Sperm morphology and motility in men: Investigations of defective male germ cell differentiation

Inaugural Dissertation

As part of the binational joint award PhD program of the Justus Liebig University Giessen and the Monash University Melbourne

submitted to the Faculty of Medicine in partial fulfillment of the requirements for the

PhD-Degree of the Faculties of Veterinary Medicine and Medicine of the Justus Liebig University Giessen

by

Shashika Dinethri Kothalawala

of

(Colombo, Sri Lanka)

Giessen (2023)

From the

Institute for Veterinary Anatomy, Histology and Embryology

Director / Chairman

Prof. Dr. Dr. Stefan Arnhold

Chair:	**Prof. Dr. Christine Wrenzycki**
First reviewer and supervisor:	**Prof. Dr. Daniela Fietz**
Second reviewer:	**Prof. Dr. Timo Strünker**
Vice-Chair and co-supervisor:	**Prof. Dr. Andreas Meinhardt**

Date of Doctoral Defence: 08 May 2024

Sperm morphology and motility in men: Investigations of defective male germ cell differentiation

Shashika Dinethri Kothalawala

*(*MSc in Cellular and Molecular Immunology, BSc in Biotechnology)

A thesis submitted for the degree (joint award) of Doctor of Philosophy at the Monash University, Australia and Justus Liebig University Giessen, Germany

in November 2023

Department of Molecular and Translational Science,

Faculty of Medicine, Nursing and Health Sciences

&

Centre for Reproductive Health,

Hudson Institute of Medical Research

Under the supervision of:

Professor Kate Loveland, Dr. Liza O'Donnell (Monash University, Australia) and Professor Daniela Christa Fietz (Justus Liebig University Giessen, Germany)

Copyright notice

Main Abstract

Abnormalities in transcriptional expression and translational modifications during spermatogenesis can lead to defective sperm morphology and motility that contribute to approximately 80% of male infertility cases. Both RNA sequencing (RNAseq) and mass spectrometry (MS) are powerful tools to identify differentially expressed genes (DEG) and proteins (DEP) that are altered in these pathological conditions. This thesis aimed to identify DEG in normal and impaired spermatogenesis, and DEP in normal and phenotypically abnormal human spermatozoa, to elucidate new mechanisms contributing to human sperm development and function.

Testis biopsies from men were classified as normal spermatogenesis (NSP), spermatid arrest (SDA, lack of elongated spermatids and spermatozoa), or Sertoli cell only (SCO, lack of all germ cells) (n=2-3 per group) and subjected to RNAseq. DEGs were filtered based on predominant expression in spermatids and functional gene annotations relevant to sperm morphology and motility. Of these, 10 DEGs were validated by qRT-PCR and the localisation of two proteins was determined in testis biopsies by immunohistochemistry (IHC). Secondly, human ejaculates classified as normozoospermia (NORM, n=3) or asthenoteratozoospermia (defective morphology and motility, AT, n=3) were subjected to MS. Proteins selected through *in silico* analysis were examined by IHC in testicular biopsies and protein changes between AT vs NORM verified by quantitative Western blotting (qWB). Finally, importin 4 (IPO4) was selected for further investigation and identification of its potential cargos (ASF1B and TNP1) using dual-label immunofluorescence (IF).

Out of 10 DEG with potential roles in sperm morphology and/or motility, six (*SPATA31E1, TEKT3, SLC9C1, PDE4A, CFAP47*, and *TNC)* were identified as novel, human spermatid-enriched genes. IHC revealed localisation of ORAI1 and SPATA31E1 in developing human germ cells, with SPATA31E1 enriched in late spermatocytes and spermatids. 35 DEP were identified between AT and NORM sperm by MS. Through *in silico* analysis, five candidates were selected for further investigation. MS data showed IPO4, ELSPBP1, and IFT57 were higher abundance

in AT whereas CCDC105 and ACTRT2 were lower. IHC revealed their localisation in human germ cells (ACTRT2, IPO4, IFT57, CCDC105) with a predominant localisation of ACTRT2, IPO4 and IFT57 in round/elongated spermatids. ELSPBP1 was identified as an epididymal-specific protein suggesting its sperm binding ability could occur in a phenotype-dependent manner during post-testicular maturation. qWB on a different set of ejaculates did not reveal statistically significant differences for ACTRT2, ELSPBP1, and IFT57 between AT and NORM. IHC/IF analysis of IPO4 showed its localisation in the Golgi apparatus, acrosome of round spermatids, and elongating spermatid nuclei. IPO4-ASF1B co-localisation was observed in the developing human acrosome in round spermatids and IPO4-TNP1 in step 3-5 spermatid nuclei, suggesting IPO4's potential role in cargo transportation, acrosome biogenesis and spermatid nucleus condensation.

In conclusion, both transcriptomic and proteomic strategies discovered novel candidate genes and proteins likely involved in human sperm morphology and motility. Further investigation will reveal their functional properties that maintain normal sperm function. Furthermore, these genes/proteins are candidates for future studies that aim to define new diagnostic markers for the differentiation of obstructive and non-obstructive azoospermia and identify new targets for male contraception.

Main abstract – German

Zusammenfassung

Anomalien in der Transkription und Translation während der Spermatogenese können zu einer gestörten Spermienmorphologie und -motilität führen, die zu etwa 80 % der männlichen Unfruchtbarkeit beitragen. Sowohl die RNA-Sequenzierung (RNAseq) als auch die Massenspektrometrie (MS) sind leistungsfähige Instrumente zur Identifizierung von differentiell exprimierten Genen (DEG) und Proteinen (DEP), die bei diesen pathologischen Zuständen verändert sind. Ziel dieser Arbeit war es, DEGs in normaler und gestörter Spermatogenese sowie DEPs in normalen und phänotypisch abnormen menschlichen Spermien zu identifizieren, um neue Mechanismen aufzuklären, die zur Entwicklung und Funktion menschlicher Spermien beitragen.

Hodenbiopsien von Männern wurden als normale Spermatogenese (NSP), Spermatidenarrest (SDA, Mangel an elongierten Spermatiden) oder Sertoli-cell only (SCO, Mangel an allen Keimzellen) klassifiziert (n=2-3 pro Gruppe) und einer RNAseq unterzogen. Die DEGs wurden auf der Grundlage der vorherrschenden Expression in Spermatiden und funktionellen Genannotationen, die für die Spermienmorphologie und -motilität relevant sind, herausgefiltert. Davon wurden 10 DEGs durch qRT-PCR validiert, und die Lokalisation von zwei Proteinen wurde in Hodenbiopsien durch Immunhistochemie (IHC) bestimmt. Zweitens wurden menschliche Ejakulate, die als Normozoospermie (NORM, n=3) oder Asthenoteratozoospermie (defekte Morphologie und Motilität, AT, n=3) klassifiziert wurden, einer MS unterzogen. Die durch In-silico-Analyse ausgewählten DEPs wurden mittels IHC in Hodenbiopsien untersucht und die Proteinveränderungen zwischen AT und NORM durch quantitativen Western Blot (qWB) überprüft. Schließlich wurde Importin 4 (IPO4) zur weiteren Untersuchung und Identifizierung seiner potenziellen Cargos (ASF1B und TNP1) mittels Dual-Label-Immunfluoreszenz (IF) ausgewählt.

Von 10 DEG mit potenzieller Bedeutung für die Spermienmorphologie und/oder -motilität wurden sechs (*SPATA31E1, TEKT3, SLC9C1, PDE4A, CFAP47 und TNC*) als neuartige, in humanen Spermatiden exprimierte Gene identifiziert. Die IHC zeigte die Lokalisation von ORAI1 und SPATA31E1 in sich entwickelnden humanen Keimzellen,

wobei SPATA31E1 in späten Spermatozyten und Spermatiden vermehrt vorkommt. 35 DEPs wurden durch MS zwischen AT- und NORM-Spermien identifiziert. Durch In-silico-Analyse wurden fünf Kandidaten für weitere Untersuchungen ausgewählt. Die MS-Daten zeigten, dass IPO4, ELSPBP1 und IFT57 in AT häufiger vorkommen, während CCDC105 und ACTRT2 weniger häufig vorkommen. Die IHC ergab, dass sie in Keimzellen lokalisiert sind (ACTRT2, IPO4, IFT57, CCDC105), wobei ACTRT2, IPO4 und IFT57 vorwiegend in runden/elongierten Spermatiden zu finden sind. ELSPBP1 wurde als Nebenhoden-spezifisches Protein identifiziert, was darauf hindeutet, dass seine Spermienbindungsfähigkeit während der post-testikulären Reifung phänotypabhängig sein könnte. qWB an einer anderen Gruppe von Ejakulaten ergab keine statistisch signifikanten Unterschiede für ACTRT2, ELSPBP1 und IFT57 zwischen AT und NORM. Die IHC/IF-Analyse von IPO4 zeigte seine Lokalisation im Golgi-Apparat, im Akrosom runder Spermatiden und in den sich elongierenden Spermatidenkernen. Eine IPO4-ASF1B-Ko-Lokalisation wurde im sich entwickelnden Akrosom in runden Spermatiden und IPO4-TNP1 in Spermatidenkernen von Step 3-5 beobachtet, was auf eine mögliche Rolle von IPO4 beim Transport von Proteinen, der Akrosom-Biogenese und der Kondensation von Spermatidenkernen hindeutet.

Zusammenfassend lässt sich sagen, dass durch transkriptomische und proteomische Strategien neue Kandidatengene und -proteine entdeckt wurden, die wahrscheinlich an der Morphologie und Motilität humaner Spermien beteiligt sind. Weitere Untersuchungen werden die funktionellen Eigenschaften dieser Gene/Proteine aufdecken, die für die Aufrechterhaltung der normalen Spermienfunktion verantwortlich sind. Darüber hinaus sind diese Gene/Proteine Kandidaten für künftige Studien, die darauf abzielen, neue diagnostische Marker für die Unterscheidung von obstruktiver und nicht-obstruktiver Azoospermie zu definieren und neue Ziele für die männliche Empfängnisverhütung zu identifizieren.

Publications during enrolment

Submitted manuscript.

Kothalawala S. D., Günther S., Schuppe H-C., Pilatz A., Wagenlehner F., Kliesch S., O'Donnell L., Fietz D. (2023). Identification of Differentially Expressed Genes in Human Testis Biopsies Showing Intact or Defective Spermatogenesis**. Submitted manuscript: Andrology. 03 July 2023**

Conference proceedings (presenter underlined)

Kothalawala S. D., Timm T., Lochnit G., Schuppe H-C., Pilatz A., Wagenlehner F., Kliesch S., O'Donnell L., Loveland K., Fietz D. (2023). Utilizing mass spectrometry to identify proteomic changes in human sperm with abnormal morphology and motility. **56th Annual Conference of Physiology and Pathology of Reproduction and 48th Joint Conference of Veterinary and Human Reproductive Medicine. German Veterinary Medical Society (DVG), 1st – 3rd March 2023, Munster, Germany**. Poster presentation (with best poster award).

Kothalawala S. D., Timm T., Lochnit G., Schuppe H-C., Pilatz A., Wagenlehner F., Kliesch S., O'Donnell L., Loveland K., Fietz D. (2022). Use of mass spectrometry to identify proteins related to impaired sperm morphology and motility in infertile men. **12th European Congress of Andrology, 19th – 21st October 2022, Barcelona, Spain.** Poster presentation.

Kothalawala S. D., Timm T., Lochnit G., Schuppe H-C., Pilatz A., Wagenlehner F., O'Donnell L., Loveland K., Fietz D. (2021). Use of mass spectrometry to identify proteins related to impaired sperm morphology and motility in infertile men. **The 14th Annual GGL conference and Second Digital Conference 2021, 29th, 30th September 2021 (Digital), Giessen, Germany**. Oral presentation.

Kothalawala S. D., Günther S., Schuppe H-C., Pilatz A., Wagenlehner F., O'Donnell L., Fietz D. (2021). Identification of Differentially Expressed Genes in Human Testis Biopsies Showing Intact or Defective Spermatogenesis. **The 21st European Testis Workshop. 30th May – 3rd June 2021 (Digital), Barcelona, Spain.** Oral presentation.

Kothalawala S. D., Günther S., Timm T., Lochnit G., Schuppe H-C., Wagenlehner F., Fietz D. (2020) Sperm morphology and motility in mice and men: Investigations of defective male germ cell differentiation. Published poster presentation at: **The 12th International Congress of Andrology, 5th – 9th December 2020 (Digital), Muster, Germany**. Poster presentation.

Kothalawala S., Günther S., Schuppe H-C., Wagenlehner F., O'Bryan M., Fietz D. (2020) Sperm morphology and motility in mice and men: Investigations of defective male germ cell differentiation. **The Thirteenth Annual GGL conference and First Digital Conference 2020, 29th September 2020 (Digital), Giessen, Germany**. Oral presentation

Declaration

This thesis is an original work of my research and contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

Print Name: **Shashika Dinethri Kothalawala**

Date: 14.11.2023

Acknowledgements

First and foremost, my sincere gratitude and appreciation goes to my supervisory team without whom this thesis would not have been a success. I would like to thank Prof. Daniela Fietz my main supervisor from JLU and who interviewed and selected me for the IRTG P1 project. It has been a very professional yet a joyful PhD journey under your supervision. Thanks for always being there for me. The same gratitude goes to Dr. Liza O'Donnell and Prof. Kate Loveland who joined the project soon after and guided me at their best and very generously shared their years of gathered experience in this field. It was an honour and pleasure to have such a wonderful supervisory team who not only guided me in science, but also were very caring when it came to personal matters. The project took a total turn closer to a year after it was started, and I am so grateful to Daniela Fietz for being courageous to take the challenge to continuing it, while being such a pillar of strength for me, and for Dr. Liza and Prof. Kate for joining the supervisory team and providing their tremendous support and care for the project. I am inspired by you all great scientists.

I would also like to extend my gratitude to the DFG - IRTG for finding this project and the whole IRTG group for creating this consortium, and for giving me this opportunity to join the group. It was a great experience to get to know and to learn from all of you. Though the difficult times with the global pandemic minimised networking opportunities and learning experiences, the group managed to continue to have retreats and online seminars which enabled continuous learning throughout this period. I extend my gratitude to Prof. Andreas Meinhardt and Prof. Kate Loveland for creating this group and for guiding the whole team throughout these years. I would like to mention the IRTG coordinators Pia Juergens (JLU) and Rose Kiarie (Monash) for the coordination done during this period and for always been available to support and guide in official matters. My gratitude also goes to Prof. Hans-Christian Schuppe, Prof. Adrian Pilatz and Prof. Florian Wagenlehner for providing the much valuable samples for the project, Dr. Stefan Gunther of Max-Planck institute, Bad-Nauheim for RNAseq work and Prof. Lochnit Gunther and Dr. Thomas Timm of Biochemistry department, JLU for providing the Mass spec facility.

I would like to especially thank the members of the Fietz lab at the Vet Med institute. It was nice to have been working in a peaceful and friendly environment. I would like

to thank you all for the warm welcome and for generously teaching me various new techniques and for appreciating the work. My thanks extend to all, and I would like to especially mention Dr. Katja Hartmann, Alexandra Hax and Dr. Michele Klaymiuk for teaching and helping me with various techniques. Also Prof. Monika Kressin for always being a supporting mentor. My sincere appreciation and gratitude to all my colleagues, Rashidul Islam, Lachlan Cauchi, Jannis Hayer, Mirium Figura, Raouda Sgaer, Sylvia Jager and Mohsen Aberoumandi. Also, my IRTG colleague Hiba Hasan. Having good friends and colleagues can definitely make the PhD journey a happy and a successful one. You all proved that and made this journey a memorable one too.

This binational PhD program was an exciting but a difficult journey, especially due to restrictions brought with the pandemic crisis. I would like to appreciate and thank Prof. Kate Loveland for making necessary arrangements for my research stay at the Hudson Institute, Monash University, Melbourne. The research visit during the pandemic restrictions would not have been possible without your effort and cooperation. Thanks for warmly welcoming me to the Loveland team and enabling me to successfully complete the tasks at Hudson. I also like to thank my colleagues at the Hudson Institute, Dr. Jennifer Hutchison, Penny Whiley, Dr. Sarah Moody, Samira Hosseini, Junlan Maa, Michael Luu, Benedict Nathaniel of the Loveland research group and Sneha Biniwale and Dr. Ai-Leen Chan for all the support given during my research stay in Melbourne. I would also like to thank the Histology scientists at Hudson, Angela Vias, Paula Gillis and Alyce Nicholls for helping me with histology work and for always being there to discuss matters related to troubleshooting.

Finally, my utmost gratitude and love go to my family. My parents who always believe in me and believe and respect my choices in life, and for unconditionally loving me with all you have. My Ammi and Thathi, this PhD is very much an achievement of yours as much as mine. Your support and encouragement have brought me thus far in my academics and career. All the life lessons you taught, including resilience, commitment and patience always enabled me to withstand hardships. My brother Shenal, thank you for being there, listening to my worries and keeping me company during this time, it means a lot. Also, my new sister Ayesha, who joined the family very recently, thank you for being there with my parents and for taking care of them so that I could fully concentrate on my academic journey. I am blessed to have such loved ones and my inspiration to thrive more in life comes from all of you.

Abbreviations (A-Z)

AT	Asthenoteratozoospermia
cAMP	cyclic Adenosine Monophosphate
cDNA	complementary Deoxyribose Nucleotide Adenosine
cWB	Chemiluminescence WB
DE	Differentially expressed
DEG	Differentially Expressed Genes
DEP	Differentially Expressed Proteins (s)
EAU	European Andrology Union
FDR	False Discovery Rate
fWB	Fluorescence Western Blotting
GDP	Guanosine diphosphate,
GO	Gene Ontology
GTP	Guanosine triphosphate
HPA	Human Protein Atlas
IF	Immunofluorescence
IFT	Intraflagella transport
IHC	Immunohistochemistry
IMT	Intra-manchette transport
KEGG	Kyoto Encyclopedia of Genes and Genomes
LC-ESI-MS	Liquid Chromatography-Electrospray Ionization-Mass Spectrometry
LC-MS/MS/MS	Liquid Chromatography and Tandem Mass Spectrometry
Log2FC	Log2 (Fold change)
MALDI-TOF	Matrix-Assisted Laser Desorption/ionization Time-Of-Flight
MGI	Mouse Genome Informatics
MMAF	multiple morphological abnormalities of the flagella
MS	Mass Spectrometry
NADPH	Nicotinamide Adenine Dinucleotide Phosphate
NCBI	National Center for Biotechnology Information

NES	Nuclear Export Signal
NGS	Next Generation Sequencing
NLS	Nuclear Localisation Signal
NOA	Non-obstructive azoospermia
NORM	Normozoospermia
NPC	Nuclear Pore Complex
NSP	Normal spermatogenesis
nTPM	normalised Transcripts Per Million
NTR	Nuclear Transporter Receptors
OA	Obstructive Azoospermia
OAT	Oligo-Astheno-Teratozoospermia
PCR	Polymerase Chain Reaction
PKA	Protein Kinase A
qRT-PCR	quantitative Reverse Transcription PCR
RNAseq	RNA sequencing
ROS	Reactive Oxygen Species
RT-PCR	Reverse Transcription PCR
sAC	Soluble Adenylyl Cyclases
SCO	Sertoli cell only
scRNAseq	single cell RNA sequencing
SDA	Spermatid arrest
SSC	Spermatogonial Stem Cell
SZA	Spermatocyte Arrest
tmAC	transmembrane Adenylyl Cyclases
WB	Western Blot
WES	Whole Exome Sequencing
WGS	Whole Genome Sequencing
WHO	World Health Organization
YAA	Young Adult Testis Atlas

Databases used for data collection

DAVID	- https://david.ncifcrf.gov/
ENSEMBL	- https://www.ensembl.org/index.html
GeneCards	- https://www.uniprot.org/
GermOnline	- https://www.uniprot.org/
InterPro	- https://www.ebi.ac.uk/interpro/
MGI	- https://www.informatics.jax.org/
NCBI	- https://www.ncbi.nlm.nih.gov/
OMIM	- https://www.omim.org/
PANTHER	- https://www.ebi.ac.uk/interpro/
PubMed	- https://pubmed.ncbi.nlm.nih.gov/
SWISS MODEL	- https://swissmodel.expasy.org/
The Human Protein Atlas	- https://www.uniprot.org/
The Young Testis Atlas	- https://humantestisatlas.shinyapps.io/humantestisatlas1/
UniProt	- https://www.uniprot.org/

Table of Contents

Chapter 1

Literature Review

1.1 General Overview

Infertility or inability to produce a child has become a global issue with an increase in prevalence in recent years. It is reported that 10 – 15% of couples suffer from fertility problems out of which 30% have reduced fertility in both partners (Tüttelmann et al., 2018). Male infertility contributes to half of all cases of couples experiencing infertility (Zhu et al., 2016). Standardised physical examination, medical history and investigation of semen parameters are needed to evaluate potential causes of male infertility (Jungwirth et al., 2012). Male fertility could be impaired for various reasons including congenital or acquired urogenital abnormalities, malignancies, urogenital tract infections, increased scrotal temperature, endocrine disturbances, genetic abnormalities or immunological factors (Jungwirth et al., 2018). There are still around 30% cases of male infertility with unknown/unclear causes, categorised as idiopathic infertility (Lin et al., 2006; Nieschlag, 2010, p. 7).

For the assessment of male fertility, the medical history and physical examination are needed. The standard physical examination includes scrotal ultra-sound and semen analysis. The World Health Organization (WHO) has published guidelines for semen analysis and processing in the WHO laboratory manual to make the results comparable between laboratories, thus standardising the diagnostics (Table 1.1) to set a starting point for the determination of (in)fertility (Jungwirth et al., 2018; Tüttelmann et al., 2018; World Health Organization, 2021). Upon analysis of semen parameters, male infertility phenotypes are broadly categorised into a) oligozoospermia (sperm count < 15 million per mL), b) asthenozoospermia (progressive motile rate < 32%), and c) teratozoospermia (sperm with normal morphology < 4%). The most common phenotype is a combination of all three, known as Oligo-Astheno-Teratozoospermia (OAT), where both qualitative and quantitative abnormalities of sperm are detected (Jungwirth et al., 2012; Tüttelmann et al., 2018).

Table 1.1: The semen parameters published by the World Health Organization (WHO) guidelines are as follows

Semen parameter	Lower limit (Range)
Semen volume mL	1.5 (1.4-1.7)
Total semen number (10^6/ejaculate)	39 (33-46)
Sperm concentration (10^6/ mL)	15 (12-16)
Total motility (PR +NP)	40 (38- 42)
Progressive motility (PR%)	32 (31- 34)
Vitality (live spermatozoa %)	58 (55-63)
Sperm morphology (normal forms %)	4 (3.0 – 4.0)

PR: progressive motility (all space-gaining motion, both linear and in large arcs). NP: non-progressive motility (motion on the spot flagellation or motion in small circles) (World Health Organization, 2021)

Another common pathology is known as azoospermia (absence of sperm in the ejaculate). An estimated percentage of 0.1% - 1% of all men worldwide and 10 – 15% of infertile men suffer from this condition. Further investigation and analysis of testicular biopsies are needed to reveal the cause of azoospermia, which could be due to either a physical obstruction in the ductal system (obstructive azoospermia – OA), or an inability of the testes to produce sperm during spermatogenesis (non-obstructive azoospermia - NOA) (Tüttelmann et al., 2018).

About 30 - 40% of the male infertility cases are caused by genetic abnormalities such as chromosomal defects or gene mutations (Tüttelmann et al., 2007; Tüttelmann et al., 2018; Vogt, 2004). Even though this fact was revealed a few decades ago, relatively few human genes have been identified to be the cause of male infertility. Large-scale whole exome sequencing (WES) and whole genome sequencing (WGS) studies are helping to identify novel candidate genes or novel mutations in genes (Coutton et al., 2018; Liu et al., 2021; Malcher et al., 2022), and offer many advantages over small-scale familial studies which have revealed a small number of mutations that cause male infertility, such as *TEX15* (Okutman et al., 2015) and *NPAS2* (Ramasamy et al., 2015) in non-obstructive azoospermia and *ADGRG2* (Patat et al., 2016) in obstructive azoospermia (Tüttelmann et al., 2018).

Research suggests that there could be up to 2000 genes involved in male gametogenesis, which reflects the complexity of the whole process (Coutton et al.,

2015; Tüttelmann et al., 2018) and infertility can arise due to mutations, polymorphisms and variants of these genes. The genetic complexity of spermatogenesis illustrates why idiopathic forms of spermatogenic defects are high in prevalence but appear in low numbers. For the past decade, scientists have used next generation sequencing (NGS) to investigate spermatogenesis, and many studies have reported novel genes or new gene mutations which are related to male infertility (Coutton et al., 2018; Liu et al., 2021). With the technological advancements in NGS, these technologies are expected to discover more genes and gene defects related to this matter in future (Coutton et al., 2015). Nevertheless, in spite of many studies, only a few 'clinically' relevant gene mutations or polymorphisms have been identified to actually cause human male infertility. Mutations in the *CFTR* gene found in obstructive azoospermic patients decades ago is a marker that is routinely performed for azoospermic patients (Tüttelmann et al., 2018). Examining novel gene mutations in a large patient cohort is essential to establish its validity prior its use in a clinical platform (Coutton et al., 2015; Tüttelmann et al., 2018). For this reason, research groups from various countries have formed units and consortiums to define and share information and advance the understanding of genetic causes of male infertility (Tüttelmann et al., 2018).

Defining the genetic cause of infertility in patients will enable the clinician to provide genetic counselling, define the best treatment strategies and offer a better understanding of the disease prognosis (Coutton et al., 2015; O'Flynn O'Brien et al., 2010). This knowledge would also provide the basis for the development of targeted personalised medicine designed to address the defective gene and or its products. This could be achieved by protein therapy to treat protein deficiencies or gene therapy by either using mRNA injections directly or by using lipid shuttles/ virus vectors (Coutton et al., 2015).

The studies in this thesis mainly focus on the causes of male infertility in relation to sperm morphology and motility, with the aim of identifying novel genes and proteins using genomic and proteomic approaches via RNA Sequencing (RNAseq) and Mass Spectrometry (MS) that may play a major role in male (in)fertility. The identification of novel genes and proteins associated with sperm morphology and motility will provide a better understanding of human spermatogenesis and spermiogenesis.

1.2 Spermatogenesis and Spermiogenesis

1.2.1 Spermatogenesis

Spermatogenesis is the term for all processes that contributes to the production of male gametes (spermatozoa) (Weinbauer et al., 2010, p. 11) In this dynamic and complex cell development and differentiation process, immature stem cells proliferate and differentiate into mature haploid spermatozoa (Jan et al., 2017).

Spermatogenesis includes mitotic (pre-meiotic phase) and meiotic divisions (meiotic phase) as well as morphological transformations (post-meiotic phase) (Jan et al., 2017; Matzuk & Lamb, 2008). The total process can be broadly divided into 4 phases (1) mitotic proliferation and differentiation of diploid germ cells (spermatogonia) (2) progression through meiosis and the division of tetraploid germ cells (spermatocytes) into haploid germ cells (spermatids), (3) the transformation of spermatids into elongated spermatids during spermiogenesis and (4) the release of sperm from the seminiferous epithelium into the tubular lumen via the process of spermiation (Weinbauer et al., 2010, p. 16). Any abnormality in these phases can result in failure of spermatogenesis and thus in- or subfertility (Jan et al., 2017).

The stages of spermatogenesis have been characterised by microscopical analysis. The process of spermatogenesis can be explained as follows: In the pre-meiotic phase, 'A_{dark}' (quiescent or testicular stem cells), 'A_{pale}' (Ap- actively dividing) and 'B' (differentiating) spermatogonia lie at the base of the seminiferous epithelium. A_{dark} spermatogonia rarely divide and are thought to have stem cell potential. Ap spermatogonia actively divide and eventually give rise to B spermatogonia while also maintaining a pool of Ap cells. 'B' spermatogonia ultimately differentiates to pre-leptotene primary spermatocytes at the commencement of meiosis I. These spermatocytes then undergo intensive mRNA synthesis as well as DNA duplication and homologous chromosome alignment, meiotic recombination, and crossover formation to increase the genetic complexity of the gametes. Towards the end of meiosis I, tetraploid diplotene spermatocytes divide to produce diploid secondary spermatocytes, and these cells quickly enter the second meiotic division to produce

haploid round spermatids that undergo spermiogenesis. Spermiogenesis involves an incredibly complex cellular transformation as explained below (Fig. 1.1) (Jan et al., 2017; Matzuk & Lamb, 2008; Weinbauer et al., 2010, pp. 17-19).

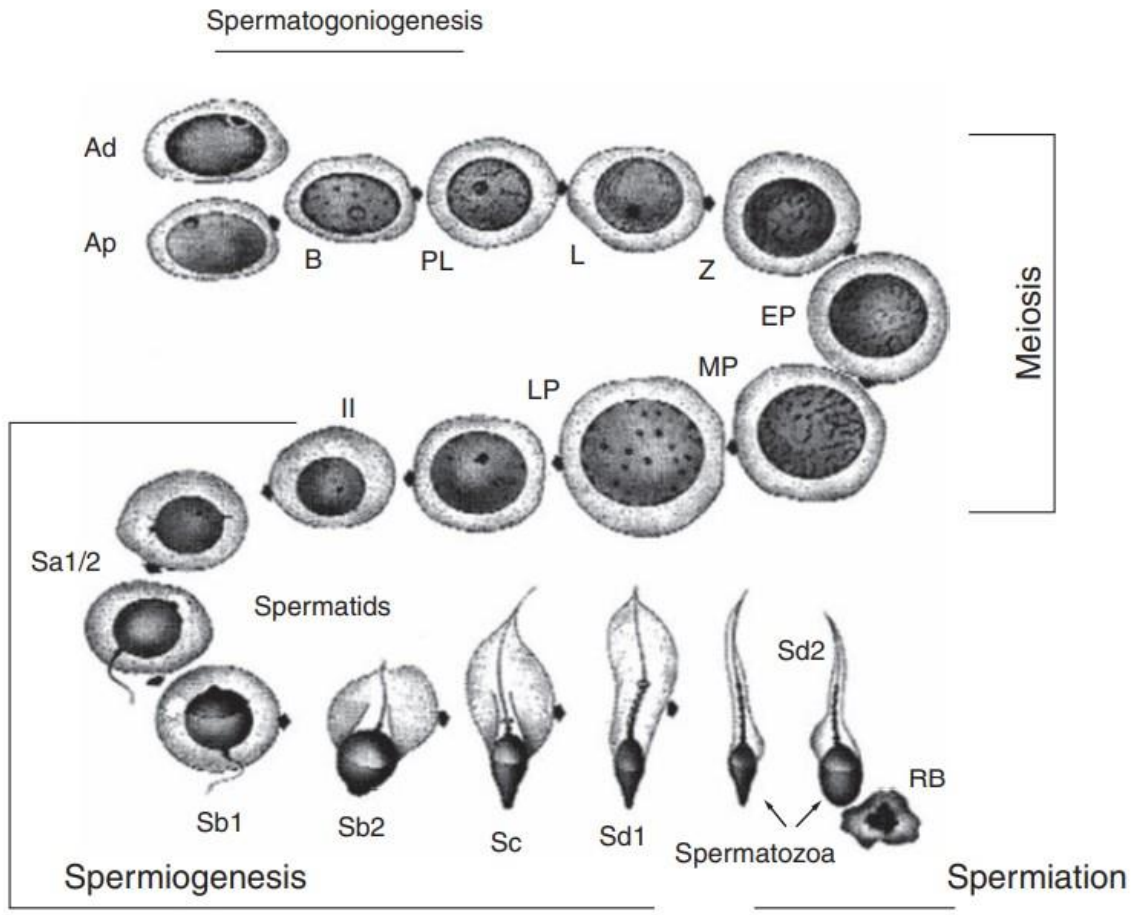


Figure 1.1: Schematic representation of all germ cell types that occur in the human seminiferous epithelium. Ap spermatogonia enters the spermatogenic process (arrow on the cell indicates direction of germ cell development). Ad spermatogonia are believed to constitute the testicular stem cells. Ad = A-dark spermatogonium, Ap = A-pale spermatogonium, B = B spermatogonium, Pl = preleptotene spermatocytes, L = leptotene spermatocytes, EP = early pachytene spermatocytes, MP = mid pachytene spermatocytes, LP = late pachytene spermatocytes, II = 2nd meiotic division, RB = residual body, Sa1 – Sd2 = developmental stages of spermatid maturation (Figure and figure legend adapted from (Weinbauer et al., 2010, p. 17).

In the germinal epithelium, somatic Sertoli cells are located in between germ cells on the basal membrane with their cytoplasm extending to the lumen of the seminiferous tubule, surrounding germ cells of all stages (Fig. 1.2). Due to this, Sertoli cells are referred to as ‘’supporting structure of the germinal epithelium’’ with morphological and

functional contact to all germ cells. Sertoli cells aid in the normal function of the testis by secreting many factors such as proteins, cytokines, growth factors etc and assist in cell division. These somatic cells play an important role in coordinating spermatogenesis in general (Weinbauer et al., 2010, p. 15).

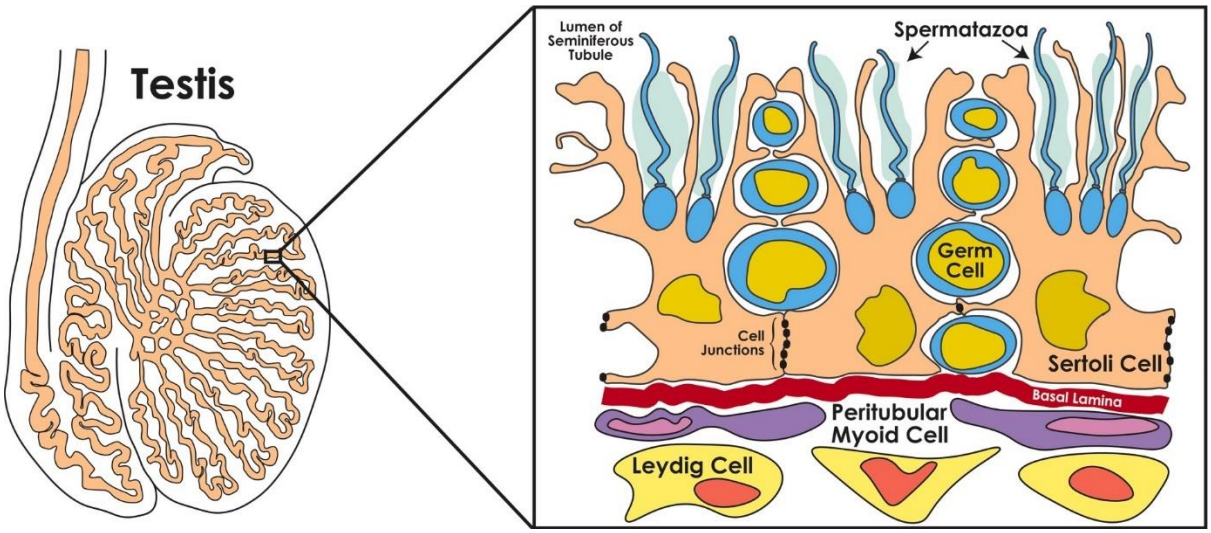


Figure 1.2: Cellular structure of the testis. Sertoli cells (yellow nuclei and pink-orange cytoplasm) are spread from basal membrane to the lumen, having close contact with germ cells of all stages. Germ cells (blue), Leydig cells (light yellow), peritubular myoid cells (purple). Figure adapted from (Washburn et al., 2022).

1.2.2 Spermiogenesis

Spermiogenesis is the 3rd phase in spermatogenesis where haploid round spermatids undergo a series of morphological and structural changes to transform into spermatozoa capable of motility. The transformation into elongated spermatids includes development of the acrosome, nuclear condensation, shaping and chromatin compaction, development of the flagellum and eventually the removal of the cytoplasm prior to spermiation (Jan et al., 2017; O'Donnell, 2014; Weinbauer et al., 2010, pp. 18-19).

Soon after round spermatids form, they start producing proteins needed for the development of the acrosome and the flagellum. Proteins destined for the acrosome are synthesised in the Golgi and are eventually deposited onto the nuclear membrane, where they spread over part of its surface. The microtubule-based axoneme, which forms the central core of the flagellum also forms and attaches to the nucleus. During mid-spermiogenesis, the nucleus starts to compact, and the spermatid chromatin starts to condense as DNA-associated histones are replaced by protamines. Gene transcription also ceases and mRNAs are stored prior to translation during the elongation phase. During this phase, the characteristic nuclear shape is achieved via a structure known as the manchette, and the complex flagellum develops (Pleuger et al., 2020). Prior to their release from the seminiferous epithelium, elongated spermatids are translocated to the luminal edge and commence the process of spermiation. During this process, their cytoplasm is condensed and removed by the Sertoli cell. The elongated spermatid eventually disengages from the Sertoli cell and is released into the lumen, and the remaining cytoplasm, now termed the residual body, is phagocytosed by the Sertoli cells (Weinbauer et al., 2010, p. 19). Impaired spermiogenesis and spermiation will result in reduced sperm number in ejaculate, abnormal sperm morphology and low or abnormal sperm motility (O'Donnell, 2014). Also, spermatogenesis is based and programmed by many specific transcriptional regulatory pathways (Krausz & Sassone-Corsi, 2005) and mishap at any such pathway could affect the process significantly.

1.3 Genetic defects and spermatogenesis

Spermatogenesis is a complex process. It is believed to be regulated by up to 2000 genes, out of which hundreds of genes (~700 – 900) are exclusively expressed in the male germ line (Chalmel et al., 2012; Tüttelmann et al., 2018). Thus, the genetics of male fertility is complex making it quite difficult to define individual causes of infertility (Tüttelmann et al., 2018). The large number of genes expressed during germ cell development suggests that the majority of genes causing male infertility are yet to be identified (Coutton et al., 2015). Even though a vast range of mutations and polymorphisms of various genes were identified in relation with nonspecific spermatogenic impairment, very few of these are in the front line for the development

of clinical diagnostics and treatment strategies (Tüttelmann et al., 2018). Though it is estimated that 30% of male infertility cases occur due to chromosomal aberrations or gene mutations involved in spermatogenesis, no treatments for genetic causes have been developed. As an example, AZF deletions in the long arm of Y chromosome were found to be one common cause of spermatogenic failure decades ago, yet there are no treatment options to repair these deletions. Chromosomal aberrations (mainly 47, XXY Klinefelter syndrome), microdeletions of the Y chromosome and cystic fibrosis transmembrane conductance regulator (*CTFR*) mutations are also associated with male infertility (Jungwirth et al., 2018).

Compared to other organs, the mammalian testis transcriptome is highly complex, particularly because spermatocytes and spermatids express more protein-coding genes (Soumillon et al., 2013). The overall understanding of the process of male fertility and infertility-causing factors is also complex because important molecular roles are played by non-coding RNAs, RNA binding proteins, transcription factors, protein modifications etc (Matzuk & Lamb, 2008) and factors that result in epigenetic alterations in the sperm genome which can then influence the phenotype of the offspring (Bianchi et al., 2019). The understanding of the molecular basis of male gamete production could be achieved by determining the dynamic transcriptional profile during human spermatogenesis, which will also uncover the underlying cause of spermatogenic dysfunction (Zhu et al., 2016). The identification of genes that regulate human sperm morphology and motility provides a better understanding of the causes of male infertility. The following table (Table 1.2) gives examples of some of these mutations and their functional consequences that are further explained in following sections.

Table 1.2: A summary of genes related to sperm morphology and motility defects; genes and proteins important in cilia and flagella function - experimental details of research in the recent past.

No	Gene	Human phenotype	Protein localisation and function in different organisms	Mutation observed in mutant models	Publications list
1	*DNAH1* Dynein axonemal heavy chain1 in chromosome	Multiple morphological abnormalities of the sperm flagella MMAF	Inner Dynein Arm Function Dynein contributing to the stabilization of the radial spoke and hence of the central doublet	In mouse - MDHC7 mutant: Reduced ciliary beating and asthenozoospermia (no MMAF-like Phenotype) Partial KO present with asthenozoospermia with no obvious flagellar defect dhc1b/dhc2 mutant: short flagella in Chlamydomonas	Human - (Ben Khelifa et al., 2014) (Wang et al., 2017) Mouse - (Neesen et al., 2001) (Pazour et al., 1999)
2	*CFAP43* Protein feature - WD repeat domains	-	Cilia associated and component of the(T/TH) complex connected to inner dynein arm. Fu et al., 2018; Kubo et al., 2018; Urbanska et al., 2018) Tetrahymena, Chlamydomonas, -flagellar protein located between doublet microtubules 5 and 6 and paraflagellar rod (Trypanosoma)	Mouse - Asthenozoospermia (MMAF like phenotype) Tetrahymena Fap43 gene deletion: Altered waveform, beat stroke and reduced swimming speed (Urbanska et al., 2018) Trypanosoma TbCfap43 (Tb.927.4.5380) RNAi mutant: Cell growth defects, Abnormal flagellum beating (axonemal disorganization), normal flagellum length	(Coutton et al., 2018) (Fu et al., 2018; Kubo et al., 2018; Urbanska et al., 2018) (Coutton et al., 2018)
3	*CFAP44* Protein feature - WD repeat domains	-	Cilia associated and component of the (T/TH) complex connected to inner dynein arm - (Fu et al., 2018 Kubo et al., 2018; Urbanska et al., 2018) (Tetrahymena, Chlamydomonas), -flagellar protein located between doublet microtubules 5 and 6 and paraflagellar rod (Coutton et al., 2018) (Trypanosoma)	Mouse Asthenozoospermia (no MMAF like phenotype) (Coutton et al., 2018) Same for trypanosome and Tetrahymena as in CFAP43	(Coutton et al., 2018) Urbanska et al., 2018 Fu et al., 2018 Kubo et al., 2018 (Fu et al., 2018; Kubo et al., 2018; Urbanska et al., 2018)

No	Gene	Human phenotype	Protein function in human	Mutation in mutant models	Publications list
4	Aurora Kinase C (*AURKC*)	Large headed multi-flagella spermatozoa	Protein Kinase. Necessary during meiosis for chromosome pairing and cytokines	Subfertility in male and female homozygous mice. Meiotic defects	Human reference – (Dieterich et al., 2007; Dieterich et al., 2009) Mouse – (Balboula & Schindler, 2014; Kimmins et al., 2007; Schindler et al., 2012)
5	dpy-19like2 *DPY19L2* 12 in chromosome 12	Globozoospermia	Transmembrane protein necessary to link the acrosome to the nucleus	Pure globozoospermia phenotype in mice	Human – (Harbuz et al., 2011; Koscinski et al., 2011) Mouse – (Yassine, Escoffier, Martinez, et al., 2015b)
6	Spermatogenesis associated protein 16 *SPATA16* in chromosome 3	Globozoospermia	Localises to the Golgi apparatus and to the pro acrosomal vesicles	Not done	Human – (Dam et al., 2007; ElInati et al., 2016) Dam et al., 2007; Elinati et al., 2016
7	Spermatogenesis associated protein 6 *SPATA6* in chromosome	Acephalic spermatozoa	Necessary for a tight head-tail connection	Pure acephalic spermatozoa phenotype in mice	(Yuan et al., 2015)
8	*CFAP69* Armadillo-type a-helical repeats	-	Cilia from olfactory sensory neurons (Talaga et al, 2017), sperm flagellum mid-piece (mammals) (Dong et al, 2018), -flagella associated protein (Chlamydomonas) (Pazour et al, 2005; Yang et al, 2006)	Mouse Olfactory defects (Talaga et al, 2017) and Asthenozoospermia (MMAF like phenotype) (Dong et al, 2018)	Mouse - (Dong et al., 2018; Talaga et al., 2017) Clamydomonas - (Pazour et al., 2005; Yang et al., 2006)
9	*FSIP2* *AKAP4* interacting domain	-	Sperm flagellum fibrous sheath (mammals) (Brown et al, 2003)		(Brown et al., 2003) FSIP2 - (Martinez et al., 2018)
10	*AK7* AK domain, coiled-coil domain, DPY30 domain	-	Cilia and sperm flagella (mammals) Fernandez-Gonzalez et al, 2009; Lores et al, 2019,	PCD (Fernandez-Gonzalez et al, 2009) and Asthenozoospermia (MMAF-like phenotype) (Lores et al, 2019)	(Fernandez-Gonzalez et al., 2009) (Lores et al., 2019)

No	Gene	Human phenotype	Protein localisation and function in different organisms	Mutation observed in mutant models	Publications list
11	*WDR66* (*CFAP251*) Calcium regulating EF-hand domain	-	Sperm flagellum (mammals) (Kherraf et al, 2018) -Flagella (Trypanosoma) (Kherraf et al, 2018) -radial spoke component (Chlamydomonas, Tertrahymena) Heuser et al, 2012 and Urbanska et al, 2015	Tetrahymena Fap251 mutant: Reduced cells swimming, Impaired ciliary beating coordination, normal length (Urbanska et al, 2015) Trypanosoma TbCfap44 (Tb.927.7.3560) RNAi mutant: Cell growth defects, Abnormal flagellum beating (axonemal disorganization), normal flagellum length (Coutton et al, 2018)	(Kherraf et al., 2018) (Heuser et al., 2012) (Urbanska et al., 2015) (Coutton et al., 2018)

Genes related to sperm motility

No	Gene	Human phenotype	function in human	Mutation in mutant models	Publications list
12	Septin12, *SEPT12*	Asthenozoospermia	GTP as a polymerizing protein constitutive of the sperm annulus	Maturation arrest and flagellar defects in mice	Human – (Kuo et al., 2012) Mouse- (Lin et al., 2009)
13	Cation channel sperm associated 1 *CATSPER 1*	Asthenozoospermia	Part of a cation channel mediating Ca2+during sperm motility and capacitation	Mouse - Absence of hyperactivated motility, increased basal cAMP content and beat frequency	Human - (Avenarius et al., 2009) Mouse- (Carlson et al., 2003; Ren et al., 2001)
14	Cation channel sperm-associated 2 *CATSPER 2*	Asthenozoospermia	Part of a cation channel mediating Ca2+during sperm motility and capacitation	Mouse - Absence of hyperactivated motility, increased basal cAMP content and beat frequency	Human – (Avidan et al., 2003; Jan et al., 2017; Zhang et al., 2007) Mouse – (Carlson et al., 2005; Quill et al., 2003)
15	Solute carrier family 26 member 8 *SLC26A8*	Asthenozoospermia	Anion transporter located in the annulus interacting and stimulating the CFTR anion channel activity	Mouse- Null males were sterile due to complete lack of sperm motility and reduced sperm fertilization potential	Human – (Dirami et al., 2013) Mouse – (Touré et al., 2007)

1.4 Morphological defects of sperm and their contribution to male infertility

Both quantitative and qualitative sperm factors contribute to fertility fitness of a male. While the number of sperm is an important determinant of male fertility, morphological or qualitative abnormalities of the sperm are also important contributors. Only a few morphological defects lead to one certain sperm phenotype such as *AURCK* mutation leading to macrozoospermia (Dieterich et al., 2007; Dieterich et al., 2009). Most aberrant sperm phenotypes are related to a plethora of different mutations – as known so far (Ray et al., 2017). Morphological abnormalities of sperm are broadly divided into defects of the sperm head and of the flagellum (tail). Severe teratozoospermia can be observed in many different specific phenotypes, including macrozoospermia (large headed multi-flagella spermatozoa), globozoospermia (round headed spermatozoa), pinheaded (headless or tailless spermatozoa) and multiple morphological abnormalities of the flagella (MMAF) (Coutton et al., 2015).

1.4.1 Abnormalities of the sperm head and their contribution to male infertility

Sperm head abnormalities can be broadly categorised into three groups as explained below.

1.4.1.1 Large-headed multi-flagellar spermatozoa

Large-headed multi-flagella spermatozoa or macrozoospermia is a rare sperm defect which is characterised by oversized irregular heads, abnormal mid-pieces and multiple flagella (Fig 1.3 B) (Ray et al., 2017). It has been hypothesised that chromosome nondisjunction during meiosis causes macrozoospermia, which is supported by several studies (Achard et al., 2007; Brahem et al., 2012; Vicari et al., 2003). A study in 2007 reported that a homozygous deletion (c.144delC) of the *AURKC* gene (Aurora kinase C) was observed in half (n=32) of the macrozoospermic patients (n=62) subjected to the study (Dieterich et al., 2009). A study conducted by Ben Khelifa and colleagues in 2011 reported that patients (n=2) with mutated *AURKC* gene (c.436-2A>G) produce macrozoospermic spermatozoa with homogenous tetraploid DNA. This was an interesting finding since it showed that spermatogenesis continued even though the two meiotic divisions were disturbed after DNA synthesis in early

spermatocytes. The inability to complete any of the two meiotic divisions results the tetraploidy. It was also suggested that the malfunctioning in the meiotic checkpoint controls could be the cause of this condition (Ben Khelifa et al., 2014; Ray et al., 2017).

1.4.1.2 Globozoospermia

Globozoospermia is characterised by the presence of round-headed spermatozoa without an acrosome. Due to the lack of an acrosome, spermatozoa are unable to penetrate the granulosa cells and the zona pellucida and thus cannot fertilize the egg (Fig: 1.3 C). This severe condition has an incident rate of 0.1% and is a rare cause of male infertility. Patients with this condition can show the phenotype in different percentages in the ejaculate. Some could have a completely globozoospermic condition (100%) while in some patients a mixture of normal and round-headed spermatozoa is observed in different proportions (Ray et al., 2017).

Following are the genes in which mutations have been identified as potentially causing globozoospermia:

- *PICK1* gene - A homozygous missense mutation (G198A) in exon 13 of the *PICK1* gene (Liu et al., 2010).
- *ZPBP1* gene - and a heterozygous missense and a splicing mutation in *ZPBP1* derived from familial studies (Yatsenko et al., 2012).
- *SPATA16* - A homozygous mutation (c.848G>A) in *SPATA16* (spermatogenesis-associated protein 16, previously named NYD-SP12) (Dam et al., 2007) was identified in a familial study. The mutation was not observed in patients of other geographical regions but another two mutations of *SPATA16* were reported (ElInati et al., 2016; Karaca et al., 2014). Since the gene is highly expressed in the human testis and the protein is localised in the Golgi and acrosomal regions, it can be safely assumed that SPATA16 has a role in spermiogenesis (Dam et al., 2007; Lu et al., 2006).
- *DPY19L2* gene - A 200-kb homozygous deletion of the *DPY19L2* gene (Harbuz et al., 2011). Many studies have shown mutations in this gene in different patient proportions. *Dpy19l2* is expressed predominantly in spermatids in the mouse, with the protein localised to the inner nuclear membrane facing the

acrosomal vesicle, where it was shown to be essential for acrosome attachment to the nucleus during spermiogenesis (Pierre et al., 2012).

1.4.1.3 Pinhead/ acephalic spermatozoa

Acephalic spermatozoa is a condition where the ejaculate mostly has headless spermatozoa (only flagella) with a small proportion of sperm heads without flagella. The condition is believed to occur due to the defects in formation of the junction between the sperm head and the flagella at the connecting piece during spermiogenesis (Chemes & Rawe, 2010).

Mutations in the following genes have been suggested to cause this phenotype: (Ray et al., 2017).

- *SUN5* bi-allelic mutations. SUN5 is found in elongating spermatids and spermatozoa and is located in the nuclear membrane at the tail/head junction (Yassine, Escoffier, Abi Nahed, et al., 2015). SUN protein forms LINC (Linker of nucleoskeleton and cytoskeleton) complexes which shows its probable importance in tail-head junctions in spermatozoa (Ray et al., 2017).
- *Spata6* – A gene knockout model in mice revealed a null phenotype with acephalic spermatozoa. The gene encodes a myosin-based microfilament protein that is localised in the connecting piece of the sperm (Yuan et al., 2015). According to the Human Protein Atlas, *SPATA6* mRNA is enriched in human testis with the highest mRNA expression in early/late spermatids (Uhlén et al., 2015).

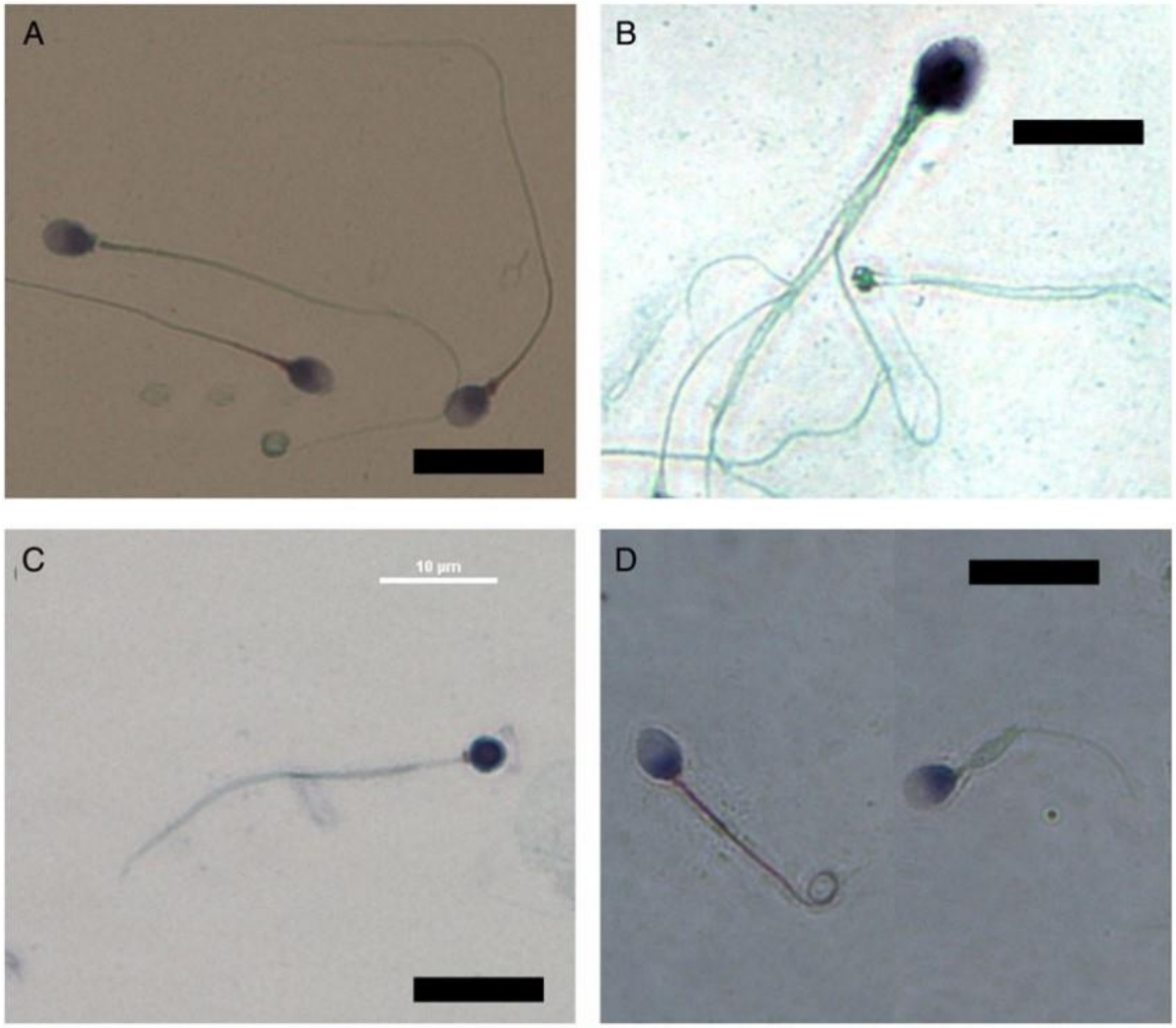


Figure 1.3: Light microscopy analysis of spermatozoa from different types of genetic infertility. Sperm morphology was assessed with Harris-Shorr staining. **(A)** Control spermatozoa present a normal-shaped head, a typical short mid piece followed by a long principal piece. **(B)** A typical macrospermic spermatozoon from a patient with a homozygous aurora kinase C (*AURKC*) c.144del mutation presents a large head and several flagella. **(C)** A typical globozoospermic spermatozoon from a dpy-19-like2 (*DPY19L2*) deleted patient presents as a round-shaped head and absence of mitochondria in the mid-piece. **(D)** Spermatozoa from a patient with a homozygous mutation in dynein, axonemal, heavy chain1 (*DNAH1*) with multiple morphological abnormalities of the flagella (MMAF) phenotype, presenting with various defects of the flagellum. Scale bars 10 mm. (Figure and figure legend adapted from (Coutton et al., 2015)).

1.4.2 Sperm flagellum, its structural abnormalities and their contribution to male infertility

The ability to progress through the female genital tract and penetrate the zona pellucida is a key feature of mature spermatozoa that determines its ability to fertilize. Defective flagella always result in asthenozoospermia and therefore reduced male fertility (Coutton et al., 2015).

The sperm flagellum is slightly different to other motile cilia and is divided into 3 main regions; the mid-piece, principal piece and the end piece (Ray et al., 2017). The mid-piece contains a central cytoskeletal structure termed the axoneme. This axoneme is covered by a helical mitochondrial sheath, containing mitochondria for energy generation during motility. The axoneme of the mid-piece consists of a central microtubule-based structure containing 9 outer microtubules doublets and 2 central singlets (9+2 structure). The microtubules are connected to each other by protein complexes containing the molecular motor protein, dynein (Fig. 1.4 a). The movement of flagella is facilitated by the sliding motion of the inner and outer dynein arms. Cilia found in epithelial cells of the airways, fallopian tubes and brain ventricles show the same axoneme structure however cilia exclusively constitute of an axoneme, whereas mammalian flagella are characterised by the specialized accessory parts such as the outer dense fibers, and the mitochondrial sheath (in the mid-piece) and fibrous sheaths (principal piece). The terminal piece does not contain a peri-axonemal structures (Coutton et al., 2015; Ray et al., 2017).

Ultra-structural analysis of immotile spermatozoa demonstrates flagella-irregularities which includes abnormalities in the axoneme pattern in association with other components of the sperm flagella (Pereira et al., 2017). Proteomic analyses have identified 700 - 1000 proteins localised exclusively in the human sperm flagella and most of these proteins are axonemal proteins (Baker et al., 2013; Coutton et al., 2015). The proteomic complexity of the flagella structure and its biogenesis mean, that defects in any of the protein coding gene, could cause flagella malformation causing teratozoospermia and asthenozoospermia (Coutton et al., 2018).

1.4.2.1 Multiple Morphological Abnormalities of the sperm Flagella (MMAF)

Multiple Morphological Abnormalities of the sperm Flagella (MMAF), a phenotype previously known as 'dysplasia of the fibrous sheath', accounts for various morphological flagella defects (Fig. 1.3 D) (Coutton et al., 2015; Nsota Mbango et al., 2019; Ray et al., 2017). Different forms of abnormalities of the sperm flagella have been reported in many studies. The figure below shows experimental evidence in the case of *DNAH1* mutation where electron microscopic images show visible structural defects in the axoneme (Fig. 1.4 b). The typical MMAF condition shows mosaic morphological abnormalities which include coiled, absent, angular, bent, irregular or short flagella, commonly defined as asthenoteratozoospermia (Ray et al., 2017).

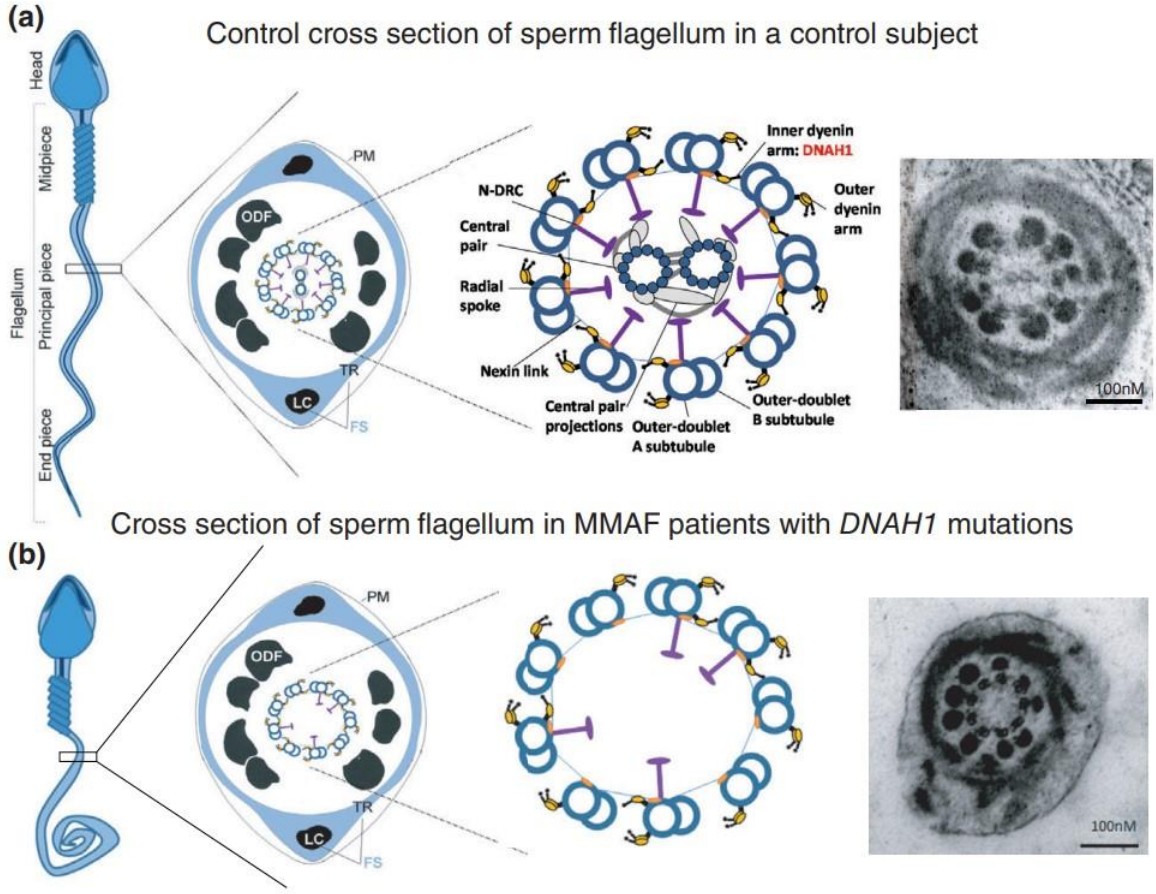


Figure 1.4: Structure of the human sperm flagellum. **(a)** A schematic representation of a human flagellum structurally divided onto: mid-piece, principal-piece and end-piece and a cross section of the principal piece depicting PM- plasma membrane, ODF – outer dense fibers, FS – fibrous sheath that is made of 2 longitudinal columns (LC) that are connected by transverse ribs (TR). The middle axoneme is enlarged to show the 9+2 structure. In the right-side corner contains the electron microscopic (EM) cross section of a normal spermatozoa. **(b)** Human sperm flagellum in a patient with *DNAH1* mutation; the schematic flagellum, cross

section of the principal piece and EM cross section depicting the changes (absence of the central pair) seen in flagellum of a *DNAH1* mutated patient (Figure adapted from (Ray et al., 2017) with modified figure legend).

1.4.2.2 Examples of gene mutations causing flagella abnormalities

The following examples highlight examples of mutations in genes that encode proteins involved in the assembly and/or structure of the flagellum:

- *DNAH1* – Null mutations of *DNAH1* result in axonemal disorganisation including mis-localisation of the peripheral microtubule doublets, an absence of the central pair and loss of the inner dynein arms, severe disorganization of the fibrous sheath, the outer dense fibers and the mitochondrial sheath. This mutation causes MMAF condition (Ben Khelifa et al., 2014; Coutton et al., 2018)
- *SEPT12* – Essential for the structural and mechanical integrity of sperm, including proper mitochondrial architecture and establishment of the annulus (Steels et al., 2007). Mutations in the gene have shown flagella defects in mice (Kissel et al., 2005; Lin et al., 2009) and teratozoospermia (Lin et al., 2012; Pereira et al., 2017).
- AKAP3 and AKAP4 - Responsible for the most abundant proteins of the fibrous sheath which was revealed in a genetic study with partial genomic deletions of these genes (Baccetti et al., 2005)
- *CFAP43* and *CFAP44* – These proteins are necessary for sperm axoneme organisation and transgenic mouse models (*Cfap43*$^{-/-}$ show infertility with 100% immotile sperms with severe flagella defects similar to human MMAF. *Cfap44*$^{-/-}$ mice showed immotile sperm causing infertility but lesser sperm defects compared to human *CFAP44* mutations (Coutton et al., 2018).
- *DZIP1* – induce asthenozoospermia through severe MMAF (Lv et al., 2020).

1.5 Sperm motility, and the contribution of reduced motility to male infertility

According to EAU and WHO guidelines, two main sperm movement patterns are considered as total sperm motility: progressive motility and non-progressive motility. Progressive motility is the forward movement of the sperm where the sperm swims in a straight line in larger circles. The non-progressive motility refers to sperm movement but not in a straight line, rather sperm can be seen moving in tight circles. Asthenozoospermia is defined as the reduction or absence of motile spermatozoa in the ejaculate. The sperm immotility is declared when the total motility cut-off is below 40% and the progressive motility rate is 32% or below (Jungwirth et al., 2018; World Health Organization, 2021). While morphological defects of the spermatozoa can affect the sperm motility, there are many molecular and physiological factors which determine the normal sperm motility. Defects in these pathways can directly influence sperm motility which are detailed in the following sections.

1.5.1 Defects in physiological processes required for sperm motility, and their contribution to male infertility.

Multiple regulatory and metabolic pathways are involved in sperm motility, and abnormalities in the expression or function of the genes and proteins involved could cause male infertility (Pereira et al., 2017). Two particularly important mechanisms that regulate sperm motility are modulation of intracellular calcium (Ca^{2+}) and pathways involving cyclic adenosine monophosphate (cAMP) and the cAMP-dependent protein kinase A (PKA). These pathways intersect in the regulation of sperm motility and involve the modulation of calcium ions, adenylyl cyclases (AC), bicarbonate ions, different membrane channels and phosphorylation (Pereira et al., 2017). However, most of the cAMP synthesis in sperm is regulated by soluble adenylyl cyclases (sAC), which plays a particularly important role in regulating flagella beat frequency (Pereira et al., 2017).

A study has shown how the inactivation of the sAC gene in mice at the embryonic stem cell state (by deleting sAC gene exon 2-4 in wild type locus) hinders the progressive motility of spermatozoa resulting in infertility. However, when cAMP was re-introduced, the forward motility was restored indicating that the cause of the immotility was the

absence of cAMP in the spermatozoa due to the lack of sACs (Esposito et al., 2004). Tyrosine phosphorylation of proteins is achieved by the cAMP-dependent pathway via the stimulation of Protein Kinase A (PKA) activity. PKA activates particular intermediate tyrosine kinases needed for sperm motility (Bajpai & Doncel, 2003). It is also worthwhile noting that the cAMP/PKA pathway also depends on the presence of reactive oxygen species (ROS) in the spermatozoa or ejaculate. The effect of ROS on sperm motility is discussed further.

1.5.2 Regulation of sperm motility by calcium (Ca^{2+}) pathways

Calcium plays a regulatory activity in sperm capacitation, the process by which the sperm gains its ability to fertilise the oocyte during its transit through the female reproductive tract. It is also involved in sperm hyperactivation and the initiation of the acrosome reaction. Intracellular calcium is needed for the proper flagella beating in spermatozoa, with the rate and symmetry of flagella beat highly responsive to local calcium levels (Pereira et al., 2017; Suarez, 2008; Suarez et al., 1993).

1.5.2.1 Importance of Ca^{2+} in the function of spermatozoa

a) Calmodulin.

Intracellular Ca^{2+} signaling in the sperm flagellum is regulated by the calcium-dependent kinase Calmodulin, which senses Ca^{2+} levels within the sperm axoneme (Ho & Suarez, 2003; Linck et al., 2014). The interaction of Calmodulin with protein kinases, phosphatases and sACs are important for the regulation of the sperm motility (Hess et al., 2005; Xie et al., 2006).

b) Ca^{2+} permeable ion channels

Ca^{2+} permeable ion channels in sperm include high-voltage-gated Ca^{2+} channels, cyclic-nucleotide-gated channels, cation channels of sperm (CatSper), and transient receptor potential channels. CATSPER 1-4 are a family of alkalinisation-activated cation channels that are activated by progesterone and are highly conserved in humans (Pereira et al., 2017). Studies have shown that mutations in these channels cause male infertility through sperm immobility, and its general importance in sperm motility in mammals (Avenarius et al., 2009; Chung et al., 2014; Hwang et al., 2021).

Mutations in genes *CATSPER*, *SLC26A8* have shown defects in capacitation processes and asthenozoospermia (Bhilawadikar et al., 2013; Dirami et al., 2013; Pereira et al., 2017; Shu et al., 2015).

c) Ca^{2+} and tyrosine phosphorylation

Protein phosphorylation is an important cellular function where a phosphate group is added to an amino acid by a protein kinase. Protein kinases thus play important roles in many signaling pathways and cellular processes. Intracellular Ca^{2+} is important for the activation of these kinases. In spermatozoa, the phosphotyrosine content of human fibrous sheath proteins are increased during capacitation highlighting the role for sperm motility (Leclerc et al., 1996; Pereira et al., 2017; Visconti & Kopf, 1998). The A-kinase anchoring (AKAP) proteins and Ca^{2+}-binding and tyrosine phosphorylation-regulated (CABYR) protein located in the fibrous sheath of human sperm are two important tyrosine-phosphorylated proteins that regulate sperm capacitation and research has shown that immotile sperm do not capacitate properly during tyrosine phosphorylation deficiency (Eddy et al., 2003; Naaby-Hansen et al., 2002).

d) Ca^{2+} and Bicarbonate ions

Bicarbonate ions are another mechanism involved in the activation of sperm motility. These ions are usually present in the female reproductive tract and enter the sperm during capacitation and required for Ca^{2+} uptake. Because local Ca^{2+} ions regulate the cAMP pathway in sperm, bicarbonate ions also influence cAMP-dependent processes. Studies have shown that bicarbonate ions facilitate Ca^{2+} entry which increases flagellar beat frequency and reduces the beat asymmetry (Luconi et al., 2005; Wennemuth et al., 2003). However, no mutations in genes associated with bicarbonate ions have been shown to be associated with asthenozoospermia.

1.5.3 Reactive oxygen species (ROS) and sperm motility

ROS are produced by diverse cells during normal metabolism, but ROS levels can be elevated during pathological situations, such as during infections accompanied by leukocyte infiltration. ROS could be produced by sperm itself in mitochondria or by nicotinamide adenine dinucleotide phosphate (NADPH) oxidation in the cytoplasm or

can arise from leukocytes' infiltration upon inflammation of the genital tract. Usually, the level of ROS is balanced by antioxidants. An excessive ROS production or low antioxidant levels could cause an imbalance known as "oxidative stress" (Pereira et al, 2017). It was shown that low levels of ROS promote the normal sperm capacitation (Aitken et al., 1997; Aitken et al., 2012).

Several studies have shown sources of excessive ROS production and their deleterious effects on sperm motility. Chronic inflammation (Pasqualotto et al., 2000), infections (Folliero et al., 2022), varicocele (Pasqualotto et al., 2008; Wang et al., 2022), unhealthy lifestyle such as smoking (Bao et al., 2019) and aging (Nguyen-Powanda & Robaire, 2020) have all shown to induce excessive ROS production with detrimental effects on the sperm motility (Pereira et al., 2017). Possible pathogenic factors are the reduction of glutathione levels or an accumulation of oxidized glutathione by inhibition of glucose-6-phosphate dehydrogenase (G6PD) involved in NADPH production. Reduced antioxidant levels lower defense mechanisms against free radicals culminating to oxidative stress (Kao et al., 2008; Pereira et al., 2017; Urata et al., 2001). High ROS levels also may result in a decrease of axoneme protein phosphorylation, which also affects sperm motility (de Lamirande & Gagnon, 1995). Moreover, excess levels of ROS can damage DNA and interfere with normal apoptotic pathways, that can cause low sperm motility.

1.5.4 Cell volume, osmolytes and sperm motility

Volume regulation mechanisms and epididymal osmolysis play a crucial role for sperm motility. Sperm face different osmolarities during maturation and ejaculation, and changes in the cell volume accordingly (Pereira et al., 2017). Spermatozoa mature in the epididymis and are stored there until ejaculation where the osmotic pressure in vas deferens is about 340 mosmol/kg and around 280±290 mosmol/kg in the female genital tract. The sperm's ability to adapt according to the environmental pressure and to change its volume appropriately is vital. Impaired volume regulation can lead to large sperm head or tail angulation ultimately affecting the movement pattern and decreasing progressive sperm motility (Cooper & Yeung, 2003; Pereira et al., 2017).

Due to limitation of cytoplasm, volume adaptation is more difficult in sperm compared to somatic cells. Regulation of cell volume in hypotonic environments mostly depends on the cytoplasmic droplet in the mid-piece of the spermatozoa. Studies have shown that sperm with abnormal or missing cytoplasmic droplets is a sign of defective spermiogenesis and that these sperms show poor motility (Xu et al., 2013). Ions such as K^+, Ca^{2+} and Cl^- are important in regulating the cell volume by their roles in different volume regulatory mechanisms (Pasantes-Morales et al., 2006; Yeung et al., 2005). Another important protein family involved in volume regulation are aquaporins (AQAP), which are thought to be involved in the cytoplasmic volume reduction during spermiogenesis (Pereira et al., 2017). Different aquaporins are present in the human flagellum: AQAP3 in principal-piece of the ejaculated sperm (Chen et al., 2011) and AQAP7 at end-piece of the ejaculated sperm (Saito et al., 2004). Reduction or absence of these aquaporins were found in sperm of infertile men and related to poor sperm migration and motility rate and these defects (Alyasin et al., 2020; Chen et al., 2011; Saito et al., 2004).

1.6 Identifying novel genes and proteins through genomic and proteogenomic methodologies.

1.6.1 The development of high-throughput technologies to measure gene expression.

DNA contains genetic information that is transcribed into RNA and is ultimately translated into proteins. The latter influences the function of cells. The transcriptome is defined as the collection of RNA molecules of a cells (or organism) that provide functional information of the genome (Kukurba & Montgomery, 2015). In order to understand the functional mRNA expression of a cell or biological sample, mRNA is reverse-transcribed to complementary DNA (cDNA) which contains exons of a transcript. The presence of a particular mRNA can be detected by converting DNA to cDNA by reverse transcription PCR (RT-PCR) and the expression level can be measured by quantitative RT-PCR. This requires designing primers according to the transcript/s sequence of a particular gene, thus can only access a limited number of genes in a given biological sample. This hurdle was overcome by micro-array technology that can target hundreds to thousands of genes simultaneously and later

was replaced by high-throughput sequencing strategies. Traditional sequencing was introduced prior to micro-array technology where the nucleotide sequence of the DNA fragment was identified (eg. Sanger sequencing). The second generation of sequencing or next-generation sequencing (NGS) opened a new horizon of transcriptome analysis by parallel processing multiple transcripts. This could either be DNA or RNA sequencing by synthesised cDNA (Kukurba & Montgomery, 2015; Querfurth et al., 2012; Wang et al., 2009).

NGS utilises many techniques including DNA-sequencing, RNA-sequencing (RNAseq), CHIP-Seq etc (Querfurth et al., 2012). The single-cell RNA-sequencing (scRNA-seq) further advanced the findings by providing genetics of a single cell which also gives information on cell and tissue development and function (Guo et al., 2018). RNA Sequencing has the ability to combine both discovery and quantification of RNA of transcripts which makes it a powerful tool (Conesa et al., 2016). The most common and major application of RNAseq is to identify differentially expressed genes between different groups (Stark et al., 2019). In combination with different biochemical assays or functional genomics, RNAseq technology has been applied to the study of many aspects of RNA biology such as RNA-protein binding, RNA structure and RNA-RNA interaction (Conesa et al., 2016; Stark et al., 2019).

For thoroughly studied, well-annotated genomes such as the human genome, the existing annotated reference transcriptome alone could be used by researchers to align their RNAseq data for analysis. Meaning, the transcripts could be mapped to the existing genome (reference genome) to identify novel transcripts and its regulation. A reference genome is a collection of previously identified and confirmed DNA sequences of genes from multiple individuals. The human reference genome HG38 is a compiled collection of all the DNA sequences scientists found. The Human genome project containing all relevant human transcripts and isoforms is available at National Center for Biotechnology Information (NCBI) and is steadily growing. There are bioinformatics repositories against which newly-sequenced DNA fragments can be aligned to identify the gene it belongs. Alongside advances in RNA sequencing technologies, advances in computational biology, bioinformatics and statistical algorithms have greatly empowered the use of NGS technology.

Even though RNAseq has greatly expanded the ability to explore and compare cell and tissue transcriptomes, some limitations exist. It was found that certain genes are not visible at the transcript level (Castellana et al., 2014; Eddy, 2001). The possibility of detecting molecular functional unit peptides (or proteins) and mapping them to an already existing proteome which is linked to genome repositories, overcame this hurdle in the recent past through quantitative mass spectrometry and proteogenomic bioinformatics (Cravatt et al, 2007).

1.6.2. The development of high-throughput methods to measure protein expression.

Mass spectrometry (MS) has developed in recent years as the preferred method for deep characterisation and understanding of proteins in biological systems (Cravatt et al., 2007; Vidova & Spacil, 2017). Similar to high-throughput RNA-sequencing methods, MS also analyses a pool of peptides simultaneously that are mapped against an already available proteome in a database or bioinformatics repository with the combination of computational biology and bioinformatics tools (Vidova & Spacil, 2017). This approach enables the identification of novel proteins that – as functional proteins – are responsible for the final phenotype. By extrapolating the peptide information mapped against the proteome and also by linking it to an established genome gives deep characterisation of the sample and predictions of the peptide identified in the sample. The comparison of protein abundances between normal vs pathological sample groups can give information on differentially expressed proteins (DEP). The differences in protein abundances are usually linked to pathological conditions. Many studies have reported this proteogenomic approach using mass spectrometry as a successful method in identifying novel protein coding regions responsible for a disease, pathway analysis etc. (Baker et al., 2013; Baldwin, 2004; Bitton et al., 2010; Castellana et al., 2014; Khan et al., 2018; Klasberg et al., 2016; Vidova & Spacil, 2017).

