Preclinical Evaluation of TEX101 Protein as a Male Infertility Biomarker and Identification of its Functional Interactome

by

Christina Schiza

A thesis submitted in conformity with the requirements for the degree of Doctorate of Philosophy
Laboratory Medicine and Pathobiology
University of Toronto

© Copyright by Christina Schiza 2017
Preclinical Evaluation of TEX101 Protein as a Male Infertility Biomarker and Identification of its Functional Interactome

Christina Schiza
Doctor of Philosophy
Laboratory Medicine and Pathobiology
University of Toronto
2017

Abstract

TEX101 is a testis-specific cell surface protein expressed exclusively in the male germ cells, and it was previously suggested as a biomarker for the differential diagnosis of male infertility. Molecular function of TEX101 is not known, but previous studies demonstrated that murine TEX101 was a cell membrane chaperone, essential for fertilization. In this work, we employed quantitative proteomic strategies to investigate the clinical and the functional aspects of human TEX101 protein.

To facilitate translation of TEX101 into clinic, we first developed and optimized an immunoassay for TEX101 measurement in biological fluids. We then evaluated the performance of TEX101 ELISA in 805 seminal plasma samples of fertile, sub-fertile, and infertile men. We demonstrated the clinical utility of TEX101 to evaluate vasectomy success, to stratify azoospermia forms and subtypes, and to predict the success of sperm retrieval in patients diagnosed with non-obstructive azoospermia.

To gain insights into human TEX101 function, we employed mass spectrometry approaches to investigate the physical and the transient TEX101 interactome. Co-immunoprecipitation-mass spectrometry revealed a testis-specific complex TEX101-DPEP3 which was validated by hybrid immunoassay in testicular tissues and spermatozoa. In addition, we discovered a TEX101 rs35033974 homozygous knockdown model, which carried a missense variation leading to near-
complete protein degradation. We then performed global proteomic analysis of *TEX101* knockdown spermatozoa, and we identified the transient interactome of TEX101.

New knowledge on TEX101 will provide insights into reproductive biology, facilitate better diagnostics of male infertility, contribute to rational selection of assisted reproduction treatments, and offer new protein targets to develop non-hormonal male contraceptives.
Acknowledgments

Every ending is a new beginning! The past five years have been a new experience for me. It was a challenge and a bet with myself. This thesis is my reward. There have been many people to guide me and to share my excitement and my disappointment throughout these years.

First, I would like to thank my supervisor, Dr. Eleftherios P. Diamandis, for giving me the opportunity to join ACDC Lab, and to pursue a PhD in one of the top universities worldwide. Dr. Diamandis’ challenging approach made me realize my inner strength, and gave me the power to fight and not give up.

I would also like to acknowledge the members of my advisory committee, Dr. Keith Jarvi and Dr. Alexander Romaschin, for their valuable advice throughout my PhD. A very special thanks to Dr. Jarvi for his positive energy, and for providing all the biological specimens for my studies. In addition, I would like to acknowledge the members of my final oral examination committee, Dr. G. Schmitt-Ulms, Dr. V. Kulasingam, and Dr. R. Oko for reviewing my thesis and for providing comments and feedback. I would also want to thank Dr. Harry Elsholtz for the support and aid for all academic needs.

I wish to express my appreciation to two great scientists and labmates, Dr. Andrei P. Drabovich and Dr. Dimitrios Kormpakis. I was very lucky to work closely with both of them throughout my PhD. They were always willing to share their knowledge and experience, and they taught me how to focus on my goals, and to be patient. They encouraged me to keep trying improving my skills, and they were a great example that hard work is always rewarded. I would also like to thank Andrei for reviewing my thesis and my manuscripts.

This journey would have been lonely if I did not have my labmates and friends next to me. Lampros, Ilijana, and Dimitris are the people that I shared moments of excitement, disappointment, stress, endless complaining about PhD life, but also fun by exploring new things in Toronto. A special thanks to Ihor Batruch for trying his best to keep our instruments alive and in good shape, but most of all for being Ihor!

Finally, a special thanks to our lab manager Antoninus for keeping the lab in balance for so many years, and to Susan Lau for recruiting all the patients, and for the administrative support.
Some 8,000 km away, but always next to me, my family and my friends. I would like to thank Christina, Myrto, Konstantina, Georgia and Cleo for being my anchors back home, and for believing in me. Special thanks to Cleo for all the technical support, and for staying up late to keep me company while I was writing my thesis.

Last but not least, I want to thank my parents, Giorgos and Anastasia, for their unconditional support and all the sacrifices they made, to give me and my sister a better future. I have to acknowledge that although it was really hard for them to have both their children away, they never doubted about our choices. Finally, I would like to thank my sister. Although she is younger than me, she has always been an example of determination, strength, and courage.
Table of Contents

Abstract.................................................................................................................................................. ii
Acknowledgments............................................................................................................................... iv
Table of Contents .................................................................................................................................. vi
List of Tables ......................................................................................................................................... xii
List of Figures ......................................................................................................................................... xiii
List of Appendices ............................................................................................................................. xv
List of Abbreviations ........................................................................................................................ xvi
1 Introduction .......................................................................................................................................... 2
  1.1 Infertility ......................................................................................................................................... 2
  1.2 The Male Reproductive System .................................................................................................. 3
    1.2.1 Anatomy and Physiology of the Male Reproductive System ........................................... 4
      1.2.1.1 Testes and spermatogenesis ......................................................................................... 4
      1.2.1.2 Genital ducts and sperm maturation ........................................................................... 6
      1.2.1.3 Accessory glands and production of seminal fluid ..................................................... 7
    1.2.2 Sperm transport and maturation ......................................................................................... 7
  1.3 Male Infertility ............................................................................................................................ 9
    1.3.1 Subtypes of male infertility ................................................................................................... 9
    1.3.2 Diagnosis of male infertility ............................................................................................... 10
    1.3.3 Current treatments ............................................................................................................. 12
    1.3.4 Emerging male infertility biomarkers ............................................................................... 13
  1.4 Innovative proteomics strategies to study clinical and functional aspects of male reproduction and infertility .................................................................................................................. 15
    1.4.1 Reagents and tools to study testis-specific proteins ......................................................... 15
    1.4.2 Protein-protein interactions ............................................................................................... 16
    1.4.3 Global and targeted mass spectrometry approaches ...................................................... 17
  1.5 TEX101 protein: a biomarker of spermatogenesis and male infertility .................................. 18
    1.5.1 TEX101 expression pattern at transcript and protein level ............................................ 18
    1.5.2 TEX101 localization during spermatogenesis and sperm maturation ......................... 20
    1.5.3 Suggested functional roles of TEX101 protein ............................................................... 21
1.5.4 TEX101 as a biomarker for the differential diagnosis of azoospermia .................................. 23

1.6 Rationale and Objectives of the present study ........................................................................... 27

1.6.1 Rationale ................................................................................................................................. 27
1.6.2 Hypothesis ............................................................................................................................. 28
1.6.3 Objectives ............................................................................................................................. 29

2 Development of ELISA for the measurement of native TEX101 in biological fluids .......... 31

2.1 Introduction ................................................................................................................................ 31

2.2 Materials and Methods .............................................................................................................. 33

2.2.1 Cloning of TEX101 cDNA ..................................................................................................... 33
2.2.2 Production of human recombinant TEX101 .......................................................................... 33
2.2.3 Purification of human TEX101 with immobilized metal ion affinity chromatography .......... 34
2.2.4 Analysis of human recombinant TEX101 by mass spectrometry ......................................... 35
2.2.5 Assessment of TEX101 glycosylation .................................................................................. 36
2.2.6 Animal handling and somatic cell fusion for monoclonal antibody production .................. 36
2.2.7 Screening for IgG secreting clones ....................................................................................... 36
2.2.8 Screening for immunogen reacting clones .......................................................................... 37
2.2.9 Immunocapture-SRM screening for clones producing antibodies against native TEX101 ................................................................. 37
2.2.10 Clone screening by Western blot ....................................................................................... 39
2.2.11 Expansion of clones and purification of anti-human TEX101 monoclonal antibodies .................................................................................................................. 39
2.2.12 Development of a sandwich immunoassay ....................................................................... 40
2.2.13 Measurement of TEX101 in seminal plasma and blood serum ......................................... 40
2.2.14 Statistical analysis .............................................................................................................. 41

2.3 Results ....................................................................................................................................... 43

2.3.1 Production, purification and analysis of recombinant human TEX101 protein .................... 43
2.3.2 Production and screening of monoclonal antibodies ............................................................ 43
2.3.3 Immunoassay development .................................................................................................. 45
2.3.4 ELISA performance and TEX101 measurement in seminal plasma ................................. 45
2.3.5 ELISA performance and TEX101 measurement in blood serum ........................................ 47

2.4 Discussion .................................................................................................................................. 48
3 Preclinical evaluation of TEX101 protein ELISA for the differential diagnosis of male infertility ........................................................................................................................................................................52

3.1 Introduction..................................................................................................................................................................................................................52

3.2 Materials and Methods......................................................................................................................................................................................................54

3.2.1 Experimental design and statistical rationale..........................................................................................................................................................54

3.2.2 Patients with inclusion and exclusion criteria......................................................................................................................................................54

3.2.3 Seminal plasma samples ...........................................................................................................................................................................................55

3.2.4 TEX101 ELISA measurements ...............................................................................................................................................................................56

3.2.5 Measurement of TEX101 isoforms in spermatozoa and seminal plasma by SRM ........................................................................................................56

3.2.6 Immunohistochemistry of TEX101 in testicular tissues ..........................................................................................................................................58

3.2.7 Size-exclusion chromatography .........................................................................................................................................................................58

3.2.8 Isolation of SMVs from seminal plasma samples ............................................................................................................................................58

3.2.9 TEX101 measurement in seminal plasma samples, SMVs and vesicle-free seminal plasma by ELISA .........................................................................................59

3.2.10 TEX101 measurement in seminal plasma samples, SMVs and vesicle-free seminal plasma by SRM ........................................................................................................................................59

3.2.11 TEX101 measurement in seminal plasma by immunocapture-SRM ........................................................................................................................................60

3.2.12 Statistics ..............................................................................................................................................................................................................60

3.3 Results .............................................................................................................................................................................................................62

3.3.1 Measurement of TEX101 in 821 seminal plasma samples using guanidine-based treatment protocol ........................................................................................................................................................................62

3.3.2 Isoform identity of TEX101 in spermatozoa and seminal plasma ..............................................................................................................................................63

3.3.3 Confirmation of TEX101 germ cell-specificity and extracellular membrane localization by immunohistochemistry .................................................................................................................................65

3.3.4 Search for soluble TEX101 complexes in seminal plasma by size-exclusion chromatography ......................................................................................................................................................................66

3.3.5 TEX101 association with SMVs in seminal plasma ......................................................................................................................................................68

3.3.6 Effect of seminal plasma pretreatment on TEX101 measurements by ELISA ..............................................................................................................................................69

3.3.7 Stability of TEX101 in semen .................................................................................................................................................................................70

3.3.8 Evaluation of TEX101 as a male infertility biomarker in a population of 805 men ..............................................................................................................................................70

3.3.9 Evaluation of TEX101 as a biomarker to differentiate between NOA and OA ............................................................................................................................................75

3.3.10 Evaluation of TEX101 as a biomarker to predict sperm retrieval in NOA patients ........................................................................................................................................................................76
6.3 Future Directions .................................................................142
References ......................................................................................145
Appendices .....................................................................................174
List of Publications ........................................................................Error! Bookmark not defined.
List of Tables

Table 3.1 Clinicopathological variables of the patient cohort .................................................. 55
Table 4.1 List of candidate interacting partners of TEX101 in human testicular tissue, and validation of candidates by co-IP SRM. ........................................................................................................ 98
Table 4.2 List of candidate interacting partners of TEX101 in human spermatozoa, and validation of candidates by co-IP SRM. ........................................................................................................ 99
Table 5.1 Clinicopathological variables of the patient cohort .......................................................... 111
Table 5.2 Clinical parameters of wild type TEX101 and homozygous rs35033974 TEX101 men (G99V^h^h^), which were subjected to differential proteomic analysis .................................................. 116
Table 5.3 List of downregulated testis-specific proteins in spermatozoa of rs35033974 homozygous men, as discovered by shotgun MS and verified by SRM. .................................................. 131
List of Figures

Figure 1.1 General algorithm for the evaluation of male infertility. ........................................... 11
Figure 1.2 Genomic, transcriptomic and proteomic organization of TEX101. ......................... 19
Figure 1.3 Two-marker algorithm for the differential diagnosis of azoospermia (OA versus NOA) with seminal plasma proteins ECM1 and TEX101................................................... 24
Figure 1.4 Model of TEX101 expression and secretion into SP in men with normal spermatogenesis (NS) and azoospermia. .................................................................................. 26
Figure 2.1 Pipeline for the production of mouse monoclonal anti-TEX101 antibodies and screening of colonies using two-step immunocapture-SRM assay........................................... 32
Figure 2.2 Screening of hybridoma colonies by immunocapture-SRM assay and Western Blot. 44
Figure 3.1 Investigation of TEX101 isoforms in spermatozoa and SP using SRM. .................... 64
Figure 3.2 Immunohistochemical staining of TEX101 protein in testicular tissue with active spermatogenesis.................................................................................................................... 65
Figure 3.3 Investigation of TEX101 analyte identity in SP and the impact of SP pretreatment. . 67
Figure 3.4 Evaluation of TEX101 as a biomarker. ........................................................................ 73
Figure 3.5 Clinical utility of TEX101 ELISA as a test to evaluate vasectomy success, differentiate between non-obstructive and obstructive azoospermia and predict the success of sperm retrieval in patients with non-obstructive azoospermia. ............................................ 74
Figure 4.1 Production of mouse monoclonal antibodies against different epitopes of native TEX101 protein. ............................................................................................................... 95
Figure 4.2 Identification of TEX101 protein interactome by co-immunoprecipitation coupled with mass spectrometry........................................................................................................... 97
Figure 4.3 Verification of human TEX101 interactome by co-IP-SRM. .......................................... 100
Figure 4.4 Detection and validation of TEX101-DPEP3 complex by hybrid immunoassay. .... 102
Figure 5.1 Identification of rs35033974 variant in TEX101 gene. .............................................. 122
Figure 5.2 Conservation of TEX101 residues............................................................................... 123
Figure 5.3 Monitoring wild type and variant G99V form of TEX101 protein in spermatozoa lysate by targeted mass-spectrometry................................................................. 127
Figure 5.4 Identification of candidate proteins downregulated in TEX101 rs35033974 homozygous men. ....................................................................................................................... 128
Figure 5.5 Verification of testis-specific membrane-bound and secreted downregulated proteins in three TEX101 rs35033974 homozygous and four TEX101 wild type men by SRM and western blotting.
List of Appendices

Appendix 2.1 Analysis of recombinant human TEX101 protein. ................................. 174
Appendix 2.2 Calibration curves and linear correlation between endogenous/spike-in TEX101 concentration and dilution in seminal plasma and female serum. ........................................ 175
Appendix 2.3 Calibration curves for the endogenous TEX101 in seminal plasma (▲) and the deglycosylated recombinant TEX101 diluted in PBS buffer (●). ................................. 176
Appendix 3.1 TEX101 levels measured in SP samples (N=821) by ELISA using GndCl-based protocol (3 M guanidine hydrochloride for 1 hour at RT). ........................................ 177
Appendix 3.2 Diagnosis, sperm count and TEX101 levels in SP of 17 infertile men with high sperm count (≥7 mln/mL). ........................................................................ 178
Appendix 3.3 Relative amounts of TEX101 captured from SP pools by commercial (mPoly) or in-house generated (23ED616.8) antibodies, as measured by SRM. ........................................ 179
Appendix 3.4 Correlation between sperm count and TEX101 levels in SP. ...................... 180
Appendix 4.1 Parameters of a multiplex scheduled SRM assay for the validation of TEX101 candidate interacting proteins in human testicular tissue. ........................................ 181
Appendix 4.2 Parameters of a multiplex scheduled SRM assay for the validation of TEX101 candidate interacting proteins in human sperm cells from normal fertile individuals. ....... 183
Appendix 4.3 Optimization of testicular tissue lysis. ......................................................... 185
Appendix 4.4 Identification of TEX101 interactome by 34ED229 mAb co-immunoprecipitation coupled with mass spectrometry. ................................................................. 186
Appendix 4.5 SDS-PAGE and Western blotting analysis of purified recombinant DPEP3 protein. ............................................................................................. 187
Appendix 4.6 Production of mouse monoclonal antibodies against different epitopes of native DPEP3 protein. ................................................................. 188
Appendix 4.7 Investigation of TEX101 sulfation in testicular tissue, seminal plasma and spermatozoa. .................................................................................. 189
Appendix 4.8 Western blotting analysis of endogenous DPEP3 protein........................... 190
Appendix 5.1 Parameters of the unscheduled PRM assay for monitoring TEX101 protein in Glu-C digested sperm cell lysate and seminal plasma. ............................................... 191
Appendix 5.2 Parameters of a multiplex scheduled SRM assay for the quantification of TEX101 and other downregulated proteins in sperm cell samples from TEX101 G99V homozygotes and control fertile individuals. ......................................................... 192
Appendix 5.3 List of proteins significantly downregulated in sperm cell samples obtained from TEX101 rs35033974 homozygous men when compared with TEX101 wild-type fertile control men. ................................................................. 194

Appendix 5.4 ADAM proteins identified. ................................................................................................................................. 196
List of Abbreviations

ABC, Ammonium bicarbonate  
ACE, Angiotensin-converting enzyme  
ACN, Acetonitrile  
ADAM, A disintegrin and metalloprotease  
AR, Acrosome reaction  
ART, Assisted Reproduction Technologies  
AUC, Area under the curve  
BCA, Bicinchoninic acid  
BMGY, Buffered complex glycerol medium  
BMMY, Buffered complex methanol  
BTB, Blood-testis barrier  
CFTR, Cystic fibrosis transmembrane conductance receptor  
CI, Confidence intervals  
COC, Cumulus-oocyte complex  
DPEP3, Dipeptidase 3  
DTT, Dithiothreitol  
ECM1, Extracellular matrix protein 1  
ELISA, Enzyme-linked immunosorbent assay  
ER, Endoplasmatic reticulum  
Fc, constant region of immunoglobulin  
FDR, False discovery rate  
FRT, Female reproductive tract  
FSH, Follicle stimulating hormone  
FWHM, Full width at half maximum  
GndCl, Guanidine hydrochloride  
GnRH, Gonadotropin-releasing hormone  
GPI, Glycophosphatidylinositol  
HAT, Hypoxanthine-aminopterin-thymidine medium  
hCG, human chorionic gonadotropin  
HPA, Human Protein Atlas  
HPLC, High Pressure liquid chromatography
HRP, Horseradish peroxidase
HS, Hyperspermatogenesis
HZA, Hemizona assay
ICl, Intracervical insemination
ICSI, Intracytoplasmic sperm injection
IUI, Intrauterine insemination
IVF, in vitro fertilization
LC-MS/MS, Liquid chromatography-tandem mass spectrometry
LFQ, Label-free quantification
LH, Luteinizing hormone
LOD, Limit of detection
LY6K, Lymphocyte antigen 6K
MA, Maturation arrest
mAb, Monoclonal antibody
MS, Mass spectrometry
NHS, N-hydroxysuccinimide
NOA, Non-obstructive azoospermia
OA, Obstructive azoospermia
PPI, Protein-protein interaction
PRM, Parallel reaction monitoring
PSA, Prostate-specific antigen
PTM, Post-translational modification
PV, Post vasectomy
ROC, Receiver operating characteristic curve
RT, Room temperature
SCO, Sertoli cell-only
SCX, Strong cation exchange chromatography
SMVs, Seminal microvesicles
SNP, Single nucleotide polymorphism
SP, Seminal plasma
SPA, Sperm penetration assay
SRM, Selected reaction monitoring
TBS, Tris-buffered saline
TESE, Testicular sperm extraction
TEX101, Testis-expressed sequence 101
UTJ, Uterotubular junction
ZP, Zona pellucida
Chapter 1
Introduction

Chapter 1

1 Introduction

1.1 Infertility

Infertility is estimated to affect 10-15% of couples in developed countries (1). According to the World Health Organization (WHO) guidelines, infertility is defined as the lack of conception, leading to pregnancy and a live birth, after one year of unprotected intercourse (2). Infertility is a medical problem with implications for male and female partner’s health, as well as for public health. The emotional issues experienced by the infertile couples, as well as financial costs of assisted reproduction, are substantial (3). Costs of a standard *in vitro* fertilization (IVF) cycle range from $2,000 to $13,000 across 32 middle and high income countries worldwide (3, 4). The direct and indirect costs of assisted reproduction technologies (ART) may not be fully covered by health care system in some countries, and thus, set barriers to ART access for several patients (4). In addition, social and cultural determinants render ART not accessible to many couples (5). It is estimated that 1 in 6 couples experience infertility in Canada, and this number has doubled since the 1980s (6). Over the past decade, there has been a significant increase in the application of ART in order to treat infertility (7-10).

Currently, a wide range of established ART procedures ensure that most of the infertile couples will be able to have a child. On the other hand, there are possible risks associated with infertility and ART: multiple pregnancies, and effects on the offspring (e.g. genetic abnormalities passed from parents) (11). One could argue that advances in ART can provide solution to most infertility cases. Therefore, investigation of the causes of male or female infertility and reproductive biology is not very important. Nevertheless, identifying the cause of infertility can be beneficial not only for the couple’s health, but also for the public health, and the health care system. Confirmation of infertility causes and associated factors can assist the clinician to provide correct advice to couples seeking fertility treatment, and thus eliminate the cost of unnecessary treatments. Furthermore, ART risks can be controlled and prevented when both partners are thoroughly examined. On the other hand, studying reproductive biology and identifying key molecular factors and mechanisms that regulate the production of male and female gametes can facilitate the development of novel non-hormonal male and female contraceptives. Traditionally, hormonal contraceptive approaches have focused on women, while hormonal male contraception using testosterone administration has been shown to
suppress spermatogenesis (12). Development of non-hormonal methods of male contraception is an emerging field in the area of male reproduction (13).

In the past, traditional societies assumed that female factor contributed mostly to a couple’s infertility. As a result, the female reproductive system and the associated disorders were extensively studied. However, further research demonstrated that the causes of infertility in couples seeking treatments can be identified in both partners. It is estimated that the male disorders, alone or combined with the female factor, contribute to approximately 50% of all infertility cases (2, 14). For the high proportion of sub-fertile or infertile men, however, the cause of infertility remains unknown (15). The complex nature of sperm production and maturation underscores the challenge to diagnose and treat male infertility (16, 17). Despite the advances in the fields of genomics and proteomics, the molecular mechanisms leading to fertility defects remain unknown.

The understanding of the molecular and genetic mechanisms that control spermatogenesis and sperm maturation will not only facilitate accurate diagnosis and treatment, but also introduce novel techniques for the identification of sperm functional defects, and furthermore tailor the application of the available treatment options. Alternatively, identifying key molecules and mechanisms of infertility can provide new tools and knowledge on how to control fertility.

Family planning has not only a huge impact on public health, but also social and economic implications. In one of the 17 sustainable development goals set by the United Nations in September 2015, reproductive health and family planning are the core targets (18). Since contraceptive methods and infertility treatment are key elements of family planning, advances in human reproduction can assist the fulfillment of this goal and promote well-being.

### 1.2 The Male Reproductive System

The human male reproductive system consists of a number of organs located outside of the body (external genital organs) and within the pelvis (internal genital organs). The male reproductive system has four primary functions: i) the production and secretion of male sex hormones (androgens); ii) the production, maintenance and transport of spermatozoa (male gametes); iii) the production of seminal fluid for the protection and nourishing of sperm; and iv) the deposition of sperm into the female reproductive tract (FRT). In the following section, we will
summarize the anatomy and the physiology of the different parts that constitute the male reproductive system.

1.2.1 Anatomy and Physiology of the Male Reproductive System

1.2.1.1 Testes and spermatogenesis

The testes are paired organs, responsible for the production of spermatozoa and male hormones. The seminiferous tubules are the functional unit of the testes where spermatogenesis takes place. They are highly convoluted tubules, lined by a stratified epithelium of two distinct cell populations: the spermatogenic and the supporting cells.

The spermatogenic cells contain spermatogonia, primary and secondary spermatocytes, spermatids and spermatozoa which are arranged in layers from the basal layer of the epithelium to the lumen of the seminiferous tubule. Spermatogonia are stem cells, located in the basal layer (19). There are two subtypes of spermatogonia: type A cells that undergo mitotic divisions to maintain the spermatogonia population, and type B cells that undergo mitotic divisions and differentiate into primary spermatocytes (20). Primary spermatocytes are the largest germ cells that undergo the first meiotic division to give rise to haploid secondary spermatocytes. Secondary spermatocytes rapidly undergo the second meiotic division to produce haploid spermatids (19). Subsequently, spermatids undergo spermiogenesis, a transformation process that results in the production of morphologically complete spermatozoa. Spermiogenesis includes condensation of the nucleus and formation of sperm head, formation of the sperm tail and the acrosome cap, and shedding of excessive spermatid cytoplasm (21).

Sertoli cells are also found in the seminiferous tubules, located on the basement membrane. Sertoli cells provide structural and metabolic support to the differentiating germ cells. It is estimated that each Sertoli cell supports ~30-50 germ cells at different stages, and that Sertoli-germ cell cross-talk coordinates different events in spermatogenesis such as mitosis, meiosis, spermiogenesis, or apoptosis (22). Furthermore, tight junctions between adjacent Sertoli cells lead to the formation of the blood-testis barrier (BTB). The BTB has multiple functions: (i) it controls the flow of nutrients, hormones, electrolytes and exogenous chemicals across the Sertoli cell epithelium into the apical compartment, creating a unique microenvironment for germ cells; (ii) it creates an immunological barrier to antigens expressed by post-meiotic germ
cells, thus eliminating the production of anti-sperm antibodies which could lead to immune response and male infertility; and (iii) it confers cell polarity in the seminiferous epithelium, assisting the proper orientation of spermatids (23).

Leydig cells reside inside the interstitial connective tissue between the seminiferous tubules, usually in small groups adjacent to the blood vessels. Leydig cells are responsible for the production of testosterone (19).

Spermatogenesis is a complex, multi-step process taking place in the seminiferous tubules. In humans, spermatogenesis is initiated at puberty under the influence of gonadotropins, follicle stimulating hormone (FSH) and luteinizing hormone (LH). LH stimulates the production of testosterone by the Leydig cells, which along with FSH target the Sertoli cells to trigger spermatogenesis (22). In humans, spermatogenesis takes approximately 64-68 days or ~4.2 seminiferous epithelial cycles to proceed from type A spermatogonium to mature sperm cells (24).

In each stage of spermatogenesis, germ cells present a distinct gene expression pattern. There are more than 1,000 genes enriched in the testis, possibly due to the fact that germs cells are the only cell type that undergoes meiosis. Analysis of tissue-specific proteins showed that more than 70% testis-specific proteins are intracellular, and they are mostly involved in intracellular events such as meiosis, histone replacement, and sperm differentiation. Interestingly, half of the testis-specific genes are expressed during late stages of spermatogenesis, while less than 10% are expressed by spermatogonia and spermatocytes (25). As it was mentioned above, the presence of the BTB does not allow the release of testis-specific proteins in circulation. Thus, these proteins are normally not detected in blood, unless a pathological tissue destruction occurs due to testicular cancer, trauma and torsion.

After the completion of spermatogenesis, fully developed spermatozoa detach from the seminiferous epithelium, and they are released in the lumen. Spermatozoa pass through the rete testis, and enter the epididymis where they are subjected to further maturation (19).
1.2.1.2 Genital ducts and sperm maturation

The genital ducts consist of the epididymis, the vas deferens, and the ejaculatory duct. After their production, spermatozoa are stored in the epididymis, and prior ejaculation they pass through the vas deferens to the ejaculatory duct, which expels them into the female genital tract by contraction (19).

Based on functional and morphological differences, and gene-specific expression patterns, the epididymis is divided into three segments: the caput (head), the corpus (body), and the cauda (tail) (26). The epididymis is important for sperm fertility, and it is involved in different processes such as sperm concentration and transport, maturation, immunoprotection, and storage (27). The sperm-concentrating ability of the epididymis is achieved through water and protein reabsorption by the epithelial cells that line the duct (26). Human sperm remains in the epididymis approximately 2 to 6 days, much shorter than other species (28). At this stage, transcription and translation are silenced. One portion of all essential proteins are acquired from sperm during spermatogenesis, and they are further processed during sperm maturation. The role of the epididymis is to facilitate the maturation of the existing proteins, and to provide the haploid sperm with additional proteins that are crucial for sperm transit and fertilization. The cross-talk between spermatozoa and epididymis, and the protein-protein interactions are of utmost importance for the production of functional sperm. The modifications that take place in the epididymis include: (i) sperm surface protein removal by proteolysis; (ii) sperm surface protein processing and maturation; (iii) redistribution of proteins on the sperm membrane; and (iv) transient or permanent integration of epididymal proteins into the sperm membrane (29). All these stages are essential for motility and fertility acquisition by the mature spermatozoa. The key element in the epididymis function is the highly specialized luminal microenvironment, which is created and maintained by the blood-epididymis barrier. The blood-epididymis barrier is also responsible for the immunoprotection of sperm by controlling the secretion of proteins and the exchange of molecules between the lumen and the circulation (30). There is also evidence that epididymal cells secrete microvesicles, named epididymosomes, which serve as transporters of highly hydrophobic and less soluble proteins (e.g. GPI-anchored proteins) from the epididymis to spermatozoa (31, 32). Finally, the epididymis is the reservoir of mature sperm. Sperm can be stored in the cauda for several weeks, protected from the immune system (33).
1.2.1.3 Accessory glands and production of seminal fluid

The accessory glands include the paired seminal vesicles, the prostate gland, and the paired bulbourethral glands (19). Each of these glands contribute with their secretions to the formation of seminal fluid, which in addition to spermatozoa constitute semen.

The seminal vesicles are paired glands, situated lateral to each vas deferens, and their activity is regulated by testosterone. The seminal vesicles secrete a viscous alkaline fluid, which enables spermatozoa to move. Secretions from seminal vesicles and the other accessory glands provide the sperm with a protective glycoprotein-coat against acidic secretions and immunity in the female reproductive tract, and infection-causing microbes (34-36). Other products of the seminal vesicles are prostaglandin, and fructose for nourishing spermatozoa (35). The prostate is the largest accessory gland in the male reproductive tract. The prostate secretes a slightly alkaline, milky fluid, which is rich in lipids, citrate and proteolytic enzymes that catalyze the liquefaction of semen after ejaculation (34). In addition, prostatic secretions contain high amount of polyamines (34), spermidine and spermine, which appear to be essential to sperm function, enhancing sperm count and motility, and they are implicated in sperm capacitation and acrosome reaction (37). The bulbourethral (Cowper’s) glands are small paired glands that provide a clear, viscous, alkaline mucous fluid which contains sialic acid and galactose. This fluid is discharged during intercourse to lubricate and neutralize the acidic lumen of the urethra, in order for the semen to pass through during ejaculation (34).

To summarize, apart from spermatozoa, semen is composed of secretions from various glands in the male reproductive tract. These gland-specific secretions are combined during sperm transit through the male reproductive tract, to assist maturation, to provide nutrition, and to protect sperm during its interaction with the female reproductive tract. Therefore, changes in the composition of seminal fluid may reflect disorders in the male reproductive system (38).

1.2.2 Sperm transport and maturation

Following the first stage of maturation in the epididymis and the other accessory glands, sperm enters the female reproductive tract, and after passing the cervix and the uterus, it eventually reaches the oviduct. During this transit, sperm interacts with a number of components of the uterine and oviduct secretions, and in this way complete sperm maturation is achieved (39).
Three major regions comprise the oviduct: the uterotubular junction (UTJ), the isthmus, and the ampulla. The UTJ connects the uterus with the oviduct and forms a natural barrier through which only spermatozoa can pass (40). Although numerous spermatozoa are initially released in the female tract, only a very small fraction will manage to migrate up to the UTJ, and enter the oviduct (41). After passing through the UTJ, sperm bind to the epithelial cells of the isthmus, through protein interactions between lectin-like sperm proteins and carbohydrate residues present on the oviductal epithelial cells (42), thus forming a sperm reservoir. Release of sperm occurs gradually, while sperm undergo capacitation, a process that includes multiple physiological and biochemical modifications (41). Sperm capacitation is associated with an efflux of cholesterol from the plasma membrane, leading to membrane fluidity, protein redistribution on the membrane, and sperm hyperactivation through increase of intracellular Ca\(^{+2}\) (43).

The sperm release from the isthmus is modulated by interaction between the cumulus-oocyte complex (COC) in the ampulla, and the epithelial cells in the oviduct (44). Gradient of temperature and chemoattractants, such as progesterone, guide the sperm towards the oocyte. Once the spermatozoa approach the COC, the sperm acrosome reaction (AR) takes place, to achieve complete maturation and acquisition of fertilizing ability. Since AR is an irreversible process, its initiation is strictly regulated spatially and temporally by the COC. Premature AR leads to sperm incompetence to cross the cumulus layer and encounter with the oocyte (45). Upon acrosome reaction, hydrolytic enzymes are released from the sperm to facilitate the dispersion of the cumulus mass, and the penetration of zona pellucida (ZP), a glycoprotein-rich membrane surrounding the plasma membrane of the oocyte (46, 47). Progesterone released from the cumulus cells is one of the factors that induce AR (48). Sperm interaction with ZP leads to ZP3 protein activation, and Ca\(^{+2}\) influx in spermatozoa, inducing AR (45). Combination of two alternative AR models suggests that both cumulus cells and zona pellucida may induce acrosome reaction. It is possible that the acrosome-reacted sperm may disperse the cumulus layer, allowing the intact sperm to pass through and bind to ZP. This means that a large number of spermatozoa have to be sacrificed to facilitate the passage of a small population, which will participate in egg fertilization (46).
1.3 Male Infertility

Recent studies report a decline in male fertility over the years, and an increase of the contribution of the male factor to a couple’s infertility (49, 50). Although male infertility is now receiving more attention, and more research has been conducted, it is still difficult to define the causes. Therefore, male infertility is still classified based on semen quality, which does not reveal the real etiology of infertility (51). In approximately 30-40% of cases, the cause of male infertility cannot be defined.

