

DISS. ETH NO.

**CHARACTERIZATION OF THE GENOMIC LANDSCAPE OF
MOUSE SPERMATOGONIAL CELLS DURING POSTNATAL
TESTIS MATURATION**

A thesis submitted to attain the degree of
DOCTOR OF SCIENCES of ETH ZURICH
(Dr. sc. ETH Zurich)

presented by
IRINA-ELENA LAZAR-CONTES
MSc – Karolinska Institute, Sweden
born on 08.04.1988
citizen of Romania

accepted on the recommendation of

Prof. Isabelle Mansuy, examiner
Prof Tuncay Baubec, co-examiner
Prof Gavin Kelsey, co-examiner

2020

Summary

Spermatogonial cells are the postnatal initiators of spermatogenesis and fulfil essential proliferative and differentiation functions across lifetime, in order to ensure successful sperm formation and fertilization potential of an organism. Furthermore, spermatogonial cells are the only stem cell population in the body which convey genetic information to the embryo. During postnatal testis maturation, spermatogonial cells are characterized by significantly different transcriptional programs. However, the underlying chromatin landscape which may contribute to such dynamic gene expression remains largely unknown.

The work presented in this thesis analyses chromatin accessibility differences between early postnatal and steady-state adult spermatogonial cells in the mouse testis. Using bioinformatics tools, chromatin accessibility profiles were correlated with newly generated and previously published gene expression profiles and histone modifications, in order to reveal novel characteristics of the age-dependent molecular phenotype of spermatogonial cells.

We found that open chromatin in spermatogonial cells undergoes significant reorganization during testis development, with predominantly more accessible regions in adult spermatogonial cells than in cells from postnatal developing testis. Motif enrichment analysis predicted specific transcription factor families in the regions of differential accessibility, suggestive of distinct age-specific regulatory networks. Biological pathways associated with more accessible regions in adult spermatogonia were consistent with RNA metabolism, maintenance of a cell stem state and cell fate priming, whereas less accessible regions displayed enrichment in pathways predominantly involved in developmental processes. By integrating chromatin accessibility, transcriptome and histone modifications data from several postnatal timepoints, we revealed 4 categories of differentially accessible regions in proximity of genes with dynamic expression profiles across age, and we describe specific patterns of activating, repressive or bivalent histone modifications at each of these categories.

By assessing chromatin accessibility specifically at transposable elements (TEs), we showed that certain subtypes, mainly components of the ERVK-LTR family, become less accessible in adult spermatogonia, whilst specific LINE L1 subtypes increase in accessibility. Exemplary LTR loci within the subtypes with altered chromatin

accessibility coincided with transcription starting sites (TSSs) of pluripotency genes, suggesting an age-dependent dynamic regulatory potential of TEs on the proliferation and differentiation capabilities of spermatogonial cells. In addition, some of the more accessible LINE L1 loci were distinctively associated with olfactory receptor gene regions within the genome, prompting further investigation into the non-random gene expression regulation that certain TE subtypes may fulfil in spermatogonial cells. The abundant number of enriched transcription factor motifs situated at the differentially accessible TEs also emphasize important regulatory role for these TE subtypes in spermatogonial cells across postnatal testis maturation.

Lastly, we investigated how an environmental insult applied in early postnatal development, during spermatogonial cell establishment and first division rounds, could alter their molecular phenotype. By investigating several levels of transcriptome regulation, we found differences in the transcript usage of several genes encoding RNA-binding proteins and known splicing factors in spermatogonial cells of exposed pups. Chromatin accessibility assessment in spermatogonial cells from control and stressed pups shortly after the end of the stress exposure revealed regions of differential accessibility between the 2 groups, and their association with pathways related to numerous metabolic and immune processes. Notably, we also found testis and blood alterations in immune signalling, implying that an altered testis milieu, potentially via systemic factors, could alter spermatogonial cells in ways described above. In the *Conclusions and Outlook* section of the thesis we explore how our findings contribute to the current knowledge about spermatogonial cell age-dependent transcriptome and epigenetic profiles, what further experiments would be needed to validate these results, and the implications of our results for assessing germline reactivity to environmental insults and lifestyle factors.

Zusammenfassung

Spermatogonien sind die postnatalen Initiatoren der Spermatogenese und erfüllen lebenslang wesentliche Proliferations- und Differenzierungsfunktionen, um die erfolgreiche Spermienbildung und das Befruchtungspotential eines Organismus zu gewährleisten. Darüber hinaus sind Spermatogonien die einzige Stammzellen im Körper, die direkt zur genetischen Information eines Nachkommens beitragen. Spermatogonien zeichnen sich durch signifikant unterschiedliche Programme zur Regulation der Gene während der postnatalen Hodenreifung aus. Die zugrundeliegenden Eigenschaften von Chromatin, welche zu einer solchen dynamischen Genexpression beitragen, bleiben jedoch weitgehend unbekannt.

Die in dieser Dissertation vorgestellte Arbeit analysiert die Unterschiede in der Zugänglichkeit zum Chromatin zwischen frühen postnatalen und reifen Spermatogonien im Hoden der Maus. Mittels Bioinformatik wurden die Zugänglichkeit zum Chromatin mit neu erstellten und zuvor publizierten Profilen von Genexpression und Histonmodifikationen korreliert, um neue Charakteristika des altersabhängigen molekularen Phänotyps von Spermatogonien aufzudecken.

Wir fanden heraus, dass die Chromatinzugänglichkeit in Spermatogonien während der Hodenentwicklung eine signifikante Reorganisation erfährt, wobei die zugänglichen Chromatinregionen in adulten Spermatogonien überwiegend besser zugänglich sind als in Spermatogonien aus sich noch postnatal entwickelnden Hoden. Die Motivanreicherungsanalyse sagte spezifische Transkriptionsfaktorfamilien in den Regionen der differenziellen Chromatinzugänglichkeit voraus, was auf ausgeprägte, altersspezifische Regulationsmechanismen schliessen lässt. Darüber hinaus waren die biologischen Stoffwechselwege, die mit leichter zugänglichen Regionen in adulten Spermatogonien assoziiert waren, konsistent mit dem RNA-Stoffwechsel, der Aufrechterhaltung eines Zellstammzustandes und der Vorbereitung des Zellschicksals, während weniger zugängliche Regionen vorwiegend mit Entwicklungsprozessen assoziiert waren. Durch die Integration von Daten der Chromatinzugänglichkeit, des Transkriptoms und Histonmodifikationen aus mehreren postnatalen Zeitpunkten, entdeckten wir 4 Kategorien veränderter Zugangsprofile des Chromatins in der Nähe von Genen mit dynamischen Genexpression. Ferner

beschreiben wir spezifische Muster von aktivierenden, repressiven oder bivalenten Histonmodifikationen in jeder dieser Kategorien.

Indem wir die Chromatinzugänglichkeit speziell an transponierbaren Elementen (TEs) im gesamten Genom untersuchen, zeigen wir auf, dass bestimmte Subtypen, hauptsächlich Komponenten der ERVK-LTR-Familie, bei adulten Spermatogonien weniger zugänglich werden, während spezifische LINE L1-Subtypen die Zugänglichkeit erhöhen. Beispielhafte ERVK-Subtypen mit veränderter Chromatinzugänglichkeit überlappten mit Transkriptionsstartstellen (TSS) von Genen welche die Pluripotenz steuern, was auf ein altersabhängiges, dynamisches Regulationspotenzial von TEs auf die Proliferations- und Differenzierungsfähigkeit von Spermatogonien hindeutet. Darüber hinaus waren einige der leichter zugänglichen LINE L1-Subtypen eindeutig mit olfaktorischen Rezeptor-Genregionen innerhalb des Genoms assoziiert, was zu weiteren Untersuchungen der nicht-zufälligen Genexpressionsregulation führte, die bestimmte TE-Subtypen in Spermatogonien erfüllen können. Die grosse Anzahl von Transkriptionsfaktormotiven, die an den differenziell zugänglichen TEs angereichert waren, weist ebenfalls auf eine wichtige regulatorische Rolle dieser TE-Subtypen in Spermatogonien während der postnatalen Hodenreifung hin.

Schliesslich untersuchten wir, wie eine Umweltbelastung in der frühen postnatalen Entwicklung von Mäusen, während der Etablierung der Spermatogonienzellen und der ersten Teilungen, deren molekularen Phänotyp verändern könnte. Durch die Untersuchung mehrerer Ebenen der Transkriptom-Regulation fanden wir Unterschiede in der Transkriptverwendung mehrerer Gene, die für RNA-bindende Proteine und bekannte Spleissfaktoren kodieren, in den Spermatogonien exponierter Jungtiere. Die Bewertung der Chromatinzugänglichkeit in Spermatogonien von gestressten und nicht-gestressten Jungtieren kurz nach dem Ende der Stressexposition zeigte Regionen mit unterschiedlicher Zugänglichkeit zwischen den beiden Gruppen auf und ihre Verbindung mit Stoffwechselwegen, die mit der Entwicklung des Zentralnervensystems zusammenhängen. In adulten Spermatogonien zeigten weniger Regionen eine unterschiedliche Zugänglichkeit zwischen den Kontroll- und den exponierten Männchen auf, jedoch bezogen sie sich immer noch auf Gene, die an der neuronalen Proliferation und der Entwicklung des Nervensystems beteiligt sind, was auf eine erhöhte Anfälligkeit dieser biologischen

Stoffwechselwege für Umweltbelastungen im postnatalen Alter hinweist. In den Schlussfolgerungen und Aussichten der Dissertation untersuchen wir, wie unsere Ergebnisse zum aktuellen Wissen über das altersabhängige Transkriptom und die epigenetischen Profile von Spermatogonien beitragen, welche weiteren Experimente erforderlich wären um diese Ergebnisse zu validieren, und welche Auswirkungen unsere Ergebnisse auf die Bewertung der Keimbahnreaktivität auf Umweltbelastungen und gesundheitliche Risikofaktoren haben.

Table of Contents

<i>Summary</i>	1
<i>Zusammenfassung</i>	3
<i>Table of Contents</i>	6
<i>Glossary</i>	9
<i>Gene / Protein Nomenclature</i>	11
1. Introduction	12
1.1 The mammalian germline	12
1.1.1 Development of the mammalian germline in the embryo	12
1.1.2 Spermatogenesis in mammals	13
1.2 Spermatogonial cells.....	15
1.2.1 Spermatogonial cell organization in the testis.....	15
1.2.2 Spermatogonial stem cell identity: fragmentation and hierarchical models.....	16
1.2.3 Gene expression dynamics in spermatogonial cells.....	18
1.2.3.1 Bulk RNA-sequencing experiments	19
1.2.3.2 Single cell RNA-sequencing experiments	19
1.2.4 Epigenetic regulation in spermatogonial cells	22
1.2.4.1 DNA methylation	23
1.2.4.2 Histone modifications.....	24
1.2.4.3 Non-coding RNAs.....	24
1.2.4.4 Chromatin accessibility and 3D organization	25
1.3 Current methods for isolating spermatogonial cells	27
1.3.1 Surface marker – based cell sorting	27
1.3.2 Genetic tagging – based cell sorting	29
1.4 Environmental insults on spermatogonial cells	30
1.4.1 Toxicant exposure.....	31
1.4.2 Other environmental exposures: heat, diet, chronic stress.....	32
1.5 Overview	33
2 Postnatal characterization of mouse spermatogonial cells reveals distinct chromatin regulatory landscapes in the developing and adult testis	34
2.1 Abstract.....	35
2.2 Introduction.....	36
2.3 Results.....	38
2.3.1 Fluorescence activated cell sorting (FACS) enriches spermatogonial cells from developing and adult mouse testis	38
2.3.2 Accessible chromatin is reorganised in adult spermatogonia compared to early postnatal stage 38	38
2.3.3 Differentially accessible chromatin regions associate with distinct gene expression dynamics and histone marks.....	41
2.3.4 Differentially accessible chromatin regions are marked by binding sites for distinct families of transcription factors (TFs)	48
2.3.5 Open chromatin at transposable elements (TEs) undergoes significant remodelling in the transition from early postnatal to adult spermatogonia.....	52
2.4 Discussion.....	58

2.5 Methods	61
2.5.1 Mouse husbandry	61
2.5.2 Germ cells isolation.....	61
2.5.3 Spermatogonial cell enrichment by FACS	62
2.5.4 Immunocytochemistry	63
2.5.5 RNA extraction and RNA-seq library preparation	63
2.5.6 Omni-ATAC library preparation	64
2.5.7 High-throughput sequencing data analysis.....	64
2.5.7.1 RNA-seq data analysis	65
2.5.7.2 Omni-ATAC data analysis	66
2.5.7.3 ChIP-seq data analysis	68
2.5.7.4 Bisulphite sequencing (BS) data analysis.....	68
2.6 Authors contributions	69
2.7 Acknowledgements	69
2.8 Supplemental figures	70
3 Early life stress alters chromatin accessibility landscape and transcript usage in spermatogonial cells during postnatal testis maturation.....	78
3.1 Abstract	79
3.2 Introduction.....	80
3.3 Results.....	82
3.3.1 Exposure to traumatic stress alters spermatogonial cell transcriptome in the developing testis	82
3.3.2 Chromatin accessibility in spermatogonial cells is altered by early postnatal exposure	88
3.3.3 Alteration in inflammatory proteins in blood of MSUS pups correlates with immune receptor activation in the developing testis	93
3.4 Discussion.....	96
3.5 Methods	100
3.5.1 Animals and MSUS paradigm	100
3.5.2 Testis dissection and spermatogonial cell enrichment	100
3.5.3 Immunocytochemistry	102
3.5.4 RNA-seq library preparation and data analysis.....	102
3.5.5 Omni-ATAC library preparation and data analysis.....	104
3.5.6 Proteomic analysis of blood plasma	105
3.5.7 RT-qPCR and western blot analyses.....	106
3.5.8 Quantification and data analysis.....	107
3.6 Acknowledgements	108
3.7 Authors contributions	108
3.8 Supplemental figures	109
4 Conclusions and Outlook	115
4.1 Challenges and limitations	115
4.2 Postnatal dynamics of spermatogonial cells	117
4.2.1 Open chromatin reorganization in postnatal spermatogonia	117
4.2.2 Unique chromatin landscape of genes differentially expressed in early postnatal and adult spermatogonia	118
4.2.3 Chromatin accessibility landscape at transposable elements	119
4.3 Implications of spermatogonial cell vulnerability to environmental insults	121
4.3.1 Importance of isoform usage quantification	121
4.3.2 Open chromatin reorganization after stress exposure	123

4.3.3	Systemic proteomic alterations and their effects on the testis	124
References	126	
Acknowledgements	139	

Glossary

A ₁₋₄	Differentiating A-type spermatogonial cells
A _{al}	Undifferentiated A aligned spermatogonial cells
AD	Alzheimer disease
A _{dark}	Undifferentiated A dark spermatogonial cells
A _p	Undifferentiated A paired spermatogonial cells
A _{pale}	Undifferentiated A pale spermatogonial cells
A _s	Undifferentiated A single spermatogonial cells
ASD	Autism spectrum disorder
ATAC-seq	Assay for transposase-accessible chromatin sequencing
BCL6B	B-cell CLL/lymphoma 6 member B protein
BMX1	Bone marrow tyrosine kinase gene in chromosome x protein
BPA	Bisphenol-A
BTB	Blood-testis-barrier
c-KIT	Tyrosine protein kinase KIT/Cluster of differentiation 117
CGI	CpG island
CTCF	CCCTC-binding factor
DBP	Di(n-Butyl) Phthalate
DMRT1	Doublesex And Mab-3 Related Transcription Factor 1
DNA	Deoxyribonucleic acid
DNAme	DNA methylation
dpc	Days post coitum
DTU	Differential transcript usage
E	Embryonic day
EDC	Endocrine disruptor compound
EE2	17 α -ethinylestradiol
ERVK	Endogenous retrovirus K-promoter
ESC	Embryonic stem cell
ETV5	ETS Variant Transcription Factor 5
FACS	Fluorescence activated cell sorting
FC	Fold change
FGF	Fibroblast growth factor
GDNF	Glial cell line-derived neurotrophic factor

GFP	Green fluorescent protein
GFRA1	Glial cell line-derived neurotrophic factor receptor α
HiC	Chromosome conformation capture-based technique
HSC	Hematopoietic stem cell
ID4	Inhibitor of DNA binding 4
LHX1	LIM Homeobox 1
lincRNA	Long intergenic non-coding RNA
lncRNA	Long non-coding RNA
LPPR3	Lipid Phosphate Phosphatase-Related Protein Type 3
LTR	Long terminal repeat
MACS	Magnetic activated cell sorting
MAGE	Melanoma-associated antigen
MAPK	Mitogen activated protein kinase
miR	microRNA
mRNA	Messenger RNA
mTOR	Mammalian target of Rapamycin
NANOS2	Nanos C2HC-Type Zinc Finger 2
NGN3	Neurogenin-3
OCT4	Octamer-binding transcription factor 4
PAX7	Paired box 7
PGC	Primordial germ cell
piRNA	Piwi-interacting RNA
PLZF	Promyelocytic leukemia zinc finger
PND	Postnatal day
PNW	Postnatal week
POU5F1	POU Class 5 Homeobox 1
RET	Ret Proto-Oncogene
RNA	Ribonucleic acid
RNA-seq	RNA sequencing
RT	Room temperature
SALL4	Spalt Like Transcription Factor 4
scRNA-seq	Single cell RNA sequencing
sncRNA	Small non-coding RNA
SSC	Spermatogonial stem cell

TAD	Topologically associated domain
TE	Transposable element
TF	Transcription factor
THY1	Thy1 cell surface antigen / Cluster of differentiation 90
Tn5	Transposase 5
VEGF	Vascular endothelial growth factor
WT	Wild type

Gene / Protein Nomenclature

Mouse gene symbols are italicised, with the first letter capitalised and the rest of the letters in lowercase (e.g., *Thy1*). Protein names are the same as the gene name, but not italicised and with all letters capitalised (e.g., THY1).

Human gene symbols are italicised with all letters capitalised (e.g., *THY1*). Protein names are the same as the gene name, but not italicised and with all letters capitalised.

1. Introduction

1.1 The mammalian germline

1.1.1 Development of the mammalian germline in the embryo

In mammals, reproductive potential across lifetime relies entirely on the population of cells capable of producing mature sperm and oocytes, collectively referred to as the germline. Germline development is initiated in the embryo, when primordial germ cells (PGCs), the earliest precursors of functional gametes, were first identified as a distinct cell population (Ginsburg et al., 1990). Most of the evidence about PGC formation, migration and specification has been obtained using the mouse as a model of choice, thanks to their accessibility and ease of use. Mouse PGCs were first detected in the embryo at embryonic day 7.25 (E7.25), in the endoderm of the yolk sac and below the primitive streak, from where they migrate to the gonadal ridges, which they colonise by E10.5 (Richardson and Lehmann, 2010; Saitou and Yamaji, 2012).

In the female embryo, proliferation of PGCs continues until ~ E13.5, when around 25,000 PGCs exit mitosis and enter prophase I of meiosis. Following entry in meiosis, female mouse PGCs continue meiotic progression until day 3-5 after birth, when all oogonia remain arrested in the diplotene stage of prophase I until hormonal stimulation during puberty (Chuma and Nakatsuji, 2000; McLaren, 2000). In contrast, in male embryos, PGCs enter mitotic arrest once they reach the gonadal ridges, and remain arrested in the G₀/G₁ phase of the cell cycle throughout embryonic development as gonocytes, until postnatal day 1-2 (PND1-2). At this stage, gonocytes exit quiescence and resume mitotic proliferation in order to form spermatogonial cells, and initiate the spermatogenic cascade (Saitou and Yamaji, 2012).

Unlike in the mouse, human PGC formation is not as well defined, mainly due to the scarcity of early human embryonic tissue. Several studies in humans have defined an overall chronology of PGC development, starting with their detection at week 3 post-fertilization, in the yolk sac near the allantois. PGCs migrate to the gonadal ridges by week 8, and proliferate in the gonadal ridges followed by differentiation into oogonia (week 9) or gonocytes (week 10-12) soon after sex determination is complete. Similar to the mouse, oogonia continue proliferating and enter meiosis during foetal development, whereas gonocytes undergo mitotic arrest and remain quiescent until

puberty (De Felici, 2013). Given the extensive research using the mouse as animal model, and the work included in this thesis, the next chapters will focus on describing the current knowledge with reference to the mouse germline, and touch upon the existing evidence in humans when available.

1.1.2 Spermatogenesis in mammals

Mouse spermatogenesis begins immediately after birth, at PND1-2, when gonocytes (also termed pro-spermatogonia in their quiescent state) migrate from the centre of the seminiferous tubules to the basement membrane of the testis, and give rise to the mitotically active spermatogonial cells. In mice, a full spermatogenic cycle in the testis lasts around 35 to 37 days, and is divided into several stages of consecutive mitotic and meiotic divisions (Figure 1-1). First, spermatogonial cells divide repeatedly in order to form a stable pool of spermatogonial stem cells (SSCs), capable of self-renewal. Out of these cells, a subpopulation of SSCs gradually loses stem cell potential and forms A type progenitor spermatogonia, which are committed to further differentiation into Intermediate and B type spermatogonia (Oatley and Brinster, 2006). At around PND9-10 the first populations of preleptotene spermatocytes are formed from the differentiating spermatogonia. Preleptotene diploid spermatocytes meiotically divide (meiosis I) and form leptotene, pachytene, zygotene and diplotene haploid spermatocytes (secondary spermatocytes) by the end of the second postnatal week. During this meiotic division, migration of spermatocytes from the basement membrane of the testis towards the lumen of the seminiferous tubules also takes place. In meiosis II, spermatocytes give rise to round spermatids, which then enter a phase of morphological restructuring, marked by nuclear condensing of chromatin, through histone-to-protamine exchange and cytoplasm reduction, to form elongating spermatids. In the last phase of testis transit, elongating spermatids are

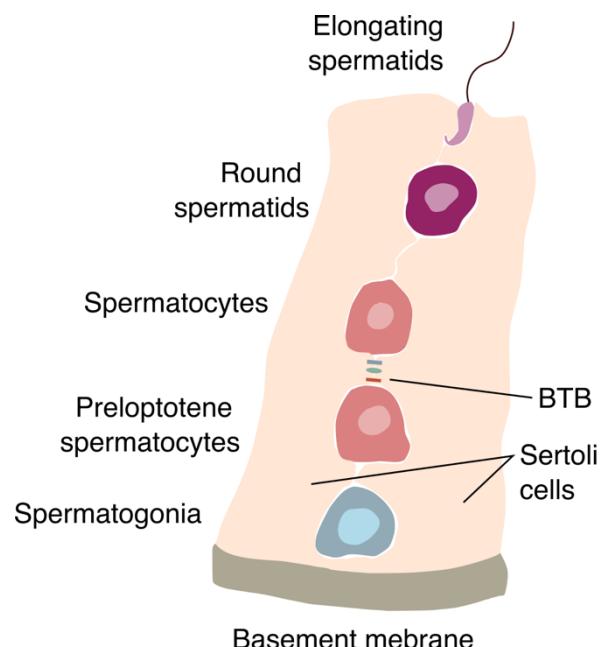


Figure 1-1 Schematic of a transversal section through a seminiferous tubule in mice

released into the lumen of the seminiferous tubules. The maturation of elongating spermatids and their transformation into fertilization-competent spermatozoa takes place during their migration through the epididymis, an independent structure in which adult spermatozoa reside until their release (La and Hobbs, 2019; Phillips et al., 2010).

Concomitant with germ cell maturation, several other important processes take place in the testis, and involve the somatic cell milieu. Sertoli cells, the only somatic cells situated in the seminiferous tubules together with the germ cells, expand from the basement membrane until the lumen of the seminiferous tubules, and are equipped with extensive cytoplasmic processes which they send out between each other and to the germ cells. Some of the cellular junctions between adjacent Sertoli cells are very tightly maintained together and form what is known as the blood-testis-barrier (BTB). The BTB develops between PND15-25, and divides the seminiferous tubules epithelium into 2 domains: the basal domain, where undifferentiated and differentiating spermatogonial cells and preleptotene spermatocytes reside, and the adluminal compartment, where the rest of the germ cells are located. This separation allows for the germ cells in the adluminal compartment to complete meiosis in an immunoprivileged compartment, while being sequestered from the circulatory and lymphatic systems (Li et al., 2012; Mruk and Cheng, 2015; Potter and De Falco, 2017; Smith and Braun, 2012). Although not entirely understood how, the BTB also allows for selective circulation of certain metabolites and molecules, while restricting the access of most nutrients (such as glucose) and hormones (Mruk and Cheng, 2015; Rondanino et al., 2017).

Outside the seminiferous tubules, in the interstitial space of the testis, other types of somatic cells reside, such as Leydig cells, resident testis macrophages and vasculature-associated cells. Leydig cells are the most important testosterone producing cells in the testis, with numerous roles in testis maturation and fertility. Their effects on spermatogenesis have been reported to occur through both direct and indirect (via other somatic cells in the testis) communication with germ cells. The rest of the somatic cells in the interstitial space, particularly resident macrophages, have also been shown to play important roles in testis morphogenesis, and spermatogenesis, through direct effects on spermatogonial cell differentiation (DeFalco et al., 2015; Heinrich and DeFalco, 2019; Potter and De Falco, 2017).

Although the main spermatogenesis milestones are conserved across mammals, large variation in timescales exists. In humans, the entire spermatogenic cycle lasts for around 76 days within the seminiferous tubules, followed by another several weeks of transport through the ductal system of the epididymis, adding up to around 3 months (Amann, 2008). In comparison to other mammals such as rodents, rabbits and bulls, the amount of sperm produced per gram of testis is much lower in humans, in part due to differences in the rate of spermatogonial cell renewal and their differentiating capacity (Ehmcke et al., 2006). Furthermore, differences in testis morphology between different individuals seem to be more pronounced in humans than in other mammalian species (Amann, 2008; De Rooij and Russell, 2000).

1.2 Spermatogonial cells

Spermatogonial cells ensure stable and continuous sperm formation throughout lifetime. Within the testis milieu, spermatogonial cells encompass only a small proportion of the total number of germ cells (around 0.3% of all testis cells are undifferentiated spermatogonia in rodents and 12% in humans) (Fayomi and Orwig, 2018). Their organization and molecular characteristics will be discussed in detail in the next subchapters.

1.2.1 Spermatogonial cell organization in the testis

Pioneering studies in the field, later confirmed through lineage tracing experiments, indicate that immediately after birth, the population of quiescent prospermatogonia follows 2 distinct trajectories: some of the prospermatogonia (approximately 70%) directly enter terminal differentiation to give rise to the first round of spermatozoa, whilst the remaining cells mitotically divide and form the undifferentiated A_s spermatogonial cells, presumably containing the real stem cell subpopulation (Griswold, 2016; Kluin and Rooij, 1981). Follow-up studies found that these 2 cell fates were distinguished by the presence (in prospermatogonia that form the initial pool of undifferentiated spermatogonial cells) or absence (in prospermatogonia that give rise to terminally differentiated spermatogonial cells) of the neurogenin-3 (NGN3) transcription factor (Yoshida et al., 2004, 2006). Undifferentiated spermatogonia comprise 3 distinct cell populations: A_{single} (A_s), A_{paired} (A_{pr}) and A_{aligned} (A_{al}) spermatogonia, with the latter 2 categories being characterized by the existence of

cellular “bridges” between the individual cells. A_{al} spermatogonial cells can further differentiate into A_{1-4} , Intermediate, and type B spermatogonia, which are irreversibly committed to meiosis, and further give rise to preleptotene spermatocytes (Figure 1-2) (Griswold, 2016; Oatley and Brinster, 2012; De Rooij, 2017).

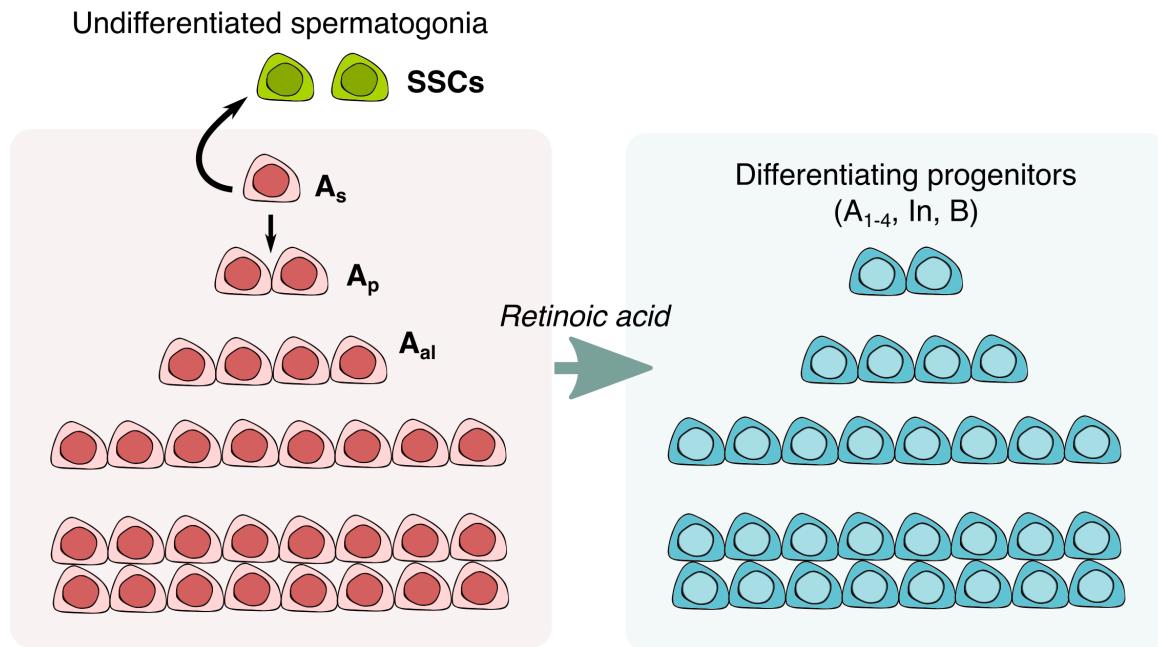


Figure 1-2 Schematic of the spermatogonial cell population in the mouse testis according to the A_s model, in which a subpopulation of A_s spermatogonia divides symmetrically to maintain the SSCs pool, while asymmetric division leads to more differentiated spermatogonia, committed to differentiation and sperm formation (adapted from (Helsel and Oatley, 2017)).

1.2.2 Spermatogonial stem cell identity: fragmentation and hierarchical models

The identity of the “true” stem cells of the spermatogonial lineage is still a topic of intense debate. Generally, there are 2 main models which support slightly different mechanisms for SSC self-renewal. The first model, known as the “fragmentation model” suggests that all undifferentiated type A spermatogonia have stem cell potential (Figure 1-3). In this model, stem cell renewal mainly takes place by fragmentation of A_{pr} and A_{al} spermatogonial cells, and rarely through self-renewal of A_s spermatogonia. As such, A_{pr} and A_{al} cells can take 2 trajectories: they can proceed with differentiation and form differentiating spermatogonial cells, while some of the cells undergo fragmentation to finally reform A_s or A_{pr} spermatogonia (Hara et al., 2014; Klein et al., 2010; Nakagawa et al., 2007, 2010).

The fragmentation model

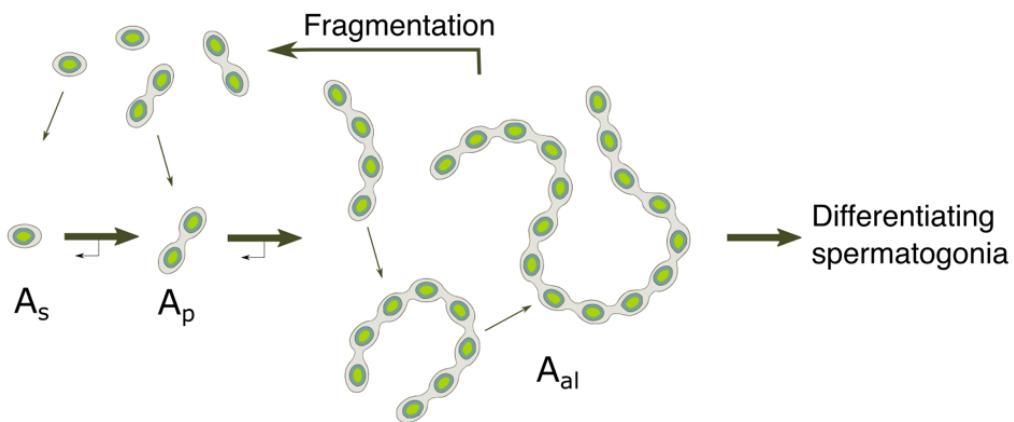


Figure 1-3 Schematics of the fragmentation model (adapted from de Rooij, 2017).

The more recent model proposed for explaining SSCs self-renewal is the “hierarchical A_s model” (Figure 1-4). In this case, only a subset of A_s spermatogonia possesses stem cell capacities and is able to self-renew, whereas most of the A_s population becomes gradually committed to form more differentiated spermatogonial cells (Aloisio et al., 2014; Chan et al., 2014; Helsel et al., 2017b, 2017a; Komai et al., 2014; Oatley et al., 2011; Sun et al., 2015). This self-renewal capacity has been tested through transplantation assays, in which SSCs potential was attributed to the subpopulation of cells capable of repopulating the recipient mouse testis. Such functional assessment led to a proposed molecular phenotype for these cells based on the concomitant expression of 3 transcription factors: ID4, BMI1 and PAX7.

The hierarchical A_s model

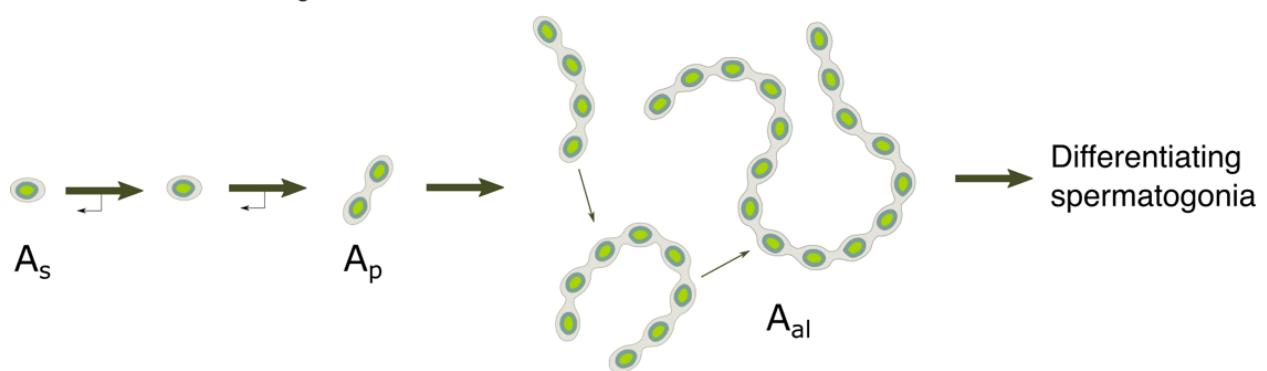


Figure 1-4 Schematics of the A_s model (adapted from de Rooij, 2017).

Presently, one of the most widely accepted markers for SSCs is the Inhibitor of DNA-binding 4 (ID4). Several transplantation studies performed by the Oatley group

revealed that only a subpopulation of A_s spermatogonia expresses high levels of ID4 ($ID4^{Bright}$), and it is this subpopulation which also possesses the highest colonization potential once transplanted into recipient testis (Chan et al., 2014; Helsel et al., 2017a; Oatley et al., 2011). We now know that most of the A_s spermatogonia express ID4, however in a very heterogeneous manner: around 20% of the cells are $ID4^{Bright}$, 40% of A_s cells have intermediate ID4 expression levels and the remaining 40% of A_s spermatogonia are $ID4^{Dim}$ (Helsel et al., 2014). Furthermore, several other factors have been attributed to the SSC subpopulation of spermatogonial cells, with enhanced colonization potential, such as GFRA1, RET, DMRT1 and ETV5 (Goertz et al., 2011; Hammoud et al., 2015; Hasegawa and Saga, 2014; Niu et al., 2011; Zhang et al., 2016b). Taken together, these studies argue for a vastly heterogeneous profile of SSCs, for which the general stem cell definition focused on self-renewal might not be sufficient to define a stem cell of the spermatogonial lineage.

Unlike in rodents, the human spermatogonial cell lineage encompasses fewer amplifying mitotic and meiotic divisions. The most primitive types of spermatogonial cells identified in non-human primates and humans are A_{dark} and A_{pale} undifferentiated spermatogonial cells, which are believed to represent the stem cell pool of the spermatogonial population, with A_{dark} cells slowly cycling and A_{pale} spermatogonia rapidly dividing. In contrast to rodents, where the transition from A_s to B-type spermatogonia entails 9 mitotic divisions, in humans one mitotic amplification delineates the transit of A_{pale} cells to B-type spermatogonia, which will further give rise to primary spermatocytes. Moreover, human spermatogonial cells only undergo a total of 5 amplifying divisions compared to 12 in mice and rats and 8 in monkeys, explaining in part the reduced number of resulting sperm cells/gram of testis (around 4 million in humans compared to 40 million in rodents and monkeys) (Fayomi and Orwig, 2018; Simorangkir et al., 2005).

1.2.3 Gene expression dynamics in spermatogonial cells

Elucidating the main transcriptional programs and the epigenetic factors that characterize spermatogonial cells across age and within subpopulations is a topic of major interest, given the essential role spermatogonial cells play in maintaining reproductive capabilities of an organism across lifetime. Furthermore, spermatogonial cells contain the only adult stem cell population capable of contributing genetic

information to the next generation, making them an attractive candidate for developmental, reproductive and inheritance studies. In the following subchapters I will discuss the main gene expression and epigenetic programs revealed by bulk and single cell sequencing studies.

1.2.3.1 Bulk RNA-sequencing experiments

Spermatogonial cells are characterized by dynamic gene expression programs depending on their undifferentiated or differentiating state. Pioneering studies in the field employed transplantation experiments to identify the first gene markers for distinguishing spermatogonial cell populations with high self-renewing capacity from spermatogonial cells committed to meiosis (Kubota et al., 2003, 2004b, 2004a). The first gene markers attributed to undifferentiated and differentiating spermatogonial cells were *Thy1* and *c-Kit*, respectively. Initial microarray analyses in pup testis identified numerous genes highly enriched in THY1⁺ cells compared to the THY1 depleted cell population (Oatley et al., 2006). Among these, *Gfra1*, *Ret*, *Lhx1*, *Pou5f1*, *Bcl6b* and *Id4* mRNA expression was 10-fold greater in the THY1⁺ cells.

Later studies using the *Id4-eGFP* transgenic reporter mouse line characterized the transcriptome profile of the ID4⁺ spermatogonial cells using bulk RNA-seq and found 1,450 genes differentially expressed between the ID4-eGFP^{Bright} and ID4-eGFP^{Dim} spermatogonial cells. This suggests a dynamic transcriptional landscape between progenitor and highly enriched stem cell subpopulation within the undifferentiated spermatogonial cell milieu (Helsel et al., 2017a). Furthermore, bulk RNA-seq from *Id4-eGFP* mice, and from several other studies using transgenic mouse lines or surface marker strategies to enrich for SSCs, confirmed the presence of genes such as *Bcl6b* (*Plzf*), *Etv5*, *Lhx1*, *T*, *Plzf*, *Gfra1*, *Pax7*, *Sall4*, *Nanos2* and *Pou5f1* as distinguishing factors for the stem cell enriched population of spermatogonia in the testis (Aloisio et al., 2014; Costoya et al., 2004; Hammoud et al., 2015; Helsel et al., 2017a; La et al.; Niu et al., 2011; Oatley et al., 2007).

1.2.3.2 Single cell RNA-sequencing experiments

Recently, an abundance of single-cell RNA-sequencing (scRNA-seq) studies have focused on compiling gene expression atlases of the mouse and human testis, without the need for prior cell selection/sorting procedure. The first published scRNA-seq dataset compared gene expression profile of THY1⁺ and ID4⁺ spermatogonial cells

from PND6 testis of WT and *Id4-eGFP* mice against a panel of 172 hand-picked genes (Hermann et al., 2015). The authors revealed distinct subsets of cells within each of these groups, and found that ID4-eGFP spermatogonia displayed less heterogeneity compared to THY1⁺ cells. Based on this initial dataset, they speculated that these subsets of cells could represent functionally distinct subgroups of undifferentiated spermatogonial cells and early differentiating progenitors (Hermann et al., 2015).

Subsequent experiments characterized the full transcriptome of ID4-eGFP spermatogonia from PND3 and PND7 pups, and confirmed that cellular heterogeneity exists within this cell population, and that only a subset of the ID4-eGFP cells are true SSCs, capable of self-renewal (Song et al., 2016). Additional studies focused on the scRNA-seq characterization of spermatogonial cell subpopulations from neonatal testis, and revealed distinct gene expression profiles characteristic to self-renewing and differentiating spermatogonial subsets (Green et al., 2018; Hermann et al., 2018; Mutoji et al., 2016). Self-renewing spermatogonia, as defined by the mRNA expression of known markers (*Gfra1*, *Etv5*, *Id4*, *Nanos2*, *Pax7*, *Tspan8*, and *Plzf*), were enriched for genes related to intracellular signalling pathways such as mTOR, MAPK cascade and genes involved in glycolysis. Differentiating spermatogonia were enriched for genes related to cell cycle control, while also upregulating mitochondrial function and oxidative phosphorylation in order to support proliferation and differentiation. More so, late stage differentiating spermatogonial cell populations were found to already transcribe meiotic genes, for rapid mRNA availability in the transition from spermatogonia to early spermatocytes (Hermann et al., 2018).

Aside from characterizing neonatal and early postnatal spermatogonial cell populations in the testis, several scRNA-seq studies have also investigated the transcriptome of adult testis cells. Interestingly, adult spermatogonial cells exhibited transcriptome features which are not found in the developing testis, indicative of changes taking place in gene expression of spermatogonia during testis maturation (Grive et al., 2019; Hermann et al., 2018). Previous bulk RNA-seq of spermatogonial cells from several developmental timepoints (PND0, 7, 12, 14 and postnatal week 20) also revealed differentially enriched gene programs between early postnatal and adult spermatogonial cells (Hammoud et al., 2015). At PND7, THY1⁺ spermatogonia exhibit enrichment in genes related to splicing and cell cycle and translation, whilst PND12 cells upregulate genes related to metabolic processes. PND14 spermatogonia already

exhibit enrichment in pathways related to meiosis and germ cell maturation, genes which remain highly expressed in adult THY1⁺ spermatogonial cells (Hammoud et al., 2015). Additionally, several transcription factors essential in stem cell maintenance and self-renewal (*Plzf*, *Taf4b*, *Etv5*, *Gfra1*) undergo a gradual decrease from PND7 to adulthood cells (Hammoud et al., 2015).

In the scRNA-seq performed by Hermann and collaborators, undifferentiated spermatogonia from PND6 testis were more heterogeneous in their gene expression profiles compared to the adult population (Hermann et al., 2018). A cluster of cells with high expression of genes involved in cell cycle regulation, proliferation and morphogenesis was unique to PND6 spermatogonia, whilst genes involved in RNA metabolism and ribosome biogenesis were highly expressed in both PND6 and adult spermatogonia. Furthermore, PND6 undifferentiated spermatogonia were enriched in pathways related to autophagy, DNA repair and genome integrity. Reciprocally, adult spermatogonia displayed elevated levels of genes expressed in mitochondrial function and oxidative phosphorylation (Hermann et al., 2018).

The most recent study to investigate the transcriptional dynamics of spermatogonial cells as mice age, revealed positive enrichment of pathways related to DNA damage response, cell cycle and phagocytosis in adult spermatogonial cells. In contrast, the authors found a downregulation of pathways related to transcriptional and translational regulation and in signalling factors such as *Kit*, *Fgf8*, *Fgfr1*, reinforcing the idea of a switch from autocrine to paracrine signalling in adult testis (Grive et al., 2019). However, some of these findings are not in agreement with the scRNA-seq study from Hermann and colleagues (Hermann et al., 2018), suggesting that even when not pre-enriched for certain sorting markers, differences in results could potentially be driven by other factors. These factors include number of sequenced cells, sequencing depth, replication strategy - technical vs biological and bioinformatic approach, and are important to consider especially in scRNA-seq. Furthermore, given the scarcity of the undifferentiated spermatogonial cell population in the adult testis, even slight changes in the experimental strategy could contribute to the differences in results we observe between some scRNA-seq studies (Suzuki et al., 2019).

Thanks to the rapid advancement of scRNA-seq, human spermatogonial cell profiling has also gained momentum in the recent years. Several studies have focused on investigating the gene expression signatures of spermatogonial cells, but also of other

spermatogenic cell types in the human testis, and defined the unique features of human spermatogenesis (Guo et al., 2017, 2018; Hermann et al., 2018; Sohni et al., 2019). One of the first scRNA-seq datasets from human spermatogonia employed SSEA4⁺ or KIT⁺ selection by FACS to enrich for undifferentiated and differentiating spermatogonial cells from adult testis, respectively (Guo et al., 2017). The authors revealed that undifferentiated human spermatogonia show upregulation of genes involved in FGF signalling, whilst the more differentiated spermatogonial progenitors upregulate genes involved in the WNT pathway. The study also included chromatin accessibility profiling of human spermatogonial cells using ATAC-sequencing (ATAC-seq), which revealed potential regulatory roles for repeat elements from the LTR class (Guo et al., 2017). scRNA-seq in infant and adult unsorted testis cells performed by the same lab validated previously found human spermatogonial cell markers and the FGF and WNT reciprocal upregulation in undifferentiated and differentiating human spermatogonia (Guo et al., 2018). Hermann and colleagues profiled unselected human spermatogonial cells using scRNA-seq, and compared their findings with mouse spermatogonial cells to identify conserved and unique transcriptome features between species (Hermann et al., 2018). Another study that performed scRNA-seq of spermatogonial cells from infants and adults revealed a new potential marker with increased selectivity for human SSCs, LPPR3 (Sohni et al., 2019). Furthermore, the authors identified a more heterogeneous profile of infant spermatogonia compared to that of the adults, with 2 unique cell populations displaying PGC-like gene signatures (Sohni et al., 2019).