1.7. Summary and relevance to this thesis

Many factors contribute to male infertility, among which the genetic factors play a major role due to the complexity of genes involved in spermatogenesis. Many molecular processes and pathways involved enhance it further. Among others, two key features of sperm are in focus of male infertility research: sperm morphology and motility. There are already some genetic aberrations described leading to an impaired sperm morphology and/or motility, but it is likely that there are more, up-to now undiscovered genes involved in sperm structural and functional aberrations. High throughput technologies such as RNA sequencing and mass spectrometry detect and measure the levels of gene and protein expression as well as in identifying novels genes and proteins. The differentially expressed genes and proteins between different biological samples can be used to test and build new hypotheses on development of pathological conditions, such as male infertility as it has been performed in this thesis. Both these high throughput assays have massively contributed to the understanding and discovery of novels genes and proteins and thus have taken proteogenomics to a revolutionised level. By combining these two methods, we aimed to achieve a deep characterisation of differentially expressed genes (DEGs) and DEPs in either impaired spermatogenesis and altered sperm morphology and motility to unveil new causes for male factor infertility.

1. 8 Main research hypothesis and Objective

1.8.1 Main Hypothesis

That novel genes and proteins can be identified in testicular tissue and ejaculated sperm using well-defined biological samples exhibiting physiological and pathological conditions. That genes/proteins can have an impact on correct germ cell development and sperm morphology and motility as well as various pathologies.

1.8.2 Main Objective

- To utilise RNA sequencing and proteomics to investigate human testis biopsies and sperm preparations from men with normal germ cell development vs

impaired spermatogenesis/sperm function in order to identify novel genes and proteins involved in sperm morphology and motility.

1.9 References

Achard, V., Paulmyer-Lacroix, O., Mercier, G., Porcu, G., Saias-Magnan, J., Metzler-Guillemain, C., & Guichaoua, M. R. (2007). Reproductive failure in patients with various percentages of macronuclear spermatozoa: high level of aneuploid and polyploid spermatozoa. *J Androl*, *28*(4), 600-606. https://doi.org/10.2164/jandrol.106.001933

Aitken, R. J., Fisher, H. M., Fulton, N., Gomez, E., Knox, W., Lewis, B., & Irvine, S. (1997). Reactive oxygen species generation by human spermatozoa is induced by exogenous NADPH and inhibited by the flavoprotein inhibitors diphenylene iodonium and quinacrine. *Mol Reprod Dev*, *47*(4), 468-482. https://doi.org/10.1002/(sici)1098-2795(199708)47:4<468::Aid-mrd14>3.0.Co;2-s

Aitken, R. J., Jones, K. T., & Robertson, S. A. (2012). Reactive oxygen species and sperm function--in sickness and in health. *J Androl*, *33*(6), 1096-1106. https://doi.org/10.2164/jandrol.112.016535

Alyasin, A., Momeni, H. R., & Mahdieh, M. (2020). Aquaporin3 expression and the potential role of aquaporins in motility and mitochondrial membrane potential in human spermatozoa. *Andrologia*, *52*(6), e13588. https://doi.org/10.1111/and.13588

Avenarius, M. R., Hildebrand, M. S., Zhang, Y., Meyer, N. C., Smith, L. L., Kahrizi, K., Najmabadi, H., & Smith, R. J. (2009). Human male infertility caused by mutations in the CATSPER1 channel protein. *Am J Hum Genet*, *84*(4), 505-510. https://doi.org/10.1016/j.ajhg.2009.03.004

Avidan, N., Tamary, H., Dgany, O., Cattan, D., Pariente, A., Thulliez, M., Borot, N., Moati, L., Barthelme, A., Shalmon, L., Krasnov, T., Ben-Asher, E., Olender, T., Khen, M., Yaniv, I., Zaizov, R., Shalev, H., Delaunay, J., Fellous, M., . . . Beckmann, J. S. (2003). CATSPER2, a human autosomal nonsyndromic male infertility gene. *Eur J Hum Genet*, *11*(7), 497-502. https://doi.org/10.1038/sj.ejhg.5200991

Baccetti, B., Collodel, G., Estenoz, M., Manca, D., Moretti, E., & Piomboni, P. (2005). Gene deletions in an infertile man with sperm fibrous sheath dysplasia. *Hum Reprod*, *20*(10), 2790-2794. https://doi.org/10.1093/humrep/dei126

Bajpai, M., & Doncel, G. F. (2003). Involvement of tyrosine kinase and cAMP-dependent kinase cross-talk in the regulation of human sperm motility. *Reproduction*, *126*(2), 183-195. https://doi.org/10.1530/rep.0.1260183

Baker, M. A., Naumovski, N., Hetherington, L., Weinberg, A., Velkov, T., & Aitken, R. J. (2013). Head and flagella subcompartmental proteomic analysis of human spermatozoa. *Proteomics*, *13*(1), 61-74. https://doi.org/10.1002/pmic.201200350

Balboula, A. Z., & Schindler, K. (2014). Selective disruption of aurora C kinase reveals distinct functions from aurora B kinase during meiosis in mouse oocytes. *PLoS Genet*, *10*(2), e1004194. https://doi.org/10.1371/journal.pgen.1004194

Baldwin, M. A. (2004). Protein identification by mass spectrometry: issues to be considered. *Mol Cell Proteomics*, *3*(1), 1-9. https://doi.org/10.1074/mcp.R300012-MCP200

Bao, H. Q., Sun, L., Yang, X. N., Ding, J., Ma, S. Y., Yang, L., & Xu, X. O. (2019). [Impact of cigarette smoking on sperm quality and seminal plasma ROS in preconception males]. *Zhonghua Nan Ke Xue*, *25*(1), 41-45.

Ben Khelifa, M., Coutton, C., Zouari, R., Karaouzene, T., Rendu, J., Bidart, M., Yassine, S., Pierre, V., Delaroche, J., Hennebicq, S., Grunwald, D., Escalier, D., Pernet-Gallay, K., Jouk, P. S., Thierry-Mieg, N., Toure, A., Arnoult, C., & Ray, P. F. (2014). Mutations in DNAH1, which encodes an inner arm heavy chain dynein, lead to male infertility from multiple morphological abnormalities of the sperm flagella. *Am J Hum Genet*, *94*(1), 95-104. https://doi.org/10.1016/j.ajhg.2013.11.017

Bhilawadikar, R., Zaveri, K., Mukadam, L., Naik, S., Kamble, K., Modi, D., & Hinduja, I. (2013). Levels of Tektin 2 and CatSper 2 in normozoospermic and oligoasthenozoospermic men and its association with motility, fertilization rate, embryo quality and pregnancy rate. *J Assist Reprod Genet*, *30*(4), 513-523. https://doi.org/10.1007/s10815-013-9972-6

Bianchi, E., Boekelheide, K., Sigman, M., Braun, J. M., Eliot, M., Hall, S. J., Dere, E., & Hwang, K. (2019). Spermatozoal large RNA content is associated with semen characteristics, sociodemographic and lifestyle factors. *PLoS One*, *14*(5), e0216584. https://doi.org/10.1371/journal.pone.0216584

Bitton, D. A., Smith, D. L., Connolly, Y., Scutt, P. J., & Miller, C. J. (2010). An integrated mass-spectrometry pipeline identifies novel protein coding-regions in the human genome. *PLoS One*, *5*(1), e8949. https://doi.org/10.1371/journal.pone.0008949

Brahem, S., Mehdi, M., Elghezal, H., & Saad, A. (2012). Study of aneuploidy rate and sperm DNA fragmentation in large-headed, multiple-tailed spermatozoa. *Andrologia*, *44*(2), 130-135. https://doi.org/10.1111/j.1439-0272.2010.01115.x

Brown, P. R., Miki, K., Harper, D. B., & Eddy, E. M. (2003). A-kinase anchoring protein 4 binding proteins in the fibrous sheath of the sperm flagellum. *Biol Reprod*, *68*(6), 2241-2248. https://doi.org/10.1095/biolreprod.102.013466

Carlson, A. E., Quill, T. A., Westenbroek, R. E., Schuh, S. M., Hille, B., & Babcock, D. F. (2005). Identical phenotypes of CatSper1 and CatSper2 null sperm. *J Biol Chem*, *280*(37), 32238-32244. https://doi.org/10.1074/jbc.M501430200

Carlson, A. E., Westenbroek, R. E., Quill, T., Ren, D., Clapham, D. E., Hille, B., Garbers, D. L., & Babcock, D. F. (2003). CatSper1 required for evoked Ca2+ entry and control of flagellar function in sperm. *Proc Natl Acad Sci U S A*, *100*(25), 14864-14868. https://doi.org/10.1073/pnas.2536658100

Castellana, N. E., Shen, Z., He, Y., Walley, J. W., Cassidy, C. J., Briggs, S. P., & Bafna, V. (2014). An automated proteogenomic method uses mass spectrometry to reveal novel genes in Zea mays. *Mol Cell Proteomics*, *13*(1), 157-167. https://doi.org/10.1074/mcp.M113.031260

Chalmel, F., Lardenois, A., Evrard, B., Mathieu, R., Feig, C., Demougin, P., Gattiker, A., Schulze, W., Jegou, B., Kirchhoff, C., & Primig, M. (2012). Global human tissue profiling and protein network analysis reveals distinct levels of transcriptional germline-specificity and identifies target genes for male infertility. *Hum Reprod*, *27*(11), 3233-3248. https://doi.org/10.1093/humrep/des301

Chemes, H. E., & Rawe, V. Y. (2010). The making of abnormal spermatozoa: cellular and molecular mechanisms underlying pathological spermiogenesis. *Cell Tissue Res*, *341*(3), 349-357. https://doi.org/10.1007/s00441-010-1007-3

Chen, Q., Peng, H., Lei, L., Zhang, Y., Kuang, H., Cao, Y., Shi, Q. X., Ma, T., & Duan, E. (2011). Aquaporin3 is a sperm water channel essential for postcopulatory sperm osmoadaptation and migration. *Cell Res*, *21*(6), 922-933. https://doi.org/10.1038/cr.2010.169

Chung, J. J., Shim, S. H., Everley, R. A., Gygi, S. P., Zhuang, X., & Clapham, D. E. (2014). Structurally distinct Ca(2+) signaling domains of sperm flagella orchestrate tyrosine phosphorylation and motility. *Cell*, *157*(4), 808-822. https://doi.org/10.1016/j.cell.2014.02.056

Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., McPherson, A., Szczesniak, M. W., Gaffney, D. J., Elo, L. L., Zhang, X., & Mortazavi, A. (2016). A survey of best practices for RNA-seq data analysis. *Genome Biol*, *17*, 13. https://doi.org/10.1186/s13059-016-0881-8

Cooper, T. G., & Yeung, C. H. (2003). Acquisition of volume regulatory response of sperm upon maturation in the epididymis and the role of the cytoplasmic droplet. *Microsc Res Tech*, *61*(1), 28-38. https://doi.org/10.1002/jemt.10314

Coutton, C., Escoffier, J., Martinez, G., Arnoult, C., & Ray, P. F. (2015). Teratozoospermia: spotlight on the main genetic actors in the human. *Hum Reprod Update*, *21*(4), 455-485. https://doi.org/10.1093/humupd/dmv020

Coutton, C., Vargas, A. S., Amiri-Yekta, A., Kherraf, Z. E., Ben Mustapha, S. F., Le Tanno, P., Wambergue-Legrand, C., Karaouzene, T., Martinez, G., Crouzy, S., Daneshipour, A., Hosseini, S. H., Mitchell, V., Halouani, L., Marrakchi, O., Makni, M., Latrous, H., Kharouf, M., Deleuze, J. F., . . . Ray, P. F. (2018). Mutations in CFAP43 and CFAP44 cause male infertility and flagellum defects in Trypanosoma and human. *Nat Commun*, *9*(1), 686. https://doi.org/10.1038/s41467-017-02792-7

Cravatt, B. F., Simon, G. M., & Yates, J. R., 3rd. (2007). The biological impact of mass-spectrometry-based proteomics. *Nature*, *450*(7172), 991-1000. https://doi.org/10.1038/nature06525

Dam, A. H., Koscinski, I., Kremer, J. A., Moutou, C., Jaeger, A. S., Oudakker, A. R., Tournaye, H., Charlet, N., Lagier-Tourenne, C., van Bokhoven, H., & Viville, S. (2007). Homozygous mutation in SPATA16 is associated with male infertility in human globozoospermia. *Am J Hum Genet*, *81*(4), 813-820. https://doi.org/10.1086/521314

de Lamirande, E., & Gagnon, C. (1995). Impact of reactive oxygen species on spermatozoa: a balancing act between beneficial and detrimental effects. *Hum Reprod*, *10 Suppl 1*, 15-21. https://doi.org/10.1093/humrep/10.suppl_1.15

Dieterich, K., Soto Rifo, R., Faure, A. K., Hennebicq, S., Ben Amar, B., Zahi, M., Perrin, J., Martinez, D., Sele, B., Jouk, P. S., Ohlmann, T., Rousseaux, S., Lunardi, J., & Ray, P. F. (2007). Homozygous mutation of AURKC yields large-headed polyploid spermatozoa and causes male infertility. *Nat Genet*, *39*(5), 661-665. https://doi.org/10.1038/ng2027

Dieterich, K., Zouari, R., Harbuz, R., Vialard, F., Martinez, D., Bellayou, H., Prisant, N., Zoghmar, A., Guichaoua, M. R., Koscinski, I., Kharouf, M., Noruzinia, M., Nadifi, S., Sefiani, A., Lornage, J., Zahi, M., Viville, S., Sele, B., Jouk, P. S., . . . Ray, P. F. (2009). The Aurora Kinase C c.144delC mutation causes meiosis I arrest in men and is frequent in the North African population. *Hum Mol Genet*, *18*(7), 1301-1309. https://doi.org/10.1093/hmg/ddp029

Dirami, T., Rode, B., Jollivet, M., Da Silva, N., Escalier, D., Gaitch, N., Norez, C., Tuffery, P., Wolf, J. P., Becq, F., Ray, P. F., Dulioust, E., Gacon, G., Bienvenu, T., & Toure, A. (2013). Missense mutations in SLC26A8, encoding a sperm-specific activator of CFTR, are associated with human asthenozoospermia. *Am J Hum Genet*, *92*(5), 760-766. https://doi.org/10.1016/j.ajhg.2013.03.016

Dong, F. N., Amiri-Yekta, A., Martinez, G., Saut, A., Tek, J., Stouvenel, L., Lores, P., Karaouzene, T., Thierry-Mieg, N., Satre, V., Brouillet, S., Daneshipour, A., Hosseini, S. H., Bonhivers, M., Gourabi, H., Dulioust, E., Arnoult, C., Toure, A., Ray, P. F., . . . Coutton, C. (2018). Absence of CFAP69 Causes Male Infertility due to Multiple Morphological Abnormalities of the Flagella in Human and Mouse. *Am J Hum Genet*, *102*(4), 636-648. https://doi.org/10.1016/j.ajhg.2018.03.007

Eddy, E. M., Toshimori, K., & O'Brien, D. A. (2003). Fibrous sheath of mammalian spermatozoa. *Microsc Res Tech*, *61*(1), 103-115. https://doi.org/10.1002/jemt.10320

Eddy, S. R. (2001). Non-coding RNA genes and the modern RNA world. *Nat Rev Genet*, *2*(12), 919-929. https://doi.org/10.1038/35103511

Ellnati, E., Fossard, C., Okutman, O., Ghedir, H., Ibala-Romdhane, S., Ray, P. F., Saad, A., Hennebicq, S., & Viville, S. (2016). A new mutation identified in SPATA16 in two globozoospermic patients. *J Assist Reprod Genet*, *33*(6), 815-820. https://doi.org/10.1007/s10815-016-0715-3

Esposito, G., Jaiswal, B. S., Xie, F., Krajnc-Franken, M. A., Robben, T. J., Strik, A. M., Kuil, C., Philipsen, R. L., van Duin, M., Conti, M., & Gossen, J. A. (2004). Mice deficient for soluble adenylyl cyclase are infertile because of a severe sperm-motility defect. *Proc Natl Acad Sci U S A*, *101*(9), 2993-2998. https://doi.org/10.1073/pnas.0400050101

Fernandez-Gonzalez, A., Kourembanas, S., Wyatt, T. A., & Mitsialis, S. A. (2009). Mutation of murine adenylate kinase 7 underlies a primary ciliary dyskinesia phenotype. *Am J Respir Cell Mol Biol*, *40*(3), 305-313. https://doi.org/10.1165/rcmb.2008-0102OC

Folliero, V., Santonastaso, M., Dell'Annunziata, F., De Franciscis, P., Boccia, G., Colacurci, N., De Filippis, A., Galdiero, M., & Franci, G. (2022). Impact of Escherichia coli Outer Membrane Vesicles on Sperm Function. *Pathogens*, *11*(7). https://doi.org/10.3390/pathogens11070782

Fu, G., Wang, Q., Phan, N., Urbanska, P., Joachimiak, E., Lin, J., Wloga, D., & Nicastro, D. (2018). The I1 dynein-associated tether and tether head complex is a conserved regulator of ciliary motility. *Mol Biol Cell*, *29*(9), 1048-1059. https://doi.org/10.1091/mbc.E18-02-0142

Guo, J., Grow, E. J., Mlcochova, H., Maher, G. J., Lindskog, C., Nie, X., Guo, Y., Takei, Y., Yun, J., Cai, L., Kim, R., Carrell, D. T., Goriely, A., Hotaling, J. M., & Cairns, B. R. (2018). The adult human testis transcriptional cell atlas. *Cell Res*, *28*(12), 1141-1157. https://doi.org/10.1038/s41422-018-0099-2

Harbuz, R., Zouari, R., Pierre, V., Ben Khelifa, M., Kharouf, M., Coutton, C., Merdassi, G., Abada, F., Escoffier, J., Nikas, Y., Vialard, F., Koscinski, I., Triki, C., Sermondade, N., Schweitzer, T., Zhioua, A., Zhioua, F., Latrous, H., Halouani, L., . . . Ray, P. F. (2011). A recurrent deletion of DPY19L2 causes infertility in man by blocking sperm head elongation and acrosome formation. *Am J Hum Genet*, *88*(3), 351-361. https://doi.org/10.1016/j.ajhg.2011.02.007

Hess, K. C., Jones, B. H., Marquez, B., Chen, Y., Ord, T. S., Kamenetsky, M., Miyamoto, C., Zippin, J. H., Kopf, G. S., Suarez, S. S., Levin, L. R., Williams,

C. J., Buck, J., & Moss, S. B. (2005). The "soluble" adenylyl cyclase in sperm mediates multiple signaling events required for fertilization. *Dev Cell*, *9*(2), 249-259. https://doi.org/10.1016/j.devcel.2005.06.007

Heuser, T., Dymek, E. E., Lin, J., Smith, E. F., & Nicastro, D. (2012). The CSC connects three major axonemal complexes involved in dynein regulation. *Mol Biol Cell*, *23*(16), 3143-3155. https://doi.org/10.1091/mbc.E12-05-0357

Ho, H. C., & Suarez, S. S. (2003). Characterization of the intracellular calcium store at the base of the sperm flagellum that regulates hyperactivated motility. *Biol Reprod*, *68*(5), 1590-1596. https://doi.org/10.1095/biolreprod.102.011320

Hwang, J. Y., Maziarz, J., Wagner, G. P., & Chung, J. J. (2021). Molecular Evolution of CatSper in Mammals and Function of Sperm Hyperactivation in Gray Short-Tailed Opossum. *Cells*, *10*(5). https://doi.org/10.3390/cells10051047

Jan, S. Z., Vormer, T. L., Jongejan, A., Roling, M. D., Silber, S. J., de Rooij, D. G., Hamer, G., Repping, S., & van Pelt, A. M. M. (2017). Unraveling transcriptome dynamics in human spermatogenesis. *Development*, *144*(20), 3659-3673. https://doi.org/10.1242/dev.152413

Jungwirth, A., Diemer, T., Kopa, Z., Krausz, C., Minhas, S., & Tournaye, H. (2018). *EAU Guidelines on Male Infertility*. European Association of Urology.

Jungwirth, A., Giwercman, A., Tournaye, H., Diemer, T., Kopa, Z., Dohle, G., Krausz, C., & European Association of Urology Working Group on Male, I. (2012). European Association of Urology guidelines on Male Infertility: the 2012 update. *Eur Urol*, *62*(2), 324-332. https://doi.org/10.1016/j.eururo.2012.04.048

Kao, S. H., Chao, H. T., Chen, H. W., Hwang, T. I. S., Liao, T. L., & Wei, Y. H. (2008). Increase of oxidative stress in human sperm with lower motility. *Fertil Steril*, *89*(5), 1183-1190. https://doi.org/10.1016/j.fertnstert.2007.05.029

Karaca, N., Yilmaz, R., Kanten, G. E., Kervancioglu, E., Solakoglu, S., & Kervancioglu, M. E. (2014). First successful pregnancy in a globozoospermic patient having homozygous mutation in SPATA16. *Fertil Steril*, *102*(1), 103-107. https://doi.org/10.1016/j.fertnstert.2014.04.002

Khan, S. Y., Ali, M., Kabir, F., Chen, R., Na, C. H., Lee, M. W., Pourmand, N., Hackett, S. F., & Riazuddin, S. A. (2018). Identification of novel transcripts and peptides in developing murine lens. *Sci Rep*, *8*(1), 11162. https://doi.org/10.1038/s41598-018-28727-w

Kherraf, Z. E., Amiri-Yekta, A., Dacheux, D., Karaouzene, T., Coutton, C., Christou-Kent, M., Martinez, G., Landrein, N., Le Tanno, P., Fourati Ben Mustapha, S., Halouani, L., Marrakchi, O., Makni, M., Latrous, H., Kharouf, M., Pernet-Gallay, K., Gourabi, H., Robinson, D. R., Crouzy, S., . . . Ray, P. F. (2018). A Homozygous Ancestral SVA-Insertion-Mediated Deletion in WDR66 Induces Multiple Morphological Abnormalities of the Sperm Flagellum and Male Infertility. *Am J Hum Genet*, *103*(3), 400-412. https://doi.org/10.1016/j.ajhg.2018.07.014

Kimmins, S., Crosio, C., Kotaja, N., Hirayama, J., Monaco, L., Höög, C., van Duin, M., Gossen, J. A., & Sassone-Corsi, P. (2007). Differential functions of the Aurora-B and Aurora-C kinases in mammalian spermatogenesis. *Mol Endocrinol*, *21*(3), 726-739. https://doi.org/10.1210/me.2006-0332

Kissel, H., Georgescu, M. M., Larisch, S., Manova, K., Hunnicutt, G. R., & Steller, H. (2005). The Sept4 septin locus is required for sperm terminal differentiation in mice. *Dev Cell*, *8*(3), 353-364. https://doi.org/10.1016/j.devcel.2005.01.021

Klasberg, S., Bitard-Feildel, T., & Mallet, L. (2016). Computational Identification of Novel Genes: Current and Future Perspectives. *Bioinform Biol Insights*, *10*, 121-131. https://doi.org/10.4137/BBI.S39950

Koscinski, I., Elinati, E., Fossard, C., Redin, C., Muller, J., Velez de la Calle, J., Schmitt, F., Ben Khelifa, M., Ray, P. F., Kilani, Z., Barratt, C. L., & Viville, S. (2011). DPY19L2 deletion as a major cause of globozoospermia. *Am J Hum Genet*, *88*(3), 344-350. https://doi.org/10.1016/j.ajhg.2011.01.018

Krausz, C., & Sassone-Corsi, P. (2005). Genetic control of spermiogenesis: insights from the CREM gene and implications for human infertility. *Reproductive biomedicine online*, *10*(1), 64-71. https://doi.org/10.1016/s1472-6483(10)60805-x

Kubo, T., Hou, Y., Cochran, D. A., Witman, G. B., & Oda, T. (2018). A microtubule-dynein tethering complex regulates the axonemal inner dynein f (I1). *Mol Biol Cell*, *29*(9), 1060-1074. https://doi.org/10.1091/mbc.E17-11-0689

Kukurba, K. R., & Montgomery, S. B. (2015). RNA Sequencing and Analysis. *Cold Spring Harb Protoc*, *2015*(11), 951-969. https://doi.org/10.1101/pdb.top084970

Kuo, Y. C., Lin, Y. H., Chen, H. I., Wang, Y. Y., Chiou, Y. W., Lin, H. H., Pan, H. A., Wu, C. M., Su, S. M., Hsu, C. C., & Kuo, P. L. (2012). SEPT12 mutations cause male infertility with defective sperm annulus. *Hum Mutat*, *33*(4), 710-719. https://doi.org/10.1002/humu.22028

Leclerc, P., de Lamirande, E., & Gagnon, C. (1996). Cyclic adenosine 3',5'monophosphate-dependent regulation of protein tyrosine phosphorylation in relation to human sperm capacitation and motility. *Biol Reprod*, *55*(3), 684-692. https://doi.org/10.1095/biolreprod55.3.684

Lin, Y. H., Lin, Y. M., Teng, Y. N., Hsieh, T. Y., Lin, Y. S., & Kuo, P. L. (2006). Identification of ten novel genes involved in human spermatogenesis by microarray analysis of testicular tissue. *Fertil Steril*, *86*(6), 1650-1658. https://doi.org/10.1016/j.fertnstert.2006.04.039

Lin, Y. H., Lin, Y. M., Wang, Y. Y., Yu, I. S., Lin, Y. W., Wang, Y. H., Wu, C. M., Pan, H. A., Chao, S. C., Yen, P. H., Lin, S. W., & Kuo, P. L. (2009). The expression level of septin12 is critical for spermiogenesis. *Am J Pathol*, *174*(5), 1857-1868. https://doi.org/10.2353/ajpath.2009.080955

Lin, Y. H., Wang, Y. Y., Chen, H. I., Kuo, Y. C., Chiou, Y. W., Lin, H. H., Wu, C. M., Hsu, C. C., Chiang, H. S., & Kuo, P. L. (2012). SEPTIN12 genetic variants confer susceptibility to teratozoospermia. *PLoS One*, *7*(3), e34011. https://doi.org/10.1371/journal.pone.0034011

Linck, R., Fu, X., Lin, J., Ouch, C., Schefter, A., Steffen, W., Warren, P., & Nicastro, D. (2014). Insights into the structure and function of ciliary and flagellar doublet microtubules: tektins, Ca2+-binding proteins, and stable protofilaments. *J Biol Chem*, *289*(25), 17427-17444. https://doi.org/10.1074/jbc.M114.568949

Liu, C., Tu, C., Wang, L., Wu, H., Houston, B. J., Mastrorosa, F. K., Zhang, W., Shen, Y., Wang, J., Tian, S., Meng, L., Cong, J., Yang, S., Jiang, Y., Tang, S., Zeng, Y., Lv, M., Lin, G., Li, J., . . . Zhang, F. (2021). Deleterious variants in X-linked CFAP47 induce asthenoteratozoospermia and primary male infertility. *Am J Hum Genet*, *108*(2), 309-323. https://doi.org/10.1016/j.ajhg.2021.01.002

Liu, G., Shi, Q. W., & Lu, G. X. (2010). A newly discovered mutation in PICK1 in a human with globozoospermia. *Asian J Androl*, *12*(4), 556-560. https://doi.org/10.1038/aja.2010.47

Lores, P., Coutton, C., Khouri, E. E., Stouvenel, L., Givelet, M., Thomas, L., Rode, B., Schmitt, A., Louis, B., Sakheli, Z., Chaudhry, M., Fernandez-Gonzales, A.,

Mitsialis, A., Dacheux, D., Wolf, J. P., Papon, J. F., Gacon, G., Escudier, E., Arnoult, C., . . . Toure, A. (2019). Corrigendum: Homozygous missense mutation L673P in adenylate kinase 7 (AK7) leads to primary male infertility and multiple morphological anomalies of the flagella but not to primary ciliary dyskinesia. *Hum Mol Genet*, *28*(6), 1052. https://doi.org/10.1093/hmg/ddy368

Lu, L., Lin, M., Xu, M., Zhou, Z. M., & Sha, J. H. (2006). Gene functional research using polyethylenimine-mediated in vivo gene transfection into mouse spermatogenic cells. *Asian J Androl*, *8*(1), 53-59. https://doi.org/10.1111/j.1745-7262.2006.00089.x

Luconi, M., Porazzi, I., Ferruzzi, P., Marchiani, S., Forti, G., & Baldi, E. (2005). Tyrosine phosphorylation of the a kinase anchoring protein 3 (AKAP3) and soluble adenylate cyclase are involved in the increase of human sperm motility by bicarbonate. *Biol Reprod*, *72*(1), 22-32. https://doi.org/10.1095/biolreprod.104.032490

Lv, M., Liu, W., Chi, W., Ni, X., Wang, J., Cheng, H., Li, W. Y., Yang, S., Wu, H., Zhang, J., Gao, Y., Liu, C., Li, C., Yang, C., Tan, Q., Tang, D., Zhang, J., Song, B., Chen, Y. J., . . . Cao, Y. (2020). Homozygous mutations in DZIP1 can induce asthenoteratospermia with severe MMAF. *J Med Genet*, *57*(7), 445-453. https://doi.org/10.1136/jmedgenet-2019-106479

Malcher, A., Stokowy, T., Berman, A., Olszewska, M., Jedrzejczak, P., Sielski, D., Nowakowski, A., Rozwadowska, N., Yatsenko, A. N., & Kurpisz, M. K. (2022). Whole-genome sequencing identifies new candidate genes for nonobstructive azoospermia. *Andrology*, *10*(8), 1605-1624. https://doi.org/10.1111/andr.13269

Martinez, G., Kherraf, Z. E., Zouari, R., Fourati Ben Mustapha, S., Saut, A., Pernet-Gallay, K., Bertrand, A., Bidart, M., Hograindleur, J. P., Amiri-Yekta, A., Kharouf, M., Karaouzène, T., Thierry-Mieg, N., Dacheux-Deschamps, D., Satre, V., Bonhivers, M., Touré, A., Arnoult, C., Ray, P. F., & Coutton, C. (2018). Whole-exome sequencing identifies mutations in FSIP2 as a recurrent cause of multiple morphological abnormalities of the sperm flagella. *Hum Reprod*, *33*(10), 1973-1984. https://doi.org/10.1093/humrep/dey264

Matzuk, M. M., & Lamb, D. J. (2008). The biology of infertility: research advances and clinical challenges. *Nat Med*, *14*(11), 1197-1213. https://doi.org/10.1038/nm.f.1895

Naaby-Hansen, S., Mandal, A., Wolkowicz, M. J., Sen, B., Westbrook, V. A., Shetty, J., Coonrod, S. A., Klotz, K. L., Kim, Y. H., Bush, L. A., Flickinger, C. J., & Herr, J. C. (2002). CABYR, a novel calcium-binding tyrosine phosphorylation-regulated fibrous sheath protein involved in capacitation. *Dev Biol*, *242*(2), 236-254. https://doi.org/10.1006/dbio.2001.0527

Neesen, J., Kirschner, R., Ochs, M., Schmiedl, A., Habermann, B., Mueller, C., Holstein, A. F., Nuesslein, T., Adham, I., & Engel, W. (2001). Disruption of an inner arm dynein heavy chain gene results in asthenozoospermia and reduced ciliary beat frequency. *Hum Mol Genet*, *10*(11), 1117-1128. https://doi.org/10.1093/hmg/10.11.1117

Nguyen-Powanda, P., & Robaire, B. (2020). Oxidative Stress and Reproductive Function in the Aging Male. *Biology (Basel)*, *9*(9). https://doi.org/10.3390/biology9090282

Nieschlag, E. (2010). Scope and Goals of Andrology. In E. Nieschlag, H. M. Hermann M. Behre, & S. Nieschlag (Eds.), *Andrology*

Male Reproductive Health and Dysfunction (3rd Editon ed., pp. 7). Springer. https://doi.org/10.1007/978-3-540-78355-8

Nsota Mbango, J. F., Coutton, C., Arnoult, C., Ray, P. F., & Toure, A. (2019). Genetic causes of male infertility: snapshot on morphological abnormalities of the sperm flagellum. *Basic Clin Androl*, *29*, 2. https://doi.org/10.1186/s12610-019-0083-9

O'Donnell, L. (2014). Mechanisms of spermiogenesis and spermiation and how they are disturbed. *Spermatogenesis*, *4*(2), e979623. https://doi.org/10.4161/21565562.2014.979623

O'Flynn O'Brien, K. L., Varghese, A. C., & Agarwal, A. (2010). The genetic causes of male factor infertility: a review. *Fertil Steril*, *93*(1), 1-12. https://doi.org/10.1016/j.fertnstert.2009.10.045

Okutman, O., Muller, J., Baert, Y., Serdarogullari, M., Gultomruk, M., Piton, A., Rombaut, C., Benkhalifa, M., Teletin, M., Skory, V., Bakircioglu, E., Goossens, E., Bahceci, M., & Viville, S. (2015). Exome sequencing reveals a nonsense mutation in TEX15 causing spermatogenic failure in a Turkish family. *Hum Mol Genet*, *24*(19), 5581-5588. https://doi.org/10.1093/hmg/ddv290

Pasantes-Morales, H., Lezama, R. A., Ramos-Mandujano, G., & Tuz, K. L. (2006). Mechanisms of cell volume regulation in hypo-osmolality. *Am J Med*, *119*(7 Suppl 1), S4-11. https://doi.org/10.1016/j.amjmed.2006.05.002

Pasqualotto, F. F., Sharma, R. K., Potts, J. M., Nelson, D. R., Thomas, A. J., & Agarwal, A. (2000). Seminal oxidative stress in patients with chronic prostatitis. *Urology*, *55*(6), 881-885. https://doi.org/10.1016/s0090-4295(99)00613-5

Pasqualotto, F. F., Sundaram, A., Sharma, R. K., Borges, E., Jr., Pasqualotto, E. B., & Agarwal, A. (2008). Semen quality and oxidative stress scores in fertile and infertile patients with varicocele. *Fertil Steril*, *89*(3), 602-607. https://doi.org/10.1016/j.fertnstert.2007.03.057

Patat, O., Pagin, A., Siegfried, A., Mitchell, V., Chassaing, N., Faguer, S., Monteil, L., Gaston, V., Bujan, L., Courtade-Saidi, M., Marcelli, F., Lalau, G., Rigot, J. M., Mieusset, R., & Bieth, E. (2016). Truncating Mutations in the Adhesion G Protein-Coupled Receptor G2 Gene ADGRG2 Cause an X-Linked Congenital Bilateral Absence of Vas Deferens. *Am J Hum Genet*, *99*(2), 437-442. https://doi.org/10.1016/j.ajhg.2016.06.012

Pazour, G. J., Agrin, N., Leszyk, J., & Witman, G. B. (2005). Proteomic analysis of a eukaryotic cilium. *J Cell Biol*, *170*(1), 103-113. https://doi.org/10.1083/jcb.200504008

Pazour, G. J., Dickert, B. L., & Witman, G. B. (1999). The DHC1b (DHC2) isoform of cytoplasmic dynein is required for flagellar assembly. *J Cell Biol*, *144*(3), 473-481. https://doi.org/10.1083/jcb.144.3.473

Pereira, R., Sa, R., Barros, A., & Sousa, M. (2017). Major regulatory mechanisms involved in sperm motility. *Asian J Androl*, *19*(1), 5-14. https://doi.org/10.4103/1008-682X.167716

Pierre, V., Martinez, G., Coutton, C., Delaroche, J., Yassine, S., Novella, C., Pernet-Gallay, K., Hennebicq, S., Ray, P. F., & Arnoult, C. (2012). Absence of Dpy19l2, a new inner nuclear membrane protein, causes globozoospermia in mice by preventing the anchoring of the acrosome to the nucleus. *Development*, *139*(16), 2955-2965. https://doi.org/10.1242/dev.077982

Pleuger, C., Lehti, M. S., Dunleavy, J. E., Fietz, D., & O'Bryan, M. K. (2020). Haploid male germ cells-the Grand Central Station of protein transport. *Hum Reprod Update*, *26*(4), 474-500. https://doi.org/10.1093/humupd/dmaa004

Querfurth, R., Fischer, A., Schweiger, M. R., Lehrach, H., & Mertes, F. (2012). Creation and application of immortalized bait libraries for targeted enrichment and next-generation sequencing. *Biotechniques*, *52*(6), 375-380. https://doi.org/10.2144/0000113877

Quill, T. A., Sugden, S. A., Rossi, K. L., Doolittle, L. K., Hammer, R. E., & Garbers, D. L. (2003). Hyperactivated sperm motility driven by CatSper2 is required for fertilization. *Proc Natl Acad Sci U S A*, *100*(25), 14869-14874. https://doi.org/10.1073/pnas.2136654100

Ramasamy, R., Bakircioglu, M. E., Cengiz, C., Karaca, E., Scovell, J., Jhangiani, S. N., Akdemir, Z. C., Bainbridge, M., Yu, Y., Huff, C., Gibbs, R. A., Lupski, J. R., & Lamb, D. J. (2015). Whole-exome sequencing identifies novel homozygous mutation in NPAS2 in family with nonobstructive azoospermia. *Fertil Steril*, *104*(2), 286-291. https://doi.org/10.1016/j.fertnstert.2015.04.001

Ray, P. F., Toure, A., Metzler-Guillemain, C., Mitchell, M. J., Arnoult, C., & Coutton, C. (2017). Genetic abnormalities leading to qualitative defects of sperm morphology or function. *Clin Genet*, *91*(2), 217-232. https://doi.org/10.1111/cge.12905

Ren, D., Navarro, B., Perez, G., Jackson, A. C., Hsu, S., Shi, Q., Tilly, J. L., & Clapham, D. E. (2001). A sperm ion channel required for sperm motility and male fertility. *Nature*, *413*(6856), 603-609. https://doi.org/10.1038/35098027

Saito, K., Kageyama, Y., Okada, Y., Kawakami, S., Kihara, K., Ishibashi, K., & Sasaki, S. (2004). Localization of aquaporin-7 in human testis and ejaculated sperm: possible involvement in maintenance of sperm quality. *J Urol*, *172*(5 Pt 1), 2073-2076. https://doi.org/10.1097/01.ju.0000141499.08650.ab

Schindler, K., Davydenko, O., Fram, B., Lampson, M. A., & Schultz, R. M. (2012). Maternally recruited Aurora C kinase is more stable than Aurora B to support mouse oocyte maturation and early development. *Proc Natl Acad Sci U S A*, *109*(33), E2215-2222. https://doi.org/10.1073/pnas.1120517109

Shu, F., Zhou, X., Li, F., Lu, D., Lei, B., Li, Q., Yang, Y., Yang, X., Shi, R., & Mao, X. (2015). Analysis of the correlation of CATSPER single nucleotide polymorphisms (SNPs) with idiopathic asthenospermia. *J Assist Reprod Genet*, *32*(11), 1643-1649. https://doi.org/10.1007/s10815-015-0548-5

Soumillon, M., Necsulea, A., Weier, M., Brawand, D., Zhang, X., Gu, H., Barthes, P., Kokkinaki, M., Nef, S., Gnirke, A., Dym, M., de Massy, B., Mikkelsen, T. S., & Kaessmann, H. (2013). Cellular source and mechanisms of high transcriptome complexity in the mammalian testis. *Cell Rep*, *3*(6), 2179-2190. https://doi.org/10.1016/j.celrep.2013.05.031

Stark, R., Grzelak, M., & Hadfield, J. (2019). RNA sequencing: the teenage years. *Nat Rev Genet*, *20*(11), 631-656. https://doi.org/10.1038/s41576-019-0150-2

Steels, J. D., Estey, M. P., Froese, C. D., Reynaud, D., Pace-Asciak, C., & Trimble, W. S. (2007). Sept12 is a component of the mammalian sperm tail annulus. *Cell Motil Cytoskeleton*, *64*(10), 794-807. https://doi.org/10.1002/cm.20224

Suarez, S. S. (2008). Control of hyperactivation in sperm. *Hum Reprod Update*, *14*(6), 647-657. https://doi.org/10.1093/humupd/dmn029

Suarez, S. S., Varosi, S. M., & Dai, X. (1993). Intracellular calcium increases with hyperactivation in intact. *Proc. Natl. Acad. Sci. USA*, *90*, 4660-4664. https://doi.org/doi: 10.1073/pnas.90.10.4660

Talaga, A. K., Dong, F. N., Reisert, J., & Zhao, H. (2017). Cilia- and Flagella-Associated Protein 69 Regulates Olfactory Transduction Kinetics in Mice. *J*

Neurosci, *37*(23), 5699-5710. https://doi.org/10.1523/JNEUROSCI.0392-17.2017

Touré, A., Lhuillier, P., Gossen, J. A., Kuil, C. W., Lhôte, D., Jégou, B., Escalier, D., & Gacon, G. (2007). The testis anion transporter 1 (Slc26a8) is required for sperm terminal differentiation and male fertility in the mouse. *Hum Mol Genet*, *16*(15), 1783-1793. https://doi.org/10.1093/hmg/ddm117

Tüttelmann, F., Rajpert-De Meyts, E., Eberhard Nieschlag, E., & Simoni, M. (2007). Gene polymorphisms and male infertility – a meta analysis and mini review. *Reproductive biomedicine online*, *15*, 643-658. https://doi.org/10.1016/s1472-6483(10)60531-7.

Tüttelmann, F., Ruckert, C., & Ropke, A. (2018). Disorders of spermatogenesis: Perspectives for novel genetic diagnostics after 20 years of unchanged routine. *Med Genet*, *30*(1), 12-20. https://doi.org/10.1007/s11825-018-0181-7

Uhlén, M., Fagerberg, L., Hallstrom, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, A., Kampf, C., Sjostedt, E., Asplund, A., Olsson, I., Edlund, K., Lundberg, E., Navani, S., Szigyarto, C. A., Odeberg, J., Djureinovic, D., Takanen, J. O., Hober, S., . . . Ponten, F. (2015). Proteomics. Tissue-based map of the human proteome. *Science*, *347*(6220), 1260419. https://doi.org/10.1126/science.1260419

Urata, K., Narahara, H., Tanaka, Y., Egashira, T., Takayama, F., & Miyakawa, I. (2001). Effect of endotoxin-induced reactive oxygen species on sperm motility. *Fertil Steril*, *76*(1), 163-166. https://doi.org/10.1016/s0015-0282(01)01850-7

Urbanska, P., Joachimiak, E., Bazan, R., Fu, G., Poprzeczko, M., Fabczak, H., Nicastro, D., & Wloga, D. (2018). Ciliary proteins Fap43 and Fap44 interact with each other and are essential for proper cilia and flagella beating. *Cell Mol Life Sci*, *75*(24), 4479-4493. https://doi.org/10.1007/s00018-018-2819-7

Urbanska, P., Song, K., Joachimiak, E., Krzemien-Ojak, L., Koprowski, P., Hennessey, T., Jerka-Dziadosz, M., Fabczak, H., Gaertig, J., Nicastro, D., & Wloga, D. (2015). The CSC proteins FAP61 and FAP251 build the basal substructures of radial spoke 3 in cilia. *Mol Biol Cell*, *26*(8), 1463-1475. https://doi.org/10.1091/mbc.E14-11-1545

Vicari, E., de Palma, A., Burrello, N., Longo, G., Grazioso, C., Barone, N., Zahi, M., D'Agata, R., & Calogero, A. E. (2003). Absolute polymorphic teratozoospermia in patients with oligo-asthenozoospermia is associated with an elevated sperm aneuploidy rate. *J Androl*, *24*(4), 598-603. https://doi.org/10.1002/j.1939-4640.2003.tb02711.x

Vidova, V., & Spacil, Z. (2017). A review on mass spectrometry-based quantitative proteomics: Targeted and data independent acquisition. *Anal Chim Acta*, *964*, 7-23. https://doi.org/10.1016/j.aca.2017.01.059

Visconti, P. E., & Kopf, G. S. (1998). Regulation of protein phosphorylation during sperm capacitation. *Biol Reprod*, *59*(1), 1-6. https://doi.org/10.1095/biolreprod59.1.1

Vogt, P. H. (2004). Molecular Genetic of Human Male Infertility: From Genes to New Therapeutic Perspectives. *Current pharmaceutical design*. https://doi.org/10.2174/1381612043453261

Wang, K., Gao, Y., Wang, C., Liang, M., Liao, Y., & Hu, K. (2022). Role of Oxidative Stress in Varicocele. *Front Genet*, *13*, 850114. https://doi.org/10.3389/fgene.2022.850114

Wang, X., Jin, H., Han, F., Cui, Y., Chen, J., Yang, C., Zhu, P., Wang, W., Jiao, G., Wang, W., Hao, C., & Gao, Z. (2017). Homozygous DNAH1 frameshift mutation causes multiple morphological anomalies of the sperm flagella in Chinese. *Clin Genet*, *91*(2), 313-321. https://doi.org/10.1111/cge.12857

Wang, Z., Gerstein, M., & Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet*, *10*(1), 57-63. https://doi.org/10.1038/nrg2484

Washburn, R. L., Hibler, T., Thompson, L. A., Kaur, G., & Dufour, J. M. (2022). Therapeutic application of Sertoli cells for treatment of various diseases. *Semin Cell Dev Biol*, *121*, 10-23. https://doi.org/10.1016/j.semcdb.2021.04.007

Weinbauer, G. F., Luetjens, C. M., Simoni, M., & Nieschlag, E. (2010). Physiology of Testicular Function. In *Andrology Male Reproductive Health and Dysfunction* (3rd Edition ed., pp. 11-60). Springer. https://doi.org/ 10.1007/978-3-540-78355-8

Wennemuth, G., Carlson, A. E., Harper, A. J., & Babcock, D. F. (2003). Bicarbonate actions on flagellar and Ca2+ -channel responses: initial events in sperm activation. *Development*, *130*(7), 1317-1326. https://doi.org/10.1242/dev.00353

World Health Organization. (2021). *WHO laboratory manual for the examination and processing of human semen*. WHO. https://www.who.int/publications/i/item/9789240030787

Xie, F., Garcia, M. A., Carlson, A. E., Schuh, S. M., Babcock, D. F., Jaiswal, B. S., Gossen, J. A., Esposito, G., van Duin, M., & Conti, M. (2006). Soluble adenylyl cyclase (sAC) is indispensable for sperm function and fertilization. *Dev Biol*, *296*(2), 353-362. https://doi.org/10.1016/j.ydbio.2006.05.038

Xu, H., Yuan, S. Q., Zheng, Z. H., & Yan, W. (2013). The cytoplasmic droplet may be indicative of sperm motility and normal spermiogenesis. *Asian J Androl*, *15*(6), 799-805. https://doi.org/10.1038/aja.2013.69

Yang, P., Diener, D. R., Yang, C., Kohno, T., Pazour, G. J., Dienes, J. M., Agrin, N. S., King, S. M., Sale, W. S., Kamiya, R., Rosenbaum, J. L., & Witman, G. B. (2006). Radial spoke proteins of Chlamydomonas flagella. *J Cell Sci*, *119*(Pt 6), 1165-1174. https://doi.org/10.1242/jcs.02811

Yassine, S., Escoffier, J., Abi Nahed, R., Pierre, V., Karaouzene, T., Ray, P. F., & Arnoult, C. (2015). Dynamics of Sun5 localization during spermatogenesis in wild type and Dpy19l2 knock-out mice indicates that Sun5 is not involved in acrosome attachment to the nuclear envelope. *PLoS One*, *10*(3), e0118698. https://doi.org/10.1371/journal.pone.0118698

Yassine, S., Escoffier, J., Martinez, G., Coutton, C., Karaouzene, T., Zouari, R., Ravanat, J. L., Metzler-Guillemain, C., Lee, H. C., Fissore, R., Hennebicq, S., Ray, P. F., & Arnoult, C. (2015). Dpy19l2-deficient globozoospermic sperm display altered genome packaging and DNA damage that compromises the initiation of embryo development. *Mol Hum Reprod*, *21*(2), 169-185. https://doi.org/10.1093/molehr/gau099

Yatsenko, A. N., O'Neil, D. S., Roy, A., Arias-Mendoza, P. A., Chen, R., Murthy, L. J., Lamb, D. J., & Matzuk, M. M. (2012). Association of mutations in the zona pellucida binding protein 1 (ZPBP1) gene with abnormal sperm head morphology in infertile men. *Mol Hum Reprod*, *18*(1), 14-21. https://doi.org/10.1093/molehr/gar057

Yeung, C. H., Barfield, J. P., & Cooper, T. G. (2005). The role of anion channels and Ca2+ in addition to K+ channels in the physiological volume regulation of

murine spermatozoa. *Mol Reprod Dev*, *71*(3), 368-379. https://doi.org/10.1002/mrd.20261

Yuan, S., Stratton, C. J., Bao, J., Zheng, H., Bhetwal, B. P., Yanagimachi, R., & Yan, W. (2015). Spata6 is required for normal assembly of the sperm connecting piece and tight head-tail conjunction. *Proc Natl Acad Sci U S A*, *112*(5), E430-439. https://doi.org/10.1073/pnas.1424648112

Zhang, Y., Malekpour, M., Al-Madani, N., Kahrizi, K., Zanganeh, M., Lohr, N. J., Mohseni, M., Mojahedi, F., Daneshi, A., Najmabadi, H., & Smith, R. J. (2007). Sensorineural deafness and male infertility: a contiguous gene deletion syndrome. *J Med Genet*, *44*(4), 233-240. https://doi.org/10.1136/jmg.2006.045765

Zhu, Z., Li, C., Yang, S., Tian, R., Wang, J., Yuan, Q., Dong, H., He, Z., Wang, S., & Li, Z. (2016). Dynamics of the Transcriptome during Human Spermatogenesis: Predicting the Potential Key Genes Regulating Male Gametes Generation. *Sci Rep*, *6*, 19069. https://doi.org/10.1038/srep19069

Chapter 2

Identification of Differentially Expressed Genes in Human Testis Biopsies with Defective Spermatogenesis

Amended version of the manuscript submitted to the journal ‘’Andrology’’. The manuscript is under revision after the first review.

Submission number - ANDR-2023-0267

Title

Identification of Differentially Expressed Genes in Human Testis Biopsies with Defective Spermatogenesis

Authors

Shashika D. Kothalawala[1,6], Stefan Günther[2], Hans-Christian Schuppe[3,4], Adrian Pilatz[3,4], Florian Wagenlehner[3,4], Sabine Kliesch[5], Liza O'Donnell[6], Daniela Fietz[1,4]

Affiliations

1 - Institute for Veterinary Anatomy, Histology and Embryology, Justus-Liebig University of Giessen, Germany.

2 - Max-Planck Institute for Heart and Lung Research, Bad Nauheim, Germany.

3 - Clinic of Urology, Pediatric Urology and Andrology, Justus-Liebig University of Giessen, Germany.

4 - Hessian Centre of Reproductive Medicine, Justus Liebig University, Giessen, Germany

5 - Centre of Reproductive Medicine and Andrology, University of Muenster, Muenster, Germany

6 - Centre for Reproductive Health, Hudson Institute of Medical Research, & Dept. of Molecular & Translational Sciences, Monash University, Clayton, Victoria, Australia

Corresponding Author

Shashika D. Kothalawala, Institute for Veterinary Anatomy, Histology and Embryology, Frankfurter Strasse 98, 35392, Giessen, Germany. Email - Shashika.kothalawala@monash.edu or Shashika.D.Kothalawala@vetmed.uni-giessen.de

Funding

The study was funded by the German Research Foundation (IRTG *DFG GRK 1871/1*)

Abstract

Background: Genetic abnormalities cause approximately 30-40% of human male infertility cases. The complex process of spermatogenesis involves thousands of genes which undergo remarkable changes in transcriptional activity at each step.

Objective: To elucidate differentially expressed transcripts in human testis tissues of normal and abnormal spermatogenesis that could reveal new genes that may regulate sperm morphology and function.

Material and methods: Human testis biopsies were collected from men with well-characterised phenotypes of normal spermatogenesis, spermatid arrest (spermatogenic arrest during spermatid development), and Sertoli cell-only phenotype (germ cell arrest) (n=2-3 per group), and transcriptional differences were quantified by RNA sequencing (RNAseq). To identify genes that could regulate sperm morphology and/or function, differentially expressed genes were filtered based on predominant expression in spermatids (highly downregulated in spermatid arrest and Sertoli cell-only phenotypes vs normal spermatogenesis and evidence for enrichment in human germ cells and spermatids based on published datasets) and gene functional annotations relevant to sperm morphology and motility, including calcium and cAMP regulation and flagella function. The differential expression of 10 selected genes was validated by qRT-PCR and the localisation of 2 proteins was determined in testis biopsies.

Results: Several genes with potential roles in sperm morphology and/or motility were identified based on their expression patterns and functions, and their changes in the RNAseq dataset were independently validated by qRT-PCR. The analysis revealed 6 genes (*SPATA31E1, TEKT3, SLC9C1, PDE4A, CFAP47*, and *TNC*) that are excellent candidates for novel genes enriched in developing human sperm. The immunohistochemical localisation of two proteins, ORAI1 and SPATA31E1, in testis biopsies verified that both are expressed in developing human germ cells, with SPATA31E1 enriched in late spermatocytes and spermatids.

Conclusion: This study identified human germ cell-enriched genes that could play functional roles in spermiogenesis and could thus be important in the development of morphologically normal, motile sperm.

Keywords – RNA sequencing; Spermatogenesis; Spermiogenesis; Sperm morphology; Sperm motility

2.1 Introduction

Infertility, i.e. the inability to conceive within one year (Jungwirth et al., 2012; Minhas et al., 2021), has become a global issue with an increase in prevalence in recent years, with 10 – 15% of couples suffering from infertility problems (Tüttelmann et al., 2018). Male factor infertility is prevalent in approximately half of infertile couples (Agarwal et al., 2015) and can be caused by various factors such as congenital or acquired urogenital abnormalities, malignancies, urogenital tract infections, endocrine disturbances, and genetic abnormalities (Minhas et al., 2021). In around 30% to 40% of cases, the cause of infertility remains unclear and is thus referred to as idiopathic infertility (Lin et al., 2006; Panner Selvam et al., 2019).

Ejaculate analysis as standardised by the World Health Organization (World Health Organization, 2021) is the gold standard method to assess male fertility. Abnormalities in sperm number are defined as oligozoospermia (low sperm count < 15 million/mL) or azoospermia (no sperm in the ejaculate). Impaired motility defects are defined as asthenozoospermia, i.e. low progressive motility, and sperm morphology defects are defined as teratozoospermia, where sperm exhibit a wide range of defined structural abnormalities. The most common diagnosis is a combination of all three abnormalities and thus classified as oligoasthenoteratozoospermia (OAT), where both qualitative and quantitative abnormalities of sperm number and function are detected (Tüttelmann et al., 2018). Morphological and functional abnormalities can be broadly divided into sperm head and flagellum defects(Ray et al., 2017). Sperm abnormalities can arise during sperm development in the testis (spermatogenesis and/or spermiogenesis) or during passage through the epididymis.

Normal sperm motility and morphology rely on a large number of genes (Pereira et al., 2017), and germ cell development, particularly of spermatocytes and spermatids, involves remarkable transcriptional complexity and diversity (Soumillon et al., 2013). Up to 2000 genes are involved in male gametogenesis (Chalmel et al., 2012; Matzuk & Lamb, 2008; Tüttelmann et al., 2018). The causes of defects in sperm number and/or function are complex, with potentially 30 to 40% of idiopathic infertility cases caused

by genetic abnormalities such as chromosomal defects or gene mutations (Houston et al., 2021; Tüttelmann et al., 2007; Tüttelmann et al., 2018; Vogt, 2004). So far, relatively a few clinically relevant genes have been identified (Krausz et al., 2020; McLachlan & O'Bryan, 2010; Tüttelmann et al., 2018).

Identifying genes that are essential for human male fertility will facilitate genetic counselling of patients, infertility diagnosis, and selection of the best treatment (Coutton et al., 2015; Krausz et al., 2020; O'Flynn O'Brien et al., 2010). Importantly, elucidating genes that are specific to male germ cells and are essential for fertility also provides new opportunities for the development of non-hormonal male contraceptive approaches. Sperm production involves many genes that are highly restricted to germ cells and are not found elsewhere in the body (Chalmel et al., 2007; Soumillon et al., 2013) thus contraceptive strategies that target genes restricted to spermatid development could be highly specific and effective, with limited off-target effects. The current study focused on identifying novel genes with a potential role in the development of human spermatozoa. The study used histologically well-defined testis samples from men with azoospermia and different levels of spermatogenic failure and analysed transcriptional differences using a defined analysis pipeline to identify novel genes with potential roles in impaired sperm morphology and motility.

2.1.1 Hypothesis

- That altered gene expression between phenotypically classified testis biopsies will reveal genes involved in defective sperm morphology and motility and thus male fertility.

2.1.2 Research Objectives

2.1.2.1 Main Objective

To identify novel genes that are involved in altered sperm morphology and motility in men.

2.1.2.2 Specific Objectives

- To identify differences in mRNA transcripts in frozen testis biopsies from azoospermic individuals with normal spermatogenesis (obstructive

azoospermia), classified as normal spermatogenesis (NSP) vs azoospermic individuals with defective spermatogenesis (non-obstructive azoospermia), classified as spermatid arrest (SDA) and Sertoli cell only (SCO) by RNAseq.

- To design a bioinformatics strategy to identify novel genes that could be responsible for sperm morphology and motility defects in infertility.
- To validate differential expression of candidate genes by reverse transcription polymerase chain reaction (RT-PCR) and quantitative reverse transcription polymerase chain reaction (qRT-PCR).
- To further analyse the protein localisation of selected genes in testis biopsies by immunochemistry (IHC).

2.1.3 Experimental Design

Fresh-frozen human testis biopsies which are histologically well-defined as normal/ intact spermatogenesis (NSP – normal spermatogenesis) and defective spermatogenesis (SDA- spermatid arrest, SCO – Sertoli cell only) were selected for the study. Extracted RNA was subjected to RNA sequencing at the Max Planck Institute, Bad Nauheim, Germany. Transcriptomics data were analysed and a preliminary gene cluster was selected (explained in the methodology). The expression of transcripts in different pathological testis tissue groups were determined by RT-PCR and subsequently qRT-PCR. Proteins encoded by two genes for which reliable antibodies were available were determined in testis biopsies by IHC (Fig. 2.1). Other than the RNA sequencing and statistical analysis, the produced lists of DEGs and all other experiments and analyses were performed by the PhD candidate.

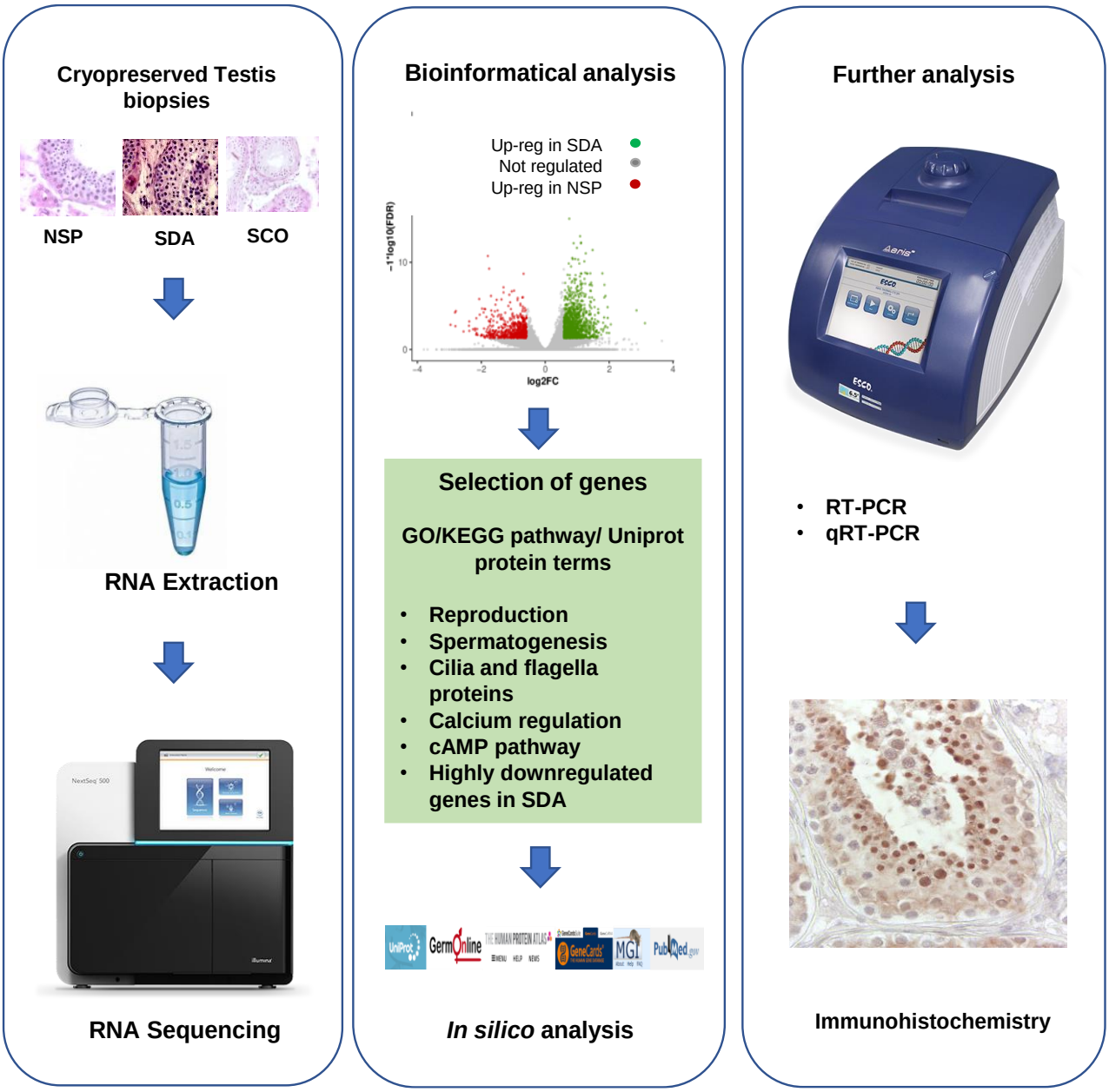


Figure 2.1: Experimental design of RNA sequencing and further analysis. The study was designed to perform RNA sequencing from cryopreserved testis biopsies of three different phenotypes, bioinformatics and the selection of potential candidate genes. In order to do so, *in silico* analysis was performed to gather information with results being validated by RT-PCR, qRT-PCR and immunohistochemistry on testis biopsies. NSP – normal spermatogenesis, SDA – spermatids arrest, SCO – Sertoli cell only

2.2 Material and Methods

2.2.1 Sample collection and RNA extraction

Human testis biopsies were collected at the Clinic of Urology, Pediatric Urology and Andrology, Justus-Liebig University of Giessen, Germany and the Centre for Reproductive Medicine and Andrology, University Clinic Muenster, Germany. Patients were classified as having either obstructive (OA) or non-obstructive azoospermia (NOA) after diagnosis by clinical parameters (sex hormones, ultrasound, and ejaculate analysis) (Jungwirth et al., 2018; World Health Organization, 2021). Surgeries were performed after written informed consent and after approval by the local Ethics Committee (AZ 26/11). Biopsies for histological evaluation were fixed with Bouin's solution overnight, stained with hematoxylin and eosin following standard protocols, and evaluated following score count analysis (Bergmann & Kliesch, 2010). Another piece of the biopsy was stored in liquid nitrogen for RNA analysis. The testis biopsies used for RNAseq were further classified into phenotypes based on histological evaluation of fixed tissues. The following samples were selected: normal/intact spermatogenesis from OA patients (normal spermatogenesis; NSP, n=3; score 10), biopsies with a phenotype of spermatid arrest from NOA patients (SDA, n=4; contain somatic cells and germ cells up to round spermatids with elongating spermatids only occasionally observed in some tubules, score 4-5), and biopsies with a phenotype of Sertoli cell-only (SCO, n=3; complete absence of germ cells, score 0). Total RNA was extracted using the PeqGOLD total RNA kit following the manufacturer's protocol (PeqGOLD total RNA kit, VWR Life Sciences). The quality and quantity of RNA were determined by Nanodrop (Eppendorf photometer). The sample quality and purity were further assessed by Labchip GX Touch (Perkin Elmer) and samples with low RIN values were excluded. Ultimately, 10ng of purified RNA (NSP, n=3), (SDA, n=2), (SCO, n=2) were subjected to RNA Sequencing.

2.2.2 RNA Sequencing and bioinformatical analysis

Libraries were prepared for RNA (barcoded separately and pooled) and run on Illumina NextSeq 500 sequencer at Max-Planck institute, Bad Nauheim, Germany. Data generated were subjected to appropriate aligner platforms and software packages for analysis. RNA and library preparation integrity were verified with LabChip Gx Touch 24 (Perkin Elmer). 10ng of total RNA was used as input for SMARTer® Stranded Total RNA-Seq Kit - Pico Input Mammalian (Takara Clontech). Sequencing was performed on the NextSeq500 instrument (Illumina) using v2 chemistry, resulting in an average of 37M reads per library with 1x75bp single end setup. The resulting raw reads were assessed for quality, adapter content and duplication rates with FastQC (Andrews, 2010). Trimmomatic version 0.39 was employed to trim reads after a quality drop below a mean of Q20 in a window of 10 nucleotides (Bolger et al., 2014). Only reads between 30 and 150 nucleotides were cleared for further analyses. Trimmed and filtered reads were aligned versus the Ensembl human genome version hg38 (GRCh38.27) using STAR 2.6.1d with the parameter “outFilterMismatchNoverLmax 0.1” to increase the maximum ratio of mismatches to mapped length to 10% (Dobin et al., 2013). The number of reads aligning to genes was counted with feature Counts 1.6.5 tool from the Subread package (Liao et al., 2014). Only reads mapping at least partially inside exons were admitted and aggregated per gene. Reads overlapping multiple genes or aligning to multiple regions were excluded. The Ensembl annotation was enriched with UniProt data (release 06.06.2014) based on Ensembl gene identifiers (Activities at the Universal Protein Resource (UniProt)). Differentially expressed genes (DEG) between groups were defined as follows: base mean >= 5, Log2 fold change (FC) <=-0.585 or >= 0.585, FDR <= 0.05. DEGs were filtered for specific terms relevant to sperm function based on Gene Ontology (GO), UniProt protein and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway terms using a bioinformatics strategy explained in the Results.

All downstream analyses were based on the DESeq normalised gene count matrix. Deseq2: p-values were based on Wald test that was corrected for multiple testing using the Benjamini-Hochberg method. The statistical significance was determined by FDR and p values <0.05. The methodology for calculating the fold change was as follows: DESeq2 = log2 of (DESeq-norm mean (CondA) + 1) / (DESeq-norm mean (CondB) + 1)). Volcano and MA plots were produced to visualise DEG expression. A

global clustering heatmap of samples was created based on the Euclidean distance of regularised log transformed gene counts. Dimension reduction analyses (PCA) were performed on regularised log transformed counts using the R packages FactoMineR (FactoMineR: A Package for Multivariate Analysis). DEGs were analysed for gene set overrepresentation analyses with KOBAS (KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases).

2.2.3 Complementary DNA (cDNA) synthesis, reverse transcription polymerase chain reaction (RT-PCR)

Primers were designed for the selected genes using NCBI and OligoExplorer software (Table 2.1). Reagents used for cDNA synthesis and RT-PCR (if not otherwise stated) are from Applied Biosystems by Thermo Fisher Scientific. cDNA was synthesised and RT-PCR was performed as described earlier (Pleuger et al., 2017). Briefly, extracted total RNA (the same sample cohort used for RNA sequencing) was adjusted to the final concentration of 200µg/mL, treated with 1U/µL DNAse 1 (Roche), 1X DNAse buffer (Roche), 1U/µL RNAse inhibitor (Invitrogen), and incubated at 37°C for 30 minutes to degrade genomic DNA. cDNA synthesis was performed using 1X Gold PCR buffer, 1 mM dNTPs, 5mM $MgCl_2$, 25µM random hexamer primers (Invitrogen), 1U/µL Multiscript reverse transcriptase, 1U/µL RNAse inhibitor, 18 µL distilled water and 9 µL of treated mRNA in a 60 µL reaction. The mix was incubated at 21°C for 8 minutes, 42°C for 15 minutes, heated at 99 °C for 5 minutes and stored at 4°C.

RT-PCR was optimised for each primer with 1X Gold PCR buffer, 1mM $MgCl_2$, 200 µM nucleotide mix, 10 pmoles of forward and reverse primers (Table 2.1), 0.65 U of Gold Amplitaq enzyme in a 25µl reaction under following conditions: 95°C for 10 minutes, 45 cycles of 94°C for 45 seconds (denaturation), 60°C for 45 seconds (annealing), 72°C for 45 seconds (extension) and final extension 72°C for 10 minutes. For *PDE4A* and *SLC8B1*, the annealing temperatures were 58°C. Distilled water was used as the negative control for each reaction. PCR products were visualised using agarose gel electrophoresis (2% agarose gels, VWR Life Sciences) with Gel Green nucleic acid stain (Biotium Inc., Fremont, USA). To check primer and product specificity, PCR products were purified and sequenced (Eurofins Genomics, Ebersberg, Germany).

2.2.4 Quantification of gene expression by qRT-PCR

TaqMan probes were selected according to the NCBI gene accession numbers and purchased from Thermo Fisher (Table 2.1). Three samples per group were used from previously synthesised cDNA. Briefly, qRT-PCR assay mix was prepared with 10 µL of 2X TaqMan gene expression master mix (Applied Biosystems by Thermo Fisher Scientific), 1 µL of 20X TaqMan gene expression assay, 8 µL of sterile distilled water and 1 µL of cDNA to a final volume of 20 µL. Triplicates were run for each gene and *RPS27* and *RPL19* were used as reference genes and distilled water as the negative control. Plates were run at 95°C for10 minutes enzyme activation and 40 cycles of 95°C for15 seconds (denaturation) and 60°C for 1 minute (annealing). Relative gene expressions as fold change was manually calculated in Excel using the $\Delta\Delta C_q$ method (Pfaffl, 2001) and graphs were plotted using the software GraphPad Prism (version 9.4.1). One-way ANOVA was used for statistical comparison.

Table 2.1: Primers used for qRT-PCR Primers used for RT-PCR and qRT-PCR

No	Gene	mRNA transcription ID	Primer Sequence (5'->3') TaqMan Gene expression probes	Product size
01	*ORAI1*	NM_032790.3	F – TGCTCTGCTGGGTCAAGTTC R – CTGTCGGTCAGTCTTATGGC Hs03046013_m1	230 bp
02	*CACNB2*	NM_201596.3	F - ATCCGATTCCGATGTATCTC R – CCCTATCCACCAGTCATTG Hs01100744_m1	241 bp
03	*SLC8B1*	NM_001358345.2	F – CGTGGCAACACACTTCACTG R – GAGGTCCCCTCATCTGTCCT Hs00905331_m1	269 bp
04	*TMEM37*	NM_183240.3	F – TCCTCCTCCTGGTGTCTTTC R –GGGTGATGCTGTTGATGTG Hs01931464_s1	174 bp
05	*CFAP47*	NM_001304548.2	F – TTCCACCATCAAGGGTGCTC R – CCAATTCCAATGTTCGCAACC Hs00379969_m1	244 bp
06	*PDE4A*	NM_001111307.2	F – CCTCATCGTAACACCATTTG R – CACGGTTCAACATCCTTTTG Hs00183479_m1	281 bp
07	*TNC*	NM_002160.4	F- CCCCAGGAATTCACACTTTC R – CCGGTTCGGCTTCTGTAAC	151 bp

			Hs01115665_m1	
08	*SPATA31E1*	NM_178828.5	F – GGAAGGAATGCCATCCACCA R – GACCCTAAGCACCCCTTGTC Hs04186440_m1	239 bp
09	*TEKT3*	NM_031898.3	F- TTCAGACGCACTTAGCAAAG R – GTCTCGGCACAACTCAATG Hs00230431_m1	161 bp
10	*SLC9C1*	NM_183061.3	F – ACGCCATACAATGGATGAGTCC R – GCGGTTAGCATGGGATCTGA Hs01590986_m1	259 bp

Reference genes used in TaqMan assay

		Gene ID	**TaqMan probe**
01	*RPS27*	6232	Hs04185005_g1
02	*RPL19*	6143	Hs02338565_gH1

2.2.5 Immunohistochemistry (IHC)

IHC was performed to localise proteins of two genes of interest (*ORAI1* and *SPATA31E1*) in in archived, Bouin's fixed-paraffin embedded tissue sections (NSP n=5, SDA n=3, SCO n=5) This sample cohort comprised of samples from a) the patients' samples used for RNA sequencing and b) other patients' archive material to increase the sample number. IHC was performed using previously described methods (Pleuger et al., 2017). Each slide contained a negative control tissue which excluded the primary antibody treatment and placenta tissue slides were included for each protein as a negative control. Briefly, 5µm thick paraffin sections were mounted on coated slides, de-waxed and rehydrated using xylol and a decreasing alcohol series, respectively. Heat-mediated antigen retrieval was performed in Tris-EDTA buffer (pH 9, SPATA31E1) or citrate buffer (pH 6, ORAI1) for 20 minutes in a microwave oven. Sections were then incubated in 3% hydrogen peroxide (H_2O_2) in Tris-EDTA buffer to inhibit endogenous peroxidase activity, followed by incubation in 1.5% Bovine serum albumin (ROTH GmbH) in Tris-EDTA buffer for 30 minutes to block non-specific binding. Sections were incubated overnight at 4°C with primary antibodies (ORAI1 1:100, SPATA31E1 1:5000, Table 2.2). The next day, sections were washed and incubated with secondary antibody (Table 2.2) at room temperature for one hour. Staining reaction was visualised using NovaRed (VECTOR Laboratories). After counterstaining with hematoxylin (MORPHISTO Laborchemikalien & Histologieservice

GmbH), sections were dehydrated and mounted using Kaiser's gelatin (Merck, Darmstadt, Germany). Leica DM750 microscope was used to observe the samples.

Table 2.2: Primary and secondary antibodies used for IHC

No.	Primary antibody		Secondary antibody	
	Manufacturer details	**Dilution (diluent Tris-EDTA buffer)**	**Manufacturer details**	**Dilution (diluent Tris-EDTA buffer)**
01	ORAI1 antibody Santa Cruz Biotechnology (sc-377281)	1:100	peroxidase conjugated AffiniPure Goat Anti-Mouse (Jackson ImmunoResearch Laboratories)	1:400
02	SPATA31E1 (Invitrogen. AB_2647802)	1:5000	Peroxidase conjugated AffiniPure Goat Anti-Rabbit IgG (Jackson ImmunoResearch Laboratories)	1:800

2.3 Results

2.3.1 RNA sequencing and selection of candidate genes

A total of 260,736,814 reads were processed from all the selected samples subjected to RNA sequencing. After further filtering 88,148,401 reads were used. A total of over 49 million reads were generated in NSP samples, over 24 million reads in SDA samples, and over 13 million reads in SCO samples. Differentially expressed genes were identified based on a mean minimum count => 5, FDR <= 0.05 and log2FC fold change of >0.585 or <-0.585. The number of significantly differentially expressed genes (DEG) between groups were as follows: NSP vs SDA n=1,873; SDA vs SCO n= 4,017; and NSP vs SCO n=10,253 (Fig. 2.2 A and Fig. 2.3 A-C). The complete RNA sequencing raw dataset has been submitted to GEO database (GEO accession GSE224929) (Suppl. Material Appendix B.1). The filtered RNAseq dataset and the full list of differentially expressed genes between groups are mentioned in Appendix B.2 and B3 (Suppl. Material - Excel sheets submitted to the journal).

An important consideration regarding the analysis of these samples is that DEGs will be altered between the testis biopsies with different phenotypes because the cellular

content differs. This fact allowed the design of a bioinformatic strategy to select genes that were likely to be predominantly expressed in germ cells compared to testicular somatic cells. The SDA samples lack/ have very few elongated spermatids and therefore genes predominantly expressed during spermiogenesis should be downregulated compared to NSP biopsies. Consistent with this, gene set enrichment analysis of Gene Ontology (GO) terms comparing SDA to NSP showed that downregulated genes in SDA were significantly associated with terms relevant for spermiogenesis including fertilisation and spermatid development (Fig. 2.3 D). Conversely, SCO biopsies will show a relative increase in the expression of somatic cell genes compared to NSP and SDA, because they lack germ cells and thus RNA from somatic cells will be over-represented and RNA from germ cells will be highly downregulated.

Gene set enrichment analysis of Gene Ontology comparing SDA vs NSP showed that there were a significant number of genes downregulated in SDA, that were associated with the terms related to male germ cell development including, “meiosis II” (3/11 genes), “male meiosis” (6/40 genes), “binding of sperm to zona pellucida” (6/35 genes), and “spermatid development” (14/121 genes). The analysis of NSP vs SCO also showed a significant number of genes downregulated in SCO were associated with GO terms associated with germ cell development including - “DNA methylation in gamete formation” (15/18), “binding of sperm to zona pellucida” (22/35), and “sperm motility” (35/59) (Figure 2.3 E).

To focus on the identification of genes associated with sperm development and function, DEGs were filtered for relevant terms in GO/ KEGG pathways, particularly those annotated as being involved in sperm motility and morphology. First, DEGs with functions known to be related to sperm morphology and motility were selected, focusing on reproduction (Fig. 2.2 B), spermatogenesis (Fig. 2.2 C), calcium (Ca2+) pathway/regulation (Fig. 2.2 E), flagella (Fig. 2.2F), and cyclic adenosine monophosphate (cAMP) dependent protein kinase A (PKA) pathway (Fig. 2.2 G). The latter two pathways were chosen based on the fact that they are known to be essential for sperm motility (Mata-Martinez et al., 2021; Pereira et al., 2017; Suarez et al., 1993). Second, highly downregulated DEGs between NSP vs SDA were focused (Figure 2.2 D) because this identifies genes highly expressed in elongating and elongated spermatids.