1.3.1 Subtypes of male infertility

The causes of male infertility can be pre-testicular (endocrine disorders originating from hypothalamus, pituitary, or adrenals), testicular (varicoceles, testicular cancer, chemotherapy, or trauma), and post-testicular (reproductive tract obstructions, disorders in ejaculation, erectile dysfunction). Pre-testicular and testicular disorders have an adverse effect on spermatogenesis, while post-testicular disorders affect the transit of sperm in the urogenital tract, and the sperm motility due to improper maturation and storage in the epididymis (52). Pre-testicular, testicular and post-testicular disorders lead to disruption of control mechanisms which affect sperm count, motility, and morphology. Based on WHO guidelines on semen analysis, and particularly sperm count, there are two types of male infertility: (i) oligospermia, with decreased sperm concentration (<15 million sperm/mL); and (ii) azoospermia, with non-detectable sperm in semen (53). Azoospermia can be further sub-classified into obstructive azoospermia (OA), caused by a physical blockage in the epididymis or vas deferens, and non-obstructive azoospermia (NOA), associated with spermatogenesis failure (54). Hypospermatogenesis (HS), maturation arrest (MA) and Sertoli cell-only syndrome (SCO) are three forms of NOA with increasing severity (55). Azoospermia affects approximately 2% of the general male population, and accounts for ~15% of all male infertility cases. NOA occurs in ~60% of men diagnosed with azoospermia (54). Additional types of male infertility, as assessed by semen analysis, are asthenospermia, characterized by poor sperm motility, and teratospermia, presented with abnormal sperm morphology (53).
1.3.2 Diagnosis of male infertility

The diagnosis of male infertility has two goals: (i) to identify the causes of infertility, and apply treatments that will allow natural conception and ARTs; and (ii) to investigate whether the causes of infertility can also pose risk to man’s health, like infections, mutations in CFTR, or other genetic disorders. The initial assessment for evaluation for male infertility includes complete medical history, physical examination, and at least two semen analyses (56). Additional tests for endocrine evaluation, genetic tests, and specialized sperm evaluation can also be completed.

Typically, semen analysis is the basic standard test to evaluate male infertility (56). The severity of infertility is defined by deviation from the following WHO guidelines and reference values (53): ~1.5 mL volume, pH >7.1, appearance and color, viscosity, and evaluation of coagulation and liquefaction (57). Microscopic parameters analysis consist of sperm concentration (cut-off 15 million sperm/mL), sperm motility and morphology, as well as presence of non-sperm cells, such as immature germ cells and leukocytes (57). Overall, a routine semen analysis can identify subtypes of male infertility, such as azoospermia and oligospermia, asthenospermia and teratospermia. Normal semen analysis does not rule out infertility. Defects in sperm function can be determined by performing additional tests, such as sperm capacitation test, presence of anti-sperm antibodies, DNA fragmentation, sperm penetration assay (SPA), and ZP binding assay (58, 59). When semen analysis indicates azoospermia, further investigation is required to differentiate between OA and NOA. Physical examination and serum levels of FSH and inhibin B can point to OA or NOA (60, 61), however testicular biopsy is currently the gold standard to reliably differentiate between the two forms (62). Besides being a highly invasive procedure with high risk of complications, testicular biopsy cannot always diagnose different subtypes of NOA due to spatial heterogeneity of spermatogenesis in the testis. On the other hand, screening for cystic fibrosis transmembrane conductance regulator (CFTR) mutations can confirm the presence of OA due to congenital bilateral absence of vas deferens or epididymis, which accounts only for 2% of OA patients (63), while karyotype abnormalities and Y-chromosome microdeletions are most of the times indicators of NOA (64) (Figure 1.1).
Figure 1.1 General algorithm for the evaluation of male infertility. Initial assessment is based on medical history, physical examination, and routine semen analysis. Additional tests are performed to confirm infertility subtype. Clinical management and possible ART options are also included.

To summarize, routine semen analysis is currently the standard test for evaluation and diagnosis of male infertility. Recent studies on male reproduction and infertility show that semen analysis has a poor predictive value, and it should be always combined with clinical examination (56). Normal semen analysis does not rule out fertility issues, thus there are more than 30% of cases classified as unexplained, or idiopathic infertility (15). In azoospermia, testicular biopsy, and additional genetic tests are currently used to confirm the different subtypes. Thus, there is a need for new diagnostic and prognostic tools, that when combined with semen analysis will be able to provide better treatment options.
1.3.3 Current treatments

In most of the cases male infertility is associated with impaired spermatogenesis. Due to the complexity of the disease, tailored treatment for every sub-fertile, or infertile man is not yet available. Application of ART is a generic treatment to most infertility cases. Semen analysis is the main determinant of the treatment that a man will receive, and the more severe the abnormality, the more comprehensive is the treatment. A general treatment approach includes improvement of lifestyle (diet, smoking, exercise, stress) (65-68). When endocrine disorders are the leading cause of male infertility, different hormonal therapies, such as testosterone, human chorionic gonadotrophin (hCG), gonadotropin-releasing hormone (GnRH), or estrogen modulator administration, can be applied to restore spermatogenesis (69, 70).

When a patient is diagnosed with azoospermia, more invasive treatment approaches can be applied. In cases of OA with known congenital or acquired epididymal or vassal obstruction, microsurgical reconstruction can be used to restore fertility (71). In this way, pregnancy can be achieved without any ART techniques, some of which are invasive, and may cause complications for the female partner. In most NOA cases, and in some OA cases that microsurgical reconstruction is not feasible, testicular sperm extraction (TESE) is performed at the time of biopsy, or separately (72, 73). Microscopic TESE (microTESE), as opposed to standard TESE, can improve sperm retrieval rates with minimal testicular tissue extraction, since it can identify pockets with active spermatogenesis (74). Success rate of microTESE is overall ~50-60%, but it varies among the three histopathological subtypes of NOA (72). After successful sperm retrieval, sperm can be directly used for ART, or it can be cryopreserved for future ART.

Assisted reproduction techniques can be applied in cases of oligospermia, azoospermia, and idiopathic infertility to achieve pregnancy. There are three major methods of assisted reproduction: artificial insemination, intracervical (ICI), or intrauterine (IUI), *in vitro* fertilization (IVF), and intracytoplasmic sperm injection (ICSI) (75). Artificial insemination is recommended in cases of moderate male factor infertility, and ovarian stimulation is optional (76). The overall success rate of IUI or ICI are lower compared to other assisted reproduction methods, therefore couples with 3-6 unsuccessful IUI should consider an alternative approach. For more severe male factor infertility, *in vitro* fertilization is usually the method of choice. In IVF, fertilization occurs outside the body, and successfully formed embryos are implanted in the
uterus. In general, IVF requires a minimum of 50,000-500,000 motile sperm (~100,000 sperm per oocyte) to achieve oocyte fertilization. Therefore, IVF is not suitable for men with severe oligospermia and azoospermia (75). Modifications to IVF procedure, led to the development of ICSI, for which a single sperm is required for oocyte fertilization. Men who fail IUI and IVF, and men who have severe abnormalities (azoospermia, or idiopathic infertility), can undergo sperm retrieval, and father children through ICSI (75, 77).

1.3.4 Emerging male infertility biomarkers

As discussed in section 1.3.2, standard semen analysis has a poor predictive value to assess sperm quality and fertility outcome, even when supplemented with additional tests, e.g. genetic tests, presence of anti-sperm antibodies, functional sperm tests. For many cases, these tests are normal or inconclusive, resulting in idiopathic infertility diagnosis. Therefore, alternative, more accurate diagnostic methods are required for the management of male sub-fertility and/or infertility. Integration of genomics, transcriptomics and proteomics may be a promising approach for more accurate diagnosis and treatment of male infertility.

Semen is a complex fluid, with cellular (spermatozoa), and non-cellular components (seminal plasma). To date, many studies have focused on spermatozoa for the discovery of male infertility biomarkers (78, 79). Nevertheless, seminal plasma can be a promising source of biomarkers for male infertility. All cellular functions are carried out by proteins, and seminal plasma consists of proteins that are relevant to sperm function, and sperm functional interactions with the male and female reproductive tract. Seminal plasma proteins are secreted from various glands of the male reproductive tract, and particularly ~10% of proteins originate from the testis and epididymis, ~65% from the seminal vesicles, ~25% from the prostate, and ~1% from the other accessory glands (38). Interestingly, more than 1,000 genes display a tissue-enriched expression pattern across male tissues (testis, epididymis and prostate) (25). Most of these genes are coding for proteins involved in spermatogenesis and sperm maturation. The majority of these proteins are intracellular, but there is still a large number of extracellular membrane-shed or secreted proteins that are present in seminal plasma (80). Alterations in seminal plasma proteomic profile may reflect disorders in the male reproductive system, and facilitate diagnosis and prediction of male fertility status.
Genomic alterations, such as single nucleotide variations (SNVs) and somatic mutations, can be initially identified in the DNA level, and then confirmed in the protein level in seminal plasma or and spermatozoa, and potentially can be used for the diagnosis of male infertility (81, 82). There are evidences that SNVs in genes that are involved in spermatogenesis and sperm maturation may be the cause of male infertility, and particularly the unexplained, or idiopathic cases (82-84). This is also supported by the fact that approximately 500 mouse models, generated by targeted testis-specific gene disruption, were phenotypically normal, but subfertile, or infertile due to defects in sperm production and maturation (81, 85). Likewise, testis-specific genes in humans may carry polymorphisms that alter the expression and function of the corresponding proteins, resulting in the disruption of spermatogenesis (86-90).

Other types of molecules that can serve as infertility biomarkers are messenger RNA (mRNA), micro RNA (miRNA), or metabolites (91-94). It should be noted, that the blood-tissue barriers in the male reproductive tract, are highly stringent and selective. Therefore, only hydrophobic molecules, such as hormones or their metabolites, are transported into blood plasma and can be detected (23, 30). Therefore, seminal plasma is the most useful proximal fluid that can be used for the discovery of novel male infertility biomarkers.

Seminal plasma has also been proposed for the non-invasive differential diagnosis of azoospermia. Although azoospermia is the most extensively studied infertility condition, there is still a need for an alternative diagnostic approach which will replace testicular biopsy. Comparative proteomic profiling of seminal plasma between fertile and azoospermic men has been performed by different groups aiming at identifying candidate markers. In one of the most comprehensive studies, sixteen proteins (e.g. TEX101, LDHC) could differentiate fertile controls from post-vasectomy (simulated OA), three could differentiate fertile controls from NOA (e.g. TEX101, LDHC), and eleven proteins were different between post-vasectomy and NOA (e.g. ECM1, CRISP1, SPAG11B). In the validation phase, two proteins, TEX101 and ECM1, outperformed in differentiating fertile control, NOA and OA cases (95-97). Those studies demonstrated the potential of seminal plasma as a source of non-invasive male infertility biomarkers. Nevertheless, the translation of seminal plasma biomarkers from bench to bedside is still slow.
1.4 Innovative proteomics strategies to study clinical and functional aspects of male reproduction and infertility

Proteomic technologies have recently been introduced to the field of reproductive biology. A wide range of proteomic approaches are available to study the male reproductive system, and to identify proteins that play a crucial role in processes like spermatogenesis, sperm maturation, and sperm-egg interaction. High-throughput proteomic technologies, such as mass spectrometry, have facilitated the proteome mapping of human testicular tissue (98), spermatozoa (79), and seminal plasma (95, 99, 100), enriching our knowledge about sperm function, and creating a pool of candidate male infertility biomarkers. Recent advances in genomics, transcriptomics and proteomics have led to the generation of publically available large-scale reference data sets that contain information for the majority of human protein-coding genes and their the expression, localization, and tissue specificity, based on RNA and protein data (101).

Transcription and translation of germ cell-expressed proteins is terminated at the stage of haploid spermatocytes, when protamines replace histones during sperm DNA compaction (102, 103). Thus, further regulation of spermatogenesis and sperm maturation is carried over exclusively through post-translational modifications and protein-protein interactions (104). Therefore, proteomics is the tool of choice for investigating the molecular mechanisms that regulate the function of sperm. Furthermore, proteomics may facilitate the identification of proteins that are responsible for the production of defective sperm, and may be implicated in the diagnosis and treatment of male infertility.

1.4.1 Reagents and tools to study testis-specific proteins

Although male tissue-specific genes comprise almost 5% of the human genome (25), most of them are not well studied and their function remains unknown. Majority of commercially available antibodies are typically generated against peptides or protein fragments and fail to bind the native protein conformations. Thus, they are mostly suitable for Western blot analysis, but not for more comprehensive assays such as ELISA, immunoprecipitation, and flow cytometry. Antibodies that capture the native form of the antigen are required for performing functional studies, and investigating protein-protein interactions (PPIs). However, these
antibodies are available only for few targets. Reagents such as monoclonal antibodies that capture the native protein conformation in biological specimens, are a valuable tool to identify proteins, to study their function, and moreover to develop sensitive tests for their detection in various pathological conditions.

Production of high quality monoclonal antibodies is crucial for proteomic studies. Western blot and solid-phase immunoassays have been the methods of choice for screening of antibody clones. The disadvantage of this approach is the selection of antibodies that may not always recognize the native endogenous proteins (105). Mass spectrometry-based approaches have been recently introduced for rapid screening of hybridoma clones, and selection of high affinity and specificity monoclonal antibodies against native form of proteins, eliminating the risk of cross-reactivity (106, 107). Monoclonal antibodies and corresponding immunoassays can have multiple applications in male reproductive biology. Monoclonal antibodies can be used for the characterization of testis-specific protein targets, and mapping of testis-specific interactomes. Such findings can elucidate the functional role of each protein in reproduction, and potential implications in male infertility. On the other hand, sensitive and target-specific immunoassays can assist the clinical validation of candidate infertility biomarkers emerge as clinical diagnostic tests.

1.4.2 Protein-protein interactions

Transcription and translation silencing at late stages of spermatogenesis and sperm maturation indicate that most of these processes are accomplished through protein-protein interactions. Studies in mice have demonstrated the importance of PPIs in the production of functional sperm, and the defects in fertilization when these interactions were disrupted (40, 108-114). There are various in vivo and in vitro techniques for identifying and characterizing PPIs (115-123). Affinity purification or co-immunoprecipitation approaches followed by mass spectrometry emerged as a technique of choice to identify direct and indirect PPIs under native physiological conditions (124-126). High resolution mass spectrometry, combined with sophisticated tools and software enable the identification of high-confidence PPIs (127-129). The availability of antibodies recognizing the native form of the proteins of interest is a cornerstone in this approach. In addition, in order to overcome the challenge of identifying transient PPIs, protein cross-linking chemistry emerged as a new technology that can “freeze”
transient protein interactions by covalently binding protein complexes (130-132). The ultimate goal in the area of protein-protein interactions is to pursue and achieve the development of approaches for global scale analysis of protein complexes.

1.4.3 Global and targeted mass spectrometry approaches

A combination of global and targeted mass spectrometry-based approaches are usually employed for the identification and quantification of proteins in various biological specimens. Global, or shotgun mass spectrometry, can provide information regarding the presence of a protein and different protein isoforms in a sample, as well as the presence of post-translational modifications. Typical targeted mass spectrometry-based approaches include selected reaction monitoring (SRM) and parallel reaction monitoring (PRM). SRM and PRM assays can be used for the relative or absolute quantification of potentially any proteoform. Global mass spectrometry analysis and comparative profiling of seminal plasma and spermatozoa in different physiological or pathological states reflect alterations which may be hardly predicted at the genome level (91, 95, 99, 133-135). Differentially expressed proteins can be further validated by using multiplex targeted mass spectrometry–based protein assays (SRM and PRM) for quantitative analysis of candidates. This is the pipeline that integrates discovery and verification of candidate biomarkers for various physiological states (136, 137). Coupling high-affinity antibodies with mass spectrometry, tandem (MS/MS) or targeted (SRM), can increase the sensitivity of the method, and drop the limit of quantification to low ng/mL (138). Immunocapture-mass spectrometry and traditional immunoassays (ELISA) can detect and quantify analytes with high sensitivity. However, immunocapture-mass spectrometry outperforms immunoassays in specificity and its potential for multiplex analysis, thus it can serve as an alternative approach.

A novel application of proteomics and particularly mass spectrometry comes from the integrative analysis of proteomic and genomic data (139). Combination of high-throughput next-generation sequencing and deep mass spectrometry analysis facilitates the identification and quantification of protein-coding variants such as mutations or polymorphisms at the protein level (140). Global mass spectrometry-based analysis with customized reference protein databases can be used to detect variants at the protein level (141). Finally, targeted mass-spectrometry assays can be developed to quantify these variant proteoforms in hundreds
biological samples. This novel approach can provide new insights into the expression and function of proteins, as well as the effect of each variant at the protein level. It can be applied in order to investigate the function of testis-specific proteins by identifying protein-coding variants that affect spermatogenesis or sperm maturation, and they are likely to be associated with male infertility (82, 84).

1.5 TEX101 protein: a biomarker of spermatogenesis and male infertility

In this work, we were interested in studying the function of a previously discovered testis-specific infertility biomarker (97). TEX101 protein was initially identified in mice, and it was proven to be essential in the production of fertile sperm (113, 142). In the quest of novel male infertility biomarkers in SP, TEX101 outperformed other proteins expressed in different tissues of the male reproductive tract, and emerged as a promising candidate biomarker for the differential diagnosis of azoospermia (97).

1.5.1 TEX101 expression pattern at transcript and protein level

TEX101 protein (encoded by testis expressed 101 gene, Tex101) was originally identified in mice (142). The Tex101 gene is located on the long arm of chromosome 7 in mouse, and on chromosome 19, position 19q13.31, in humans (Figure 1.2a). Alignment of TEX101 protein sequences shows 55% homology between the two species. Mouse Tex101 gene consists of five translated exons, and forms three major transcripts under the regulation of two distinct promoters (143). Based on nucleotide and amino acid sequence predictions (www.uniprot.org), alternative splicing of human TEX101 results in two isoforms: an extracellular membrane-bound isoform Q9BY14-1 (249 amino acids) and a cytosolic, intracellular isoform Q9BY14-2 (267 amino acids) (Figure 1.2b). According to the Human Protein Atlas (HPA) (www.proteinatlas.org), human TEX101 is exclusively expressed by germ cells, and not by any other human tissue or cell type, including Sertoli and Leydig cells of the testicular tissue.

The 3D structure of TEX101 protein remains unresolved. Based on the amino acid sequence, TEX101 possesses a Ly-6/uPAR domain (Lymphocyte Antigen 6/urokinase Plasminogen Activator Receptor), defined by 8-10 distinct cysteine residues that are highly conserved among
mammalian species (Figure 1.2c). Therefore its 3D structure may resemble that of uPAR. A number of proteins, comprising the Ly-6 family, share this single Ly-6/uPAR domain, and they are mapped to chromosome 8. However, the uPAR locus, PLAUR, encoding three tandem Ly-6/uPAR domains, is found at 19q13.3 (144). Interestingly, a few members of the Ly-6 family are mapped at the same locus, among which is TEX101 (19q13.31) and SAMP14 (19q.31.33). Based on this fact, it was speculated that these proteins are more closely related to uPAR, and may share the same or similar functions (145).

Figure 1.2 Genomic, transcriptomic and proteomic organization of TEX101. (a) Chromosomal location, orientation and position on chromosome of TEX101 gene. (b) Structure of TEX101 gene and its splice variants. Protein-coding exons are presented with dark grey boxes. (c) TEX101 protein structure. Signal peptide, glycosylation sites, C-terminal Propeptide, and uPAR domain are highlighted with red, orange green, and blue, respectively. Protter tool for proteoforms visualization was used (146).
TEX101 is predicted to have three potential post-translation modification sites; two N-glycosylation sites at positions Asn-45 and Asn-159, and one lipidation site at position Asn-222 ([www.uniprot.org](http://www.uniprot.org)). Lipidation results in the attachment of a glycophosphatidylinositol (GPI) group to the C-terminus of the protein, and its trafficking to the extracellular plasma membrane (147). Mouse and human TEX101 proteins migrate around ~35 kDa (148). It has been shown that TEX101 protein is glycosylated at all potential sites in mice (149), while in humans the molecular weight of the mature bare protein is 21 kDa, and glycosylation leads to molecular masses around 29-35 kDa.

Glycosylation plays a critical role in the folding, the conformational stability, and the function of glycoproteins (150), and it is implicated in protein-protein interaction (151). It is estimated that more than 50% of all human proteins are glycoproteins (152). Glycosylation is essential in the process of spermatogenesis and extracellular quality control of sperm (153). Several testis-specific and sperm-specific glycoproteins are involved in sperm migration and capacitation, cumulus penetration, zona-pellucida binding, and sperm-egg membrane fusion (154). Human TEX101 is a glycosylated extracellular membrane-bound protein, however the specific biological role of TEX101 glycosylation remains unknown.

1.5.2 TEX101 localization during spermatogenesis and sperm maturation

As mentioned in the previous section, human TEX101 is exclusively expressed by male germ cells and not by any other cell type or human tissue, as assessed by mRNA expression and protein immunohistochemistry staining provided by HPA. Based on prediction algorithms, two forms of TEX101, a membrane-bound and a secreted form, are present in the testicular tissue. Looking into the expression of TEX101 during spermatogenesis, there is evidence that TEX101 is absent in diploid spermatogonia stem cells, and its expression starts at the stage of primary spermatocytes after mitosis. Thereafter, haploid secondary spermatocytes, haploid spermatids, and haploid testicular sperm express TEX101 protein, mostly located on the plasma membrane, and only weakly present in the cytoplasm (HPA).

As a GPI-anchored protein, TEX101 is localized in specialized membrane microdomains, called lipid rafts (155). Lipid rafts differ from plasma membranes in their lipid composition, and they are enriched in cholesterol, sphingolipids, such as sphingomyelin, and GPI-anchored proteins.
Proteins that are localized in sperm lipid rafts are implicated in sperm-zona pellucida and sperm-oocyte recognition (158). GPI-anchored proteins expressed on the sperm membrane are most of the times shed from the surface for two reasons: (i) they are activated after being released from sperm surface; (ii) their cleavage facilitates the sperm-zona pellucida binding (159). TEX101 is one of the GPI-anchored proteins which are cleaved from the sperm surface and released into seminal plasma during epididymal maturation (160). It was demonstrated that TEX101 is cleaved by a testis-specific isoform of angiotensin-converting enzyme (tACE) during post-testicular maturation (113, 159, 160). Testicular ACE isoform is also a GPI-anchored protein, and it is shed from sperm surface in the epididymis (161). Additional in vivo and in vitro experiments indicated that the removal of TEX101 by tACE occurs in a substrate-specific manner, while other testis-specific GPI-anchored proteins, like SPACA4, remain unaffected (145). The human GPI-ase enzyme, that is responsible for the cleavage of TEX101 from sperm, is still not identified.

Other studies provide evidence that TEX101 is only partially cleaved from the surface of testicular spermatozoa during epididymal maturation, and it is still present in the lipid rafts of mature sperm (155). Immunofluorescence analysis of mouse and human sperm showed that TEX101 is localized to the post-acrosomal region of mature sperm (148), a region involved in sperm-egg interaction (162).

1.5.3 Suggested functional roles of TEX101 protein

To date, most available literature data on TEX101 is based on mouse but not human data. Targeted Tex101 gene disruption in mice revealed its direct involvement in fertilization (113, 163). Two independent studies demonstrated that Tex101\(^{-/-}\) mice had normal mating ability, but they were not able to produce offspring, confirming the infertile phenotype. Interestingly, there was no significant difference in the weight and the histology of the testis between Tex101\(^{-/-}\) and Tex101\(^{+/+}\) mice. In addition, sperm count, acrosome reaction efficiency, and sperm motility and viability parameters were examined, and no significant differences were found. Notably, heterozygous Tex101\(^{+/-}\) mutant mice were fertile, while disruption of Tex101 gene in female mice had no effect on the pregnancy rate.
To elucidate the functional role of TEX101, further *in vivo* and *in vitro* studies were conducted in mice. It was shown that *Tex101* knockout mouse model had identical reproductive phenotype with the well-studied *Adam3* knockout model. In both cases, sperm was produced, but it was incompetent for fertilization (113, 163). The critical role of ADAM3 protein in fertilization is exerted through the formation of complexes with other ADAM proteins, such as ADAM2, ADAM4, ADAM5 and ADAM6 (108, 164, 165). It has been shown that these ADAM-containing complexes which are present in the sperm membrane are involved in the sperm-egg plasma membrane adhesion and fusion, sperm migration from the uterus to the oviduct, as well as sperm-sperm association (109, 166-172). In the absence of TEX101 protein, ADAM proteins were expressed and targeted to the sperm membrane, however, they were subjected to non-specific degradation from proteases present in the epididymal fluid during post-testicular sperm maturation (113). Impaired maturation of ADAM proteins led to sperm fertilization defects (113, 163). A combination of immunoprecipitation, western blotting and mass spectrometry analysis, revealed that TEX101 interacts with these testis-specific members of the ADAM family on the surface of testicular germ cells (113, 163). These findings suggested that TEX101-ADAMs protein interactions occur post-translationally on the outer cell membrane, and unlike other proteins (43), TEX101 does not mediate trafficking of ADAMs to the sperm surface. Moreover, when ACE gene was disrupted, both TEX101 and ADAM3 proteins remained on sperm (113). This led to aberrant localization of ADAM3 on epididymal sperm, and male mice were infertile (173).

TEX101 has been also identified in the cumulus-sperm interaction during the acrosome reaction (148). In contrast to previous studies, demonstrating that TEX101 disappears after epididymal maturation (113, 160, 163), it was shown that TEX101 is detected on mature spermatozoa (148). It was suggested that the remaining TEX101 protein on mature sperm participates in the sperm-cumulus cross-talk. Immunostaining of mouse cumulus cells with anti-TEX101 antibody was positive, after incubation of cumulus cells with mouse sperm (148). According to the proposed model, TEX101 cleavage and binding to the surface of cumulus cells leads to Ca^{2+} mobilization and progesterone production by the cumulus (148). These events are known to induce acrosome reaction, and facilitate the penetration of the cumulus layer. It was suggested that TEX101 may bind to the uPA/uPAR complex expressed on the surface of cumulus cells (174), triggering Ca^{2+} mobilization and progesterone release required for acrosome reaction. An independent study in cancer cells showed that TEX101 could bind to uPA/uPAR complex and interfere with uPA
activity (175). The uPA system is known to be involved in signal transduction, apart from its serine protease activity, and is associated with Ca\(^{2+}\) mobilization (176). However, there was no follow-up study to explain the mechanism behind this proposed model.

Overall, there is evidence that TEX101 protein exerts its function through interactions with other testis-specific proteins on the surface of germ cells. Interestingly, the genes that were reported to be involved in the formation of complexes with *Tex101* on male germ cells in mice, have no orthologs in the human genome and the corresponding proteins are not expressed in the human testis. Notably, transcriptomic studies have shown that there are fundamental differences in the dynamics of gene expression between human germ cells and that of rodents (177). Therefore, it may not be possible for results obtained from mouse models to be directly translated into human reproduction studies.

**1.5.4 TEX101 as a biomarker for the differential diagnosis of azoospermia**

In quest for male infertility biomarkers, the SP proteome project was previously launched for the identification of proteins by using shotgun mass spectrometry (99). In the first phase, more than 3,000 proteins were identified in SP samples of fertile and in fertile men (95, 99). Proteins that were expressed in the male urogenital tract, particularly in the testis and epididymis, and they showed differential abundances in the three clinical groups (normal spermatogenesis, NOA and OA/post-vasectomy (PV)) were selected as promising candidates for the differential diagnosis of azoospermia (96). After the verification and validation of the top candidate proteins by mass spectrometry-based SRM and immuno-SRM, two proteins (ECM1 and TEX101) emerged as biomarkers for the differential diagnosis of NOA versus OA (95-97). Simply by measuring the two proteins, the testis-specific protein TEX101 and the epididymis-expressed protein ECM1, in SP, a two-marker algorithm was introduced, and potentially applied for the non-invasive differential diagnosis of azoospermia with high sensitivity and specificity (Figure 1.3) (97). Furthermore, TEX101 protein levels in testicular tissue and SP could differentiate between NOA subtypes. TEX101 levels ≥120 ng/mL indicated normal spermatogenesis, while levels of 5-120 ng/mL were associated with HS or MA, and levels below 5 ng/mL (theoretically zero) represented SCO syndrome (97).
Figure 1.3 Two-marker algorithm for the differential diagnosis of azoospermia (OA versus NOA) with seminal plasma proteins ECM1 and TEX101. When azoospermia is diagnosed by semen analysis, low SP levels of ECM1 (<2.3 µg/mL) and TEX101 (<5 ng/mL) suggest OA, while high SP levels of ECM1 (>2.3 µg/mL) suggest NOA. SP concentration of TEX101 protein may also discriminate between NOA subtypes of Sertoli cell-only (5 ng/mL) and hypospermatogenesis or maturation arrest (5-120 ng/mL). Men with OA have good chances of sperm retrieval by TESE, while for men with Sertoli cell-only, sperm retrieval is unlikely and TESE can be avoided (97). From Drabovich, A.P. et al., Sci. Transl. Med. 5, 212ra160 (2013). Reprinted with permission from AAAS.
TEX101 immunohistochemistry and quantitative mass spectrometry data were combined, as well as previous knowledge about the biology of TEX101 protein, and a theory about TEX101 expression and secretion into SP was generated, to explain the variability of TEX101 concentrations in different forms and subtypes of azoospermia (Figure 1.4) (97). This theory stems from the following facts: (i) TEX101 is a GPI-anchored membrane protein exclusively expressed in germ cells, but not any other cell type or tissue; (ii) TEX101 is shed from the surface of spermatozoa during epididymal sperm maturation (113, 160); and (iii) soluble TEX101 is implicated in the interaction of sperm with cumulus cells and the acrosome reaction of sperm (46, 148). Thus, in cases of physical obstruction of vas deferens, or absence of germ cells in the testis, TEX101 levels are undetectable (theoretically zero) in SP obtained from patients with OA, PV and SCO syndrome. In MA and HS cases, germ cells are present in the testis, and TEX101 is expressed. However, spermatocytes maturation is arrested, and sperm never pass through the epididymis to allow cleavage of TEX101. The low amounts of TEX101 (<120 ng/mL) that are detected in SP, may come from non-specific cleavage of TEX101 from the surface of immature sperm inside the testis. Taking into account all existing data, we speculate that the concentration of TEX101 in SP may correlate with the number of germ cells in the testis, and reflect the extent of spermatogenesis. As a result, TEX101 could be an informative biomarker for the male fertility status, ranging from azoospermia to oligospermia and normal spermatogenesis. In addition, TEX101 could be used to predict the success of sperm retrieval by TESE in NOA cases, and thus eliminate the testicular biopsies.
Figure 1.4 Model of TEX101 expression and secretion into SP in men with normal spermatogenesis (NS) and azoospermia. In tissues with NS, TEX101 is expressed in spermatocytes, which mature into sperm cells and move to epididymis. In epididymis, GPI-anchored TEX101 is cleaved from the sperm and is released into SP, at a concentration of about 2 µg/mL. In men with OA, spermatogenesis occurs, but TEX101 is not detectable in SP because of the physical obstruction. In NOA with HS and MA, spermatocytes do not mature and spermatozoa never pass through the epididymis to allow TEX101 cleavage. Non-specific shedding of TEX101 from spermatocytes in the seminiferous tubules results in its low levels in SP (<120 ng/mL). In NOA with SCO, germ cells are absent, so TEX101 is not expressed and detected in SP (97). From Drabovich, A.P. et al., Sci. Transl. Med. 5, 212ra160 (2013). Reprinted with permission from AAAS.
1.6 Rationale and Objectives of the present study

1.6.1 Rationale

There are more than 1,000 human genes identified as testis-specific, with ~15% of genes associated with reproduction, and ~10% with spermatogenesis (25). Although, male reproductive processes, such as spermatogenesis and sperm maturation, have been studied more or less extensively in various species, including humans, there is still a large number of male tissue-specific genes for which the functional role is yet unknown. Understanding the role of each gene in the reproductive process will facilitate the elucidation of complex interactions under physiological and pathological conditions. This knowledge is essential for the identification of potential biomarkers and the development of treatments for a wide range of infertility cases.

The Testis Expressed 101 (TEX101) protein is a typical example of a testis-enriched gene that has been previously associated with spermatogenesis and fertilization in mice (113, 163). Although there are several studies on murine TEX101 and reproduction, our knowledge about the human protein is still very limited. Human TEX101 has been previously studied only in the context of male infertility biomarker discovery, and it was reported as a biomarker for azoospermia (97). However, further investigation were required in order to determine if TEX101 levels in SP: (i) could reflect the rate of spermatogenesis, particularly in different groups of NOA patients; and (ii) correlate with sperm concentration in semen in various subfertility groups.

Previous studies in mice demonstrated that TEX101 exerted its function through protein-protein interactions with other testis-specific proteins on the surface of germ cells and sperm (113, 163). We assumed that the function of TEX101 in humans may be similar to the one in mice, however, we expect that human and mouse TEX101 interactome may differ, due to the lack of some orthologs and expression of different genes in each species. Until now, functional studies on TEX101 in humans have not been conducted, thus this area remains unexplored. Discovery of human TEX101 interactome may reveal proteins that are essential for human reproduction, like ADAM3 in mice. Furthermore, in the lack of human in vitro germ cell lines, identification of human TEX101 knockouts will aid in the better understanding of TEX101 function.
With respect to further evaluating the potential of TEX101 as a male infertility biomarker, and understanding the functional role of human TEX101 in the production of fertile spermatozoa, we used mass spectrometry-based approaches. To evaluate TEX101 as a male infertility biomarker, we developed a simple, sensitive and higher throughput antibody-based ELISA, which was then used to measure TEX101 concentration in a large number of seminal plasma samples from men with different fertility status. To investigate the functional role of human TEX101, we used high-resolution mass spectrometry to identify physical and functional TEX101-interactions in human biological samples. To overcome the lack of human cell lines or established in vitro models for human germ cells, we performed our studies by using human testicular tissue and semen samples. In our attempt to understand the functional role of human TEX101, we sought for a human knockout or knockdown model, in which TEX101 gene expression or protein turn-over would be disrupted due to a missense amino acid substitution.

1.6.2 Hypothesis

Exclusive testis-specificity of TEX101 protein suggests its unique biological function in spermatogenesis and fertilization. However, previous data about TEX101 functional interactome in mice, cannot be translated into human reproduction due to lack of TEX101-interacting orthologous genes and proteins in humans. In addition, disruption of Tex101 gene leads to absolute sterility in mice. Studies on human TEX101 showed that it is a promising non-invasive biomarker for the differential diagnosis of azoospermia, and can distinguish between SCO cases from HS and MA. Here we undertook a two-stage approach. First, from a clinical perspective, we hypothesized that a sensitive TEX101 antibody-based test will facilitate the high-throughput measurement (>200 samples) of TEX101 concentration in SP. Using this test, we will evaluate the power of TEX101 biomarker: (i) to discriminate better among different NOA groups, and predict sperm retrieval; and (ii) to distinguish oligospermia from normal spermatogenesis, and sub-classify oligospermia cases in mild and severe oligospermia. Additionally, we hypothesized that measurement of TEX101 concentration in a large number of samples may lead to the discovery of human natural TEX101 knockouts or knockdowns, which can be used to study TEX101 function in humans. Second, from a functional perspective, we hypothesized that humanTEX101 function is exerted through physical and transient PPIs with cell-surface testis-
specific proteins, and it may be involved in the trafficking and maturation of proteins crucial for fertilization.