1.2.4 Epigenetic regulation in spermatogonial cells

The heterogenous patterns of gene expression revealed by RNA-seq studies point towards distinct epigenetic mechanisms that may underscore the transcriptional differences which characterize spermatogonial cells (Green et al., 2018; Grive et al., 2019; Hammoud et al., 2014, 2015; Hermann et al., 2015, 2018). The following sections will describe current advances in the profiling of the main epigenetic marks in spermatogonial cells, both within distinct subpopulations, as well as in the transition from early postnatal to adult state.

1.2.4.1 DNA methylation

Although most of the DNA methylation (DNAm) patterns in spermatogonial cells are acquired before birth at the gonocytes stage, several studies investigating DNAm in spermatogonial cells from neonate and adult testis have suggested that patterns of DNAm keep on being acquired until the beginning of meiosis (Cheng et al., 2020; Hammoud et al., 2014, 2015; Oakes et al., 2007). Supportive of this theory are the high levels of the *de novo* DNA methyltransferase enzymes, DNMT3a and DNMT3b, which were found in spermatogonial cells (La Salle and Trasler, 2006; Shima et al., 2004).

In PND8 undifferentiated spermatogonial cells, Oakes and collaborators found that *de novo* methylation takes place in non-repetitive, non-GCI sequences in the genome, whereas some intergenic LTR families such as ERVKs are demethylated (Oakes et al., 2007). Furthermore, the authors found that some of the paternally imprinted genes still acquired DNAm at the spermatogonial stage, and finalized this process by the pachytene stage (Oakes et al., 2007). Similarly, findings from Hammoud and colleagues reinforced the idea that imprinting patterns are still being finalized after birth in spermatogonial cells (Hammoud et al., 2015). Additionally, the authors also found a progressive demethylation taking place at promoters of meiotic genes, olfactory receptors, protocadherins, cytokines and melanocortin receptors between PND0 and PND7/PND12 undifferentiated spermatogonia (Hammoud et al., 2015). When investigating DNAm profiles of adult undifferentiated and differentiating spermatogonia, Hammoud and colleagues found strikingly similar DNAm patterns at promoters of genes, with only a handful of meiotic genes exhibiting a loss in DNAm at their promoters (Hammoud et al., 2014).

Recently, Cheng and colleagues performed a comparison of DNAm patterns between 4 subpopulations of spermatogonia: PND6 ID4-eGFP spermatogonia enriched for SSCs, PND6 ID4-eGFP spermatogonia enriched in progenitors, adult undifferentiated and adult differentiating spermatogonia (Cheng et al., 2020). The study found that most CG-rich regions across the genome were highly methylated across all 4 cell subpopulations of spermatogonia, and that most of the differences in DNAm were evident in the lowly methylated regions. Furthermore, adult spermatogonia exhibited more pronounced differences between the 2 cell subpopulations than PND6 spermatogonia (Cheng et al., 2020). Taken together, the

available data suggest that DNAm is relatively stable in spermatogonial cells between undifferentiated and differentiating subpopulations, and across postnatal life.

1.2.4.2 Histone modifications

Histone modifications in spermatogonial cells have usually been studied together with DNAm, in order to infer consequences on the active or repressive status of chromatin regions. Hammoud and colleagues identified regions of bivalent chromatin, characterized by the presence of an activating H3K4me3 and a repressive H3K27me3, at numerous promoter regions of developmental genes, in both undifferentiated and differentiating spermatogonial subpopulations in the adult testis (Hammoud et al., 2014). Additionally, these bivalent chromatin regions were also marked by DNA hypomethylation. The authors hypothesized that such bivalent histone profiles are important for allowing transfer of these genes to the embryo in a transcriptionally competent state, while maintaining their silenced status in the germline (Hammoud et al., 2014). Furthermore, similar bivalent status of the promoter of developmental genes was found in undifferentiated spermatogonial cells in PND7 testis, suggesting this is a conserved mechanism across age in spermatogonial cells (Hammoud et al., 2014). Interestingly, a more recent study assessed the bivalent status of developmental genes specifically in the undifferentiated spermatogonial population, and found that certain gene promoters gained this bivalent status only at the progenitor stage, and did not have it in the SSC-enriched fraction (Cheng et al., 2020).

1.2.4.3 Non-coding RNAs

Few studies have investigated the small or long non-coding RNA content of spermatogonial cells in the mouse testis. Given the scarcity of these cells in the testis milieu, especially in adult mice, and the difficulty to obtain sufficient cell numbers for performing small RNA profiling, most of the studies to date used PND6-8 spermatogonial cells, when the ratio of spermatogonial cells to other cell types in the testis is maximal. One of the first studies to investigate small non-coding RNA (sncRNA) content of spermatogonial cells was performed by the Brinster lab, and specifically profiled microRNA (miRs) species in THY1⁺ spermatogonia enriched from PND6 pup testis (Niu et al., 2011). The authors compared the miRNA content of THY1⁺ spermatogonia and THY1⁻ somatic testis cells and found 139 differentially expressed miRs, with several members of the miR-291 family showing the highest difference in

expression. Furthermore, the study showed that miR-21, may play a role in the SSC population maintenance through an ETV5-mediated regulation of apoptosis (Niu et al., 2011). Further studies have identified roles in stem cell self-renewal and maintenance for other miRs in spermatogonial cells such as miR-20, miR-106a, miR-202 and miR-221 and 222 (Chen et al., 2017; He et al., 2013; Yang et al., 2013).

Prepachytene piRNAs are another species of small non-coding RNAs expressed in spermatogonial cells. These small non-coding RNAs are mostly transcribed from transposable element (TE) regions in the genome and are known to regulate TE silencing in spermatogonial cells (Rojas-Ríos and Simonelig, 2018; Saxe and Lin, 2011; Tóth et al., 2016; Watanabe et al., 2011).

Long non-coding RNAs (lncRNAs) have emerged in the past years as important regulators of gene expression in various tissues in the body, including the testis (Sahlu et al., 2020; Sun et al., 2013; Wichman et al., 2017). At PND6, pup testis exhibits more than 3000 differentially expressed lncRNAs compared to adult 8-week old mouse testis, with almost half of the lncRNA genes situated in intergenic regions (lincRNAs) (Sun et al., 2013). Furthermore, GO enrichment of the differentially expressed lncRNA-associated proteins revealed pathways related to transcription regulation as the most enriched ones (Sun et al., 2013). A more refined look at lncRNAs expressed in spermatogonial cells, revealed a high number of lncRNAs expressed in SSC-enriched spermatogonia and type A spermatogonia from pup testis. Out of these, 241 were specific to SSC- and 3639 to type A-enriched spermatogonia (Liang et al., 2014). Furthermore, parallel to the dynamic mRNA expression that takes place during spermatogenesis, lncRNAs also display a dynamic expression, with 1630 lncRNAs differentially expressed between spermatogonial cells and pachytene spermatocytes (Wichman et al., 2017). Interestingly, lncRNAs have also emerged as important regulators of pluripotency-associated genes in mouse embryonic stem cells (mESCs) (Bergmann et al., 2015). Given that some of these pluripotency pathways are shared by mESCs and spermatogonial cells, investigating how lncRNAs may contribute to pluripotency potential of spermatogonial subpopulations is of increasing interest.

1.2.4.4 Chromatin accessibility and 3D organization

Chromatin accessibility dynamics in spermatogonial cells has so far mainly been inferred from studies of histone marks, known for their activating or repressive

potential (Hammoud et al., 2014, 2015). However, a genome-wide direct assessment of chromatin accessibility in spermatogonial cells has only been reported in 3 studies (Cheng et al., 2020; Luo et al., 2020; Maezawa et al., 2018).

In the first study, the authors profiled chromatin accessibility changes in the mitosis-to-meiosis transition using the Assay for Transposase-Accessible Chromatin sequencing (ATAC-seq) and found that the overall number of ATAC-seq peaks (reflective of open chromatin, accessible to the transposase Tn5) decreased from undifferentiated spermatogonia to the round spermatids stage (Maezawa et al., 2018). Additionally, open chromatin regions localized in intronic and intergenic regions in PND8 differentiating spermatogonial cells (KIT^+) became inaccessible in pachytene spermatocytes, while there was little change in chromatin accessibility in regions around promoter of genes. However, this chromatin reorganization also comprised *de novo* formation of open chromatin regions in pachytene spermatocytes compared to differentiating spermatogonia (Maezawa et al., 2018). Similarly, another recent study revealed major chromatin accessibility reorganization taking place between spermatogonial and pachytene stages, with over 6,000 differentially accessible chromatin regions (Luo et al., 2020). With the spermatogonial cell population, a study from the Oatley lab assessed chromatin accessibility differences between Id4-eGFP^{Bright} and Id4-eGFP^{Dim} PND6 spermatogonia, and found an increase in accessibility in the Id4-GFP^{Dim} subpopulation (Cheng et al., 2020). Interestingly, this is in opposition to the findings from Maezawa and colleagues, which reported an overall decrease in accessibility taking place once spermatogonia advance to a more differentiated (Maezawa et al., 2018). Although the 2 studies used different techniques to isolate and enrich for spermatogonial cells, such contrasting findings will need further clarification from future studies employing ATAC-seq in mouse spermatogonial cells.

Very recently, 2 studies have employed chromosome conformation capture sequencing (HiC) in mouse spermatogonial cells to investigate the 3D spatial organization of chromatin across spermatogenesis (Luo et al., 2020; Vara et al., 2019). Notably, the studies used different FACS strategies for enriching spermatogonial cells and used PND6-8 pups and adult male mice, respectively. Even so, both studies found that chromatin compartmentalization declines in the transition from undifferentiated spermatogonia to differentiating spermatogonial cells and reaches a minimal state in

pachytene spermatocytes, before increasing again in round spermatids, and reaching the highest level in adult spermatozoa. In one of the studies, the authors suggested that the loss of these chromatin loops in the initial phase of meiosis was CTCF- and cohesin-independent, as both proteins had conserved abundance in all spermatogenic cell types investigated. Chromatin accessibility at the topological associated domains (TADs) of the chromatin loops was also not different between spermatogonial cells and pachytene spermatocytes, suggesting that 3D chromatin organization might not be functionally linked to chromatin accessibility at TADs. Lastly, meiosis-related genes and certain piRNAs were located in chromatin compartments that were inactive in spermatogonial cells, and became active in pachytene spermatocytes, a switch that correlated with the mRNA expression changes of these genes (Luo et al., 2020). Although these findings shed new light on the chromatin landscape of germ cells throughout the distinct phases of spermatogenesis, the impact of age and postnatal testis maturation on chromatin organization in the spermatogonial cells, has not yet been investigated.

1.3 Current methods for isolating spermatogonial cells

In the adult, spermatogonial cells represent a minority of the cells making up the testis. Thus, studies investigating the molecular characteristics of spermatogonia have usually made use of testis from prepubertal pups, in order to maximize the ratio of spermatogonial cells to other testis cells. In order to obtain highly enriched populations of undifferentiated spermatogonial cells, and even more so, of A_s spermatogonia and the SSCs they comprise, several techniques are routinely used in this field of research.

1.3.1 Surface marker – based cell sorting

During the past two decades, surface proteins specifically expressed by different spermatogonial cell populations have been identified. Making use of these expression patterns by using antibody-coated magnetic beads – magnetic activated cell sorting (MACS) or fluorescent antibodies against the specific surface proteins – fluorescence activated cell sorting (FACS) are the two most commonly cell sorting techniques presently used for the enrichment of spermatogonial cells from mice.

In MACS, antibody-coated magnetic beads are incubated with cells, and the separation of the positive cells from the negative populations takes place in a magnetic

field. The magnetic beads used in MACS are very small and physiologically inert, so that they do not pose any mechanical pressure on the cells, nor interfere with the normal activity or viability of the cells. Furthermore, MACS is less harsh on the cells than FACS, as the pressure and the speed that the cells are exposed to while passing through the magnetic column are not as high, making it better suited for enrichment procedures which precede *in vitro* culturing of the spermatogonial cells (Miltenyi et al., 1990). The Brinster lab originally used MACS using anti-THY1 magnetic beads to enrich for spermatogonial cells from PND6-9 pup testis prior to culturing (Kubota et al., 2003). Presently, MACS is still routinely used to enrich for spermatogonial cells from the initial testis cell suspension prior to culturing. However, unlike FACS, MACS is more time-consuming for multiparameter cell sorting and for high-throughput cell sorting, in which the number of samples required for subsequent molecular analyses is higher. Furthermore, FACS allows for differentiation between various cell populations which exhibit distinct levels of certain surface proteins, as is the case for human spermatogonial cells in which THY1^{Bright} and THY1^{Dim} spermatogonia have distinct molecular properties and stem cell capabilities (Hermann et al., 2009; Valli et al., 2014).

The usage of FACS for spermatogonial cell enrichment has been introduced by the Brinster group more than a decade ago. The experimental set-up assumed SSC-enrichment by FACS from pup and adult mouse testis, after staining with several combinations of surface proteins, followed by transplantation into recipient busulfan-treated testis (to deplete endogenous germ cells) to assess colonization potential (Kubota et al., 2003, 2004b, 2004a; Niu et al., 2011). The number of colonies formed by the transplanted cells was used as an indicator of the purity and viability of the sorted spermatogonia. As none of the surface proteins identified so far is uniquely expressed on the surface of spermatogonial cells, a multiparametric approach in which several surface proteins are used to perform FACS is preferred. The experiments by Brinster and his group identified that the MHC-I-Thy1⁺α6-integrin⁺αV-integrin^{-/Dim} surface phenotype is common across postnatal spermatogonial cells, and leads to high enrichment of undifferentiated spermatogonial cells in the sorted cell populations (Kubota et al., 2004b).

1.3.2 Genetic tagging – based cell sorting

In the last couple of years, significant advances in genetic labelling of the subpopulation of spermatogonial cells with stem cell capabilities have been made. Inhibitor of DNA-binding 4 (ID4) is a transcriptional repressor which has been initially identified by the Oatley group as being expressed by a subpopulation of spermatogonia within the A_s population in pup and adult testis (Oatley et al., 2011). Follow-up studies from the same group found that ID4-eGFP A_s spermatogonia sorted from an *Id4-eGfp* transgenic mouse line, exhibit enhanced colonization capabilities compared to ID4-eGFP⁻ spermatogonial cells, when transplanted into recipient testis. Furthermore, the regenerative capabilities of the cells were maximal in the ID4-eGFP^{Bright} spermatogonia compared to the ID4-eGFP^{Dim} population. Transcriptome characterization revealed distinct transcriptome profiles between these 2 subpopulations. Markers of undifferentiated spermatogonia such as *Lhx1*, *Plzf*, *Gfra1*, *Sall4*, *Nanos2*, and *Pou5f1* were significantly upregulated in the ID4-eGFP^{Bright} cells, whilst ID4-GFP^{Dim} spermatogonia showed an increased expression of genes associated with progenitor status such as *Ngn3* and *c-Kit* (Chan et al., 2014; Helsel et al., 2014; Law et al., 2019).

Several other mouse transgenic lines have been developed to study the subpopulations of undifferentiated spermatogonia which encompass the stem cells of the germline. OCT4 is a transcription factor with important roles in self-renewal of mESCs, which was also investigated in the context of SSCs (Liao et al., 2019a; Zeineddine et al., 2014). OCT4-eGFP⁺ cells in the testis localize to a broader subpopulation of type A undifferentiated spermatogonia compared to ID4-eGFP cells, which comprises a mix of spermatogonial cells during their transition from gonocytes to spermatogonia, self-renewing spermatogonia and spermatogonial cells committed to differentiation (Liao et al., 2019a).

One interesting characteristic of undifferentiated spermatogonial cells which was discovered and further used to develop a reporter mouse line is their high telomerase expression. Based on this characteristic, the telomerase reverse transcriptase (*Tert*) reporter mouse line was developed and employed to characterize the transcriptional landscape of these cells (Garbuzov et al., 2018; Pech et al., 2015).

The role of promyelocytic leukaemia zinc-finger (*Plzf*) in SSC maintenance is well established. PLZF is a transcriptional repressor highly expressed in undifferentiated spermatogonial cells of the A_s, A_{pr} and A_{al} subpopulations, and becomes less expressed during spermatogonial cells transition to more differentiated progenitors (Costoya et al., 2004). Recently, a tamoxifen-inducible *Plzf-mC/CreER* mouse line was generated (La and Hobbs, 2019). In this study, the authors argue that the expression of ID4 was not restricted to the population of A_s spermatogonia with stem cell capabilities in adult testis, and instead propose transcription factor PDX1 to specifically mark adult spermatogonia with potent stem cell potential (La and Hobbs, 2019).

Taken together, all of the studies described in this section provide valuable insight into the molecular characteristics of spermatogonial cells. They, however, do not reach a consensus regarding the molecular signature of distinct spermatogonial subpopulations. One important reason for this could be the existence of spermatogonia as a continuum of cell states, in which stem cell potential may be dictated by the interplay between several factors, such as niche-derived cues, postnatal age and environmental stimuli.

1.4 Environmental insults on spermatogonial cells

The impact of environmental cues on the development of various tissues in the body is unquestionably important. Although in the past two decades a growing number of studies have explored the potential effects of environmental influences on the germline, our understanding of the underlying molecular mechanisms remains limited (Anway et al., 2005; Bohacek and Mansuy, 2015; Carone et al., 2010; Chytrova et al., 2010; Gapp et al., 2014; Rodgers et al., 2015; Sharma et al., 2018). The majority of these studies investigated the potential epigenome alterations induced by various exogenous factors such as diet, heat, toxicant exposure and stress in adult spermatozoa, however little research has focused on the effects on spermatogonial cells (De Castro Barbosa et al., 2016; De and Kassis, 2017; Minkina and Hunter, 2017). Given the unique properties of spermatogonial cells, as the only adult stem cells to contribute with genetic (and potentially epigenetic) information to the next generation, an in-depth understanding of how environmental cues could alter the spermatogonial milieu is essential. In the upcoming subchapters I will briefly

summarize the most well-characterized effects of different environmental factors on spermatogonial cells and their implications for spermatogenesis and sperm integrity.

1.4.1 Toxicant exposure

One of the most extensively researched class of factors altering the germline are endocrine disrupting chemicals (EDCs). Correlative evidence from humans and further mechanistic insight from rodent models of EDC exposure suggest that the effects on developing testis could in part explain the decline in human fertility in both males and females, which has been reported in the past several decades (Carlsen et al., 1992; Sharpe and Skakkebaek, 1993). Prenatal exposure of rats and mice to phthalates led to alterations of both developing germ cells and the somatic testis milieu of the developing testis (Ferrara et al., 2006; Gaido et al., 2007; Mahood et al., 2007).

Ferrara and colleagues showed that exposure to Di(n-Butyl) Phthalate (DBP) from embryonic day 13.5 to 21.5 leads to delayed entry of gonocytes into quiescence, and to a decrease in the number of spermatogonial cells during the first 2 postnatal weeks (Ferrara et al., 2006). However, these effects were not visible anymore in the adult testis of the exposed rats. Porro and colleagues made use of *Oct-eGFP* transgenic mice to investigate the effects of another EDC, 17 α -ethinylestradiol (EE2) on gonocytes and spermatogonial cell development (Porro et al., 2015). Maternal administration of EE2 between day post coitum (dpc) 5.5 and PND7 led to an increased expression of *Oct4* in spermatogonial cells at PND7, and decreased sperm counts in adult offspring. These findings prompted the authors to conclude that exposure to EED could favour a stem cell profile of spermatogonia over differentiation, leading to a decreased sperm count (Porro et al., 2015). Other studies that focused particularly on prenatal exposure to DBP found altered proliferation and morphology of developing germ cells before birth, however postnatal and adult timepoints were not investigated (Gaido et al., 2007; Mahood et al., 2007).

The first study to provide evidence that early postnatal exposure to EDCs alters spermatogonial cells came from Vrooman and collaborators (Vrooman et al., 2015). The authors treated male pups from 2 different mouse strains (CD-1 and C57Bl/6J) with either Bisphenol-A (BPA) or ethinylestradiol between PND1-12, and assessed the effects of these 2 toxicants on spermatogonial cells in the developing and adult testis. One first important finding from the study was that aberrant meiotic recombination

leading to pachytene spermatocytes death takes place in exposed juvenile males, and persists into adulthood. Interestingly, sensitivity to the exposure was dependent on the genetic background of the mice, with CD-1 mice being sensitive to both BPA and EE2, while C57Bl/6J mice did not exhibit any germline abnormalities. To test if the alterations in meiotic recombination are due to germ cells or the somatic milieu, the authors performed spermatogonial cell transplantation from exposed and control juvenile mice, and found that the recombination phenotype was persistent, suggestive of alterations in the spermatogonial, and not in the somatic supportive cells (Vrooman et al., 2015). These findings highlight the importance of a healthy spermatogonial cell population for the integrity of future germ cells, and in adult sperm formation. Furthermore, some of these results suggest that even brief exposures in early postnatal life could have long-lasting effects on the germline.

Ethanol toxicity has also been investigated in connection to potential effects on spermatogonial cell proliferative capabilities. A single subcutaneous injection of ethanol in PND12, but not in PND7 male mice led to increased apoptosis of spermatogonial cells (Caires et al., 2012). Gene expression assessment at PND7, 12 and PND42 revealed an altered gene expression profile for several genes important in GDNF and VEGF signalling. Similarly, an earlier study in adult rats found an increased apoptosis in spermatogonia and spermatocytes following repeated oral intake of an ethanol-containing liquid (Zhu et al., 2000).

1.4.2 Other environmental exposures: heat, diet, chronic stress

Investigating the effects of other types of environmental cues than toxicants on the germline has so far been limited. Recently, a class of genes called melanoma antigens (MAGE), highly expressed in spermatogonial cells, has been described to play an important role in protecting spermatogonial cells against nutrient, heat and genotoxic stress (Fon Tacer et al., 2019; Lee et al., 2020). *Mage-a* KO in spermatogonial cells of mice subjected to reduced food intake led to an increased seminiferous tubules damage (Fon Tacer et al., 2019). Furthermore, *Mage-b4* KO mice exposed to heat stress displayed decreased testis size, reduced fertility and an accumulation of stress granules in the undifferentiated spermatogonial cells, but did not show any increased sensitivity to busulfan treatment, suggesting a protective effect of *Mage-b4* only against a specific type of stressor (Lee et al., 2020).

Effects of traumatic stress exposure on spermatogonial cells have so far remained largely unexplored. In rats, repeated immobilization stress led to increased apoptosis of spermatogonial cells and spermatocytes 24h after the end of the stress paradigm, potentially caused by the transient testosterone reduction measured during the immobilization period (Lee et al., 2020). In mice, chronic unpredictable stress performed daily for 4 weeks also led to increased apoptosis of spermatogonial cells in the exposed mice, and to a mild altered expression of Kisspeptin granules in the arcuate hypothalamic nucleus (Hirano et al., 2014).

1.5 Overview

Spermatogonial cells play a central role in the postnatal establishment of germ cells and in sperm production across life. An overview of mammalian spermatogenesis, as well as the molecular transitions that define spermatogonial cell states across postnatal life, have been discussed in **Chapter 1**.

While distinct transcriptomes have been described for spermatogonial cells from early postnatal and adult testis, the underlying chromatin landscape has remained mostly unexplored. In **Chapter 2** we characterise the accessibility dynamics of open chromatin between early postnatal and adult spermatogonia, and the potential implications for gene expression regulation.

Environmental stressors during lifetime can change the epigenome of germ cells, with potential consequences for the offspring phenotype. In **Chapter 3** we investigate the possibility that early postnatal stress affects spermatogonial cells, by comparing the transcriptome and chromatin landscape of spermatogonial cells from exposed and control mice.

Lastly, the findings of this thesis are summarized and discussed in **Chapter 4**. In this chapter we also propose additional experiments to further investigate the regulatory roles of chromatin accessibility dynamics in spermatogonial cells, across age and following environmental experiences.

2 Postnatal characterization of mouse spermatogonial cells reveals distinct chromatin regulatory landscapes in the developing and adult testis

Irina Lazar-Contes^{1,2,*}, Deepak K. Tanwar^{1,2,3,*}, Pierre-Luc Germain^{2,3}, Isabelle M. Mansuy^{1,2,#}

1. Laboratory of Neuroepigenetics, Medical Faculty of the University of Zurich and Department of Health Science and Technology of the ETH Zurich, Brain Research Institute, Zurich, Switzerland
2. Institute for Neuroscience, Swiss Federal Institute of Technology, Zurich, Switzerland
3. Statistical Bioinformatics Group, Swiss Institute of Bioinformatics, Zurich, Switzerland

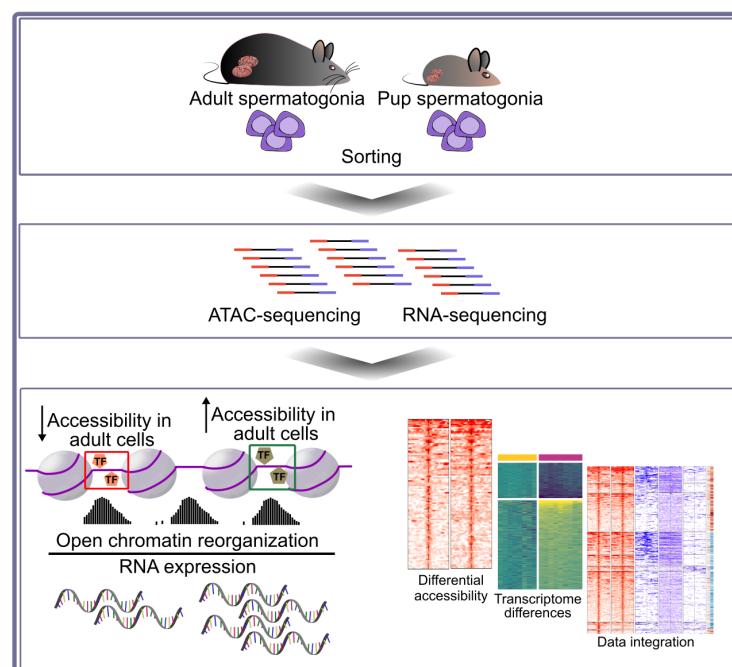
*Authors share equal contribution

#Correspondence: mansuy@hifo.uzh.ch

Available on *BioRxiv*, doi: <https://doi.org/10.1101/2020.08.20.259374>

In preparation for *Developmental Cell*

Graphical Abstract



Highlights

Open chromatin reorganization in spermatogonial cells between early postnatal and adult stage

Data integration reveals distinct chromatin landscapes at genes important for spermatogonial cell maintenance and proliferation

Regions of differential accessibility are enriched in numerous transcription factor binding sites

Less accessible LTR-ERVKs and more accessible LINE L1s associate with distinct gene families in adult spermatogonial cells

2.1 Abstract

In mammals, spermatogonial cells are undifferentiated male germ cells that exit quiescence after birth, and self-renew or differentiate to produce spermatogenic cells and functional sperm across life. Recent rodent studies showed that spermatogonial cells have a highly dynamic transcriptome between early postnatal life and adulthood. However, what drives this developmental transition at the chromatin level is not fully understood. We characterized chromatin accessibility in early postnatal and adult spermatogonial cells in mice using ATAC-seq, and integrated the data with a range of transcriptomic and epigenomic features across age. We show that extensive chromatin remodelling occurs in spermatogonial cells across postnatal age, and that this correlates with distinct biological pathways and transcription factor (TF) motif enrichment. We further identify genomic regions with significantly different chromatin accessibility that are marked by distinct histone modifications, and are situated in proximity of transcription start sites (TSSs) of genes important for spermatogonial cell maintenance and proliferation. Some of the regions with increased accessibility correspond to transposable elements (TEs) enriched in multiple TF motifs, and with increased RNA expression in adult spermatogonia. Taken together, our results underscore the dynamic nature of open chromatin in spermatogonial cells across postnatal life, and reveal novel profiles of chromatin organization, histone modifications and gene expression between developing and adult spermatogonia.

2.2 Introduction

Spermatogonial cells are the initiators and supporting foundation of spermatogenesis in various species, including mammals. In mice, they become active one to two days after birth, when they exit mitotic arrest and start dividing, to populate the basement membrane of seminiferous tubules. During the first week of postnatal life, a subpopulation of spermatogonial cells proliferates and gives rise to undifferentiated A_{single} (A_s), A_{paired} (A_{pr}) and A_{aligned} (A_{al}) cells. The remaining spermatogonia differentiate to form chains of daughter cells that become primary and secondary spermatocytes around postnatal day (PND) 10 to 12. Spermatocytes undergo meiosis and give rise to haploid spermatids that develop into spermatozoa. Spermatozoa are then released in the lumen of the seminiferous tubules, and continue to mature in the epididymis until becoming capable of fertilization at PND 42-48 (Kubota and Brinster, 2018; Oatley and Griswold, 2017; De Rooij, 2017).

Recent work showed that distinct transcriptional profiles characterize spermatogonial cells in early postnatal life (Green et al., 2018; Hammoud et al., 2014, 2015; Hermann et al., 2018; Law et al., 2019). During the first week of postnatal development, spermatogonia display unique features necessary for the rapid establishment and expansion of the cell population along the basement membrane. These include high expression of genes involved in cell cycle regulation, stem cell proliferation, transcription and RNA processing (Grive et al., 2019). In comparison, in the adult testis, the focus lies in the maintenance of a steady cell population, which balances proliferation and differentiation capabilities to ensure sperm formation across life. Previous reports have revealed that adult spermatogonial cells prioritize pathways related to paracrine signalling and niche communication, as well as mitochondrial function and oxidative phosphorylation (Grive et al., 2019; Hermann et al., 2018).

Concomitant with gene expression changes, histone tail modifications and DNA methylation differences in spermatogonial cells from distinct postnatal stages have also been described (Hammoud et al., 2014, 2015). However, little is known about the accessible chromatin landscape of spermatogonial cells during the transition from early postnatal to adult stage, and how it could facilitate these dynamic transcriptional changes.

To investigate open chromatin reorganization in the transition from early postnatal to adult spermatogonia, we employed the Omni-ATAC protocol, an improved version of the Assay for Transposase-Accessible Chromatin with high-throughput sequencing (ATAC-seq) (Corces et al., 2017). Our results revealed extensive chromatin remodelling, in particular an increase in accessibility, of open chromatin regions in adult spermatogonia compared to PND15 cells. We further characterized novel histone signatures at regions of differentially accessible chromatin associated with genes dynamically expressed across postnatal age, by integrating published chromatin immunoprecipitation sequencing (ChIP-seq), bisulphite sequencing (BS) and transcriptome data. Lastly, by investigating chromatin accessibility at transposable elements (TEs), we described previously uncharacterized changes in accessibility at LTR and LINE L1 subtypes between developing and adult spermatogonial cells. Taken together our results suggest an important contribution of open chromatin reorganization to the diverse transcriptome of developing and adult spermatogonial cells.

2.3 Results

2.3.1 Fluorescence activated cell sorting (FACS) enriches spermatogonial cells from developing and adult mouse testis

We collected testes from 8- and 15-days old pups (PND8 and 15), and from adult males at postnatal week (PNW) 20, and prepared cell suspensions by enzymatic digestion of the testes. To achieve high enrichment of spermatogonial cells in our cell populations, we employed fluorescence-activated cell sorting (FACS) using the surface phenotype established by Kubota et al. (Figure S 2-1A) (Kubota et al., 2004b). The enriched spermatogonial populations were further used for Omni-ATAC ($n = 6$ PND15 and $n = 5$ adult samples) and RNA-seq ($n = 9$ PND8 and $n = 8$ PND15 samples) approaches. To evaluate the purity of our sorted samples, we performed immunocytochemistry using PLZF, a well-established marker of undifferentiated spermatogonia (Costoya et al., 2004). This revealed that our FACS strategy generates cell populations that are 85-95% PLZF⁺, compared to only 3-6% PLZF⁺ cells identified in our samples before FACS (Figure S 2-1B). Using our newly generated RNA-seq data, we also examined the expression of known pluripotency, spermatogonial stem cell and somatic markers (Leydig and Sertoli cells). We observed that cells in our sorted preparations expressed high level of stem cell and undifferentiated spermatogonial markers, but low level of somatic cells markers, confirming spermatogonial cell enrichment (Figure S 2-1C).

2.3.2 Accessible chromatin is reorganised in adult spermatogonia compared to early postnatal stage

Previous RNA-seq analyses in rodents showed that spermatogonial cell transcriptome changes during postnatal life, with distinct transcriptional programs revealed in neonatal, juvenile and adult stages (Grive et al., 2019; Hammoud et al., 2015). To determine if changes in chromatin organization underlie the dynamic transcriptome of spermatogonial cells as mice age, we profiled chromatin accessibility of spermatogonial cells at PND15, and from adult testis. We relied on the Omni-ATAC protocol, which shows a higher signal-to-noise ratio compared to the original ATAC-seq, and allows a lower number of starting cells (Corces et al., 2017). Accessible regions were identified by peak-calling on the merged nucleosome-free fragments (NFF) from all PND15 and adult samples. Following the removal of lowly enriched

regions, we included 158,978 regions in our downstream analyses (see Methods section for details). Most of the Tn5-accessible regions were intergenic (38%), located in gene bodies (33%) or in proximity of a gene transcription starting site (TSS) (+/-1 kb from TSS, 28%) (Figure S 2-2A). Differential accessibility analysis revealed 3212 differentially accessible regions between PND15 and adult spermatogonia (FDR \leq 0.05, abs Log₂ fold change (FC) \geq 1), with the majority of the regions showing a gain in accessibility in adult stage (Figure 2-1A and Table S1). The regions of differential accessibility were predominantly localized in intergenic regions and introns (34% in introns and 45% in intergenic regions). Only 15% of all differentially accessible regions resided +/- 1kb from the TSS of a gene (Figure 2-1B).

To investigate the biological significance of the regions with differential accessibility, we performed Gene Ontology (GO) analysis using the Genomic Regions Enrichment of Annotations Tool (GREAT) (McLean et al., 2010). Overall, regions of increased accessibility in adult spermatogonia were enriched in pathways associated with cell fate and stem cell population maintenance, protein metabolism and RNA metabolic processes (Figure 2-1C and Table S1). Given the high number of more accessible regions situated at different genomic locations, we asked if this genomic distribution is associated with distinct biological functions. We found that regions located in gene bodies (mainly introns) were specifically enriched for GO terms related to reproduction and protein metabolism, whilst regions close to (+/- 1 kb), or overlapping TSSs of genes, were enriched for cell fate specification and tissue morphogenesis (Figure S 2-2B and Table S1). We observed that the regions of less accessibility in adult spermatogonial cells were predominantly localized in intergenic regions (Table S1). GO enrichment in less accessible chromatin regions revealed enrichment of pathways related to embryonic development (Figure 2-1C and Table S1).

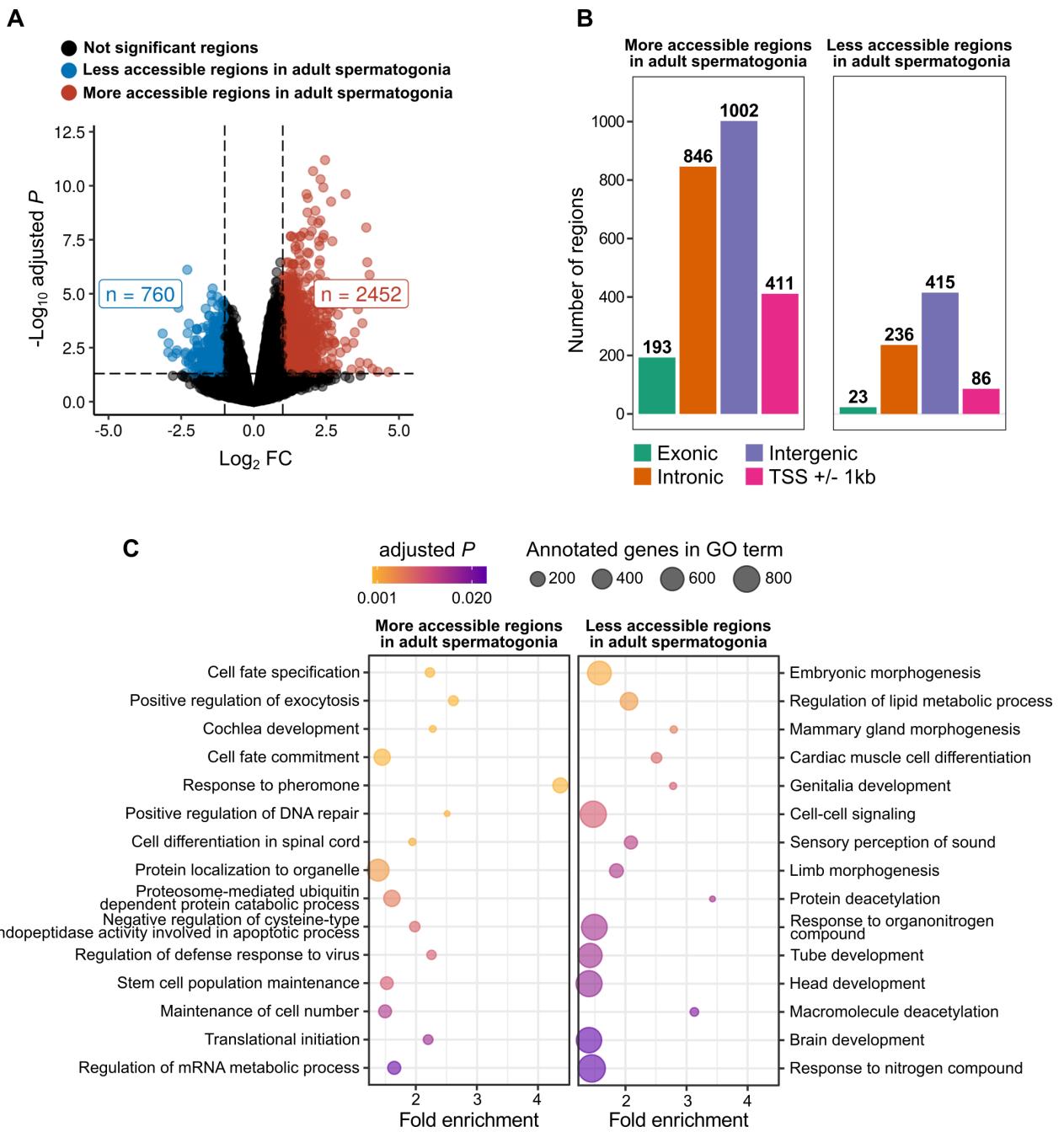


Figure 2-1 Regions of differential chromatin accessibility from PND15 to adult spermatogonial cells associate with distinct gene pathways.

(A) Volcano plot of differentially accessible ATAC-seq regions (adjusted $P \leq 0.05$ and absolute Log₂ FC ≥ 1) between adult and PND15 spermatogonial cells;

(B) Bar plot illustrating the genomic distribution of differentially accessible ATAC-seq regions between adult and PND15 spermatogonial cells. Genomic regions were categorized in intronic, exonic, intergenic and +/- 1kb from the TSS of a gene;

(C) Dot plots of top enriched GO biological processes (BP) terms for regions with increased and decreased chromatin accessibility in adult spermatogonia compared to PND15. The size of the dot indicates the number of genes in the term, and the colour of each dot corresponds to the adjusted P value of the term's enrichment.

2.3.3 Differentially accessible chromatin regions associate with distinct gene expression dynamics and histone marks

To better understand the chromatin accessibility – transcriptome relationship across postnatal age, we conducted RNA-seq on spermatogonial cells from PND8 and PND15 testis, and further used previously generated literature datasets from PND14 and postnatal week (PNW) 8 THY1⁺ spermatogonial cells which were categorized as adult spermatogonia (Hammoud et al., 2014, 2015). We found 719 differentially expressed genes ($FDR \leq 0.05$, $\text{abs Log}_2 FC \geq 1$) between PND8 and PND15 cells, indicative of dynamic transcriptional programs during early phases of spermatogonial cell proliferation and differentiation (Figure 2-2A and Table S2).

The comparison between PND14 and adult spermatogonia transcriptomes, while not affording the same type of analysis due to the low sample size, suggested even broader transcriptional changes taking place between early postnatal and steady - state adult spermatogonial cells ($\text{Log}_2 CPM \geq 1$ and $\text{abs Log}_2 FC \geq 1$) (Figure 2-2B and Table S2). To identify the biological significance of these dynamic gene expression profiles, we conducted pathway analysis using Fast Gene Set Enrichment Analysis (FGSEA) (Korotkevich et al., 2016; Subramanian et al., 2005). Between PND8 and PND15 spermatogonia we revealed a downregulation of pathways related to RNA processing and splicing, cell cycle, redox homeostasis and protein catabolism, and an upregulation of terms associated with cellular transport, exocytosis and signal transduction (Figure 2-2C and Table S2).

GSEA from PND14 and adult spermatogonia revealed that, similarly to the transition from PND8 to PND15, pathways related to RNA processing, ribosome biogenesis, and cell cycle were downregulated in adult spermatogonia (Figure 2-2D and Table S2). We also found a downregulation of pathways related to developmental programs and mitochondrial functions. In contrast, upregulated pathways were related to spermatogenesis, as well as numerous processes involving cytokine signalling (Figure 2-2D and Table S2).

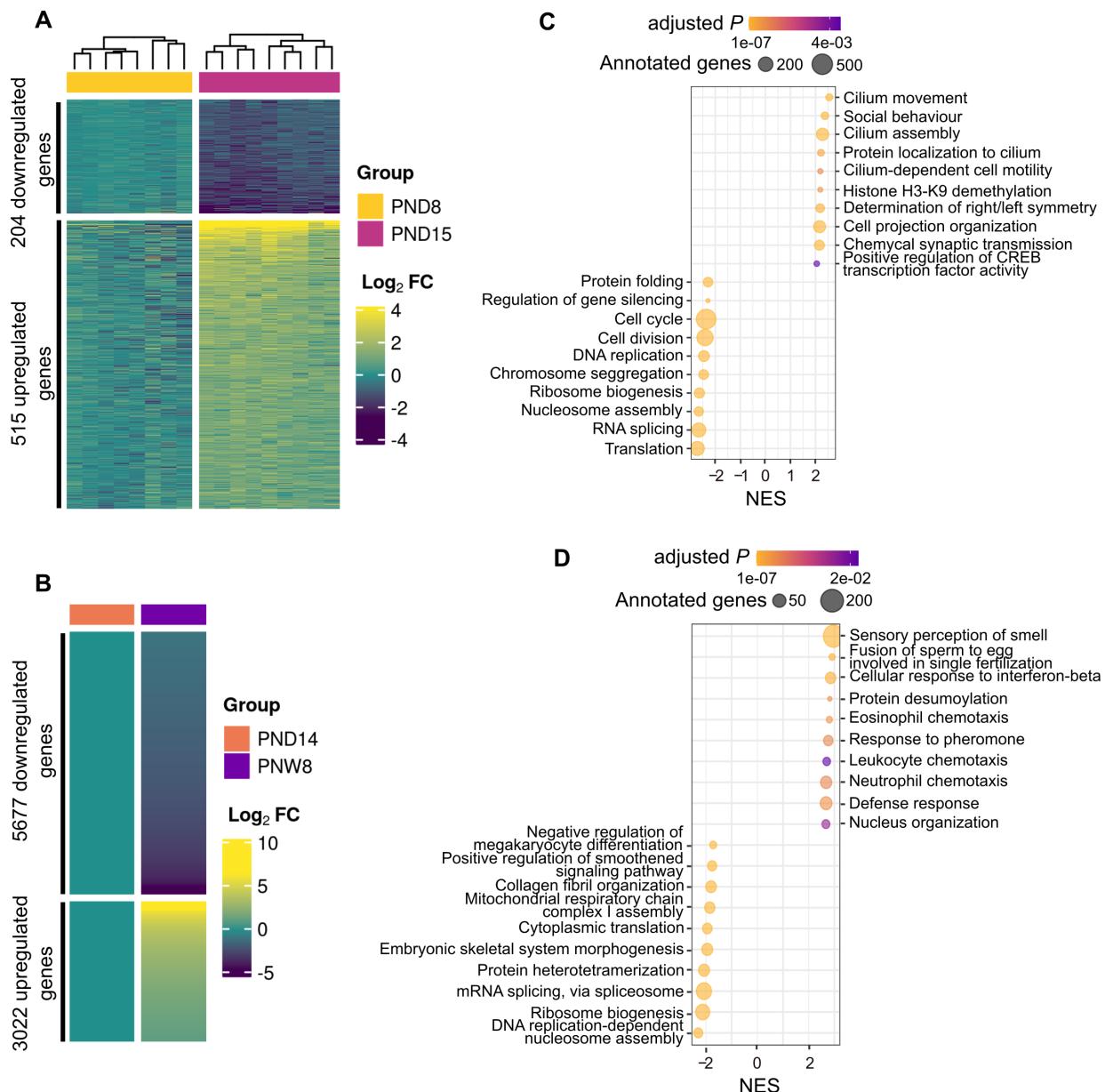


Figure 2-2 Transcriptome dynamics between early postnatal and adult stage spermatogonial cells.

