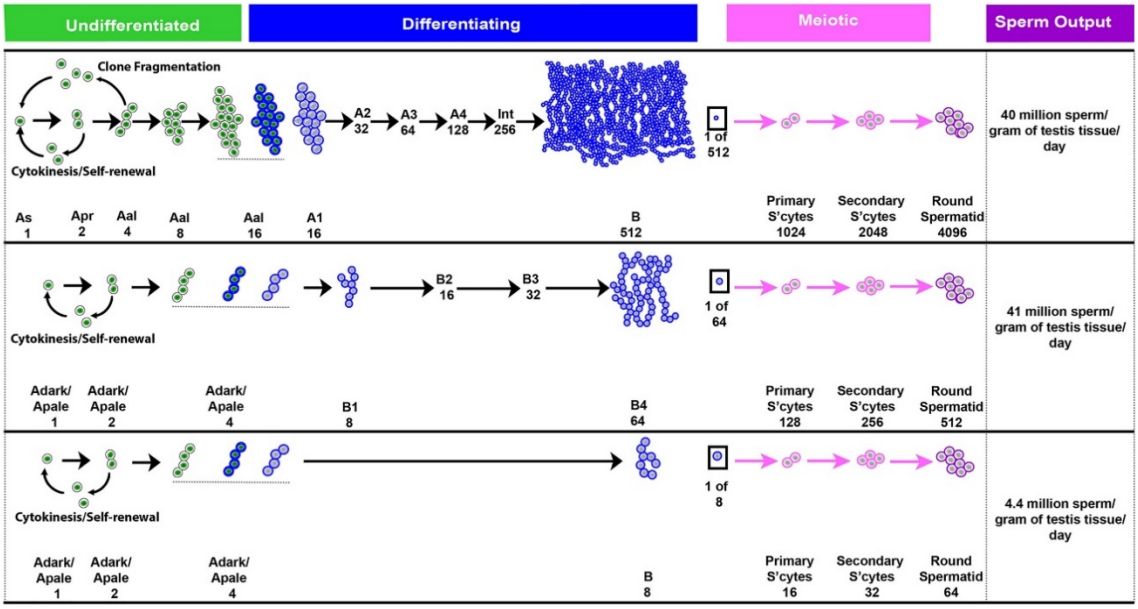


Figure 5: Comparison of spermatogonia. (A) *In rodents spermatogonia form chains, which fragment. Some of the cells maintain the stem cell pool, while others undergo multiple rounds of mitosis and commit towards meiosis.* ***(B)*** *In monkeys A_{dark} and A_{pale} spermatogonia maintain the stem cell pool, with the spermatogonia committing to meiosis undergoing a limited number of mitotic divisions.* ***(C)*** *In human there is the same A_{dark} A_{pale} system for maintaining the stem cells. The A spermatogonia undergo even fewer mitotic divisions than in monkey prior to meiosis. Fayomi and Orwig, 2018.*

Spermatogonial stem cells are among the spermatogonia, which through their mitotic divisions, prior to committing to meiosis, fuel the entire process of spermatogenesis with new cells[111,114]. The maintenance of the stem cell compartment among the spermatogonia is organized differently depending on the species[111,114-117]. In primates spermatogonia are classified either as A_{dark}, or A_{pale}[115]. Some models suggest the A_{dark} spermatogonia as the quiescent reserve stem cell compartment with A_{pale} spermatogonia as the active and quickly dividing cells[118]. Other models suggest A_{dark} and A_{pale} spermatogonia as the same cell type in different stages of the cell cycle[119]. In rodents, single A type spermatogonia form syncytial clones through partial divisions[111,117]. These divisions result in A_{paired} and $A_{aligned}$

spermatogonia, which have the appearance of spermatogonia chains. These chains fragment and among these a fraction of the cells retain their identity as A type spermatogonia, which maintain the stem cell compartment. The differences in the number of premeiotic divisions has a strong impact in the eventual output of sperm[111,117].

After mitotic divisions of the A type spermatogonia, the spermatogonia which commit to meiosis differentiate into B type spermatogonia, which then divide to become primary spermatocytes[79].

Spermatocytes are the testicular cell type, which undergoes meiosis[79]. During meiosis, homologous chromosomes align in order for crossing overs to appear[77]. In mammals, males have heteromorphic sex chromosomes[120,121], which have only small regions of homology, known as pseudoautosomal regions (PARs), and large regions which are not homologous, known as sexually differentiated regions (SDRs)[122-124]. The lack of homology between X and Y prevents them from aligning in zygotene, which results in the process of meiotic sex chromosome inactivation (MSCI)[125-127]. During MSCI the X and Y chromosome become rapidly silenced and sequestered in the sex body[127]. This process is mediated by BRCA1 DNA repair associated (BRCA1) recruiting ATR serine/threonine kinase (ATR) specifically on the unsynapsed sex chromosomes[128,129]. ATR then phosphorylates histone H2AX in the chromatin of the X and Y chromosomes, through which the silencing is mediated[83,130,131]. The H2AX phosphorylation is accompanied by further post-translational modification such as ubiquitylation of H2A[132,133]. After meiosis is completed the silencing of the sex chromosomes is partially reversed and they become transcriptionally active again[134]. Interestingly, the epigenetic hallmarks of MSCI have not been observed in monotremes[135].

Unlike in female, meiosis progresses without an arrest phase[79]. Through meiosis one diploid preleptotene spermatocyte divides and becomes four haploid spermatids[77].

The spermatids undergo morphological changes from round cells to the typical sperm shape. This process is also called spermiogenesis[136]. The interspecies differences in sperm morphology, which make sperm one of the most diverse cell types in animals, become readily apparent as a result of these changes[137]. In mammals alone we can observe diverse sperm shapes like the human round headed spermatids, the hooked head for mice, and sperm pairs of the opossum[37,138,139].

The differentiation process results in three distinct regions in the spermatid. The head, midpiece, and tail[79]. Development of the sperm tail starts from the centriole at one cell pole

and is composed of a microtubular structure called the axoneme[140,141]. The nucleus becomes elongated, by the formation of the manchette[142,143]. Further the reshaping of the nucleus is facilitated by the exchange of histones into protamines[144]. The protamines allow DNA to be packed tightly in a paracrystalline state[145]. The transition through this DNA condensation is enabled by transition proteins[146-149]. Spermatids are further compacted through the removal of cytoplasm[140,150]. Another important feature for fertility is the development of the acrosome, which contains hydrolytic enzymes required for oocyte penetration[151,152].

After the spermatid has finished the differentiation from round spermatid to elongated spermatid, it enters the lumen of the seminiferous tubule. From there it exits the testis through the rete testis and matures in the epididymis into a mature spermatozoon[150].

1.3.4 Coordination of spermatogenesis

The process of spermatogenesis is highly coordinated. At a specific region of the seminiferous epithelium spermatogonia simultaneously become committed to progress in spermatogenesis mediated by RA[73,153]. This commitment happens in a species-specific interval, e.g. 16 days in humans[76]. As the duration of spermatogenesis is longer (74 days in human) than the initiation interval and the tight temporal regulation of progression through spermatogenesis, specific germ cell associations can be observed (Figure 6)[37,75,76,153,154]. These associations are used as the basis to define stages through which the seminiferous tubules cycle. In mice for example 12 stages have been described[37]. Adjacent segments along the linear seminiferous tubules are in adjacent stages of the seminiferous cycle. This results in the wave of the seminiferous tubule[31,154]. This asynchrony of seminiferous cycle along the tubule is needed for the maintenance of continuous sperm production[37]. While a given seminiferous cycle stage can be observed for an entire cross-section of a mouse seminiferous tubule, in primates multiple stages are observed for one cross-section and the wave appears to be organized in a helical manner along the tubule[37,155].

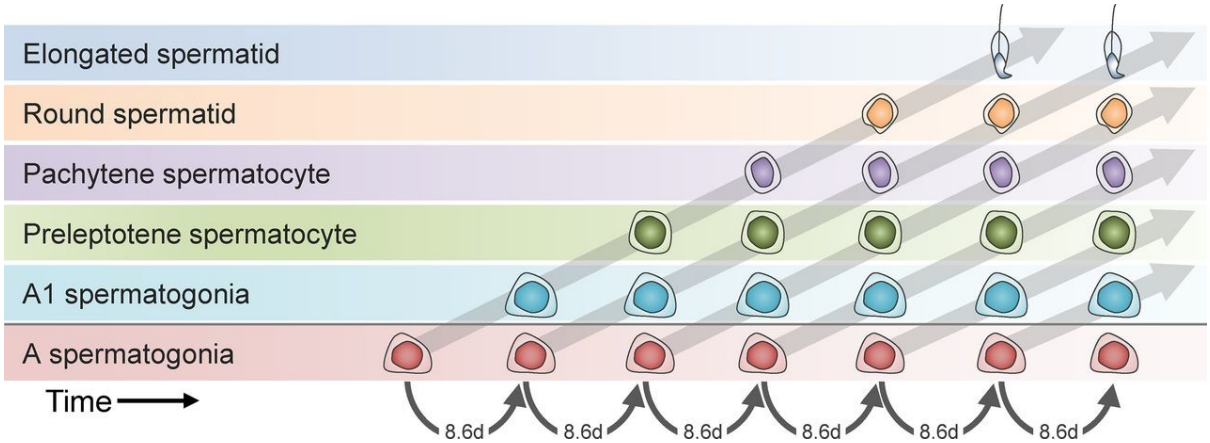


Figure 6: Germ cell associations. *In mouse every 8.6 days a new seminiferous cycle is initiated. This leads to the appearance of specific germ cell combinations in cross sections of seminiferous tubules. Adapted from Griswold (2015)*

1.4 Testicular evolution

The testis is a rapidly evolving tissue, which can be observed as a large variation of relative testis size, sperm production rates, sperm parameters, sperm motility, and sperm morphology between species[156-162]. An explanation for this can be found in the strong selective force of sperm competition, which is an evolutionary pressure on males to achieve reproductive success[160-165]. An example for the effects of this can be the influence of mating system on the relative testis size in primates (Figure 7)[161]. In primate species, in which the females mate with multiple males, such as chimpanzees, the relative size of the testis compared to the body increases (~0.27 % of the body mass)[166]. In primate species, in which the females mate infrequently and with few, or only one male, such as gorillas, the relative testis size decreases (~0.05 % of the body mass)[166].

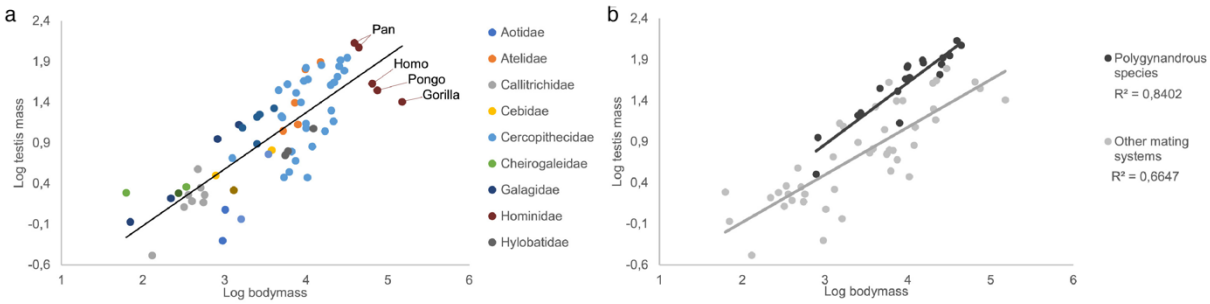


Figure 7: Testicular size changes in primates. a) *correlation of testis size and bodymass for primates.* ***b)*** *correlation of testis size and body mass for polyandrous species. Adapted from Martinez and Garcia (2020).*

The basis for phenotypic changes, are changes in the underlying genotype. There are two major classes of mutations as a source for the phenotypic changes[167]. 1) Mutations of a genes coding sequence. This change in turn then changes the gene products function[168]. 2) Mutations in regulatory sequences. These can affect in which tissue at which developmental

times a transcript becomes expressed and affect transcript abundance[169,170]. Further the regulatory change could be an impact on post translational transcript dynamics[171,172].

While initially it was not clear, how large the impact of either class of mutation was on phenotypic evolution the advent of high-throughput RNA-sequencing technologies, was instrumental in dissecting the molecular source[170,173-176]. In mammals these endeavors were driven by bulk tissue RNA-seq data for different tissues and developmental time points[157,158,177-181].

Through this data, it became apparent that, while there are cases in which the change in coding sequence became fixed resulting in an associated functional change, the majority of phenotypic changes can be attributed to a continuous accumulation of regulatory changes[167]. Furthermore, the data show that the observed rapid phenotypic evolution of the testis is reflected at the molecular level, with testis showing the highest rate of gene expression changes across mammalian organs[157,158]. In addition, testis expressed genes and genes associated with spermatogenic functions show increased signs of positive selection[182-185].

These data also show another particularity of testicular transcriptomes, which had already been observed in microarray data. The testicular transcriptome is especially diverse and transcripts for most genes can be found[186-188]. Based on cell type enriched data, the majority of this promiscuous transcription can be attributed to pachytene spermatocytes and round spermatids[188]. The “out of the testis” hypothesis, according to which new genes become initially transcribed in the testis, is based on these observations[189]. This promiscuous transcription is facilitated by the extensive remodeling of chromatin through histone exchanges and the permissive chromatin state that has been shown for the testis (Figure 8)[144,188].

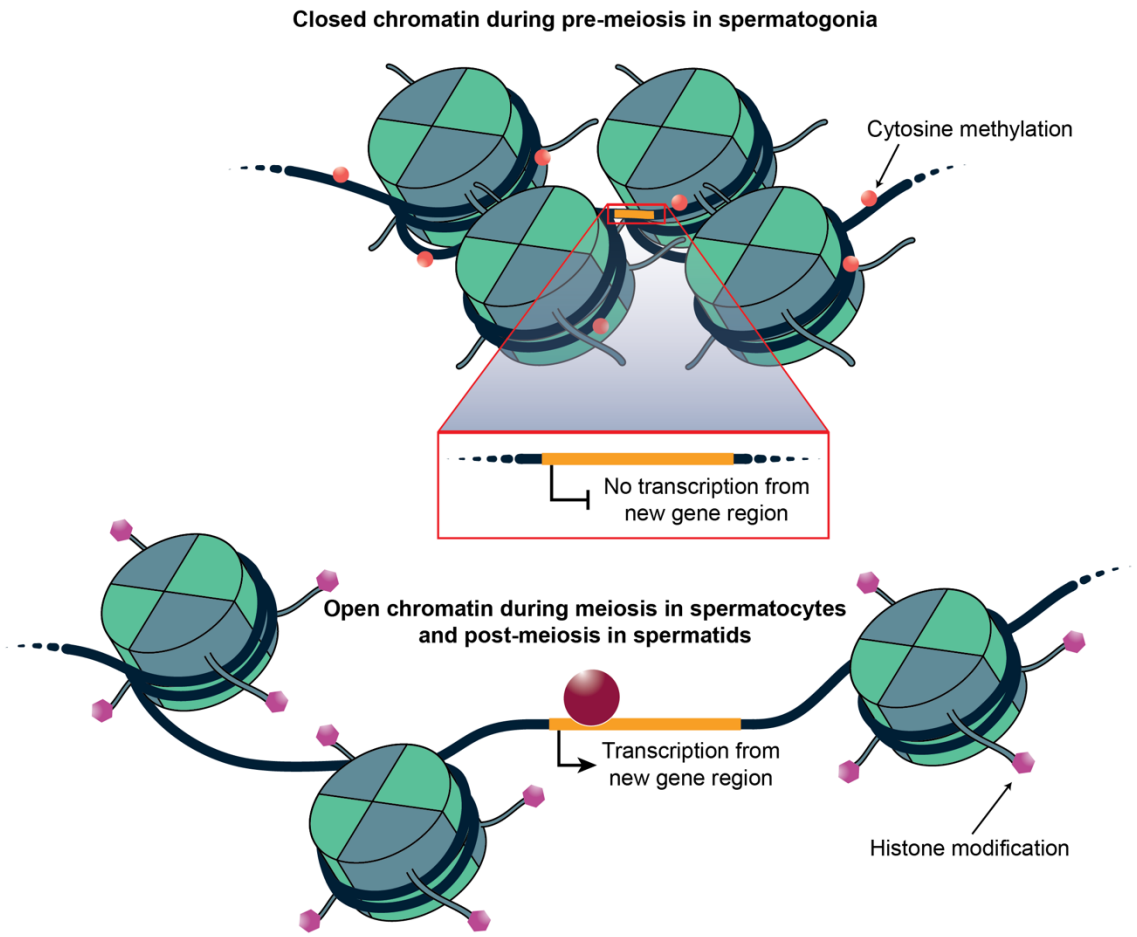


Figure 8: Permissible chromatin. *The chromatin in postmeiotic germ cells is in a state permissible for transcription. Adapted from Trost et al., 2023.*

The phenotypic impact of these diverse transcripts is buffered on the translational layer, as testis expressed genes, especially those expressed in later stages of spermatogenesis, show a lower translational efficiency, as shown by Ribo-seq data[159]. Despite this translational buffering, the testicular translatome also diverges faster than the translatomes for other tissues (Figure 9)[159,167].

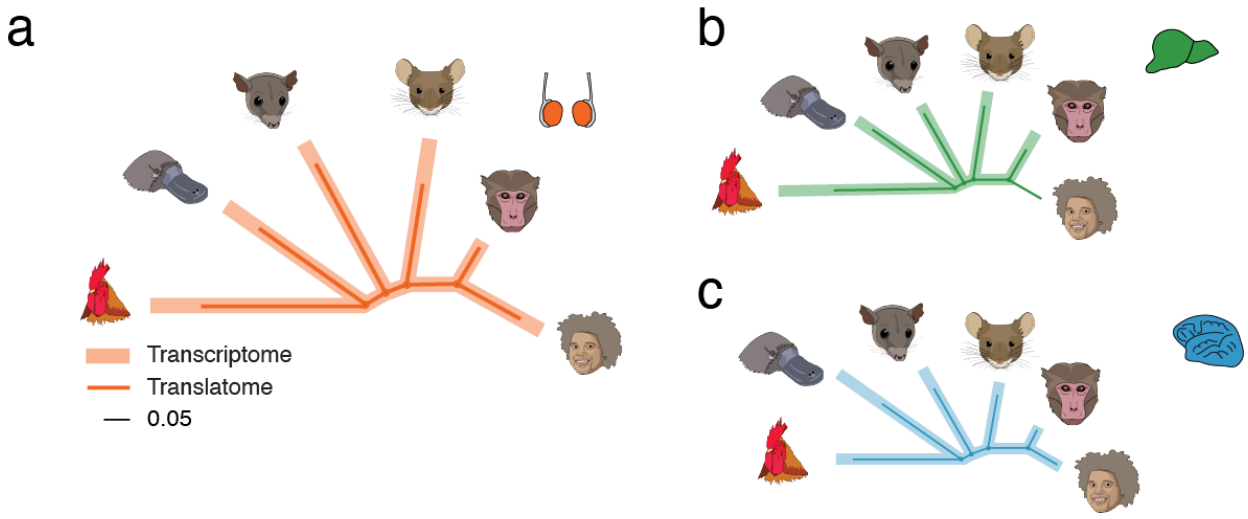


Figure 9: Multilayer evolution of gene expression. *Phylogenetic analyses of the translatome (thin and dark) and transcriptome (thick and light) for* ***a)*** *Testis,* ***b)*** *Lliver, and* ***c)*** *Brain. The branch lengths indicate evolutionary changes in expression levels. Adapted from Trost et al. (2023) based on data from Wang et al. (2020)*

2 Aim of the research

The testis is a rapidly evolving tissue. Yet, the cellular sources and how which cell types contribute to these observations in the adult mammalian testis has not been comprehensively studied. Furthermore, comparative work studying the transcriptome evolution of female gametogenesis at single cell resolution is limited[190]. Given the importance of understanding these processes for both evolution and their impact on health, my work aims at closing these knowledge gaps. Specifically, the aims are as follows:

- **Generate a comprehensive transcriptome atlas of the adult mammalian testis:** As the snRNA-seq atlas of the adult mammalian testis is the basis for the testicular analyses, the first aim is to ensure the generated data is of sufficient quality and assess, which cell types are reflected in the data. Furthermore, this atlas provides a resource for understanding testicular transcriptomes across the main mammalian lineages.
- **Understand the cellular source for the rapid testicular evolution:** The testis is a rapidly evolving tissue[157-162]. To dissect, which of the testicular cell types contribute how to this observation, we explore transcriptomic divergence and assess selective pressures shaping testicular evolution.
- **Identify core gene expression programs of the testis:** Despite the rapid evolution, the testis has conserved functions in reproduction and as a gland. To understand this, we assess which genes share expression patterns across the sampled species and evaluate their functional roles.
- **Identify lineage-specific innovations of the testicular transcriptome:** Males are in competition for reproductive success[162]. To understand, which changes occurred on specific lineages, we assess lineage specific changes in gene expression patterns. Furthermore, we assess the expression of lineage specific genes.
- **Probe the dynamics of sex chromosome expression during spermatogenesis:** During spermatogenesis the sex chromosomes become transcriptionally repressed through MSCI[125,126]. Furthermore, as male mammals are the hemizygous sex, the expression of sex-linked genes is of special interest[191]. To understand the dynamics of sex chromosome expression, we dissect the expression patterns of sex-linked genes.
- **Generate a transcriptome atlas of the developing mammalian ovary:** The ovary plays a pivotal role in reproduction. Seeing that the ovarian reserve is established in development, a developmental atlas of the mammalian ovary is of interest. For this I annotate snRNA-seq libraries of developing glires ovaries.

- **Understand the cellular particularities of female developing germ cells:** The germ cells are the cell types, which convey the reproductive functions of the gonads. As such, I explore their transcriptomes.

Overall, the aim of my research is to further our understanding of the molecular evolution of both male and female gametogenesis.

3 Results

3.1 Adult testis transcriptome across eleven species

The project regarding the transcriptome of the adult mammalian testis was a collaborative project. My main contribution to this project was establishing the experimental protocol and generation of the data set, while the bioinformatic analysis was conducted by Dr. Florent Murat. It is important to note that while the code for the bioinformatic analysis was implemented by Dr. Murat, I contributed during this stage by discussing, which analyses to perform, assessing intermediate results to direct follow up analyses, and selection of the genes exemplifying key findings. My contribution to the success of this project is reflected in the shared first-authorship of "The molecular evolution of spermatogenesis across mammals" published in Nature[192]. Based on this, I will be using "we" when describing the results of this project, which comprises of Dr. Murat and me.

The adult testis is a complex organ responsible for the continuous generation of gametes and the secretion of androgens. Beyond this, it is of particular interest to evolutionary biologists, because of its pivotal role in reproduction and the rapid testicular evolution[156-162]. Because of these attributes, we wanted to investigate the transcriptomes of testicular cell types. With these we aimed to dissect, which cell types contribute how to these observations. For this, I generated single nuclei RNA-sequencing data (snRNA-seq) of adult testis samples for eleven species (Figure 10).

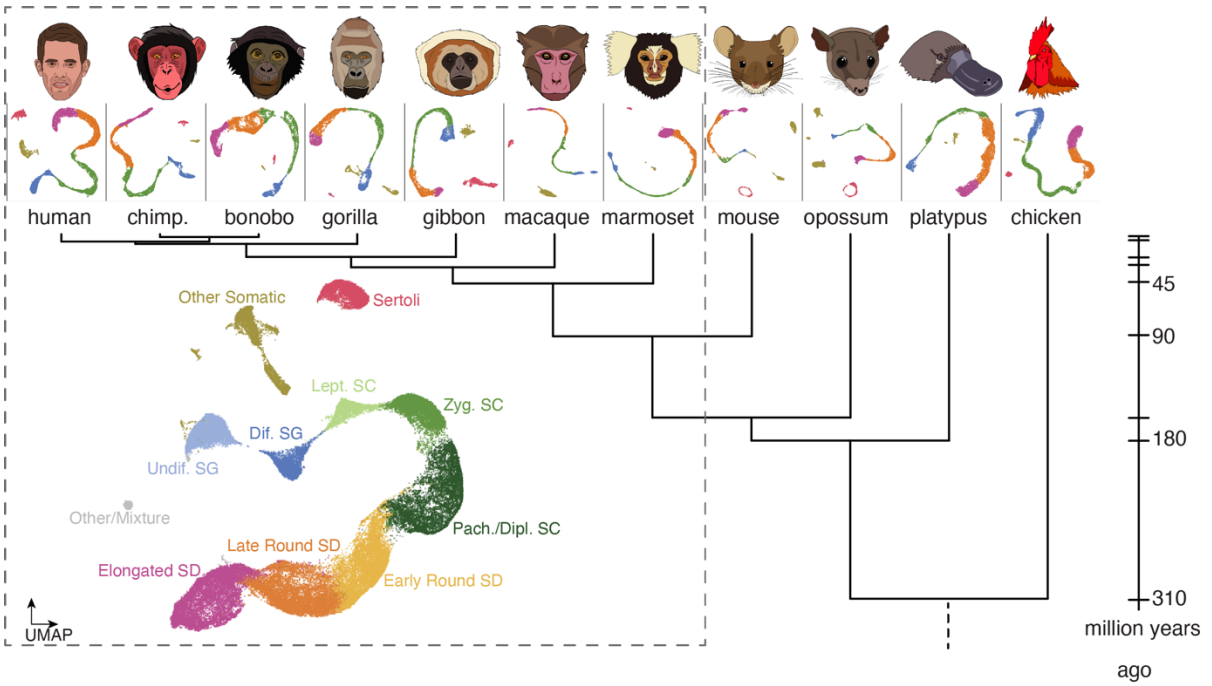


Figure 10: Species phylogeny. *The phylogeny of species from which I used samples to generate snRNA-seq data of the adult testis and the UMAP projection. Dashed box shows integrated primate dataset. Shown are undifferentiated and differentiated spermatogonia (undif. SG and dif. SG, respectively), leptotene, zygotene, pachytene and diplotene spermatocytes (lept. SC, zyg. SC, pach. SC and dipl. SC, respectively), spermatids (SD) and somatic cell types. Adapted from Murat et al. (2023).*

The transcriptomic data I generated cover key species for simian (anthropoid) evolution, including four of the five extant great ape species (human, chimpanzee, bonobo, gorilla) as well as representatives from apes (gibbon), cercopithecidae (macaque), ceboidea (marmoset; in this text I avoid the nomenclature of old and new world monkey on purpose). Furthermore, the three main mammalian lineages of eutherians (simians and mouse), marsupials (opossum), and monotremes (platypus) are covered. The data generated for red jungle fowl (hereafter referred to as chicken) was used as an evolutionary outgroup. In total, I generated 25 snRNA-seq libraries for these species, in addition two libraries from adult human testis samples have previously been prepared in the lab. Overall, the resulting dataset is comprised of one to three biological replicates and at least two technical replicates per species. The number of biological replicates was restricted, depending on sample availability.

Given the challenge of obtaining fresh samples from such diverse species, I worked with deep frozen tissue samples, as these were the only viable option. While fresh samples were not feasible for this study, it is important to note that deep-freezing compromises cell membrane integrity, rendering the isolation of intact single cells, required for single cell RNA-seq difficult[193]. Despite this, intact nuclei can still be isolated from deep-frozen tissue. While the

nuclei only contain RNA transcripts, which have been retained in the nucleus, like the long non-coding RNA MALAT1, or have been freshly transcribed and not yet been exported from the nucleus, the nuclear transcriptome has been shown to be a good proxy for the cellular transcriptome [194,195]. To gain single cell resolution transcriptome data for the adult testicular transcriptome, I established protocols for extraction of intact nuclei, which I used as input for library generation. Thus, I generated snRNA-seq data instead of single cell RNA-seq data.
Within these libraries we obtained 97,521 nuclei, which passed the filtering process. These nuclei have a median number of 1,338 genes expressed and a median number of 1,856 unique molecular identifiers (UMIs) per nucleus. The number of UMIs corresponds to unique RNA molecules detected in a nucleus.

3.1.1 Cell type assignment

For the analysis, we first projected the data into low dimensional space using Uniform Manifold Approximation and Projection (UMAP), a technique for visualizing high-dimensional data by reducing its dimensionality while preserving the structure of the data[196]. The resulting UMAP plots across the eleven species share the attribute of having a large continuity of cells with several distinct clusters separate from this continuity (Figure 11). To assign cell types to the nuclei, we used genes which have previously been described as expressed in specific cell types. Together, based on the expression of the literature cell type marker genes and the coordinates in the low dimensional projection, we found that the cell type continuity reflects the unidirectional differentiation process of spermatogenesis, while the separate clusters correspond to the somatic cell types (Figures 10, 11).

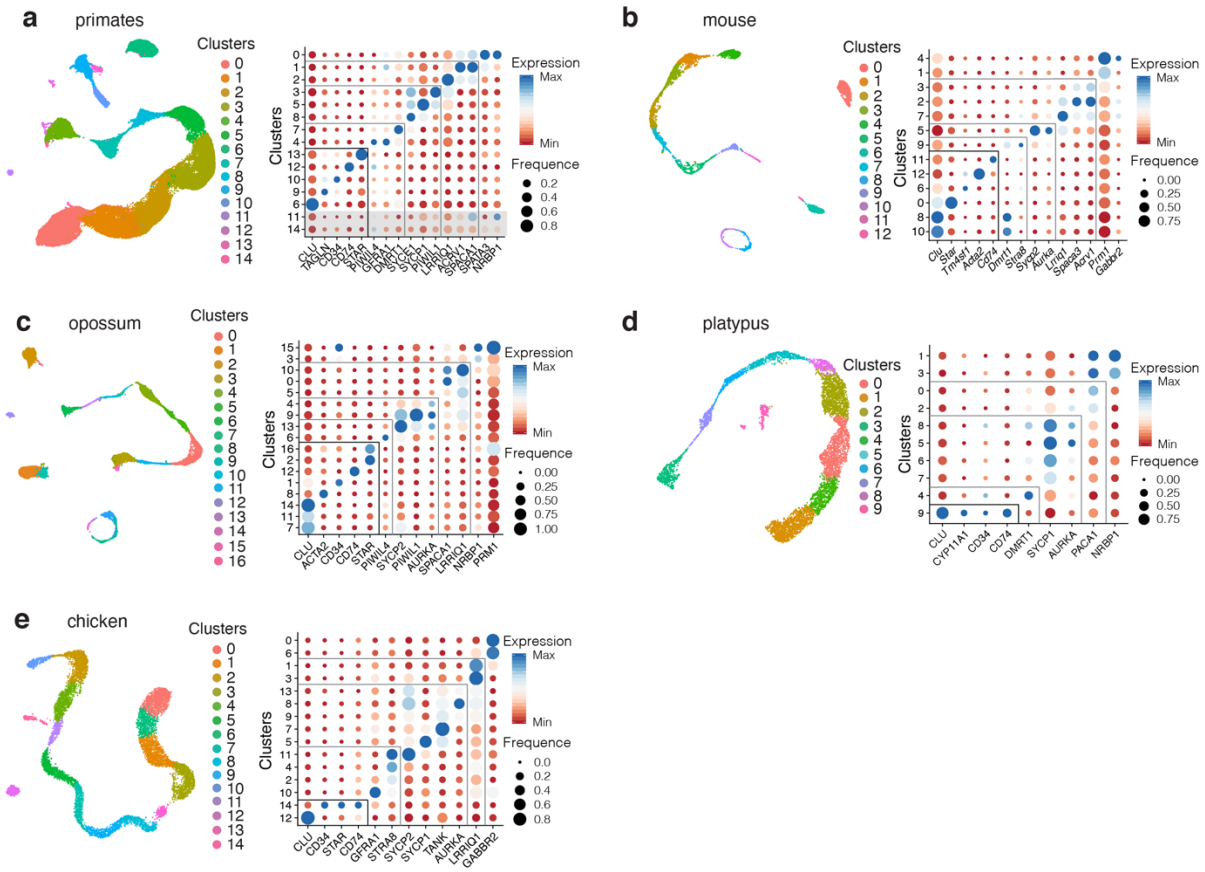


Figure 11: Expression of marker genes. a-e*) We assigned the cell type based on the expression of known marker genes. Adapted from Murat et al. (2023).*

The earliest cells in spermatogenesis, spermatogonia, are marked by the expression of genes such as *GDNF family receptor alpha 1* (*GFRA1*), *Piwi like RNA-mediated gene silencing 4* (*PIWIL4*), *Doublesex and mab-3 related transcription factor 1* (*DMRT1*), and *STRA8*[153,197-202]. We found the spermatogonia at one end of the projected cell continuity. The stem cell compartment fueling spermatogenesis is found among the spermatogonia [111,117].

The spermatogonia committing to meiosis first differentiate into B type spermatogonia and then into the meiotic spermatocytes. The spermatocytes are marked by the expression of synaptonemal complex components such as *SYCP1* and *SYCP2*, *PIWIL1*, *TANK*, or *AURKA*[202-205]. As the spermatocytes are the cells undergoing meiosis, we also found the expression of genes introducing double strand breaks such as *SPO11*, as well as genes required for the homologous recombination repair machinery such as *DMC1*[87,88,206]. In the UMAP plot we found the spermatocytes located directly after the spermatogonia on the cell continuity. After the two meiotic divisions each spermatocyte gives rise to four round spermatids.