1.6.3 Objectives

We present immunoassay- and mass spectrometry-based approaches to study the clinical and functional aspects of human TEX101 protein.

The objectives of our study were:

1. Production of mouse monoclonal antibodies against native TEX101 protein, and development of a sensitive TEX101 immunoassay (ELISA) and measurement of TEX101 in biological fluids.

2. Characterization of TEX101 protein as an analyte in seminal plasma, and evaluation of TEX101 ELISA performance in a large population of fertile, subfertile and infertile men.

3. Development of a quantitative co-immunoprecipitation-MS (co-IP-MS) platform for the discovery and verification of human TEX101 physical interactions in human testicular tissue and spermatozoa.

Chapter 2
Development of ELISA for the measurement of native TEX101 in biological fluids

This chapter includes data published in *Molecular & Cellular Proteomics*: “Immunocapture-selected reaction monitoring screening facilitates the development of ELISA for the measurement of native TEX101 protein in biological fluids” by Korbakis D.†, Brinc D.†, Schiza C.†, Soosaipillai A., Jarvi K., Drabovich A.P., and Diamandis E.P. *Molecular & Cellular Proteomics* 2015, 14 (6): 1517-26. †authors contributed equally

A link to the published paper can be found at
doi: 10.1074/mcp.M114.047571

Copyright permission has been granted.
Chapter 2
2 Development of ELISA for the measurement of native TEX101 in biological fluids

2.1 Introduction

Monoclonal antibodies that bind the native form of a protein are indispensable for the development of sensitive immunoassays, production of therapeutic antibodies and for studying protein interaction networks by affinity purification-mass spectrometry \((126, 178)\). Large-scale purification of native proteins from biological samples may be challenging, so recombinant proteins or protein fragments are often used for antibody production. Antibodies produced against short peptides, protein fragments or even full length recombinant proteins, however, may not bind the native protein conformation present in biological fluids, thus limiting the utility of antibodies. Rapid screening of antibody-producing hybridoma clones for native proteins requires highly specific and sensitive assays, performed under non-denaturing conditions. Here, we report the capability of an immunocapture-SRM assay to facilitate fast screening of hybridoma cultures for monoclonal antibodies that recognize the native conformation of TEX101 protein in biological fluids.

TEX101 emerged as a non-invasive biomarker for the differential diagnosis of azoospermia. Concentration of TEX101 was initially measured in SP by mass spectrometry-based SRM and immuno-SRM assays, with limits of detection of 120 and 5 ng/mL, respectively \((179, 180)\). However, due to the ultra-wide range of TEX101 concentrations in seminal plasma of infertile and healthy men \((0.5 \text{ ng/mL to } 50,000 \text{ ng/mL})\) and theoretically zero levels for some azoospermic patients, a sensitive TEX101 immunoassay is required to develop a clinical laboratory test. Since TEX101 may emerge as a novel biomarker of male infertility, in this work we focused on the development of an ELISA for sensitive measurement of TEX101 in seminal plasma and serum.

Our initial efforts to develop a TEX101 immunoassay using commercially available polyclonal antibodies were not successful. We found that commercial antibodies recognized only the denatured form of TEX101 and were useful for immunohistochemistry and Western blots, but not for the analysis of native TEX101 in seminal plasma. Here, we describe the production of mouse monoclonal antibodies against native TEX101, screening of antibody-producing clones
by the two-step immunocapture and SRM assay, development of a sensitive ELISA and measurement of TEX101 in seminal plasma and serum (Figure 2.1).

**Figure 2.1 Pipeline for the production of mouse monoclonal anti-TEX101 antibodies and screening of colonies using two-step immunocapture-SRM assay.** Screening included the coating of microtiter plates with sheep anti-mouse IgG antibodies, the addition of hybridoma cell supernatants, incubation with seminal plasma containing the native form of TEX101 followed by trypsin digestion and SRM analysis. Two-step immunocapture followed by SRM detection facilitated rapid screening of antibody-producing colonies. Eventually, all positive clones were expanded and a sensitive immunofluorescent assay for TEX101 was developed in seminal plasma and serum.
2.2 Materials and Methods

2.2.1 Cloning of TEX101 cDNA

A commercial *Pichia* Expression Kit (Invitrogen, Waltham, MA) was used for production of recombinant TEX101. Based on the published *TEX101* cDNA sequence (transcript variant 2, NM_001130011.1), a set of oligonucleotide primers (forward 5’-GAAGAAGGGGTATCTCTCGAGAAAGACTGTATTGTCAAAAGGGTCTGTCCAT-3’ and reverse 5’-TAGGGAATTCTTTAATGGTGATGGTGATGATGATTTTCAGTCTTTGAGGTTGA-3’) were designed for PCR amplification of the fragment coding for the mature form of TEX101 present in seminal plasma (aa 26-222). Primers facilitated the generation of compatible restriction ends for ligation into the pPIC9 vector, as well as the incorporation of a C-terminus polyhistidine tag for protein purification. *TEX101* human cDNA ORF Clone (RC225319; Origene, Rockville, MD) was used as a template. PCR was performed in a 20 μL reaction mixture, supplemented with 0.4 μL of cDNA (5 ng/μL final concentration), 4 μL of 5x Phusion GC Buffer, which contained 7.5 mM MgCl₂, and provided 1.5 mM MgCl₂ in the final reaction, 200 μM deoxynucleoside triphosphates, 250 nM of the primers and 0.4 U of Phusion High-Fidelity DNA polymerase (New England BioLabs, Ipswich, MA) on an Eppendorf Mastercycler thermal cycler. The PCR conditions were 98°C for 30 s, followed by 26 cycles of 98°C for 10 s, 68°C for 20 s, and 72°C for 20 s, with a final extension at 72°C for 7 min. In-frame cloning of the PCR product into the yeast expression vector pPIC9 was accomplished through double digestion, using *XhoI* and *EcoRI* restriction enzymes, and ligation of the two DNA fragments. The sequence of the construct was confirmed by DNA sequencing.

2.2.2 Production of human recombinant TEX101

Prior to transformation of yeast cells, pPIC9 vector containing the *TEX101* cDNA was linearized with *SacI* restriction enzyme to favor the integration of the construct in *P. pastoris* genome via homologous recombination. The linearized construct was introduced into the yeast strains GS115 and KM71 by electroporation. A stable clone was selected from the GS115 strain according to the manufacturer’s recommendations (Invitrogen). Stable yeast clones were grown in the buffered complex glycerol medium until the culture reached log-phase (OD₆₀₀=2-6).
Following that, the cell pellet was resuspended in the buffered complex methanol (BMMY) to an OD$_{600}$ of 1.0 and was grown at 30°C with shaking. TEX101 production was induced with 10 mL/L methanol over 4 days. Yeast culture containing secreted TEX101 was centrifuged and the supernatant was concentrated 100-fold initially by positive pressure ultrafiltration in an Amicon™ stirring chamber (Millipore, Billerica, MA) with a 10-kDa cutoff regenerated cellulose membrane (Millipore), followed by Amicon™ centrifugal filter tubes Ultracel 3K (Millipore). A rabbit polyclonal anti-TEX101 antibody HPA041915 (Sigma-Aldrich, St. Louis, MO) and a mouse monoclonal anti- His antibody (Cat# A00186-100, GenScript, Piscataway, NJ) were used to monitor TEX101 production by Western blot analysis.

2.2.3 Purification of human TEX101 with immobilized metal ion affinity chromatography

The recombinant TEX101 was purified from yeast culture supernatants by immobilized metal ion affinity chromatography. HIS-Select Nickel Affinity gel (Sigma-Aldrich) selective for recombinant proteins with histidine tags was used to purify TEX101 from yeast culture according to the manufacturer’s recommendations. In summary, the nickel affinity gel was first washed with 1-2 volumes of de-ionized water to remove ethanol, and then equilibrated with 3-5 volumes of equilibration buffer (10 mM imidazole in 50 mM NaH$_2$PO$_4$, 0.3 M NaCl, pH 8.0). Prior to application on the affinity gel, the recombinant protein sample was clarified by centrifugation to obtain a pH between 7.0 and 8.0. Recombinant protein solution was incubated with affinity gel, which was subsequently washed with equilibration buffer. TEX101 was eluted with 250 mM imidazole in 50 mM NaH$_2$PO$_4$, 0.3 M NaCl, pH 8.0 at room temperature. The presence of TEX101 in various fractions was determined with Western blotting by using rabbit polyclonal anti-TEX101 antibody HPA041915 (Sigma-Aldrich). The purity and the molecular mass of TEX101 were determined by SDS-PAGE stained with Coomassie Blue. The purified TEX101 protein concentration was determined by the bicinchoninic acid assay (Pierce Biotechnology, Rockford, IL).
2.2.4 Analysis of human recombinant TEX101 by mass spectrometry

Following SDS-PAGE analysis, all visible gel bands were excised and analyzed by LC-MS/MS. An in-gel digestion protocol was followed, as described elsewhere (181). In all cases, peptides were extracted from solution using C18 OMIX tips (Varian Inc., Lake Forest, CA) and eluted in 5 µL of elution buffer B (65% acetonitrile, 0.1% formic acid). Buffer A (80 µL of 0.1% formic acid) was added to sample tubes and transferred to a 96-well microplate (Axygen, Union City, CA). Using a 96-well microplate autosampler, 40 µL of each sample was loaded onto a 3 cm C18 trap column (inner diameter 150 µm; New Objective, Woburn, MA) that was packed in-house with 5 µm Pursuit C18 (Varian Inc.). An increasing concentration of Buffer B (0.1% formic acid in acetonitrile) was used to elute the peptides from the trap column onto a resolving analytical 5-cm PicoTip Emitter Column (inner diameter 75 µm, 8 µm tip; New Objective). This column was packed in-house using 3 µm Pursuit C18 (Varian). The EASY-nLC system (Proxeon Biosystems, Odense, Denmark) was coupled online to an LTQ-Orbitrap XL hybrid mass spectrometer (Thermo Fisher Scientific, San Jose, CA) and a nanoelectrospray ionization source (Proxeon) was used with a spray voltage of 2 kV and temperature of 160 °C. A data-dependent mode was utilized to analyze samples and a full MS1 scan was acquired from 450-1450 m/z in the mass analyzer (resolution of 60,000). This was followed by MS2 scan acquisition of the top six parent ions in the LTQ mass analyzer. The subsequent parameters were enabled: dynamic exclusion, charge state screening and monoisotopic precursor selection. Ions with charge states of +1, ≥ +4 and unassigned charge states did not undergo MS2 fragmentation.

For protein identification and data analysis, XCalibur software (v. 2.0.5; Thermo Fisher) was used to generate RAW files of each MS run. RAW files were subsequently used to generate Mascot Generic Files (MGF) on Mascot Daemon (version 2.2.2). Once generated, MGFs were searched with Mascot (Matrix Science, London, UK; version 2.2). Protein searches were performed against the non-redundant human UniProtKB/Swiss-Prot database (version 10, October 2013) using the following parameters: fully tryptic cleavages, 7 ppm precursor ion mass tolerance, 0.4 Da fragment ion mass tolerance, allowance of one missed cleavage and fixed modifications of carbamidomethylation of cysteines. Variable modifications included oxidation of methionine, pyro-Glu from glutamine of the N-terminus-carbamoylmethylcystein cyclization at N terminus, deamidation of glutamine, oxidation of tryptophan, and acetylation of the N-terminus.
2.2.5 Assessment of TEX101 glycosylation

The TEX101 protein glycosylation was assessed by treatment of purified recombinant TEX101 protein with the deglycosylation enzyme PNGase F (Roche, Mannheim, Germany). The mixture was incubated at 37°C for 3 hours. PNGase F treated and non-treated TEX101 were subjected to SDS-PAGE stained with Coomassie Blue.

2.2.6 Animal handling and somatic cell fusion for monoclonal antibody production

Female BALB/c mice were obtained from the Toronto Centre for Phenogenomics (TCP). All animal research (Animal Use Protocol# 14-04-0119a-H) was approved by TCP Animal Care Committee. Mice were inoculated subcutaneously with 100 μg of deglycosylated recombinant TEX101 protein, mixed (1:1) with Sigma Adjuvant System (Sigma-Aldrich). Two subsequent booster injections with 25 μg of antigen in adjuvant were performed at three-week intervals. Final boost was an intraperitoneal injection of 25 μg of antigen in phosphate-buffered saline (137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, 1.8 mM KH₂PO₄). Three days later, mouse spleen was excised aseptically and homogenized. Extracted spleen cells were fused with NSO murine myeloma cells (5:1 ratio) using polyethylene glycol (Sigma-Aldrich). Successfully fused cells were selected using HAT media (Invitrogen), supplemented with 20% fetal bovine serum (Hyclone, Thermo-Fisher Scientific, Waltman, MA).

2.2.7 Screening for IgG secreting clones

Cell culture supernatants were screened for the presence of IgG and IgM antibodies by using the following immunoassay protocol. 96-well microtiter plates were coated with goat anti-mouse IgG+IgM (H+L) antibody (Jackson ImmunoResearch Laboratories, Inc. – West Grove, PA) diluted (1:1,000) in sodium carbonate-bicarbonate buffer (0.2 M Na₂CO₃, 0.2 M NaHCO₃, pH 9.2). Plates were washed twice with PBST (0.05% Tween 20 in 137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, 1.8 mM KH₂PO₄, pH 7.4) and 100 μL of 5% milk in PBST were added per well. Plates were incubated for one hour at RT. After washing (3x), 100 μL of each hybridoma supernatant and the appropriate controls were added in two adjacent wells and incubated for an
hour at room temperature. Following one more round of washing (3x), 100 μL of HRP-conjugated Goat anti-mouse IgG (Fc Fraction; Jackson ImmunoResearch) and 100 μL of goat anti-mouse IgM (μ chain) (Jackson ImmunoResearch) antibodies (in 1% milk/PBST), were added on each one of the paired wells, respectively. Following a final wash (3x), 100 μL of 3,3,5,5'-tetramethylbenzidine substrate solution were added and plates were incubated for 15 min at 37 °C with gentle shaking. Fifty μL of stop solution (2M H₂SO₄) were added on top. Absorbance was measured with the Wallac EnVision 2103 Multilabel Reader (Perkin Elmer, Waltham, MA) at 450 nm, with a reference wavelength of 620 nm. IgG positive colonies were transferred in 48-well culture plates.

2.2.8 Screening for immunogen reacting clones

IgG positive clones were screened for reaction with the immunogen using an indirect immunoassay protocol. Recombinant TEX101 protein was immobilized on 96-well microtiter plates (100 ng per well) diluted in coating buffer. Plates were washed (2x) and 100 μL of 5% milk in PBST were added per well. Plates were incubated for one hour at RT, followed by wash (3x). 100 μL of hybridoma supernatant and the appropriate controls were added and incubated for an hour at RT. Then, plates were washed (3x) and 100 μL of HRP-conjugated goat anti-mouse IgG (Fc Fraction) (Jackson ImmunoResearch) antibody (in 1% milk in PBST) were added on the plates. Following final wash (3x), 100 μL of 3,3,5,5’-tetramethylbenzidine substrate solution were added and plates were incubated for 15 min at 37 °C with gentle shaking. Fifty μL of stop solution (2M H₂SO₄) were added on top. Absorbance was measured with the Wallac EnVision 2103 Multilabel Reader (Perkin Elmer) at 450 nm, with a reference wavelength of 620 nm.

2.2.9 Immunocapture-SRM screening for clones producing antibodies against native TEX101

White 96-well microtiter plates were coated with 500 ng/well of sheep anti-mouse IgG-Fcγ fragment-specific antibody (Jackson ImmunoResearch) in TBS buffer. Plates were washed twice with PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, 1.8 mM KH₂PO₄, pH 7.4). Cell culture supernatants at two dilutions in 1% (w/v) BSA in PBS buffer were applied to the plates
and incubated for 2 hours at RT with gentle shaking. A protein G–purified mouse polyclonal anti-TEX101 antibody (ab69522; Abcam, Cambridge, MA) was used as a positive control for the assay. Plates were washed (6x) with PBS and 100 μL of 100-fold diluted seminal plasma, from a normal donor, in 6% (w/v) BSA and PBS buffer were added. Following 2 hours of incubation at RT with gentle shaking, wells were once again washed 3x with PBS and 3x with 50mM ammonium bicarbonate. Fifty mM ammonium bicarbonate, 50 mM dithiothreitol (Sigma-Aldrich), 50 fmoles of heavy isotope-labeled TEX101 proteotypic peptide GALCQETILIIK tagged with a trypsin-cleavable tag (SpikeTides™_TQL, JPT Peptide Technologies GmbH, Germany) and 0.05% RapiGest SF (Waters) were mixed and ninety four μL of this mix were added to each well and kept for 15 minutes at RT. Then, 5 μL of 100 mM iodoacetamide were added and samples were kept for 40 minutes in the dark at RT. Samples were then digested by addition of 5 μL of 0.05 μg/μL of sequencing-grade modified porcine trypsin (Promega Cat# V5111, Madison, WI) in 50 mM ABC. Trypsin inactivation and cleavage of RapiGest SF was achieved with the addition of 1% trifluoroacetic acid. C18 microextraction and desalting of peptides was done as described above. EASY-nLC system (Proxeon) was coupled online to a Quantiva triple-quadrupole mass spectrometer (Thermo Fisher Scientific) using a nanoelectrospray ionization source. All transitions being monitored were scheduled within 1.5-min intervals during a 30-min LC gradient. Four unique TEX101 peptides were monitored with the scheduled SRM mode, with one used for quantification and the rest used for qualitative analysis. Relative abundance of TEX101 in each sample was estimated as a ratio to the spiked-in tagged heavy isotope-labeled peptide internal standard GALC[cm]QETILIIK*-JPTtag. The SRM method had the following parameters: optimized collision energy (CE) values; mass/charge ratio (m/z) scan width, 0.010; scan time, 0.015 to 0.040 s; FWHM resolution of the first quadrupole (Q1), 0.4; FWHM resolution of the third quadrupole (Q3), 0.7; pressure of the second quadrupole, 1.5 mtorr; tuned S-lens values; declustering voltage, +1 V. RAW files recorded for each sample were analyzed with the Pinpoint software, and peptide areas were used to calculate light-to-heavy peptide ratios and protein concentrations in each sample.
2.2.10 Clone screening by Western blot

We confirmed results of immunocapture-SRM screening by Western blot. Seminal plasma samples (10 μg per well) were prepared in 50 mM DTT and 31.25 mM Tris-HCl pH 6.8, 12.5% glycerol, 1% SDS, 0.005% Bromophenol Blue (Laemmli sample buffer; BioRad Cat# 161-0737, Hercules, CA) and applied to 4-15% Mini-PROTEAN® TGX™ Precast Gels Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (BioRad Cat# 456-1021). Gel electrophoresis was performed at 200 V for approximately 30 minutes. Gels were then stained with SimplyBlueTM SafeStain Coomassie® G-250 stain (Invitrogen). For Western blotting, gels were transferred using a Trans-Blot® Turbo Blotting System (BioRad) and a Trans-Blot® Turbo Transfer Pack (0.2 μm PVDF membrane, BioRad). After blocking in 5% milk in TBST (0.05% Tween 20 in 50 mM Tris, 150 mM NaCl, pH 7.5) for approximately 2 hours at 22°C, membranes were attached to a Mini-PROTEAN® II Multiscreen Apparatus (BioRad Cat# 456-1021). Hybridoma culture supernatants were added on each lane and membranes were incubated for 1.5 hours at room temperature. A rabbit polyclonal anti-TEX101 antibody HPA041915 (Sigma-Aldrich) was used as a positive control. After incubation, the membrane was washed (3x) in TBST, followed by the addition of alkaline phosphatase-conjugated AffiniPure goat anti-rabbit or anti-mouse IgG (0.03 μg/mL in 1% milk in TBST; Jackson ImmunoResearch) and incubation at 22°C for 45 minutes. The membrane was then extensively washed with TBST, dried and 125 μL of chemiluminescence substrate (Siemens, Los Angeles, CA) per square centimeter was added. The membrane was then placed into an autoradiography cassette and exposed and developed using Radiomat™ B Plus-Full Speed Blue sensitive x-ray film (8 x 10 inches, AGFA X-Ray Film, Mortsel, Belgium).

2.2.11 Expansion of clones and purification of anti-human TEX101 monoclonal antibodies

Subsequently, cells were further grown and transferred in serum-free medium (CD-1 medium; Invitrogen), containing 8 mM L-Glutamine. Supernatants were collected and purified using a protein G column, according to the manufacturer’s protocol (GammaBind Plus, GE Healthcare). Culture supernatants were diluted two-fold with the binding buffer (10 mM Na2HPO4/NaH2PO4, 150 mM NaCl, 10 mM EDTA, pH 7.0) and loaded on the column. The column was then washed with binding buffer and antibodies were eluted with 0.5 M acetic acid at pH 3.0.
2.2.12 Development of a sandwich immunoassay

White 96-well ELISA plates were coated with 500 ng/well of mouse monoclonal anti-TEX101 antibody 23-ED-616 in 50 mM Tris buffer, pH 7.8. Plates were washed (2x) with 0.05% Tween 20 in 20 mM Tris, 150 mM NaCl, pH 7.4). Treated TEX101 calibrators and samples (see below for treatment) were added into each well (100 μL/well) and incubated for 2 hours with gentle shaking. Plates were then washed (2x) with the washing buffer. A biotinylated mouse monoclonal anti-TEX101 antibody 23-ED-155, diluted in a solution containing 60 g/L BSA, 25 mL/L normal mouse serum, 100 mL/L normal goat serum, and 10 g/L bovine IgG in 50 mM Tris, pH 7.8, were added (250 ng of antibody per 100 μL of solution per well) and incubated for 1 hour. Plates were washed (6x) and alkaline phosphatase-conjugated streptavidin was added in the wells (100 μL per well). Incubation was for 20 minutes at RT with gentle shaking, followed by a final wash (6x). Diflunisal phosphate (DFP) solution was prepared in substrate buffer (0.1 M NaCl, 1 mM MgCl₂ in 0.1 M Tris, pH 9.1), added on the plate (100 μL per well) and incubated for 10 minutes at RT with gentle shaking. Subsequently, the developing solution (1 M Tris, 0.4 M NaOH, 2 mM TbCl₃ and 3 mM EDTA) was added on top and mixed for 1 minute. Time-resolved fluorescence was measured with the Wallac EnVision 2103 Multilabel Reader (Perkin Elmer), as previously described (182).

2.2.13 Measurement of TEX101 in seminal plasma and blood serum

Matched pre- and post-vasectomy seminal fluid samples (N=9), as well as male (N=17) and female (N=17) serum samples, were obtained after informed consent and institutional review board approval (Mount Sinai Hospital, Toronto, ON, Canada). Seminal fluid was allowed to liquefy at room temperature for 1 h after collection, aliquoted in 1-mL portions and centrifuged (3x) at 13,000 g for 15 minutes at room temperature to separate plasma from cells and cellular debris. The supernatant seminal plasma was then frozen at 80°C until use.

Several seminal plasma samples were pooled and mixed to prepare the calibrators. TEX101 concentration in the pool was calculated using a quantitative SRM method, similar to the aforementioned one. Heavy isotope-labeled peptide GALC[cm]QETILIIE with a trypsin-cleavable JPT tag (JPT Peptide Technologies GmbH) was used as the internal standard for the absolute quantification of endogenous TEX101 protein. Calibration curve was prepared by spiking
increasing amounts of the internal standard (0.1 to 3,000 fmoles) into 10 μL of 10-fold diluted seminal plasma pool (1 μL equivalent), before proteomic sample preparation and trypsin digestion. In parallel, different volumes of the seminal plasma pool (10-fold diluted 0.6, 1, 2 and 6 μL of seminal plasma) were supplemented with 600 fmoles of the internal standard and digested with three full process replicates and measured by SRM with two injections each. Dilution-adjusted ratios were used to calculate TEX101 concentrations in the seminal plasma pool. Furthermore, numerous aliquots (volume of 20 μL each) of the pool were prepared and stored at -20°C.

Calibrators for the seminal plasma assay were prepared by mixing (1:1) one aliquot of SP with reagent mix (16 μL 7.7M guanidine-HCl, 2 μL 2M NaOH and 2 μL dH₂O). The mix was incubated for one hour at RT, and then diluted (50x) with assay diluent (60 g/L BSA, 25 mL/L normal mouse serum, 100 mL/L normal goat serum, and 10 g/L bovine IgG in 50 mM Tris, pH 7.8). Serial dilutions of the treated calibrator sample were prepared in a two-fold dilution step (ranging from ~ 47ng/mL to 0 ng/mL) and added on ELISA plate. Seminal plasma samples from individuals were mixed (1:1) with the reagent mix, followed by one-hour incubation at room temperature. Samples were further diluted 10-fold with assay diluent before loading on the plate.

Calibrators for the serum assay were prepared by diluting SP in a female serum pool (1:10) and 60 μL of this mixture was supplemented with 10% sodium deoxycholate [5% final]. The mixture was incubated for 1 hour at 63°C and further diluted (5x) with assay diluent (60 g/L BSA, 25 mL/L normal mouse serum, 100 mL/L normal goat serum, and 10 g/L bovine IgG in 50 mM Tris, pH 7.8). Once again, serial dilutions of the treated calibrator sample were prepared in a two-fold dilution step (ranging from ~ 47ng/mL to 0 ng/mL) and added on ELISA plate. Serum samples from individuals were mixed (1:1) with 10% sodium deoxycholate, followed by one-hour incubation at 63°C. Samples were further diluted 3- to 6-fold with assay diluent before loading on the plate.

2.2.14 Statistical analysis

To assess assay’s linearity, a linear regression model was built using the log-transformed values of endogenous TEX101 concentration and sample dilutions, within CV≤15%. Statistical
analysis and plots were prepared using R statistical software v 2.15.2 (available from www.Rproject.org).
2.3 Results

2.3.1 Production, purification and analysis of recombinant human TEX101 protein

The cDNA coding for mature TEX101 protein (amino acids 26 to 222) was cloned into a methylotrophic *P. pastoris* yeast expression system and secreted TEX101 was purified by nickel affinity chromatography from the yeast culture supernatant. Yeast mainly expressed two forms of TEX101 protein (~30 kDa and ~90 kDa), as assessed by SDS-PAGE and mass spectrometry (Appendix 2.1a). Higher molecular weight forms were produced due to the variable glycosylation of recombinant proteins in *P. pastoris* (183). Treatment of recombinant TEX101 with PNGase F reduced their molecular weight to the expected 27–35 kDa (Appendix 2.1b and c).

2.3.2 Production and screening of monoclonal antibodies

Deglycosylated recombinant TEX101 protein was used as an immunogen to generate mouse monoclonal antibodies. The fusion of murine splenocytes with murine myeloma cells resulted in the generation of 167 IgG-secreting hybridoma colonies, with 60 antibodies reacting with recombinant TEX101. These 60 antibodies were further screened for reaction with the native TEX101 protein in seminal plasma. Immunocapture-SRM revealed that 18 colonies produced antibodies which could bind to the native form of TEX101 present in seminal plasma (Figure 2.2a). Western blot, however, showed positivity of 26 colonies for denatured seminal plasma TEX101 (Figure 2.2b). Seventeen of the 18 aforementioned colonies were also identified by Western blot, while one colony was identified exclusively by the immunocapture-SRM. Thus, 18 positive hybridoma colonies were eventually expanded in the serum-free media and purified using protein G columns.
Figure 2.2 Screening of hybridoma colonies by immunocapture-SRM assay and Western Blot. (a) Relative abundance of the native form of TEX101 immunocaptured from seminal plasma using 60 IgG-secreting hybridoma colonies, as measured by a TEX101 SRM assay. Lanes (-) and (+) denote the negative control (no mouse anti-TEX101 antibodies) and the positive control (anti-TEX101 mouse polyclonal antibody ab69522), respectively. Asterisks mark the clones that were also positive on Western blot. (b) Representative Western blot of colonies (lanes 1-8 in duplicates) screened for reaction with TEX101 in seminal plasma. Lane (+) represents the positive control (rabbit polyclonal antibody HPA041915).
2.3.3 Immunoassay development

Two mouse monoclonal antibodies (23ED616 and 23ED155) targeting different TEX101 epitopes were used to develop an immunofluorometric assay. Even though both antibodies could capture the native form of TEX101 in seminal plasma, we thoroughly investigated a wide variety of conditions which would further increase ELISA signal and thus allow for measuring TEX101 concentrations in the low pg/mL range. Incubation of seminal plasma with guanidine hydrochloride at pH 12 or treatment with sodium deoxycholate at 63°C emerged as the most efficient procedures. We suspect that only a fraction of endogenous TEX101 is present in seminal plasma in the soluble form, while another fraction is either bound to other proteins, embedded into membranous vesicles such as epididymosomes or encapsulated into membrane fragments of destroyed spermatozoa. For instance, it has been previously reported that the majority of plasma membranes of the disrupted spermatozoa heads are released in the form of unilamellar vesicles (184). These vesicles are typically pelleted by centrifugation at 285,000 g (184) and thus most probably are present in our seminal plasma samples prepared by centrifugation of semen at 13,000 g. Opening of such membrane vesicles and dissociation of intravesicular proteins requires high pH treatment, such as incubation with 100 mM Na₂CO₃ at pH 11.6 (185). Thus, treatment of seminal plasma with guanidinium at pH 12 or deoxycholate at 63°C may lead to the release of TEX101 from such vesicles, making it available for ELISA measurements.

Assay calibrators were prepared by mixing seminal plasma samples obtained from several dozen healthy fertile men. The concentration of TEX101 protein in the pool was assessed by a quantitative SRM assay. The endogenous proteotypic peptide and spike-in internal standard were measured in 4 different volumes of seminal plasma, with 3 full process replicates for each volume and 2 injections for each replicate. Assuming that one mole of the proteotypic peptide represents one mole of protein, the mean TEX101 protein concentration was derived as 4.7±1.5 µg/mL.

2.3.4 ELISA performance and TEX101 measurement in seminal plasma

Initially, the assay was developed for TEX101 measurements in seminal plasma, the fluid of choice for diagnosis of male infertility. Two monoclonal antibodies (23ED616 and 23ED155)
were paired in the sandwich immunoassay. The endogenous TEX101 protein present in pooled seminal plasma was used to calibrate the assay. Since LOD of the immunoassay without seminal plasma pre-treatment was not high enough, we examined literature (186-196) to identify conditions which would improve the assay sensitivity in seminal plasma. Thus, seminal plasma was subjected to pretreatment with the following detergents, salts and solvents: 1% SDS, 2-5% Triton X-100, 1% CHAPS, 2% sodium deoxycholate, 1 M urea, 3 M guanidine chloride, 10% dimethyl sulfoxide, 10% dimethylformamide, 10% glycerol, 10% methanol, 10% acetonitrile. Additional treatments included heating between 40-70 °C and the use of high pH (10-12). As a result, the use of guanidinium at pH 12 and room temperature allowed a significant increase of ELISA signal and pg/mL sensitivity of TEX101 in seminal plasma (Appendix 2.2a).

The limit of blank (LOB) of the optimized immunoassay, measured with 2% (w/v) BSA (N=7) was established as 6.5 pg/mL. The corresponding within-run limit of detection (LOB + 1.64*SD) was determined as 20.1 pg/mL, with a linear range spanning from 33 pg/mL to 19.4 ng/mL (regression coefficient β₁=0.900, P<0.0001) (Appendix 2.2b). Aliquots of seminal plasma were measured over four days in order to calculate within-run (N=7 aliquots, each in 7 replicates) and total (N=4 aliquots, each in 2 replicates) imprecision. According to results, within-run and total imprecisions were generally low for the broad range of TEX101 concentrations (<5% and 11%, respectively), only getting worse when reaching the LOD. Within-run limit of quantification (LOQ) was set at the TEX101 concentration showing CV≤20% – which in this case was close to the assay’s LOD (<30.0 pg/mL). Stability of TEX101 was assessed by daily measurements of TEX101 concentrations in seminal plasma stored at 4°C and 22°C for seven days. As a result, concentration of TEX101 in the untreated seminal plasma changed marginally. When measured in the treated seminal plasma stored at 4°C, TEX101 concentration slightly increased after 24 hours from 50.25±1.91 to 67.84±0.33 ng/mL and then remained stable.

TEX101 was measured in 9 pairs of pre- and post-vasectomy seminal plasma samples. TEX101 concentrations in pre-vasectomy samples ranged from 37.1±0.05 μg/mL to 0.82±0.01 μg/mL, while TEX101 in all post-vasectomy samples was below the lowest calibrator (0.047 ng/mL) loaded on plate.
2.3.5 ELISA performance and TEX101 measurement in blood serum

To investigate TEX101 presence in the systemic circulation, we modified our protocol to allow for TEX101 measurements in male and female serum samples. TEX101 calibrators were prepared by diluting aliquots of seminal plasma in female serum. Treatment of serum with 10% sodium deoxycholate (5% final) at 63°C was found to be more effective than treatment with guanidine chloride at pH 12. A representative calibration curve for the TEX101 assay in female serum is shown in Appendix 2.2c. Within-run LOB of the optimized immunoassay measured with female serum samples (N=7) was established as 20.7 pg/mL. The corresponding within-run LOD (LOB + 1.64*SD) was determined as 40.0 pg/mL, with a linear range spanning from 45.7 pg/mL to 28.9 ng/mL (regression coefficient \( \beta_1 = 1.034, P<0.0001 \)) (Appendix 2.2d). Within-run (N=4 aliquots, each in 7 replicates) and total (N=4 aliquots, each in 2 replicates) imprecisions assessed over 4 days were found low for the broad range of TEX101 concentrations (<6% and 13%, respectively), only deteriorating when reaching assay's LOD. Within-run limit of quantification (LOQ) was set at the LOD (40.0 pg/mL).

Recovery of endogenous TEX101 from serum was assessed by spiking seminal plasma pool (of TEX101 concentration: 4.7±1.5 µg/mL) into female serum samples. According to results, endogenous TEX101 recovery was calculated as 90%, by comparing the measured value with the theoretical one.

Measurement of serum samples from 17 healthy females and 17 healthy males revealed undetectable levels of TEX101 in the systemic circulation (< 40 pg/mL). The apparent absence of TEX101 in serum can be explained by its high tissue specificity and the stringency of testis-blood barriers in healthy males.
2.4 Discussion

Despite rapid progress in proteomics and mass spectrometry, monoclonal antibodies remain indispensable analytical tools for the ultrasensitive analysis of proteins in biological fluids. Likewise, antibody-based ELISAs are still the assays of choice for clinical laboratory diagnostics and large-scale biomarker validation studies due to their simplicity, high sensitivity and specificity, high throughput and low cost (197). One of the most important aspects of ELISA development is the availability of antibodies which bind the native form of proteins.