(A) Heatmap of differentially expressed genes (adjusted *P* ≤ 0.05 and abs Log₂ FC ≥ 1) between PND15 (*n* = 8) and PND8 (*n* = 9) spermatogonial cells. Shown are the Log₂ FC with respect to the average of the PND8;

(B) Heatmap of differentially expressed genes (Log₂ CPM ≥ 1 and abs Log₂ FC ≥ 1) between adult (PNW8) spermatogonia (*n* = 1) and PND14 (*n* = 1) from literature RNA-seq datasets;

(A, B) Genes are ordered by Principal Component Analysis (PCA) method using seriation (R package);

(C, D) Dot plots of top 10 enriched GO BP terms (adjusted *P* ≤ 0.05) from GSEA analysis of PND15 vs PND8 and PNW8 vs PND14 comparison, respectively. GO terms are summarized by REVIGO and ordered by their normalized enrichment scores (NES). The size of the dot indicates the number of expressed genes annotated in the GO term, and the colour corresponds to the adjusted *P* value.

Next, we integrated the chromatin accessibility and transcriptome findings from early postnatal to adult spermatogonial stage. Additionally, we mined previously published ChIP-seq and BS data from THY1⁺ spermatogonia and investigated histone mark (H3K4me3, H3K27ac, H3K27me3) distribution and DNAm patterns (Hammoud et al., 2014, 2015) at the regions of differentially accessible chromatin. Notably, ChIP-seq data was only available for adult spermatogonial cells (PNW8), whilst DNAm was available for PND7, PND14 and adult stages.

For integrating our ATAC-seq data with the RNA-seq, ChIP-seq and BS datasets, we first divided differentially accessible regions into proximal (situated less than +/- 2.5 kb from a TSS) and distal (situated more than +/- 2.5 kb from a TSS), following ENCODE practice (Harrow et al., 2012a; Myers et al., 2011; Thurman et al., 2012). We further grouped proximal regions based on the change in chromatin accessibility and expression of the nearest gene between PND14 and adult spermatogonia, which led to 6 distinct categories. The first 2 most abundant categories contained regions with increased chromatin accessibility and upregulated nearby genes - Category 1, and increased chromatin accessibility and downregulated nearby genes - Category 2. Category 3 and 4 comprised proximal regions of decreased chromatin accessibility for which gene expression was either downregulated or upregulated, respectively. Categories 5 and 6 were proximal regions with increased and decreased chromatin accessibility respectively, for which the expression of the nearest gene was not detected in spermatogonial cells (Figure 2-3A and Table S3).

For a subset of regions in Category 1, ChIP-seq overlap revealed the presence of active H3K4me3, H3K27ac or dual H3K4me3/H3K27ac, and an overall lack of H3K27me3 (Figure 2-3B and Table S3). Notably, several of the genes in Category 1 for which chromatin was also marked by histone modifications, are known regulators of stem cell potential. *Pdk1* promoter region was marked by a dual H3K4me3/H3K27ac, while *Pdk1* mRNA showed a slight upregulation in adult spermatogonia (Figure S 2-3A). *Pdk1* (phosphoinositide-dependent protein kinase 1) is a glycolysis factor, important for spermatogonial stem cell self-renewal (Chen et al., 2020; Kanatsu-Shinohara et al., 2016). In contrast, *Gata2* promoter region was marked by a bivalent H3K4me3/H3K27me3 mark, while its expression showed a slight upregulation across testis maturation (Figure S 2-3A). *Gata2* is a known target of

NANOS2, an essential regulator of spermatogonial stem cell potential (Barrios et al., 2010; Sada et al., 2009). Other exemplary genes in Category 1 included pyruvate cellular carriers *Slc25a18*, *Slc23a1* and *Slc2a5*, suggesting differences in glycolysis regulation in adult spermatogonial cells (Table S3). Notably, we found an increased chromatin accessibility at the TSS of GDNF receptor *Gfra2* (Figure 2-3C). At mRNA level, *Gfra2* displayed a marked upregulation in adult spermatogonial cells, indicating an increased utilization of GFRA2 receptors in adult spermatogonial cells compared to early postnatal stages, in which GFRA1-mediated signalling is dominant (Figure 2-3C) (Grive et al., 2019; Hammoud et al., 2015).

Interestingly, the highest number of differentially accessible chromatin regions were in Category 2, and included proximal regions with increased chromatin accessibility and decreased expression of nearby genes in adult spermatogonia, indicative of active repression taking place (Figure 2-3A and Table S3). Members of this category included developmental genes such as *Tbx4*, *Satb1* and *Hmx1* (Figure 2-3B and Figure S 2-3B). ChIP-seq overlap revealed that some of the regions nearby developmental genes displayed a poised H3K27me3/H3K4me3 mark (Figure S 2-3B). In addition, GO enrichment analysis on genes in Category 2 found an enrichment for pathways related to regulation of cell cycle, RNA processing, DNA repair and cell division (Table S4). An exemplary gene is *Fgf8*, important for maintenance of undifferentiated spermatogonia in the testis, via FGFR1 signalling (Hasegawa and Saga, 2014). *Fgf8* showed increased chromatin accessibility at the TSS and a downregulated expression in adult spermatogonia, in agreement with recent findings from scRNA-seq data which reported a downregulation of FGFR1-mediated signalling with age (Figure S 2-3B) (Grive et al., 2019; Hasegawa and Saga, 2014).

Regions in Category 3 displayed decreased chromatin accessibility and a downregulation of nearby genes in adult spermatogonia and were mostly depleted of any of the 3 histone marks investigated (Figure 2-3B). GO enrichment on genes in this category genes revealed enrichment for developmental pathways and WNT signalling (Table S4). A notable example we identified in this category is *Pdgfra*, a gene involved in the hepatic stellate cell activation pathway, which was recently identified by scRNA-seq to be upregulated in spermatogonial stem cells in the immature testis compared to adult stage (Hermann et al., 2018). *Pdgfra* displayed a marked downregulation in adult spermatogonia, and a decrease in chromatin accessibility overlapping the TSS

(Figure 2-3C). Another example of a gene important for early postnatal spermatogonial cell establishment is *Dap2ip*, which displayed a marked decrease in chromatin accessibility at its TSS and a decreased expression in adult spermatogonial cells (Figure S 2-3C). Surprisingly, we also identified a category of proximal regions (Category 4) with decreased accessibility (<20 regions) at genes which were upregulated in adult spermatogonial cells and with previously uncharacterized roles in spermatogonial cells (Figure 2-3A and Table S3).

Notably, DNAm profiles across postnatal stages did not show drastic changes in any of the 6 categories of proximal regions, suggesting a relatively stable DNAm profile in the transition from early postnatal to adult stage (Figure S 2-4). Aside from proximal regions, we also identified numerous distal regions with differential chromatin accessibility between PND15 and adult (Figure S 2-5A and B). Similar to proximal regions, accessibility in distal regions mainly increased in adult spermatogonial cells compared to early postnatal stage. When integrating the literature ChIP-seq data, we observed enrichment for H3K4me3, H3K27ac and H3K27me3 at a small number of the differentially accessible distal regions, indicative of potential regulatory roles (Figure S 2-5A and B and Table S3). Similar to proximal regions, DNAm levels did not display major changes in distal regions (Figure S 2-5A and B).

Taken together, our data integration reveals novel associations between open chromatin regions of differential accessibility, histone marks and dynamically expressed genes in spermatogonial cells during testis maturation, and reveal how chromatin accessibility may contribute to the differential utilization of signalling pathways across age.

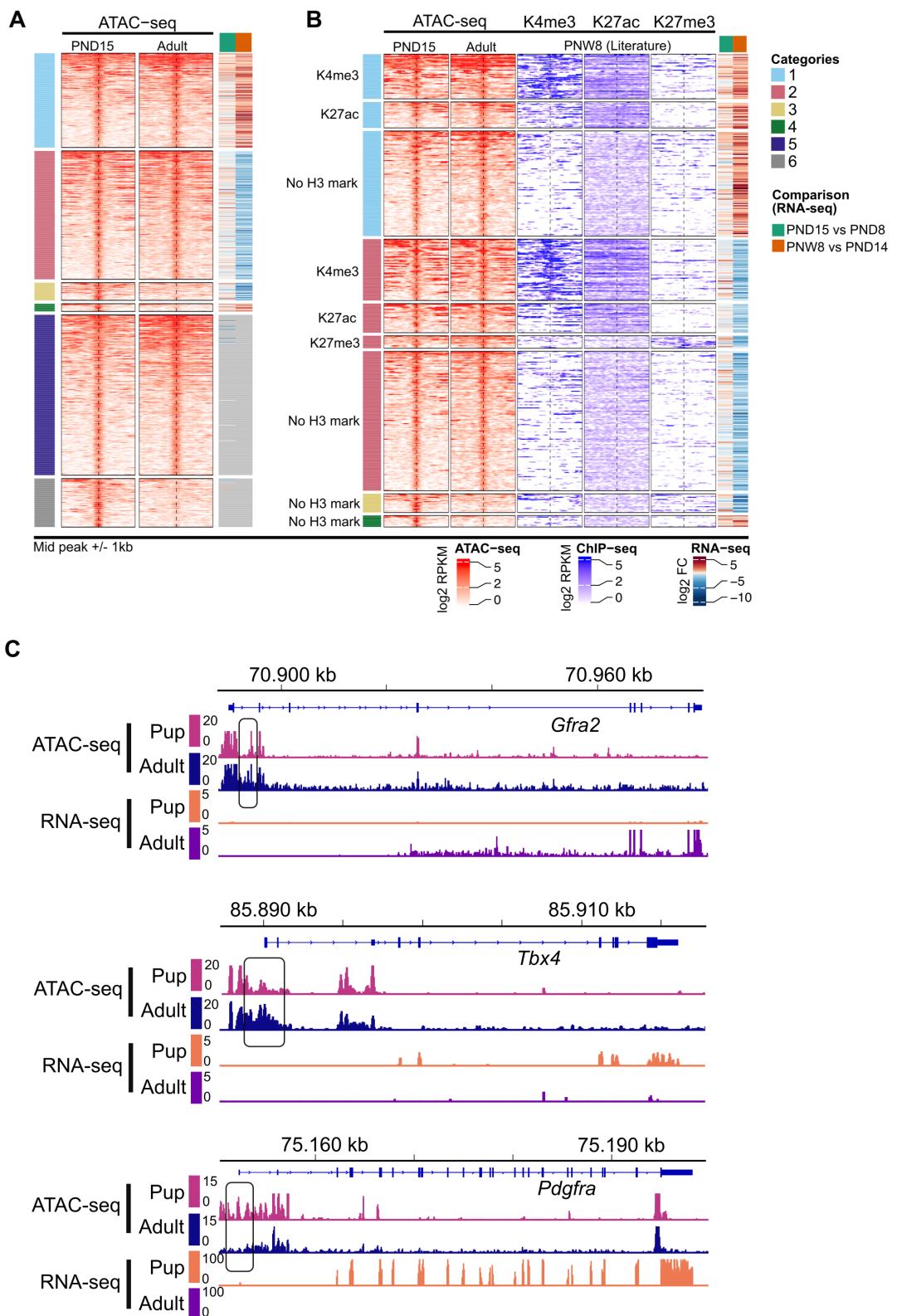


Figure 2-3 Chromatin accessibility and histone modifications at proximal regions of genes dynamically expressed between adult and PND15 spermatogonial cells.

(A) Heatmap of chromatin accessibility differences deduced by Omni-ATAC between PND15 and adult spermatogonial cells. Differentially accessible regions were grouped based on the correlation with the

expression of the nearest gene between PND14 and adult spermatogonia in 6 distinct categories: proximal regions of increased chromatin accessibility and increased gene expression (Category 1, n = 171), increased chromatin accessibility and decreased gene expression (Category 2, n = 233), decreased chromatin accessibility and decreased gene expression (Category 3, n = 32) and decreased chromatin accessibility and increased gene expression (Category 4, n = 14). Proximal inactive regions were defined as regions of increased accessibility (Category 5, n = 291) or decreased accessibility (Category 6, n = 88) for which the nearest gene expression was not detected from RNA-seq;

(B) Heatmaps showing the overlap between Category 1-4 regions and genes, and literature ChIP-seq data in PNW8 spermatogonia for H3K4me3, H3K27ac and H3K27me3. For each of the Category 1-4 the following sub-categorization was applied: regions that are enriched for H3K4me3 (with or w/o H3K27ac and/or H3K27me3), regions that are enriched for H3K27ac (and lack both H3K4me3 and H3K27me3) and regions that are enriched for H3K27me3 (and lack both H3K4me3 and H3K27ac);

(A, B) Each line represents a peak region and the regions are ordered within a category by the ATAC-seq signal. Mid-x-axis corresponds to the middle of a peak region and is extended to +/- 1 kb. The colour-key of the ATAC-seq heatmap represents the ATAC-seq signal in Log₂ Reads Per Kilobase per Million (RPKM) reads sequenced. For RNA-seq, log₂ FC is shown from PND15 vs PND8 and PNW8 vs PND14 comparisons;

(C) Genomic snapshots from the Integrative Genomics Viewer (IGV, Broad Institute) of exemplary genes from Category 1 (*Gfra2*), Category 2 (*Hmx1*) and Category 3 (*Pdgfra*) showing relative abundance of transcripts from RNA-seq and chromatin accessibility from ATAC-seq. RNA-seq data corresponds to literature PND14 and adult (PNW8) spermatogonial cells and ATAC-seq data to PND15 and adult (PNW20) spermatogonial cells, respectively.

2.3.4 Differentially accessible chromatin regions are marked by binding sites for distinct families of transcription factors (TFs)

TFs are important in establishing and maintaining distinct transcriptional signatures across the developmental trajectory of a cell population (Fushan et al., 2015; Shavlakadze et al., 2019). TF predominantly bind regions of open chromatin throughout the genome and allow for a dynamic regulation of gene expression (Klemm et al., 2019). To investigate if the regions of open chromatin with differential accessibility between PND15 and adult spermatogonia are enriched in TF binding motifs, we performed motif enrichment analysis using the Hypergeometric Optimization of Motif EnRichment (HOMER) tool (Heinz et al., 2010).

In regions with increased chromatin accessibility, we identified 41 enriched TF motifs ($q\text{-value} \leq 0.05$) (Figure 2-4A and Table S4). The top candidate list ($q\text{-value} \leq 0.0001$) was dominated by members of the Fos/Jun family (FOS, FOSB, FOSL1 and FOSL2, JUN, JUNB and JUND) (Figure 2-4B). Notably, at mRNA level, some of the TFs displayed age-specific differences ($\text{Log}_2\text{CPM} \geq 1$ and $\text{abs Log}_2\text{FC} \geq 1$): *Fos*, *Junb* and *Jund* were downregulated in adult spermatogonial cells (Figure 2-4C). JUN, FOS and CREB are part of the AP-1 (activating protein-1) superfamily, and play an important role in regulating cell proliferation and death, by mediating the senescence-associated chromatin and transcriptional landscape (Martínez-Zamudio et al., 2020; Shaulian and Karin, 2002). JUND and c-FOS specifically promote the proliferative potential of spermatogonial stem cells (He et al., 2008; Wang et al., 2018). USF1 and POU3F1, 2 important factors in the maintenance of the spermatogonial stem cell pool displayed enriched motifs in the regions of increased chromatin accessibility and a decrease in expression (Figure 2-4A and C). POU3F1 is a GDNF-regulated TF, which has been shown to play an important role in promoting spermatogonial cell self-renewal capacity (Niu et al., 2011; Wu et al., 2010). Our analysis also revealed enriched binding sites for retinoic acid receptors such as RXR α and RAR α (Figure 2-4A). Recent reports have revealed that *Rxra* and *Rara* utilization in spermatogonial cells is vastly dependent on the niche microenvironment (Lord et al., 2018).

To check if some of TF binding motifs are preferentially enriched in certain genomic locations, we performed motif enrichment analysis for more accessible chromatin regions situated in gene bodies, intergenic regions and in regions +/- 1kb from TSS. We identified several TF motifs specifically enriched in intergenic regions, specifically

members of the ubiquitously expressed NF-Y complex, NF-YA, NF-YB and NF-YC (Figure 2-4D). In mESCs, NF-Y TFs facilitate a permissive chromatin conformation, and play an important role in the expression of core ESC pluripotency genes (Oldfield et al., 2014). Furthermore, NF-YA/B motif enrichment has also been found in regions of open chromatin in human spermatogonial cells (Guo et al., 2017).

Less accessible chromatin regions in adult spermatogonial cells also displayed a high number of enriched TF binding motifs (Figure 2-4A and Table S4). Notably, almost all of these TF motifs were uniquely enriched in the regions of decreased chromatin accessibility and predominantly associated with developmental functions. Top hits included members of the FOX family (FOXO1, FOXO3, FOXP2, FOXK1, FOXA2) and members of the ETS and ETS-related TFs (ETS1, GABPA, ETV4, ELF1, ELF3) (Figure 2-4B). Expression levels of most of these TFs were decreased in adult spermatogonial cells (Figure 2-4A). FOXO1 is also a pivotal regulator of the self-renewal and differentiation of spermatogonial stem cells, via the PI3K-Akt signalling pathway (Chan et al., 2014; Goertz et al., 2011). The roles of ETS-related TFs in spermatogonial cells have yet to be clarified, however recently published data found high expression of *Etv4* in the stem-cell enriched fraction of the spermatogonial population, particularly during the spermatogonial stem cell pool establishment, immediately after birth (Cheng et al., 2020; Law et al., 2019). Motif enrichment analysis on the regions with decreased chromatin accessibility situated in gene body and intergenic regions revealed that TFs important in numerous developmental processes (FOXC1, FOXJ2, FOXM1, LHX6) were specifically enriched in intergenic regions of decreased chromatin accessibility (Figure 2-4D). This is consistent with the enrichment in developmental pathways that we found at the regions of decreased chromatin accessibility in adult spermatogonia (Figure 2-1C). Taken together, our TF motif analyses reveal that a shifting repertoire of TFs accompanies the chromatin reorganization taking place from early postnatal to adult spermatogonia.

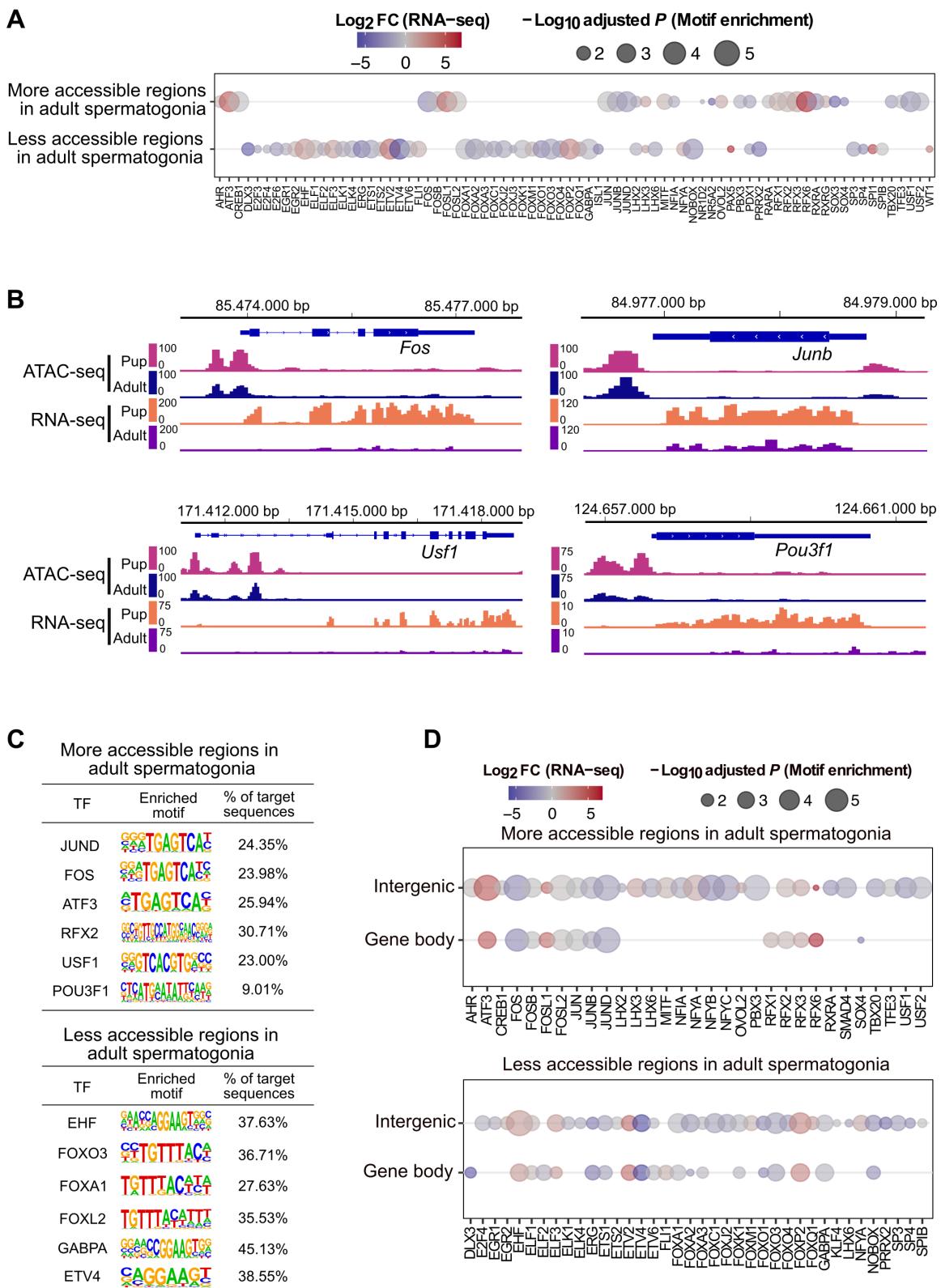


Figure 2-4 Transcription factor dynamics at differentially accessible regions as predicted by motif enrichment in adult spermatogonial cells.

(A) Dot plot of all transcription factor motifs enriched in the regions of decreased and increased accessibility between PND15 and adult spermatogonia;

- (B) Genomic snapshots from the Integrative Genomics Viewer (IGV, Broad Institute) of mRNA expression levels of representative enriched TFs in the regions of increased chromatin accessibility. RNA-seq data corresponds to literature PND14 and adult (PNW8) spermatogonial cells and ATAC-seq data to PND15 and adult (PNW20) spermatogonial cells, respectively;
- (C) HOMER extracted consensus sequences for each transcription factor motif. Representative examples from the most enriched transcription factor families are depicted;
- (D) Dot plots of top transcription factor motifs enriched in differentially accessible chromatin regions situated in gene bodies and in intergenic areas of the genome;
- (A, D) Each dot corresponds to a motif. The differential gene expression of each transcription factor was extracted from the PND14 and adult spermatogonia RNA-seq from literature, and is shown as colour-coded Log₂ FC. The size of the dot indicates the HOMER motif enrichment adjusted *P* of each motif.

2.3.5 Open chromatin at transposable elements (TEs) undergoes significant remodelling in the transition from early postnatal to adult spermatogonia

Recent evidence suggests an important role for long terminal repeat (LTR) - type elements, specifically for ERVKs, in the transcriptional regulation of mRNAs and long non-coding RNAs (lncRNAs) during spermatogenesis (Davis et al., 2017; Sakashita et al., 2020). Accessibility analysis at LTRs in mitotic and meiotic germ cells, revealed a unique chromatin accessibility landscape in spermatogonial cells, compared to the rest of germ cells in the testis (Sakashita et al., 2020).

To explore potential differences in TEs regulation driven by postnatal age, we compared the accessibility of TEs in PND15 and adult spermatogonia. For this purpose, we quantified the ATAC-seq reads overlapping TEs defined by UCSC RepeatMasker, and performed differential accessibility analysis at the subtype level (see Methods section). Our results revealed that the transition from PND15 to adult stage is accompanied by significant chromatin accessibility differences at 135 TE subtypes (FDR ≤ 0.05 , abs Log₂ FC ≥ 0.5) (Figure 2-5A and B and Table S5). Although most of the differentially accessible TE subtypes displayed a decrease in chromatin accessibility between PND15 and adult spermatogonia (68.9%, 93/135) (Figure 2-5A), we also observed 42 TE subtypes which increased in accessibility in adult spermatogonia (Figure 2-5B). Of note, the increase in accessibility correlated with an increased expression in adult spermatogonial cells compared to early postnatal stage (Figure 2-5B). TE loci within the subtypes harbouring changes in chromatin accessibility were predominantly situated in intergenic and intronic regions (68% intergenic and 25% intronic), and only 6% were located in proximity of a gene (+/- 1kb from a TSS) (Figure 2-5C).

LTRs were the most abundant TEs with altered chromatin accessibility, specifically ERVK and ERV1 subtypes (Figure 2-5A and B). Exemplary ERVK subtypes harbouring less accessible chromatin included RLTR17, RLTR9A3, RLTR12B and RMER17B (Table S5). Enrichment of RLTR17 and RLTR9 repeats has been reported previously in mESCs, specifically at TFs important for pluripotency maintenance such as *Oct4* and *Nanog* (Fort et al., 2014). Interestingly, we identified the promoter region of the lncRNA *Lncenc1*, an important regulator of pluripotency in mESCs (Fort et al., 2014; Sun et al., 2018b), harbouring several LTR loci with decreased accessibility in adult spermatogonia. One of these LTR loci, RLTR17, overlapped the TSS of *Lncenc1*.

This decrease in accessibility correlated with a marked decrease in expression of *Lncenc1* in adult spermatogonia (Figure 2-5D). *Lncenc1* (also known as *Platr18*) is part of the pluripotency-associated transcript (*Platr*) family of lncRNAs which were recently identified as potential regulators of the pluripotency-associated genes *Oct4*, *Nanog* and *Zfp42* in mESCs (Bergmann et al., 2015; Dann et al., 2008; Wu et al., 2018). We also identified several other *Platr* genes, such as *Platr27* and *Platr14*, for which the TSS overlapped LTRs with reduced accessibility, RLTR17 and RLTR16B_MM, respectively (Figure 2-5D and Table S5). *Platr27* and *Platr14* also showed a decrease in mRNA expression in adult spermatogonia, while their expression was unchanged between PND8 and PND15 (Figure 2-5D and Table S5). The remaining LTR subtypes with decreased accessibility in adult spermatogonia belonged to the ERV1, ERVL and MaLR families (Figure 2-5A). Only very few other non-LTR TEs showed a decrease in chromatin accessibility, with 7 DNA element subtypes, 2 Satellite subtypes and 1 LINE subtype, respectively (Table S5).

Emerging evidence suggests an important contribution of TEs in providing tissue-specific substrates for TF binding (Fort et al., 2014; Sundaram and Wysocka, 2020; Sundaram et al., 2014). To investigate the regulatory potential of the less accessible LTR subtypes, we assessed the enrichment of TF motifs in these regions using HOMER (Table S6). To do so, we focused on the family level and grouped together all LTR subtypes coming from one family (EVK, ERV1, ERVL and ERVL-MaLR families). Among the less accessible LTR families, ERVKs showed the highest number of enriched TF motifs in adult spermatogonial cells. Top hits included TFs with known regulatory roles in cell proliferation and differentiation such as FOXL1 and FOXQ1, stem cell maintenance factors ELF1, EBF1 and THAP11 and TFs important in spermatogenesis PBX3, ZNF143 and NFYA/B (Figure 2-5E and Figure S 2-6A). ERVLs displayed motif enrichment for very few TFs, among which the previously undescribed ETV2, newly reported spermatogonial stem cell factor ZBTB7A and the testis-specific CTCF paralog CTCFL (Figure 2-5E) (Green et al., 2018).

Among the TE subtypes which increased in accessibility, members of the ERVK, ERVL and ERV1 families were predominant (57,1%, 24/42) (Figure 2-5B). Interestingly, we also found a considerable number of LINE L1 subtypes with increased chromatin accessibility in adult spermatogonial cells (Figure 2-5B). When parsing the data for more accessible loci within the L1 subtypes, we found several L1

loci situated less than +/- 5 kb from the TSS of numerous olfactory (*Olfr*) genes. Most of them were located in *Olfr* gene clusters on chromosome 2, 7 and 11 (Table S5). Furthermore, the increase in accessibility of the L1 loci correlated with an increase in mRNA expression of the nearby *Olfr* gene in adult spermatogonial cells (Figure 2-6A). Representative examples were *Olfr362* and *Olfr1307*, both situated in the *Olfr* gene cluster on Chr2 (Figure 2-6B). Interestingly, when visualizing the data in IGV, we also observed that the *Olfr* gene cluster on chromosome 2 exhibited a higher density of L1 loci compared to neighbouring regions (Figure S 2-6B).

Similar to before, we performed TF motif enrichment analysis at the family level by grouping together the more accessible TE subtypes coming from one family (Table S6). More accessible LINE L1s were highly enriched in TF motifs, particularly in multiple members of the ETS, E2F and FOX families (Figure 2-6C). The most significant motifs belonged to spermatogonial stem cell maintenance and stem cell potential regulators FOXO1 and ZEB1, as well as TFs which have been recently associated with active enhancers of the stem cell-enriched population of spermatogonia such as ZBTB17 and KLF5 (Figure 2-6C and Figure S 2-6A) (Cheng et al., 2020). More accessible ERV1s also displayed enrichment of several TF binding sites, including spermatogenesis-related TFs (PBX3, PRDM1, NFYA/B), hypoxia inducible HIF1A and cytokine regulators STAT5A/B, suggestive of different spermatogonial cell metabolic demands between early postnatal and adult stage (Figure 2-6C and Figure S 2-6A). Overall, we provide an extensive characterization of the chromatin accessibility landscape of TEs in PND15 and adult spermatogonia, reveal differences in accessibility and TF motif landscape at distinct subtypes of TEs between these 2 timepoints, and suggest potential gene programs that may be regulated by these changes.

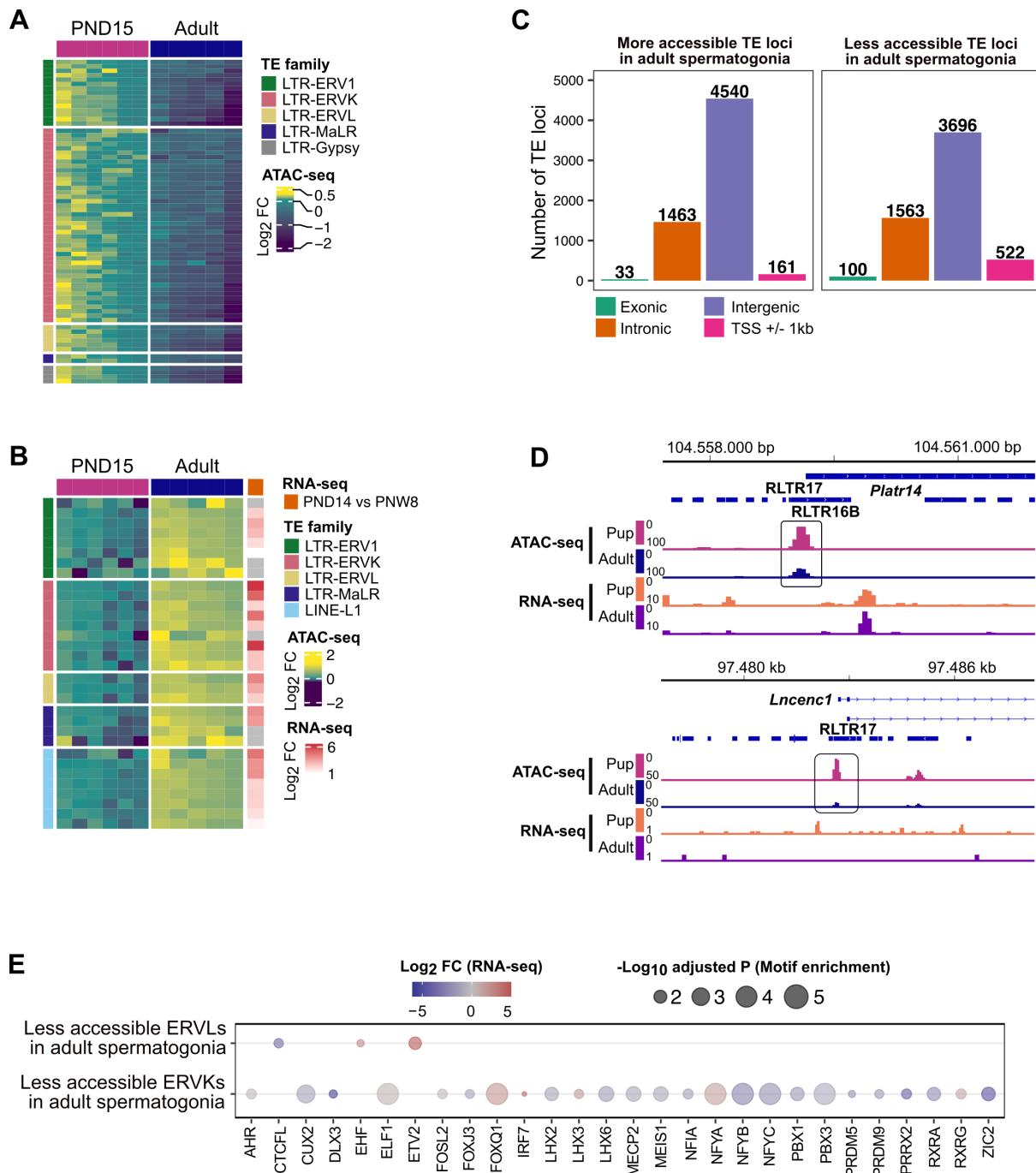


Figure 2-5 Differential chromatin accessibility at transposable elements (TEs) in adult spermatogonial cells compared to PND15

(A) Heatmap of the LTR and LINE subtypes with decreased accessibility between adult and PND15 spermatogonia extracted from the Omni-ATAC data (adjusted $P \leq 0.05$ and $\text{Log}_2 \text{ FC} \geq 0.5$);

(B) Heatmap of the LTR and LINE subtypes with increased accessibility between adult and PND15 spermatogonia extracted from the Omni-ATAC data (adjusted $P \leq 0.05$ and $\text{Log}_2 \text{ FC} \geq 0.5$). Expression changes of these subtypes between PND14 and adult spermatogonia RNA-seq from literature is represented as $\text{Log}_2 \text{ FC}$;

(A, B) For both accessibility heatmaps, Log₂ FC are shown with respect to the average of the PND15 samples. Samples are clustered using Ward's method. Subtypes are ordered by principal component analysis (PCA) method using seriation (R package);

(C) Bar plot illustrating the genomic distribution of differentially accessible TEs between adult and PND15 spermatogonial cells;

(D) Genomic snapshots from the Integrative Genomics Viewer (IGV, Broad Institute) of *Lncenc1* and *Platr14* showing LTRs from RepeatMasker and the average normalized RNA-seq and ATAC-seq coverage (RPKM). LTR loci were extracted using RepeatMasker. RNA-seq data corresponds to literature PND14 and adult (PNW8) spermatogonial cells and ATAC-seq data to PND15 and adult (PNW20) spermatogonial cells, respectively;

(E) Dot plots of top transcription factor motifs enriched in the less accessible ERVKs and ERVL subtypes. Each dot corresponds to a motif. The differential gene expression of each transcription factor was extracted from the PNW8 vs PND14 comparison from literature RNA-seq, and is shown as colour-coded Log₂ FC. The size of the dot indicates the HOMER motif enrichment adjusted *P* of each motif.

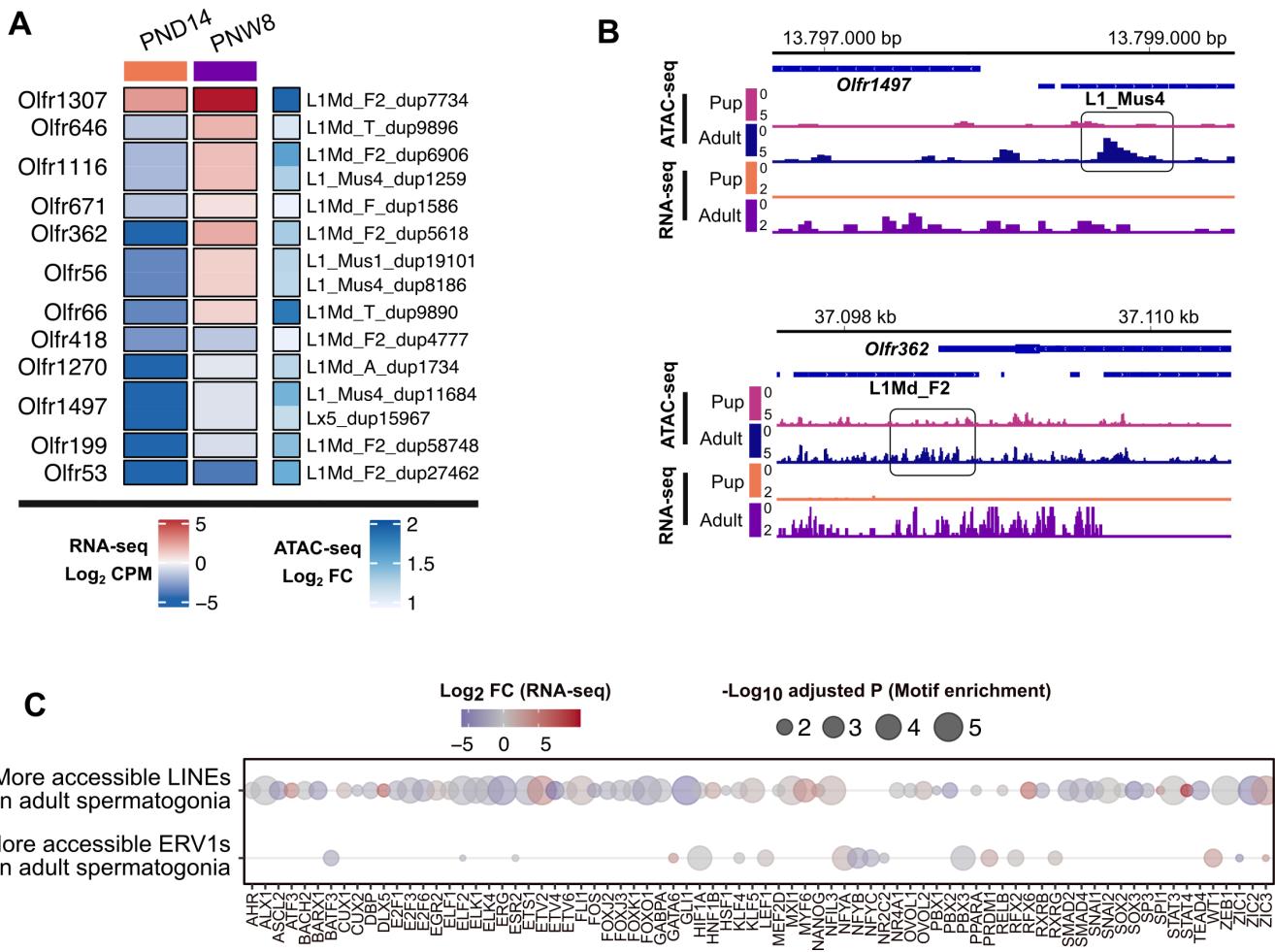


Figure 2-6 Increased accessibility at LINE L1 subtypes located near *Olfr* gene clusters

(A) Heatmap of the *Olfr* genes for which we identified an upregulated expression from PND14 to PNW8 timepoints and an increase in accessibility at a nearby L1 locus. RNA expression levels are expressed as Log₂ CPM at each timepoint. Accessibility changes at each of the corresponding L1 locus situated within +/- 5kbp from the gene are expressed as Log₂ FC calculated from the ATAC-seq analysis of the differentially accessible TEs between adult and PND15 spermatogonial cells;

(B) Genomic snapshots from the Integrative Genomics Viewer (IGV, Broad Institute) of exemplary genes *Olfr1497* and *Olfr362* showing relative abundance of transcripts from RNA-seq and chromatin accessibility from ATAC-seq. LINE loci were extracted using Repeat Masker. RNA-seq data corresponds to literature PND14 and adult (PNW8) spermatogonial cells and ATAC-seq data to PND15 and adult (PNW20) spermatogonial cells, respectively;

(C) Dot plots of top transcription factor motifs enriched in the more accessible L1 and ERV1 subtypes. Each dot corresponds to a motif. The differential gene expression of each transcription factor was extracted from the PNW8 vs PND14 comparison from literature RNA-seq, and is shown as colour-coded Log₂ FC. The size of the dot indicates the HOMER motif enrichment adjusted *P* of each motif.

2.4 Discussion

As initiators of the spermatogenic cascade, spermatogonial cells are essential in germ cell proliferation and differentiation throughout postnatal life. Although recent studies employing bulk and scRNA-seq have revealed distinct transcriptional signatures of spermatogonial cells across postnatal life, very few have focused on describing the underlying landscape of open chromatin, and the extent to which it can contribute to the gene expression dynamics (Grive et al., 2019; Hammoud et al., 2015; Hermann et al., 2018).

Our ATAC-seq revealed a reorganization of open chromatin in adult spermatogonia compared to the PND15 cell population. We found that the regions of differential accessibility were associated with distinct gene pathways, with morphogenesis and developmental pathways enriched in regions of decreased chromatin accessibility, while regions of increased chromatin openness in adult spermatogonia were enriched for DNA repair pathways, stem cell maintenance, RNA processing and protein metabolism.

Notably, distinct families of TFs were enriched at the regions of increased chromatin accessibility compared to the regions which became less accessible in adult spermatogonial cells. As such, TFs known to regulate spermatogonial cell proliferation, pluripotency potential and niche-dependent signalling such as JUND and c-FOS, POU3F1, and retinoic acid receptors, displayed enriched binding sites in the regions of increased chromatin accessibility. In contrast, FOX and ETS TF motifs, known regulators of developmental pathways, mainly mapped to regions which decreased in accessibility in adult spermatogonia. For some of the enriched TF motifs, we observed a preference for certain genomic locations: NF-YA and B binding sites exhibited enrichment specifically in intergenic regions of more accessible chromatin, which interestingly, were also associated with spermatogenesis-related pathways. NF-YA/B also localize in intergenic regions of open chromatin in human spermatogonial cells (Guo et al., 2017), prompting additional investigation of their roles in regulating spermatogonial cell programs, with potential consequences for sperm formation. Altogether, our pathway and TF analyses reveal potential regulatory elements for the age-specific gene expression of spermatogonial cells.

Our comparison of the gene expression changes from PND8 to PND15 spermatogonial cells, together with the reanalysis of literature RNA-seq data from PND14 and adult THY1+ spermatogonia, confirmed the dynamic transcriptome associated with postnatal spermatogonial cell states (Grive et al., 2019; Hammoud et al., 2015). The age-dependent upregulation of pathways associated with signal transduction, cellular communication and cytokine signalling, supports previous findings suggesting that, as the testis matures and the somatic niche develops, spermatogonial cells rely more on paracrine signalling, and undergo vast changes in gene expression. To obtain a comprehensive profile of the chromatin landscape and the transcriptome differences between early postnatal and adult spermatogonial cells, we integrated the chromatin accessibility and gene expression profiles with known histone H3 modifications and global DNAmc patterns of THY1+ spermatogonial cells from literature (Hammoud et al., 2014, 2015). This allowed us to identify several distinct categories of differentially accessible chromatin regions for which the nearest gene was dynamically expressed between early postnatal and adult stage.

In the category of upregulated genes with increased nearby chromatin accessibility, we identified several factors associated with redox processes, mitochondria function and cell proliferation. Interestingly, a similar number of genes displayed an increase in chromatin accessibility and a downregulated expression in adult spermatogonia, suggesting that active repression is taking place at these genes (Starks et al., 2019). This category comprised factors important for cell cycle and RNA processing, as well as developmental genes. For some of the developmental genes, more accessible chromatin was also marked by a bivalent H3K4me3/H3K27me3, indicative of a poised state. Notably, previous findings in THY1+ adult spermatogonial cells and in sperm, also revealed a poised state at promoters of developmental genes (Erkek et al., 2013; Hammoud et al., 2014; Jung et al., 2017). Therefore, our findings suggest that open chromatin reorganization may play a role in regulating the poised status of certain developmental genes in the germline. We also identified a category of regions for which the decrease in chromatin accessibility correlated with a decreased expression, which also included developmental factors. The stable methylation patterns we detected at the differentially accessible chromatin regions, suggest a minimal impact for DNAmc in regulating gene expression dynamics of spermatogonial cells across postnatal age.

Lastly, by investigating chromatin accessibility specifically at TEs, we revealed that distinct TE subtypes undergo changes in chromatin accessibility between PND15 and adult spermatogonia. ERVK and ERV1 subtypes were the largest categories of TEs to become less accessible in adult spermatogonia, whilst LINE L1 subtypes displayed an increase in chromatin accessibility. Although the majority of these TEs resided in intergenic and intronic regions, we identified specific loci belonging to the differentially accessible ERVK and LINE L1 subtypes, which localized nearby TSS of distinct gene families.

RLTR17, one of the LTR subtypes with decreased chromatin accessibility in adult spermatogonial cells, overlapped the TSS of several downregulated long-non coding RNAs from the *Platr* family. *Platr* genes, including the ones identified in our study, *Lncenc1* and *Platr14*, are LTR-associated long non-coding RNAs important for pluripotency potential of mouse and human embryonic stem cells (Bergmann et al., 2015; Fort et al., 2014). In mouse embryonic stem cells, RLTR17 is highly expressed and is enriched in open chromatin regions, and provides binding sites for pluripotency factors OCT4 and NANOG (Fort et al., 2014). On the basis of these findings, we suggest that RLTR17 chromatin organization may play a significant role in regulating pluripotency programs between early postnatal and adult spermatogonial cells.