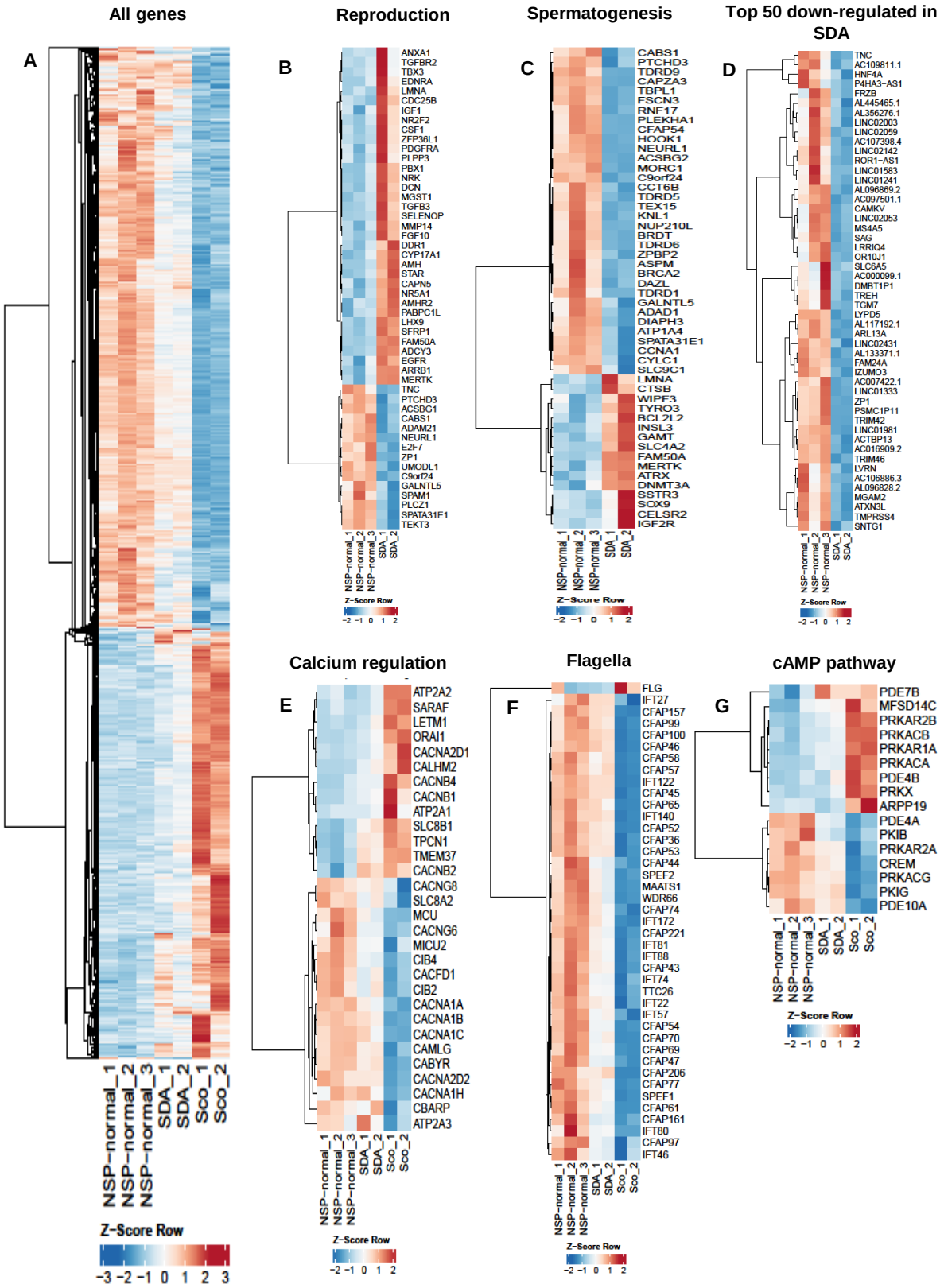


Figure 2.2 Heat maps of DEGs in the RNAseq. **A** – Heat map of all differentially expressed genes. **B** – DEGs annotated as involved in reproduction according to GO/KEGG terms; **C** – DEGs annotated as involved in spermatogenesis according to GO/KEGG terms; **D** – The top

50 most downregulated genes in SDA compared to NSP; **E** – DEGs annotated as involved in calcium pathway regulation according to GO/KEGG terms; **F** – DEGs annotated as associated with flagella according to GO/KEGG terms. **G** – DEGs annotated as involved in cAMP pathway regulation according to GO/KEGG terms. Colours on the heat maps correspond to a z-score which is a measure of relative expression. Blue indicates low relative expression and red colour high relative expression. NSP – Normal spermatogenesis, SDA – Spermatid arrest, SCO – Sertoli cell only. GO – gene ontology, DEG – differentially expressed genes

The list of DEGs were refined using various published datasets to focus on genes that a) are likely to be involved in human spermatogenesis due to their expression pattern, b) have functions relevant to sperm development, morphology, or motility and c) are highly enriched or specifically expressed in the human testis. DEGs identified in functional categories were analysed in ENSEMBL, Human Protein Atlas (HPA), GermOnline, UniProt, and PubMed. Genes were further refined based on limited information on a potential role in male fertility in PubMed, to facilitate the selection of genes with novel roles in spermatogenesis (Suppl. Table 2.1; Excel sheet of Suppl. Material - Appendix B.4 submitted to the journal)

Using this approach, six genes were selected with strong evidence to support a novel role in human sperm development, morphology and/or motility. These genes were downregulated in SDA vs NSP (Table 2.3). *CFAP47* (Cilia and Flagella Associated Protein 4; log2FC -0.83 in SDA) was selected from the gene cluster annotated with UniProt term, flagella proteins. *PDE4A* (Phosphodiesterase 4A; log2FC -0.79 in SDA) was selected from the genes clustered under cAMP pathway, *SPATA31E1* (SPATA31 subfamily E member 1; log2FC -.165), *SLC9C1* (Solute Carrier Family 9 Member C1; log2FC -0.79), and *TEKT3* (Tektin 3; log2FC -0.81) were selected from the spermatogenesis gene cluster, and *TNC* (Tenascin C; log2Fc -.1.75) was selected from the highly downregulated gene cluster in SDA vs NSP. Four genes, *ORAI1* (ORAI Calcium Release-Activated Calcium Modulator 1)*, CACNB2* (Calcium Voltage-Gated Channel Auxiliary Subunit Beta 2)*, SLC8B1* (Solute Carrier Family 8 Member B1), and *TMEM37* (Transmembrane Protein 37) were selected based on their roles in calcium regulation, which is known to be important for multiple aspects of spermatogenesis as well as for sperm function.

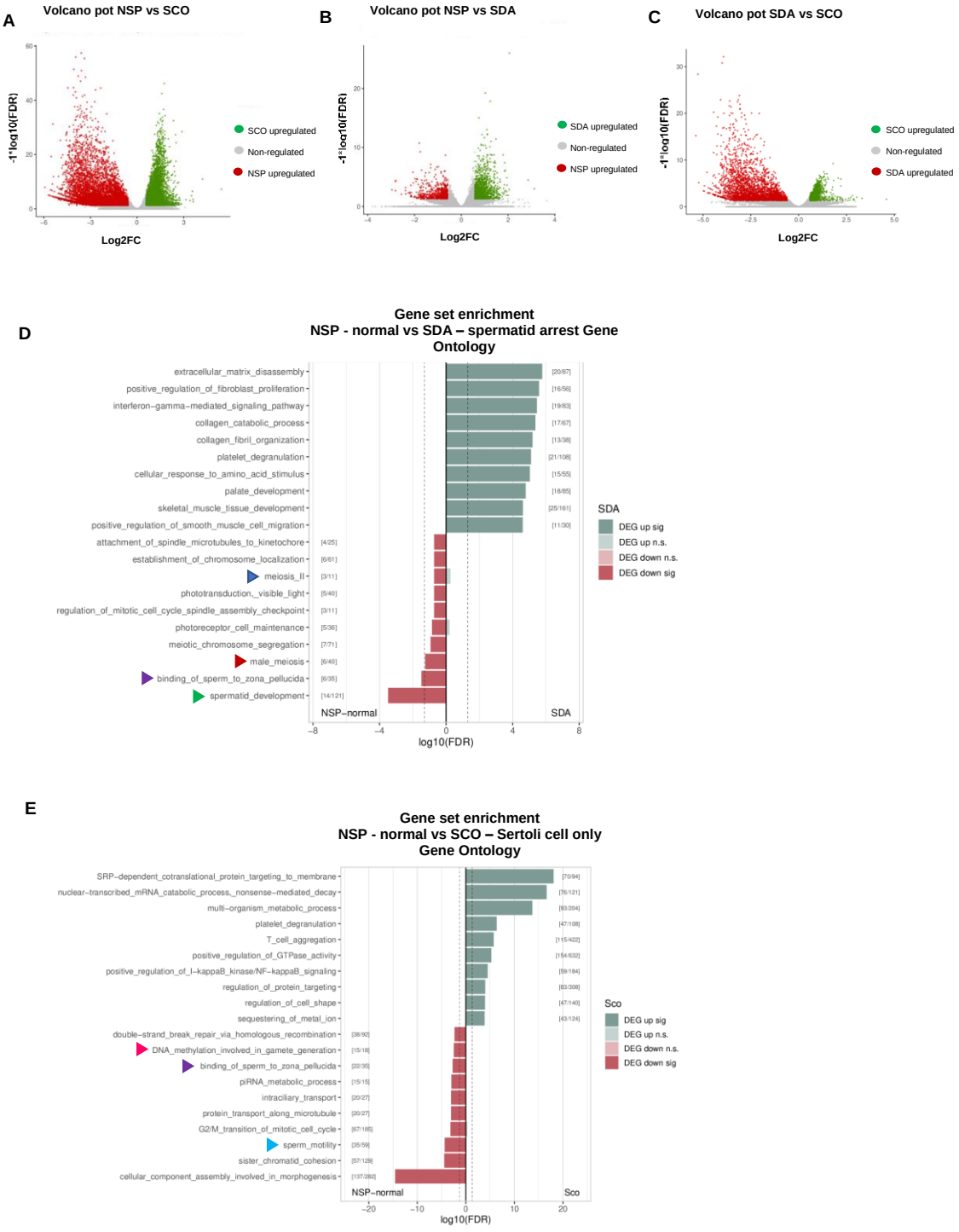


Figure 2.3 Representation of RNAseq data. A – The volcano plot showing upregulated genes of NSP and SCO. Each dot corresponds to a gene. Red dots are genes upregulated in the NSP group and green dots are upregulated genes in the SCO group. **B** – The volcano plot showing upregulated genes of NSP and SDA where red dots correspond to genes upregulated in the NSP group and green dots are genes upregulated in the SDA group. **C** – The volcano plot showing upregulated genes of SDA and SCO, where red dots are genes upregulated in

the SDA group and green dots are genes upregulated in the SCO group. Volcano plots were made against Log2FC vs -1*log10(FDR) with statistical regulation: counts>5, log2FC <>+-0.585, FDR <= 0.05. **D** – The gene set enrichment plot of NSP vs SDA showing the top 10 up or down regulated genes clustered under relative gene ontology (GO) terms. A number of genes were down regulated in the spermatid arrest group compared to normal spermatogenesis which are categorized under meiosis II (3/11) highlighted with a blue arrow, male meiosis (6/40, highlighted with a red arrow), binding of sperm to zona pellucida (6/35, highlighted with a purple arrow) and spermatid development (14/121, highlighted with a green arrow). **E** – The gene set enrichment plot of NSP vs SCO showing the top 10 up or down regulated genes clustered under relative gene ontology (GO) terms. A number of genes were down regulated in the Sertoli cell only group compared to normal spermatogenesis which are categorized under DNA methylation in gamete formation (15/18, highlighted with a pink arrow), binding of sperm to zona pellucida (22/35, highlighted with a purple arrow) and sperm motility (35/59, highlighted with light blue arrow). NSP – normal spermatogenesis, SDA – spermatid arrest, SCO – Sertoli cell only, log2FC – the fold change, FDR – false discovery rate

Table 2.3: Genes selected and their fold change values according to different groups

	Gene	Filter Condition – related GO /KEGG pathway/ UniProt protein terms	log2FC SDA/NSP	log2FC SCO/SDA	log2FC SCO/NSP	Conditions checked
01	*CFAP47*	Flagella proteins	-0.83	-2.54	-3.46	▪ Protein coding ▪ UniProt gene/protein ▪ Base mean and p values ▪ KEGG pathway terms ▪ GO terms ▪ Gene information on NCBI and PubMed ▪ Ensembl protein profiles ▪ Information on the human protein atlas, GermOnline
02	*ORAI1*	Calcium regulation	0.79	1.05	1.78	
03	*CACNB2*		1.12	Not DE	1.09	
04	*SLC8B1*		1.14	Not DE	1.47	
05	*TMEM37*		1.53	Not DE	1.92	
06	*PDE4A*	cAMP pathway	-0.79	-.1.41	-2.26	
07	*SPATA31E1*	spermatogenesis	-1.65	-2.46	-4.19	
08	*SLC9C1*		-0.79	-2.34	-3.21	
09	*TEKT3*	reproduction	-0.81	-3.39	-4.29	
10	*TNC*	Highly downregulated in SDA	-1.75	Not DE	-2.62	

DE – differentially expressed, NSP – Normal spermatogenesis, SDA – Spermatid arrest, SCO – Sertoli cell only, GO – Gene Ontology, KEGG – Kyoto Encyclopedia of Genes and Genomes, NCBI – National Center for Biotechnology Information

2.3.2 Analysis of candidate gene expression in biopsies from men with intact vs impaired spermatogenesis

mRNA transcripts for selected genes were detected in all NSP samples by RT-PCR. Genes *SPATA31E1, TEKT3,* and *CFAP47* were detected in NSP and SDA patients but not in SCO, confirming expression in germ cells but not in somatic cells (Fig. 2.4). Further, quantitative gene expression changes were analysed by qRT-PCR. *SPATA31E1, ORAI1, SLC9C, TNC, TEKT3, CFAP47,* and *PDE4A* were reduced in SDA compared to NSP biopsies, and the levels were further reduced in SCO biopsies confirming the enrichment of these genes in germ cells. *SLC8B1, TMEM37,* and *CACNB2* were selected based on likely functional roles in calcium regulation which is important for sperm function, however their expression was increased in SDA and SCO compared to NSP, suggesting that they are not enriched in germ cells (Fig. 2.5).

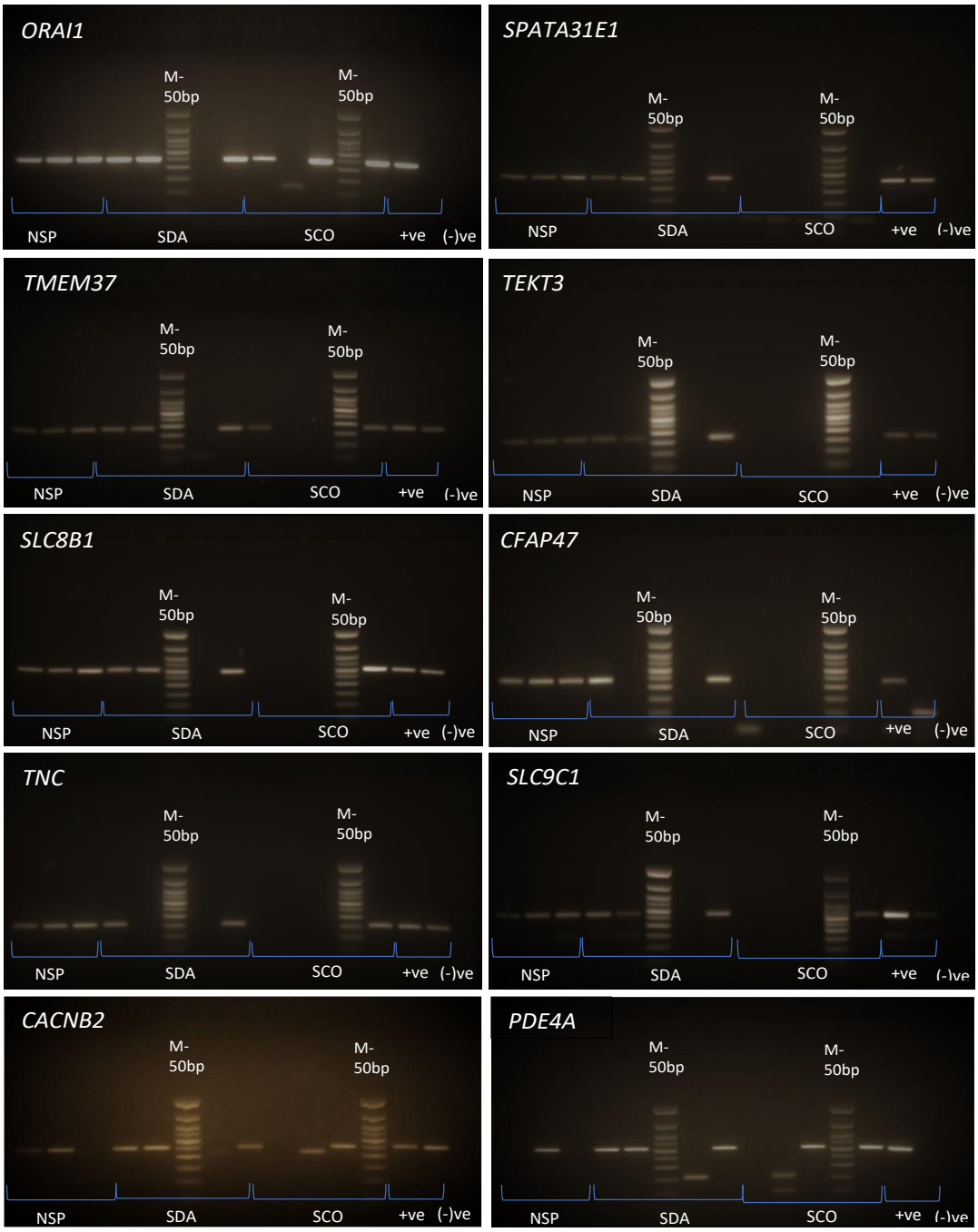


Figure 2.4: RT-PCR results of selected 10 genes. RT-PCR was conducted for 10 genes selected and agarose gel electrophoresis was performed to visualise DNA bands produced. NSP (n=3), SDA (n=4) and SCO (n=3) samples were used with 2 known testis RNA samples as positive controls and distilled water as the negative control. All the negative controls were free from DNA (last well of each gel) and at least one positive control produced the respective DNA band for genes, *ORAI1* - 230 bp, *SPATA31E1* - 239bp, *TMEM37* – 174bp, *TEKT3* – 161bp, *SLC8B1* – 269bp, *CFAP47* – 244bp, *TNC* – 151bp, *SLC9C1* – 259bp, *CACNB2* – 241bp, *PDE4A* – 281bp (Results of each sample is summarized in the table 3 (supplementary document). NSP -Normal spermatogenesis, SDA- spermatid arrest, SCO- Sertoli cell only, M – molecular marker (50 bp), +ve – positive control

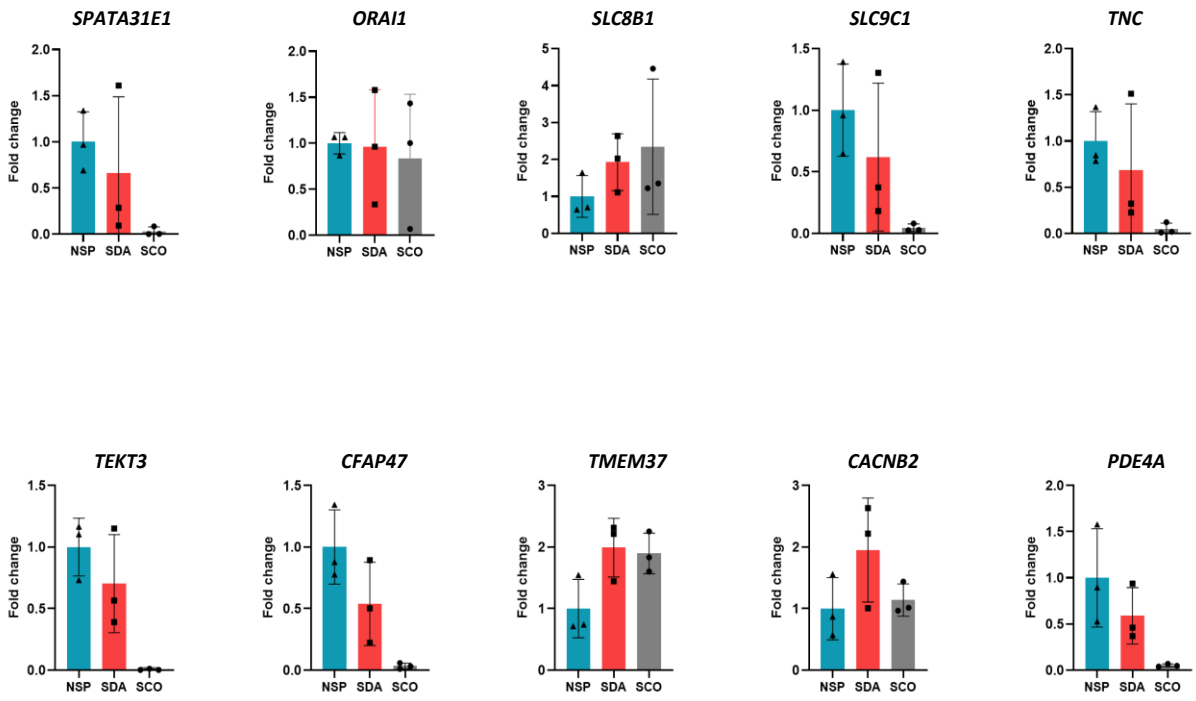


Figure 2.5: Relative gene expression of each candidate gene selected. qRT-PCR was performed to determine the changes in mRNA expression in selected genes in biopsies with normal vs abnormal spermatogenesis. RNA from testis samples of normal spermatogenesis (NSP, n=3), spermatid arrest (SDA, n=3) and Sertoli cell only (SCO, n=3) (each sample was assessed in triplicate). Two housekeeping genes were used for each sample. All the negative controls showed no signal. Each dot represents the average of triplicate measures for a single patient sample. The average value for NSP is set at 1, while SDA and SCO values represent fold-change from this. Data is shown as mean ± SD, n=3 samples per group. NSP -normal spermatogenesis, SDA- spermatid arrest, SCO- Sertoli cell only, SD – standard deviation

2.3.3 Protein localisation of SPATA31E1 and ORAI1 by IHC

Two genes were selected for further analysis by immunohistochemistry to verify protein localisation in human germ cells. In NSP, SPATA31E1 was not detected in spermatogonia and early spermatocytes but in nuclei of pachytene spermatocytes from approximately stage IV of spermatogenesis (Clermont, 1963). This nuclear staining remained detectable in spermatids until step 4-5. Cytoplasmic staining of round spermatids was apparent in some sections where the staining intensity was high. Elongating spermatid cytoplasm was generally negative (Fig. 2.6 A). Both lightly

stained and unstained Sertoli cell nuclei were observed in NSP; this seems likely due to cyclic variation in Sertoli cells (Fig. 2.6 A and Fig. 2.7 NSP). Basal Sertoli cell cytoplasm staining could also be observed in tubules with the greatest staining intensity in each patient (Fig. 2.7 NSP). Within these tubules, the staining seemed to vary according to the stage of the spermatogenic cycle. In the tubules/ tissues with lower staining intensity, Sertoli cell cytoplasmic staining was not evident (Fig. 2.7 NSP).

As in NSP, SDA biopsies showed SPATA31E1 staining in nuclei of late pachytene spermatocytes and early round spermatids with mature spermatids absent in these biopsies. Both stained and unstained Sertoli cell nuclei were evident as observed in NSP biopsies (Fig. 2.6 B). Some tubules also showed evidence of Sertoli cell cytoplasmic staining (Fig. 2.7 SDA). In SCO, a greater heterogeneity between patients was noted, with some patients showing very little immunostaining (n=2) and others showing clear Sertoli cell nuclear and light cytoplasmic immunostaining (n=3) (Figure 2.6 C, Fig. 2.8). Interestingly, both stained and unstained Sertoli cell nuclei were observed in the same tubules of some patients, suggesting that the cyclic variation in Sertoli cell nuclear SPATA31E1 immunostaining observed in NSP and SDA biopsies is preserved in SCO patients (Fig. 2.6 C, Fig. 2.8). The specificity of immunostaining was confirmed by obtaining unstained negative control testis tissues in each slide and unstained placenta tissue slides.

For ORAI1, staining in NSP samples was detected in the nuclei of all germ cells except for A_{dark} spermatogonia. Nuclear staining was observed in Sertoli cells, however some Sertoli cell nuclei were unstained suggesting nuclear localisation could be stage-specific (Fig. 2.6 D), similar to SPATA31E1. In SDA samples, ORAI1 was also observed in nuclei of all germ cells present (up to round spermatids) except A_{dark} spermatogonia. Sertoli cell nuclei also showed immunostaining, with some immuno-negative nuclei (Fig. 2.6 E) suggesting that stage-specificity of Sertoli cell staining was preserved in SDA biopsies. In SCO samples, both stained and unstained Sertoli cell nuclei were observed (Figure 4F). No germ cell cytoplasm staining was observed for ORAI1 in all three phenotypes analysed. The specificity of immunostaining was confirmed by obtaining unstained negative control testis tissues in each slide and unstained placenta tissue slides.

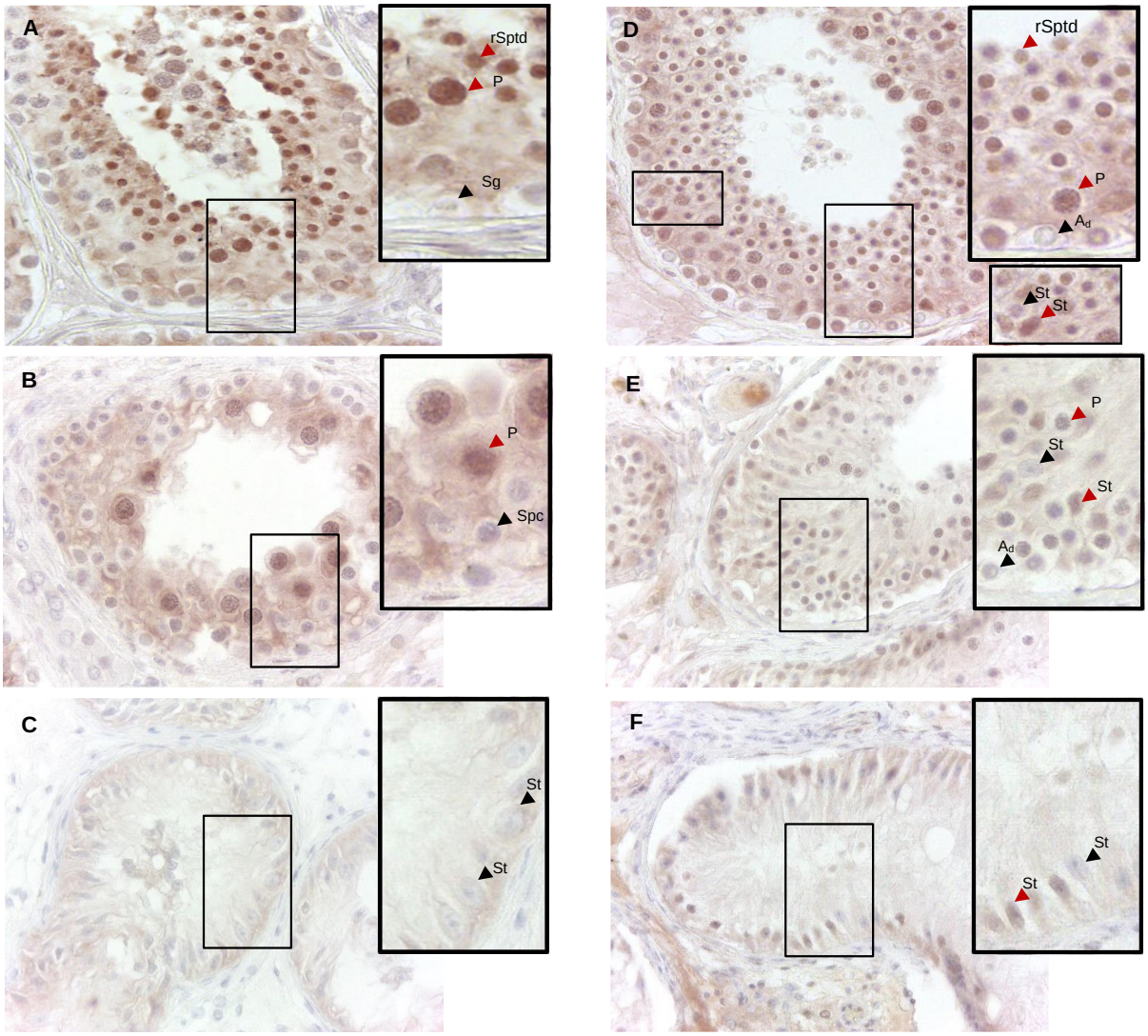


Figure 2.6: Immunohistochemistry of SPATA31E1 and ORAI1 in human testis biopsies. A-C. SPATA31E1, **D-F. ORAI1**. Immunostaining was detected using NovaRed, and sections were counterstained with hematoxylin. **A** – SPATA31E1 staining in NSP. *Inset*: black arrowhead indicates immuno-negative spermatogonia (Sg) and red arrowheads indicate nuclear staining in pachytene spermatocytes (P) and round spermatids (rSptd). **B** – SPATA31E1 staining in SDA. *Inset*: black arrowhead indicates immuno-negative early spermatocytes (Spc) and red arrowhead nuclear staining in a late pachytene spermatocyte (P). **C** – SPATA31E1 staining in SCO. *Inset*: black arrows indicate immuno-negative Sertoli cell nuclei (St). **D** – ORAI1 staining in NSP. *Inset:* red arrows indicate nuclear staining of pachytene spermatocytes (P), round spermatids (rSptd) and Sertoli cell (St) and black arrows indicate immuno-negative A_{dark} spermatogonia (A_d) and some immuno-negative Sertoli cell nuclei (St). **E** – ORAI1 staining in SDA. *Inset:* red arrows indicate stained pachytenes (P) and Sertoli cell nuclei (St) and black arrows indicate immuno-negative A_{dark} spermatogonia (A_d) and Sertoli cells nuclei (St). **F** – ORAI1 staining in SCO. *Inset:* red arrow indicates

immunostaining of Sertoli cell nucleus (St) and black arrow indicates immuno-negative Sertoli cell nucleus (St).

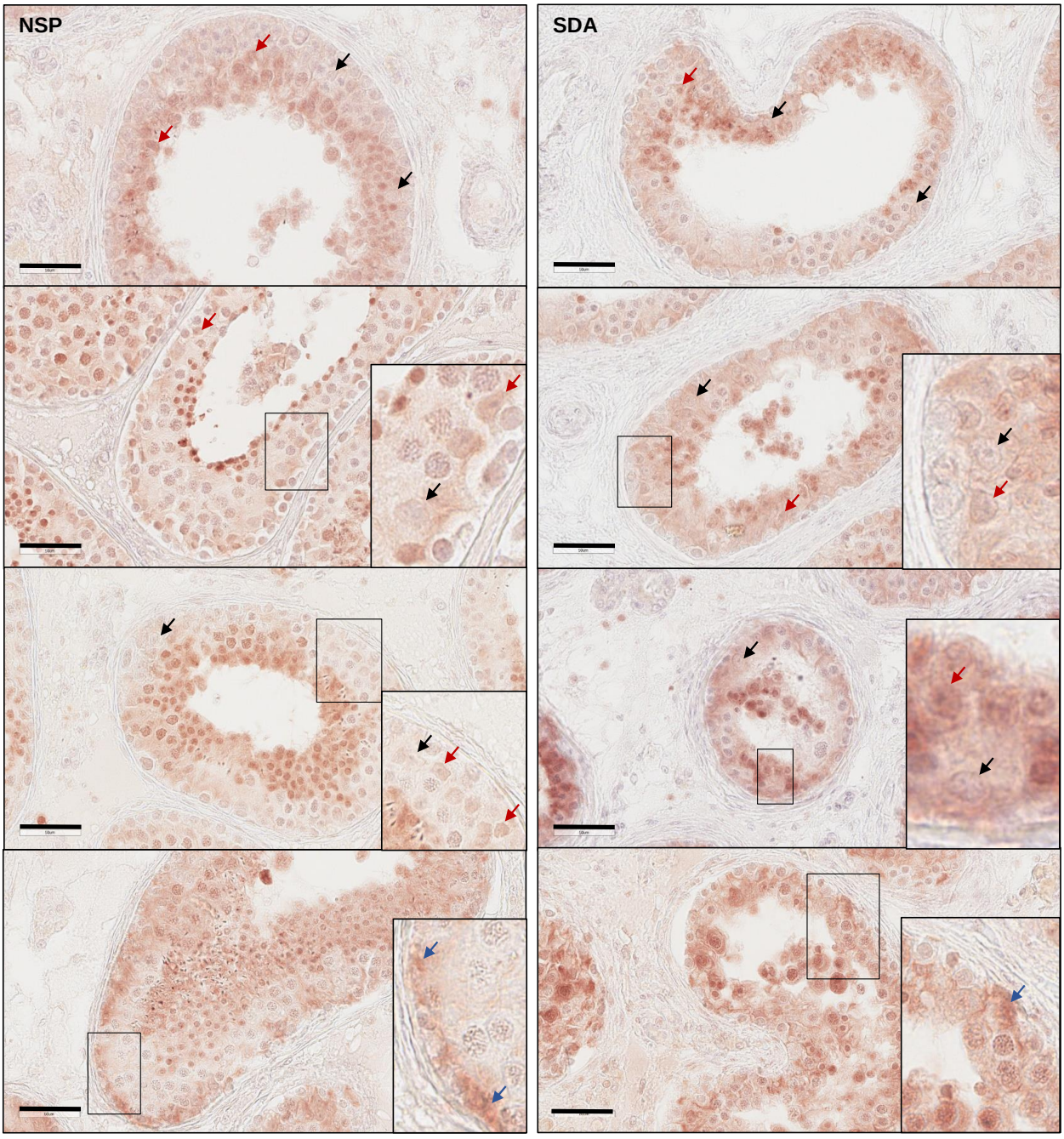


Figure 2.7: SPATA31E1 images depicting patient heterogeneity in Sertoli cells in NSP and SDA samples. Immunostaining was detected using NovaRed, and sections were counterstained with hematoxylin. *Left panel:* SPATA31E1 staining in NSP. *Right panel:* SPATA31E1 staining in SDA. *Inset:* Black arrows indicate immuno-negative and red arrows indicate immuno-positive Sertoli cell nuclear staining. The blue arrows indicate the Sertoli cell cytoplasmic staining. The bar indicates 50µM. NSP – normal spermatogenesis, SDA – spermatid arrest.

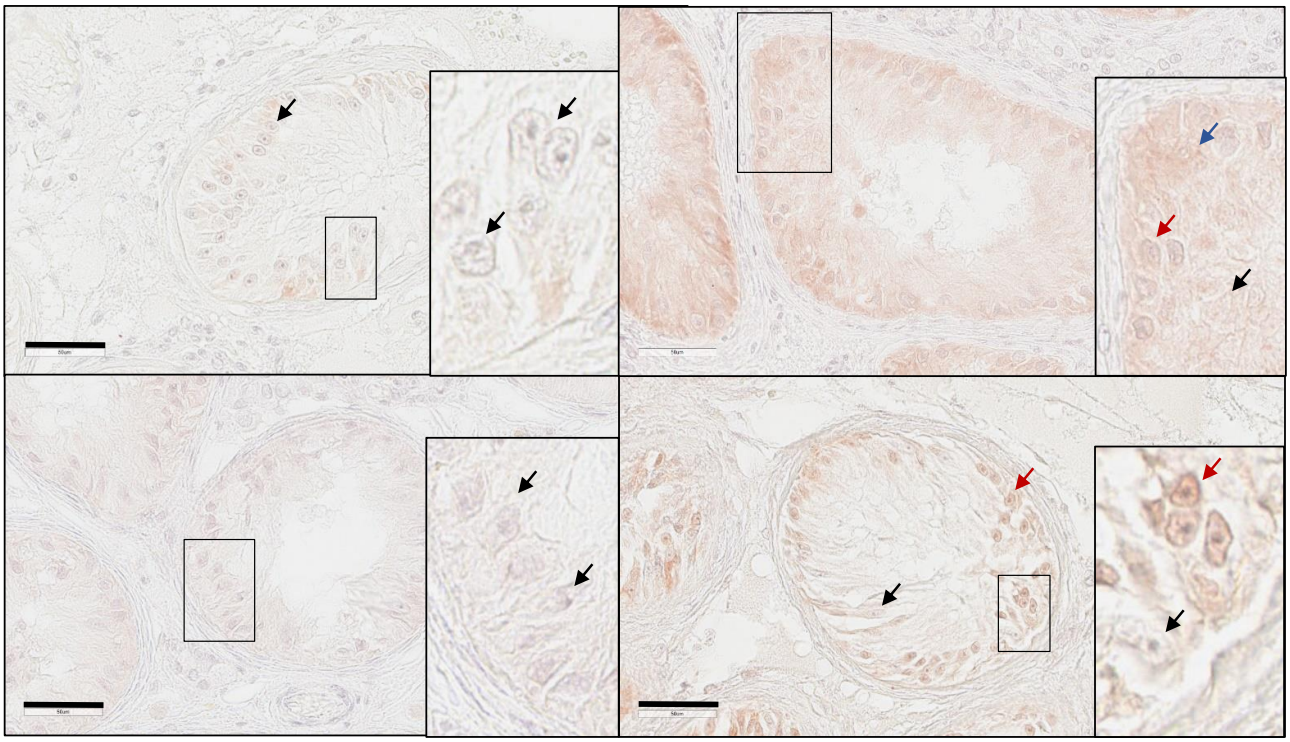


Figure 2.8: SPATA31E1 Images depicting the heterogeneity in Sertoli cells in SCO samples. Immunostaining was detected using NovaRed, and sections were counterstained with hematoxylin. *Inset:* Black arrows indicate immuno-negative and red arrows indicate immuno-positive Sertoli cell nuclear staining. Blue arrow indicates a faint cytoplasmic staining. The bar indicates 50μM. SCO – Sertoli cell only.

2.4 Discussion

Next generation sequencing and RNAseq have been applied for the identification of altered gene expression in abnormal testicular histology and function, and to identify novel genes involved in disease development (Dado-Senn et al., 2018; Robertson et al., 2020). The application of RNAseq technology in male reproductive research has contributed to the understanding of RNA complexity and molecular mechanisms of impaired male fertility (Bianchi et al., 2019). Spermatogenesis is believed to be regulated by up to 2000 genes, of which approximately 700 – 900 genes are exclusively expressed in the male germ line (Chalmel et al., 2012; Matzuk & Lamb, 2008; Tüttelmann et al., 2018). Since there are so many genes likely involved in sperm

production, many genes causing male infertility are probably yet to be identified (Coutton et al., 2015).

The current study focused on discovering genes that could be linked to human sperm development, morphology and/or motility. In an untargeted approach, bulk RNAseq analysis was used to quantify gene expression levels in testis biopsies from men with obstructive azoospermia, with intact spermatogenesis, and non-obstructive azoospermia due to spermatid arrest (SDA) or germ cell arrest (Sertoli cell-only, SCO). By definitive histological characterisation of tissue specimens, the cellular components of the starting material are precisely known in contrast to tissue samples used without prior histological assessment. Additionally, knowledge on cellular composition is essential for the interpretation of DEGs between the sample groups. Many genes will be differentially expressed simply due to the different cell compositions of biopsies; for example, in SDA where spermatids are lost, a marked downregulation of genes enriched in spermatids will result. Biopsies with SCO compared to NSP will show an upregulation of somatic cell genes and a downregulation of germ cell-specific genes. Careful selection of genes with particular expression patterns and functional annotations can be used to identify novel genes and corresponding proteins that could play important roles in human spermatogenesis. The study particularly focused on the comparison of NSP (i.e. intact spermatogenesis with all germ cell stages, including elongated spermatids or “testicular sperm” present) and SDA (i.e. spermatid arrest; germ cells present until the phase of round spermatid with elongating spermatids only occasionally observed in some tubules, suggesting a defective spermiogenesis) samples.

To further re-fine target gene selection, the list of differentially expressed genes (DEGs) was classified based on their functions, expression patterns, and the published literature. Genes associated with calcium and cAMP pathways were particularly focused since these are important for sperm motility (Mata-Martinez et al., 2021; Pereira et al., 2017; Suarez, 2008; Suarez et al., 1993). Abnormalities in the flagella result in asthenozoospermia which impairs male fertility (Coutton et al., 2015) and thus genes likely associated with flagella function were also selected based on gene annotations. Also, genes annotated as having an involvement in spermatogenesis and/or spermiogenesis were selected and focused on genes that were highly downregulated in biopsies showing spermatid arrest biopsies compared

to NSP, as these were likely to be predominantly expressed in elongating or elongated spermatids. Finally, genes that had little published information but were enriched in germ cells were selected. Through this approach, the list was narrowed down to 10 genes with putative roles in sperm development and/or function. The data suggested that *SPATA31E1, TEKT3, SLC9C1, PDE4A, CFAP47,* and *TNC* are predominantly expressed in human germ cells. Thus, this study successfully identified several genes that may play previously unrecognised roles in human sperm production and function. Immunohistochemical analyses of two proteins, i.e. SPATA31E1 and ORAI1, confirmed the localisation to human germ cells, which have not previously been described.

Gene *CFAP47* was selected based on annotations and expression pattern suggesting a role in sperm flagella function. *CFAP47* mRNA was reduced in SDA and SCO testis samples by qRT-PCR, suggesting it is expressed in post-meiotic germ cells. Analysis of *CFAP47* expression patterns in the Human Protein Atlas (HPA) confirmed *CFAP47* is expressed in early and late spermatids (Uhlén et al., 2015) and annotations in UniProt suggest the protein is highly expressed in spermatozoa (Liu et al., 2021). A recent study that was published while these experiments were being performed, revealed that deleterious X-linked variants of *CFAP47* were associated with asthenoteratozoospermia in a whole exome sequencing study (Liu et al., 2021), thus providing confidence in the approach used for the selection of genes that are likely to have a functional role in spermiogenesis.

PDE4A was selected because of its likely involvement in the regulation of cAMP, which is known to be involved in sperm motility (Pereira et al., 2017), yet PDE4A's role in human spermatozoa is not clear. RNAseq and qRT-PCR data suggest that *PDE4A* is enriched in human spermatids, and this proposition is confirmed by *PDEA4A* tissue expression data in Human Protein Atlas. This protein has been shown to be enriched in rat spermatids (Salanova et al., 1999) and thus it is likely to be conserved in spermatids in both rodents and humans.

TEKT3, *TNC,* and *SLC9C1* were selected based on their annotations suggesting an involvement in reproduction and spermatogenesis. Results showed that these genes are detected in normal spermatogenesis but are reduced in biopsies with spermatid arrest and SCO, suggesting they are predominantly expressed in human germ cells,

particularly in early/late spermatids. Consistent with these findings, *TEKT3* mRNA is enriched in human spermatids according to the HPA (Uhlén et al., 2015), and its protein is detected in sperm flagella according to UniProt. *TEKT3* has shown an association with sperm motility in null mice by showing reduced progressive motility and structural defects in the sperm (Roy et al., 2009). *TNC* mRNA is also enriched in human spermatids compared to other germ cells (Uhlén et al., 2015), yet little is known of its role in these germ cells. According to the Human Protein Atlas, *SLC9C1* mRNA is highly enriched in human testis compared to other tissues, with high expression in spermatids (Uhlén et al., 2015), suggesting its expression is mainly restricted to developing spermatids in humans. *SLC9C1* was recently published as a causative gene for asthenozoospermia (Cavarocchi et al., 2021), again confirming that this approach to gene selection is able to reveal genes with functional significance in human spermatids. In summary, the RNAseq and bioinformatics approach used in this study proved to be useful in identifying proteins that are expressed in developing sperm and thus potentially play a role in sperm development or function.

This study focused on *SPATA31E1,* because little is known about this gene and the molecular function has not been identified. According to HPA (Pontén et al., 2009), SPATA31E1 is highly testis-specific, making it an attractive target for contraception. The role of *SPATA31E1* in male fertility is not well studied, other than a publication in stallions suggesting that SNPs in *SPATA31E1* may have a deleterious effect on fertility (Schrimpf et al., 2016). *SPATA31E1* was recently mentioned as a human spermatid-specific gene and a potential drug target (Robertson et al., 2020). Analyses in the current study showed *SPATA31E1* is expressed in intact human spermatogenesis and is reduced in SDA and SCO biopsies. Tissue expression analyses in HPA suggest that *SPATA31E1* mRNA and protein is highly enriched in the testis, restricted to early and late spermatids, with some mRNA (but not protein) in brain (Uhlén et al., 2015). This expression pattern is consistent with data of the current study, showing it is reduced in testis biopsies from patients with spermatid arrest and shows very low expression in SCO. Immunohistochemistry was used to show that SPATA31E1 is localised to germ cell nuclei from late pachytene spermatocytes (stages IV, V) suggesting this protein could play a role in the later stages of meiosis. It was also detected in early round spermatids and elongating spermatids suggesting a role during

spermiogenesis. Further studies are needed to investigate the expression of SPATA31E1 in human spermatozoa.

SPATA31E1 was also detectable in human Sertoli cell nuclei, and both immuno-positive and immuno-negative Sertoli cell nuclei were noted in NSP and SDA, suggesting the protein expression in Sertoli cell nuclei depends on the stage of the spermatogenic cycle. Because of the helical nature of human spermatogenesis and the fact that multiple stages are observed in a single tubule cross-section, it was difficult to distinguish the precise stages at which Sertoli cell nuclear expression of SPATA31E1 varied. Intriguingly, apparent stage-specificity was also observed of SPATA31E1 nuclear (and to some degree, cytoplasmic) Sertoli cell localisation in seminiferous tubules lacking germ cells in SCO biopsies. Stage-specific Sertoli cell protein expression can be observed in the mouse embryonic testis, even in the absence of germ cells, leading to speculation that fetal Sertoli cells, in the presence or absence of germ cells, show innate cyclic variation in their function that does not rely on the presence of postnatal germ cells (Timmons et al., 2002). Thus, it is tempting to speculate that the variation in SPATA31E1 nuclear localisation in SCO reflects the preservation of human Sertoli cell cyclic activity in the absence of germ cells. Heterogeneity between SCO patients were also noted in terms of whether SPATA31E1 was expressed in the nuclei or not. This could potentially reflect Sertoli cell maturity or function that is influenced by the genetic, epigenetic and/or hormonal status of the patient. Further studies should investigate whether SPATA31E1 Sertoli cell nuclear staining is related to Sertoli stage, maturity, and functional status.

The protein expression of ORAI1 in the testis was analysed because its localisation has not been described previously. ORAI1 is a calcium-selective ion channel-calcium transporter class and is a membrane calcium subunit that is activated when calcium stores are depleted (Frischauf et al., 2016; Lunz et al., 2019; Mercer et al., 2006). ORAI1 was expressed in various organs including the testis (Pontén et al., 2009). Immunohistochemical data demonstrated that ORAI1 was expressed in germ cells and Sertoli cells, suggesting multiple roles in spermatogenesis. It was also stage specific in Sertoli cell nuclei in all samples, including SCO. *Orai1*-/- mice were sterile showing defects in elongating spermatid development with markedly reduced motility and low sperm count, suggesting a functional role in germ cell development (Davis et al.,

2016). ORAI1 has also been suggested to be involved in the turnover of Sertoli cell intercellular junctions in the rat testis (Lyon et al., 2017).

In conclusion, this study analysed transcriptional changes in human testis biopsies with well characterised phenotypes, comparing gene expression in biopsies from testes with normal spermatogenesis, spermatid arrest and those lacking germ cells. Various bioinformatic approaches were used to select 10 differentially expressed genes that were likely to be testis and/or spermatid specific, and thus potentially play a functional role, in developing sperm. Of these genes, *SPATA31E1, TEKT3, SLC9C1, PDE4A, CFAP47,* and *TNC* showed excellent evidence of enrichment in human germ cells in the testis. The study also identified ORAI1 as a protein with potential roles in Sertoli and germ cell function, and SPATA31E1 as a testis-specific protein expressed in spermatocytes, spermatids and Sertoli cells that could be further explored for its utility as a contraceptive target due to its restricted expression. Data generated in this study provides proteins that could be explored for their utility as targets for contraception, predictive markers for successful surgical sperm retrieval, or for an involvement in human male infertility.

2.5 References

Agarwal, A., Mulgund, A., Hamada, A., & Chyatte, M. R. (2015). A unique view on male infertility around the globe. *Reprod Biol Endocrinol*, *13*, 37. https://doi.org/10.1186/s12958-015-0032-1

Andrews, S. (2010). *A quality control tool for high throughput sequence data*. http://www.bioinformatics.babraham.ac.uk/projects/fastqc

Bergmann, M., & Kliesch, S. (2010). Testicular biopsy and histology. In H. M. B. E. Nieschlag, S. Nieschlag (Ed.), *Andrology: Male reproductive health and dysfunction* (3rd edition ed., pp. 158-165). Springer Verlag.

Bianchi, E., Boekelheide, K., Sigman, M., Braun, J. M., Eliot, M., Hall, S. J., Dere, E., & Hwang, K. (2019). Spermatozoal large RNA content is associated with semen characteristics, sociodemographic and lifestyle factors. *PLoS One*, *14*(5), e0216584. https://doi.org/10.1371/journal.pone.0216584

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114-2120. https://doi.org/10.1093/bioinformatics/btu170

Cavarocchi, E., Whitfield, M., Chargui, A., Stouvenel, L., Lores, P., Coutton, C., Arnoult, C., Santulli, P., Patrat, C., Thierry-Mieg, N., Ray, P. F., Dulioust, E., & Toure, A. (2021). The sodium/proton exchanger SLC9C1 (sNHE) is essential

for human sperm motility and fertility. *Clin Genet*, *99*(5), 684-693. https://doi.org/10.1111/cge.13927

Chalmel, F., Lardenois, A., Evrard, B., Mathieu, R., Feig, C., Demougin, P., Gattiker, A., Schulze, W., Jegou, B., Kirchhoff, C., & Primig, M. (2012). Global human tissue profiling and protein network analysis reveals distinct levels of transcriptional germline-specificity and identifies target genes for male infertility. *Hum Reprod*, *27*(11), 3233-3248. https://doi.org/10.1093/humrep/des301

Chalmel, F., Rolland, A. D., Niederhauser-Wiederkehr, C. C., S. S. W., Demougin, P., Gattiker, A., Moore, J., Patard, J., Wolgemuth, D. J., Je´gou, B., & Primig, M. (2007). The conserved transcriptome in human and rodent male gametogenesis. *PNAS*, *104*, 8346–8351. https://doi.org/10.1073/pnas.0701883104.

Clermont, Y. (1963). The cycle of the seminiferous epithelium in man. *Am J Anat*, *112*, 35-51. https://doi.org/10.1002/aja.1001120103

Coutton, C., Escoffier, J., Martinez, G., Arnoult, C., & Ray, P. F. (2015). Teratozoospermia: spotlight on the main genetic actors in the human. *Hum Reprod Update*, *21*(4), 455-485. https://doi.org/10.1093/humupd/dmv020

Dado-Senn, B., Skibiel, A. L., Fabris, T. F., Zhang, Y., Dahl, G. E., Penagaricano, F., & Laporta, J. (2018). RNA-Seq reveals novel genes and pathways involved in bovine mammary involution during the dry period and under environmental heat stress. *Sci Rep*, *8*(1), 11096. https://doi.org/10.1038/s41598-018-29420-8

Davis, F. M., Goulding, E. H., D'Agostin, D. M., Janardhan, K. S., Cummings, C. A., Bird, G. S., Eddy, E. M., & Putney, J. W. (2016). Male infertility in mice lacking the store-operated Ca(2+) channel Orai1. *Cell Calcium*, *59*(4), 189-197. https://doi.org/10.1016/j.ceca.2016.02.007

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, *29*(1), 15-21. https://doi.org/10.1093/bioinformatics/bts635

Frischauf, I., Zayats, V., Deix, M., Hochreiter, A., Polo, I. J., Muik, M., Lackner, B., Svobodová, B., Pammer, T., Litviňuková, M., Sridhar, A. A., Derler, I., Bogeski, I., Romanin, C., Ettrich, R. H., & Schindl, R. (2016). A calcium-accumulating region, CAR, in the channel Orai1 enhances Ca2+ permeation and SOCE-induced gene transcription. *Sci Signal*. https://doi.org/10.1126/scisignal.aab1901

Guo, J., Grow, E. J., Mlcochova, H., Maher, G. J., Lindskog, C., Nie, X., Guo, Y., Takei, Y., Yun, J., Cai, L., Kim, R., Carrell, D. T., Goriely, A., Hotaling, J. M., & Cairns, B. R. (2018). The adult human testis transcriptional cell atlas. *Cell Res*, *28*(12), 1141-1157. https://doi.org/10.1038/s41422-018-0099-2

Houston, B. J., Riera-Escamilla, A., Wyrwoll, M. J., Salas-Huetos, A., Xavier, M. J., Nagirnaja, L., Friedrich, C., Conrad, D. F., Aston, K. I., Krausz, C., Tuttelmann, F., O'Bryan, M. K., Veltman, J. A., & Oud, M. S. (2021). A systematic review of the validated monogenic causes of human male infertility: 2020 update and a discussion of emerging gene-disease relationships. *Hum Reprod Update*, *28*(1), 15-29. https://doi.org/10.1093/humupd/dmab030

Jungwirth, A., Diemer, T., Kopa, Z., Krausz, C., Minhas, S., & Tournaye, H. (2018). *EAU Guidelines on Male Infertility*. European Association of Urology.

Jungwirth, A., Giwercman, A., Tournaye, H., Diemer, T., Kopa, Z., Dohle, G., Krausz, C., & European Association of Urology Working Group on Male, I. (2012).

European Association of Urology guidelines on Male Infertility: the 2012 update. *Eur Urol*, *62*(2), 324-332. https://doi.org/10.1016/j.eururo.2012.04.048

Krausz, C., Riera-Escamilla, A., Moreno-Mendoza, D., Holleman, K., Cioppi, F., Algaba, F., Pybus, M., Friedrich, C., Wyrwoll, M. J., Casamonti, E., Pietroforte, S., Nagirnaja, L., Lopes, A. M., Kliesch, S., Pilatz, A., Carrell, D. T., Conrad, D. F., Ars, E., Ruiz-Castane, E., . . . Tuttelmann, F. (2020). Genetic dissection of spermatogenic arrest through exome analysis: clinical implications for the management of azoospermic men. *Genet Med*, *22*(12), 1956-1966. https://doi.org/10.1038/s41436-020-0907-1

Liao, Y., Smyth, G. K., & Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, *30*(7), 923-930. https://doi.org/10.1093/bioinformatics/btt656

Lin, Y. H., Lin, Y. M., Teng, Y. N., Hsieh, T. Y., Lin, Y. S., & Kuo, P. L. (2006). Identification of ten novel genes involved in human spermatogenesis by microarray analysis of testicular tissue. *Fertil Steril*, *86*(6), 1650-1658. https://doi.org/10.1016/j.fertnstert.2006.04.039

Liu, C., Tu, C., Wang, L., Wu, H., Houston, B. J., Mastrorosa, F. K., Zhang, W., Shen, Y., Wang, J., Tian, S., Meng, L., Cong, J., Yang, S., Jiang, Y., Tang, S., Zeng, Y., Lv, M., Lin, G., Li, J., . . . Zhang, F. (2021). Deleterious variants in X-linked CFAP47 induce asthenoteratozoospermia and primary male infertility. *Am J Hum Genet*, *108*(2), 309-323. https://doi.org/10.1016/j.ajhg.2021.01.002

Lunz, V., Romanin, C., & Frischauf, I. (2019). STIM1 activation of Orai1. *Cell Calcium*, *77*, 29-38. https://doi.org/10.1016/j.ceca.2018.11.009

Lyon, K., Adams, A., Piva, M., Asghari, P., Moore, E. D., & Vogl, A. W. (2017). Ca2+ signaling machinery is present at intercellular junctions and structures associated with junction turnover in rat Sertoli cells. *Biol Reprod*, *96*(6), 1288-1302. https://doi.org/10.1093/biolre/iox042

Mata-Martinez, E., Sanchez-Cardenas, C., Chavez, J. C., Guerrero, A., Trevino, C. L., Corkidi, G., Montoya, F., Hernandez-Herrera, P., Buffone, M. G., Balestrini, P. A., & Darszon, A. (2021). Role of calcium oscillations in sperm physiology. *Biosystems*, *209*, 104524. https://doi.org/10.1016/j.biosystems.2021.104524

Matzuk, M. M., & Lamb, D. J. (2008). The biology of infertility: research advances and clinical challenges. *Nat Med*, *14*(11), 1197-1213. https://doi.org/10.1038/nm.f.1895

McLachlan, R. I., & O'Bryan, M. K. (2010). Clinical Review#: State of the art for genetic testing of infertile men. *J Clin Endocrinol Metab*, *95*(3), 1013-1024. https://doi.org/10.1210/jc.2009-1925

Mercer, J. C., DeHaven, W. I., Smyth, J. T., Wedel, B., Boyles, R. R., Bird, G. S., & Putney, J. J. W. (2006). Large Store-Operated Calcium-Selective Currents Due To Co-Expression Of Orai1 Or Orai2 With The Intracellular Calcium Sensor, Stim1. *Journal of Biological Chemistry* 24979–24990. https://doi.org/doi: 10.1074/jbc.M604589200

Minhas, S., Bettocchi, C., Boeri, L., Capogrosso, P., Carvalho, J., Cilesiz, N. C., Cocci, A., Corona, G., Dimitropoulos, K., Gul, M., Hatzichristodoulou, G., Jones, T. H., Kadioglu, A., Martinez Salamanca, J. I., Milenkovic, U., Modgil, V., Russo, G. I., Serefoglu, E. C., Tharakan, T., . . . Reproductive, H. (2021). European Association of Urology Guidelines on Male Sexual and Reproductive Health: 2021 Update on Male Infertility. *Eur Urol*, *80*(5), 603-620. https://doi.org/10.1016/j.eururo.2021.08.014

Norreen-Thorsen, M., Struck, E. C., Öling, S., Zwahlen, M., Von Feilitzen, K., Odeberg, J., Lindskog, C., Pontén, F., Uhlén, M., Dusart, P. J., & Butler, L. M. (2022). A human adipose tissue cell-type transcriptome atlas. *Cell Rep*, *40*(2), 111046. https://doi.org/10.1016/j.celrep.2022.111046

O'Flynn O'Brien, K. L., Varghese, A. C., & Agarwal, A. (2010). The genetic causes of male factor infertility: a review. *Fertil Steril*, *93*(1), 1-12. https://doi.org/10.1016/j.fertnstert.2009.10.045

Panner Selvam, M. K., Agarwal, A., Peter, P. N., Baskaran, S., & Bendou, H. (2019). Sperm Proteome Analysis and Identification of Fertility-Associated Biomarkers in Unexplained Male Infertility. *Genes*. https://doi.org/doi:10.3390

Pereira, R., Sa, R., Barros, A., & Sousa, M. (2017). Major regulatory mechanisms involved in sperm motility. *Asian J Androl*, *19*(1), 5-14. https://doi.org/10.4103/1008-682X.167716

Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Research 29*, 2002-2007.

Pleuger, C., Fietz, D., Hartmann, K., Schuppe, H. C., Weidner, W., Kliesch, S., Baker, M., O'Bryan, M. K., & Bergmann, M. (2017). Expression of ciliated bronchial epithelium 1 during human spermatogenesis. *Fertil Steril*, *108*(1), 47-54. https://doi.org/10.1016/j.fertnstert.2017.05.019

Pontén, F., Gry, M., Fagerberg, L., Lundberg, E., Asplund, A., Berglund, L., Oksvold, P., Bjorling, E., Hober, S., Kampf, C., Navani, S., Nilsson, P., Ottosson, J., Persson, A., Wernerus, H., Wester, K., & Uhlen, M. (2009). A global view of protein expression in human cells, tissues, and organs. *Mol Syst Biol*, *5*, 337. https://doi.org/10.1038/msb.2009.93

Ray, P. F., Toure, A., Metzler-Guillemain, C., Mitchell, M. J., Arnoult, C., & Coutton, C. (2017). Genetic abnormalities leading to qualitative defects of sperm morphology or function. *Clin Genet*, *91*(2), 217-232. https://doi.org/10.1111/cge.12905

Robertson, M. J., Kent, K., Tharp, N., Nozawa, K., Dean, L., Mathew, M., Grimm, S. L., Yu, Z., Legare, C., Fujihara, Y., Ikawa, M., Sullivan, R., Coarfa, C., Matzuk, M. M., & Garcia, T. X. (2020). Large-scale discovery of male reproductive tract-specific genes through analysis of RNA-seq datasets. *BMC Biol*, *18*(1), 103. https://doi.org/10.1186/s12915-020-00826-z

Roy, A., Lin, Y. N., Agno, J. E., DeMayo, F. J., & Matzuk, M. M. (2009). Tektin 3 is required for progressive sperm motility in mice. *Mol Reprod Dev*, *76*(5), 453-459. https://doi.org/10.1002/mrd.20957

Salanova, M., Chun, S. Y., Iona, S., Puri, C., Stefanini, M., & Conti, M. (1999). Type 4 cyclic adenosine monophosphate-specific phosphodiesterases are expressed in discrete subcellular compartments during rat spermiogenesis. *Endocrinology*, *140*(5), 2297-2306. https://doi.org/10.1210/endo.140.5.6686

Schrimpf, R., Gottschalk, M., Metzger, J., Martinsson, G., Sieme, H., & Distl, O. (2016). Screening of whole genome sequences identified high-impact variants for stallion fertility. *BMC Genomics*, *17*, 288. https://doi.org/10.1186/s12864-016-2608-3

Soumillon, M., Necsulea, A., Weier, M., Brawand, D., Zhang, X., Gu, H., Barthes, P., Kokkinaki, M., Nef, S., Gnirke, A., Dym, M., de Massy, B., Mikkelsen, T. S., & Kaessmann, H. (2013). Cellular source and mechanisms of high transcriptome complexity in the mammalian testis. *Cell Rep*, *3*(6), 2179-2190. https://doi.org/10.1016/j.celrep.2013.05.031

Suarez, S. S. (2008). Control of hyperactivation in sperm. *Hum Reprod Update*, *14*(6), 647-657. https://doi.org/10.1093/humupd/dmn029

Suarez, S. S., Varosi, S. M., & Dai, X. (1993). Intracellular calcium increases with hyperactivation in intact. *Proc. Natl. Acad. Sci. USA*, *90*, 4660-4664. https://doi.org/doi: 10.1073/pnas.90.10.4660

Timmons, P. M., Rigby, P. W. J., & Poirier, F. (2002). The murine seminiferous epithelial cycle is pre-figured in the Sertoli cells of embyonic testis. *Development*, *129*, 635-647. https://doi.org/doi: 10.1242/dev.129.3.635.

Tüttelmann, F., Rajpert-De Meyts, E., Eberhard Nieschlag, E., & Simoni, M. (2007). Gene polymorphisms and male infertility – a meta analysis and mini review. *Reproductive biomedicine online*, *15*, 643-658. https://doi.org/10.1016/s1472-6483(10)60531-7.

Tüttelmann, F., Ruckert, C., & Ropke, A. (2018). Disorders of spermatogenesis: Perspectives for novel genetic diagnostics after 20 years of unchanged routine. *Med Genet*, *30*(1), 12-20. https://doi.org/10.1007/s11825-018-0181-7

Uhlén, M., Fagerberg, L., Hallstrom, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, A., Kampf, C., Sjostedt, E., Asplund, A., Olsson, I., Edlund, K., Lundberg, E., Navani, S., Szigyarto, C. A., Odeberg, J., Djureinovic, D., Takanen, J. O., Hober, S., . . . Ponten, F. (2015). Proteomics. Tissue-based map of the human proteome. *Science*, *347*(6220), 1260419. https://doi.org/10.1126/science.1260419

Vogt, P. H. (2004). Molecular Genetic of Human Male Infertility: From Genes to New Therapeutic Perspectives. *Current pharmaceutical design*. https://doi.org/10.2174/1381612043453261

World Health Organization. (2021). *WHO laboratory manual for the examination and processing of human semen*. WHO. https://www.who.int/publications/i/item/9789240030787

2.6 Supplementary Material

(APPENDIX B – B.1, B.2, B3 contains raw datasets that were submitted to the journal. Appendix B.4 is the data mining summary, and a tabular version is presented as Suppl. Table 2.1)

Appendix B.1 – Raw dataset submitted to GEO – (GEO accession GSE224929) https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE224929

Token - cfsbeeiwvrafjyz

Appendix B.2 – RNA sequencing dataset – Bigmatrix

The Excel sheet contains the overall counts per each sample and the statistics.

Appendix B.3 –Differentially expressed genes in between groups

The Excel sheet contains all differentially expressed genes compared in between the three sample groups.

Appendix B.4 - Data mining summary of selected genes

The Excel sheet contains the summary of data collected from different portals. This includes the selected genes and their selection criteria, publications and expression of the gene in different cells and protein detection in tissues.

A tabular version is presented as Suppl. Table 2.1

Suppl. Table 2.1: Data mining summary of selected genes

Human Gene Symbol	PubMed references until at the time of selection	Gene found in Human and Mouse or Rat	The Human Protein Atlas (HPA)					
			Function/ description	Protein in organs	RNA expression in organs	RNAseq results	Predicted location	IHC
CFAP47 **(Selected under flagella category)**	12 references. None on spermatogenesis.	H, M	Transmembrane transport (mainly). Tissue enhanced (brain, choroid plexus, fallopian tube)	No data	Brain>Fallopian >Pituitary>testis	Endometrial ciliated cells (highest). E.Sptd>Spc>L. Sptds	Intracellular and Membrane	late pachytene, round spermatocytes stained. Some early pachytenes are lightly stained
PDE4A **(Selected from the cAMP group)**	78 references. None on spermatogenesis. Gene shows activity in spermatogenesis, present in spermatogonia, germ cells	H, M	Hydrolyzes the second messenger cAMP, which is a key regulator of many important physiological processes.	Low tissue specificity. Overview not available	Skeletal muscle>cerebral cortex>testis	L.Sptd (highest)>>>E. Sptd>Spc>Spg	Intracellular	Leydig cells (high), Seminiferous cells (medium)
ORAI1 **(Selected from the calcium group)**	Many - 2 on spermatogenesis. 1 study shows sterility in knockout mice (Davis et al,2016 - only mice samples)	H, M	The protein encoded by this gene is a membrane calcium channel subunit that is activated by the calcium sensor STIM1 when calcium stores are depleted	Testis (highest)	Skeletal Mus>Spleen>>>testis	Immune cell >>>Late Sptd>Spg>E.Sptd>Spc	Membrane	staining - high in Leydig cells, low in seminiferous cells
CACNB2 **(Selected from the**	76 references. None on spermatogenesis	H, M	Calcium transport, Ion transport, Transport. Regulatory subunit	Low tissue specificity. High in testis	Very low in testis	Low -Spg (RNA sig cell - Jintao and other)	Intracellular	Medium detection in cells in seminiferous duct, not

calcium group)			of L-type calcium channels					detected in leydig cells
SLC8B1 **(Selected from the calcium group)**	17 references. None on spermatogenesis. A study related to mitochondrial activity	H, M	Gene- SLC24A6 belongs to a family of potassium-dependent sodium/calcium exchangers that maintain cellular calcium homeostasis through the electrogenic countertransport of 4 sodium ions for 1 calcium ion and 1 potassium ion	Low specificity	Adrenal gland (highest). Low expression in testis	very low in germ cells	Intracellular, Membrane (different isoforms)	Stained leidig cells, and seminiferous cells are not stained. But expressed in testis
TMEM37 **(Selected from the calcium group)**	5 references. None on spermatogenesis	H, M	Protein - Calcium transport, Ion transport, Transport	Testis (low tissue specificity)	Low in testis	Adipocytes>>>Spg>Sertoli	Nucleoplasm, Nucleoli fibrillar center, Cytosol (sub cellular)	Low stained seminiferous cells and medium stained leydig cells
TNC **(Highly downregulated in SDA)**	264 references. Study on OAT hormone treatment	H, M	Extracellular matrix protein implicated in guidance of migrating neurons as well as axons during development, synaptic plasticity as well as neuronal regeneration	Not given	Smooth muscles	L.Sptd>E.Sptd (RNA tissue, Jintao dataset shows highest in arystroglial cells)	Intracellular, Secreted (different isoforms)	Not detected in cells in SD. Medium stained leydig cells. Antibody HPA004823 stains SD cells well
SPATA31E1 **(Selected from reproduction and**	3 references. None on spermatogenesis but the UniProt gene is for spermatogenesis	H, R	May play a role in spermatogenesis	Only in testis	Testis, brain	L.Sptd>E.Sptd (RNA tissue) in Jintao dataset E. Sptd>L.Sptd)	Membrane	Late spermatocytes – late pachytenes and round spermatids

spermatogenesis group)								
TEKT3 **Selected from reproduction group)**	13 references. 1 study shows the expression in male germ cells	H, M	May be a structural component of the sperm flagellum. Required for normal sperm mobility	Falopian > bronchus > testis	Testis (highest)	E.Sptd=L.sptd	Nucleoplasm, Plasma membrane	Highly stained Leydig cells and medium stained SD cells with HPA073565 antibody
SLC9C1 **Selected from reproduction and spermatogenesis group)**	8 references. 1 study sperm motility (Wang et al, 2003)	H, M	Sperm-specific sodium/hydrogen exchanger involved in intracellular pH regulation of spermatozoa. Required for sperm motility and fertility.	Not given	Testis (highest)	E.Sptd, L.Sptd	Membrane	No data on IHC in the atlas

Abbreviations HPA – The Human Protein Atlas. RNAseq – RNA sequencing. Spc = spermatocytes, Spg = spermatogonia, E. Sptd = early spermatids, L. Sptd = late spermatids. H = human, M = mouse, R = rat. RNAseq data – (Guo et al., 2018; Norreen-Thorsen et al., 2022; Uhlén et al., 2015). Protein in organs - (Pontén et al., 2009)

Chapter 3

Use of mass spectrometry to identify proteins related to impaired sperm morphology and motility in infertile men.

3.1 Introduction

Asthenozoospermia and teratozoospermia are two common ejaculated sperm phenotypes that contribute to male infertility. Asthenozoospermia is defined as low or zero progressive motility of sperm (< 32%) by the World Health Organization (WHO) (World Health Organization, 2021). It is one of the most prevalent conditions and is diagnosed in up to 80% of male infertility cases (Tu et al., 2020). Asthenozoospermia can be caused by structural defects in the sperm flagellum, e.g. in the core axoneme or the peri-axonemal structures (mitochondrial sheath, fibrous sheath), as well as structural deficits in the sperm, metabolic or functional defects in the testis and/or epididymis or sex glands can also contribute to asthenozoospermia (Cao et al., 2018; Tu et al., 2020). In teratozoospermia, morphologically defective sperm are observed in an ejaculate (morphologically normal sperm <4%) (World Health Organization, 2021) (WHO 2021). Both astheno- and teratozoospermia can often be observed together as asthenoteratozoospermia (AT), as many morphological abnormalities in the sperm flagella also affect sperm motility (Tu et al., 2020).

The molecular basis of sperm motility is not fully understood (Cao et al., 2018). Many genes (and ultimately proteins) operate during spermiogenesis to enable the morphological and functional differentiation of the spermatozoa. Testicular spermatozoa are morphologically fully differentiated but functionally immature and incapable of motility until they transit the epididymis. Proteins produced in the epididymis and transferred to spermatozoa as they traverse the epididymal duct influence sperm maturation and function (Eskandari-Shahraki et al., 2020). Contact with fluids secreted by the epididymis, seminal vesicles and prostate (termed seminal plasma) also influence sperm function (Barrachina et al., 2022). The proteins in the ejaculate originate from spermatozoa, non-spermatogenic cells, and seminal plasma proteins (Agarwal et al., 2020). Out of these, the structural proteins of a spermatozoa can directly influence the integrity, morphology and motility of the sperm (Tu et al.,

2020). Thus, there are many proteins detected in the human ejaculate and these can influence, and reflect, sperm function (Agarwal et al., 2020).

Identification of the sperm proteome has previously been performed using conventional methods such as 2D gel electrophoresis combined with matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) (Martinez-Heredia et al., 2006). The identification of differentially expressed proteins (DEP) between different sample types was achieved using an advanced version of 2D gel electrophoresis; i.e. difference gel electrophoresis (DIGE) and by using image software to visualise the colour intensities of different staining dyes (Gupta et al., 2014; Hamada et al., 2013). The current most advanced proteomic techniques are MALDI-TOF and liquid chromatography-tandem mass spectrometry (LC/MS/MS) (Agarwal et al., 2020). Mass spectrometry has developed in recent years as the preferred method for deep characterisation and understanding of proteins in biological systems (Cravatt et al., 2007). Mass spectrometry is used to identify the proteome of a sample and can also identify novel protein sequences in different protein isoforms (Agarwal et al., 2020; Baldwin, 2004; Bitton et al., 2010; Castellana et al., 2014; Khan et al., 2018; Klasberg et al., 2016). Advancements in mass spectrometry, bioinformatics and computational analysis have been applied to the study of human spermatozoa to understand the molecular pathways involved in sperm function, protein transportation, post-translational modifications, and the role of proteins in various pathogenic conditions leading to male infertility (Agarwal et al., 2020).

The first human sperm proteome study using mass spectrometry in 2005 reported 1760 proteins in human spermatozoa (Johnston et al., 2005). Later studies reported a higher number of proteins with known roles in spermatozoa and metabolic pathways involved in sperm function (Agarwal et al, 2020). Approximately 1000 proteins detected in the sperm flagellum likely play a role in maintaining its structure and thus play crucial roles in sperm motility (Amaral et al., 2014; Baker et al., 2013). Sperm flagellum proteins also contribute to two transport systems involved in the development of the sperm tail, i.e. intra-manchette transport (IMT) and intraflagella transport (IFT). Abnormalities in flagellar structural proteins or the disturbances in the transport systems (IMT/IFT) can severely affect the proper functioning of the spermatozoa (Tu et al., 2020). Proteomic studies conducted on asthenozoospermic samples have identified DEPs in five categories: energy and metabolism, structural organisation,

antioxidant activity, transportation and signaling and stress management (Agarwal et al., 2020; Cao et al., 2018).

Previous studies have suggested that, due to the complexity of spermatogenesis, the abnormalities observed in asthenozoospermia and teratozoospermia are most likely not induced by single gene defects, although several previous studies revealed monogenic causes for these conditions (see chapter 1 eg: *SPATA16, DPY19L2* in globozoospermia). Gene mutations or differential gene and protein expression, particularly in the spermiogenesis process, can influence the structural and functional aberrations of the sperm (Tu et al., 2020). Secondly, post-testicular modifications play a major role in the ultimate function of mature sperm (Björkgren & Sipilä, 2019). There are many genes and proteins involved in sperm morphology and function and a lack of knowledge regarding the exact molecular biological pathways involved (Tu et al., 2020). By analysing clinically well-defined samples and isolated sub-compartmental proteins, i.e. sperm head and tail separately, proteins that are important in spermatozoal structural integrity can be revealed (Baker et al., 2013). Determining the localisation of DEPs to specific germ cell stages and/or specific components of mature spermatozoa can further confirm the importance of identified proteins in sperm function.

Identification of proteins dysregulated in abnormal (AT) sperm is important because it will provide information as to the causes of this phenotype but can also identify proteins that could be useful as biomarkers for the diagnosis of sperm phenotypes. Furthermore, identifying testis-specific proteins that have specific roles in sperm function is important for the identification of potential targets for non-hormonal male contraceptive strategies. The aim of this study was to identify novel proteins that are involved in sperm morphology and motility by comparing well-defined asthenoteratozoospermic (AT) samples that have no known genetic cause, with a well-defined normozoospermic (NORM) control group by using high throughput mass spectrometry and bioinformatics. By using whole ejaculates without separating spermatozoa, the study analysed all proteins that are derived from motile and immotile spermatozoa, seminal fluids and proteins acquired during post testicular modifications. Also, the study was designed to use a bioinformatics strategy and data mining to specifically target novel proteins related to this aim.

3.1.1 Hypotheses

- That altered protein expression/localisation in functionally and phenotypically abnormal human spermatozoa could contribute to defective sperm morphology and motility and thus male infertility.
- That the identification of proteins that are altered in AT could reveal new strategies for the design of therapies or the design of male contraceptives.

3.1.2 Research Objectives

3.1.2.1 Main objective

To identify novel proteins that are involved in (a) altered sperm morphology and motility and/ or (b) epididymal maturation of sperm in infertile men.

3.1.2.2 Specific Objectives

- To identify proteins that differ in ejaculated sperm with AT compared to normal sperm by mass spectrometry.
- To use *in silico* analysis for the significantly differentially expressed proteins (DEPs) between the two groups in order to select a cluster of most promising proteins for further analysis.
- To further investigate the localisation of differentially expressed proteins (DEPs) in testis biopsies from men with normal or defective spermatogenesis to understand how these proteins are involved in germ cell / spermatozoa development.
- To verify the changes in differentially expressed proteins by quantitative Western blotting.

3.1.3 Experimental design

Human ejaculates determined as normal or abnormal (asthenoteratozoospermia - AT) according to WHO guidelines were selected for the study. Frozen ejaculates of normal (NORM, n=3) and AT (n=3) were subjected to mass spectrometry (MS). Protein abundances and statistical values were obtained by the software packages used at the Institute of Biochemistry. Proteins that were significantly different ($p < 0.05$)

between the two phenotypes were classified as Differentially Expressed Proteins (DEP). The potential functional roles of DEPs in sperm development and function were assessed using online datasets including UniProt, GermOnline, ENSEMBL, The Human Protein Atlas (HPA), The Young Adult Testis Atlas, MGI, GeneCards and PubMed (Fig. 3.1). The localisation of five selected proteins in testis biopsies was determined by immunohistochemistry (IHC) to investigate their involvement in germ cell development and quantitative Western blotting was performed to verify the changes observed by mass spectrometry (Fig. 3.1). All the experiments except for mass spectrometry were performed by the PhD candidate.