The spermatids undergo morphological changes from round to elongated with a clear head and tail in the process called spermiogenesis [136]. The spermatids are marked by the expression

of protamines like *Protamine 1* (*PRM1*) required for the compaction of the nuclei, *SPATA3*, *NRBP1*, and *GABBR2*[145,207,208]. In the spermatids we further found expression of genes required for the function of sperm in fertilization such as *ACRV1*, a component of the acrosome[209]. In the UMAP plot the spermatids are located at the end of the cell continuity.

The separate clusters in the low dimensional projections correspond to the somatic cell types of the testis. We identified them based on expression of *Clusterin* (*CLU*) in Sertoli cells, which we unambiguously identify in all species except for platypus[210]. For the classification of the other somatic cell types, we used various markers. Expression of *Steroidogenic Acute Regulatory Protein* (*STAR*) and *Cholesterol side-chain cleavage enzyme* (*CYP11A1*) as markers for Leydig cells. Expression of *Transgelin* (*TAGLN*) as marker for peritubular muscle cells. Expression of *Actin Alpha 2* (*ACTA2*) as marker for smooth muscle cells. Expression of *CD34 and Transmembrane 4 L6 family member 1* (*TM4SF1*) as markers for endothelial cells. Expression of *Apolipoprotein E* (*APOE*) and *CD74* as markers for macrophages[211-218]. However, we did not as unambiguously identify the other somatic cell types, as the Sertoli cells. Because of this, we classified the remaining somatic cells together as “other somatic”.

The oldest divergence among simians for which we generated data was around 45 million years ago[219]. We leveraged this close evolutionary relationship, to integrate the data across the simian datasets. Based on the increased number of nuclei and the additional information they afford, we refined the assignment of germ cell types and identified sub-cell types corresponding to intermediate cell states during spermatogenesis (Figure 10 dashed box).

We distinguished undifferentiated from differentiated spermatogonia based on *GFRA1* and *DMRT1* expression respectively[197,198].

Among the spermatocytes we distinguished the stages of meiotic prophase I. For this, we used a combination of both the expression of marker genes, as well as the relative levels of gene expression during the different stages. As markers for spermatocytes in leptotene we used *SYCE1* and *SPO11*; as marker for zygotene spermatocytes we used *SYCP1*; and pachytene/diplotene spermatocytes are marked by *PIWIL1* expression[79,203,206]. We did not capture nuclei in diakinesis as during this stage the nuclear membrane is dissolved in preparation of the cellular division.

After assigning cell types, we identified which genes are expressed differentially between the cell types. In total we identified 133,222 genes differentially expressed across all species and cell types. This overall represents all the cell type marker genes we could identify in our

dataset. Among these genes are well established cell type marker genes, which we have also used in the cell type assignment. Further we find a host of novel cell type markers for all testicular cell types, which have previously not been identified as testicular cell type markers, or not been verified in all species assessed in this project. Fitting with a continuous differentiation process, the expression of marker genes during spermatogenesis shows a gradual change of cell type markers and no clear cell type boundary (Figure 12).

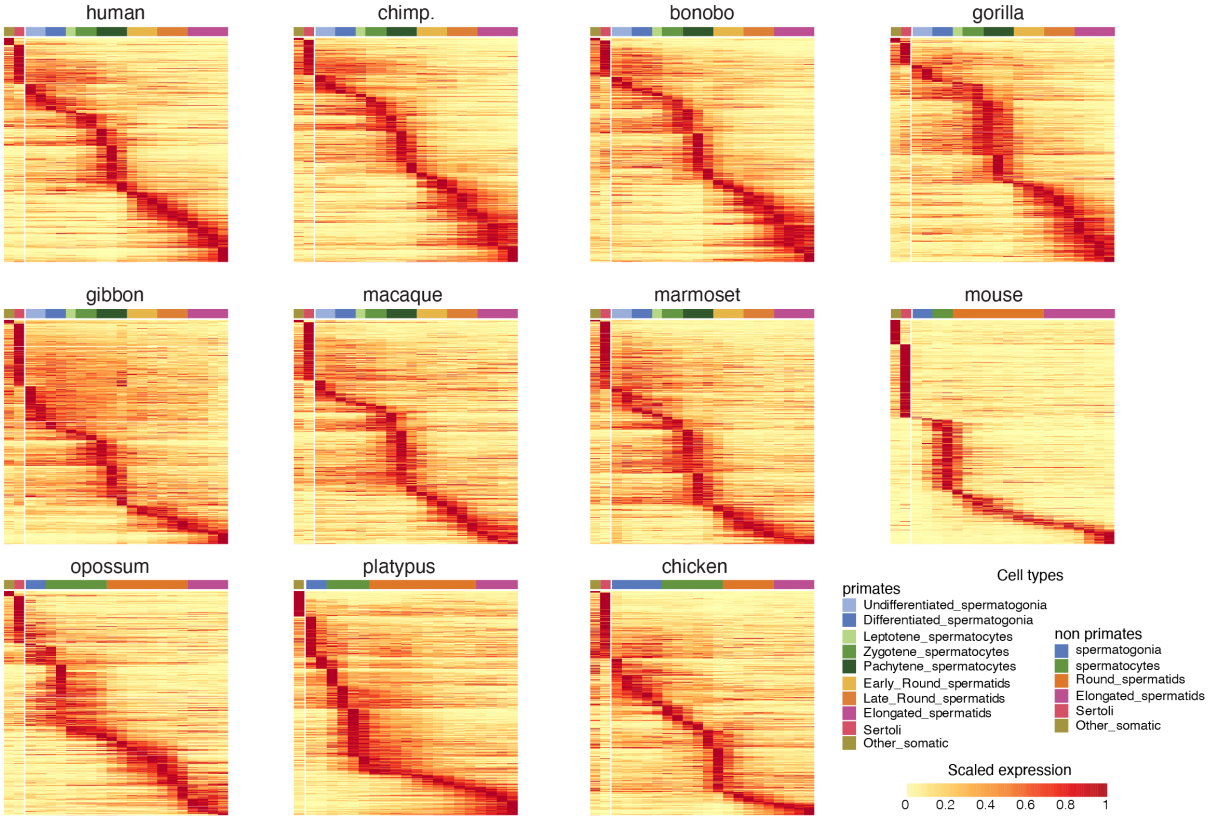


Figure 12: Testicular cell type marker genes. *The expression profiles of all identified cell type marker genes. Adapted from Murat et al. (2023).*

Taken together, we used the expression of cell type marker genes from literature in combination with the coordinates in low dimensional space to assign the cell types present in the data. Based on the assigned cell types we then further expanded the repertoire of testicular cell type marker genes. Among the germ cells the observed expression profiles of literature and novel marker genes reflect the continuous process of spermatogenesis.

3.1.2 Data overview and quality

As a next step, we wanted to gauge the integrity and quality of the generated dataset. Based on the assigned cell types we found the fraction of captured nuclei per cell type varies between the species. Especially mouse shows a low fraction of spermatogonial cells, which is in agreement with previous studies showing a higher fraction of postmeiotic germ cells in

mouse[201]. In comparison to the inter species differences, we observed a low variation within species, with the exception of chimpanzee. Overall, we found the vast majority of captured nuclei across species were classified as germ cells, reflecting the highly productive process of spermatogenesis (Figure 13a).

Next, we assessed per cell type metrics. Through this we found, the testicular cell types vary in their number of expressed genes and RNA content. Especially the large Sertoli cells, as well as pachytene spermatocytes, display a high RNA content. This is reflected by both the median number of UMIs and median number of genes captured per nucleus (Figure 13b,c).

To account for the observed differences in RNA content and nuclei numbers per cell type, we generated pseudo-bulk transcriptomes for each cell type, by aggregating all transcripts within a given cell type. In this analysis we found that despite the differences in RNA content and nuclei number, a similar number of around 15,000 expressed protein-coding genes can be observed across species and cell types per cell type (Figure 13b,c).

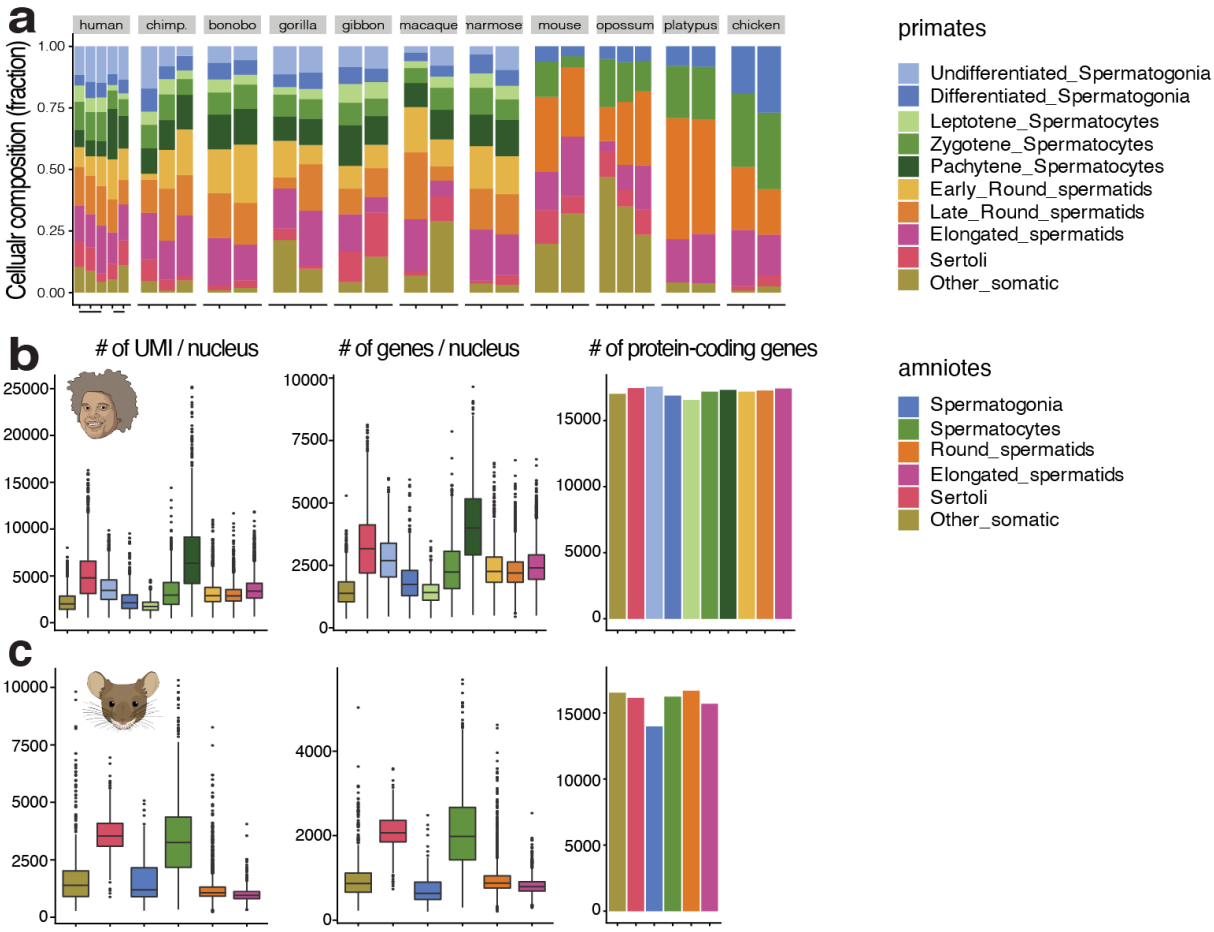


Figure 13: Nuclei metrics. a) *Cellular composition of libraries, replicates from the same individual indicated by black line.* ***b, c)*** *Median UMI count per nucleus, median number of genes per nucleus, and sum of expressed protein coding genes separated by cell type in human and in mouse, respectively. Adapted from Murat et al. (2023)*

Using these pseudo-bulk cell type transcriptomes, we also assessed cell type relations across species via principal component analysis (PCA). The progression of cell types through spermatogenesis is covered by the first principal component, which explains 11 % of the observed variance. This indicates, that shared aspects of the spermatogenesis gene expression programs across mammals and amniotes are captured within the dataset we created. This further suggests that cell type differences reflect a continuous accumulation of changes. The second component explaining 8 % of variance reflects the species-specific divergences of spermatogenesis. Together the first two principal components are sufficient to separate somatic cell types from the spermatogenic cell types and also separate the species (Figure 14). Furthermore, we found for each cell type the replicates per species group closely together. Indicating the robustness of the data

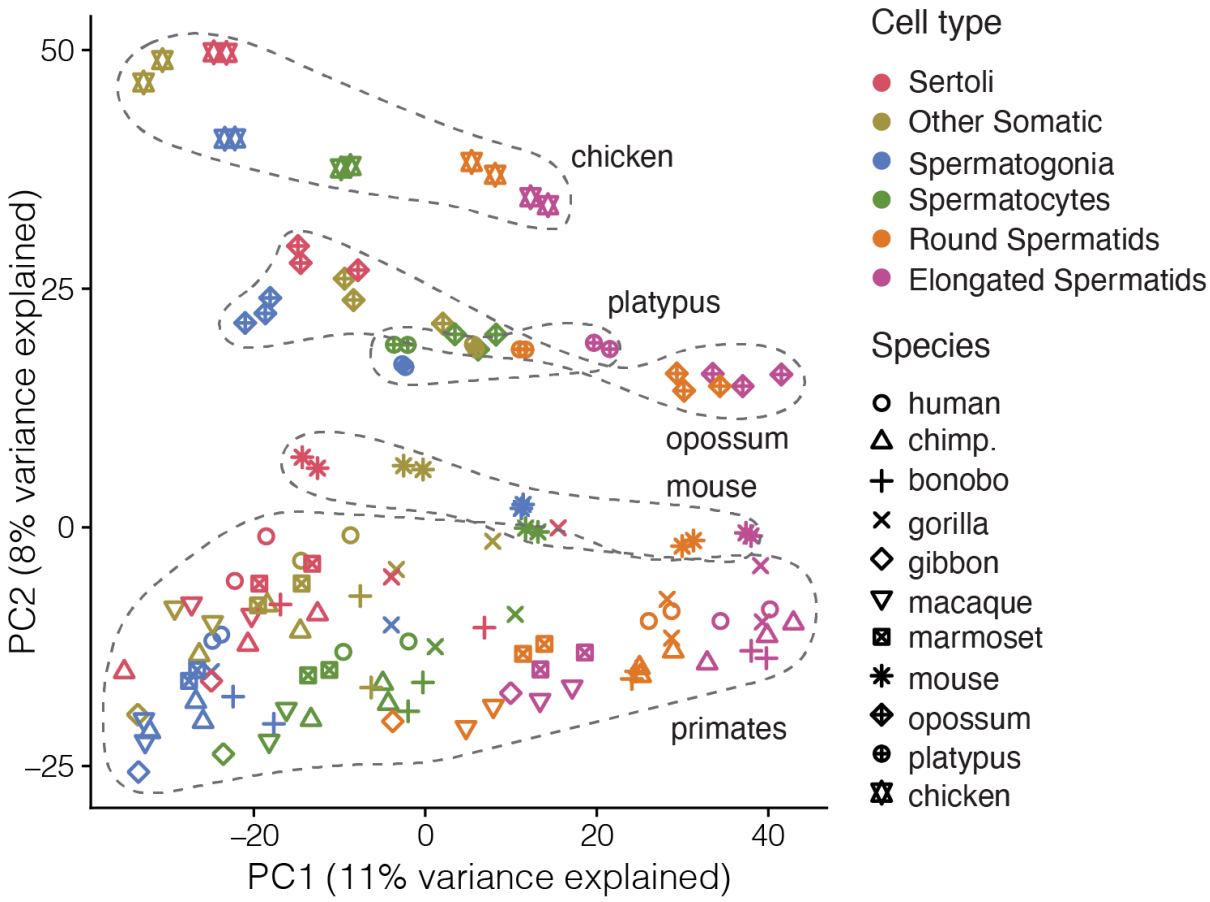


Figure 14: Principal component analysis of pseudo-bulks. *Each symbol represents one pseudo-bulk transcriptome for one replicate. Dashed lines encircle species or lineages. Chimp., chimpanzee. Adapted from Murat et al. (2023).*

3.1.3 Rates of testicular cell type evolution

Phenotypic innovation can have two major classes of mutations as a source. One class, are mutations in regulatory sequences[167]. These can affect a transcripts expression profile, as well

as impact post translational processing and the dynamics of RNA degradation. The other class of mutations are mutations directly affecting the coding sequence. Changes in sequence can directly change the final gene product and through this change the function of the gene product.

Previous work supports the view that continuous accumulation of regulatory changes over evolutionary time is reflected in present day RNA abundances[157,158]. To trace the divergence of gene expression profiles between species, we constructed gene expression trees using pseudo-bulk transcriptomes for the whole testis (Figure 15a). In these trees, samples are grouped based on the similarity of RNA levels and longer branches reflect greater differences. The resulting branching pattern thus reflects the degree of divergence between species. Overall, this tree recapitulates the mammalian phylogeny. The eutherians (primates and mouse) are grouped to the exclusion of the marsupial (opossum). The placentals (eutherians and marsupials) are grouped to the exclusion of the monotreme (platypus). All mammals are grouped together, separate from the evolutionary outgroup chicken. Also, the eutherian lineages of rodents and primates are resolved. Within the primates the grouping of the ape gibbon with the cercopithecid macaque does not follow established phylogenies. The ape is expected to group together with the great apes (chimpanzee, bonobo, gorilla, and human) to the exclusion of the cercopithecid. Seeing that we only had access to one gibbon individual, the grouping is most likely due to the sample and does not reflect a biological reality. Furthermore, branching among the great apes does not follow the established phylogeny of human grouping with chimpanzee and bonobo to the exclusion of gorilla. Although, it follows branching observed in previous bulk tissue RNA-seq data for great ape testis, which was explained as a reflection of male mating patterns and physiology evolution[157]. We tested the reliability of the generated expression trees through bootstrap analysis (1,000 samples with replacement). Overall, the branching pattern in the expression tree is robust, seeing that we obtained high bootstrap values, except within great apes. The robustness of branching among the great apes, could beyond reflecting mating pattern differences also be a reflection of incomplete lineage sorting [220].

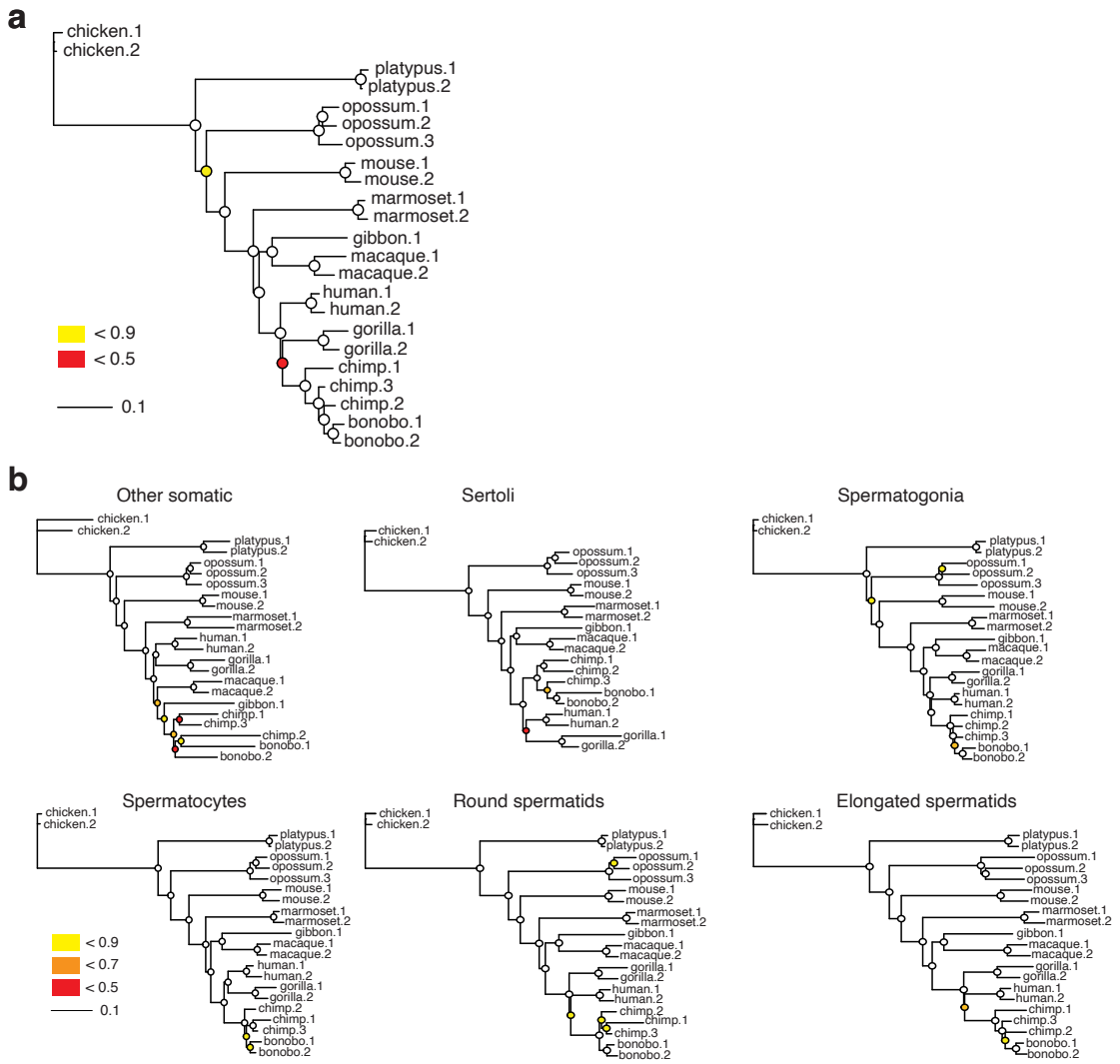


Figure 15: Gene expression trees. a) *Gene expression phylogeny based on whole testis pseudo-bulk transcriptomes* ***b)*** *Gene expression phylogeny based on cell type pseudo-bulk transcriptomes.* ***a*** *and* ***b*** *based on 4,498 1:1 orthologous genes. Bootstrap values (1,000 samples with replacement) indicated by circles, white fill ≥0.9. Adapted from Murat et al. (2023).*

To leverage the cell type resolution in our data and to get an understanding which cell types contribute how to the observed rapid testicular evolution, we also constructed gene expression trees for pseudo-bulk cell type transcriptomes (Figure 15b). These trees are in agreement with the whole testis tree, except for the tree corresponding to the other somatic cell types. In this tree, the branching is less robust and the grouping of (apes and macaque) to the exclusion of human is consistent with neither phylogeny nor the branching pattern observed for other cell types. This could be a result of differences in cell types grouped together here as other somatic.

Continuing with this line of analysis, we calculated the total branch length per cell type. In the gene expression trees we generated a longer branch length corresponds to more

accumulated change (Figure 16). Through this it becomes readily apparent, that especially the trees for the postmeiotic cell types haven an increased branch length. This can be observed both, when using expression trees across amniotes, as well as when generating the expression trees for primates only.

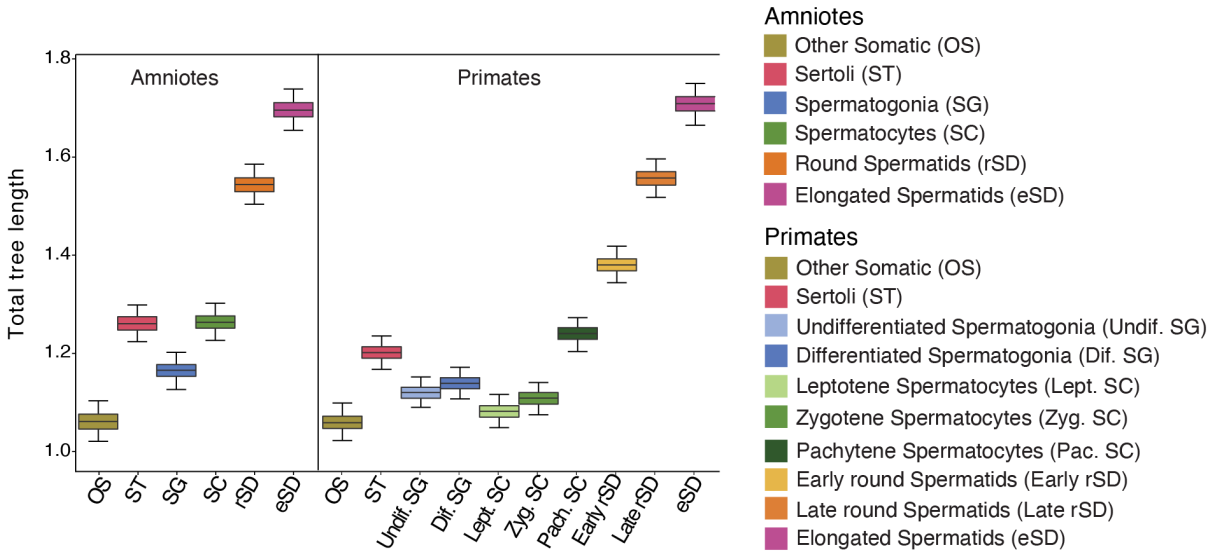


Figure 16: Branch length analysis. *The total tree length for amniote (left) and primate (right) cell type pseudo-bulk expression trees. Adapted from Murat et al. (2023).*

To further our understanding about the divergence of the cell types, we then used the pseudo-bulk cell type transcriptomes to perform pairwise species comparisons of the cell type transcriptomes (Figure 17). As expected, we found that correlation decreases with increasing evolutionary distance between the compared species. Further, the decrease for the postmeiotic cell types is greater than for premeiotic germ cells and somatic cells. Indicating, that overall testicular divergence is driven by postmeiotic cell types. In the display of the correlation, we also observed the correlation between human and platypus as well as between human and chicken is similar. This is despite the chicken lineage branching from the human lineage ~140 million years earlier than the platypus lineage. This is similar to an observation made for bulk tissue RNA-seq data and could reflect the core transcriptomic program necessary for organ function[157]. To ensure our observed pairwise divergences are not driven by differences in nuclei or UMI count, we also performed a downsampling analysis. Similarly, to the full sample, the postmeiotic cell types are more diverged than the premeiotic germ cells and the somatic cell types.

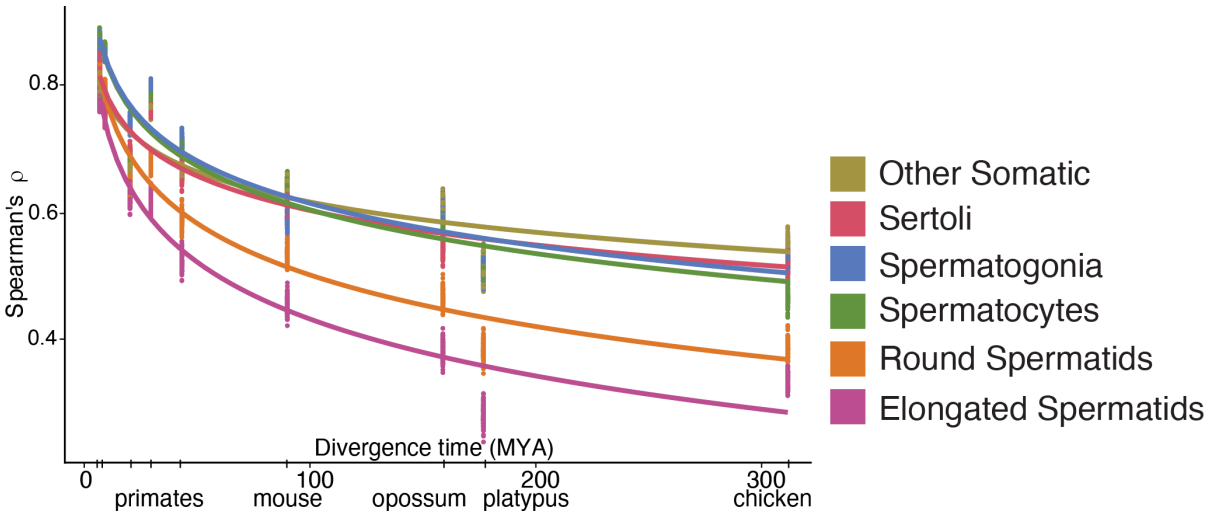


Figure 17: Pairwise comparison. *Pairwise correlations between humans and other species. Lines correspond to linear regression trends (after log transformation of the time axis. MYA, millions of years ago. Adapted from Murat et al. (2023).*

Taken together, we leveraged the cell type resolution in the data to dissect the contribution of the testicular cell types to the observed rapid evolution of the testis. We find that especially the postmeiotic germ cells are a driver of the divergence. Further, we show that the mitotic spermatogonia display similar divergence as the testicular somatic cell types.

3.1.4 Evolutionary forces

The rapid evolution of the testis is driven by the rapid evolution of late spermatogenic cell types. This observation can be explained by two non-mutually exclusive patterns of natural selection. One explanation is an increase in positive selection, an increased rate of fixation for adaptive changes. The other is a decrease in negative selection, a reduction in the removal of, potentially deleterious, mutations. To explore this, we leveraged the per cell resolution of our dataset and assessed various metrics to determine how the transcriptome of individual cells is shaped by evolutionary forces. To improve clarity, we present these results as a per cell type summary.

3.1.4.1 Negative selection and gene expression constraint

Cells expressing more genes refractory to change are expected to be under stronger negative selection. As both regulatory and sequence changes for these genes are more likely to be eliminated. In human, we assessed the degree to which expressed genes are refractory to change using the probability of being loss-of-function intolerant (pLI) score[221]. This score, based on exome data from ~61k human individuals, ranges from 0 to 1. A higher score corresponds to genes, which are not tolerant to loss-of-function mutations, while a lower score corresponds to increased loss-of-function tolerance. Displayed as percentiles, a higher

percentile thus corresponds to a greater loss of function tolerance (Figure 18a). Sertoli cells and other somatic cells showed the lowest tolerance to loss-of-function mutations. In the germ cells we found a steady increase in loss-of-function tolerance starting in differentiated spermatogonia with a peak of functional tolerance in early round spermatids. This indicates a reduction of negative selection with progression of spermatogenesis in human.

For mice, we followed a similar analysis path using data provided by the international mouse phenotyping consortium (IMPC)[222]. This consortium provides a database of neutrally ascertained mouse knockout phenotypes. Genes that produce a lethal phenotype when knocked out are under strong functional constraint. Consequently, cells expressing more genes with a lethal phenotype when knocked out are expected to be under increased negative selection. Plotting the percentage of expressed genes with a lethal phenotype when knocked out among expressed genes, which have been tested for a lethal phenotype when knocked out, we found a slightly increased expression of lethal genes in spermatogonia compared to the somatic cell types (Figure 18b). This elevated expression decreases with progression in spermatogenesis, already falling below the somatic level of expressed lethal genes in spermatocytes. Round and elongated spermatids showed similar levels of expressed lethal genes compared to each other. This indicates a reduction of negative selection with progression of spermatogenesis in mice.

3.1.4.2 Positive selection and sequence evolution

Beyond the effect of negative selection on which genes are expressed, purifying selection also shapes the sequence of expressed genes. In primates, we assessed this by calculating the mean for the rate of normalized non synonymous mutations over synonymous mutations (dN/dS) (Figure 18c) [223,224]. The rate of mutations is normalized by the mutation rate in neutrally evolving regions. For the estimation of dN and dS values we used PAML with the M0 site model of aligned homologous sequences across primates[225]. A higher dN/dS ratio corresponds to more amino acid altering substitutions becoming fixed. Across primates, we observed the highest dN/dS values in round spermatids, with a continuous increase from spermatogonia. The testicular somatic cells showed a dN/dS value similar to spermatogonia and also similar to cell types from another somatic tissue.

The increased fixation of changes is not necessarily due to a relaxation of negative selection i.e., a less efficient removal of changes. The increased fixation of changes could also be driven by an increase in positive selection.

To analyze, how positively selected genes contribute to the observed transcriptomes, we used a list of 331 genes, which in a comparison across nine simian genomes showed signs of positive selection (among 11,096 tested genes, Figure 18d)[226]. The percentage of expressed positively selected genes among expressed genes tested for positive selection increases throughout spermatogenesis. The somatic cell types showed an expression of positively selected genes similar to spermatogonia.

Similarly, when we analyzed the expression of genes for which signs of positive selection have been observed in eutherians (544 genes out of 16,419 genes tested for positive selection), we observe a similar pattern across all species[182].