In contrast to the decreased accessibility of LTRs, LINE L1 subtypes displayed an increase in chromatin accessibility in adult spermatogonial cells. Some of these L1 loci were situated in the vicinity of *Oifr* genes with upregulated mRNA expression in adult spermatogonia. Recent findings in mouse and human embryonic stem cells have suggested a non-random genomic localization for L1 elements, specifically at genes which encode proteins with specialized functions (Lu et al., 2020). Among these, the *Oifr* gene family was the most enriched in L1 elements (Lu et al., 2020). Although their role in spermatogonial cells is currently not established, *Oifr* proteins have been implicated in the swimming behaviour of sperm (Fukuda and Touhara, 2005; Vanderhaeghen et al., 1997). Given the dynamic chromatin profile of LINE L1 elements that we found at *Oifr* genes between early postnatal and adult spermatogonia, we speculate that *Oifr* genes could play additional roles in spermatogenesis, other than in sperm physiology. In addition, we found a high number of enriched TF motifs present at the differentially accessible ERVKs and LINE L1 families. These findings reveal that chromatin accessibility at TEs is reorganized in

spermatogonia cells during the transition from developing to adult stages, and may contribute to distinct gene regulatory networks (Sundaram and Wysocka, 2020; Sundaram et al., 2014).

One limitation of our study is the incomplete purification achieved using FACS, which doesn't fully remove other testis cell types from our cell preparations. Therefore, we cannot entirely exclude the influence of contaminating cells on some of the transcriptome and chromatin accessibility data interpretation. Secondly, differences can also arise from the literature datasets which involve similar, but not identically enriched populations of spermatogonial cells. Nevertheless, by comparing open chromatin landscape between developing and adult spermatogonial cells, our results reveal for the first time that there is an age-dependent dynamic reorganization of chromatin accessibility in spermatogonial cells. By integrating this newly generated data with gene expression profiles and known histone modifications, we provide novel insight into the chromatin - transcriptome dynamics of mouse spermatogonial cells between developing and adult stages, and compile an information-rich resource for further germline studies.

2.5 Methods

2.5.1 Mouse husbandry

Male C57Bl/6J mice were purchased from Janvier Laboratories (France) and bred in-house to generate male mice used for experiments. All animals were kept on a reversed 12-h light/12-h dark cycle in a temperature- and humidity-controlled facility, with food (M/R Haltung Extrudat, Provimi Kliba SA, Switzerland) and water provided ad libitum. Cages were changed once weekly. Animals from 2 independent breedings were used for the experiments.

2.5.2 Germ cells isolation

Germ cells were isolated from male mice at postnatal day (PND) 8 or 15 for RNA-seq and ATAC-seq experiments, and adults at 20 weeks of age (PNW20) for ATAC-seq. Testicular single-cell suspensions were prepared as previously described with slight modifications (Kubota et al., 2004b). For preparations using PND8 and PND15 pups, testes from 2 animals were pooled for each sample. Pup testes were collected in sterile

HBSS on ice. Tunica albuginea was gently removed from each testis, making sure to keep the seminiferous tubules as intact as possible. Tubules were enzymatically digested in 0.25% trypsin-EDTA (ThermoFisher Scientific) and 7mg/ml DNase I (Sigma-Aldrich) solution for 5 min at 37°C. The suspension was vigorously pipetted up and down 10 times and incubated again for 3 min at 37°C. The digestion was stopped by adding 10% foetal bovine serum (ThermoFisher Scientific) and the cells were passed through a 20µm-pore-size cell strainer (Miltenyi Biotec) and pelleted by centrifugation at 600g for 7 min at 4°C. Cells were resuspended in PBS-S (PBS with 1% PBS, 10 mM HEPES, 1 mM pyruvate, 1mg/ml glucose, 50 units/ml penicillin and 50 µg/ml streptomycin) and used for sorting. For preparations from adult testis, one adult male was used for each sample. The tunica was removed and seminiferous tubules were digested in 2 steps. The first consisted in an incubation in 1mg/ml collagenase type IV (Sigma-Aldrich) for 5 min at 37°C and vigorous swirling until the tubules were completely separated. Then tubules were placed on ice for 5 min to sediment, the supernatant removed and washed with HBSS. Washing/sedimentation steps were repeated 3 times and were necessary to remove interstitial cells. After the last washing step, sedimented tubule fragments were digested again with 0.25% trypsin-EDTA and 7mg/ml DNase I solution, and the digestion was stopped by adding 10% FBS. The resulting single-cell suspension was filtered through a 20µm strainer (Corning Life Sciences) and washed with HBSS. After centrifugation at 600g for 7 min at 4°C, the cells were resuspended in PBS-S, layered on a 30% Percoll solution (Sigma-Aldrich) and centrifuged at 600g for 8 min at 4°C without braking. The top 2 layers (HBSS and Percoll) were removed and the cell pellets resuspended in PBS-S and used for sorting.

2.5.3 Spermatogonial cell enrichment by FACS

For pup testis, dissociated cells were stained with BV421-conjugated anti-β2M, biotin-conjugated anti-THY1 (53-2.1), and PE-conjugated anti-αv-integrin (RMV-7) antibodies. THY1 was detected by staining with Alexa Fluor 488-Sav. For adult testes, cells were stained with anti-α6-integrin (CD49f; GoH3), BV421-conjugated anti-β2 microglobulin (β2M; S19.8), and R-phycoerythrin (PE)-conjugated anti-THY1 (CD90.2; 30H-12) antibodies. α6-Integrin was detected by Alexa Fluor 488-SAv after staining with biotin-conjugated rat anti-mouse IgG1/2a (G28-5) antibody. Prior to FACS, 1

$\mu\text{g}/\text{ml}$ propidium iodide (Sigma) was added to the cell suspensions to discriminate dead cells. All antibody incubations were performed in PBS-S for at least 30 min at 4°C followed by washing in PBS-S excess. Antibodies were obtained from BD Biosciences (San Jose, United States) unless otherwise stated. Cell sorting was performed on a FACS Aria III 5L, at 4°C and using an 85 μm nozzle, at the Cytometry Facility of University of Zurich. For RNA-seq on PND8 and PND15 spermatogonia, cells were collected in 1.5 ml Eppendorf tubes in 500 μL PBS-S, immediately pelleted by centrifugation and snap frozen in liquid N₂. Cells pellets were stored at -80°C until RNA extraction. For OmniATAC on PND15 spermatogonia, 25'000 cells were collected in a separate tube, pelleted by centrifugation and immediately processed using the OmniATAC library preparation protocol (Corces et al., 2017). For OmniATAC on adult spermatogonia, 5000 cells from each animal were collected in a separate tube and further processed using the same protocol.

2.5.4 Immunocytochemistry

The protocol used for assessing spermatogonial cell enrichment after sorting was kindly provided by the Oatley Lab at Washington State University, Pullman, USA (Yang et al., 2013). Briefly, 30,000-50,000 cells were adhered to poly-L-Lysine coated coverslips (Corning Life Sciences) in 24-well plates for 1 h. Cells were fixed in freshly prepared 4% PFA for 10 min at room temperature then washed in PBS with 0.1% Triton X-100 (PBS-T). Non-specific antibody binding was blocked by incubation with 10% normal goat serum for 1 h at room temperature. Cells were incubated overnight at 4°C with mouse anti-PLZF (0.2 $\mu\text{g}/\text{ml}$, Active Motif, clone 2A9) primary antibody. Alexa488 goat anti-mouse IgG (1 $\mu\text{g}/\text{mL}$, ThermoFisher Scientific) was used for secondary labelling at 4°C for 1 h. Coverslips were washed 3x and mounted onto glass slides with VectaShield mounting medium containing DAPI (Vector Laboratories) and examined by fluorescence microscopy. Stem cell enrichment was determined by counting PLZF⁺ cells in 10 random fields of view from each coverslip and dividing by the total number of cells present in the field of view (DAPI-stained nuclei).

2.5.5 RNA extraction and RNA-seq library preparation

For RNA-seq on PND8 and PND15 spermatogonial cells, total RNA was extracted from sorted cells using the AllPrep RNA/DNA Micro kit (Qiagen). RNA quality was

assessed using a Bioanalyzer 2100 (Agilent Technologies). Samples were quantified using Qubit RNA HS Assay (ThermoFisher Scientific). 10 ng of total RNA from each sample were used to prepare total long RNA sequencing libraries using the SMARTer® Stranded Total RNA-Seq Kit v2 - Pico Input Mammalian (Takara Bio USA, Inc.) at the Functional Genomics Centre (FGC) Zurich, according to the manufacturer's instructions. The number of samples sequenced at each timepoint was 8 for PND8 and 9 for PND15 spermatogonia (each sample was representative of 4 testes from 2 pups).

2.5.6 Omni-ATAC library preparation

Chromatin accessibility was profiled from PND15 and adult spermatogonial cells. The libraries were prepared according to the Omni-ATAC protocol, starting from 25 000 PND15 and 5000 adult sorted spermatogonia, respectively (Corces et al., 2017). Briefly, sorted cells were lysed in cold lysis buffer (10 mM Tris-HCl pH 7.4, 10 mM NaCl, 3 mM MgCl₂, 0.1% NP40, 0.1% Tween-20, and 0.01% digitonin) and the nuclei were pelleted and transposed using the Nextera Tn5 (Illumina) for 30 min at 37°C in a thermomixer with shaking at 1000 rpm. The transposed fragments were purified using the MinElute Reaction Cleanup Kit (Qiagen). Following purification, libraries were generated by PCR amplification using the NEBNext High-Fidelity 2X PCR Master Mix (New England Biolabs), and purified using Agencourt AMPure XP magnetic beads (Beckman Coulter) in order to remove primer dimers (78bp) and large fragments of 1000-10,000 bp in length. Library quality was assessed on an Agilent High Sensitivity DNA chip using the Bioanalyzer 2100 (Agilent Technologies). 6 samples were sequenced from PND15 and 5 from adult spermatogonial cells.

2.5.7 High-throughput sequencing data analysis

Data availability: the datasets used in this study are available from the following GEO accessions: **GSE_____**, GSE49621, GSE49622, GSE62355, and GSE49623. An overview of the datasets included in the study is shown in the following table:

Source	*Seq	Stages (n)
GSE_____	RNA-seq	PND8 (8), PND15 (9)
GSM1525703	RNA-seq	PND14 (1)

GSM1415671	RNA-seq	PNW8 (1)
GSE_____	ATAC-seq	PND15 (6), PNW20 (5)
GSM1202705	ChIP-seq (H3K4me3)	PNW8 (1)
GSM1202708	ChIP-seq (H3K27me3)	PNW8 (1)
GSM1202713	ChIP-seq (H3K27ac)	PNW8 (1)
GSM1202723	ChIP-seq (Input)	PNW8 (1)
GSE49623	BS-seq	PND7 (1), PND14 (1), PNW8 (7)

2.5.7.1 RNA-seq data analysis

Quality control and alignment: Single-end (SE) sequencing was performed on the PND8, and PND15 spermatogonia samples on the Illumina HiSeq4000 at the FGC Zurich. PND8 raw data (FASTQ files) was merged from two individual runs. For the analysis of literature RNA-seq data, FASTQ files of PND14 and PNW8 samples were obtained using fastq-dump (version 2.10.8). Quality assessment of the FASTQ files was performed using FastQC (Andrews et al., 2012) (version 0.11.8). TrimGalore (Krueger, 2015) (version 0.6.2) was used to trim adapters and low-quality ends from reads with Phred score less than 30 (-q 30), and for discarding trimmed reads shorter than 30 bp (--length 30). Trimmed reads were then pseudo-aligned using Salmon (Patro et al., 2017) (version 0.9.1) with automatic detection of the library type (-l A), correcting for sequence-specific bias (--seqBias) and correcting for fragment GC bias correction (--gcBias) on a transcript index prepared for the Mouse genome (GRCm38) from GENCODE (Harrow et al., 2012a) (version M18) with additional piRNA precursors, and transposable elements (concatenated by family) from Repeat Masker as in (Gapp et al., 2018).

Downstream analysis: the downstream analysis was performed using R (R Core Team, [2020](#)) (version 3.6.2), using packages from The Comprehensive R Archive Network (CRAN) (<https://cran.r-project.org>) and Bioconductor (Huber et al., 2015). Pre-filtering of genes was performed using the filterByExpr function from edgeR (Robinson et al., 2009) (version 3.28.1) with a design matrix and requiring at least 15

counts (min.counts = 15). Normalization factors were obtained using TMM normalization (Robinson and Oshlack, 2010) from the edgeR package and differential gene expression (DGE) analysis was performed using the limma-voom (Law et al., 2014) pipeline from limma (Ritchie et al., 2015) (version 3.42.2). The Log₂ fold change between PNW8 and PND14 spermatogonial samples was calculated by subtracting Log₂ normalized expression values of PND14 from PNW8 samples. Gene ontology (GO) analysis was performed on expressed genes with fGSEA (version 1.15.2), using fGSEAMultilevel function on sets with 10 to 1000 annotated genes (minSize = 10, maxSize = 1000), and p-values boundary of 1E-100 (eps = 1e-100) (Korotkevich et al., 2016). For the PND15 vs PND8 comparison, genes were pre-ranked using their t-statistic; for the PNW8 vs PND14 comparison the Log₂ fold change was used due to the lack of multiple replicates for each timepoint. REVIGO was used to summarize the GO terms obtained following fGSEA (Supek et al., 2011).

2.5.7.2 Omni-ATAC data analysis

Quality control, alignment, and peak calling: Paired-end (PE) sequencing was performed on PND15 and adult spermatogonial cell samples on the Illumina HiSeq2500 platform at the FGC Zurich. FASTQ files were assessed for quality using FastQC (Andrews et al., 2012) (version 0.11.8). Quality control (QC) was performed using TrimGalore (Krueger, 2015) (version 0.6.2) in PE mode (--paired), trimming adapters, low-quality ends (-q 30) and discarding reads that become shorter than 30 bp after trimming (--length 30). Alignment on the GRCm38 genome was performed using Bowtie2 (Langmead and Salzberg, 2012) (version 2.3.5) with the following parameters: fragments up to 2 kb were allowed to align (-X 2000), entire read alignment (--end-to-end), suppressing unpaired alignments for paired reads (--no-mixed), suppressing discordant alignments for paired reads (--no-discordant) and minimum acceptable alignment score with respect to the read length (--score-min L,-0.4,-0.4). Using alignmentSieve (version 3.3.1) from deepTools (Ramirez et al., 2016) (version 3.4.3), aligned data (BAM files) were adjusted for the read start sites to represent the centre of the transposon cutting event (--ATACshift), and filtered for reads with a high mapping quality (--minMappingQuality 30). Reads mapping to the mitochondrial chromosome and ENCODE blacklisted regions (ENCODE accession ENCF547MET), were filtered out. To call nucleosome-free regions, all aligned files were merged within groups (PND15 and adult), sorted, and indexed using SAMtools

(Li et al., 2009) (version 0.1.19), and nucleosome-free fragments (NFFs) were obtained by selecting alignments with a template length between 40 and 140 inclusively. Peak calling (identifying areas in a genome that have been enriched for transcription factors) on the NFFs was performed using MACS2 (Zhang et al., 2008) (version 2.2.7.1) with mouse genome size (-g 2744254612) and PE BAM file format (-f BAMPE).

Differential accessibility analysis: The downstream analysis was performed in R (version 3.6.2), using packages from CRAN (<https://cran.r-project.org>) and Bioconductor (Huber et al., 2015). The peaks were annotated based on overlap with GENCODE (Harrow et al., 2012b) (version M18) transcript and/or the distance to the nearest transcription start site (available at the following link: https://github.com/mansuylab/SC_postnatal_adult/bin/annoPeaks.R). The number of extended reads overlapping in the peak regions was calculated using the csaw package (Lun and Smyth, 2015) (version 1.20.0). Peak regions which did not have at least 15 reads in at least 40% of the samples were filtered out. Normalization factors were obtained on the filtered peak regions using the TMM normalization method (Robinson and Oshlack, 2010) and differential analysis on the peaks (adults vs PND15) was performed using the Genewise Negative Binomial Generalized Linear Models with Quasi-likelihood (glmQLFit) Tests from the edgeR package (Robinson et al., 2009) (version 3.28.1). Peak regions which had an absolute Log₂ fold change ≥ 1 and an FDR ≤ 0.05 were categorized as differentially accessible regions (DARs). GO analysis was performed on DARs with the rGREAT package (Zuguang, 2020) (version 1.18.0), which is a wrapper around the GREAT tool (McLean et al., 2010) (version 4.0). Transcription factor motif enrichment analysis was performed using the marge package (Amezquita, 2018) (version 0.0.4.9999), which is a wrapper around the Homer tool (Heinz et al., 2010) (version 4.11.1).

Differential accessibility analysis at transposable elements: TE gene transfer format (GTF) file was obtained from http://labshare.cshl.edu/shares/mhammelllab/www-data/TEtranscripts/TE_GTF/mm10_rmsk_TE.gtf.gz on 03.02.2020. The GTF file provides hierarchical information about TEs: **Class** (level 1, e.g. LTR), **Family** (level 2, e.g. LTR → L1), **Subtype** (level 3, e.g. LTR → L1 → L1_Rod), and **Locus** (level 4, e.g. LTR → L1 → L1_Rod → L1_Rod_dup1). TE loci were annotated based on overlap with GENCODE (version M18) as described above for ATAC-seq peaks. Filtered BAM

files (without reads mapping to blacklisted or mitochondrial regions) were used for analysing TEs. Mapped reads were assigned to TEs using featureCounts from the R package Rsubread (Liao et al., 2019b) (version 2.0.1) and were summarized to Subtypes (level 3), allowing for multi-overlap with fractional counts, while ignoring duplicates. The number of extended reads overlapping at the TE loci were obtained using the csaaw package (Lun and Smyth, 2015) (version 1.20.0). Subtypes which did not have at least 15 reads, and loci which did not have at least 5 reads in at least 40% of the samples, were filtered out. Normalization and differential accessibility analysis were performed as described above. Subtypes which had an absolute Log_2 fold change ≥ 0.5 and an FDR ≤ 0.05 were categorized as differentially accessible subtypes and the loci with an absolute Log_2 fold change ≥ 1 and an FDR ≤ 0.05 were categorized as differentially accessible loci. For further downstream data analysis, only the differentially accessible loci of differentially accessible subtypes were considered. GO and motif enrichment analysis were performed as described above.

2.5.7.3 ChIP-seq data analysis

Quality control, alignment, and peak calling: ChIP-Seq SE data for PNDW8 (adults) were obtained from GEO accession GSE49621 (Hammoud et al., 2014). FASTQ files were obtained using fastq-dump (version 2.10.8), and different runs were merged. The FASTQ files were assessed for quality using FastQC (Andrews et al., 2012) (version 0.11.8). Quality control (QC) was performed using TrimGalore (Andrews et al., 2012) (version 0.6.0), trimming adapters, low-quality ends (-q 30) and discarding trimmed reads shorter than 30 bp (--length 30). Alignment to the GRCh38 genome was performed using Bowtie2 (Langmead and Salzberg, 2012) (version 2.3.5). Reads with more than 3 mismatches were removed from the aligned data, as suggested in (Royo et al., 2016), and reads with low mapping quality (--minMappingQuality 30) or mapping to the mitochondrial chromosome or aforementioned blacklisted regions were filtered out. Peak calling was performed using MACS2 (Zhang et al., 2008) (version 2.2.7.1) with mouse genome size (-g 2744254612) and SE BAM file format (-f BAM).

2.5.7.4 Bisulphite sequencing (BS) data analysis

Quality control and alignment: BS paired-end data for PND7, PND14, and PNW8 (adults) were obtained from GEO accession GSE49623 (Hammoud et al., 2015). FASTQ files were obtained using fastq-dump (version 2.10.8), and different runs were

merged. FASTQ files were assessed for quality using FastQC (Andrews et al., 2012) (version 0.11.8). QC was performed using TrimGalore (Krueger, 2015) (version 0.6.4_dev) in PE mode (--paired), trimming adapters, low-quality ends (-q 30) and discarding trimmed reads shorter than 30 bp (--length 30). Alignment of the QC data was performed using Bismark (Krueger and Andrews, 2011) (version 0.22.3) on a GRCm38 index built using bismark_genome_preparation (version 0.17.0). Methylation information for individual cytosines was extracted using the bismark_methylation_extractor tool from the Bismark package (version 0.22.3).

2.6 Authors contributions

ILC and IMM designed and conceived the study. ILC performed all RNA-seq, ICC and ATAC-seq experiments. DKT analysed the RNA-seq, ATAC-seq, ChIP-seq and WGBS data, with significant support from PLG. ILC and DKT prepared all the figures. ILC interpreted the data with significant input from PLG and IMM. ILC and DKT wrote the manuscript, with significant input from PLG and IMM. All authors read and accepted the final version of the manuscript.

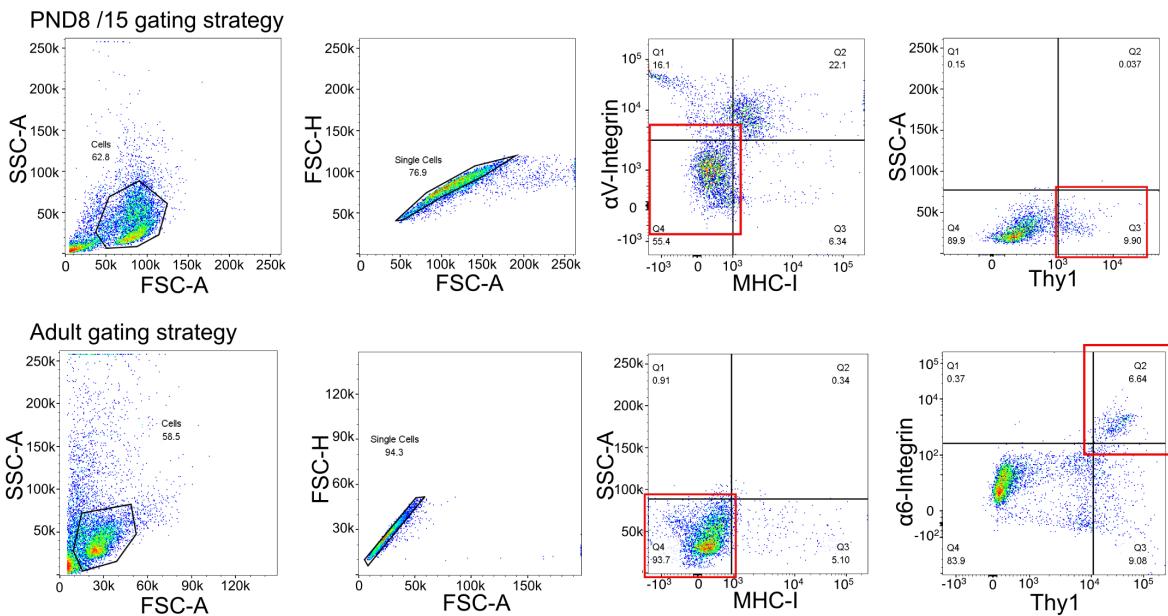
2.7 Acknowledgements

We thank Martin Roszkowski and Francesca Manuella for assisting with the breedings, Yvonne Zipfel for animal care in Zurich, Silvia Schelbert and Alberto Corcoba for taking care of the animal licenses and lab organization in Zurich. We thank Niharika Gaur for technical support with the Omni-ATAC protocol. We thank Rodrigo Arzate for conceptual support and critical reading of the manuscript. We thank Catherine Aquino and Emilio Yángüez from the FGC (ETH/UZH) Zurich for support and advice with sequencing and library preparation. We thank Jon Oatley, Tessa Lord and Nathan Law for advice and for providing detailed protocols of the testis dissection and preparation and for the immunocytochemistry of spermatogonial cells. We thank Zuguang Gu for his support with the heatmaps generation. We thank S3IT of UZH (www.s3it.uzh.ch) for computational infrastructure.

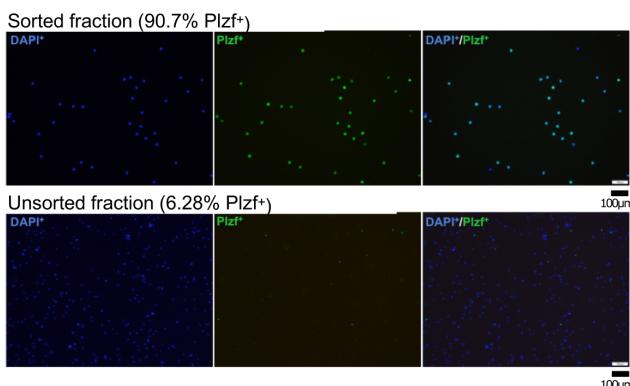
2.8 Supplemental figures

Figure S1

A



B



C

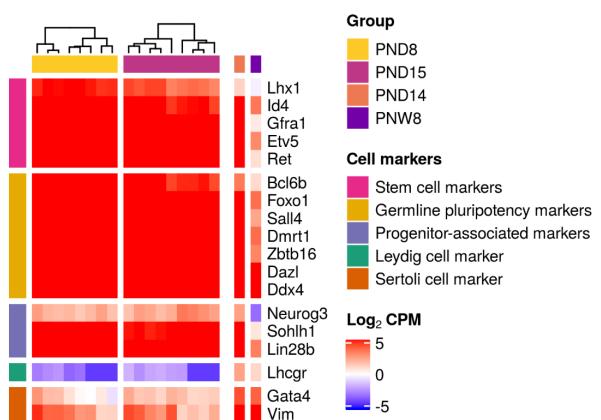


Figure S 2-1 FACS of spermatogonial cells from pup testis leads to high enrichment of PLZF⁺ cells.

- (A) Representative dot plots of the sorting strategy for spermatogonial cell enrichment. Gating based on side scatter/forward scatter (SSC-A/FSC-A) and forward scatter – height/ forward scatter – area (FSC-H/FSC-A) was performed to exclude cell debris and cell clumps;
- (B) PLZF⁺ cells are enriched following FACS, illustrated by immunocytochemistry on unsorted and sorted cell samples. Immunocytochemistry of PND15 unsorted and sorted testis cell suspension was performed by fixating the cells on poly-L-lysine coated slides. Cells were stained with anti-PLZF antibody (S19 clone, Active Motif) and VECTASHIELD (with DAPI) antifade mounting medium was used for mounting. Cells were visualized under a fluorescence

microscope and counted in 10 different fields of view / slide. The number of PLZF⁺ and PLZF⁻ cells from 10 different fields of view was averaged;

- (C) Heatmap of the expression profile of selected markers of spermatogonial and different testicular somatic cells extracted from the RNA-seq data on PND8, PND15 samples and on literature PND14 and PNW8 samples. Key genes for stem cell potential, stem and progenitor spermatogonia, and Leydig and Sertoli cells were chosen to assess the enrichment of spermatogonial cells in the sorted cell populations. Gene expression is represented in Log₂CPM (counts per million).

Figure S2

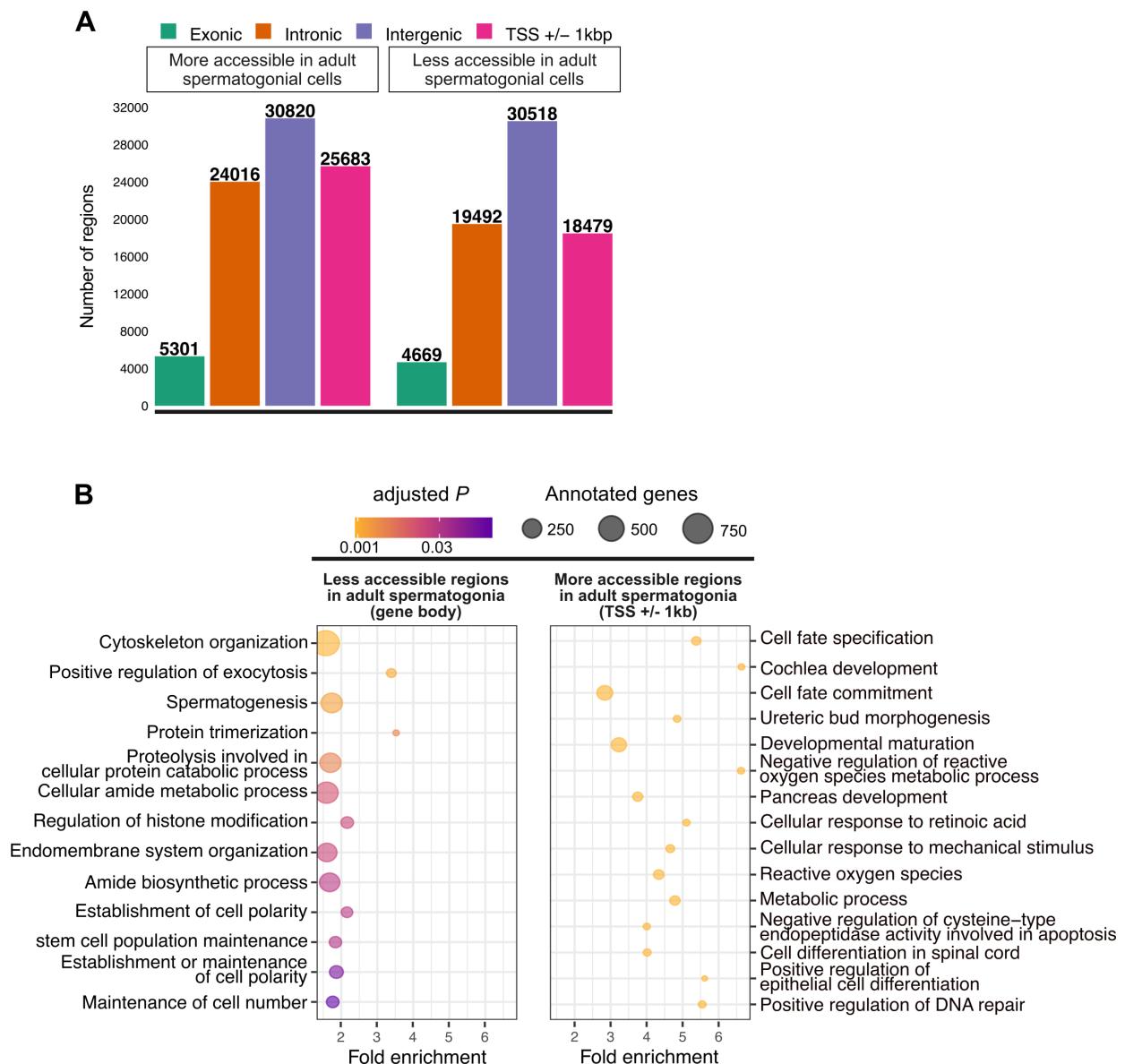


Figure S 2-2 Omni-ATAC profiles of PND15 adult spermatogonial cell samples and their genomic distribution.

(A) Genomic distribution of the 158, 978 Omni-ATAC regions identified;

(B) Dot plots of top enriched GO biological processes for regions with increased chromatin accessibility in adult spermatogonia, within gene bodies and around transcription starting sites (TSSs) of nearby genes (TSS +/- 1kb). The size of dots indicates the number of genes in the term and the color of each dot corresponds to the adjusted P value of the term's enrichment.

Figure S3

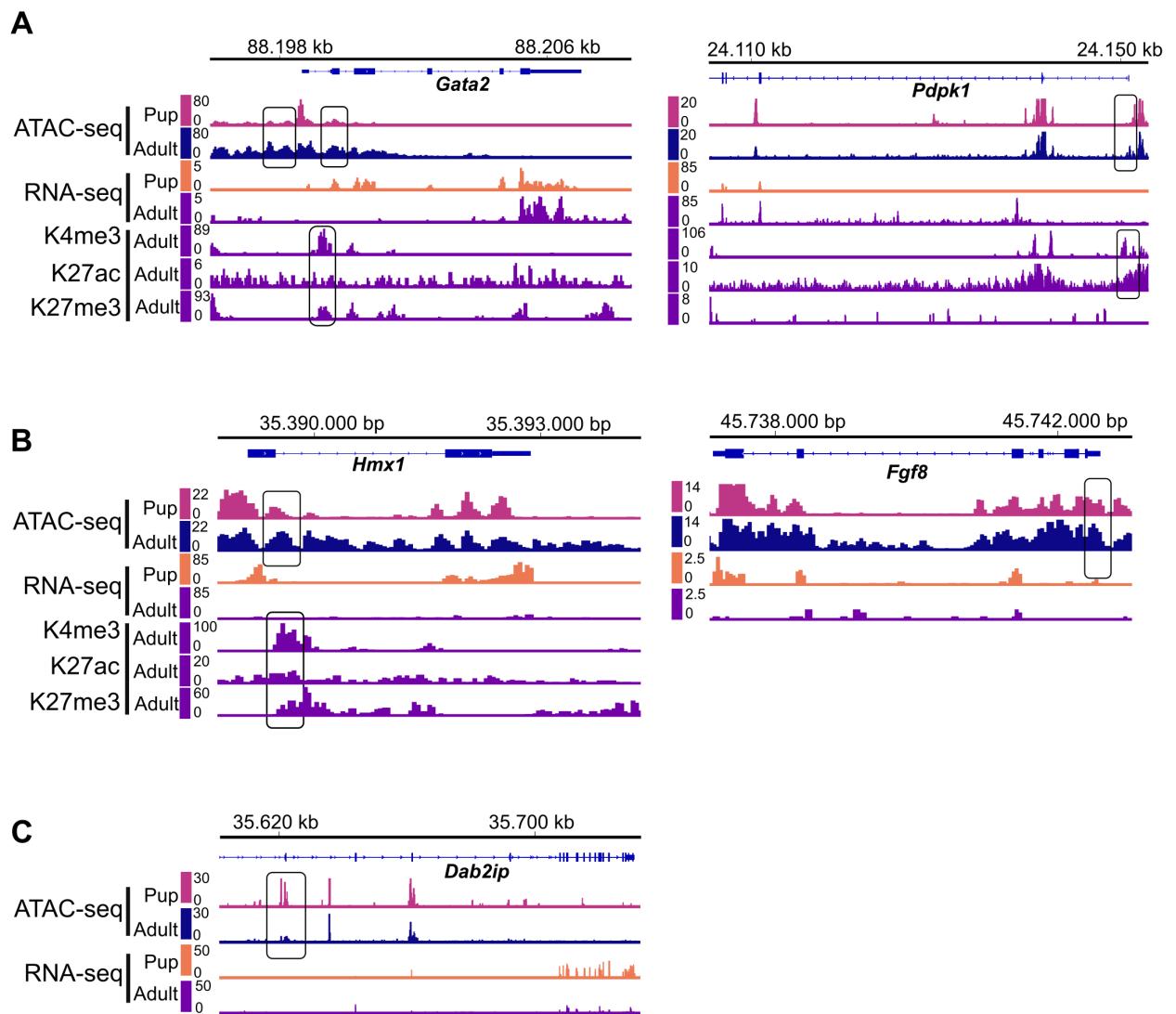


Figure S 2-3 Representative examples from Categories 1-3 resulted from the overlap of chromatin accessibility, gene expression and histone profiling datasets.

(A-C) Genomic snapshots from the Integrative Genomics Viewer (IGV, Broad Institute) of exemplary genes from Category 1 (*Gata2* and *Pdpk1*), Category 2 (*Hmx1* and *Fgf8*) and Category 3 (*Dab2ip*) showing relative abundance of: transcripts from RNA-seq, chromatin accessibility from ATAC-seq and enrichment of 3 different histone marks (H3K27ac, H3K4me3 and H3K27me3) from ChIP-seq. RNA-seq data corresponds to literature PND14 and adult (PNW8) spermatogonial cells and ATAC-seq data to PND15 and adult (PNW20) spermatogonial cells, respectively.

Figure S4

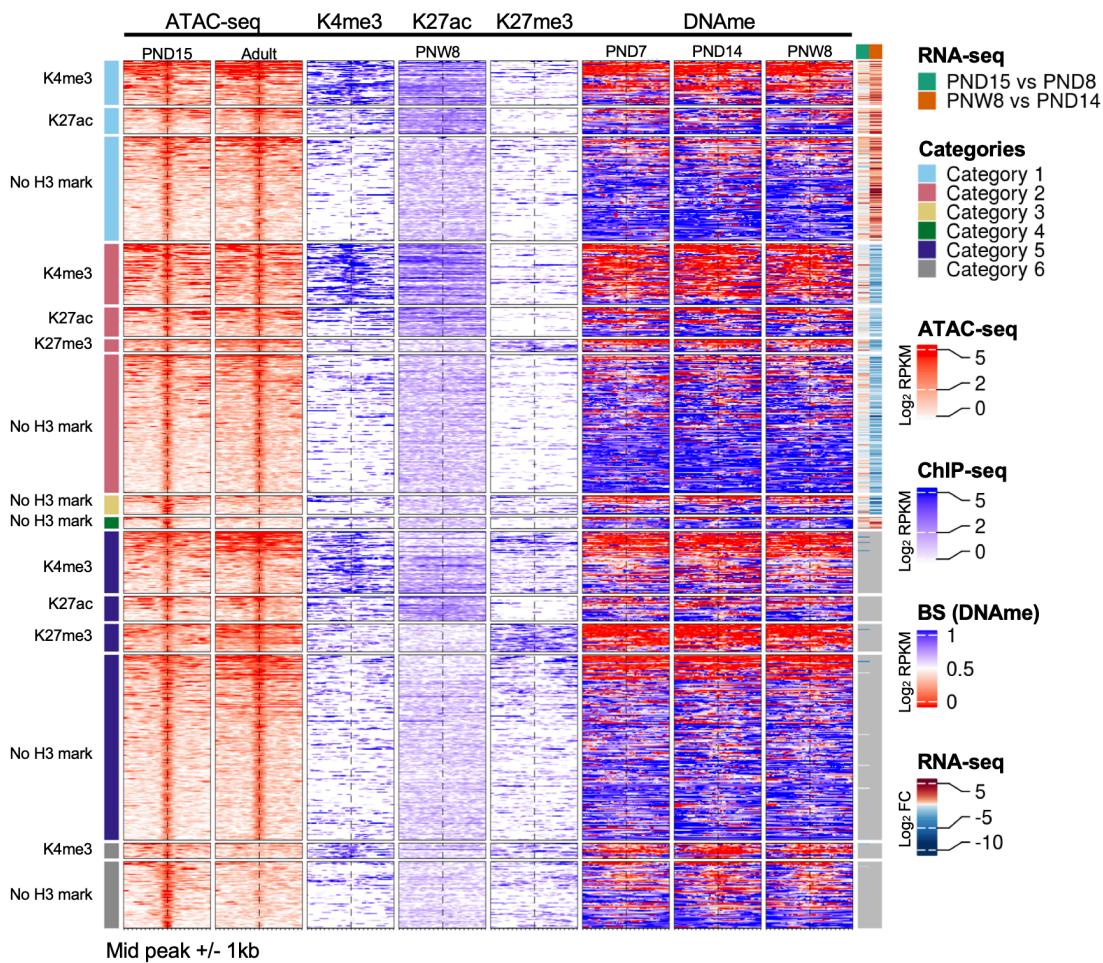


Figure S 2-4 DNAme profiles of spermatogonial cells do not show major changes over the period of testis postnatal maturation.

Enriched heatmaps showing the overlap between Category 1-4 regions, literature ChIP-seq data in PNW8 spermatogonia for H3K4me3, H3K27ac and H3K27me3, and DNAme data from BS in PND7, PND14 and PNW8 spermatogonia. For each of the Category 1-4 the following sub-categorization was applied:

- regions that are enriched for H3K4me3 (with or w/o H3K27ac and/or H3K27me3)
- regions that are enriched for H3K27ac (and lack both H3K4me3 and H3K27me3)
- regions that are enriched for H3K27me3 (and lack both H3K4me3 and H3K27ac)

Each line represents a peak region and the regions are ordered by the ATAC-seq signal. Mid-x-axis corresponds to the middle of a peak region and is extended to +/- 1 kbp. The color-key of the ATAC-seq, ChIP-seq and BS heatmaps represent ATAC-seq, ChIP-seq and BS signal, respectively. For RNA-seq, log₂ FC is shown from PND8 vs PND15 and PND14 vs PNW8 comparisons. For BS, the level of DNAme is between 0 and 1, with 0 representing a completely unmethylated locus and 1 a fully methylated locus, respectively.

Figure S5

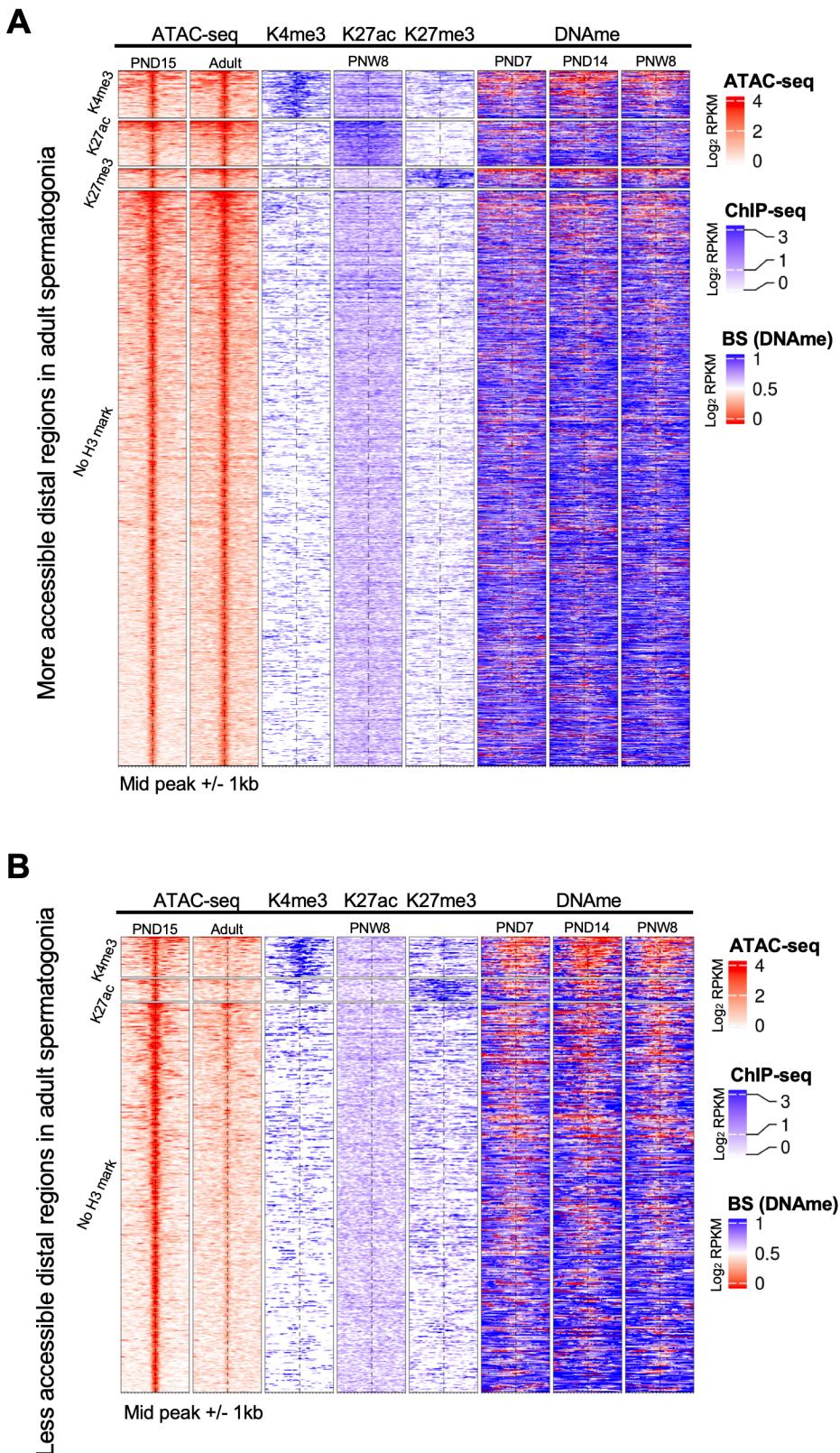


Figure S 2-5 Distinct chromatin profiles between PND15 and adult spermatogonia are present at distal regions across the genome. (legend continues on the next page)

(A-B) Enriched heatmaps showing the overlap between more accessible regions (A) and less accessible regions (B) situated in distal regions in spermatogonial cells, and literature ChIP-seq and DNAme data in PNW8 spermatogonia. The following sub-categorization was applied:

- regions that are enriched for H3K4me3 (with or w/o H3K27ac and/or H3K27me3)
- regions that are enriched for H3K27ac (and lack both H3K4me3 and H3K27me3)
- regions that are enriched for H3K27me3 (and lack both H3K4me3 and H3K27ac)

Each line represents a peak region and the regions are ordered by the ATAC-seq signal. Mid-x-axis corresponds to the middle of a peak region and is extended to +/- 1 kbp. The color-key of the ATAC-seq, ChIP-seq and BS heatmaps represent ATAC-seq, ChIP-seq and BS signal, respectively. For BS, the level of DNAme is between 0 and 1, with 0 representing a completely unmethylated locus and 1 a fully methylated locus, respectively.