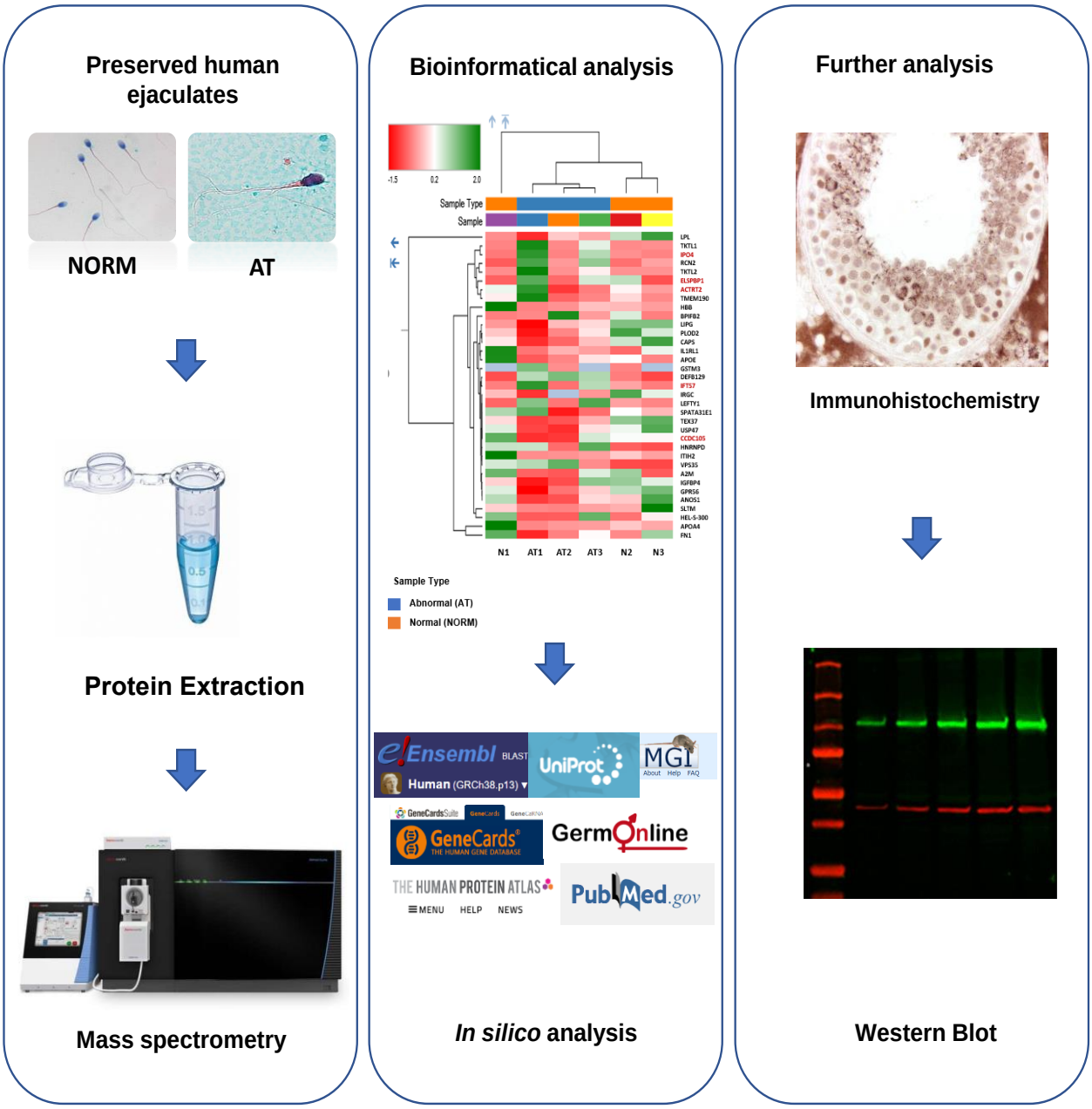


Figure 3.1: Figure legend in the next page

Figure 3.1: The experimental design of the study. The study was designed to perform mass spectrometry on human ejaculates with two different phenotypes – NORM – normal, AT – asthenoteratozoospermia. Once differentially expressed proteins were identified, *in silico* approaches were used to investigate their potential functional relevance and to select candidate proteins. Candidate proteins were then further investigated by immunohistochemistry in testis biopsies to determine their potential functions. Finally, quantitative Western blots were used to validate the mass spec results in a different set of samples.

3.2 Material and Methods

3.2.1 Sample collection

Human ejaculates were collected at the Clinic of Urology, Paediatric Urology and Andrology, Justus- Liebig University of Giessen, Germany. Ejaculates were analysed according to the WHO guidelines (World Health Organization, 2021) and categorised as normal (n=3) or asthenoteratozoospermia (AT, n=3) with morphological defects (e.g. head and tail defects) and zero motility, without any known genetic cause. All the samples had a sperm count of > 15 million/mL defined as normal by the WHO guidelines. Frozen ejaculates from normal (NORM, n=3) and AT (n=3) samples were subjected to mass spectrometry analysis at the Institute of Biochemistry, JLU. Ejaculates were used without a density gradient separation to obtain information on proteins related to spermatozoa and epididymal maturation. The data was analysed using custom statistical and bioinformatics software packages used at the Institute of Biochemistry, JLU.

Testis biopsies were collected at the Clinic of Urology, Paediatric Urology and Andrology, Justus-Liebig University of Giessen, Germany and the Centre of Reproductive Medicine and Andrology (CeRA), University Clinic Munster, Germany. Patients were diagnosed as obstructive or non-obstructive azoospermia after diagnosis by clinical parameters (sex hormones, ultrasound and ejaculate analysis) (Jungwirth et al., 2018; World Health Organization, 2021). Surgeries were performed after written informed consent and after approval by the local Ethics Committee (AZ 26/11). Biopsies for histological evaluation were fixed in Bouin’s solution overnight, embedded into paraffin, 5 µM sections were prepared and stained with hematoxylin

and eosin following standard protocols, and the phenotype was evaluated by clinical score count analysis (Fietz & Kliesch, 2023). Samples used in this study were classified as intact spermatogenesis (normal spermatogenesis, n=5; contains a full complement of germ cells; score = 10;), spermatid arrest/+mixed atrophy (SDA, n=3; containing germ cells up to round spermatids with occasionally observed elongating spermatids, and somatic cells; score = 4-5), spermatocyte arrest (SZA, n=3; containing germ cells up to pachytene spermatocytes and somatic cells, score = 0), and Sertoli cell only (SCO, n=5; absence of germ cells; score =0).

3.2.2 Mass Spectrometry

3.2.2.1 Sample extraction and digestion

Lyophilised samples were extracted with lysis buffer consisting of 6 M urea (Sigma, Taufkirchen, Germany), 2 M thiourea (Sigma), 4 % 3-3´-(Cholamidopropyl)-3,3-dimethylammoniumpropylsulfate (CHAPS; Roth, Karlsruhe, Germany), and 30 mM dithiothreitol (DTT; Fluka, Seelze, Germany). For each sample, 100 µg of dissolved proteins were reduced with 5 mM DTT (30 min at 50°C) and cysteines were modified with 10 mM iodoacetamide (30 min at 24°C; Sigma). The reaction was quenched with an excess of cysteine (Sigma) and trypsin (sequencing grade, Promega, Mannheim, Germany) added at a protein-enzyme ratio of 40:1 in a final volume of 100 µl 25 mM ammonium bicarbonate (Fluka). After incubation for 16 hours at 37°C, the digestion was stopped by the addition trifluoroacetic acid (TFA, Applied Biosystems™, Warrington, UK) to a final concentration of 1%.

3.2.2.2 Peptide desalting and purification

The acidified samples were applied to activated Chromabond C18ec cartridges (Macharey-Nagel, Germany; 100 mg adsorbent weight, 1 mL bed volume). Retained peptides were washed with 0.1% TFA and eluted with 80% ACN in 0.1% TFA. The eluate was lyophilised overnight, re-dissolved in 50 µL of HEPES buffer (50 mM, pH 8.5; Sigma) and protein concentrations were determined using NanoDrop 2000c (Thermo Scientific, MA, USA).

3.2.2.3 TMT labelling of peptides

For isobaric labelling of the six individual samples, 17 µg of the tryptic peptides of each sample (dissolved in 50 µl of 50 mM HEPES, pH 8.0) was incubated at room temperature with 100 µg of one label of the TMT6plexIsobaric Label Reagent Set (Thermo Scientific, MA, USA) dissolved in 20 µL of ACN for 60 minutes. Unused labelling reagents were quenched by the addition of 4.2 µL of 5% cysteine prepared in HEPES buffer. Finally, the six samples were mixed together to yield one master sample and dried under vacuum.

3.2.2.4 High pH reversed-phase peptide fractionation

To increase the number of peptide identifications in the Liquid Chromatography-Electrospray Ionization-Mass Spectrometry (LC-ESI-MS) analysis, the labelled master sample was fractionated using the High pH Reversed-Phase Peptide Fractionation Kit (Pierce™, Thermo Scientific, MA, USA), according to the manufacturer's protocol. Briefly, the dried labelled peptides were resuspended in 300 µL of 0.1% TFA and loaded on to the pre-conditioned C18 spin column. After washing once with water and once with 5% ACN in 0.1% trimethylamine (TEA), labelled peptides were eluted by eight steps of increasing concentration of ACN in 0.1% TEA (10%, 12.5%, 15%, 17.5%, 20%, 22.5%, 25% and 50%). Each fraction was then lyophilised overnight and reconstituted with 20 µL of 0.1% TFA and analysed by LC-MS.

3.2.2.5 Liquid Chromatography and Tandem Mass Spectrometry (LC-MS/MS/MS)

One µg of each of the eight fractions was loaded onto a 50 cm µPAC™ C18 column (Pharma Fluidics, Gent, Belgium) in 0.1% formic acid (Fluka) at 35°C. Peptides were eluted with a linear gradient of ACN from 3% to 44% over 240 min followed by a wash with 72% acetonitrile at a constant flow rate of 300 nl/min (ThermoScientific™UltiMate™3000RSLCnano) and infused via an Advion TriVersa NanoMate (Advion BioSciences, Inc. New York, USA) into an Orbitrap Eclipse Tribrid mass spectrometer (Thermo Scientific). The mass spectrometer was operating in positive-ionization mode with a spray voltage of the NanoMate system set to 1.6 kV and source temperature at 300°C. Using the data-dependent acquisition mode, the

instrument performed full MS scans every 2.5 seconds over a mass range of m/z 400–1600, with the resolution of the Orbitrap set to 120000. The RF lens was set to 30%, auto gain control (AGC) was set to standard with a maximum injection time of 50 ms. In each cycle the most intense ions (charge state 2-6) above a threshold ion count of 5.000 were selected with an isolation window of 0.7 m/z for CID-fragmentation at normalized collision energy of 35% and an activation time of 10 ms. Fragment ion spectra were acquired in the linear IT with a scan rate set to rapid and mass range to normal and a maximum injection time of 100 ms. After fragmentation, the selected precursor ions were excluded for 20s for further fragmentation. From each MS/MS cycle, up to ten fragment ions were selected with an isolation window of 3 m/z for HCD-fragmentation at normalised collision energy of 55%. MS3-fragment ion spectra were acquired in the Orbitrap with a resolution of 50.000. The mass range was set to 100 – 500 m/z, maximum injection time to 100 ms and AGC to 300.

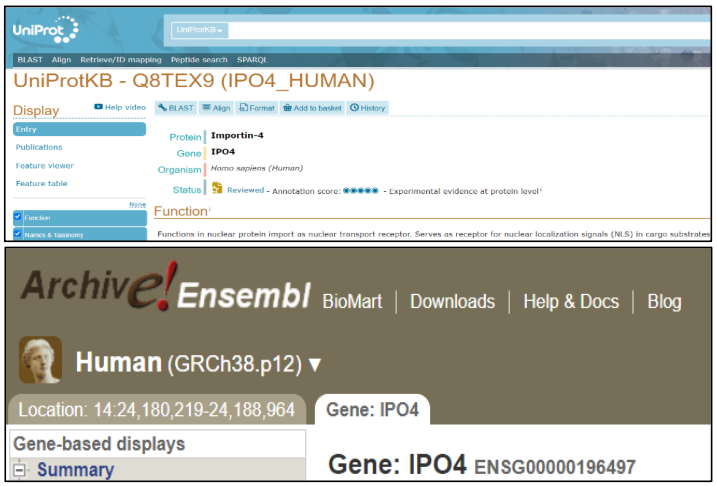
3.2.2.6 Protein identification and quantitation

Data were acquired with Xcalibur 4.3.73.11. (Thermo Fisher Scientific). The eight individual result files were assembled into one master result file and analysed with Proteome Discoverer 2.4.0.305 (Thermo Fisher Scientific). Sequest HT (Proteome Discoverer version 2.5.0.400; Thermo Fisher) was used to search against UniProt human database (*20201007, 70502241 residues, 214889 sequences*) which also extracted the quantitation data from the 6 TMT tags and performed the label-based quantitative analysis. A precursor ion mass tolerance of 10 ppm was used, and one missed cleavage was allowed. Carbamidomethylation of cysteines and labelling with TMT10plex at peptide N-termini and lysines side chains were defined as static modifications with optional oxidation of methionine. The fragment ion mass tolerance was set to 0.6 Da for the linear IT MS2 detection. The FDR (False discovery rate) for peptide identification was limited to 0.01 by using a decoy database. Protein identifications were accepted with a decoy FDR cutoff of less than 1%. Protein identifications were accepted if they could be established with at least 1 unique identified peptide of a length between 6 and 144 amino acids. TMT reporter ion values were quantified from MS3 scans with an integration tolerance of 20 ppm. Proteins were

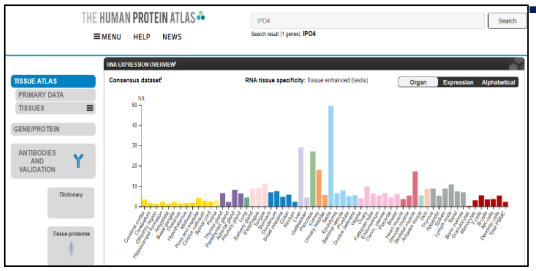
defined as differentially expressed (DEP) when the fold-change (FC) between different groups was > 0.58 or < − 0.58 and the p-value of the student's t-test was < 0.05.

3.2.3 *In silico* analysis

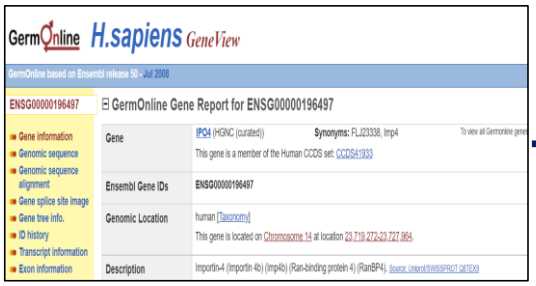
DEPs were subjected to *in silico* analysis by browsing and collecting information from publicly available online databases: The Human Protein Atlas (HPA), UniProt, GermOnline, The Young Adult Testis Atlas (YAA), Mouse Genome Informatics (MGI), and GeneCards (Fig. 3.2). PubMed was used to investigate DEPs reported in the published literature. The data collected through this datamining process and the rationale behind the selection of the proteins are summarised in the Results section. Proteins with previously defined functional roles in sperm development were omitted from further investigation since the objective of the study was to identify novel proteins involved in human sperm morphology and motility.



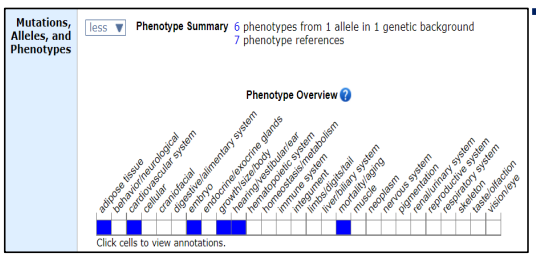
Cross check transcript IDs or protein accession number obtained from Mass Spec data sheets. Obtain transcript variants and GO biological function



- **Predicted location**
- **Cellular component**
- **Subcellular summary**
- **Tissue enhanced**
- **Cell type enhanced**
- **Protein expression in organs**
- **RNA expression in organs**
- **RNA expression in different cell types**
- **IHC results in testis**



- **Main function, chromosome location, mouse ortholog, microarray tissue profiling**



- **Mouse ortholog, availability of knock-out mice, phenotype annotations related to the biological system**

- **Related findings in published literature**

Figure 3.2: Overview of the workflow used to investigate the functional significance of DEPs in spermatogenesis and/or sperm function.

3.2.4 Immunohistochemistry (IHC)

IHC was performed to localise proteins selected in testis biopsy sections classified as NSP (n=5), SDA (n=3), SZA (n=3), and SCO (n=5), using previously described methods (Pleuger et al., 2017). Specific primary and secondary antibodies are listed in Table 3.1. Briefly, 5 µM thick paraffin sections were mounted on coated slides, de-waxed and rehydrated using xylol and a decreasing alcohol series (100%, 96%, 80%, 70%), respectively. Heat-mediated antigen retrieval was performed in citrate buffer (pH 6) for 20 minutes in a microwave oven (5 minutes at high heat- 800W, and 15 minutes at medium heat - 520W). Sections were then incubated in 3% hydrogen peroxide (H_2O_2) in Tris-EDTA buffer to inhibit endogenous peroxidase activity, followed by incubation in 1.5% Bovine serum albumin (BSA) (ROTH GmbH) in Tri-EDTA, for 30 minutes to block non-specific binding. Sections were incubated overnight at 4°C with primary antibodies in optimised dilutions (Table 3.1). One tissue section of each slide was used as the negative control where primary antibody was omitted and distilled water was added. The next day, sections were washed and incubated with the respective horseradish peroxidase-labelled secondary antibodies (Table 3.1) at room temperature for one hour. Proteins were detected using the horseradish peroxidase substrate NovaRed which produces red staining (VECTOR Laboratories). After counterstaining with hematoxylin (MORPHISTO Laborchemikalien & Histologieservice GmbH), sections were dehydrated and mounted using Kaiser's gelatin (Merck, Darmstadt, Germany). A Leica DM750 microscope was used to observe the samples and slides were scanned using the VS 120 virtual slide microscope (OLYMPUS). For the comparison of immunostaining between testis biopsies from different phenotypical groups, NSP, SDA, SZA and SCO biopsies were included in the same experiments. A suitable tissue biopsy (other than testis) was used as a negative control for each protein.

Table 3.1 – Antibodies used for immunohistochemistry.

No.	Primary antibody		Secondary antibody	
	Manufacturer details	**Dilution (diluent – Tris-EDTA buffer)**	**Manufacturer details**	**Dilution (diluent – Tris-EDTA buffer)**
01	Anti-IPO4 antibody Sigma life science Prestige antibodies provided by Atlas antibodies. (HPA039043)	1:500	Peroxidase conjugated AffiniPure Goat Anti-Rabbit (Jackson Immuno Research Laboratories)	1:400
02	Anti-ELSPBP1 Sigma life science Prestige antibodies provided by Atlas antibodies. (HPA044256)	1:1000	Peroxidase conjugated AffiniPure Goat Anti-Rabbit IgG (Jackson Immuno Research Laboratories)	1:400
03	Anti-ACTRT2 Sigma life science Prestige antibodies provided by Atlas antibodies. (HPA025079)	1:600	Peroxidase conjugated AffiniPure Goat Anti-Rabbit (Jackson Immuno Research Laboratories)	1:400
04	Anti-CCDC105 Sigma life science Prestige antibodies provided by Atlas antibodies. (HPA058585)	1:400	Peroxidase conjugated AffiniPure Goat Anti-Rabbit (Jackson Immuno Research Laboratories)	1:400
05	Anti-IFT57 Sigma life science Prestige antibodies provided by Atlas antibodies. (HPA035514)	1:200	Peroxidase conjugated AffiniPure Goat Anti-Rabbit (Jackson Immuno Research Laboratories)	1:400

3.2.5 Protein extraction

Previously collected human ejaculates stored at -20°C were thawed on ice and washed with 400 µL of ice-cold 1X Phosphate Buffered Saline (PBS) buffer by centrifuging at 4000 rpm for 5 minutes. The supernatant was discarded and 500 µL of ice-cold RIPA buffer and 5 µL of Halt-Protease inhibitor cocktail (Thermo Fisher Scientific) was added (except 200 µL of RIPA and 2 µL of protease inhibitor was added to those ejaculates that had comparatively a lower sperm count). The samples were homogenised using 19G and 21G needle-syringe units. If needed, the homogenising rod was used to break the pellets. For the human testis tissue sample used as a control, 500µL of ice-cold RIPA and 5µL protease inhibitor were added and the tissue was homogenised using the homogenising rod. The mixtures were then briefly

centrifuged at 13000 rpm for 2 minutes to reduce the foam and kept at 4°C for 2 hours on a shaker plate. Samples were then centrifuged at 16000 rpm for 20 minutes at 4°C, and the final supernatant was stored at -20°C. All the steps were carried out on ice to reduce protein degradation.

3.2.6 Determination of the protein concentration

Protein concentration was determined using the DC protein assay (DC Protein Assay, BioRad laboratories) and the assay was carried out according to the manufacturer's protocol. Briefly, 2000mg/mL of Bovine Serum Albumin (BSA) was prepared, first by diluting the required amount of BSA powder in RIPA buffer, and subsequently preparing a 1:1 dilution series. Three dilutions (1:1) per extracted protein sample were made by serially diluting them in RIPA buffer (Annex doc). In a flat bottom culture plate (FALCON), 5µL of each dilution (standard and sample) were added in triplicates. The solution A+S was made by mixing reagent A and S in appropriate volumes (20µL of reagent S to each 1000µL of reagent A and kit reagents were added according to the manufacturer's protocol) and 25 µL of this mixture was added to each well. Finally, 200 µL of reagent B was added to each well and the plate was incubated for 15 minutes at room temperature. The absorbance was read at 750 nm using the CLARIOstar® plate reader (BMG LABTECH). A standard curve was plotted in Excel and the unknown concentrations were extrapolated using the formula obtained from the standard curve (Fig. 3.3).

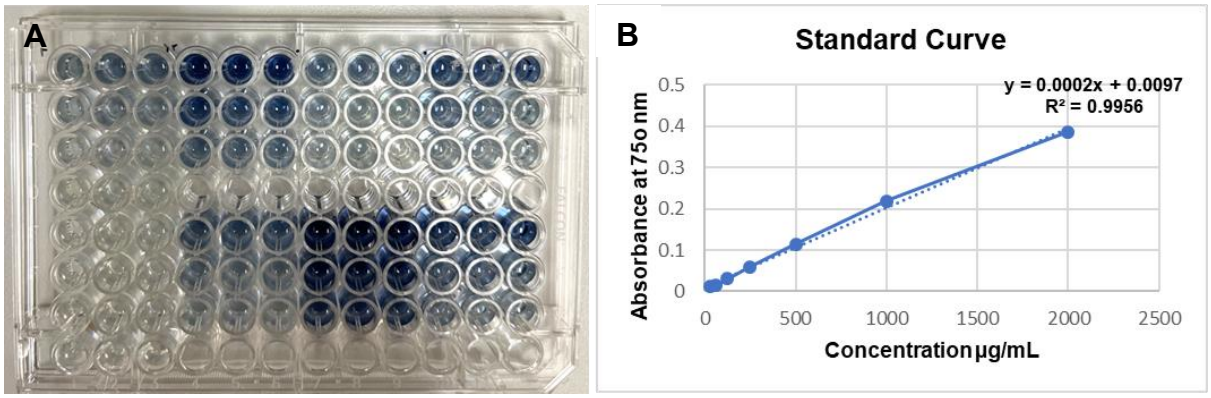


Figure 3.3 – Determination of protein concentration. The tissue culture plate with the dilution series of the standard curve and proteins **(A)** and the standard curve extrapolated **(B).**

3.2.7 Western Blotting – The Fluorescent method

The precast gels (BioRad Mini PROTEAN TGX Stain-free Precast Gels 10%) were placed in the gel chamber and filled with 1X running buffer (Annex doc). The combs were carefully taken out and the wells were checked for any distortions. Protein lysates stored at -20°C were thawed on ice and the required volume was taken to load 30µg of protein per well. A mixture of 2X sample buffer (BioRad) and β-Mercaptoethanol (BME, Sigma Aldrich) (in 1:10 dilution – 50µL of BME added to 450µL of 2X SB) was added to the required amount of the sample in a 1:1 ratio. The mixture was then heated for 5 minutes at 70°C. Briefly, cooled samples and a molecular weight marker (Page Ruler Pre-Stained Protein Ladder, Thermo Scientific) were loaded into the wells (according to the gel plan and a negative control lane was maintained with no protein product) of the precast gels and run at 110V for 1.30 – 1.45 hours (BioRad). After the run, the gel was carefully removed from the precast and transferred onto a Nitrocellulose (NC) membrane. The transfer apparatus was arranged as shown in Supplementary Fig. 3.1 The apparatus was filled with 1X transfer buffer (Annex doc) and the overnight transfer was done at 30V (BioRad) and 4°C. The next day, the NC membranes were cleaned with running distilled water and 1X TBS buffer and treated with Ponceau staining for 15 minutes to determine the successful running and blotting of the samples. The membranes were cut accordingly, washed 4 times in TBST (1X TBS and 0.1% Tween 20) and blocked with 5% skim milk for 1 hour. After washing off the blocking solution, membranes were incubated with the appropriate primary antibody in 3% milk solution for overnight at 4°C. The next day, membranes were

incubated with the anti-alpha tubulin monoclonal antibody (Table 3.2) to detect alpha tubulin protein as the loading control, and then with the appropriate fluorescence-tagged secondary antibodies (Table 3.2) diluted in 3% milk at room temperature. Membranes were washed 4 times in TBST after each antibody incubations. Finally, the membranes were transferred to 1X TBS solution and images were taken using the Odyssey scanner (Odyssey® CLx Infrared Imaging System by LI-COR). Images were saved in Tiff format and the quantitation values of the reference protein, and the target protein were determined using ImageJ software. The relative quantitation plots were created using GraphPad Prism (version 9.4.1). All negative control (NC) lanes in the fluorescent WB blots were clear.

3.2.8 Western Blotting – the Chemiluminescence method

The precast gels (BioRad Mini PROTEAN TGX Stain-free Precast Gels 10%) were set in the gel apparatus appropriately and the inner and outer chambers were filled with MOPS buffer (after adding the antioxidant) (Annex doc). Samples mixed with the reducing agent and sample buffer in calculated volumes and 5µL of the marker (Thermo Fisher) were loaded in the wells of the precast gels (according to the gel plan and a negative control blot was maintained with no protein product). Samples were run at 200V for 1 hour. After the run, gels and PVDF membranes were arranged in the transfer apparatus (Suppl. Fig. 3.1), filled with 1X Transfer buffer (Annex doc), and transferred for 1 hour at 30V, 4°C. The membranes were cleaned with distilled water, cut appropriately, and blocked for 1 hour in 5% milk. Membranes were treated overnight at 4°C with respective primary antibodies diluted in TBST (Table 3.2). The next day, membranes were washed in TBST and blocked for 30 minutes with goat serum. Membranes were then treated with appropriate HRP-conjugated secondary antibodies for 1 hour at room temperature. Membranes were washed 4 times in TBST after each treatment, except for the anti-goat serum blocking step. Membranes were incubated for 2 minutes with ECL select (Thermo Scientific™ Pierce™ ECL Plus Western Blotting Substrate), the excess was tipped off and the membranes placed on X-ray films. The activated X-ray film was incubated in the developer solution for 2-3 minutes (until bands appeared) then about 30 seconds in the stop solution, washed in distilled water and dried. Alternatively, blots after the ECL treatment were imaged using

the iBright machine (Invitrogen iBright gel documentation system by Thermo Fisher Scientific) to visualise the relevant protein bands. The quantitation values were obtained using ImageJ software and the relative quantitation plots were created using the software GraphPad Prism (version 9.4.1). The negative control (NC) lanes were clear in all Westerns, except for the re-probed chemiluminescence blot of IFT57 which showing a non-specific band at a higher molecular weight in the NC lane.

Table 3.2 – Antibodies used for Western Blotting

	Primary antibody			
	Manufacturer details	**Host**	**Dilution In fluorescent WB**	**Dilution in Chemi. WB**
01	Anti-ACTRT2 Sigma life science Prestige antibodies provided by Atlas antibodies (HPA025079)	Rabbit	1:1500	1:5000
02	Anti-ELSPBP1 Sigma life science Prestige antibodies provided by Atlas antibodies (HPA044256)	Rabbit	1:2000	1:5000
03	Anti-IFT57 Sigma life science Prestige antibodies provided by Atlas antibodies (HPA035514)	Rabbit	1:500	1:500
04	Anti-IPO4 antibody Sigma life science Prestige antibodies provided by Atlas antibodies (HPA039043)	Rabbit	1:1500	1:5000
05	Anti-CCDC105 Sigma life science Prestige antibodies provided by Atlas antibodies (HPA058585)	Rabbit	1:500	1:2000
	Loading control			
04	Monoclonal Anti-a-Tubulin Clone B-5-1-2 produced in mouse, ascites fluid Sigma-Aldrich T5168	Mouse	1:5000	-
	Secondary antibodies			

05	Goat anti-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor™ 680 from Thermo Fisher Scientific, catalog # A-21058	Goat	1:10000	-
06	Goat anti-Rabbit IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor™ 790 from Thermo Fisher Scientific, catalog # A11369	Goat	1:5000	-
07	Peroxidase conjugated AffiniPure Goat Anti-Rabbit IgG (Jackson Immuno Research Laboratories)	Goat	-	1:40,000

WB – Western Blotting

3.3 Results

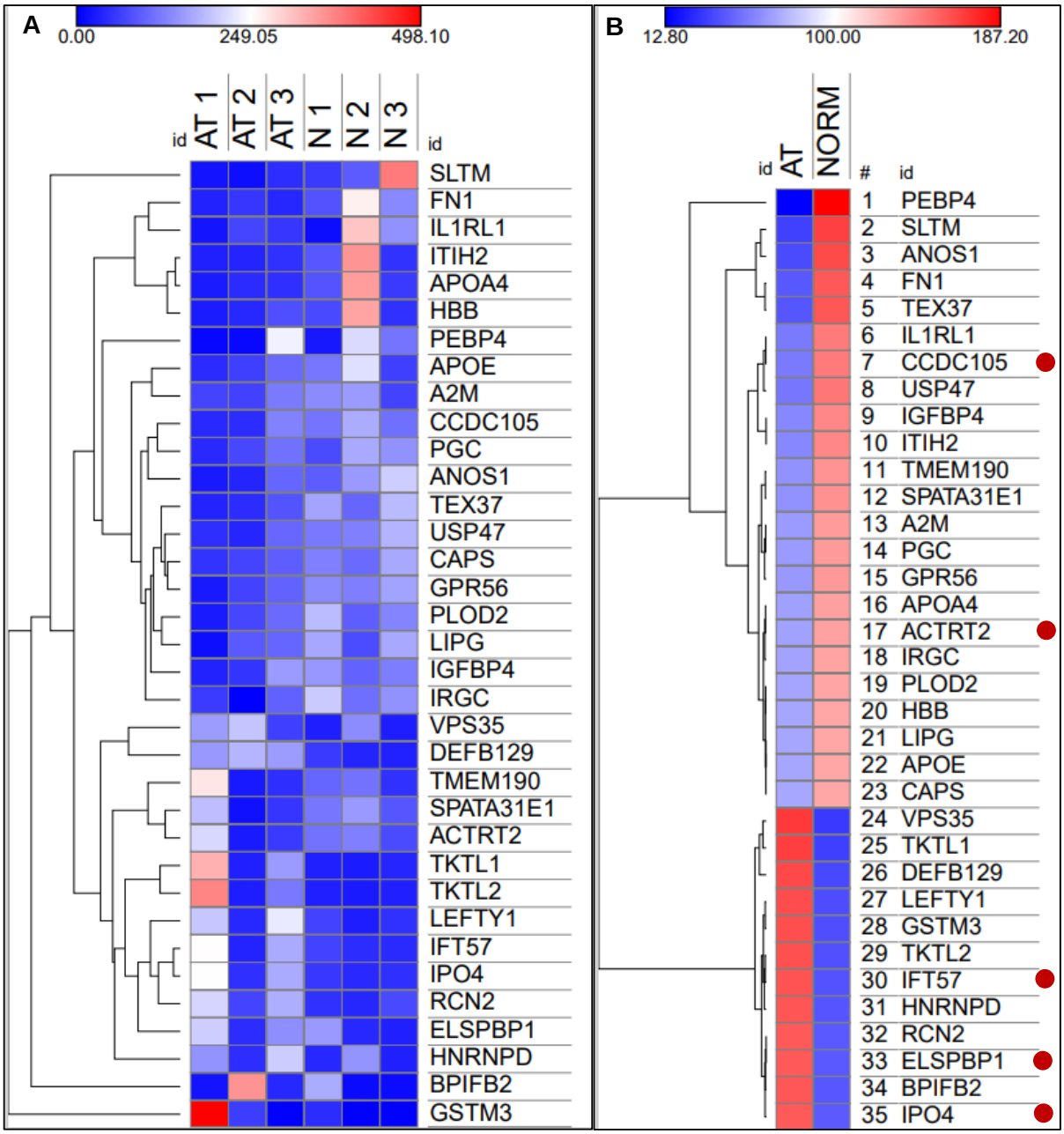
3.3.1 Proteins selected from Mass spectrometry and *in silico* analysis

The mass spectrometry analysis detected a total of 2720 proteins in NORM and AT samples and 1725 proteins were quantitatively compared between groups. The proteins below the cut off of P<0.05 and fold change of 0.58 =< log2FC <= -0.58 were considered as DEPs and selected for further analysis. Statistical analyses revealed that the abundance of 35 proteins was significantly different between NORM and AT samples (Table 3.3), although there was between-sample variability as is commonly observed in human semen samples (Fig 3.4). These 35 proteins were further examined as explained in Figure 3.2. Of these, 25 proteins were significantly decreased, and 10 proteins were significantly increased in the AT group compared to NORM (Table. 3.3). The results of the bioinformatical analysis are summarised in Tables 3.4, 3.5, 3.6. Furthermore, the DEP list was submitted to DAVID Bioinformatics portal (Huang da et al., 2009; Sherman et al., 2022) to obtain functional annotation clustering and table. There were 21 proteins that clustered under ‘signal’’, 16 proteins under ‘secreted’ and 15 proteins that clustered under 'extracellular region’’ among other annotation terms. Out of the 35 DEPs, two proteins were epididymis-secreted proteins (ELSPBP1, and DEFB129) and some were known testicular markers (e.g. TKTL1). The significantly expressed DEPs were first searched in the RNAseq dataset

(Chapter 2 and Table 3.3) to gain an understanding of whether the mRNA was likely to be predominantly expressed in human germ cells or somatic cells. Only SPATA31E1 which was selected from the previous RNAseq dataset (Chapter 2), was detected among the significant DEP set obtained by MS. SPATA31E1 was low in abundance in AT compared to NORM and previously showed downregulated mRNA in spermatid arrest samples compared to normal spermatogenesis biopsies (Chapter 2). The UniProt protein database and published literature of all DEPs were investigated to collect information on functional role of the protein in spermiogenesis and/or sperm function. Information such as a) predicted location of protein and RNA, b) tissue enrichment, c) cell enrichment, d) organ localisation, e) tissue-specific RNA expression, f) cell-specific RNA expression, and g) IHC results in the testis (if applicable) were collected from the Human Protein Atlas (HPA). scRNAseq datasets provided by the Young Adult testis Atlas (YAA) were used to assess whether mRNA transcripts of selected DEPs are expressed in meiotic and/or post-meiotic germ cells, i.e spermatocytes or early/late spermatids (Table 3.4). Human tissue and testis microarray datasets available at GermOnline were used to investigate relative mRNA expression in testis cells (Chalmel et al., 2007) and various tissues compared to testis (Chalmel et al., 2012) (Table 3.5). Finally, information on DEPs was investigated in a relevant mouse dataset via GermOnline (Chalmel et al., 2007) and Mouse Genome Informatics (MGI) to collect information on mouse ortholog of the gene and available mouse models (Table 3.6). Each DEP was also searched in PubMed. PubMed searches were used to categorise proteins as either a) not previously published, b) important in mouse testis biology but not investigated in human testis biology, or c) have already confirmed roles in sperm production and function. Proteins with a confirmed or suggested contribution to human sperm morphology and motility and its defects were excluded since this chapter aimed to find novel proteins. Out of the list of 35 DEPs, 23 had previous publications that were related to male fertility in mammals, with 4 proteins (TKTL1, DEFB129, IRGC, ACTRT2) having been previously investigated using human samples. The other 12 proteins did not show any previous investigations related to male infertility (SLTM, IL1RL1, VPS35, USP47, CCDC105, ITIH2, A2M, TKTL2, APOA4, CAPS, RCN2, BPIFB2).

This approach enabled the identification of proteins with potential roles in human spermatozoa which could be further investigated for their role in human

spermatogenesis using IHC in testis biopsies. Data in HPA was used to determine the level of gene expression in testis compared to other organs in order to determine whether the protein was likely enriched in, or specific to, the testis, and RNAseq data was assessed to determine the level of mRNA expression in germ cells, particularly in spermiogenesis (round spermatids to late spermatids) (Pontén et al., 2009; Uhlén et al., 2015). Based on the analyses of all DEPs (Tables 3.3 - 3.6), 10 DEPs that showed high expression in testis and spermatids but were not well defined in human spermatogenesis related to sperm morphology and motility, were further investigated (Table 3.7). Of these proteins, n=6 showed higher levels in AT compared to NORM sperm and n=4 showed lower levels. All showed high abundance in human testis and in round/elongating spermatids except for the epididymis-specific protein in the list. Ultimately, 5 DEPs (see shaded proteins in Table 3.7) were chosen for further analysis based on their novelty and the availability of antibodies suitable for IHC in Bouin's fixed, paraffin-embedded human testis sections. These proteins were: 1) human testis- and germ cell-specific (ACTRT2) or, 2) enriched in testis and with potential roles in human sperm morphology and motility based on localisation and function (IPO4, IFT57), or 3) epididymis-specific (ELSPBP1), or 4) no information in the literature but enriched in the testis and human spermatids (CCDC105).



AT – Abnormal/ Asthenoteratozoospermic samples

N/ NORM – Normal/ control – Normozoospermic samples

Figure 3.4: Hierarchical clustering and heat map of significantly differentially expressed proteins in Mass spec, individual-samples (A) and grouped-samples (B). The absolute scaled abundances of each sample (A) and average abundances of the groups (B) were used to create heatmaps. The Euclidean hierarchical clustering was used. The scale indicates protein abundance in the samples, with red indicating high abundance and blue indicating lower abundance. The proteins selected for further analysis are marked with a red dot on the

right-hand side. AT indicates the abnormal/ asthenoteratozoospermic test samples while N/ NORM indicates the control/ normal / normozoospermic samples.

Table 3.3: Mass Spectrometry results showing the list of differentially expressed proteins (DEP) between normal (Norm) and asthenoteratozoospermic (AT) ejaculates

Mass Spec results							Comparison with RNAseq data (Chapter 2)		
No.	Protein Accession No.	Gene Symbol	Description	Increased or decreased in AT sperm compared to NORM sperm	Proteins Abundance ratio AT/NORM	Mass Spec Log2FC (AT/NORM)	Log2FC		
							SCO/NSP	SDA/NSP	SCO/SDA
1	V9HW76	PEBP4/ HEL-S-300	Epididymis secretory protein Li 300	Decreased	0.068	-3.88	Not DE	Not DE	Not DE
2	H7BXE3	SLTM	SAFB-like transcription modulator	Decreased	0.21	-2.25	Not DE	Not DE	Not DE
3	A0A024R462	FN1	Fibronectin	Decreased	0.274	-1.87	0.96	0.86	not DE
4	P23352	KAL1; ANOS1	Anosmin-1	Decreased	0.234	-2.09	Not DE	Not DE	1.26
5	Q96LM6	TEX37	Testis-expressed sequence 37 protein	Decreased	0.268	-1.9	-4.27	Not DE	-3.59
6	Q01638	IL1RL1	Interleukin-1 receptor-like 1	Decreased	0.375	-1.41	1.69	Not DE	Not DE
7	B7Z7I0	TKTL1	cDNA FLJ55544, highly similar to Homo sapiens transketolase-like 1 (TKTL1),	Increased	4.831	2.27	-3.67	Not DE	-3.35
8	Q96K76	USP47	Ubiquitin carboxyl-terminal hydrolase 47	Decreased	0.36	-1.47	Not DE	Not DE	Not DE
9	Q8IYK2	CCDC105	Coiled-coil domain-containing protein 105	Decreased	0.374	-1.42	-4.81	Not DE	-4.49
10	Q8WZ59	TMEM190	Transmembrane protein 190	Decreased	0.458	-1.13	-4.09	Not DE	Not DE
11	A0A024R1U8	IGFBP4	Insulin-like growth factor-binding protein 4	Decreased	0.419	-1.25	1.73	1.25	Not DE
12	B4DM79	ITIH2	cDNA FLJ53848, highly similar to Inter-alpha-trypsin inhibitor heavy chain H2	Decreased	0.419	-1.25	Not DE	Not DE	Not DE
13	Q9H1M3	DEFB129	Beta-defensin 129 OS=Homo sapiens	Increased	4.345	2.12	Not DE	Not DE	Not DE

14	B4E2J2	GSTM3	cDNA FLJ51755, highly similar to Glutathione S-transferase Mu 3	Increased	4.052	2.02	1.01	Not DE	0.74
15	P20142	PGC	Gastricsin	Decreased	0.488	-1.03	Not DE	Not DE	Not DE
16	P01023	A2M	Alpha-2-macroglobulin	Decreased	0.491	-1.03	1.05	0.57	0.55
17	K9JA26	ELSPBP1	Epididymis luminal secretory protein 149	Increased	3.641	1.86	Not DE	Not DE	Not DE
18	Q9H0I9	TKTL2	Transketolase-like protein 2	Increased	3.958	1.98	-4.20	Not DE	Not DE
19	O75610	LEFTY1	Left-right determination factor 1	Increased	4.128	2.05	Not DE	Not DE	Not DE
20	P06727	APOA4	Apolipoprotein A-IV	Decreased	0.508	-0.98	Not DE	Not DE	Not DE
21	Q8TDY3	ACTRT2	Actin-related protein T2	Decreased	0.513	-0.96	-4.00	-0.57	Not DE
22	A0A0S2Z517	GPR56; ADGRG1	ADGRG1 C-terminal fragment (Fragment)	Decreased	0.48	-1.06	1.33	0.54	0.85
23	Q9NWB7	IFT57	Intraflagellar transport protein 57	Increased	3.886	1.96	-1.10	Not DE	-0.79
24	E7ETU9	PLOD2	Procollagen-lysine 5-dioxygenase	Decreased	0.528	-0.92	Not DE	Not DE	Not DE
25	C8C504	HBB	Beta-globin	Decreased	0.531	-0.91	Not DE	Not DE	Not DE
26	A0A024R2B5	LIPG	Phospholipase A1	Decreased	0.534	-0.91	1.06	not DE	0.91
27	P02649	APOE	Apolipoprotein E	Decreased	0.535	-0.9	1.27	0.83	Not DE
28	Q13938	CAPS	Calcyphosin	Decreased	0.536	-0.9	Not DE	Not DE	Not DE
29	Q14257	RCN2	Reticulocalbin-2	Increased	3.639	1.86	-0.96	-0.45	Not DE
30	H0Y8G5	HNRNPD	Heterogeneous nuclear ribonucleoprotein D0 (Fragment)	Increased	3.791	1.92	0.65	Not DE	Not DE
31	Q8TEX9	IPO4	Importin 4	Increased	3.587	1.84	-2.05	Not DE	Not DE
32	A0A1W2PP10	VPS35	Vacuolar protein sorting-associated protein 35	Increased	5.201	2.38	0.49	Not DE	Not DE
33	Q6ZUB1	SPATA31E1	Spermatogenesis-associated protein 31E1	Decreased	0.452	-1.15	-4.19	-1.65	-2.46
34	Q6NXR0	IRGC	Interferon-inducible GTPase 5	Decreased	0.519	-0.95	-3.57	Not DE	-3.03
35	Q8N4F0	BPIFB2	Lipoprotein lipase	Increased	3.696	1.89	Not DE	Not DE	Not DE

Abbreviations: RNAseq –Mass spec – mass spectrometry, RNA sequencing, Log2FC – log 2-fold change, DE- differentially expressed, NSP – normal spermatogenesis, SDA – spermatid arrest, SCO – Sertoli cell only.

Table 3.4: Data collected from the Human Protein Atlas and Young Adult Testis Atlas

	Accession	Gene Symbol	The Human Young Adult Testis Atlas		The Human Protein Atlas			
			Availability in the dataset	Cluster/s	RNA expression in organs	RNA/ Protein enriched in	Protein expression in organs	IHC data in testis tissue
1	V9HW76	PEBP4	X	X	Highest -Skeletal muscle. 2nd - Epididymis nTPM - 590.5. testis - low - nTPM - 15.5	RNA- Tissue enhanced (epididymis, skeletal muscle, tongue)	No data	High – rSptd, Pach Low - Leyd
2	H7BXE3	SLTM	√	SSCs, L/p Spc	Low organ specificity Testis – nTPM - 27 epididymis - 25.2 seminal vesicles - 45.6	Expressed in many organs	Highly expressed in many	High - Pach, Spc- rSptd-Leyd > Spg - pLep Neg - Elo.Sptd, Sert (HPA040256). High – Pach Spc- rSptd, Neg. Elo.Sptd, Sert, Pert, Spg, pLep Low - Leyd (HPA040381)
3	A0A024R462	FN1	√	Endothelial	Highest - placenta. Testis low. Avg. nTPM - 64.8	Tissue enhanced (ductus deferens, placenta). Medium expression in late spermatids and Leydig cells	Highest - Kidney. Medium - Testis. Medium expression in 10 others	Medium stained - Elo Sptd and Leyd (HPA027066)
4	P23352	KAL1; ANOS1	√	Sertoli	Highest - lung. Testis - nTPM - 9.5 Epididymis - 4.7, seminiferous - 7.3	No data	No data	No data
5	Q96LM6	TEX37	√	E/p.Spc, L/p. Spc., rSptd, Elo.Sptd	High, only in testis - nTPM - 106.9	Protein and RNA - Expressed only in testis	High expression in testis	High - E. Sptd-Spg> pLep. Low - rSptd-Pach

6	Q01638	IL1RL1	X	X	Highest - placenta. Testis - low nTPM - 1	Low protein and RNA expression in Testis	No data	Medium in cells in seminiferous duct (HPA007406)
7	B7Z7I0	TKTL1	√	Spg, eSpc, rSptd, sperm	Highest - Testis nTPM - 292.3	Group enriched (blood, testis). Single cell atlas - highest in Spermatogonia	High and only in Testis	High - Spg-pLep > Elo.Sptd. Medium -rSptd-Pach (HPA000505). CAB032191 - strong staining in cells in SD
8	Q96K76	USP47	√	E/p.Spc	Highest - skeletal muscle. Testis - Avg. nTPM - 53.9	Low cell and tissue type specificity	High in testis with 23 others. Medium in epididymis	High - Leyd. Medium - Sertoli. Low - Elo.Sptd
9	Q8IYK2	CCDC105	√	rSptd	Expressed only in Testis. Avg. nTPM - 15.2	Tissue enriched - Testis. Cell - Early spermatids	Expressed in a few. Medium expression in testis with others	Medium - Pach, rSptd
10	Q8WZ59	TMEM190	√	Spg, e/p Spc, L/p Spc, rSptd, Elo.Sptd,	Highest - Fallopian tubes. Testis - low avg. nTPM - 34.6	Tissue enhanced (fallopian tube, ovary, testis). Cell - Early spermatids, Ciliated cells, Late spermatids	No data	No data
11	A0A024R1U8	IGFBP4	√	Spg, e/p Spc, L/p Spc, rSptd, Elo.Sptd, sperm, Sertoli, Leyd	Highest - ovary. Testis - low -Avg. nTPM - 131.1	Tissue enriched - ovary. Cell - Ito cells, hepatocytes	Not expressed in Testis. Low expression in 6 others	Not detected in cells in seminiferous duct and Leydig cells
12	B4DM79	ITIH2	X	X	Only in liver	Tissue enriched (liver). Cell - hepatocytes	No data	Not detected in cells in seminiferous duct and Leydig cells
13	Q9H1M3	DEFB129	X	X	Only in epididymis. nTPM - 2972.7	Tissue enriched (epididymis)	Only in epididymis	Not detected in testis. High expression in glandular cells in epididymis
14	B4E2J2	GSTM3	√	SSC, Spg, e/p Spc, rSptd, sperm	Highest in seminal vesicle then testis (nTPM - 934.4)	Tissue enhanced (seminal vesicle, testis). Cell type enhanced (Late spermatids, Early	Medium expression in testis, epididymis, seminal vesicles and Kidney	Medium expressed in Elo.Sptd, Leyd (HPA035190)

						spermatids, Spermatocytes)		
15	P20142	PGC	X	X	Highest in stomach. Not expressed in testis. Seminal vesicle - nTPM 749.8	Group enriched (ductus deferens, pancreas, stomach). Cell type enhanced (Pancreatic endocrine cells, Alveolar cells type 2, Cholangiocytes)	Highest in stomach, not expressed in testis	Medium in cells in SD (HPA031718)
16	P01023	A2M	√	Spg, eSpc, LSpc, rSptd, Sperm, Sert	Highest in lung and liver. Testis low - nTPM - 141.8, epididymis 354.6	Tissue enhanced (liver, lung). Cell type enhanced (Horizontal cells, Hepatocytes, Ito cells, Endothelial cells, Sertoli cells)	Medium expression in 5. Testis - low -	Low in Leyd. Not detected in cells in SD HPA002265
17	K9JA26	ELSPBP1	X	X	Highest in epididymis (nTPM - 4567.1. Low in Testis - nTPM - 2.3	Tissue enriched (epididymis). Cell type enhanced (Rod photoreceptor cells, Early spermatids, Horizontal cells, Muller glia cells, Cone photoreceptor cells, Late spermatids)	Only in Epididymis	Medium - Leyd, Low - cells in SD (HPA044256)
18	Q9H0I9	TKTL2	√	Lp Spc, rSptd	Highly expressed in testis (nTPM - 134.7	Cancer pro. Tissue enriched – testis. Group enriched (Spermatocytes, Early spermatids, Late spermatids	No data on protein level	No data
19	O75610	LEFTY1	X	X	Highest - pancreas (nTPM - 25.1. Testis -0.7	Group enriched (blood, epididymis, intestine, pancreas)	No data	Medium - Leyd. Not detected in Cells in SD (HPA047883)
20	P06727	APOA4	X	X	Highest in Small intestine. Not expressed in testis, Epididymis	Tissue enriched (intestine). Cell type enriched (Hepatocytes)	Highest - Duodenum, small intestine. Not expressed in testis, Epi	Low - Leyd (CAB068250>CAB068252 -HPA002549>CAB068251
21	Q8TDY3	ACTRT2	√	SSCs, Spg, eSpc,	Expressed only in testis nTPM - 133.3	Tissue enriched - testis. Group enriched (Late	Highest - testis.	High -Elo.Sptd-rSptd > Spg-pLep-Pach> Sertoli

				LSpc, Elo.Sptd, Sperm		spermatids, Early spermatids)	Low - stomach, small intestine, placenta	
22	A0A0S2Z517	GPR56; ADGRG1	√	Sertoli	Highest - Thyroid and kidney (nTPM - 134.5). Testis 108.9	Low tissue specificity. Cell type enhanced (Cytotrophoblasts, Syncytiotrophoblasts, Exocrine glandular cells, Proximal tubular cells)	High in 6. medium testis and 23 others	Medium - Leyd, Cells in SD (HPA046065)
23	Q9NWB7	IFT57	√	SSCs, Spg, eSpc, LSpc, rSptds, Elo. Sptd, sperm	Expressed in all. Highest in Testis (nTPM - 87)	Low tissue specificity. Cell type enhanced (Late spermatids, Early spermatids, Ciliated cells)	Expressed in all organs. High expression in testis along with 17 other organs	Two antibodies detect differently in the same patient
24	E7ETU9	PLOD2	X	X	Highest in liver. Testis - nTPM 20.9 Epididymis - nTPM-20.4,	Tissue enhanced (liver). Cell type enhanced (Extravillous trophoblasts, Pancreatic endocrine cells)	Medium in Testis and 23 others (incl. Epididymis)	Medium - Leyd>Sert. Low - Elo.Sptd
25	C8C504	HBB	X	X	Highest bone marrow. Not expressed in testis/epididymis	Tissue enriched (bone marrow). Cell type enriched (Erythroid cells)	Highest - placenta. Not expressed in Testis/ epididymis	Not detected in any (<25% Leyd with CAB009526)
26	A0A024R2B5	LIPG	X	X	Highest - thyroid gland, Testis - nTPM - 17.1. Epididymis - nTPM- 4.4	Tissue enhanced (liver, placenta, thyroid gland). Cell type enhanced (Hepatocytes, Extra villous trophoblasts, Cone photoreceptor cells)	No data	Not detected. Weak <25% with HPA016966
27	P02649	APOE	√	Spg, eSpc, L. Spc, Elo.Sptd, Sperm, Sert, Leyd	Highest - liver. Testis - low. nTPM - 174.1. Epididymis - nTPM - 30.5	Tissue enhanced (adrenal gland, liver). Cell type enhanced (Hepatocytes, Muller glia cells, Hofbauer cells, Proximal tubular cells, Peritubular cells)	High in cerebral cortex and 4 others. Low expression in Testis	No detection in cells in seminiferous duct. Low - high detection in Leydig cells

28	Q13938	CAPS	X	X	Highest - Fallopian tube. Testis - nTPM -5.1, Epididymis - nTPM - 6.9	Tissue enhanced (brain, fallopian tube). Cell type enriched (Ciliated cells)	Highest - fallopian tube, nasopharynx and bronchus. Not expressed in testis	Not detected in any (HPA043520)
29	Q14257	RCN2	√	e.Spc, L.Spc, Elo.Sptd, Sperm	Highest - epididymis - nTPM - 491.3. Testis - nTPM - 218.5	Tissue enhanced (epididymis). Cell type enhanced (Spermatocytes	High in testis, epididymis and 4 others	High - Pach, rSptd> Elo.Sptd (HPA030694). High - Pach, rSptd>Elo.Sptd. > pLep-Spg (HPA030695)
30	H0Y8G5	HNRNPD	√	Spg, eSpc, Elo.Sptd, Sperm	Highest - Thymus. Testis- nTPM - 235.1. Epididymis - nTPM - 209.5	Low tissue and cell specificity	High in testis, Epididymis and 31 others	High- Spg-pLep-Pach>rSptd (HPA004911)
31	Q8TEX9	IPO4	X	X	Highest - testis - nTPM - 11.3. Epididymis - nTPM - 0.4	Tissue enhanced (testis). Cell type enhanced (Late spermatids, Early spermatids, Spermatocytes)	No data	High- cells in SD, Leyd (HPA039043)
32	A0A1W2PP10	VPS35	√	e/p Spc, L/p Spc, Sperm	Expressed in all. Low in testis	Not mentioned	Low tissue specificity for protein (detected in testis)	High - Spg, r. Sptd, early Sptd, Sert. Others - not detected
33	Q6ZUB1	SPATA31E1	X	X	Testis nTPM - 7.5 , brain	Tissue enriched – testis May play a role in spermatogenesis	Only in testis	High - Pach, r. Sptd, E. Sptd, L. Sptd – Others - not detected
34	Q6NXR0	IRGC	√	Spg, e/p Spc L/p Spc Elo. Sptd, Sperm, Macrophages, Endothelia, Myoid, Leydig	Only in testis – nTPM – 145.5	Tissue enriched. Testis - Spermatogenesis (mainly)	Only in testis	Medium - cells in SD. Leydig - not detected
35	Q8N4F0	BPIFB2	X	X	Only in salivary glands	Tissue enriched - salivary glands	Only in salivary glands	Not detected

√ Present, X – not found. Abbreviations: SSC = spermatogonial stem cells, Spg = spermatogonia e/p Spc = Early primary spermatocytes, L/p Spc = late primary spermatocytes, e Spc = early spermatocytes, L Spc = late spermatocytes, pLep = pre leptotene spermatocytes, Pach = pachytene spermatocytes, rSptd = round spermatids, E Sptd = early spermatids, Elo.Sptd = elongating spermatids, L. Sptd = late spermatids. Sert = Sertoli cells, Leyd = Leydig cells. SD – seminiferous duct, nTPM = normalised transcripts per million. Tissue RNA expression (Uhlén et al., 2015). Tissue protein expression – (Pontén et al., 2009).

Table 3.5: Information on human testicular mRNA expression of DEPs collected from GermOnline

	Accession	Gene Symbol	GermOnline	
			The conserved transcriptome in human male gametogenesis	
			Microarray expression data - Human testis cells (Chalmel et al., 2007)	Microarray expression data – human tissue profiling (Chalmel et al., 2012)
1	V9HW76	PEBP4	spermatocytes (highest, normalized intensity 94)> spermatids (63.8)>tubules>total testis	Highest - skeletal muscle - intensity - 1058, testis - 60.64
2	H7BXE3	SLTM	Spc - 143.9> total 114> tubules -111>Sptd - 54.8	Highest - Thymus - 684, Testis – 227
3	A0A024R462	FN1	Tub (647) >Total> Spc > Sptd (39.87)	Highest - Synovial membrane (13214). Testis – 443
4	P23352	KAL1; ANOS1	Total (93.6) >Tub >Spc >Sptd (8.2)	Highest - JS2(700-850)>Lung>. Testis - 139.7
5	Q96LM6	TEX37	Sptd (1175) > Tub >Spc	Highest -testis -48.3>JS10
6	Q01638	IL1RL1	Sptd (26.9) > Spc >Tubules	Highest - Vulva - 102. Testis – 55
7	B7Z7I0	TKTL1	Spc (2172) >Tub > Sptd (390)	Highest -JS5(7000-9000). Testis – 3134
8	Q96K76	USP47	Tub (197) > Spc > Sptd (107)	Highest –Muscle (1061). Testis (459)
9	Q8IYK2	CCDC105	Spc (222) >Tub> Sptd (117)	Highest - Testis (223)
10	Q8WZ59	TMEM190	Sptd (277) > Spc > Tub	Highest - Fallopian (626) Testis (171)
11	A0A024R1U8	IGFBP4	Tub (617) > Spc > Sptd (44)	Highest - Liver (3422). Testis -382
12	B4DM79	ITIH2	Sptd (20) > Tub > Spc	Highest - liver (4188). No expression in others
13	Q9H1M3	DEFB129	Spc – Sptd (3.9)- Tub-low in all	Highest-testis-33.8. low expression in all
14	B4E2J2	GSTM3	Spc (3154) > Tub > Sptd (2671)	Highest-JS10(7500). Testis (7103)
15	P20142	PGC	Sptd (23) > Spc >Tub	Highest – Stomach (14569). No expression in others
16	P01023	A2M	Tub (800) > Sptd (200) >Spc	Highest – Lung (8862). Testis (662). JS1-2000
17	K9JA26	ELSPBP1	Sptd (15) >tub-Spc	Highest- testis (76)
18	Q9H0I9	TKTL2	no data	no data
19	O75610	LEFTY1	The profile given for gene symbol LEFTY1 is of a different ENSGO number	
20	P06727	APOA4	Sptd (45) > Spc-Tub	Highest-small intenstine-307. Testis-59. JS1-JS10-similar values
21	Q8TDY3	ACTRT2	Sptd (2907) > Tub > Spc	Highest-Testis-768>JS10
22	A0A0S2Z517	GPR56; ADGRG1	Tub (3030) > Sptd (205) > Spc	Highest-Kidney (1548). Testis - 928
23	Q9NWB7	IFT57	Sptd (1103) > Tub > Spc	Highest-JS10(2200). Testis-1445
24	E7ETU9	PLOD2	Tub (40) > Spc > Sptd (5)	Highest- fetal cartilage (1387). Testis-89
25	C8C504	HBB	Does not appear in GermOnline	
26	A0A024R2B5	LIPG	Tub (52) > Spc > Sptd (10)	Highest - Huve cells (856). Testis – 99

27	P02649	APOE	Tub (595) > Sptd (279) > Spc	Highest-Liver-7676.Testis - 450.JS1-1777
28	Q13938	CAPS	Sptd (46) > Spc >Tub	Highest-Fallopian tube (213). Testis - 68. JS1-JS10-60-80
29	Q14257	RCN2	Spc (959) >Tub> Sptd (288)	Highest-JS10(1290). Testis-1004
30	H0Y8G5	HNRNPD	Tub (480)> Spc> Sptd (165)	Highest-Thymus (1945). Testis-1312.JS1-JS10-1200-2500
31	Q8TEX9	IPO4	Sptd (561) > Spc > Tub	Highest-testis-588
32	A0A1W2PP10	VPS35	Tub > Spc > Sptd	Many
33	Q6ZUB1	SPATA31E1	Does not appear in GermOnline	
34	Q6NXR0	IRGC	Sptd > Spc - Tub > total testis	Highest – testis
35	Q8N4F0	BPIFB2	Does not appear in GermOnline	

Data in brackets are normalised intensity values as reported under microarray data in GermOnline (Chalmel et al., 2012; Chalmel et al., 2007). Abbreviations: Spg = spermatogonia, Spc = spermatocytes, Sptd = spermatids, Sert = Sertoli cells, Tub = tubules

Table 3.6: Information on mouse ortholog mRNA expression in testis collected from GermOnline

	Accession	**Gene Symbol**	**GermOnline**	
			The conserved transcriptome in human and rodent male gametogenesis – Mouse ortholog data	
			Microarray expression data - mouse testis cells (Chalmel et al., 2007)	**Microarray expression data – mouse tissue profiling (Chalmel et al., 2007)**
1	V9HW76	PEBP4	Spc (NI 5700 >Tub (4600) > Sptd (4500)	Testis - 1947. The ovary shows a very slight expression. Others - no
2	H7BXE3	SLTM	Spc 2095.6>total>Sptd (1653)>Tub	Highest - CD4+naïve cell>Testis
3	A0A024R462	FN1	Sertoli > Spg > Tub > total > Sptd (378) >Spc	Highest - Aorta -. Testis – 557
4	P23352	KAL1; ANOS1	No mouse ortholog	
5	Q96LM6	TEX37	Tub (1609) > Sptd (1147)>Spc>Spg>Sert	Highest - Testis (834)
6	Q01638	IL1RL1	Spg (124) > Sert (114) > Sptd (91) > Tub > Spc	Highest - 592. Testis – 203
7	B7Z7I0	TKTL1	Spg (800) >Tub>Sptd (146)>Spc	Highest - Testis -135
8	Q96K76	USP47	Spc (2288) >Tub>Sert>Sptd (1190)>Spg>	Highest - muscle retina (1500). Testis (1151)
9	Q8IYK2	CCDC105	Sptd(2132)>Tub>Spc>Spg-Sert (24)	Highest - Testis (278)
10	Q8WZ59	TMEM190	Sptd(3300)>Tub>Spc>Sert>Spg	Highest – Testis (1291)
11	A0A024R1U8	IGFBP4	Spg(3667)>Sert>Sptd(298)>Tub>Spc	Highest-Liver 9967. Testis - 1010
12	B4DM79	ITIH2	Tub(143)-Spg>Sptd(85)-Spc>Sert	Highest- Liver (4807). Testis - 165
13	Q9H1M3	DEFB129	No data	No data
14	B4E2J2	GSTM3	No mouse ortholog	
15	P20142	PGC	Sptd(205)>Tub-Spc>Sert>Spg	Highest - stomach (10227)

16	P01023	A2M	Sert(176)>Sptd(143)>Tub>Spc-Spg	Highest-Embryo Day 7
17	K9JA26	ELSPBP1	No ortholog prediction only genomic alignment with mouse	
18	Q9H0I9	TKTL2	Spc(755)>Tub>Sptd(671)>Spg>Sert	Highest – Testis (170)
19	O75610	LEFTY1	Not available	
20	P06727	APOA4	Sptd (144) > Tub (130)-Spg-Sert-Spc	Highest-Testis-10507
21	Q8TDY3	ACTRT2	Sptd (5024) >Tub>Spc>Spg>Sert	Highest-Testis-2867.very low expression in others
22	A0A0S2Z517	GPR56; ADGRG1	Sert (2220) > Spg > Tub> Sptd (220) >Spc	Highest-Kidney (1353). Testis-243
23	Q9NWB7	IFT57	Sptd (1300)>Spc>tub>Spg>Sert	Highest-Testis-667
24	E7ETU9	PLOD2	Sert(4700)>Spg>Tub>Spc> Sptd (48)	Highest - ovary (2735). Testis-102
25	C8C504	HBB	Not available	
26	A0A024R2B5	LIPG	Sert(3000)>Spc>tub>Spg> Sptd (14)	Highest-Embryo Day 7 (2909) > Ovary. Testis-87
27	P02649	APOE	Sert(1945)>Spg>Tub> Sptd (427)>Spc	Highest-Liver (17213). Testis-1400
28	Q13938	CAPS	Not available	
29	Q14257	RCN2	Spg(2200)>Sert>Spc>tub> Sptd (390)	Highest- Brain (2428). Testis-893
30	H0Y8G5	HNRNPD	Spg(2780)>Sert>Spc>tub>Sptd (645)	Highest-CD4+naïve cells (2185). Testis-684
31	Q8TEX9	IPO4	Sptd(3750)>tub>Spc>Spg>Sert (1400)	Highest-testis-3191
32	A0A1W2PP10	VPS35	Sert (3000)> Spc (2000)> Spg (1800)> Sptd (1300)	Highest – brain
33	Q6ZUB1	SPATA31E1	No data in GermOnline	
34	Q6NXR0	IRGC	Sptd (7700) >Tub (6500) >Spc (1650)	Highest - Testis (2048)
35	Q8N4F0	BPIFB2	Not available	

Data in brackets are normalised intensity values as reported under microarray data in GermOnline (Chalmel et al., 2007). Abbreviations: Spg = spermatogonia, Spc = spermatocytes, Sptd = spermatids, Sert = Sertoli cells, Tub = tubules

Table 3.7: Short-listed differentially expressed proteins

Protein	Findings from PubMed and other comments	Protein increased or decreased in AT compared to NORM sperm	RNAseq data (HPA)	Overall summary and category
CCDC105	• No publications in PubMed at the time of selection.	Decreased	E. Sptd > L. Sptd > Spc	Novel
TMEM190	• 3 papers. KO mice are fertile. Plays a role along with IZUMO1 as an acrosomal protein in mouse cauda epididymal sperm (Nishimura et al., 2011) • Not investigated with human samples.	Decreased	Respiratory Ciliated cells> E. Sptd > Endometrial ciliated cells >	No defined role in human spermatogenesis
IFT57	• 54 papers. Many are related to fertility but not defined in human spermatogenesis. • *Ift172* - Shows less developed sperm with reduced motility in homozygous mutant mice (Zhang et al., 2020) • IFT57 reduced in *Ift81*, *Ift74* and *Ift72* homozygous mutant mice (Qu et al., 2020; Shi et al., 2019; Zhang et al., 2020)	Increased	L.Sptd>E.Sptd>Spc	No publications in human spermatogenesis. Novel for human fertility
TEX37	• mRNA in elongated spermatids • Total 4 papers in PubMed. Also known as TCS21. 1st reported as a testis-specific novel gene in 2007. • KO mice were fertile (Khan et al., 2018b) • Sperm-specific in humans and has recently been shown to be decreased in infertile human spermatozoa and identified as a potential new marker (Greither et al., 2020).	Decreased	L.Sptd > E. Sptd	Testis specific (protein)
ACTRT2	• 4 papers. 2 on male infertility. • Reduced in obesity-induced asthenozoospermic men (Liu et al., 2015). • Localised in the post-acrosomal area and the mid-piece of human spermatozoa (Liu et al., 2015) • Heat-induced reversible reduction in sperm concentration and motility and have the potential to be the biomarkers and clinical targets for scrotal heat treatment induced male infertility (Wu et al., 2020)	Decreased	L. Sptd > E. Sptd	Testis specific (mRNA) Highly enriched in testis (protein)
TKTL2	• 10 papers. Mostly cancer related.	Increased	Spc > E. Sptd > L. Sptd	Testis-specific (mRNA)

	• None on male infertility.			
IPO4	• 12 papers. Mostly cancer related. • None on human male fertility but other importins are reported to be involved with nuclear-cytoplasmic transportation in germ cells (Bernardes et al., 2022; Loveland et al., 2015; Miyamoto et al., 2020; Nathaniel et al., 2022; Whiley et al., 2012; Young et al., 2013)	Increased	L. Sptd > E. Sptd > Spc > Spg	Testis-enhanced. No publications related to human sperm morphology and motility.
RCN2	• 46 papers. • None on infertility. • Gene is related to calcineurin	Increased	Low specificity among other cells Spc > Spg > L. Sptd > E. Sptd	High in Testis and Epididymis (Protein) mRNA – highest in epididymis No papers on testis function
ELSPBP1	• 13 papers. • Mostly related to fertility of other animals	Increased	Rod photoreceptor cells > L. Sptd > Bipolar cells > E. Sptd	Epididymis specific (mRNA and protein)
GSTM3	• Many publications, mainly related to cancer. 3 mention male infertility. • A study in pigs shows the importance in epididymal maturation and sperm morphology (Llavanera et al., 2020). • Shows a link with non-obstructive azoospermia (Cui et al., 2018)	Increased	L. Sptd > E. Sptd > Spc	Proteins in testis and epididymis

Abbreviations HPA – The Human Protein Atlas. RNAseq – RNA sequencing. Spc = spermatocytes, Spg = spermatogonia, E. Sptd = early spermatids, L. Sptd = late spermatids. Orange shaded box –proteins selected for further analysis.

RNAseq data – (Guo et al., 2018; Norreen-Thorsen et al., 2022; Uhlén et al., 2015)

Importin 4 (IPO4) levels were 3.5-fold higher in human sperm in the AT group compared to NORM (Table 3.3). IPO4 belongs to a protein group involved in the RanGTPase cycle and has functions in nuclear protein import (further discussed in Chapter 4). According to HPA, the predicted location of the protein is intracellular and present in cytoplasm and nucleus (Table 3.5). *IPO4* is testis-enriched and is present in spermatocytes and early and late spermatids (highest) according to HPA's RNAseq data and GermOnline's microarray data (Tables 3.4 and 3.5). *IPO4* shows highest expression in the testis compared to all other tissues according to HPA (Table 3.4). The gene is located on chromosome 14. A mouse ortholog gene on chromosome 14 was reported in MGI - MGI:1923001 (Fig. 3.5C). The functional role of IPO4 in testis biology is unknown, but its function in cancer biology and nuclear transportation is reported (Table 3.7). This made IPO4 a promising candidate for further study.

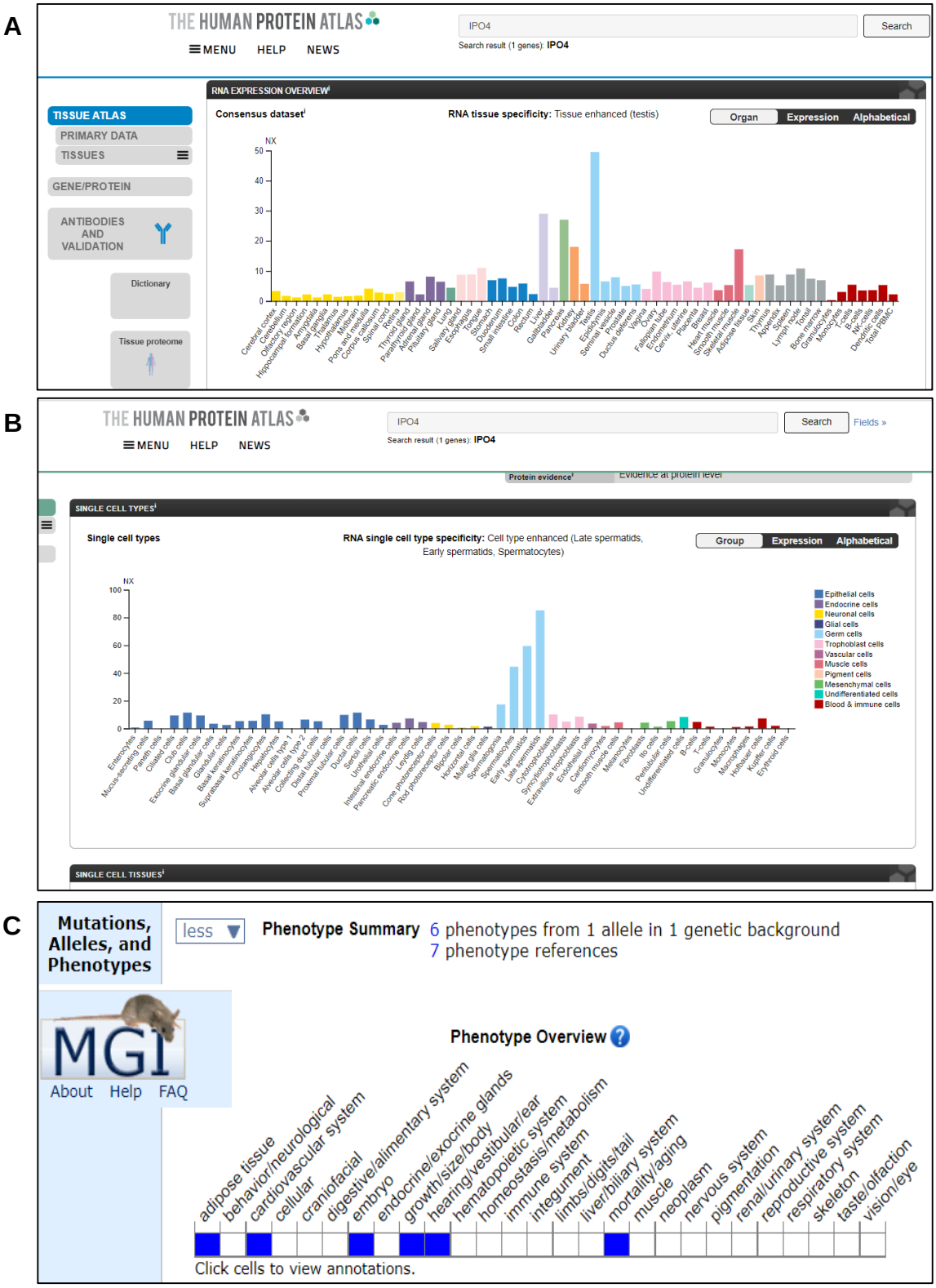


Figure 3.5: *In silico* analysis data of IPO4. A –RNA tissue overview in the HPA. B- RNA single cell overview in the HPA. C – the phenotype summary of the mouse ortholog in MGI. HPA – Human Protein Atlas, MGI – Mouse Genome Informatics.

Actin-related protein T2 (ACTRT2) levels were two-fold lower in ejaculated AT sperm compared to NORM. According to UniProt, the protein is localised in the post-acrosomal region and the sperm flagellum mid-piece (Table 3.7). It is predicted to be involved in cytoskeletal organisation and its predicted localisation is intracellular. According to HPA, the protein is classified as testis-enriched, and its mRNA was testis-specific (Fig 3.6, Table 3.4). GermOnline microarray data show that the RNA is enriched in early and late spermatids (highest) (Table 3.5). In a scRNAseq dataset from Young Human Testis Atlas (Human Testis Atlas Browser by Cairns Lab @ Utah), *ACTRT2* was detected in spermatogonial stem cells (SSCs), spermatogonia, early and late spermatocytes, elongated spermatids, and sperm (Table 3.4). *ACTRT2* is located on chromosome 1 in men with the mouse ortholog located on chromosome 4 - MGI:1920603 (Fig. 3.6). ACTRT2 was selected due to its testis-specificity, the availability of antibodies and its enrichment in early and late spermatids suggesting it is expressed during human spermiogenesis and could play a functional role in sperm.

Epididymal sperm-binding protein 1 precursor (ELSPBP1) levels were four-fold higher in AT sperm compared to NORM. *ELSPBP1* is located on chromosome 19. UniProt annotations predict that this protein binds to spermatozoa upon ejaculation and predicts it could function in capacitation however these predictions are based on similarity to other proteins, rather than published information. According to HPA, the mRNA is expressed at highest levels in epididymis and the protein is detected only in the epididymis (Fig. 3.7, Table 3.4 and 3.7). There was no mouse ortholog reported in MGI database and GermOnline (Table 3.6). ELSPBP1 was of interest because it could be a human sperm protein that originates from the epididymis instead of the testis that could have a role in post-testicular modification of sperm. ELSPBP1 has been described in bovine sperm (Table 3.7), however, at the time of its selection, there were no publications regarding its importance in human sperm.

Figure 3.6: *In silico* analysis data of ACTRT2. A – protein overview in the HPA. B- RNA tissue overview in the HPA. C – RNA single cell overview in the HPA. HPA – Human Protein Atlas.