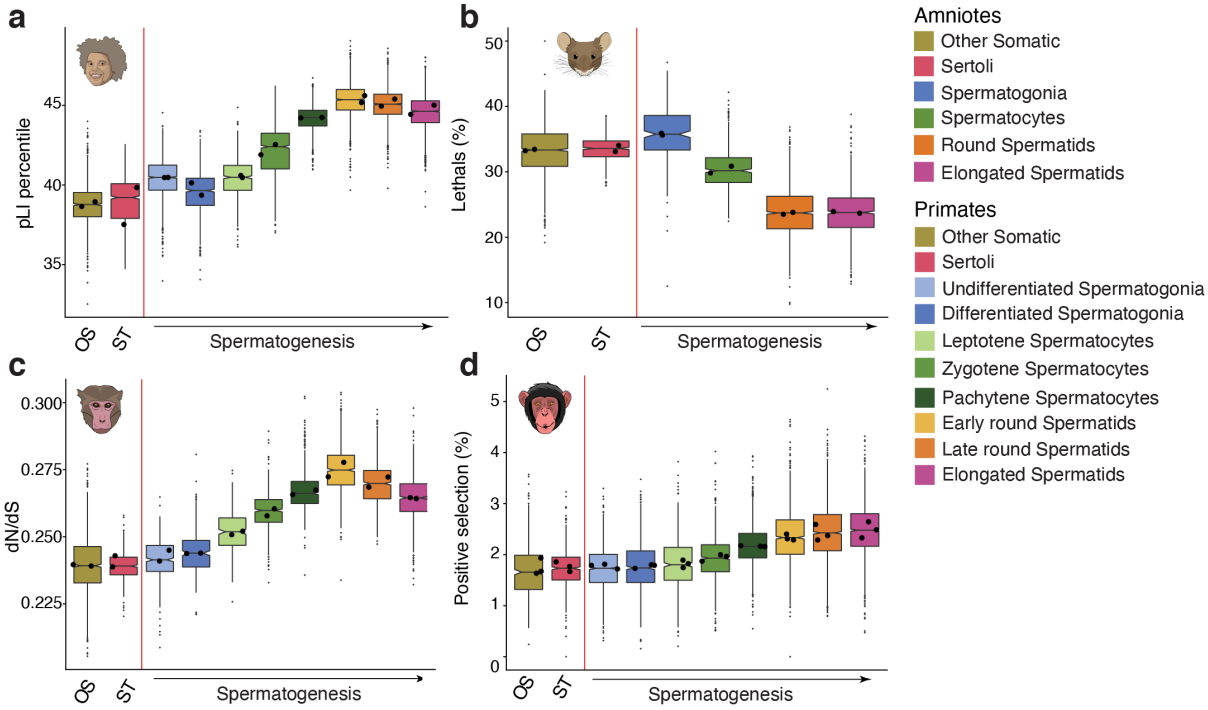


Figure 18: Functional constraint and selection. a) *mean pLi score percentile of expressed genes in human.* ***b)*** *percentage of expressed genes leading to a lethal phenotype in mouse when knocked out.* ***c)*** *mean dN/dS of expressed genes in macaque.* ***d)*** *percentage of expressed genes under positive selection in chimpanzee. All data based on value per cell. Adapted from Murat et al. (2023).*

3.1.4.3 Expression pleiotropy and evolutionary constraint

Another key constraint for which mutations are permissible for a gene is the breadth of expression, across both tissues as well as developmental processes[176]. This is also termed expression pleiotropy. A narrow spatiotemporal expression profile, i.e., expression in few tissues at a specific stage of an animal's lifecycle, corresponds to low pleiotropy. The lower

pleiotropy is concomitant with a reduction in functional constraint, as constraints from other developmental stages and other tissues are not applicable. This results in increased adaptability.

We used spatiotemporal gene expression profiles based on bulk tissue RNA-seq data, which have previously been generated, to assess both the tissue specificity as well as the time specificity profiles of testicular cells (Figure 19a,b)[158]. In this analysis we observed across species, that the somatic cell types express the genes with the broadest expression profiles. Furthermore, among germ cells the spatiotemporal specificity increases along spermatogenesis. Among the primates the undifferentiated spermatogonia display a higher specificity than the differentiated spermatogonia. This showed that pleiotropy decreases with progression of spermatogenesis.

Phylogenetically younger genes, i.e., genes which have emerged on a more recent branch, show less pleiotropy than older genes[227]. To assess the age of testicular transcriptomes, we used a previously established phylogenetic age score (Figure 19c). In this score a higher value corresponds to a younger transcriptome[158]. We found, the transcriptomic age of the testicular somatic cell types is slightly older than that of other somatic tissues and similar to the transcriptomic age of spermatogonia. Across species, in germ cells the expressed genes become progressively younger throughout spermatogenesis.

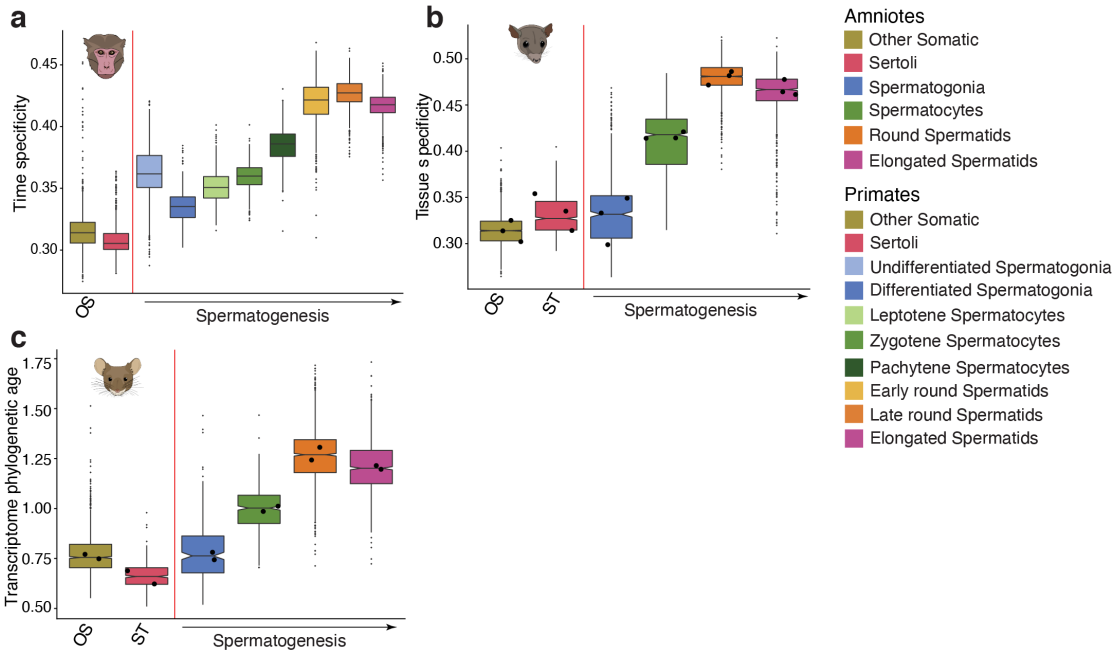


Figure 19: Pleiotropic constraint. a) *Mean time specificity in macaque.* ***b)*** *Mean tissue specificity in opossum.* ***c)*** *Mean transcriptome age in mouse. All data based on per cell value of expressed genes. Adapted from Murat et al. (2023)*

3.1.4.4 Chromatin accessibility and expression of new genes

The emergence of new genes expressed in the testis has previously been described and led to the notion of the testis as the "birthplace of new genes"[189]. This notion is based on previous bulk cell type analyses of RNA-seq, DNA methylome, and CHIP-seq and supported by the high degree of chromatin remodeling due to histone exchanges and chromatin compaction around protamines[144,188]. These data showed that especially spermatocytes and spermatids have a permissive chromatin state, which facilitates transcription[188,228].

To get a better understanding of the chromatin state in germ cells during spermatogenesis, I generated an additional multimodal dataset of adult mouse testis, which yields both snRNA-seq and snATAC-seq data from the same nucleus. These data give insight into which chromatin regions are accessible to transposase and thus illuminate, which chromatin regions and which cis regulatory elements are accessible. In this data, I identified the testicular cell types based on the snRNA-seq modality. I could not assign cell type based on the snATAC-seq modality alone. To get an understanding, of which genes are accessible for transcription, I analyzed the enrichment of peaks in transcription start sites (Figure 20a). In this analysis I observed that peaks are less enriched in TSS in the postmeiotic cells. This decrease of enrichment of signal in the TSS is accompanied by a more widespread alignment of reads across the genome. This could be reflective of the globally accessible chromatin state, diluting the signal at specific sites.

The open chromatin state can facilitate the transcription of novel genomic elements. To assess this, we utilized expanded testis specific transcriptome annotations, we previously generated based on published and newly made bulk tissue RNA-seq libraries for adult testis of the elven investigated species (Figure 20b). These new annotations increased the number of features and expanded previously annotated features. In this analysis we found, that across species intergenic elements contribute increasingly to the transcriptome of germ cells as they progress through spermatogenesis, plateauing in round spermatids. In the somatic cell types, the contribution of intergenic elements to the transcriptome is similar to the contribution we observe in spermatogonia. Supporting the notion of promiscuous transcription in later spermatogenic stages.

The widespread transcriptional activity could lead to the expression of transcripts with deleterious gene products. To buffer the phenotypic impact of these gene products, translation of transcripts in cells with widespread transcription could be attenuated. To asses

this, we used translational efficiency values, which have been generated previously based on bulk tissue RNA-seq and Ribo-seq data of adult testis (Figure 20c)[159]. To account for the bulk tissue origin of this data, we calculated per cell the median translational efficiency of expressed genes, which have more than 60 % of that gene's total transcripts in the cells cell type. In this analysis we found that transcripts are more efficiently translated in the somatic cell types than in the germ cells. Furthermore, the translational efficiency decreases during spermatogenesis with a nadir of translational efficiency in round spermatids, buffering the effects on transcriptional divergence.

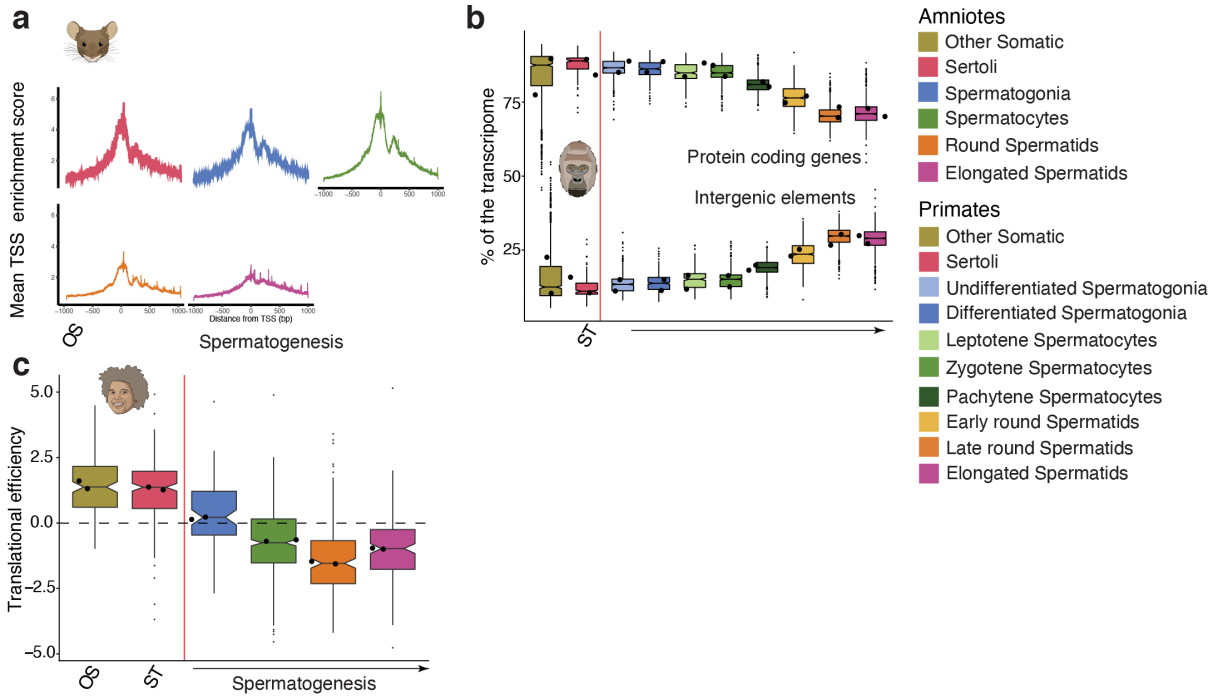


Figure 20: Transcriptionally permissive landscape. a) *Mean transcriptional start site (TSS) enrichment in snATAC-seq data from human.* ***b)*** *Percentage of UMIs mapping to protein coding genes or inter genic regions in gorilla.* ***c)*** *Translational efficiency of genes with predominant expression in a given cell type. b,c Adapted from Murat et al. (2023).*

3.1.4.5 Functional impact on fertility

The expression profiles we generated have so far shown that the genes expressed in the later stages of spermatogenesis have narrower spatiotemporal expression profile than those expressed early. Furthermore, they are less likely to be associated with a lethal phenotype when knocked out and less refractory to loss of function mutations. To better understand the impact of the late spermatogenic transcriptomes on tissue function, especially focusing on the testicular role in reproduction, we assessed the expression of genes with previously described fertility phenotypes in IMPC (173 out of 3,252 genes tested for infertility) (Figure

21a)[222]. In mice, we observed peak expression of genes associated with infertility when knocked out in spermatocytes. Furthermore, we observed higher amounts of infertility associated genes in germ cells than in the somatic cell types.

Tracing the evolutionary emergence of these genes, we identified two rodent specific genes, *D1Pas1* and *H2al2 (Figure 21b,c)*[229]. Both of these originated from the X chromosome through RNA-based retroduplication[230,231]. As they both show an increase in expression levels with progression through spermatogenesis and have an impact on mouse fertility, this shows that not only are more younger genes expressed during later stages of spermatogenesis, but also that these new genes convey a functional role, especially during later stages of spermatogenesis.

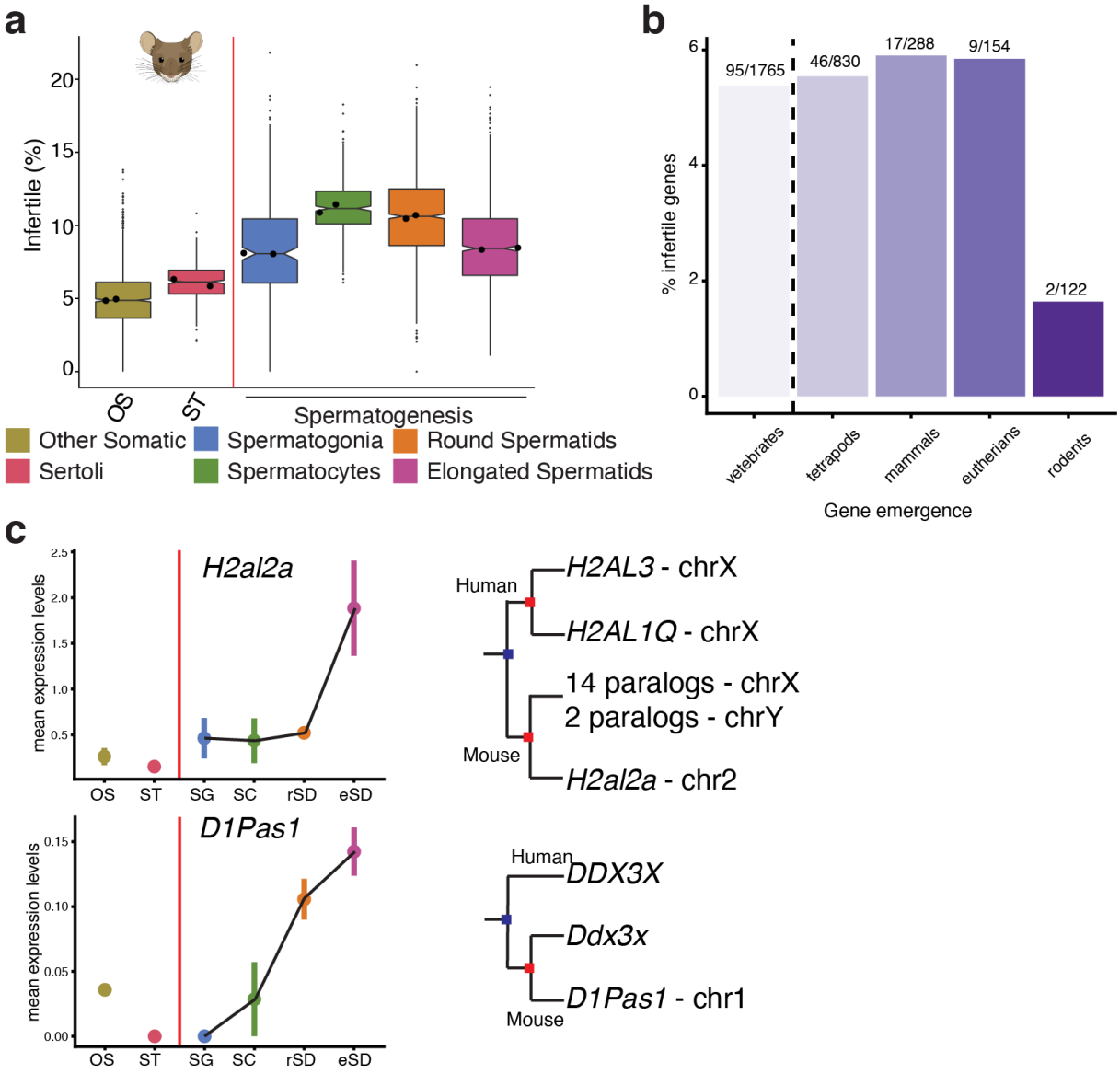


Figure 21: Fertility. a) *percentage of genes with an infertility phenotype when knocked out in mouse.* ***b)*** *Percentages of mouse fertility genes among genes newly emerged.* ***c)*** *expression profile of H2al2a (top) and D1Pas1 (bottom) and paralog tree. Adapted from Murat et al. (2023).*

Taken together, our findings provide evidence for postmeiotic germ cells being both under less purifying selection as well as increased positive selection. Further, the genes expressed in the later spermatogenic cell types are under less constraint through expression pleiotropy. Overall, the transcriptomes become younger with progression in spermatogenesis and a widespread transcription of intergenic elements can be observed. The phenotypic impact of these transcripts is buffered on the translational layer. While the highly specific transcriptomes of later spermatogenic cell types have a low impact on overall organismal survival, they have an increased impact on reproductive function. This functional impact of late expressed genes is also carried by lineage specific, newly emerged genes.

3.1.5 Conserved and diverged amniote gene expression

The analyses per cell gave us insights in the global patterns of gene expression evolution. To also dissect, which genes contributed how to the observed properties of germ cells and also understand, what are conserved and diverged aspects of the testicular transcriptome in amniote evolution we scrutinized gene expression trajectories in germ cells across amniotes. In this analysis we compared the relative expression levels of the 2,927 one-to-one orthologous genes along spermatogenic cell types, which are expressed in all investigated species. To maximize the number of orthologous genes, we used only human as a representative for the primate lineage. For the comparison of the expression trajectories, we applied a soft clustering approach using Mfuzz (Figure 22a)[232].

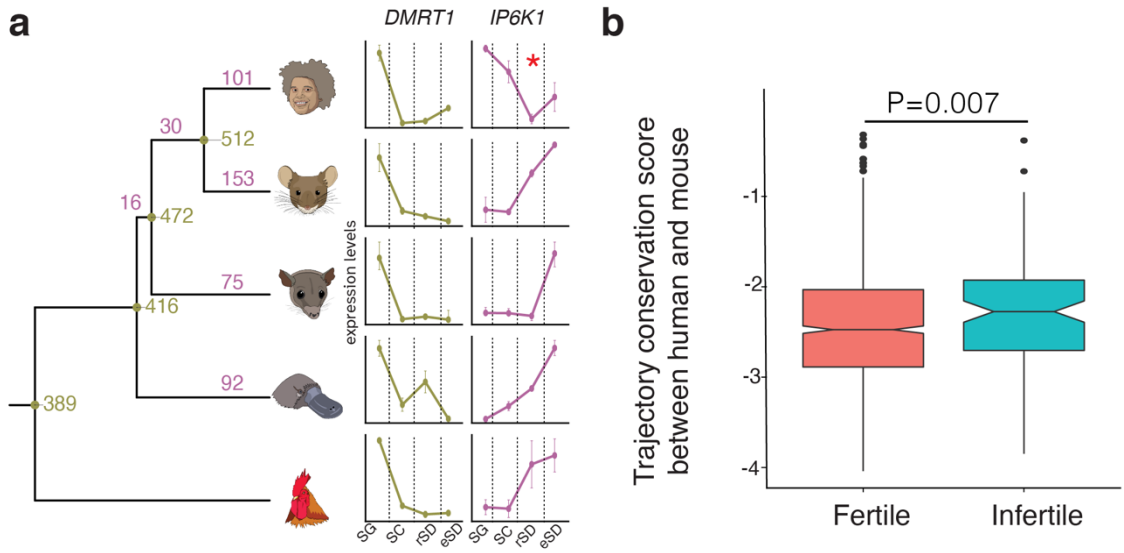


Figure 22: Amniote trajectories. a) *changed (purple) and conserved (olive) trajectories in amniotes.* ***b)*** *trajectory conservation score of genes with or without fertility phenotype. Adapted from Murat et al. (2023).*

Among these tested genes, we found 389 genes with a conserved expression pattern across amniotes. Among these were genes in each of the trajectory clusters. Notably, these conserved genes include examples with important fertility phenotypes, showing peak expression in each of the different spermatogenic cell types. Among these are *DMRT1*, which in mouse is required for maintenance and replenishment of spermatogonial stem cells[200], *PARG* which in premeiotic cells is involved in DNA double strand breaks and repair[233], *TLE3* which is specifically not expressed in spermatocytes and plays a role in the communication with Sertoli cells[234], *BBX* which is expressed in postmeiotic cells and in rat is associated with infertility[235].

To test, whether genes with a fertility phenotype show more conserved expression trajectories than genes without a fertility phenotype, we devised a trajectory conservation score, which reflects the probability an orthologous gene set is part of the same trajectory cluster across all species. A higher conservation score is indicative of a higher expression trajectory similarity. In this analysis we found, that genes associated with an infertility phenotype when knocked out in mice have a significantly higher trajectory conservation score between human and mice than genes for which no fertility phenotype has been described (Figure22b).

The trajectory analysis further allowed us to identify gene expression trajectory changes and the branch on which they occurred. We found between 16 (between therians and monotremes) and 153 (between human and mouse) of these gene expression trajectory changes. The identification of these is important, as the changed context in which the gene is expressed could be indicative of a change in the functional role of the gene product. One example for an expression change is *Inositol Hexakisphosphate Kinase 1* (*IP6K1*). The knockout of this gene in mice results in infertility[236]. Across amniotes, the expression pattern with increasing expression towards later spermatogenic stages is conserved, except in primates. In these, expression of *IP6K1* is high in spermatogonia and lowest in spermatocytes or round spermatids.

We wanted to further dissect the role of the genes with conserved expression trajectory in the cellular function. For this, we performed a GO term enrichment analysis for genes with peak expression in each of the spermatogenic cell types. In this analysis we found, conserved genes for each cell type are enriched for terms associated with the biological function of that cell type (Figure 23a-d). For example, meiotic spermatocytes are enriched for the term “meiotic cell cycle”. For the genes with changed expression trajectories, we cannot indicate a cell type of predominant expression, as that varies between species. Consequently, we found broader GO terms more associated with typical metabolical processes, when analyzing the changed genes (Figure 23e).

In agreement with the notion of gene expression pleiotropy as constraint rendering genes refractory to expression changes, we found genes for which we have called a trajectory change to be significantly more tissue- and time-specific, than genes we have called as conserved (Figure 23f).

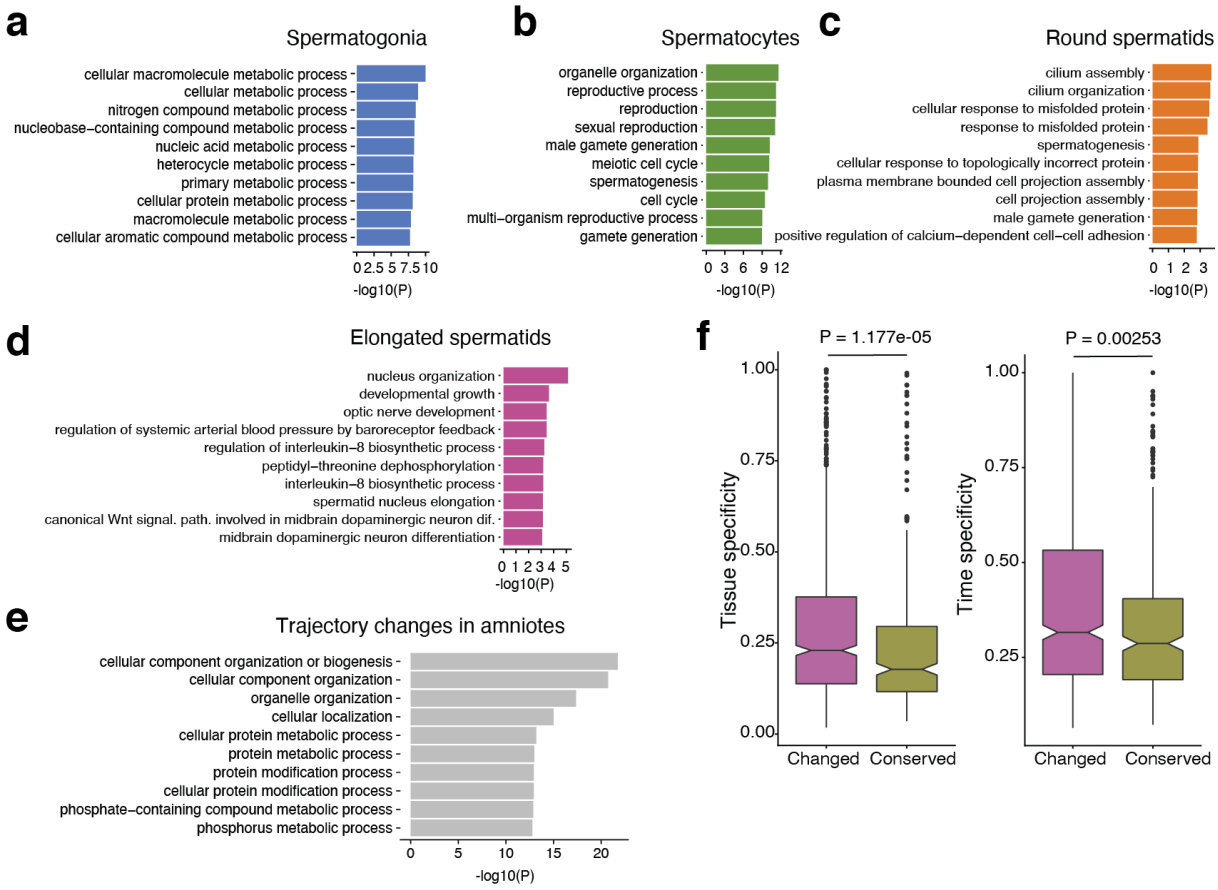


Figure 23: ***GO analysis.*** *Enriched GO-terms for genes predominantly expressed in* ***a)*** *spermatogonia,* ***b)*** *spermatocytes,* ***c)*** *round spermatids,* ***d)*** *elongated spermatids, or* ***e)*** *with trajectory changes.* ***f)*** *Spatiotemporal specificity of genes with changed or conserved expression trajectories. Adapted from Murat et al. (2023).*

The progression of germ cells through spermatogenesis is highly coordinated to ensure continuous production of germ cells, which is reflected in the seminiferous cycle[31,153]. The coordination hinges on inter cellular communication - both between the different germ cell types and, crucially, between germ cells and Sertoli cells, which nurture the process of spermatogenesis. In this communication the Sertoli cells also need to communicate with germ cells on both sides of the blood-testis barrier[36]. To investigate this, we used CellPhoneDB, which based on a provided receptor-ligand interaction list, identifies cell type pairs for which a receptor-ligand pair is significantly expressed[237]. Our analysis revealed, that the majority of interactions can be found among the somatic cell types and between somatic cells and premeiotic spermatogonia (Figure 24a). To further improve our understanding of the seminiferous cycle, we specifically investigated significant receptor-ligand interactions between Sertoli cells and germ cells (Figure 24b,c). To account for the approaches proneness to false positive results, we focused the analysis on conserved aspects, which are shared among species. In this analysis we found that across all amniotes Sertoli cells communicate

with spermatogonia, spermatocytes, and round spermatids through the cell adhesion molecule CADM1, which previously was thought to not be expressed in Sertoli cells[238]. Our data also supports the notion of NECTIN2-NECTIN3 complex mediated communication between Sertoli cells and spermatids in humans[239,240].

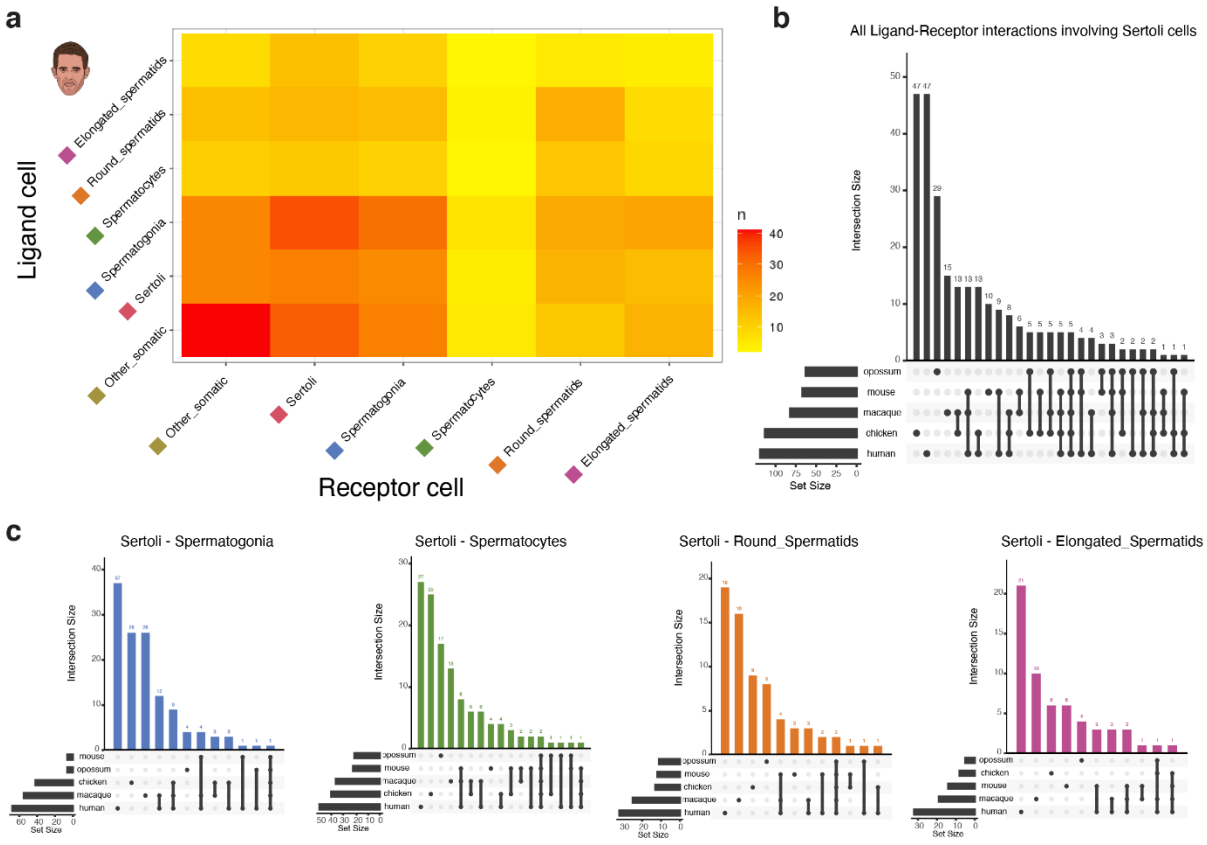


Figure 24: Cell communication. a) *significant ligand receptor interactions per cell type in human.* ***b)*** *significant interactions between Sertoli cells and all germ cell types for amniotes.* ***c)*** *significant interactions between Sertoli cells and germ cells separated by cell type. Adapted from Murat et al. (2023).*

Taken together, we found genes with conserved expression trajectories in germ cells across amniotes. Among these are gene with a well described impact on fertility. We also found, genes with a conserved expression trajectory are under stronger pleiotropic constraint. We further found divergent aspects of gene expression profiles between amniotes. Among these, we too identified genes with previously described fertility phenotypes. Our analysis of GO term enrichment revealed the per cell type conserved gene expression to be reflective of the cell type function. Our cell communication analysis highlighted the importance of interactions between cells prior to the blood-testis barrier and revealed the across amniote role of CADM1 in Sertoli cell communication.

3.1.6 Conserved and diverged primate gene expression

The seven simian species in the dataset are closely related. We already leveraged this relationship in the integrated embedding of the primate nuclei, which was the basis for the additionally assigned germ cell types. Only simians, there are 11,948 one-to-one orthologous genes - more than the 8,045 one-to-one orthologous genes available for the cross amniote analysis. Of the simian orthologs, we found 4,559 are expressed in all the samples in sufficient abundance for trajectory analysis. With this additional detail, we wanted to dissect conserved and diverged aspects of the simian testis transcriptome evolution (Figure 25).