Figure S6

A

Enriched TF motifs in less accessible ERVKs in adult spermatogonia			Enriched TF motifs in less accessible LINE L1s in adult spermatogonia		
TF	Enriched motif	% of target sequences	TF	Enriched motif	% of target sequences
PBX3		24.15%	FOXO1		15.02%
THAP11		22.25%	ZEB1		14.7%
ZNF143		27.21%	E2F3		13.79%
FOXL1		24.53%	KLF5		12.67%
NF-YA		25.88%	ZBTB17		12.74%

B

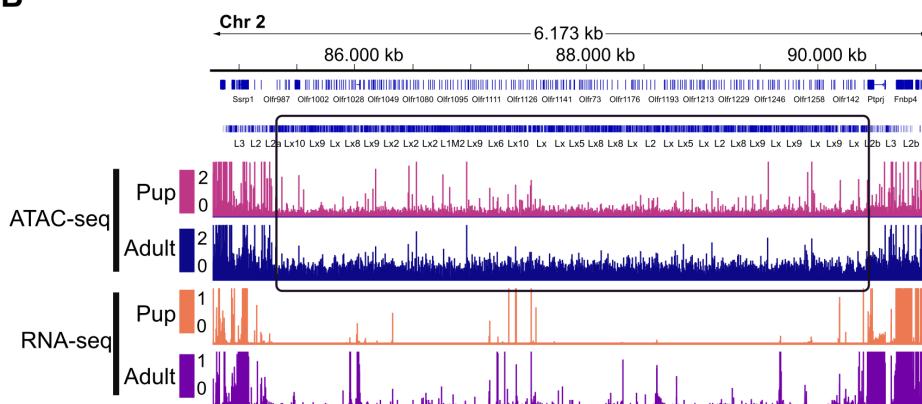


Figure S 2-6 Differentially accessible TEs exhibit enriched TF motifs and correspond to regions nearby non-random gene families

- (A) HOMER extracted consensus sequences for TF motifs enriched in the less accessible ERVK family and the more accessible LINE L1 family, respectively. Representative examples from the most enriched transcription factor families are depicted;
- (B) Genomic snapshots from the Integrative Genomics Viewer (IGV, Broad Institute) of a cluster of *Oifr* genes on Chr2 overlapping an increased density of LINE elements relative to the neighboring regions. The relative abundance of: transcripts from RNA-seq and chromatin accessibility from ATAC-seq are shown. RNA-seq data corresponds to literature PND14 and adult (PNW8) spermatogonial cells and ATAC-seq data to PND15 and adult (PNW20) spermatogonial cells, respectively.

3 Early life stress alters chromatin accessibility landscape and transcript usage in spermatogonial cells during postnatal testis maturation

Irina Lazar-Contes^{1,2}, Gretchen van Steenwyk^{1,2}, Deepak K Tanwar^{1,2,3}, Pierre-Luc Germain^{2,3}, Francesca Manuella^{1,2}, Martin Roszkowski^{1,2}, Niharika Gaur¹, Isabelle M Mansuy^{1,2,#}

¹ Laboratory of Neuroepigenetics, University of Zürich and Swiss Federal Institute of Technology, Brain Research Institute, Winterthurerstrasse 190, CH-8057, Zurich, Switzerland

²Institute for Neuroscience, Swiss Federal Institute of Technology, Zurich, Switzerland

³ Statistical Bioinformatics Group, Swiss Institute of Bioinformatics, CH-8057 Zurich, Switzerland

#Correspondence: mansuy@hifo.uzh.ch

Research paper in preparation

3.1 Abstract

Early postnatal environment is an important time period in the development of numerous cell types in the body, including germline cells, which exit quiescence and begin proliferating soon after birth. Little is known about the impact of environmental factors on spermatogonial cells, the most primitive germ cells in the testis, which encompass the germline stem cell population. In this study we show that the spermatogonial cells of exposed pups exhibit differential transcript usage and altered chromatin accessibility following an early life traumatic stress exposure. More than half of the genes with differential transcript usage encoded RNA-binding proteins with potential roles in chromatin remodelling, and known splicing factors. At the chromatin level, we identify regions of differential accessibility in spermatogonial cells from exposed pups, enriched in pathways related to neuronal development and differentiation, as well as pathways related to immune function and fatty acid metabolism. Finally, we show that the effects of stress exposure are not restricted to the germ stem cell milieu, and alter circulating proteins involved in the complement cascade and the innate immune response, as well as their corresponding receptors in the developing testis. These findings reveal that environmental insults can alter the spermatogonial cell transcriptome and chromatin landscape, and underscore the potential of environmentally-driven alterations to perturb early postnatal germ cell development.

3.2 Introduction

Exposure to environmental insults in rodents (toxicants, endocrine disruptors, dietary challenges, chronic stress) alters the transcriptome and epigenome of numerous cell types in the body, including resident stem cells in different postnatal tissues (Boku et al., 2015; Daun et al., 2020; Liu et al., 2015; Vrooman et al., 2015; Woo et al., 2011). Effects of chronic stressful experiences in postnatal life including maternal separation, limited nesting material or social deprivation have been mainly studied in the context of neural stem cell proliferation and neurogenesis (Boku et al., 2015; Daun et al., 2020; Suri et al., 2013). Conversely, dietary challenges or toxicants such as bisphenol A or heavy metals affect stem cells in several tissues including muscle, adipose or germline (Dias et al., 2018; Liu et al., 2015; Vrooman et al., 2015; Woo et al., 2011).

In mice, early postnatal life is a critical time for male germline development, with spermatogonial cells in the gonads exiting the quiescent state at postnatal day (PND) 1-2 after birth, and starting to populate the basement membrane of the testis. Following 35-37 days of successive mitotic divisions and 2 meiotic phases, spermatogonial cells give rise to adult sperm cells, capable of fertilization (Griswold and Oatley, 2013; Oatley and Brinster, 2006, 2012). We now know, based on findings from studies published in the past decade, that adult spermatozoa can be altered by various exposure types, by perturbing the epigenetic milieu of the cell (Chen et al., 2016; Dietz et al., 2011; Gapp et al., 2014, 2018; Rodgers et al.; Weiss et al., 2011). However, investigating the effects of exogenous factors on spermatogonial cells, the precursor cells of adult sperm has so far been limited (Caires et al., 2012; Forand et al., 2009; Houston et al., 2018; Vrooman et al., 2015; Zhu et al., 2000).

In this study we examine if an environmental exposure in early postnatal life can affect the transcriptome and epigenome of spermatogonial cells, the most primitive population of germ cells in the testis, which contains the stem cell population of the germline (Oatley and Brinster, 2006; Rooij, 2017). We used a mouse model of early postnatal chronic stress, in which changes in mature sperm are well documented, and have been causally linked to the effects of exposure (Franklin et al., 2010; Gapp et al., 2014, 2018; Steenwyk et al., 2019). We investigate the effects of this exposure on spermatogonia from control and exposed males in the developing and adult testis, to

be able to detect and compare immediate and long-lasting effects of our stress paradigm.

Transcriptomic profiling of spermatogonial cells in prepubertal testis at PND8 revealed alterations in isoform usage of genes linked to RNA processing and transport, including genes encoding RNA-binding proteins and several members of the spliceosome machinery. Alterations in biological pathways related to RNA splicing, DNA conformation change and transcription factor binding were also present at PND15, but not in adult spermatogonial cells transcriptome. To test the potential chromatin regulatory landscape underlying these transcriptome alterations, we performed Omni-ATAC in PND15 and adult spermatogonial cells from control and exposed males. At PND15 we identified a decrease in accessibility at regions enriched for immune processes, neuronal development and splicing. Pathways related to stress response, fatty acid and hormone metabolism, as well as motif enrichment for TFs important in lipid and immune signalling were enriched in the regions of increased accessibility in PND15 spermatogonia from exposed pups.

Further proteomic analyses of PND15 plasma revealed increased immune activation in the blood of exposed pups, and western blot validation confirmed the upregulation of proinflammatory high mobility group box protein 1 (HMGB1). Finally, we found that one of HMGB1 main receptors, receptor for advanced glycation end products (RAGE), was upregulated in prepubertal testis of exposed pups. Altogether, our findings reveal that postnatal environmental exposure affects germline development by altering transcript usage and chromatin accessibility of spermatogonial cells in the developing testis. We bring additional evidence that this may take place via the activation of immune-mediated signalling in the testis and in systemic circulation. Our study lies the foundation for further investigation of exposure-dependent transcriptome and epigenetic alterations in spermatogonial cells, and the consequences on adult sperm formation and embryo health.

3.3 Results

3.3.1 Exposure to traumatic stress alters spermatogonial cell transcriptome in the developing testis

To investigate if environmental factors affect spermatogonial postnatal development we conducted transcriptomic and chromatin accessibility profiling of spermatogonia from developing and adult testis, using a well-established mouse model of chronic traumatic stress. Exposure consisted in daily unpredictable maternal separation combined with unpredictable maternal stress (MSUS) from PND1 to PND14, resulting in metabolic and behavioural symptoms in adulthood (Figure S 3-1A-C) (Franklin et al., 2010, 2011; Gapp et al., 2014; Weiss et al., 2011). Enrichment of spermatogonial cells from pup and adult testis was achieved by FACS, using a previously established protocol (see Figure S 2-1A in Chapter 2, and Figure S 3-2A) (Kubota et al., 2004b). To validate the identity of our sorted cell populations, we performed immunocytochemistry using PLZF, a validated undifferentiated spermatogonial cell marker (Costoya et al., 2004), and found on average, that more than 85-90% of our cells were PLZF⁺ following FACS (see Figure S 2-1B in Chapter 2). Furthermore, gene expression profiles extracted from the RNA-seq data at PND8, PND15 and adult stages revealed high expression of undifferentiated spermatogonia and pluripotency markers (Figure S 3-2B). We note that low expression levels for some differentiating spermatogonial cell markers and somatic cell markers were still present following FACS (Figure S 3-2B).

First, we performed differential gene expression (DGE) analysis on RNA sequencing (RNA-seq) data from 8 control and 9 MSUS biological replicates (4 testes from 2 pups pooled for each biological replicate) from PND8. We found that the overall gene expression levels between control and MSUS pups in spermatogonial cells weren't significantly altered at PND8 (Figure S 3-3A). Although representative of the overall transcriptome state, gene-level analysis by itself does not offer information about the finer levels of transcriptome organization such as differential usage of transcripts (Baralle and Giudice, 2017; Wang, 2008). Therefore, we used our RNA-seq data to analyse differential transcript usage (DTU) between MSUS and control pups using IsoformSwitchAnalyzeR (Vitting-Seerup and Sandelin, 2017, 2019). We identified 55 genes with DTU in spermatogonial cells at PND8 (Figure 3-1A and Table S1). A subset of these genes (23 genes) displayed an isoform switch, defined as a pair of transcripts

coming from the same gene which are altered in the opposite direction (Figure 3-1B). The remaining genes (22) exhibited differential usage of a single transcript (Table S1). The most common differences in splicing events predicted using IsoformSwitchAnalyzeR, were due to alternative transcription starting sites (ATSSs) and alternative transcription termination sites (ATTs), followed by alternative 3' or 5' splice sites or exon skipping events (Figure 3-1C). The majority of isoform switches identified were between a non-coding and a coding variant, leading to the loss or gain of a functional protein domain (Figure 3-1D). We then performed overrepresentation analysis (ORA) of the genes with differential isoform usage using gProfiler and identified a significant enrichment ($FDR < 0.05$) for terms related to ribonucleoprotein complexes, translation, RNA processing and post-transcriptional regulation of gene-expression (Figure 3-2A and Table S1). Interestingly, more than half of the genes with DTU encoded RNA-binding proteins, with known roles in RNA processing, alternative splicing and chromatin organization and DNA binding (Table S1) (Han et al., 2017; He et al., 2016; Kwon et al., 2013). Exemplary genes in these categories include core members of the splicing machinery such as *U2surp* and *Hnrnpdl*, which showed DTU of the main protein-coding transcript equipped with an RNA-binding domain (Figure 3-2B). Previous knock-down experiments of *U2surp* and *Hnrnpdl* mainly affected gene networks involved in RNA processing, RNA splicing and transcription regulation (Li et al., 2019; De Maio et al., 2018). Other representative RNA-binding protein genes which displayed differential isoform switches at PND8 included *Cct2*, *Chmp4b* and *Hspa8*, molecular chaperones involved in spermatogenesis and sperm maturation (Dun et al., 2012). For these genes we observed an increased usage of the protein coding transcript equipped with the RNA-binding chaperonin domain (Figure S 3-3B and Table S1).

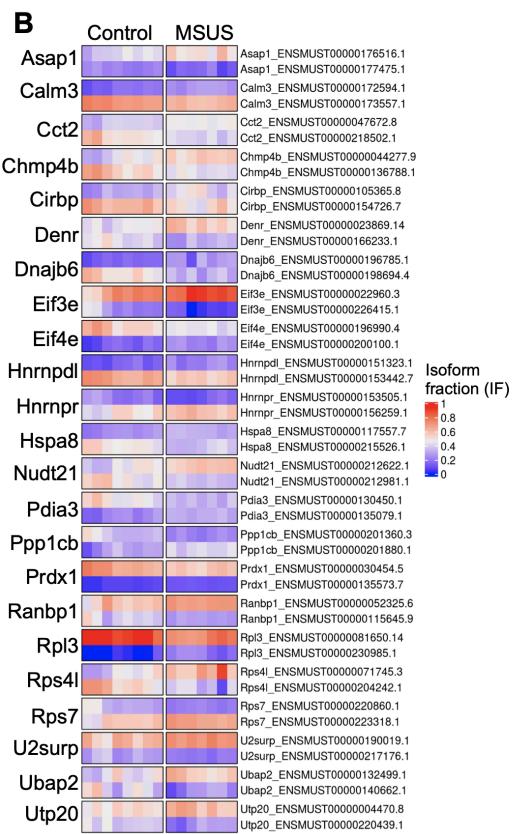
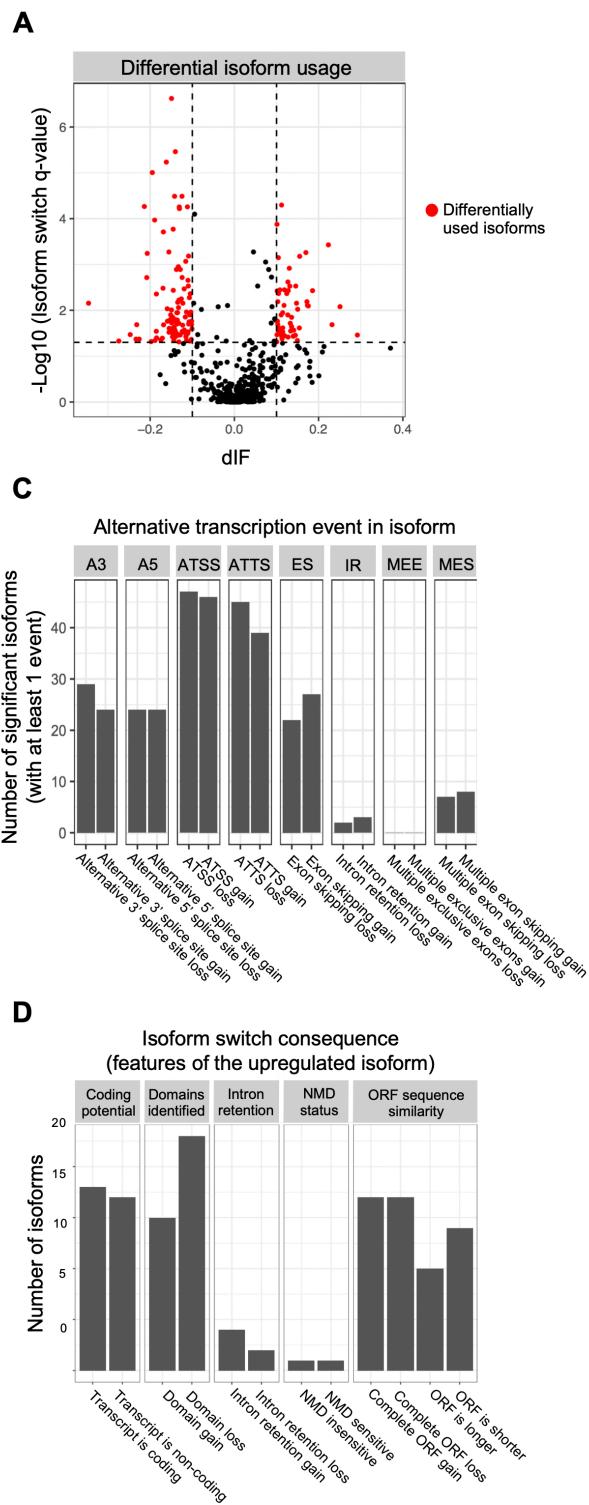


Figure 3-1 RNA-seq reveals altered transcript usage in PND8 spermatogonial cells from MSUS pups.

(A) Volcano plot of differentially used transcript between control and MSUS pups. Differentially spliced isoforms are highlighted in red ($q\text{-value} \leq 0.05$, $\text{dIF} > 0.1$, $\text{dIF} = \text{absolute change in isoform fraction (IF) between control and MSUS group}$, $n = 8$ control and 7 MSUS samples);

(B) Heatmap of the genes with isoform switches between control and MSUS pups (Ensembl GRCm38.p6 transcript IDs, only switches are shown). Each cell represents the isoform fraction (IF) value per sample and per condition (control and MSUS);

(C) Distribution of splicing events leading to alternative isoform usage (ATSS = Alternative transcription starting site, ATTS = Alternative transcription termination site);

(D) Distribution of switch consequences (NMD = Non-sense mediated decay, ORF = Open reading frame).

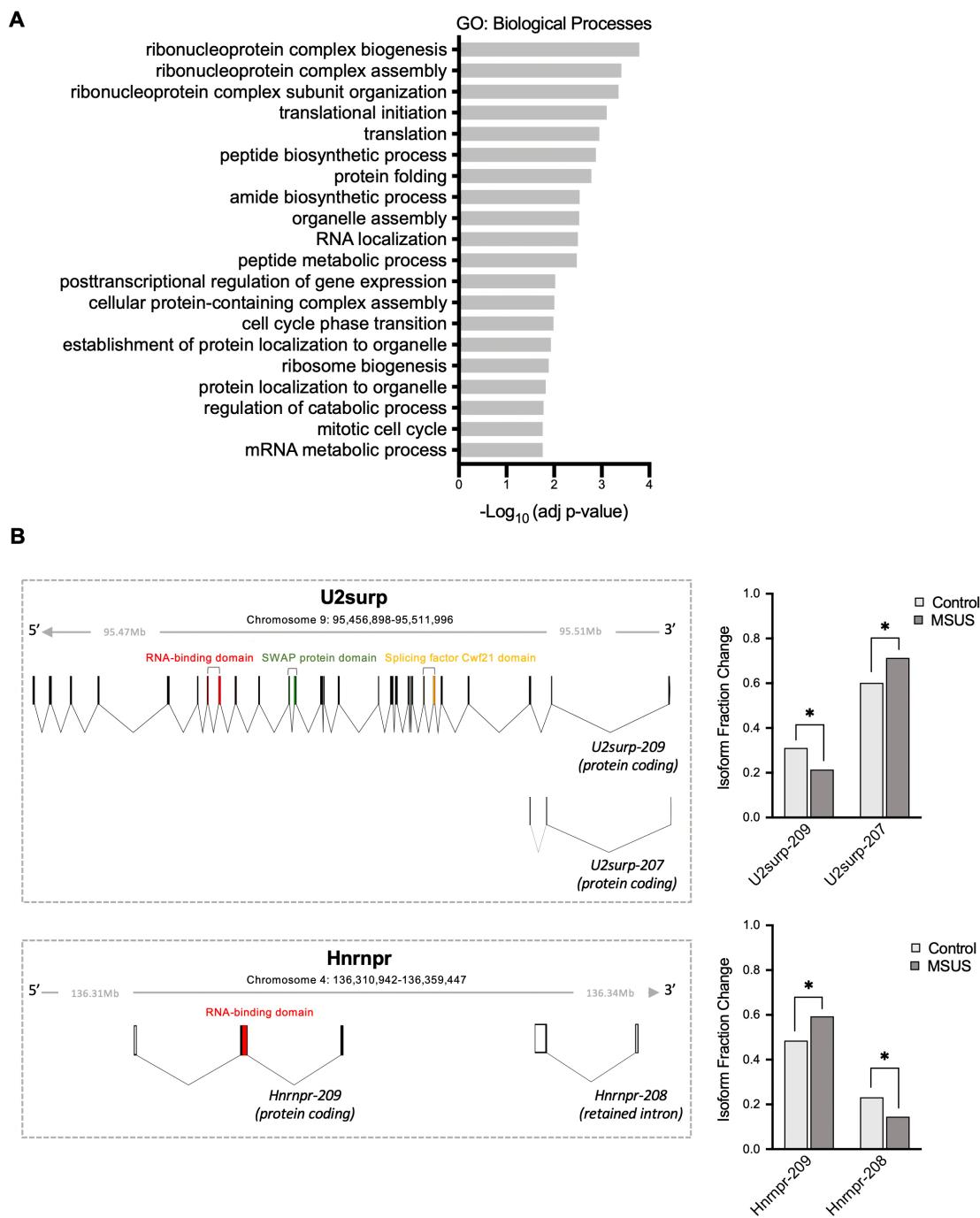


Figure 3-2 Differentially spliced genes associate with GO pathways related to RNA processing, transport and metabolism.

(A) Overrepresentation analysis (ORA) of the candidate spliced genes in PND8 spermatogonia, against GO database (Benjamini-Hochberg corrected p-value ≤ 0.05 , GO term size set to min 10 and max 1000 genes, min 5 genes/intersection);

(B) Exemplary spliceosome components U2surp and Hnrnpr are shown, both genes display an isoform switch between a coding and a non-coding transcript variant, leading to loss of the isoform with an RNA-binding domain (Colour coding of predicted protein domains by Pfam).

Next, we performed RNA-seq in PND15 spermatogonial cells from 9 control and 8 MSUS biological replicates (4 testes from 2 pups pooled for each biological replicate). PND15 corresponds to the immediate endpoint of stress exposure, therefore an immediate effect on spermatogonial cells was assessed at this timepoint. We did not detect any significant changes between the 2 groups at the individual gene level or in DTU of genes. Investigate if there is any change in the biological pathways utilized by spermatogonial cells in MSUS pups we performed Gene Set Enrichment Analysis (GSEA) on the t-ranked genes with dynamic expression between control and MSUS spermatogonial cells. Indeed, we found a number of negatively enriched pathways, in spermatogonial cells of MSUS pups compared to controls (Figure 3-3A). Top hits included gene programs related to RNA localization, RNA 3'-end processing and ribonucleoprotein complex biogenesis, as well as GO terms associated with chromatin organization such as DNA conformation change, DNA-protein binding and nucleus organization (Figure 3-3A and B and Table S2).

In contrast with the transcriptome alterations that we found at PND8 and PND15, no transcriptional differences at gene, transcript or pathway level were observed between adult spermatogonial cells from MSUS and control males (Figure S 3-3C). Collectively, our RNA-seq findings suggest that exposure to MSUS is able to induce changes at the transcriptional level in spermatogonial cells, during and immediately after chronic stress exposure, indicative of transcriptional susceptibility of early postnatal spermatogonia to an environmental insult.

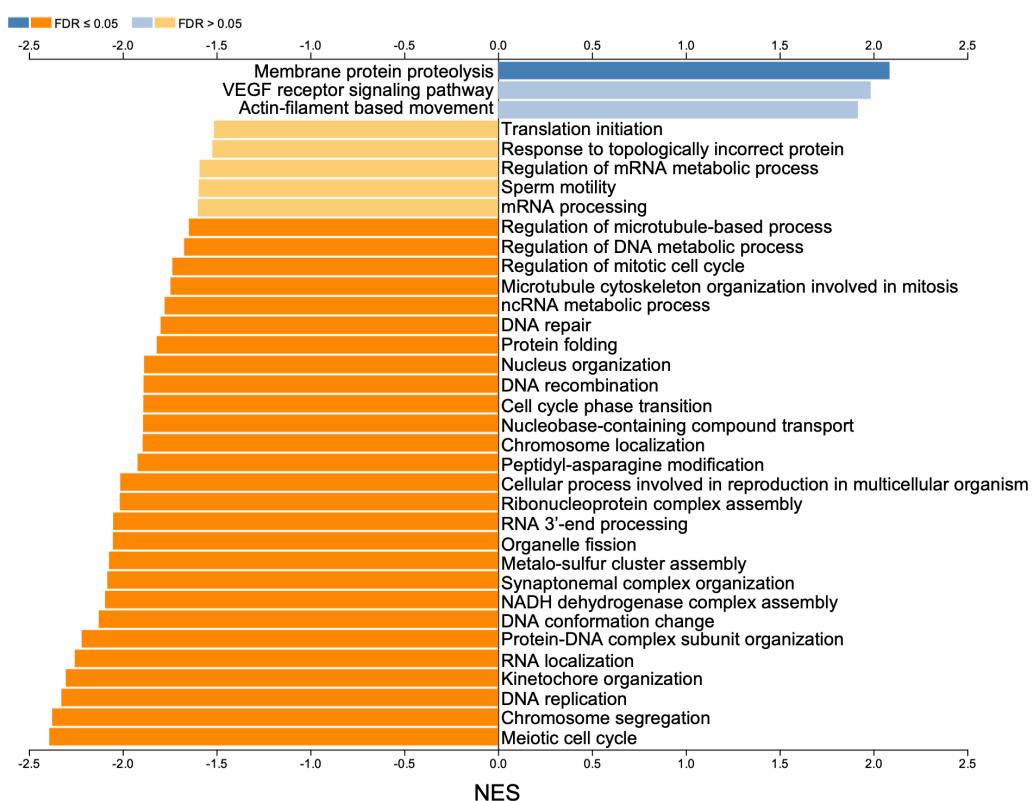
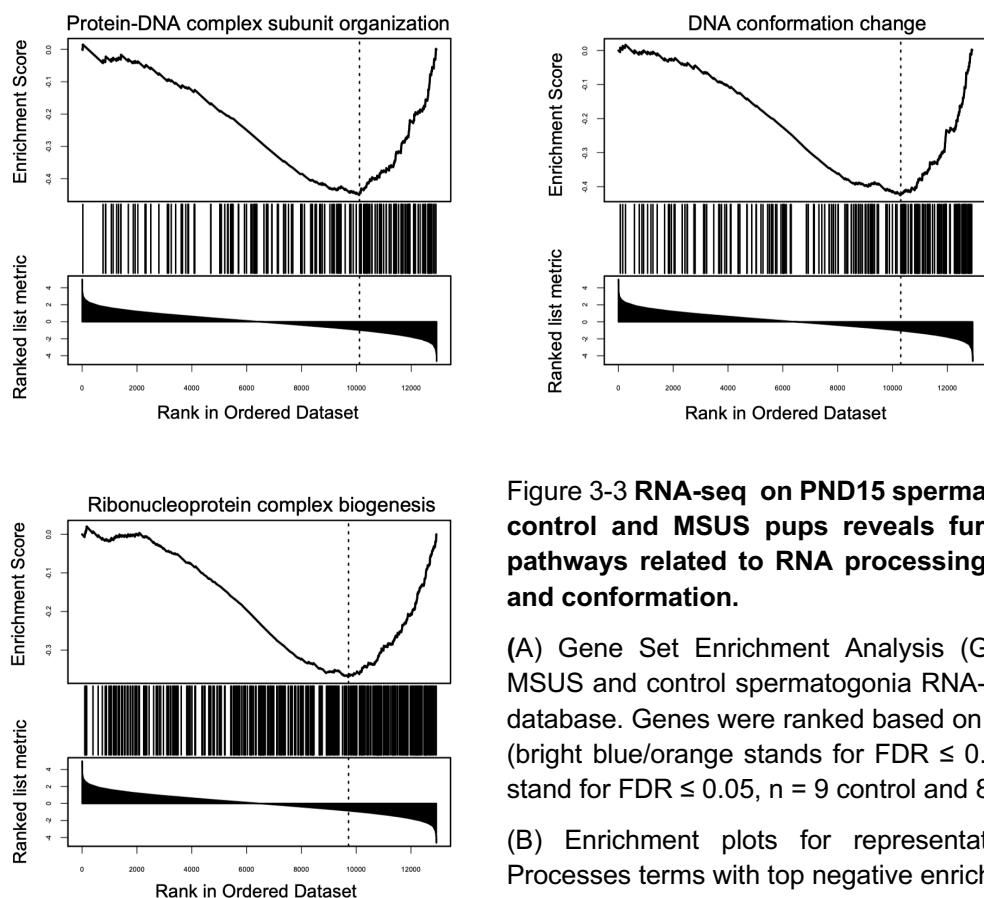
A**B**

Figure 3-3 RNA-seq on PND15 spermatogonial cells from control and MSUS pups reveals further alterations in pathways related to RNA processing and DNA binding and conformation.

(A) Gene Set Enrichment Analysis (GSEA) from PND15 MSUS and control spermatogonia RNA-seq ran against GO database. Genes were ranked based on calculated t-statistic (bright blue/orange stands for $FDR \leq 0.1$, dark blue/orange stand for $FDR \leq 0.05$, $n = 9$ control and 8 MSUS samples);

(B) Enrichment plots for representative GO Biological Processes terms with top negative enrichment scores.

3.3.2 Chromatin accessibility in spermatogonial cells is altered by early postnatal exposure

Open chromatin regions in the genome encompass 90% of the transcription factor (TF) binding sites and are involved in the dynamic regulation of gene expression (B. Teif and G. Cherstvy, 2016; Klemm et al., 2019; Thurman et al., 2012). Furthermore, previous studies have revealed that a vast majority of RNA-binding proteins are transcriptional regulators and chromatin modifiers (Fiszbein et al., 2016; He et al., 2016; Iwamori et al., 2016). We employed Omni-ATAC in PND15 spermatogonial cells from 6 control and 6 MSUS replicates (2 mice were pooled for each replicate) to investigate if the alterations in splicing and in pathways related to chromatin organization that we identified in PND8 and PND15 spermatogonial cells of MSUS pups are paralleled by changes at the level of open chromatin (Corces et al., 2017).

Accessible regions were identified by peak-calling on the merged nucleosome-free fragments (NFF) from all PND15 control and MSUS samples. Following the removal of lowly enriched regions, we included 237,110 regions in our downstream analyses (Table S3) (see Methods section for details). The overall genomic distribution revealed that most regions were localized in distal intergenic regions (37%), followed by introns and promoter regions (30% and 22%, respectively) (Figure 3-4A). Differential accessibility analysis revealed 282 less accessible and 310 more accessible regions in spermatogonial cells of MSUS pups ($\text{Abs Log}_2\text{FC} > 0.5$ and $p\text{-value} < 0.005$) (Figure 3-4B and Table S3).

To gain more insight into the biological relevance of these regions we employed Genomic Regions Enrichment of Annotations Tool (GREAT) using the GO biological processes (BP) database (McLean et al., 2010). Chromatin regions of decreased accessibility in PND15 MSUS spermatogonia associated with nervous system functioning such as axonal transport, neuronal migration and axo-dendritic transport. Furthermore, we also identified terms related to splicing, cytokine production and immune function among the top enriched terms (Figure 3-4C and Table S3). GO terms associated with more accessible chromatin regions were related to stress-responsive signalling pathways such as MAPK and mTORC, and to hormone metabolism and fatty acid oxidation (Figure 3-4C and Table S3).

Given the high number of GO BP terms that we found enriched in the regions of differential accessibility in MSUS spermatogonial cells from PND15 pups, we

continued by performing MEME ChIP analysis and identified transcription factor (TF) motifs enriched in these regions (Machanick and Bailey, 2011). In the regions of decreased accessibility, we identified motifs enriched for NFY-A, MTF1 and 1 novel motif which was not attributed to any known TFs (Figure 3-4D). NFY-A is an ubiquitously expressed TF, with known roles in regulating gene expression in mESCs, by maintaining chromatin open regions around TSSs of genes and recruiting members of the transcriptional machinery (Oldfield et al., 2019). Interestingly, a similar role has also been suggested for NFY-A in 2-cell embryos (Lu et al., 2016). In the regions of increased accessibility, we identified 3 enriched TF motifs, among which the Histone-lysine N-methyltransferase PRDM16 (Figure 3-4D). PRDM16 is a known regulator of brown adipose tissue formation, by interacting with PPAR- γ and activating its transcriptional activity. Several studies have revealed how PRDM16 promotes browning of white adipose tissue and can protect against obesity in mice (Kajimura et al., 2009; Lodhi et al., 2017; Seale et al., 2008). In addition, a role for PRDM16 in maintaining stem cells in different tissues in the mouse has also been suggested, including neural stem cells and hematopoietic stem cells by regulating oxidative stress (Chuikov et al., 2010; Shimada et al., 2017).

Lastly, we used the UCSC Data Integrator tool to investigate the predicted regulatory potential of our identified differentially accessible regions, by comparing them with the annotated cis-Regulatory Elements from ENCODE (Dunham et al., 2012; Myers et al., 2011). We found that 65 of the 282 regions of decreased accessibility in PND15 spermatogonia from MSUS pups displayed regulatory signatures, with distal enhancer-like signatures being the most predominant (Figure 3-4E and Table S3). Similarly, a third of the regions with increased accessibility in MSUS spermatogonial cells (101/310 regions) also displayed regulatory signatures, with the majority of these regions displaying signatures of distal or proximal enhancers (Figure 3-4E and Table S3).

Although we did not detect any significant transcriptome alteration in adult spermatogonial cells of MSUS males, we asked if changes in chromatin accessibility could represent a latent effect of MSUS exposure. Omni-ATAC in adult spermatogonial cells followed by differential accessibility analysis revealed very few changes in chromatin accessibility between MSUS and control males, with 26 less accessible regions and 46 regions gaining in accessibility (Abs Log₂FC > 0.5 and p-

value ≤ 0.01) (Figure S 3-4A and Table S4). GREAT analysis on all the differentially accessible regions revealed enrichment in GO terms mainly related to regulation of neuronal development, DNA binding and protein dephosphorylation (Figure S 3-4B and Table S4).

Altogether, our results revealed that early life chronic stress leads to chromatin accessibility changes in early postnatal spermatogonial cells, with only mild alterations detectable in adult spermatogonia. Notably, the open chromatin reorganization taking place in PND15 spermatogonial cells of MSUS pups relates to specific gene pathways and TF motifs implicated in immune function, lipid metabolism and stress response pathways. These findings, together with the transcriptome alterations we detected in early postnatal spermatogonial cells of MSUS pups, support the concept of environmentally-driven alterations to the germline, at an earlier cellular stage than the adult sperm (Chen et al., 2016; Gapp et al., 2014; Lambrot et al., 2013; Rodgers et al.; Sun et al., 2018a; Tyebji et al., 2020). In addition, our results suggest that these changes are more poignant in early postnatal development, during and immediately after exposure, prompting further investigation of their consequences on future progenitor cells during spermatogenesis.

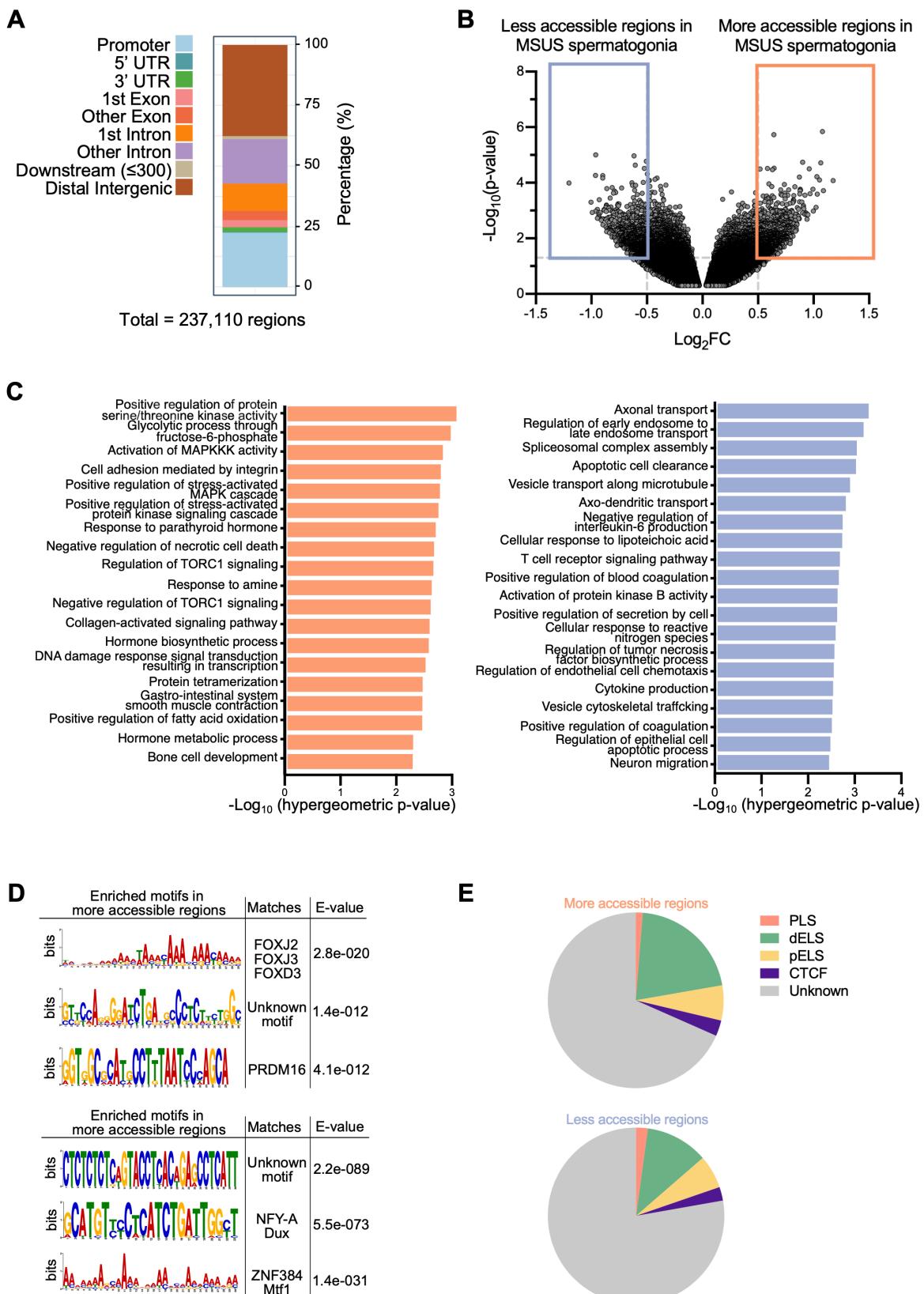


Figure 3-4 PND15 spermatogonial cells from MSUS pups display changes in chromatin accessibility.

- (A) Genomic distribution of all identified regions after merging NFFs from control and MSUS spermatogonia;
- (B) Volcano plot of differentially accessible regions between control and exposed pups at PND15. Points highlighted in blue denote decreased accessibility regions and orange points denote increased accessibility regions ($\text{Log}_2\text{FC} > 0.5$ and $p\text{-value} \leq 0.001$; $n = 6$ biological samples per group);
- (C) GREAT analysis of more and less accessible region proximal genes ($p \leq 0.01$, at least 3 gene hits per region);
- (D) HOMER transcription factor motif analysis of more and less accessible regions ($p \leq 0.05$);
- (E) SCREEN search of cis-regulatory element signatures by ENCODE (PLS = promoter-like signature, dELS = distal enhancer-like signature, pELS = proximal enhancer-like signature, CTCF = CTCF bound region).

3.3.3 Alteration in inflammatory proteins in blood of MSUS pups correlates with immune receptor activation in the developing testis

Spermatogonial cells reside within the seminiferous tubules, near the basement membrane and outside of the blood-testis-barrier (BTB) that is formed by the tight junctions of the Sertoli cells (Oatley and Brinster, 2006). The most primitive population consisting of undifferentiated spermatogonial cells, specifically locate in the vicinity of the blood vasculature network and interstitial cells, and spread throughout the basement membrane of the seminiferous tubules only after differentiation (Potter and De Falco, 2017; Yoshida et al., 2007). Since the testis vasculature provides a local source of signalling molecules, and potentially delivers systemic regulatory factors to undifferentiated spermatogonia nearby, we asked if systemic factors in blood are altered following early postnatal stress.

We performed LC-MS/MS-based proteomics of plasma from PND8 and PND15 control and MSUS pups ($n = 5$ animals / group), time window at which spermatogonial cells from exposed pups displayed the alterations in DTU and chromatin accessibility. Differential expression analysis of the identified proteins in blood plasma revealed 30 protein candidates significantly altered at PND8 ($p\text{-value} \leq 0.05$) (Figure S 3-5A and Table S5) and 31 protein candidates at PND15 ($p\text{-value} \leq 0.05$) (Figure 3-5A and Table S5). To better understand the biological roles of these differentially altered proteins we performed Overrepresentation Analysis (ORA) using gProfiler (Raudvere et al., 2019). At PND8 we found enrichment for pathways associated with the innate immune response and with lipid transport (Figure S 3-5B and Table S5). Top 10 GO pathways included lipid and vitamin binding, cytolysis and complement activation, with the majority of the proteins in the pathways being downregulated in MSUS pups (Figure S 3-5C and Figure S 3-5D). At PND15, enriched GO terms were also predominantly related to inflammatory response and pathways activated in response to stress and external stimuli (Figure 3-5B and Supplementary Table S5). In contrast to PND8, candidate proteins associated with these pathways were upregulated at PND15, pointing towards an exacerbated immune response in blood of MSUS pups (Figure S 3-6A-C).

In order to validate our proteomic findings, we performed western blot on plasma samples from a different cohort of PND15 MSUS and control pups ($n = 5$ animals / group). We confirmed the upregulation of several candidate proteins including

HMGB1, a ubiquitously expressed protein with multifaceted functions (Figure 3-5C). When released in the extracellular space HMGB1 can act as a proinflammatory cytokine and activate downstream proinflammatory receptors such as RAGE and TLR4, but can also function as a nuclear chromatin-remodelling factor and a critical regulator of mitochondrial function (Bertheloot and Latz, 2017; Kierdorf and Fritz, 2013). To investigate if HMGB1 receptors in the testis are altered, we performed RT-qPCR analysis on PND15 testis and found a significant upregulation of *Rage* mRNA in PND15 MSUS pup testis ($p = 0.0198$), the main downstream receptor of HMGB1 (Fig. 5D). We also found an upregulation of mitochondrial factors *Tfam* ($p = 0.0224$) and *Nrf2* ($p = 0.0311$) in the testis of PND15 MSUS pups (Figure 3-5D). TFAM is a structural and functional homolog of HMGB1 within the mitochondrial space, responsible for mitochondrial gene expression (Julian et al., 2012). Deficiency in *Tfam* leads to mitochondrial electron transport chain proteome imbalance, thermogenesis defects and metabolic disorders (Julian et al., 2012; Little et al., 2014; Masand et al., 2018). Additionally, under stress conditions, TFAM can be released from cells similarly to HMGB1, and act as a proinflammatory signalling factor via RAGE activation (Julian et al., 2012). Interestingly, PRDM16, which displayed enrichment of binding sites in the regions of more accessible chromatin in PND15 spermatogonial cells of MSUS pups, is also part of the extensive network of mitochondrial factors, together with TFAM and NRF2, which are responsible for metabolic fitness and adaptive responses to environmental stressors (Dominy and Puigserver, 2013). Together, our results suggest significant changes in factors related to inflammation, lipid metabolism and mitochondrial function in blood of MSUS pups, which correlate with an upregulation of corresponding receptors and binding partners in the testis.

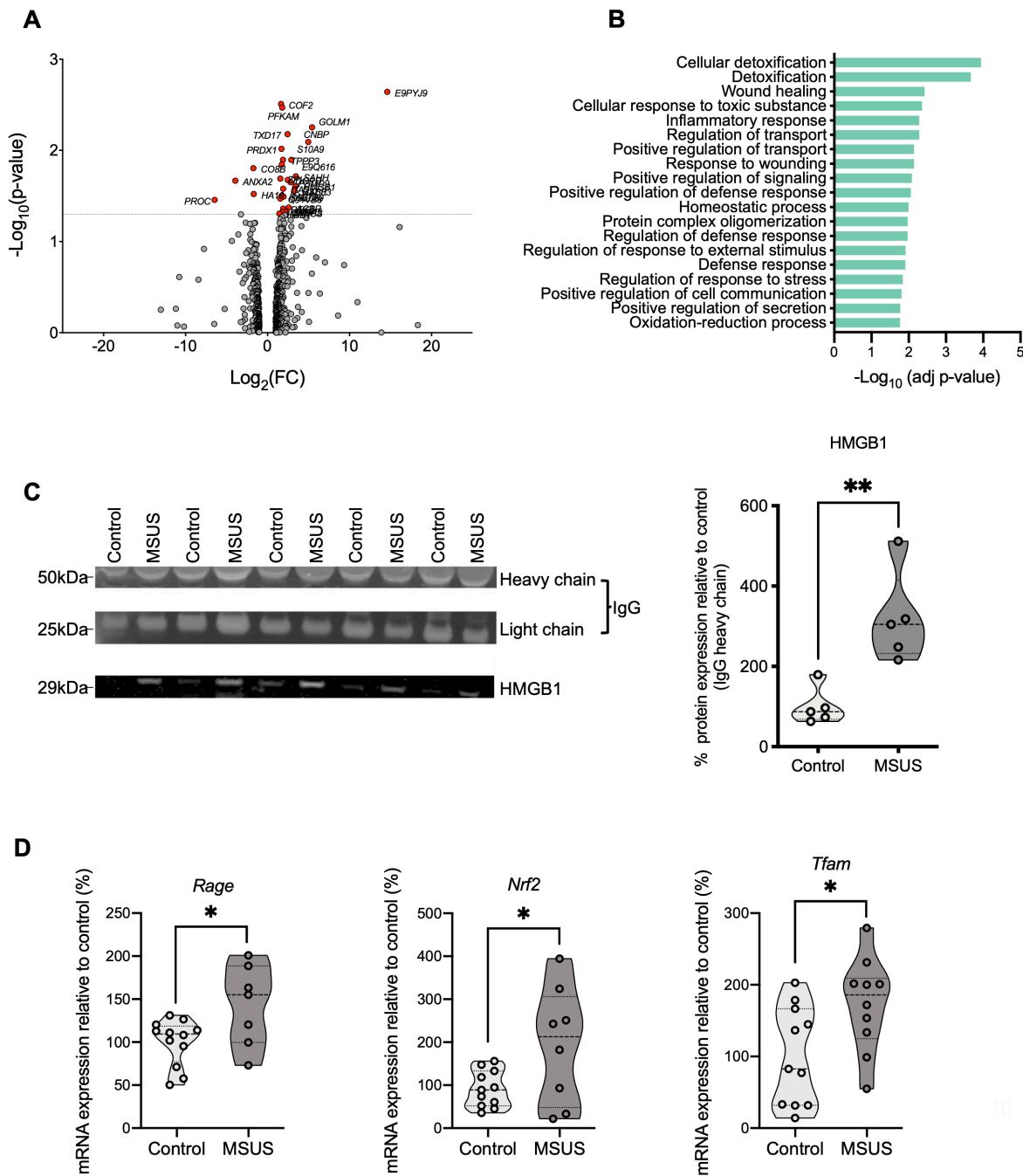


Figure 3-5 Exposure alters protein signatures in blood and their corresponding receptor mRNA in testis of pups at PND15.