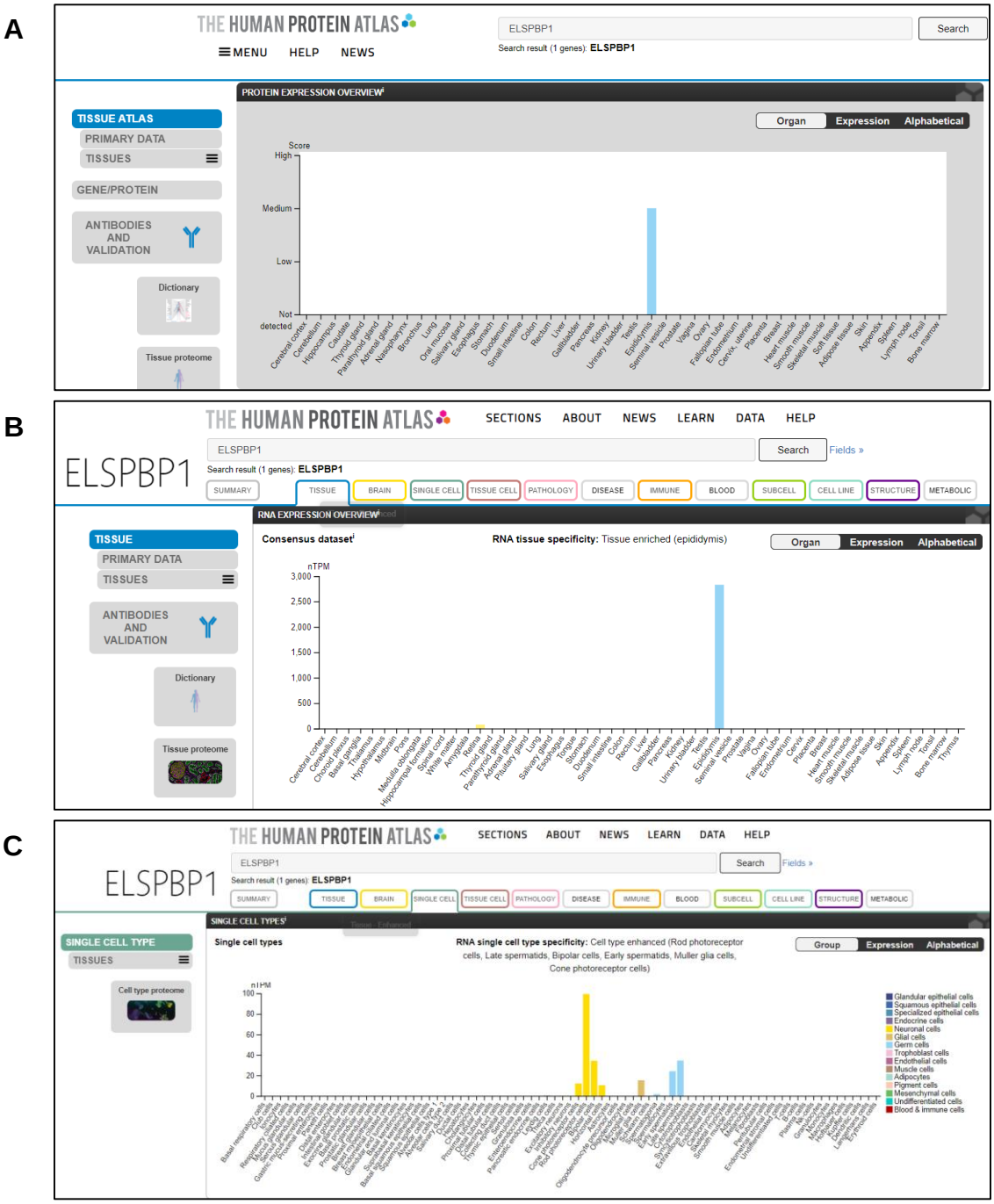


Figure 3.7: *In silico* analysis data of ELSPBP1. A – protein overview in the HPA. B- RNA tissue overview in the HPA. C – RNA single cell type overview in the HPA. HPA – Human Protein Atlas.

Coiled-coil domain-containing protein 105 (CCDC105) was two-fold lower in AT sperm compared to NORM. CCDC105 is a novel protein with no previous publications found in PubMed at the time of its selection. According to UniProt, this is an uncharacterised protein coding gene, and the function of the protein is not clear. According to HPA, the predicted location of the protein is intracellular (Table 3.4). *CCDC105* expression is testis-specific with the highest expression in human early and late spermatids (Fig 3.8, Tables 3.4, 3.5, 3.7). Human *CCDC105* is located on chromosome 19 and the mouse ortholog is found on chromosome 10 - MGI:1918226 (Fig. 3.8). CCDC105 was chosen for further study because it is spermatid-specific and there were no reports of the protein in the literature at the time of selection. However, a very recent study has reported that CCDC105 is detected in bull sperm (Leung et al., 2023).

Intraflagellar transporter protein 57 (IFT57) levels were almost four-fold higher in AT sperm compared to NORM. This protein is required for the formation of cilia and flagella and is predicted to play an indirect role in sonic hedgehog signalling, with a pro-apoptotic function. The predicted location of the protein is intracellular with a subcellular localisation to nuclear speckles and mitochondria. *IFT57*/IFT57 are expressed in many organs according to HPA with the highest RNA expression in testis and epididymis (Fig 3.9, Tables 3.4, 3.7). In human testis cells, it is localised in spermatocytes, early and late spermatids; with highest expression in spermatids (Fig 3.9, Tables 3.4, 3.5, 3.7). The gene is described by YAA in SSCs clusters, spermatogonia, early and late spermatocytes, round and elongated spermatids, and sperm (Table 3.4). The location of the human gene is on chromosome 3 and the mouse ortholog gene is found on chromosome 16 - MGI:1921166 (Fig. 3.9). This protein was chosen for further examination as it has not been described in human sperm. Yet given its role in the formation of cilia and published information of other intraflagella proteins on mouse sperm, it is likely to play a role in the formation and/or function of the human sperm flagellum.

Figure 3.8: *In silico* analysis data of CCDC105. A – protein overview in the HPA. B- RNA tissue overview in the HPA. C – RNA single cell type overview in the HPA. HPA – Human Protein Atlas.

Figure 3.9 – *In silico* analysis data of IFT57. A – protein overview in the HPA. B- RNA tissue overview in the HPA. C – RNA single cell type overview in the HPA. HPA – Human Protein Atlas.

3.3.2 Immunohistochemical localisation of selected proteins in human testes biopsies

3.3.2.1 Localisation of IPO4 in human testes

In order to understand the cellular localisation of IPO4 during human spermatogenesis, immunostaining was performed in NSP testis biopsies. Overall, the staining was more concentrated in the nucleus and cytoplasm of germ cells towards the lumen (Fig. 3.10 A, B). Spermatogonia (Sg) and early spermatocytes (eSp) exhibited lightly stained nuclei but little to no cytoplasmic staining (Fig. 3.10 A, red arrow Sg showing only nuclear staining and purple arrow eSp). Nuclear IPO4 immunostaining was more intense from pachytene spermatocytes onwards (Fig. 3.10 A, red arrow PS). From approximately stage III of spermatogenesis onwards, pachytene spermatocytes exhibited both nuclear and cytoplasmic staining (Fig. 3.10 B, red arrow PS), and this localisation continued in round spermatids (Fig. 3.10 A, red arrow rSptd). In some round spermatids, the cytoplasmic staining was concentrated to a vesicle on one side of the nucleus, likely consistent with staining in the Golgi apparatus (Fig. 3.10 A inset; red arrow rSptd). In later steps of spermiogenesis, immunostaining was observed in what appears to be the newly formed acrosome (Fig. 3.10 A inset, blue arrow rSptd). The majority of Sertoli cell nuclei were unstained (Fig. 3.10 B, black arrows SC) but some were immuno-positive, suggesting potential stage specific Sertoli cell staining (Fig. 3.10 B, red arrow SC).

In SDA biopsies, the cellular localisation was broadly similar with nuclear localisation in some spermatogonia (Fig. 3.10 C, red arrow Sg), but not all (Fig. 3.10 C, black arrow Sg). Early spermatocytes were immuno-negative (Fig. 3.10 D, eSp). Early pachytene spermatocytes were immuno-negative (Fig. 3.10 D, black arrow PS) while late pachytene spermatocytes showed light nuclear staining with unstained cytoplasm (Fig. 3.10 C inset, red arrow PS). Round spermatids showed lightly stained nuclei and unstained cytoplasm (Fig. 3.10 C, red arrow rSptd). Some Sertoli cell nuclei were immuno-positive (Fig. 3.10 D, red arrow SC) and others were not (Fig. 3.10 D, black arrow SC) suggesting stage-specific staining as observed in NSP biopsies.

SZA samples were overall immuno-negative (Fig. 3.10 E, F). In SCO biopsies, patient heterogeneity was observed. In some samples and tubules, lightly stained Sertoli cell

nuclei and cytoplasm were observed (Fig. 3.10 G, H, red arrows indicate immunostaining in some Sertoli cells SC and blue arrow indicate cytoplasmic staining, black arrows indicate no immunostaining in other Sertoli cells SC), suggesting possible stage-specificity of localisation in Sertoli cells. All negative control testis tissues in each slide and placenta tissue used as a negative control were unstained.

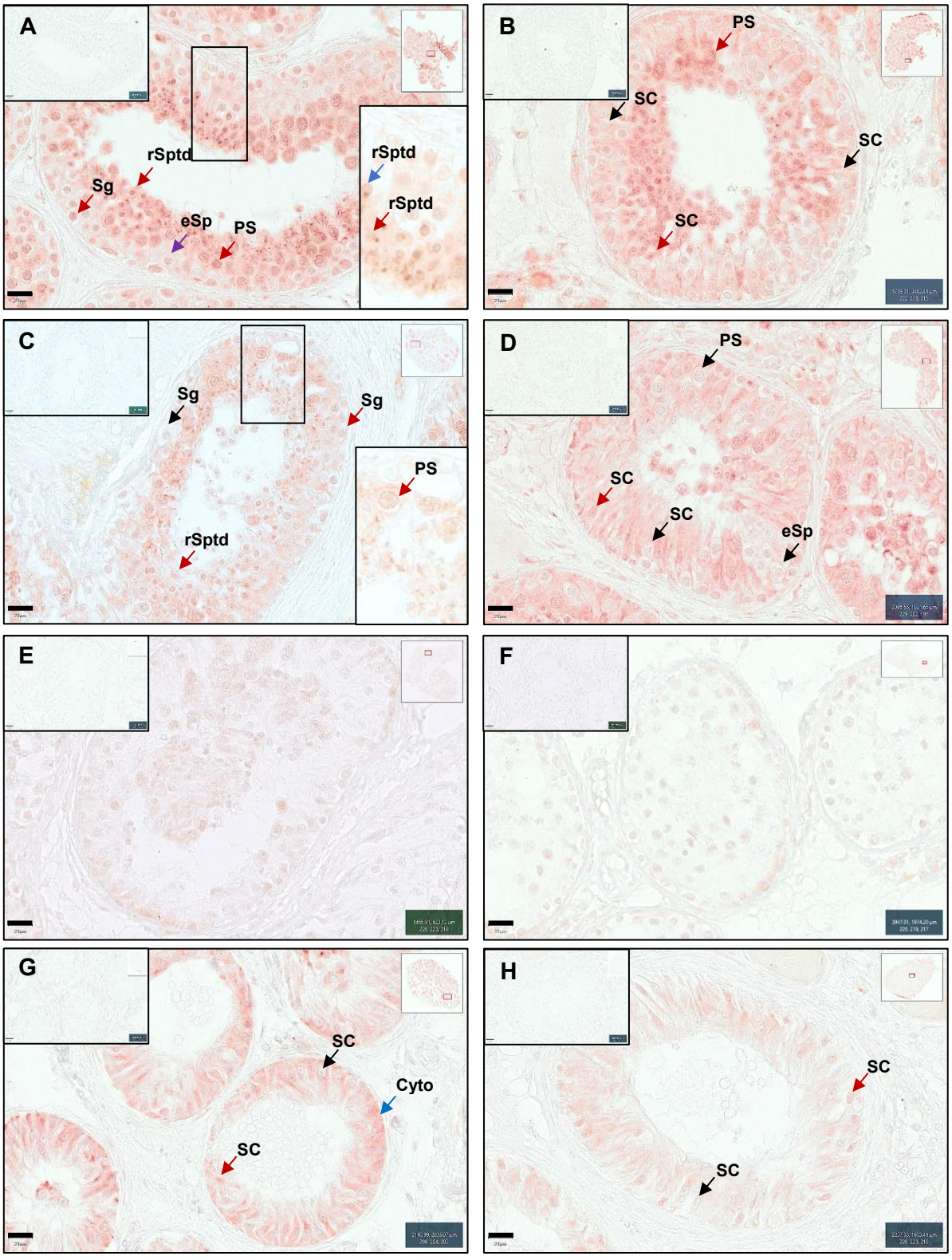


Figure 3.10: Figure legend in next page

Figure 3.10: Immunohistochemistry of IPO4 in human testis biopsies. Immunostaining was detected using NovaRed, and sections were counterstained with hematoxylin. Each phenotype contains biopsy images from two different individuals. **A, B** IPO4 immunostaining in NSP testis biopsies. **A.** Red arrows indicate nuclear staining in spermatogonia (Sg), early pachytene spermatocytes (PS) and round spermatids (rSptd). The purple arrow indicates an early spermatocyte with little staining (eSp). **A** *Inset*: red arrow indicate a positively stained round spermatid showing the Golgi (rSptd) and blue arrow indicates round spermatid with developing acrosome (rSptd). **B.** Black arrows indicate Sertoli cell (SC) nuclei with very little, if any, staining and the red arrow indicates nuclear staining in other Sertoli cells (SC) and both nuclear and cytoplasmic stained pachytene spermatocyte (PS). **C, D** - IPO4 immunostaining in SDA testis biopsies. **C.** Red arrows indicate positively stained spermatogonia (Sg) and round spermatids (rSptd) and the black arrow indicates some immuno-negative spermatogonia (Sg). *Inset*: red arrow indicates only nuclear staining in pachytene spermatocytes (PS). **D.** Black arrows indicate immuno-negative early spermatocytes (eSp), pachytene spermatocytes (PS) and Sertoli cells (SC). Red arrow indicates positive staining in Sertoli cell nuclei (SC). **E, F.** IPO4 immunostaining in SZA testis biopsies showing immuno-negative tubules. **G, H**. IPO4 immunostaining in SCO testis biopsies. Red arrows indicate immunostaining of Sertoli cell nuclei (SC) and black arrows indicate immuno-negative Sertoli cell nuclei (SC). The blue arrow indicates cytoplasmic staining in some Sertoli cells (Cyto). Respective negative control tissues are inserted in the upper left panels for each phenotype. Scale bars are 20μM. NSP – normal spermatogenesis, SDA – spermatid arrest, SZA – spermatocyte arrest, SCO – Sertoli cell only.

3.3.2.2 Localisation of ACTRT2 in human testes

In NSP biopsies, ACTRT2 staining was strongly concentrated towards the adluminal part of the tubules, consistent with the described localisation in *in silico* analysis (Fig. 3.11 A, B). Spermatogonia and early spermatocytes showed no obvious immunostaining (Fig. 3.11 B black arrow Sg and 3.11 A, black arrow eSp, respectively). ACTRT2 protein expression was first evident in late pachytene spermatocytes (stages IV - V) with light positive staining in the nuclei (Fig. 3.11 A inset, purple arrow PS). Round spermatids showed ACTRT2 in the nucleus (Fig. 3.11 A inset, red arrow rSptd). As spermiogenesis progressed, the staining intensity in the nucleus increased and cytoplasmic staining became apparent in step 3-4 spermatids (Fig. 3.11 B, red arrows rSptd and Sptd). Both nuclear and cytoplasmic staining was observed in elongating spermatids (Fig. 3.11 A inset and B, red arrow Sptd). Sertoli cells were immuno-negative (Fig. 3.11 B, black arrow SC).

In SDA biopsies, immunostaining was only evident in some round spermatids (Fig. 3.11 C, red arrow rSptd). Round spermatids in most of the areas in the tubule were

immuno-negative (Fig. 3.11 C, black arrow rSptd). Sertoli cells were immuno-negative in all tubules (Fig. 3.11 C, black arrow SC). No immunostaining was observed in SZA (Fig. 3.11 E, F) or SCO biopsies (Fig. 3.11 G, H). All negative controls (testis and placenta) were unstained.

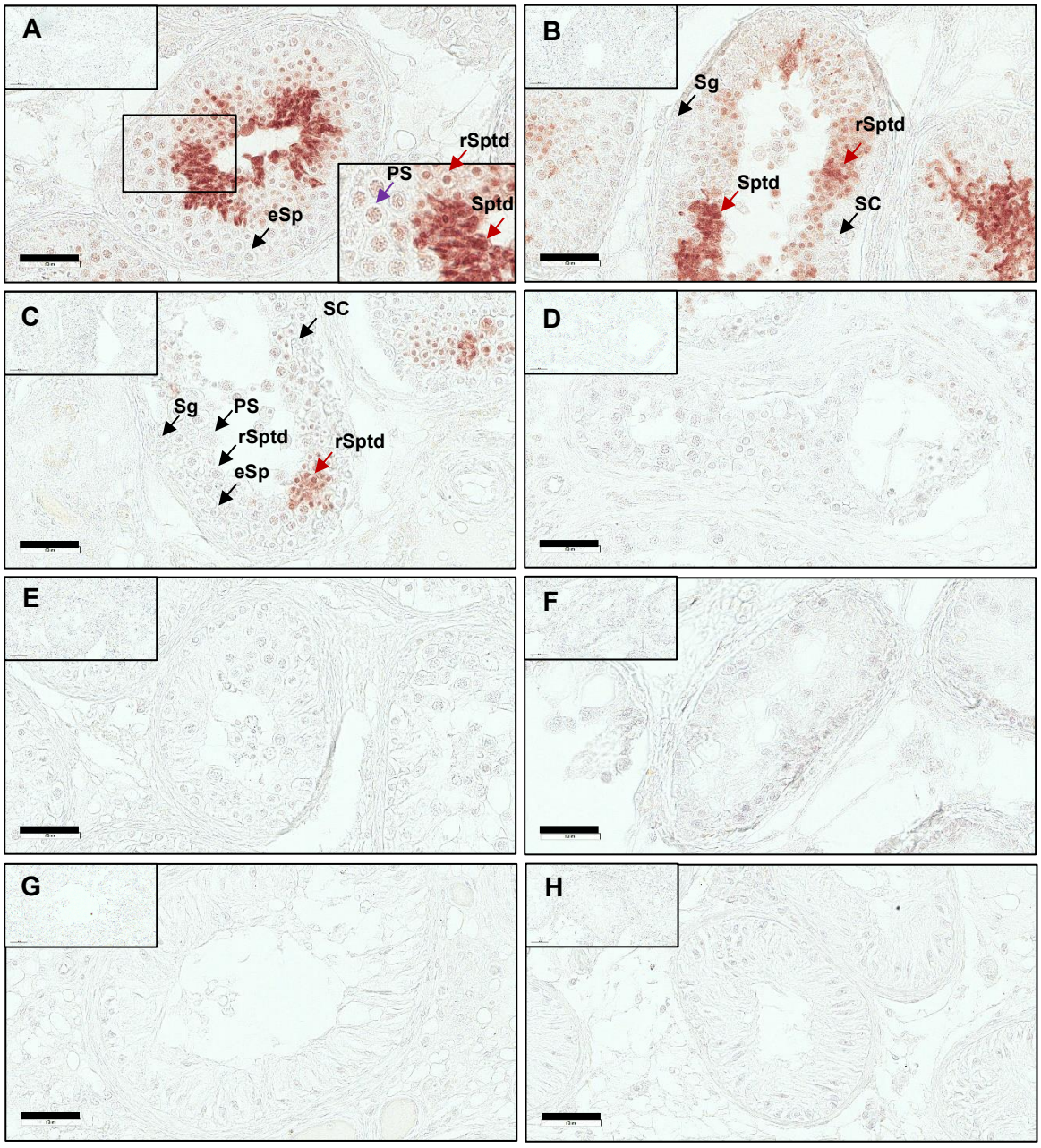


Figure 3.11: Immunohistochemistry of ACTRT2 in human testis biopsies. Immunostaining was detected using NovaRed, and sections were counterstained with hematoxylin. Each phenotype contains biopsy images from two different individuals. **A, B** ACTRT2 immunostaining in NSP testis biopsies. **A.** Black arrow indicates immuno-negative early primary spermatocytes (eSp). *Inset*: purple arrow indicates lightly stained pachytene spermatocyte nucleus (PS), and red arrows indicate nuclear staining in round spermatids

(rSptd) and nuclear and cytoplasmic staining of elongating spermatids (Sptd). **B.** Black arrows indicate immuno-negative spermatogonia (Sg) and Sertoli cells (SC) and red arrow indicates nuclear and cytoplasmic staining of round spermatids (rSptd) and nuclear and cytoplasmic staining of elongating spermatids (Sptd). **C, D** ACTRT2 immunostaining in SDA testis biopsies. **C.** black arrows indicate immuno-negative spermatogonia (Sg), early spermatocytes (eSp), pachytene spermatocytes (PS), round spermatid (rSptd) and Sertoli cell (SC). Red arrow indicates positive immunostaining in some round spermatids (rSptd). **D.** completely unstained SDA tubule. **E, F** ACTRT2 immunostaining in SZA testis biopsies; no specific staining is observed in any tubules. **G, H** ACTRT2 immunostaining in SCO testis biopsies; no specific immunostaining is observed in any tubules. Respective negative control tissues are inserted in the upper left panels phenotype. Scale bars are 50μM. NSP – normal spermatogenesis, SDA – spermatid arrest, SZA – spermatocyte arrest, SCO – Sertoli cell only.

3.3.2.3 ELSPBP1 in human testes tissues

All testis samples regardless of the spermatogenesis state were immuno-negative (Fig. 3.12 A-F). ELSPBP1 was immunolocalised in principle epithelial cells of the epididymal ducts (Fig. 3.12 G, red arrow EC) while the placenta tissue used as negative control was unstained (Fig. 3.12 H).

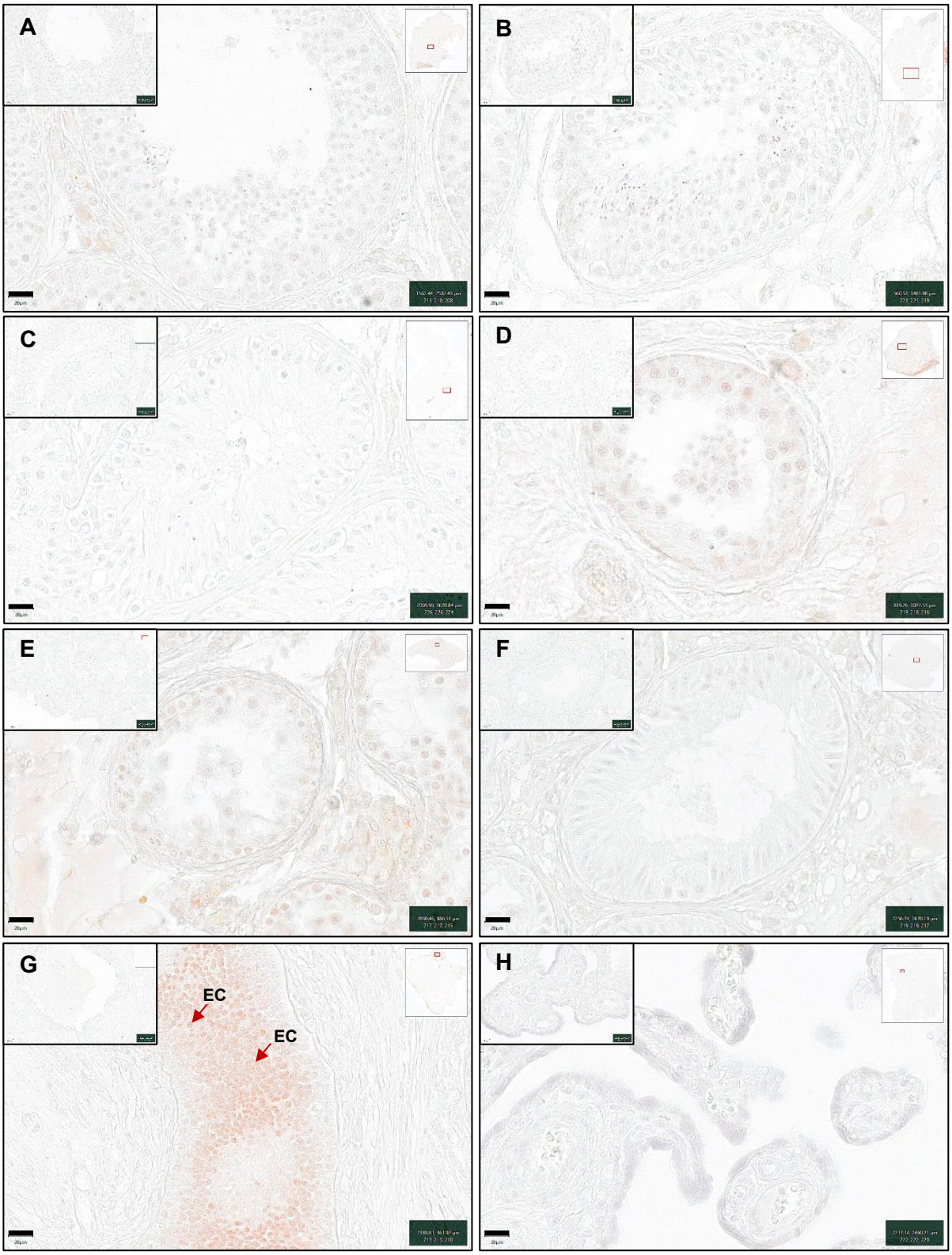


Figure 3.12: Immunohistochemistry of ELSPBP1 in human testis biopsies, epididymis and placenta tissues. Immunostaining was detected using NovaRed, and sections were counterstained with hematoxylin. No ELSPBP1 staining in NSP (**A, B**), SDA (**C, D**), SZA (**E**), or SCO testis biopsies (**F**). **G.** ELSPBP1 was observed in epididymis and was primarily observed in epithelial cells (EC) of the epididymal ducts, with the surrounding stroma immuno-

negative. Red arrows indicate epithelial cells (EC). **H.** Lack of ELSPBP1 staining in placenta as the negative control tissue. Respective negative control tissues are inserted in upper left panels for each phenotype. Scale bars are 20μM. NSP – normal spermatogenesis, SDA – spermatid arrest, SZA – spermatocyte arrest, SCO – Sertoli cell only.

3.3.2.4 Localisation of CCDC105 in human testis

In NSP biopsies, CCDC105 immunostaining was observed in the nuclei of all germ cells (Fig. 3.13 A, B). Staining was observed at low levels in the nuclei of spermatogonia and early spermatocytes (Fig. 3.13 A purple arrow eSp and inset purple arrow Sg). Nuclear staining persisted and became slightly more intense during the development of pachytene spermatocytes (Fig. 3.13 A inset, red arrow PS). As germ cells progressed through meiosis to spermiogenesis, more intense staining was observed in the nuclei of round spermatids and persisted to elongating and elongated spermatids (Fig. 3.13 B, purple arrow rSptd and 3.13 A inset, red arrow Sptd). Sertoli cells were immuno-negative (Fig. 3.13 A, B black arrow SC). The staining intensity of NSP biopsies varied between patients (Fig. 3.13, panels A vs B).

In SDA biopsies, the overall staining intensity of CCDC105 was much lower compared to NSP, with most SDA biopsies showing no immunostaining (Fig. 3.13 C, D). SZA (Fig. 3.13 E, F) and SCO (Fig. 3.13 G, H) samples showed no immunostaining. All negative controls (testis and placenta) were unstained.

Figure 3.13: Immunohistochemistry of CCDC105 in human testis biopsies. Immunostaining was detected using NovaRed, and sections were counterstained with hematoxylin. Each phenotype contains biopsy images from two different individuals. **A, B** CCDC105 immunostaining in NSP testis biopsies. **A.** purple arrow indicates lightly stained early spermatocytes (eSp), and black arrow indicates immuno-negative Sertoli cells (SC) *inset:* purple arrow indicates lightly stained spermatogonia's nucleus (Sg) and red arrows indicate nuclear immunostaining in pachytene spermatocytes (PS) and spermatids (Sptd). **B -** black arrow indicates immuno-negative Sertoli cells (SC) and purple arrow indicates lightly nuclear stained round spermatid (rSptd). **C**, **D** – CCDC105 immunostaining in SDA testis biopsies showing immuno-negative tubules **E, F** CCDC105 immunostaining in SZA testis biopsies; no specific staining is observed in any tubules. **G, H** CCDC105 immunostaining in

SCO testis biopsies; no specific staining is observed in any tubules. Respective negative control tissues are inserted in upper left panels for each phenotype. Scale bars are 50μM. NSP – normal spermatogenesis, SDA – spermatid arrest, SZA – spermatocyte arrest, SCO – Sertoli cell only.

3.3.2.5 Localisation of IFT57 in human testis

In NSP biopsies, A_{pale} spermatogonia showed IFT57 immunostaining in their nuclei (Fig. 3.14 A inset, red arrow SgA_p), while A_{dark} were unstained (Fig. 3.14 A and Fig. 3.14 B inset, black arrow SgA_d). Nuclear immunostaining was evident in type B spermatogonia and early spermatocytes (Fig. 3.14 A, red arrow eSp). Pachytene spermatocytes showed light to variable nuclear immunostaining which increased in intensity as they progressed through meiosis. In stage V, pachytene spermatocytes showed granulated immunostaining in the nuclei with areas of dark, granular cytoplasmic staining depending on the stage of development (Fig. 3.14 B inset, red arrow PS). This accumulation was not observed in early pachytene spermatocytes (Fig. 3.14 A inset, purple arrow PS). Round spermatids up to step 3 showed light nuclear and cytoplasmic staining (Fig. 3.14 B inset, red arrows rSptd). In steps 4-5 spermatids, IFT57 was concentrated in the nucleus (Fig. 3.14 A inset, red arrow Sptd). Sertoli cell cytoplasm was immuno-negative, but nuclei were mostly stained (Fig. 3.14 A, red arrow SC), although some immuno-negative nuclei were also observed, suggesting stage-specific nuclear localisation (Fig. 3.14 A inset, black arrow SC). Interstitial staining was observed in NSP samples with Leydig cells often showing granular cytoplasmic staining (Fig. 3.14 B, green arrow Leyd).

In SDA tissues (Fig. 3.14 C, D), the staining pattern was broadly similar to NSP biopsies in terms of spermatogonia and spermatocyte immunostaining (Fig. 3.14 C, red arrow SgA_p and black arrows SgA_d, eSp). Most pachytene spermatocytes showed light nuclear staining (Fig. 3.14 D, purple arrow PS), some with the granular cytoplasmic staining (Fig. 3.14 C, D, red arrows PS). Round spermatid nuclei were variably stained, some showed no staining (Fig. 3.14 C, black arrow rSptd), but others showed lightly stained nuclei (Fig 3.14 C, purple arrow rSptd). Sertoli cells were mostly unstained (Fig. 3.14 D, black arrow SC) with some stained nuclei (Fig. 3.14 D, red arrow SC).

In SZA biopsies, differences in staining intensity were observed between samples from different patients (Fig. 3.14 E vs F). Biopsies showed the same staining pattern as in NSP and SDA (Fig. 3.14 E, red arrow SgA_p. Fig. 3.14 F, black arrow SgA_d). Early primary spermatocytes showed lightly stained nuclei and unstained cytoplasm (Fig. 3.14 E, F, purple arrows eSp). Pachytene spermatocytes' nuclei were homogeneously stained without showing the granulated cytoplasm as was seen in NSP and SDA (Fig. 3.14 E, red arrow PS). Mostly immuno-negative Sertoli cells were observed (Fig. 3.14 E, F black arrows SC) with occasional immuno-positive nuclei (Fig. 3.14 E red arrow SC). In SCO biopsies, Sertoli cell nuclei showed considerable staining heterogeneity, with both stained and unstained nuclei observed (Fig. 3.14 G, H red and black arrows, SC). All negative controls (testis and placenta) were unstained.

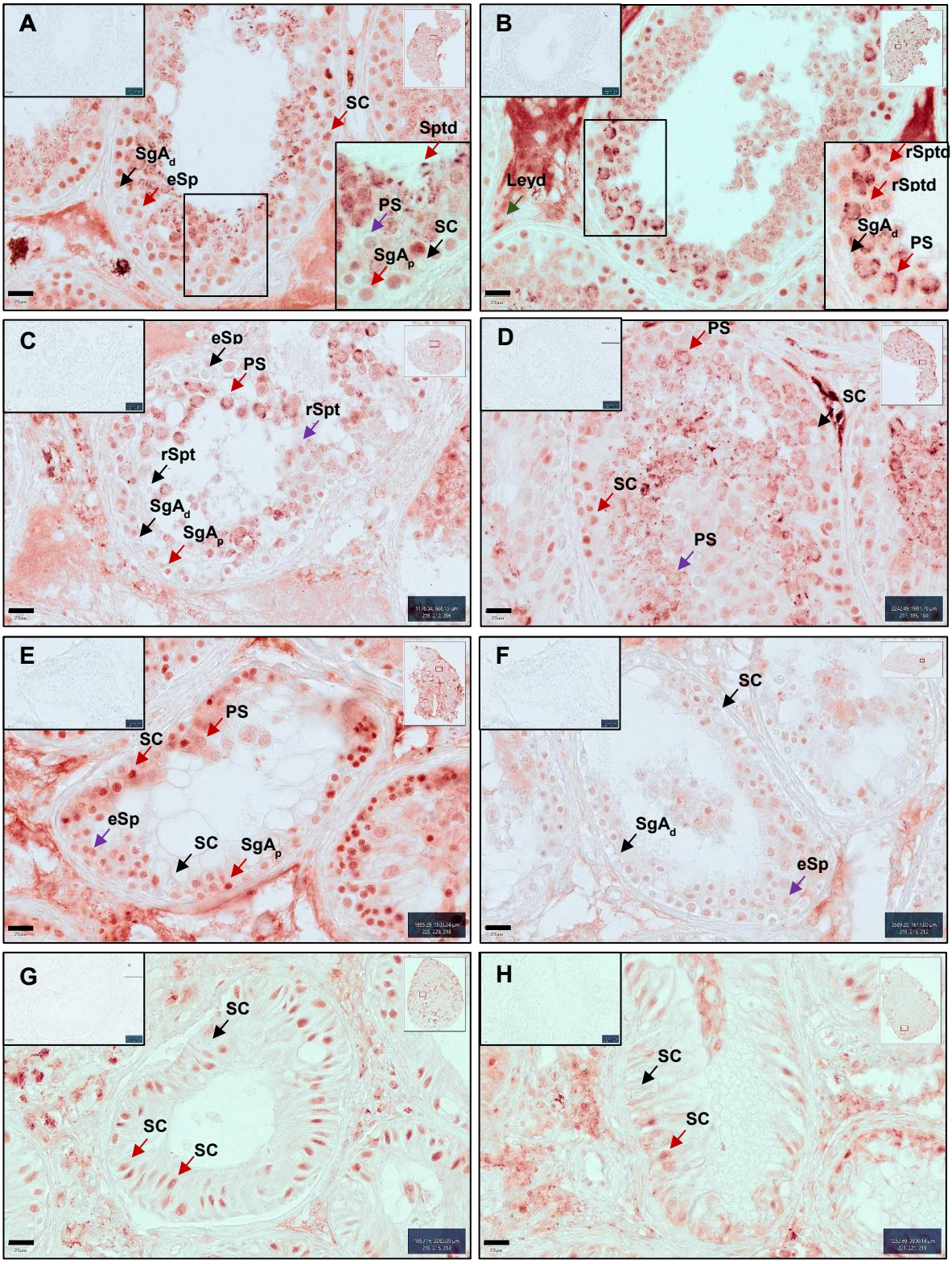


Figure 3.14: Immunohistochemistry of IFT57 in human testis biopsies. Immunostaining was detected using NovaRed, and sections were counterstained with hematoxylin. Each phenotype contains biopsy images from two different individuals. **A, B.** IFT57 immunostaining in NSP biopsies. **A.** Black arrows indicate immuno-negative A_{dark} spermatogonia (SgA_d) and red arrows indicate immunostaining of early primary spermatocytes (eSp) and Sertoli cells (SC). **A** *Inset*: black arrow indicates immuno-negative Sertoli cells (SC) and red arrows indicates nuclear staining in A_{pale} spermatogonia (SgA_p) and spermatids (Sptd). The purple

arrow indicates some pachytene spermatocytes with light immunostaining (PS). **B.** NSP biopsy showing granulated pachytene spermatocytes. Green arrow indicates stained Leydig cell (Leyd). *Inset:* black arrow indicates immuno-negative A_{dark} spermatogonia (SgA_d) and red arrows indicate pachytene spermatocytes with light nuclear staining and condensed cytoplasmic staining (PS) and nuclear staining in round spermatids (rSptd). **C, D.** IFT57 immunostaining in SDA. **C.** Black arrows indicate immuno-negative A_{dark} spermatogonia (SgA_d), early primary spermatocytes (eSp) and round spermatids (rSptd). The purple arrow indicates light nuclear staining in round spermatids (rSptd) and red arrows indicate nuclear staining in A_{pale} spermatogonia (SgA_p) and pachytene spermatocytes (PS). **D.** Black arrow indicates immuno-negative Sertoli cells (SC), red arrows indicate cytoplasmic staining in pachytenes (PS) and nuclear staining in Sertoli cells (SC). The purple arrow indicates lightly nuclear stained pachytene spermatocytes (PS). **E, F.** IFT57 immunostaining in SZA biopsies. **E –** black arrow indicates immuno-negative Sertoli cells (SC) and red arrows indicate some Sertoli cells with nuclear staining (SC), nuclear staining in A_{pale} spermatogonia (SgA_p) and pachytene spermatocytes (PS). Purple arrow indicates light nuclear staining in early spermatocytes (eSp). **F.** Black arrows indicate immuno-negative A_{dark} spermatogonia (SgA_d) and Sertoli cells (SC) and the purple arrow indicates light nuclear staining in early spermatocytes (eSp). **G, H.** IFT57 immunostaining in SCO biopsies, red arrows indicating nuclear staining in Sertoli cells (SC) and black arrows indicating some immuno-negative Sertoli cells (SC). Respective negative control tissues are shown in the upper left panel for each phenotype. Scale bars are 20µM. NSP – normal spermatogenesis, SDA – spermatid arrest, SZA – spermatocyte arrest, SCO – Sertoli cell only.

3.3.3 Protein quantification in human semen samples

Antibodies for the five selected protein were first tested in human testis lysates using chemiluminescence Western Blot (WB). Antibodies for all five proteins gave the expected molecular weight protein bands, and the negative control blot was clear (Fig. 3.15), indicating that these antibodies can be used to detect the candidate proteins in human testis.

For the relative quantification of the protein amounts in ejaculates from patients with AT sperm compared to NORM, the Western blot system was changed to fluorescence detection. The Ponceau image confirmed the completed run of proteins (Fig. 3.16 A). IPO4 and CCDC105 did not give a specific protein band in fluorescent WB although specific bands were detected using the chemiluminescence method (Fig 3.15). Due to time constraints, IPO4 and CCDC105 were not analysed further.

Quantification of each protein in WB was determined in relation to alpha tubulin loading control. Each fluorescent WB was repeated twice for reproducibility and Ponceau images of each run confirmed successful protein separation (Suppl. Fig. 3.2). The mean value of relative expression for each sample was obtained as described in the

methods. Statistical differences were determined between NORM vs AT using the student's t test (GraphPad Prism). The number of selected samples used for quantitative computation differed for each protein (see Discussion).

Using the fluorescent WB method, bands of the appropriate molecular weight were observed in all semen samples; 41kDa for ACTRT2, 26k – 31kDa for ELSPBP1 and 49kDa for IFT57 (Fig. 3.16 B, C). Quantitation of band intensity revealed that ACTRT2 (Fig. 3.17) did not show a difference in the band intensity between AT and NORM. ELSPBP1 (Fig. 3.18) and IFT57 (Fig. 3.19) showed a trend to increase as suggested by the mass spec analysis (Table 3.3), but the difference was not statistically significant (P>0.05) between AT and NORM ejaculates.

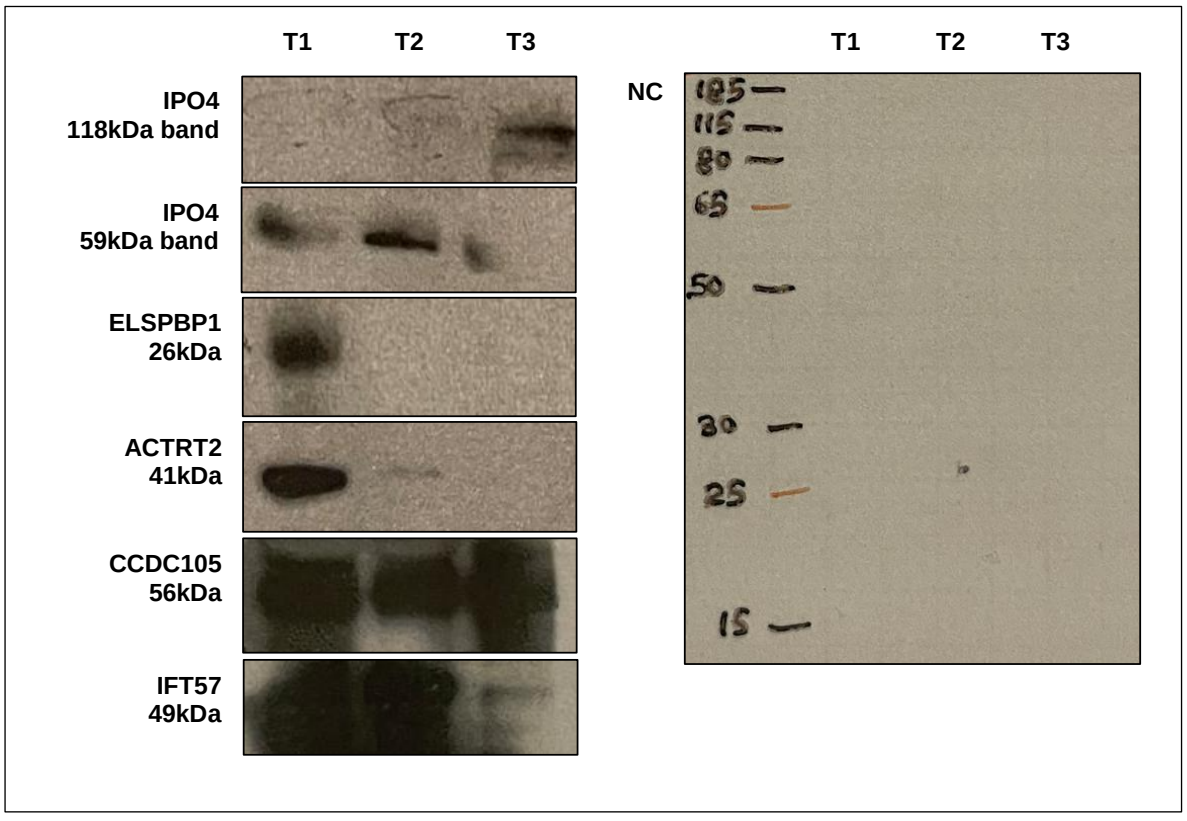


Figure 3.15: Chemiluminescence Western blot optimisation. Testis lysates were analysed by Western blots for all 5 proteins. T1 = testis lysates extracted by AP-40, T2 = testis lysates extracted by T-Per, and T3 = testis lysates extracted by Trizol. All five proteins were detected at the appropriate molecular weights: IPO4 gave 59kDa and 118kDa bands, ELSPBP1 – 26kDa, ACTRT2 – 41kDa, CCDC105- 56kDa and IFT57 – 49kDa. The negative control blot was free from bands. NC – negative control blot.

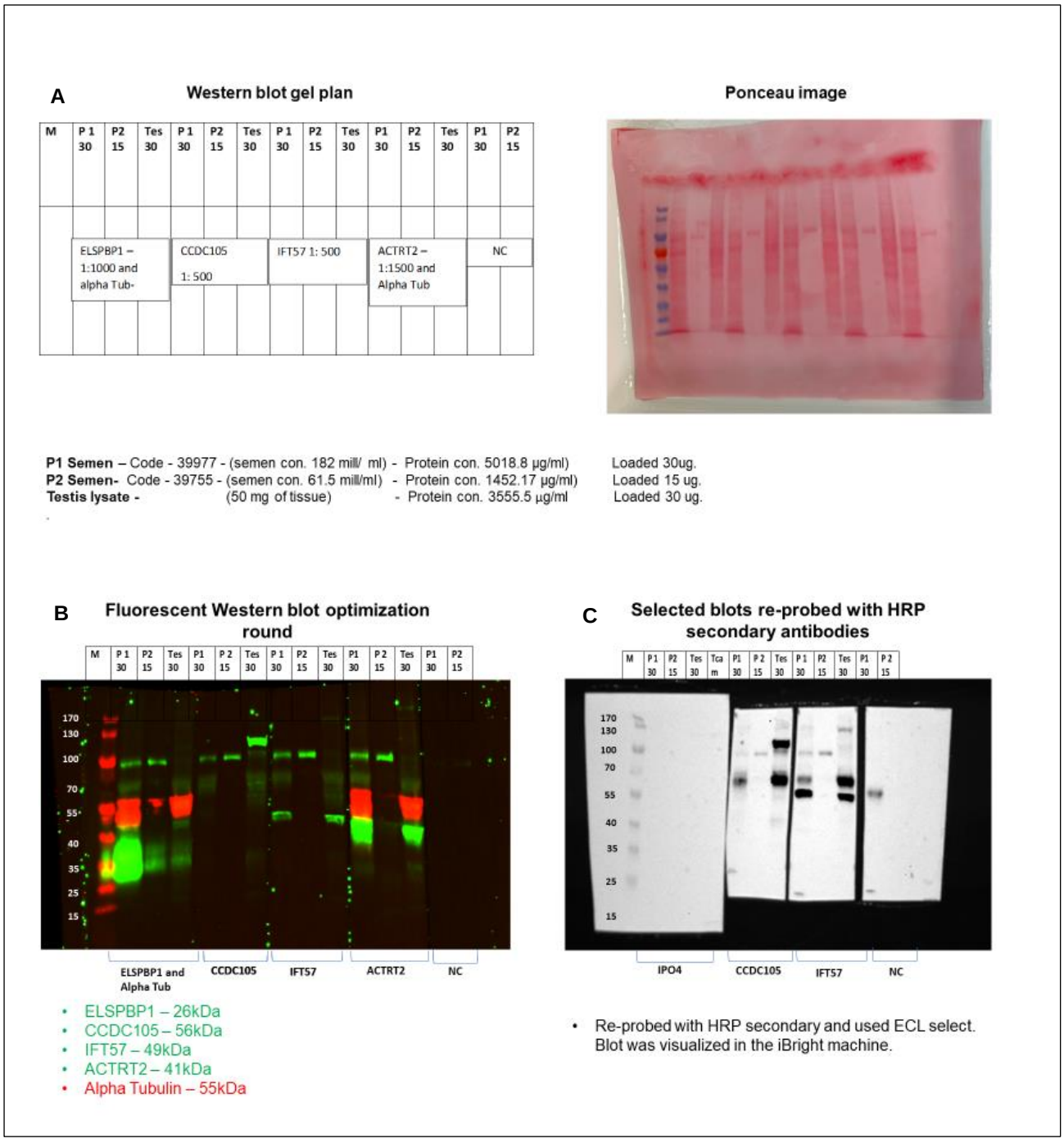


Figure 3.16: Fluorescent Western blot optimisation for ELSPBP1, CCDC105, IFT57, and ACTRT2 and re-probing of selected blots with HRP conjugated secondary antibodies. **A** – Gel plan for the 4 proteins and the Ponceau images depicting the successful run of proteins in each well. **B** – Fluorescent image of the blot. The green bands are the target protein signal and the red bands around 55kDa are the alpha tubulin reference protein control. A strong signal showed around 26-40kDa for ELSPBP1, 49kDa band for IFT57 and 41kDa band for ACTRT2. **C** – The re-probed blot using the chemiluminescence method. CCDC105, IFT57 and NC blots were re-probed with HRP conjugated secondary antibodies, treated with ECL select and imaged in the iBright machine. A 56kDa band for CCDC105, 49kDa band for IFT57. M – marker, P1, P2. Respective patient protein lysates, N/NORM – semen samples from normozoospermic patients, AT – semen samples from patients with a defined phenotype of

asthenoteratozoospermia, 30 – 30µg of proteins loaded per well, Tes- testis lysate as a control, NC – negative control.

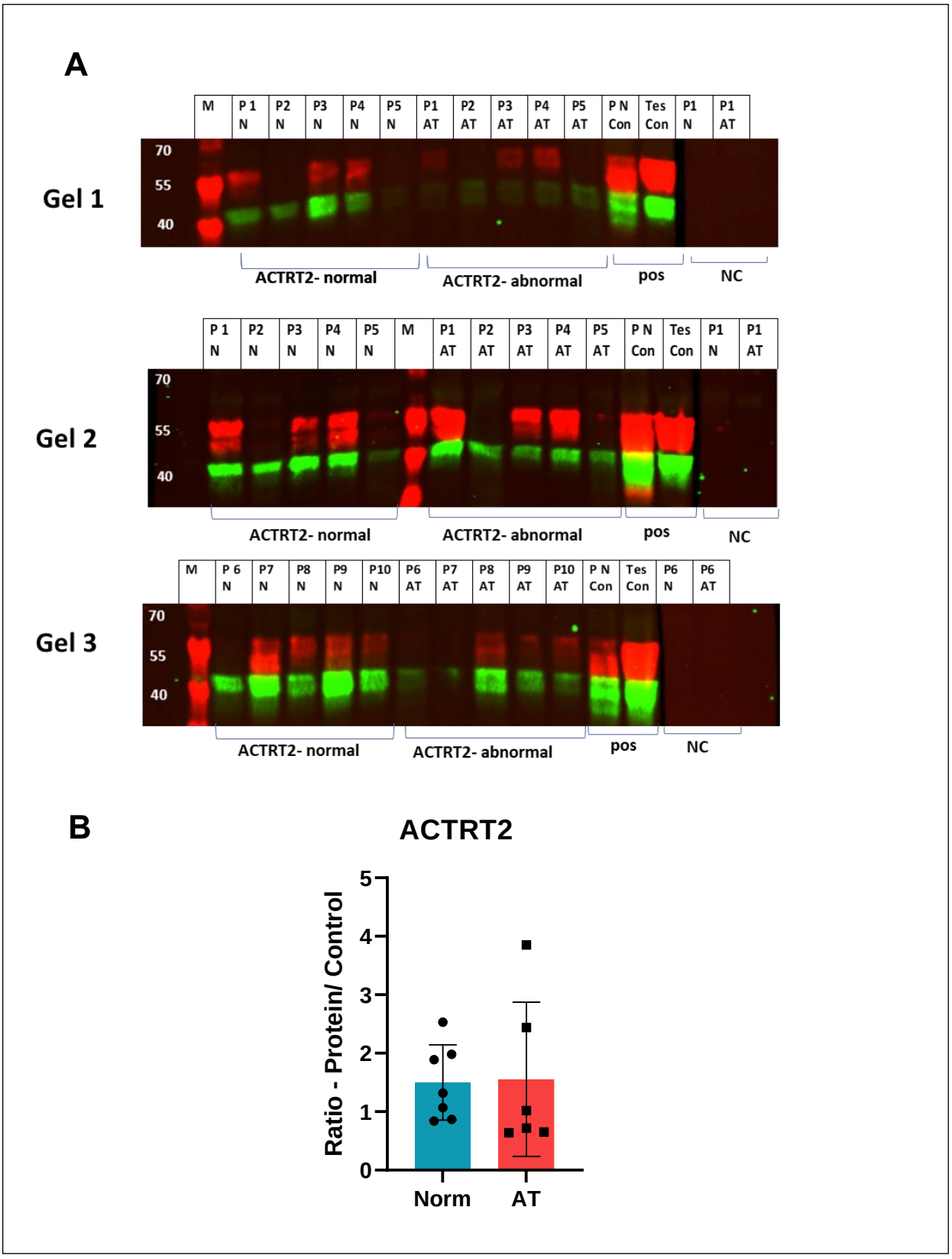


Figure 3.17: Levels of ACTR2 protein in semen samples from men with normal semen vs men with asthenoteratozoospermia. (A) The fluorescent images of ACTRT2 Western blots and **(B)** the relative protein quantification**.** M – marker (protein ladder), P1-P10 depicts the different patient codes, N/NORM – normal semen, semen samples from normozoospermic

patients (n=7), AT – semen samples from patients with a defined phenotype of asthenoteratozoospermia (n=6) p > 0.05. Tes – testis lysate as a control, Pos – positive control blot, NC – negative control blot, Norm – normal semen.

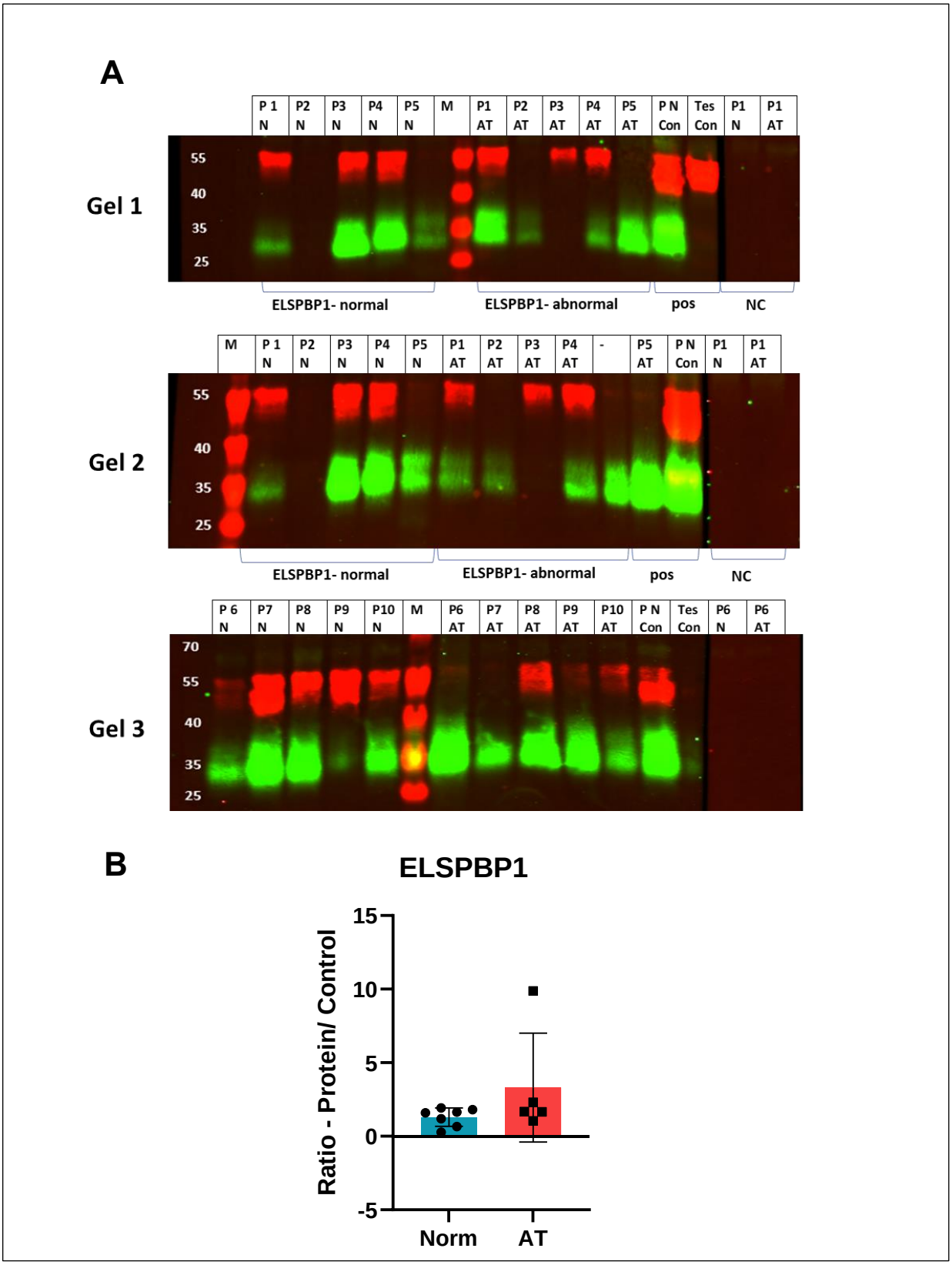


Figure 3.18: Figure legend in next page

Figure 3.18: Levels of ELSPBP1 protein in semen samples from men with normal semen vs men with asthenoteratozoospermia. (A) The fluorescent images of ELSPBP1 Western blots and **(B)** the relative protein quantification. M – marker (protein ladder), P1-P10 depicts the different patient codes, N/Norm – semen samples from normozoospermic patients (n=7), normal semen, AT – semen samples from patients with a defined phenotype of asthenoteratozoospermia (n=5), $p > 0.05$ Tes – testis lysate as a control, Pos – positive control blot, NC – negative control blot, Norm – normal semen

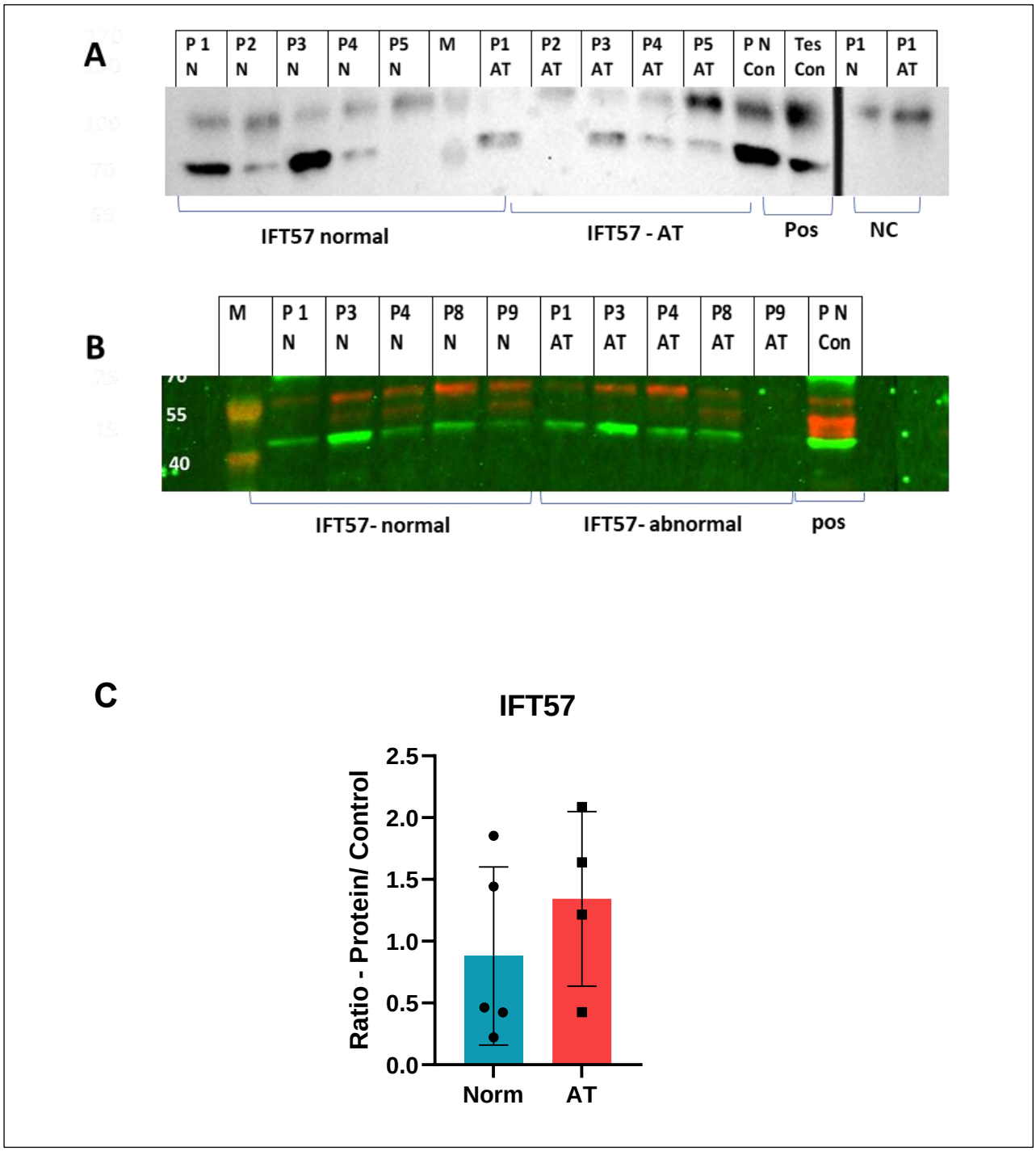


Figure 3.19: Figure legend in next page

Figure 3.19 Levels of IFT57 protein in semen samples from men with normal semen vs men with asthenoteratozoospermia. (A). The image of chemiluminescence Western blot image of IFT57, **(B)** the fluorescent image of IFT57 Western blot and **(C)** the relative quantification. The blot re-probed with HRP conjugated secondary antibodies, treated with ECL select and images using iBright machine and selected samples were repeated in fluorescent Western blot. M – marker (protein ladder), P1-P10 depicts the different patient codes, N/Norm – semen samples from normozoospermic patients (n=5), AT – semen samples from patients with a defined phenotype of asthenoteratozoospermia (n=4), $p>0.05$ Tes – testis lysate as a control, Pos – positive control blot, NC – negative control blot, Norm – normal semen

3.4 Discussion

Mass spectrometry (MS) plays a major role in the characterisation and understanding of proteins in biological systems (Cravatt et al., 2007). Advancements in machinery, databases, bioinformatics, and statistical tools have expanded the sensitivity, capability and utility of MS (Castellana & Bafna, 2010). Mass spectrometry is applicable to identifying the proteome of a sample, finding novel protein coding genes and gene regions responsible for a disease, and to identifying proteins that are differentially expressed between samples (Baker et al., 2013; Baldwin, 2004; Bitton et al., 2010; Castellana et al., 2014; Khan et al., 2018; Klasberg et al., 2016).

This chapter focused on identifying proteins that were differentially expressed in human ejaculates with normal (normal motility and morphology, NORM) and altered morphology and motility (i.e. asthenoteratozoospermia, AT). MS was used to identify novel proteins that could play a significant role in sperm morphology and motility and this study unveiled 35 proteins that showed a significant difference between NORM and AT sperm. The semen samples used had a well characterised phenotype, as per WHO guidelines. A bioinformatic approach using portals such as UniProt, GermOnline, the Human Protein Atlas (HPA), the Young Adult Testis Atlas and GeneCards enabled the investigation of these 35 proteins in terms of organ/tissue specificity and function. Previously published data on these proteins were accessed using PubMed to identify proteins which have limited previous findings related to human sperm morphology and motility. The majority of the DEPs (n=23) have been previously reported to be associated with fertility or conditions in either human or other mammals. For example, IRGC was decreased in AT ejaculates in this study and has previously been shown to be downregulated in semen from asthenozoospermic men (Li et al., 2023). DEFB129

was increased in AT ejaculates in this study and has previously been shown to be increased (mRNA) in caput epididymal tissue retrieved from non-obstructive azoospermic men (Dube et al., 2008). The remaining proteins were assessed according to their functional annotations, expression patterns in different organs, specificity to testis/epididymis, expression pattern in germ cells, expression pattern of mRNA in the testis and ten proteins were determined to be likely candidates that could be important in human spermatogenesis and/or sperm function). Of these, five proteins were selected for further analysis based on availability of antibodies suitable for immunohistochemistry (IHC) and Western blot analysis. These candidates included proteins that were testis-specific, predominantly expressed in human testis and germ cells or epididymis-specific and one protein, CCDC105, was chosen based on its likely localisation in human spermatids but no descriptions of the protein in the published literature at the time of selection. The localisation of these proteins in testis biopsies of different pathologies was determined by IHC, and the levels in normal vs AT sperm were quantified by Western blot analysis.

CCDC105 was significantly decreased in sperm samples from patients with AT samples. There was no published literature found in PubMed for CCDC105 at the time of its selection and there was no description of its functional role in the various databases investigated as described above. According to UniProt, *CCDC105* is an uncharacterised protein-coding gene and, according to HPA, its mRNA expression is restricted to the testis, with the highest expression in late spermatids (Uhlén et al., 2015). Thus, human CCDC105 is likely only expressed in developing sperm. The cellular localisation of CCDC105 protein was then assessed in human testis biopsies by IHC. CCDC105 was localised in all germ cell nuclei with increased intensities observed in nuclei from late pachytene spermatocytes to spermatids. These results are supported by IHC data from HPA, where CCDC105 was localised in late pachytene spermatocytes and spermatids, although the immunostaining was very faint. Interestingly, CCDC105 localisation was very low to undetectable in SDA, SZA, and SCO biopsies because these cells lack spermatids where CCDC105 is primarily found. Although CCDC105 was able to be detected in human testis lysates by Western blotting, it was not able to be quantified in human sperm lysates by quantitative fluorescence Western blot, and thus further studies are required to validate the

difference in protein levels of CCDC105 in AT sperm compared to NORM, as was indicated by the mass spec analysis.

Analysis of protein differences between sperm from normozoospermic patients compared to patients with a phenotype of AT revealed that several epididymis-specific proteins differed between the two groups (ELSPBP1, DEFBP129) and some that showed high expression in epididymis (RCN2, HNRNPD). It is well known that testicular spermatozoa leave the testis and enter the epididymis in an immotile state, however they gain their motility and ability to fertilise an oocyte as they traverse the epididymis. Spermatozoa pass through the caput, corpus and cauda sections of the epididymis, with functional maturation taking place in the caput and corpus, while the cauda primarily acts as the storage place for functionally mature sperm (Cornwall, 2009). During the migration through the epididymis, spermatozoa encounter many proteins that are secreted by the epididymal cells, and these proteins are taken up by the maturing spermatozoa (Björkgren & Sipilä, 2019; Cornwall, 2009). Proteins that are already present in the sperm can undergo post-translational modifications and functional maturation during epididymal transit, such as phosphorylation, glycolysation and acetylation that takes place during the epididymal maturation, and that interactions with epididymal proteins can affect the activity of sperm proteins (Björkgren & Sipilä, 2019). Ultimately, these interactions assist in regulating sperm motility, acrosome reaction and capacitation (Björkgren & Sipilä, 2019). Thus, post-testicular sperm maturation is accomplished in the epididymis and is essential for male fertility (Nixon et al., 2020). A number of proteins that were altered between AT and NORM ejaculates are also epididymal specific (ELSPBP1, DEFBP129) or highly expressed in the epididymis according to HPA. Given that these proteins differed in ejaculates depending on sperm morphology and motility, it seems likely that these proteins are acquired by the sperm as they traverse the epididymis, and that the ability of the sperm to acquire these proteins depends on the sperm phenotype.

ELSPBP1 was thus further investigated because it was likely an epididymal-derived protein that was higher in abundance in sperm with a phenotype of AT. The absence of ELSPBP1 in the testis and its presence in the epithelial cells lining the epididymal ducts was confirmed by IHC, suggesting that ELSPBP1 is acquired by sperm as they transit through the epididymis. The increase in ELSPBP1 protein levels in AT sperm identified by mass spectrometry was investigated in a larger number of sperm samples

by quantitative Western blot analysis, however a significant difference was not observed, perhaps due to high variability between patients. An issue was encountered for all Western blots where there was between-sample variation in the loading control (alpha tubulin) despite all samples containing adequate protein concentrations. Thus, only those samples that produced both the test protein and loading control bands were selected for quantification. This variation could be due to the fact that reference protein controls used in Western blot show a between-patient variability as reported before (Feng et al., 2021). Also, that the semen samples had been previously frozen, as frozen samples show a higher between-patient variability (Feng et al., 2021). In future studies, fresh semen samples could be used to reduce between-patient variability and increase the chance of observing statistically significant differences. Whether ELSBP1 is higher in ejaculates from AT patients, needs to be confirmed in a larger patient cohort. The molecular function of ELSPBP1 is not well characterised but a study showed that ELSPBP1 was bound to the post-acrosomal and mid-piece regions of the dead bovine spermatozoa and that it tends to bind more strongly to dead vs motile spermatozoa (D'Amours et al., 2012b). Interestingly, a recent study showed that ELSPBP1 protein levels were high in human sperm with higher DNA fragmentation (Belardin et al., 2022). Taken together, these results suggest that ELSPBP1 is produced in the epididymis and is loaded onto the sperm as they transit through the ducts. However, the amount of ELSPBP1 loaded onto sperm increases when sperm are abnormal or have reduced viability. These findings suggest that ELSPBP1 could preferentially bind to morphologically defective human spermatozoa, however the mechanism and whether it binds to specific sites on the sperm remains to be investigated. Future studies could also investigate whether the level of ELSPBP1 in human sperm correlates with abnormal sperm function, motility, and morphology and thus whether it could be useful as a biomarker to aid in the diagnosis of abnormal sperm morphology.

ACTRT2 is an Actin-related protein that was determined by MS as being lower abundance in AT semen compared to NORM samples. According to HPA, both ACTRT2 mRNA and protein show the highest expression in the human testis compared to other tissues, and this is due to its mRNA expression being restricted to germ cells, particularly elongated spermatids (Pontén et al., 2009; Uhlén et al., 2015), suggesting a specific role in spermiogenesis. ACTRT2 is localised in the post-

acrosomal area and the mid-piece of human spermatozoa (Liu et al., 2015) and is detected in a multimeric complex localised to the subacrosomal region in mouse spermatids (Zhang et al., 2022). Interestingly, the levels of ACTRT2 in spermatozoa were reduced in sperm from patients with asthenozoospermia and obesity (Liu et al., 2015) and were reduced in mouse and human sperm under conditions of scrotal hypothermia (Wu et al., 2020). Although the significant reduction in ACTRT2 in AT sperm observed in the mass spectrometry analysis was not confirmed by Western blot, overall, the data are consistent with a role of ACTRT2 in human sperm and that its levels are reduced in sperm with abnormal morphology and/or function. IHC analyses in this chapter confirmed that ACTRT2 is mainly expressed during spermiogenesis and is localised to both the nucleus and cytoplasm of developing spermatids. The apparent reduction of ACTRT2 in round spermatids in SDA testis biopsies needs to be confirmed in a larger set of samples, however it is possible that ACTRT2 protein production during spermiogenesis depends on the normal complement of germ cells and/or on signals from Sertoli cells. In summary, the results suggest that ACTRT2 plays a role during spermiogenesis, is functionally important in human spermatozoa, and that reduced ACTRT2 protein levels could be indicative of abnormal sperm development and function.

IFT57, expressed in various tissues, showed higher abundance in AT samples. IFT57 is an intraflagella transporter protein. According to HPA, the protein is detectable in all organs (Pontén et al., 2009), but the highest mRNA expression is evident in human spermatids (Uhlén et al., 2015). The function of IFT57 in the human testis and sperm has not been investigated. Studies in mice have shown that the deletion of other intraflagellar transport proteins such as *Ift172* (Zhang et al., 2020), *Ift81* (Qu et al., 2020) and *Ift74* (Shi et al., 2019) lead to defects in spermiogenesis and the development and function of the sperm flagellum, and also to a reduction in IFT57 levels in *Ift72* and *Ift74* deficient mice, highlighting its likely role in the sperm flagellum. The IHC results in this chapter confirmed that IFT57 is expressed in the human testis, in germ cells as well as Sertoli cells. The localisation in Sertoli cell nuclei was stage-specific, suggesting an involvement in cyclic Sertoli cell function. The IHC results suggested that the subcellular localisation of IFT57 in germ cells changes during meiosis, by showing darkly stained cytoplasmic staining in late pachytene spermatocytes and spermiogenesis, with, its localisation in elongating spermatids.

These results suggest that IFT57 could play multiple roles in germ cell development and contribute to ejaculated sperm sperm function.

The transporter protein IPO4 showed higher abundance in AT compared to control semen samples. IPO4 belongs to the family of importins that are involved in the RanGTPase cycle and function as nuclear transport receptors. Some importins have been shown to be involved in spermatogenesis (Bernardes et al., 2022; Loveland et al., 2015; Miyamoto et al., 2020; Nathaniel et al., 2022; Whiley et al., 2012; Young et al., 2013), however, IPO4 has not been described in the human testis. IPO4 plays various roles in cancer biology (Miao et al., 2023; Xu et al., 2019; Yang et al., 2020; Zhou et al., 2020) and it is known to exist in a complex with histones and ASF1 in cells, where it transports histones to the nucleus for chromatin assembly (Bernardes et al., 2022; Yoon et al., 2018). Due to technical limitations of the IPO4 antibody used in this thesis, quantitation of IPO4 in semen samples was not able to be performed. Therefore, the increased abundance of IPO4 in AT as suggested by the mass spec results was not able to be independently confirmed. A protein overview of IPO4 was not available in the HPA but the highest mRNA expression was found in the testis, with significantly lower but detectable expression in other organs and tissues (Uhlén et al., 2015). The RNAseq datasets given by HPA (Uhlén et al., 2015) and microarray data from GermOnline showed the highest expression of *IPO4* in 1) late, 2) early spermatids and 3) spermatocytes, suggesting an involvement in meiosis and spermiogenesis. Investigation of IPO4 localisation in the human testis by IHC revealed a stage-specific localisation during germ cell development. Its localisation in the cytoplasm of germ cells increased in intensity from late pachytenes to round spermatids. The results suggested IPO4 could be localised to the Golgi vesicle in round spermatids and later what appears as the acrosome in round and elongating spermatids, but this proposition requires further investigation. Immunocytochemical staining of IPO4 in ejaculated sperm smears would provide important information on its subcellular localisation in human sperm, and how it is altered in sperm with abnormal morphology and function. It also noted that IPO4 localisation in germ cells was not observed in SZA biopsies (with an arrest in meiosis I), despite the protein being clearly visible in spermatogonia and spermatocytes in NSP biopsies, suggesting IPO4 levels could be reduced in different infertility phenotypes. However, it is important to note that there was heterogeneity between patient biopsies in terms of

immunostaining intensities. This heterogeneity could be due to handling of the samples during and after surgery and the timing of fixation in Bouin's solution. Whether there is a true reduction in IPO4 levels in germ cells of men with defective spermatogenesis requires confirmation in a larger cohort of samples and quantitative methods to assess differences in IPO4 protein levels. Nevertheless, the localisation of IPO4 in different germ cells in NSP biopsies and the increased abundance in functionally and morphologically abnormal sperm points to an important role for IPO4 in spermiogenesis and sperm function. The localisation pattern of IPO4, particularly its high levels in early and late spermatids is consistent with the hypothesis that it could be actively involved in transporting cargoes during spermiogenesis which could be important for sperm development and function. To further understand its potential role in human spermatogenesis, identifying its cargo proteins that are transported to the nucleus during spermiogenesis will be of interest (see Chapter 4).

In conclusion, this chapter investigated alterations in the sperm proteome from men with a phenotype of asthenoteratozoospermia compared to normal semen samples. More than 2000 proteins were detected, however only 35 proteins were significantly altered, suggesting these could play key roles in sperm development and function. Of these proteins, 23 have been previously shown to be involved in sperm development and function or fertility-related conditions, highlighting the utility of this approach in identifying proteins involved in male fertility. The analyses identified 10 proteins that are likely to play novel roles in human spermiogenesis. Five proteins were then selected for further analysis. The bioinformatic analyses predicted that IPO4, ACTRT2, CCDC105, and IFT57 would be expressed in human germ cells, and IHC studies verified this hypothesis. These proteins showed distinct localisation in testis biopsies with intact and impaired spermatogenesis suggesting that they may have important functions in spermiogenesis, and that altered expression levels could be linked to abnormalities in sperm function and thus male infertility. This chapter verified that the previously uncharacterised, testis-specific protein CCDC105 is detected in human germ cells from normal spermatogenesis biopsies, however it was not evident in biopsies of other phenotypes. This chapter also verified that the epididymal protein ELSPBP1 is likely to bind to sperm during their passage through the epididymis and suggests that this protein could be taken up by the sperm in a phenotype-dependent manner. Overall, the findings point to a range of functional roles for the identified

proteins in human spermatogenesis and sperm function. The results suggest that the levels of these proteins in ejaculated sperm could be functionally important in sperm and could also be further explored for their utility as biomarkers of abnormal sperm in the ejaculate.

3.5 References

Agarwal, A., Panner Selvam, M. K., & Baskaran, S. (2020). Proteomic Analyses of Human Sperm Cells: Understanding the Role of Proteins and Molecular Pathways Affecting Male Reproductive Health. *Int J Mol Sci*, *21*(5). https://doi.org/10.3390/ijms21051621

Amaral, A., Paiva, C., Attardo Parrinello, C., Estanyol, J. M., Ballesca, J. L., Ramalho-Santos, J., & Oliva, R. (2014). Identification of proteins involved in human sperm motility using high-throughput differential proteomics. *J Proteome Res*, *13*(12), 5670-5684. https://doi.org/10.1021/pr500652y

Baker, M. A., Naumovski, N., Hetherington, L., Weinberg, A., Velkov, T., & Aitken, R. J. (2013). Head and flagella subcompartmental proteomic analysis of human spermatozoa. *Proteomics*, *13*(1), 61-74. https://doi.org/10.1002/pmic.201200350

Baldwin, M. A. (2004). Protein identification by mass spectrometry: issues to be considered. *Mol Cell Proteomics*, *3*(1), 1-9. https://doi.org/10.1074/mcp.R300012-MCP200

Barrachina, F., Battistone, M. A., Castillo, J., Mallofre, C., Jodar, M., Breton, S., & Oliva, R. (2022). Sperm acquire epididymis-derived proteins through epididymosomes. *Hum Reprod*, *37*(4), 651-668. https://doi.org/10.1093/humrep/deac015

Belardin, L. B., Antoniassi, M. P., Camargo, M., Intasqui, P., & Bertolla, R. P. (2022). Separating the chaff from the wheat: antibody-based removal of DNA-fragmented sperm. *Hum Reprod*. https://doi.org/10.1093/humrep/deac260

Bernardes, N. E., Fung, H. Y. J., Li, Y., Chen, Z., & Chook, Y. M. (2022). Structure of IMPORTIN-4 bound to the H3-H4-ASF1 histone-histone chaperone complex. *Proc Natl Acad Sci U S A*, *119*(38), e2207177119. https://doi.org/10.1073/pnas.2207177119

Bitton, D. A., Smith, D. L., Connolly, Y., Scutt, P. J., & Miller, C. J. (2010). An integrated mass-spectrometry pipeline identifies novel protein coding-regions in the human genome. *PLoS One*, *5*(1), e8949. https://doi.org/10.1371/journal.pone.0008949

Björkgren, I., & Sipilä, P. (2019). The impact of epididymal proteins on sperm function. *Reproduction*, *158*(5), R155-r167. https://doi.org/10.1530/rep-18-0589

Cao, X., Cui, Y., Zhang, X., Lou, J., Zhou, J., Bei, H., & Wei, R. (2018). Proteomic profile of human spermatozoa in healthy and asthenozoospermic individuals. *Reprod Biol Endocrinol*, *16*(1), 16. https://doi.org/10.1186/s12958-018-0334-1

Castellana, N., & Bafna, V. (2010). Proteogenomics to discover the full coding content of genomes: a computational perspective. *J Proteomics*, *73*(11), 2124-2135. https://doi.org/10.1016/j.jprot.2010.06.007

Castellana, N. E., Shen, Z., He, Y., Walley, J. W., Cassidy, C. J., Briggs, S. P., & Bafna, V. (2014). An automated proteogenomic method uses mass spectrometry to reveal novel genes in Zea mays. *Mol Cell Proteomics*, *13*(1), 157-167. https://doi.org/10.1074/mcp.M113.031260

Chalmel, F., Lardenois, A., Evrard, B., Mathieu, R., Feig, C., Demougin, P., Gattiker, A., Schulze, W., Jegou, B., Kirchhoff, C., & Primig, M. (2012). Global human tissue profiling and protein network analysis reveals distinct levels of transcriptional germline-specificity and identifies target genes for male infertility. *Hum Reprod*, *27*(11), 3233-3248. https://doi.org/10.1093/humrep/des301

Chalmel, F., Rolland, A. D., Niederhauser-Wiederkehr, C. C., S. S. W., Demougin, P., Gattiker, A., Moore, J., Patard, J., Wolgemuth, D. J., Je´gou, B., & Primig, M. (2007). The conserved transcriptome in human and rodent male gametogenesis. *PNAS*, *104*, 8346–8351. https://doi.org/10.1073/pnas.0701883104.