We developed an immunoassay for TEX101 protein, due to the importance of this analyte for diagnosis of male infertility (180). Our initial efforts to produce an immunoassay using six different commercially available antibodies were not successful. Immunohistochemistry, prototype sandwich assays, Western blots and immuno-SRM assays revealed that commercial antibodies recognized only the denatured form of TEX101, but they could not bind the native form present in seminal plasma. Only a mouse polyclonal antibody (ab69522 from Abcam) could capture the native form of TEX101, as revealed by immuno-SRM (180). Due to its low purity and high cost, ab69522 was suitable as a positive control for colony screening, but not for ELISA development. Thus, we proceeded with production of new TEX101 monoclonal antibodies.

To obtain native immunogen, we first attempted to purify TEX101 from seminal plasma. Multi-step chromatography included anion exchange, size exclusion and reverse phase separations of seminal plasma from fertile men, with TEX101 being monitored by shotgun and SRM mass spectrometry. Multi-step separations enriched TEX101 from the initial abundance of ~0.01% (5 µg/mL) to ~5% of total protein. Such purity, however, was not sufficient for successful mouse immunization. We also attempted to produce TEX101 in E. coli; however, expression levels were low and most of TEX101 was lost due to protein aggregation. Expression in the P. pastoris yeast system was more successful, and the recombinant TEX101 was used for mouse immunization. Due to TEX101 hyperglycosylation in P. pastoris, however, the produced antibodies recognized only recombinant TEX101, but not the endogenous TEX101 in seminal plasma, as confirmed by prototype sandwich ELISA and immuno-SRM assays. The deglycosylated form of recombinant TEX101 elicited a strong immune response and led to dozens of hybridoma colonies. Screening by immunocapture-SRM assay revealed 18 colonies which produced antibodies against the native form of TEX101. In order to further improve the sensitivity of the developed sandwich immunoassay, we tested a variety of sample treatment
reagents, with 3 M guanidine chloride at pH 12 and room temperature or 5% sodium deoxycholate at 63°C being particularly effective for seminal plasma. Assay calibration was accomplished with the endogenous TEX101 protein present in a pool of seminal plasma samples (concentration of TEX101: 4.7±1.5 µg/mL). We also tried using the deglycosylated recombinant TEX101 diluted in PBS, whose performance matched the endogenous following treatment with guanidine at pH 12 (Appendix 2.3). Recombinant TEX101 was also recovered from seminal plasma samples (min. 112% and max. 191%) and female serum samples (min. 102% and max. 243%), with results indicating potential interference within both treated matrices; unlike the endogenous TEX101 protein.

Until recently, screening of antibody clones was accomplished either by Western blot, under denaturing conditions, or by solid-phase ELISA. Mass spectrometry- and proteomics-based approaches can revolutionize antibody production through the rapid screening of hybridoma clones (106). Immunocapture-mass spectrometry detection was previously demonstrated by Schoenherr et al., who proposed a SISCAPA-based pipeline for the screening of hybridoma supernatants for anti-peptide antibodies (107). To our knowledge, the use of two-step immunocapture followed by SRM detection to select hybridoma clones recognizing the native form of proteins present in biological fluids has not been previously demonstrated.

Two-step purification using sheep anti-mouse antibodies is a crucial step of our approach. Since dozens of IgG-producing colonies have to be tested, rapid screening is feasible only if colonies are grown in the FBS-supplemented media using a small-scale cell culture. To screen supernatants produced by each colony, mouse antibodies against TEX101 are first purified with sheep anti-mouse antibodies and then incubated with seminal plasma containing the native form of TEX101. Such two-step immunocapture, followed by SRM quantification, allows identification of hybridoma colonies which secrete antibodies against the native form of TEX101.

Our approach can be extended to develop high-quality antibodies against a variety of proteins with unknown function (198). For example, many testis- and prostate-specific proteins have not as yet been characterized and some of them have not been previously identified, due to their sequestration from the systemic circulation by blood-tissue barriers. Even though the Human Protein Atlas initiative (199) produced polyclonal antibodies against denatured forms of many testis- and prostate-specific proteins, monoclonal antibodies against the native protein forms of such proteins and corresponding immunoassays are still not available (105).
Coupling high-affinity antibodies with mass spectrometry-based quantitative analysis could complement the traditional immunoassays in the biomarker verification process (200-202). There are numerous studies on immunocapture-MS/MS or SRM, using either anti-peptide antibodies in complex digests (203, 204) or antibodies that capture the intact protein (138, 205), with quantification reaching low ng/mL. According to these studies, the advantages of immunocapture MS/MS or SRM over ELISA are associated with their high specificity and its great potential for multiplex analyses, thus proving to be a powerful alternative to immunoassays. We believe that 0.5 ng/mL or even lower LOD for TEX101 can be achieved by immunoaffinity-SRM assay with a sample volume as large as 1,000 µL of seminal plasma (reasonable volume from the clinical perspective). In future, if executed with a high throughput (>100 samples per day), a multiplex immunoaffinity-SRM assay for TEX101 and ECM1 markers may eventually replace ELISA assays for differential diagnosis of azoospermia.

To conclude, we here propose that immunocapture-SRM screening can facilitate the development of monoclonal antibodies and immunoassays against the native forms of challenging protein targets, such as membrane-bound testis-specific proteins. With more rigorous clinical validation, the developed TEX101 ELISA may emerge as a novel non-invasive clinical laboratory test for the differential diagnosis of male infertility. Native protein capturing antibodies will allow mapping of the human TEX101 interactome and reveal its functional role in reproduction and male infertility.
Chapter 3
Preclinical evaluation of TEX101 protein ELISA for the differential diagnosis of male infertility

This chapter includes data published in *BMC Medicine*: “Preclinical evaluation of TEX101 protein ELISA for the differential diagnosis of male infertility” by Korbakis D.†, Schiza C.†, Brinc D., Soosaipillai A., Légaré C., Sullivan R., Mullen B., Jarvi K., Diamandis E.P., and Drabovich A.P. *BMC Medicine* 2017, 15:60. † authors contributed equally

A link to the published paper can be found at

doi: 10.1186/s12916-017-0817-5

Copyright permission has been granted.
Chapter 3

3 Preclinical evaluation of TEX101 protein ELISA for the differential diagnosis of male infertility

3.1 Introduction

Infertility is a common medical condition with an estimated prevalence of nearly 15% in the general population (206). The disorder affects both men and women, while the male factor, exclusive or combined with female abnormalities, contributes to approximately 50% of all cases. The clinical categories of male infertility range from lowered production of sperm, or oligospermia, to severe cases of azoospermia with non-measurable levels of sperm in semen (207). Azoospermia is diagnosed in nearly 2% of the general population and has two major forms, non-obstructive (NOA) and obstructive azoospermia (OA). Diagnostic testicular biopsy remains a standard tool for differential diagnosis of azoospermia, by evaluating the rate of spermatogenesis and the presence of sperm in the testis (62). However, it is an invasive surgical procedure with potential complications.

In quest of alternative, non-invasive approaches for differential diagnosis of male infertility, seminal plasma was proposed as a clinical sample enriched with testis-derived proteins, mRNA and metabolites (38, 208, 209). SP contains up to 3,200 proteins secreted by testis, epididymis, prostate, seminal vesicles and Cowper’s glands (95, 99). Numerous proteins found in SP are directly involved in the production and maturation of sperm or interaction with zona pellucida and fusion with oocytes (96). Since testis-specific biomarkers are not found in blood serum due to stringent blood-testis and blood-epididymis barriers, semen and SP remain the only viable fluids for the non-invasive diagnosis of male infertility (23, 30).

Previously, the clinical utility of TEX101 was proposed as a biomarker based on its measurements in SP by mass spectrometry (97). To translate TEX101 biomarker into the clinic, we developed a simple ELISA test (210). In this work, we focused on the full characterization of TEX101 as an analyte in SP, pre-clinical evaluation of the performance of TEX101 ELISA in a large cohort of fertile, subfertile and infertile men and validation of TEX101 as a prognostic biomarker of male infertility and a predictive biomarker of sperm retrieval in NOA patients. Our objective was also to evaluate TEX101 ELISA as a test for the differential diagnosis of the most common clinical conditions of male infertility (unexplained infertility, oligospermia and
azoospermia) and propose simple decision trees for use in the clinic. We hypothesized that TEX101 levels in SP would vary in patients with different categories of male infertility. We would suggest that TEX101 test is to be offered after the standard protocols for the initial evaluation of infertility (semen analysis and measurement of motility, morphology and levels of reproductive hormones) in patients admitted to the urology clinics, but prior to diagnostic testicular biopsies.
3.2 Materials and Methods

3.2.1 Experimental design and statistical rationale

Based on our previous measurements of TEX101 by ELISA (210), the minimal sample size required to validate TEX101 performance in pre- and post-vasectomy groups was 18 samples (two-tailed Mann Whitney U test, \(\alpha=0.05\), 80% power, allocation ratio of 1). Even though the minimal sample size was small, we decided to measure TEX101 in all SP samples available in our biobank, in order to establish more accurate clinical cut-off values.

3.2.2 Patients with inclusion and exclusion criteria

Healthy fertile men pre- and post-vasectomy, and men referred to Mount Sinai Hospital for clinical infertility evaluation were included in the study. Initial patient evaluation included computer-assisted semen analysis and measurement of reproductive hormones (testosterone, estradiol, follicle-stimulating hormone, luteinizing hormone and prolactin). Sperm concentration, ejaculate volume, motility, and morphology were graded based on World Health Organization 5th edition criteria (53). Azoospermia was defined as no sperm found on initial semen analysis, and oligospermia included men with spermatozoa present at concentrations <15 million/mL. Unexplained infertility was defined as inability to conceive after one year of regular unprotected intercourse, spermatozoa concentration >15 million/mL and normal hormonal parameters. Note that we do not have female factor data for this group. There are a number of female infertility factors, such as tubal obstruction, endometriosis, polycystic ovary syndrome, among many others, that make female factor infertility a determinant or co-determinant. Thus, some patients in this group may be in fact healthy fertile men. Clinical reference standards for vasectomy were sperm counting, while clinical reference standards for OA, NOA and sperm retrieval were sperm counting, diagnostic testicular biopsies and TESE. Clinical cut-off values were reported based on the randomized blind measurements of TEX101 in retrospectively collected 805 SP samples using ELISA with the DOC-based protocol. The analysis did not include 45 patients which SP samples were fully consumed in the preliminary experiments.
3.2.3 Seminal plasma samples

Semen samples were obtained with the informed consent from patients by masturbation with 2–5 days of abstinence before collection. Eight hundred and fifty (N=850) semen samples were obtained from healthy fertile men and patients diagnosed with unexplained infertility, oligospermia and azoospermia (Table 3.1). Sample collection was approved by the institutional review boards of Mount Sinai Hospital (approval #08-117-E) and University Health Network (#09-0830-AE). Samples were collected as a convenience series and analysed retrospectively. The time difference between initial sperm count measurements in semen and TEX101 measurements in SP varied from several months to up to five years. Following collection, semen was left to liquefy at RT for 1 hour, then aliquoted and centrifuged three times at 13,000 g for 15 min. SP was separated from cells and cellular debris and stored at -80°C.

Table 3.1 Clinicopathological variables of the patient cohort

<table>
<thead>
<tr>
<th>Clinical Parameters</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Number of subjects</strong></td>
<td>850</td>
<td>100</td>
</tr>
<tr>
<td><strong>Age – Median (Range), years</strong></td>
<td>41</td>
<td>(19-63)</td>
</tr>
<tr>
<td><strong>Ethnic Group</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>African-Canadian</td>
<td>28</td>
<td>3.3</td>
</tr>
<tr>
<td>Asian</td>
<td>68</td>
<td>8.0</td>
</tr>
<tr>
<td>Caucasian</td>
<td>271</td>
<td>31.9</td>
</tr>
<tr>
<td>Hispanic</td>
<td>11</td>
<td>1.3</td>
</tr>
<tr>
<td>Indo-Canadian</td>
<td>11</td>
<td>1.3</td>
</tr>
<tr>
<td>Middle Eastern</td>
<td>27</td>
<td>3.2</td>
</tr>
<tr>
<td>Native-Canadian</td>
<td>11</td>
<td>1.3</td>
</tr>
<tr>
<td>Unspecified/Unavailable</td>
<td>423</td>
<td>49.8</td>
</tr>
<tr>
<td><strong>Diagnosis (sperm count in mln/mL)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-vasectomy</td>
<td>67 (&gt; 15)</td>
<td>7.9</td>
</tr>
<tr>
<td>Post-vasectomy</td>
<td>63 (0)</td>
<td>7.4</td>
</tr>
<tr>
<td>Unexplained infertility</td>
<td>283 (15-360)</td>
<td>33.3</td>
</tr>
<tr>
<td>Oligospermia</td>
<td>276 (0.01-15)</td>
<td>32.5</td>
</tr>
<tr>
<td>Azoospermia</td>
<td>20 (0-0.5)</td>
<td>2.3</td>
</tr>
<tr>
<td>Non-obstructive azoospermia (NOA)</td>
<td>100 (0-0.24)</td>
<td>11.8</td>
</tr>
<tr>
<td>Obstructive azoospermia (OA)</td>
<td>39 (0-0.18)</td>
<td>4.6</td>
</tr>
<tr>
<td>Unknown diagnosis</td>
<td>2</td>
<td>0.2</td>
</tr>
</tbody>
</table>
3.2.4 TEX101 ELISA measurements

The 96-well ELISA plates were coated with 500 ng/well of mouse monoclonal anti-TEX101 antibody 23ED616.8 in 50 mM Tris-HCl buffer at pH 7.8. Plates were washed twice with the washing buffer (0.05% Tween 20 in 20 mM Tris-HCl and 150 mM NaCl at pH 7.4). Assay calibrators were prepared as previously described (210). Briefly, several dozen SP samples from fertile pre-vasectomy men were pooled, and endogenous TEX101 concentration (4.7 ± 1.5 µg/mL) was measured by SRM. Multiple 20 µL aliquots of the pool were stored at -20°C. For ELISA measurements, calibrators and patient samples were thawed and mixed (1:1) with either Reagent mixture #1 (6 M GndCl at pH 12; 1 hour incubation at RT) or Reagent mixture #2 (4% DOC in water; incubation for 1 hour at 63°C). Following treatment, calibration samples were diluted 100-fold with the assay diluent (60 g/L BSA, 25 mL/L normal mouse serum, 100 mL/L normal goat serum and 10 g/L bovine IgG in 50 mM Tris-HCl at pH 7.8). Subsequently, serial dilutions of the treated calibrator (0.5 to 50 ng/mL, 100 µL/well) were prepared with 4-fold dilution steps. Similarly, patient SP samples (20 µL) were treated with Reagent mixture #1 or 2 (1:1), further diluted 10-, 100-, 1000- and 10000-fold with the assay diluent and added on ELISA plates (100 µL/well). Following 2 hour incubation with gentle shaking, plates were washed twice with the washing buffer. A biotinylated mouse monoclonal anti-TEX101 antibody 23ED660.7 in the assay diluent (250 ng in 100 µL per well) was added and incubated for 1 hour. Plates were then washed six times, and streptavidin-conjugated alkaline phosphatase was added for 15 min with gentle shaking. After the final wash (6x), diflunisal phosphate solution was prepared in the substrate buffer (0.1 M NaCl, 1 mM MgCl₂ in 0.1 M Tris at pH 9.1), added to the plate (100 µL per well) and incubated for 10 min at RT with gentle shaking. Finally, the developing solution (1 M Tris-HCl, 0.4 M NaOH, 2 mM TbCl₃ and 3 mM EDTA) was added and mixed for 1 min. Time-resolved fluorescence was measured with the Wallac EnVision 2103 Multilabel Reader (Perkin Elmer), as previously described (182).

3.2.5 Measurement of TEX101 isoforms in spermatozoa and seminal plasma by SRM

SP and spermatozoa of patients with the normal sperm count (median 26 mln/mL; N=17) were prepared. Spermatozoa were lysed by RapiGest SF. Total protein was measured by BCA assay,
and 5 and 10 µg of total protein of spermatozoa lysate and SP, respectively, were mixed with 50 mM ammonium bicarbonate. SRM assays were developed using previously described selectivity and reproducibility criteria (211, 212). Two hundred femtomoles of the heavy isotope-labeled TEX101 peptide AGTETAILATK (present in both membrane/secreted isoform Q9BY14-1 and intracellular isoform Q9BY14-2) and 500 femtomoles of heavy isotope-labeled peptide QIQTSSQTSPEAMGTPR (present exclusively in the intracellular isoform Q9BY14-2) were spiked before trypsin digestion. Heavy isotope-labeled peptides (SpikeTides™ TQL, JPT Peptide Technologies GmbH, Berlin, Germany) included trypsin-cleavable quantifying JPT tags (serine-alanine-[3-nitro]tyrosine-glycine). Five millimoles dithiothreitol (DTT) (Sigma-Aldrich, St. Louis, MO) and 0.05% RapiGest SF surfactant (Waters Corp., Milford, MA) were added, and samples were incubated at 60°C for 30 min. Following that, samples were alkylated in the dark with 10 mM iodoacetamide (Sigma-Aldrich) for 40 min and digested overnight at 37°C with proteomics-grade porcine trypsin (Sigma-Aldrich, #T6567). Trypsin inactivation and cleavage of RapiGest SF was achieved with the addition of trifluoroacetic acid (1% final). Additionally, 5 mM of L-methionine (Sigma-Aldrich) were added to each digest, in order to limit the oxidation of methionine residues during sample preparation or storage. Finally, digests were loaded on C18 OMIX tips (Varian Inc., Lake Forest, CA), and bound peptides were eluted in 3 µL of 65% acetonitrile in water with 0.1% formic acid. Water with 0.1% formic acid was added (60 µL final volume), and samples were transferred to the 96-well microplate (Axygen, Union City, CA). Using a 96-well microplate autosampler, 18 µL of each sample were loaded onto a 3 cm trap column (inner diameter 150 µm; New Objective, Woburn, MA) packed in-house with 5 µm Pursuit C18 (Varian Inc.). An increasing concentration of Buffer B (0.1% formic acid in acetonitrile) was used to elute the peptides from the trap column onto a resolving analytical 5-cm PicoTip emitter column (inner diameter 75 µm, 8 µm tip; New Objective) packed in-house with 3 µm Pursuit C18 (Varian). The EASY-nLC system (Proxeon Biosystems) was coupled online to TSQ Quantiva™ triple quadrupole mass spectrometer (Thermo Scientific, San Jose, CA) with a nanoelectrospray ionization source. The SRM parameters were the following: positive polarity, declustering and entrance potentials 150 and 10 V, respectively; ion transfer tube temperature 300°C; optimized collision energy values; scan time 40 ms; 0.4 and 0.7 Da full width at half maximum (FWHM) resolution settings for the first and third quadrupoles, respectively; and 1.5 mTorr argon pressure in the second quadrupole. TEX101 peptides were monitored in a non-scheduled SRM mode during a 30-min LC gradient. The relative abundance of TEX101 in each sample was estimated as a ratio of the endogenous peptide to the spiked
heavy isotope-labeled standards. Raw files for each sample were recorded and analyzed with Skyline software (v3.1.0.7382, MacCoss Lab Software, Seattle, WA, USA).

3.2.6 Immunohistochemistry of TEX101 in testicular tissues

In-house generated mouse monoclonal anti-TEX101 antibodies 23ED660, 23ED11 and 23ED228 were used to stain testicular tissue samples fixed with 10% buffered formalin. Samples were incubated with the antibody solutions for 1 hour at RT. Multiple dilutions (400-to 5,000-fold) were tested. Heat-induced epitope retrieval was performed in citrate buffer at pH 6.0. Vectastain Elite ABC Kit (Vector Laboratories Inc. Burlingame, CA 94010), 3,3'-diaminobenzidine substrate (Sigma-Aldrich) and LabVision 720 autostainer (Thermo Fisher Scientific Inc.) were used for detection.

3.2.7 Size-exclusion chromatography

SP samples from pre-vasectomy men were centrifuged at 4,000 g for 20 min and pooled (total protein 39 mg/mL). Vesicle-free pool (total protein 22 mg/mL) was obtained by ultracentrifugation at 120,000 g. Five hundred micrograms of total protein from both pools were diluted to 500 µl with the running buffer (0.1 M NaH2PO4-Na2HPO4 and 0.15 M NaCl, pH 7.0) and loaded on a TSKgel G3000SW size exclusion column (Tosoh Bioscience LLC, King of Prussia, PA). Pools were run at 1 mL/min for 35 min, and fractions were collected every 0.5 min from 8 to 27 min. Presence of TEX101 in each fraction was measured by ELISA with DOC-based treatment.

3.2.8 Isolation of SMVs from seminal plasma samples

Pre- and post-vasectomy SP samples, as well as four samples obtained from the group of infertile individuals with moderate-to-high sperm count, were individually pooled. SMVs were isolated with a method adapted from Fabiani et al (213). Semen samples were centrifuged at 4,000 g for 20 min at 4°C to remove cells and cell debris. The remaining supernatants were
diluted 1:2 in a solution containing 30 mM Tris and 130 mM NaCl at pH 7.5 and centrifuged one more time at 4,000 g for 20 min at 4°C. Following that, supernatants were ultracentrifuged at 120,000 g for 2 hours at 4°C, pellets were re-suspended in 30 mM Tris and 130 mM NaCl at pH 7.5 and subjected to size-exclusion chromatography on Sephacryl S-500 HR (15×85 mm; Pharmacia Canada Ltd, Dorval, QC, Canada) for membranous vesicle purification. The eluate was collected into 18 fractions (0.5 mL). The SMV-positive fractions were detected by elevated absorbance at 280 nm. The fractions 8 to 15 were pooled (4 mL) and ultracentrifuged at 120,000 g for 2 hours at 4°C to pellet the SMVs. The pellets were resuspended in PBS at pH 7.4 and stored with matched vesicle-free SP samples at −80°C. The amount of total protein was assessed by the bicinchoninic acid (BCA) assay (Pierce Biotechnology, Rockford, IL).

3.2.9 TEX101 measurement in seminal plasma samples, SMVs and vesicle-free seminal plasma by ELISA

SMVs, vesicle-free SP and the original pooled pre- or post-vasectomy or infertile SP samples were mixed with corresponding reagent mixes (1:1) and incubated for 1 hour either at RT or at 63°C. Mixtures were further diluted with ELISA diluents before loading on the plate. Analysis was accomplished as described above.

3.2.10 TEX101 measurement in seminal plasma samples, SMVs and vesicle-free seminal plasma by SRM

TEX101 concentrations in the original SP pools, SMVs and vesicle-free fractions were calculated using the SRM assay described above. Ten micrograms of total protein from each sample were subjected to trypsin digestion. Heavy isotope-labeled peptide with a trypsin-cleavable tag AGTETAILATK*-JPTtag was used as an internal standard for absolute quantification of TEX101 protein.
**3.2.11 TEX101 measurement in seminal plasma by immunocapture-SRM**

Immunocapture-SRM assay was used to investigate the effect of SP pretreatment on antibody-antigen interactions. Initially, a pool of SP was prepared and was subjected to various types of treatment prior to analysis. Treatment options included: (i) mixing (1:1) with 6 M GndCl (pH 12) and incubation at RT for 1 hour, ii) mixing (1:1) with 4% DOC and incubation at RT for 1 hour, (iii) mixing (1:1) with 6 M GndCl (pH 12) and incubation at 63°C for 1 hour, (iv) mixing (1:1) with 4% DOC and incubation at 63°C for 1 hour, and (v) incubation at 63°C for 1 hour. Non-treated SP was also included in the analysis. According to the established protocol (210), white 96-well microtiter plates were coated with 500 ng per well of purified mouse immunoglobulins in 50 mM Tris-HCl, pH 7.8. Antibodies used for coating included a commercial mouse polyclonal anti-TEX101 antibody (ab69522; Abcam, Cambridge, MA), an in-house generated mouse monoclonal anti-TEX101 antibody (23ED616.8) and a mouse IgG (Equitech-Bio, Inc., Cat. #M60) as an isotype control. Following incubation overnight at RT, plates were washed twice with PBS. Three dilutions (10x, 100x, 1000x) of pre-treated SP samples in 6% BSA in PBS were loaded onto the plate (100 μL per well) and incubated for 2 hours at RT with gentle shaking. Following that, plates were washed three times with PBS and three times with 50 mM ammonium bicarbonate. A mix containing 50 mM ammonium bicarbonate, 5 mM DTT and 100 fmoles of heavy isotope-labeled TEX101 peptide AGTETAILATK*-JPTtag was added to each well and incubated for 30 min at RT. Subsequently, 10 mM iodoacetamide was added and samples were kept for 40 min in the dark at RT. L-methionine (5mM final) was added to prevent methionine oxidation in tryptic peptides. Extraction of peptides from solution and SRM quantification were accomplished as mentioned above. Raw files for each sample were recorded and analyzed with Skyline software, and peptide areas were used to calculate light-to-heavy ratios and TEX101 concentration in each sample.

**3.2.12 Statistics**

Power calculations were done with G*Power software (version 3.1.7, Heinrich Heine University Dusseldorf). GraphPad Prism (v4.0; GraphPad Software, San Diego, CA, USA) was used to generate scatterplots, perform statistical analysis and calculate the area under the Receiver
Operating Characteristic curve (ROC AUC) and diagnostic sensitivity and specificity. Comparisons for two groups were made using non-parametric Mann Whitney U test, while multiple groups were analysed by Kruskal-Wallis test, followed by Dunn's multiple comparison test. Since only 30 pre-vasectomy samples were matched to post-vasectomy samples, unpaired Mann-Whitney U analysis was applied. All hypotheses testing was two-tailed, and P-values <0.05 were considered statistically significant. The assay’s reference interval was estimated using the pre-vasectomy SP samples. TEX101 values were log_{10}-transformed and the lower and upper 95% CI of the arithmetic mean were calculated. Correlations between TEX101 concentration and other continuous variables were assessed by Spearman correlation coefficients ($r_s$). For the first and second intended uses (evaluation of vasectomy efficiency and differential diagnosis of azoospermia forms) the cut-offs and sensitivities were determined based on 100% specificity. For the third intended use (prediction of sperm retrieval), the cut-off, sensitivity and specificity were exploratory. Samples with missing data were not used in calculations.
3.3 Results

3.3.1 Measurement of TEX101 in 821 seminal plasma samples using guanidine-based treatment protocol

We initially measured TEX101 in 821 SP samples obtained from healthy fertile men pre- and post-vasectomy (PV), as well patients with unexplained infertility, oligospermia, and OA and NOA. SP samples and ELISA standards were treated with 3 M GndCl at pH 12 for 1 hour at RT, before analysis.

Results revealed high TEX101 concentrations in SP of fertile pre-vasectomy men (median 3,433 ng/mL; N=65), while it was undetectable in PV men (median 0.5 ng/mL; N=61). Similarly, TEX101 values were high in the group of men with unexplained infertility (median 2,875 ng/mL; N=276) and were significantly reduced in oligospermia (median: 270.5 ng/ml, Mann Whitney U test P<0.0001, N=269) and azoospermia (median 0.5 ng/mL, Mann Whitney U test P<0.0001, N=150) samples (Appendix 3.1). Since TEX101 is a germ cell-specific protein, we expected a strong correlation between its concentration in SP and the number of germ cells or spermatozoa in semen. Correlation between measured TEX101 concentration (ng/mL) and the sperm count (mln/mL), however, was of moderate strength (r=0.74). In addition, we noticed that a small fraction of patients (N=17) with unexplained infertility (sperm count >15 mln/mL) and oligospermia (sperm count >7 mln/mL) had undetectable levels of TEX101 (<0.5 ng/mL) in SP. It should be mentioned that previous studies of Tex101−/− mouse knockout models revealed normal phenotypes, normal sperm morphology and high sperm count, but absolute sterility of male mice (113). Thus, we decided to examine in detail samples from these 17 patients with alternative methods, such as SRM mass spectrometry.

We thus measured TEX101 by SRM in SP and matched spermatozoa of 17 patients with the high sperm count and undetectable TEX101 protein (Appendix 3.2). TEX101 protein was detected in both SP and spermatozoa in all 17 patients. This allowed us to exclude the hypothesis of TEX101 gene knock-outs in those patients. Additional examination of recent genomic data for loss-of-function mutations in TEX101 gene supported our conclusions. According to the ExAC Browser (http://exac.broadinstitute.org), stop gain, frameshift and splice donor mutations in the TEX101 gene were rare and detected with the allele frequency <0.004% in the general population of 60, 692 individuals.
3.3.2 Isoform identity of TEX101 in spermatozoa and seminal plasma

According to UniProt (www.uniprot.org), alternative splicing of the human TEX101 gene results in two isoforms: an extracellular membrane-bound isoform Q9BY14-1 (249 amino acids) and an intracellular isoform Q9BY14-2 (267 amino acids). To further investigate the identity of TEX101 in SP, we developed an SRM assay for the peptide QIQTSSQTSPEAMGTPR which was unique for the intracellular isoform Q9BY14-2. We hypothesized that the intracellular isoform Q9BY14-2 could be exclusively expressed in those 17 samples, but was not captured by our monoclonal antibody (generated against the extracellular isoform Q9BY14-1) and measured by ELISA. SRM measurements in SP and spermatozoa in these 17 samples revealed the absence of intracellular isoform Q9BY14-2 and the exclusive presence of the extracellular membrane-bound isoform Q9BY14-1 (Figure 3.1).
Figure 3.1 Investigation of TEX101 isoforms in spermatozoa and SP using SRM. Semen samples from men with high sperm count (≥7 mln/mL) and unexplained infertility and oligospermia (N=17) were centrifuged to separate spermatozoa from SP, which were then analysed by SRM. Peptide AGTETAILATK common for both isoforms was used to quantify total TEX101, while unique peptide QIQTSSQTSPEEMGTPR was used to detect the presumed intracellular TEX101 isoform Q9BY14-2. Corresponding heavy isotope-labeled peptides were used as internal standards. Intracellular isoform Q9BY14-2 was not detected in either spermatozoa or SP, so total TEX101 was assumed to be a 249 aa extracellular membrane isoform Q9BY14-1. Checkmark indicates detection and X absence of the peptides shown.
3.3.3 Confirmation of TEX101 germ cell-specificity and extracellular membrane localization by immunohistochemistry

Previously, TEX101 protein expression in testicular tissues was studied using rabbit polyclonal antibodies generated against TEX101 peptide fragments (Atlas Antibodies HPA041915 and HPA042513). Here, we stained testicular tissues with normal spermatogenesis using our monoclonal antibody 23ED228, generated against the full TEX101 protein. We observed more intense staining of germ cells and lower background with our antibody at a final concentration of 80 ng/mL, versus 400 ng/mL for HPA041915 antibody. We observed no staining of Leydig, Sertoli and spermatogonia cells, very weak cytoplasmic and membrane staining in primary spermatocytes and very intense extracellular/membrane staining in secondary spermatocytes, spermatids and testicular spermatozoa. Immunohistochemistry confirmed the exclusive expression of TEX101 in germ cells and its localization to the extracellular membrane (Figure 3.2).

Figure 3.2 Immunohistochemical staining of TEX101 protein in testicular tissue with active spermatogenesis. (a) Testicular tissue stained with our mAb 23ED228 (final concentration: 80 ng/ml). Cell types presented include Sertoli cells (negative staining), Leydig cells (negative staining) and germ cells at different stages of spermatogenesis, such as spermatogonia (negative), primary spermatocytes (positive cytoplasm), secondary spermatocytes (positive membrane), spermatids (positive) and spermatozoa. (b) Negative control (no primary Ab added). (c) H & E staining of testicular tissue showing nucleus (purple) and cytoplasm (pink).
3.3.4 Search for soluble TEX101 complexes in seminal plasma by size-exclusion chromatography

Previously, it was suggested that mouse TEX101 was involved in protein-protein interactions prior to its cleavage from the cell surface and release into SP (113, 163). Since potential complexes of human TEX101 might hamper its capture or detection by antibodies with ELISA test, we used size-exclusion chromatography to investigate the presence of soluble TEX101 complexes in the pool of pre-vasectomy SP. Following sodium deoxycholate (DOC)-63°C pretreatment of SP, TEX101 in each fraction was measured by ELISA. Results revealed that TEX101 was eluted as a single peak with 28 kDa size, as estimated by size-exclusion chromatography standards (Figure 3.3a). This molecular weight was in a good agreement with our previous estimates by Western Blot (~30 kDa for the glycosylated form and ~20 kDa for a deglycosylated form after PNGaseF treatment) (210).

We assumed that putative TEX101 complexes with other proteins would elute within fractions corresponding to higher molecular weights (i.e.>30 kDa). However, we detected negligible amounts of TEX101 (<8%) in those fractions (Figure 3.3a, before ultracentrifugation) and concluded that the majority of TEX101 was present as an unbound free soluble form. We also noted that some TEX101 (~14% of total) was found in fractions corresponding to high-molecular weight molecules (>500 kDa) and the void volume. Ultracentrifugation at 120,000 g of the same pre-vasectomy SP pool followed by size-exclusion chromatography (Figure 3.3a, after ultracentrifugation) resulted in non-detectable TEX101 in those high-molecular weight fractions. We thus hypothesized that a fraction of TEX101 might be associated with debris of spermatozoa cellular membranes, exosomes or seminal microvesicles (SMVs).
Figure 3.3 Investigation of TEX101 analyte identity in SP and the impact of SP pretreatment. (a) Size-exclusion chromatography was used to investigate TEX101 association with protein complexes and SMVs in the pre-vasectomy SP. We estimated the molecular weight of free soluble TEX101 as 28 kDa. No major complexes with other proteins were detected in the range 70-210 kDa. Ultracentrifugation at 120,000 g of the same pre-vasectomy SP pool followed by size-exclusion chromatography resulted in the non-detectable TEX101 in the high-molecular weight fractions. (b) Effect of GndCl- and DOC-based treatments on TEX101 measurements by ELISA were estimated for SP pools of pre-vasectomy, post-vasectomy or infertile men, as well as for their corresponding vesicle-free fractions and SMVs. Dotted lines represented ELISA limit of detection. Both GndCl- and DOC-based treatments were efficient in the pre-vasectomy pools with high TEX101, with substantial amounts of TEX101 found associated with SMVs. GndCl-based treatment, however, was not efficient in the pool of infertility samples with low TEX101, suggesting potential interaction between SP matrix and SMVs.
3.3.5 TEX101 association with SMVs in seminal plasma

All the above-mentioned experiments suggested that the failure of ELISA to measure low amounts of TEX101 in some SP samples (17 samples from unexplained infertility and oligospermia groups) was not related to TEX101 mutant forms, intracellular isoforms or soluble complexes. To investigate if TEX101 was associated with SMVs, we pooled 4 of these 17 SP samples (2 from each group) and isolated SMVs by centrifugation at 120,000 g. As a control, we also isolated SMVs from pools of pre- or post-vasectomy SP samples. SMVs were denatured with Rapigest SF at 65°C, proteins were digested by trypsin, and TEX101 was quantified by the antibody-independent SRM assay in SMVs, matched vesicle-free supernatants and original SP pools. Results confirmed the presence of TEX101 in SMVs and vesicle-free fractions of the 4-sample infertility pool, as well as of the pre-vasectomy pool. TEX101 levels in SMVs, matched vesicle-free supernatants and initial SP of the PV pool were below the limit of detection (27 pg/µg of total protein).