(A) Volcano plot showing fold-change of protein expression in exposed pups compared to control pups. Proteins with significantly different expression between the two groups are highlighted in red ($\text{Log}_2\text{FC} \geq 1$, $p\text{-value} \leq 0.05$, $n = 5$ biological samples per group);

(B) ORA on candidate proteins with differential expression between the 2 groups at PND15 (Benjamini-Hochberg corrected $p\text{-value} \leq 0.05$, GO term size set to min 10 and max 1000 genes, min 3 genes intersection). The complete list of GO terms is available in Table S5;

(C) Western blot validation of HMGB1 protein upregulation in PND15 blood plasma from exposed pups (two-tailed t-test, $t=3.968$, $df=8$, ** $p\text{-value} < 0.01$, $n = 5$ samples per group);

(D) RT-qPCR of *Rage*, *Nrf2* and *Tfam* mRNA in PND15 testis (two-tailed t-test, $t=2.572$, $df=17$, * $p\text{-value} < 0.05$, $n = 12$ control and 7 MSUS samples).

3.4 Discussion

In this study we show that environmental exposure in early postnatal life alters the transcriptome and epigenome of developing germ cells in the testis. Spermatogonial cells encompass the stem cell population in the germline, which only accounts for around 0.03% of the total germ cells in the testis. Although intensely researched, a distinct molecular profile of spermatogonial stem cells, which would allow complete purification, has yet to be determined (Chan et al., 2014; Law et al., 2019; Oatley and Brinster, 2012). In our study, we employed a previously established FACS strategy, using a combination of known surface markers, which allowed the isolation of an 80-90% pure undifferentiated spermatogonial population. Admittedly, our sorted cell populations still exhibited low expression levels of differentiating spermatogonial cell markers.

Spermatogonial cell proliferation starts immediately after birth in rodents, and is characterized by intense transcriptome and epigenome reprogramming (Hammoud et al., 2014, 2015; Maezawa et al., 2018). Currently, little is known about the effects of environmental exposure on such early stages of postnatal germ cell development in the testis. By investigating gene expression and transcript usage, 2 distinct levels of transcriptome organization, we find changes in the transcript usage of a number of genes in spermatogonial cells at PND8, in the absence of global changes in gene expression. Unlike gene-level analysis, transcript usage is an underexplored regulatory mechanism. Recently, advances in bioinformatic tools have allowed the study of genome-wide differential transcript usage from RNA-seq data, and have helped investigating how transcript dysregulation contributes to various types of disorders (Coelho et al., 2020; Hentze et al., 2018; MM Scotti, 2016). Although specific targets and functions of the differentially used transcripts we have identified are largely unknown, GO enrichment analysis revealed an association to functions such as RNA processing, metabolism and transport.

By performing RNA-seq at PND15, we also identified differential enrichment of terms associated with RNA processing and splicing pathways, but also in pathways associated with DNA binding, chromatin organization and DNA conformation. However, we did not detect any further differences in transcript usage or in the global gene expression levels. Notably, our spermatogonial cell population encompasses a

heterogenous mix of stem cells, undifferentiated progenitors and potentially scarce populations of differentiating spermatogonia, therefore a finer dissection of these individual cell populations could be needed to investigate in more depth the transcriptional response to early life postnatal stress exposure.

Previous studies have identified a role for numerous RNA-binding proteins not only in splicing regulation, but also in chromatin organization (Han et al., 2017; Kwon et al., 2013; Rappsilber et al., 2002). In an extensive proteomic characterization of mESCs and neural cells, Han et. al found that half of the RNA-binding proteins they identified were enriched for chromatin-associated functions, thereby suggesting an extensive crosstalk between RNA processing and chromatin regulatory networks (Han et al., 2017). Furthermore, open chromatin regions display a dynamic reorganization during spermatogenesis. The transition from spermatogonial cells to spermatids has been recently found to be accompanied by vast changes in chromatin accessibility (Maezawa et al., 2018). Therefore, changes at the chromatin level in spermatogonial cells have the potential to influence later stages of germ cell development, with long-lasting consequences on the spermatogenic process. By employing Omni-ATAC in PND15 spermatogonial cells we identified hundreds of differentially accessible chromatin regions between control and MSUS pups. Notably, these regions were enriched in distinct biological pathways. Regions of decreased chromatin accessibility related to neuronal function and inflammatory pathways, while regions of increased accessibility displayed a predominant association with stress response and metabolic pathways such as fatty-acid oxidation and hormone metabolism. Furthermore, distinct TF motifs were enriched between more accessible and less accessible chromatin regions. To our knowledge, this is the first time that early life traumatic stress is shown to alter chromatin accessibility in spermatogonial cells.

Previous work from our lab has shown that fatty acid metabolism and arachidonic acid precursors are altered in the blood of male mice exposed to early-life traumatic stress (Steenwyk et al., 2019). Interestingly, unlike adult sperm cells which reside in an immunoprivileged environment ensured by the BTB, undifferentiated spermatogonia are located outside the BTB, and display preferential localization in seminiferous tubule areas which are adjacent to testis blood vessels (Mrak and Cheng, 2015; Potter and De Falco, 2017; Yoshida et al., 2007). This allows for signalling from adjacent

interstitial cells and resident immune cells of the testis, as well as from blood circulating factors (DeFalco et al., 2015; Heinrich and DeFalco, 2019; Potter and De Falco, 2017).

On the basis of this evidence, we further investigated the effects of early stress exposure on circulating factors by profiling the proteome of PND8 and PND15 control and MSUS pups. We observed consistent changes in expression of proteins involved in the complement cascade and mitochondrial function across early postnatal age. Complement cascade and innate immune system activation are known regulators of neuronal proliferation and gonad development (Coulthard et al., 2017; Lee et al., 2017). Upregulated proteins included master regulators of numerous cellular inflammatory and redox processes such as S100A9 and HMGB1, two circulating alarmins with extracellular and intracellular roles (Bertheloot and Latz, 2017). We further showed that HMGB1 upregulation correlated with the upregulation of *Rage* mRNA expression in PND15 testis of MSUS pups, one of its major cellular receptors. In addition, *Nrf2* and *Tfam*, 2 important mitochondrial factors, were also upregulated in the testis of MSUS pups. Notably, HMGB1, NRF2 and TFAM all rely on RAGE binding for activation of downstream signalling pathways (Dominy and Puigserver, 2013; Tang et al., 2011; Yan et al., 2009), indicative of a potential common avenue through which circulating factors could alter the testis milieu, with potential consequences on germ cells. Indeed, 2 of these downstream signalling pathways, MAPK and mTOR were associated specifically with the chromatin regions of increased accessibility in spermatogonial cells of MSUS pups. Furthermore, enrichment of PRDM16 motif, a TF important for the control of genes that regulate oxidative stress levels in mouse hematopoietic and neural stem cells, was also enriched at the regions of increased accessibility in MSUS spermatogonia.

Very few other studies have investigated the effects of an environmental stressor such as chronic unpredictable stress or immobilization stress on spermatogonial cells in the testis (Hirano et al., 2014; Yazawa et al., 1999). Both types of stressors led to an increased number of damaged or apoptotic spermatogonial cells and of spermatocytes, the more differentiated progenitor cells within the testis. However, none of the studies investigated the molecular underpinnings of this apoptotic phenotype.

In this study we bring evidence that open chromatin regions associated with pathways important for inflammatory processes, redox homeostasis and metabolic regulation

are altered in spermatogonial cells following chronic stress exposure. We further show that these changes are not solitary, but accompanied by alterations in similar pathways at the testis level and in the blood of these pups. This correlative evidence suggests that blood factors could contribute to an altered testis milieu, and that these perturbations may affect spermatogonial cells. Understanding the molecular mechanisms of this inter-tissue communication is essential for shedding light on how environmentally-driven alterations can be passed on to the germline.

3.5 Methods

3.5.1 Animals and MSUS paradigm

Animal experiments were approved by the veterinary authority of the Canton of Zurich. All procedures complied with the relevant ethical regulations for animal testing and research. C57BL/6 J male and female mice aged 2 months were purchased from Janvier Labs (France) and kept in a temperature and humidity-controlled facility under a 12h reverse light–dark cycle (lights off at 8:00, lights on at 20:00) with food and water provided *ad libitum*. Cages and bedding were changed once per week.

MSUS paradigm was performed as previously described (Franklin et al., 2010). Briefly, primiparous C57BL/6J female and male mice aged 2.5-3 months were mated, and randomly selected dams and litters were assigned to the control or MSUS groups. MSUS dams and litters were subjected to 3 hours of unpredictable maternal separation combined with unpredictable maternal stress for 14 days, starting at PND1. The maternal stress was performed at different times during the 3 hours of separation, and consisted from 5 min of an acute swim in cold water (18 °C for 5 min) or 20 min restraint in a plexiglass tube. Control animals were left undisturbed apart from cage changes once per week. At PND21, all pups were weaned and housed (3-5 mice/cage) by gender and treatment. To avoid litter effects, siblings were distributed in different cages of the same treatment group.

Behavioural assessment of adult male mice was performed using the elevated plus maze. The setup consisted of a platform with two open and two closed arms (dark grey PVC, 30 x 5 cm) at 60 cm elevation from the floor. Mice were placed in the centre of the platform facing a closed arm and video recorded (Viewpoint tracking software) for 5 min. The latency to first enter an open arm and the total time spent in the open arms were manually scored. The experimenter was blind to treatment.

3.5.2 Testis dissection and spermatogonial cell enrichment

Testes were collected from pups at PND8, PND15 and adults at 5 months of age from control and MSUS groups. Testicular single-cell suspensions were prepared as previously described with slight modifications (Kubota et al., 2004b). For pup spermatogonia preparations, 4 testes from 2 pups were pooled for each biological replicate, to obtain enough enriched spermatogonial cell populations for subsequent

analyses. Following gentle removal of the tunica albuginea, the seminiferous tubules were digested using 0.25% trypsin-EDTA (ThermoFisher Scientific) and 7mg/ml DNase I (Sigma-Aldrich) solution for 5 min at 37°C. The suspension was pipetted up and down 10 times and incubated again for 3 min at 37°C. The digestion was stopped by adding 10% foetal bovine serum (ThermoFisher Scientific) and the cells were passed through a 20µm pore-size cell strainer (Miltenyi Biotec). Cells were resuspended in PBS-S (PBS with 1% PBS, 10 mM HEPES, 1 mM pyruvate, 1mg/ml glucose, 50 units/ml penicillin and 50 µg/ml streptomycin) and used for fluorescence activated cell sorting (FACS).

For adult testis digestion, a 2-steps enzymatic procedure was performed. Briefly, each testis was treated with collagenase type IV (Sigma-Aldrich) for 5 min at 37°C and vigorous swirling until the tubules were completely separated. The tubules were placed on ice until they sedimented completely, the supernatant removed and the remaining tubules were washed with HBSS. Washing/sedimentation steps were repeated 3 times and were necessary to remove interstitial cells. Tubule fragments were digested next with 0.25% trypsin-EDTA and 7mg/ml DNase I solution. The resulting single-cell suspension was filtered through a 40µm strainer (Corning Life Sciences) and washed with HBSS. Cells were resuspended in PBS-S, layered on a 30% Percoll solution (Sigma-Aldrich) and centrifuged at 600g for 8 min at 4°C without braking. The top 2 layers (HBSS and Percoll) were removed and the cell pellets were resuspended in PBS-S and used for sorting.

FACS was performed by incubating PND8 and 15 testis cell suspensions with BV421-conjugated anti-β2microglobulin (β2M; S19.8), biotin-conjugated anti-THY1 (CD90.2; 53-2.1), and R-phycoerythrin (PE)-conjugated anti-αv-integrin (CD51; RMV-7) antibodies. THY1 was detected by staining with Streptavidin-Alexa Fluor 488. For adult testes, cells were stained with anti-α6-integrin (CD49f; GoH3), BV421-conjugated anti-β2M (S19.8), and PE-conjugated anti-THY1 (CD90.2; 30H-12) antibodies. All antibodies were obtained from BD Biosciences (San Jose, United States). Prior to sorting, 1 µg/ml propidium iodide (Sigma-Aldrich) was added to the cell suspensions to discriminate dead cells. Cell sorting was performed on a FACS Aria III 5L, at 4°C using an 85um nozzle, at the Cytometry Facility of University of Zurich. Cells were collected in 1.5 ml Eppendorf tubes for RNA-sequencing and ATAC-sequencing library preparations.

3.5.3 Immunocytochemistry

The protocol used for assessing spermatogonial cell enrichment after sorting was kindly provided by the Oatley Lab at Washington State University, Pullman, USA (Yang et al., 2013). Briefly, 30,000-50,000 cells were adhered to poly-L-Lysine coated coverslips (Corning Life Sciences) in 24-well plates for 1 h. Cells were fixed in freshly prepared 4% PFA for 10 min at room temperature then washed in PBS with 0.1% Triton X-100 (PBS-T). Non-specific antibody binding was blocked by incubation with 10% normal goat serum for 1 h at room temperature. Cells were incubated overnight at 4°C with mouse anti-PLZF (0.2 µg/ml, Active Motif, clone 2A9) primary antibody. Alexa488 goat anti-mouse IgG (1 µg/mL, ThermoFisher Scientific) was used for secondary labelling at 4°C for 1 h. Coverslips were washed 3x and mounted onto glass slides with VectaShield mounting medium containing DAPI (Vector Laboratories) and examined by fluorescence microscopy. Stem cell enrichment was determined by counting PLZF+ cells in 10 random fields of view from each coverslip and dividing by the total number of cells present in the field of view (DAPI-stained nuclei).

3.5.4 RNA-seq library preparation and data analysis

Library preparation: 25,000 - 50,000 spermatogonial cells from PND8 and PND15 pup testis were used for RNA extraction using the AllPrep RNA/DNA Micro Kit (Qiagen). Library preparation was performed from 10 ng of total RNA / sample using the SMARTer® Stranded Total RNA-Seq Kit v2 - Pico Input Mammalian (Takara Bio USA, Inc.) according to the manufacturer's instructions. For adult spermatogonia, 1000 cells were collected by FACS, directly lysed in 10 µL buffer RLT Plus (Qiagen) and processed following the SmartSeq2 protocol (Picelli et al., 2014). RNA and library quality were assessed using the Bioanalyzer 2100 (Agilent Technologies). Libraries were prepared at the Functional Genomics Centre Zurich and sequenced on an Illumina HiSeq 4000 using v4 reagents and 100 bp single-end read length. The number of samples sequenced at each timepoint was 8 control and 7 MSUS for PND8, 9 control and 8 MSUS for PND15, and 6 control and 6 MSUS for adult mice.

Data analysis: Sequenced data (FASTQ files) was assessed for quality using FastQC (version 0.11.8). Adapter sequences and short reads (length < 30bp) were removed, and bases with low quality ($Q < 30$) were trimmed using TrimGalore (version 0.6.4_dev). Trimmed data was pseudo-aligned with Salmon (version 0.11.2) with

automatic library type detection (-l A) on a transcript index prepared from a) the GENCODE annotation (version M18), b) piRNA precursors, and c) transposable elements (TEs) from repeat masker (concatenated by family). All downstream analysis was performed using R (version 3.6.0).

For differential gene expression analysis, all the quantified RNA species were aggregated at the gene level or at the repeat family level for TEs. Features with less than 15 reads in at least 40% of the samples were removed. Normalization was performed using the Trimmed Mean of *M*-values (TMM) method (edgeR package). Differential analysis was performed using the voom/dupCor methods of the limma R package (version 3.34.9) and the significance threshold was set at FDR ≤ 0.05 . Gene Set Enrichment Analysis (GSEA) was performed on GO Biological Processes using WebGestalt on all genes identified from PND15 RNA-seq ranked based on their t-scores (Liao et al., 2019c). Only terms containing less than 1000 genes were assessed, and an FDR ≤ 0.1 was used to account for multiple comparisons.

Differential Transcript Usage (DTU) was performed using IsoformSwitchAnalyzeR (Vitting-Seerup and Sandelin, 2017, 2019). Briefly, isoform switches were identified, isoforms were annotated via integration of a wide range of (predicted) annotations for the isoforms involved in the identified switches. The significance threshold for differentially used isoforms was set at FDR ≤ 0.05 . Pfam was used for prediction of protein domains, CPC2 (version 2.0) for calculation of the coding potential, SignalP (version 5.0) for prediction of Signal Peptides, and IUPred2A was used to predict Intrinsically Disordered Region (IDRs) and Intrinsically Disordered Binding Regions (IDBRs). The results from Pfam, CPC2, IUPred2A, and SignalP were used with IsoformSwitchAnalyzeR for annotation and prediction of functional consequences for the identified isoform changes.

Overrepresentation Analysis (ORA) was performed on candidate genes from DTU analysis using gProfiler (Raudvere et al., 2019). The query was run on *Mus Musculus* background, with GO terms of minimum 10 and maximum 1000 genes and an intersection size of minimum 3 genes. Benjamini-Hochberg correction at an FDR ≤ 0.05 was performed to account for multiple comparisons.

3.5.5 Omni-ATAC library preparation and data analysis

Library preparation: 25,000 spermatogonial cells from PND15 testis and 5,000 spermatogonial cells from adult testis were used as starting material for the Omni-ATAC protocol (Corces et al., 2017). Briefly, sorted cells were lysed in cold buffer and the cell nuclei were pelleted and transposed using Nextera Tn5 (Illumina) for 30 min at 37°C and 1000 rpm. Transposed DNA was purified using the MinElute Reaction Cleanup Kit (Qiagen), and the libraries were generated using the NEBNext High-Fidelity 2X PCR Master Mix (New England Biolabs). Primer dimers and fragments bigger than 1000 bp were removed using Agencourt AMPure XP magnetic beads (Beckman Coulter). Library quality was assessed using an Agilent High Sensitivity DNA Chip on the Bioanalyzer 2100 (Agilent Technologies). Libraries were sequenced on an Illumina HiSeq 4000 using v4 reagents and 100 paired-end read length. At PND15, 6 control and 6 MSUS samples were sequenced, and 5 control and 5 MSUS samples from adult testis.

Data analysis: FASTQ files were assessed for quality using FastQC (Andrews et al., 2012) (version 0.11.8). Quality control (QC) was performed using TrimGalore (Krueger, 2012) (version 0.6.2) in PE mode (--paired), trimming adapters, low-quality ends (-q 30) and discarding reads that become shorter than 30 bp after trimming (--length 30). Alignment on the GRCm38 genome was performed using Bowtie2 (Langmead and Salzberg, 2012) (version 2.3.5) with the following parameters: fragments up to 2 kb were allowed to align (-X 2000), entire read alignment (--end-to-end), suppressing unpaired alignments for paired reads (--no-mixed), suppressing discordant alignments for paired reads (--no-discordant) and minimum acceptable alignment score with respect to the read length (--score-min L,-0.4,-0.4). Using alignmentSieve (version 3.3.1) from deepTools (Ramirez et al., 2016) (version 3.4.3), aligned data (BAM files) were adjusted for the read start sites to represent the centre of the transposon cutting event (--ATACshift), and filtered for reads with a high mapping quality (--minMappingQuality 30). Reads mapping to the mitochondrial chromosome and ENCODE blacklisted regions (ENCODE accession ENCF547MET), were filtered out.

To call nucleosome-free regions, all aligned files were merged within groups (PND15 and adult), sorted, and indexed using SAMtools (Li et al., 2009) (version 0.1.19), and

nucleosome-free fragments (NFFs) were obtained by selecting alignments with a template length between 40 and 140 inclusively. Peak calling (identifying areas in a genome that have been enriched for transcription factors) on the NFFs was performed using MACS2 (Zhang et al., 2008) (version 2.2.7.1) with mouse genome size (-g 2744254612) and PE BAM file format (-f BAMPE). The downstream analysis was performed in R (version 3.6.2), using packages from CRAN (<https://cran.r-project.org>) and Bioconductor (Huber et al., 2015). The peaks were annotated based on overlap with GENCODE (Harrow et al., 2012a) (version M18) transcript and/or the distance to the nearest transcription start site. The number of extended reads overlapping in the peak regions was calculated using the csaW package (Lun and Smyth, 2015) (version 1.20.0). Peak regions which did not have at least 15 reads in at least 40% of the samples were filtered out. Normalization factors were obtained on the filtered peak regions using the TMM normalization method (Robinson and Oshlack, 2010) and differential analysis on the peaks (control vs MSUS at PND15 and at adult stage) was performed using the Genewise Negative Binomial Generalized Linear Models with Quasi-likelihood (glmQLFit) Tests from the edgeR package (Robinson et al., 2009) (version 3.28.1). Peak regions which had an absolute Log₂FC ≥ 0.5 and a p-value ≤ 0.005 were categorized as differentially accessible regions between MSUS and control groups for PND15 samples. Peak regions with an absolute Log₂FC ≥ 0.5 and a p-value ≤ 0.01 were categorized as differentially accessible between MSUS and control groups at adult stage.

GO enrichment analysis was performed using GREAT, using as a background all identified regions at PND15 and in adult stage respectively (McLean et al., 2010). TF motif enrichment was assessed for the differentially accessible regions at PND15 and at adult stage using MEME ChIP, in classic mode (Machanick and Bailey, 2011).

3.5.6 Proteomic analysis of blood plasma

Blood was collected in EDTA coated tubes (Microvette, Sarstedt) from n = 5 per group at PND8 and from n = 5 per group at PND15. Samples were stored at 4 °C for 1-3 hours and centrifuged at 2,000g for 10 min. Plasma was collected and stored at -80 °C until processed for proteomics analysis. Pups from different litters were chosen to avoid potential litter effects.

Plasma samples were enriched for small regulatory proteins using protein depletion columns (Seppro Mouse, Sigma Aldrich), and quantification was performed using the Qubit protein assay kit (ThermoFisher Scientific) according to manufacturer's instructions. Samples were purified by TCA precipitation and processed with a filter-assisted sample preparation protocol (FASP). First, 20 µg of protein was resuspended in 30 µl SDS denaturation buffer (4% SDS (w/v), 100 mM Tris/HCl pH 8.2, 0.1 M DTT), and incubated at 95 °C for 5 min. Samples were further diluted with 200 µl UA buffer (8 M 25 urea, 100 mM Tris/HCl pH 8.2) and spun at 35 °C at 14000 x g for 20 min in regenerated cellulose centrifugal filter units (Microcon 30, Merck Millipore). Samples were washed once with 200 µl UA buffer and spun again at 35 °C and 14000 x g for 20 min. Cysteines were blocked with 100 µl IAA solution (0.05 M iodoacetamide in UA buffer) and incubated for 1 min at room temperature (RT) at 600 rpm followed by a centrifugation at 35 °C and 14000 x g for 15 min. Filter units were washed with 100 µl of UA buffer 3x and then 2x with 0.5 M NaCl. Proteins were digested overnight at room temperature with a 1:50 ratio of trypsin (0.4 µg) in 130 µl TEAB (0.05 M Triethylammonium bicarbonate in water). After protein digestion, the peptide solutions were centrifuged at 35 °C and 14000g for 15 min and acidified with 3 µl of 20% TFA (Trifluoroacetic acid). Peptides were cleaned using Sep-Pak C18 silica columns (Waters Corporation) activated with 10 ml methanol and washed with a solution 1 ml of 60% ACN (acetonitrile) and 0.1% TFA. Columns were equilibrated with 3 x 1 ml of 3% ACN 0,1% TFA. The samples were diluted in 800 µl of 3% ACN 0.1% TFA and loaded onto the silica columns, then washed with 4 x 1 ml 3% ACN 0.1% TFA and eluted with 60% ACN 0.1% TFA. Samples were lyophilized in a vacuum concentrator and resolubilized in 19 µl 3% CAN 0.1% FA (Formic acid). 1 µl of synthetic peptides (Biognosys AG) were added to each sample for retention time calibration. Peptides were analysed using LCMS/MS (Orbitrap FusionTM TribidTM MS) at the proteomics facility of the Functional Genomics Centre (FGC) Zurich. Raw data was quantitatively and qualitatively analysed by Mascot and Progenesis QI platforms.

3.5.7 RT-qPCR and western blot analyses

RT-qPCR: Testes from PND15 pups were dissected and RNA extracted using a phenol/chloroform extraction method (Trizol, Thermo Fischer Scientific). Reverse transcription was performed using the SuperScript III RT kit (Qiagen) according to

manufacturer's instructions. Reactions were run in triplicate on a Light Cycler II 480 (Roche) using SYBR Green I Mastermix (Roche). Primers for *Rage*, *Nrf2* and *Tfam* were designed using Primer Blast (NCBI). Melting curve analysis following the reaction confirmed the specificity of our primer pairs. *Tubd1* was used as an endogenous control.

Western blot: Western blot validation was performed on PND15 blood plasma from an independent breeding than the one used for proteomics analysis (n = 5 control and 5 MSUS samples). Plasma was collected from trunk blood in EDTA coated tubes (Microvette, Sarstedt). Samples were stored at 4 °C for 1-3 hours and centrifuged at 2,000g for 10 min. Raw blood plasma samples containing 5 ug protein were diluted in water and denatured using 4x Laemmli Sample Buffer (Bio-Rad, Hercules, CA, USA) and 10% of mercaptoethanol for 5 min at 95 °C. Samples were then loaded onto 4%–20% gradient Mini-PROTEAN TGX gel (Bio-Rad, Hercules, CA, USA), and separated at 70 V for 30 min and then 140 V for 1 hr using SDS running buffer (25 mM Tris, 192 mM Glycin, and 0.1% SDS). Proteins were transferred using a Trans-BlotTurboBlotting System and a Trans-BlotTurboMini PVDF transfer pack (both from Bio-Rad, Hercules, CA, USA). Membranes were blocked using 5% of non-fat dry milk in TBS-Tween for 1 hr at room temperature (RT), and they were incubated overnight at RT with 5ml 0.9ug/ml anti-HMGB1 primary antibody (Abcam, ab18256). Incubation with the secondary antibody was performed for 1hr at RT, using 1:10,000 IRDye 800CW anti-Rabbit IgG (926-32211, LI-COR, Lincoln, NE, USA). Membranes were washed with TB-Tween 3x for 10 min between incubations. Imaging of the membranes was performed using an Odyssey Infrared Imaging System (LI-COR, Lincoln, NE, USA) and quantified using ImageJ v1.41.

3.5.8 Quantification and data analysis

Initial sample size was estimated based on our previous experiments using the MSUS paradigm (Gapp et al., 2014, 2018; Steenwyk et al., 2019). For behavioural analysis RT-qPCR and western blot, a two-tailed Student t test was used to assess statistical significance between control and MSUS groups, with statistical significance set at $p \leq 0.05$. ROUT test was employed to identify statistical outliers and exclude them from the analyses. All statistics of behavioural, RT-qPCR and western blot analyses were computed with Prism 8. Figures in this study were generated using Prism8 and base

plotting in R. All reported replicates were biological replicates and representative of 1 animal/replicate, except PND8 and PND15 replicates, where 2 pups were pooled for each sample.

3.6 Acknowledgements

We thank Yvonne Zipfel for animal care in Zurich, Silvia Schelbert and Alberto Corcoba for taking care of the animal licenses and lab organization in Zurich. We thank Alberto Corcoba for critical reading of the manuscript. We thank Catherine Aquino and Emilio Yángüez from the FGC Zurich for library preparation, support and advice with sequencing and library preparation. We thank Jon Oatley, Melissa Oatley, Tessa Lord and Nathan Law for advice and conceptual feedback, as well as for providing detailed protocols of the testis dissection and preparation and for the immunocytochemistry of spermatogonial cells.

3.7 Authors contributions

ILC and IMM designed and conceived the study. ILC performed all RNA-seq, ICC and ATAC-seq experiments, with significant support from MR and NG. GvS performed the proteomics and western blot experiments. DKT analysed the RNA-seq and ATAC-seq data, with significant support from PLG. ILC, GvS and DKT prepared the figures. ILC interpreted the data with significant input from GvS and IMM. ILC wrote the manuscript, with significant input from GvS, PLG and IMM. All authors read and accepted the final version of the manuscript.

3.8 Supplemental figures

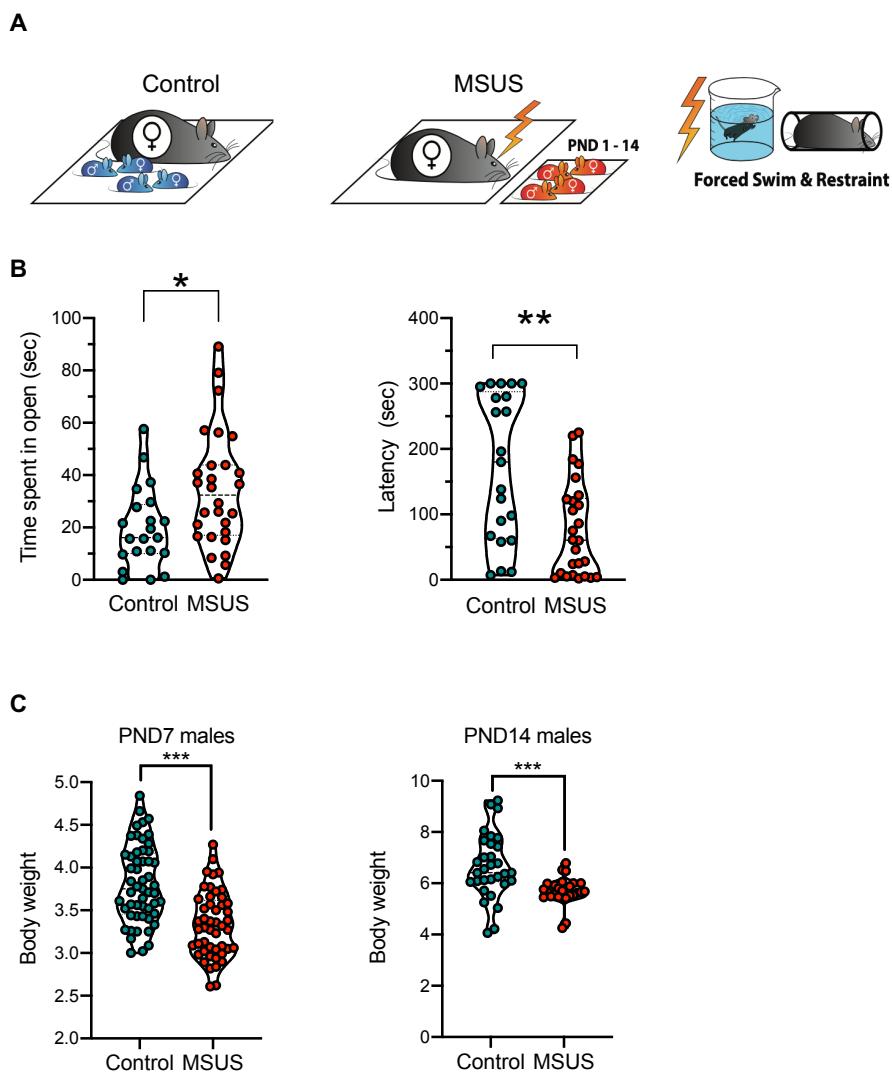


Figure S 3-1 MSUS induces behavioural and metabolic alterations in exposed males.

(A) Schematic representation of the MSUS paradigm (see Methods section for detailed description of the paradigm). All rights for the figure belong to Martin Roszkowski and were used with his consent;

(B) MSUS adult males exhibit increased time spent in open arm and decreased latency to enter the open arm, indicative of risk-taking behaviours, when tested on the elevated plus maze (EPM), n = 22 control and 28 MSUS males tested. ROUT test was used to remove statistical outliers. Experiment was performed in single-blind;

(C) MSUS pups exhibit lower weight at PND7 and PND14 ($n = 53$ control and 48 MSUS at PND7, $n = 31$ control and 27 MSUS at PND14).

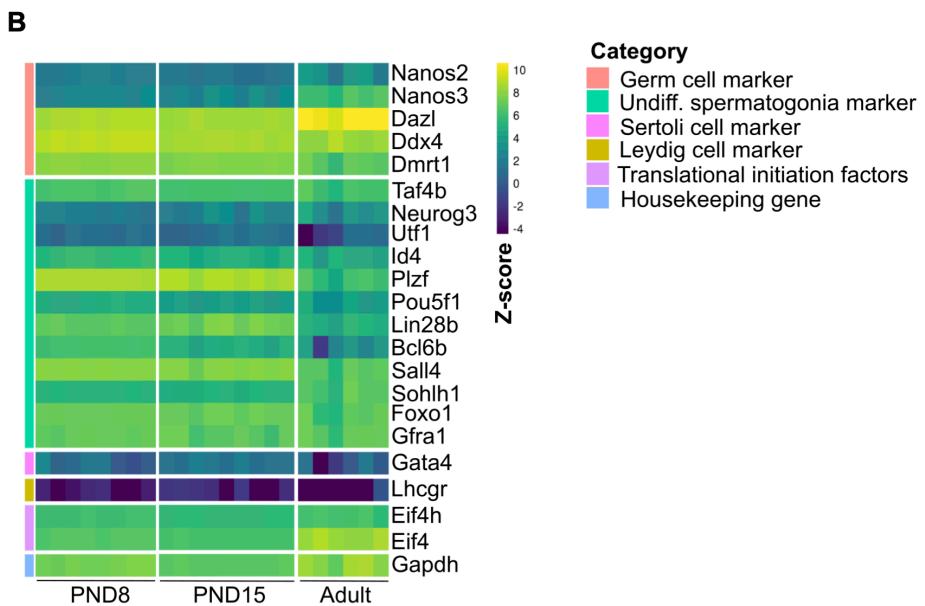
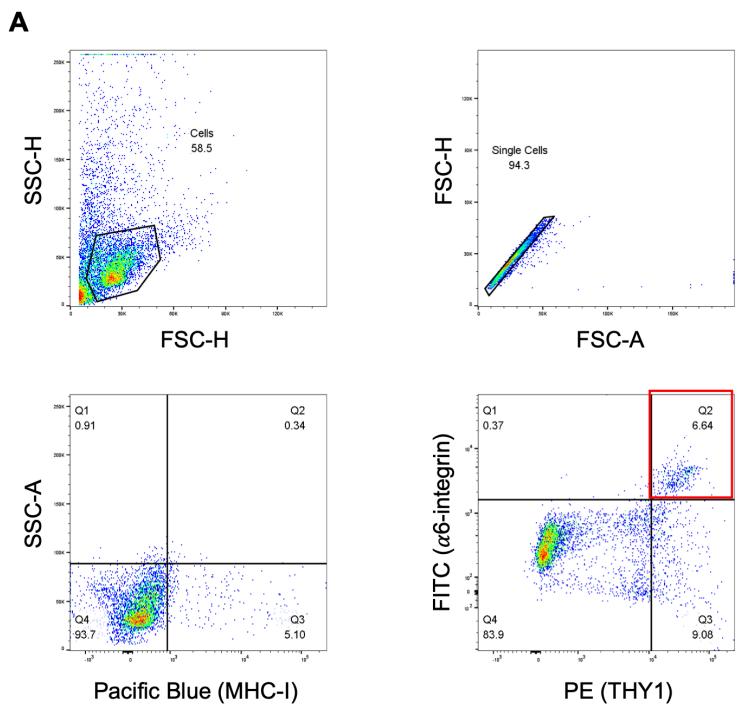


Figure S 3-2 FACS strategy for enrichment of spermatogonial cells from adult testis.

(A) Testicular cell suspension was passed through an 85um nozzle following enzymatic digestion. Representative dot plots for sorting strategy of adult spermatogonial cells. Following gating for live and single cells based on forward/side scatters, MHC⁺/ α 6-integrin⁺/THY1⁺ cells were sorted (Pacific Blue⁻/FITC⁺/PE⁺). Cells in red quadrant represent the sorted spermatogonial cells;

(B) Heatmap of representative gene markers mRNA expression in the sorted spermatogonial cell populations at PND8, PND15 and at adult stage. Genes characteristic to germ cells, stem cell potential and somatic cells were chosen to verify the enrichment procedure using FACS.

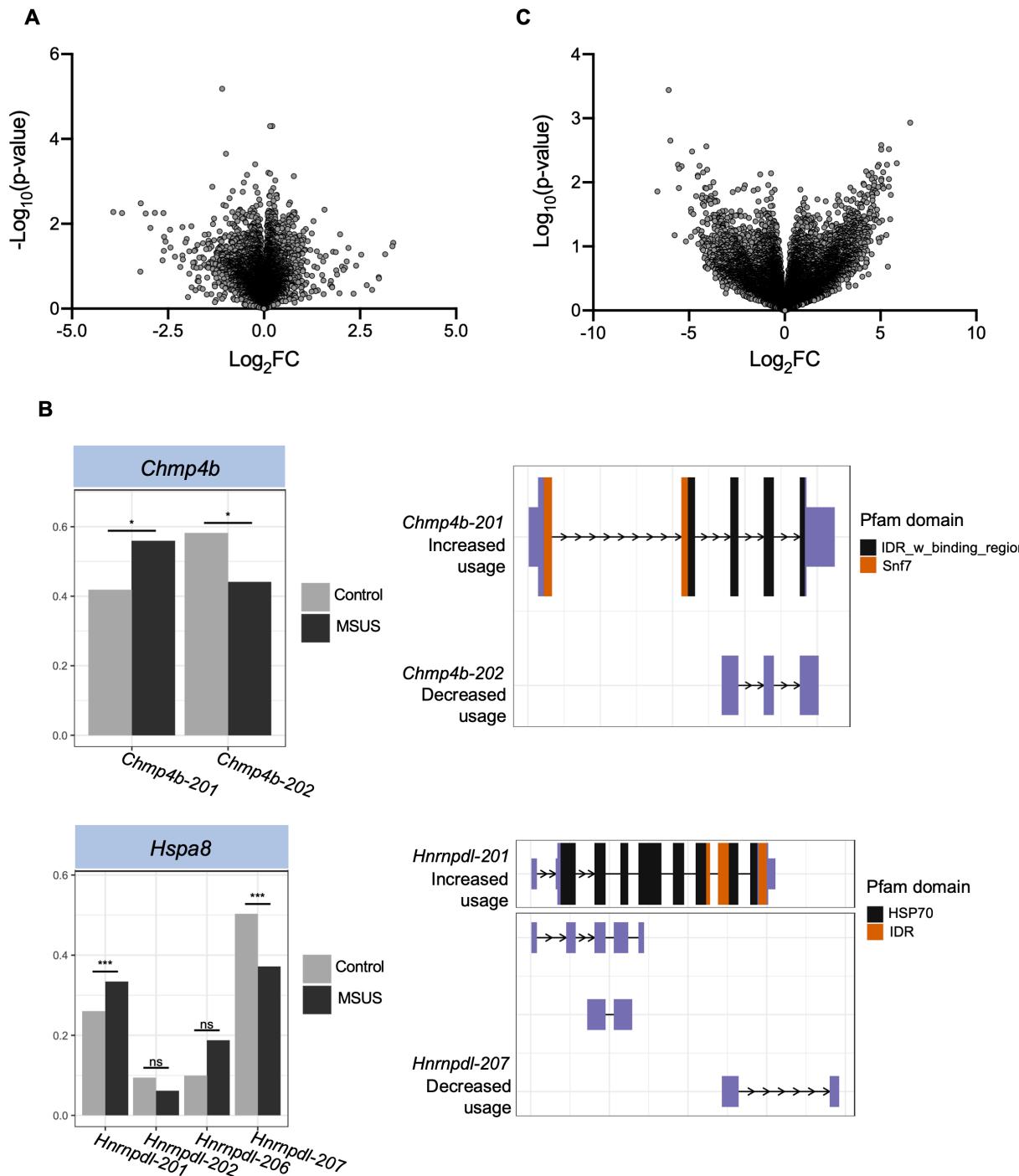


Figure S 3-3 RNA-seq reveals differential transcript usage in early postnatal spermatogonia but no changes at gene level.

(A) Differential gene expression analysis at PND8 reveals no significantly altered genes (FDR > 0.05, n = 8 control and 9 MSUS samples);

(B) Differential transcript usage of *Chmp4b* and *Hspa8*, in spermatogonial cells from PND8 MSUS and control pups (* FDR ≤ 0.05, *** FDR ≤ 0.001). Splicing plots shown are extracted directly from the IsoformSwitchAnalyzer and Ensembl transcript IDs are used for each transcript annotation. Plots on the right display the predicted Pfam protein domains for each of the gene transcripts;

(C) Differential gene expression analysis in adult spermatogonia from control and MSUS males reveals no significantly altered genes (FDR > 0.05, n = 6 samples / group).

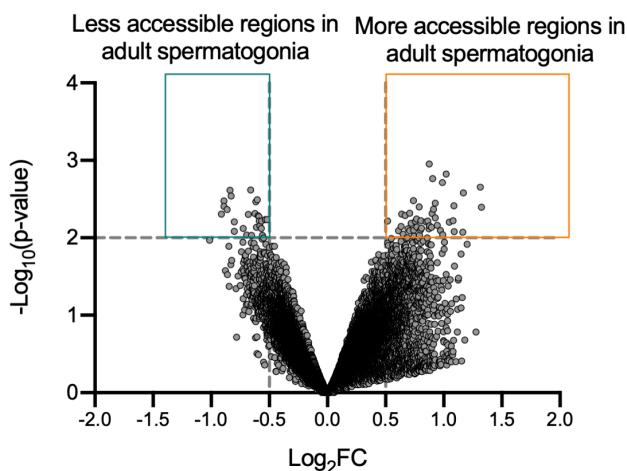
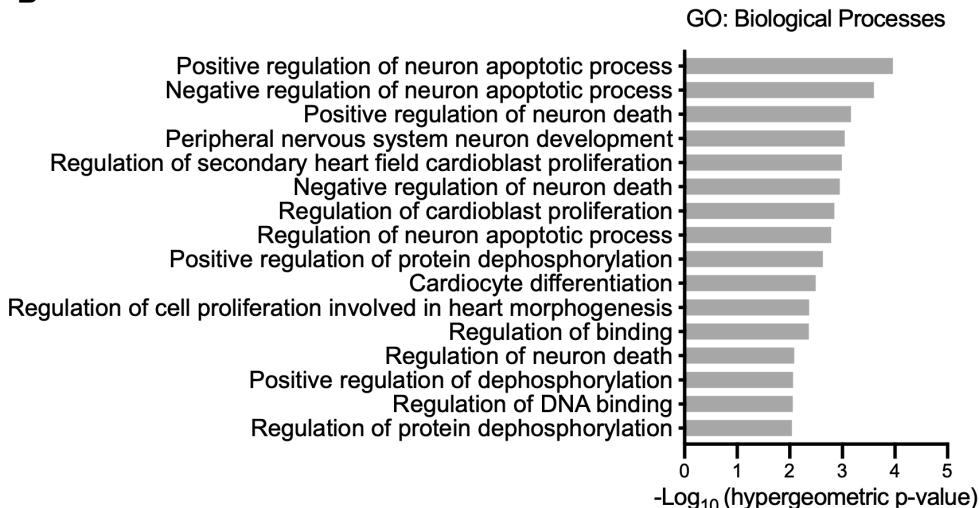
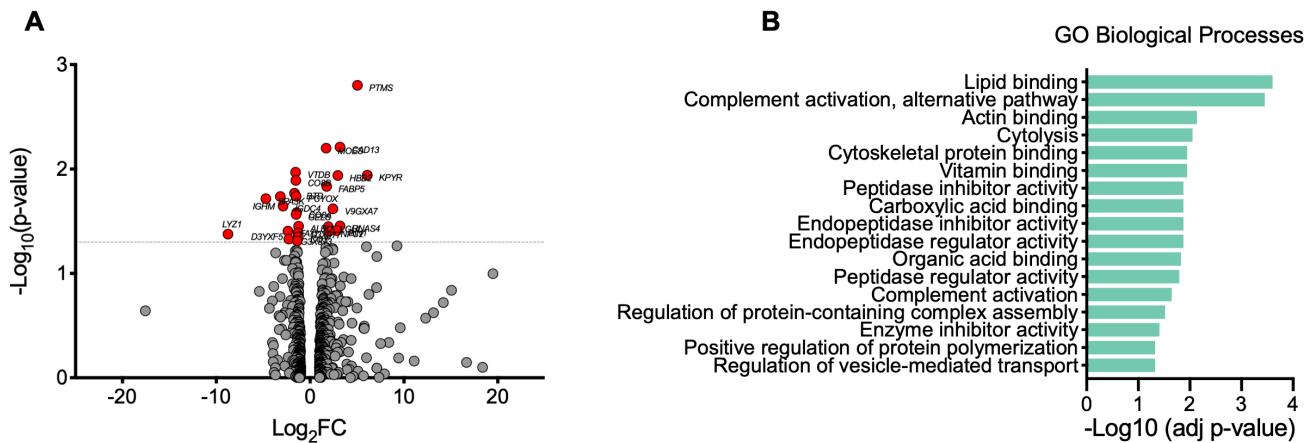
A**B**

Figure S 3-4 Adult spermatogonial cells from MSUS males display altered chromatin accessibility at regions linked to nervous system development.