Cornwall, G. A. (2009). New insights into epididymal biology and function. *Hum Reprod Update*, *15*(2), 213-227. https://doi.org/10.1093/humupd/dmn055

Cravatt, B. F., Simon, G. M., & Yates, J. R., 3rd. (2007). The biological impact of mass-spectrometry-based proteomics. *Nature*, *450*(7172), 991-1000. https://doi.org/10.1038/nature06525

Cui, Z., Agarwal, A., da Silva, B. F., Sharma, R., & Sabanegh, E. (2018). Evaluation of seminal plasma proteomics and relevance of FSH in identification of nonobstructive azoospermia: A preliminary study. *Andrologia*, *50*(5), e12999. https://doi.org/10.1111/and.12999

D'Amours, O., Frenette, G., Bordeleau, L. J., Allard, N., Leclerc, P., Blondin, P., & Sullivan, R. (2012b). Epididymosomes transfer epididymal sperm binding protein 1 (ELSPBP1) to dead spermatozoa during epididymal transit in bovine. *Biol Reprod*, *87*(4), 94. https://doi.org/10.1095/biolreprod.112.100990

Dube, E., Hermo, L., Chan, P. T., & Cyr, D. G. (2008). Alterations in gene expression in the caput epididymides of nonobstructive azoospermic men. *Biol Reprod*, *78*(2), 342-351. https://doi.org/10.1095/biolreprod.107.062760

Eskandari-Shahraki, M., Prud'homme, B., Bergeron, F., & Manjunath, P. (2020). Epididymal proteins Binder of SPerm Homologs 1 and 2 (BSPH1/2) are dispensable for male fertility and sperm motility in mice. *Sci Rep*, *10*(1), 8982. https://doi.org/10.1038/s41598-020-66017-6

Feng, Y., Wang, R., Su, D., Zhai, Y., Wang, L., Yu, L., Zhang, Y., Ma, X., & Ma, F. (2021). Identifying new sperm Western blot loading controls. *Andrologia*, *53*(11), e14226. https://doi.org/10.1111/and.14226

Fietz, D., & Kliesch, S. (2023). Biopsie und Histologie des Hodens. In E. Nieschlag, H. M. Behre, S. Kliesch, & S. Nieschlag (Eds.), *Andrologie: Grundlagen und Klinik der reproduktiven Gesundheit des Mannes* (pp. 195-210). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-61901-8_11

Greither, T., Schumacher, J., Dejung, M., Behre, H. M., Zischler, H., Butter, F., & Herlyn, H. (2020). Fertility Relevance Probability Analysis Shortlists Genetic Markers for Male Fertility Impairment. *Cytogenet Genome Res*, *160*(9), 506-522. https://doi.org/10.1159/000511117

Guo, J., Grow, E. J., Mlcochova, H., Maher, G. J., Lindskog, C., Nie, X., Guo, Y., Takei, Y., Yun, J., Cai, L., Kim, R., Carrell, D. T., Goriely, A., Hotaling, J. M., & Cairns, B. R. (2018). The adult human testis transcriptional cell atlas. *Cell Res*, *28*(12), 1141-1157. https://doi.org/10.1038/s41422-018-0099-2

Gupta, S., Ghulmiyyah, J., Sharma, R., Halabi, J., & Agarwal, A. (2014). Power of proteomics in linking oxidative stress and female infertility. *Biomed Res Int*, *2014*, 916212. https://doi.org/10.1155/2014/916212

Hamada, A., Sharma, R., du Plessis, S. S., Willard, B., Yadav, S. P., Sabanegh, E., & Agarwal, A. (2013). Two-dimensional differential in-gel electrophoresis-based proteomics of male gametes in relation to oxidative stress. *Fertil Steril*, *99*(5), 1216-1226.e1212. https://doi.org/10.1016/j.fertnstert.2012.11.046

Huang da, W., Sherman, B. T., & Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*, *4*(1), 44-57. https://doi.org/10.1038/nprot.2008.211

Johnston, D. S., Wooters, J., Kopf, G. S., Qiu, Y., & Roberts, K. P. (2005). Analysis of the human sperm proteome. *Ann N Y Acad Sci*, *1061*, 190-202. https://doi.org/10.1196/annals.1336.021

Jungwirth, A., Diemer, T., Kopa, Z., Krausz, C., Minhas, S., & Tournaye, H. (2018). *EAU Guidelines on Male Infertility*. European Association of Urology.

Khan, M., Jabeen, N., Khan, T., Hussain, H. M. J., Ali, A., Khan, R., Jiang, L., Li, T., Tao, Q., Zhang, X., Yin, H., Yu, C., Jiang, X., & Shi, Q. (2018b). The evolutionarily conserved genes: Tex37, Ccdc73, Prss55 and Nxt2 are dispensable for fertility in mice. *Sci Rep*, *8*(1), 4975. https://doi.org/10.1038/s41598-018-23176-x

Khan, S. Y., Ali, M., Kabir, F., Chen, R., Na, C. H., Lee, M. W., Pourmand, N., Hackett, S. F., & Riazuddin, S. A. (2018). Identification of novel transcripts and peptides in developing murine lens. *Sci Rep*, *8*(1), 11162. https://doi.org/10.1038/s41598-018-28727-w

Klasberg, S., Bitard-Feildel, T., & Mallet, L. (2016). Computational Identification of Novel Genes: Current and Future Perspectives. *Bioinform Biol Insights*, *10*, 121-131. https://doi.org/10.4137/BBI.S39950

Leung, M. R., Zeng, J., Wang, X., Roelofs, M. C., Huang, W., Zenezini Chiozzi, R., Hevler, J. F., Heck, A. J. R., Dutcher, S. K., Brown, A., Zhang, R., & Zeev-Ben-Mordehai, T. (2023). Structural specializations of the sperm tail. *Cell*, *186*(13), 2880-2896 e2817. https://doi.org/10.1016/j.cell.2023.05.026

Li, J., Xu, X., Liu, J., Zhang, S., Wang, T., Liu, Y., & Wang, Z. (2023). The immunity-related GTPase IRGC mediates interaction between lipid droplets and mitochondria to facilitate sperm motility. *FEBS Lett*, *597*(12), 1595-1605. https://doi.org/10.1002/1873-3468.14640

Liu, Y., Guo, Y., Song, N., Fan, Y., Li, K., Teng, X., Guo, Q., & Ding, Z. (2015). Proteomic pattern changes associated with obesity-induced asthenozoospermia. *Andrology*, *3*(2), 247-259. https://doi.org/10.1111/andr.289

Llavanera, M., Delgado-Bermúdez, A., Mateo-Otero, Y., Padilla, L., Romeu, X., Roca, J., Barranco, I., & Yeste, M. (2020). Exploring Seminal Plasma GSTM3 as a Quality and In Vivo Fertility Biomarker in Pigs-Relationship with Sperm Morphology. *Antioxidants (Basel)*, *9*(8). https://doi.org/10.3390/antiox9080741

Loveland, K. L., Major, A. T., Butler, R., Young, J. C., Jans, D. A., & Miyamoto, Y. (2015). Putting things in place for fertilization: discovering roles for importin proteins in cell fate and spermatogenesis. *Asian J Androl*, *17*(4), 537-544. https://doi.org/10.4103/1008-682X.154310

Martinez-Heredia, J., Estanyol, J. M., Ballesca, J. L., & Oliva, R. (2006). Proteomic identification of human sperm proteins. *Proteomics*, *6*(15), 4356-4369. https://doi.org/10.1002/pmic.200600094

Miao, J., Yang, Z., Guo, W., Liu, L., Song, P., Ding, C., & Guan, W. (2023). Integrative analysis of the proteome and transcriptome in gastric cancer identified LRP1B as a potential biomarker. *Biomark Med*. https://doi.org/10.2217/bmm-2022-0288

Miyamoto, Y., Sasaki, M., Miyata, H., Monobe, Y., Nagai, M., Otani, M., Whiley, P. A. F., Morohoshi, A., Oki, S., Matsuda, J., Akagi, K. I., Adachi, J., Okabe, M., Ikawa, M., Yoneda, Y., Loveland, K. L., & Oka, M. (2020). Genetic loss of importin alpha4 causes abnormal sperm morphology and impacts on male fertility in mouse. *FASEB J*, *34*(12), 16224-16242. https://doi.org/10.1096/fj.202000768RR

Nathaniel, B., Whiley, P. A. F., Miyamoto, Y., & Loveland, K. L. (2022). Importins: Diverse roles in male fertility. *Semin Cell Dev Biol*, *121*, 82-98. https://doi.org/10.1016/j.semcdb.2021.08.002

Nishimura, H., Gupta, S., Myles, D. G., & Primakoff, P. (2011). Characterization of mouse sperm TMEM190, a small transmembrane protein with the trefoil domain: evidence for co-localization with IZUMO1 and complex formation with other sperm proteins. *Reproduction*, *141*(4), 437-451. https://doi.org/10.1530/rep-10-0391

Nixon, B., Cafe, S. L., Eamens, A. L., De Iuliis, G. N., Bromfield, E. G., Martin, J. H., Skerrett-Byrne, D. A., & Dun, M. D. (2020). Molecular insights into the divergence and diversity of post-testicular maturation strategies. *Mol Cell Endocrinol*, *517*, 110955. https://doi.org/10.1016/j.mce.2020.110955

Norreen-Thorsen, M., Struck, E. C., Öling, S., Zwahlen, M., Von Feilitzen, K., Odeberg, J., Lindskog, C., Pontén, F., Uhlén, M., Dusart, P. J., & Butler, L. M. (2022). A human adipose tissue cell-type transcriptome atlas. *Cell Rep*, *40*(2), 111046. https://doi.org/10.1016/j.celrep.2022.111046

Pleuger, C., Fietz, D., Hartmann, K., Schuppe, H. C., Weidner, W., Kliesch, S., Baker, M., O'Bryan, M. K., & Bergmann, M. (2017). Expression of ciliated bronchial epithelium 1 during human spermatogenesis. *Fertil Steril*, *108*(1), 47-54. https://doi.org/10.1016/j.fertnstert.2017.05.019

Pontén, F., Gry, M., Fagerberg, L., Lundberg, E., Asplund, A., Berglund, L., Oksvold, P., Bjorling, E., Hober, S., Kampf, C., Navani, S., Nilsson, P., Ottosson, J., Persson, A., Wernerus, H., Wester, K., & Uhlen, M. (2009). A global view of protein expression in human cells, tissues, and organs. *Mol Syst Biol*, *5*, 337. https://doi.org/10.1038/msb.2009.93

Qu, W., Yuan, S., Quan, C., Huang, Q., Zhou, Q., Yap, Y., Shi, L., Zhang, D., Guest, T., Li, W., Yee, S. P., Zhang, L., Cazin, C., Hess, R. A., Ray, P. F., Kherraf, Z. E., & Zhang, Z. (2020). The essential role of intraflagellar transport protein IFT81 in male mice spermiogenesis and fertility. *Am J Physiol Cell Physiol*, *318*(6), C1092-C1106. https://doi.org/10.1152/ajpcell.00450.2019

Sherman, B. T., Hao, M., Qiu, J., Jiao, X., Baseler, M. W., Lane, H. C., Imamichi, T., & Chang, W. (2022). DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Res*, *50*(W1), W216-w221. https://doi.org/10.1093/nar/gkac194

Shi, L., Zhou, T., Huang, Q., Zhang, S., Li, W., Zhang, L., Hess, R. A., Pazour, G. J., & Zhang, Z. (2019). Intraflagellar transport protein 74 is essential for spermatogenesis and male fertility in micedagger. *Biol Reprod*, *101*(1), 188-199. https://doi.org/10.1093/biolre/ioz071

Tu, C., Wang, W., Hu, T., Lu, G., Lin, G., & Tan, Y. Q. (2020). Genetic underpinnings of asthenozoospermia. *Best Pract Res Clin Endocrinol Metab*, *34*(6), 101472. https://doi.org/10.1016/j.beem.2020.101472

Uhlén, M., Fagerberg, L., Hallstrom, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, A., Kampf, C., Sjostedt, E., Asplund, A., Olsson, I., Edlund, K., Lundberg, E., Navani, S., Szigyarto, C. A., Odeberg, J., Djureinovic, D., Takanen, J. O., Hober, S., . . . Ponten, F. (2015). Proteomics. Tissue-based map of the human proteome. *Science*, *347*(6220), 1260419. https://doi.org/10.1126/science.1260419

Whiley, P. A., Miyamoto, Y., McLachlan, R. I., Jans, D. A., & Loveland, K. L. (2012). Changing subcellular localization of nuclear transport factors during human spermatogenesis. *Int J Androl*, *35*(2), 158-169. https://doi.org/10.1111/j.1365-2605.2011.01202.x

World Health Organization. (2021). *WHO laboratory manual for the examination and processing of human semen*. WHO. https://www.who.int/publications/i/item/9789240030787

Wu, Y. Q., Rao, M., Hu, S. F., Ke, D. D., Zhu, C. H., & Xia, W. (2020). Effect of transient scrotal hyperthermia on human sperm: an iTRAQ-based proteomic analysis. *Reprod Biol Endocrinol*, *18*(1), 83. https://doi.org/10.1186/s12958-020-00640-w

Xu, X., Zhang, X., Xing, H., Liu, Z., Jia, J., Jin, C., & Zhang, Y. (2019). Importin-4 functions as a driving force in human primary gastric cancer. *J Cell Biochem*, *120*(8), 12638-12646. https://doi.org/10.1002/jcb.28530

Yang, F., Liu, H., Zhao, J., Ma, X., & Qi, W. (2020). POLR1B is upregulated and promotes cell proliferation in non-small cell lung cancer. *Oncol Lett*, *19*(1), 671-680. https://doi.org/10.3892/ol.2019.11136

Yoon, J., Kim, S. J., An, S., Cho, S., Leitner, A., Jung, T., Aebersold, R., Hebert, H., Cho, U. S., & Song, J. J. (2018). Integrative Structural Investigation on the Architecture of Human Importin4_Histone H3/H4_Asf1a Complex and Its Histone H3 Tail Binding. *J Mol Biol*, *430*(6), 822-841. https://doi.org/10.1016/j.jmb.2018.01.015

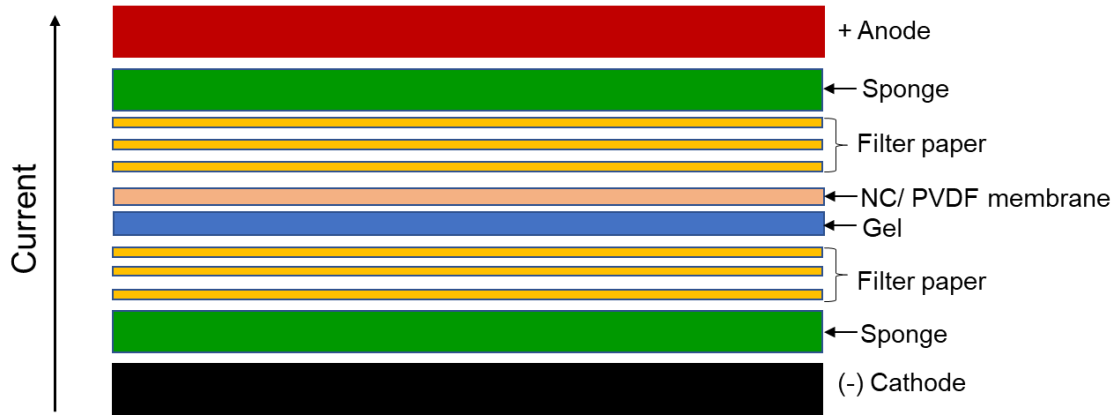
Young, J. C., Ly-Huynh, J. D., Lescesen, H., Miyamoto, Y., Browne, C., Yoneda, Y., Koopman, P., Loveland, K. L., & Jans, D. A. (2013). The nuclear import factor importin alpha4 can protect against oxidative stress. *Biochim Biophys Acta*, *1833*(10), 2348-2356. https://doi.org/10.1016/j.bbamcr.2013.06.007

Zhang, S., Liu, Y., Huang, Q., Yuan, S., Liu, H., Shi, L., Yap, Y. T., Li, W., Zhen, J., Zhang, L., Hess, R. A., & Zhang, Z. (2020). Murine germ cell-specific disruption of Ift172 causes defects in spermiogenesis and male fertility. *Reproduction*, *159*(4), 409-421. https://doi.org/10.1530/REP-17-0789

Zhang, X. Z., Wei, L. L., Zhang, X. H., Jin, H. J., & Chen, S. R. (2022). Loss of perinuclear theca ACTRT1 causes acrosome detachment and severe male subfertility in mice. *Development*, *149*(12). https://doi.org/10.1242/dev.200489

Zhou, Y., Liu, F., Xu, Q., Yang, B., Li, X., Jiang, S., Hu, L., Zhang, X., Zhu, L., Li, Q., Zhu, X., Shao, H., Dai, M., Shen, Y., Ni, B., Wang, S., Zhang, Z., & Teng, Y. (2020). Inhibiting Importin 4-mediated nuclear import of CEBPD enhances chemosensitivity by repression of PRKDC-driven DNA damage repair in cervical cancer. *Oncogene*, *39*(34), 5633-5648. https://doi.org/10.1038/s41388-020-1384-3

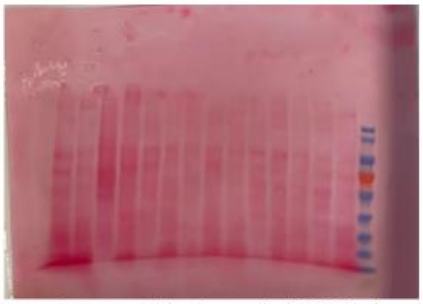
3.6 Supplementary Figures



Supplementary Figure 3.1: Schematic figure showing the arrangement in the transfer apparatus.

Gel plan for ACTRT2

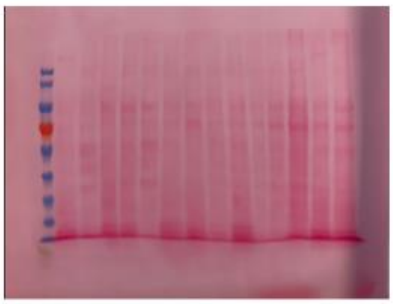
M	P 1 N 30	P2 N 30	P3 N 30	P 4 N 30	P5 N 30	P1 AT 30	P 2 AT 30	P3 AT 30	P4 AT 30	P5 AT 30	P Con 30	Tes 30	P1 N 30	P1 AT 15
	ACTRT2 – 1:1500 and Alpha Tub (1:5000)												NC	



Ponceau staining image of a ACTRT2 gel

Gel plan for ELSPBP1

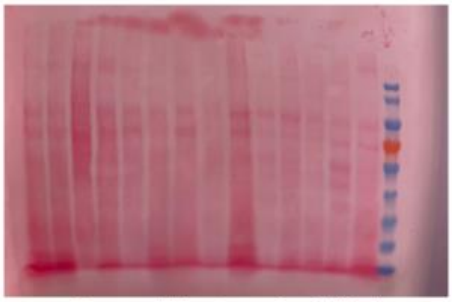
M	P 1 N 30	P2 N 30	P3 N 30	P 4 N 30	P5 N 30	P1 AT 30	P 2 AT 30	P3 AT 30	P4 AT 30	P5 AT 30	P Con 30	Tes 30	P1 N 30	P1 AT 30
	ELSPBP1 – 1:2000 and Alpha Tub (1:5000)												NC	



Ponceau staining image of an ELSPBP1 gel

Gel plan for IFT57

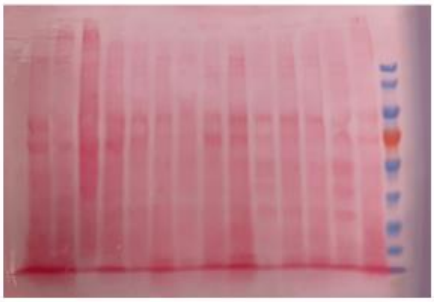
M	P 1 N 30	P2 N 30	P3 N 30	P 4 N 30	P5 N 30	P1 AT 30	P 2 AT 30	P3 AT 30	P4 AT 30	P5 AT 30	P Con 30	Tes 30	P1 N 30	P1 AT 30
	IFT57 – 1:500 and Alpha Tub (1:5000)												NC	



Ponceau staining image of an IFT57 gel

Gel plan for CCDC105

M	P 1 N 30	P2 N 30	P3 N 30	P 4 N 30	P5 N 30	P1 AT 30	P 2 AT 30	P3 AT 30	P4 AT 30	P5 AT 30	P Con 30	Tes 30	P1 N 30	P1 AT 30
	CCDC105 – 1:500												NC	



Ponceau staining image of a CCDC105 gel

Supplementary Figure 3.2: Gel plans and respective Ponceau images of ACTRT2, ELSPBP1, IFT57, and CCDC105. The NC membranes showing Ponceau staining can be reverse imaged orientation for photos, thus will not correspond to the gel plan as shown. NC negative control

Chapter 4

Significance of importin 4 in human male fertility

4.1 Introduction

Importins are a type of karyopherin, which are proteins involved in the active transportation of cargo proteins from the cytoplasm to the nucleus. Importins can be broadly categorised as alpha and beta importins based on structural and functional features. Seven importin alpha proteins and nearly 20 importin beta proteins have been identified in humans (Harel & Forbes, 2004; Kelley et al., 2010; Whiley et al., 2012). Many studies have shown that some importins play functional roles in spermatogenesis, see (Nathaniel et al., 2022; Young et al., 2013) for reviews. Studies in rodents have revealed that some importins play a role in the normal development and morphology of sperm and thus possibly impacting male fertility (Miyamoto et al., 2020). This Chapter focuses on importin 4, an importin beta protein that was detected in high abundance in morphologically abnormal ejaculated spermatozoa in Chapter 3. IPO4 has not been well studied in the human testis yet.

The transport of proteins between the cytoplasm and nucleus can be both active and passive; molecules smaller than 40kDa can diffuse through the nuclear pore complex (NPC) whereas molecules larger than 40kDa require an active transportation mechanism. Active transportation requires a transporter protein that recognises its cargo. These transporters are collectively known as nuclear transporter receptors (NTR) that can identify localisation signals on the cargo proteins requiring transport between the cytoplasm and nucleus (Mackmull et al., 2017). These NTRs include importin alphas (karyopherin alpha) or betas (karyopherin beta) (Mackmull et al., 2017). Karyopherin beta acts as a transporter that mediates most of the active transportation between the cytoplasm and nucleus in both directions. The karyopherin beta family is comprised of importins and exportins (Wing et al., 2022). The terms importin and exportins are coined as per the direction of the cargo transportation, with importins transporting cargos from cytoplasm to nucleus, and exportins transporting in the opposite direction. Transporter proteins recognise the respective cargo by a nuclear localisation signal (NLS) or nuclear export signal (NES) sequences of the

cargo, that consist of one or two motifs of 8-10 basic amino acids. Since the transportation and importin-cargo binding depends on the recognition of NLS motifs, recent studies have focused on identifying these motifs through various approaches (Panagiotopoulos et al, 2022). This Chapter mainly focuses on the nuclear importation of cargo proteins.

Although not fully understood, the recognition of the cargos by the NTRs is believed to occur via different mechanisms. Importin beta can bind to its cargos directly or in some cases via the mediation of alpha importins as an adaptor. The importin alpha family proteins contain 10 armadillo repeat motifs (ARM, a protein domain with 40 residues and a helical structure) and an importin beta binding (IBB) domain which can act as an adaptor to importin beta by assisting the binding of importin beta to its cargo (Mackmull et al., 2017). Importin alpha recognises the NLS of the cargo, while importin beta binds to alpha through the IBB and is responsible for binding to the NPC and movement through the pore. Recognition of cargos can also be achieved by importin beta alone. Importin beta proteins contain repeated motifs called HEAT repeats (19 – 24 repeats, structural motifs with two anti-parallel alpha helices connected by a loop) in a spherical structure (Wing et al., 2022). These motifs undergo structural changes upon binding to the cargo, phenylalanine glycine nucleoproteins (FG-nups) in the pore of NPC and RAN-GTP and assist in protein recognition (Loveland et al., 2015; Major et al., 2011). The groves and loops of the helical structure of HEAT repeats are used for this recognition (Wing et al., 2022). Upon binding of the cargo protein to the importin beta subunit, the complex moves from the cytoplasm through the NPC to the nucleus by binding to the NPC. The recognition and binding of RAN-GTP by HEAT repeats enables the dissociation of the cargo inside the nucleus (Fig.4.1). Another important factor that regulates transportation is the concentration gradient of RAN-GTP that maintains high levels of RAN-GTP in the nucleus and RAN-GDP in the cytoplasm (Major et al., 2011). Yet, the precise mechanisms involving recognition of each cargo by the respective importin are not fully understood (Mackmull et al., 2017).

The direction of importin-mediated protein transportation is governed by the GDP- or GTP-bound version of the RAN protein in the cytoplasm and nucleus, respectively. The asymmetric proportions, or gradient, of these versions in the cytoplasm and nucleus are important for both the binding and release of the cargo protein by importins. In the cytoplasm, importins recognise and bind to cargoes in the presence

of RAN-GDP or, more precisely, in the absence of RAN-GTP. Proteins such as RanGTPase activating protein (RANGAP) and RAN binding protein (RanBP1) in the cytoplasm assist in maintaining a GDP-bound RAN state in the cytoplasm by activating GTP hydrolysation (Wing et al., 2022). This enables the dimeric importin complex to bind to its cargo protein in the cytoplasm which then enters the nucleus meeting RAN-GTP there. These mechanisms ensure the presence of RAN-GTP inside the nucleus by hindering the hydrolysis of GTP to GDP. For example, the regulator of chromosome condensation protein (RCC1) regulates the high levels of RAN-GTP inside the nucleus by binding to histones in nuclear chromatin which enables the dissociation of GDP from RAN and promoting GTP binding to RAN (Kierszenbaum et al., 2002). When the importin-NLS-cargo protein complex enters the nucleus, it binds to RAN-GTP, which promotes the release of the cargo due to the conformational changes that take place in the importin beta subunit upon the formation of the new importin-RAN-GTP complex. After the release of the cargo, the importin-RAN-GTP complex moves back to the cytoplasm again. The hydrolysation of RAN-GTP to RAN-GDP by RanGTPase in the cytoplasm releases importin, enabling it to transport more molecules to the nucleus (Fig.4.1) (Kierszenbaum et al., 2002; Loveland et al., 2015).

Some importin subunits have been shown to play key roles in spermatogenesis. A potential functional relevance of importin 4 (IPO4), a beta importin protein, was identified in Chapter 3. Mass spectrometry analyses of human spermatozoa was used to identify proteins that were differentially expressed in human sperm with a normal phenotype (NORM) compared to sperm with a phenotype of impaired morphology and motility (aka asthenoteratozoospermia, AT) (Chapter 3). Despite more than 2700 proteins being identified, only 35 proteins were significantly altered in AT compared to NORM. It is thus possible that these 35 proteins could be important for sperm formation and motility and/or could contribute to the altered sperm phenotype in AT patients. The results in Chapter 3 demonstrated that IPO4 was significantly higher in abundance (3.5-fold, $p=0.03$) in AT sperm compared to those from normozoospermic patients.

An investigation of IPO4 function in various mouse and human databases suggested a possible involvement in spermatogenesis (Chapter 3). According to the Human Protein Atlas (HPA), IPO4 is expressed at a low level in many tissues but shows the highest protein expression in the testis (Pontén et al., 2009). In isolated human

testicular cells, *IPO4* expression is lowest in spermatogonia and increases as spermatogenesis progresses, with highest expression in round and elongated spermatids (Uhlén et al., 2015). IPO4 therefore could participate in the nuclear import of proteins during spermatogenesis, and perhaps spermiogenesis in particular. The predominant expression of *IPO4* in germ cells compared to testicular somatic cells was confirmed by a previous study where *IPO4* expression in haploid cells was shown to be 38 times higher than in diploid cells in mice (Pradeepa et al., 2008). A search of the Mouse Genome Informatics (MGI) database revealed that a whole-body knockout (KO) of *Ipo4* results in a complete penetrance of prenatal weaning phenotype, abnormal embryo size and embryonic growth retardation. KO pups die after weaning and therefore fertility has not been assessed in the complete absence of *Ipo4*. The localisation of IPO4 to germ cells in humans has not yet been independently confirmed.

The immunohistochemical (IHC) localisation of IPO4 in normal human spermatogenesis biopsies (NSP, n=5) (Chapter 3) revealed that IPO4 is present in the nuclei of diploid germ cells (light staining in spermatogonia and early spermatocytes). However, during meiosis the protein was also observed in the cytoplasm, with IPO4 immunostaining visible in both the nucleus and cytoplasm in pachytene spermatocytes and round spermatids. These results suggest that IPO4 could participate in cytoplasmic to nuclear protein transport during late meiosis and early spermiogenesis. Sertoli cells also showed weak immunohistochemical staining in both cytoplasm and nuclei. In spermatid arrest (SDA, n=3) samples, the localisation of the protein in cytoplasm of post-meiotic cells was not evident, even though these samples did contain a range of round spermatids that showed IPO4 staining in nuclei. This observation suggests that IPO4 may not translocate as effectively from the nucleus to the cytoplasm (after the release of the cargo inside the nucleus, for recycling) which could influence IPO4-mediated protein transportation, by accumulating in the nucleus. This would also mean that IPO4 is not recycled to the cytosol to transport other important cargos to the nucleus. In addition, the higher abundance of IPO4 in AT sperm measured by mass spectrometry also suggests an alteration in IPO4 function which could be associated with infertility.

The altered localisation of IPO4 round spermatids in SDA biopsies and from AT sperm observed in Chapter 3 led to the hypothesis that IPO4 may regulate the nuclear import

of key proteins required for spermatogenesis/ spermiogenesis and normal sperm morphology and function. However, no publications are available to date that have investigated potential functions of IPO4 in mouse or human spermiogenesis. An understanding of the potential functional role of IPO4 in spermiogenesis relies on identifying the cargo proteins it may transport. This Chapter investigates IPO4 and its potential cargo proteins in normal and impaired spermatogenesis. Given that GDT/GTP bound RAN is involved in importin functions, this Chapter also investigated a novel RAN isoform in human spermatogenesis.

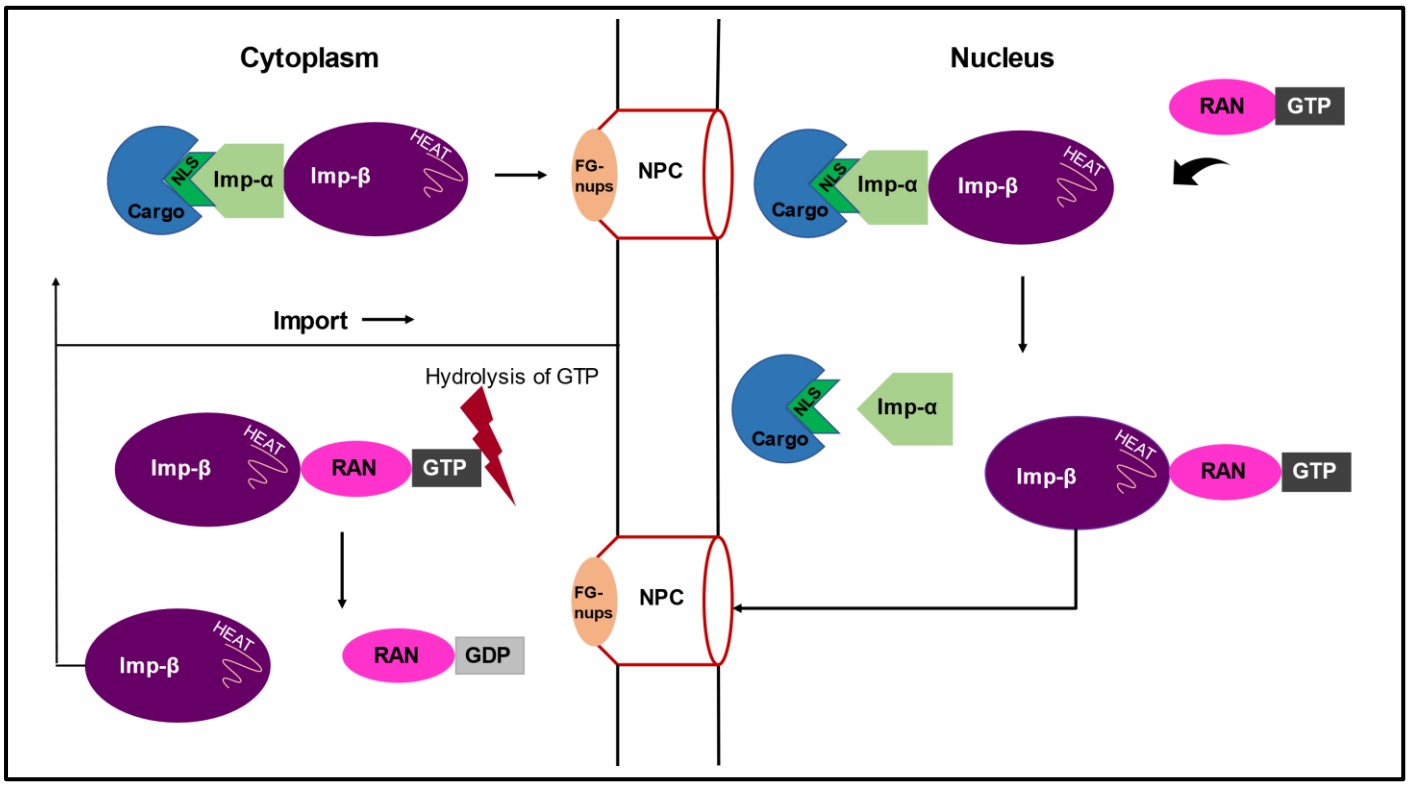


Figure 4.1: Nucleocytoplasmic transportation of cargoes by importins beta. The diagram depicts the import of cargoes from the cytoplasm to the nucleus by importins. The arrows show the direction of events. Briefly, cargo proteins bind to the importin α-β complex and are actively transported from the cytoplasm to nucleus through the NPC. Inside the nucleus, RAN-GTP binds to Imp-β which causes the cargo to dissociate and be released into the nucleus. The Imp β – RAN-GTP complex then moves back to the cytoplasm through the NPC and RAN-GTP is hydrolysed to RAN-GDP releasing Imp β. NLS – nuclear localisation signal, Imp-α- importin alpha, Imp-β – importin beta, HEAT- structural motifs with two anti-parallel alpha helices connected by a loop, found in importin beta, NPC – nuclear pore complex, FC-nups – phenylalanine glycine nucleoproteins present in the core of NPC, RAN – nuclear protein Ran, GDP – Guanosine diphosphate, GTP – Guanosine triphosphate,

4.1.1 Hypothesis:

- That IPO4 is involved in human spermiogenesis and in normal sperm development, function and motility, and the subcellular localisation of IPO4 could be important for optimal sperm function.

4.1.2 Research Objectives

4.1.2.1 General Objective:

To investigate the localisation of importin 4 and target cargo proteins in human spermiogenesis.

4.1.2.2 Specific Objectives:

1. To determine the cellular and sub-cellular localisation of IPO4 in human testis tissue samples (NSP and SDA) by immunofluorescence.
2. To identify potential binding partners/cargos of IPO4 and whether they are expressed in human round spermatids.
3. To validate whether candidate cargo proteins co-localise with IPO4 in human testis tissue.
4. To determine whether the co-localisation of IPO4 and candidate cargo proteins changes in testis biopsies with SDA.

4.2 Material and Methods

4.2.1 Sample collection

Testis biopsies were collected at the Clinic of Urology, Paediatric Urology and Andrology, Justus-Liebig University of Giessen, Germany and the Centre of Reproductive Medicine and Andrology (CeRA), University Clinic Munster, Germany. Testicular biopsies were indicated because of obstructive or non-obstructive azoospermia diagnosed by clinical examination, ultrasound, hormone status, and ejaculate analysis (Jungwirth et al., 2018; World Health Organization, 2021). Surgeries were performed after written informed consent and after approval by local Ethics Committee (AZ 26/11). Biopsies for histological evaluation were fixed with Bouin's solution overnight, cut, and stained with hematoxylin and eosin following standard

protocols, and evaluated following score count analysis according to latest standard protocols (Bergmann & Kliesch, 2010; Fietz & Kliesch, 2023). Samples used in this study were classified as intact spermatogenesis (normal spermatogenesis, n=5; contains germ cells up to elongated spermatids, i.e. testicular sperm and somatic cells; score = 10;), spermatid arrest (SDA, n=3; contains germ cells up to round spermatid with occasionally observed elongating spermatids only in some tubules, and somatic cells; score = 4-5), and Sertoli cell only (SCO, n=5; complete absence of germ cells; score =0). All following experiments were performed by the PhD candidate.

4.2.2 *In silico* analysis to identify candidate cargo proteins of IPO4

A thorough search was conducted using published literature and online databases to gather information on probable cargo proteins that could be transported by IPO4. Briefly, data were gathered from previously published papers on importins and predicted information in portals such as UniProt, The Human Protein Atlas (HPA), Ensembl, GeneCards, and STRING interaction network. Firstly, published literature on candidate cargoes and importin were gathered. Data from GeneCards was used to investigate transcription factors that could interact with IPO4 and the STRING interaction image was used to understand the possible binding links of IPO4 with other proteins. A list of spermatid-enriched genes was retrieved from HPA to investigate whether candidate IPO4 cargos are expressed in human spermatids (further explained in the results section). RNAseq and Mass spec datasets as detailed in Chapters 2 and 3 were used to investigate the expression levels of selected genes and proteins related to nucleocytoplasmic transportation. Two proteins that fulfilled these criteria (ASF1B and TNP1) were selected for further assessment based on the availability of antibodies for immunofluorescent assays. Information regarding RAN protein structure was retrieved from the NCBI and SWISS-MODEL databases.

4.2.3 Immunofluorescence (IF)

IF was performed at the Centre for Reproductive Health, Hudson Institute of Medical Research, Melbourne University, Victoria, Melbourne. Paraffin-embedded tissue slides were cut at 5 µM thickness and baked at 60°C for 20 minutes to facilitate the de-paraffinisation and the rehydration was done by immersing the slides twice in 100% Histosol solutions for 10 minutes and then in a gradient of alcohol series at 100%, 80% and 70% on a shaker plate. Finally, slides were kept for 5 minutes in distilled water. Antigen retrieval was facilitated by boiling the slides in Citrate buffer plus (pH 6). Triton X (0.1%, Sigma-Aldrich) was added to the Citrate buffer solution to increase permeability. Slides were boiled in a microwave at 800W for 4 minutes followed by 450W for 6 minutes and slides were allowed to cool for 20 minutes. Slides were then washed thrice in PBST (Phosphate buffered saline + 0.1% Triton X), and unspecific binding was blocked using 1.5% Bovine Serum Albumin (BSA, Sigma Aldrich) dissolved in PBST solution, and 5% donkey serum (Sigma Aldrich) for 1 hour at room temperature. The blocking solution was tapped off and the respective primary antibody diluted in 1% PB (1% BSA in 1X PBS, PB = PBS-BSA) was added in the optimised dilution (Table 4.1), and 1% PB was used on the negative control tissue sections. Slides were incubated at 4°C for overnight in a humid dark chamber. The next day, the primary antibody was tipped off and slides were washed thrice in PBST solution on a shaker plate. The respective secondary antibody (Table 4.1) in the optimised dilution (diluted in 1% PB) was applied for 1 hour in a humidified dark chamber at room temperature. Slides were washed (in a dark container) 4 times in PBST and 4',6-Diamidin-2-phenylindol (DAPI, Invitrogen) diluted in 1:1000 with 1X PBS (Gibco PBS pH 7.4) was added and incubated for 10 minutes. Slides were washed again 4 times in PBST. A few drops of Sudan Black (Sigma) were added for 30 seconds, the sections were then washed with distilled water and slides were kept in distilled water until mounting. Sides were mounted using ProLong Gold (ProLong™ Gold antifade reagent, Invitrogen), dried and imaged on a VSI 120 fluorescent microscope (Olympus Life Science).

Table 4.1: Antibodies used for immunofluorescence assays

No.	Primary antibody		Secondary antibody	
	Manufacturer details	**Dilution (diluent – 1% PB)**	**Manufacturer details**	**Dilution (diluent – 1% PB)**
01	Polyclonal Anti-IPO4 antibody Sigma life science Prestige antibodies provided by Atlas antibodies. (HPA039043)	1:200	Alexa Fluor™ 555 Donkey anti-Rabbit IgG (H+L) Invitrogen by Thermo Fisher Scientific	1:500
02	Monoclonal Anti-ASF1B Sigma -Aldrich (SAB1405157)	1:200	Alexa Fluor™ 488 Donkey anti-Mouse IgG (H+L) Invitrogen by Thermo Fisher Scientific	1:500
03	Monoclonal Anti-TNP1 Sigma-Aldrich (WH0007141M1)	1:400	Alexa Fluor™ 488 Donkey anti-Mouse IgG (H+L) Invitrogen by Thermo Fisher Scientific	1:500

4.3 Results

4.3.1 Identification of putative cargo proteins of IPO4 in human spermiogenesis

During spermiogenesis, spermatids actively transcribe genes important for spermatid development, however transcription ceases as the spermatid nucleus starts to condense and the chromatin becomes highly compacted. Therefore, the GeneCards database was interrogated (Fishilevich et al., 2017) for transcription factors that contain binding sites for IPO4. This analysis identified 328 transcription factors that are predicted to bind to IPO4. Next, genes known to be enriched in human round spermatids according to a dataset in the Tissue Cell tab of Human Protein Atlas were identified (Norreen-Thorsen et al., 2022) and n= 951 genes were deemed to be enriched in human spermatids compared to other testicular cells. Further, experimental data from UniProtKB and Swiss-Prot showing an association with IPO4 or an association with protein condensation in spermatids were considered; this identified various potential cargo proteins such as histones, RPS3A, VDR and TNP2 (Jäkel et al., 2002; Miyauchi et al., 2005; Pradeepa et al., 2008) as well as histone H3-H4 dimer in a complex with ASF1A or ASF1B (Alvarez et al., 2011; Yoon et al., 2018).

Using all available information, RPS3A, ASF1A, ASF1B, TNP1, TNP2, CDAN1, and DDX20 (Table 4.2) were selected for further investigation based on a) the potential

affinity to importin beta based on publicly available information (see Materials and Methods) and b) the mRNA being highly enriched in, or specific to, human round spermatids. This information indicated these proteins were putative IPO4 cargo proteins in human round spermatids (Table 4.3). The availability of antibodies to investigate the protein localisation in human paraffin-embedded testis biopsies was then assessed.

Based on the availability of suitable antibodies, ultimately two proteins were selected for further analysis: 1) ASF1B based on a) it has been proven to bind to histone H3-H4 dimers, and histones are predicted to bind to IPO4 based on STRING interactions, and b) it is primarily expressed in human germ cells, from spermatocytes to late spermatids; and 2) TNP1 is a) predicted to bind to IPO4 based on previous findings showing a high affinity towards importin betas, and b) shows mRNA expression and protein localisation in spermatids during spermiogenesis. Both TNP1 and TNP2 were of interest because they are involved in remodeling of spermatid chromatin during the elongation phase of spermiogenesis and are therefore considered essential for spermiogenesis (Pradeepa et al., 2008; Steger et al., 1998). TNP1 was selected for further assessment due to the availability of suitable antibodies.

ASF1B was chosen for further assessment due to its localisation in spermatids and its likely role in spermiogenesis. ASF1B is a histone chaperone that regulates nucleosome assembly and disassembly by depositing, removing or exchanging histones as reported in UniProt based on published data. ASF1B is also involved in the nuclear import of the histone H3-H4 dimer with IPO4 (Alvarez et al., 2011; Campos et al., 2010; Campos et al., 2015). According to HPA and GeneCards, ASF1B is localised to the nucleoplasm. Lymphoid organs, testis, and germ cells reveal an enrichment of *ASF1B* mRNA according to human testis RNAseq data in HPA. The highest mRNA expression in testicular cells was detected in spermatocytes followed by spermatogonia, late spermatids and early spermatids (Uhlén et al., 2015). The mRNA transcripts for many proteins important for spermatid elongation and development are first transcribed in late spermatocytes and early spermatids (Steger et al., 1998), thus the mRNA expression pattern fit to a protein involved in spermatid elongation. The expression pattern of both ASF1B mRNA and protein in germ cells and the data predicting it could be a cargo of IPO4 suggested that ASF1B could be an

excellent candidate to further investigate as a potential cargo protein of IPO4 during human spermiogenesis.

TNP1 was also chosen for assessment as a potential cargo of IPO4 in human spermatids. As round spermatid nuclei start to condense during the elongation phase of spermiogenesis, the spermatid chromatin is remodelled via the replacement of histones with transition proteins (TNP1 and TNP2) and then finally protamines (Pradeepa et al., 2008; Steger et al., 1998; Zhao et al., 2004). Thus, the nuclear localisation of TNP1 and 2 are essential for spermiogenesis and involvement in male fertility has been shown for TNP1 previously (Jedrzejczak et al., 2007; Pradeepa et al., 2008); yet it is not known whether TNP1 is translocated to the nucleus by IPO4.

According to HPA, the TNP1 protein is specific to testis tissue and human testis RNAseq data shows that the mRNA is specific to spermatids (highest in late spermatids followed by early spermatids (Table. 4.3) (Uhlén et al., 2015). The subcellular localisation reported in GeneCards demonstrating its presence in both cytoplasm and nucleus but with a predominance in the nucleus. The high specificity of the protein in testis, mRNA enrichment in spermatids and the experimental finding of a high affinity of TNP1 to importin betas (Pradeepa et al., 2008), suggested that TNP1 is a putative cargo protein of IPO4 during human spermiogenesis.

Table 4.2: Selected putative cargo proteins of IPO4

Gene name	UniProt ID	Protein	Function	Detected in Mass Spec of ejaculated spermatozoa (Chapter 3)
RPS3A	P61247	40S ribosomal protein S3a	May play a role during erythropoiesis through regulation of transcription factor DDIT3.	Not detected
ASF1B	Q9NVP2	Anti-silencing function 1B histone chaperone	Required for spermatogenesis. This gene encodes a member of the H3/H4 family of histone chaperone proteins and is similar to the anti-silencing function-1 gene in yeast. May play a key role in modulating the nucleosome structure of chromatin by ensuring a constant supply of histones at sites of nucleosome assembly	Not detected
TNP2	Q05952	Transition protein 2	Plays a key role in the replacement of histones to protamine in elongating spermatids of mammals. In condensing spermatids, it is loaded onto the nucleosomes and promotes the recruitment and processing of protamines, which are responsible for histone eviction	Not detected
TNP1	P09430	Transition protein 1	Plays a key role in the replacement of histones to protamine in the elongating spermatids of mammals. In condensing spermatids, it is loaded onto the nucleosomes and promotes the recruitment and processing of protamines, which are responsible for histone eviction. Transition protein-1 is a spermatid-specific product of the haploid genome which replaces histone and is itself replaced in the mature sperm by the protamines (Gene summary by OMIM)	Not detected
ASF1A	Q9Y294	Anti-silencing function 1A histone chaperone	Histone chaperone that facilitates histone deposition and histone exchange and removal during nucleosome assembly and disassembly	Not detected
CDAN1	Q8IWY9	Codanin 1	May act as a negative regulator of ASF1 in chromatin assembly. Gene belongs to ''testis-chromatic organization' cluster.	Detected
DDX20	Q9UHI6	DEAD-box helicase 20 (HPA). Probable ATP-dependent RNA helicase DDX20 (UniProt)	The SMN complex plays a catalyst role in the assembly of small nuclear ribonucleoproteins (snRNPs), the building blocks of the spliceosome. Thus plays a role in the splicing of cellular pre-mRNAs	Not detected

Function – information gathered from public databases such as UniProt, GeneCards, OMIM (see Materials and Methods).

Table 4.3: Data collected of the selected putative cargo proteins of IPO4

Gene name	UniProt ID	HPA and GeneCards				
		Tissue enhanced (Protein)	**Cell/s enhanced**	**RNA**	**Subcellular localisation**	**IHC**
RPS3A	P61247	Low specificity	Low specificity	L.Sptd>E.Sptd>Spc>Spg	Localised to the Nucleoli (enhanced), Endoplasmic reticulum (enhanced), Cytosol (enhanced)	Medium - Leyd > Elo.Sptd-Spg
ASF1B	Q9NVP2	Lymphoid/ testis	Spermatogonia , Spermatocytes, Erythroid cells, Early spermatids	Spc>Spg>L.Sptd>E. Sptd	Nucleoplasm (HPA) GeneCards – nucleus – 5, cytoplasm - 4	Medium - Spg. Low - r/e Sptd. Others not detected (HPA069385)
TNP2	Q05952	Testis (only)	Late Spermatids	L.Sptd>>>E.Sptd	HPA - Not given. GeneCards - nucleus 5, Cytosol - 2	Not shown
TNP1	P09430	Testis (only)	Late Sptd, E. Sptd	L.Sptd>>>E.Sptd	HPA-not given. GeneCards - Nucleus - 5, Cytosol - 2	High - E/L. Sptd. Shows stage specificity
ASF1A	Q9Y294	Low specificity	Low specificity	Spg>E.Sptd>Spc>L.Sptd	Nucleoplasm	Not shown
CDAN1	Q8IWY9	Low specificity	Sertoli (RNA)	Sert>>Spc>Spg>L.Sptd>E. Sptd	Cytosol	High- Leyd, r/e. Sptd, Medium - El. Sptd. Not detected - Sertoli, Spg, other Spc
DDX20	Q9UHI6	Testis (highest)	r/early spermatids	E. Sptd>L.Sptd>Spc>Spg	Nucleoplasm, nuclear bodies, cytosol (HPA)	Medium-r/e.Sptd>Elo.Sptd

Abbreviations - Sptd – spermatids, E. Sptd –early spermatids, r/e – round/early, L. Sptd – late spermatids, Spc – spermatocytes, Spg – spermatogonia, Leyd – Leydig cells, Elo. Sptd – elongating spermatids. Sert – Sertoli cells, HPA – Human Protein Atlas, IHC - Immunohistochemistry

RNAseq data - (Guo et al., 2018; Norreen-Thorsen et al., 2022; Uhlén et al., 2015). Subcellular localisation - predicted subcellular locations listed in HPA and GeneCards. IHC – gathered information after analysing IHC images of testis tissues published in HPA.

4.3.2 Detection of IPO4 and ASF1B in human testis biopsies by immunofluorescence.

Before establishing the co-localisation of IPO4 and its putative cargo protein ASF1B using dual-label immunofluorescence, the immunofluorescent (IF) detection of IPO4 and ASF1B in testis sections was first optimised using NSP biopsies. Compared to conventional light immunohistochemistry, immunofluorescence is more applicable to investigating the subcellular localisation of proteins, such as assessing nuclear vs cytoplasmic localisation. A stage-specific immunofluorescent signal for IPO4 was observed in germ cells, particularly those located towards the lumen of the tubule, with both nuclear and cytoplasmic staining observed (Fig. 4.2 A-D); this cell-specific localisation of IPO4 is discussed in more detail below. The respective negative control was unstained (Fig. 4.2 E).

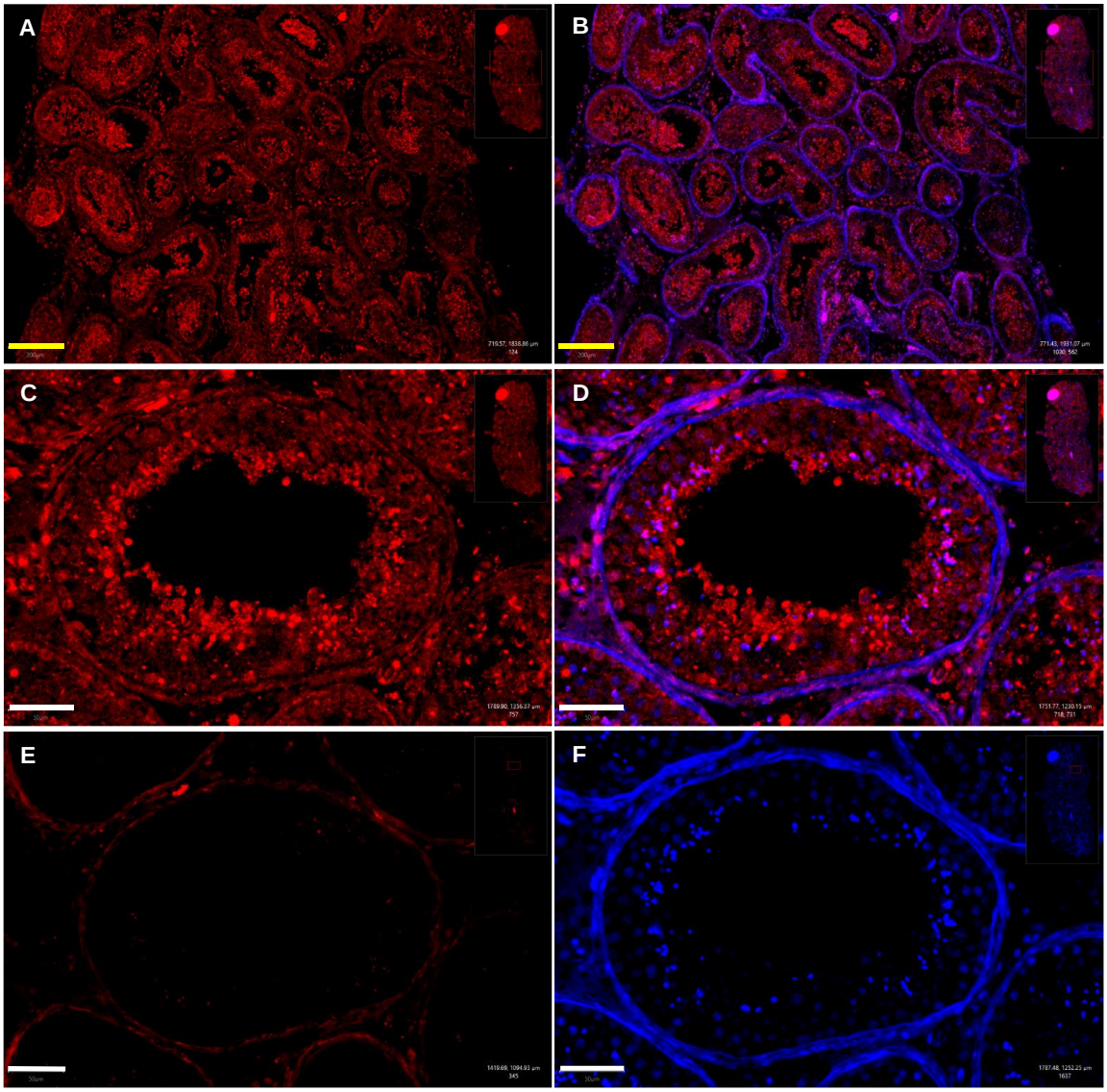


Figure 4.2: A general view of localisation of IPO4 with DAPI staining in NSP biopsies. Immunostaining was detected using Alexa Fluor 555 secondary antibody and DAPI. Slides were scanned with a VS 120 scan microscope, signal detected using QuPath software and representative images are shown. NSP – normal spermatogenesis. *Inset:* the total tissue scanned. **A** – low magnification of NSP tubules immunostained for IPO4. **B** –low magnification of NSP tubules with IPO4 and DAPI staining. **C** – a higher magnification of a tubule stained in IPO4. **D** – a higher magnification of a NSP tubule with IPO4 and DAPI staining. **E**- the negative control of tubule C. **F**– DAPI staining of tubule C. The yellow scale bars are 200μM and the white scale bars are 50μM.

4.3.2.1 The expression of IPO4 in different testis pathologies

The immunofluorescent localisation of IPO4 in biopsies of different testis pathologies was first assessed by investigating the single channel fluorescent images of IPO4. In dual-labelled NSP biopsies, IPO4 was detected as a faint signal in the nuclei of spermatogonia in all stages. However, the signal was comparatively stronger in spermatogonia nuclei in stages V and VI (Fig. 4.3 D yellow arrow, Sg). Most pachytene spermatocyte nuclei also showed a faint signal in stages I-VI (Fig. 4.3 D, white arrow, PS). Pachytene spermatocytes in stages IV-V also showed light cytoplasmic staining (Fig. 4.3 D yellow arrow, PS). In early round spermatids, IPO4 was localised in the Golgi apparatus and in the acrosome as it developed and spread across the nucleus in spermatids at steps 3 and 4 of development (Fig. 4.3, C-F yellow arrows). Focal dots of IPO4 were also observed in the cytoplasm of round spermatids (Fig. 4.3 F pink arrow, Cyto). Some round spermatids in steps 1-3 showed focal dots of IPO4 in the nucleus Fig. 4.3 E yellow arrow, rSptd) (Suppl. Table 4.1).

The localisation of IPO4 in germ cells was then assessed in biopsies with different pathologies. In SDA biopsies (Fig. 4.4), the immunolocalisation pattern was broadly similar to NSP. Focal areas of IPO4 were observed in spermatogonia's nuclei and, to some extent, cytoplasmatic localisation of IPO4 appeared in spermatogonia (Fig. 4.4 C, yellow arrow Sg and Fig. 4.4 D pink arrow, Cyto). The pattern of nuclear localisation was comparable to that observed in NSP, with IPO4 signal observed in spermatogonia (Fig. 4.4 C yellow arrow – Sg), a light signal in pachytene spermatocytes (Fig. 4.4 C, D yellow arrows, PS and Fig. 4.4 F white arrow, PS) and different subcellular regions of round spermatids, consistent with a localisation in the developing acrosome and in focal dots in the cytoplasm (Fig. 4.4 D yellow arrow, rSptd). In SCO biopsies (Fig. 4.5), a very faint IPO4 expression was detected in parts of Sertoli cell nuclei (Fig. 4.5 C, D yellow arrow SC) and cytoplasm (Fig. 4.5 E, pink arrow Cyto).

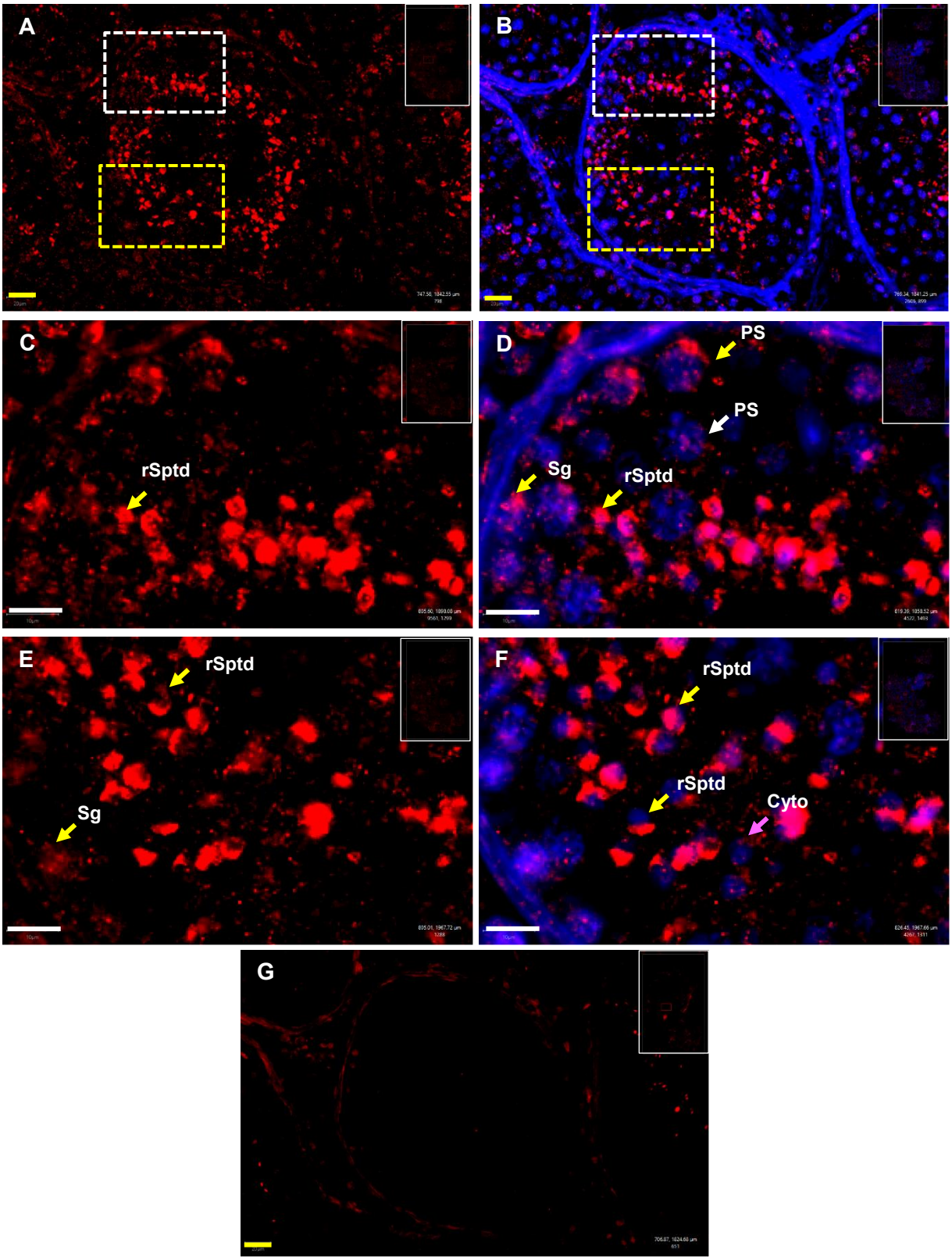


Figure 4.3: Detailed localisation of IPO4 in NSP biopsies. IPO4 immunostaining was detected using Alexa Fluor 555 secondary antibody and nuclei stained with DAPI. Slides were scanned with a VS 120 scan microscope, signal detected using QuPath software and

representative images are shown. NSP – normal spermatogenesis. *Inset:* the total tissue scanned. IPO4 immunofluorescence (red) and DAPI nuclear staining (blue) in testis biopsies. **A** – a lower magnification of a NSP tubule. **B** – tubule A with IPO4-DAPI. **C, D** - the higher magnification images of the white dotted area in **A** and **B** are represented in **C** (only IPO4) and **D** (with IPO4 and DAPI channels). **C** – the yellow arrow shows stained round spermatid (rSptd) **D** – yellow arrows show stained spermatogonia (Sg), pachytene spermatocyte (PS) with cytoplasmic staining, round spermatids (rSptd) and white arrow shows lightly stained pachytene spermatocyte (PS). **E, F** – the higher magnification images of the yellow dotted areas of **A** and **B** are represented in **E** (only IPO4) and **F** (with IPO4 and DAPI channels). **E -** the yellow arrows show stained spermatogonia (Sg) and light nuclear staining in round spermatid (rSptd). **F** –yellow arrows show stained round spermatids (rSptd) and pink arrow shows cytoplasmic staining (Cyto) in a round spermatid. **G** – negative control of tubule A. The yellow scale bars are 20μM and white scale bars are 10μM.

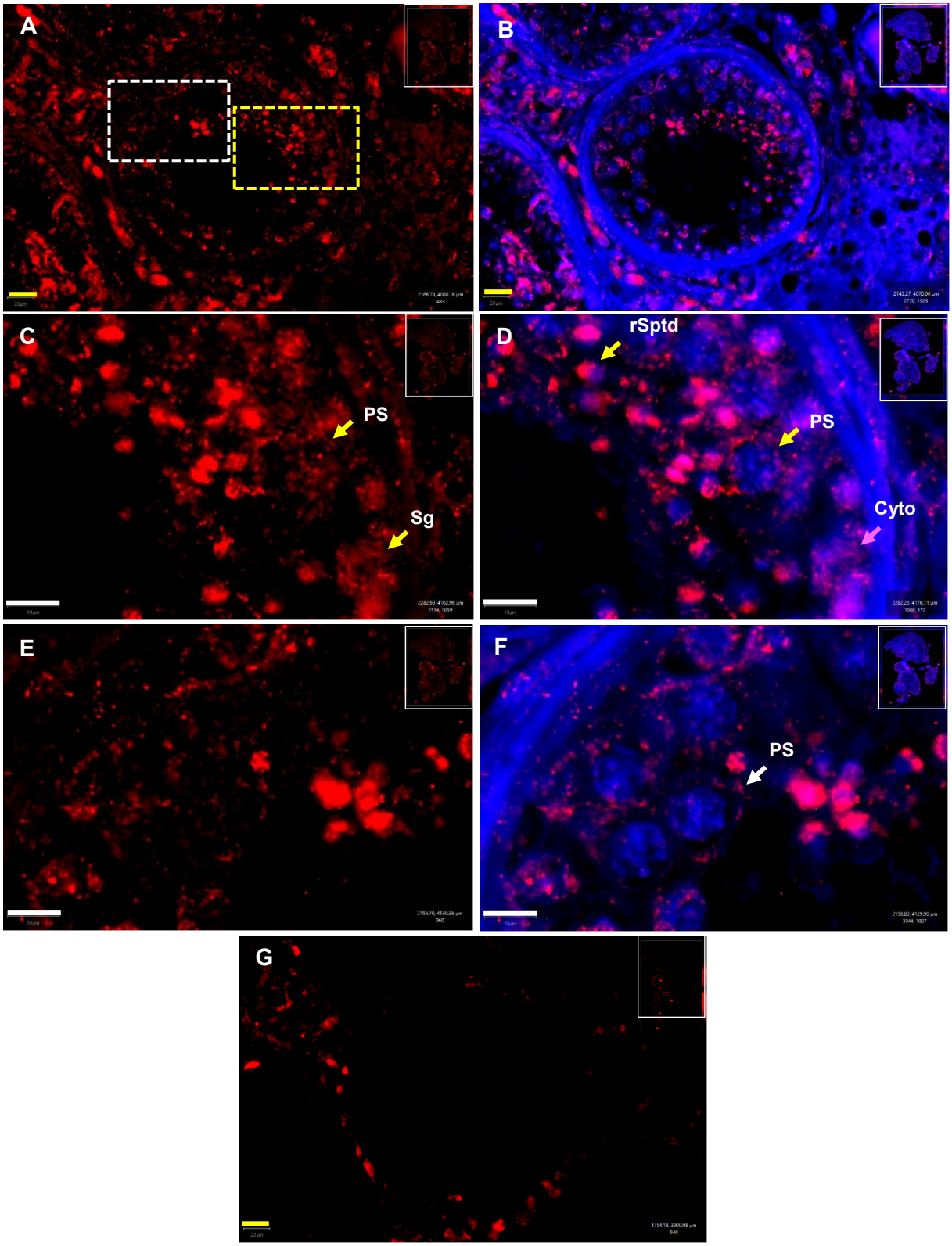


Figure 4.4: Localisation of IPO4 in SDA biopsies. IPO4 immunostaining was detected using Alexa Fluor 555 secondary antibody and nuclei stained with DAPI. Slides were scanned with a VS 120 scan microscope, signal detected using QuPath software and representative images are shown. SDA – spermatid arrest. *Inset:* the total tissue scanned. IPO4 immunofluorescence (red) and DAPI nuclear staining (blue) in testis biopsies. **A** – a lower magnification of a tubule

in SDA biopsy. **B** – tubule A with IPO4-DAPI**. C, D** - the higher magnification images of the yellow dotted area in **A** are represented in **C** (only IPO4) and **D** (with IPO4 and DAPI channels). **C** – the yellow arrow shows stained spermatogonia (Sg) and white arrow shows lightly stained pachytene spermatocytes (PS). **D** – pink arrow shows cytoplasmic staining (Cyto) in spermatogonia, yellow arrows show nuclear and cytoplasmic IPO4 localisation in a late pachytene spermatocyte (PS) and the acrosomal staining in round spermatids (rSptd). **E, F** – the higher magnification images of the white dotted area of **A** are represented as **E** (only IPO4) and **F** (with IPO4 and DAPI channels). **E –** image showing a general view of the light IPO4 cytoplasmic staining. **F –** white arrow shows lightly stained pachytene spermatocyte nucleus with light cytoplasmic staining (PS). **G** – negative control of tubule A. The yellow scale bars are 20μM and white scale bars are 10μM.

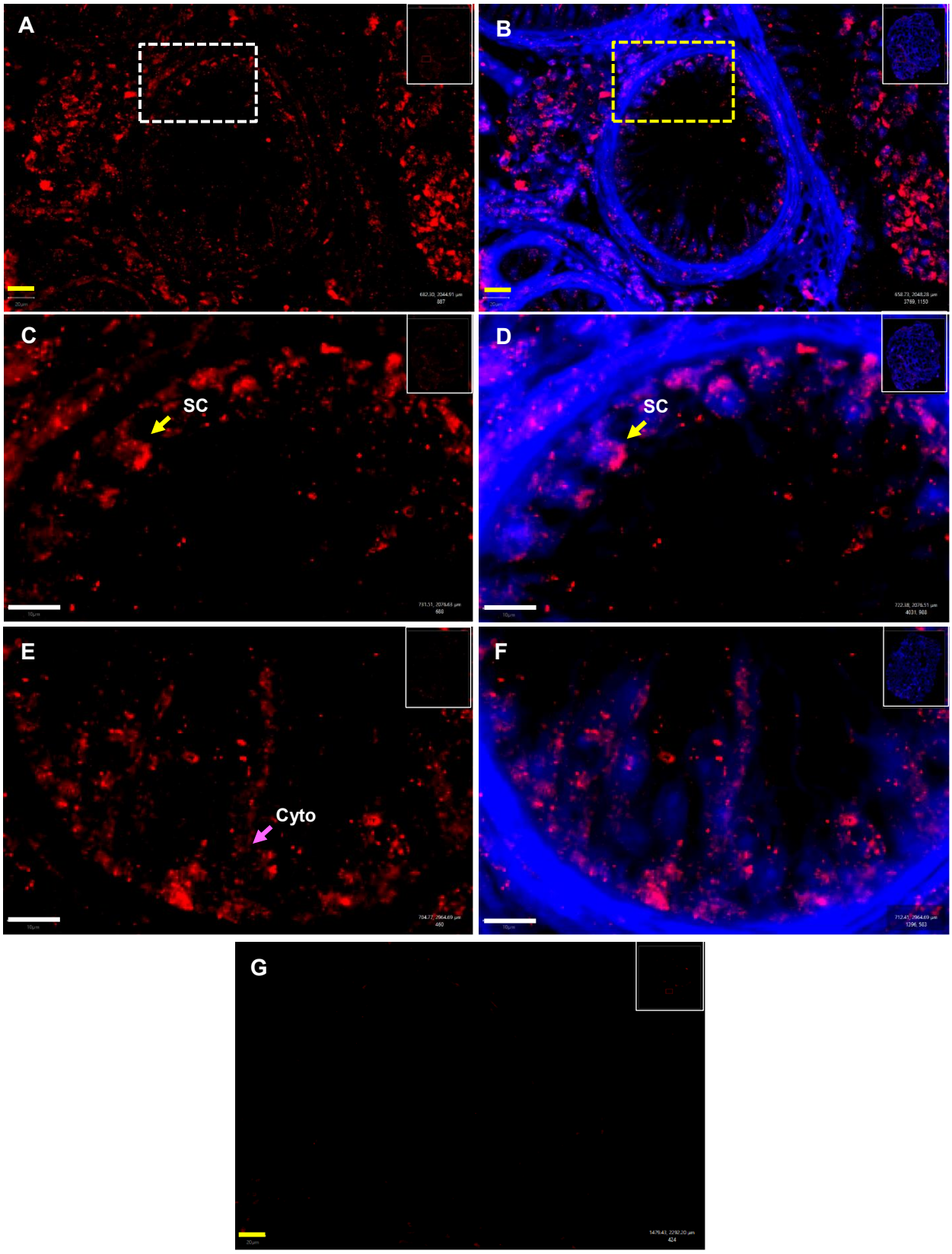


Figure 4.5: Localisation of IPO4 in SCO biopsies. IPO4 immunostaining was detected using Alexa Fluor 555 secondary antibody and nuclei stained with DAPI. Slides were scanned with a VS 120 scan microscope, signal detected using QuPath software and representative images

are shown. SCO – Sertoli cell only phenotype. *Inset:* the total tissue scanned. IPO4 immunofluorescence (red) and DAPI nuclear staining (blue) in testis biopsies. **A** – a lower magnification of a SCO biopsy P. **B** –tubule A with IPO4-DAPI**. C, D** - the higher magnification images of the white dotted area in **A** are represented in **C** (only IPO4) and yellow dotted area represented in **D** (with IPO4 and DAPI channels). **C, D** – the yellow arrow shows a faintly stained Sertoli cell nucleus (SC). **E, F** – the higher magnification images of a part of a SCO tubule from another patient, represented in **E** (IPO4-only) and **F** (with IPO4-DAPI). **E –** pink arrow shows light cytoplasmic staining. **F** – image showing light cytoplasmic staining near Sertoli cells. **G** – negative control of tubule A. The yellow scale bars are 20μM and white scale bars are 10μM.

4.3.2.2 Expression of ASF1B in NSP tissues

Immunofluorescence detection of ASF1B was also established in NSP testis biopsies and then analysed using dual labelling with IPO4 (Fig. 4.6). ASF1B showed nuclear immunostaining in spermatogonia (I - IV) (Fig. 4.6 C, D yellow arrows Sg). Pachytene spermatocytes in stage I-III showed very little staining (Fig. 4.6 C, D white arrows PS), however nuclear localisation was more evident in later pachytene spermatocytes in stages IV to VI (Fig. 4.6 G, H yellow arrows PS). Round spermatids showed immunostaining in the nucleus in stages I to III (Fig. 4.6 C yellow arrow rSptd); however, as they commenced elongation, nuclear staining was not evident (Fig. 4.6 D, H white arrows, Sptd). Lightly stained Sertoli cell nuclei were also observed (Fig. 4.6 C white arrow SC) (Suppl. Table 4.2). The unstained negative control verified the specificity of the ASF1B signal (not shown).

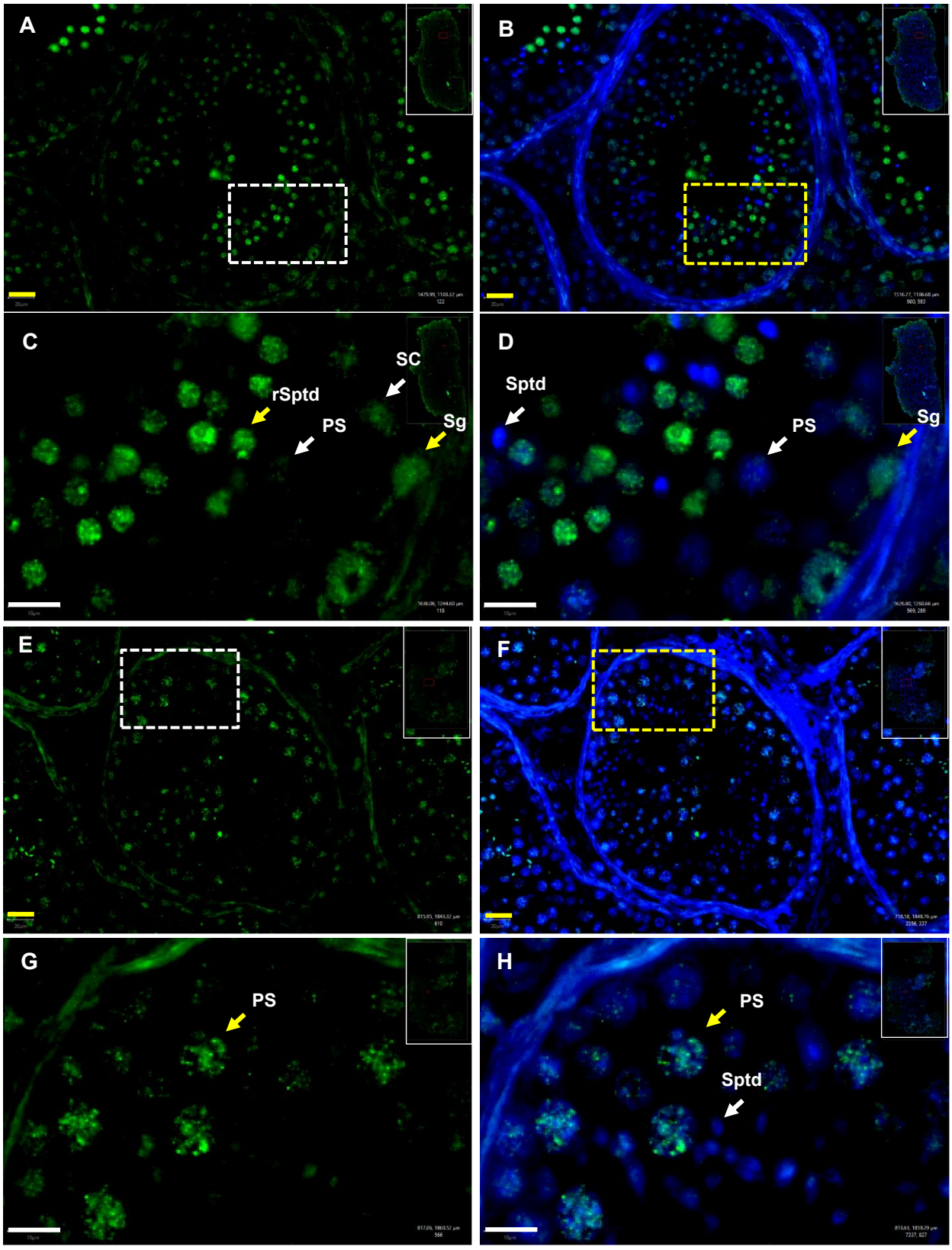


Figure 4.6: Localisation of ASF1B in NSP biopsies. Immunostaining was detected using Alexa Fluor 488 secondary antibody and DAPI. Slides were scanned with a VS 120 scan microscope, signal detected using QuPath software and representative images are shown.