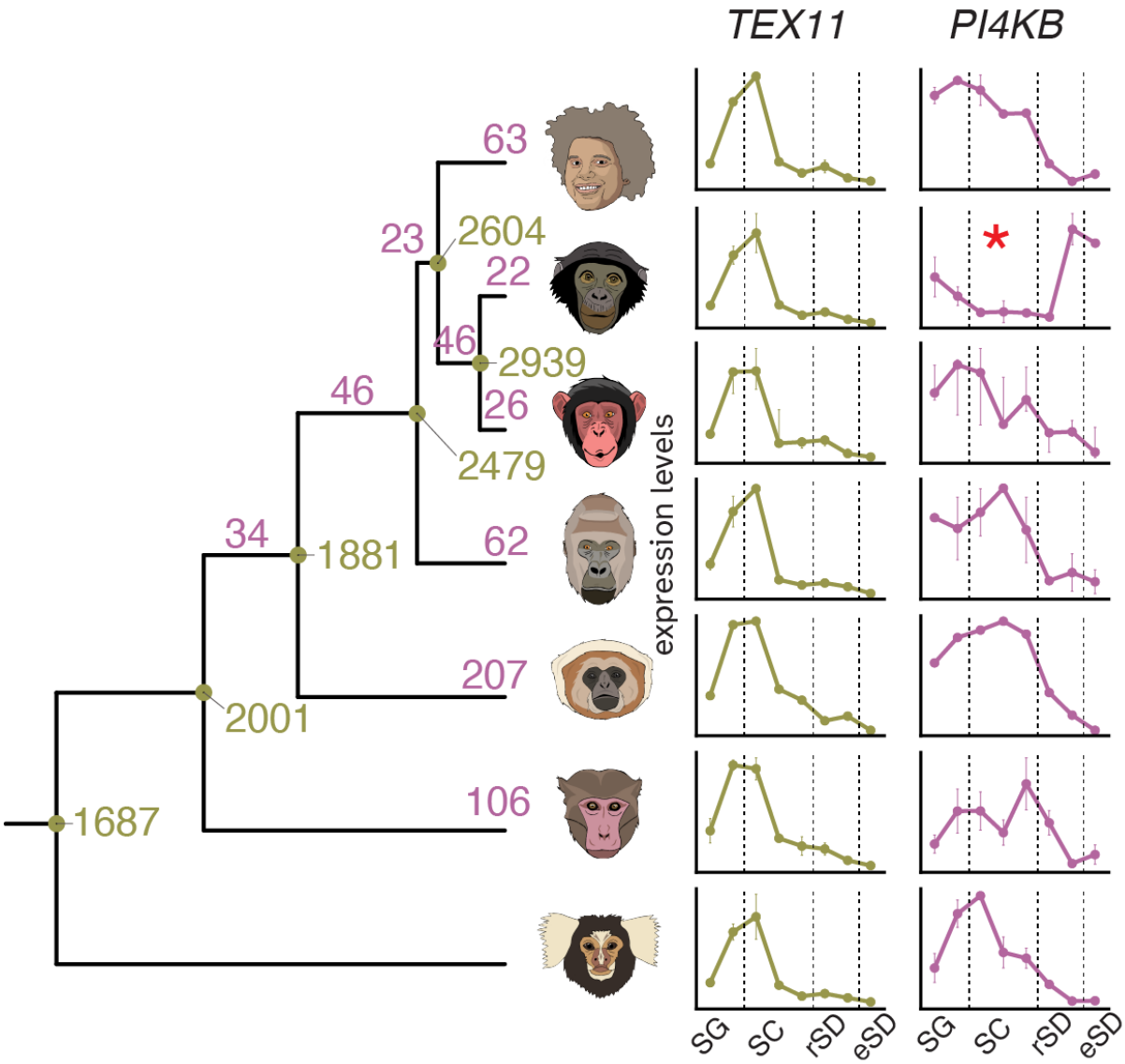


Figure 25: Primate trajectories. a) *changed (purple) and conserved (olive) trajectories in primates. Adapted from Murat et al. (2023).*

Our analysis revealed 1,687 genes (~38 % of the tested genes) with a conserved expression trajectory across simian germ cells. Analogous to the observation in the trajectory analysis in

amniotes, we found genes with important functions for fertility in all of the expression clusters. Among these are *Anosmin1* (*Anos1*) with conserved highest expression in undifferentiated spermatogonia. Mutations in *Anos1* can cause Type 1 Kallmann syndrome[241]. *N-Myristoyltransferase 2* (*NMT2*), which is putatively associated with hypogonadotropism[242,243] has peak expression in differentiated spermatogonia.

RNA binding motif protein 6 (*RBM6*) has peak expression in Leptotene spermatocytes. It functions in the homologous repair of DSBs[244]. *WD Repeat Domain 62* (*WDR62*), which plays a role in centriole duplication during meiosis has peak expression in Zygotene spermatocytes[245]. *Zinc finger MYND-type containing 10* (*ZMYND10)* is needed for correct cilia formation, which affects sperm motility has peak expression in Pachytene spermatocytes[246,247]. *Sirtuin 4* (*SIRT4)*, which is associated with infertility in human has peak expression in early round spermatids[248]. *Spermatogenesis associated 20* (*SPATA20)*, which is associated with sperm head defects in human, has peak expression in late round spermatids[249,250]. *Cystatin 8* (*CST8)*, which leads to capacitation defects when knocked out has peak expression in elongated spermatids[251].

Beyond the conserved aspects of simian germ cell transcriptomes, we also identified significant gene expression trajectory divergences. Overall, we detected between 22 and 207 genes with a significantly changed expression profile on a specific branch of the simian species tree, potentially indicating a new functional role of the gene product. To validate our observation of lineage specific gene expression trajectory changes, we selected three candidates with a significant expression change within the great apes. Our collaborators Dr. Sofia Boeg Winge and Dr. Kristian Almstrup used these as targets for single molecule RNA in situ hybridization experiments (smISH) in great ape testis cross sections. The three validated genes were *Rubicon like autophagy enhancer* (*RUBCNL)* (Figure 26) and *ADAM metallopeptidase with thrombospondin type 1 motif 17* (*ADAMTS17)* (Figure 27a) with expression changes on the human branch and *Myosin IIIB* (*MYO3B)* (Figure 27b) with an expression change on the pan branch. The *RUBCNL* stainings were evaluated by calculating stained pixel intensity and normalizing by observed number of cells in the corresponding area. For this we stratified the tubular areas into areas predominantly occupied by spermatogonia, spermatocytes, or spermatids. For *ADAMTS17* and *MYO3B* the fraction of spermatogonia and round spermatids with at least staining dot were counted. For all three candidates the

evaluation of the smISH results followed the trend observed in the expression trajectory analysis, validating our *in silico* results *in situ*.

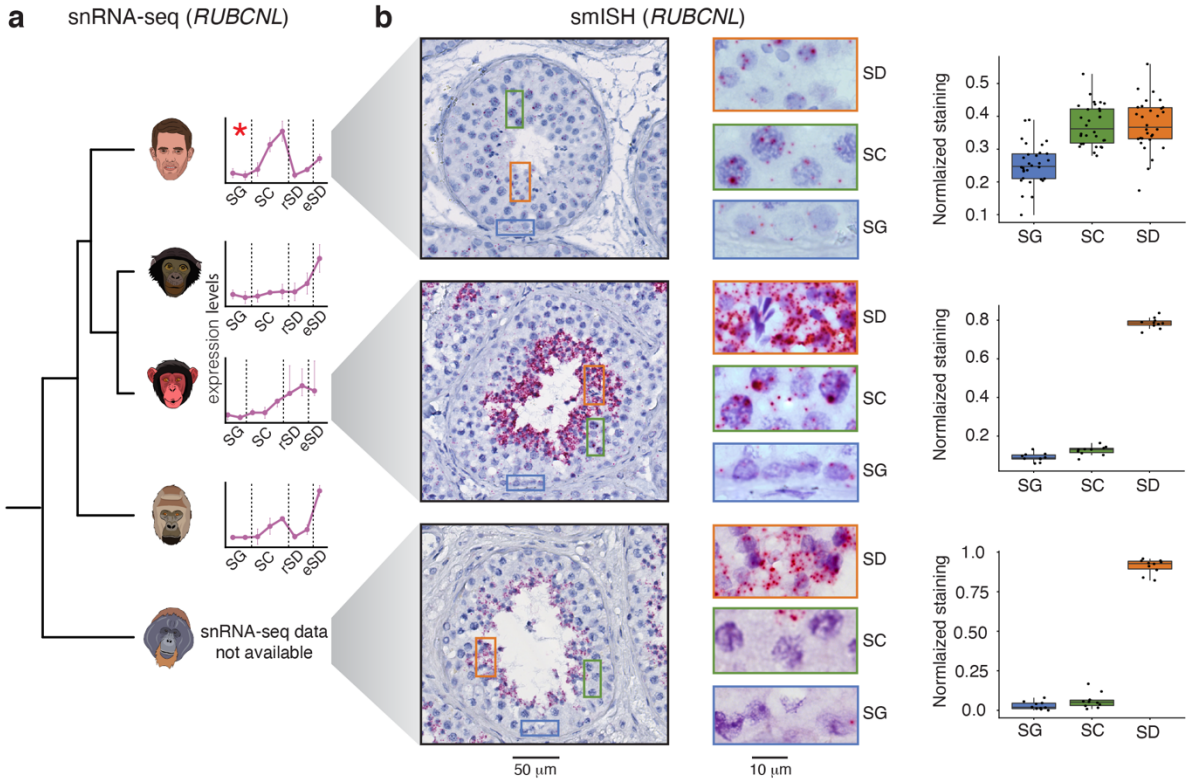


Figure 26: Trajectory validation. a) *snRNA-seq expression trajectory of RUBCNL.* ***b)*** *detection and quantification of RUBCNL using RNAScope Adapted from Murat et al. (2023).*

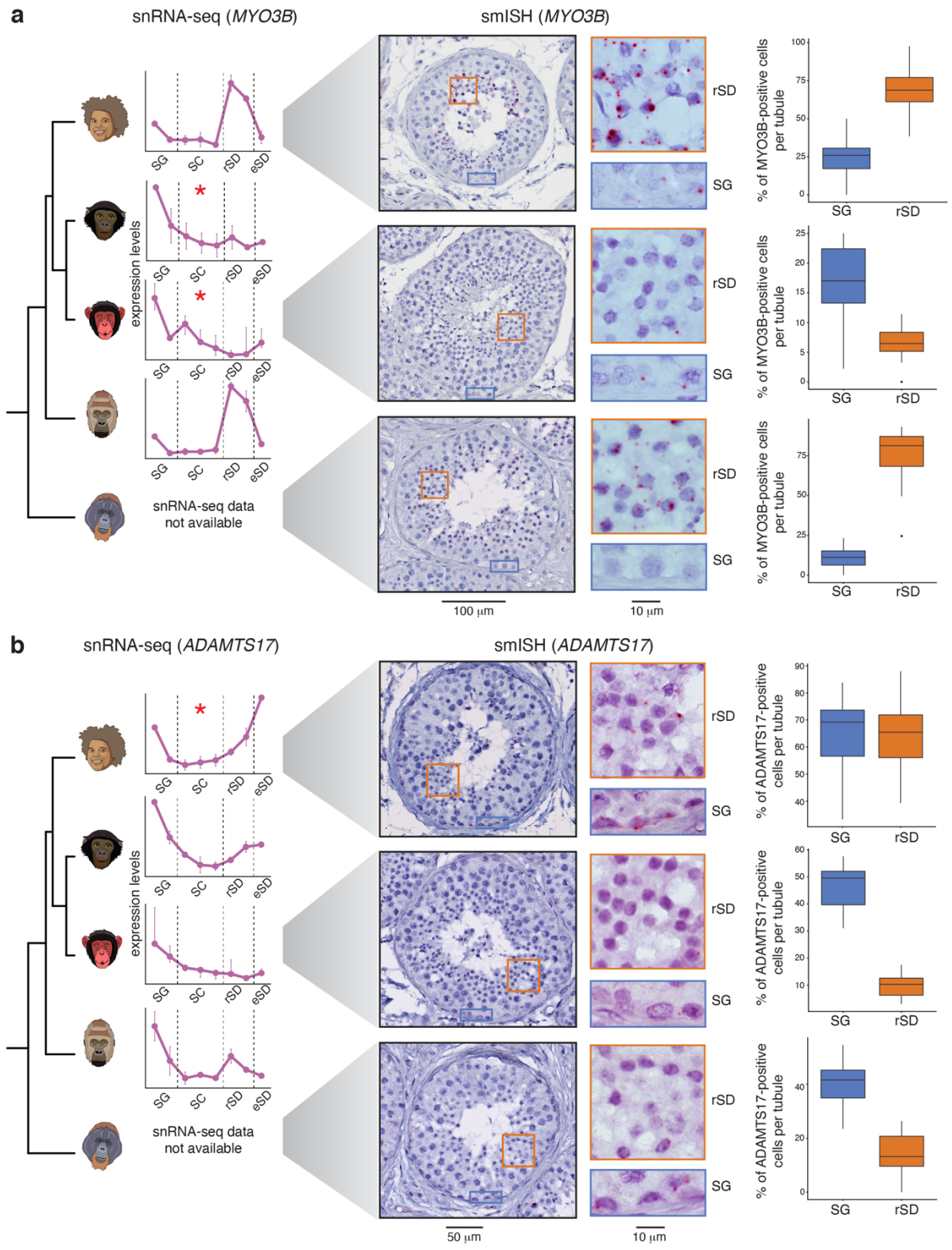


Figure 27: Trajectory validation. a) *expression trajectory and RNAScope staining for MYO3B.* ***b)*** *expression trajectory and RNAScope staining for ADAMTS17. Adapted from Murat et al. (2023).*

Similar to the amniote trajectory analyses, we performed a GO term enrichment analysis of conserved genes with peak expression in a specific cell type. The enriched terms correspond to expected cell type functions, such as “spermatid development” in late round spermatids

(Figure 28a). Furthermore, the analysis of GO terms for the genes with changed expression trajectories, again showed an enrichment of broader terms again (Figure 28b).

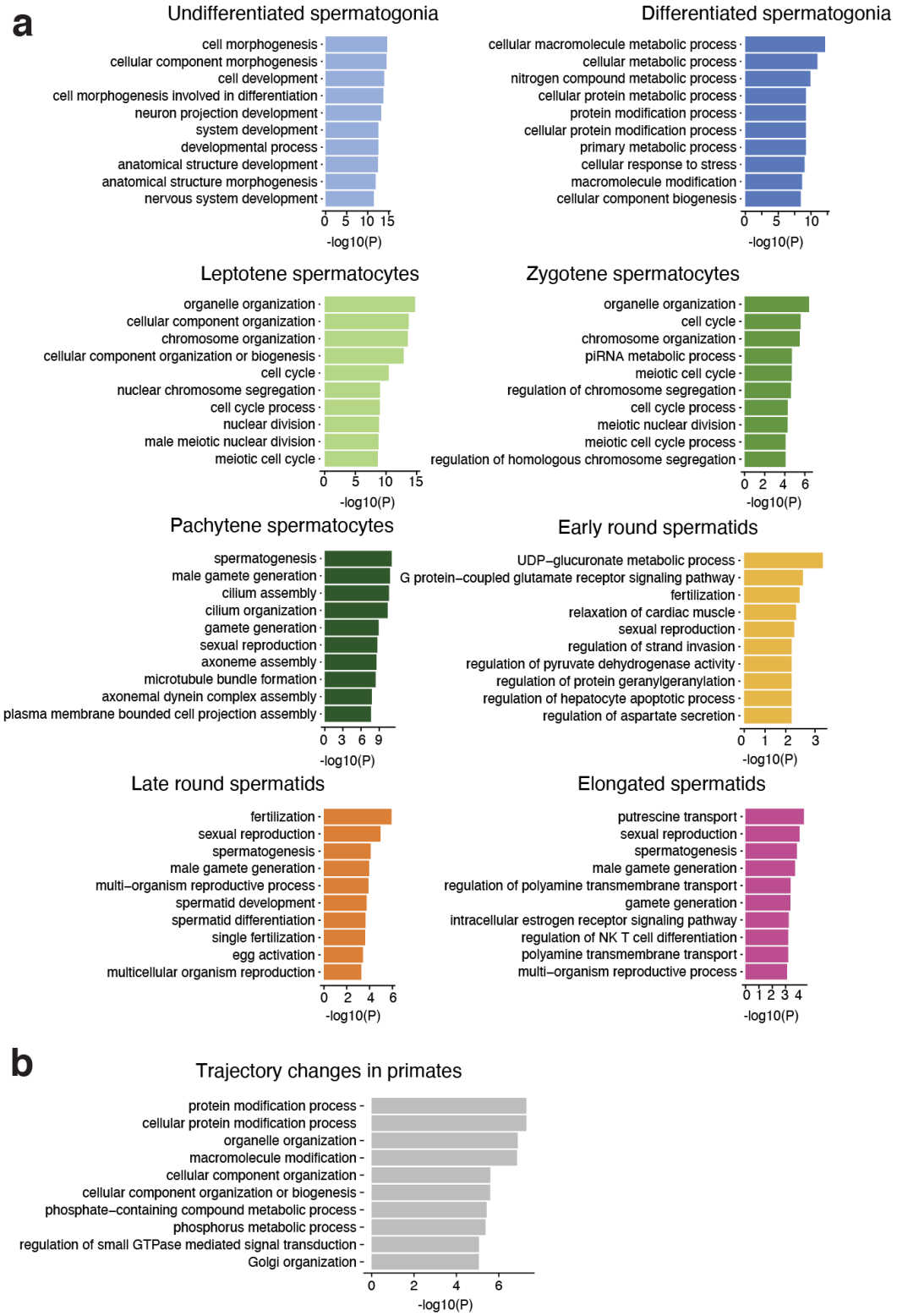


Figure 28: Primate GO analysis. a) *enriched GO-terms for genes with conserved predominant expression in a specific cell type.* ***b)*** *enriched GO-terms for genes with an expression trajectory change. Adapted from Murat et al. (2023).*

Taken together, we found conserved and diverged aspects of simian testicular evolution. Among the one -to-one orthologous genes we found a high fraction with a conserved gene expression trajectory. For three cases of expression trajectory change in the great apes we successfully validated the results using smISH, further supporting the robustness of our transcriptomic analyses.

3.1.7 Sex chromosomes

In mammals, sex chromosomes have emerged independently twice from different sets of autosomal chromosomes[121,252-254]. Previous work has shown the therian Y chromosome emerged slightly more than 180 million years ago, shortly before the split of eutherians and marsupials[121]. Consequently, marsupial and eutherian sex chromosomes have evolved mostly independently from each other[121,252]. The first monotreme Y chromosome also emerged at a similar time ~180 million years ago [121]. Since then, monotreme sex chromosomes have expanded to five pairs of X and Y[254].

Differentiation of the sex chromosomes occurred through recombination arrest between X and Y[255]. The first region on the Eutherian Y chromosome, stratum 1 (S1), which stopped recombining with the X chromosome contains four genes in present day eutherians, among these is *SRY*, a key gene in male sex determination[256-258]. Over evolutionary time the sex chromosomes differentiated further, as can be observed from the extensive gene content remodeling[167].

As the testis is an exclusively male tissue, it is of particular interest to understand the expression patterns of the genes present on the sex chromosomes, the sex-linked genes. Previous work in mice for example showed an abundance of X-linked genes specifically being expressed in spermatogonia, based on cDNA subtraction experiments[259].

3.1.7.1 X-linked gene expression

To gain insight in sex-linked gene expression, we analyzed expression of X-linked genes across mammals. To identify genes which have accumulated on the X chromosome for testicular function, we restricted our analysis to genes with testis specific expression according to bulk tissue RNA-seq data[158]. Using our single cell resolution, we further dissected, which cell type is the predominant source for the observed expression. In this analysis we found an excess of X-linked genes expressed in spermatogonia across all mammals and, in therians, also in Sertoli cells, when we compare it to the expected number of expressed X-linked genes, based on chromosomal gene content (Figure 29a). In platypus, this excess is especially pronounced on

X1 and X5, which have large sexually differentiated regions (SDRs) (Figure 29b). For human and macaque, we also find an excess of X-linked transcripts in leptotene spermatocytes. However, gene expression trajectory analysis suggests that these transcripts were already expressed in undifferentiated spermatogonia, indicating transcript carry over as a potential source for these.

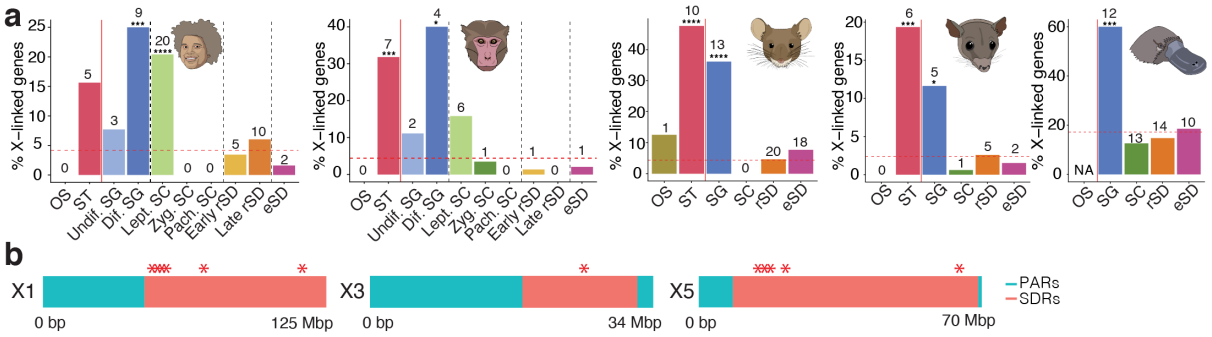


Figure 29: X-linked expression. a) *expression of testis and cell type specific X-linked genes per cell type. Dashed line marks expected number of genes.* ***b)*** *distribution of testis specific expressed in spermatogonia on platypus X1, X3, and X5. Pseudoautosomal regions, PARs; sexually differentiated regions, SDRs. Adapted from Murat et al. (2023)*

3.1.7.2 Differentiation and differences of X- and Y-bearing spermatids

The reductive meiotic divisions of spermatocytes result in two populations of spermatids, which can be distinguished based on their sex chromosomal content. The X-bearing spermatids and Y-bearing spermatids. We managed to distinguish these populations in our data through use of a Gaussian mixture model based on X transcript content of spermatids. By specifically adding sex-linked genes to the parameters used for low dimensional representation we also resolved these two populations in the UMAP projections across mammals (Figure 30). The X-bearing spermatids we identified have a significantly increased expression of X-linked transcripts, compared to the Y-bearing spermatids. In single cell testicular data sets which have previously been published, we could not distinguish the X- from the Y-bearing spermatid population, likely due to transcript exchange through cytoplasmic bridges (Figure 30 a,b)[201,260].

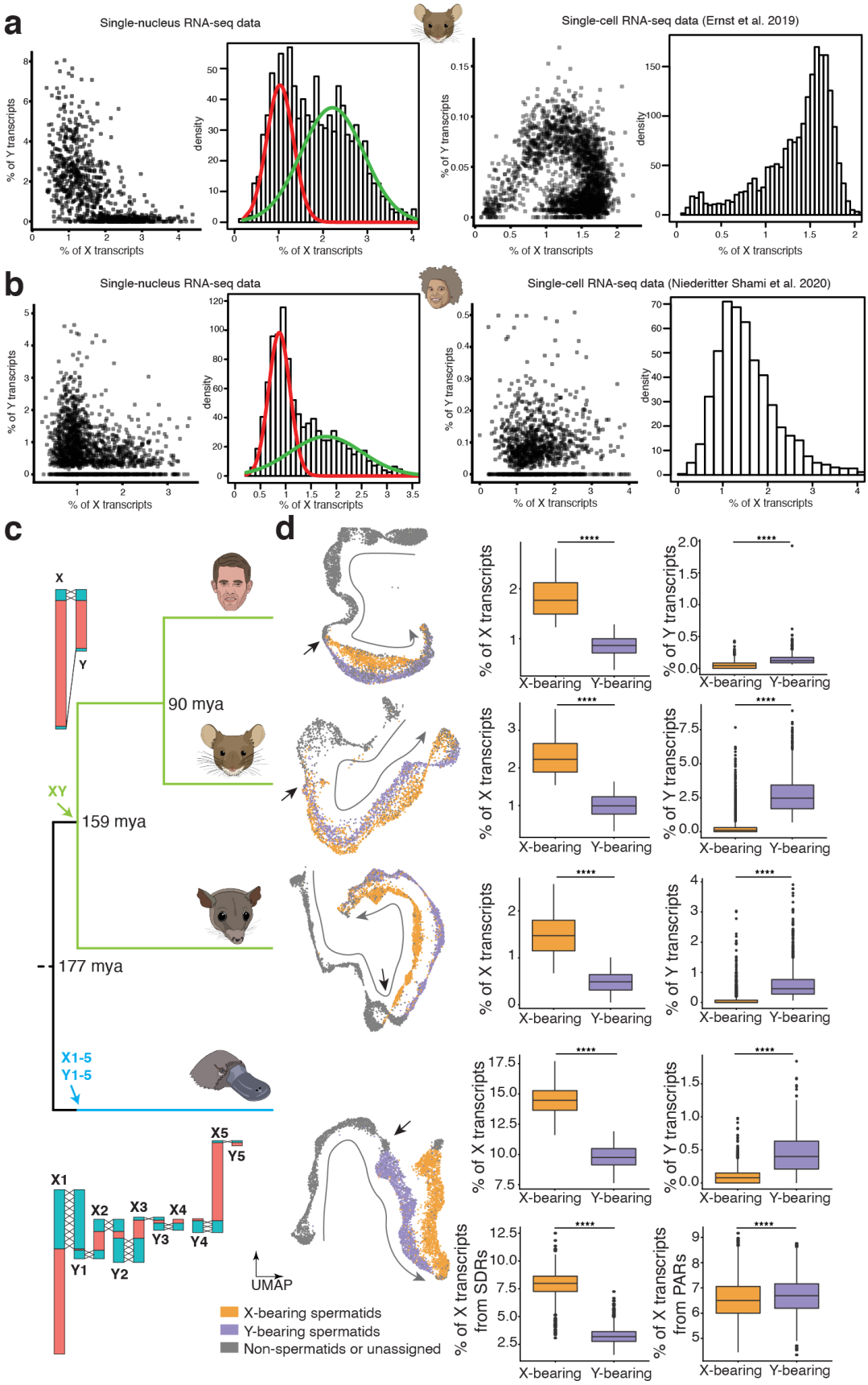


Figure 30: X- and Y-bearing spermatids. *Comparison of snRNA classification with single cell data from mouse* ***(a)*** *and human* ***(b)****. **c)** mammalian lineage with originations of sex chromosome systems. **d)** UMAP representation of separated spermatids (left) and expression of sex linked transcripts (right). Adapted from Murat et al. (2023).*

Based on this distinction between X- and Y-bearing spermatid population, we wanted to further dissect, what the transcriptional differences between these two populations are. For this, we performed a differential expression analysis. We considered genes with significant expression differences according to a Wilcoxon rank-sum test with Bonferroni correction and which are expressed in at least twofold more cells in one population as differentially expressed between the spermatid populations.

In this analysis we found 3 (in mice) to 71 (in humans) differentially expressed genes between the spermatid populations (Figure 31). Overall, the majority of these genes are located on the X chromosome and show an increased expression in the X-bearing spermatids. Among the differentially expressed genes are also gametologs, such as *DDX3X* and *DDX3Y*, which are X-Y homologous genes. This is expected, as the sex chromosomes are the main differentiating characteristic between the two populations and X chromosome has a higher gene content than the Y chromosome. Interestingly, we not only found sex-linked genes differentially expressed, but also a total of 20 autosomal genes across species. Among these are genes such as the lncRNAs *FAM230C* (from chromosome 13) and *FAM230F* (from chromosome 22), with an enriched expression in human Y-bearing spermatids.

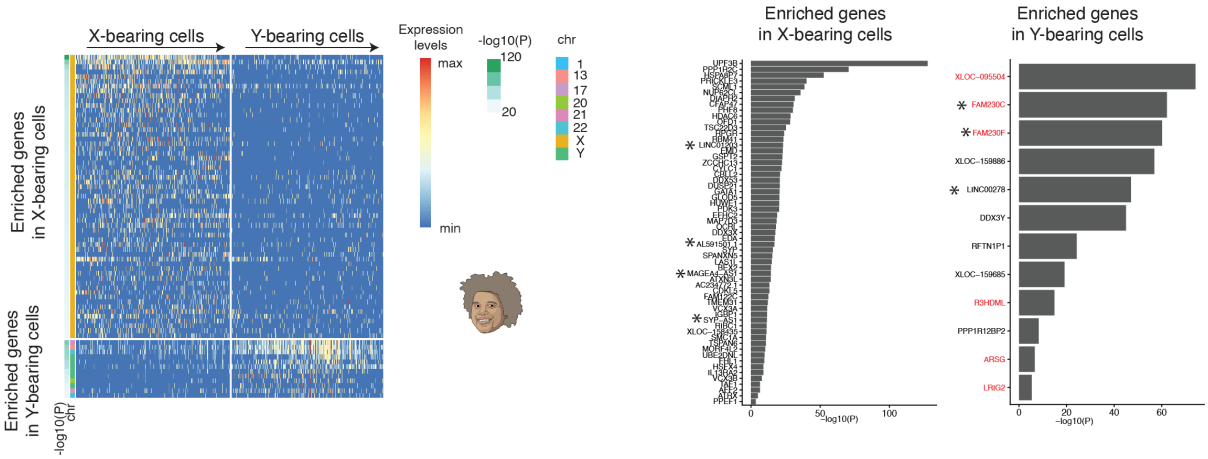


Figure 31: Differentially expressed genes in X- and Y-bearing spermatids. *Significantly enriched genes in either population. Autosomal genes are marked by an asterisk Adapted from Murat et al. (2023).*

3.1.7.3 Meiotic sex chromosome inactivation (MSCI)

During spermatogenesis, it is not only the expression from the sex chromosomes which is of interest, but also the specific repression of transcription from the sex chromosomes in the process called meiotic sex chromosome inactivation (MSCI). This has been described in therian males[125]. Transcription from the sex chromosomes becomes inhibited through their sequestration in the sex body[132]. This process is accompanied by specific chromatin marks

such as phosphorylation H2AX[125]. Previous work on MSCI did not find these therian MSCI chromatin marks in monotremes[135]. To analyze MSCI, we analyzed the relative expression of X-linked genes in comparison to autosomal genes per nucleus and resolved this for progression through spermatogenesis (Figure 32a). In this analysis we found a clear decrease of X-linked transcript contribution in spermatocytes across mammals, in accordance with MSCI in spermatocytes. To understand the decrease in X-linked transcription in the platypus, despite the previously described lack of MSCI associated chromatin marks, we further dissected X-linked transcription in platypus. For this we leveraged a new platypus genome, which had become available[123]. This genome for the first time contained detailed definitions of the sexually differentiated and the pseudoautosomal regions of platypus sex chromosomes. We found that genes from the PARs show a similar expression pattern to genes from autosomes. The decrease of sex-linked transcription in spermatocytes is restricted to genes in the SDRs, showing that platypus MSCI is restricted to the sex chromosome regions, which cannot pair homologously (Figure 32a).

Seeing that MSCI, at least in regards to chromatin context, is achieved through a different process between monotremes and therians, we wanted to probe how efficiently the SDRs become inactivated in monotremes compared to therians. For this, we scrutinized transcript abundance changes between spermatogonia and spermatocytes in the mammalian species using differentially expressed genes between spermatogonia and spermatocytes and a $\log_e$-fold change of more than 0.25. Using this analysis, we found one MSCI escapee in mouse, opossum, and platypus, while we found none in human. This indicates a similar level of MSCI tightness across mammalian species (Figure 32b).

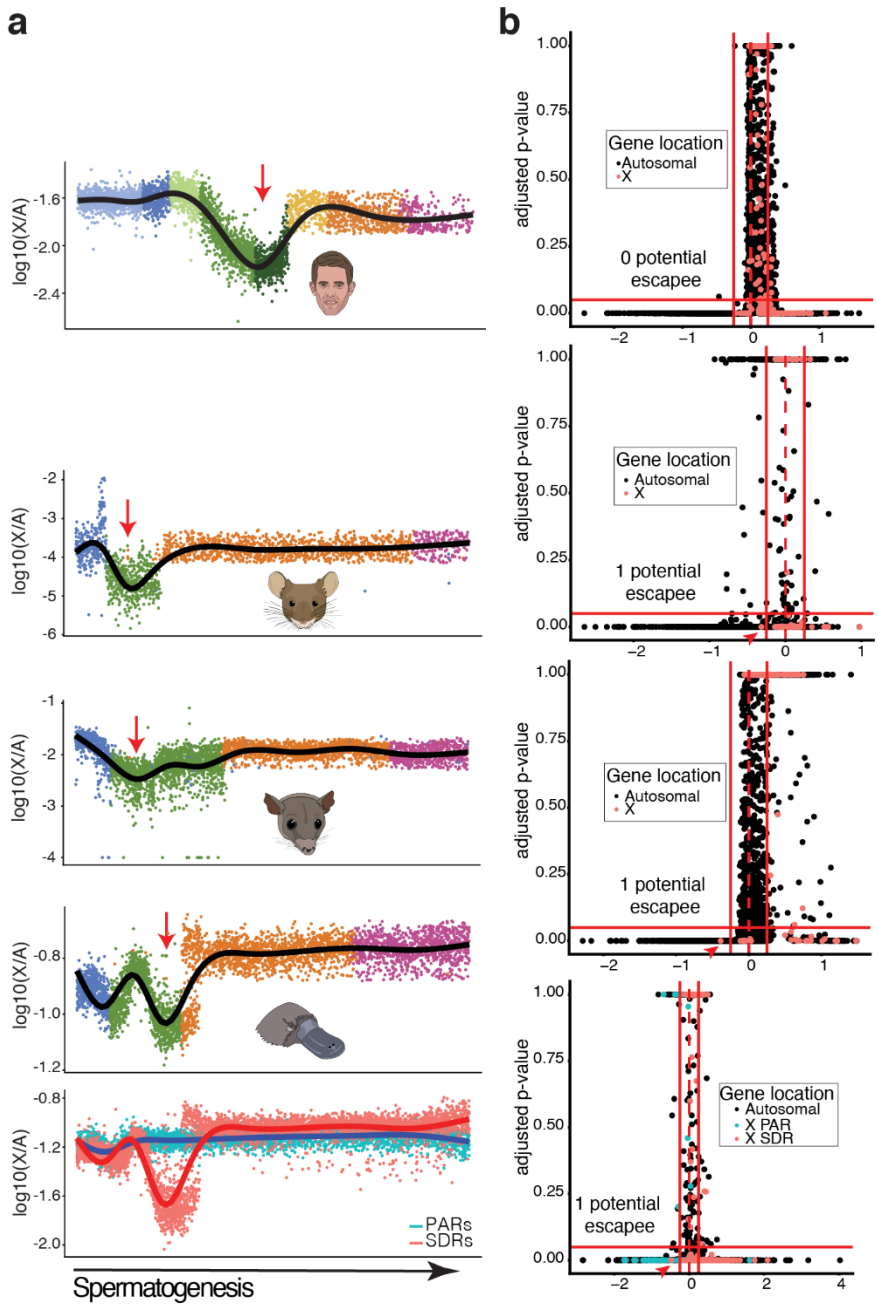


Figure 32: MSCI in adult testis. a) *X-to-autosome ratio of germ cells during spermatogenesis for human, mouse, opossum, and platypus.* ***b)*** *log_e-fold expression change between spermatocyte and spermatogonia of genes on the x-axis and adjusted p-value of differential expression analysis on the y-axis. Adapted from Murat et al. (2023).*

In summary, we showed the accumulation of X-linked transcripts in both Sertoli cells as well as spermatogonia across mammals. Furthermore, we used the resolution afforded by our snRNA-seq data to separate X- and Y-bearing spermatids. These two populations are not only distinguished by sex-linked gene expression, but also by expression originating from the autosomes. Beyond this, our analysis of X-linked expression along spermatogenesis captured MSCI across therians and also for monotremes. Despite mechanistic differences this process is similarly efficient across branches.