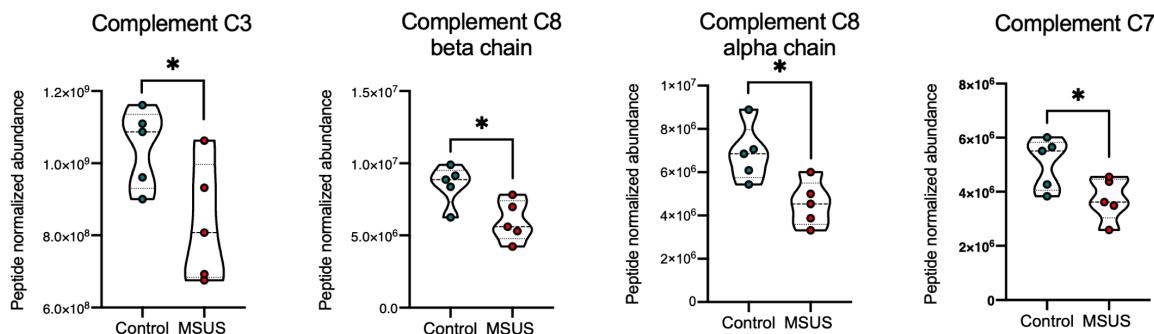
(A) Differential accessibility analysis reveals regions of increased and decreased accessibility in adult spermatogonial cells from MSUS males compared to control animals ($\text{Abs Log}_2\text{FC} > 0.5$, $\text{p-value} \leq 0.01$, $n = 6$ samples / group);

(B) GREAT analysis of differentially accessible regions reveals nervous system development related terms among the most enriched ones ($p \leq 0.05$, minimum 10 and maximum 1000 genes per GO term included, biological processes associated with minimum 3 regions were considered).



C

Complement cascade



Lipid binding

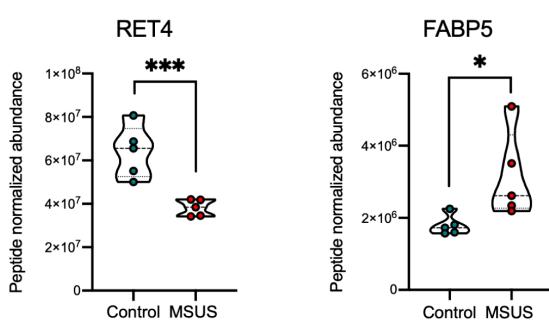


Figure S 3-5 Altered proteins in plasma samples from MSUS pups are associated with innate immune response and lipid binding pathways.

(A) Volcano plot showing $\log_2\text{FC}$ of protein expression in PND8 exposed pups compared to controls. Proteins with significantly different expression between the two groups are highlighted in red ($\log_2\text{FC} \geq 1$, $p\text{-value} \leq 0.05$, $n = 5$ biological samples / group);

(B) Overrepresentation analysis on candidate proteins with differential expression between the 2 groups at PND8 (Benjamini-Hochberg corrected $p\text{-value} \leq 0.05$, term size set to min 10 and max 1000 genes, min 3 genes intersection). The complete list of GO terms is available in Table S1;

(C) Peptide normalised abundance of exemplary proteins associated with the most enriched GO terms (* $p \leq 0.05$; *** $p \leq 0.001$).

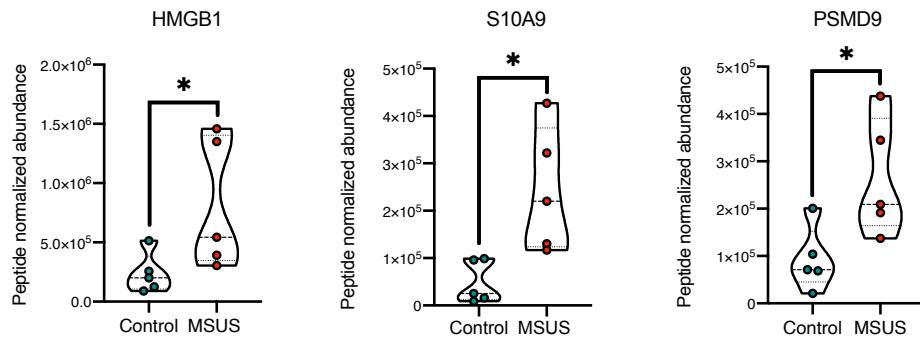
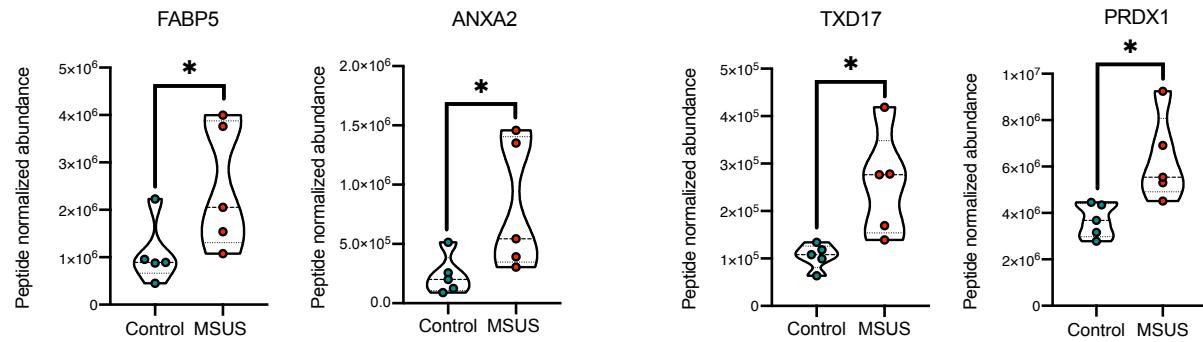
A**Inflammatory proteins****B****Lipid-binding proteins****C****Redox proteins**

Figure S 3-6 Inflammation, lipid-related and redox proteins are upregulated in plasma of MSUS pups compared to controls.

(A - C) Peptide normalised abundance of exemplary proteins altered in plasma samples from PND15 MSUS pups (* = p ≤ 0.05).

4 Conclusions and Outlook

This thesis provides evidence that the chromatin accessibility landscape of mouse spermatogonial cells exhibits a dynamic reorganization between early postnatal and adult stage. To gain a better understanding of the distinct gene expression – epigenome relationship in pup and adult spermatogonial cells, these results were complemented by transcriptome data and histone and DNA methylation profiles. In addition, the immediate and long-lasting effects of an exogenous insult, chronic stress exposure, applied during early postnatal spermatogonial cell development, were investigated.

In the following section I will summarize the results obtained in **Chapter 2** and **Chapter 3** and discuss the rationale for our approach, with the challenges and limitations that it encompasses. I will also propose further experiments that could be performed to test the causal link between chromatin reorganization and gene expression changes. Finally, I will discuss the overall contribution of these results for fully understanding spermatogonial cell regulation across postnatal life, and in response to an environmental insult.

4.1 Challenges and limitations

The enrichment method used in this thesis relies on a multiparametric FACS strategy, established by Brinster and collaborators following serial transplantation experiments (Kubota et al., 2004b, 2004a). Accordingly, spermatogonial cells from pup and adult male testis were sorted based on their MHC-I-Thy1⁺αV-integrin^{-/Dim}α6-integrin⁺ surface profile, as this cell population was reported to contain the majority of stem cells of the spermatogonial cell population across ages (Kubota et al., 2004b, 2004a; Oatley and Brinster, 2012; Zhang et al., 2016a). To assess the enrichment proportion in our sorted cell populations, we stained our sorted and unsorted cell samples using PLZF, a widely used undifferentiated spermatogonial marker. We found that, on average, more than 90% of our sorted cells were PLZF⁺, confirming the efficiency of our sorting strategy. Furthermore, genes which are known to be expressed by undifferentiated spermatogonial cells with stem cell potential (*Id4*, *Etv5*, *Lhx1*, *Tspan8*, *Gfra1*) were highly expressed in our enriched cell populations, while differentiating spermatogonial markers (*Stra8*, *Kit*, *Ngn3*) and somatic cell markers (*Vim*, *Gata4*) were very lowly

expressed or not expressed at all. However, as FACS is able to enrich, but does not completely purify spermatogonial cells from a testicular cell suspension, we cannot exclude the influence of contaminating cells on the interpretation of our findings. Additionally, our approach omits potential differences which could exist within the pool of undifferentiated spermatogonial cells, between A_s, A_{pr} and A_{al} subpopulations.

A second limitation of our approach is the multistep protocol employed for sorting the undifferentiated spermatogonial cells from adult testis. The adult testis cellular milieu is highly complex, as it encompasses abundant populations of differentiating spermatogenic cells such as spermatocytes and spermatids, as well as several types of somatic cells (Sertoli cells, Leydig cells, interstitial cells, resident macrophages). Therefore, a double enzymatic digestion and several rounds of sedimentation and washing steps need to be performed in order to remove most of the other germ and somatic cells. Furthermore, a centrifugation step in 30% Percoll is also employed, to concentrate spermatogonial cells prior to FACS. All of these lengthy steps invariably lead to increased cell loss compared to the straightforward sorting protocol employed from pup testis.

Lastly, slight differences in the age of the mice from which the cells have been collected exist in the comparison between our chromatin accessibility profiles with literature transcriptome, histone and DNAm datasets. Our pup ATAC-seq libraries have been generated using PND15 pups, whilst the literature RNA-seq and DNAm data were obtained from PND14 pup spermatogonia. However, previous reports found similar transcriptome and histone modifications in PND12 and PND14 spermatogonial cells (Hammoud et al., 2015), suggesting a similar genomic landscape in closely timed juvenile spermatogonia. Additionally, by the end of the second week of postnatal life the maturation of Sertoli cells within the seminiferous tubules is complete, so that differences between PND14 and PND15 spermatogonia imposed by the surrounding niche should be minimal compared to the fast growth taking place earlier on (Flickinger, 1967; Mruk and Cheng, 2015).

For investigating the chromatin accessibility landscape of adult spermatogonial cells, we have used PNW20 mice and the available literature datasets for the RNA-seq, ChIP-seq and DNAm were generated from PNW8 males. During adulthood and up to 1 year of age, differences in spermatogonial cell numbers and colonization potential following transplantation were found to be minimal, suggestive of a steady-state adult

spermatogonial population (Kubota and Brinster, 2018). In conclusion, we chose the available literature datasets which best matched our cell sorting, library preparation and sequencing strategies, and for which age-driven differences in molecular characteristics of spermatogonia are known to be minimal.

4.2 Postnatal dynamics of spermatogonial cells

4.2.1 Open chromatin reorganization in postnatal spermatogonia

To our knowledge, this is the first time that chromatin accessibility between early postnatal and adult spermatogonia has been compared, as well as the first attempt to characterize chromatin accessibility landscape in adult spermatogonial cells. We found more than 3000 differentially accessible chromatin regions between PND15 and adult spermatogonial cells, indicative of distinct chromatin accessibility landscapes between pup and adult stages. Although only a minority of these regions were situated in proximity to a gene TSS, these regions displayed enrichment for numerous TF motifs, particularly in intergenic regions and introns. As discussed in more detail in Chapter 2, specific TF families were associated with the chromatin regions which displayed an increase or a decrease in accessibility. Motifs for TFs which belong to the Fos/Jun family, nuclear receptors (RXRs, RAR, RFXs) and pluripotency factors USF1/2 and POU3F1, were enriched in more accessible chromatin regions, whilst regions of decreased accessibility in adult spermatogonia revealed enrichment for TFs with widely known developmental roles.

Although open chromatin represents only a minority of the total chromatin content of a cell (approximately 2-3%), it usually binds most of the transcription factors with roles in gene expression regulation (Klemm et al., 2019). We hypothesise that the high number of differentially accessible regions we find in open chromatin is a strong indicator of the distinct gene networks employed by spermatogonial cells in early postnatal stage and in the adult testis. Indeed, we found that more accessible regions were associated with gene pathways important for RNA processing, stem cell maintenance and spermatogenesis, whilst regions which were less accessible in adult spermatogonia were mainly associated with developmental pathways.

By performing TF motif enrichment at distinct genomic locations, we are able to find a clear preference of NF-Y and LHX motifs for intergenic regions. At present, the role of

NF-Y TFs in spermatogonial cells hasn't been characterized, raising exciting new questions about the involvement of this TF family in the germline. The presence of NF-YA and NF-YB motifs was also recently described in open chromatin regions in human spermatogonial cells, where they co-localized with more accessible LTR subtypes, indicative of their potential regulatory role in the human germline (Guo et al., 2017). Further ChIP-seq experiments for NF-Y TFs would allow to specifically identify and investigate their target regions and genes in mouse spermatogonial cells.

4.2.2 Unique chromatin landscape of genes differentially expressed in early postnatal and adult spermatogonia

The majority of differentially accessible regions from PND15 to adult spermatogonia resided distally from TSSs. However, we found that a number of regions were situated +/- 2.5 kb from the TSS of a gene. We augmented the chromatin accessibility data with RNA-seq and ChIP-seq datasets from previously published reports, and revealed novel patterns of chromatin accessibility and histone modifications at genes with differential expression between early postnatal and adult spermatogonia.

To our surprise, regions which increased in chromatin accessibility in adult spermatogonia did not only overlap genes with an upregulated expression, but also a comparable number of genes which were downregulated in adult spermatogonia. Genes in the latter category were developmental factors, some of which exhibited enrichment for the repressive H3K27me3 or the bivalent H3K4me3/HeK27me3. The presence of a more accessible chromatin region proximal to a gene which decreases in expression may indicate that active repression mechanisms are in place in adult spermatogonial cells. In contrast, the genes with upregulated mRNA expression and marked by an increase in chromatin accessibility in adult stage were important for spermatogonial stem cell maintenance, and were generally depleted of the inactivating H3K27me3.

Genes important for spermatogonial cell population establishment in the neonatal testis were downregulated in adult spermatogonia, and were marked by a decrease in chromatin accessibility in proximity to their TSS. Collectively, our combinatorial strategy revealed novel chromatin-associated mechanisms which may be used by spermatogonial cells in order to transition from an early postnatal gene expression landscape to the adult, steady-state phenotype.

4.2.3 Chromatin accessibility landscape at transposable elements

One major finding of this thesis is that open chromatin undergoes reorganization at distinct TEs, in the transition from PND15 to adult stage spermatogonial cells. An increasing number of studies suggest specific regulatory roles of certain TE subtypes in transcription regulation (Chuong et al., 2017; Davis et al., 2017; Fort et al., 2014; Lu et al., 2020; Sundaram and Wysocka, 2020). In mESCs for example, LTRs, specifically ERVks, have been shown to contribute with binding substrates to TFs important for pluripotency potential such as OCT4 and NANOG (Fort et al., 2014). Among the identified ERVK subtypes, BGLII, RLTR9E and RLTR17 were the most significantly enriched, and the highest expressed. Another recent study reported that there is a non-random association between certain TE families and distinct gene categories in mESCs (Lu et al., 2020). For example, LINE L1 elements were specifically enriched at olfactory receptor genes, odorant binding and pheromone receptors, whilst SINEs were mainly associated with housekeeping genes.

In the light of these recent findings, we investigated the possibility that TEs could play a role in the differential transcriptome landscape of early postnatal and adult spermatogonia. We assessed chromatin accessibility at all annotated TEs across the genome, by using our ATAC-seq data from PND15 and adult spermatogonial cells and revealed more than 100 differentially accessible TE subtypes. Interestingly, we observed that the majority of differentially accessible TEs belonged to the LTR and LINE families. Most of the LTRs became less accessible in adult spermatogonia, with ERVks subtypes being by far the most abundant, whilst LINE L1 subtypes increased in accessibility in adult spermatogonia.

We next focused our attention on TE loci from the differentially accessible subtypes which overlapped TSSs of genes, or were situated in their proximity (less than 5 kbp up or downstream). As such, we identified several less accessible RLTR17 loci situated at the TSS of pluripotency factors from the *Platr* family (pluripotency associated transcripts). Some of these genes, *Lncenc1* (*Platr* 18), *Platr27* and *Platr14* respectively, also displayed a decrease in expression between pup and adult spermatogonial cells. In contrast, more accessible LINE L1 subtypes were situated at genes with specialized function. Several loci belonging to L1Md_2 and L1_Mus4 were situated in proximity of *Oifr* genes with increased expression in adult spermatogonia.

In addition, the differentially accessible TEs were also marked by a high number of enriched TF motifs, particularly at less accessible ERVKs and more accessible LINE L1.

Collectively, our results suggest that early postnatal and adult spermatogonia could employ distinct TE-associated mechanisms in the regulation of certain gene categories. This would assume that certain TEs escape silencing in the germline, and undergo chromatin reorganization with consequences for pluripotency program or certain specialized gene functions, as described above. Of note, one study found that specific ERVK subtypes, specifically LTR12C, LTR12D and LTR12E, were hypomethylated and more accessible in human spermatogonial cells compared to mESCs (Guo et al., 2017). These TEs were also enriched in NF-YA/B motifs (Guo et al., 2017), suggesting that TEs could play more important regulatory roles in the mammalian germline, than previously believed. How this relates to the transition from early postnatal to adult stage is a question that needs further investigation. We hypothesize that TEs not only display cell type-specific expression in germ cells during spermatogenesis, as recently shown using scRNA-seq of the whole human testis (Guo et al., 2018), but also have age-specific regulation within certain germ cell populations such as spermatogonia. This could in turn have consequences on their target gene expression.

Arguably, the correlation between chromatin reorganization at TEs and the change in expression of their nearby genes should be complemented by further functional experiments. To test the effects of chromatin reorganization at TEs on gene expression in spermatogonial cells, we propose to perturb chromatin accessibility using a Cas9-mediated reprogramming system. The most attractive candidates for such a follow-up experiment are the more accessible LINE L1 loci identified in the vicinity of *Oifr* genes. By using the dCas9-KRAB complex, a decrease in chromatin accessibility at these genomic sites can be achieved, either in cultured primary spermatogonial cells, or using a spermatogonial cell line (Gilbert et al., 2013; Thakore et al., 2015). As a read-out, the consequences on *Oifr* gene expression can be assessed. Similarly, an activating dCas9-SunTag or dCas9-VPR system could be employed to target the less accessible RLTR17 loci situated in the vicinity of pluripotency associated genes from the *Platr* family (Chavez et al., 2015; Gilbert et al., 2013; Tanenbaum et al., 2014). In this case, evaluating spermatogonial cell

proliferation and pluripotency potential could also be considered as a phenotypic read-out.

4.3 Implications of spermatogonial cell vulnerability to environmental insults

Mounting evidence suggests that exogenous factors can affect the sperm epigenome, and contribute to offspring's susceptibility to disease (Anway et al., 2005; Golding et al., 2017; Lambrot et al., 2013; Öst et al., 2014; Rassoulzadegan et al., 2006). However, adult sperm formation is initiated and maintained by spermatogonial cells in the testis. Therefore, environmentally-driven alterations to the transcriptome or epigenome of the spermatogonial cells could in fact precede, and even contribute, to the remodelling of sperm's epigenetic blueprint.

The number of studies focusing on the impact of environmental factors on spermatogonial cells is relatively small. Most of the existent evidence revealed increased apoptosis of spermatogonial cells in response to high doses of toxicant exposure, in some instances with overall consequences for sperm production (Vrooman et al., 2015). However, effects of diet or chronic stress on spermatogonial cells have so far been understudied, despite the fact that they are more likely to happen in real life and are closer to physiological circumstances (Hirano et al., 2014; Yazawa et al., 1999).

In this thesis we provide evidence that chronic stress exposure in early postnatal development is able to alter the transcriptome and chromatin accessibility landscape of spermatogonial cells. We hypothesized that the first 2 weeks of postnatal development could also underly a period of spermatogonial cell vulnerability, in which changes to the transcriptome and the epigenome may occur. This is because within this timeframe, the developing testis displays unique features, such as the lack of a blood-testis-barrier, and an ongoing maturation and proliferation of the Sertoli cells, and of other testis cells within the interstitial space (Flickinger, 1967; Heinrich and DeFalco, 2019).

4.3.1 Importance of isoform usage quantification

Our results indicate that the transcriptome of spermatogonial cells undergoes alterations not at the gene expression level, but in the transcript usage of genes, which was detectable after the first week of chronic stress. These genes mainly associated

with RNA processing, ribonucleoprotein complex assembly and translation. Genetic variation impacting isoform distribution of a gene can take place naturally, without causing any detrimental effects (Pai and Luca, 2019). However, in case of environmental insults, a more pronounced effect on the isoform distribution of a gene could potentially lead to dysregulation of the signalling pathways that the gene product is important for. For example, an isoform switch between a non-protein coding and a protein coding transcript of a gene involved in splicing or chromatin remodelling has the potential to disturb these cellular processes and lead to disease phenotypes.

Recently, several other studies have assessed not only the differential gene expression, in which a gene is quantified based on all expressed transcripts, but also the dynamics of individual isoform usage (Coelho et al., 2020; Gandal et al., 2018). Gandal and collaborators performed RNA-seq from frontal cortical samples of patients and identified 64 genes with DTU between autism spectrum disorder (ASD) and schizophrenia patients. Of note, only a small proportion of these genes also displayed differential expression at the gene level or at transcript level (Gandal et al., 2018). Similarly, Marques Coelho and collaborators investigated gene expression and isoform usage in the frontal and temporal lobes of patients with Alzheimers disease (AD) and found only a small proportion of the genes with differential expression to overlap with the genes found to exhibit DTU (Coelho et al., 2020).

These recent findings point to a complementary role of isoform switches to the disease phenotype, alongside genes with differential expression, and underline the importance of assessing transcriptome changes at different organization levels, including gene expression, transcript expression and transcript usage. Although advances in bioinformatic tools are already enabling the investigation of isoform usage and splicing differences from short-read RNA-seq, the certainty of these findings is still debatable. Therefore, employing long-read and single-molecule sequencing platforms such as Pac-Bio would allow a conclusive assessment of these potential transcriptome alterations. Alternatively, we are planning to experimentally validate our isoform usage findings by performing RT-qPCR using isoform-specific primers, and calculate the ratios between the isoform pairs in naïve and exposed groups, similarly to how it is described in (Gandal et al., 2018).

4.3.2 Open chromatin reorganization after stress exposure

Considering the dynamic regulation of transcription that open chromatin ensures throughout the genome, we sought to investigate if potential changes at the chromatin level in spermatogonial cells could result from exposure to early life stress. Immediate and long-lasting effects of stress exposure on chromatin accessibility were assessed from Omni-ATAC data analysis of PND15 and adult spermatogonial cells from control and MSUS males.

As discussed in more detail in Chapter 3, open chromatin reorganization was observed predominantly at PND15, indicative of immediate effects on chromatin following environmental exposure. We found that the differentially accessible regions targeted distinct biological pathways. More accessible chromatin regions were enriched in stress-response pathways related to MAPK and mTOR cascades and metabolic pathways, whilst less accessible regions associated with pathways of the immune system. These observations open up significant follow-up question on how these chromatin accessibility changes could impact the subsequent developing germ cells within the testis.

First, spermatogonial cells differentiate into spermatocytes and spermatids during spermatogenesis. Each of this germ cell stage is characterized by distinct transcriptome and chromatin transitions, as recently reported (Green et al., 2018; Hermann et al., 2018; Maezawa et al., 2018). Therefore, it would be desired to perform a stage-specific characterization of the open chromatin regions potentially affected by early life stress in each of these germ cell types.

Second, adult sperm only retains about 7.5% histones following the histone-to-protamine exchange that takes place at the spermatid stage (Gold et al., 2018). However, the retained histone regions are situated in enhancer regions, as defined by the presence of histone marks from other adult or embryonic tissues (Gold et al., 2018; Jung et al., 2017). Investigating chromatin accessibility in sperm from males exposed to MSUS, and direct comparison with the changes identified in spermatogonial cells, is an essential step for elucidating the consequences of our current findings on fertilization-competent sperm cells.

Third, spermatogonial cells encompass the only stem cell population in the adult body which directly contributes information to the next generation. Modifications to their

transcriptome, epigenome or chromatin landscape have the potential to not only impact the generation exposed directly to the environmental stimuli, but also to be carried on to the next generation. Linking the spermatogonial alterations with sperm and embryo analyses, ideally from the same lineage, is essential to investigate the mechanisms of intergenerational transmission of acquired traits.

4.3.3 Systemic proteomic alterations and their effects on the testis

Previous findings from our lab have revealed that systemic alterations in serum metabolites from blood of pup and adult MSUS males correlate with changes in sperm long and small non-coding RNAs (Gapp et al., 2018; Steenwyk et al., 2019). However, sperm is situated in the lumen of the seminiferous tubules, with limited access to systemic factors from the blood. In contrast, spermatogonial cells lie outside the BTB, in close vicinity of testis vasculature and the interstitial cells, and are therefore more likely to be influenced by perturbations in circulating factors reaching the testis (Yoshida et al., 2007).

Our plasma proteomic analyses at PND8 and PND15 revealed a considerable number of inflammatory and redox proteins altered between control and MSUS pups. Furthermore, we found an increased expression of the corresponding receptors of some of these proteins in PND15 testis of MSUS pups, suggesting that alterations in testis gene expression may be a result of the altered plasma proteome. It remains to be further investigated how these findings associate with the enrichment in biological pathways related to immune function and stress response that we found in spermatogonial cells of MSUS pups. Studies in *Drosophila* and *C. elegans* have already suggested that inter-tissue communication has the potential to alter germline homeostasis, and that exogenous exposures such as diet can initiate such alterations (Ables et al., 2012; Nono et al., 2020). However, bringing mechanistic evidence for environmentally-induced inter-tissue alterations in mammals remains challenging, mainly due to the higher complexity of the biological networks involved.

Initially, a more detailed investigation of the testis gene expression and protein expression milieu could be performed. Next, a targeted modulation of factors which we found to be altered in plasma of MSUS pups should be considered. This could be achieved using small-molecule modulators of specific inflammatory factors identified by our plasma proteomic analysis. As a read-out we suggest to assess chromatin

accessibility in spermatogonial cells, as well as the expression of previously identified candidates in the testis of MSUS pups. This would strengthen our findings, and provide additional evidence that an exogenous factor is able to induce alterations to the germline, specifically at the level of open chromatin organization.

To conclude, the findings discussed in this thesis contribute to the understanding of chromatin transitions during postnatal age in spermatogonial cells, and their impact on gene expression. Additionally, this thesis reveals that environmental experiences induce spermatogonial cell alterations in early postnatal life, and contribute novel information to the understanding of how the germline can be affected by exogenous stimuli. Taken together, the extensive description of spermatogonial cell states in naïve conditions and probing their sensitivity to exogenous factors are a valuable addition to the field of spermatogonial cell research and to germline studies in general.

References

- Ables, E.T., Laws, K.M., and Drummond-Barbosa, D. (2012). Control of adult stem cells in vivo by a dynamic physiological environment: Diet-dependent systemic factors in *Drosophila* and beyond. *Wiley Interdiscip. Rev. Dev. Biol.* 1, 657–674.
- Aloisio, G.M., Nakada, Y., Saatcioglu, H.D., Peña, C.G., Baker, M.D., Tarnawa, E.D., Mukherjee, J., Manjunath, H., Bugde, A., Sengupta, A.L., et al. (2014). PAX7 expression defines germline stem cells in the adult testis. *J. Clin. Invest.* 124, 3929–3944.
- Amann, R.P. (2008). The cycle of the seminiferous epithelium in humans: A need to revisit? *J. Androl.* 29, 469–487.
- Amezquita, R.A. (2018). marge: An API for Analysis of Motifs Using HOMER in R.
- Andrews, S., Krueger, F., Segonds-Pichon, A., Biggins, L., Krueger, C., and Wingett, S. (2012). FastQC. A quality control tool for high throughput sequence data.
- Anway, M.D., Cupp, A.S., Uzumcu, M., and Skinner, M.K. (2005). Epigenetic transgenerational actions of endocrine disruptors and male fertility. *Science* 308, 1466–1469.
- B. Teif, V., and G. Cherstvy, A. (2016). Chromatin and epigenetics: current biophysical views. *AIMS Biophys.* 3, 88–98.
- Baralle, F.E., and Giudice, J. (2017). Alternative splicing as a regulator of development and tissue identity. *Nat. Rev. Mol. Cell Biol.* 18, 437–451.
- Barrios, F., Filipponi, D., Pellegrini, M., Paronetto, M.P., Di Siena, S., Geremia, R., Rossi, P., De Felici, M., Jannini, E.A., and Dolci, S. (2010). Opposing effects of retinoic acid and FGF9 on Nanos2 expression and meiotic entry of mouse germ cells. *J. Cell Sci.* 123, 871–880.
- Bergmann, J.H., Li, J., Eckersley-Maslin, M.A., Rigo, F., Freier, S.M., and Spector, D.L. (2015). Regulation of the ESC transcriptome by nuclear long noncoding RNAs. *Genome Res.* 25, 1336–1346.
- Bertheloot, D., and Latz, E. (2017). HMGB1, IL-1 α , IL-33 and S100 proteins: Dual-function alarmins. *Cell. Mol. Immunol.* 14, 43–64.
- Bohacek, J., and Mansuy, I.M. (2015). Molecular insights into transgenerational non-genetic inheritance of acquired behaviours. *Nat. Rev. Genet.* 16, 641–652.
- Boku, S., Toda, H., Nakagawa, S., Kato, A., Inoue, T., Koyama, T., Hiroi, N., and Kusumi, I. (2015). Neonatal maternal separation alters the capacity of adult neural precursor cells to differentiate into neurons via methylation of retinoic acid receptor gene promoter. *Biol. Psychiatry* 77, 335–344.
- Caires, K.C., Shima, C.M., De Avila, J., and McLean, D.J. (2012). Acute ethanol exposure affects spermatogonial stem cell homeostasis in pre-pubertal mice. *Reprod. Toxicol.* 33, 76–84.
- Carlsen, E., Giwercman, A., Keiding, N., and Skakkebaek, N.E. (1992). Evidence for decreasing quality of semen during past 50 years. *Br. Med. J.* 305, 609–613.
- Carone, B.R., Fauquier, L., Habib, N., Shea, J.M., Hart, C.E., Li, R., Bock, C., Li, C., Gu, H., Zamore, P.D., et al. (2010). Paternally Induced Transgenerational Environmental Reprogramming of Metabolic Gene Expression in Mammals. *Cell* 143, 1084–1096.
- De Castro Barbosa, T., Ingerslev, L.R., Alm, P.S., Versteyhe, S., Massart, J., Rasmussen, M., Donkin, I., Sjögren, R., Mudry, J.M., Vetterli, L., et al. (2016). High-fat diet reprograms the epigenome of rat spermatozoa and transgenerationally affects metabolism of the offspring.
- Chan, F., Oatley, M.J., Kaucher, A. V, Yang, Q.-E., Bieberich, C.J., Shashikant, C.S., and Oatley, J.M. (2014). Functional and molecular features of the Id4+ germline stem cell population in mouse testes. *Genes Dev.* 28, 1351–1362.
- Chavez, A., Scheiman, J., Vora, S., Pruitt, B.W., Tuttle, M., P R Iyer, E., Lin, S., Kiani, S., Guzman, C.D., Wiegand, D.J., et al. (2015). Highly efficient Cas9-mediated transcriptional programming. *Nat. Methods* 12, 326–328.
- Chen, J., Cai, T., Zheng, C., Lin, X., Wang, G., Liao, S., Wang, X., Gan, H., Zhang, D., Hu, X., et al. (2017). MicroRNA-202 maintains spermatogonial stem cells by inhibiting cell cycle regulators and RNA

binding proteins. *Nucleic Acids Res.* 45, 4142–4157.

Chen, Q., Yan, M., Cao, Z., Li, X., Zhang, Y., Shi, J., Feng, G.H., Peng, H., Zhang, X., Zhang, Y., et al. (2016). Sperm tsRNAs contribute to intergenerational inheritance of an acquired metabolic disorder. *Science* (80-). 351, 397–400.

Chen, W., Zhang, Z., Chang, C., Yang, Z., Wang, P., Fu, H., Wei, X., Chen, E., Tan, S., Huang, W., et al. (2020). A bioenergetic shift is required for spermatogonial differentiation. *Cell Discov.* 6, 1–17.

Cheng, K., Chen, I.-C., Eric Cheng, C.-H., Mutoji, K., Hale, B.J., Hermann, B.P., Geyer, C.B., Oatley, J.M., and McCarrey, J.R. (2020). Unique Epigenetic Programming Distinguishes Regenerative Spermatogonial Stem Cells in the Developing Mouse Testis. *IScience* 23, 101596.

Chuikov, S., Levi, B.P., Smith, M.L., and Morrison, S.J. (2010). Prdm16 promotes stem cell maintenance in multiple tissues, partly by regulating oxidative stress. *Nat. Cell Biol.* 12, 999–1006.

Chuma, S., and Nakatsuji, N. (2000). Autonomous Transition into Meiosis of Mouse Fetal Germ Cells in Vitro and Its Inhibition by gp130-Mediated Signaling.

Chuong, E.B., Elde, N.C., and Feschotte, C. (2017). Regulatory activities of transposable elements: From conflicts to benefits. *Nat. Rev. Genet.* 18, 71–86.

Chytrava, G., Ying, Z., and Gomez-Pinilla, F. (2010). Exercise contributes to the effects of DHA dietary supplementation by acting on membrane-related synaptic systems. *Brain Res.* 1341, 32–40.

Coelho, D.M., Carvalho, L.I. da C., Melo-de-Farias, A.R., Lambert, J.-C., and Costa, M.R. (2020). Differential transcript usage unravels gene expression alterations in Alzheimer's disease human brains. *MedRxiv* 2020.03.19.20038703.

Corces, M.R., Trevino, A.E., Hamilton, E.G., Greenside, P.G., Sinnott-Armstrong, N.A., Vesuna, S., Satpathy, A.T., Rubin, A.J., Montine, K.S., Wu, B., et al. (2017). An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat. Methods* 14, 959–962.

Costoya, J.A., Hobbs, R.M., Barna, M., Cattoretti, G., Manova, K., Sukhwani, M., Orwig, K.E., Wolgemuth, D.J., and Pandolfi, P.P. (2004). Essential role of Plzf in maintenance of spermatogonial stem cells. *Nat. Genet.* 36, 653–659.

Coulthard, L.G., Hawksworth, O.A., Li, R., Balachandran, A., Lee, J.D., Sepehrband, F., Kurniawan, N., Jeanes, A., Simmons, D.G., Wolvetang, E., et al. (2017). Complement C5aR1 signaling promotes polarization and proliferation of embryonic neural progenitor cells through PKC ζ . *J. Neurosci.* 37, 5395–5407.

Dann, C.T., Alvarado, A.L., Molyneux, L.A., Denard, B.S., Garbers, D.L., and Porteus, M.H. (2008). Spermatogonial Stem Cell Self-Renewal Requires OCT4, a Factor Downregulated During Retinoic Acid-Induced Differentiation. *Stem Cells* 26, 2928–2937.

Daun, K.A., Fuchigami, T., Koyama, N., Maruta, N., Ikenaka, K., and Hitoshi, S. (2020). Early Maternal and Social Deprivation Expands Neural Stem Cell Population Size and Reduces Hippocampus/Amygdala-Dependent Fear Memory. *Front. Neurosci.* 14.

Davis, M.P., Carrieri, C., Saini, H.K., Dongen, S., Leonardi, T., Bussotti, G., Monahan, J.M., Auchyannikava, T., Bitetti, A., Rappaport, J., et al. (2017). Transposon-driven transcription is a conserved feature of vertebrate spermatogenesis and transcript evolution. *EMBO Rep.* 18, 1231–1247.

De, S., and Kassis, J.A. (2017). Passing epigenetic silence to the next generation. *Science* (80-). 356.

DeFalco, T., Potter, S.J., Williams, A. V., Waller, B., Kan, M.J., and Capel, B. (2015). Macrophages Contribute to the Spermatogonial Niche in the Adult Testis. *Cell Rep.* 12, 1107–1119.

Dias, I., Salviano, I., Mencalha, A., de Carvalho, S.N., Thole, A.A., Carvalho, L., Cortez, E., and Stumbo, A.C. (2018). Neonatal overfeeding impairs differentiation potential of mice subcutaneous adipose mesenchymal stem cells. *Stem Cell Rev. Reports* 14, 535–545.

Dietz, D.M., LaPlant, Q., Watts, E.L., Hodes, G.E., Russo, S.J., Feng, J., Oosting, R.S., Vialou, V., and Nestler, E.J. (2011). Paternal Transmission of Stress-Induced Pathologies. *Biol. Psychiatry* 70, 408–414.

Dominy, J.E., and Puigserver, P. (2013). Mitochondrial biogenesis through activation of nuclear signaling proteins. *Cold Spring Harb. Perspect. Biol.* 5, a015008.

- Dun, M.D., Aitken, R.J., and Nixon, B. (2012). The role of molecular chaperones in spermatogenesis and the post-testicular maturation of mammalian spermatozoa. *Hum. Reprod. Update* 18, 420–435.
- Dunham, I., Kundaje, A., Aldred, S.F., Collins, P.J., Davis, C.A., Doyle, F., Epstein, C.B., Frietze, S., Harrow, J., Kaul, R., et al. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature* 489, 57–74.
- Ehmcke, J., Wistuba, J., and Schlatt, S. (2006). Spermatogonial stem cells: Questions, models and perspectives. *Hum. Reprod. Update* 12, 275–282.
- Erkek, S., Hisano, M., Liang, C.Y., Gill, M., Murr, R., Dieker, J., Schübeler, D., Vlag, J. Van Der, Stadler, M.B., and Peters, A.H.F.M. (2013). Molecular determinants of nucleosome retention at CpG-rich sequences in mouse spermatozoa. *Nat. Struct. Mol. Biol.* 20, 868–875.
- Fayomi, A.P., and Orwig, K.E. (2018). Spermatogonial stem cells and spermatogenesis in mice, monkeys and men. *Stem Cell Res.* 29, 207–214.
- De Felici, M. (2013). Origin, migration, and proliferation of human primordial germ cells. In *Oogenesis*, (Springer-Verlag London Ltd), pp. 19–37.
- Ferrara, D., Hallmark, N., Scott, H., Brown, R., McKinnell, C., Mahood, I.K., and Sharpe, R.M. (2006). Acute and Long-Term Effects of *in Utero* Exposure of Rats to Di(*n*-Butyl) Phthalate on Testicular Germ Cell Development and Proliferation. *Endocrinology* 147, 5352–5362.
- Fiszbein, A., Giono, L.E., Quaglino, A., Berardino, B.G., Sigaut, L., von Bilderling, C., Schor, I.E., Steinberg, J.H.E., Rossi, M., Pietrasanta, L.I., et al. (2016). Alternative Splicing of G9a Regulates Neuronal Differentiation. *Cell Rep.* 14, 2797–2808.
- Flickinger, C.J. (1967). The postnatal development of the Sertoli cells of the mouse. *Zeitschrift Für Zellforsch. Und Mikroskopische Anat.* 78, 92–113.
- Fon Tacer, K., Montoya, M.C., Oatley, M.J., Lord, T., Oatley, J.M., Klein, J., Ravichandran, R., Tillman, H., Kim, M., Connelly, J.P., et al. (2019). MAGE cancer-testis antigens protect the mammalian germline under environmental stress. *Sci. Adv.* 5, eaav4832.
- Forand, A., Fouchet, P., Lahaye, J.B., Chicheportiche, A., Habert, R., and Bernardino-Sgherri, J. (2009). Similarities and differences in the *in vivo* response of mouse neonatal gonocytes and spermatogonia to genotoxic stress. *Biol. Reprod.* 80, 860–873.
- Fort, A., Hashimoto, K., Yamada, D., Salimullah, M., Keya, C.A., Saxena, A., Bonetti, A., Voineagu, I., Bertin, N., Kratz, A., et al. (2014). Deep transcriptome profiling of mammalian stem cells supports a regulatory role for retrotransposons in pluripotency maintenance. *Nat. Genet.* 46, 558–566.
- Franklin, T.B., Russig, H., Weiss, I.C., Gräff, J., Linder, N., Michalon, A., Vizi, S., and Mansuy, I.M. (2010). Epigenetic Transmission of the Impact of Early Stress Across Generations. *Biol. Psychiatry* 68, 408–415.
- Franklin, T.B., Linder, N., Russig, H., Thöny, B., and Mansuy, I.M. (2011). Influence of Early Stress on Social Abilities and Serotonergic Functions across Generations in Mice. *PLoS One* 6, e21842.
- Fukuda, N., and Touhara, K. (2005). Developmental expression patterns of testicular olfactory receptor genes during mouse spermatogenesis. *Genes to Cells* 11, 71–81.
- Fushan, A.A., Turanov, A.A., Lee, S.G., Kim, E.B., Lobanov, A. V., Yim, S.H., Buffenstein, R., Lee, S.R., Chang, K.T., Rhee, H., et al. (2015). Gene expression defines natural changes in mammalian lifespan. *Aging Cell* 14, 352–365.
- Gaido, K.W., Hensley, J.B., Liu, D., Wallace, D.G., Borghoff, S., Johnson, K.J., Hall, S.J., and Boekelheide, K. (2007). Fetal mouse phthalate exposure shows that gonocyte multinucleation is not associated with decreased testicular testosterone. *Toxicol. Sci.* 97, 491–503.
- Gandal, M.J., Zhang, P., Hadjimichael, E., Walker, R.L., Chen, C., Liu, S., Won, H., Van Bakel, H., Varghese, M., Wang, Y., et al. (2018). Transcriptome-wide isoform-level dysregulation in ASD, schizophrenia, and bipolar disorder. *Science* (80-.). 362.
- Gapp, K., Jawaid, A., Sarkies, P., Bohacek, J., Pelczar, P., Prados, J., Farinelli, L., Miska, E., and Mansuy, I.M. (2014). Implication of sperm RNAs in transgenerational inheritance of the effects of early trauma in mice. *Nat. Neurosci.* 17, 667–669.

- Gapp, K., van Steenwyk, G., Germain, P.L., Matsushima, W., Rudolph, K.L.M., Manuella, F., Roszkowski, M., Vernaz, G., Ghosh, T., Pelczar, P., et al. (2018). Alterations in sperm long RNA contribute to the epigenetic inheritance of the effects of postnatal trauma. *Mol. Psychiatry* 25, 2162–2174.
- Garbuzov, A., Pech, M.F., Hasegawa, K., Sukhwani, M., Zhang, R.J., Orwig, K.E., and Artandi, S.E. (2018). Purification of GFR α 1+ and GFR α 1– Spermatogonial Stem Cells Reveals a Niche-Dependent Mechanism for Fate Determination. *Stem Cell Reports* 10, 553–567.
- Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A., et al. (2013). XCRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* 154, 442.
- Ginsburg, M., Snow, M.H., and McLaren, A. (1990). Primordial germ cells in the mouse embryo during gastrulation. *Development* 110.
- Goertz, M.J., Wu, Z., Gallardo, T.D., Hamra, F.K., and Castrillon, D.H. (2011). Foxo1 is required in mouse spermatogonial stem cells for their maintenance and the initiation of spermatogenesis. *J. Clin. Invest.* 121, 3456–3466.
- Gold, H.B., Jung, H., and Corces, V.G. (2018). Not just heads and tails: the complexity of the sperm epigenome Running Title: The sperm epigenome. 5.
- Golding, J., Ellis, G., Gregory, S., Birmingham, K., Iles-Caven, Y., Rai, D., and Pembrey, M. (2017). Grand-maternal smoking in pregnancy and grandchild's autistic traits and diagnosed autism. *Sci. Rep.* 7, 46179.
- Green, C.D., Ma, Q., Manske, G.L., Shami, A.N., Zheng, X., Marini, S., Moritz, L., Sultan, C., Gurczynski, S.J., Moore, B.B., et al. (2018). A Comprehensive Roadmap of Murine Spermatogenesis Defined by Single-Cell RNA-Seq. *Dev. Cell* 46, 651–667.e10.
- Griswold, M.D. (2016). Spermatogenesis: The Commitment to Meiosis. *Physiol. Rev.* 96, 1–17.
- Griswold, M.D., and Oatley, J.M. (2013). Concise review: Defining characteristics of mammalian spermatogenic stem cells. *Stem Cells* 31, 8–11.
- Grive, K.J., Hu, Y., Shu, E., Grimson, A., Elemento, O., Grenier, J.K., and Cohen, P.E. (2019). Dynamic transcriptome profiles within spermatogonial and spermatocyte populations during postnatal testis maturation revealed by single-cell sequencing. *PLoS Genet.* 15, e1007810.
- Guo, J., Grow, E.J., Yi, C., Goriely, A., Hotaling, J.M., and Cairns Correspondence, B.R. (2017). Chromatin and Single-Cell RNA-Seq Profiling Reveal Dynamic Signaling and Metabolic Transitions during Human Spermatogonial Stem Cell Development. *21*, 533–546.
- Guo, J., Grow, E.J., Mlcochova, H., Maher, G.J., Lindskog, C., Nie, X., Guo, Y., Takei, Y., Yun, J., Cai, L., et al. (2018). The adult human testis transcriptional cell atlas. *Cell Res.* 28, 1141–1157.
- Hammoud, S.S., Low, D.H.P., Yi, C., Carrell, D.T., Guccione, E., and Cairns, B.R. (2014). Chromatin and Transcription Transitions of Mammalian Adult Germline Stem Cells and Spermatogenesis. *Cell Stem Cell* 15, 239–253.
- Hammoud, S.S., Low, D.H.P., Yi, C., Lee, C.L., Oatley, J.M., Payne, C.J., Carrell, D.T., Guccione, E., and Cairns, B.R. (2015). Transcription and imprinting dynamics in developing postnatal male germline stem cells. *Genes Dev.* 29, 2312–2324.
- Han, H., Braunschweig, U., Gonatopoulos-Pournatzis, T., Weatheritt, R.J., Hirsch, C.L., Ha, K.C.H., Radovani, E., Nabeel-Shah, S., Sterne-Weiler, T., Wang, J., et al. (2017). Multilayered Control of Alternative Splicing Regulatory Networks by Transcription Factors. *Mol. Cell* 65, 539–553.e7.
- Hara, K., Nakagawa, T., Enomoto, H., Suzuki, M., Yamamoto, M., Simons, B.D., and Yoshida, S. (2014). Mouse spermatogenic stem cells continually interconvert between equipotent singly isolated and syncytial states. *Cell Stem Cell* 14, 658–672.
- Harrow, J., Frankish, A., Gonzalez, J.M., Tapanari, E., Diekhans, M., Kokocinski, F., Aken, B.L., Barrell, D., Zadissa, A., Searle, S., et al. (2012a). GENCODE: The reference human genome annotation for the ENCODE project. *Genome Res.* 22, 1760–1774.
- Harrow, J., Frankish, A., Gonzalez, J.M., Tapanari, E., Diekhans, M., Kokocinski, F., Aken, B.L., Barrell, D., Zadissa, A., Searle, S., et al. (2012b). GENCODE: The reference human genome annotation for

The ENCODE Project. *Genome Res.* 22, 1760–1774.