NSP – normal spermatogenesis. *Inset:* the total tissue scanned. **A** – NSP tissue showing the general ASF1B (green) localisation in germ cells. **B** – ASF1B (green) and DAPI (blue) stained NSP tissue. **C** - the higher magnification of the white dotted area in image **A** (ASF1B only). Yellow arrows indicate ASF1B stained spermatogonia (Sg) and round spermatids (rSptd) and white arrows indicates lightly stained Sertoli cell nucleus (SC) and pachytene spermatocyte (PS) with very little staining. **D** –the higher magnification of the yellow dotted area in image **B** (ASF1B-DAPI). The yellow arrow indicates ASF1B stained spermatogonia (Sg) and white arrows indicate pachytene spermatocytes (PS) with very little staining and unstained spermatids (Sptd). **E** – A different NSP tubule with ASF1B staining. **F** – Image **E** with ASF1B and DAPI channels. **G** – the higher magnification image of the white dotted part in image **E** (ASF1B-only). The yellow arrow indicates ASF1B stained pachytene spermatocyte (PS). **H**– the higher magnification of the yellow dotted part of image **F** (with ASF1B-DAPI). The yellow arrow indicates ASF1B stained pachytene spermatocyte (PS) and the white arrow indicates unstained spermatids (Sptd). The yellow scale bars are 20μM and the white scale bars are 10μM.

4.3.2.3 Co-localisation of IPO4 and ASF1B in NSP

Dual-label immunofluorescence was performed in NSP tissues to examine the association between IPO4 and ASF1B during spermiogenesis and to investigate the hypothesis that IPO4 could be involved in nuclear transport of ASF1B during human spermiogenesis (Fig. 4.7). ASF1B was observed in spermatogonia nuclei and IPO4 was apparent as focal dots in the nuclei (Fig. 4.7 C yellow arrows Sg). In early pachytene spermatocytes, very little ASF1B or IPO4 staining was observed (Fig. 4.7 B, white arrows PS). As meiosis progressed, ASF1B was observed in pachytene spermatocyte nuclei and focal areas of IPO4 were apparent in the cytoplasm (Fig. 4.7 C, E and F, white arrows PS). Early round spermatids (steps 1-2) showed ASF1B in the nucleus and IPO4 mainly in the Golgi apparatus (Fig. 4.7 A, yellow arrow, rSptd) and IPO4 at the site of the developing acrosome (Fig. 4.7 A, blue arrow, rSptd). As round spermatid development progressed, IPO4 in the acrosome spread across the nucleus and some of the IPO4 staining co-localised with ASF1B as noted by yellow fluorescence (Fig. 4.7 C, yellow arrows, rSptd); this localisation is consistent with co-localisation of IPO4 and ASF1B in the acroplaxome which is located between the acrosome membrane and the nuclear membrane (Fig. 4.7 C, blue arrow). In elongating spermatids from approximately steps 4-5 onwards, nuclear localisation of ASF1B abruptly ceased, and IPO4 was evident in both the acrosome and nucleus (Fig. 4.7 E, F, yellow arrow, Sptd) (Suppl. Table 4.3).

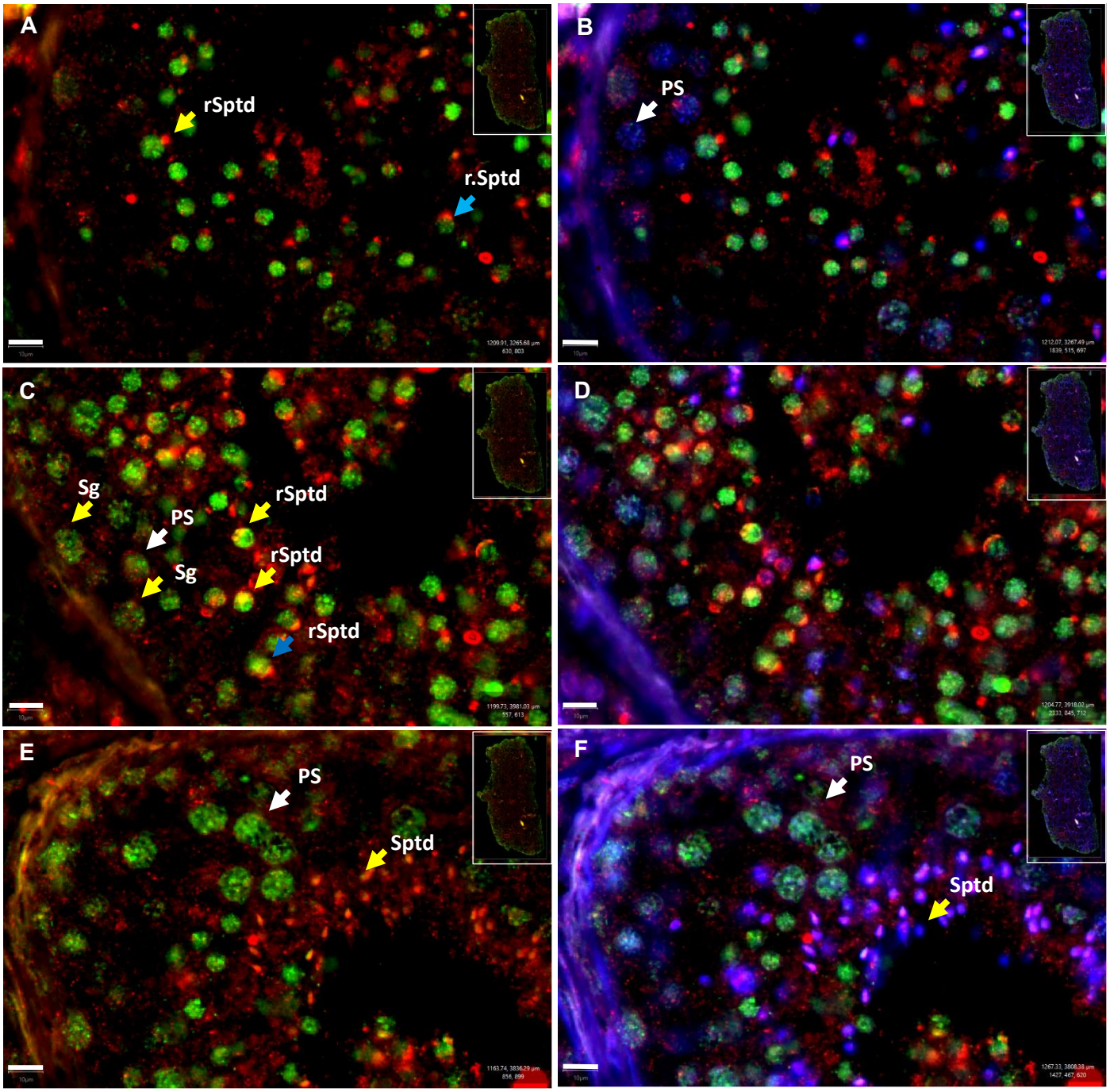


Figure 4.7: Dual-labelled immunofluorescence images of IPO4-ASF1B in NSP biopsies. Immunostaining was detected using secondary antibodies Alexa Fluor 555 for IPO4, Alexa Fluor 488 for ASF1B and nuclei stained with DAPI. Slides were scanned with a VS 120 scan microscope, signal detected using QuPath software and representative images are shown. NSP – normal spermatogenesis. *Inset:* the total tissue scanned. IPO4 immunofluorescence (red) ASF1B (green) and DAPI nuclear staining (blue) in testis biopsies. **A** – Yellow arrow indicates round spermatid with ASF1B staining in the nucleus and IPO4 staining in the Golgi (rSptd), and blue arrow indicates ASF1B in the nucleus and IPO4 predominantly localised in the acrosomal vesicle as it spreads across the nucleus of round spermatids (rSptd). **B** – Image A with DAPI (blue) channel. White arrow indicates very faintly stained pachytene

spermatocytes (PS) with only DAPI staining. **C** – White arrow indicates ASF1B (green) lightly stained pachytene nucleus (PS). The blue arrow indicates a round spermatid showing the IPO4-ASF1B co-localisation in the acrosome at the beginning of acrosome development. Yellow arrows indicate IPO4-ASF1B co-localisation in round spermatids at approximately step 3-4 of development. Spermatogonia (Sg) show focal IPO4 and ASF1B staining but no obvious co-localisation. **D** – image C with DAPI. **E** – White arrow shows pachytene spermatocyte (PS) with nuclear staining of ASF1B and IPO4 focal dots mainly in the cytoplasm. The yellow arrow indicates IPO4 stained spermatid (Sptd), **F** – image E with the DAPI channel. White arrow shows pachytene spermatocyte (PS) with nuclear staining of ASF1B and IPO4 focal dots mainly in the cytoplasm. The yellow arrow shows IPO4 stained spermatid (Sptd). White scale bars are 10μM.

4.3.2.4 Co-localisation of IPO4-ASF1B in SDA.

In dual-labelled SDA biopsies, IPO4 and ASF1B were observed in spermatogonia, pachytene spermatocytes and round spermatids (Fig. 4.8 and Fig. 4.9). Spermatogonia generally showed very little staining. A faint ASF1B signal was seen in nuclei of spermatogonia with IPO4 observed in the cytoplasm (Fig. 4.8 C yellow arrow, Sg). Pachytene spermatocytes had ASF1B labelling in the nucleus and focal areas of IPO4 immunostaining in the cytoplasm (Fig. 4.8 C and 4.9 C yellow and pink arrows, PS). The ASF1B signal in pachytene spermatocytes' nuclei was more evident in later spermatocytes compared to earlier stage pachytenes (Fig. 4.8 C, E, yellow arrows PS vs Fig. 4.9 F, blue arrow PS). In early round spermatids, ASF1B was detected in the nucleus and focal staining of IPO4 was observed in the acrosomal vesicle (Fig. 4.9 F yellow arrows, rSptd) with some co-localisation observed in the developing acrosome, consistent with a localisation to the acroplaxome (4.9 E, F, white arrows rSptd). Elongating spermatids were only occasionally observed in SDA biopsies. However, when present, it was evident that ASF1B was absent in the elongating spermatid nuclei, but IPO4 remained. (Fig 4.8 C, D blue arrows, Sptd). The negative controls were unstained (Fig. 4.8 B and Fig.4.9 B).

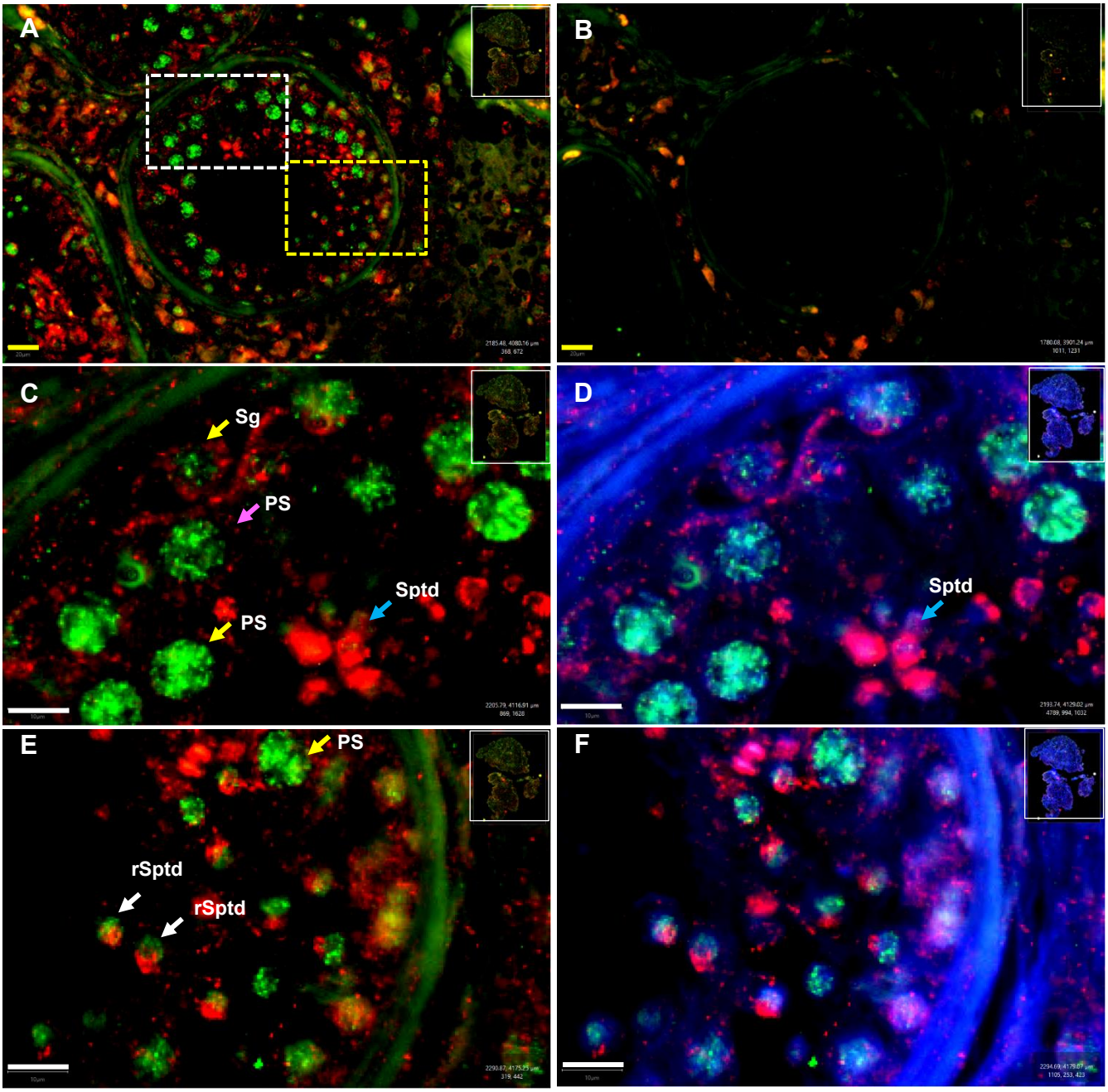


Figure 4.8: Dual-labelled immunofluorescence images of IPO4-ASF1B in SDA sample 1. Immunostaining was detected using Alexa Fluor 555, 488 secondary antibodies and DAPI. Slides were scanned with a VS 120 scan microscope, signal detected using QuPath software and representative images are shown. SDA – spermatid arrest. *Inset:* the total tissue scanned. IPO4 immunofluorescence (red) ASF1B (green) and DAPI nuclear staining (blue) in SDA sample 1. **A** – a lower magnification of a tubule in SDA sample 1. **B** – the corresponding negative control of tubule A. **C, D** - the higher magnification images of the white dotted area in **A** are represented as **C** (with IPO4 and ASF1B channels) and **D** (with IPO4, ASF1B and DAPI channels). **C** – yellow arrows show ASF1B nuclear stained and IPO4 cytoplasmic stained spermatogonia (Sg) and ASF1B stained pachytene spermatocyte nucleus (PS), pink arrow shows IPO4 staining in the cytoplasm of pachytene spermatocyte (PS) and the blue arrow indicates IPO4 (only) stained spermatids. **D** – blue arrow indicates IPO4 (only) stained spermatids. **E, F** – the higher magnification images of the yellow dotted area of **A** are

represented as **E** (with IPO4 and ASF1B channels) and **F** (with IPO4, ASF1B and DAPI channels). **E** – yellow arrow indicates ASF1B stained pachytene spermatocyte (PS) and white arrows indicate round spermatids (rSptd) showing ASF1B stained nuclei and IPO4 stained acrosome. The yellow scale bars are 20μM and white scale bars are 10μM.

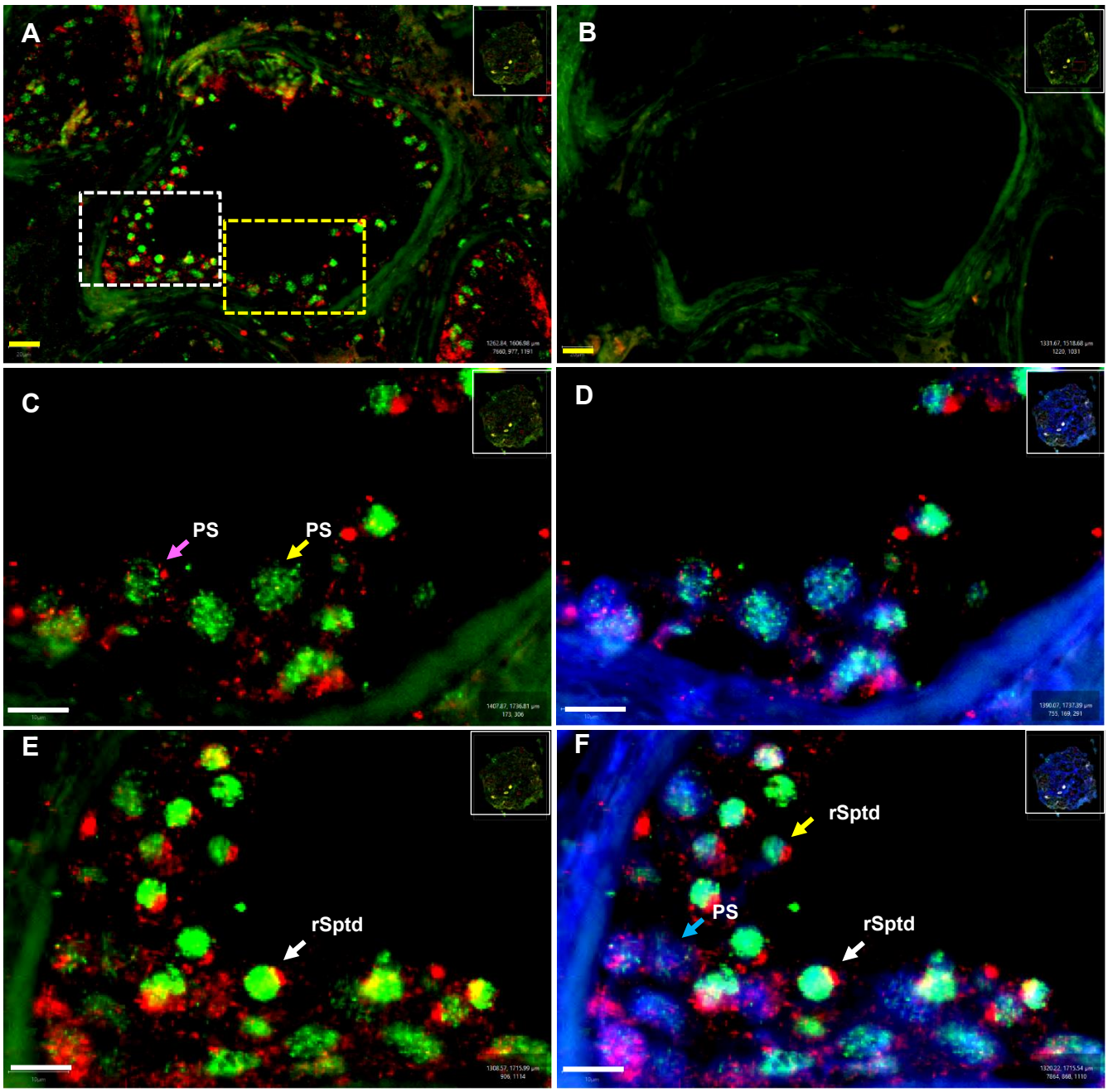


Figure 4.9: Dual-labelled immunofluorescence images of IPO4-ASF1B in SDA sample 2. Immunostaining was detected using Alexa Fluor 555, 488 secondary antibodies and DAPI. Slides were scanned with a VS 120 scan microscope, signal detected using QuPath software and representative images are shown. SDA – spermatid arrest. *Inset:* the total tissue scanned. IPO4 immunofluorescence (red) ASF1B (green) and DAPI nuclear staining (blue) in testis SDA sample 2. **A** – a lower magnification of a tubule in SDA sample 2. **B** – the corresponding negative control of tubule A. **C, D** - the higher magnification images of the yellow dotted area in **A** are represented as **C** (with IPO4 and ASF1B channels) and **D** (with IPO4, ASF1B and

DAPI channels). **C** – the yellow arrow shows ASF1B nuclear stained pachytene spermatocytes (PS) and the pink arrow shows an ASF1B nuclear stained and IPO4 cytoplasmic stained pachytene spermatocyte (PS). **E, F** – the higher magnification images of the white dotted area of **A** are represented as **E** (with IPO4 and ASF1B channels) and **F** (with IPO4, ASF1B and DAPI channels). **E** – the white arrow indicates a round spermatid with IPO4 and ASF1B staining in the developing acrosome, with yellow co-localisation seen at a site consistent with the acroplaxome. **F** –the blue arrow shows a pachytene spermatocyte with very faint ASF1B nuclear staining, yellow arrow shows an early round spermatid with IPO4 and ASF1B staining (rSptd) and the white arrow shows the corresponding round spermatid (rSptd) shown in E. The yellow scale bars are 20μM and white scale bars are 10μM.

4.3.2.5 Co-localisation of IPO4-ASF1B in SCO.

In dual-labelled SCO biopsies, both IPO4 and ASF1B were detected at low levels in Sertoli cell nuclei. ASF1B was detected in the majority of Sertoli cell nuclei and IPO4 was detected only in some (Fig. 4.10 C, D, yellow arrows SC), with unstained nuclei occasionally observed (Fig. 4.10 E, F, white arrows). Co-localisation (as evident by yellow fluorescence) was not observed in Sertoli cell nuclei (Fig. 4.10 A - D). Interstitial staining was apparent for both IPO4 and ASF1B (Fig. 4.10 A). As this staining was also observed in the negative control slides (Fig. 4.10 G), it is considered non-specific. The seminiferous tubules in the negative controls were unstained (Fig. 4. 10 G, H).

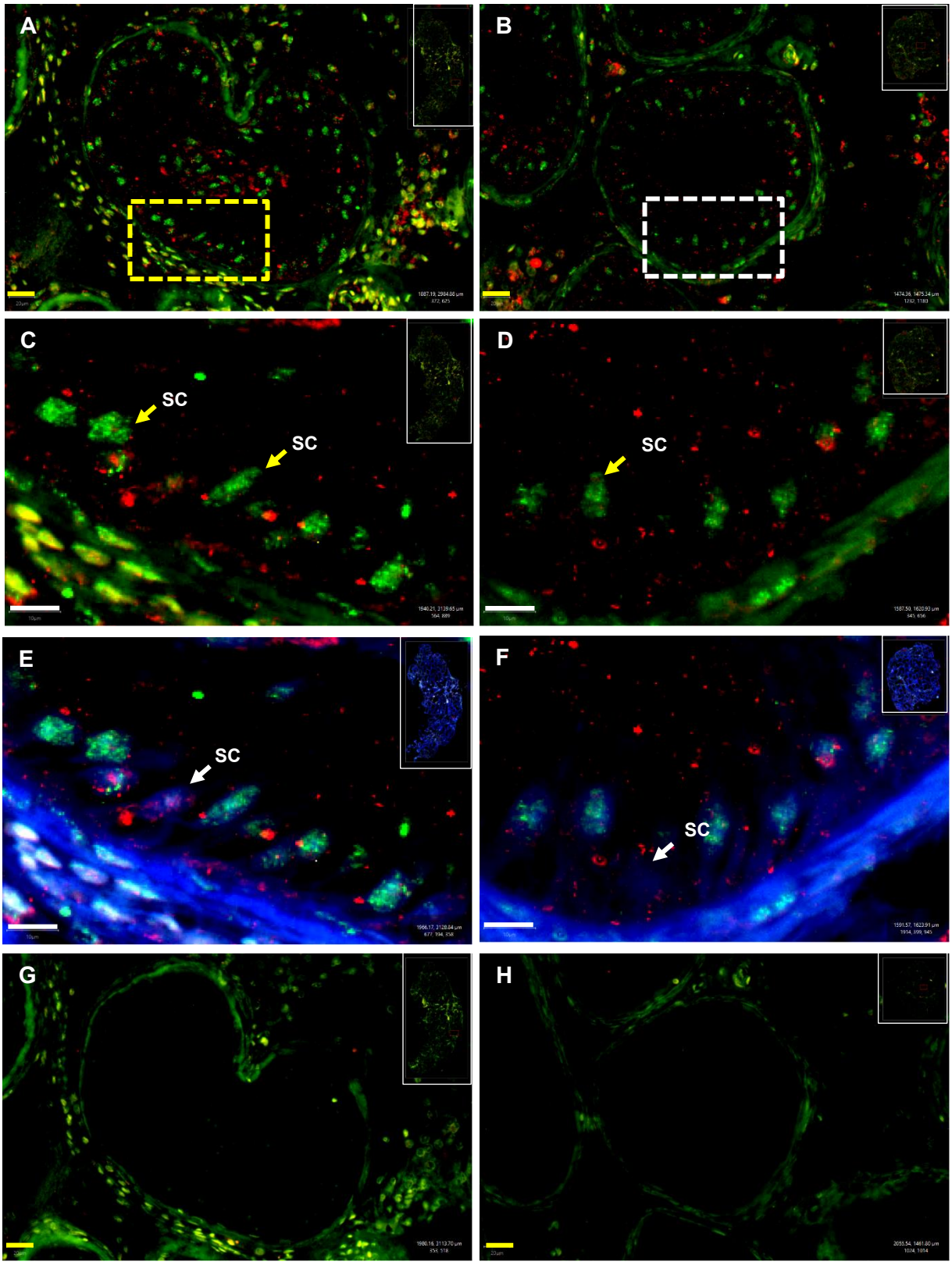


Figure 4.10: Dual-labelled immunofluorescence images of IPO4-ASF1B in SCO biopsies. Immunostaining was detected using Alexa Fluor 555, 488 secondary antibodies and DAPI. Slides were scanned with a VS 120 scan microscope, signal detected using QuPath software and representative images are shown. SCO – Sertoli cell only. *Inset:* the total tissue

scanned. IPO4 immunofluorescence (red) ASF1B (green) and DAPI nuclear staining (blue) in testis. The figure shows tubules of two different SCO samples **A** and **B**. The higher magnification images of the yellow dotted area in **A** are shown in **C** (with IPO4 and ASF1B channels) and **E** (with IPO4, ASF1B and DAPI channels). **C-** yellow arrows show Sertoli cell nuclei with lightly stained IPO4 and ASF1B but without apparent co-localisation. **E** – white arrow shows unstained Sertoli cell nucleus. The higher magnification images of the white dotted area of **B** are shown in **D** (with IPO4 and ASF1B channels) and **F** (with IPO4, ASF1B and DAPI channels). **D** - yellow arrow shows lightly stained (ASF1B) Sertoli cell nucleus. **F** - white arrow shows unstained Sertoli cell nucleus. **G**– the negative control of the tubule A. **H** – the negative control of the tubule B. The yellow scale bars are 20μM and white scale bars are 10μM.

4.3.4 Co-localisation of IPO4-TNP1 in NSP.

Prior to co-localisation experiments, TNP1 was investigated in NSP biopsies alone to define optimal conditions and to broadly assess its localisation (Fig. 4.11). TNP1 localisation was observed to be highly stage-specific (Fig. 4.11 B) (Suppl. Table 4.4). Co-localisation of IPO4 and TNP1 was then assessed in NSP using dual-labelled immunofluorescence to examine whether TNP1 could be a potential cargo of IPO4 (Fig. 4.12 and Fig. 4.13). TNP1 was barely detectable in germ cells prior to step 3 spermatids and was detected in high intensity only in spermatids of steps 3-5 (Fig. 4.12 B and F, white arrows, Sptd). IPO4 and TNP1 were clearly co-localised in the nuclei of steps 3-5 spermatids (Fig. 4.12 C, G yellow arrows Sptd and Fig. 4.13 A pink arrow Sptd). However, IPO4-TNP1 co-localisation was not visible in earlier round spermatids where IPO4 was visible in the developing acrosome (Fig. 4.13 C, D white arrows rSptd, only IPO4) (Suppl. Table 4.5). The unstained negative control confirmed the specificity of the TNP1 antibody (Fig. 4.11 C, D). Due to some practical hindrances and time constraints, SDA and SCO samples were not analysed for IPO4-TNP1.

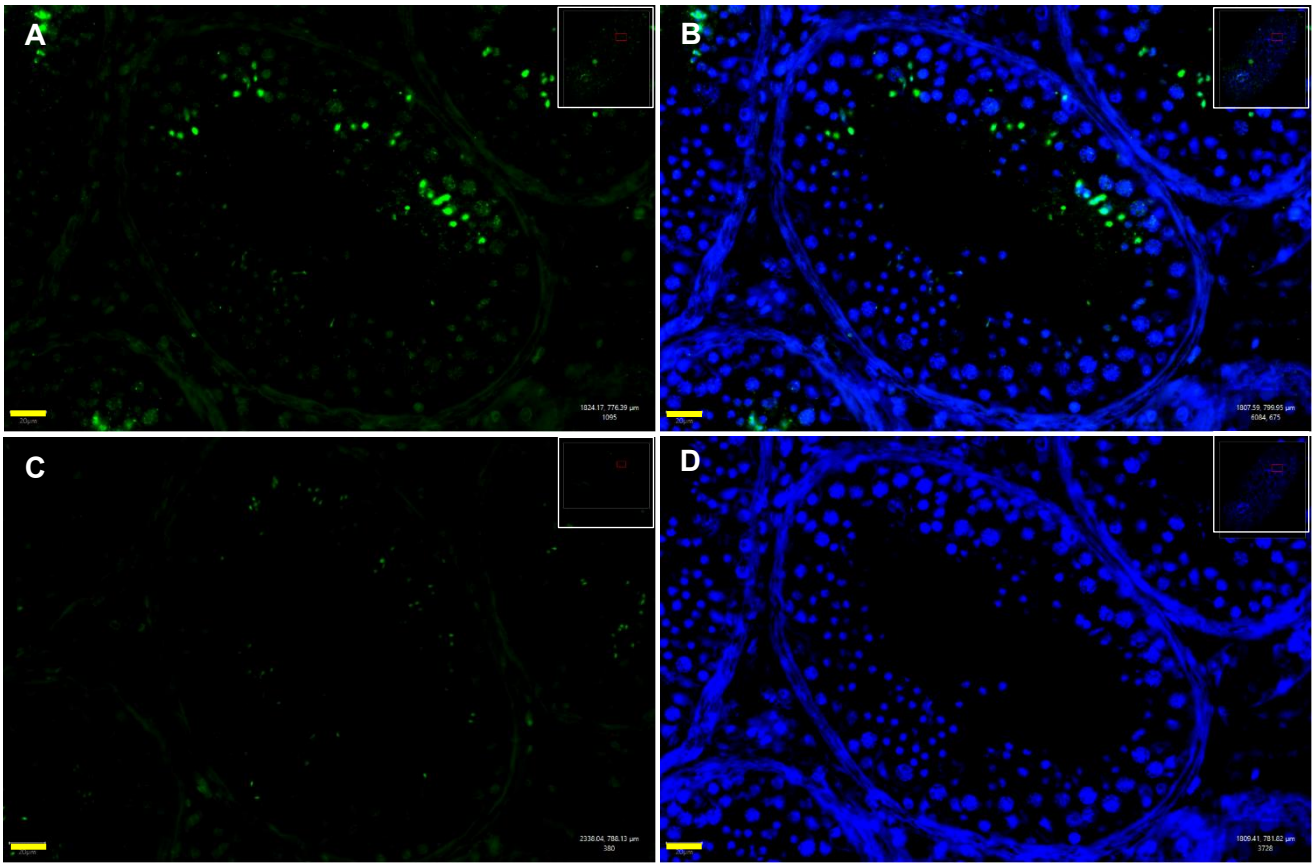


Figure 4.11: A general view of the localisation of TNP1 in NSP biopsies. Immunostaining was detected using Alexa Fluor 488 secondary antibody and DAPI. Slides were scanned with a VS 120 scan microscope, signal detected using QuPath software and representative images are shown. NSP – normal spermatogenesis. *Inset:* the total tissue scanned. **A** – NSP tissue showing the TNP1 (green) localisation in spermatids. **B** – TNP1 (green) and DAPI (blue) stained NSP tissue. **C** – negative control section. **D** – NSP negative control tubule with only DAPI (blue). The grey scale bars are 20μM.

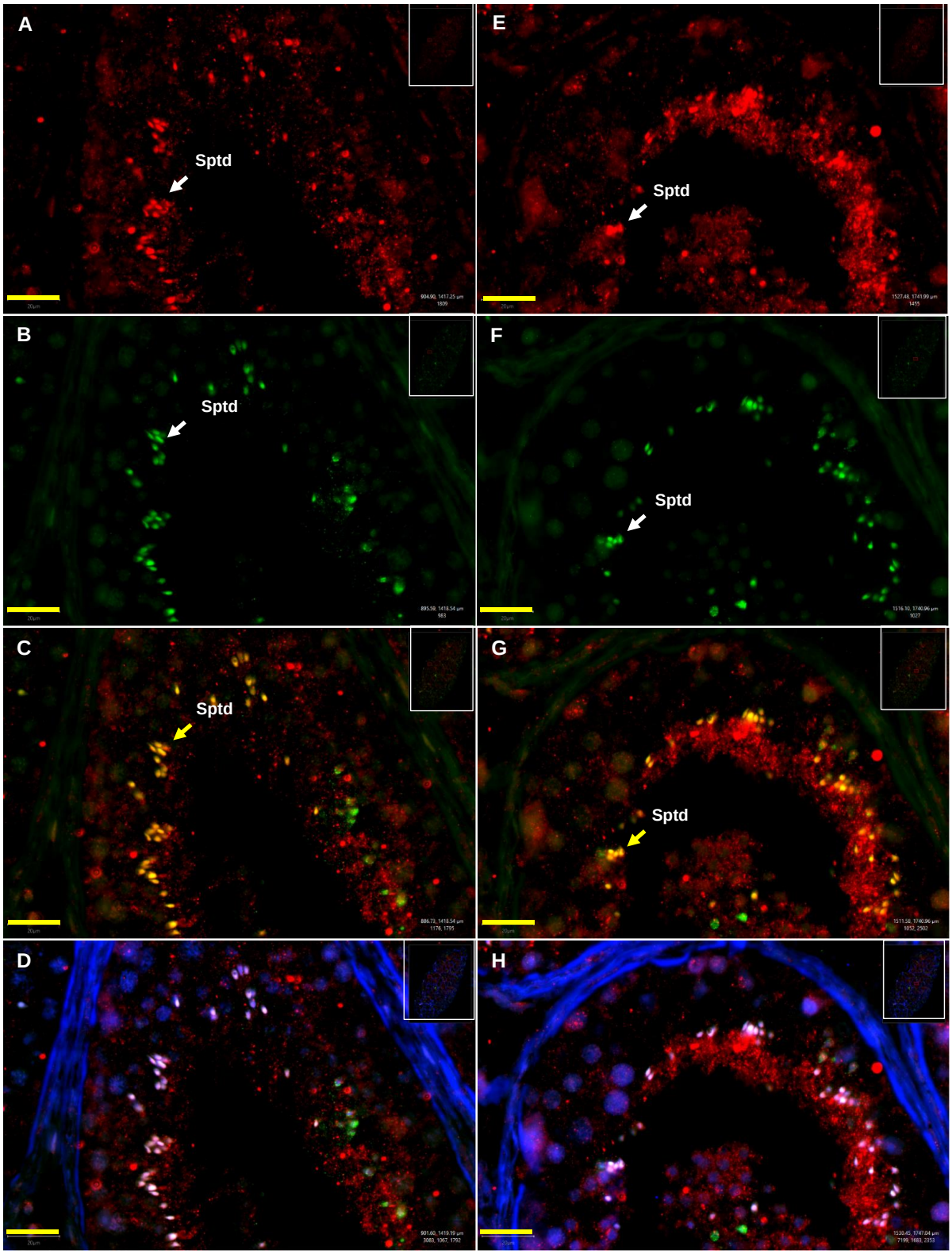


Figure 4.12 Co-localisation of IPO4 - TNP1 in NSP, sample 1. Immunostaining was detected using Alexa Fluor 488, 555 secondary antibodies and DAPI. Slides were scanned with a VS 120 scan microscope, signal detected using QuPath software and representative images are shown. NSP – normal spermatogenesis. *Inset:* the total tissue scanned. IPO4 immunofluorescence (red) TNP1 (green) and DAPI nuclear staining (blue) in testis. The image

panel shows two different NSP samples as **A-D** - one tubule with different channels and **E-H** another tubule with different channels. **A, E** – IPO4 localisation in NSP**. B, F** – TNP1 localisation in NSP, **C, G** – IPO4-TNP1 co-localisation in NSP. **D, H** – IPO4- TNP1 co-localisation with DAPI channel. **A, E** –white arrows indicate IPO4 (red) localisation in spermatids (Sptd). **B, F**– white arrows indicate TNP1 (green) localisation in spermatids (Sptd). **C, G**– yellow arrows indicate the co-localisation of IPO4-TNP1 (yellow) in spermatids (Sptd), **D, H** – The co-localisation of IPO4-TNP1 with DAPI nuclear staining (blue) visible as white staining when all three channels are combined. The yellow scale bars are 20µM.

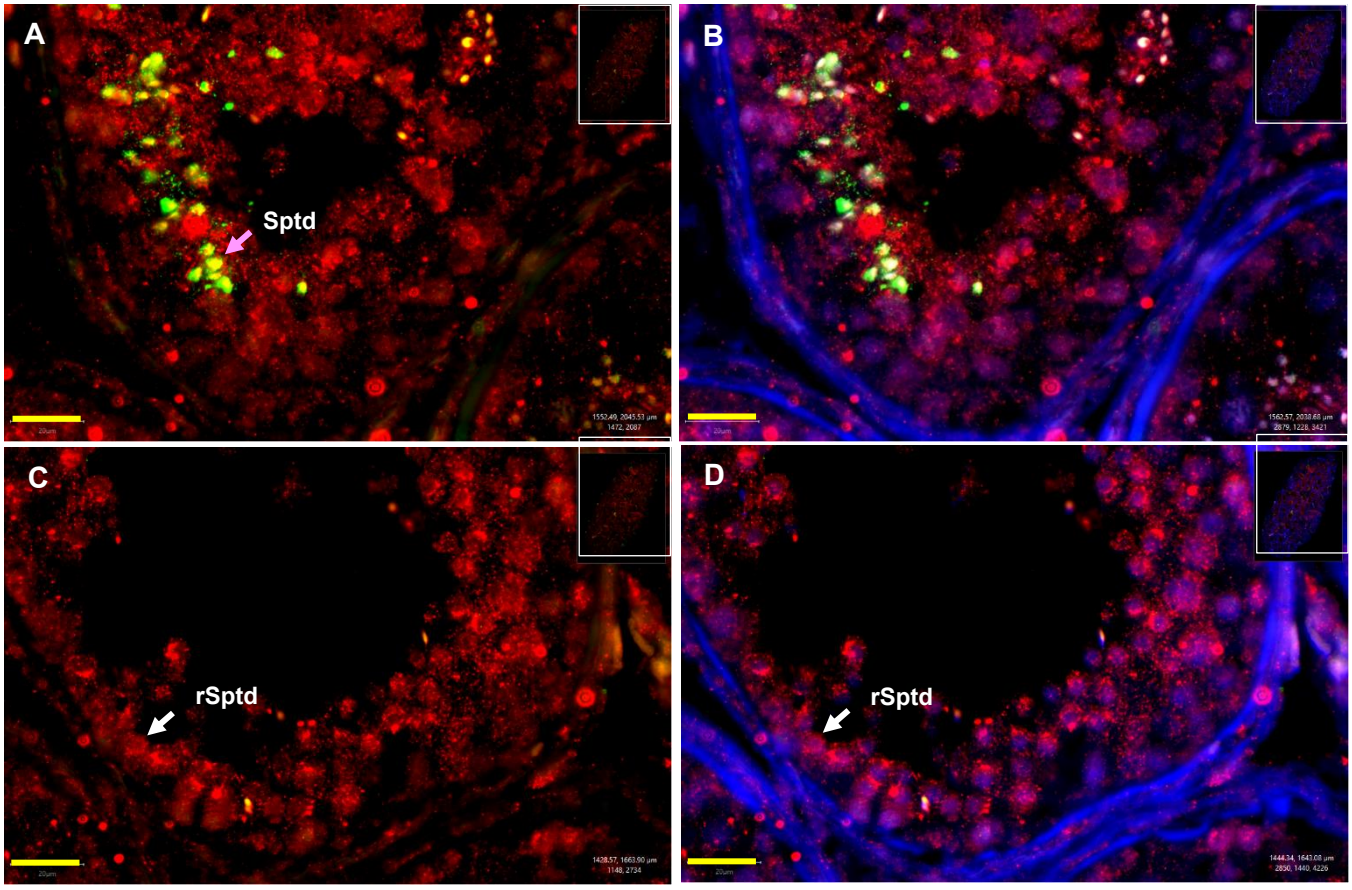


Figure 4.13: Co-localisation of IPO4 - TNP1 in NSP, sample 2. Immunostaining was detected using Alexa Fluor 488, 555 secondary antibodies and DAPI. Slides were scanned with a VS 120 scan microscope, signal detected using QuPath software and representative images are shown. NSP – normal spermatogenesis. *Inset:* the total tissue scanned. **A –** IPO4 and TNP1 channels**.** The pink arrow indicates dual labelled step 5 spermatids (Sptd)**. B -** the tubule A with IPO4-TNP1 and DAPI channels. **C –** A different tubule with IPO4-TNP1 channels. The white arrow indicates a round spermatid (rSptd) in the developing acrosome stage immunostained only with IPO4. **D –** the tubule in C with IPO4-TNP1 and DAPI channels. The white arrow indicates the IPO4 (red) immunolocalisation in round spermatid acrosome (rSptd) but TNP1 immunolocalisation is not evident. The yellow scale bars are 20µM.

4.3.4 Ran protein isoforms in human spermatids and their potential relevance for importin activity during spermiogenesis

In the cytoplasm of cells, importins bind to their cargo protein in the presence of RAN-GDP and release their cargo inside the nucleus in the presence of RAN-GTP (Loveland et al., 2015). Thus, both forms of the RAN protein play a crucial role in nuclear transportation that made it an interesting protein to investigate in relation to IPO4 nucleocytoplasmic transportation. During analysis of mass spectrometry data of human sperm in Chapter 3, it was noted that the UniProt annotation of RAN detected was different to canonical RAN. In the human sperm proteome described in Chapter 3, the RAN-GTP was detected in human sperm, but, unlike IPO4, it was not significantly different between normozoospermic samples and samples with asthenoteratozoospermia (AT). However, the RAN protein detected in human sperm and annotated as GTP-bound RAN had the UniProt accession B5MDF5. This accession number differs from the canonical human RAN protein which has an accession number P62826. RAN-P62826 protein is well described in various online databases including GermOnline and GeneCards but RAN-B5MDF5 is not well described, suggesting it could be a unique isoform of RAN in human sperm.

Spermatids are well known to show high gene function diversity and to express unique variants of proteins which likely contribute to their specialised functions (Chalmel et al., 2007; Soumillon et al., 2013). Given that this Chapter aimed to investigate novel importin functions during spermiogenesis, the potential for a novel RAN protein in human sperm was of interest and was explored further. Because the RAN-B5MDF5 protein isoform could be important in human sperm, yet it is not well characterised, a bioinformatic analysis of its sequence and relationship to the canonical RAN- P62826 was undertaken.

Both RAN UniProt accession numbers are annotated as ‘GTP binding protein RAN’ and have the same ENSEMBL ID (ENSG00000132341). To determine whether these two protein annotations arose from different transcripts, the ENSEMBL database was investigated. Both protein accessions B5MDF5 (ENST00000392367.4) and P62826 (ENST00000392369.6) are annotated as transcripts termed RAN-201 and RAN-202 respectively, suggesting that the two proteins arise from different mRNA transcripts

and thus are splice variants of the same gene. To understand potential structural/functional differences between the two isoforms, the protein sequences of B5MDF5 and P62826 were compared in BLASTp. The analyses revealed an additional amino acid sequence in the RAN-B5MDF5 protein producing a protein of 233aa compared to 216aa in the canonical RAN-P62826 protein (Table 4.4 and Fig. 4.14). Thus, RAN-B5MDF5 could be a novel RAN isoform in human sperm that has not been described previously.

Table 4.4 Comparison between the two RAN proteins identified in human sperm

Accession number	**P62826**	**B5MDF5**
ENSEMBL ID	ENSG00000132341	ENSG00000132341
gene tree under phylogenomic database	ENSGT00940000153786	ENSGT00940000153786
Transcript ID	ENST00000392369.6	ENST00000392367.4
ENSEMBL protein ID	ENSP00000376176.2	ENSP00000376174
Statistics (ENSEMBL)	**Exons:** 6 **Coding exons:** 6 **Transcript length:** 1,530 bps **Translation length:** 216 residues	**Exons:** 8 **Coding exons:** 7 **Transcript length:** 780 bps **Translation length:** 233 residues
Gene ID		
Gene name as per the transcript	RAN-202	RAN-201
Biotype	Protein coding	Protein coding
Proteomics DB	57429	6173
Amino acids	216aa	233aa

RAN – P62826

```
MAAQGEPQVQ  FKLVLVGDGG  TGKTTFVKRH  LTGEFEKKYV  ATLGVEVHPL
        60          70          80          90         100
VFHTNRGPIK  FNVWDTAGQE  KFGGLRDGYY  IQAQCAIIMF  DVTSRVTYKN
       110         120         130         140         150
VPNWHRDLVR  VCENIPIVLC  GNKVDIKDRK  VKAKSIVFHR  KKNLQYYDIS
       160         170         180         190         200
AKSNYNFEKP  FLWLARKLIG  DPNLEFVAMP  ALAPPEVVMD  PALAAQYEHD
       210
LEVAQTTALP DEDDDL
```

RAN – B5MDF5

```
MAAQGEPQVQ  FKLVLVGDGG  TGKTTFVKRH  LTGEFEKKYV  ATLGVEVHPL
        60          70          80          90         100
VFHTNRGPIK  FNVWDTAGQE  KFGGLRDGYY  IQEMGIHHVA  HAGLELMSSA
       110         120         130         140         150
QCAIIMFDVT  SRVTYKNVPN  WHRDLVRVCE  NIPIVLCGNK  VDIKDRKVKA
       160         170         180         190         200
KSIVFHRKKN  LQYYDISAKS  NYNFEKPFLW  LARKLIGDPN  LEFVAMPALA
       210                     220                     230
PPEVVMDPAL AAQYEHDLEV AQTTALPDED DDL
```

Figure 4.14: Comparison of the protein sequences of RAN-P62826 and RAN-B5MDF5. The **bold red** part in the B5MDF5 is the additional amino acid sequence which is not found in P62826.

Given the EMSEMBL analysis predicted that RAN-B5MDF5 could be a novel RAN isoform expressed in human sperm, the structures of the two different RAN isoforms were investigated in the SWISS-MODEL portal. In this portal, the canonical RAN-P62826 216aa sequence corresponds to the template ‘SMTL ID: 7mo0.A’ and the protein RAN-B5MDF5 corresponds to the template ‘SMTL ID: 7mnw.2’. Analyses of the ribbon structures of the two proteins in the SWISS-MODEL revealed important differences in the overall conformation of the two proteins with distinct differences in the two structures (Fig. 4.15 A, different colour coded areas). A closer view of the amino acids in the two different sequences revealed clear differences in the ribbon structures (Fig. 4.15 B, C). Importantly, the analyses revealed that binding sites for

‘importin (transcription factor) interaction’ and GTP correspond to different residues, as reported in InterPro (RAN-B5MDF5 – https://www.ebi.ac.uk/interpro/protein/UniProt/B5MDF5/ and RAN-P62826 https://www.ebi.ac.uk/interpro/protein/UniProt/P62826/). These analyses thus predict that RAN-B5MDF5 could contain fewer residues that would bind to GTP. Overall, the analysis predicts that the novel RAN-B5MDF5 in human sperm could interact with importins and GTP differently to the canonical RAN-P62826 protein, thus suggesting that human sperm contain a functionally different RAN protein that could be important for specialised sperm functions.

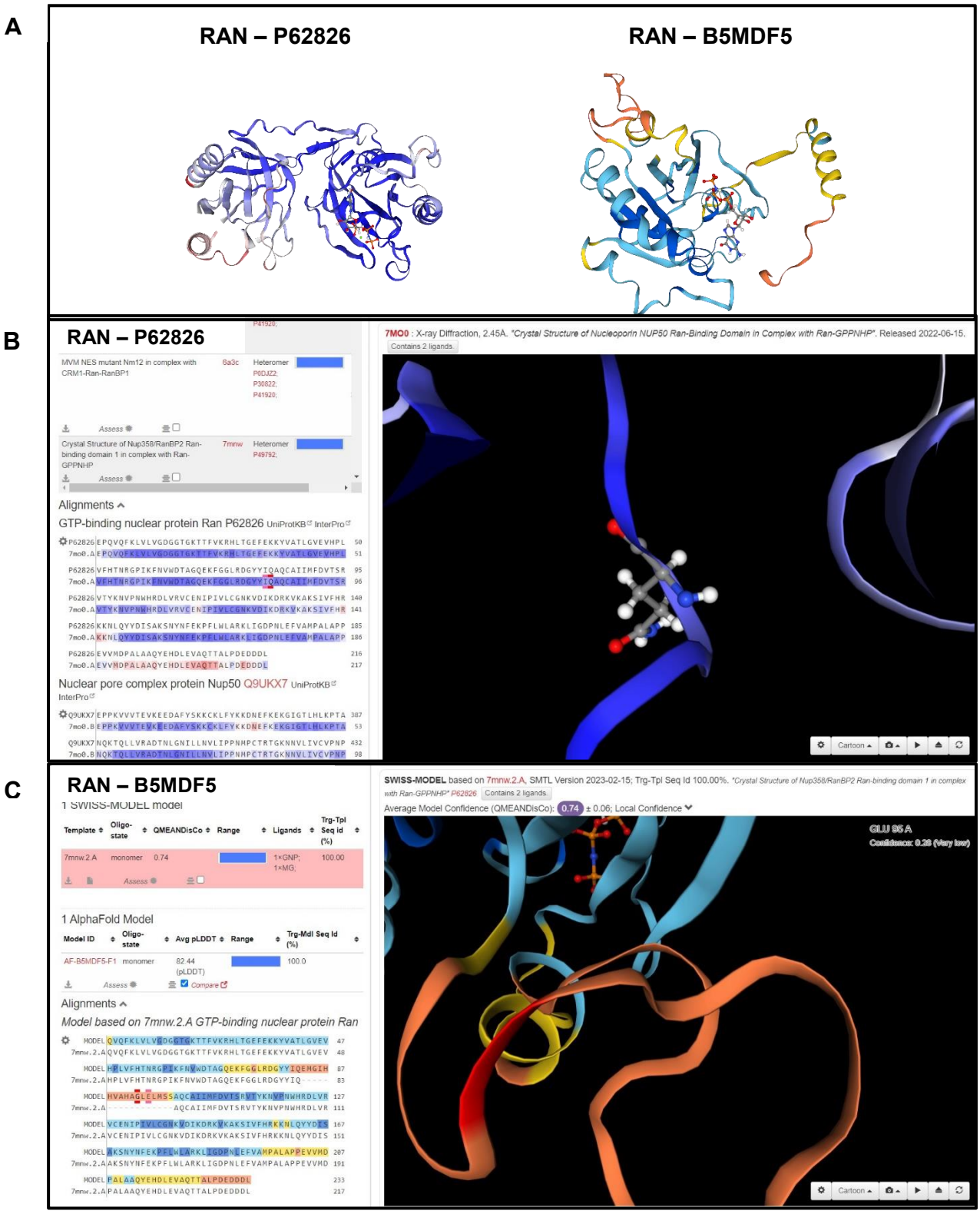


Figure 4.15: The ribbon structures of RAN–P62826 and RAN–B5MDF5 adopted from the SWISS-MODEL. Panel A shows the full structures of the two versions of RAN. Panels B and C show a closer view of structures in RAN-P62826 and RAN-B5MDF5, respectively. The dark

orange band in (C) corresponds to the additional amino acid sequence in RAN-B5MDF5 that is not present in RAN-P62826 in its amino acid sequence shown in (B).

4.4 Discussion

Importins are a group of well-studied transporter proteins that play an important role in many cellular processes, and some importins have been shown to be involved in spermatogenesis and in nuclearcytoplasmic transportation in germ cell (Bernardes et al., 2022; Loveland et al., 2015; Miyamoto et al., 2020; Nathaniel et al., 2022; Whiley et al., 2012; Young et al., 2013). The nucleocytoplasmic transportation of cargo proteins and transcription factors to the nucleus at the correct time is crucial for the proper development and function of germ cells (Loveland et al., 2015). This Chapter investigated importin 4 (IPO4), a beta importin, and its potential involvement in spermatogenesis/spermiogenesis. Studies in Chapter 3 showed that IPO4 was detected at higher levels in human sperm with abnormal sperm morphology and motility. This Chapter investigated the potential role of IPO4 in human spermiogenesis and particularly focused on predicting potential IPO4 cargo proteins through *in silico* analyses. As a beta importin, IPO4 is predicted to bind to a cargo protein in the cytoplasm and actively transport it into the nucleus. In the nucleus, the cargo is released in the presence of RAN-GTP and the IPO4-RAN-GTP complex returns to the cytoplasm so that the transport cycle can start over. So far, the presence and function of IPO4 has not been described in human spermatids albeit there are studies on IPO4 in human and non-human cell lines (Alvarez et al., 2011; Apta-Smith et al., 2018; Bernardes et al., 2022), leaving the open question, which cargos could play a role in sperm development and function. Therefore, this Chapter developed a bioinformatic approach to identify potential IPO4 cargo proteins predominantly detected in human spermatids and to analyse their co-localisation with IPO4 to further investigate the role of IPO4 in spermiogenesis. Ultimately, two proteins were selected for further analysis.

As shown by IHC in Chapter 3, IPO4 was localised to the nuclei of different germ cells and its cytoplasmic expression increased from late pachytenes to round spermatids and elongating spermatids in NSP samples with a clear accumulation in round spermatids, which was not as apparent in SDA samples. Further information on IPO4's subcellular localisation in human germ cells was gained from immunofluorescent

localisation studies. The immunofluorescence localisation in NSP biopsies further verified the low staining in nuclei of spermatogonia and early spermatocytes but a stronger staining in round spermatids and elongating spermatids nuclei with a clear accumulation of IPO4 in the acrosomal areas of round spermatids which is novel. The accumulation of IPO4 in both nuclei and cytoplasm of round spermatids and its apparent nuclear localisation from step 4 elongating spermatids suggests its probable importance in protein transportation during spermatid elongation. IPO4 was evident in acrosome development and in spermatid nuclei throughout the process of spermatid elongation and maturation suggesting an essential role of IPO4's cargos in nuclear remodeling and formation of the sperm head. Thus, it was intriguing to investigate the potential cargo proteins that IPO4 could be transporting during spermiogenesis.

ASF1B was selected as a potential cargo protein due to its binding affinity to histones H3-H4 dimers and a likely interaction with IPO4 (Alvarez et al., 2011; Yoon et al., 2018). Histones are needed for the proper DNA packaging, to regulate the binding activity and for expression of DNA via chemical modifications. For example, acetylation of H3-H4 has shown to facilitate binding of transcription factors and thus activating transcription (Gunes & Kulac, 2013). A previous study using human cervical carcinoma (HeLA) cell lines showed that H3-H4 interacts with the chaperone ASF1A (paralog of ASF1B), and the histones are then bound to IPO4 to be transported to the nucleus. The same study also reported IPO4-H3-H4-ASF1B in the cytosol, but with different post-translational modifications in H3, suggesting the H3-H4-ASF1B complex, too, could be transported to the nucleus by IPO4 (Alvarez et al., 2011). A recent study using cryo-electron microscopy showed that IPO4-H3-H4-ASF1 exists in a complex and that releasing of H3-H4-ASF1 complex in the nucleus is due to conformational changes in IPO4 while binding to RAN-GTP (Bernardes et al., 2022). Thus, this Chapter investigated whether a similar mechanism could be operating during human spermiogenesis involving ASF1B.

The results of this Chapter revealed co-localisation of IPO4-ASF1B in the developing spermatid acrosome, more precisely at the acroplaxome area. At the beginning of spermiogenesis, an accumulation of IPO4 protein in round spermatids was observed as a staining in the area of the Golgi apparatus in early step 1 round spermatids and an acrosome cap-like formation in round spermatids (steps 2-3) of NSP. This suggests IPO4 could be involved in protein transport during acrosome biogenesis. As the

acrosome developed, ASF1B and IPO4 co-localised first at the acroplaxome area and then in the acrosomal granule, whereas IPO4 was detected within the entire acrosome. The reason for the co-localisation of ASF1B and IPO4 in the acrosomal granule is not clear but may suggest that IPO4 could transport ASF1B during acrosome formation. Proacrosomal granules contain many important proteins relevant for acrosome biogenesis; these granules are first formed in the Golgi apparatus and later fuse on the nuclear surface to build the acrosome granule (Khawar et al., 2019). The formation of these vesicles and their transportation to the nuclear surface are important for acrosome biogenesis (Khawar et al., 2019). The immunolocalisation suggests that IPO4 is present in the Golgi at the earliest phase of acrosome biogenesis, when proacrosomal vesicles are forming. As acrosome biogenesis progresses, IPO4 can be seen throughout the entire acrosome as it spreads along the nuclear surface. Dual-labelled immunofluorescence for IPO4 and ASF1B revealed discrete foci of co-localisation in the central acrosomal vesicle and at the site where the acrosome meets the nuclear surface, consistent with a localisation in the acroplaxome. The localisation of ASF1B to the acrosome is a novel observation. Previous studies have shown localisation of histones in subacrosomal area in spermatozoa (Tran et al., 2012) and binding links of ASF1B with histones in cell lines (Alvarez et al., 2011). Thus, it can be speculated that ASF1B could exist in a complex with histones during acrosome biogenesis. Previous studies have also suggested the localisation of alpha importins in the acrosome, proposing that these importins could be assisting in acrosome vesicle trafficking (Loveland et al., 2015; Tran et al., 2012). Thus, the findings in this Chapter can also suggest that IPO4 could be transporting important cargos during the acrosome biogenesis and assist in acrosome vesicle trafficking, and that ASF1B is a likely cargo protein of IPO4 in the human acrosome. After assessing the co-localisation of IPO4-ASF1B in normal spermatogenesis, the study also addressed the question in samples with an impaired spermatogenesis phenotype and chose SDA for this study. There, IPO4-ASF1B was observed in round spermatids in the acrosome/ acroplaxome and granule areas, too, confirming that these proteins function in the normal manner up to this developmental stage. In samples with no germ cells, i.e. SCO, IPO4 and ASF1B were localised both inside the Sertoli cell nuclei but without any sign of co-localisation. These observations point to potential roles for IPO4 and ASF1B in Sertoli cells.

Another interesting finding was that the localisation of ASF1B in the spermatid nucleus abruptly ceased around step 4 as the spermatids commenced elongation, while IPO4 remains inside spermatid nuclei. The reason for the disappearance of ASF1B in late spermatid nuclei is not known. Previous studies have shown that ASF1B binds to H3-H4 histones (Alvarez et al., 2011), however histones in the spermatid nuclei are replaced by transition proteins during spermatid elongation (Steger et al, 2000), thus perhaps ASF1B disappearance from the nucleus reflects the disappearance of histones during elongation. Thus, it could be that conformational changes occur in ASF1B due to the replacement of histones, making it undetectable by the same antibody at this stage. This speculation can be further analysed in a cryo-electron microscopy study to evaluate structural differences of IPO4-ASF1B-H3-H4 at different spermatogenesis phases. In summary, these results can conclude that IPO4 likely has roles in the acrosome biogenesis and in the nucleus of elongating spermatids at a specific window of development in steps 4-6. The co-localisation of IPO4-ASF1B in the acroplaxome and acrosomal granule suggests that ASF1B and IPO4 together could have roles in acrosome biogenesis.

As spermatids elongate and their nucleus condenses, there are major changes in chromatin packaging (Qian et al., 2013). These alterations in chromatin packaging involve the replacement of histone residues, bound to the outside of chromatin, with protamines, which enable the chromatin to be tightly coiled. The replacement of histones is facilitated by transition proteins in round spermatids which prevents the transcription and translation in elongating spermatids since protamine-bound DNA are tightly compacted DNA that cease the transcription and translation (Steger et al., 2000). The chromatin is further re-structured in elongating spermatids, when transition proteins are replaced by protamines (Steger et al., 2000). Both transition protein 1 (TNP1) and 2 (TNP2) play crucial roles in spermiogenesis since they are required for normal chromatin condensation and constitute 90% of basic chromatin proteins useful in steps in between histone removal and replacement by protamines (Meistrich et al., 2003; Zhao et al., 2004). Other than replacing histones, these transition proteins are believed to play roles in obtaining proper nuclear shaping of the sperm, chromatin condensation and maintaining the DNA integrity. Experiments performed on *Tnp1* or *Tnp2* null mice showed preserved fertility (Pradeepa et al., 2008), suggesting the remaining transition protein to compensate for the loss of the other. Nevertheless,

delayed and abnormal DNA condensation was observed, also resulting in a higher level of DNA breaks, structural abnormalities in the head and mid-piece with reduced motility showing its importance in spermiogenesis (Meistrich et al., 2003). Previously, it was found that IPO4 has a higher affinity to TNP2 in pull-down assays compared to other importins (Pradeepa et al., 2008), however due a lack of a mouse-derived antibody for TNP2, TNP1 was selected to be investigated for its co-localisation with IPO4.TNP1 was previously observed in elongating and elongated spermatids (Steger et al., 1998). The current study detected IPO4-TNP1 co-localisation in elongating and elongated spermatids' nuclei (spermatid step 3-5). The localisation analysis could not detect an immunofluorescent signal of TNP1 in round spermatid cytoplasm prior to its localisation in the nucleus. Although a clear demonstration of transport is not possible in IHC/IF, the co-localisation could suggest that TNP1 is a cargo of IPO4 for nuclear import during spermiogenesis.

RAN proteins and RAN binding proteins play a critical role in nuclear transportation and have also been investigated in male infertility (Bao et al., 2014). The importance of RAN-GTP in nuclear transportation and its involvement with importins are well known (Loveland et al., 2015). A previous study showed the presence of RAN-GTPase in the nucleus of round spermatids; and the fact that importin 1 was located to elongating spermatids and the manchette, suggested its involvement in nucleocytoplasmic transportation and ultimately in sperm development (Kierszenbaum et al., 2002). The presence of GTP-bound RAN is suggested to be important for IPO4 to release its cargo protein inside the nucleus. An equilibrium of RAN-GDP in the cytosol and RAN-GTP inside the nucleus is crucial for importins to perform nucleocytoplasmic transportation. Importins can only bind to their cargo in the cytosol when RAN-GDP is present and releases the cargo inside the nucleus upon binding of RAN-GTP (Major et al., 2011). RAN-GTP is not automatically or easily hydrolysed inside the nucleus as Tyr39 inhibits the intrinsic (natural) hydrolysis. In the nucleus, Tyr39 at its right position prevents the positioning of the attacking water which helps in maintaining the RAN-GTP state, that enables importins to dissociate from their cargo (Brucker et al., 2010). In this Chapter, the isoforms of RAN proteins in human sperm were investigated. In mass spectrometry, a previously uncharacterised RAN isoform (B5MDF5; RAN-B5) with an amino acid sequence that is not found in the canonical GTP-bound RAN protein identified in UniProt as P62826 (RAN-P6) was

detected, although not statistically different between NORM and AT groups. It would be interesting to investigate whether RAN-B5 has Tyr39 in the correct position. If Tyr39 was absent, there would be a high possibility of an increased RAN-GTP hydrolysis in the nucleus preventing the dissociation of the cargo from importins. Thus, the functional significance of this novel RAN isoform in nuclearcytoplasmic transportation in germ cells and its involvement in spermatogenesis needs to be investigated.

In conclusion, this study has identified putative cargo proteins of IPO4 and investigated the co-localisation of two of these proteins with IPO4 during spermiogenesis. The study showed a potential role of IPO4 in acrosome biogenesis by showing IPO4 localisation in both Golgi and acrosome, indicating that it could play a role during acrosome biogenesis. A co-localisation of IPO4 and ASF1B in the acrosome granule suggests IPO4 may transport ASF1B during acrosome formation. The co-localisation of IPO4-TNP1 in spermatid nuclei suggests that IPO4 could transport TNP1 into the nucleus from mid-spermiogenesis and thus this transport is likely to be important for spermiogenesis, given the role of transition proteins in chromatin compaction in elongating spermatid nuclei. The fact IPO4 was evident in the acrosome and in spermatid nuclei throughout the elongation phase could point to a key role of IPO4 in development of the functional spermatozoa. This Chapter also provides preliminary evidence that human sperm express a unique isoform of RAN, a protein involved in cytoplasmic-nuclear transport, that is likely functionally different to the canonical RAN protein. Taken together, these results suggest that cytoplasm-nuclear transport during spermiogenesis is an important aspect of human sperm development, and that defects in these processes could contribute to the development of morphologically and functionally abnormal spermatozoa.

4.5 References

Alvarez, F., Munoz, F., Schilcher, P., Imhof, A., Almouzni, G., & Loyola, A. (2011). Sequential establishment of marks on soluble histones H3 and H4. *J Biol Chem*, *286*(20), 17714-17721. https://doi.org/10.1074/jbc.M111.223453

Apta-Smith, M. J., Hernandez-Fernaud, J. R., & Bowman, A. J. (2018). Evidence for the nuclear import of histones H3.1 and H4 as monomers. *EMBO J*, *37*(19). https://doi.org/10.15252/embj.201798714

Bao, J., Tang, C., Li, J., Zhang, Y., Bhetwal, B. P., Zheng, H., & Yan, W. (2014). RAN-binding protein 9 is involved in alternative splicing and is critical for male germ cell development and male fertility. *PLoS Genet*, *10*(12), e1004825. https://doi.org/10.1371/journal.pgen.1004825

Bergmann, M., & Kliesch, S. (2010). Testicular biopsy and histology. In H. M. B. E. Nieschlag, S. Nieschlag (Ed.), *Andrology: Male reproductive health and dysfunction* (3rd edition ed., pp. 158-165). Springer Verlag.

Bernardes, N. E., Fung, H. Y. J., Li, Y., Chen, Z., & Chook, Y. M. (2022). Structure of IMPORTIN-4 bound to the H3-H4-ASF1 histone-histone chaperone complex. *Proc Natl Acad Sci U S A*, *119*(38), e2207177119. https://doi.org/10.1073/pnas.2207177119

Brucker, S., Gerwert, K., & Kotting, C. (2010). Tyr39 of ran preserves the Ran.GTP gradient by inhibiting GTP hydrolysis. *J Mol Biol*, *401*(1), 1-6. https://doi.org/10.1016/j.jmb.2010.05.068

Campos, E. I., Fillingham, J., Li, G., Zheng, H., Voigt, P., Kuo, W. H., Seepany, H., Gao, Z., Day, L. A., Greenblatt, J. F., & Reinberg, D. (2010). The program for processing newly synthesized histones H3.1 and H4. *Nat Struct Mol Biol*, *17*(11), 1343-1351. https://doi.org/10.1038/nsmb.1911

Campos, E. I., Smits, A. H., Kang, Y. H., Landry, S., Escobar, T. M., Nayak, S., Ueberheide, B. M., Durocher, D., Vermeulen, M., Hurwitz, J., & Reinberg, D. (2015). Analysis of the Histone H3.1 Interactome: A Suitable Chaperone for the Right Event. *Mol Cell*, *60*(4), 697-709. https://doi.org/10.1016/j.molcel.2015.08.005

Chalmel, F., Rolland, A. D., Niederhauser-Wiederkehr, C. C., S. S. W., Demougin, P., Gattiker, A., Moore, J., Patard, J., Wolgemuth, D. J., Je´gou, B., & Primig, M. (2007). The conserved transcriptome in human and rodent male gametogenesis. *PNAS*, *104*, 8346–8351. https://doi.org/10.1073/pnas.0701883104.

Fietz, D., & Kliesch, S. (2023). Biopsie und Histologie des Hodens. In E. Nieschlag, H. M. Behre, S. Kliesch, & S. Nieschlag (Eds.), *Andrologie: Grundlagen und Klinik der reproduktiven Gesundheit des Mannes* (pp. 195-210). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-61901-8_11

Fishilevich, S., Nudel, R., Rappaport, N., Hadar, R., Plaschkes, I., Iny Stein, T., Rosen, N., Kohn, A., Twik, M., Safran, M., Lancet, D., & Cohen, D. (2017). GeneHancer: genome-wide integration of enhancers and target genes in GeneCards. *Database (Oxford)*, *2017*. https://doi.org/10.1093/database/bax028

Gunes, S., & Kulac, T. (2013). The role of epigenetics in spermatogenesis. *Turk J Urol*, *39*(3), 181-187. https://doi.org/10.5152/tud.2013.037

Guo, J., Grow, E. J., Mlcochova, H., Maher, G. J., Lindskog, C., Nie, X., Guo, Y., Takei, Y., Yun, J., Cai, L., Kim, R., Carrell, D. T., Goriely, A., Hotaling, J. M., & Cairns, B. R. (2018). The adult human testis transcriptional cell atlas. *Cell Res*, *28*(12), 1141-1157. https://doi.org/10.1038/s41422-018-0099-2

Harel, A., & Forbes, D. J. (2004). Importin beta: conducting a much larger cellular symphony. *Mol Cell*, *16*(3), 319-330. https://doi.org/10.1016/j.molcel.2004.10.026

Jäkel, S., Mingot, J. M., Schwarzmaier, P., Hartmann, E., & Görlich, D. (2002). Importins fulfil a dual function as nuclear import receptors and cytoplasmic chaperones for exposed basic domains. *EMBO J*, *21*(3), 377-386. https://doi.org/10.1093/emboj/21.3.377

Jedrzejczak, P., Kempisty, B., Bryja, A., Mostowska, M., Depa-Martynow, M., Pawelczyk, L., & Jagodzinski, P. P. (2007). Quantitative assessment of transition proteins 1, 2 spermatid-specific linker histone H1-like protein transcripts in spermatozoa from normozoospermic and asthenozoospermic men. *Arch Androl*, *53*(4), 199-205. https://doi.org/10.1080/01485010701426430

Jungwirth, A., Diemer, T., Kopa, Z., Krausz, C., Minhas, S., & Tournaye, H. (2018). *EAU Guidelines on Male Infertility*. European Association of Urology.

Kelley, J. B., Talley, A. M., Spencer, A., Gioeli, D., & Paschal, B. M. (2010). Karyopherin alpha7 (KPNA7), a divergent member of the importin alpha family of nuclear import receptors. *BMC Cell Biol*, *11*, 63. https://doi.org/10.1186/1471-2121-11-63

Khawar, M. B., Gao, H., & Li, W. (2019). Mechanism of Acrosome Biogenesis in Mammals. *Front Cell Dev Biol*, *7*, 195. https://doi.org/10.3389/fcell.2019.00195

Kierszenbaum, A. L., Gil, M., Rivkin, E., & Tres, L. L. (2002). Ran, a GTP-binding protein involved in nucleocytoplasmic transport and microtubule nucleation, relocates from the manchette to the centrosome region during rat spermiogenesis. *Mol Reprod Dev*, *63*(1), 131-140. https://doi.org/10.1002/mrd.10164

Loveland, K. L., Major, A. T., Butler, R., Young, J. C., Jans, D. A., & Miyamoto, Y. (2015). Putting things in place for fertilization: discovering roles for importin proteins in cell fate and spermatogenesis. *Asian J Androl*, *17*(4), 537-544. https://doi.org/10.4103/1008-682X.154310

Mackmull, M. T., Klaus, B., Heinze, I., Chokkalingam, M., Beyer, A., Russell, R. B., Ori, A., & Beck, M. (2017). Landscape of nuclear transport receptor cargo specificity. *Mol Syst Biol*, *13*(12), 962. https://doi.org/10.15252/msb.20177608

Major, A. T., Whiley, P. A., & Loveland, K. L. (2011). Expression of nucleocytoplasmic transport machinery: clues to regulation of spermatogenic development. *Biochim Biophys Acta*, *1813*(9), 1668-1688. https://doi.org/10.1016/j.bbamcr.2011.03.008

Meistrich, M. L., Mohapatra, B., Shirley, C. R., & Zhao, M. (2003). Roles of transition nuclear proteins in spermiogenesis. *Chromosoma*, *111*(8), 483-488. https://doi.org/10.1007/s00412-002-0227-z

Miyamoto, Y., Sasaki, M., Miyata, H., Monobe, Y., Nagai, M., Otani, M., Whiley, P. A. F., Morohoshi, A., Oki, S., Matsuda, J., Akagi, K. I., Adachi, J., Okabe, M., Ikawa, M., Yoneda, Y., Loveland, K. L., & Oka, M. (2020). Genetic loss of importin alpha4 causes abnormal sperm morphology and impacts on male fertility in mouse. *FASEB J*, *34*(12), 16224-16242. https://doi.org/10.1096/fj.202000768RR

Miyauchi, Y., Michigami, T., Sakaguchi, N., Sekimoto, T., Yoneda, Y., Pike, J. W., Yamagata, M., & Ozono, K. (2005). Importin 4 is responsible for ligand-independent nuclear translocation of vitamin D receptor. *J Biol Chem*, *280*(49), 40901-40908. https://doi.org/10.1074/jbc.M509347200

Nathaniel, B., Whiley, P. A. F., Miyamoto, Y., & Loveland, K. L. (2022). Importins: Diverse roles in male fertility. *Semin Cell Dev Biol*, *121*, 82-98. https://doi.org/10.1016/j.semcdb.2021.08.002

Norreen-Thorsen, M., Struck, E. C., Öling, S., Zwahlen, M., Von Feilitzen, K., Odeberg, J., Lindskog, C., Pontén, F., Uhlén, M., Dusart, P. J., & Butler, L. M. (2022). A human adipose tissue cell-type transcriptome atlas. *Cell Rep*, *40*(2), 111046. https://doi.org/10.1016/j.celrep.2022.111046

Panagiotopoulos, A. A., Kalyvianaki, K., Tsodoulou, P. K., Darivianaki, M. N., Dellis, D., Notas, G., Daskalakis, V., Theodoropoulos, P. A., Panagiotidis, C., Castanas, E., & Kampa, M. (2022). Recognition motifs for importin 4 [(L)PPRS(G/P)P] and importin 5 [KP(K/Y)LV] binding, identified by bio-informatic simulation and experimental in vitro validation. *Comput Struct Biotechnol J*, *20*, 5952-5961. https://doi.org/10.1016/j.csbj.2022.10.015

Pontén, F., Gry, M., Fagerberg, L., Lundberg, E., Asplund, A., Berglund, L., Oksvold, P., Bjorling, E., Hober, S., Kampf, C., Navani, S., Nilsson, P., Ottosson, J., Persson, A., Wernerus, H., Wester, K., & Uhlen, M. (2009). A global view of protein expression in human cells, tissues, and organs. *Mol Syst Biol*, *5*, 337. https://doi.org/10.1038/msb.2009.93

Pradeepa, M. M., Manjunatha, S., Sathish, V., Agrawal, S., & Rao, M. R. (2008). Involvement of importin-4 in the transport of transition protein 2 into the spermatid nucleus. *Mol Cell Biol*, *28*(13), 4331-4341. https://doi.org/10.1128/MCB.00519-07

Qian, M. X., Pang, Y., Liu, C. H., Haratake, K., Du, B. Y., Ji, D. Y., Wang, G. F., Zhu, Q. Q., Song, W., Yu, Y., Zhang, X. X., Huang, H. T., Miao, S., Chen, L. B., Zhang, Z. H., Liang, Y. N., Liu, S., Cha, H., Yang, D., . . . Qiu, X. B. (2013). Acetylation-mediated proteasomal degradation of core histones during DNA repair and spermatogenesis. *Cell*, *153*(5), 1012-1024. https://doi.org/10.1016/j.cell.2013.04.032

Soumillon, M., Necsulea, A., Weier, M., Brawand, D., Zhang, X., Gu, H., Barthes, P., Kokkinaki, M., Nef, S., Gnirke, A., Dym, M., de Massy, B., Mikkelsen, T. S., & Kaessmann, H. (2013). Cellular source and mechanisms of high transcriptome complexity in the mammalian testis. *Cell Rep*, *3*(6), 2179-2190. https://doi.org/10.1016/j.celrep.2013.05.031

Steger, K., Klonisch, T., Gavenis, K., Drabent, B., Doenecke, D., & Bergmann, M. (1998). Expression of mRNA and protein of nucleoproteins during human spermiogenesis. *Mol Hum Reprod*, *4*(10), 939-945. https://doi.org/10.1093/molehr/4.10.939

Steger, K., Pauls, K., Klonisch, T., Franke, F. E., & Bergmann, M. (2000). Expression of protamine-1 and -2 mRNA during human spermiogenesis. *Mol Hum Reprod*, *6*(3), 219-225. https://doi.org/10.1093/molehr/6.3.219

Tran, M. H., Aul, R. B., Xu, W., van der Hoorn, F. A., & Oko, R. (2012). Involvement of classical bipartite/karyopherin nuclear import pathway components in acrosomal trafficking and assembly during bovine and murid spermiogenesis. *Biol Reprod*, *86*(3), 84. https://doi.org/10.1095/biolreprod.111.096842

Uhlén, M., Fagerberg, L., Hallstrom, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, A., Kampf, C., Sjostedt, E., Asplund, A., Olsson, I., Edlund, K., Lundberg, E., Navani, S., Szigyarto, C. A., Odeberg, J., Djureinovic, D., Takanen, J. O., Hober, S., . . . Ponten, F. (2015). Proteomics. Tissue-based map of the human proteome. *Science*, *347*(6220), 1260419. https://doi.org/10.1126/science.1260419

Whiley, P. A., Miyamoto, Y., McLachlan, R. I., Jans, D. A., & Loveland, K. L. (2012). Changing subcellular localization of nuclear transport factors during human spermatogenesis. *Int J Androl*, *35*(2), 158-169. https://doi.org/10.1111/j.1365-2605.2011.01202.x

Wing, C. E., Fung, H. Y. J., & Chook, Y. M. (2022). Karyopherin-mediated nucleocytoplasmic transport. *Nat Rev Mol Cell Biol*, *23*(5), 307-328. https://doi.org/10.1038/s41580-021-00446-7

World Health Organization. (2021). *WHO laboratory manual for the examination and processing of human semen*. WHO. https://www.who.int/publications/i/item/9789240030787

Yoon, J., Kim, S. J., An, S., Cho, S., Leitner, A., Jung, T., Aebersold, R., Hebert, H., Cho, U. S., & Song, J. J. (2018). Integrative Structural Investigation on the Architecture of Human Importin4_Histone H3/H4_Asf1a Complex and Its Histone H3 Tail Binding. *J Mol Biol*, *430*(6), 822-841. https://doi.org/10.1016/j.jmb.2018.01.015

Young, J. C., Ly-Huynh, J. D., Lescesen, H., Miyamoto, Y., Browne, C., Yoneda, Y., Koopman, P., Loveland, K. L., & Jans, D. A. (2013). The nuclear import factor importin alpha4 can protect against oxidative stress. *Biochim Biophys Acta*, *1833*(10), 2348-2356. https://doi.org/10.1016/j.bbamcr.2013.06.007

Zhao, M., Shirley, C. R., Mounsey, S., & Meistrich, M. L. (2004). Nucleoprotein transitions during spermiogenesis in mice with transition nuclear protein Tnp1 and Tnp2 mutations. *Biol Reprod*, *71*(3), 1016-1025. https://doi.org/10.1095/biolreprod.104.028191

4.6 Supplementary material

Suppl. Table 4.1: IPO4 localisation in germ cells according to the spermatogenic stages detected in normal spermatogenesis (NSP) biopsies.