We also measured by SRM in the pre-vasectomy SP and corresponding SMVs three highly tissue-specific secreted proteins which represented major glands in the male urogenital system. We detected high amounts of KLK3 (prostate-specific protein; 4.14 ng/µg of total protein), SEMG1 (seminal vesicle-specific protein; 37.3 ng/µg), ADAM7 (epididymis-specific protein; 0.45 ng/µg) and TEX101 (0.14 ng/µg) in the digest of SMVs. Interestingly, relative ratios of SEMG1, ADAM7 and TEX101 versus KLK3 were enriched 3.3, 40 and 12-fold in SMVs relative to the initial SP. No enrichment of these proteins (e.g. same ratios versus KLK3 in SMVs and SP) would indicate that SMVs absorbed non-specifically KLK3, SEMG1, ADAM7 and TEX101. Substantial enrichment of ADAM7 and TEX101 in SMVs thus suggested the specific mechanism of their association with SMVs. Such mechanism was previously demonstrated for ADAM7 protein (214). Overall, the presence of prostate-, seminal vesicle-, testis- and epididymis-specific proteins indicated that SMVs were produced by multiple glands in the male urogenital system, as previously discussed (32).

We then investigated which of two SP pre-treatment protocols (GndCl- or DOC-based) would result in more complete SMV lysis and release of TEX101 and thus facilitate accurate measurement by ELISA. As a result, the same levels of TEX101 in the pre-vasectomy SP pool were measured with both treatment protocols in SMVs and vesicle-free SP (Figure 3.3b). Substantial amounts of TEX101 were associated with vesicles, as previously shown by SRM.
Amounts of TEX101 in the PV SMVs and vesicle-free SP pool were below detection (<0.5 ng/mL) in all fractions (Figure 3.3b).

Surprisingly, the pool of 4 infertility samples revealed the discrepancy between two pre-treatment protocols. In the infertile males, incubation of the original pool and the vesicle-free fraction with sodium DOC at 63°C identified significantly higher amount of TEX101 compared to GndCl at RT. GndCl treatment was thus less effective in the initial SP compared to the SMV fraction, suggesting some interaction between SP matrix and SMVs (Figure 3.3b). Based on these findings, we decided to re-examine the efficiency of treatments with GndCl at RT or DOC at 63°C and investigate any parameters that could affect the ELISA performance.

### 3.3.6 Effect of seminal plasma pretreatment on TEX101 measurements by ELISA

Immunocapture-SRM was used to investigate the impact of SP pretreatment on the efficiency of TEX101 capture by the in-house generated monoclonal 23ED616.8 and the commercial polyclonal ab69522 antibodies. Aliquots of pooled SP samples were mixed (1:1) with 6 M GndCl (pH 12) or 4% DOC and incubated at RT or 63°C. Samples were analysed in triplicate at 10-, 100- and 1,000-fold dilutions (Appendix 3.3).

Results confirmed our earlier observations that ab69522 captured the native form of TEX101 in SP. Incubation of SP with GndCl or DOC disrupted the capture of TEX101 by ab69522. As we showed previously (210), 23ED616.8 possessed lower affinity for native TEX101 than ab69522. Sample pretreatment with detergents at 63°C, however, resulted in more efficient capture, at least 10-fold higher than TEX101 capture after denaturation at 63°C without detergents. We also noted that GndCl at 63°C resulted in a slightly lower signal, possibly due to a higher impact on disruption of TEX101-antibody complexes.

Both GndCl- and DOC-based protocols worked equally well in SP with high amounts of TEX101 (~5,000 ng/mL, pre-vasectomy samples). The GndCl-based protocol was compromised only in some infertility samples with low TEX101 levels (<300 ng/mL). Since detection of even very low TEX101 levels would indicate some residual spermatogenesis in testis and increased chances for sperm retrieval, it was critical to demonstrate the robust performance of the DOC-
based protocol in all samples, including samples with very low amounts of TEX101 (~0.5 ng/mL). Considering all of the above, we re-measured our entire clinical cohort of SP samples by ELISA and DOC-based protocol.

3.3.7 Stability of TEX101 in semen

We previously demonstrated high stability of TEX101 protein in SP (210). Since the use of our test may involve SP collection and storage at home, and then transportation of the whole semen prior to measurements in the clinical lab, we assessed TEX101 stability in the whole semen stored at +4°C for up to 14 days. A semen sample from a patient with unexplained infertility was obtained at the infertility clinic, let to liquefy at room temperature (RT) and aliquoted. One aliquot was centrifuged right away, and TEX101 was measured in SP by ELISA using the DOC-based protocol. Other four aliquots of the whole semen were stored at +4°C for 5, 6, 9 and 14 days, centrifuged and measured by TEX101 ELISA. As a result, TEX101 concentrations in SP slightly varied from day to day (2.4±0.4 µg/mL, CV=16%), but no particular trends indicating TEX101 degradation were observed. We thus concluded that TEX101 protein was stable in both SP and whole semen.

3.3.8 Evaluation of TEX101 as a male infertility biomarker in a population of 805 men

TEX101 levels were measured in 805 SP samples using the DOC-based pretreatment protocol (Figure 3.4a). Assay’s LOD was calculated as 0.5 ng/mL. TEX101 was detected in high amounts in all pre-vasectomy samples (N=64; median: 5,436 ng/mL), while it was undetectable in the PV SP (N=57). It should be noted that a fraction of PV samples showed higher than usual background fluorescence (resulting in TEX101 levels between 0.5 and 0.9 ng/mL). Such elevated background, however, did not affect the proposed clinical cut-offs. TEX101 levels were high in the unexplained infertility group (N=277; median: 4,967 ng/mL) and significantly lower in the oligospermic (N=270; median: 450 ng/mL) and azoospermic (N=137; median: 0.54 ng/mL) groups (Kruskal-Wallis test P<0.0001). Likewise, differences between paired groups were also significant (Dunn's multiple comparison test P-values <0.001), apart from the pre-
vasectomy versus unexplained infertility group ($P > 0.05$) and PV versus azoospermia ($P > 0.05$). Interestingly, comparison of 30 matched pre-vasectomy (median 7,014 ng/mL) and PV men (median 0.5 ng/mL) revealed that median TEX101 concentration decreased at least 13,500-fold after vasectomy, with the maximum decrease of 113,000 fold.

Based on 64 pre-vasectomy samples, the normal range of TEX101 in the healthy fertile men was estimated between 515.1 ng/mL (95% CI: 386.8-686.0) and 50,360 ng/mL (95% CI: 37,817-67,065). Our data confirmed that TEX101 was a highly informative biomarker to predict the success of vasectomy or vasectomy reversal (96). TEX101 at >0.9 ng/mL differentiated between pre- and post-vasectomy samples with 100% sensitivity at 100% specificity and ROC AUC 1.00 [95% CI 1.00-1.00]. We thus suggest that TEX101 has a clinical utility to non-invasively evaluate the success of vasectomy or vasectomy reversal (Figure 3.5).

TEX101 was not a useful marker of unexplained infertility (AUC=0.56 [95% CI 0.48-0.63], $P > 0.05$), but performed well in oligospermia (AUC=0.88 [95% CI 0.84-0.92], $P < 0.001$) and azoospermia (AUC=0.99 [95% CI 0.98-1.00], $P < 0.001$). A cut-off value 496 ng/mL provided 96% sensitivity at 98% specificity for distinguishing fertile pre-vasectomy men from patients with azoospermia.

TEX101 levels were broadly distributed in the azoospermia group with the unknown form (median 0.6 ng/mL), NOA with the unknown histological subtype (0.6 ng/mL), as well as NOA-HS (70.4 ng/mL) and NOA-MA (1.9 ng/mL). Levels in the NOA-SCO subtype and OA were mainly below LOD of 0.5 ng/mL (Figure 3.4b). Based on our small set of NOA samples with the biopsy-confirmed histological subtypes, TEX101 could differentiate between HS and SCO (Mann-Whitney U test $P = 0.0336$), but not between MA and SCO ($P = 0.10$). Additional studies with a larger sample size are required to investigate if TEX101 can differentiate between histological subtypes of NOA.

Finally, the correlation between TEX101 levels and sperm count improved ($r_s = 0.83$, $P < 0.001$ versus $r_s = 0.74$ with GndCl-based protocol), but was still not very strong (Appendix 3.4). This fact suggested that (i) TEX101 expression per germ cell may vary in different individuals; (ii) the fraction of TEX101 cleaved from the surface may vary, or (iii) TEX101 was released into SP not only by epididymal spermatozoa, but also by testicular germ cells.
Fertile pre-vasectomy (N=64)
Fertile post-vasectomy (N=57)
Unexplained infertility with SC≥15 mln/mL (N=277)*
Oligospermia with SC<15 mln/mL (N=270)
Azoospermia with SC=0 (N=137)
TEX101 in SP, as measured by ELISA, ng/ml

Pre-V Post-V
Idiopathic
Oligospermia Azoospermia
1
10
100
1,000
10,000
100,000

Azoospermia
NOA (HS)
NOA (MA)
NOA (SCO)
NOA (N/A)
OA
1
10
100
1,000
10,000

NOA-ns (11)
NOA-ss (15)

TEX101 in SP, as measured by ELISA, ng/ml

No sperm or spermatids on TESE (N=11)
Some sperm or spermatids on TESE (N=15)
P<0.0001 P<0.001 P<0.001
**Figure 3.4 Evaluation of TEX101 as a biomarker.** (a) TEX101 levels in SP of healthy fertile pre- and post-vasectomy men and patients with unexplained infertility, oligospermia and azoospermia, as measured by ELISA, using DOC-based treatment. Median values for each group are presented as horizontal lines. TEX101 differentiated pre-vasectomy samples from post-vasectomy samples (Kruskal-Wallis test with the Dunn's multiple comparison test $P$-value<0.0001), oligospermia ($P$<0.001) and azoospermia ($P$<0.001), but not unexplained infertility ($P$>0.05). (b) In the azoospermia group, TEX101 could differentiate between HS and SCO (Mann-Whitney U test $P$=0.0336), but not between MA and SCO ($P$=0.10). SC, sperm count; NOA, non-obstructive azoospermia; OA, obstructive azoospermia; HS, hypospermatogenesis; MA, maturation arrest; SCO, Sertoli cell-only syndrome; mln, million. (c) Prediction of spermatozoa or spermatids retrieval in NOA patients using TEX101 ELISA and DOC-based treatment of SP. TESE, testicular sperm extraction. *Female factor is not known for this group.
Figure 3.5 Clinical utility of TEX101 ELISA as a test to evaluate vasectomy success, differentiate between non-obstructive and obstructive azoospermia and predict the success of sperm retrieval in patients with non-obstructive azoospermia. A cut-off value of 0.9 ng/mL provided 100% sensitivity at 100% specificity, with ROC AUC=1.00 [95% CI 1.00-1.00], for distinguishing pre- and post-vasectomy men. Combination of TEX101≥0.9 ng/mL with epididymis-specific protein ECM1≥2.3 µg/mL provided 81% sensitivity at 100% specificity to differentiate between non-obstructive and obstructive azoospermia. A TEX101 cut-off value of ≥0.6 ng/mL provided 73% sensitivity at 64% specificity and ROC AUC=0.69 [95% CI 0.48-0.89] to predict sperm retrieval in patients with non-obstructive azoospermia.
3.3.9 Evaluation of TEX101 as a biomarker to differentiate between NOA and OA

TEX101≥0.9 ng/mL differentiated between NOA (N=81) and OA/PV (N=93) with AUC=0.67 (95% CI 0.59-0.75; P=0.00012) and 32% sensitivity at 99% specificity. Thus, TEX101 alone was not a strong marker for the non-invasive differentiation between NOA and OA. Previously, we proposed that combination of epididymis-specific protein ECM1 (measured by ELISA) and TEX101 (measured by mass spectrometry) could differentiate between NOA and OA and thus eliminate the majority of diagnostic testicular biopsies (97).

Here, we tested the combination of ECM1 and TEX101, both measured by ELISA, in NOA (N=42) and OA/PV (N=70) samples. TEX101≥0.9 ng/mL detected five additional NOA cases missed by ECM1≥2.3 µg/mL and thus increased sensitivity to detect NOA from 69% (ECM1 alone) to 81%, at 100% specificity. Assuming 20% prevalence of OA in the azoospermic population, combination of ECM1 and TEX101 provided 100% positive and 57% negative predictive values. All azoospermia cases diagnosed as NOA by ECM1≥2.3 µg/mL and TEX101≥0.9 ng/mL will thus be correct (32 patients), while 19% of patients diagnosed as OA based on ECM1<2.3 µg/mL and TEX101<0.9 ng/mL (8 patients in the current set) will be actually NOA (false negatives). We believe that false negatives are acceptable since presumed OA patients would be followed up with sperm retrieval which will re-classify false OA as NOA. That was our rationale to select biomarker cut-offs which provided 100% specificity at the expense of lower sensitivity. In combination with ECM1, non-invasive differentiation between NOA and OA is another clinical utility of TEX101 (Figure 3.5). Combined ECM1 and TEX101 test may eliminate the majority of diagnostic testicular biopsies.
3.3.10 Evaluation of TEX101 as a biomarker to predict sperm retrieval in NOA patients

Twenty six NOA patients in our cohort were treated with testicular sperm extraction (TESE), a surgical procedure to retrieve spermatozoa or spermatids from the testis and use for *in vitro* fertilization. Here, we evaluated the clinical utility of TEX101 to predict spermatozoa or spermatids retrieval by TESE in NOA patients. Note that useful biomarkers to predict sperm retrieval are currently not available. Testicular biopsy followed by histological subtyping thus remains the only practical diagnostic procedure.

Overall, the rate of successful sperm retrieval in NOA patients was previously estimated as 53% (215). Rates varied for different histological subtypes of NOA: 81% for HS, 21% for MA and 31% for SCO subtypes (215). In our cohort of NOA samples with known histological subtypes and measured TEX101 (20 out of 26 samples in Figure 5C), success rates for the corresponding subtypes were 100%, 55 and 0%, respectively. Previously proposed biomarkers, such as follicle-stimulating hormone in blood, had moderate predictive efficiency (AUC=0.72) with 70% sensitivity at 62% specificity (216).

Analysis of our clinical cohort (Figure 3.4c) revealed TEX101 AUC=0.69 [95% CI 0.48-0.89]. With the cut-off of ≥0.6 ng/mL, TEX101 had 73% sensitivity, 64% specificity, 70% positive and 68% negative predictive values (Figure 3.5). Based on our small study with 26 patients, TEX101 alone had moderate diagnostic value as a predictor for spermatozoa or spermatid retrieval rate in patients with NOA. Due to its germ cell-specificity, however, we suggest that TEX101 should be thoroughly evaluated in a larger cohort of NOA patients.
3.4 Discussion

Immunoassays, such as sandwich ELISA, are indispensable tools for quantification of proteins in biological and clinical samples. Availability of ELISAs facilitates the clinical validation of putative protein biomarkers and enables their translation into diagnostic laboratory tests. The clear advantages of ELISAs over other methods for protein quantification include high analytical sensitivity, selectivity in complex biological matrices, high-throughput analysis, low reagent costs, simple execution and straightforward interpretation of results (217).

Novel protein assays should be thoroughly evaluated prior to their use in the clinic. Clinical evidence of a novel diagnostic test includes scientific evidence (association of an analyte with the clinical condition), analytical performance (analytical sensitivity, selectivity, LOD, linearity and reproducibility) and clinical performance (data to support reference ranges). It should be emphasized that different protein assay platforms, such as ELISA or mass spectrometry, may result in different reference values and different diagnostic performance for the same protein biomarker due to different analyte identities (peptide or protein, free or bound form, linear or conformational epitope) and the use of different calibration standards (136).

Our previous work on TEX101 as a biomarker of azoospermia (97) motivated us to generate monoclonal antibodies and develop a first-of-a-kind TEX101 ELISA (210). It should be noted that SP is not a conventional fluid for clinical diagnostics and has some distinct differences, such as fast protease-mediated liquefaction, high viscosity and abundance of SMVs, which can hamper the performance of immunoassays. We realized that SP might need additional treatment procedures to facilitate protein quantification with high sensitivity. To improve ELISA sensitivity, we tested multiple combinations of detergents, temperatures and pH values, to select two SP pretreatment protocols: GndCl (pH 12) at RT and DOC at 63°C (210). Successful analysis of 10 pre- and 10 post-vasectomy samples and the scaled-up production of monoclonal antibodies encouraged us to design the large pre-clinical validation of TEX101 ELISA in a cohort of more than 800 SP samples available in our biobank.

We first implemented our assay with GndCl-based protocol due to its simplicity and pre-treatment at RT. Overall, the performance of the test was impressive, with absolute discrimination between healthy pre-vasectomy men and patients with OA or PV. While reviewing the results, however, we noticed that 17 samples from the unexplained infertility and
oligospermia groups with substantial sperm count (≥7 mln/mL) exhibited undetectable levels of TEX101 in SP (Appendix 3.1).

Comprehensive characterization of the analyte identity of our TEX101 ELISA was thus required. Mass spectrometry analysis of TEX101 in SP and spermatozoa suggested the exclusive presence of the extracellular membrane isoform Q9BY14-1 and the absence of intracellular isoform Q9BY14-2 which could potentially be undetectable by ELISA. IHC staining of testicular tissues confirmed the exclusive expression of TEX101 in germ cells and its extracellular membrane localization. Our initial hypothesis for mutant TEX101 protein in the infertile men was rejected by detection of TEX101 in both SP and corresponding spermatozoa by mass spectrometry. Size-exclusion chromatography confirmed the molecular weight of TEX101 (28 kDa), did not reveal any additional TEX101-protein complexes in SP, but suggested association of a small fraction of TEX101 with high-molecular weight structures.

Numerous studies previously suggested the presence of SMVs, such as prostasomes and epididymosomes, in semen (32, 218-220). SMVs are produced by glands of the male urogenital tract and modulate spermatozoa maturation, transport and capacitation. Even though TEX101 protein is not expressed in epididymis, it was previously identified in the epididymosome fractions (31). However, it is still not clear if soluble TEX101 is simply absorbed by epididymis-secreted vesicles or is present in the vesicles produced in the lumen of seminiferous tubules (such as residual cytoplasm of spermatids) and co-purified with epididymosomes.

ELISA results revealed that both GndCl- and DOC-based protocols performed equally well in the pre-vasectomy pools including initial SP, vesicle-free SP and the SMVs fraction. However, the superior efficiency with the infertility pool was found for the DOC-, but not GndCl-based treatment or denaturation by heat only. We thus believe that when TEX101 levels are very high (~5,000 ng/mL in the pre-vasectomy pool), the impact of vesicles is negligible. However, when TEX101 levels are very low (<300 ng/mL in the infertility pool), the impact of SMVs becomes significant. Our results suggested that unlike DOC-based protocol, GndCl-based protocol might not fully release TEX101 associated with SMVs in the matrix of SP. We would like to emphasize that these effects were noticeable only when TEX101 levels in SP were very low.

Finally, we re-measured our SP samples using DOC-based pretreatment protocol. Note that the range of measured TEX101 concentrations in SP exceeded 127,000-fold (0.5 - 63,825 ng/mL).
This is an unprecedented range of protein concentration in biological fluids, exceeding the 50,000-fold range for C-reactive protein (221) and 70,000-fold range for human chorionic gonadotropin in blood serum (222). We defined the reference intervals of TEX101 in the healthy pre-vasectomy population and demonstrated clinical utility of TEX101 as a biomarker to predict the success of vasectomy or vasectomy reversal (Figure 3.5). TEX101 differentiated between pre- and post-vasectomy samples with 100% sensitivity and 100% specificity, which is an extraordinary performance for a protein biomarker in biological fluids. In future, TEX101 assay may be implemented as a home-based test and replace hospital-based sperm counting for patients after vasectomy or vasectomy reversal.

TEX101 alone was not a highly informative marker for the non-invasive differentiation between NOA and OA (32% sensitivity at 99% specificity). However, combination of TEX101 with ECM1 increases sensitivity to detect NOA from 69% (ECM1 alone) to 81%, both at 100% specificity. With our two marker test, a fraction of NOA patients could be misdiagnosed as OA. Since TESE is recommended for all OA patients due to the very high chances of sperm retrieval, all misdiagnosed patients will be followed by TESE and thus re-classified. Thus, it is acceptable to misdiagnose some NOA patients as OA, but not vice versa. The decision to proceed with TESE is critical for NOA patients, for which the chances of TESE are not known and may be quite low. OA patients misdiagnosed as NOA may choose to avoid TESE, even though being fertile using TESE and assisted reproduction techniques. We thus suggest that differentiation between NOA and OA with ECM1 and TEX101 is another clinical utility of TEX101. Such test is non-invasive and will eliminate the majority of diagnostic testicular biopsies.

Our test also revealed that TEX101 may be an informative biomarker to predict sperm or spermatid retrieval (AUC=0.69; 73% sensitivity at 64% specificity). It should be noted that the outcome of sperm retrieval by TESE, which is the clinical reference standard in this case, may depend on the length of surgery and be different in different clinics. It is still possible that some rare focal spermatogenesis would be present in some seminiferous tubules, but would be missed during surgery. TEX101 levels in SP may reflect the cumulative yield of spermatogenesis and thus be useful to detect rare focal spermatogenesis.

Prediction of sperm retrieval by TEX101 was comparable to other biomarkers to predict sperm retrieval including seminal protein LGALS3BP (AUC=0.76; 100% sensitivity at 45% specificity) (223) and seminal leptin (AUC=0.59; 43% sensitivity at 75% specificity) (224),
seminal ESX1 mRNA (84% sensitivity at 28% specificity) (225) and blood serum FSH protein (AUC=0.62; 71% sensitivity at 68% specificity) (224). Better biomarkers to predict sperm retrieval in NOA patients are still required. Ultimate germ cell specificity of TEX101 protein warrants its thorough validation in a large cohort of NOA patients.

To conclude, in this study we presented an optimized TEX101 immunoassay and its pre-clinical evaluation in a large set of SP samples. We propose to implement our TEX101 ELISA as a clinical test to evaluate vasectomy success, stratify azoospermia forms and subtypes and predict the success of sperm retrieval in NOA patients. It should be noted that SP is a promising but unconventional fluid for clinical diagnostics. Our work revealed potential issues with SP as a fluid for clinical analyses, demonstrated solutions for the measurement of protein analytes in SP and paved the road to translation of SP-based diagnostic tests into clinic.
Chapter 4

Identification of a physical TEX101 interactome and validation of a testis-specific protein complex TEX101-DPEP3
Chapter 4

4 Identification of a physical TEX101 interactome and validation of a testis-specific protein complex TEX101-DPEP3

4.1 Introduction

Spermatogenesis is a highly organized process involving intricate cell cycle progression and differentiation of spermatogonial stem cells and their transformation into mature spermatozoa. The necessity to silence transcription and translation at the late stages of spermatogenesis resulted in evolution of epididymis, in which spermatozoa are activated by epidydimis-secreted proteins through numerous proteolytic cascades and protein-protein interactions (PPIs). With no cell culture models of human germ cells available as yet, human spermatogenesis remains one of the least studied developmental processes.

Numerous studies in mice emphasized the importance of protein-protein interactions for the production of fertile spermatozoa. Null mice models of selected testis-specific genes presented male infertility phenotypes, presumably through disrupted PPIs and improper processing of proteins during spermatogenesis and sperm maturation (108-114, 166, 226, 227). Early studies discovered the essential role of numerous cell surface proteins for sperm-oocyte interaction and fusion (40). Some of the most critical factors included metalloprotease-disintegrin ADAM2 (228), the cell adhesion tetraspanin CD9 (229) and the sperm-egg fusion protein IZUMO1 (230). Recent discovery of the cell surface recognition complex between IZUMO1 and the sperm-egg fusion protein JUNO revealed detailed mechanisms on gamete recognition and sperm-oocyte fusion (231, 232). Of 1,035 highly testis-enriched proteins in the human proteome (101), nearly 160 proteins are membrane-bound and may be directly involved in remodeling of spermatozoa cell surface during spermatogenesis, sperm transit and sperm-oocyte interaction. Elucidation of the exact roles of germ cell-specific proteins and their PPIs in human reproduction is thus still an emerging field.

Proteomics and mass spectrometry emerged as techniques of choice to discover PPIs and elucidate molecular functions of proteins (233). Affinity purification or co-immunoprecipitation (co-IP) approaches followed by mass spectrometry identify direct and indirect PPIs under native physiological conditions (124-126). Advances in sensitivity and throughput of mass
spectrometry facilitated mapping of interactomes of bacteria (234), yeast (235, 236), insects (237) and human cells (238). High resolution mass spectrometry combined with label-free quantification approaches enabled identification of high-confidence PPIs after a single step of affinity purification (129).

In our study, we focused on the testis-specific protein TEX101 which is exclusively expressed on the surface of testicular germ cells (25) and was suggested as a cell-surface chaperone involved in trafficking and maturation of certain cell surface proteins essential for fertilization in mice (113, 239). Four TEX101-interacting proteins (ADAM3-6) previously discovered in mice (113, 163) appeared to be pseudogenes in human. While human TEX101 has been validated as a prominent biomarker of male infertility (97, 240), its functional role in human reproduction is not known, and interactome of human TEX101 has never been profiled. In this work, we established a quantitative co-IP-MS platform to discover human TEX101 interactome and elucidate its role in PPIs of testis-specific proteins.
4.2 Materials and Methods

4.2.1 Expression and purification of recombinant proteins TEX101 and DPEP3

Human recombinant proteins TEX101 and DPEP3 were individually produced in Expi293F transient expression system according to manufacturer’s recommendations (Invitrogen). Briefly, DNA coding for the mature forms of TEX101 and DPEP3 (aa 26-222 and 36-463, respectively) were cloned into a pcDNA3.4 plasmid for mammalian protein expression (GeneArt™ Gene Synthesis, Invitrogen). Expi293 cells were grown in suspension culture on a shaker platform (125 rpm) inside a CO₂ (8%) cell culture incubator, and cell culture containing secreted TEX101 and DPEP3 proteins were collected and centrifuged 72 and 96 hours post-transfection, respectively. TEX101 and DPEP3 protein production was assessed by Western blot analysis with a rabbit polyclonal anti-TEX101 antibody HPA041915 (Sigma-Aldrich, St. Louis, MO) and a rabbit polyclonal anti-DPEP3 antibody HPA058607 (Sigma-Aldrich), respectively. Anion exchange chromatography was employed for the purification of recombinant TEX101 and DPEP3 from Expi293 cell culture supernatants. Purification was performed with an automated AKTA FPLC system on a pre-equilibrated 5-mL anion-exchange HiTrap Mono Q™ Sepharose high performance column (GE Healthcare). Concentrated culture supernatant was diluted 4-fold with 50 mM Tris, pH 9.0 (running buffer A), and following binding and washing steps, recombinant TEX101 was eluted in 4-mL fractions with a linear gradient of 50 mM Tris, 1 M NaCl, pH 9.0 (buffer B) at a flow rate of 2 mL/min, as follows: (i) 5% B for 10 min; (ii) continuous gradient from 5 to 40% for 44 min; and (iii) 100% B for 15 min. TEX101 was eluted with 250 mM NaCl in 50 mM Tris, pH 9.0. Likewise, culture supernatant containing DPEP3 was diluted 4-fold with 50 mM Tris, pH 9.5 (buffer A), and recombinant DPEP3 was eluted in 4-mL fractions using the linear gradient of 50 mM Tris, 1 M NaCl, pH 9.5 (buffer B), as described above. The concentration of TEX101 in various fractions was measured by an in-house TEX101 ELISA (210), while DPEP3 concentration was determined by a selected reaction monitoring (SRM) assay, by using a heavy isotope-labeled peptide SWSEELQGVL* with a trypsin-cleavable quantifying JPT tag (serine-alanine-[3-nitro]tyrosine-glycine; SpikeTides™_TQL peptides) JPT Peptide Technologies GmbH, Berlin, Germany). The positive fractions, for TEX101 and DPEP3, were pooled and the purity and the molecular mass of recombinant proteins, were determined by SDS-PAGE stained with Coomassie Blue. Following
SDS-PAGE, all visible gel bands were excised, subjected to in-gel digestion and analyzed by LC-MS/MS in Q Exactive™ Plus Hybrid Quadrupole-Orbitrap™ Mass Spectrometer (Thermo Scientific). For protein identification, the LC-MS/MS raw files were analyzed using the MaxQuant software (version 1.5.2.8) with the human UniProtKB/Swiss-Prot database (HUMAN5640_sProt-012016).

4.2.2 Monoclonal antibody production against human TEX101 and DPEP3

Female BALB/c mice were obtained from the Toronto Centre for Phenogenomics (TCP). All animal research was approved by TCP Animal Care Committee (Animal Use Protocol #14-04-0119aH). Monoclonal antibody (mAb) production was performed as previously described (210). Briefly, mice were inoculated with purified human recombinant proteins TEX101 or DPEP3, three booster injections were performed at 3-week intervals. Extracted spleen cells were fused with NSO murine myeloma cells (5:1 ratio), and successfully fused cells were selected using HAT media (Invitrogen) supplemented with 20% fetal bovine serum (HyClone, Thermo Fischer Scientific, Waltman, MA). Cell culture supernatants were tested for IgG and IgM antibody secretion using an immunoassay protocol described elsewhere (210).

4.2.3 Immunocapture-SRM screening for clones producing antibodies against native TEX101 and DPEP3 proteins

According to our established protocol (210), 96-well polystyrene plates were coated with 500 ng/well of sheep anti-mouse IgG-Fcγ fragment-specific antibody (Jackson ImmunoResearch) in 50 mM Tris buffer (pH 7.8). After overnight incubation, plates were washed twice with PBS, and hybridoma culture supernatants were added to the plates, and incubated for 2 hours at RT with gentle shaking. A commercial mouse polyclonal anti-TEX101 antibody (ab69522; Abcam, Cambridge, MA) was used as a positive control. Plates were washed 6 times with phosphate-buffered saline (PBS; 137 mM NaCl, 2.7 mM KCl, 10 mM Na2HPO4, 1.8 mM KH2PO4), and 100 μl of 100-fold diluted testicular tissue lysate pool (obtained from normal donors or patients with testicular cancer, with confirmed active spermatogenesis), in 6% BSA were added in each well. After 2 hour incubation at RT, plates were washed 3x with PBS, and 3x with 50 mM
ammonium bicarbonate (ABC). Prior to MS analysis, a mix of 50 mM ABC, 5 mM dithiothreitol (DTT) (Sigma-Aldrich), 100 fmoles of heavy isotope-labeled TEX101 proteotypic peptide AGTETAILATK*-JPTtag with a trypsin-cleavable tag, and 0.05% RapiGest SF was added to each well. Following 30 min incubation at RT, 10 mM iodoacetamide were added and samples were kept for 40 min in the dark at RT. Samples were digested overnight at 37°C after the addition of proteomics-grade porcine trypsin (Sigma-Aldrich, #T6567). After digestion, trypsin inactivation and RapiGest cleavage were achieved by adding trifluoroacetic acid (TFA) at final concentration 1%. L-methionine (5 mM final) was also added to the digests to prevent the oxidation of methionine residues.

Regarding DPEP3 antibodies, a two-step IP-SRM, as described above, was used for the screening of hybridoma culture supernatants for mAbs against recombinant human DPEP3, and native DPEP3 protein. Ten-fold diluted seminal plasma pool was added in each well. Serum of immunized mice was used as a positive control. Following 2 hours of incubation at RT and two wash steps with PBS and 50 mM ABC, 300 fmoles of heavy isotope-labeled DPEP3 proteotypic peptide SWSEEELQGVR*-JPTtag with a trypsin-cleavable tag, were added on the plate, and samples were prepared for mass spectrometric analysis, as described above.

TEX101 or DPEP3 peptides were extracted from solution using C18 OMIX tips (Varian Inc., Lake Forest, CA) and were monitored in a non-scheduled SRM mode during a 30 min LC gradient in TSQ Quantiva™ triple quadrupole mass spectrometer (Thermo Scientific). Raw files for each sample were analyzed with Skyline software (v3.6.0.10493), and relative abundance of TEX101 and DPEP3 were calculated using the ratio of endogenous versus internal standard peptides. Hybridoma cultures, positive for antibody secretion against native TEX101 and DPEP3, were grown and transferred in serum-free media (Invitrogen). Supernatants were harvested and purified using protein G according to the manufacturer’s protocol (GammaBind Plus, GE Healthcare).
4.2.4 Pairing of anti-TEX101 monoclonal antibodies in a sandwich immunoassay

The 96-well microtiter plates were coated with 500 ng/well of purified mouse monoclonal anti-TEX101 antibodies (34ED229, 34ED233, 34ED470, 34ED556, 34ED604 and 34ED629) in 50 mM Tris-HCl buffer at pH 7.8. Plates were washed (0.05% Tween 20 in 20 mM Tris-HCl and 150 mM NaCl at pH 7.4) and 100 µL of SP pool sample diluted 50-fold in assay diluent (60 g/L BSA, 25 mL/L normal mouse serum, 100 mL/L normal goat serum and 10 g/L bovine IgG in 50 mM Tris-HCl at pH 7.8) were loaded on the plates. Assay diluent was also added to the plates as a negative control. After 2 hours of incubation with gentle shaking, plates were washed 3 times. Biotinylated mouse monoclonal anti-TEX101 antibodies in the assay diluent (250 ng in 100 µL per well) were added to the plate and incubated for 1 hour. All mAbs were paired with each other in a sandwich format, generating 36 combinations (6x6). Plates were washed three times, and streptavidin-conjugated alkaline phosphatase was added for 15 min with gentle shaking. Wells were washed six times, and 100 µL of diflunisal phosphate (DFP) solution in substrate buffer (0.1 M NaCl, 1 mM MgCl₂ in 0.1 M Tris at pH 9.1) were added and incubated for 10 min with gentle shaking. Lastly, the developing solution (1 M Tris-HCl, 0.4 M NaOH, 2 mM TbCl₃ and 3 mM EDTA) was mixed for 1 min, and time-resolved fluorescence was measured with the Wallac EnVision 2130 Multilabel Reader (Perkin Elmer).