3.2 Ovary

The ovary is the female organ of gametogenesis. Interestingly, the reservoir of germ cells is already established prenatally, but the germ cells then remain “locked” in meiosis I, until reproductive age is reached[96-99]. Per reproductive cycle then one to few germ cells progress to meiosis II, which is only completed after fertilization[104,105]. To get a better understanding of how the female germline is established and dissect germline evolution, I generated snRNA-seq data for developing ovaries. For this analysis, I focused on the evolution of glires and thus generated data for both glires lineages, lagomorpha and rodentia. Among lagomorpha, I generated data for *Oryctolagus cuniculus* rabbit. Among rodents I generated data for two murinae (*Mus musculus* mouse and *Rattus norwegicus* rat). Further I generated data for chicken as an evolutionary outgroup.

Developmental data from human has shown, that multiple stages of germ cell development can be found in a single sample from PCW 17-20 ovary[94]. To account for heterochronies in development and maximize the chance for capturing diverse germ cell types in samples of only one developmental stage, I assessed previously published bulk RNA-seq expression profiles of developmental ovaries[158].

In this analysis I found, expression of marker genes for mitotically active oogonia (*Kit*, *Nanog*), Retinoic acid responsive oogonia (*Stra8*), Oogonia committed to meiosis (*Dmc1*, *Sycp3*), and Oogenesis (*Zp3*, *Wee2*)[94]. These markers are simultaneously expressed in E14.5 mouse, E17 rat, newborn rabbit, and E20 chicken samples (Supplementary Figure 1).

Based on this, I generated seven snRNA-seq libraries for these species, including at least one biological replicate per species. Within these libraries 54,243 cells passed the filtering process.

3.2.1 Cell type assignment

I assigned the cell types based on the expression of marker genes. For differentiation of somatic cells from germ cells I used the expression of *Wt1* and *Ddx4* respectively[261,262]. The somatic cells I further defined based on the expression of additional markers. Such as *Foxl2* for granulosa cells, *Lama2* for stromal cells, *Synpo2* for adrenal gland cells, *Vwf* for endothelial cells, *Epcam* and *Upk3b* for epithelial cells[14,51,94] (Figure 33).

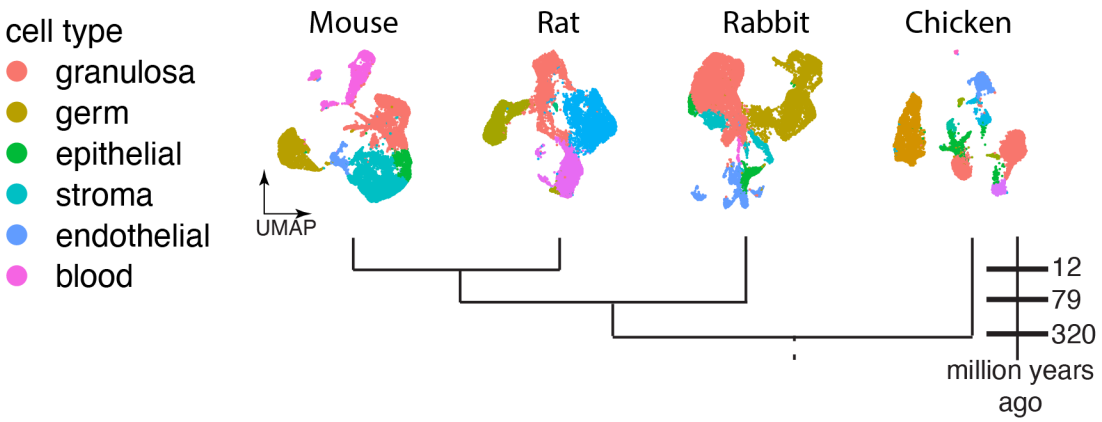


Figure 33: Developmental atlas of mammalian ovary. *UMAP embedding of developmental ovary for mouse, rat, rabbit, and chicken (left to right). Branch length on indicated phylogeny not to scale.*

In comparison to the testicular data, I found a lower fraction of germ cells among the captured nuclei (11,725 germ cells out of 54,243 total cells; ~22 %). This is expected, as in the developing ovary, the majority of cells are somatic and in concordance with other developmental ovary data [14,94]. Furthermore, there is no continuous production of oocytes, which would skew this ratio towards germ cells.

While I found expression of genes corresponding to multiple germ cell stages in the germ cells, the differences between the nuclei are not sufficient to clearly distinguish them in glires, (Figure 34).

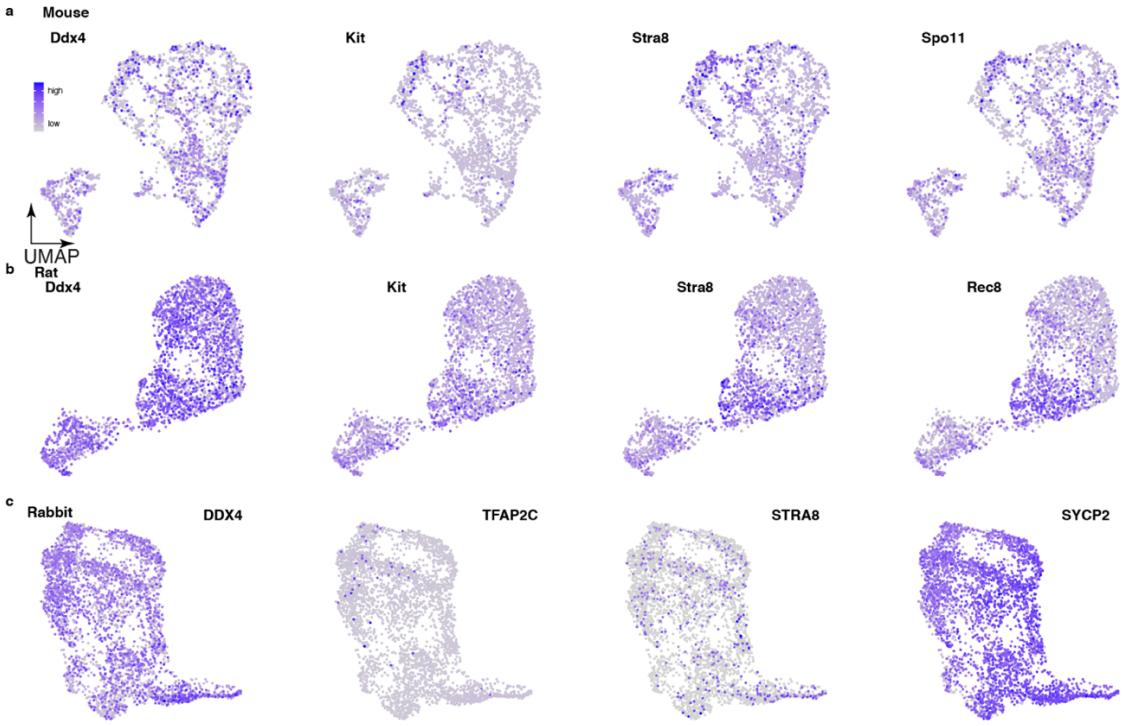


Figure 34:Ovarian germ cells. a-c) *Expression of germ cell marker genes in mouse, rat, and rabbit (top to bottom).*

Seeing the well-established pattern of germ cell commitment to meiosis over several weeks in human, this result highlights lineage specific differences in regulating commitment to meiosis[94]. In the chicken dataset, the resolution is sufficient and I identified mitotic oogonia, retinoic acid responsive oogonia, and oogonia committed to meiosis, based on the expression of stage specific marker genes (Figure 35).

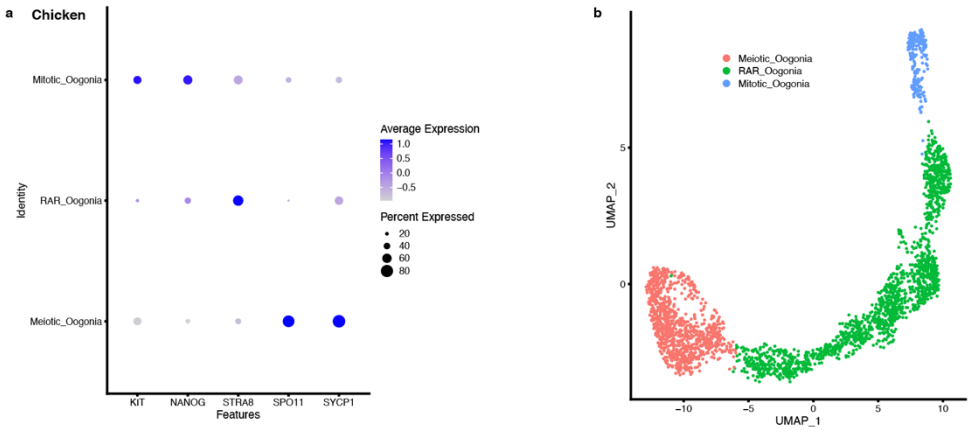


Figure 35: Chicken germ cells. a) *Expression of marker genes in chicken germ cell subtypes.* ***b)*** *UMAP embedding of chicken germ cells.*

3.2.2 Divergence of ovarian cell types

To test, whether a difference in divergence can be observed for ovarian cell types, I performed pairwise comparison between pseudo-bulks of the mouse cell types and the same cell type in the other species. In this analysis, I observed a faster divergence of the germ cells compared to the somatic cell types. While this trend is based on few species, it overall mirrors the trend observed in the testis. This supports the notion, that germ cells in general evolve faster than somatic cell types (Figure 36).

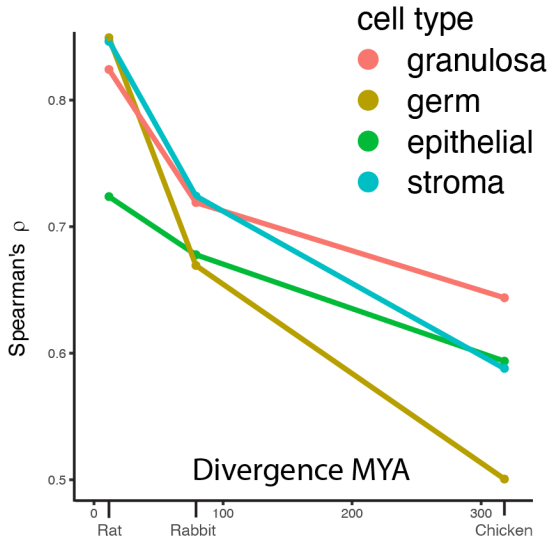


Figure 36: Divergence of ovarian cell types. *Pairwise comparison of mouse with rat, rabbit, or chicken pseudobulk cell type transcriptomes.*

3.2.3 MSCI in chicken germ cells

Seeing that I was able to separate multiple substages of chicken germ cells, I also assessed the relative abundance of Z-linked transcription during oogenesis. Similar to mammalian spermatogenesis, in which MSCI reduces sex-linked transcription in the heterogametic sex during meiosis, the relative contribution of Z-linked transcripts is decreased in chicken oogenesis (Figure 37a). This is in accord with previous work, indicating the presence of MSCI in chicken germ cells[263]. I find a similar trend, when I look at the absolute counts (Figure 37b).

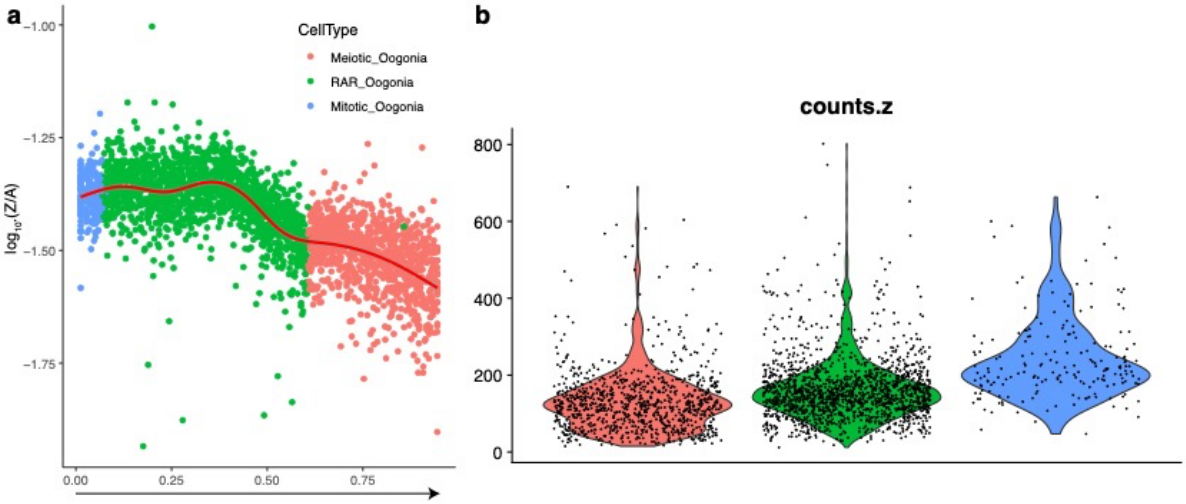


Figure 37: Z-linked expression in chicken germ cells. a) *Z-to-autosome ratio of chicken germ cells progressing through oogenesis (indicated by arrow).* ***b)*** *Number of expressed Z-linked genes in the germ cell subtypes.*

4 Discussion

For my thesis, I presented the snRNA-seq atlas of the adult mammalian testis transcriptome, we created based on the data I generated. These data cover the three main mammalian lineages - eutherians, marsupials, and monotremes – as well as chicken as an evolutionary outgroup. I showed, how we systematically investigated this data to gain insights into the evolution of the mammalian testis and spermatogenesis. I presented, how we based on these analyses, showed that especially the postmeiotic cell types drive the rapid molecular evolution of the testis. Furthermore, we found conserved and diverged aspects of testis expressed genes for both the amniote and simian testis and connected these results to functional aspects. Beyond this, in our analysis of sex-linked gene expression, we found the X chromosome is enriched for genes expressed in specifically in early spermatogenesis and we, for the first time, provided transcriptome evidence for MSCI in platypus.

In an attempt to further the understanding of evolution female gametogenesis in mammals, I also generated snRNA-seq data for developing ovaries. This data comprises of three glires species and chicken as the outgroup. Based on these data, I generated a snRNA-seq atlas of developmental mammalian ovary. In this data I found the somatic cell types diverge slower than the germ cells. In the chicken germ cells, I found support for the presence of MSCI.

In this section of my thesis, I will summarize the main findings from the results section and discuss them in regards to previously published work.

4.1 Testis

Based on morphological and bulk tissue transcriptomic data, the testis has been described as a rapidly evolving tissue[157-162,166]. With the advent of single cell transcriptomics technologies further progress in understanding the testicular transcriptome and the evolution of it was made[201,202,260,264]. While there are multiple publications on individual mammalian species and few works comparing few species, a comprehensive multispecies work on the testicular evolution of mammals was lacking so far [201,202,260,264]. We closed this gap with the presented dataset, by covering the mammalian lineages of monotremes, marsupials, and eutherians. We used chicken as an evolutionary outgroup. Furthermore, we densely sampled simians to gain insights into the testicular evolution of these closely related species with varied mating systems. Among the primates sampled, we included human. Overall, with ~100,000 nuclei

passing filter and data from eleven species, to my knowledge we generated the most comprehensive single nuclei transcriptome dataset for adult amniote testis[202].

Based on the sampling pattern, we not only identified aspects of testicular biology which are conserved across mammals, but also particularities of the individual lineages. Taken together, the data presented furthers our understanding of testicular evolution.

4.1.1 Testicular cell types

Upon assignment of the cell types, we consistently observed across all species that germ cells form a continuity of cells separated from the somatic cell types. Taken together with our identification of continuously changing marker gene expression, our data highlights and reinforces spermatogenesis as a developmental process of gradual changes, even when drastic changes in the biological processes take place (e.g., commitment of mitotically dividing spermatogonia towards meiosis)[31]. The list of 133,222 cell type marker genes across amniotes can be useful in connecting the cell type specific gene expression to functional relevance for biological processes. E.g., across amniotes we find expression of *SYCP2* as marker gene for spermatocytes. In the meiotic spermatocytes SYCP2 plays a pivotal role in establishing/maintaining the synaptonemal complex[203]. On the other hand, in cases when single species experiments would ascribe a function of a gene in a basic biological process, we can cross-reference our list of marker genes to explore whether it is a species or lineage specific function. In these cases, when expression is only present in one or few lineages, further experiments would be required, as the lack of observed expression could be caused by a dropout event in our snRNA-seq data. Based on the methodology of the marker gene identification, marker genes can be markers for multiple cell types simultaneously, thus the list of cell type exclusive marker genes is shorter than the list of all identified 133,322 cell type marker genes.

Based on the assigned cell types, we also observed commonalities and differences in the testicular cell type composition between species. Across amniotes the germ cells are more prevalent than somatic cells in the dataset I generated. This is consistent with the expectation for fertile individuals[265,266]. Among germ cells we found the composition to be changing between species. For mouse we found the lowest fraction of spermatogonia among germ cells compared to the other species. This is an expected observation, as depending on species, there are differences in the strategy used for mitotic expansion of germ cells[111,114-117]. The

mouse system with long chains separating commits a large number of spermatogonia simultaneously to become meiotic spermatocytes. Thus, a larger fraction of germ cells is in meiosis or postmeiosis. These observed differences also highlight a further point. While mice are a suitable model organism for understanding biological principals, studies based on mouse testis can only yield limited insights into the maintenance of the stem cell niche of other lineages, due to these differences.

While the cell type composition is consistent within species, conclusions drawn from this assessment need to be carefully evaluated and should be confirmed by further experiments, for example by histological sections and imaging. This is because the data captured reflects the nuclei input into library generation. The testicular cell types differ vastly in their role and thus in the composition of the cellular membranes. E.g., Sertoli cells are large cells with a basal and a luminal section and are spanning the blood-testis barrier, while elongated spermatids have very compacted cytoplasm and the plasma membrane is suited for short term survival outside the organism[31,36]. The lysis conditions for both of these and all the other testicular cell types should be optimized, so the eventually captured nuclei reflect the true composition. The consistent capture of nuclei for most testicular cell types across amniotes shows, I was overall successful with the used lysis conditions. Yet, the lack of unambiguously assigned Sertoli cells in platypus testis shows the captured nuclei do not reflect the cellular composition of the source tissue in its entirety, as Sertoli cells are present in the testis of platypus[113].

Among the testicular cell types, we observed differences in mRNA content. Especially the nuclei of pachytene spermatocytes have a high content of mRNAs, while differentiating spermatogonia are low in mRNA content. This is the expected pattern for mRNA content, as it has also previously been observed in mouse and other mammalian species[188]. Despite the differences in cellular composition and mRNA content, we found similar numbers of annotated protein coding genes per cell type, when we assessed the cell type pseudobulk transcriptomes. This observation indicates that our cross-cell type comparisons are robust, as further supported by downsampling analysis. While a higher diversity in the transcriptome of the postmeiotic cells is expected, due to the permissible chromatin state and the widespread transcription, a large fraction contributing to this diversity are transcripts not coding for proteins. Furthermore, while we reached a sequencing depth of ~250 million reads per

library, the libraries have not been sequenced to comparable depths as previous deep sequencing projects aimed at illuminating the testicular transcriptome diversity[188,267]. Together, these factors can explain the observation of the similar numbers of protein coding genes in the pseudobulk cell type transcriptomes, even though a higher number of distinct transcripts would be expected for the pachytene spermatocytes[188].

4.1.2 Gene expression evolution

In previous bulk tissue transcriptome analyses of mammalian tissues, using principal component analysis, the first principal component was dominated by the tissue signal and not the lineage signal[157]. In our dataset we recapitulate a similar observation, albeit on the cell type level. This indicates that gene expression programs of spermatogenesis, which have been ancestrally shared across mammals and amniotes are captured within the dataset we created and that cell type differences reflect a continuous accumulation of changes. This is, despite the rapid testicular evolution and the long divergence times across species covered within this work.

The continuous accumulation of changes is also reflected in the gene expression trees we created for both whole tissue pseudo-bulk transcriptomes as well as the per cell type pseudo-bulk transcriptomes. The unexpected grouping of the gibbon could be attributed to the individual used in our work. This we could not assess further, due to a lack of a secondary biological replicate. The unexpected lineage grouping in the gene expression tree for the other somatic cells, I would attribute to differences in the cell types captured for the different species and the abundances.

4.1.3 Rapid evolution of late spermatogenic cell types

Based on the branch lengths as well as the pairwise comparison of cell type transcriptomes we found that especially the late spermatogenic cell types show signatures of the rapid evolution associated with the testis. We also made this observation in the per cell analyses. We found the cells in later spermatogenic stages express genes with both signatures of increased positive selection, which is reflected as an increased fixation rate of changes, as well as less negative selection due to pleiotropic constraints, which is reflected as a lowered rate of removal of changes. The transcriptomes of the later spermatogenic cell types are also characterized by being younger, more lineage specific, as well as showing an increase of intergenic transcripts. Based on this, we can now attribute the observations of rapid testicular

evolution made from bulk tissue transcriptome data and previous single cell analyses, to the late spermatogenic cell types for the main mammalian lineages and chicken.

This observation of rapid evolution in the late spermatogenic cell types highlights the importance of the testis in evolutionary innovation and points to these cell types, when the testis has been described as “the birthplace of new genes” in previous work[189]. There are factors, which presumably contribute to this. One factor is, the spatiotemporal profile of the expressed genes is very specific. Thus, the testis expressed genes are under less pleiotropic constraint making them more susceptible to change[158,268]. Another factor is that postmeiotic spermatids are haploid, thus are affected by haploid selection. This means all alleles become visible to selection. Thus, both positive as well as purifying selection can act more efficiently on haploid cells[191]. This also can give rise to a genomic conflict within the individual between the diploid somatic cell types and the haploid germ cells, furthering the fixation of changes[191,269-272]. A third factor is the chromatin state observed in later spermatogenic cell types, which is characterized by open chromatin permissible for transcription, prior to the histones being exchanged for the protamines and the tight paracrystalline chromatin arrangement in the spermatid head[144,145,149,188,228]. While the permissible chromatin state allows for many genetic elements to be transcribed, the phenotypic effects of these transcripts are buffered translationally through decreased translational efficiency[159,167,188,273]. Another aspect of sperm biology further helps in containing deleterious phenotypes at this stage, the removal of spermatids through apoptosis[274-276]. These factors together with the strong selective force exerted by sperm competition, the postcopulatory competition between individuals for fertilization, likely have driven the rapid evolution of especially the late spermatogenic cell types[160-162,164,165].

In these analyses we also found an increase of genes with a fertility phenotype expressed in later stages of spermatogenesis, together with a decreased expression of genes which are lethal when knocked out. Among the fertility genes are also lineage specific genes like *H2al2a*, a mouse specific histone variant which promotes apoptosis in spermatocytes[277]. Another mouse specific fertility gene is *D1pas1*, which is a helicase required for correct piRNA maturation in spermatocytes[230,278]. Interestingly, these genes originated as a retrocopy from X chromosome linked genes onto an autosome[230,231,279]. The move from sex chromosome

onto an autosome, ensures that expression is possible in all postmeiotic spermatids, even those not bearing the specific sex-chromosome.

4.1.4 Testicular gene expression trajectories

So far, I have discussed insights we gained from analyzing the global expression patterns in the dataset I generated. To get an understanding of which genes are contributing how to the observations we made so far, we then traced the expression patterns of individual genes across spermatogenesis. Through this we found 389 genes for amniotes and 1,687 genes for simians with conserved expression trajectories. These genes most likely reflect the core transcriptional programs for testicular cell types for these lineages. This is supported by our GO term enrichment analysis, in which we found the conserved genes with peak expression in a specific cell type are enriched for terms associated with the given cell type. Furthermore, we found genes with a fertility phenotype when knocked out are more conserved than genes, which do not show a fertility phenotype when knocked out. This supports the assumption of the conserved genes reflecting a core transcriptional program, as the generation of fertile sperm is a central function of the testis. Expanding on this, the list of genes with conserved expression trajectories could be used for the identification of genes required for reproductive success in amniotes and primates. This could contribute to increasing yields in an agricultural context as well as be helpful in alleviating problems with decreasing fertility rates among humans[280-283].

Using the gene expression trajectories, we also identified genes with expression changes on specific lineages. The changed context in which the gene is expressed could be indicative of neofunctionalization of the gene product in a lineage specific manner[284,285]. Our collaborators Dr. Sofia Boeg Winge and Dr. Kristian Almstrup verified three of the trajectory changes we have called in simians through smFISH. The shift of *RUBCNL* expression from spermatid only expression in great apes, to expression also in spermatocytes in human is one of the cases. *RUBCNL* is an inducer of autophagy and the additional expression in an earlier cell type could be indicative of a stricter quality control mechanism in human compared to apes, which is also reflected in the reduced daily sperm production rate of human compared to other primates[160,286]. *ADAMTS17* is a member of a family of metalloproteases, which are associated with fertility and play role in modulation of the extracellular matrix[287]. In human the expression has changed to not only being expressed in spermatogonia, but also in round

spermatids. This could hint at a role in cell-to-cell communication in the adluminal parts of the seminiferous tubule, beyond the blood-testis barrier[288]. *MYO3B* is a case for a change on the pan lineage, in which expression in round spermatids was lost. As a myosin, a functional connection to intra cellular vesicle trafficking is thinkable, with the pan lineage under stronger sperm competition than the other primate lineages, due to mating system differences, adapting a different trafficking pathway[155,289]. While the results of the smFISH experiments overall confirm our observations from the transcriptome data, the quantitative differences could be explained by the differences in the technique used. The snRNA-seq data only captures transcripts in the nuclei, while the smFISH approach also captures transcripts in the cytoplasm, as they were performed in situ within cross sections of seminiferous tubules. While the change of expression gives an indication that there is a functional change, we did not perform a functional assessment of the tested genes, thus my hypotheses in regards to the functional impact remain speculative.

4.1.5 Cellular communication

As indicated by both the continuously changing marker gene set and the gene expression trajectory analyses of spermatogenic cell types, spermatogenesis is a highly coordinated process in which at specific stages specific gene expression programs have to be activated[153]. Furthermore, this process is coordinated within the seminiferous tubules in the seminiferous cycle[153]. Our cell-to-cell communication analysis confirms the notion observed in previous work, that especially the premeiotic germ cells are communicating among each other and the surrounding somatic cells[260]. At this stage in spermatogenesis, the stem cell niche for the spermatogonia needs to be maintained and the fate decision towards committing to meiosis needs to be made[117]. The decrease in observed interaction among postmeiotic cells, could also reflect a consequence of the blood-testis barrier, formed by Sertoli cells[36,288]. This barrier is important for the maintenance of the differences in the intercellular fluids in the two areas, as well as preventing immunogenic reactions to the spermatids[36,288,290]. While overall, we have identified conserved aspects of cellular communication, such as the presence of CADM1 in Sertoli cells across amniotes, the approach we used cannot give us insights into the full repertoire of such interactions. The receptor-ligand interactions are based on a provided database. As such we can only find interactions for annotated pairs. Further the receptor-ligand pairs need to be found sufficiently expressed in both cell types for which

communication is assessed[237]. To ensure this is not affected by dropouts, a sufficient number of nuclei to a sufficient sequencing depth need to be analyzed. Furthermore, this is only a reflection of expressed receptors and ligands. While this gives an indication the cells might communicate, the spatial arrangement of the cells within the tissue might prevent such communication. As our work was focused on the transcriptomic aspects, we did not verify the interactions we found through secondary experiments.

4.1.6 Sex chromosome dynamics in the testis

In mammals, which all have the XY-system for sex determination, the male is the heterogametic sex[121]. The sex chromosomes arose twice in parallel from different sets of autosomes in therians and monotremes[121,252-254]. Since then, the sex chromosomes have undergone extensive changes of their structure, as well as their gene content. This remodeling is linked the sex-related selective forces[121,167]. As females carry twice as many copies of X chromosomes as the male, the X chromosome has spent 2/3 of its evolutionary time in females and only 1/3 in males[191,291]. This difference favors the accumulation of female beneficial genes on the X chromosome, even when they are slightly deleterious to males. An opposing driver lies in the hemizygosity of the males. As the males only carry one copy of the X chromosome, male beneficial genes are always visible to selection, thus leading to an accumulation of male specific genes on the X chromosomes. In previous studies of somatic tissues, this male specific accumulation has not been observed[292]. Yet, the accumulation has been shown specifically in mouse spermatogonia and later in human[259,293]. We show, the hypotheses of specifically spermatogonia expressed genes accumulating on the X chromosome holds true across mammals. The unexpected observation of excess X-linked genes expressed in leptotene spermatocytes can be explained by transcript carryover, as expression for these can mainly already be observed in differentiated spermatogonia.

Another aspect of sex-linked gene content is genes with a role during meiosis or in postmeiotic cell types. In these cells, expression from the sex chromosomes is prevented either by meiotic silencing of sex chromosomes, or the absence of either the X or Y chromosome due to the meiotic divisions. These genes could maintain their expression by moving onto an autosome[294,295]. Earlier I discussed the expression pattern for *H2al2* in mouse, which is one such case of a retrocopy of a X-linked gene[231].

As indicated, the postmeiotic cells differ in their DNA content. They carry either an X or Y chromosome, except for aneuploidies, never both. In the dataset I generated, we were able to segregate the two populations of spermatids, which in other testis datasets was not possible, because they relied on whole single cells for data generation. This difference in observation between single nucleus and single cell data is because the four spermatids originating from one spermatocyte remain connected through cytoplasmic bridges[296]. While not all transcripts are exchanged among the spermatids via these bridges, significant exchange can still be observed. Due to this, the haploid spermatids can become effectively diploid in regards to these transcripts[296].

The analysis of the separated spermatid groups, revealed that the groups not only differ in the expression of sex-linked genes, but are also distinguished by the expression of autosomal genes. For human, we found among these the lncRNAs *FAM230C* and *FAM230F*, which are linked to genomic disorders due to non-allelic recombination[297]. The most frequent among these microdeletion disorders is DiGeorge syndrome[297]. Despite the difference in expression of FAM230C and FAM230F between X- and Y-bearing spermatids, the prevalence of DiGeroge syndrome in male and female offspring is similar, which could hint at a different function of these lncRNAs in Y-bearing spermatids[298]. The differential transcriptomes between X- and Y-bearing spermatids could open up new avenues for investigating male contribution to the sex ratios of offspring[299].

4.1.6.1 MSCI in mammalian testis

Meiotic sex chromosome inactivation (MSCI) is the process through which mammalian sex chromosomes become transcriptionally inert and sequestered in the sex body during meiosis[125-127]. Previous to our work MSCI hallmarks, which had been established based on work in mice, could not be observed in platypus and thus the presence of MSCI in platypus was questioned[135]. Based on our transcriptomic data together with the newer platypus genome, which is annotated for the sex chromosomes and the extent of both SDRs and PARs, we showed that in platypus sex chromosomes specifically the SDRs become silenced during meiosis, while the PARs maintain transcription[123]. This is, despite the large evolutionary distance between platypus and other extant mammalian lineages. Further, this is observed despite the sex chromosomes originating from different sets of autosomes[121]. While this shows that MSCI can be observed in all mammalian lineages, the observed differences in

associated hallmarks indicates, MSCI as a derived state from an underlying common mechanism. We believe, MSCI is derived from the mechanism of meiotic silencing of unsynapsed chromatin (MSUC). MSUC is an ancestral control mechanism against aneuploidy in the postmeiotic germ cells[125]. Observation of MSCI is also not singular to the mammalian lineage and can be observed for example in insects and fungi[300,301].

4.2 Ovary

As the female tissue producing gametes through oogenesis, the ovary plays a crucial role in reproduction. Seeing this, it is of great interest to understand how evolution shaped the ovarian transcriptome. While there is comparative single cell transcriptome data of adult mammalian ovaries, there are still gaps in regards to the later developmental stages[105].