Hasegawa, K., and Saga, Y. (2014). FGF8-FGFR1 signaling acts as a niche factor for maintaining undifferentiated spermatogonia in the mouse. *Biol. Reprod.* 91, 145–146.

He, C., Sidoli, S., Warneford-Thomson, R., Tatomer, D.C., Wilusz, J.E., Garcia, B.A., and Bonasio, R. (2016). High-Resolution Mapping of RNA-Binding Regions in the Nuclear Proteome of Embryonic Stem Cells. *Mol. Cell* 64, 416–430.

He, Z., Jiang, J., Kokkinaki, M., Golestaneh, N., Hofmann, M.-C., and Dym, M. (2008). Gdnf Upregulates c-Fos Transcription via the Ras/Erk1/2 Pathway to Promote Mouse Spermatogonial Stem Cell Proliferation. *Stem Cells* 26, 266–278.

He, Z., Jiang, J., Kokkinaki, M., Tang, L., Zeng, W., Gallicano, I., Dobrinski, I., and Dym, M. (2013). MiRNA-20 and MiRNA-106a Regulate Spermatogonial Stem Cell Renewal at the Post-Transcriptional Level via Targeting STAT3 and Ccnd1. *Stem Cells* 31, 2205–2217.

Heinrich, A., and DeFalco, T. (2019). Essential roles of interstitial cells in testicular development and function. *Andrology*.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Mol. Cell* 38, 576–589.

Helsel, A.R., and Oatley, J.M. (2017). Transplantation as a quantitative assay to study Mammalian male germline stem cells. In *Methods in Molecular Biology*, (Humana Press Inc.), pp. 155–172.

Helsel, A.R., Yang, Q.-E., Oatley, M.J., Lord, T., Sablitzky, F., and Oatley, J.M. (2014). ID4 levels dictate the stem cell state in mouse spermatogonia. *Genes Dev.*

Helsel, A.R., Yang, Q.-E., Oatley, M.J., Lord, T., Sablitzky, F., and Oatley, J.M. (2017a). ID4 levels dictate the stem cell state in mouse spermatogonia. *Development* 144, 624–634.

Helsel, A.R., Oatley, M.J., and Oatley, J.M. (2017b). Glycolysis-Optimized Conditions Enhance Maintenance of Regenerative Integrity in Mouse Spermatogonial Stem Cells during Long-Term Culture. *Stem Cell Reports* 8, 1430–1441.

Hentze, M.W., Castello, A., Schwarzl, T., and Preiss, T. (2018). A brave new world of RNA-binding proteins. *Nat. Rev. Mol. Cell Biol.* 19, 327–341.

Hermann, B.P., Sukhwani, M., Simorangkir, D.R., Chu, T., Plant, T.M., and Orwig, K.E. (2009). Molecular dissection of the male germ cell lineage identifies putative spermatogonial stem cells in rhesus macaques. *Hum. Reprod.* 24, 1704–1716.

Hermann, B.P., Mutoji, K.N., Velte, E.K., Ko, D., Oatley, J.M., Geyer, C.B., and McCarrey, J.R. (2015). Transcriptional and Translational Heterogeneity among Neonatal Mouse Spermatogonia1. *Biol. Reprod.* 92, 1–12.

Hermann, B.P., Cheng, K., Singh, A., Roa-De La Cruz, L., Mutoji, K.N., Chen, I.-C., Gildersleeve, H., Lehle, J.D., Mayo, M., Westernströer, B., et al. (2018). The Mammalian Spermatogenesis Single-Cell Transcriptome, from Spermatogonial Stem Cells to Spermatids. *Cell Rep.* 25, 1650–1667.e8.

Hirano, T., Kobayashi, Y., Omotehara, T., Tatsumi, A., Hashimoto, R., Umemura, Y., Nagahara, D., Mantani, Y., Yokoyama, T., Kitagawa, H., et al. (2014). Unpredictable Chronic Stress-Induced Reproductive Suppression Associated with the Decrease of Kisspeptin Immunoreactivity in Male Mice. *J. Vet. Med. Sci* 76, 1201–1208.

Houston, B.J., Nixon, B., Martin, J.H., De Iuliis, G.N., Trigg, N.A., Bromfield, E.G., McEwan, K.E., and Aitken, R.J. (2018). Heat exposure induces oxidative stress and DNA damage in the male germ line. *Biol. Reprod.* 98, 593–606.

Huber, W., Carey, V.J., Gentleman, R., Anders, S., Carlson, M., Carvalho, B.S., Bravo, H.C., Davis, S., Gatto, L., Girke, T., et al. (2015). Orchestrating high-throughput genomic analysis with Bioconductor. *Nat. Methods* 12, 115–121.

Iwamori, N., Tominaga, K., Sato, T., Riehle, K., Iwamori, T., Ohkawa, Y., Coarfa, C., Ono, E., and Matzuk, M.M. (2016). MRG15 is required for pre-mRNA splicing and spermatogenesis. *Proc. Natl. Acad. Sci. U. S. A.* 113, E5408–E5415.

- Julian, M.W., Shao, G., Bao, S., Knoell, D.L., Papenfuss, T.L., VanGundy, Z.C., and Crouser, E.D. (2012). Mitochondrial Transcription Factor A Serves as a Danger Signal by Augmenting Plasmacytoid Dendritic Cell Responses to DNA. *J. Immunol.* **189**, 433–443.
- Jung, Y.H., Sauria, M.E.G., Lyu, X., Cheema, M.S., Ausio, J., Taylor, J., and Corces, V.G. (2017). Chromatin States in Mouse Sperm Correlate with Embryonic and Adult Regulatory Landscapes. *Cell Rep.* **18**, 1366–1382.
- Kajimura, S., Seale, P., Kubota, K., Lunsford, E., Frangioni, J. V., Gygi, S.P., and Spiegelman, B.M. (2009). Initiation of myoblast to brown fat switch by a PRDM16-C/EBP- β transcriptional complex. *Nature* **460**, 1154–1158.
- Kanatsu-Shinohara, M., Tanaka, T., Ogonuki, N., Ogura, A., Morimoto, H., Cheng, P.F., Eisenman, R.N., Trumpp, A., and Shinohara, T. (2016). Myc/Mycn-mediated glycolysis enhances mouse spermatogonial stem cell self-renewal. *Genes Dev.* **30**, 2637–2648.
- Kierdorf, K., and Fritz, G. (2013). RAGE regulation and signaling in inflammation and beyond. *J. Leukoc. Biol.* **94**, 55–68.
- Klein, A.M., Nakagawa, T., Ichikawa, R., Yoshida, S., and Simons, B.D. (2010). Mouse germ line stem cells undergo rapid and stochastic turnover. *Cell Stem Cell* **7**, 214–224.
- Klemm, S.L., Shipony, Z., and Greenleaf, W.J. (2019). Chromatin accessibility and the regulatory epigenome. *Nat. Rev. Genet.* **20**, 207–220.
- Kluin, P.M., and Rooij, D.G. (1981). A comparison between the morphology and cell kinetics of gonocytes and adult type undifferentiated spermatogonia in the mouse. *Int. J. Androl.* **4**, 475–493.
- Komai, Y., Tanaka, T., Tokuyama, Y., Yanai, H., Ohe, S., Omachi, T., Atsumi, N., Yoshida, N., Kumano, K., Hisha, H., et al. (2014). Bmi1 expression in long-term germ stem cells. *Sci. Rep.* **4**, 1–12.
- Korotkevich, G., Sukhov, V., and Sergushichev, A. (2016). Fast gene set enrichment analysis. *BioRxiv* 060012.
- Krueger, F. (2015). Trim Galore. A wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files, www.bioinformatics.babraham.ac.uk/projects/trim_galore/.
- Krueger, F., and Andrews, S.R. (2011). Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* **27**, 1571–1572.
- Kubota, H., and Brinster, R.L. (2018). Spermatogonial stem cells. *Biol. Reprod.* **99**, 52–74.
- Kubota, H., Avarbock, M.R., and Brinster, R.L. (2003). Spermatogonial stem cells share some, but not all, phenotypic and functional characteristics with other stem cells. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 6487–6492.
- Kubota, H., Avarbock, M.R., and Brinster, R.L. (2004a). Growth factors essential for self-renewal and expansion of mouse spermatogonial stem cells. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 16489–16494.
- Kubota, H., Avarbock, M.R., and Brinster, R.L. (2004b). Culture Conditions and Single Growth Factors Affect Fate Determination of Mouse Spermatogonial Stem Cells. *Biol. Reprod.* **71**, 722–731.
- Kwon, S.C., Yi, H., Eichelbaum, K., Föhr, S., Fischer, B., You, K.T., Castello, A., Krijgsveld, J., Hentze, M.W., and Kim, V.N. (2013). The RNA-binding protein repertoire of embryonic stem cells. *Nat. Struct. Mol. Biol.* **20**, 1122–1130.
- La, H.M., and Hobbs, R.M. (2019). Mechanisms regulating mammalian spermatogenesis and fertility recovery following germ cell depletion. *Cell. Mol. Life Sci.* **76**, 4071–4102.
- La, H.M., Mäkelä, J.-A., Chan, A.-L., Rossello, F.J., Nefzger, C.M., Legrand, J.M.D., De Seram, M., Polo, J.M., and Hobbs, R.M. Identification of dynamic undifferentiated cell states within the male germline.
- Lambrot, R., Xu, C., Saint-Phar, S., Chountalos, G., Cohen, T., Paquet, M., Suderman, M., Hallett, M., and Kimmins, S. (2013). Low paternal dietary folate alters the mouse sperm epigenome and is associated with negative pregnancy outcomes. *Nat. Commun.* **4**, 1–13.
- Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359.

- Law, C.W., Chen, Y., Shi, W., and Smyth, G.K. (2014). voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* 15, R29.
- Law, N.C., Oatley, M.J., and Oatley, J.M. (2019). Developmental kinetics and transcriptome dynamics of stem cell specification in the spermatogenic lineage. *Nat. Commun.* 10, 2787.
- Lee, A., Rusch, J., Usmani, A., Lima, A., Wong, W., Huang, N., Lepamets, M., Vigh-Conrad, K., Worthington, R., Magi, R., et al. (2017). The complement system supports normal postnatal development and gonadal function in both sexes.
- Lee, A., Klein, J., Tacer, K.F., Lord, T., Oatley, M., Oatley, J., Porter, S., Pruett-Miller, S., Tikhonova, E., Karamyshev, A., et al. (2020). Enhanced stress tolerance through reduction of G3BP and suppression of stress granules. *BioRxiv* 2020.02.03.925677.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079.
- Li, N., Wang, T., and Han, D. (2012). Structural, cellular and molecular aspects of immune privilege in the testis. *Front. Immunol.* 3, 152.
- Li, R.Z., Hou, J., Wei, Y., Luo, X., Ye, Y., and Zhang, Y. (2019). hnRNPDL extensively regulates transcription and alternative splicing. *Gene* 687, 125–134.
- Liang, M., Li, W., Tian, H., Hu, T., Wang, L., Lin, Y., Li, Y., Huang, H., and Sun, F. (2014). Sequential expression of long noncoding RNA as mRNA gene expression in specific stages of mouse spermatogenesis. *Sci. Rep.* 4, 1–6.
- Liao, J., Ng, S.H., Luk, A.C., Suen, H.C., Qian, Y., Wing, A., Lee, T., Tu, J., Chak, J., Fung, L., et al. (2019a). STEM CELLS AND REGENERATION Revealing cellular and molecular transitions in neonatal germ cell differentiation using single cell RNA sequencing.
- Liao, Y., Smyth, G.K., and Shi, W. (2019b). The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Res.* 47, e47–e47.
- Liao, Y., Wang, J., Jaehnig, E.J., Shi, Z., and Zhang, B. (2019c). WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. *Nucleic Acids Res.* 47, W199–W205.
- Little, J.P., Simtchouk, S., Schindler, S.M., Villanueva, E.B., Gill, N.E., Walker, D.G., Wolthers, K.R., and Klegeris, A. (2014). Mitochondrial transcription factor A (Tfam) is a pro-inflammatory extracellular signaling molecule recognized by brain microglia. *Mol. Cell. Neurosci.* 60, 88–96.
- Liu, J.T., Chen, B.Y., Zhang, J.Q., Kuang, F., and Chen, L.W. (2015). Lead exposure induced microgliosis and astrogliosis in hippocampus of young mice potentially by triggering TLR4-MyD88-NF κ B signaling cascades. *Toxicol. Lett.* 239, 97–107.
- Lodhi, I.J., Dean, J.M., Song, H., Hsu, F.-F., Semenkovich Correspondence, C.F., He, A., Park, H., Tan, M., Feng, C., and Semenkovich, C.F. (2017). PexRAP Inhibits PRDM16-Mediated Thermogenic Gene Expression. *Cell Rep.* 20.
- Lord, T., Oatley, M.J., and Oatley, J.M. (2018). Testicular Architecture Is Critical for Mediation of Retinoic Acid Responsiveness by Undifferentiated Spermatogonial Subtypes in the Mouse. *Stem Cell Reports* 10, 538–552.
- Lu, F., Liu, Y., Inoue, A., Suzuki, T., Zhao, K., and Zhang, Y. (2016). Establishing chromatin regulatory landscape during mouse preimplantation development. *Cell* 165, 1375–1388.
- Lu, J.Y., Shao, W., Chang, L., Yin, Y., Li, T., Zhang, H., Hong, Y., Percharde, M., Guo, L., Wu, Z., et al. (2020). Genomic Repeats Categorize Genes with Distinct Functions for Orchestrated Regulation. *Cell Rep.* 30, 3296–3311.e5.
- Lun, A.T.L., and Smyth, G.K. (2015). csa: a Bioconductor package for differential binding analysis of ChIP-seq data using sliding windows. *Nucleic Acids Res.* 44, e45.
- Luo, Z., Wang, X., Jiang, H., Wang, R., Chen, J., Chen, Y., Xu, Q., Cao, J., Gong, X., Wu, J., et al. (2020). Reorganized 3D Genome Structures Support Transcriptional Regulation in Mouse Spermatogenesis. *IScience* 23, 101034.
- Machanick, P., and Bailey, T.L. (2011). MEME-ChIP: Motif analysis of large DNA datasets. *Bioinformatics* 27, 1696–1697.

- Maezawa, S., Yukawa, M., Alavattam, K.G., Barski, A., and Namekawa, S.H. (2018). Dynamic reorganization of open chromatin underlies diverse transcriptomes during spermatogenesis. *Nucleic Acids Res.* 46, 593–608.
- Mahood, I.K., Scott, H.M., Brown, R., Hallmark, N., Walker, M., and Sharpe, R.M. (2007). In utero exposure to Di(n-butyl) phthalate and testicular dysgenesis: Comparison of fetal and adult end points and their dose sensitivity. *Environ. Health Perspect.* 115, 55–61.
- De Maio, A., Yalamanchili, H.K., Adamski, C.J., Gennarino, V.A., Liu, Z., Qin, J., Jung, S.Y., Richman, R., Orr, H., and Zoghbi, H.Y. (2018). RBM17 Interacts with U2SURP and CHERP to Regulate Expression and Splicing of RNA-Processing Proteins. *Cell Rep.* 25, 726–736.e7.
- Martínez-Zamudio, R.I., Roux, P.F., de Freitas, J.A.N.L.F., Robinson, L., Doré, G., Sun, B., Belenki, D., Milanovic, M., Herbig, U., Schmitt, C.A., et al. (2020). AP-1 imprints a reversible transcriptional programme of senescent cells. *Nat. Cell Biol.* 22, 842–855.
- Masand, R., Paulo, E., Wu, D., Wang, Y., Swaney, D.L., Jimenez-Morales, D., Krogan, N.J., and Wang, B. (2018). Proteome Imbalance of Mitochondrial Electron Transport Chain in Brown Adipocytes Leads to Metabolic Benefits. *Cell Metab.* 27, 616–629.e4.
- McLaren, A. (2000). Germ and somatic cell lineages in the developing gonad. *Mol. Cell. Endocrinol.* 163, 3–9.
- McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. *Nat. Biotechnol.* 28, 495–501.
- Miltenyi, S., Müller, W., Weichel, W., and Radbruch, A. (1990). High gradient magnetic cell separation with MACS. *Cytometry* 11, 231–238.
- Minkina, O., and Hunter, C.P. (2017). Stable Heritable Germline Silencing Directs Somatic Silencing at an Endogenous Locus. *Mol. Cell* 65, 659–670.e5.
- MM Scotti, M.S. (2016). RNA mis-splicing in disease. *Nat. Rev. Genet.* 17, 19–32.
- Mruk, D.D., and Cheng, C.Y. (2015). The mammalian blood-testis barrier: Its biology and regulation. *Endocr. Rev.* 36, 564–591.
- Mutoji, K., Singh, A., Nguyen, T., Gildersleeve, H., Kaucher, A. V., Oatley, M.J., Oatley, J.M., Velte, E.K., Geyer, C.B., Cheng, K., et al. (2016). TSPAN8 expression distinguishes spermatogonial stem cells in the prepubertal mouse testis. *Biol. Reprod.* 95, 1–14.
- Myers, R.M., Stamatoyannopoulos, J., Snyder, M., Dunham, I., Hardison, R.C., Bernstein, B.E., Gingeras, T.R., Kent, W.J., Birney, E., Wold, B., et al. (2011). A user's guide to the Encyclopedia of DNA elements (ENCODE). *PLoS Biol.* 9, e1001046.
- Nakagawa, T., Nabeshima, Y. ichi, and Yoshida, S. (2007). Functional Identification of the Actual and Potential Stem Cell Compartments in Mouse Spermatogenesis. *Dev. Cell* 12, 195–206.
- Nakagawa, T., Sharma, M., Nabeshima, Y.I., Braun, R.E., and Yoshida, S. (2010). Functional hierarchy and reversibility within the murine spermatogenic stem cell compartment. *Science* (80-.). 328, 62–67.
- Niu, Z., Goodyear, S.M., Rao, S., Wu, X., Tobias, J.W., Avarbock, M.R., and Brinster, R.L. (2011). MicroRNA-21 regulates the self-renewal of mouse spermatogonial stem cells. *Proc. Natl. Acad. Sci. U. S. A.* 108, 12740–12745.
- Nono, M., Kishimoto, S., Sato-Carlton, A., Carlton, P.M., Nishida, E., and Uno, M. (2020). Intestine-to-Germline Transmission of Epigenetic Information Intergenerationally Ensures Systemic Stress Resistance in *C. elegans*. *Cell Rep.* 30, 3207–3217.e4.
- Oakes, C.C., La Salle, S., Smiraglia, D.J., Robaire, B., and Trasler, J.M. (2007). Developmental acquisition of genome-wide DNA methylation occurs prior to meiosis in male germ cells. *Dev. Biol.* 307, 368–379.
- Oatley, J.M., and Brinster, R.L. (2006). Spermatogonial Stem Cells. pp. 259–282.
- Oatley, J.M., and Brinster, R.L. (2012). The germline stem cell niche unit in mammalian testes. *Physiol. Rev.* 92, 577–595.
- Oatley, J.M., and Griswold, M.D. (2017). The biology of mammalian spermatogonia.

- Oatley, J.M., Avarbock, M.R., Telaranta, A.I., Fearon, D.T., and Brinster, R.L. (2006). Identifying genes important for spermatogonial stem cell self-renewal and survival. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 9524–9529.
- Oatley, J.M., Avarbock, M.R., and Brinster, R.L. (2007). Glial cell line-derived neurotrophic factor regulation of genes essential for self-renewal of mouse spermatogonial stem cells is dependent on Src family kinase signaling. *J. Biol. Chem.* **282**, 25842–25851.
- Oatley, M.J., Kaucher, A. V., Racicot, K.E., and Oatley, J.M. (2011). Inhibitor of DNA binding 4 is expressed selectively by single spermatogonia in the male germline and regulates the self-renewal of spermatogonial stem cells in mice. *Biol. Reprod.* **85**, 347–356.
- Oldfield, A.J., Yang, P., Conway, A.E., Cinghu, S., Freudenberg, J.M., Yellaboina, S., and Jothi, R. (2014). Histone-Fold Domain Protein NF-Y Promotes Chromatin Accessibility for Cell Type-Specific Master Transcription Factors. *Mol. Cell* **55**, 708–722.
- Oldfield, A.J., Henriques, T., Kumar, D., Burkholder, A.B., Cinghu, S., Paulet, D., Bennett, B.D., Yang, P., Scruggs, B.S., Lavender, C.A., et al. (2019). NF-Y controls fidelity of transcription initiation at gene promoters through maintenance of the nucleosome-depleted region. *Nat. Commun.* **10**, 1–12.
- Öst, A., Lempradl, A., Casas, E., Weigert, M., Tiko, T., Deniz, M., Pantano, L., Boenisch, U., Itskov, P.M., Stoeckius, M., et al. (2014). Paternal Diet Defines Offspring Chromatin State and Intergenerational Obesity. *Cell* **159**, 1352–1364.
- Pai, A.A., and Luca, F. (2019). Environmental influences on RNA processing: Biochemical, molecular and genetic regulators of cellular response. *Wiley Interdiscip. Rev. RNA* **10**, e1503.
- Patro, R., Duggal, G., Love, M.I., Irizarry, R.A., and Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**, 417–419.
- Pech, M.F., Garbuzov, A., Hasegawa, K., Sukhwani, M., Zhang, R.J., Benayoun, B.A., Brockman, S.A., Lin, S., Brunet, A., Orwig, K.E., et al. (2015). High telomerase is a hallmark of undifferentiated spermatogonia and is required for maintenance of male germline stem cells. *Genes Dev.* **29**, 2420–2434.
- Phillips, B.T., Gassei, K., and Orwig, K.E. (2010). Spermatogonial stem cell regulation and spermatogenesis. *Philos. Trans. R. Soc. B Biol. Sci.* **365**, 1663–1678.
- Picelli, S., Faridani, O.R., Björklund, Å.K., Winberg, G., Sagasser, S., and Sandberg, R. (2014). Full-length RNA-seq from single cells using Smart-seq2. *Nat. Protoc.* **9**, 171–181.
- Porro, V., Pagotto, R., Harreguy, M.B., Ramírez, S., Crispo, M., Santamaría, C., Luque, E.H., Rodríguez, H.A., and Bollati-Fogolín, M. (2015). Characterization of Oct4-GFP transgenic mice as a model to study the effect of environmental estrogens on the maturation of male germ cells by using flow cytometry. *J. Steroid Biochem. Mol. Biol.* **154**, 53–61.
- Potter, S.J., and De Falco, T. (2017). Role of the testis interstitial compartment in spermatogonial stem cell function. *Reproduction* **153**, R151–R162.
- Ramirez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165.
- Rappsilber, J., Ryder, U., Lamond, A.I., and Mann, M. (2002). Large-scale proteomic analysis of the human spliceosome. *Genome Res.* **12**, 1231–1245.
- Rassoulzadegan, M., Grandjean, V., Gounon, P., Vincent, S., Gillot, I., and Cuzin, F. (2006). RNA-mediated non-mendelian inheritance of an epigenetic change in the mouse. *Nature* **441**, 469–474.
- Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., and Vilo, J. (2019). G:Profiler: A web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res.* **47**, W191–W198.
- Richardson, B.E., and Lehmann, R. (2010). Mechanisms guiding primordial germ cell migration: Strategies from different organisms. *Nat. Rev. Mol. Cell Biol.* **11**, 37–49.
- Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47–e47.

- Robinson, M.D., and Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* 11, R25.
- Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2009). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140.
- Rodgers, A.B., Morgan, C.P., Bronson, S.L., Revello, S., and Bale, T.L. Paternal stress exposure alters sperm microRNA content and reprograms offspring HPA stress axis regulation.
- Rodgers, A.B., Morgan, C.P., Leu, N.A., and Bale, T.L. (2015). Transgenerational epigenetic programming via sperm microRNA recapitulates effects of paternal stress. *Proc. Natl. Acad. Sci.* 112, 13699–13704.
- Rojas-Ríos, P., and Simonelig, M. (2018). piRNAs and PIWI proteins: regulators of gene expression in development and stem cells.
- Rondanino, C., Maouche, A., Dumont, L., Oblette, A., and Rives, N. (2017). Establishment, maintenance and functional integrity of the blood-testis barrier in organotypic cultures of fresh and frozen/thawed prepubertal mouse testes. *Mol. Hum. Reprod.* 23, 304–320.
- Rooij, D.G. de (2017). The nature and dynamics of spermatogonial stem cells. *Development* 144, 3022–3030.
- De Rooij, D.G. (2017). The nature and dynamics of spermatogonial stem cells. *Dev.* 144, 3022–3030.
- De Rooij, D.G., and Russell, L.D. (2000). All you wanted to know about spermatogonia but were afraid to ask. *J. Androl.* 21, 776–798.
- Royo, H., Stadler, M., and Peters, A. (2016). Alternative Computational Analysis Shows No Evidence for Nucleosome Enrichment at Repetitive Sequences in Mammalian Spermatozoa. *Dev. Cell* 37, 98–104.
- Sada, A., Suzuki, A., Suzuki, H., and Saga, Y. (2009). The RNA-binding protein NANOS2 is required to maintain murine spermatogonia! *Stem Cells. Science* (80-). 325, 1394–1398.
- Sahlu, B.W., Zhao, S., Wang, X., Umer, S., Zou, H., Huang, J., and Zhu, H. (2020). Long noncoding RNAs: new insights in modulating mammalian spermatogenesis. *J. Anim. Sci. Biotechnol.* 11, 16.
- Saitou, M., and Yamaji, M. (2012). Primordial germ cells in mice. *Cold Spring Harb. Perspect. Biol.* 4, a008375.
- Sakashita, A., Maezawa, S., Alavattam, K., Yukawa, M., Barski, A., Pavlicev, M., and Namekawa, S. (2020). Endogenous retroviruses drive species-specific germline transcriptomes in mammals. *BioRxiv* 2020.03.11.987230.
- La Salle, S., and Trasler, J.M. (2006). Dynamic expression of DNMT3a and DNMT3b isoforms during male germ cell development in the mouse. *Dev. Biol.* 296, 71–82.
- Saxe, J.P., and Lin, H. (2011). Small noncoding RNAs in the germline. *Cold Spring Harb. Perspect. Biol.* 3, a002717.
- Seale, P., Bjork, B., Yang, W., Kajimura, S., Chin, S., Kuang, S., Scimè, A., Devarakonda, S., Conroe, H.M., Erdjument-Bromage, H., et al. (2008). PRDM16 controls a brown fat/skeletal muscle switch. *Nature* 454, 961–967.
- Sharma, U., Sun, F., Conine, C.C., Reichholz, B., Kukreja, S., Herzog, V.A., Ameres, S.L., and Rando, O.J. (2018). Small RNAs Are Trafficked from the Epididymis to Developing Mammalian Sperm. *Dev. Cell* 46, 481–494.e6.
- Sharpe, R.M., and Skakkebaek, N.E. (1993). Are oestrogens involved in falling sperm counts and disorders of the male reproductive tract? *Lancet* 341, 1392–1396.
- Shaulian, E., and Karin, M. (2002). AP-1 as a regulator of cell life and death. *Nat. Cell Biol.* 4, E131–E136.
- Shavlakadze, T., Morris, M., Fang, J., Wang, S.X., Zhu, J., Zhou, W., Tse, H.W., Mondragon-Gonzalez, R., Roma, G., and Glass, D.J. (2019). Age-Related Gene Expression Signature in Rats Demonstrate Early, Late, and Linear Transcriptional Changes from Multiple Tissues. *Cell Rep.* 28, 3263–3273.e3.
- Shima, J.E., McLean, D.J., McCarrey, J.R., and Griswold, M.D. (2004). The Murine Testicular

Transcriptome: Characterizing Gene Expression in the Testis During the Progression of Spermatogenesis1. *Biol. Reprod.* 71, 319–330.

Shimada, I.S., Acar, M., Burgess, R.J., Zhao, Z., and Morrison, S.J. (2017). Prdm16 is required for the maintenance of neural stem cells in the postnatal forebrain and their differentiation into ependymal cells. *Genes Dev.* 31, 1134–1146.

Simorangkir, D.R., Marshall, G.R., Ehmcke, J., Schlatt, S., and Plant, T.M. (2005). Prepubertal Expansion of Dark and Pale Type A Spermatogonia in the Rhesus Monkey (*Macaca mulatta*) Results from Proliferation During Infantile and Juvenile Development in a Relatively Gonadotropin Independent Manner1. *Biol. Reprod.* 73, 1109–1115.

Smith, B.E., and Braun, R.E. (2012). Germ cell migration across sertoli cell tight junctions. *Science* (80-.). 338, 798–802.

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

Song, H.W., Bettegowda, A., Lake, B.B., Zhao, A.H., Skarbrevik, D., Babajanian, E., Sukhwani, M., Shum, E.Y., Phan, M.H., Plank, T.D.M., et al. (2016). The Homeobox Transcription Factor RHOX10 Drives Mouse Spermatogonial Stem Cell Establishment. *Cell Rep.* 17, 149–164.

Starks, R.R., Biswas, A., Jain, A., and Tuteja, G. (2019). Combined analysis of dissimilar promoter accessibility and gene expression profiles identifies tissue-specific genes and actively repressed networks. *Epigenetics and Chromatin* 12, 16.

Steenwyk, G. van, Gapp, K., Jawaid, A., Germain, P.-L., Manuella, F., Tanwar, D.K., Zamboni, N., Gaur, N., Efimova, A., Thumfart, K., et al. (2019). A novel mode of communication between blood and the germline for the inheritance of paternal experiences. *BioRxiv* 653865.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., et al. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* 102, 15545–15550.

Sun, F., Xu, Q., Zhao, D., and Degui Chen, C. (2015). Id4 Marks Spermatogonial Stem Cells in the Mouse Testis. *Sci. Rep.* 5, 17594.

Sun, J., Lin, Y., and Wu, J. (2013). Long Non-Coding RNA Expression Profiling of Mouse Testis during Postnatal Development. *PLoS One* 8, 75750.

Sun, W., Dong, H., Becker, A.S., Dapito, D.H., Modica, S., Grandl, G., Opitz, L., Efthymiou, V., Straub, L.G., Sarker, G., et al. (2018a). Cold-induced epigenetic programming of the sperm enhances brown adipose tissue activity in the offspring. *Nat. Med.* 24, 1372–1383.

Sun, Z., Zhu, M., Lv, P., Cheng, L., Wang, Q., Tian, P., Yan, Z., and Wen, B. (2018b). The Long Noncoding RNA Lncenc1 Maintains Naive States of Mouse ESCs by Promoting the Glycolysis Pathway. *Stem Cell Reports* 11, 741–755.

Sundaram, V., and Wysocka, J. (2020). Transposable elements as a potent source of diverse cis-regulatory sequences in mammalian genomes. *Philos. Trans. R. Soc. B Biol. Sci.* 375.

Sundaram, V., Cheng, Y., Ma, Z., Li, D., Xing, X., Edge, P., Snyder, M.P., and Wang, T. (2014). Widespread contribution of transposable elements to the innovation of gene regulatory networks. *Genome Res.* 24, 1963–1976.

Supek, F., Bošnjak, M., Škunca, N., and Šmuc, T. (2011). REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms. *PLoS One* 6, e21800.

Suri, D., Veenit, V., Sarkar, A., Thiagarajan, D., Kumar, A., Nestler, E.J., Galande, S., and Vaidya, V.A. (2013). Early stress evokes age-dependent biphasic changes in hippocampal neurogenesis, Bdnf expression, and cognition. *Biol. Psychiatry* 73, 658–666.

Suzuki, S., Diaz, V.D., Hermann, B.P., and Wang, P.J. (2019). What has single-cell RNA-seq taught us about mammalian spermatogenesis? *Biol. Reprod.* 101, 617–634.

Tanenbaum, M.E., Gilbert, L.A., Qi, L.S., Weissman, J.S., and Vale, R.D. (2014). A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell* 159, 635–646.

- Tang, D., Kang, R., Zeh, H.J., and Lotze, M.T. (2011). High-mobility group box 1, oxidative stress, and disease. *Antioxidants Redox Signal.* 14, 1315–1335.
- Thakore, P.I., D'Ippolito, A.M., Song, L., Safi, A., Shivakumar, N.K., Kabadi, A.M., Reddy, T.E., Crawford, G.E., and Gersbach, C.A. (2015). Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. *Nat. Methods* 12, 1143–1149.
- Thurman, R.E., Rynes, E., Humbert, R., Vierstra, J., Maurano, M.T., Haugen, E., Sheffield, N.C., Stergachis, A.B., Wang, H., Vernot, B., et al. (2012). The accessible chromatin landscape of the human genome. *Nature* 489, 75–82.
- Tóth, K.F., Pezic, D., Stuwe, E., and Webster, A. (2016). The piRNA Pathway Guards the Germline Genome Against Transposable Elements. *Adv. Exp. Med. Biol.* 886, 51–77.
- Tyebji, S., Hannan, A.J., and Tonkin, C.J. (2020). Pathogenic Infection in Male Mice Changes Sperm Small RNA Profiles and Transgenerationally Alters Offspring Behavior. *Cell Rep.* 31, 107573.
- Valli, H., Sukhwani, M., Dovey, S.L., Peters, K.A., Donohue, J., Castro, C.A., Chu, T., Marshall, G.R., and Orwig, K.E. (2014). Fluorescence- and magnetic-activated cell sorting strategies to isolate and enrich human spermatogonial stem cells. *Fertil. Steril.* 102, 566.
- Vanderhaeghen, P., Schurmans, S., Vassart, G., and Parmentier, M. (1997). Specific repertoire of olfactory receptor genes in the male germ cells of several mammalian species. *Genomics* 39, 239–246.
- Vara, C., Paytuví-Gallart, A., Cuartero, Y., Le Dily, F., Garcia, F., Salvà-Castro, J., Gómez-H, L., Julià, E., Moutinho, C., Aiese Cigliano, R., et al. (2019). Three-Dimensional Genomic Structure and Cohesin Occupancy Correlate with Transcriptional Activity during Spermatogenesis. *Cell Rep.* 28, 352-367.e9.
- Vitting-Seerup, K., and Sandelin, A. (2017). The landscape of isoform switches in human cancers. *Mol. Cancer Res.* 15, 1206–1220.
- Vitting-Seerup, K., and Sandelin, A. (2019). IsoformSwitchAnalyzeR: analysis of changes in genome-wide patterns of alternative splicing and its functional consequences. *Bioinformatics* 35, 4469–4471.
- Vrooman, L.A., Oatley, J.M., Griswold, J.E., Hassold, T.J., and Hunt, P.A. (2015). Estrogenic Exposure Alters the Spermatogonial Stem Cells in the Developing Testis, Permanently Reducing Crossover Levels in the Adult. *PLOS Genet.* 11, e1004949.
- Wang, E. (2008). Alternative isoform regulation in human tissue transcriptomes. *Nature* 456, 470–476.
- Wang, M., Liu, X., Chang, G., Chen, Y., An, G., Yan, L., Gao, S., Xu, Y., Cui, Y., Dong, J., et al. (2018). Single-Cell RNA Sequencing Analysis Reveals Sequential Cell Fate Transition during Human Spermatogenesis. *Cell Stem Cell* 23, 599-614.e4.
- Watanabe, T., Tomizawa, S., Mitsuya, K., Totoki, Y., Yamamoto, Y., Kuramochi-Miyagawa, S., Iida, N., Hoki, Y., Murphy, P.J., Toyoda, A., et al. (2011). Role for piRNAs and noncoding RNA in de novo DNA methylation of the imprinted mouse Rasgrf1 locus. *Science* 332, 848–852.
- Weiss, I.C., Franklin, T.B., Vizi, S., and Mansuy, I.M. (2011). Inheritable effect of unpredictable maternal separation on behavioral responses in mice. *Front. Behav. Neurosci.* 5, 3.
- Wichman, L., Somasundaram, S., Breindel, C., Valerio, D.M., McCarrey, J.R., Hodges, C.A., and Khalil, A.M. (2017). Dynamic expression of long noncoding RNAs reveals their potential roles in spermatogenesis and fertility. *Biol. Reprod.* 97, 313–323.
- Woo, M., Isganaitis, E., Cerletti, M., Fitzpatrick, C., Wagers, A.J., Jimenez-Chillaron, J., and Patti, M.E. (2011). Early life nutrition modulates muscle stem cell number: Implications for muscle mass and repair. *Stem Cells Dev.* 20, 1763–1769.
- Wu, F., Liu, Y., Wu, Q., Li, D., Zhang, L., Wu, X., Wang, R., Zhang, D., Gao, S., and Li, W. (2018). Long non-coding RNAs potentially function synergistically in the cellular reprogramming of SCNT embryos. *BMC Genomics* 19, 631.
- Wu, X., Oatley, J.M., Oatley, M.J., Kaucher, A. V., Avarbock, M.R., and Brinster, R.L. (2010). The POU Domain Transcription Factor POU3F1 Is an Important Intrinsic Regulator of GDNF-Induced Survival and Self-Renewal of Mouse Spermatogonial Stem Cells1. *Biol. Reprod.* 82, 1103–1111.
- Yan, S.F., Ramasamy, R., and Schmidt, A.M. (2009). Receptor for AGE (RAGE) and its ligands—cast into leading roles in diabetes and the inflammatory response. *J. Mol. Med.* 87, 235–247.

- Yang, Q.-E., Racicot, K.E., Kaucher, A. V, Oatley, M.J., and Oatley, J.M. (2013). MicroRNAs 221 and 222 regulate the undifferentiated state in mammalian male germ cells. *Development* 140, 280–290.
- Yazawa, H., Sasagawa, I., Ishigooka, M., and Nakada, T. (1999). Effect of immobilization stress on testicular germ cell apoptosis in rats. *Hum. Reprod.* 14, 1806–1810.
- Yoshida, S., Takakura, A., Ohbo, K., Abe, K., Wakabayashi, J., Yamamoto, M., Suda, T., and Nabeshima, Y. (2004). Neurogenin3 delineates the earliest stages of spermatogenesis in the mouse testis. *Dev. Biol.* 269, 447–458.
- Yoshida, S., Sukeno, M., Nakagawa, T., Ohbo, K., Nagamatsu, G., Suda, T., and Nabeshima, Y. (2006). The first round of mouse spermatogenesis is a distinctive program that lacks the self-renewing spermatogonia stage. *Development* 133, 1495–1505.
- Yoshida, S., Sukeno, M., and Nabeshima, Y.I. (2007). A vasculature-associated niche for undifferentiated spermatogonia in the mouse testis. *Science* (80-.). 317, 1722–1726.
- Zeineddine, D., Hammoud, A.A., Mortada, M., and Boeuf, H. (2014). The Oct4 protein: More than a magic stemness marker. *Am. J. Stem Cells* 3, 74–82.
- Zhang, R., Sun, J., and Zou, K. (2016a). Advances in Isolation Methods for Spermatogonial Stem Cells. *Stem Cell Rev. Reports* 12, 15–25.
- Zhang, T., Oatley, J., Bardwell, V.J., Zarkower, D., and Jain, S. (2016b). DMRT1 Is Required for Mouse Spermatogonial Stem Cell Maintenance and Replenishment. *PLOS Genet.* 12, e1006293.
- Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nussbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based Analysis of ChIP-Seq (MACS). *Genome Biol.* 9, R137.
- Zhu, Q., Meisinger, J., Emanuele, N. V., Emanuele, M.A., LaPaglia, N., and Thiel, D.H. (2000). Ethanol Exposure Enhances Apoptosis Within the Testes. *Alcohol. Clin. Exp. Res.* 24, 1550–1556.
- Zuguang, G. (2020). No Title. [Https://Github.Com/Jokergoo/RGREAT](https://Github.Com/Jokergoo/RGREAT).

Acknowledgements

I. My first and foremost thank you goes to Isabelle for offering me the chance to pursue this doctoral thesis under her supervision. The freedom I enjoyed while working on my project did not come easy at times, but has shaped my abilities to learn, understand and develop my own research, and turned me into a more skilled and overall a better scientist. I also want to thank Isabelle for always being available if I had any questions or I needed guidance in how to proceed when facing challenges moments in my project. Although we did not always agree on certain aspects of the project, I was always given the chance to express my opinions, argument my decisions and reach common ground on how to proceed further. I think such discussions have also hugely contributed to shaping my soft skills and turned me into a better communicator, a more patient listener and helped me grow as a professional.

I would like to thank Tuncay and Gavin for accepting to co-supervise my project, for participating in my annual committee meeting, and for providing valuable feedback each time I needed it. Their expertise and critical assessment of my work has contributed immensely to better focusing my project and keeping an objective mindset.

I would like to thank Pierre-Luc for just being his awesome self and for (almost) always being available for a scientific or “not-so-scientific” conversation. His expertise in bioinformatics, as well as his immense conceptual wisdom have had unreplaceable roles in shaping this project. Furthermore, his ideas and his input have unarguably lifted the quality of this thesis by many times. I would also like to thank Deepak for playing an essential role in analysing all the high-throughput sequencing data that was generated for this project. I know we had our disagreements, but we were always able to find common ground and push the project forward.

My thanks also go to the “Master” spermatogonia lab at WSU: Jon, Mellissa, Tessa, Nathan and Breeana welcomed me in the Oatley lab for 3 weeks, sharing their expertise with me, as well as being so open and friendly and making my stay in the lab a truly memorable one.

Next, I would like to thank 2 very special colleagues and above all, amazing friends, with whom I have shared my most memorable moments in the lab (and outside of it): Martin and Gretchen, also known as the “MSUS support team”. You welcomed me in the lab, played an essential role in my training as a scientist, but also were the most

supportive buddies one could have hoped for in times of “despair”. We’ve shared many long days (and evenings) of experiments, Spotify playlists and HIFO happy hours and I truly hope that our friendship will continue to be as strong in the future. Martin, you’ve finally got rid of my nagging! ☺

Furthermore, I would like to thank all the past members of the Mansuy lab: Eloise, Lukas, Johannes, Silvia, Niha, Ali and Fabio. I was able to find a very welcoming and friendly environment when I joined the lab, and share many exciting moments with them, both at work and outside of work. I would also like to thank the present members of the lab for maintaining such a nice atmosphere and stimulating environment: Maria, Franci, Kristina, Lola, Andy, Alberto, Anastasia and Anara and all the recent colleagues who have joined the lab. The daily interactions and the friendly attitude of everyone around have made this a great place to work in!

Last, but not least, I would like to thank everyone at HIFO for contributing to a vibrant and collaborative research environment, where getting help and troubleshooting advice are just one question away the majority of times.

II. The second part of my acknowledgements is dedicated to the people around me who haven’t directly contributed to this thesis, but were equally important for its success and are my strength, moral compass and my biggest achievement so far:

My deepest gratitude goes to my family. My parents and my aunt have supported me since I can remember and are the main reason for why I have had the chance to get so far in my educational and professional life:

Mama, esti si vei fi mereu stanga mea: efortul si motivatia pe care le pui in tot ceea ce faci sunt o continua sursa de inspiratie! Iti multumesc pentru toate sacrificiile pe care le-ai facut pentru mine, iti datorez mai mult decat pot exprima in cuvinte. Tata, energia si pofta ta de viata nu inceteaza niciodata sa ma uimeasca. Iti multumesc pentru ca esti mereu acolo cand am nevoie de tine! Rodi, ai avut mereu atata grija de mine, mi-ai fost mama, tata, matusa si prieten de cand ma stiu. Multumesc pentru ca esti cea mai “cool” matusa pe care as fi putut sa mi-o doresc vreodata!

This past few years of living in Switzerland have exponentially became better since I met “my partner in crime”. Benjamin, thank technology we met half-way through our PhDs! Your support and love are the main reason why I was able to pull this off. Writing

our theses together was the weirdest, coolest, best support system I could have ever asked for. I cannot wait to start a new adventure together with you!!!

Ursi, Eri, Fränzi, Lukas, Anh, Lena und Jakob: Danke, dass ihr mir das Gefühl gebt, zu Hause zu sein und meine Schweizer Familie zu sein. Danke auch dafür, dass ihr mich immer mit dem besten Essen versorgt, das man sich erhoffen konnte.

Luciana, Sinziana, Denisa, Didi, Alex, you are the best friends one could ever wish for, thank you so much for being there for me throughout the years! I hope we'll always keep this friendship alive regardless of the distance.

Weirdos: Ioana (the best flatmate and friend one could have asked for when starting all over in a new country), Jo, Tamara, Dilay, Dimitris, Vlad, Andrea, Galli – you guys have been my support system and I hope I've also been yours! To many more years of friendship!