4.2.5 Testicular tissue, spermatozoa and seminal plasma samples

Testicular tissues with active spermatogenesis (confirmed by histological examination) were obtained by orchiectomy from men with scrotal pain, testicular masses or testicular cancer (adjacent normal tissues). Sample collection was approved by the institutional review board of Mount Sinai Hospital (approval #09-0156-E). Upon removal, testicular tissues were subjected to snap-freezing, and stored in liquid nitrogen. Semen was collected from healthy fertile pre-vasectomy patients at Mount Sinai Hospital (#08-117-E) and University Health Network (#09-0830-AE). Semen was allowed to liquefy at RT for 1 hour, and then aliquoted and centrifuged 3 times at 13,000 g for 15 min at RT. The SP and sperm cells were separated and stored at -80°C.
4.2.6 Preparation of testicular tissue and spermatozoa lysates

Testicular tissue lysis and solubilisation of protein complexes was performed under optimized lysis conditions. Cryogenic tissue lysis was followed by suspension of the frozen sample powder in the lysis buffer containing 20 mM HEPES (pH 7.5), 100 mM NaCl, 1% w/v 3-[(3-cholamidopropyl)-dimethylammonio]-1-propanesulfonate (CHAPS), and 1% v/v protease inhibitor cocktail [1:10 (w/v) ratio of tissue to lysis buffer]. After overnight incubation at 4°C, testicular tissue lysates were centrifuged at 15,000 g for 10 min at 4°C. Testicular tissue lysate supernatants were pooled, and total protein concentration was determined by the bicinchoninic acid assay (BCA assay). Spermatozoa lysis and protein complexes extraction was performed in a similar way. Several sperm cell samples were pooled and incubated with lysis buffer overnight at 4°C. Spermatozoa lysate was centrifuged at 15,000 g for 10 min at 4°C, and total protein concentration in lysate supernatant was measured by BCA assay. Testicular tissue and sperm cell lysate were stored at -20°C.

4.2.7 Immobilization of IgG antibodies on N-hydroxysuccinimide (NHS)-activated sepharose beads

Two in-house generated mouse monoclonal anti-TEX101 antibodies (34ED556 and 34ED229) that recognized different epitopes, and a non-specific mouse IgG (isotype control) were diluted to a final concentration of 1 mg/mL in coupling buffer (0.2 M NaHCO3, 0.5 M NaCl, pH 8.3). NHS-activated Sepharose 4 Fast Flow (GE Healthcare; 1 mL of bead slurry) was washed six times in ice-cold equilibration buffer (1 mM HCl), and was incubated with 1 mg/mL antibody solution for 2 hours at RT on a rotator. Sepharose beads were washed once with blocking buffer A (50 mM Tris-HCl, 1 M NaCl, pH 8.0), and twice with blocking buffer B (50 mM glycine, 1 M NaCl, pH 3.0). Samples were subsequently incubated in blocking buffer A for 2 hours at RT on a rotator, followed by serial washes with buffer A and B, and finally three times with binding buffer 1x TBS (50 mM Tris, 150 mM NaCl, pH 7.5).
4.2.8 Co-IP of TEX101 complexes in testicular tissue, spermatozoa and seminal plasma

Co-IP of TEX101 complexes from testicular tissue lysate (600 μg of total protein) was performed in triplicates with anti-TEX101 antibodies 34ED556 or 34ED229 and non-specific mouse IgG coupled to NHS-activated beads (50 μL). Co-IP of TEX101 complexes from spermatozoa lysate (120 μg total protein) was performed in triplicates with 34ED556 and non-specific mouse IgG coupled to NHS-activated beads (30 μL). Co-IP of TEX101 complexes from SP (600 μg total protein) was performed in triplicates with 34ED556 and non-specific mouse IgG coupled to NHS-activated beads (50 μL). Following binding for 2 hours at RT with shaking, all beads were washed with TBS binding buffer (50 mM Tris, 150 mM NaCl, pH 7.5) and 50 mM ammonium bicarbonate, re-suspended in 50 mM ammonium bicarbonate with DTT (5 mM final) and incubated for 1 hour at 60°C. Cysteines were then alkylated with 10 mM iodoacetamide for 1 hour at RT in the dark. Proteins were digested overnight by trypsin (0.5 μg) at 37°C, supernatants were collected and remaining beads were incubated again with 30% acetonitrile at RT for 10 min. First and second supernatants were pooled, and trypsin was inactivated by 1% TFA. L-methionine sodium salt (5 mM; Sigma-Aldrich) was finally added to each digest, to suppress oxidation of endogenous methionines.

4.2.9 Identification of TEX101 complexes by liquid chromatography - tandem mass spectrometry

Following digestion, peptides were extracted with C18 OMIX tips, eluted in 3 μL of 65% ACN in water with 0.1% formic acid and diluted with 57 μL water - 0.1% formic acid. Samples were loaded to the 96-well microplate (Axygen, Union City, CA), and were analysed by an EASY-nLC 1000 system (Thermo Fischer Scientific) coupled online to a Q Exactive™ Plus Hybrid Quadrupole-Orbitrap™ Mass Spectrometer (Thermo Fischer Scientific). Each immunoprecipitation full-process replicate was analyzed with an 18 μL single injection. Peptides in each sample were loaded and separated with a 15 cm C18 analytical column (inner diameter 75 μm, tip diameter 8 μm) using a 60-min LC gradient with increasing percentage of Buffer B at flow rate 300 nL/min (Buffer B: 0.1% formic acid in acetonitrile). A data-dependent mode was utilized to acquire a full MS1 scan from 400 to 1500 m/z in the mass analyzer at
resolving power of 70,000, followed by 12 precursor ions data-dependent MS2 scans at 17,500 resolution. Ions with charge states of +1, ≥ +4 and unassigned charge states were excluded from MS2 fragmentation.

4.2.10 Data processing and analysis

Raw files for each MS run were generated by XCalibur software (v. 2.0.6; Thermo Fischer Scientific). For protein identification and data analysis, mass spectra were processed with MaxQuant software (version 1.5.2.8). Protein search was performed against the non-redundant Human UniProtKB/Swiss-Prot database (HUMAN5640_sProt-012016). Search parameters included: trypsin enzyme specificity, 2 missed cleavage, minimum peptide length of 8 amino acids, minimum of 1 unique peptide, top 8 MS/MS peaks per 100 Da, peptide mass tolerance of 20 ppm for precursor ion and MS/MS tolerance of 0.5 Da, fixed modification of cysteines by carbamidomethylation and variable modification of methionine oxidation and N-terminal protein acetylation. False-discovery rate (FDR) was set to 1% both at the protein and the peptide levels. Label-free relative quantification of identified proteins was achieved by the MaxLFQ algorithm integrated into MaxQuant (241). The ‘proteinGroups.txt’ file, generated by MaxQuant, was uploaded to Perseus software (version 1.5.5.3) for further statistical analysis of the data (242). In the first step of data processing, protein identifications classified as “Only identified by site”, “Reverse”, and “Contaminants” were excluded. LFQ intensities were log2-transformed, and two groups with three replicates each were compared (LFQ-anti-TEX101 and LFQ-mouse IgG). Proteins with less than three valid values in at least one group were filtered out. Missing LFQ values were imputed with values representing a normal distribution to enable statistical analysis. A two-sample t-test (Benjamini-Hochberg false-discovery rate-adjusted p values) was applied to determine proteins statistically enriched by anti-TEX101 versus non-specific mouse IgG. We performed variance correction (s0) for each comparison, and we applied FDR of 1% for candidate selection. The constant for variance correction (s0) was incorporated in order to control the relative importance of the t-test p value and the enrichment (fold change). Volcano plots were generated for each comparison to facilitate data visualization. The list of putative TEX101-interacting proteins was merged with the Human Protein Atlas (v.13) secretome (n=2928) and membrane-bound proteome (n=5463), to select secreted and membrane-bound proteins expressed in testis (101). Expression and localization of each
candidate protein was manually assessed using Human Protein Atlas immunohistochemistry data and the NeXtProt database.

**4.2.11 Development of SRM assays for putative TEX101-interacting proteins**

Targeted SRM assays were developed for putative TEX101-interacting proteins as previously described (211, 212, 243). Our MS and MS/MS identification data (including potential post-translational modifications) was used to select proteotypic peptides. Peptides with 7-20 aa and without oxidation, deamidation or potential missed cleavages were selected. Selected peptides were also confirmed with SRM Atlas database (www.srmatlas.org). To facilitate accurate relative quantification, synthetic heavy isotope-labeled peptides were obtained for all proteins. Survey unscheduled SRM assays with all possible y- and b-ion fragments were prepared for light and heavy peptides and monitored in testicular tissue or spermatozoa lysates on triple quadrupole mass spectrometer (TSQ Quantiva™). Intensity and interferences were assessed for each transition, and three most intense transitions were selected for each heavy and light forms. Two separate multiplex SRM assays were finally developed for candidates identified in testicular tissues (20 heavy and light peptides for 9 candidates, and TEX101) and spermatozoa (20 heavy and light peptides for 9 candidates, and TEX101). All peptides were scheduled within 2-min intervals during a 30-min gradient (Appendix 4.1 and Appendix 4.2).

**4.2.12 Verification of TEX101-interacting proteins**

TEX101-interacting proteins were verified in pools of independent testicular tissue and spermatozoa lysates. Prior to trypsin digestion, 500 fmoles of heavy isotope-labeled TEX101 and DPEP3 proteotypic peptides with a trypsin-cleavable tag (AGTETAILATK*-JPTtag and SWSEEELQGVLRL*-JPTtag) were added to all samples. Eight heavy isotope-labeled peptides for TEX101 interactome in testicular tissue, and eight heavy peptides for TEX101 interactome in spermatozoa, were pooled and diluted to a final concentration of 100 fmol/µL. Five µL of the heavy peptide pool were spiked to each sample after digestion. Ten micrograms of total protein from the initial testicular tissue and spermatozoa lysate were digested, and heavy isotope-
labeled peptides were spiked in the digest in order to calculate the recovery of each protein after co-IP. Digests were desalted and loaded on a 96-well plate. Using an autosampler, 18 µL of each sample were injected into a C18 column, peptides were separated with a 30-min gradient and quantified by TSQ Quantiva™ mass spectrometer. Each sample was analysed in duplicates, and raw files were analyzed with Skyline software (v3.6.0.10493). The relative abundance of each endogenous peptide and corresponding protein was calculated according to the heavy-to-light ratio and the amount of the heavy peptides spiked in each sample.

4.2.13 Hybrid ELISA for the detection of TEX101-DPEP3 complex

Microtiter plates (96-well) were coated with anti-TEX101 antibodies (34ED556 or 34ED229; 500 ng per well). Following overnight incubation, plates were washed 3 times, and 100 µL of testicular tissue or spermatozoa lysates (prepared as previously described), or SP, were loaded on the plate. Two dilutions (10x and 4x for testicular tissue and spermatozoa lysate, and 100x and 10x for SP) in duplicates were used for each sample and each combination of antibodies. After 2 hours incubation with gentle shaking, plates were washed 3 times with PBS, and 100 µL of biotinylated anti-DPEP3 antibodies (40ED139 or 41ED68) were added to each well, and incubated for 1 hour. The plates were then washed with PBS, and streptavidin-conjugated alkaline phosphatase was added for 15 min. After the final 6-times wash with PBS, 100 µL of DFP solution in substrate buffer were added and incubated for 10 min with gentle shaking. Finally, 100 µL of developing solution were added in each well for 1 min, and time-resolved fluorescence was measured with Wallac EnVision 2103 Multilabel Reader (Perkin Elmer).

Reversed hybrid ELISAs were also performed simultaneously using anti-DPEP3 antibodies (40ED139 and 41ED68) for capture and biotinylated anti-TEX101 antibodies for the detection of TEX101-DPEP3 complexes. In addition, control experiments with non-specific mouse IgG (500 ng per well) for capture and all biotinylated anti-TEX101 or anti-DPEP3 antibodies for detection were performed simultaneously.
4.2.14 Assessment of O-sulfotyrosine modification in TEX101 protein

TEX101 protein was purified from testicular tissue lysate, spermatozoa lysate and SP using 34ED556 or non-specific mouse IgG coupled to beads. Beads were washed three times and were re-suspended in SDS-PAGE loading buffer (2x; BioRad, #1610737, Hercules, CA) and 5% β-mercaptoethanol. Samples were heated at 95°C for 15 min and centrifuged at 17,000 g for 10 min. Following that, proteins were resolved by 4-15% Mini-PROTEAN® TGX™ Precast Gels Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (BioRad, #4561021), and then transferred to 0.2 µm PVDF membranes (BioRad, #1620176) using the TransBlot Turbo Blotting System (BioRad). Original unpurified testicular tissue lysate, spermatozoa lysate and SP (10 µg total protein) were also included. After blocking, western blot analysis was performed with TEX101 (HPA041915, Sigma-Aldrich), and sulfotyrosine (sulfo-1C-A2) (Abcam, # ab136481) antibodies.

4.2.15 Immunocapture – LC-MS/MS with mouse monoclonal anti-sulfotyrosine antibody

Polystyrene plates were incubated overnight with 500 ng/well of mouse monoclonal anti-sulfotyrosine antibody (sulfo-1C-A2) in 50 mM Tris buffer (pH 7.8). Anti-TEX101 antibody 34ED556 and non-specific mouse IgG were also used as positive and negative controls, respectively. Plates were washed, and 100 µL of 10-fold diluted testicular tissue lysate (in 6% BSA), 10-fold diluted spermatozoa lysate or 100-fold diluted SP were added to each well. Samples were incubated for 2 hours at RT with gentle shaking, and plates were then washed with PBS (3x) and 50 mM ABC (3x). Afterwards, samples were prepared for mass spectrometric analysis. Peptides were extracted from solution and were analysed an EASY-nLC 1000 system coupled to Q Exactive™ Plus Mass Spectrometry, as described above. Raw files were processed with MaxQuant software (version 1.5.2.8), and protein search was performed against the Human UniProtKB/Swiss-Prot database (HUMAN5640_sProt-012016).
4.3 Results

4.3.1 Production of TEX101 protein and mouse monoclonal antibodies

The mature form of human TEX101 protein was expressed by transfected Expi293F cells. The culture supernatant was collected 72 hours post-transfection, when the highest protein yield was achieved (as assessed by Western blot analysis). The expression and purity of TEX101 protein was assessed by Coomassie staining SDS-PAGE and Western blot analysis using a commercial anti-TEX101 rabbit polyclonal antibody (Figure 4.1a), and were also confirmed by mass spectrometry. The majority of the expressed protein migrated around ~29kDa (band b) and ~35 kDa (band c), corresponding to the glycosylated TEX101 form, while the deglycosylated form was detected in lower molecular weight (~20 kDa, band a), in agreement with the predicted molecular weight based on the amino acid sequence (Figure 4.1a). The purified recombinant TEX101 was quantified by TEX101 ELISA, previously developed in house (210).

Mice immunization with the purified mature form of TEX101 generated 24 hybridoma colonies positive for IgG antibody secretion. Hybridoma screening by immunocapture-SRM revealed twelve hybridoma colonies producing antibodies that could capture native TEX101 in normal testicular tissue lysate. Six out of twelve colonies produced antibodies with higher affinity for native TEX101 protein (Figure 4.1b), and were subsequently expanded in serum-free media and purified using protein G columns. We showed previously that the commercial polyclonal ab69522 antibody could captured native TEX101 (210, 240). Immunocapture-SRM results confirmed our previous findings, and furthermore demonstrated that mouse monoclonal anti-TEX101 antibodies produced in house possessed higher affinity for native TEX101 than ab69522 (Figure 4.1b). To investigate if in house anti-TEX101 mAbs were directed against different epitopes on the target protein, we tested all possible antibody pairs in a sandwich format immunoassay. Results revealed two groups of matching antibody pairs, with each group targeting a different epitope on native TEX101 protein (Figure 4.1c). Production of high affinity mAbs against multiple epitopes on the native endogenous TEX101 protein allowed us to develop a coIP-MS platform, and investigate the human TEX101 interactome with increased coverage.
Figure 4.1 Production of mouse monoclonal antibodies against different epitopes of native TEX101 protein. (a) Recombinant TEX101 protein was expressed by Expi293F cells. SDS-PAGE and MS analysis confirmed the presence of purified TEX101 in the excised bands, marked by arrows (a-d). Majority of TEX101 migrated around 29-35 kDa. (b) Immunocapture-SRM facilitated the fast screening of hybridoma colonies and the selection of mouse monoclonal antibodies against native TEX101 protein, present in normal testicular tissue lysate. Lane (+) indicates the positive control (anti-TEX101 mouse polyclonal antibody ab69522). Asterisks mark the clones that had strong positive reaction with native TEX101. (c) Six mouse monoclonal anti-TEX101 antibodies were paired in a sandwich immunoassay, and two different epitopes for native TEX101 protein were revealed. Antibodies 34ED233, 34ED470 and 34ED556 were directed against Epitope A, while antibodies 34ED229, 34ED629 and 34ED604 were directed against Epitope B. Dotted lines in red represent the background signal of sandwich immunoassay.
4.3.2 Identification of TEX101 interactome by co-IP-MS

To develop a stringent procedure for identification of TEX101 interactome, we optimized our sample preparation protocol and included mAbs against different epitopes of TEX101 and negative control non-specific mouse IgGs. TEX101 interactomes were identified in testicular tissues, spermatozoa and SP.

Mild non-denaturing nonionic (NP-40 and Triton X-100) and zwitterionic (CHAPS) detergents previously used for the solubilisation of membrane proteins in PPI studies (113, 227, 244, 245) were tested for TEX101 isolation from testicular tissues. Following cryolysis, the highest recovery of TEX101 was achieved using CHAPS (1% w/v) for lysis and protein solubilisation, as assessed by ELISA (Appendix 4.3). Antibodies were coupled to NHS-activated sepharose beads which previously revealed higher yields and lower non-specific binding in IP experiments (246). Since antibodies and TEX101-interacting proteins could compete for the same epitope, we selected two mAbs 34ED229 and 34ED556 generated against different TEX101 epitopes, as assessed by ELISA pairing (Figure 4.1c). Co-IP-MS experiments resulted in identification and relative quantification of several hundred proteins in testicular tissues, spermatozoa and SP. Proteins identified with FDR≤1.0% were selected as putative TEX101-interacting proteins. Comparison of 34ED229 (160-fold enrichment of TEX101) and 34ED556 (616-fold enrichment of TEX101) in the testicular tissue lysate revealed the higher enrichment efficiency and higher yield of interacting proteins for 34ED556 (Figure 4.2a, Appendix 4.4). Thus, 34ED556 antibody was used for the enrichment of TEX101 complexes from spermatozoa and SP.

Overall, 108 proteins were identified in testicular tissues with 34ED229 at FDR≤1.0% and s0=0.27 (Appendix 4.4), and 135 proteins were identified with 34ED556 at FDR≤1.0% and s0=0.29 (Figure 4.2a). Lists of candidates were filtered for secreted and membrane-bound proteins using HPA and NextProt databases (39 and 75 proteins for 34ED229 and 34ED556, respectively). Examination of candidate expression in testicular germ cells narrowed down the number of proteins to 7 proteins for 34ED229 (Appendix 4.4) and 9 proteins for 34ED556 (Figure 4.2a and Table 4.1). Seven proteins (except DPEP3 and NT5E) were found in common for 34ED229 and 34ED556.
**Figure 4.2 Identification of TEX101 protein interactome by co-immunoprecipitation coupled with mass spectrometry.** Volcano plots (-log10(p-value) versus log2(fold change)) revealed proteins co-enriched with target protein TEX101 by anti-TEX101 monoclonal antibody (34ED556) compared to the mouse IgG negative control, from normal testicular tissue with active spermatogenesis (a), spermatozoa obtained from fertile individuals (b), and pre-vasectomy seminal plasma (c). Statistical analysis was performed from 3 biological replicates in Perseus. Fold change and p-value cut-offs are indicated by the hyperbolic selection curve in black (fold change was assessed for each dataset by the variance correction (s0), FDR-adjusted t test p values <0.01). Target protein TEX101 is plotted in red, membrane/secreted proteins significantly enriched are plotted in black, and significantly enriched membrane/secreted proteins expressed by testicular germ cells (candidate interactors of TEX101) are plotted in blue.
Table 4.1 List of candidate interacting partners of TEX101 in human testicular tissue, and validation of candidates by co-IP SRM.

**Candidate interacting partners of TEX101 in human testicular tissue**

<table>
<thead>
<tr>
<th>UniProt Accession</th>
<th>Gene Name</th>
<th>Shotgun Log2 fold change</th>
<th>SRM Log2 fold change</th>
<th>Tissue specificity</th>
<th>Localization</th>
<th>Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q9BY14</td>
<td>TEX101</td>
<td>9.3</td>
<td>7.6</td>
<td>Tissue-enriched</td>
<td>GPI-anchored</td>
<td>Yes</td>
</tr>
<tr>
<td>Q9BYF1</td>
<td>ACE2</td>
<td>5.4</td>
<td>4.2</td>
<td>Group-enriched</td>
<td>Transmembrane</td>
<td>Yes</td>
</tr>
<tr>
<td>P21589</td>
<td>NT5E</td>
<td>3.9</td>
<td>0.8</td>
<td>Tissue-enhanced</td>
<td>GPI-anchored</td>
<td>No</td>
</tr>
<tr>
<td>P05556</td>
<td>ITGB1</td>
<td>3.5</td>
<td>4.1</td>
<td>Expressed in all</td>
<td>Transmembrane</td>
<td>Yes</td>
</tr>
<tr>
<td>Q9H4B8</td>
<td>DPEP3</td>
<td>2.9</td>
<td>3.7</td>
<td>Tissue-enriched</td>
<td>GPI-anchored</td>
<td>Yes</td>
</tr>
<tr>
<td>Q13508</td>
<td>ART3</td>
<td>2.3</td>
<td>0.3</td>
<td>Group-enriched</td>
<td>GPI-anchored</td>
<td>No</td>
</tr>
<tr>
<td>P11279</td>
<td>LAMP1</td>
<td>1.9</td>
<td>4.1</td>
<td>Expressed in all</td>
<td>Transmembrane</td>
<td>Yes</td>
</tr>
<tr>
<td>P13987</td>
<td>CD59</td>
<td>1.8</td>
<td>3.5</td>
<td>Expressed in all</td>
<td>GPI-anchored</td>
<td>Yes</td>
</tr>
<tr>
<td>Q9UKY0</td>
<td>PRND</td>
<td>1.4</td>
<td>3.2</td>
<td>Tissue-enriched</td>
<td>GPI-anchored</td>
<td>Yes</td>
</tr>
<tr>
<td>P60033</td>
<td>CD81</td>
<td>1.4</td>
<td>4.4</td>
<td>Expressed in all</td>
<td>Transmembrane</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Co-IP-MS in spermatozoa using 34ED556 enriched TEX101 by 1,000 fold and identified 74 proteins at FDR≤1.0% and s0=0.60 (Figure 4.2b). Finally, 9 secreted and membrane-bound proteins were selected (Table 4.2). DPEP3, CD59 and LAMP1 proteins were common for tissues and spermatozoa lysates enriched with 34ED556 antibody. Comparison of candidates derived from tissues and spermatozoa suggested that 34ED229 antibody may share an epitope with TEX101-interacting proteins, and this competition could lead to the disruption of TEX101 complexes.

Co-IP-MS of soluble complexes in SP using 34ED556 enriched TEX101 by 282 fold and identified 7 secreted and membrane-bound proteins at FDR≤1.0% and s0=0.58 (Figure 4.2c). Additional examination of these proteins revealed that 3 proteins were of epididymal origin, while 4 proteins were localized to intracellular membrane compartments. None of these 7 proteins were found in testicular tissues and spermatozoa. We thus assumed that TEX101 complexes are disrupted in SP and TEX101 in SP is present as a monomer, which is in agreement with our previous findings (240).
Table 4.2 List of candidate interacting partners of TEX101 in human spermatozoa, and validation of candidates by co-IP SRM.

### Candidate interacting partners of TEX101 in human sperm cells

<table>
<thead>
<tr>
<th>UniProt Accession</th>
<th>Gene Name</th>
<th>Shotgun Log2 fold change</th>
<th>SRM Log2 fold change</th>
<th>Tissue specificity</th>
<th>Localization</th>
<th>Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q9BY14</td>
<td>TEX101</td>
<td>10.2</td>
<td>8.8</td>
<td>Tissue-enriched</td>
<td>GPI-anchored</td>
<td>Yes</td>
</tr>
<tr>
<td>Q9H4B8</td>
<td>DPEP3</td>
<td>6.0</td>
<td>1.6</td>
<td>Tissue-enriched</td>
<td>GPI-anchored</td>
<td>Yes</td>
</tr>
<tr>
<td>P12821</td>
<td>ACE</td>
<td>4.9</td>
<td>0.3</td>
<td>Tissue-enriched</td>
<td>Transmembrane</td>
<td>No</td>
</tr>
<tr>
<td>P13987</td>
<td>CD59</td>
<td>4.1</td>
<td>1.6</td>
<td>Expressed in all</td>
<td>GPI-anchored</td>
<td>No</td>
</tr>
<tr>
<td>P11279</td>
<td>LAMP1</td>
<td>4.1</td>
<td>2.0</td>
<td>Expressed in all</td>
<td>Transmembrane</td>
<td>Yes</td>
</tr>
<tr>
<td>Q1ZYL8</td>
<td>IZUMO4</td>
<td>3.9</td>
<td>0.8</td>
<td>Tissue-enriched</td>
<td>Secreted</td>
<td>No</td>
</tr>
<tr>
<td>Q9BS86</td>
<td>ZPBP</td>
<td>3.8</td>
<td>1.3</td>
<td>Tissue-enriched</td>
<td>Secreted</td>
<td>No</td>
</tr>
<tr>
<td>Q9HBV2</td>
<td>SPACA1</td>
<td>3.0</td>
<td>1.3</td>
<td>Tissue-enriched</td>
<td>Transmembrane</td>
<td>No</td>
</tr>
<tr>
<td>Q8TDB8</td>
<td>SLC2A14</td>
<td>2.2</td>
<td>0.3</td>
<td>Tissue-enriched</td>
<td>Transmembrane</td>
<td>No</td>
</tr>
<tr>
<td>Q8TDM5</td>
<td>SPACA4</td>
<td>1.2</td>
<td>1.5</td>
<td>Tissue-enriched</td>
<td>GPI-anchored</td>
<td>Yes</td>
</tr>
</tbody>
</table>

4.3.3 Verification of TEX101 interactors by co-IP-SRM

For the verification of TEX101 interactors that emerged by our discovery experiments, we used a targeted mass spectrometry-based approach. Two multiplexed SRM assays coupled to co-IP were developed for monitoring the candidate proteins in testicular tissue and spermatozoa, respectively. We used an independent set of testicular tissue samples obtained from individuals with active spermatogenesis, and an independent set of spermatozoa samples from fertile individuals. Sample lysis and co-IP were performed as previously described in the discovery phase. We measured by SRM TEX101 protein, 9 candidate interacting proteins in testicular tissue and 9 candidate proteins in spermatozoa, before and after immunoprecipitation by anti-TEX101 34ED556 mAb. Overall, 7 out of the 9 candidate interacting proteins were confirmed to be significantly co-enriched with TEX101 in testicular tissue (Figure 4.3a and Table 4.1), and 3 out of 9 candidates were confirmed in spermatozoa (Figure 4.3b and Table 4.2), by applying previously used cut-off values (fold change≥2, and p-value<0.01). Calculation of TEX101 concentration in the samples before and after immunoprecipitation showed nearly 55% and 70% recovery of TEX101 protein with 34ED556 antibody in testicular tissue and spermatozoa, respectively.
Figure 4.3 Verification of human TEX101 interactome by co-IP-SRM. Two multiple-SRM assays were developed for targeted relative quantification of candidate interacting proteins in testicular tissue (a), and spermatozoa (b). Plotting -log10(p-value) versus log2fold change revealed false-positive proteins and real interactors co-immunoprecipitated with TEX101 by anti-TEX101 monoclonal antibody (34ED556) compared to the mouse IgG negative control. Fold change cut-off of >2 and significance p-value of <0.01, as indicated by horizontal and vertical dotted lines in red, classified candidates as true interacting partners and false-positive identifications (two-tailed t-test).

4.3.4 Production of DPEP3 protein and mouse monoclonal antibodies against its different epitopes

Following careful examination of candidate proteins, we focused on dipeptidase 3 (DPEP3), a testis-specific GPI-anchored protein localized at the cell surface of testicular germ cells. DPEP3 expression pattern in human testicular germ cells was similar to TEX101, as assessed by HPA immunohistochemistry data (http://www.proteinatlas.org/ENSG00000141096-DPEP3/tissue).

The mature form of human DPEP3 (amino acids 36 to 463) was expressed in Expi293F cells, and DPEP3 expression and purity were assessed by mass spectrometry, Coomassie staining, SDS-PAGE and Western blot analyses with a commercial anti-DPEP3 rabbit polyclonal antibody (Appendix 4.5). Purified rhDPEP3 was quantified by SRM, and was used as an immunogen for the production of mouse mAbs. Resulting 8 IgG-secreting clones were screened by IP-SRM for their ability to capture rhDPEP3 and native DPEP3 in SP, and two clones were
selected (Appendix 4.6a), expanded in serum-free media and purified with protein G columns. Pairing these two anti-DPEP3 mAbs (40ED139 and 41ED68) in a sandwich format immunoassay showed that each antibody recognized a unique epitope of DPEP3 (Appendix 4.6b).

4.3.5 Validation of TEX101-DPEP3 complex by hybrid immunoassay

TEX101-DPEP3 hybrid immunoassay was developed to confirm the existence of TEX101-DPEP3 complexes. Two anti-TEX101 (34ED556 and 34ED229) and two anti-DPEP3 (40ED139 and 41ED68) mAbs directed against different epitopes were used as capture and detection antibodies, and *vice versa*. Hybrid ELISA confirmed TEX101-DPEP3 complexes in the testicular tissue and spermatozoa used for interactome discovery and in the independent pools of testicular tissues and spermatozoa. Hybrid ELISA also confirmed the absence of TEX101-DPEP3 complexes in SP.

Based on signal intensity, the most efficient pair consisted of 34ED556 (anti-TEX101) and 41ED68 (anti-DPEP3) antibodies (Figure 4.4a; 2 and 7). Combination of 34ED229 and 41ED68 resulted in a lower signal (Figure 4.4a; 4 and 8). Interestingly, combination of 34ED556 or 34ED229 (anti-TEX101) with 40ED139 (anti-DPEP3) resulted in the loss of specific fluorescent signal (Figure 4.4a; 1 and 5, 3 and 6). Hybrid ELISA with capture by non-specific mouse IgG and detection by all four biotinylated antibodies also revealed no specific fluorescent signal. Thus, hybrid ELISA confirmed the existence of TEX101-DPEP3 complexes in testicular tissues and spermatozoa, but not SP.
Figure 4.4 Detection and validation of TEX101-DPEP3 complex by hybrid immunoassay.

(a) Relative abundance of TEX101/DPEP3 complex in an independent set of testicular tissue, spermatozoa and seminal plasma samples, as determined by TEX101-DPEP3 hybrid ELISA. Various combinations of capture and detection mAbs were used in a sandwich format: (1) 34ED556–40ED139; (2) 34ED556–41ED68; (3) 34ED229–40ED139; (4) 34ED229–41ED68; (5) 40ED139–34ED556; (6) 40ED139–34ED229; (7) 41ED68 – 34ED556; (8) 41ED68–34ED229. The mean fluorescence signal of the two replicates was calculated for each dilution, 4-fold and 10-fold for testicular tissue and spermatozoa lysate, and 10-fold and 100-fold for seminal plasma. Error bars represent the standard deviation of the two replicates, and dotted lines in red represent the background signal of TEX101-DPEP3 hybrid ELISA (6% BSA, negative control).

(b) Schematic representation of monoclonal antibody binding to target proteins, TEX101 and DPEP3. (i) Hybrid ELISA by pairing 34ED556 and 41ED68 mAbs can efficiently capture TEX101-DPEP3 complex. (ii) Hybrid ELISA by pairing 34ED229 and 41ED68 mAbs demonstrates moderate efficiency in detecting TEX101-DPEP3 complex, possibly due to partially accessible TEX101 epitope by 34ED229 mAb. (iii) Capture or detection by 40ED139 mAb leads to absolute deterioration of signal, which is at the level of the background. This suggests that 40ED139 mAb binds to the same DPEP3 epitope with TEX101.
4.3.6 Assessment of tyrosine O-sulfation of TEX101 protein

Tyrosine O-sulfation has been previously identified as a common PTM of secreted and membrane-bound proteins involved in PPIs and improving PPI affinity (247). Sulfation has been previously implicated in the formation of ADAM protein complexes on sperm (112), which were also degraded in Tex101 knockout mice (113, 163). We hypothesized that same PTM may strengthen TEX101-DPEP3 interaction, and control the dynamics of TEX101 PPIs. Thus, the presence or absence of this PTM could explain the absence of a GPI-anchored TEX101/DPEP3 complex in SP.

We assessed tyrosine O-sulfation by IP of TEX101 from testicular tissues, spermatozoa and SP followed by immunoblotting with anti-sulfotyrosine or anti-TEX101 antibodies (Appendix 4.7). As a result, TEX101 was enriched, but not detected by anti-sulfotyrosine antibody. In addition, IP-MS using anti-sulfotyrosine antibodies did not identify TEX101 in testicular tissues, spermatozoa or SP. We thus concluded that TEX101 was not modified by tyrosine O-sulfation. Further investigation of proteins modified by tyrosine O-sulfation may reveal the role of this PTM in spermatogenesis and male fertility (248).
4.4 Discussion

During past decades, numerous knockout mouse models have been generated to investigate genes pertinent to spermatogenesis and fertilization (40). In many cases, male infertility phenotypes were associated with altered or disrupted PPIs involved in spermatogenesis, sperm maturation, migration, zona pellucida binding and sperm-oocyte fusion (113, 114, 168, 173, 226, 230).

However, little of this knowledge on mouse testis-specific proteins has been translated into human reproduction studies (232, 249), often due to the absence of human orthologs. For example, examination of testis-specific genes of ADAM family (involved in spermatogenesis in mice) revealed only 6 genes with human orthologs (Adam2, Adam18, Adam21, Adam29, Adam30, Adam32), while 12 mouse testis-specific genes (Adam1a, Adam1b, Adam3, Adam4, Adam5, Adam6a, Adam6b, Adam24, Adam25, Adam26a, Adam26b, Adam34) either did not have human orthologs or were pseudogenes which did not express proteins (250). Such discrepancy justifies the importance of studies on human testis-specific genes and proteins.

TEX101 is a prominent example of a highly testis-specific protein which is crucial for production of competent sperm and fertilization (113, 163). TEX101 function, as identified in mice studies, is exerted through PPIs with numerous cell-surface testis-specific proteins; the most prominent mouse TEX101-interacting proteins, however, are pseudogenes in human (Adam3, Adam4, Adam5, Adam6a and Adam6b). In present study, we focused on elucidating interactome of human TEX101 in testicular tissues, spermatozoa and SP.