To close this knowledge gap, I generated snRNA-seq data of the developmental glires ovary, using chicken as the outgroup species. With the aim to specifically investigate the transcriptomes of the germline cells. Seeing that only a minor fraction of captured nuclei in this data are germ cells and the majority are somatic, and furthermore, the germ cells not resolving into subtypes, I did not thoroughly pursue this line of analysis.

In the data I generated, I identified the major somatic cell types of the ovary. This makes my dataset a valuable resource for understanding the somatic compartment of the glires ovary. While this is an exciting starting point for further analysis, the data for each species is restricted to one timepoint. To capture differences in developmental dynamics, further timepoints would be needed.

One interesting result, I gathered from the ovary data is that female germ cells diverge faster than somatic cells. Seeing that mammals have internal fertilization, I would not expect a strong evolutionary force, similar to sperm competition, to shape female germ cell evolution. The accelerated transcript divergence could potentially be reflective of cryptic female choice, in which the female chooses through pre- and post-copulatory mechanisms, which sperm fertilizes their egg[302]. While the male sperm fix lineage specific adaptive changes to avoid cryptic female choice, the female germ cells counteract this avoidance through adaptive changes, resulting in a “tug-of-war” and an observed faster divergence of the germ cell transcriptomes.

The main aim of the ovary work, was to further understand the female mammalian germline evolution. The majority of the germ cell differentiation, from primordial germ cells up to pachytene oocytes arrested in dictyate, happens prenatally[78,95]. Furthermore, this differentiation process is initiated asynchronously and in single human developing ovary samples, all prenatal germ cell states can be observed[94]. Based on this, in combination with expression data from previous bulk tissue RNA-seq work, I chose the timepoints for data generation. Despite this, I could not resolve the germ cells I captured into multiple stages. While it is known that glires germ cells commit at a similar timepoint to meiosis, some degree of asynchronicity is expected, which is explained by diffusion of retinoic acid, the trigger for meiotic commitment, through the ovary[7]. Yet, seeing the degree of asynchronicity of human germ cell commitment to meiosis, further mechanisms are required, to explain this observation. The differences in meiosis commitment dynamics highlights again the importance of evolutionary comparative work, for distinguishing lineage specific properties from shared properties. Especially, when results from model organism species are extrapolated to other species.

4.2.1 Chicken sex chromosome dynamics

While MSCI in all mammalian lineages is well established, there are contradictory studies regarding MSCI in female birds[263,303]. In my data I found a decrease of Z-linked transcription during progression in oogenesis, indicating the presence of MSCI in birds. This would further support the notion of MSUC as the underlying mechanism, which leads to the evolution of MSCI. This could also be indicative of differences in meiotic checkpoint control differences between mammals and chicken. In mammals, the majority of germ line aneuploidies are maternally contributed, apparently due to differences in meiotic checkpoint stringency between males and females [304-306]. The presence of chicken MSCI would point at a reduced amenability to unsynapsed chromatin.

Taken together, the ovary dataset I generated is a good starting point for further analyses and fits well into the existing datasets of human, mouse, and chicken datasets of developing ovaries. Now allowing for more dedicated analyses regarding specificities of the glires branch. Yet, I did not achieve my aim of systematically expanding the knowledge about female germ cell evolution.

4.3 Outlook

4.3.1 Testis

Overall, the dataset I generated is comprehensive in regards to covering the main mammalian lineages and the simian lineages. Closing the gap of orangutan testis transcriptome data would be of importance to better understand the great ape testis evolution as this is the only extant great ape, for which neither I nor, to my knowledge, anybody else has generated testicular transcriptome data with single cell resolution. Yet, sourcing adequate sample for this species is difficult, due to ethical and regulatory barriers[307]. In the species and sample choice for this work, we also decided to forego another interesting aspect of testis biology. The samples we chose are either from continuous breeders, or in the case of the seasonally breeding platypus, from a sample in breeding season. For seasonal breeders the testis can vary vastly in composition between seasons[113,308,309]. Expanding the dataset for seasonal breeders, with samples from different timepoints, would be an exciting avenue for understanding how fertility is regularly reestablished and, which factors are needed for coordinating the initiation of the seminiferous wave.

4.3.1.1 Chromatin accessibility in germ cells

During spermatogenesis germ cells exchange their histones[144]. Furthermore, in late spermatocytes and in round spermatids the chromatin is in a state, which is permissible for transcription[188]. To get a better understanding of these dynamics it would be of great interest, to generate single cell resolution data, which could capture these attributes, such as snATAC-seq. Unfortunately, deep frozen testicular tissue appears to be not well suited for snATAC-seq data generation. Such a dataset would further be of interest in illuminating, in how far the permissible transcriptome state is truly global, or if there are lineage specific particularities, in which certain regions remain repressed in a subset of cell types. Furthermore, the accessibility of chromatin in the elongated spermatids could potentially yield insights into how the condensation of chromatin around the protamines is coordinated. Together with transcriptome data from early zygotes, we could further our understanding of the paternal contribution to early zygote development, i.e. we find a peak which is conserved accessible across assessed species and in the corresponding transcriptome dataset of the zygote find strong expression of a transcript associated with this peak, which would indicate a strong paternal contribution.

4.3.1.2 Translational control in the testis

As we have shown with our data, in concordance with previous work, the testis is a hotbed for evolution. With the high fraction of new lineage specific genes and also the high fraction of intergenic elements transcribed in the postmeiotic cells, a systematic analysis of how translation of these new elements is regulated would be exciting. So far, there is translatome data for adult testis, yet not from small tissue sources or with single cell resolution[159]. This data is based on sequencing of ribosome protected mRNA fragments. While there are some advances towards single cell translatomics, I think, full single cell translatome data generation from deep frozen tissue will be difficult to generate, as the translating ribosomes are located in the cytoplasm and tissue dissociation for obtaining singular cell would likely be accompanied by, at least some cytoplasmic leakage.

Generation of translatome data from small deep frozen tissue pieces, is also complicated due to the low amount of input material and the complexity of the protocol. I have contributed to such a protocol, which by now has been established in the laboratory.

4.3.2 Ovary

4.3.2.1 Developmental mammalian ovary atlas

Among mammalian species there is so far data covering developmental ovaries of eutherians[14]. Yet, data for marsupials and monotremes is missing. Data for these lineages would be invaluable, as they could close the knowledge gap in regards to mammalian ovarian evolution. Monotreme data is of especial interest, as they lay eggs. Understanding, which attributes of ovary development are shared among all mammals, and which attributes are shared between chicken and monotremes, could be a key in shaping our understanding of the transition from egg-laying to vivipary. Data for marsupials is of special interest, as a major phase of development happens postnatally. Together the main mammalian lineages would be covered, yielding insights into which aspects of the developmental mammalian ovary are shared and diverged.

Another aspect, in which the data I have generated so far is lacking, is the coverage of different developmental timepoints. Seeing that multiple drastic changes happen over short periods of time, e.g., the epithelial-mesenchymal fate transition of migrating CE cells, coverage of multiple developmental timepoints is important. Furthermore, these additional timepoints would allow for cell type specific assessment of developmental heterochrony. In

addition, the additional timepoints would yield the multiple substages of germ cells, I attempted to capture, allowing for the dedicated analysis of germ cells, I had hoped to perform.

4.3.3 Gonads and germline

The evolution of gonads and the germline are often investigated in a sex specific manner [190,192,201]. Either through assessing the testis and testicular cell types, or assessing ovary and ovarian cell types. Many of the studies focusing on both together are often restricted to early developmental stages until the stage of sexual differentiation[14]. Investigating both sexes together with a focus on germ cells, including data from adult ovary, could yield exciting insights into the "tug of war" between the sexes, and how lineage specific particularities have been shaped. In this the single cell resolution of data would be of great help, to understand in which specific cell type these specificities manifest. Furthermore, these comparisons could sharpen our understanding of sex specific differences in shared biological processes, such as meiosis and aid in generating hypotheses regarding gonad evolution.

5 Material and methods

5.1 Material

For the adult mammalian testis project, I generated all the sequencing libraries except for two human libraries, which had been generated before. While the majority of the data processing was performed by Dr. Florent Murat, I participated in the data analysis side of this project through discussing intermediate results, performing analyses in parallel, and scrutinizing the analytical outcomes.

The data generation and analysis for the developmental ovary project was performed by me.

5.1.1 Adult testis samples

The adult testis samples had been previously collected in laboratory. I generated libraries for two human (*Homo sapiens*) individuals, three chimpanzee (*Pan troglodytes*) individuals, two bonobo (*Pan paniscus*) individuals, one gibbon (*Hylobates lar*) individual, two macaque (*Macaca mulatta*) individual, two marmoset (*Callithrix jacchus*) individuals, two mouse (Mus musculus) individuals, three opossum (*Monodelphis domestica*) individuals, two platypus (*Ornithorhynchus anatin*us) individuals, and two chicken (*Gallus gallus*) individuals. Two additional libraries from human adult testis were already prepared in the laboratory.

5.1.2 Ovary samples

For the ovary project I generated snRNA-seq libraries for two mouse (*Mus musculus*) individuals embryonic day 14.5, two rat (*Rattus norwegicus*) individuals from embryonic day 17, two newborn rabbit (*Oryctolagus cuniculus*) individuals, and three chicken (*Gallus gallus*) individuals from embryonic days 17, 19, and 19.5. The chicken samples for day 17 and 19 were collected by Dr. Fernando Garcia-Moreno, while the E19.5 sample was collected by me. For all three samples the left ovary was collected. The other samples had previously been collected in the laboratory.

5.2 Methods

5.2.1 Nuclei preparations for single nuclei RNA-sequencing

I isolated nuclei and generated snRNA-seq libraries as described in Murat et al., 2023[192]. Using a nuclei isolation protocol depending on species and tissue used. I established the nuclei isolation protocols based on Krishnaswami[195] and Attar[310].

5.2.2 Nuclei isolation for therian testis samples

From a deep-frozen testis sample, I cut a tissue piece weighing around 5 mg on dry ice. I avoided taking pieces from the tunica albigunea or the rete testis. Then I homogenized the piece in 100-150 µL / mg of HEPES lysis buffer (250 mM sucrose, 25 mM KCl, 5 mM $MgCl_2$, 10 mM HEPES pH 8, 1% BSA, 0.1% IGEPAL and freshly added 1 µM DTT, 0.4 U/µl murine RNase Inhibitor (New England BioLabs), 0.2 U/µl SUPERasIn (ThermoFischer Scientific)) by use of a micropestle and let lysis continue for 5 min on ice. I removed unlysed tissue and debris by centrifugation for 1 min at 100 g and 4 °C and transferring the supernatant to a now reaction tube. Then I pelleted the cells by centrifugation for 5 min at 500 g and 4 °C and fixed the nuclei by resuspending the pellet in 0.67 vol of freshly made fixation solution (1 mg/ml dithio-*bis*(succinimidyl propionate) in PBS) and incubation for 30 min at room temperature. To quench the fixation, I added Tris-HCl pH 8 to a final concentration of 20 mM. To then wash the nuclei, I pelleted the cells by centrifugation for 5 min at 500 g and 4 °C and resuspended in 0.67 vol wash buffer (250 mM sucrose, 25 mM KCl, 5 mM $MgCl_2$, 10 mM Tris-HCl pH 8, 1% BSA and freshly added 1 µM DTT, 0.4 U/µl murine RNase Inhibitor, 0.2 U/µl SUPERasIn). To get the final nuclei suspension, I pelleted the cells by centrifugation for 5 min at 500 g and 4 °C and resuspended in 0.5 vol PBS. I strained these nuclei using a 40 µm Flowmi strainer (Sigma). To estimate the nuclei concentration, I stained an aliquot of this with Hoechst and counted using a Countess II FL Automated Cell Counter (ThermoFischer Scientific). Unless stated I performed this work on ice and carefully resuspended the nuclei immediately before proceeding to a new step.

I also prepared the nuclei for the multiome library according to the protocol above, with final resuspension in 1X Nuclei Buffer (10x Genomics) using testis from mouse.

5.2.3 Nuclei isolation for the other samples

The nuclei isolation I performed for the other species and the ovaries follows the same principal, but fixation was not required. For the developmental ovaries, I used the entire collected tissue sample. I homogenized the piece in 100-150 µL / mg of TRIS lysis buffer (250 mM sucrose, 25 mM KCl, 5 mM $MgCl_2$, 10 mM Tris-HCl pH 8, 1% BSA, 0.1% IGEPAL and freshly added 1 µM DTT, 0.4 U/µl murine RNase Inhibitor (New England BioLabs), 0.2 U/µl SUPERasIn (ThermoFischer Scientific)) by use of a micropestle and let lysis continue for 5 min on ice. I removed unlysed tissue and debris by centrifugation for 1 min at 100 g and 4 °C and transferring the supernatant to a now reaction tube. To then wash the nuclei, I pelleted the

cells by centrifugation for 5 min at 500 g and 4 °C and resuspended in 0.67 vol wash buffer (250 mM sucrose, 25 mM KCl, 5 mM $MgCl_2$, 10 mM Tris-HCl pH 8, 1% BSA and freshly added 1 µM DTT, 0.4 U/µl murine RNase Inhibitor, 0.2 U/µl SUPERasIn). To get the final nuclei suspension, I pelleted the cells by centrifugation for 5 min at 500 g and 4 °C and resuspended in 0.5 vol PBS. I strained these nuclei using a 40 µm Flowmi strainer (Sigma). To estimate the nuclei concentration, I stained an aliquot of this with Hoechst and counted using a Countess II FL Automated Cell Counter (ThermoFischer Scientific). I performed this work on ice and carefully resuspended the nuclei immediately before proceeding to a new step.

5.2.4 Library construction and sequencing

I generated the therian adult testis snRNA-seq libraries using Single-Cell 3' Gel Bead and Library v2 kits from 10x Genomics. For the other snRNA-seq libraries I used the Single-Cell 3' Gel Bead and Library v3 kits from 10x Genomics. For the multiome library I used a Multiome ATAC + Gene Expression Kit from 10x Genomics. For all libraries I followed the manufacturer's protocol. I aimed at an input of 17 000 nuclei per library, corresponding to ~5 000 captured nuclei. I quantified the resulting libraries using a Qubit Fluorometer (Thermo Fisher Scientific) and determined the size with a Fragment Analyzer (Agilent).

I sequenced pooled samples using the NextSeq500/550 High Output Kit v2.5 (75 cycles) (Illumina) according to manufacturer's protocol, with a loading concentration of 2.8 pM. For sequencing I used an Illumina NextSeq 550 system either with 26 cycles for Read 1, 57 cycles for Read2, 8 cycles for index 1, no cycles for index 2 for v2 libraries, or with 28 cycles for Read 1, 56 cycles for Read2, 8 cycles for index 1, no cycles for index 2 for v3 libraries. For both I used a pairwise sequencing protocol. I sequenced each library to a depth of approximately 250 million reads.

5.2.5 Analysis of testis data

The analysis of the testis data was done as described in Murat et al. 2023 and performed by Dr. Florent Murat.

5.2.5.1 Processing of raw read

We used CellRanger v.3.0.2 for processing the platypus and chicken sequencing data and CellRanger v.2.1.1 for the other species. We used the default settings for the mkref function together with genomics and extended annotation files we had previously generated. We also generated references with concatenated exonic and intronic features for use in the expression quantification.

We used the count function with default settings.

5.2.5.2 Identification of useable droplets

To identify usable droplets, we combined kneepoint approach, in which we sorted barcodes by the associated number of reads and identified the inflexion point of the cumulative number of reads together with the fraction of intronic reads, which identifies unprocessed pre-mRNA reads in the nuclei. If present in the given genome, we further refined this identification based on the expression of the nuclear enricher lnc-RNA MALAT1.

5.2.5.3 Per cluster filtering

To account for cell type differences in UMI and mitochondrial RNA content, we generated Seurat objects with the UMI tables of the usable droplets using Seurat R package v3.1.4 with min.cells = 3. We normalized (normalization.method = "LogNormalize", scale.factor = 10 000), found variable features (selevection.method = "vst", nfeature = 10 000), and scaled the data. With this we performed PCA, calculated Louvain clusters (FindNeighbors with 20 dimensions, FindClusters with 20 dimensions and 0.5 resolution. We then removed, independently per cluster, each droplet with lower than first quartile – 1.5 x interquartile range and higher than third quartile + 1.5 x interquartile range values for both content of UMIs and fraction of mitochondrial RNA. This we followed by removing droplets containing multiple nuclei with DoubletFinder v2.0.1 using doubletFinder (PCs = 1:20, pN = 0.25, nEXP = 5 % of present nuclei, pk was calculated using paramSweep, sumarizeSweep and find.pK).

The remaining droplets correspond to the nuclei used for the following analyses.

5.2.5.4 Species dataset integration

We subset the raw count tables based on the identified nuclei and generated Seurat objects, normalized, and identified variable genes as before for each sample separately. We integrated the data for each species independently using FindIntegrationAnchors and IntegrateData (both dims = 1:20).

We ran normalization, scaling, and PCA for each integrated dataset. We performed Louvain clustering, and UMAP projections with the parameters from tale1.

Table 1: Parameters used for clustering and UMAP projection.

Step	Louvain Clustering		UMAP embedding	
Species	dims	Resolution	dims	min_dist
Human	1:20	0.5	1:20	0.3
Chimpanzee	1:20	0.5	1:20	0.3

Bonobo	1:20	0.5	1:20	0.1
Gorilla	1:20	0.5	1:20	0.1
Gibbon	1:20	0.5	1:20	0.3
Macaque	1:20	0.5	1:20	0.3
Marmoset	1:20	0.5	1:20	0.3
Mouse	1:17	0.5	1:17	0.1
Opossum	1:8	0.5	1:10	0.2
Platypus	1:10	0.3	1:10	0.3
Chicken	1:10	0.5	1:10	0.6

5.2.5.5 Primate integration

We integrated the primate data using LIGER v0.5.0 with 1:1 primate orthologues genes from Ensemble release 87. We factorized the joint matrix using optimizeALS (k = 20). The integrated UMAP was generated using n_neighbors = 100 and min_dist = 0.2.

5.2.5.6 Cell type assignment

We assigned cell types based on known marker genes for testicular cell type populations. Sertoli cells are marked by *CLU*, other somatic cell types by *TAGLN*, *ACTA2*, *CD34*, *TM4SF1*, *APOE*, *CD74*, *STAR*, and *CYP11A1*. Spermatogonia are marked by *GFRA1*, *PIWIL4* (undifferentiated), *DMRT1* (differentiated), and *STRA8*. Spermatocytes are marked by *SYCE1* (leptotene), *SYCP1* (zygotene), *PIWIL1* (pachytene), *SYCP2*, *TANK*, and *AURKA*. Round spermatids are marked by *LRRIQ1* (early) and *ACRV1*, *SPACA1* (late). Elongated spermatids are marked by *SPATA3*, *NRBP1*, *PRM1*, and *GABBR2*. We used other metrics such as UMAP coordinates, or UMI count to reinforce this assignment.

5.2.5.7 Pseudotime

We calculated pseudotimes of the germ cells using slingshot v1.2.0. We provided start and end based on the previous cell type assignment. We then ranked the obtained values and normalized them to values between 0 and 1.

5.2.5.8 Identification of marker genes

We separated the germ cells into 20 equal sized bins along spermatogenesis, based on the previously obtained pseudotime values. Then we applied the FindAllMarkers function (only.pos = T, min.pct = 0.25, logfc.threshold = 0.25, return.thresh = 0.05) on the 20 germ cell bins as well as the two somatic cell types.

5.2.5.9 Phylogenetic tree

We used TimeTree for the phylogenetic trees and divergence time calculation (v5 http://timetree.org)

5.2.5.10 Orthologous gene sets

We extracted orthologous genes from Ensembl release 87. With this we generated three orthologous genes sets, with orthologous genes expressed by all species in each set. For analyses across all eleven species, we used 4 498 1:1 orthologous genes. For analyses across the seven primates we used 8451 1:1 orthologous genes. For the cell communication analysis with Sertoli cells, we used the 35 186 genes expressed in human testis and the 1:1 orthologous genes to macaque (13 090), mouse (14 302), opossum (10 865), or chicken (10 515).

5.2.5.11 Pseudo-bulk samples

We generated the pseudo-bulk samples using the Seurat function AverageExpression per cell type with the 4 498 1:1 orthologous genes of amniotes.

We used these pseudo-bulk samples for PCA using prcomp.

We used the cor function to calculate Spearman's ρ correlation between corresponding pseudo-bulks for each cell type across species. To generate neighbor joining trees, we calculated the distance 1- ρ between samples and used the ape R v5.3. We performed bootstrap analysis sampling 1 000 with replacement to test reliability. To remove intra-species variability, we calculated the total tree length until the node between individuals.

5.2.5.12 pli score

We used the pLI score from Dickinson et al., 2016[221] and plotted for each nucleus the median pLI score across genes with at least one UMI.

5.2.5.13 Lethals

We used the phenotype information from IMPC[222]. 4 742 genes have been tested for lethality when knocked out. Among them are 1 139 knockouts, which resulted in lethality. For each nucleus we calculated the fraction of expressed genes with a lethal phenotype over expressed genes tested for lethality (i.e. 4 472 genes). For each nucleus we plotted this as a percentage.

5.2.5.14 dN/dS

We used 11 791 1:1 orthologues protein coding genes across six primates (human, chimpanzee, gorilla, gibbon, macaque, and marmoset)[311] and used for each species and gene the longest transcript to align orthologous protein sequences using clustalo v1.2.4. We then used pal2nal v14 to produce codon alignments. We estimated dN/dS ratios using codeml from

the PAML package71 v4.9 (model = 0, NSites = 0). For each nucleus we plotted the mean dN/dS of expressed for expressed genes for which we had estimated dN/dS.

5.2.5.15 positive selection

We used gene sets previously described as carrying evidence for positive selection (in primates: 331 positively selected genes out of 11 170 tested genes; in mammals: 544 out of 16 419). For each nucleus we calculated the fraction of expressed genes with a signature of positive selection over expressed genes tested for signature of positive selection. For each nucleus we plotted this as a percentage.

5.2.5.16 Spatiotemporal specificity

We used specificity scores from Cardoso-moreira et al., 2019[158] and plotted for each nucleus the median specificity score across expressed genes.

5.2.5.17 transcriptomic age

We used the phylogenetic age score as described[158], with gene ages obtained from GenTree (http://gentree.ioz.ac.cn/)[229]. We plotted the median score per nucleus.

5.2.5.18 TSS enrichment

I used the Seurat workflow for the processing of the RNA modality, using dimensions 1:50 for the PCA. For the ATAC modality I used the Signac package v1.9.0. For the UMAP reduction I used lsi with dimensions 2:50. For the integrated projection, I used the SLM algorithm. I used TSSEnrichment with standard paramteres for TSS enrichment calculation.

5.2.5.19 Intergenic elements

We extracted gene biotypes from Ensembl[311]. We grouped all not protein coding transcripts as intergenic elements together. For each nucleus we plotted the percentage out of expressed genes.

5.2.5.20 Translational efficiency

We used translational efficiency values from Wang et al., 2020[159]. To have cell type specific TEs, we only used genes for, which at least 60 % of transcripts for that gene are found in a given cell type. We calculated per nucleus TE for the expressed genes we found as specific.

5.2.5.21 infertile

We used the phenotype information from IMPC[221]. 3 252 genes have been tested for lethality when knocked out. Among them are 173 knockouts, which resulted in infertility. For each nucleus we calculated the fraction of expressed genes with a infertility phenotype over expressed genes tested for fertility. For each nucleus we plotted this as a percentage.

5.2.5.22 Gene expression trajectories

We compared gene expression trajectories for germ cell expressed genes across human, mouse, opossum, platypus, and chicken based on 2 927 1:1 orthologs from Ensembl 100[312]. Or across primates based on 4 459 1:1 orthologs from Ensembl 87[311]. We only considered genes expressed, when it is expressed in all species in at least 5 % of the cells in one cluster. We used the mfuzz package v2.44.0 for clustering, with Dmin and mestimate to estimate cluster numbers and fuzzification parameters. We used a phylogenetic framework to test for probability of changes in trajectories in specific branches with a 5 % probability cutoff.

5.2.5.23 Trajectory conservation score

We calculated across species conservation scores for 1:1 orthologous gene sets. For this we calculated the $\log_2$-transformed probability of all members falling into the same mfuzz trajectory cluster.

5.2.5.24 Gene Ontology

We used the limma package v3.40.6 and the goanna function with default parameters for GO-term enrichment analysis.

5.2.5.25 Cell communication

We used CellPhoneDB v2 using the human ligand-receptor interaction database. In this analysis we only considered genes ligands or receptors, which are expressed in more than 10 % of cells in each cell type. We used UpSetR v1.4.0 to plot significant interactions across species.

5.2.5.26 Testis and cell type specific genes

We obtained testis specific genes from Cardoso-Moreira et al., 2019[158] (RPKM ≥ 1 in testis and RPKM < 1 in brain, cerebellum, heart, kidney, and liver). We identified gene specificity for somatic cell types using the Seurat function FindAllMarkers function (only.pos = T, min.pct = 0.05, logfc.threshold = 0.25, return.thresh = 0.05). For germ cell genes we used the fuzzy clusters from the trajectory analysis and assigned them based on peak expression. We then plotted the percentage of expressed testis and cell type specific genes which are X-linked and contrasted with the percentage of genes located on a given chromosome among all genes in the genome. We then used the exact binominal test to compare the percentage of testis-specific genes per cell type and chromosome with the percentage of genes per chromosome in the genome.

5.2.5.27 Classification of X- and Y-bearing spermatids

We independently for each replicate, fitted a Gaussian Mixture Model for a bimodal distribution on spermatids based on the fraction of X-linked genes. For this we used the mixtools v1.2.0 package and function normalmixEM. To generate the UMAP projections, we added X- and Y-linked genes to the set of highly variable genes used as gene set for the associated PCA. We identified the differentially expressed genes using the Seurat function FindAllMarkers with default parameters. We calculated Bonferroni corrected P values using the Wilcoxon rank-sums test. We only considered genes expressed at least 0.25-fold higher (log_2-scale) in one of the populations, a P value below 0.01, and they are expressed in at least 10 % of the nuclei of one of the populations.

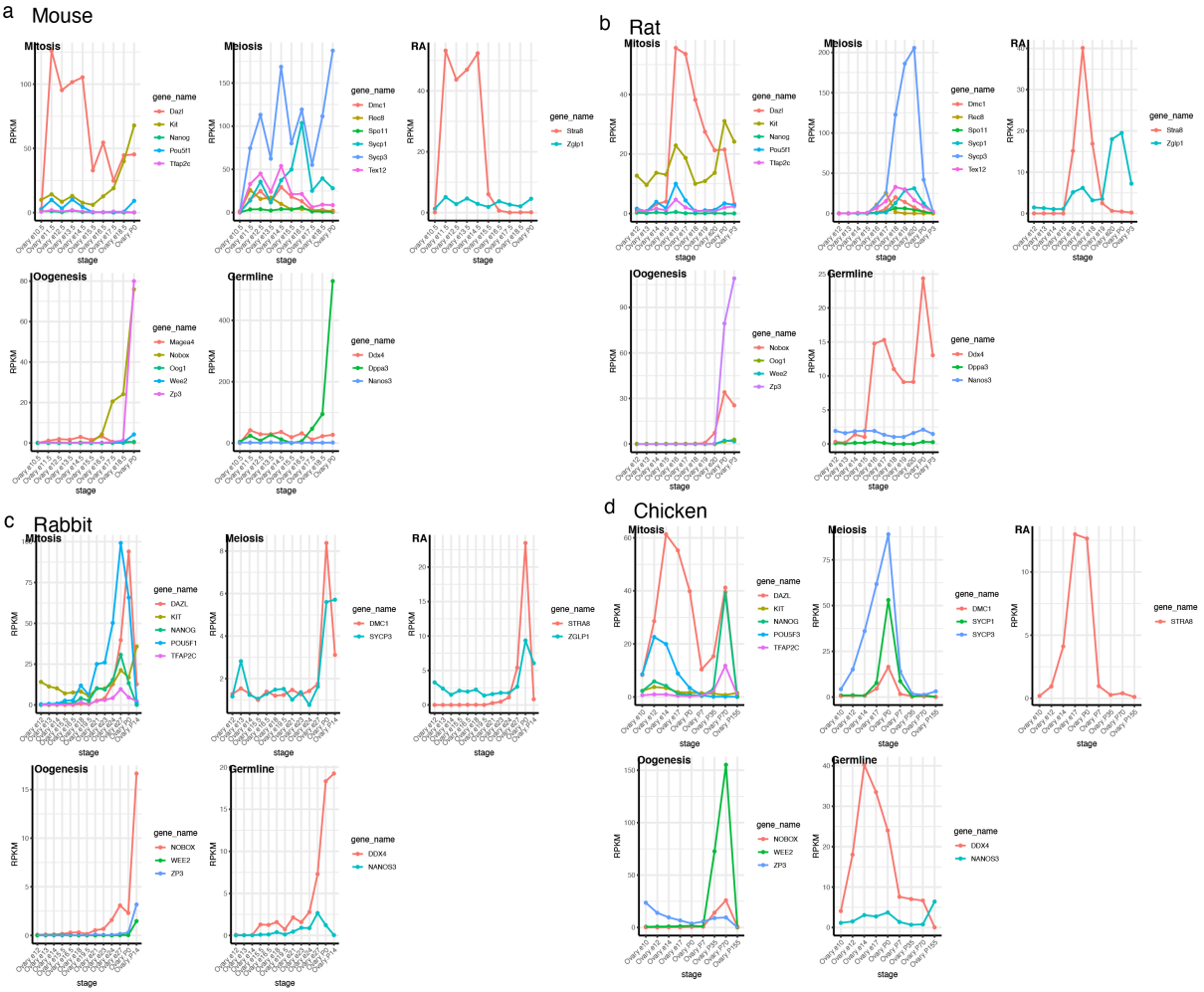
5.2.5.28 MSCI completeness

We identified the differentially expressed genes between spermatocytes and spermatogonia using the Seurat function FindAllMarkers with default parameters. We calculated Bonferroni corrected P values using the Wilcoxon rank-sums test. We only considered genes expressed at least 0.25-fold difference (log_e-scale) between two groups, a P value below 0.05, and they are expressed in at least 10 % of the nuclei of one of the populations.

5.2.5.29 Analysis of ovary data

In the ovary analysis, I followed the principles from the testis analysis. I generated mapping references using cellranger 6.0.2 and ensembl version 103[313]. During Seurat processing, I used FindClusters with 20 dimensions and a resolution of 0.5. For RunUMAP, I used 20 dimensions and 0.3 as minimum distance

6 Supplementary Figures



Supplementary Figure 1: Ovarian germ cells*. Expression of germ cell marker genes in mouse (a), rat (b), rabbit (c), and chicken (d). Data from Cardoso-Moreira et al., 2019.*

7 References

1 Kocer, A., Reichmann, J., Best, D. & Adams, I. R. Germ cell sex determination in mammals. *Mol Hum Reprod* **15**, 205-213, doi:10.1093/molehr/gap008 (2009).

2 Kleiman, M. & Tannenbaum, E. Diploidy and the selective advantage for sexual reproduction in unicellular organisms. *Theory Biosci* **128**, 249-285, doi:10.1007/s12064-009-0077-9 (2009).

3 Edson, M. A., Nagaraja, A. K. & Matzuk, M. M. The mammalian ovary from genesis to revelation. *Endocr Rev* **30**, 624-712, doi:10.1210/er.2009-0012 (2009).

4 Lawrence, B. M., O'Donnell, L., Smith, L. B. & Rebourcet, D. New Insights into Testosterone Biosynthesis: Novel Observations from HSD17B3 Deficient Mice. *Int J Mol Sci* **23**, doi:10.3390/ijms232415555 (2022).

5 Spiller, C., Koopman, P. & Bowles, J. Sex Determination in the Mammalian Germline. *Annu Rev Genet* **51**, 265-285, doi:10.1146/annurev-genet-120215-035449 (2017).

6 Imaimatsu, K., Uchida, A., Hiramatsu, R. & Kanai, Y. Gonadal Sex Differentiation and Ovarian Organogenesis along the Cortical-Medullary Axis in Mammals. *Int J Mol Sci* **23**, doi:10.3390/ijms232113373 (2022).

7 Feng, C. W., Bowles, J. & Koopman, P. Control of mammalian germ cell entry into meiosis. *Mol Cell Endocrinol* **382**, 488-497, doi:10.1016/j.mce.2013.09.026 (2014).

8 Stevant, I. & Nef, S. Genetic Control of Gonadal Sex Determination and Development. *Trends Genet* **35**, 346-358, doi:10.1016/j.tig.2019.02.004 (2019).

9 Chassot, A. A. *et al.* WNT4 and RSPO1 together are required for cell proliferation in the early mouse gonad. *Development* **139**, 4461-4472, doi:10.1242/dev.078972 (2012).

10 Stevant, I. *et al.* Dissecting Cell Lineage Specification and Sex Fate Determination in Gonadal Somatic Cells Using Single-Cell Transcriptomics. *Cell Rep* **26**, 3272-3283 e3273, doi:10.1016/j.celrep.2019.02.069 (2019).

11 Viger, R. S., Mertineit, C., Trasler, J. M. & Nemer, M. Transcription factor GATA-4 is expressed in a sexually dimorphic pattern during mouse gonadal development and is a potent activator of the Mullerian inhibiting substance promoter. *Development* **125**, 2665-2675, doi:10.1242/dev.125.14.2665 (1998).