Stage/ Cell type	**I**	**II**	**III**	**IV**	**V**	**VI**
Spg	+ (light nuclear) A and B	+ (light nuclear) A and B	Halo (not clear) A	v.light A	+ (light) A	+ (light) A
Early Spc	-	-	pL	+ in L	+ (light) in Z	+ in Z
Pach	v. light	+ (light)	+ (light)	- v. light to no staining Cyto - light	v.light to no staining Cyto +	+ Cyto +
rSptd	Step 1 + light (Golgi and/or acrosomal vesicle)	Step 2 + light (Acrosome)	Step 3 + light (Acrosome)			
Elongating or elongated Sptd	step 7 – not clearly evident	step 8 – not clearly evident		Step 4 + And acrosome	Step 5 +	Step 6 +

+ = stained nuclei - = unstained nuclei, Cyto + = cytoplasmic staining. Any other subcellular staining is mentioned if present. Abbreviation = Spg – spermatogonia, A – type A spermatogonia, B – type B spermatogonia, Spc – spermatocyte, pL – pre leptotene, L – leptotene, Z – zygotene, Pach – pachytene spermatocytes, rSptd – round spermatid, Sptd – spermatid. Grey shade – the germ cell type is absent in the particular stage

Supple. Table 4.2: ASF1B – localisation in germ cells according to the spermatogenic stage detected in normal spermatogenesis (NSP) biopsies.

Stage/ Cell type	**I**	**II**	**III**	**IV**	**V**	**VI**
Spg	+ (v.light)	+ (v.light)	Halo + A	+ (v.light) A	+ (light) A	+ A
Early Spc	-	-	-	-	-	+ in Z
Pach	+ (v.light) and unstained ones	+ (v. light) And unstained ones	-	+ (v.light) to light	+ (light)	+
rSptd	Step 1 + (light)	Step 2 + (light)	Step 3 +			
Elongating and elongated Sptd	Step 7 -	Step 8 -		Step 4 + (v light. Some parts are stained)	Step 5 +	Step 6 -

+ = stained nuclei - = unstained nuclei, Cyto + = cytoplasmic staining. Any other subcellular staining is mentioned if present. Abbreviation = Spg – spermatogonia, A – type A spermatogonia, B – type B spermatogonia, Spc – spermatocyte, pL – pre leptotene, L – leptotene, Z – zygotene, Pach – pachytene spermatocytes, rSptd – round spermatid, Sptd – spermatid. Grey shade – the germ cell type is absent in the particular stage.

Supple Table 4.3: IPO4-ASF1B co-localisation in germ cells according to the spermatogenic stage detected in normal spermatogenesis (NSP) biopsies.

Stage/ Cell type	I	II	III	IV	V	VI
Spg	-	-	-	-	+ (light) IPO4-ASF1B	+ (light) IPO4-ASF1B
Early Spc	-	-	-	-	-	-
Pach	-	-	-	-	+ ASF1B (light) + IPO4 in Cyto.	+ ASF1B (light) + IPO4 in Cyto
rSptd	Step 1 + Prominent ASF1B stained nuclei and IPO4 in Golgi)	Step 2 + Prominent ASF1B stained nuclei and IPO4 in Golgi and acrosome. Beginning of IPO4-ASF1B co-localisation in the acrosome area + light IPO4-ASF1B staining in the nucleus. +Cyto – IPO4	Step 3 + Prominent ASF1B stained nuclei and IPO4 in Golgi and acrosome. co-localisation in the acrosome granule + light IPO4-ASF1B staining in the nucleus. +Cyto – IPO4			
Elongating or elongated Sptd	Step 7 Not clearly evident	Step 8 Not clearly evident		Step 4 ASF1B stained nuclei with IPO4 stained acrosome with co-localisation in the acrosome granule +Cyto – IPO4	Step 5 Prominent IPO4 in the acrosome and nucleus	Step 6 Prominent IPO4 in the acrosome and nucleus

+ = stained nuclei - = unstained nuclei, +Cyto = cytoplasmic staining. Any other subcellular staining is mentioned if present. Abbreviation = Spg – spermatogonia, Spc – spermatocyte, Pach – pachytene spermatocytes, rSptd – round spermatid, Sptd – spermatid. Grey shade – the germ cell type is absent in the particular stage.

Supple Table 4.4: TNP1 localisation in germ cells according to the spermatogenic stage detected in normal spermatogenesis (NSP) biopsies

Stage/ Cell type	**I**	**II**	**III**	**IV**	**V**	**VI**
Spermatid type	Step 1 Step 7	2 8	3	4	5	6
Spg	-	-	-	+ v. light	+ v. light	-
Early Spc	-	-	-	+ v.v Light	+ v.v light	+ v.v light
Pach	-	-	-	+ (v. light)	+ (v. light)	+ (v. light)
rSptd	Step 1 -	Step 2 -	Step 3 +			
Elongating or elongated Sptd	Step 7 Not evident	Step 8 Not evident		Step 4 +	Step 5 +	Step 6 Did not see

+ = stained nuclei - = unstained nuclei. Abbreviation = Spg – spermatogonia, Spc – spermatocyte, Pach – pachytene spermatocytes, rSptd – round spermatid, Sptd – spermatid. Grey shade – the germ cell type is absent in the particular stage.

Supple Table 4.5: IPO4- TNP1 localisation in germ cells according to the spermatogenic stage detected in normal spermatogenesis (NSP) biopsies

Stage/ Cell type	**I**	**II**	**III**	**IV**	**V**	**VI**
Spg	-	-	-	-	-	-
Early Spc	-	-	-	-	-	-
Pach	-	-	-	-	-	-
rSptd	Step 1 -	Step 2 -	Step 3 +			
Elongating or elongated Sptd	Step 7 -	Step 8 -		Step 4 +	Step 5 +	Step 6 Did not see

+ = IPO4-TNP1 stained nuclei - = IPO4-TNP1 unstained nuclei. Abbreviation = Spg – spermatogonia, Spc – spermatocyte, Pach – pachytene spermatocytes, rSptd – round spermatid, Sptd – spermatid. Grey shade – the germ cell type is absent in the particular stage.

Chapter 5

Final Discussion and Outlook

This thesis aimed to discover genes and proteins that could be involved in human sperm development, morphology and/or motility. The project used both genomic and proteomic approaches to address the main research hypotheses; that novel genes and proteins can be identified in human a) testicular tissue and b) ejaculated sperm using well-characterised biological samples with defined physiological and pathological conditions.

Idiopathic infertility accounts for around 30% of male infertility cases (Lin et al., 2006) and although a large number of genes is expressed during spermatogenesis (Coutton et al., 2015; Tüttelmann et al., 2018), the functional roles of the majority of these genes and their respective proteins in promoting normal sperm morphology and motility has not yet been defined. Identifying novel genes and proteins that are involved in human sperm production is important to understand the causes and phenotypes of male infertility and will also have application for the design of non-hormone-based male contraceptives that could specifically inhibit sperm production or function.

In this thesis, it was hypothesised that, by utilising different sample types and technologies, a range of novel genes and proteins with likely roles in human spermiogenesis and sperm function could be identified. Chapter 2 used bulk RNA sequencing (RNAseq) analysis to identify genes that were differentially expressed between human testis biopsies with well-characterised phenotypes and – by exploiting the differing cellular composition of these biopsies - allowed the identification of novel genes involved in spermiogenesis. Whole-testis biopsies with either normal spermatogenesis, arrest of spermatogenesis at the level of round spermatids (spermatid arrest), and a complete loss of germ cells (Sertoli cell only phenotype) were subjected to RNAseq. Differentially expressed genes (DEGs) were analysed via a series of bioinformatic strategies to identify 10 genes of interest. In Chapter 3, quantitative mass spectrometry was used to identify proteins that were significantly different in terms of abundance between ejaculate samples showing normal (normozoospermia, NORM) or abnormal morphology and motility (asthenoteratozoospermia, AT), respectively. Proteins differentially expressed in human AT sperm are excellent candidates for having a functional role in sperm

development and/or function. N=35 differentially expressed proteins (DEPs) were identified between the groups and 5 of them were chosen for further assessment and immunohistochemical analysis. Finally, Chapter 4 investigated IPO4, a nucleocytoplasmic transporter protein that was more abundant in AT sperm in Chapter 3. Based on its known and predicted functions, it was hypothesised that IPO4 may play an important role in sperm development by shuttling proteins between the cytoplasm and the nucleus during spermatid elongation. The results revealed novel roles for IPO4 and two of its putative cargo proteins in acrosome biogenesis and spermatid development. Thus, this thesis has revealed novel genes and proteins that could be important for sperm motility and morphology and could contribute to male infertility.

5.1 RNAseq analysis of human testis biopsies revealed novel genes related to sperm morphology and function.

RNAseq is a powerful tool that is used not only to discover RNA transcripts, but also to compare mRNA expression (Conesa et al., 2016) and identify DEGs between different phenotypes (Stark et al., 2019). Even though next-generation sequencing has evolved from whole-tissue (bulk) RNAseq to single-cell RNAseq (scRNAseq), bulk RNAseq remains one of the most used tools due to its high reproducibility compared to scRNAseq, high sequence coverage and the usage of reliable statistical power and bioinformatical tools (Jeon et al., 2023). The cost-effective nature, the wide availability in many labs and the ease of data interpretation make it more affordable and feasible than scRNAseq.

Chapter 2 took advantage of the availability of well-characterised human testis biopsies revealing different spermatogenic phenotypes. RNAseq was used to identify DEGs between normal spermatogenesis (NSP) samples and two pathological conditions. Biopsies from patients with obstructive azoospermia were categorised as NSP, whereas biopsies from non-obstructive azoospermic patients were identified as spermatid arrest (SDA) or Sertoli cell-only phenotype (SCO). In SDA biopsies, germ cells are present until the phase of round spermatid with elongating spermatids only occasionally observed in some tubules, and thus in whole testis samples, downregulated genes are likely to be mainly comprised of genes predominantly

expressed during human spermiogenesis. Consistent with this, analysis of the RNAseq dataset revealed more than 1000 DEGs between NSP and SDA biopsies, supporting previous studies stating that human spermatogenesis is regulated by more than 2000 genes (Chalmel et al., 2012; Matzuk & Lamb, 2008; Tüttelmann et al., 2018).

This Chapter also analysed Sertoli cell-only (SCO) biopsies, where no germ cells are present. Here, gene expression changes would be expected to arise due to the loss of germ cell genes (likely to be downregulated) and functional changes in other testis somatic cells due to the lack of germ cells (Dutta et al., 2022; O'Donnell et al., 2022; Smith et al., 2015). Taking the complexity of testicular somatic cell populations (Sertoli cells, Leydig cells, peritubular myoid cells, vascular endothelial cells, and immune cells) into account (Dutta et al., 2022; O'Donnell et al., 2022; Washburn & Dufour, 2023; Zhao et al., 2014; Zhou et al., 2019), it was not surprising that DEGs between NSP and SCO accounted for more than 10,000 genes. This list can therefore be used to identify novel germ cell-specific genes as well as to investigate testicular somatic cells during a lack or loss, respectively, of germ cells.

The thesis' main objective was to identify genes associated with sperm morphology and motility. Given that more than 1000 DEGs were found between NSP and SDA, it was necessary to design a bioinformatic strategy to select genes for further analysis. Genes were clustered according to GO/KEGG pathways and UniProt terms related to sperm development, morphology, and motility. Genes related to calcium and cAMP pathways were focused on since these mechanisms are important in cilia and flagella and sperm motility (Mata-Martinez et al., 2021; Pereira et al., 2017; Suarez, 2008; Suarez et al., 1993). Genes annotated as flagella proteins were also selected since abnormalities in flagella usually result in asthenozoospermia (reduced ratio of motile to immotile sperm) (Coutton et al., 2015). Bioinformatic analyses revealed that genes that were downregulated in SDA compared to NSP were associated with relevant functional terms such as 'reproduction', 'spermatogenesis', and 'spermiogenesis' suggesting that downregulated DEGs in SDA are likely predominantly expressed in elongating or elongated spermatids.

Investigation of the DEGs in high-quality databases including ENSEMBL, Human Protein Atlas (HPA), UniProt and GeneCards were used to further short-list genes that

satisfied the criteria as a) being highly enriched or specifically expressed in human spermatids, b) having a likely functional role in spermatozoa of other mammals and c) not having a proven functional role in human spermatozoa (i.e. to identify novel genes). This strategy of filtering genes by cellular expression pattern and functional categories resulted in a list of 10 genes that were selected for further assessment. The copy number expressed in the samples and the degree of fold-change were also considered when selecting genes. The data from the RNAseq analysis, *in silico* analysis utilising public portals and DEG validation by RT-PCR and qRT-PCR revealed that *SPATA31E1, TEKT3, SLC9C1, PDE4A, CFAP47,* and *TNC* are predominantly expressed in human germ cells, particularly in spermatids. Spermiogenesis is a crucial phase in spermatogenesis during which round spermatids undergo a major morphological transformation into elongated spermatids, a process that is accompanied by major changes in gene and protein expression (Jan et al., 2017; O'Donnell, 2014). Thus, genes particularly expressed in spermatids during spermiogenesis are likely to contribute to the structural and functional changes that occur as spermatids develop, ultimately influencing sperm morphology and motility.

Some of the candidate genes that were further analysed for a role in sperm morphology and motility are summarised here. The gene *CFAP47* was considered a promising candidate due to its annotation as a flagella protein and its mRNA expression pattern suggesting a role in sperm function. *CFAP47* mRNA was downregulated in SDA compared to NSP and showed reduced levels in SDA and SCO testis samples by qRT-PCR, suggesting it is expressed in post-meiotic germ cells. Analysis of *CFAP47* expression patterns in the Human Protein Atlas (HPA) confirmed *CFAP47* to be expressed in early and late spermatids (Uhlén et al., 2015) or spermatozoa (Liu et al., 2021), respectively. After this gene was selected for further analysis, a published study revealed that deleterious X-linked variants of CFAP47 are associated with asthenoteratozoospermia (Liu et al., 2021), thus providing confidence that the method used in this thesis for selecting genes associated with sperm function is appropriate. *TEKT3* mRNA is enriched in human spermatids according to HPA (Uhlén et al., 2015), and its protein is detected in sperm flagella according to UniProt. *Tekt3* has shown an association with sperm motility in mice (Roy et al., 2009). *SLC9C1* mRNA is highly enriched in human testis compared to other tissues, with high expression in spermatids (Uhlén et al., 2015), suggesting its expression is mainly

restricted to developing spermatids in humans. After selection of *SLC9C1* as a candidate gene, it was also recently identified as a causative gene for asthenozoospermia (Cavarocchi et al., 2021). Thus, the RNAseq and bioinformatics approach proved to be useful in identifying genes that are expressed during spermiogenesis and are good candidates for a functional role in human sperm development and/or function.

Chapter 2 then focused on further investigating the candidate proteins SPATA31E1 and ORA1 by assessing their localisation in human testis by immunohistochemistry, as this has not been reported previously. SPATA31E1 was particularly detected in post-meiotic germ cells. The summary and implications of SPATA31E1 localisation are discussed in Section 5.3 below. ORAI1 was selected as it has not been reported in human reproduction in the literature previously, although some studies exist in other species. ORAI1 is a calcium-channel transporter class protein, and it is known that calcium regulation is important for sperm motility and function (Pereira et al., 2017; Suarez et al., 1993). A previous study showed that *Orai1*$^{-/-}$ mice were sterile with reduced motility, low sperm count and defects in spermatid elongation (Davis et al., 2016), suggesting its importance in germ cell development. Thus, it was interesting to investigate the localisation of ORAI1 during human spermatogenesis. ORAI1 was localised to human germ cells and Sertoli cells suggesting multiple roles in testis physiology. Its role in Sertoli cells is also of interest, as a previous study in rat testis suggested an involvement of ORAI1 in Sertoli cell intercellular junctions (Lyon et al., 2017). ORAI1 was also detected in the acrosomal, nuclei and mid-piece regions of chicken spermatozoa (Nguyen et al., 2016). These findings in other species, together with the expression and localisation of ORAI1 in human germ cells suggest that ORAI1 could be important in human spermatozoa development, structure, and function.

It is important to note that both SPATA3E1 and ORAI1 were localised in Sertoli cells and, interestingly, showed stage-specific staining in these cells. This apparent cyclic variation in protein localisation in human Sertoli cells was also observed in SCO biopsies. Stage-specific protein expression in human SCO is not well understood, however a previous study in mouse fetal testis demonstrated that Sertoli cells express proteins in a stage-dependent manner even in the absence of germ cells (Timmons et al., 2002). Thus, the findings in this thesis suggest that some proteins in human Sertoli cells could be expressed in a stage-dependent manner in the absence of germ

cells. The fact that these proteins are expressed in both Sertoli cells and germ cells implies they are multifunctional proteins with roles in somatic and germ cells, and/or that different functional isoforms are expressed in the different cell types.

A low sample number was one major limitation in this study. This limitation is difficult to overcome due to the rarity of histologically well-characterised human testis biopsies available not only as archive material for histology/immunohistochemistry but also as fresh-frozen samples for molecular biological analyses. Even though the Testis Biopsy Repository hosted at the Institute of Veterinary Anatomy, Justus-Liebig-University Giessen (JLU) holds an exceptional number of archived human testis biopsies, obtaining high quality fresh-frozen samples of certain pathologies is quite difficult. RNAseq study was conducted with these limited sample numbers due to budget and time constraints. Future studies could use larger sample numbers to verify the changes observed.

Further studies would be required to verify the mRNA expression changes in candidate proteins in a larger cohort of samples, as human spermatogenic phenotypes can be highly variable. Performing IHC for all selected proteins would also be required to better understand their potential roles in human spermiogenesis. In order to identify the functional significance of each of these proteins in spermiogenesis, the most practical approach would be to utilise conditional knock-out mouse models (Smith et al., 2015). For example, genes could be conditionally knocked out during early meiosis and during spermiogenesis, and a careful analysis of their germ cell phenotypes performed to identify their cellular roles (Smith et al., 2015). Adult Sertoli cell conditional knockout models (Giese et al., 2012; Zhou et al., 2011) would also be useful to identify the roles for SPATA31E1 and ORAI1 in these cells during spermatogenesis. Furthermore, it would be interesting to investigate SPATA31E1, ORAI1, TEKT3, and SLC9C1 in human spermatozoa by using electron microscopy to evaluate their subcellular localisation in particular regions of human sperm.

In summary, the RNAseq and bioinformatics approach used in Chapter 2 identified 10 genes, out of which 6 proved to be enriched in human germ cells, particularly in spermatids. Protein localisation of SPATA31E1 and ORAI1 in germ cells, spermatids and Sertoli cells revealed likely multiple roles in spermatogenesis. SPATA31E1 was deemed to be an interesting candidate as a contraceptive drug target because its

cellular expression is restricted to germ cells of the testis, however further functional studies are required to determine whether it is essential for male fertility. The transcriptomic approach used in this chapter was found to be useful in identifying novel genes, potentially playing a role in sperm development and/or function. The dataset generated using well-characterised testis biopsies gives a plethora of data that can be used in future studies to identify genes related to spermatogenesis.

5.2 Mass spectrometry analysis of ejaculated human spermatozoa reveals novel proteins with potential functional roles in sperm morphology and motility

Chapter 3 focused on the hypothesis that using ejaculates with defined physiological and pathological conditions will reveal novel proteins linked to sperm motility/morphology and thus function. For this study, normal human spermatozoa (normozoospermia, NORM) and morphologically defective, immotile spermatozoa (asthenoteratozoospermia, AT) were subjected to mass spectrometry (MS) analysis. A total of 2722 proteins were detected, and 1725 proteins were quantitatively compared between groups. Of these, 35 proteins were identified as significantly different between groups, and were thus classified as differentially expressed proteins (DEPs).

Of the 35 DEPs identified, 10 were up- and 25 were downregulated in AT compared to NORM. Each protein was subjected to *in silico* analysis to gather more information including potential functions in sperm, what is known about their role in the testis, and what is known in terms of their tissue-specific mRNA expression. The results revealed that proteins with known roles in spermatogenesis were significantly different in AT sperm. Out of the DEPs, 23 have been previously investigated in relation to male infertility, e.g. LEFTY1 (Souquet et al., 2012) and TEX37 (Jumeau et al., 2015). Four DEPs have been investigated previously in human samples, e.g. IRGC was decreased in AT in this thesis and was also downregulated in sperm from men with asthenozoospermic (Li et al., 2023) and DEFB129 was increased in AT, and as previously reported, *DEFB129* expression was upregulated in epididymal tissue retrieved from men with non-obstructive azoospermia (Dube et al., 2008).

Five proteins were selected for further analysis based on being a) predominantly or specifically expressed in testis, b) epididymis-specific, c) and/or novel proteins with

limited information in human spermiogenesis. CCDC105 was selected as a novel protein with limited information on its molecular functional role in public portals and no publications in PubMed at the time of selection. UniProt described the gene as an uncharacterised protein coding gene without a known functional role of the protein. However, two very recent studies had reported CCDC105 was detected in bull sperm (Leung et al., 2023) and mouse sperm (Chen et al., 2023). Importantly though, *CCDC105* was highly testis-specific in human, with the highest mRNA expression in late spermatids (Uhlén et al., 2015). Immunohistochemistry results in Chapter 3 confirmed that CCDC105 was expressed in nuclei of all germ cells but not in Sertoli cells. CCDC105 is also known as Tektin-Like 1 (TEKTL1). Tektins are proteins that are localised in centrosomes, centrioles, mitotic spindles, and basal bodies in many species including humans. They are one of the two classes of proteins in ‘’ribbons’’ that are important for axonemal doublet microtubule stability (Linck et al., 2014). The function of Tektins was revealed in knockout mouse models that showed a lack of inner arm dyneins affecting cilia and flagella formation in mice (Tanaka et al., 2004). Also, several studies have shown the importance of different Tektins in sperm motility in mice and men (Bhilawadikar et al., 2013; Roy et al., 2009; Zhang et al., 2016). Since CCDC105 is a Tektin-like protein, it would seem likely to play a role in human sperm flagella motility. This proposition is supported by the fact that CCDC105 abundance was reduced in AT sperm, suggesting that it could be an important contributor to normal human sperm function. Further studies on the role of this protein in sperm development and function are important, including the generation of transgenic mouse models with ablated CCDC105 protein.

ELSPBP1 was chosen for further analysis because it was predicted as epididymis-specific and to be localised in human spermatozoa. This localisation was confirmed in Chapter 3 by IHC using human testis and epididymal tissue. It was of interest to select an epididymis-specific protein due to the important post-testicular maturation that takes place in the epididymis, enabling the testicular spermatozoa to gain its motility during the migration through the epididymis (Cornwall, 2009). Sperm functionally mature as they migrate through the epididymis. This maturation involves the uptake of proteins that are secreted by epididymal cells and by the post-translational modification of spermatozoa proteins upon interactions with epididymal proteins (Björkgren & Sipilä, 2019; Cornwall, 2009). ELSPBP1 is a fibronectin type 2 protein

that is secreted in the epididymis and binds to spermatozoa during epididymal maturation (D'Amours et al., 2012a; Sahin et al., 2009). Thus, this protein could be involved in the post-testicular maturation and function of human sperm. There is limited knowledge on ELSPBP1 in human spermatozoa. However, in cattle, ELSPBP1 was detected in the post-acrosomal, and mid-piece regions of spermatozoa retrieved from the cauda epididymis (D'Amours et al., 2012b) suggesting a conserved role for ELSPBP1 in post-testicular sperm maturation and function. Interestingly, ELSPBP1 levels were elevated in human sperm with high DNA fragmentation (Belardin et al., 2022) and ELSPBP1 tends to bind to dead bovine spermatozoa retrieved from cauda epididymis (D'Amours et al., 2012a; D'Amours et al., 2012b). These latter findings, together with the data produced in this thesis showing that ELSPBP1 is elevated in AT sperm compared to NORM, suggests that ELSPBP1 is produced in the epididymis but differentially binds to spermatozoa depending on their phenotype and/or functional state. Whether ELSPBP1 has a higher tendency to bind to dead and/or deformed human spermatozoa should be confirmed, as this property could be useful as a marker of sperm dysfunction. The measurement of ELSPBP1 in percol-separated morphologically abnormal and immotile spermatozoa vs normal spermatozoa by flow cytometry would reveal whether the levels of this protein in sperm correlates with sperm dysfunction. The levels of ELSPBP1 could also be quantified by ELISA in a larger sample number of patients to investigate whether ELSPBP1 could be used as a diagnostic marker of human sperm morphology. This could lead to the possibility that the level of the protein in a semen sample revealing the underlying pathogenesis of the patient. Also, confocal microscopy could be used to examine binding of ELSPBP1 to different regions of normal and pathological spermatozoa, thus providing more information on the functional roles of this protein in male infertility.

ACTRT2 was significantly reduced in AT and was further investigated because its mRNA and protein are highly enriched in the human testis. IHC in Chapter 3 confirmed its specific localisation in round and elongating spermatids in the normal human testis. This finding is supported by HPA data showing *ACTRT2* is restricted to late spermatids (Uhlén et al., 2015). Other studies have shown that ACTRT2 is localised to the post-acrosomal region in human spermatozoa (Liu et al., 2015) and to the sub-acrosomal region in mouse spermatids (Zhang et al., 2022). Reduced expression of ACTRT2 was reported in sperm from men with obesity-induced asthenozoospermia (Liu et al., 2015)

and sperm from men undergoing scrotal hypothermia (Wu et al., 2020). Taken together, the findings suggest that a reduction in ACTRT2 in AT sperm could either reflect, or be causative for, sperm dysfunction. The molecular function of ACTRT2 during sperm development could be elucidated in transgenic mouse knockout models and using ACTRT2-inhibiting antibodies or pharmacological agents in *in vitro* models of human spermatozoa. ACTRT2 levels in human sperm could also be explored for their utility in being able to discriminate different phenotypes of human spermatozoa during fertility assessment.

IFT57 was selected because it belongs to the group of intraflagella transport proteins that are required for the assembly and maintenance of cilia and flagella (Ge et al., 2021; Zhang et al., 2020) and the normal development of sperm flagellum is crucial for sperm motility. IFT57 was more abundant in AT sperm. IHC analysis revealed a distinct protein pattern of IFT57 in late pachytenes and round spermatids. IFT57 was located at different subcellular sites in human germ cells, particularly evident in the cytoplasm of late pachytene spermatocytes. During spermiogenesis, IFT57 showed a translocation from the cytoplasm to the nucleus around step 3-5 at the commencement of spermatid elongation, suggesting multiple roles during spermatid development. Though the molecular function of IFT57 in human sperm has not been investigated, many other intra-flagellar proteins play important roles in flagella formation and function in mouse spermatogenesis. For example, *Ift81* and *Ift72* are involved in mouse spermiogenesis and fertility (Qu et al., 2020; Zhang et al., 2020) and *Ift74* knock-out mice displayed low sperm counts and immotile sperm (Shi et al., 2019). Reduced IFT57 levels were reported in *Ift72* (Zhang et al., 2020) and *Ift74* deficient mice (Shi et al., 2019) although the functional significance of this decrease was not defined. In order to define the function of IFT57 in human spermiogenesis, a conditional transgenic mouse model, specifically deleting this gene in germ cells, would be required as complete *Ift57*$^{-/-}$ knockout mice show embryonic lethality (Houde et al., 2006). IFT57 was shown in all organs in HPA (Pontén et al., 2009) suggesting that it is expressed in many cells, potentially involving primary cilia, but the highest mRNA expression was detected in spermatids (Uhlén et al., 2015). Based on the protein localisation of IFT57 in human germ cells, a deletion of IFT57 in germ cells would likely result in disordered sperm development, morphology and/or motility. Further studies could use confocal microscopy or immunoelectron microscopy to

examine the subcellular localisation of IFT57 and identify the functional significance of its elevation in AT sperm.

5.3 The role of SPATA31E1 in sperm morphology and motility as identified by both transcriptomic and proteomic approaches

SPATA31E1 was identified in Chapter 2 as being downregulated in testis biopsies with a phenotype of SDA and was thus predicted to have a role in spermatogenesis. However, its involvement in human sperm development/function has not been described. IHC analysis of SPATA31E1 in testis biopsies revealed stage-specific staining of pachytene spermatocyte nuclei and staining of elongating spermatids until step 4-5, underlining the potential significance in human spermatogenesis. The localisation of SPATA31E1 in spermatozoa has not previously been reported, however it was noted that mature spermatids prior to sperm release did not show any immunostaining. However, SPATA31E1 was identified in the human ejaculated sperm proteome (Chapter 3), suggesting this protein remains in spermatozoa during their entire development. SPATA31E1 was also one of the 35 DEPs identified in Chapter 3 and was reduced 2-fold in AT sperm. Thus, SPATA31E1 likely plays an important role in human sperm function.

SPATA31E1 is not well described, and the molecular function is not understood. A study in stallions addressing single nucleotide polymorphisms (SNPs) in *Spata31e1* suggested certain SNPs may be associated with a deleterious effect on fertility (Schrimpf et al., 2016). An RNAseq study listed *SPATA31E1* as a spermatid-enriched gene and a potential drug target (Robertson et al., 2020). A recent transcriptome-wide analysis of expression quantitative trait loci (eQTLs) using available RNAseq datasets, listed *SPATA31E1* as a gene related to spermatogenesis (An & Lee, 2022). The results described in this thesis are the first to suggest that SPATA31E1 could be involved in human sperm development and function.

As was predicted in GeneCards, *SPATA31E1* likely plays a role in spermatogenesis. UniProt did not have any information on its function at the beginning of this study and it had no functional annotations to determine potential functions of the protein, other than predicting its biological processes as cell differentiation and spermatogenesis. However, recent UniProt updates show that it is predicted to a) be a glycosylated

membrane protein and a member of the transmembrane helix family proteins b) contain domains with homology to the SPATA31F-like protein in InterPro, and c) an acrosome-specific protein 1 (PTHR21859) as listed in PANTHER. SPATA31E1 is a relatively large protein with 1445 amino acids showing a helical structure. However, functional domains and protein binding sites have not been described. Also, no phenotypes are mentioned in the Mouse Genome Informatics (MGI) database with no records of strains or lines available in the International Mouse Strain Resource (IMSR).

Importantly, a very recent study has reported that SPATA31E1 showed significant changes in abundance in sperm from men with normospermia, oligozoospermia and oligoasthenoteratozoospermia (OAT), where SPATA31E1 was downregulated in sperm from patients with oligospermia or OAT compared to normospermic patients (Greither et al., 2023). These recent findings, together with the data in this thesis showing that SPATA31E1 protein is significantly less abundant in sperm with an AT phenotype, suggests that this protein likely plays a key role in human sperm function. *SPATA31E1* was enriched in germ cells and the protein localisation was prominent in germ cells during spermiogenesis (Chapter 2) and HPA reports both mRNA (Uhlén et al., 2015) and protein (Pontén et al., 2009) are testis specific. Further studies investigating its subcellular localisation in human sperm to determine whether this localisation is altered in sperm with abnormal morphology would be useful to understand its functional role in spermatozoa. SPATA31E1 could also be investigated as a target for the development of a male contraceptive, since recent approaches to male contraceptive development tend to focus on proteins that are essential and specific for sperm production and/or function which limits the likelihood of off-target effects (Long et al., 2021). However, the molecular function of SPATA31E1 needs to be elucidated and its cellular specificity needs to be further confirmed before it is considered as a contraceptive target. Functional characterisation of SPATA31E1 during spermatogenesis and sperm function could be explored using a conditional knockout mouse model where SPATA31E1 is conditionally deleted in germ cells prior to spermiogenesis. If SPATA31E1 is essential for male fertility (e.g. using transgenic mouse models), then further studies should investigate whether SPATA31E1 function in the testis or spermatozoa can be inhibited via pharmacological agents.

5.4 The role of IPO4 in spermiogenesis

Chapter 4 focused on a specific protein, IPO4, that was identified as being more abundant in sperm with a phenotype of AT and was shown to be localised to human germ cells in Chapter 3. Chapter 4 hypothesised that elevated IPO4 levels detected in AT sperm could reflect a disturbance in IPO4 cytoplasmic to nuclear transport during spermiogenesis, and this could involve disordered transport of IPO4 cargo proteins. The molecular function of IPO4 is poorly understood in human male fertility, however other importins are known to have functional roles in spermatogenesis and male fertility as specific cargo proteins (Bernardes et al., 2022; Loveland et al., 2015; Miyamoto et al., 2020; Nathaniel et al., 2022; Whiley et al., 2012; Young et al., 2013). IPO4 is a beta importin involved in the transport of proteins from the cytoplasm to the nucleus. IHC results suggested that IPO4 could be important in spermiogenesis, as it was localised in the nucleus and cytoplasm of round and elongating spermatids. Chapter 4 investigated potential roles of IPO4 in sperm development during spermiogenesis by focusing on its co-localisation with two putative cargo proteins (ASF1B and TNP1) in human germ cells. Dual-label immunofluorescence and confocal microscopy was used to examine the cellular and subcellular sites at which IPO4 and these cargo proteins co-localised in order to understand at which points during germ cell development these cargo proteins could be reliant on IPO4-mediated transport and its biological significance in spermatogenesis.

ASF1B was previously shown in a complex with IPO4 along with histones H3 and H4 in HeLa cell cytosol extractions (Alvarez et al., 2011), suggesting ASF1B could be a cargo of IPO4. ASF1B is a histone chaperone protein with roles in promoting tumor growth in many cancers. It is linked to functions in regulating cell cycle, DNA replication and replication stress, and ASF1B deficiency causes apoptosis and cell cycle arrest (Jasencakova et al., 2010; Liu et al., 2020; Ouyang et al., 2021; Tamburini et al., 2005; Wang et al., 2022; Zhang et al., 2021). It is also involved in oocyte meiosis signaling (Ouyang et al., 2021) and chromatin re-organisation (Evans et al., 2014). *ASF1B* is known to be expressed in human spermatocytes and spermatids (Uhlén et al., 2015). However, its localisation in human spermatogenesis has not yet been described.

The detection of IPO4 and ASF1B at various sites during spermiogenesis provides evidence that both could be important for the development of functionally normal

sperm. ASF1B binds to histone H3-H4 dimers (Alvarez et al., 2011) and various histone dimers are involved in the nuclear changes during germ cell development (Moritz & Hammoud, 2022; Steger et al., 1998). Consistent with this proposition, immunohistochemical localisation studies in Chapter 4 demonstrated that ASF1B is detected in the nuclei of spermatids as they transition from round to elongating spermatids. These observations suggest that ASF1B could play a role in the development of the sperm head. Dual-label immunohistochemistry did not reveal convincing evidence that ASF1B is primarily transported to the nucleus by IPO4, because nuclear co-localisation was not obvious. It is possible though that discrete areas of IPO4 localisation in the spermatid nucleus could have been obscured by the strong nuclear staining of ASF1B. ASF1B immunostaining disappeared from the spermatid nucleus around step 4, although IPO4 immunostaining remained. Two recent studies have shown that IPO4 binds to individual H3 and H4 monomers in cytosol in human cell lines (Apta-Smith et al., 2018), that IPO4-H3-H4-ASF1 exists in a complex, and that binding of RAN-GTP to IPO4 facilitates the release of H3-H4-ASF1 in the nuclei of *Escherichia coli* BL21 cells (Bernardes et al., 2022). Histones in the spermatid nuclei are replaced by transition proteins during spermatid elongation (Steger et al., 2000). Thus, the disappearance of ASF1B from the nucleus during spermatid elongation could reflect the disappearance of histones as they are replaced by transition proteins, and ultimately, protamines.

Chapter 4 revealed the co-localisation of IPO4-ASF1B using dual-label immunofluorescence in round spermatids during acrosome formation, suggesting a novel role for these proteins in acrosome biogenesis. IPO4 and ASF1B co-localised in the developing human acrosome, first in the acroplaxome area where the acrosomal vesicle contacts the nucleus, and then more broadly in the acrosome as it spread across the nucleus. This is the first demonstration of both IPO4 and ASF1B in the acrosome, and the first suggestion that IPO4 may transport ASF1B as a cargo protein during acrosome development. These observations have several implications. First, that IPO4-mediated protein transport may occur in the acrosome. Acrosome formation involves protein transport from the Golgi apparatus to the acroplaxome and acrosomal granules via microtubules (O'Donnell & O'Bryan, 2014). Thus, it is possible that IPO4 transport is not limited to cytoplasmic-to-nuclear transport but could also involve transport proteins along microtubule tracks. Second, it suggests a role for ASF1B in

acrosome formation, however the role of ASF1B in acrosome biogenesis remains to be investigated.

There is some data on roles for importins in the developing acrosome. A previous study localised histone H2B variant (SubH2B) in the sub-acrosomal compartment of bovine spermatozoa and the importin protein karyopherin alpha in acrosomal vesicles (Tran et al., 2012). In this previous study, SubH2B was overexpressed in somatic cells to observe whether it could be imported to the nucleus since it contains NLS motifs, but it only showed an interaction with importin alpha 6 (Impα6) without entering the nucleus. Impα6 was observed first on the outer surface of acrosome vesicles and then between the inner acrosomal membrane and the nuclear envelope. These findings suggested that Impα6 could assist in the transportation of acrosomal vesicles from the Golgi to the nuclear envelope (Tran et al., 2012). These findings are supported by other studies suggesting various importin alphas are involved in acrosome vesicle trafficking (Arjomand et al., 2014; Loveland et al., 2015). IPO4 localisation alone and co-localisation of IPO4-ASF1B in the spermatid acrosome in Chapter 4 provide evidence to suggest that IPO4 could have a role in acrosome vesicle trafficking and acrosome biogenesis. The biological significance of ASF1B localisation in the acrosome is not known. However, considering previous studies on histone localisation in the acrosome (Tran et al., 2012) and the importance of SETD2, a methyl transferase of histone H3, in spermiogenesis and acrosome formation in mice (Zuo et al., 2018), and binding links between histone H3, H4 with ASF1B (Alvarez et al., 2011; Apta-Smith et al., 2018), it could be hypothesised that ASF1B together with IPO4 are involved in histone transportation to the developing acrosome during spermiogenesis.

Taken altogether, these findings suggest that ASF1B is likely to be transported into the spermatid nucleus by other mechanisms, but it remains possible that IPO4 could transport histones to the nucleus by forming a complex with ASF1B. Thus, ASF1B may have multiple roles in spermatid development, including an interaction with IPO4 that could be important for histone nuclear transport. Histones H3 and H4 play very important roles in DNA condensation in the nucleus (Qian et al., 2013; Steger et al., 2000) and thus this IPO4-mediated transportation could be important for sperm development. As a next step, a conditional deletion of IPO4 in round spermatids would be helpful to determine the differences in the localisation of ASF1B when IPO4 is absent. In order to determine whether IPO4 is needed for the transportation of ASF1B-

H3-H4 complex to the nucleus and whether H3-H4 dimer needs ASF1B to be transported by IPO4, investigation of ASF1B knockout and IPO4 knockout germ cell lines would be needed to quantify the H3-H4, IPO4, and/or ASF1B levels in the nucleus.

Chapter 4 also investigated the co-localisation of IPO4 and transition protein 1, TNP1. Transition proteins (TNPs) are involved in chromatin condensation in the nucleus during spermiogenesis; they replace the histones in round spermatid chromatin as the nucleus begins to condense, and are subsequently replaced by protamines that further condense the chromatin during the elongation phase (Hao et al., 2019; Liu et al., 2019; Qian et al., 2013). There are major changes in chromatin packaging during this phase (Qian et al., 2013; Steger et al., 2000). Apart from replacement of histones, TNPs are important in proper nuclear shaping of the sperm and maintaining the DNA integrity along with protamines (Lüke et al., 2014; Meistrich et al., 2003). Studies using null *Tnp1* and *Tnp2* mice have shown delayed and abnormal DNA condensation with high rate of DNA breaks and structural abnormalities in sperm head mid-piece resulting reduced motility (Meistrich et al., 2003). As transition proteins play such important roles during spermiogenesis, it was interesting to investigate whether transition proteins could be delivered to the nucleus by IPO4. Both TNP1 and TNP2 showed high binding intensity to IPO4 in mice (Pradeepa et al., 2008) and TNP1 is specific to the testis (Pontén et al., 2009) and shows restricted mRNA expression in spermatids (Uhlén et al., 2015) underlining its importance for nuclear condensation during spermiogenesis. Thus, TNP1 was selected as a candidate cargo protein of IPO4 during human spermiogenesis. Immunofluorescence analysis confirmed that TNP1 is highly restricted to human elongating spermatids, only evident in spermatids of steps 3- 5, and this timing is consistent with a role for IPO4-mediated transport to the nucleus. These results suggest that the localisation of TNP1 to the nucleus during these discrete steps of human spermatid development could be facilitated by IPO4 cytoplasmic to nuclear transport. This proposition requires confirmation, such as by investigating TNP1 and 2 localisations to the nucleus in IPO4 deficient human cell lines. Also, a conditional knockout of IPO4 in spermatids using a spermatid-specific Cre transgene would be useful to investigate whether TNP1 and TNP2 localisation in spermatid nuclei is disrupted in the absence of IPO4.

5.5 Conclusions

In conclusion, this thesis succeeded in utilising RNAseq and mass spectrometry, and well characterised human testis biopsy and sperm samples to identify novel genes and proteins that could have important roles in human sperm development, particularly in relation to sperm morphology and motility. A number of novel genes and proteins that are of likely importance for human spermiogenesis and sperm function were identified. This thesis identified six genes (*SPATA31E1, CFAP47, TEKT3, TNC, PDE4A*, and *SLC9C1*) that are highly expressed in human spermatids. SPATA31E1 was further identified as a key testis-specific protein of unknown function that could play a role in human spermatozoa. This thesis identified 35 proteins that were altered in morphologically abnormal, immotile ejaculated spermatozoa. ACTRT2, IFT57, ELSPBP1, CCDC105, and IPO4 were further investigated and their localisation during human germ cell development was defined. ACTRT2, IFT57, and IPO4 were localised to the developing spermatids and likely play multiple roles in spermatogenesis. ELSPBP1 was identified as an epididymal-derived protein that could bind to sperm in a morphology-dependent manner and thus could be explored as a diagnostic marker for the identification of abnormal spermatozoa. This thesis identified IPO4 as a protein that is likely important in spermiogenesis, as it was expressed during spermatid development in the human testis and was more abundant in AT spermatozoa. Since IPO4 is a transporter protein involved in the transportation of proteins from the cytoplasm to the nucleus, this study used confocal microscopy to map its localisation with two putative cargo proteins to understand their potential roles in human spermiogenesis. The results revealed that IPO4 and TNP1 co-localised in the nucleus of human spermatids during nuclear condensation, consistent with a role of IPO4 in transporting TNP1 to the nucleus to initiate the process of condensation. Studies on IPO4 and ASF1B revealed an unexpected co-localisation in the developing acrosome, suggesting that IPO4 may participate in other transport mechanisms and could be involved in acrosome biogenesis. Overall, this thesis contributes to a better understanding of human sperm development and reveals novel candidates that could be further explored for their utility as diagnostic markers of human sperm function, for their role in human male infertility and for their utility as targets for male contraception.

5.6 References

Alvarez, F., Munoz, F., Schilcher, P., Imhof, A., Almouzni, G., & Loyola, A. (2011). Sequential establishment of marks on soluble histones H3 and H4. *J Biol Chem*, *286*(20), 17714-17721. https://doi.org/10.1074/jbc.M111.223453

An, Y., & Lee, C. (2022). Mixed model-based eQTL analysis reveals lncRNAs associated with regulation of genes involved in sex determination and spermatogenesis: The key to understanding human gender imbalance. *Comput Biol Chem*, *99*, 107713. https://doi.org/10.1016/j.compbiolchem.2022.107713

Apta-Smith, M. J., Hernandez-Fernaud, J. R., & Bowman, A. J. (2018). Evidence for the nuclear import of histones H3.1 and H4 as monomers. *EMBO J*, *37*(19). https://doi.org/10.15252/embj.201798714

Arjomand, A., Baker, M. A., Li, C., Buckle, A. M., Jans, D. A., Loveland, K. L., & Miyamoto, Y. (2014). The alpha-importome of mammalian germ cell maturation provides novel insights for importin biology. *FASEB J*, *28*(8), 3480-3493. https://doi.org/10.1096/fj.13-244913

Belardin, L. B., Antoniassi, M. P., Camargo, M., Intasqui, P., & Bertolla, R. P. (2022). Separating the chaff from the wheat: antibody-based removal of DNA-fragmented sperm. *Hum Reprod*. https://doi.org/10.1093/humrep/deac260

Bernardes, N. E., Fung, H. Y. J., Li, Y., Chen, Z., & Chook, Y. M. (2022). Structure of IMPORTIN-4 bound to the H3-H4-ASF1 histone-histone chaperone complex. *Proc Natl Acad Sci U S A*, *119*(38), e2207177119. https://doi.org/10.1073/pnas.2207177119

Bhilawadikar, R., Zaveri, K., Mukadam, L., Naik, S., Kamble, K., Modi, D., & Hinduja, I. (2013). Levels of Tektin 2 and CatSper 2 in normozoospermic and oligoasthenozoospermic men and its association with motility, fertilization rate, embryo quality and pregnancy rate. *J Assist Reprod Genet*, *30*(4), 513-523. https://doi.org/10.1007/s10815-013-9972-6

Björkgren, I., & Sipilä, P. (2019). The impact of epididymal proteins on sperm function. *Reproduction*, *158*(5), R155-r167. https://doi.org/10.1530/rep-18-0589

Cavarocchi, E., Whitfield, M., Chargui, A., Stouvenel, L., Lores, P., Coutton, C., Arnoult, C., Santulli, P., Patrat, C., Thierry-Mieg, N., Ray, P. F., Dulioust, E., & Toure, A. (2021). The sodium/proton exchanger SLC9C1 (sNHE) is essential for human sperm motility and fertility. *Clin Genet*, *99*(5), 684-693. https://doi.org/10.1111/cge.13927

Chalmel, F., Lardenois, A., Evrard, B., Mathieu, R., Feig, C., Demougin, P., Gattiker, A., Schulze, W., Jegou, B., Kirchhoff, C., & Primig, M. (2012). Global human tissue profiling and protein network analysis reveals distinct levels of transcriptional germline-specificity and identifies target genes for male infertility. *Hum Reprod*, *27*(11), 3233-3248. https://doi.org/10.1093/humrep/des301

Chen, Z., Shiozaki, M., Haas, K. M., Skinner, W. M., Zhao, S., Guo, C., Polacco, B. J., Yu, Z., Krogan, N. J., Lishko, P. V., Kaake, R. M., Vale, R. D., & Agard, D. A. (2023). De novo protein identification in mammalian sperm using in situ cryoelectron tomography and AlphaFold2 docking. *Cell*. https://doi.org/10.1016/j.cell.2023.09.017

Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., McPherson, A., Szczesniak, M. W., Gaffney, D. J., Elo, L. L., Zhang, X., & Mortazavi, A. (2016). A survey of best practices for RNA-seq data analysis. *Genome Biol*, *17*, 13. https://doi.org/10.1186/s13059-016-0881-8

Cornwall, G. A. (2009). New insights into epididymal biology and function. *Hum Reprod Update*, *15*(2), 213-227. https://doi.org/10.1093/humupd/dmn055

Coutton, C., Escoffier, J., Martinez, G., Arnoult, C., & Ray, P. F. (2015). Teratozoospermia: spotlight on the main genetic actors in the human. *Hum Reprod Update*, *21*(4), 455-485. https://doi.org/10.1093/humupd/dmv020

D'Amours, O., Bordeleau, L. J., Frenette, G., Blondin, P., Leclerc, P., & Sullivan, R. (2012a). Binder of sperm 1 and epididymal sperm binding protein 1 are associated with different bull sperm subpopulations. *Reproduction*, *143*(6), 759-771. https://doi.org/10.1530/REP-11-0392

D'Amours, O., Frenette, G., Bordeleau, L. J., Allard, N., Leclerc, P., Blondin, P., & Sullivan, R. (2012b). Epididymosomes transfer epididymal sperm binding protein 1 (ELSPBP1) to dead spermatozoa during epididymal transit in bovine. *Biol Reprod*, *87*(4), 94. https://doi.org/10.1095/biolreprod.112.100990

Davis, F. M., Goulding, E. H., D'Agostin, D. M., Janardhan, K. S., Cummings, C. A., Bird, G. S., Eddy, E. M., & Putney, J. W. (2016). Male infertility in mice lacking the store-operated Ca(2+) channel Orai1. *Cell Calcium*, *59*(4), 189-197. https://doi.org/10.1016/j.ceca.2016.02.007

Dube, E., Hermo, L., Chan, P. T., & Cyr, D. G. (2008). Alterations in gene expression in the caput epididymides of nonobstructive azoospermic men. *Biol Reprod*, *78*(2), 342-351. https://doi.org/10.1095/biolreprod.107.062760

Dutta, S., Sandhu, N., Sengupta, P., Alves, M. G., Henkel, R., & Agarwal, A. (2022). Somatic-Immune Cells Crosstalk In-The-Making of Testicular Immune Privilege. *Reprod Sci*, *29*(10), 2707-2718. https://doi.org/10.1007/s43032-021-00721-0

Evans, E., Hogarth, C., Mitchell, D., & Griswold, M. (2014). Riding the spermatogenic wave: profiling gene expression within neonatal germ and sertoli cells during a synchronized initial wave of spermatogenesis in mice. *Biol Reprod*, *90*(5), 108. https://doi.org/10.1095/biolreprod.114.118034

Ge, T. T., Yuan, L., Xu, W. H., & Zheng, Y. (2021). Role and mechanism of intraflagellar transport in mammalian spermiogenesis. *Yi Chuan*, *43*(11), 1038-1049. https://doi.org/10.16288/j.yczz.21-206

Giese, S., Hossain, H., Markmann, M., Chakraborty, T., Tchatalbachev, S., Guillou, F., Bergmann, M., Failing, K., Weider, K., & Brehm, R. (2012). Sertoli-cell-specific knockout of connexin 43 leads to multiple alterations in testicular gene expression in prepubertal mice. *Dis Model Mech*, *5*(6), 895-913. https://doi.org/10.1242/dmm.008649

Greither, T., Dejung, M., Behre, H. M., Butter, F., & Herlyn, H. (2023). The human sperm proteome-Toward a panel for male fertility testing. *Andrology*, *11*(7), 1418-1436. https://doi.org/10.1111/andr.13431

Hao, S. L., Ni, F. D., & Yang, W. X. (2019). The dynamics and regulation of chromatin remodeling during spermiogenesis. *Gene*, *706*, 201-210. https://doi.org/10.1016/j.gene.2019.05.027

Houde, C., Dickinson, R. J., Houtzager, V. M., Cullum, R., Montpetit, R., Metzler, M., Simpson, E. M., Roy, S., Hayden, M. R., Hoodless, P. A., & Nicholson, D. W. (2006). Hippi is essential for node cilia assembly and Sonic hedgehog signaling. *Dev Biol*, *300*(2), 523-533. https://doi.org/10.1016/j.ydbio.2006.09.001

Jan, S. Z., Vormer, T. L., Jongejan, A., Roling, M. D., Silber, S. J., de Rooij, D. G., Hamer, G., Repping, S., & van Pelt, A. M. M. (2017). Unraveling transcriptome dynamics in human spermatogenesis. *Development*, *144*(20), 3659-3673. https://doi.org/10.1242/dev.152413

Jasencakova, Z., Scharf, A. N., Ask, K., Corpet, A., Imhof, A., Almouzni, G., & Groth, A. (2010). Replication stress interferes with histone recycling and predeposition marking of new histones. *Mol Cell*, *37*(5), 736-743. https://doi.org/10.1016/j.molcel.2010.01.033

Jeon, H., Xie, J., Jeon, Y., Jung, K. J., Gupta, A., Chang, W., & Chung, D. (2023). Statistical Power Analysis for Designing Bulk, Single-Cell, and Spatial Transcriptomics Experiments: Review, Tutorial, and Perspectives. *Biomolecules*, *13*(2). https://doi.org/10.3390/biom13020221

Jumeau, F., Com, E., Lane, L., Duek, P., Lagarrigue, M., Lavigne, R., Guillot, L., Rondel, K., Gateau, A., Melaine, N., Guével, B., Sergeant, N., Mitchell, V., & Pineau, C. (2015). Human Spermatozoa as a Model for Detecting Missing Proteins in the Context of the Chromosome-Centric Human Proteome Project. *J Proteome Res*, *14*(9), 3606-3620. https://doi.org/10.1021/acs.jproteome.5b00170

Leung, M. R., Zeng, J., Wang, X., Roelofs, M. C., Huang, W., Zenezini Chiozzi, R., Hevler, J. F., Heck, A. J. R., Dutcher, S. K., Brown, A., Zhang, R., & Zeev-Ben-Mordehai, T. (2023). Structural specializations of the sperm tail. *Cell*, *186*(13), 2880-2896 e2817. https://doi.org/10.1016/j.cell.2023.05.026

Li, J., Xu, X., Liu, J., Zhang, S., Wang, T., Liu, Y., & Wang, Z. (2023). The immunity-related GTPase IRGC mediates interaction between lipid droplets and mitochondria to facilitate sperm motility. *FEBS Lett*, *597*(12), 1595-1605. https://doi.org/10.1002/1873-3468.14640

Lin, Y. H., Lin, Y. M., Teng, Y. N., Hsieh, T. Y., Lin, Y. S., & Kuo, P. L. (2006). Identification of ten novel genes involved in human spermatogenesis by microarray analysis of testicular tissue. *Fertil Steril*, *86*(6), 1650-1658. https://doi.org/10.1016/j.fertnstert.2006.04.039

Linck, R., Fu, X., Lin, J., Ouch, C., Schefter, A., Steffen, W., Warren, P., & Nicastro, D. (2014). Insights into the structure and function of ciliary and flagellar doublet microtubules: tektins, Ca2+-binding proteins, and stable protofilaments. *J Biol Chem*, *289*(25), 17427-17444. https://doi.org/10.1074/jbc.M114.568949

Liu, C., Tu, C., Wang, L., Wu, H., Houston, B. J., Mastrorosa, F. K., Zhang, W., Shen, Y., Wang, J., Tian, S., Meng, L., Cong, J., Yang, S., Jiang, Y., Tang, S., Zeng, Y., Lv, M., Lin, G., Li, J., . . . Zhang, F. (2021). Deleterious variants in X-linked CFAP47 induce asthenoteratozoospermia and primary male infertility. *Am J Hum Genet*, *108*(2), 309-323. https://doi.org/10.1016/j.ajhg.2021.01.002

Liu, X., Song, J., Zhang, Y., Wang, H., Sun, H., Feng, X., Hou, M., Chen, G., Tang, Q., & Ji, M. (2020). ASF1B promotes cervical cancer progression through stabilization of CDK9. *Cell Death Dis*, *11*(8), 705. https://doi.org/10.1038/s41419-020-02872-5

Liu, Y., Guo, Y., Song, N., Fan, Y., Li, K., Teng, X., Guo, Q., & Ding, Z. (2015). Proteomic pattern changes associated with obesity-induced asthenozoospermia. *Andrology*, *3*(2), 247-259. https://doi.org/10.1111/andr.289

Liu, Y., Zhang, Y., Yin, J., Gao, Y., Li, Y., Bai, D., He, W., Li, X., Zhang, P., Li, R., Zhang, L., Jia, Y., Zhang, Y., Lin, J., Zheng, Y., Wang, H., Gao, S., Zeng, W., & Liu, W. (2019). Distinct H3K9me3 and DNA methylation modifications during mouse spermatogenesis. *J Biol Chem*, *294*(49), 18714-18725. https://doi.org/10.1074/jbc.RA119.010496

Long, J. E., Lee, M. S., & Blithe, D. L. (2021). Update on Novel Hormonal and Nonhormonal Male Contraceptive Development. *J Clin Endocrinol Metab*, *106*(6), e2381-e2392. https://doi.org/10.1210/clinem/dgab034

Loveland, K. L., Major, A. T., Butler, R., Young, J. C., Jans, D. A., & Miyamoto, Y. (2015). Putting things in place for fertilization: discovering roles for importin proteins in cell fate and spermatogenesis. *Asian J Androl*, *17*(4), 537-544. https://doi.org/10.4103/1008-682X.154310

Lüke, L., Campbell, P., Varea Sánchez, M., Nachman, M. W., & Roldan, E. R. (2014). Sexual selection on protamine and transition nuclear protein expression in mouse species. *Proc Biol Sci*, *281*(1783), 20133359. https://doi.org/10.1098/rspb.2013.3359

Lyon, K., Adams, A., Piva, M., Asghari, P., Moore, E. D., & Vogl, A. W. (2017). Ca2+ signaling machinery is present at intercellular junctions and structures associated with junction turnover in rat Sertoli cells. *Biol Reprod*, *96*(6), 1288-1302. https://doi.org/10.1093/biolre/iox042

Mata-Martinez, E., Sanchez-Cardenas, C., Chavez, J. C., Guerrero, A., Trevino, C. L., Corkidi, G., Montoya, F., Hernandez-Herrera, P., Buffone, M. G., Balestrini, P. A., & Darszon, A. (2021). Role of calcium oscillations in sperm physiology. *Biosystems*, *209*, 104524. https://doi.org/10.1016/j.biosystems.2021.104524

Matzuk, M. M., & Lamb, D. J. (2008). The biology of infertility: research advances and clinical challenges. *Nat Med*, *14*(11), 1197-1213. https://doi.org/10.1038/nm.f.1895

Meistrich, M. L., Mohapatra, B., Shirley, C. R., & Zhao, M. (2003). Roles of transition nuclear proteins in spermiogenesis. *Chromosoma*, *111*(8), 483-488. https://doi.org/10.1007/s00412-002-0227-z

Miyamoto, Y., Sasaki, M., Miyata, H., Monobe, Y., Nagai, M., Otani, M., Whiley, P. A. F., Morohoshi, A., Oki, S., Matsuda, J., Akagi, K. I., Adachi, J., Okabe, M., Ikawa, M., Yoneda, Y., Loveland, K. L., & Oka, M. (2020). Genetic loss of importin alpha4 causes abnormal sperm morphology and impacts on male fertility in mouse. *FASEB J*, *34*(12), 16224-16242. https://doi.org/10.1096/fj.202000768RR

Moritz, L., & Hammoud, S. S. (2022). The Art of Packaging the Sperm Genome: Molecular and Structural Basis of the Histone-To-Protamine Exchange. *Front Endocrinol (Lausanne)*, *13*, 895502. https://doi.org/10.3389/fendo.2022.895502

Nathaniel, B., Whiley, P. A. F., Miyamoto, Y., & Loveland, K. L. (2022). Importins: Diverse roles in male fertility. *Semin Cell Dev Biol*, *121*, 82-98. https://doi.org/10.1016/j.semcdb.2021.08.002

Nguyen, T. M., Duittoz, A., Praud, C., Combarnous, Y., & Blesbois, E. (2016). Calcium channels in chicken sperm regulate motility and the acrosome reaction. *FEBS J*, *283*(10), 1902-1920. https://doi.org/10.1111/febs.13710

O'Donnell, L. (2014). Mechanisms of spermiogenesis and spermiation and how they are disturbed. *Spermatogenesis*, *4*(2), e979623. https://doi.org/10.4161/21565562.2014.979623

O'Donnell, L., & O'Bryan, M. K. (2014). Microtubules and spermatogenesis. *Semin Cell Dev Biol*, *30*, 45-54. https://doi.org/10.1016/j.semcdb.2014.01.003

O'Donnell, L., Smith, L. B., & Rebourcet, D. (2022). Sertoli cells as key drivers of testis function. *Semin Cell Dev Biol*, *121*, 2-9. https://doi.org/10.1016/j.semcdb.2021.06.016

Ouyang, X., Lv, L., Zhao, Y., Zhang, F., Hu, Q., Li, Z., Zhu, D., & Li, L. (2021). ASF1B Serves as a Potential Therapeutic Target by Influencing Cell Cycle and Proliferation in Hepatocellular Carcinoma. *Front Oncol*, *11*, 801506. https://doi.org/10.3389/fonc.2021.801506

Pereira, R., Sa, R., Barros, A., & Sousa, M. (2017). Major regulatory mechanisms involved in sperm motility. *Asian J Androl*, *19*(1), 5-14. https://doi.org/10.4103/1008-682X.167716

Pontén, F., Gry, M., Fagerberg, L., Lundberg, E., Asplund, A., Berglund, L., Oksvold, P., Bjorling, E., Hober, S., Kampf, C., Navani, S., Nilsson, P., Ottosson, J., Persson, A., Wernerus, H., Wester, K., & Uhlen, M. (2009). A global view of protein expression in human cells, tissues, and organs. *Mol Syst Biol*, *5*, 337. https://doi.org/10.1038/msb.2009.93

Pradeepa, M. M., Manjunatha, S., Sathish, V., Agrawal, S., & Rao, M. R. (2008). Involvement of importin-4 in the transport of transition protein 2 into the spermatid nucleus. *Mol Cell Biol*, *28*(13), 4331-4341. https://doi.org/10.1128/MCB.00519-07

Qian, M. X., Pang, Y., Liu, C. H., Haratake, K., Du, B. Y., Ji, D. Y., Wang, G. F., Zhu, Q. Q., Song, W., Yu, Y., Zhang, X. X., Huang, H. T., Miao, S., Chen, L. B., Zhang, Z. H., Liang, Y. N., Liu, S., Cha, H., Yang, D., . . . Qiu, X. B. (2013). Acetylation-mediated proteasomal degradation of core histones during DNA repair and spermatogenesis. *Cell*, *153*(5), 1012-1024. https://doi.org/10.1016/j.cell.2013.04.032

Qu, W., Yuan, S., Quan, C., Huang, Q., Zhou, Q., Yap, Y., Shi, L., Zhang, D., Guest, T., Li, W., Yee, S. P., Zhang, L., Cazin, C., Hess, R. A., Ray, P. F., Kherraf, Z. E., & Zhang, Z. (2020). The essential role of intraflagellar transport protein IFT81 in male mice spermiogenesis and fertility. *Am J Physiol Cell Physiol*, *318*(6), C1092-C1106. https://doi.org/10.1152/ajpcell.00450.2019

Robertson, M. J., Kent, K., Tharp, N., Nozawa, K., Dean, L., Mathew, M., Grimm, S. L., Yu, Z., Legare, C., Fujihara, Y., Ikawa, M., Sullivan, R., Coarfa, C., Matzuk, M. M., & Garcia, T. X. (2020). Large-scale discovery of male reproductive tract-specific genes through analysis of RNA-seq datasets. *BMC Biol*, *18*(1), 103. https://doi.org/10.1186/s12915-020-00826-z

Roy, A., Lin, Y. N., Agno, J. E., DeMayo, F. J., & Matzuk, M. M. (2009). Tektin 3 is required for progressive sperm motility in mice. *Mol Reprod Dev*, *76*(5), 453-459. https://doi.org/10.1002/mrd.20957

Sahin, E., Petrunkina, A. M., Ekhlasi-Hundrieser, M., Hettel, C., Waberski, D., Harrison, R. A., & Töpfer-Petersen, E. (2009). Fibronectin type II-module proteins in the bovine genital tract and their putative role in cell volume control during sperm maturation. *Reprod Fertil Dev*, *21*(3), 479-488. https://doi.org/10.1071/rd08209

Schrimpf, R., Gottschalk, M., Metzger, J., Martinsson, G., Sieme, H., & Distl, O. (2016). Screening of whole genome sequences identified high-impact variants for stallion fertility. *BMC Genomics*, *17*, 288. https://doi.org/10.1186/s12864-016-2608-3

Shi, L., Zhou, T., Huang, Q., Zhang, S., Li, W., Zhang, L., Hess, R. A., Pazour, G. J., & Zhang, Z. (2019). Intraflagellar transport protein 74 is essential for spermatogenesis and male fertility in micedagger. *Biol Reprod*, *101*(1), 188-199. https://doi.org/10.1093/biolre/ioz071

Smith, L. B., O'Shaughnessy, P. J., & Rebourcet, D. (2015). Cell-specific ablation in the testis: what have we learned? *Andrology*, *3*(6), 1035-1049. https://doi.org/10.1111/andr.12107

Souquet, B., Tourpin, S., Messiaen, S., Moison, D., Habert, R., & Livera, G. (2012). Nodal signaling regulates the entry into meiosis in fetal germ cells. *Endocrinology*, *153*(5), 2466-2473. https://doi.org/10.1210/en.2011-2056

Stark, R., Grzelak, M., & Hadfield, J. (2019). RNA sequencing: the teenage years. *Nat Rev Genet*, *20*(11), 631-656. https://doi.org/10.1038/s41576-019-0150-2

Steger, K., Klonisch, T., Gavenis, K., Drabent, B., Doenecke, D., & Bergmann, M. (1998). Expression of mRNA and protein of nucleoproteins during human spermiogenesis. *Mol Hum Reprod*, *4*(10), 939-945. https://doi.org/10.1093/molehr/4.10.939

Steger, K., Pauls, K., Klonisch, T., Franke, F. E., & Bergmann, M. (2000). Expression of protamine-1 and -2 mRNA during human spermiogenesis. *Mol Hum Reprod*, *6*(3), 219-225. https://doi.org/10.1093/molehr/6.3.219

Suarez, S. S. (2008). Control of hyperactivation in sperm. *Hum Reprod Update*, *14*(6), 647-657. https://doi.org/10.1093/humupd/dmn029

Suarez, S. S., Varosi, S. M., & Dai, X. (1993). Intracellular calcium increases with hyperactivation in intact. *Proc. Natl. Acad. Sci. USA*, *90*, 4660-4664. https://doi.org/doi: 10.1073/pnas.90.10.4660

Tamburini, B. A., Carson, J. J., Adkins, M. W., & Tyler, J. K. (2005). Functional conservation and specialization among eukaryotic anti-silencing function 1 histone chaperones. *Eukaryot Cell*, *4*(9), 1583-1590. https://doi.org/10.1128/ec.4.9.1583-1590.2005

Tanaka, H., Iguchi, N., Toyama, Y., Kitamura, K., Takahashi, T., Kaseda, K., Maekawa, M., & Nishimune, Y. (2004). Mice deficient in the axonemal protein Tektin-t exhibit male infertility and immotile-cilium syndrome due to impaired inner arm dynein function. *Mol Cell Biol*, *24*(18), 7958-7964. https://doi.org/10.1128/mcb.24.18.7958-7964.2004

Timmons, P. M., Rigby, P. W. J., & Poirier, F. (2002). The murine seminiferous epithelial cycle is pre-figured in the Sertoli cells of embyonic testis. *Development*, *129*, 635-647. https://doi.org/doi: 10.1242/dev.129.3.635.

Tran, M. H., Aul, R. B., Xu, W., van der Hoorn, F. A., & Oko, R. (2012). Involvement of classical bipartite/karyopherin nuclear import pathway components in acrosomal trafficking and assembly during bovine and murid spermiogenesis. *Biol Reprod*, *86*(3), 84. https://doi.org/10.1095/biolreprod.111.096842

Tüttelmann, F., Ruckert, C., & Ropke, A. (2018). Disorders of spermatogenesis: Perspectives for novel genetic diagnostics after 20 years of unchanged routine. *Med Genet*, *30*(1), 12-20. https://doi.org/10.1007/s11825-018-0181-7

Uhlén, M., Fagerberg, L., Hallstrom, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, A., Kampf, C., Sjostedt, E., Asplund, A., Olsson, I., Edlund, K., Lundberg, E., Navani, S., Szigyarto, C. A., Odeberg, J., Djureinovic, D., Takanen, J. O., Hober, S., . . . Ponten, F. (2015). Proteomics. Tissue-based map of the human proteome. *Science*, *347*(6220), 1260419. https://doi.org/10.1126/science.1260419

Wang, W., Xiao, L., Pan, D., & Hu, L. (2022). ASF1B enhances migration and invasion of lung cancers cell via regulating the P53-mediated epithelial-mesenchymal transformation (EMT) signaling pathway. *Neoplasma*, *69*(2), 361-369. https://doi.org/10.4149/neo_2021_210818N1181

Washburn, R. L., & Dufour, J. M. (2023). Complementing Testicular Immune Regulation: The Relationship between Sertoli Cells, Complement, and the Immune Response. *Int J Mol Sci*, *24*(4). https://doi.org/10.3390/ijms24043371

Whiley, P. A., Miyamoto, Y., McLachlan, R. I., Jans, D. A., & Loveland, K. L. (2012). Changing subcellular localization of nuclear transport factors during human spermatogenesis. *Int J Androl*, *35*(2), 158-169. https://doi.org/10.1111/j.1365-2605.2011.01202.x

Wu, Y. Q., Rao, M., Hu, S. F., Ke, D. D., Zhu, C. H., & Xia, W. (2020). Effect of transient scrotal hyperthermia on human sperm: an iTRAQ-based proteomic analysis. *Reprod Biol Endocrinol*, *18*(1), 83. https://doi.org/10.1186/s12958-020-00640-w

Young, J. C., Ly-Huynh, J. D., Lescesen, H., Miyamoto, Y., Browne, C., Yoneda, Y., Koopman, P., Loveland, K. L., & Jans, D. A. (2013). The nuclear import factor importin alpha4 can protect against oxidative stress. *Biochim Biophys Acta*, *1833*(10), 2348-2356. https://doi.org/10.1016/j.bbamcr.2013.06.007

Zhang, S., Liu, Y., Huang, Q., Yuan, S., Liu, H., Shi, L., Yap, Y. T., Li, W., Zhen, J., Zhang, L., Hess, R. A., & Zhang, Z. (2020). Murine germ cell-specific disruption of Ift172 causes defects in spermiogenesis and male fertility. *Reproduction*, *159*(4), 409-421. https://doi.org/10.1530/REP-17-0789

Zhang, S. H., Zhang, J. H., Ding, X. P., Zhang, S., Chen, H. H., & Jing, Y. L. (2016). Association of polymorphisms in tektin-t gene with idiopathic asthenozoospermia in Sichuan, China. *J Assist Reprod Genet*, *33*(2), 181-187. https://doi.org/10.1007/s10815-015-0617-9

Zhang, W., Gao, Z., Guan, M., Liu, N., Meng, F., & Wang, G. (2021). ASF1B Promotes Oncogenesis in Lung Adenocarcinoma and Other Cancer Types. *Front Oncol*, *11*, 731547. https://doi.org/10.3389/fonc.2021.731547

Zhang, X. Z., Wei, L. L., Zhang, X. H., Jin, H. J., & Chen, S. R. (2022). Loss of perinuclear theca ACTRT1 causes acrosome detachment and severe male subfertility in mice. *Development*, *149*(12). https://doi.org/10.1242/dev.200489

Zhao, S., Zhu, W., Xue, S., & Han, D. (2014). Testicular defense systems: immune privilege and innate immunity. *Cell Mol Immunol*, *11*(5), 428-437. https://doi.org/10.1038/cmi.2014.38

Zhou, R., Wu, J., Liu, B., Jiang, Y., Chen, W., Li, J., He, Q., & He, Z. (2019). The roles and mechanisms of Leydig cells and myoid cells in regulating spermatogenesis. *Cell Mol Life Sci*, *76*(14), 2681-2695. https://doi.org/10.1007/s00018-019-03101-9

Zhou, W., Wang, G., Small, C. L., Liu, Z., Weng, C. C., Yang, L., Griswold, M. D., & Meistrich, M. L. (2011). Gene expression alterations by conditional knockout of androgen receptor in adult Sertoli cells of Utp14b jsd/jsd (jsd) mice. *Biol Reprod*, *84*(2), 400-408. https://doi.org/10.1095/biolreprod.110.090530

Zuo, X., Rong, B., Li, L., Lv, R., Lan, F., & Tong, M. H. (2018). The histone methyltransferase SETD2 is required for expression of acrosin-binding protein 1 and protamines and essential for spermiogenesis in mice. *J Biol Chem*, *293*(24), 9188-9197. https://doi.org/10.1074/jbc.RA118.002851

6. Appendix Document

Immunohistochemistry (IHC)

1. TRIS stock solution (10X)

60.5 g TRIS ($C_4H_{11}NO_3$, Roth) diluted in 900 mL double distilled water
Adjusted to pH 7.6 by adding hydrochloric acid (HCl, Roth).
Fill up to 1 L with distilled water., before adding 90 g sodium chloride (NaCl, Merck).
Control pH again.

2. TRIS washing buffer (TBST wash buffer)

100 mL of 10XTRIS stock solution
900 mL distilled water.
1000 µL Triton-X-100

3. Citrate buffer stock solutions

Stock solution A: 4.2 g Citric acid ($C_6H_8O_7$ x H_2O, Merck) diluted in 200ml of distilled water.
Stock solution B: 29.41 g Tri-Sodium citrate dihydrate ($C_6H_5O_7Na_3$ x 2 H_2O, Merck) diluted in 1 L of distilled water

4. Citrate buffer (working solution)

7 mL stock solution A
41 mL stock solution B
Fill up to 500 mL with distilled water.

5. Tris EDTA Buffer (pH 9)

25 mL of 20X TRIS stock solution
250 mL of Triton X
Make up to 500 mL with distilled water

6. 3% Hydrogen peroxide, H_2O_2

6 mL 30% H_2O_2, Roth
Fill up to 60 mL with 54 mL TBST wash buffer.

7. BSA buffer (1.5%)

1 g BSA (8076.2, Roth)

Fill up with 70 mL TBST buffer solution, dissolve with magnetic stirrer.

Immunofluorescence assay (IF)

1. Citrate Buffer (pH 6)

2.21 g Citric acid
1 L of MilliQ water
Adjust the pH to 6.0

2. PBST (PBS + Triton X)

1X PBS buffer + 0.1% Triton X (1000 mL 1X PBS + 1 mL Triton X)

3. 1% PB (1% BSA in 1X PBS)

0.1g of BSA in 10 mL of 1X PBS

Protein Extractions

1. 1X Phosphate Buffered Saline (PBS)

Diluted 10 g NaCl, 0.25 g KCl, 0.25 g KH_2PO_4, 1.75 g Na_2HPO_4 (all chemicals Merck) in 250 ml double distilled water.
Adjust to pH 7.3, then dilute 1:5.

2. RIPA Buffer

150 mM NaCl
1%NP-40 or Triton X-100
0.5% sodium deoxycholate
0.1% SDS
50 mM Tris, pH 8

3. Protein inhibitor cocktail

Halt's Inihibitor cocktail (x100) + EDTA (x100)

Western Blot (WB) - Fluorescence WB

1. 10X Tris-Glycine

30.3 g of Tris base
144.0 g of Glycine
1 L of MiliQ water

2. 1X Running Buffer (2L)

200 mL of Tris-Glycine
2 g of SDS
Up to 2 L of MilliQ water
Adjust pH to 8.3
Stored at 4 ºC

3. 1X Transfer Buffer (2L)

200 mL of Tris-Glycine
200 mL of Ethanol
1600 mL of MilliQ water
Store at 4 ºC

4. 10X TBS Buffer (2L)

121.0 g of Tris Base (0.5M)
175.0 g of Sodium Chloride (1.5M)
Adjust pH to 7.5.
Top up to 2 L with MilliQ water

5. 1X TBS Buffer (1 L)
   100 mL of 10X TBS + 900 mL of distilled water/ MilliQ water

6. TBST Buffer
   1000 mL of 1X TBS buffer + 1mL of Triton X

7. 5% Milk
   2.5 g of non-fat milk powder in 50 mL of TBST

Chemiluminescence WB

1. Sample preparation for the gel

	1x (1 Slot)	2x	3x	4x	5x
Probe (protein)	11,7 µL	23,4	35,1	46,8	58,5
Sample Buffer	4,5 µL	9	13,5	18	22,5
Reducing Agent	1,8 µL	3,6	5,4	7,2	9
Distilled water	-	-	-	-	-
Total	**18 µl**				

Load 15 µl of the sample per well

2. MOPS Buffer

950 mL of distilled water + 50 mL 20x MOPS

3. Transfer Buffer

For 1 gel	For 2 gels
50 mL NuPAGETransfer-Buffer 1 mL NuPAGE Antioxidant 100 mL Methanol 849 mL distilled water	50 mL NuPAGETransfer-Buffer 1 mL NuPAGE Antioxidant 200 mL Methanol 749 mL distilled water