We first optimized our co-IP-MS approach to ensure stringent identification of TEX101-interacting proteins. It should be noted that isolation of membrane proteins and their complexes, especially GPI-anchored proteins localized to cholesterol-enriched lipid rafts, may be challenging due to high hydrophobicity and detergent resistance (155). Here, we combined cryomilling of clinical samples with soft detergent-mediated lysis in order to release cellular components without disrupting the protein complexes. Substantially higher recovery of TEX101 from testicular tissues was achieved with CHAPS, as compared to Triton X-100 or NP-40 (Appendix 4.3). CHAPS sterol moiety could facilitate more efficient disruption of cholesterol-enriched lipid rafts and release of GPI-anchored complexes (156, 157).
Since commercial anti-TEX101 antibodies and our previously generated in-house antibodies against human TEX101 (210) could not efficiently enrich non-denatured TEX101 from tissues and SP, we produced here the second generation of anti-TEX101 mouse mAbs. Our co-IP-MS approach identified and validated physical protein interactions of TEX101 in testicular tissues and in mature spermatozoa. In addition, investigation of our candidates using the Contaminant Repository for Affinity Purification (251), showed that none of our candidate proteins were background contaminants. Interestingly, none of the testis-specific ADAM proteins (ADAM18, ADAM29 and ADAM30), potential orthologs of mouse ADAM3-6 proteins, were found to interact physically with TEX101, which may suggests either transient interactions between TEX101 and ADAMs, or alternative mechanisms of spermatozoa maturation in human. Identification of a human TEX101 knockout, or knockdown model, as well as more robust PPI studies involving protein cross-linking coupled to mass spectrometry may be used to capture transient PPIs missed by co-IP-MS approaches (252, 253).

Interestingly, none of candidates emerged as TEX101-interacting proteins after co-IP-MS of SP (Figure 4.2b), while all our candidates, including DPEP3, were previously identified in SP (99). This observation was in agreement with our previous size-exclusion chromatography data which revealed that TEX101 was present in the free soluble form in human SP (240).

Such difference could be the result of (i) slightly alkaline pH (pH ~7.8 - 8.0) of SP (254), thus weakening electrostatic interactions, (ii) loss of PTMs or altered localization, and (iii) proteolytic degradation of TEX101-interacting proteins. Here, we demonstrated the absence of TEX101 tyrosine-O-sulfation, a recognized PTM in extracellular PPIs (255) and interactions of testis-specific membrane proteins (112). We also generated mAbs against human DPEP3, and we developed a TEX101-DPEP3 hybrid ELISA, which was used to validate the presence of TEX101-DPEP3 complexes in testicular tissues and spermatozoa and confirmed the absence of complex in SP (Figure 4.4a).

Literature review on our TEX101-interacting proteins revealed that our top candidate DPEP3 has previously been shown to co-localize and form a physical complex with TEX101 on the surface of murine testicular germ cells (256). DPEP3 is the testis-specific membrane-bound protein of dipeptidase family (257). Similarly to TEX101, DPEP3 is a GPI-anchored protein expressed by testicular germ cells during spermatogenesis, and the majority of DPEP3 is shed into seminal plasma during sperm maturation (256). It was demonstrated that a fraction of
murine DPEP3 formed homodimers within the testis (256). Western blot analysis of human testicular tissue, spermatozoa, and SP detected two forms of DPEP3, a monomer and a homodimer, in testicular tissue and spermatozoa, while in SP, DPEP3 was mostly present as a homodimer (Appendix 4.8). The enzymatic activity of DPEP3 was demonstrated in vitro (257). The physiological role of TEX101-DPEP3 complex in mice remains unknown. However, numerous factors suggest the role of this complex in spermatogenesis: (i) TEX101 and DPEP3 form a physical complex on the surface of male germ cells; (ii) both TEX101 and DPEP3 are proteins with the ultimate tissue and male germ cell specificity; (iii) both TEX101 and DPEP3 are GPI-anchored proteins localized in the lipid rafts; (iv) Tex101 knockout mice are sterile, while Dpep3 knockout mice are not available as yet; (v) TEX101 has been shown as a crucial protein for maturation and processing of ADAM proteins directly involved in sperm transit and sperm-egg interaction. The latter function of TEX101 may be exerted by TEX101-DPEP3 complex, in which DPEP3 acts as a peptidase enzyme, while TEX101 acts as the modulator of DPEP3 maturation, localization or activity. Whether TEX101 protein forms a complex with the DPEP3 monomer, or homodimer, is still not clear and requires further investigation.

In addition, TEX101-DPEP3 hybrid ELISA demonstrated that not all mAbs against TEX101 and DPEP3 could capture TEX101-DPEP3 complex with the same efficiency. Combination of 34ED556 and 41ED68 mAbs was shown to be the most efficient for the detection of TEX101-DPEP3 complex (Figure 4.4b (i)). Similarly, when paring 34ED229 with 41ED68 mAb, TEX101-DPEP3 complex was detectable, although fluorescence was approximately 50% lower compared to 34ED556 and 41ED68 pair. Hybrid ELISA signal was lost when 40ED139 anti-DPEP3 mAb was used to capture, or to detect the complex. We speculate that mAbs like 34ED229 and 40ED139 may not have full access to TEX101 and DPEP3 epitopes, respectively, due to the formation of TEX101-DPEP3 complex (Figure 4.4b (ii) and (iii)). In future, TEX101-DPEP3 hybrid immunoassay can emerge as a simple, sensitive and high-throughput platform to screen for molecules which disrupt TEX101-DPEP3 complexes and may emerge as male contraceptives. As a proof of concept, we will use our hybrid ELISA to perform competitive binding assays, so as to assess the ability of mAbs against different TEX101 and DPEP3 epitopes, to disrupt the TEX101-DPEP3 complex.

To conclude, taking into account literature data and our present results, we suggest that molecules which disrupt TEX101-DPEP3 complexes may facilitate better understanding of the
role of TEX101-DPEP3 in spermatogenesis and fertilization, and lead to development of male contraceptives. With no male contraceptives available at the moment, the race for non-hormonal contraceptives continues. With only few targets and compounds investigated (13, 258-260), the most promising compounds were abandoned due to their severe side effects. We believe that our study may provide an additional target to develop male contraceptives. Disruption of PPIs by peptides and small molecules is challenging, but not impossible (261). Exclusive male germ cell specificity of TEX101 and DPEP3 may decrease the chance of side effects.
Chapter 5
Identification of a damaging variant rs35033974 associated with idiopathic male infertility through degradation of TEX101 protein and its transient interactome
Chapter 5

5 Identification of a damaging variant rs35033974 associated with idiopathic male infertility through degradation of TEX101 protein and its transient interactome

5.1 Introduction

It is estimated that 30-40% of infertile men are diagnosed with idiopathic infertility (15). A routine semen analysis only determines abnormalities of sperm morphology, motility and concentration (254). Genetic abnormalities contribute to 15-30% of the male infertility cases (262), therefore, additional karyotype, Y chromosome deletion and CFTR mutation analyses are offered to infertile patients (64). Besides large chromosomal aberrations, single nucleotide variations (SNVs) in genes that regulate spermatogenesis, sperm maturation or sperm-oocyte interaction may negatively affect male fertility and result in its idiopathic forms (82-84).

Approximately 500 infertile or subfertile mouse models have been generated and studied to date, with a large proportion of them being phenotypically normal. Targeted gene disruption in mice revealed defects in spermatogenesis, spermiogenesis, and sperm maturation, leading to defective function of mature spermatozoa and male infertility (81, 85). The mouse models revealed genes and mechanisms that regulate the reproductive function, and they have been thereafter pursued for their potential involvement in human infertility (86-90). However, only a fraction of these findings can be translated to humans. There are more than 1,000 highly testis-specific proteins in the human genome (101). Each corresponding testis-specific gene can carry multiple variants in the protein-coding regions, and some of which may relate to compromised sperm production, quality and function, thus explaining the high frequency of idiopathic infertility cases.

Recent advances in high-throughput DNA sequencing technologies have allowed the sequencing of the whole genome or exome of hundreds of thousands of individuals. This led to the generation of publicly available large-scale reference data sets of human genetic variations (263-265). The Genome Aggregation Database (gnomAD) (www.gnomad.broadinstitute.org) is currently the most comprehensive source of exome (123,136 individuals) and whole genome (15,496 individuals) sequencing data. Meanwhile, substantial improvement has been achieved in
mass spectrometry-based proteomic technologies. Global and targeted mass spectrometry-based approaches can be used to provide protein-level evidence of protein-coding variants (140). Detection of the variants at the protein level can define their effect on protein expression and function, and furthermore confirm the classification of a variant as ‘benign’ or ‘damaging’, and provide the clinical relevance of this finding.

Targeted Tex101 gene disruption in mice resulted in absolute male sterility, with no characteristic phenotypic sperm abnormalities, demonstrating its essential role in fertilization (113, 239). The presence of TEX101 was crucial for the trafficking and processing of other cell surface proteins essential for fertilization in mice, for example four members of the ADAM family (ADAM3-6) (113, 163, 239). However, the hypothesis that TEX101 is an ubiquitous cell-surface chaperone cannot be translated into humans, due to the absence of ADAM3-6 genes in the human genome (266). The major limitation when studying human reproduction is the lack of human germ cell lines which can be used to generate knockout models. Therefore, in order to study the function of TEX101 protein in human reproduction, we hypothesized that our cohort of patients may have a natural human TEX101 knockout or men with significantly down-regulated TEX101 expression (potential knockdown models). We have previously shown that TEX101 is low in a small population of men with high sperm count (Chapter 3). Here, we investigated the presence of missense TEX101 variants with negative effect on TEX101 protein expression and function, expecting to identify individuals to study the function of TEX101. We hypothesized that human TEX101 is involved in trafficking and maturation of other testis-specific proteins, thus its knockdown will result in the disruption of these interactions, misfolding or mislocalization and subsequent degradation. This approach may facilitate the identification of the transient interactome of TEX101, and the discovery of the human orthologs of ADAM3-6, and thus, elucidate molecular mechanisms behind some idiopathic infertility cases.
5.2 Materials and Methods

5.2.1 Study population and sample collection

A total number of 390 semen samples were collected with the informed consent from patients with the approval of the institutional review boards of Mount Sinai (approval #08-117-E) and University Health Network (#09-0830-AE). Samples were obtained from healthy fertile individuals before vasectomy, and individuals diagnosed with unexplained infertility and oligospermia. Clinical parameters are summarized in Table 5.1. Unexplained infertility group included men who were not able to father a pregnancy after one year of regular unprotected intercourse, despite the normal sperm concentration (>15 million/mL). It should be noted that there was no access to female factor evaluation. Oligospermia group included men with sperm concentration below 15 million/mL. After liquefaction, semen samples were centrifuged three times at 13,000 g for 15 min at RT. The SP and sperm cells fractions were separated, and stored at -80°C. Samples were analyzed retrospectively.

<table>
<thead>
<tr>
<th>Table 5.1 Clinicopathological variables of the patient cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Clinical Parameters</strong></td>
</tr>
<tr>
<td>Number of subjects</td>
</tr>
<tr>
<td>Age – Median (Range), years</td>
</tr>
<tr>
<td>Ethnic Group</td>
</tr>
<tr>
<td>African-Canadian</td>
</tr>
<tr>
<td>Asian</td>
</tr>
<tr>
<td>European</td>
</tr>
<tr>
<td>Hispanic</td>
</tr>
<tr>
<td>Indo-Canadian</td>
</tr>
<tr>
<td>Middle Eastern</td>
</tr>
<tr>
<td>Native-Canadian</td>
</tr>
<tr>
<td>Unspecified/Unavailable</td>
</tr>
<tr>
<td>Diagnosis (sperm count in mln/mL)</td>
</tr>
<tr>
<td>Pre-vasectomy</td>
</tr>
<tr>
<td>Unexplained infertility</td>
</tr>
<tr>
<td>Oligospermia</td>
</tr>
</tbody>
</table>
5.2.2 Extraction of sperm genomic DNA and TEX101 genotyping

Genomic DNA was extracted from sperm cells with QIAamp DNA Mini Kit (Qiagen Inc.) by following the manufacturer’s instructions. Briefly, sperm cells were washed twice with phosphate-buffered saline (PBS) (137 mM NaCl, 2.7 mM KCl, 10mM Na$_2$HPO$_4$, 1.8 mM KH$_2$PO$_4$). Cells were lysed in the presence of proteinase K. After binding to the membrane of QIAamp Mini spin columns, DNA was washed to remove residual contaminants and detergents, and was finally eluted. DNA purity and concentration were determined by using a NanoDrop 8000 Spectrophotometer (Thermo Scientific). DNA samples were amplified by conventional polymerase chain reaction (PCR) as follows: A set of oligonucleotide primers (forward 5’-ACAGGACTGAGACAGCCAT-3’ and reverse 5’-TCCAGGGTACCTGTGGTCTC-3’) were designed to amplify a fragment of TEX101 gene containing the polymorphism of interest (c.296G>T), and generate a 197 bp product. PCR was performed in a 60 µL reaction mix, by adding 50 ng of genomic DNA, 0.6 µL (1.2 U) of Phusion High-Fidelity DNA polymerase (Thermo Scientific), 10 µL of 5x Phusion HF Buffer, 200 µM deoxynucleoside triphosphates, and 0.5 µM of each primer on an Eppendorf Mastercycler thermal cycler. The PCR was performed with an initial denaturation step at 98°C for 1 min, followed by 40 cycles of denaturation at 98°C for 10 s, annealing at 64°C for 30 s, extension at 72°C for 30 s, with a final extension at 72°C for 7 min. The PCR products were confirmed by running on 1.5% agarose gel some representative sample and controls for each round of PCR. The resulting products were purified with QIAquick PCR Purification Kit (Qiagen) according to the manufacturer’s instructions. Sequencing of PCR products was employed for the genotyping analysis of Tex101 gene c.296G>T polymorphism in all 390 sperm cell samples for the identification of heterozygous and homozygous individuals. Sequencing was performed by The Centre for Applied Genomics, The Hospital for Sick Children, Toronto, Canada.

5.2.3 Sample preparation and Glu-C digestion

Mammalian recombinant TEX101 protein, previously produced in the Expi293F transient expression system and purified by anion exchange chromatography (Chapter 4), SP and spermatozoa samples were subjected to protein digestion for mass spectrometry analysis. Prior to digestion, sperm cell pellets were washed twice with PBS, were lysed with 0.1% RapiGest SF
(Waters, Milford, MA) in 50 mM ammonium bicarbonate (ABC), and were sonicated three times for 30 s. Cell lysates were then centrifuged at 15,000 g for 15 min at 4°C. Total protein of each sample was measured by bicinchoninic acid assay (BCA assay). Ten µg of total protein per sample were prepared in 50 mM ABC for proteomic analysis. Proteins were denatured and disulfide bonds were reduced by treating the samples with 0.05% RapiGest SF and 5 mM dithiothreitol (DTT) at 65°C for 30 min. Proteins were then alkylated with 10 mM iodoacetamide in the dark for 40 min at RT. Protein digestion was performed overnight at 37°C in the presence of sequencing grade Glu-C enzyme (Promega) (1:20, Glu-C: total protein) and 5% acetonitrile (ACN) for enhanced Glu-C activity. Digestion in the presence of ABC (pH 7.8) provided enzyme specificity and cleavage at the glutamic residues. RapiGest SF cleavage and Glu-C inactivation was achieved with 1% trifluoroacetic acid (TFA). L-methionine (5 mM final) was also added to the digests to prevent the oxidation of methionine residues. Two heavy isotope-labeled TEX101 peptides, AITIVQHSSPPGLIV*TSYSNYCE and AITIVQHSSPPVLIV*TSYSNYCE, for the wild-type and the G99V form, were pooled and diluted to a final concentration of 100 fmol/µL and 250 fmol/µL, respectively. Two µL of the heavy peptide pool were spiked in each sample after digestion. Digests were desalted and peptides were extracted from solution by C18 OMIX tips (Varian Inc., Lake Forest, CA). Peptides were eluted in 3 µL of 70% ACN in water with 0.1% formic acid, and 57 µL water with 0.1% formic acid were added in each sample. Samples were analyzed by an EASY-nLC 1000 system (Thermo Fischer Scientific) coupled online to a Q Exactive™ Plus Hybrid Quadrupole-Orbitrap™ Mass Spectrometer (Thermo Fischer Scientific).

5.2.4 Development of Parallel Reaction Monitoring assay for the analysis of wild type and G99V TEX101 protein forms

To monitor wild-type and G99V TEX101 protein forms, we selected two peptides which contain the amino acid substitution and correspond to each form. Peptides were generated by Glu-C protein digestion, and were the following: AITIVQHSSPPGLIVTSYSNYCE (m/z=1268.6) for TEX101 wild-type form, and AITIVQHSSPPVLIVTSYSNYCE (m/z= 1289.6) for TEX101 G99V form. For the accurate peptide identification and the exclusion of possible interferences, heavy isotope-labeled peptide internal standards were synthesized. The protein Basic Local
Alignment Search Tool (BLAST) (http://blast.ncbi.nlm.nih.gov/Blast.cgi) was used to confirm the uniqueness of the peptides of interest for TEX101 protein.

Initially, we used Glu-C digested mammalian recombinant TEX101 protein which was analyzed in a typical shotgun proteomics experiment on Q Exactive™ Plus Hybrid Quadrupole-Orbitrap™ Mass Spectrometer. Eighteen µL of sample were loaded and separated with a 15 cm C18 analytical column (inner diameter 75 µm, tip diameter 8 µm) using a 60-min liquid chromatography (LC) gradient with increasing percentage of Buffer B at flow rate 300 nL/min (Buffer B: 0.1% formic acid in acetonitrile). Full MS1 scan from 400 to 1500 m/z in the mass analyzer at resolving power of 70,000 was acquired in a data-dependent mode, followed by 12 precursor ions data-dependent MS2 scans at 17,500 resolution. Ions with charge states of +1, ≥+4, and unassigned charge states were excluded from MS2 fragmentation. The LC-MS/MS raw files were analyzed using the Proteome Discoverer™ software (Thermo Scientific, version 1.4.1.14) with SEQUEST and Mascot search algorithm and the human UniProtKB/Swiss-Prot database (HUMAN_sProt-07092014). Search parameters included: Glu-C enzyme specificity, 2 missed cleavage, fixed modification of cysteines by carbamidomethylation, and variable modification of methionine oxidation and N-terminal protein acetylation. The false discovery rate (FDR) was set to 1%. Subsequently, Glu-C digested recombinant TEX101 was analyzed by an unscheduled targeted PRM 60-min gradient method, in which all transitions of the endogenous peptide corresponding to the wild-type and G99V form of TEX101 were monitored. The same PRM analysis was repeated for a Glu-C digested SP pool to confirm the detection of the peptides of interest in the biological sample. In the next step, Glu-C digested SP pool was analyzed by an unscheduled PRM method with a 16-min gradient elution at 300 nL/min flow rate. A four-step gradient was used: 20% to 40% of buffer B for 8 min, 40% to 65% for 2 min, 65% to 100% for 2 min, and 100% for 4 min. In the final step of PRM method development, the selected heavy isotope-labeled peptides were spiked in SP pool sample, after digestion with Glu-C, and heavy and light endogenous peptides were included in an unscheduled PRM 16-min gradient method on Q Exactive™ Plus mass spectrometer. Calibration curves were built for the two heavy isotope-labeled peptides to determine the limit of detection. The final PRM assay settings were the following: in-source CID was set to 3.0 eV, MS transitions in the orbitrap were acquired with 17,500 resolving power at 200 m/z, AGC target was set to 3×10^6 with a maximum injection time of 100 ms, the isolation window was set to 2.0 m/z, the optimized collision energy was set to 27 and scan times were set to 100 ms. An additional endogenous TEX101
peptide, TAILATKGCICE (m/z=637.3), generated by Glu-C digestion, was monitored as a control for total protein analysis, sample handling, and digestion efficiency (Appendix 5.1).

5.2.5 Immunocapture-PRM measurements of TEX101 G99V in seminal plasma and spermatozoa

For the enrichment of total TEX101 protein from SP and spermatozoa lysate samples, an immunocapture protocol, using an in house anti-TEX101 mouse monoclonal antibody (mAb) (34ED556), was applied as previously described (Chapter 4). Briefly, protein G purified 34ED556 mAb, was immobilized on NHS-activated Sepharose 4 Fast Flow (GE Healthcare). Fifty µL of 34ED556-coupled beads were incubated with SP or spermatozoa lysate, diluted in 0.1% BSA, overnight at 4°C on a rotator. After binding, beads were washed (3x) with binding buffer 1x TBS (50 mM Tris, 150 mM NaCl, pH 7.5), followed by washing with 50 mM ABC. Then, proteins were subjected to Glu-C on-bead digestion, following the protocol described above. After overnight incubation, supernatants were transferred to clean tubes, and beads were resuspended in 30% ACN in water with 0.1% formic acid. Supernatants were combined, and samples were acidified by 1% TFA. Additionally, 5 mM of L-methionine and buffer A (water with 0.1% formic acid) were added to each digest to limit the oxidation of methionine residues and to adjust the final % ACN concentration, respectively. Two µL of the heavy peptide pool were spiked to each sample after digestion. Digests were desalted, and peptides were monitored using the optimized PRM method described above. Raw files were analyzed with Skyline software (v3.6.0.10493), and relative abundance of each endogenous TEX101 forms was calculated using the light-to-heavy peptide ratio and the concentration of the spiked in heavy peptides.

5.2.6 Human sperm sample selection and preparation for differential proteomic analysis

Sperm cell samples were obtained from two men diagnosed with oligospermia and two men diagnosed with unexplained infertility, all homozygous for TEX101 c.296G>T variant, expressing only the G99V TEX101 form. Fertile control men age matched were selected from
the pre-vasectomy group, and were all homozygous for wild-type TEX101 genotype, expressing the wild-type form of the protein. The clinical parameters of the 8 samples are summarized in Table 5.2. Before cell lysis, sperm pellets were washed (3x) with PBS. Cells were resuspended in 25 mM ABC with 0.1% RapiGest, and sonicated three times for 30 s. Lysates were centrifuged at 15,000 g for 15 min at 4°C to remove debris. Total protein concentration was measured using BCA assay. Proteins (225 µg total protein per sample) were denatured and reduced with 5 mM DTT at 65°C for 30 min, and then alkylated with 10 mM iodoacetamide in the dark at RT for 40 min. Samples were trypsin (Sigma-Aldrich) digested at a ratio of 1:20 (trypsin:total protein) overnight at 37°C. RapiGest was cleaved with TFA at 1% final concentration, and samples were centrifuged at 17,000 g for 15 min.

Table 5.2 Clinical parameters of wild type TEX101 and homozygous rs35033974 TEX101 men (G99V<sup>hh</sup>), which were subjected to differential proteomic analysis

<table>
<thead>
<tr>
<th></th>
<th>TEX101 Wild-type</th>
<th>TEX101 G99V (hh)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of subjects, N</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Age – Median (Range), years</td>
<td>36.5 (31 – 40)</td>
<td>29.5 (29 – 40)</td>
</tr>
<tr>
<td>Diagnosis</td>
<td>Pre-Vasectomy</td>
<td>Oligospermia Unexplained infertility</td>
</tr>
<tr>
<td>Fertility status</td>
<td>Fathered &gt;1 pregnancies</td>
<td>Assisted Reproduction (unknown outcome)</td>
</tr>
<tr>
<td>Sperm concentration (mln/mL)</td>
<td>&gt;15</td>
<td>2 - 30</td>
</tr>
<tr>
<td>TEX101 in SP (ng/mL)</td>
<td>8,000 – 12,500</td>
<td>3.5 - 47</td>
</tr>
</tbody>
</table>
5.2.7 Strong cation exchange chromatography (SCX) sample fractionation

Strong cation exchange chromatography was employed for peptide fractionation, to reduce sample complexity and achieve deep proteome analysis. Tryptic peptides were diluted in mobile phase A buffer (0.26 M formic acid in 10% ACN; pH 2-3) to a final volume of 500 µL, and then, they were loaded onto a 500 µL loop coupled to a PolySULFOETHYL A™ column (2.1 mm ID×200 mm, 5 µm, 200 Å, The Nest Group Inc., MA). Peptide fractionation was performed on an HPLC system (Agilent 1100) with an optimized 60-min three step gradient, and eluted peptides were monitored at 280 nm. Elution SCX buffer contained all components of mobile phase A with the addition of 1 M ammonium formate, and it was introduced at 5 min and increased to 15% at 25 min, then to 25% at 35 min, and finally to 100% at 50 min, with a flow rate of 200 µL/min. Twenty-seven fractions (400 µL each) were collected for each sample. Based on the absorbance profile, out of the 27 fractions, 13 (some fractions were pooled) were analysed by shotgun mass spectrometry. Peptides were extracted using OMIX C18 tips, and eluted in 3 µL 65% ACN in water with 0.1% formic acid, and 57 µL water with 0.1% formic acid were added in each sample.

5.2.8 Protein identification by liquid chromatography – tandem mass spectrometry (LC-MS/MS)

A total of 104 SCX fractions from eight individual sperm cell samples were loaded to a 96-well microplate (Axygen, Union City, CA), and were analyzed by an EASY-nLC 1000 system coupled to a Q Exactive™ Plus Hybrid Quadrupole-Orbitrap™ Mass Spectrometer. Two 18 µL injections were analyzed for each fraction. As described above, peptides were separated with a 15 cm C18 analytical column using a 90-min LC gradient with increasing percentage of Buffer B at flow rate 300 nL/min, as follows: 2 min at 1% - 5% Buffer B, 77 min at 5% - 35% Buffer B, 3 min at 35% - 65% Buffer B, 1 min at 65% - 100% Buffer B, and 3 min at 100% Buffer B. Full MS1 scan from 400 to 1500 m/z in the Orbitrap was acquired at resolving power of 70,000 in a data-dependent mode, followed by 12 precursor ions data-dependent MS2 scans at 17,500 resolution. Charge states of +1, ≥+4, and unassigned charge states were not subjected to MS2 fragmentation.
5.2.9 Data processing and analysis

XCalibur software (v. 2.0.6; Thermo Fisher Scientific) was utilized to generate raw files for each MS run. For protein identification and relative quantification, raw files of mass spectra were loaded and analyzed with MaxQuant software (version 1.5.2.8). The MaxQuant searches were performed against the non-redundant Human UniprotKB/Swiss-Prot database (HUMAN5640_sProt-072016). All peptide and protein identifications were filtered applying a false discovery rate of 1%, and all false positive entries were removed. Search parameters included: trypsin enzyme specificity, two missed cleavages permitted, minimum peptide length of 7 amino acids, minimum identification of one razor peptide, fixed modification of cysteines by carbamidomethylation, and variable modification of methionine oxidation and N-terminal protein acetylation. The mass tolerance of precursor ions was set to 20 ppm, and for fragment ions to 0.5 Da, and top 12 MS/MS peaks per 100 Da. MaxLFQ algorithm, integrated into MaxQuant, facilitated the label-free relative quantification of identified proteins (241). Parameters included: minimum count ratio of one, LFQ minimum number of neighbors was set to 3, and the LFQ average number of neighbors to 6 (default values). The generated ‘proteinGroups.txt’ file was uploaded to Perseus software (version 1.5.5.3) for further statistical analysis of the data (242). Initially, proteins classified as ‘Only identified by site’, ‘Reverse’, and ‘Contaminants’ were filtered out, and the LFQ intensities were log2-transformed. Subsequently, missing LFQ values were imputed from the normal distribution (down shift of 1.8 and distribution width of 0.45), and then the average LFQ intensity was calculated for the two technical replicates for each sample to enable statistical analysis. Two groups, TEX101 G99V and TEX101 wild type, were defined for further analysis with four replicates in each group. A two-sample t-test (Benjamini-Hochberg false discovery rate-adjusted p values) was applied to determine proteins statistically downregulated in the TEX101 G99V homozygotes group compared to the fertile control group. The constant for variance correction (s0) was calculated and set to 0.4, and FDR was set at 5%. Data visualization was achieved by the generation of a volcano plot. Furthermore, the Human Protein Atlas (HPA) (version 13) (101) secretome (n=2928) and membrane proteome (n=5463) (www.proteinatlas.org) were merged with the list of significantly downregulated proteins in TEX101 G99V group, previously generated by Perseus software. As a final filtering step, we selected proteins which were classified as testicular tissue enriched extracellular transmembrane, GPI-anchored or secreted, after manual search of HPA database.
5.2.10 Development of SRM assays

Our full MS data dependent MS/MS identification data, processed by MaxQuant, along with publically available experimental data in the SRM Atlas database (www.srmatlas.org), were used to select tryptic peptides for each protein of interest. Peptides with 7-20 amino acids, not prone to modifications (oxidation, deamidation), and not bearing missed cleavages were chosen. All peptides were analyzed with BLAST to ensure that all peptides were unique to each protein. Heavy isotope-labeled peptide internal standards were synthesized for each protein, to ensure accurate peptide identification and elimination of possible interferences. Initially, heavy isotope-labeled peptides were spiked in a trypsin digested sperm cell pool lysate, and heavy and light endogenous peptides were included into several survey SRM assays. In the first step of method development, all possible y and b ion fragments were monitored in unscheduled targeted mass spectrometry 30-min gradient method on TSQ Quantiva™ triple quadrupole mass spectrometer. The three most intense transitions were selected for each heavy and light peptide pair for the final SRM assay. In total, 20 heavy and light peptides were scheduled within 2 min intervals during a 30-min gradient in a single multiplex SRM assay for sperm cell lysate. SRM assays for TEX101 (240), and DPEP3 (Chapter 4) were incorporated in the aforementioned multiplex SRM assay (Appendix 5.2).

5.2.11 Protein quantification by SRM

The statistically significant downregulated proteins that were selected after applying specific criteria (as described above) were quantified by SRM in three sperm cell samples obtained from TEX101 G99V homozygous men and four fertile control men, used in the discovery phase. The experiment did not include one TEX101 G99V homozygote sperm sample which was fully consumed in the discovery experiments. Sperm cell sample preparation was performed as described above. Briefly, sperm cell lysate samples (ten µg of total protein) were subjected to protein denaturation and reduction with 0.05% RapiGest SF and 5 mM DTT at 65°C for 30 min, and alkylation with 10 mM iodoacetamide in the dark at RT for 40 min. After overnight digestion with trypsin at a ratio of 1:20 (trypsin: total protein), 1% TFA was added to the digest for trypsin inactivation and RapiGest cleavage. Five hundred fmoles of heavy isotope-labeled TEX101 and DPEP3 proteotypic peptides with a trypsin-cleavable tag (AGTETAILATK*-

119
JPTtag and SWSEEELQGVLR*-JPTtag) were added to each sample prior digestion. Eight heavy isotope-labeled peptides, corresponding to e proteins, were pooled and diluted to a final concentration of 100 fmol/µL. Five µL of the heavy peptide pool were spiked to each sample after digestion. Peptides were extracted from the solution and were loaded on a 96-well plate. Peptides were monitored in a single multiplex SRM assay with a 30-min LC gradient, with a 2 min intervals, in TSQ Quantiva™ mass spectrometer (Thermo Scientific). The parameters for SRM assay included: positive polarity, declustering and entrance potentials 150 and 10 V, respectively; ion transfer tube temperature 300°C; optimized collision energy values; scan time 20 ms; 0.4 and 0.7 Da full width at half maximum (FWHM) resolution settings for the first and third quadrupoles, respectively; and 1.5 mTorr argon pressure in the second quadrupole. Each sperm cell lysate was analyzed in two process technical replicates, and raw files were analyzed with Skyline software (v3.6.0.10493). The relative abundance of each endogenous peptide and corresponding protein was calculated according to the heavy-to-light ratio and the amount of the heavy peptides spiked in each sample.

5.2.12 Western blotting

Protein levels of TEX101, LY6K, ADAM29, and DPEP3 were assessed by Western blot analysis. Twenty µg of total protein from sperm cell lysate of one TEX101 c.296G>T homozygote, and one wild-type homozygote were loaded onto an SDS-PAGE gel (4% - 15%, Bio-Rad), and transferred onto PVDF membranes (Bio-Rad). Membranes were incubated with 5% blocking solution (skim milk powder in Tris buffer with 0.1% Tween) for 2 hours at RT. After blocking, membranes were incubated with rabbit polyclonal antibody against TEX101 and DPEP3 (Atlas Antibodies), sheep polyclonal antibody against LY6K (R&D Systems), and mouse monoclonal antibody against ADAM29 (Abnova Corporation), overnight at 4°C. Expression of GAPDH (Thermo Fisher Scientific) was used as an internal standard. The membranes were washed three times (8 min each wash) with TBST solution (Tris buffer with 0.1% Tween), and were then incubated with goat anti-rabbit, donkey anti-sheep, and goat anti-mouse secondary antibody conjugated to horseradish peroxidase (Jackson ImmunoResearch) for 1 hour at RT. After six times wash with TBST solution, proteins were detected with chemiluminescence substrate (GE Healthcare Life Sciences).
5.3 Results

5.3.1 Searching for missense and loss of function variants in TEX101 gene

The Genome Aggregation Database (gnomAD) (version 2.0) (http://exac.broadinstitute.org/) an extended version of Exome Aggregation Consortium (ExAC) (http://exac.broadinstitute.org/) (265), was used to examine the presence of human ‘knockout’ variants or missense and loss of function protein-coding variants in TEX101. The gnomAD data set provided a comprehensive catalogue of human variants by DNA sequence data of 123,136 exomes and 15,496 genomes from unrelated individuals of diverse geographic ancestry. A total of 153 missense and loss of function variants in TEX101 were previously identified and available in gnomAD. Among these variants, there were 131 missense SNVs leading to amino acid substitution in the protein sequence, and possible impact on protein structure and function. Interestingly, the TEX101 variant rs35033974, predicted as ‘damaging’, was present in the general population with an allele frequency of 8.4%, while all the other identified ‘damaging’ variants were classified as rare with frequencies <0.04%. The allele frequency spectrum of the rs35033974 variant varied among the different populations. It was more common for individuals of European ancestry (12.4% and 11.6% for non-Finnish and Finnish, respectively), less common for Latino (5.2%), Ashkenazi Jewish (5.8%), South Asian (3.3%) and African (2%), and rare for East Asian population (<0.00001%). Human TEX101 locus consisted of five coding exons (transcript ENST00000598265). The missense variant rs35033974 (c. 296 G>T) was localized within exon 4, and led to the substitution of glycine for valine at position 99 in the TEX101 protein sequence (Figure 5.1a). The impact of the G99V amino acid substitution on the stability and the function of TEX101 protein was assessed by two independent prediction algorithms, Polymorphism Phenotyping v2 (PolyPhen-2) (267) and Sorting Intolerant from Tolerant (SIFT) algorithm (268). TEX101 G99V variant was classified as ‘deleterious’ or ‘probably damaging’, and it could affect protein folding, function, or stability. In addition, alignment of TEX101 protein sequences from various mammals demonstrated conservation of the glycine residue (Figure 5.2), thus indicating that this position may be intolerant to substitutions. Prediction algorithm for aggregating regions in misfolded proteins (TANGO) indicates increased beta-aggregation tendency due to the presence of rs35033974 variant in TEX101 gene. Such aggregates are quickly directed to proteasomal degradation.
Figure 5.1 Identification of rs35033974 variant in TEX101 gene. (a) Schematic representation of TEX101 gene and protein sequence, showing the missense variant c.296 G>T, and the amino acid substitution p.99 G>V. (b) TEX101 DNA sequencing chromatograms of three representative individuals. Wild type (normal) TEX101 genotype (left panel); heterozygous TEX101 rs35033974 variant genotype (middle panel); homozygous TEX101 rs35033974 variant genotype (right panel). Red asterisk indicates the position of the polymorphism. (c) Scatter plot sperm concentration and TEX101 levels in SP of 189 sub-fertile men (diagnosed with oligospermia or unexplained infertility) of European origin. Wild type homozygous TEX101, heterozygous rs35033974 TEX101, and homozygous rs35033974 TEX101 individuals are plotted in grey, green and red, respectively. Sperm concentration (15 million/mL) and TEX101 concentration (65 ng/mL) cut-off values are indicated by dotted lines in black. Variant rs35033974 allele frequency is higher for the groups with TEX101<65 ng/mL (58.3% for upper left group, and 21% for lower left group), and lower for the groups with TEX101≥65 ng/mL (9.8% for lower right group, and 8.7% for upper right group).
Figure 5.2 Conservation of TEX101 residues. Alignment of TEX101 protein sequences from various mammals revealed that Glycine at position 99 (in red box) is highly conserved, and is present in all species, except Oryctolagus cuniculus (rabbit) and Bos taurus (bovine) (red asterisk).