12 Hu, Y. C., Okumura, L. M. & Page, D. C. Gata4 is required for formation of the genital ridge in mice. *PLoS Genet* **9**, e1003629, doi:10.1371/journal.pgen.1003629 (2013).

13 Parker, K. L., Ikeda, Y. & Luo, X. The roles of steroidogenic factor-1 in reproductive function. *Steroids* **61**, 161-165, doi:10.1016/0039-128x(96)00006-2 (1996).

14 Garcia-Alonso, L. *et al.* Single-cell roadmap of human gonadal development. *Nature* **607**, 540-547, doi:10.1038/s41586-022-04918-4 (2022).

15 Karl, J. & Capel, B. Sertoli cells of the mouse testis originate from the coelomic epithelium. *Dev Biol* **203**, 323-333, doi:10.1006/dbio.1998.9068 (1998).

16 Liu, C. F., Liu, C. & Yao, H. H. Building pathways for ovary organogenesis in the mouse embryo. *Curr Top Dev Biol* **90**, 263-290, doi:10.1016/S0070-2153(10)90007-0 (2010).

17 Rotgers, E., Jorgensen, A. & Yao, H. H. At the Crossroads of Fate-Somatic Cell Lineage Specification in the Fetal Gonad. *Endocr Rev* **39**, 739-759, doi:10.1210/er.2018-00010 (2018).

18 Smith, P., Wilhelm, D. & Rodgers, R. J. Development of mammalian ovary. *J Endocrinol* **221**, R145-161, doi:10.1530/JOE-14-0062 (2014).

19 Goodfellow, P. N. & Lovell-Badge, R. SRY and sex determination in mammals. *Annu Rev Genet* **27**, 71-92, doi:10.1146/annurev.ge.27.120193.000443 (1993).
20 Page, D. C., de la Chapelle, A. & Weissenbach, J. Chromosome Y-specific DNA in related human XX males. *Nature* **315**, 224-226, doi:10.1038/315224a0 (1985).
21 Koopman, P., Gubbay, J., Vivian, N., Goodfellow, P. & Lovell-Badge, R. Male development of chromosomally female mice transgenic for Sry. *Nature* **351**, 117-121, doi:10.1038/351117a0 (1991).
22 Albrecht, K. H. & Eicher, E. M. Evidence that Sry is expressed in pre-Sertoli cells and Sertoli and granulosa cells have a common precursor. *Dev Biol* **240**, 92-107, doi:10.1006/dbio.2001.0438 (2001).
23 Bullejos, M. & Koopman, P. Spatially dynamic expression of Sry in mouse genital ridges. *Dev Dyn* **221**, 201-205, doi:10.1002/dvdy.1134 (2001).
24 Hiramatsu, R. *et al.* A critical time window of Sry action in gonadal sex determination in mice. *Development* **136**, 129-138, doi:10.1242/dev.029587 (2009).
25 Koopman, P. Sry and Sox9: mammalian testis-determining genes. *Cell Mol Life Sci* **55**, 839-856, doi:10.1007/pl00013200 (1999).
26 Harley, V. R., Clarkson, M. J. & Argentaro, A. The molecular action and regulation of the testis-determining factors, SRY (sex-determining region on the Y chromosome) and SOX9 [SRY-related high-mobility group (HMG) box 9]. *Endocr Rev* **24**, 466-487, doi:10.1210/er.2002-0025 (2003).
27 Munger, S. C., Natarajan, A., Looger, L. L., Ohler, U. & Capel, B. Fine time course expression analysis identifies cascades of activation and repression and maps a putative regulator of mammalian sex determination. *PLoS Genet* **9**, e1003630, doi:10.1371/journal.pgen.1003630 (2013).
28 Tilmann, C. & Capel, B. Mesonephric cell migration induces testis cord formation and Sertoli cell differentiation in the mammalian gonad. *Development* **126**, 2883-2890, doi:10.1242/dev.126.13.2883 (1999).
29 Brennan, J., Tilmann, C. & Capel, B. Pdgfr-alpha mediates testis cord organization and fetal Leydig cell development in the XY gonad. *Genes Dev* **17**, 800-810, doi:10.1101/gad.1052503 (2003).
30 Leeson, T. S. & Cookson, F. B. The mammalian testicular capsule and its muscle elements. *J Morphol* **144**, 237-253, doi:10.1002/jmor.1051440208 (1974).
31 Hermo, L., Pelletier, R. M., Cyr, D. G. & Smith, C. E. Surfing the wave, cycle, life history, and genes/proteins expressed by testicular germ cells. Part 1: background to spermatogenesis, spermatogonia, and spermatocytes. *Microsc Res Tech* **73**, 241-278, doi:10.1002/jemt.20783 (2010).
32 Clermont, Y. Kinetics of spermatogenesis in mammals: seminiferous epithelium cycle and spermatogonial renewal. *Physiol Rev* **52**, 198-236, doi:10.1152/physrev.1972.52.1.198 (1972).
33 Roosen-Runge, E. C. & Holstein, A. F. The human rete testis. *Cell Tissue Res* **189**, 409-433, doi:10.1007/BF00209130 (1978).
34 Maekawa, M., Kamimura, K. & Nagano, T. Peritubular myoid cells in the testis: their structure and function. *Arch Histol Cytol* **59**, 1-13, doi:10.1679/aohc.59.1 (1996).
35 Griswold, M. D. The central role of Sertoli cells in spermatogenesis. *Semin Cell Dev Biol* **9**, 411-416, doi:10.1006/scdb.1998.0203 (1998).
36 Cheng, C. Y. & Mruk, D. D. The blood-testis barrier and its implications for male contraception. *Pharmacol Rev* **64**, 16-64, doi:10.1124/pr.110.002790 (2012).

37 Oakberg, E. F. Duration of spermatogenesis in the mouse and timing of stages of the cycle of the seminiferous epithelium. *Am J Anat* **99**, 507-516, doi:10.1002/aja.1000990307 (1956).

38 Zirkin, B. R. & Papadopoulos, V. Leydig cells: formation, function, and regulation. *Biol Reprod* **99**, 101-111, doi:10.1093/biolre/ioy059 (2018).

39 Heinrich, A. & DeFalco, T. Essential roles of interstitial cells in testicular development and function. *Andrology* **8**, 903-914, doi:10.1111/andr.12703 (2020).

40 Munger, S. C. *et al.* Elucidation of the transcription network governing mammalian sex determination by exploiting strain-specific susceptibility to sex reversal. *Genes Dev* **23**, 2521-2536, doi:10.1101/gad.1835809 (2009).

41 Jameson, S. A. *et al.* Temporal transcriptional profiling of somatic and germ cells reveals biased lineage priming of sexual fate in the fetal mouse gonad. *PLoS Genet* **8**, e1002575, doi:10.1371/journal.pgen.1002575 (2012).

42 Nef, S. *et al.* Gene expression during sex determination reveals a robust female genetic program at the onset of ovarian development. *Dev Biol* **287**, 361-377, doi:10.1016/j.ydbio.2005.09.008 (2005).

43 Chen, H. *et al.* Identification of novel markers of mouse fetal ovary development. *PLoS One* **7**, e41683, doi:10.1371/journal.pone.0041683 (2012).

44 Schmidt, D. *et al.* The murine winged-helix transcription factor Foxl2 is required for granulosa cell differentiation and ovary maintenance. *Development* **131**, 933-942, doi:10.1242/dev.00969 (2004).

45 Uda, M. *et al.* Foxl2 disruption causes mouse ovarian failure by pervasive blockage of follicle development. *Hum Mol Genet* **13**, 1171-1181, doi:10.1093/hmg/ddh124 (2004).

46 Uhlenhaut, N. H. *et al.* Somatic sex reprogramming of adult ovaries to testes by FOXL2 ablation. *Cell* **139**, 1130-1142, doi:10.1016/j.cell.2009.11.021 (2009).

47 Ottolenghi, C. *et al.* Foxl2 is required for commitment to ovary differentiation. *Hum Mol Genet* **14**, 2053-2062, doi:10.1093/hmg/ddi210 (2005).

48 Boulanger, L. *et al.* FOXL2 is a female sex-determining gene in the goat. *Curr Biol* **24**, 404-408, doi:10.1016/j.cub.2013.12.039 (2014).

49 Pailhoux, E. *et al.* A 11.7-kb deletion triggers intersexuality and polledness in goats. *Nat Genet* **29**, 453-458, doi:10.1038/ng769 (2001).

50 Nicol, B., Estermann, M. A., Yao, H. H. & Mellouk, N. Becoming female: Ovarian differentiation from an evolutionary perspective. *Front Cell Dev Biol* **10**, 944776, doi:10.3389/fcell.2022.944776 (2022).

51 Gustin, S. E. *et al.* WNT/beta-catenin and p27/FOXL2 differentially regulate supporting cell proliferation in the developing ovary. *Dev Biol* **412**, 250-260, doi:10.1016/j.ydbio.2016.02.024 (2016).

52 Rastetter, R. H. *et al.* Marker genes identify three somatic cell types in the fetal mouse ovary. *Dev Biol* **394**, 242-252, doi:10.1016/j.ydbio.2014.08.013 (2014).

53 Niu, W. & Spradling, A. C. Two distinct pathways of pregranulosa cell differentiation support follicle formation in the mouse ovary. *Proc Natl Acad Sci U S A* **117**, 20015-20026, doi:10.1073/pnas.2005570117 (2020).

54 Fujino, A., Pieretti-Vanmarcke, R., Wong, A., Donahoe, P. K. & Arango, N. A. Sexual dimorphism of G-protein subunit Gng13 expression in the cortical region of the developing mouse ovary. *Dev Dyn* **236**, 1991-1996, doi:10.1002/dvdy.21183 (2007).

55 Mork, L. *et al.* Temporal differences in granulosa cell specification in the ovary reflect distinct follicle fates in mice. *Biol Reprod* **86**, 37, doi:10.1095/biolreprod.111.095208 (2012).

56 Zheng, W. *et al.* Two classes of ovarian primordial follicles exhibit distinct developmental dynamics and physiological functions. *Hum Mol Genet* **23**, 920-928, doi:10.1093/hmg/ddt486 (2014).

57 Tam, P. P. & Snow, M. H. Proliferation and migration of primordial germ cells during compensatory growth in mouse embryos. *J Embryol Exp Morphol* **64**, 133-147 (1981).

58 Lei, L. & Spradling, A. C. Mouse primordial germ cells produce cysts that partially fragment prior to meiosis. *Development* **140**, 2075-2081, doi:10.1242/dev.093864 (2013).

59 Pepling, M. E. & Spradling, A. C. Mouse ovarian germ cell cysts undergo programmed breakdown to form primordial follicles. *Dev Biol* **234**, 339-351, doi:10.1006/dbio.2001.0269 (2001).

60 Fortune, J. E., Cushman, R. A., Wahl, C. M. & Kito, S. The primordial to primary follicle transition. *Mol Cell Endocrinol* **163**, 53-60, doi:10.1016/s0303-7207(99)00240-3 (2000).

61 Rolaki, A., Drakakis, P., Millingos, S., Loutradis, D. & Makrigiannakis, A. Novel trends in follicular development, atresia and corpus luteum regression: a role for apoptosis. *Reprod Biomed Online* **11**, 93-103, doi:10.1016/s1472-6483(10)61304-1 (2005).

62 Soede, N. M., Langendijk, P. & Kemp, B. Reproductive cycles in pigs. *Anim Reprod Sci* **124**, 251-258, doi:10.1016/j.anireprosci.2011.02.025 (2011).

63 Browne, P. *et al.* Endocrine differentiation of fetal ovaries and testes of the spotted hyena (Crocuta crocuta): timing of androgen-independent versus androgen-driven genital development. *Reproduction* **132**, 649-659, doi:10.1530/rep.1.01120 (2006).

64 Cinalli, R. M., Rangan, P. & Lehmann, R. Germ cells are forever. *Cell* **132**, 559-562, doi:10.1016/j.cell.2008.02.003 (2008).

65 Saitou, M., Barton, S. C. & Surani, M. A. A molecular programme for the specification of germ cell fate in mice. *Nature* **418**, 293-300, doi:10.1038/nature00927 (2002).

66 Sato, M. *et al.* Identification of PGC7, a new gene expressed specifically in preimplantation embryos and germ cells. *Mech Dev* **113**, 91-94, doi:10.1016/s0925-4773(02)00002-3 (2002).

67 Ginsburg, M., Snow, M. H. & McLaren, A. Primordial germ cells in the mouse embryo during gastrulation. *Development* **110**, 521-528, doi:10.1242/dev.110.2.521 (1990).

68 Molyneaux, K. A., Stallock, J., Schaible, K. & Wylie, C. Time-lapse analysis of living mouse germ cell migration. *Dev Biol* **240**, 488-498, doi:10.1006/dbio.2001.0436 (2001).

69 Lawson, K. A. *et al.* Bmp4 is required for the generation of primordial germ cells in the mouse embryo. *Genes Dev* **13**, 424-436, doi:10.1101/gad.13.4.424 (1999).

70 Lawson, K. A. & Hage, W. J. Clonal analysis of the origin of primordial germ cells in the mouse. *Ciba Found Symp* **182**, 68-84; discussion 84-91, doi:10.1002/9780470514573.ch5 (1994).

71 Tam, P. P. & Zhou, S. X. The allocation of epiblast cells to ectodermal and germ-line lineages is influenced by the position of the cells in the gastrulating mouse embryo. *Dev Biol* **178**, 124-132, doi:10.1006/dbio.1996.0203 (1996).

72 Ewen, K. A. & Koopman, P. Mouse germ cell development: from specification to sex determination. *Mol Cell Endocrinol* **323**, 76-93, doi:10.1016/j.mce.2009.12.013 (2010).

73 Lin, Y., Gill, M. E., Koubova, J. & Page, D. C. Germ cell-intrinsic and -extrinsic factors govern meiotic initiation in mouse embryos. *Science* **322**, 1685-1687, doi:10.1126/science.1166340 (2008).

74 Gill, M. E., Hu, Y. C., Lin, Y. & Page, D. C. Licensing of gametogenesis, dependent on RNA binding protein DAZL, as a gateway to sexual differentiation of fetal germ cells. *Proc Natl Acad Sci U S A* **108**, 7443-7448, doi:10.1073/pnas.1104501108 (2011).

75 Clermont, Y. & Leblond, C. P. Renewal of spermatogonia in the rat. *Am J Anat* **93**, 475-501, doi:10.1002/aja.1000930308 (1953).

76 Amann, R. P. The cycle of the seminiferous epithelium in humans: a need to revisit? *J Androl* **29**, 469-487, doi:10.2164/jandrol.107.004655 (2008).

77 Handel, M. A. & Schimenti, J. C. Genetics of mammalian meiosis: regulation, dynamics and impact on fertility. *Nat Rev Genet* **11**, 124-136, doi:10.1038/nrg2723 (2010).

78 Wang, X. & Pepling, M. E. Regulation of Meiotic Prophase One in Mammalian Oocytes. *Front Cell Dev Biol* **9**, 667306, doi:10.3389/fcell.2021.667306 (2021).

79 Jan, S. Z. *et al.* Molecular control of rodent spermatogenesis. *Biochim Biophys Acta* **1822**, 1838-1850, doi:10.1016/j.bbadis.2012.02.008 (2012).

80 Zickler, D. & Kleckner, N. Meiotic chromosomes: integrating structure and function. *Annu Rev Genet* **33**, 603-754, doi:10.1146/annurev.genet.33.1.603 (1999).

81 Fraune, J., Schramm, S., Alsheimer, M. & Benavente, R. The mammalian synaptonemal complex: protein components, assembly and role in meiotic recombination. *Exp Cell Res* **318**, 1340-1346, doi:10.1016/j.yexcr.2012.02.018 (2012).

82 Baudat, F., Manova, K., Yuen, J. P., Jasin, M. & Keeney, S. Chromosome synapsis defects and sexually dimorphic meiotic progression in mice lacking Spo11. *Mol Cell* **6**, 989-998, doi:10.1016/s1097-2765(00)00098-8 (2000).

83 Mahadevaiah, S. K. *et al.* Recombinational DNA double-strand breaks in mice precede synapsis. *Nat Genet* **27**, 271-276, doi:10.1038/85830 (2001).

84 Keeney, S. & Neale, M. J. Initiation of meiotic recombination by formation of DNA double-strand breaks: mechanism and regulation. *Biochem Soc Trans* **34**, 523-525, doi:10.1042/BST0340523 (2006).

85 de Vries, F. A. *et al.* Mouse Sycp1 functions in synaptonemal complex assembly, meiotic recombination, and XY body formation. *Genes Dev* **19**, 1376-1389, doi:10.1101/gad.329705 (2005).

86 Cromie, G. A. & Smith, G. R. Branching out: meiotic recombination and its regulation. *Trends Cell Biol* **17**, 448-455, doi:10.1016/j.tcb.2007.07.007 (2007).

87 Yoshida, K. *et al.* The mouse RecA-like gene Dmc1 is required for homologous chromosome synapsis during meiosis. *Mol Cell* **1**, 707-718, doi:10.1016/s1097-2765(00)80070-2 (1998).

88 Slotman, J. A. *et al.* Super-resolution imaging of RAD51 and DMC1 in DNA repair foci reveals dynamic distribution patterns in meiotic prophase. *PLoS Genet* **16**, e1008595, doi:10.1371/journal.pgen.1008595 (2020).

89 Bowles, J. *et al.* Retinoid signaling determines germ cell fate in mice. *Science* **312**, 596-600, doi:10.1126/science.1125691 (2006).

90 Koubova, J. *et al.* Retinoic acid regulates sex-specific timing of meiotic initiation in mice. *Proc Natl Acad Sci U S A* **103**, 2474-2479, doi:10.1073/pnas.0510813103 (2006).

91 Yao, H. H., DiNapoli, L. & Capel, B. Meiotic germ cells antagonize mesonephric cell migration and testis cord formation in mouse gonads. *Development* **130**, 5895-5902, doi:10.1242/dev.00836 (2003).

92 Menke, D. B., Koubova, J. & Page, D. C. Sexual differentiation of germ cells in XX mouse gonads occurs in an anterior-to-posterior wave. *Dev Biol* **262**, 303-312, doi:10.1016/s0012-1606(03)00391-9 (2003).

93 Bullejos, M. & Koopman, P. Germ cells enter meiosis in a rostro-caudal wave during development of the mouse ovary. *Mol Reprod Dev* **68**, 422-428, doi:10.1002/mrd.20105 (2004).

94 Li, L. *et al.* Single-Cell RNA-Seq Analysis Maps Development of Human Germline Cells and Gonadal Niche Interactions. *Cell Stem Cell* **20**, 858-873 e854, doi:10.1016/j.stem.2017.03.007 (2017).

95 Borum, K. Oogenesis in the mouse. A study of the meiotic prophase. *Exp Cell Res* **24**, 495-507, doi:10.1016/0014-4827(61)90449-9 (1961).

96 Ge, W., Li, L., Dyce, P. W., De Felici, M. & Shen, W. Establishment and depletion of the ovarian reserve: physiology and impact of environmental chemicals. *Cell Mol Life Sci* **76**, 1729-1746, doi:10.1007/s00018-019-03028-1 (2019).

97 De Felici, M. *et al.* Establishment of oocyte population in the fetal ovary: primordial germ cell proliferation and oocyte programmed cell death. *Reprod Biomed Online* **10**, 182-191, doi:10.1016/s1472-6483(10)60939-x (2005).

98 Monget, P. *et al.* The ovarian reserve in mammals: a functional and evolutionary perspective. *Mol Cell Endocrinol* **356**, 2-12, doi:10.1016/j.mce.2011.07.046 (2012).

99 Grive, K. J. & Freiman, R. N. The developmental origins of the mammalian ovarian reserve. *Development* **142**, 2554-2563, doi:10.1242/dev.125211 (2015).

100 Krajnik, K., Mietkiewska, K., Skowronska, A., Kordowitzki, P. & Skowronski, M. T. Oogenesis in Women: From Molecular Regulatory Pathways and Maternal Age to Stem Cells. *Int J Mol Sci* **24**, doi:10.3390/ijms24076837 (2023).

101 Liu, K. *et al.* Control of mammalian oocyte growth and early follicular development by the oocyte PI3 kinase pathway: new roles for an old timer. *Dev Biol* **299**, 1-11, doi:10.1016/j.ydbio.2006.07.038 (2006).

102 Reddy, P. *et al.* Activation of Akt (PKB) and suppression of FKHRL1 in mouse and rat oocytes by stem cell factor during follicular activation and development. *Dev Biol* **281**, 160-170, doi:10.1016/j.ydbio.2005.02.013 (2005).

103 Reddy, P. *et al.* PDK1 signaling in oocytes controls reproductive aging and lifespan by manipulating the survival of primordial follicles. *Hum Mol Genet* **18**, 2813-2824, doi:10.1093/hmg/ddp217 (2009).

104 Adhikari, D. & Liu, K. Molecular mechanisms underlying the activation of mammalian primordial follicles. *Endocr Rev* **30**, 438-464, doi:10.1210/er.2008-0048 (2009).

105 Zhang, T. *et al.* Mechanisms of primordial follicle activation and new pregnancy opportunity for premature ovarian failure patients. *Front Physiol* **14**, 1113684, doi:10.3389/fphys.2023.1113684 (2023).

106 Bolcun-Filas, E. & Handel, M. A. Meiosis: the chromosomal foundation of reproduction. *Biol Reprod* **99**, 112-126, doi:10.1093/biolre/ioy021 (2018).

107 Wang, Q., Racowsky, C. & Deng, M. Mechanism of the chromosome-induced polar body extrusion in mouse eggs. *Cell Div* **6**, 17, doi:10.1186/1747-1028-6-17 (2011).

108 Masui, Y. & Markert, C. L. Cytoplasmic control of nuclear behavior during meiotic maturation of frog oocytes. *J Exp Zool* **177**, 129-145, doi:10.1002/jez.1401770202 (1971).

109 MacLean, G., Li, H., Metzger, D., Chambon, P. & Petkovich, M. Apoptotic extinction of germ cells in testes of Cyp26b1 knockout mice. *Endocrinology* **148**, 4560-4567, doi:10.1210/en.2007-0492 (2007).

110 de Kretser, D. M., Loveland, K. L., Meinhardt, A., Simorangkir, D. & Wreford, N. Spermatogenesis. *Hum Reprod* **13 Suppl 1**, 1-8, doi:10.1093/humrep/13.suppl_1.1 (1998).

111 Fayomi, A. P. & Orwig, K. E. Spermatogonial stem cells and spermatogenesis in mice, monkeys and men. *Stem Cell Res.* **29**, 207-214, doi:10.1016/j.scr.2018.04.009 (2018).

112 Simeunovic, B., Strbenc, M. & Bavdek, S. V. Position and histological structure of the testes in the brown hare (Lepus europaeus) during seasonal regression and recrudescence. *Anat Histol Embryol* **29**, 73-82, doi:10.1046/j.1439-0264.2000.00233.x (2000).

113 Temple-Smith, P. D. M. *Seasonal breeding biology of the platypus, Ornithorhynchus anatinus (Shaw, 1799), with special reference to the male*. (The Australian National University (Australia), 1973).

114 Phillips, B. T., Gassei, K. & Orwig, K. E. Spermatogonial stem cell regulation and spermatogenesis. *Philos Trans R Soc Lond B Biol Sci* **365**, 1663-1678, doi:10.1098/rstb.2010.0026 (2010).

115 Clermont, Y. & Leblond, C. P. Differentiation and renewal of spermatogonia in the monkey, Macacus rhesus. *Am J Anat* **104**, 237-273, doi:10.1002/aja.1001040204 (1959).

116 de Rooij, D. G. Spermatogonial stem cell renewal in the mouse. I. Normal situation. *Cell Tissue Kinet* **6**, 281-287, doi:10.1111/j.1365-2184.1973.tb01617.x (1973).

117 Boitani, C., Di Persio, S., Esposito, V. & Vicini, E. Spermatogonial cells: mouse, monkey and man comparison. *Semin Cell Dev Biol* **59**, 79-88, doi:10.1016/j.semcdb.2016.03.002 (2016).

118 Clermont, Y. Two classes of spermatogonial stem cells in the monkey (Cercopithecus aethiops). *Am J Anat* **126**, 57-71, doi:10.1002/aja.1001260106 (1969).

119 Hermann, B. P., Sukhwani, M., Hansel, M. C. & Orwig, K. E. Spermatogonial stem cells in higher primates: are there differences from those in rodents? *Reproduction* **139**, 479-493, doi:10.1530/REP-09-0255 (2010).

120 Marshall Graves, J. A. & Shetty, S. Sex from W to Z: evolution of vertebrate sex chromosomes and sex determining genes. *J Exp Zool* **290**, 449-462, doi:10.1002/jez.1088 (2001).

121 Cortez, D. *et al.* Origins and functional evolution of Y chromosomes across mammals. *Nature* **508**, 488-493, doi:10.1038/nature13151 (2014).

122 Graves, J. A., Wakefield, M. J. & Toder, R. The origin and evolution of the pseudoautosomal regions of human sex chromosomes. *Hum Mol Genet* **7**, 1991-1996, doi:10.1093/hmg/7.13.1991 (1998).

123 Zhou, Y. *et al.* Platypus and echidna genomes reveal mammalian biology and evolution. *Nature* **592**, 756-762, doi:10.1038/s41586-020-03039-0 (2021).

124 Carey, S. B. *et al.* Representing sex chromosomes in genome assemblies. *Cell Genom* **2**, doi:10.1016/j.xgen.2022.100132 (2022).

125 Turner, J. M. Meiotic Silencing in Mammals. *Annu. Rev. Genet.* **49**, 395-412, doi:10.1146/annurev-genet-112414-055145 (2015).
126 Turner, J. M. Meiotic sex chromosome inactivation. *Development* **134**, 1823-1831, doi:10.1242/dev.000018 (2007).
127 McKee, B. D. & Handel, M. A. Sex chromosomes, recombination, and chromatin conformation. *Chromosoma* **102**, 71-80, doi:10.1007/BF00356023 (1993).
128 Boulton, S. J. Cellular functions of the BRCA tumour-suppressor proteins. *Biochem Soc Trans* **34**, 633-645, doi:10.1042/BST0340633 (2006).
129 Traven, A. & Heierhorst, J. SQ/TQ cluster domains: concentrated ATM/ATR kinase phosphorylation site regions in DNA-damage-response proteins. *Bioessays* **27**, 397-407, doi:10.1002/bies.20204 (2005).
130 Fillingham, J., Keogh, M. C. & Krogan, N. J. GammaH2AX and its role in DNA double-strand break repair. *Biochem Cell Biol* **84**, 568-577, doi:10.1139/o06-072 (2006).
131 Turner, J. M. *et al.* Silencing of unsynapsed meiotic chromosomes in the mouse. *Nat Genet* **37**, 41-47, doi:10.1038/ng1484 (2005).
132 Hoyer-Fender, S. Molecular aspects of XY body formation. *Cytogenet Genome Res* **103**, 245-255, doi:10.1159/000076810 (2003).
133 Baarends, W. M. *et al.* Histone ubiquitination and chromatin remodeling in mouse spermatogenesis. *Dev Biol* **207**, 322-333, doi:10.1006/dbio.1998.9155 (1999).
134 Larson, E. L., Kopania, E. E. K. & Good, J. M. Spermatogenesis and the Evolution of Mammalian Sex Chromosomes. *Trends Genet* **34**, 722-732, doi:10.1016/j.tig.2018.06.003 (2018).
135 Daish, T. J., Casey, A. E. & Grutzner, F. Lack of sex chromosome specific meiotic silencing in platypus reveals origin of MSCI in therian mammals. *BMC Biol* **13**, 106, doi:10.1186/s12915-015-0215-4 (2015).
136 Clermont, Y. & Leblond, C. P. Spermiogenesis of man, monkey, ram and other mammals as shown by the periodic acid-Schiff technique. *Am J Anat* **96**, 229-253, doi:10.1002/aja.1000960203 (1955).
137 Fitzpatrick, J. L., Kahrl, A. F. & Snook, R. R. SpermTree, a species-level database of sperm morphology spanning the animal tree of life. *Sci Data* **9**, 30, doi:10.1038/s41597-022-01131-w (2022).
138 Auger, J., Jouannet, P. & Eustache, F. Another look at human sperm morphology. *Hum Reprod* **31**, 10-23, doi:10.1093/humrep/dev251 (2016).
139 Moore, H. D. Gamete biology of the new world marsupial, the grey short-tailed opossum, Monodelphis domestica. *Reprod Fertil Dev* **8**, 605-615, doi:10.1071/rd9960605 (1996).
140 Hikim, R. L. E. R. S. & ED, A. C. Histological and histopathological evaluation of the testis. *Clearwater: Cache River Press, Clearwater, Florida* (1990).
141 Fawcett, D. W. The mammalian spermatozoon. *Dev Biol* **44**, 394-436, doi:10.1016/0012-1606(75)90411-x (1975).
142 Fawcett, D. W. & Phillips, D. M. The fine structure and development of the neck region of the mammalian spermatozoon. *The Anatomical Record* **165**, 153-183 (1969).
143 Kierszenbaum, A. L. & Tres, L. L. The acrosome-acroplaxome-manchette complex and the shaping of the spermatid head. *Archives of histology and cytology* **67**, 271-284 (2004).
144 Kimmins, S. & Sassone-Corsi, P. Chromatin remodelling and epigenetic features of germ cells. *Nature* **434**, 583-589, doi:10.1038/nature03368 (2005).

145 Balhorn, R. The protamine family of sperm nuclear proteins. *Genome Biol* **8**, 227, doi:10.1186/gb-2007-8-9-227 (2007).
146 Lu, L.-Y. *et al.* RNF8-dependent histone modifications regulate nucleosome removal during spermatogenesis. *Developmental cell* **18**, 371-384 (2010).
147 Nair, M. *et al.* Nuclear regulator Pygo2 controls spermiogenesis and histone H3 acetylation. *Developmental biology* **320**, 446-455 (2008).
148 Okada, Y., Scott, G., Ray, M. K., Mishina, Y. & Zhang, Y. Histone demethylase JHDM2A is critical for Tnp1 and Prm1 transcription and spermatogenesis. *Nature* **450**, 119-123 (2007).
149 Meistrich, M. L., Mohapatra, B., Shirley, C. R. & Zhao, M. Roles of transition nuclear proteins in spermiogenesis. *Chromosoma* **111**, 483-488, doi:10.1007/s00412-002-0227-z (2003).
150 O'Donnell, L., Nicholls, P. K., O'Bryan, M. K., McLachlan, R. I. & Stanton, P. G. Spermiation: the process of sperm release. *Spermatogenesis* **1**, 14-35 (2011).
151 Berruti, G. & Paiardi, C. Acrosome biogenesis: Revisiting old questions to yield new insights. *Spermatogenesis* **1**, 95-98, doi:10.4161/spmg.1.2.16820 (2011).
152 Yanagimachi, R. Mammalian fertilization (1994) The physiology of reproduction. *Raven Press, New York*, 189-317.
153 Griswold, M. D. Spermatogenesis: The Commitment to Meiosis. *Physiol Rev* **96**, 1-17, doi:10.1152/physrev.00013.2015 (2016).
154 Leblond, C. P. & Clermont, Y. Definition of the stages of the cycle of the seminiferous epithelium in the rat. *Ann N Y Acad Sci* **55**, 548-573, doi:10.1111/j.1749-6632.1952.tb26576.x (1952).
155 Wistuba, J. *et al.* Organization of seminiferous epithelium in primates: relationship to spermatogenic efficiency, phylogeny, and mating system. *Biol Reprod* **69**, 582-591, doi:10.1095/biolreprod.103.015925 (2003).
156 Birkhead, T. R. & Pizzari, T. Postcopulatory sexual selection. *Nat. Rev. Genet.* **3**, 262-273, doi:10.1038/nrg774 (2002).
157 Brawand, D. *et al.* The evolution of gene expression levels in mammalian organs. *Nature* **478**, 343-348, doi:10.1038/nature10532 (2011).
158 Cardoso-Moreira, M. *et al.* Gene expression across mammalian organ development. *Nature* **571**, 505-509, doi:10.1038/s41586-019-1338-5 (2019).
159 Wang, Z. Y. *et al.* Transcriptome and translatome co-evolution in mammals. *Nature* **588**, 642-647, doi:10.1038/s41586-020-2899-z (2020).
160 Ramm, S. A., Scharer, L., Ehmcke, J. & Wistuba, J. Sperm competition and the evolution of spermatogenesis. *Mol. Hum. Reprod.* **20**, 1169-1179, doi:10.1093/molehr/gau070 (2014).
161 Martinez, G. & Garcia, C. Sexual selection and sperm diversity in primates. *Mol Cell Endocrinol* **518**, 110974, doi:10.1016/j.mce.2020.110974 (2020).
162 Simmons, L. W. & Fitzpatrick, J. L. Sperm wars and the evolution of male fertility. *Reproduction* **144**, 519-534, doi:10.1530/REP-12-0285 (2012).
163 Parker, G. A. Sperm competition and its evolutionary consequences in the insects. *Biological reviews* **45**, 525-567 (1970).
164 Parker, G. Sperm competition and the evolution of ejaculates: towards a theory base. *Sperm competition and sexual selection (ed. TR Birkhead & AP Moller)*, 3-54 (1998).
165 Parker, G., Birkhead, T. & Møller, A. 3-54 (Academic Press San Diego, 1998).

166 Harcourt, A. H., Harvey, P. H., Larson, S. G. & Short, R. V. Testis weight, body weight and breeding system in primates. *Nature* **293**, 55-57, doi:10.1038/293055a0 (1981).
167 Necsulea, A. & Kaessmann, H. Evolutionary dynamics of coding and non-coding transcriptomes. *Nat Rev Genet* **15**, 734-748, doi:10.1038/nrg3802 (2014).
168 Niefind, K. & Schomburg, D. Amino acid similarity coefficients for protein modeling and sequence alignment derived from main-chain folding angles. *J Mol Biol* **219**, 481-497, doi:10.1016/0022-2836(91)90188-c (1991).
169 Rice, G. & Rebeiz, M. Evolution: How Many Phenotypes Do Regulatory Mutations Affect? *Curr Biol* **29**, R21-R23, doi:10.1016/j.cub.2018.11.027 (2019).
170 Wray, G. A. The evolutionary significance of cis-regulatory mutations. *Nat Rev Genet* **8**, 206-216, doi:10.1038/nrg2063 (2007).
171 Roos, D. & de Boer, M. Mutations in cis that affect mRNA synthesis, processing and translation. *Biochim Biophys Acta Mol Basis Dis* **1867**, 166166, doi:10.1016/j.bbadis.2021.166166 (2021).
172 Castro-Mondragon, J. A. *et al.* Cis-regulatory mutations associate with transcriptional and post-transcriptional deregulation of gene regulatory programs in cancers. *Nucleic Acids Res* **50**, 12131-12148, doi:10.1093/nar/gkac1143 (2022).
173 King, M. C. & Wilson, A. C. Evolution at two levels in humans and chimpanzees. *Science* **188**, 107-116, doi:10.1126/science.1090005 (1975).
174 Britten, R. J. & Davidson, E. H. Gene regulation for higher cells: a theory. *Science* **165**, 349-357, doi:10.1126/science.165.3891.349 (1969).
175 Hoekstra, H. E. & Coyne, J. A. The locus of evolution: evo devo and the genetics of adaptation. *Evolution* **61**, 995-1016, doi:10.1111/j.1558-5646.2007.00105.x (2007).
176 Carroll, S. B. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* **134**, 25-36, doi:10.1016/j.cell.2008.06.030 (2008).
177 Ponting, C. P. The functional repertoires of metazoan genomes. *Nat Rev Genet* **9**, 689-698, doi:10.1038/nrg2413 (2008).
178 Haygood, R., Fedrigo, O., Hanson, B., Yokoyama, K. D. & Wray, G. A. Promoter regions of many neural- and nutrition-related genes have experienced positive selection during human evolution. *Nat Genet* **39**, 1140-1144, doi:10.1038/ng2104 (2007).
179 Enard, W. *et al.* Intra- and interspecific variation in primate gene expression patterns. *Science* **296**, 340-343, doi:10.1126/science.1068996 (2002).
180 Khaitovich, P. *et al.* Parallel patterns of evolution in the genomes and transcriptomes of humans and chimpanzees. *Science* **309**, 1850-1854, doi:10.1126/science.1108296 (2005).
181 Hawkins, R. D., Hon, G. C. & Ren, B. Next-generation genomics: an integrative approach. *Nat Rev Genet* **11**, 476-486, doi:10.1038/nrg2795 (2010).
182 Kosiol, C. *et al.* Patterns of Positive Selection in Six Mammalian Genomes. *PLoS Genet.* **4**, e1000144-e1000144, doi:10.1371/journal.pgen.1000144 (2008).
183 Good, J. M. & Nachman, M. W. Rates of protein evolution are positively correlated with developmental timing of expression during mouse spermatogenesis. *Mol Biol Evol* **22**, 1044-1052, doi:10.1093/molbev/msi087 (2005).
184 Khaitovich, P., Enard, W., Lachmann, M. & Paabo, S. Evolution of primate gene expression. *Nat Rev Genet* **7**, 693-702, doi:10.1038/nrg1940 (2006).
185 Nielsen, R. *et al.* A scan for positively selected genes in the genomes of humans and chimpanzees. *PLoS Biol* **3**, e170, doi:10.1371/journal.pbio.0030170 (2005).

186 Ramskold, D., Wang, E. T., Burge, C. B. & Sandberg, R. An abundance of ubiquitously expressed genes revealed by tissue transcriptome sequence data. *PLoS Comput Biol* **5**, e1000598, doi:10.1371/journal.pcbi.1000598 (2009).

187 Kleene, K. C. A possible meiotic function of the peculiar patterns of gene expression in mammalian spermatogenic cells. *Mech Dev* **106**, 3-23, doi:10.1016/s0925-4773(01)00413-0 (2001).

188 Soumillon, M. *et al.* Cellular source and mechanisms of high transcriptome complexity in the mammalian testis. *Cell Rep.* **3**, 2179-2190, doi:10.1016/j.celrep.2013.05.031 (2013).

189 Kaessmann, H. Origins, evolution, and phenotypic impact of new genes. *Genome Res.* **20**, 1313-1326, doi:10.1101/gr.101386.109 (2010).

190 Zhang, Q. *et al.* The evolution of ovarian somatic cells characterized by transcriptome and chromatin accessibility across rodents, monkeys, and humans. *Life Medicine* **3**, doi:10.1093/lifemedi/lnae028 (2024).

191 Immler, S. Haploid selection in “diploid” organisms. *Annu. Rev. Ecol. Evol. Syst.* **50**, 219-236 (2019).

192 Murat, F. *et al.* The molecular evolution of spermatogenesis across mammals. *Nature* **613**, 308-316, doi:10.1038/s41586-022-05547-7 (2023).

193 Muldrew, K. & McGann, L. E. Mechanisms of intracellular ice formation. *Biophys J* **57**, 525-532, doi:10.1016/S0006-3495(90)82568-6 (1990).

194 Grindberg, R. V. *et al.* RNA-sequencing from single nuclei. *Proc Natl Acad Sci U S A* **110**, 19802-19807, doi:10.1073/pnas.1319700110 (2013).

195 Krishnaswami, S. R. *et al.* Using single nuclei for RNA-seq to capture the transcriptome of postmortem neurons. *Nat. Protoc.* **11**, 499-524, doi:10.1038/nprot.2016.015 (2016).

196 McInnes, L., Healy, J. & Melville, J. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426* (2018).

197 von Kopylow, K. & Spiess, A. N. Human spermatogonial markers. *Stem Cell Res* **25**, 300-309, doi:10.1016/j.scr.2017.11.011 (2017).

198 He, Z., Jiang, J., Hofmann, M. C. & Dym, M. Gfra1 silencing in mouse spermatogonial stem cells results in their differentiation via the inactivation of RET tyrosine kinase. *Biol Reprod* **77**, 723-733, doi:10.1095/biolreprod.107.062513 (2007).

199 Carmell, M. A. *et al.* MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline. *Dev Cell* **12**, 503-514, doi:10.1016/j.devcel.2007.03.001 (2007).

200 Zhang, T., Oatley, J., Bardwell, V. J. & Zarkower, D. DMRT1 Is Required for Mouse Spermatogonial Stem Cell Maintenance and Replenishment. *PLoS Genet.* **12**, e1006293, doi:10.1371/journal.pgen.1006293 (2016).

201 Ernst, C., Eling, N., Martinez-Jimenez, C. P., Marioni, J. C. & Odom, D. T. Staged developmental mapping and X chromosome transcriptional dynamics during mouse spermatogenesis. *Nat. Commun.* **10**, 1251, doi:10.1038/s41467-019-09182-1 (2019).

202 Tirumalasetty, M. B., Bhattacharya, I., Mohiuddin, M. S., Baki, V. B. & Choubey, M. Understanding testicular single cell transcriptional atlas: from developmental complications to male infertility. *Front Endocrinol (Lausanne)* **15**, 1394812, doi:10.3389/fendo.2024.1394812 (2024).

203 Kouznetsova, A., Novak, I., Jessberger, R. & Hoog, C. SYCP2 and SYCP3 are required for cohesin core integrity at diplotene but not for centromere cohesion at the first meiotic division. *J Cell Sci* **118**, 2271-2278, doi:10.1242/jcs.02362 (2005).

204 Vourekas, A. *et al.* Mili and Miwi target RNA repertoire reveals piRNA biogenesis and function of Miwi in spermiogenesis. *Nat Struct Mol Biol* **19**, 773-781, doi:10.1038/nsmb.2347 (2012).

205 Johnson, M. L., Wang, R. & Sperry, A. O. Novel localization of Aurora A kinase in mouse testis suggests multiple roles in spermatogenesis. *Biochem Biophys Res Commun* **503**, 51-55, doi:10.1016/j.bbrc.2018.05.170 (2018).

206 Romanienko, P. J. & Camerini-Otero, R. D. The mouse Spo11 gene is required for meiotic chromosome synapsis. *Mol Cell* **6**, 975-987, doi:10.1016/s1097-2765(00)00097-6 (2000).

207 Vatin, M. *et al.* Identification of a New QTL Region on Mouse Chromosome 1 Responsible for Male Hypofertility: Phenotype Characterization and Candidate Genes. *Int J Mol Sci* **21**, doi:10.3390/ijms21228506 (2020).

208 He, X., Zhang, Y., Yan, Y., Li, Y. & Koide, S. S. Identification of GABABR2 in rat testis and sperm. *J Reprod Dev* **49**, 397-402, doi:10.1262/jrd.49.397 (2003).

209 Herr, J. C., Flickinger, C. J., Homyk, M., Klotz, K. & John, E. Biochemical and morphological characterization of the intra-acrosomal antigen SP-10 from human sperm. *Biol Reprod* **42**, 181-193, doi:10.1095/biolreprod42.1.181 (1990).

210 O'Donnell, L. *et al.* Sertoli cell-enriched proteins in mouse and human testicular interstitial fluid. *PLoS One* **18**, e0290846, doi:10.1371/journal.pone.0290846 (2023).

211 Epstein, L. F. & Orme-Johnson, N. R. Acute action of luteinizing hormone on mouse Leydig cells: accumulation of mitochondrial phosphoproteins and stimulation of testosterone synthesis. *Mol Cell Endocrinol* **81**, 113-126, doi:10.1016/0303-7207(91)90210-j (1991).

212 Landreh, L. *et al.* Human testicular peritubular cells host putative stem Leydig cells with steroidogenic capacity. *J Clin Endocrinol Metab* **99**, E1227-1235, doi:10.1210/jc.2013-4199 (2014).

213 Li, L., Miano, J. M., Cserjesi, P. & Olson, E. N. SM22 alpha, a marker of adult smooth muscle, is expressed in multiple myogenic lineages during embryogenesis. *Circ Res* **78**, 188-195, doi:10.1161/01.res.78.2.188 (1996).

214 Fatigati, V. & Murphy, R. A. Actin and tropomyosin variants in smooth muscles. Dependence on tissue type. *J Biol Chem* **259**, 14383-14388 (1984).

215 Arakelian, L. *et al.* Endothelial CD34 expression and regulation of immune cell response in-vitro. *Sci Rep* **13**, 13512, doi:10.1038/s41598-023-40622-7 (2023).

216 Shih, S. C. *et al.* The L6 protein TM4SF1 is critical for endothelial cell function and tumor angiogenesis. *Cancer Res* **69**, 3272-3277, doi:10.1158/0008-5472.CAN-08-4886 (2009).

217 Greenow, K., Pearce, N. J. & Ramji, D. P. The key role of apolipoprotein E in atherosclerosis. *J Mol Med (Berl)* **83**, 329-342, doi:10.1007/s00109-004-0631-3 (2005).

218 Farr, L., Ghosh, S. & Moonah, S. Role of MIF Cytokine/CD74 Receptor Pathway in Protecting Against Injury and Promoting Repair. *Front Immunol* **11**, 1273, doi:10.3389/fimmu.2020.01273 (2020).

219 Kumar, S., Stecher, G., Suleski, M. & Hedges, S. B. TimeTree: A Resource for Timelines, Timetrees, and Divergence Times. *Mol Biol Evol* **34**, 1812-1819, doi:10.1093/molbev/msx116 (2017).

220 Mailund, T., Munch, K. & Schierup, M. H. Lineage sorting in apes. *Annu Rev Genet* **48**, 519-535, doi:10.1146/annurev-genet-120213-092532 (2014).

221 Dickinson, M. E. *et al.* High-throughput discovery of novel developmental phenotypes. *Nature* **537**, 508-514, doi:10.1038/nature19356 (2016).

222 Koscielny, G. *et al.* The International Mouse Phenotyping Consortium Web Portal, a unified point of access for knockout mice and related phenotyping data. *Nucleic Acids Res.* **42**, 802-809, doi:10.1093/nar/gkt977 (2014).

223 Miyata, T. & Yasunaga, T. Molecular evolution of mRNA: a method for estimating evolutionary rates of synonymous and amino acid substitutions from homologous nucleotide sequences and its application. *J Mol Evol* **16**, 23-36, doi:10.1007/BF01732067 (1980).

224 Nei, M. & Gojobori, T. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol Biol Evol* **3**, 418-426, doi:10.1093/oxfordjournals.molbev.a040410 (1986).

225 Yang, Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol* **24**, 1586-1591, doi:10.1093/molbev/msm088 (2007).

226 Van Der Lee, R., Wiel, L., Van Dam, T. J. P. & Huynen, M. A. Genome-scale detection of positive selection in nine primates predicts human-virus evolutionary conflicts. *Nucleic Acids Res.* **45**, 10634-10648, doi:10.1093/nar/gkx704 (2017).

227 Long, M., VanKuren, N. W., Chen, S. & Vibranovski, M. D. New gene evolution: little did we know. *Annu Rev Genet* **47**, 307-333, doi:10.1146/annurev-genet-111212-133301 (2013).

228 Trost, N., Mbengue, N. & Kaessmann, H. The molecular evolution of mammalian spermatogenesis. *Cells Dev* **175**, 203865, doi:10.1016/j.cdev.2023.203865 (2023).

229 Shao, Y. *et al.* GenTree, an integrated resource for analyzing the evolution and function of primate-specific coding genes. *Genome Res.* **29**, 682-696, doi:10.1101/gr.238733.118 (2019).

230 Session, D. R., Lee, G. S. & Wolgemuth, D. J. Characterization of D1Pas1, a mouse autosomal homologue of the human AZFa region DBY, as a nuclear protein in spermatogenic cells. *Fertil Steril* **76**, 804-811, doi:10.1016/s0015-0282(01)01996-3 (2001).

231 Molaro, A., Young, J. M. & Malik, H. S. Evolutionary origins and diversification of testis-specific short histone H2A variants in mammals. *Genome Res* **28**, 460-473, doi:10.1101/gr.229799.117 (2018).

232 Futschik, M. E. & Carlisle, B. Noise-robust soft clustering of gene expression time-course data. *J Bioinform Comput Biol* **3**, 965-988, doi:10.1142/s0219720005001375 (2005).

233 Janisiw, E. *et al.* Poly(ADP-ribose) glycohydrolase coordinates meiotic DNA double-strand break induction and repair independent of its catalytic activity. *Nat Commun* **11**, 4869, doi:10.1038/s41467-020-18693-1 (2020).

234 Lee, S. *et al.* Differential Regulation of TLE3 in Sertoli Cells of the Testes during Postnatal Development. *Cells* **8**, doi:10.3390/cells8101156 (2019).

235 Wang, C. Y. *et al.* PiggyBac Transposon-Mediated Mutagenesis in Rats Reveals a Crucial Role of Bbx in Growth and Male Fertility. *Biol Reprod* **95**, 51, doi:10.1095/biolreprod.116.141739 (2016).

236 Malla, A. B. & Bhandari, R. IP6K1 is essential for chromatoid body formation and temporal regulation of Tnp2 and Prm2 expression in mouse spermatids. *J Cell Sci* **130**, 2854-2866, doi:10.1242/jcs.204966 (2017).

237 Efremova, M., Vento-Tormo, M., Teichmann, S. A. & Vento-Tormo, R. CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. *Nat Protoc* **15**, 1484-1506, doi:10.1038/s41596-020-0292-x (2020).

238 Wakayama, T. & Iseki, S. Role of the spermatogenic-Sertoli cell interaction through cell adhesion molecule-1 (CADM1) in spermatogenesis. *Anat. Sci. Int.* **84**, 112-121, doi:10.1007/s12565-009-0034-1 (2009).

239 Ozaki-Kuroda, K. *et al.* Nectin couples cell-cell adhesion and the actin scaffold at heterotypic testicular junctions. *Curr. Biol.* **12**, 1145-1150, doi:10.1016/s0960-9822(02)00922-3 (2002).

240 Bronson, R., Mikhailik, A., Schwedes, J., Gnatenko, D. & Hatchwell, E. Detection of candidate nectin gene mutations in infertile men with severe teratospermia. *J. Assist. Reprod. Genet.* **34**, 1295-1302, doi:10.1007/s10815-017-0985-4 (2017).

241 Salenave, S. *et al.* Kallmann's syndrome: a comparison of the reproductive phenotypes in men carrying KAL1 and FGFR1/KAL2 mutations. *J Clin Endocrinol Metab* **93**, 758-763, doi:10.1210/jc.2007-1168 (2008).

242 Quintero-Rivera, F. *et al.* Is the disruption of an N-myristoyltransferase (NMT2) associated with hypoplastic testes? *Am J Med Genet A* **143A**, 1796-1798, doi:10.1002/ajmg.a.31799 (2007).

243 Fan, Y. *et al.* Diagnostic Application of Targeted Next-Generation Sequencing of 80 Genes Associated with Disorders of Sexual Development. *Sci Rep* **7**, 44536, doi:10.1038/srep44536 (2017).

244 Awwad, S. W., Darawshe, M. M., Machour, F. E., Arman, I. & Ayoub, N. Recruitment of RBM6 to DNA Double-Strand Breaks Fosters Homologous Recombination Repair. *Mol Cell Biol* **43**, 130-142, doi:10.1080/10985549.2023.2187105 (2023).

245 Ho, U. Y. *et al.* WDR62 is required for centriole duplication in spermatogenesis and manchette removal in spermiogenesis. *Commun Biol* **4**, 645, doi:10.1038/s42003-021-02171-5 (2021).

246 Moore, D. J. *et al.* Mutations in ZMYND10, a gene essential for proper axonemal assembly of inner and outer dynein arms in humans and flies, cause primary ciliary dyskinesia. *Am J Hum Genet* **93**, 346-356, doi:10.1016/j.ajhg.2013.07.009 (2013).

247 Shi, L., Shen, X., Chi, Y. & Shen, Y. Primary ciliary dyskinesia relative protein ZMYND10 is involved in regulating ciliary function and intraflagellar transport in Paramecium tetraurelia. *Eur J Protistol* **77**, 125756, doi:10.1016/j.ejop.2020.125756 (2021).

248 Bello, J. H. *et al.* Dysregulation of mitochondrial sirtuin genes is associated with human male infertility. *Andrologia* **54**, e14274, doi:10.1111/and.14274 (2022).

249 Martinez, G. *et al.* Expanding the sperm phenotype caused by mutations in SPATA20: A novel splicing mutation in an infertile patient with partial globozoospermia. *Clin Genet* **103**, 612-614, doi:10.1111/cge.14284 (2023).

250 Wang, X. *et al.* Identification of nonfunctional SPATA20 causing acephalic spermatozoa syndrome in humans. *Clin Genet* **103**, 310-319, doi:10.1111/cge.14268 (2023).
251 Chau, K. M. & Cornwall, G. A. Reduced fertility in vitro in mice lacking the cystatin CRES (cystatin-related epididymal spermatogenic): rescue by exposure of spermatozoa to dibutyryl cAMP and isobutylmethylxanthine. *Biol Reprod* **84**, 140-152, doi:10.1095/biolreprod.110.084855 (2011).
252 Potrzebowski, L. *et al.* Chromosomal gene movements reflect the recent origin and biology of therian sex chromosomes. *PLoS Biol* **6**, e80, doi:10.1371/journal.pbio.0060080 (2008).
253 Veyrunes, F. *et al.* Bird-like sex chromosomes of platypus imply recent origin of mammal sex chromosomes. *Genome Res* **18**, 965-973, doi:10.1101/gr.7101908 (2008).
254 Grutzner, F. *et al.* In the platypus a meiotic chain of ten sex chromosomes shares genes with the bird Z and mammal X chromosomes. *Nature* **432**, 913-917, doi:10.1038/nature03021 (2004).
255 Charlesworth, B. & Charlesworth, D. The degeneration of Y chromosomes. *Philos Trans R Soc Lond B Biol Sci* **355**, 1563-1572, doi:10.1098/rstb.2000.0717 (2000).
256 Hughes, J. F. *et al.* Strict evolutionary conservation followed rapid gene loss on human and rhesus Y chromosomes. *Nature* **483**, 82-86, doi:10.1038/nature10843 (2012).
257 Ross, M. T. *et al.* The DNA sequence of the human X chromosome. *Nature* **434**, 325-337, doi:10.1038/nature03440 (2005).
258 Kashimada, K. & Koopman, P. Sry: the master switch in mammalian sex determination. *Development* **137**, 3921-3930, doi:10.1242/dev.048983 (2010).
259 Wang, P. J., McCarrey, J. R., Yang, F. & Page, D. C. An abundance of X-linked genes expressed in spermatogonia. *Nat Genet* **27**, 422-426, doi:10.1038/86927 (2001).
260 Shami, A. *et al.* Single-cell RNA sequencing of human, macaque, and mouse testes uncovers conserved and divergent features of mammalian spermatogenesis. *Dev. Cell* **54**, 529-547 e512, doi:10.1101/2020.03.17.994509 (2020).
261 Wagner, K. D., Wagner, N. & Schedl, A. The complex life of WT1. *J Cell Sci* **116**, 1653-1658, doi:10.1242/jcs.00405 (2003).
262 Raz, E. The function and regulation of vasa-like genes in germ-cell development. *Genome Biol* **1**, REVIEWS1017, doi:10.1186/gb-2000-1-3-reviews1017 (2000).
263 Schoenmakers, S. *et al.* Female meiotic sex chromosome inactivation in chicken. *PLoS Genet* **5**, e1000466, doi:10.1371/journal.pgen.1000466 (2009).
264 Lau, X., Munusamy, P., Ng, M. J. & Sangrithi, M. Single-Cell RNA Sequencing of the Cynomolgus Macaque Testis Reveals Conserved Transcriptional Profiles during Mammalian Spermatogenesis. *Dev. Cell* **54**, 548-566.e547, doi:10.1016/j.devcel.2020.07.018 (2020).
265 Rebourcet, D. *et al.* Sertoli Cell Number Defines and Predicts Germ and Leydig Cell Population Sizes in the Adult Mouse Testis. *Endocrinology* **158**, 2955-2969, doi:10.1210/en.2017-00196 (2017).
266 Orth, J. M., Gunsalus, G. L. & Lamperti, A. A. Evidence from Sertoli cell-depleted rats indicates that spermatid number in adults depends on numbers of Sertoli cells produced during perinatal development. *Endocrinology* **122**, 787-794, doi:10.1210/endo-122-3-787 (1988).

267 Lecluze, E., Jegou, B., Rolland, A. D. & Chalmel, F. New transcriptomic tools to understand testis development and functions. *Mol Cell Endocrinol* **468**, 47-59, doi:10.1016/j.mce.2018.02.019 (2018).

268 Carroll, S. B. Evolution at two levels: on genes and form. *PLoS Biol* **3**, e245, doi:10.1371/journal.pbio.0030245 (2005).

269 Ewing, E. P. Selection at the haploid and diploid phases: cyclical variation. *Genetics* **87**, 195-207, doi:10.1093/genetics/87.1.195 (1977).

270 Immler, S., Arnqvist, G. & Otto, S. P. Ploidally antagonistic selection maintains stable genetic polymorphism. *Evolution* **66**, 55-65, doi:10.1111/j.1558-5646.2011.01399.x (2012).

271 Larracuente, A. M. & Presgraves, D. C. The selfish Segregation Distorter gene complex of Drosophila melanogaster. *Genetics* **192**, 33-53, doi:10.1534/genetics.112.141390 (2012).

272 Bergero, R. *et al.* Meiosis and beyond - understanding the mechanistic and evolutionary processes shaping the germline genome. *Biol Rev Camb Philos Soc* **96**, 822-841, doi:10.1111/brv.12680 (2021).

273 Khan, Z. *et al.* Primate transcript and protein expression levels evolve under compensatory selection pressures. *Science* **342**, 1100-1104, doi:10.1126/science.1242379 (2013).

274 Martincic, D. S., Virant Klun, I., Zorn, B. & Vrtovec, H. M. Germ cell apoptosis in the human testis. *Pflugers Arch* **442**, R159-160, doi:10.1007/s004240100007 (2001).

275 Shaha, C., Tripathi, R. & Mishra, D. P. Male germ cell apoptosis: regulation and biology. *Philos Trans R Soc Lond B Biol Sci* **365**, 1501-1515, doi:10.1098/rstb.2009.0124 (2010).

276 Dunkel, L., Hirvonen, V. & Erkkila, K. Clinical aspects of male germ cell apoptosis during testis development and spermatogenesis. *Cell Death Differ* **4**, 171-179, doi:10.1038/sj.cdd.4400234 (1997).

277 Gu, C. H. *et al.* Over-expression of testis-specific expressed gene 1 attenuates the proliferation and induces apoptosis of GC-1spg cells. *J Huazhong Univ Sci Technolog Med Sci* **34**, 535-541, doi:10.1007/s11596-014-1311-3 (2014).

278 Inoue, H. *et al.* Mouse D1Pas1, a DEAD-box RNA helicase, is required for the completion of first meiotic prophase in male germ cells. *Biochem Biophys Res Commun* **478**, 592-598, doi:10.1016/j.bbrc.2016.07.109 (2016).

279 Choi, H., Zhou, L., Zhao, Y. & Dean, J. RNA helicase D1PAS1 resolves R-loops and forms a complex for mouse pachytene piRNA biogenesis required for male fertility. *Nucleic Acids Res* **52**, 11973-11994, doi:10.1093/nar/gkae712 (2024).

280 Hayes, B. J., Bowman, P. J., Chamberlain, A. J. & Goddard, M. E. Invited review: Genomic selection in dairy cattle: progress and challenges. *J Dairy Sci* **92**, 433-443, doi:10.3168/jds.2008-1646 (2009).

281 Souza, F. M., Lopes, F. B., Rosa, G. J. M. & Magnabosco, C. U. Economic values of reproductive, growth, feed efficiency and carcass traits in Nellore cattle. *J Anim Breed Genet* **139**, 170-180, doi:10.1111/jbg.12652 (2022).

282 Skakkebaek, N. E. *et al.* Environmental factors in declining human fertility. *Nat Rev Endocrinol* **18**, 139-157, doi:10.1038/s41574-021-00598-8 (2022).

283 Fauser, B. *et al.* Declining global fertility rates and the implications for family planning and family building: an IFFS consensus document based on a narrative review of the literature. *Hum Reprod Update* **30**, 153-173, doi:10.1093/humupd/dmad028 (2024).

284 Zattara, E. E., Busey, H. A., Linz, D. M., Tomoyasu, Y. & Moczek, A. P. Neofunctionalization of embryonic head patterning genes facilitates the positioning of novel traits on the dorsal head of adult beetles. *Proc Biol Sci* **283**, doi:10.1098/rspb.2016.0824 (2016).

285 Huminiecki, L. & Wolfe, K. H. Divergence of spatial gene expression profiles following species-specific gene duplications in human and mouse. *Genome Res* **14**, 1870-1879, doi:10.1101/gr.2705204 (2004).

286 Cheng, X. & Sun, Q. RUBCNL/Pacer and RUBCN/Rubicon in regulation of autolysosome formation and lipid metabolism. *Autophagy* **15**, 1120-1121, doi:10.1080/15548627.2019.1596500 (2019).

287 Russell, D. L., Brown, H. M. & Dunning, K. R. ADAMTS proteases in fertility. *Matrix Biol* **44-46**, 54-63, doi:10.1016/j.matbio.2015.03.007 (2015).

288 Wen, Q. *et al.* Signaling pathways regulating blood-tissue barriers - Lesson from the testis. *Biochim Biophys Acta Biomembr* **1860**, 141-153, doi:10.1016/j.bbamem.2017.04.020 (2018).

289 Zakrzewski, P., Lenartowska, M. & Buss, F. Diverse functions of myosin VI in spermiogenesis. *Histochem Cell Biol* **155**, 323-340, doi:10.1007/s00418-020-01954-x (2021).

290 Mital, P., Hinton, B. T. & Dufour, J. M. The blood-testis and blood-epididymis barriers are more than just their tight junctions. *Biol Reprod* **84**, 851-858, doi:10.1095/biolreprod.110.087452 (2011).

291 Gurbich, T. A. & Bachtrog, D. Gene content evolution on the X chromosome. *Curr Opin Genet Dev* **18**, 493-498, doi:10.1016/j.gde.2008.09.006 (2008).

292 Khil, P. P., Smirnova, N. A., Romanienko, P. J. & Camerini-Otero, R. D. The mouse X chromosome is enriched for sex-biased genes not subject to selection by meiotic sex chromosome inactivation. *Nat Genet* **36**, 642-646, doi:10.1038/ng1368 (2004).

293 Koslowski, M., Sahin, U., Huber, C. & Tureci, O. The human X chromosome is enriched for germline genes expressed in premeiotic germ cells of both sexes. *Hum Mol Genet* **15**, 2392-2399, doi:10.1093/hmg/ddl163 (2006).

294 Emerson, J. J., Kaessmann, H., Betran, E. & Long, M. Extensive gene traffic on the mammalian X chromosome. *Science* **303**, 537-540, doi:10.1126/science.1090042 (2004).

295 Kaessmann, H., Vinckenbosch, N. & Long, M. RNA-based gene duplication: mechanistic and evolutionary insights. *Nat Rev Genet* **10**, 19-31, doi:10.1038/nrg2487 (2009).

296 Braun, R. E., Behringer, R. R., Peschon, J. J., Brinster, R. L. & Palmiter, R. D. Genetically haploid spermatids are phenotypically diploid. *Nature* **337**, 373-376, doi:10.1038/337373a0 (1989).

297 Babcock, M., Yatsenko, S., Stankiewicz, P., Lupski, J. R. & Morrow, B. E. AT-rich repeats associated with chromosome 22q11.2 rearrangement disorders shape human genome architecture on Yq12. *Genome Res.* **17**, 451-460, doi:10.1101/gr.5651507 (2007).

298 Botto, L. D. *et al.* A population-based study of the 22q11.2 deletion: phenotype, incidence, and contribution to major birth defects in the population. *Pediatrics* **112**, 101-107, doi:10.1542/peds.112.1.101 (2003).

299 Zietsch, B. P., Walum, H., Lichtenstein, P., Verweij, K. J. H. & Kuja-Halkola, R. No genetic contribution to variation in human offspring sex ratio: a total population study of 4.7 million births. *Proc Biol Sci* **287**, 20192849, doi:10.1098/rspb.2019.2849 (2020).

300 Shiu, P. K., Raju, N. B., Zickler, D. & Metzenberg, R. L. Meiotic silencing by unpaired DNA. *Cell* **107**, 905-916, doi:10.1016/s0092-8674(01)00609-2 (2001).

301 Bean, C. J., Schaner, C. E. & Kelly, W. G. Meiotic pairing and imprinted X chromatin assembly in Caenorhabditis elegans. *Nat Genet* **36**, 100-105, doi:10.1038/ng1283 (2004).

302 EBERHARD, W. G. Copulatory courtship and cryptic female choice in insects. *Biological Reviews* **66**, 1-31 (1991).

303 Guioli, S., Lovell-Badge, R. & Turner, J. M. Error-prone ZW pairing and no evidence for meiotic sex chromosome inactivation in the chicken germ line. *PLoS Genet* **8**, e1002560, doi:10.1371/journal.pgen.1002560 (2012).

304 Hassold, T. & Hunt, P. To err (meiotically) is human: the genesis of human aneuploidy. *Nat Rev Genet* **2**, 280-291, doi:10.1038/35066065 (2001).

305 Hunt, P. A. & Hassold, T. J. Sex matters in meiosis. *Science* **296**, 2181-2183, doi:10.1126/science.1071907 (2002).

306 Lane, S. & Kauppi, L. Meiotic spindle assembly checkpoint and aneuploidy in males versus females. *Cell Mol Life Sci* **76**, 1135-1150, doi:10.1007/s00018-018-2986-6 (2019).

307 Gagneux, P., Moore, J. J. & Varki, A. The ethics of research on great apes. *Nature* **437**, 27-29, doi:10.1038/437027a (2005).

308 Dufour, J. J., Fahmy, M. H. & Minvielle, F. Seasonal changes in breeding activity, testicular size, testosterone concentration and seminal characteristics in rams with long or short breeding season. *J Anim Sci* **58**, 416-422, doi:10.2527/jas1984.582416x (1984).

309 Jimenez, R., Burgos, M. & Barrionuevo, F. J. Circannual Testis Changes in Seasonally Breeding Mammals. *Sex Dev* **9**, 205-215, doi:10.1159/000439039 (2015).

310 Attar, M. *et al.* A practical solution for preserving single cells for RNA sequencing. *Sci. Rep.* **8**, 2151, doi:10.1038/s41598-018-20372-7 (2018).

311 Yates, A. *et al.* Ensembl 2016. *Nucleic Acids Res.* **44**, D710-D716, doi:10.1093/nar/gkv1157 (2016).

312 Yates, A. D. *et al.* Ensembl 2020. *Nucleic Acids Res* **48**, D682-D688, doi:10.1093/nar/gkz966 (2020).

313 Howe, K. L. *et al.* Ensembl 2021. *Nucleic Acids Res* **49**, D884-D891, doi:10.1093/nar/gkaa942 (2021).