5.3.2 TEX101 c.296 G>T variant is enriched in sub-fertile men with low levels of TEX101 protein in seminal plasma

Genotypes for the TEX101 variant rs35033974 (c. 296 G>T) in 390 fertile and sub-fertile men were determined by genomic DNA amplification and sequencing analysis (Table 5.1 and Figure 5.1b). TEX101 rs35033974 variant was identified in all three groups that were analyzed by sequencing. In the pre-vasectomy fertile group (N=37), we identified 4 men with heterozygous (GT) TEX101, in the unexplained infertility group (N=175) we identified 23 men with heterozygous (GT) and 2 men with homozygous (TT) TEX101, and in the oligospermic group (N=178) we identified 25 men with heterozygous (GT) and 2 men with homozygous (TT) TEX101. Overall, the variant allele frequency was higher in the sub-fertile group of oligospermia and unexplained infertility (8%) compared to the pre-vasectomy fertile group (5.4%).

Our cohort consisted mostly of men of European descent (~60%), while all the other populations were underrepresented. Therefore, we calculated the variant allele frequency for all European fertile and sub-fertile men, and we examined potential impact of this variant allele on
the concentration of TEX101 in SP, as previously assessed by our in house TEX101 ELISA (240). The TEX101 variant rs35033974 allele frequency was 12.5% for the European sub-fertile group, and 5.4% for the European fertile group. Based on the 37 normal fertile samples (all of European origin), we set the cut-off values at 15 million/mL and 65 ng/mL for sperm concentration and TEX101 protein concentration in SP, respectively. Based on sperm concentration and TEX101 concentration in SP, all European sub-fertile individuals were further sub-classified into four groups (Figure 5.1c): (i) sperm≥15 million/mL and TEX101<65 ng/mL; (ii) sperm<15 million/mL and TEX101<65 ng/mL; (iii) sperm<15 million/mL and TEX101≥65 ng/mL; (iv) sperm≥15 million/mL and TEX101 ≥65 ng/mL. Calculation of the variant rs35033974 allele frequency for each group demonstrated 58.3% and 21.0% variant allele frequency for the groups with TEX101<65 ng/mL and sperm concentration ≥, or <15 million/mL, respectively. Variant allele frequency was lower, 9.8% and 8.7%, for the groups with TEX101≥65 ng/mL and sperm concentration <, or ≥15 million/mL, respectively. We thus speculated that the presence of TEX101 variant rs35033974 allele may be the reason for low TEX101 concentration, especially in men who were homozygous for the variant allele.

5.3.3 Variant rs35033974 results in near-complete degradation of TEX101 G99V protein in germ cells

We developed a targeted mass spectrometry-based approach to monitor the wild type and the variant G99V form of TEX101 protein in SP and spermatozoa lysate in wild type normal, G99V heterozygotes and G99V homozygotes. Conventional protein digestion with trypsin was not suitable, due to the generation of a long tryptic peptide (38 aa). Therefore, we explored alternative proteases, such as endopeptidase Glu-C and neutrophil elasase. Eventually we optimized the digestion protocol by using endopeptidase Glu-C enzyme, and we developed a short PRM assay to monitor wild type and variant TEX101 peptides, and one additional endogenous TEX101 peptide as a control. Our method was sensitive enough to detect TEX101 protein in SP when the concentration of the protein was ≥1.5 µg/mL, as previously assessed by ELISA (240). To improve the sensitivity, we developed an immuno-PRM assay, based on the immunoenrichment of TEX101 protein by using an in house anti-TEX101 mouse monoclonal antibody (34ED556) (Chapter4) prior to PRM analysis (Figure 5.3a). Higher enrichment efficiency was achieved for TEX101 in spermatozoa than in SP. Monitoring an additional
endogenous TEX101 control peptide confirmed that the anti-TEX101 monoclonal antibody could bind equally well the wild type and the G99V form of TEX101 protein.

The immuno-PRM assay was then used to detect the wild type and variant G99V forms of TEX101 protein in spermatozoa lysate of the three individuals: one with wild type homozygous TEX101 (GG), one with heterozygous TEX101 (GT), and one with homozygous (TT) TEX101 (Figure 5.3b). As expected, only the wild type TEX101 peptide and the variant G99V peptide were detected in wild type homozygous (GG) and the variant homozygous (TT) spermatozoa samples, respectively. Immuno-PRM analysis of spermatozoa lysate obtained from an individual with heterozygous TEX101 (GT) led to the detection of both protein forms, wild type and variant G99V. Interestingly, the relative abundances of the two forms were not comparable, with the level of the G99V variant form being significantly lower than the level of wild type TEX101. The relative abundance of the wild type and variant endogenous peptides was calculated according to the heavy-to-light ratio and the amount of heavy peptides spiked in the heterozygous TEX101 sample. It was estimated that the variant G99V TEX101 form represented <3% of the total TEX101 protein in this spermatozoa sample. It should be noted that rs35033974 is a germline variant, therefore both alleles were expected to be equally expressed to corresponding protein forms. We thus assumed that the G99V variant TEX101 form is misfolded (may be even aggregated) and then destroyed by proteasomal degradation. This could explain the low levels of TEX101 in men who were homozygotes for the rs35033974 variant.
a

(Sperm lysate) → TEX101 Immuno-enrichment → PRM Q Exactive Plus

Wild type (GG)   Heterozygous (GT)   Homozygous (TT)

- Endogenous wild type peptide (m/z=1288.6230) and internal standard (m/z=1271.6300)

- Endogenous G99V peptide (m/z=1289.6465) and internal standard (m/z=1292.6534)

- Endogenous control peptide (m/z=637.3447) for wild-type and G99V protein forms
5.3.4 Proteomic profiling of spermatozoa reveals testis-specific proteins downregulated in TEX101 rs35033974 homozygous men

The four sub-fertile men, homozygous for the rs35033974 TEX101 variant, were considered as TEX101 knockdown models. To investigate whether other spermatozoa-expressed proteins were affected by the ‘near-absence’ of TEX101 protein, we performed global proteomic analysis of four homozygous rs35033974 men, and four homozygous wild type TEX101 men with confirmed fertility (Figure 5.4a). To reduce sample complexity, and achieve deep proteome coverage, the samples were subjected to peptide fractionation by SCX prior to mass spectrometry analysis. Using MaxQuant software, we were able to identify and quantify 8,046 proteins with one peptide hit, and FDR≤1.0%. After statistical analysis, 100 proteins were found to be downregulated (FDR≤5.0% and s0=0.4) (Appendix 5.3). Filtering for membrane-bound and secreted proteins using HPA and NextProt databases resulted in a list of 51 proteins. Further examination of protein expression in testicular germ cells narrowed down the number of proteins to 14 testis-specific proteins (Figure 5.4b). Likewise, 99 proteins were upregulated (FDR≤5.0% and s0=0.4) in the TEX101 rs35033974 homozygotes, however, none was
membrane, or secreted and testis-specific. We focused on testis-specific proteins as the most interesting candidates potentially directly involved in sperm-egg interaction.

**Figure 5.4** Identification of candidate proteins downregulated in *TEX101 rs35033974 homozygous men*. (a) Spermatozoa samples from four homozygous rs35033974 men and four homozygous wild type TEX101 men with confirmed fertility were digested, fractionated by SCX-HPLC, and analysed by LC-MS/MS. (b) In total 8,046 proteins were identified and quantified. Volcano plot (-log10(p-value) versus log2(fold change)) reveals proteins downregulated in the ‘near-absence’ of TEX101 protein in the G99V homozygote group compared to the control group. Fold change and p-value cut-offs are indicated by the hyperbolic selection curve in black (fold change was assessed by the variance correction (s0=0.4), FDR-adjusted t-test p values <0.05). Target protein is plotted in red, testis-specific, membrane-bound proteins are plotted in blue, and testis-specific, secreted proteins are plotted in green. Proteins found significantly upregulated in G99V homozygote group were intracellular proteins, therefore we assume that they are false positive.
In our spermatozoa proteome, we identified all seven human testis-specific ADAM proteins, four of which were known to have adhesion activity (ADAM2, ADAM18, ADAM29 and ADAM32) (269). It should be noted that *Tex101* knockout mice were infertile due to the disruption and degradation of ADAM protein complexes involved in sperm transit and sperm-egg interaction (113, 163). We hypothesized that proteomic analysis of the *TEX101* knockoutdown individuals (homozygous for rs35033974), would reveal proteins which are degraded in the absence of TEX101 protein. Our goal was also to discover the human orthologs of ADAM3-6, assuming that, likewise murine ADAMs, human orthologs would be degraded in *TEX101* knockdown sperm cells. Out of the four identified ADAMs in our human sperm proteome, ADAM18, ADAM29, and ADAM32 were potential candidate orthologs for murine ADAMs. ADAM2 protein is also expressed in mouse germ cells, and is crucial for fertilization (109), but it was not affected by *Tex101* knockout (113, 114, 163). Based on data analysis and applied selection criteria, none of the three candidate human ortholog ADAM proteins were identified among the 14 testis-specific downregulated candidate proteins. Examination of each one of the ADAM proteins showed that only ADAM29 level was affected by the presence of rs35033974 variant (Appendix 5.4). Although ADAM29 was not significant (*p* value=0.003 and Fold change=2.5), it was included in the list of candidate proteins to be verified by SRM and Western blot analysis.

### 5.3.5 Verification of testis-specific downregulated proteins in *TEX101 rs35033974* homozygous men by SRM and Western blotting

After assessing the 14 testis-specific protein candidate list, we selected 8 proteins that were previously studied and involved in spermatozoa migration (114), zona pellucida binding/penetration (270-272), and sperm–oocyte fusion (230, 273). To verify the candidate proteins, we used a targeted mass spectrometry assay. An SRM assay was developed to monitor TEX101 and 8 candidate proteins identified after statistical analysis. DPEP3 protein, which was previously shown to interact with TEX101 (Chapter 4), but it was not affected by *TEX101* rs35033974 variant, was also included in the SRM method. Overall, 7 out of the 8 candidate proteins were confirmed to be downregulated in homozygous rs35033974 men, while DPEP3 level was not significantly different between homozygous rs35033974 and homozygous wild type men (Figure 5.5a and Table 5.3). Lymphocyte antigen 6 complex, locus K (LY6K) was the protein
that was significantly \((p \text{ value}<0.05, \text{Fold change}=14)\) affected by the absence of TEX101 in homozygous rs35033974 men. LY6K is a testis-specific GPI-anchored protein localized at the cell surface of testicular germ cells, with similar expression pattern to TEX101, as assessed by immunohistochemistry data in HPA. Thus, we decided to further validate LY6K \textit{in vitro} using western blot analysis. Western blot confirmed undetectable levels of TEX101 and LY6K proteins in homozygous rs35033974, while DPEP3 expression was maintained at normal levels (Figure 5.5b).

**Figure 5.5** Verification of testis-specific membrane-bound and secreted downregulated proteins in three TEX101 rs35033974 homozygous and four TEX101 wild type men by SRM and western blotting. (a) A multiplex SRM assay was developed for the targeted relative quantification of TEX101 and 8 candidate downregulated proteins. Plotting -log10 (p-value) versus -log2 ratio revealed false-positive and true-positive downregulated proteins in the ‘near-absence’ of TEX101 protein. Fold change cut-off of >2 and significance p-value of <0.05, are indicated by horizontal and vertical dotted lines in red (one-tailed t-test). (b) Western blot analysis of one TEX101 wild type sperm sample, and one homozygous G99V sperm sample confirmed the reduced levels of TEX101, LY6K and ADAM29, while DPEP3 protein was present as a monomer and homodimer, and its level was not affected. GAPDH was used as a loading control for total protein.
Table 5.3  List of downregulated testis-specific proteins in spermatozoa of rs35033974 homozygous men, as discovered by shotgun MS and verified by SRM.

Downregulated testis-specific proteins in TEX101 G99V<sup>hh</sup> men

<table>
<thead>
<tr>
<th>UniProt Accession</th>
<th>Gene Name</th>
<th>Shotgun Log2 fold change</th>
<th>SRM Log2 fold change</th>
<th>Tissue specificity</th>
<th>Localization</th>
<th>Verification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q9BY14</td>
<td>TEX101</td>
<td>-5.0</td>
<td>-4.2</td>
<td>Tissue-enriched</td>
<td>GPI-anchored</td>
<td>Yes</td>
</tr>
<tr>
<td>Q17RY6</td>
<td>LY6K</td>
<td>-5.5</td>
<td>-3.8</td>
<td>Tissue-enriched</td>
<td>GPI-anchored</td>
<td>Yes</td>
</tr>
<tr>
<td>P10323</td>
<td>ACR</td>
<td>-2.9</td>
<td>-3.6</td>
<td>Tissue-enriched</td>
<td>Extracellular/secreted</td>
<td>Yes</td>
</tr>
<tr>
<td>Q6X784</td>
<td>ZPB2</td>
<td>-2.9</td>
<td>-1.2</td>
<td>Tissue-enriched</td>
<td>Extracellular/secreted</td>
<td>No</td>
</tr>
<tr>
<td>Q8IXA5</td>
<td>SPACA3</td>
<td>-2.6</td>
<td>-2.1</td>
<td>Tissue-enriched</td>
<td>Extracellular/Transmembrane</td>
<td>Yes</td>
</tr>
<tr>
<td>Q5VZ72</td>
<td>IZUMO3</td>
<td>-2.6</td>
<td>-1.2</td>
<td>Tissue-enriched</td>
<td>Extracellular/Transmembrane</td>
<td>Yes</td>
</tr>
<tr>
<td>P38567</td>
<td>SPAM1</td>
<td>-2.6</td>
<td>-3.3</td>
<td>Tissue-enriched</td>
<td>GPI-anchored</td>
<td>Yes</td>
</tr>
<tr>
<td>Q8NEB7</td>
<td>ACRBP</td>
<td>-2.6</td>
<td>-2.9</td>
<td>Tissue-enriched</td>
<td>Extracellular/secreted</td>
<td>Yes</td>
</tr>
<tr>
<td>Q9UKF5</td>
<td>ADAM29</td>
<td>-1.3</td>
<td>-1.2</td>
<td>Tissue-enriched</td>
<td>Extracellular/Transmembrane</td>
<td>Yes</td>
</tr>
<tr>
<td>Q9H4B8</td>
<td>DPEP3</td>
<td>-0.2</td>
<td>0.4</td>
<td>Tissue-enriched</td>
<td>GPI-anchored</td>
<td></td>
</tr>
</tbody>
</table>
5.4 Discussion

Advances in high-throughput DNA sequencing technologies have facilitated the generation of large public catalogues of human genotypes and variations. Genome-wide association studies have been employed to identify variants that are implicated in complex diseases, like type 2 diabetes (274, 275), Crohn disease (274, 276), Alzheimer’s (277) and Parkinson’s disease (278, 279), and rheumatoid arthritis (280). Although several gene-targeted sequencing studies have been conducted, and various SNVs were associated with male infertility (86-88, 281-284), most of the associations failed further validation. This is mostly due to two reasons. First, most studies focus on a single gene, while there is a large number of highly testis-specific genes (>1,000) which are involved in spermatogenesis (25), indicating the male infertility is a multi-factor disease. Second, ethnic variability, heterogeneity in groups, and small sample size lead to underpowered studies. Despite these limitations, SNVs in MTHFR, GSTM1, FSHB, and AHRR genes have been reproducibly associated with male infertility across multiple studies (285-287). However, SNVs in these genes provided only associations with male infertility, but not detailed molecular mechanisms.

The established genetic causes of male infertility include the azoospermia factor (AZF) deletions of the Y chromosome (288), the Klinefelter’s syndrome (289), and mutations in the CFTR gene (290). The first two genetic tests can identify men with spermatogenic impairment (82), while the third test can identify men with physical obstruction in the reproductive tract, due to congenital defects in epididymis and vas deferens (54). Although these tests can identify a minority of infertile men, they still have very high predictive value for the success rate of sperm retrieval (64, 291, 292). Overall, only a small percentage of men display severe infertile phenotypes. Genomic analysis of sub-fertility is more complex, and SNVs in multiple genes may be involved in spermatogenesis.

In this study, we focused on a testis-specific gene, TEX101, which was previously shown to be essential for the production of fertilization-competent sperm in mice (113, 163, 256, 293, 294). It is likely that disruption of human TEX101 gene may lead to the expression of non-functional protein with potential effect on sperm quality and function. However, TEX101 SNVs have never been evaluated for association with male infertility or sub-fertility in previous genome-wide or gene-targeted association studies. Our study is the first to investigate the presence of missense variants in TEX101, and their effect on the expression and function of TEX101 protein. Our
ultimate goal was to discover a human TEX101 knockout or knockdown model that could be further used to investigate the function of human TEX101 protein, since in vitro human germ cell models are not available. Towards our aim, we first identified human TEX101 SNVs in publically available databases, and following in our cohort of fertile and sub-fertile individuals, and we reported an increase in the variant allele frequency in the group of men with low TEX101 concentration (<65 ng/mL).

Unlike previous studies that focused only on the identification of SNVs with significant susceptibility factor for different types of infertility, we investigated the effect of TEX101 variant rs35033974 on the translated protein of spermatozoa. Immuno-PRM analysis of spermatozoa from a TEX101 rs35033974 heterozygote individual confirmed the presence of wild type and G99V TEX101, with the variant G99V form being 3% or less of the total TEX101 present in the sperm sample. It has been shown elsewhere, that SNVs can lead to protein misfolding and aggregation, rapid degradation in the endoplasmatic reticulum (ER), and lower concentration of the protein at its destination (295-297). Typical example is the CFTR molecules that are misfolded due to the presence of mutations that interfere with CFTR protein maturation in the ER, leading to cystic fibrosis (298). Similar mechanisms are activated in protein aggregation diseases (Huntington’s, Alzheimer’s, and Parkinson’s diseases), in order to destroy aberrant proteins (299). Proper folding and trafficking is crucial for GPI-anchored proteins to be fully functional and targeted to the plasma membrane (300). In our case, introduction of valine, with a larger and more hydrophobic side chain compared to glycine, may compromise the required flexibility for TEX101 protein to acquire its native conformation (301, 302) and also lead to aggregation of hydrophobic domains. The fact that glycine at position 99 is highly conserved in mammals indicates that it may not be tolerant for substitutions. In contrast to TEX101 rs35033974 homozygotes, TEX101 protein level was >65 ng/mL for 70% of the TEX101 rs35033974 heterozygous men, showing that the expression of the wild type allele can compensate for the degradation of G99V protein form. Notably, heterozygous Tex101 +/- knockout mice were fertile, in contrast to homozygous Tex101/- knockout mice, which were sterile (113).

TEX101 rs35033974 homozygous individuals were selected as human TEX101 knockdown model for the identification of the transient interactions of TEX101 that may occur post-translation, during protein trafficking to the extracellular membrane, or during sperm maturation
in the epididymis. In-depth proteomic analysis of spermatozoa ensured the identification of low-abundance candidate proteins that might have been missed with the co-IP approach. This highlights the advantage of having a human TEX101 knockdown model to study the function of the protein.

Combination of literature and our generated data revealed LY6K protein as the top candidate to be affected in the TEX101 knockdown sperm. LY6K protein, similarly to TEX101, is expressed by testicular germ cells during spermatogenesis, and it is shed into seminal plasma during sperm maturation (114). It has been previously shown in mice that TEX101 and LY6K physically interacted in testis (294), and this interaction was crucial for the trafficking and the post-translational processing of both GPI-anchored proteins (239). Tex101−/− and Ly6k−/− mice were infertile due to compromised migration of sperm in the oviduct (113, 114, 239). Interestingly, DPEP3, a testis-specific GPI-anchored protein that was previously shown to form a physical complex with TEX101 (Chapter 4), was not affected in the TEX101 knockdown sperm of rs35033974 homozygotes. It is likely that the role of TEX101-DPEP3 complex is not to target both proteins to the membrane, and regulate their post-translational processing, but it may be involved in sperm maturation after the two proteins are anchored and associated on the cell surface.

Our human sperm proteome contained seven testis-specific ADAM proteins with adhesion (ADAM2, ADAM18, ADAM29, ADAM32) and metalloprotease (ADAM20, ADAM21, ADAM28, ADAM30) activity (269). We showed that ADAM29 protein level decreased in the human TEX101 knockdown sperm by SRM and western blot analysis. It is likely that ADAM29 is the human ortholog of one of the ADAMs, ADAM3-6, however there is not as yet data supporting that ADAM29 is involved in human spermatogenesis and fertilization, since its function is still unknown.

In the present study, we were not able to associate the rs35033974 TEX101 variant to male infertility due to insufficient clinical data. We demonstrated that the variant allele frequency is higher in the subfertile population, and we identified four men with homozygous rs35033974 TEX101 and low TEX101 concentration. All four individuals were diagnosed with either unexplained infertility or oligospermia, and subsequently they were referred to a fertility clinic, where the couples received assisted reproductive treatment, such as intrauterine insemination (IUI), in vitro fertilization (IVF), or intracytoplasmic sperm injection (ICSI). Due to the lack of
assisted reproduction outcomes for this set of samples, semen samples are currently prospectively collected, accompanied by clinical data for both partners and pregnancy outcomes. Two diagnostic tests, hemizona assay (HZA) (303) and sperm penetration assay (SPA) (304), can be used to predict the fertilization potential of TEX101 knockdown sperm compared to the TEX101 wild type sperm, as well as the outcomes of IUI and IVF (305).

To conclude, in the present study we identified a ‘damaging’ TEX101 missense variant in a cohort of fertile and sub-fertile men, and we demonstrated G99V TEX101 protein degradation. We used this human TEX101 knockdown model to investigate the transient interactions of TEX101, and we identified and verified two testis-specific proteins, LY6K and ADAM29, with significant reduction in their concentrations in TEX101 knockdown spermatozoa. The concentration of G99V TEX101 protein in SP of homozygotes varied (3.5 - 47 ng/mL), showing that the degradation rate of the variant form of TEX101 may differ among individuals. It is still not clear if low level of G99V TEX101 protein can maintain fertility. Thus, the concentration of residual G99V TEX101 in homozygotes may be a true biomarker to predict the success of fertilization. Measurement of TEX101 protein concentration, and rs35033974 TEX101 genotyping could be implemented in the evaluation of unexplained infertility and mild oligospermia (10-15 million sperm/mL) cases, in which routine semen analysis is normal. Finally, if disruption of TEX101 gene in humans results in infertility, similarly to mice, TEX101 protein may be a future target for the development of male contraceptives.
Chapter 6
Summary and Future Directions
Chapter 6

6 Summary and Future Directions

6.1 Overview

This thesis presents a comprehensive investigation of a human testis-specific protein, TEX101, and incorporates basic research and translational components. The overall objective was to examine the clinical and functional aspects of TEX101 protein in humans. From a clinical perspective, TEX101 was previously shown to be a non-invasive biomarker of azoospermia, and its concentration in SP was able to differentiate NOA from OA patients when it was combined with the concentration of ECM1 protein (97). The present work is a follow-up study that aims to translate this discovery into routine clinic practice. First, we generated monoclonal antibodies, and we developed a first-of-a-kind TEX101 ELISA for measuring TEX101 concentration in biological fluids (Chapter 2). Further, we evaluated the performance of TEX101 ELISA in a large cohort of SP samples from the whole range of male infertility conditions, from severe azoospermia to oligospermia and normal spermatogenesis, and we proposed the clinical utility of TEX101 as a non-invasive SP biomarker (Chapter 3). From a functional perspective, this is the first study to investigate the function of human TEX101 protein. We used two mass spectrometry-based approaches to identify candidate interacting partners of TEX101 in human testicular tissue and spermatozoa. Our first approach using a co-immunoprecipitation coupled to mass spectrometry platform allowed us to identify candidate proteins that interacted with TEX101 on the cell surface of testicular germ cells and spermatozoa. Specifically, DPEP3, a testis-specific dipeptidase, was found to form a GPI-anchored complex with TEX101 in germ cells and spermatozoa. The presence of the complex was further validated by a mass spectrometry-independent approach (Chapter 4). Our second approach utilized a human TEX101 knockdown model, which was identified in this study. The presence of the homozygous missense TEX101 variant rs35033974 led to TEX101 protein degradation, generating a human TEX101 knockdown model, an essential and unique tool to identify the transient interactome of TEX101, and to study its function. This approach utilized global proteomic analysis of the human TEX101 knockdown spermatozoa, and identified testis-specific membrane-bound, or secreted proteins that were downregulated in the absence or significantly low amounts of TEX101 (Chapter 5).
6.2 Summary of Key Findings

The main goals of this study were:

1) Translate TEX101 infertility biomarker into clinical practice, by developing a sensitive TEX101 immunoassay, and evaluate the clinical utility of TEX101 diagnostic test in the context of male infertility;

2) Identify and validate the physical interactome of human TEX101 protein in testicular tissue, SP, and spermatozoa; and

3) Identify and verify the transient interactome of human TEX101 in spermatozoa, utilizing a human TEX101 knockdown model.

An overview of methodology and the key findings for each objective are summarized below:

1) Translation of a previously discovered TEX101 infertility biomarker into clinical practice, by developing a sensitive TEX101 immunoassay, and evaluation of the clinical utility of TEX101 diagnostic test in the context of male infertility.

- Antibody-based ELISA is the method of choice for large-scale validation of biomarkers that were previously discovered by mass spectrometry-based approaches. Due to their simplicity, high throughput, and high sensitivity and specificity, ELISA diagnostic tests are easily translated into clinical practice. Therefore, we expressed human recombinant TEX101 protein in the P. pastoris yeast system, which was used as an immunogen for the production of anti-TEX101 monoclonal antibodies.

- For the first time, we demonstrated that the use of a two-step immunocapture followed by SRM detection facilitated the selection of hybridoma clones that secrete antibodies against the native form of the protein of interest in biological fluids. In our case, screening of antibody-secreting clones by immunocapture-SRM facilitated the selection of antibodies against native TEX101 and the development of a sensitive TEX101 ELISA (LOD = 0.5 ng/mL).

- Seminal plasma has been previously proposed as a clinical sample for non-invasive diagnostics. However, its distinctive features, such as high viscosity, fast protease-
mediated liquefaction, and abundance of SMVs, may affect the performance of the immunoassay. Particularly for TEX101 ELISA, we demonstrated that SP required additional treatment for protein quantification with high sensitivity. DOC-based treatment protocol outperformed GndCl-based protocol in the full range of measured TEX101 concentrations in SP (0.5 – 63,825 ng/mL), therefore it was the protocol of choice for the TEX101 measurement in the large cohort of samples (N=805).

- We used mass spectrometry, size-exclusion chromatography, ultracentrifugation, and immunohistochemistry to characterize TEX101 as an analyte in SP, and we confirmed the exclusive presence of the extracellular membrane isoform Q9BY14-1, its extracellular membrane localization in germ cells, the presence of its free, unbound form, or the SMV-associated form in SP.

- Evaluation of TEX101 as a male infertility biomarker in a cohort of 805 men showed that TEX101 levels differentiated pre-vasectomy samples from post-vasectomy (P<0.0001), oligospermia (P<0.001), and azoospermia (P<0.001), but not unexplained infertility (P>0.05). In the azoospermia group, TEX101 differentiated between HS and SCO (P=0.0336), but not between MA and SCO (P=0.10).

- Evaluation of TEX101 ELISA performance demonstrated the following clinical utility for TEX101: (i) Evaluation of vasectomy success by setting a TEX101 cut-off value at 0.9 ng/mL with 100% sensitivity at 100% specificity; (ii) Differentiation between OA and NOA, only when combined with ECM1. A combination of TEX101≥0.9 ng/mL with ECM1≥2.3 µg/mL provides 81% sensitivity at 100% specificity; and (iii) Prediction of sperm retrieval in NOA patients. TEX101 has moderate diagnostic value by setting a cut-off value at 0.6 ng/mL with 73% sensitivity at 64% specificity.

2) Identify and validate the physical interactome of human TEX101 protein in testicular tissue, SP, and spermatozoa.

- We developed and optimized a co-IP-MS approach to ensure stringent identification of TEX101-interacting proteins. For this approach, mammalian recombinant TEX101 was expressed, and high affinity mouse mAbs were generated against native TEX101 protein. Testicular tissue and spermatozoa lysis, protein complexes solubilization were achieved by cryomilling and CHAPS.
Co-IP-MS identified several hundred proteins, of which 9 proteins in testicular tissues and 9 proteins in spermatozoa were significantly co-enriched with TEX101 by anti-TEX101 mAb, and passed our stringent selection criteria. No soluble complexes of TEX101 were identified in SP.

Co-IP-SRM validated 7 and 3 candidates in an independent set of testicular tissue and spermatozoa samples, respectively.

The presence of TEX101-DPEP3 testis-specific protein complex was confirmed in testicular tissue and spermatozoa by co-IP-SRM and an in-house developed hybrid immunoassay.

We demonstrated that tyrosine-O-sulfation, a known PTM in extracellular PPIs of testis-specific membrane proteins, was not present in TEX101 protein, and thus it is not involved in preserving TEX101 interactions.

3) Identify and verify the transient interactome of human TEX101 in spermatozoa, utilizing a human TEX101 knockdown model.

In search of a human TEX101 knockout or knockdown, we identified a deleterious missense TEX101 variant (rs35033974) in publically available databases, and following in our cohort of fertile and sub-fertile individuals.

We developed PRM and an immuno-PRM assays for monitoring the wild type and G99V TEX101 forms in SP and spermatozoa. We demonstrated that only residual amounts of TEX101 G99V form was detected in rs35033974 heterozygotes and homozygotes, possibly due to degradation in the ER.

Four identified TEX101 rs35033974 homozygotes presented very low levels of G99V TEX101 protein (3.5 - 47 ng/mL in SP), and they were utilized as human TEX101 knockdown model.

We performed a comprehensive proteomic analysis of TEX101 wild type and TEX101 knockdown spermatozoa. Out of 8,046 identified proteins, 8 testis-specific membrane-bound, or secreted proteins exhibited significantly reduced expression in TEX101 knockdown sperm. Seven out of 8 candidates were verified by SRM.
• Two testis-specific proteins, LY6K and ADAM29, were verified by SRM and western blot analysis.

• DPEP3, a testis-specific GPI-anchored protein that was previously shown to form a complex with TEX101 in germ cells and mature spermatozoa, was not affected in the TEX101 knockdown sperm of rs35033974 homozygotes.

Overall, the goals of this study were to use innovative proteomic approaches to study clinical and functional aspects of testis-specific protein in the context of male reproduction. As a result, we demonstrated the clinical utility of TEX101, and elucidated some aspects of molecular function of TEX101 protein in humans, as we identified proteins that may physically or functionally interact with TEX101, and may be implicated in the production of competent sperm and fertilization. From a clinical point of view, TEX101 is an excellent marker to evaluate success of vasectomy, and a moderate marker to predict sperm retrieval in NOA patients. From a functional point of view, our findings suggest that DPEP3, LY6K, and ADAM29 interact with TEX101 post-translation, either on the sperm surface or during trafficking to the membrane. It is evident that sperm production and maturation depend on multiple gene expression, as several processes need to be orchestrated. As such, one of the main strengths of our approaches is that by using high-throughput methods, we can map the global interactome of TEX101, which gives us a complete picture of TEX101 physiological function, and potential implications in infertility.

As with many studies that investigate endogenous PPIs, we do recognize that there are limitations in our approaches. The main two limitations are: (i) the lack other known testis-specific PPIs in human germ cells and spermatozoa, which could serve as positive controls for the optimization co-IP; and (ii) the lack of established in vitro human germ cell lines, which could provide the option of gene manipulation, like RNA interference and gene knockout, and investigate TEX101 function in more details. To account for this, we sought for human TEX101 natural knockout, or knockdown models, and we identified a missense variant (rs35033974) that leads to TEX101 knockdown and near-complete absence of TEX101 protein. To conclude, this work completed the evaluation of TEX101 as a biomarker of male infertility, and the TEX101 ELISA as a clinical diagnostic tool, and also laid the groundwork for further investigation of the function of TEX101 in human spermatogenesis and fertilization.
6.3 Future Directions

This dissertation focused on TEX101, one the 1,000 testis-specific genes, and shed new insights into its involvement in male reproduction. Research on testis-specific proteins and their function has two main clinical applications: (i) to discover innovative and effective contraceptive strategies; and (ii) to improve the selection of competent sperm and ensure successful outcomes for ART (306, 307). Promising targets for non-hormonal male contraceptive drug development should possess the following properties: (i) exclusive testis-, or epididymis-specificity and surface orientation; (ii) function that affect sperm production, maturation, or motility; (iii) evidence that mutations of the gene in mouse or human results in infertility; and (iv) suitability for rapid, efficient, but reversible contraception (13, 85, 308, 309). No male contraceptives are available at the moment, and only few targets and compounds have been investigated (13, 258-260), while the most promising compounds were abandoned due to their severe side effects.

Based on the existing evidence, TEX101 protein holds promise as a male contraception target. In this study, we contributed to the current knowledge by identifying several testis-specific TEX101-interacting partners. However, further investigation of these PPIs is required in order to understand their physiological role in sperm production, and the effect of their disruption on male fertility. Recent studies have shown that PPIs can be targets for drug development (261). Since PPIs play pivotal role in the production and maturation of haploid spermatids and spermatozoa, targeting and disrupting these PPIs with small molecules or peptides may lead to the discovery of novel non-hormonal male contraceptive. Additionally, in this work we developed a TEX101-DPEP3 hybrid immunoassay for the mass spectrometry-independent validation of TEX101-DPEP3 complex in testis and spermatozoa. We propose the future use of this assay as a simple, sensitive and high-throughput in vitro platform to screen for molecules that disrupt TEX101-DPEP3 complex. This would complement our findings, and potentially provide more information about the function of the TEX101-DPEP3 complex. The use of this type of hybrid immunoassay platform could be expanded for the detection of other protein complexes.

Approximately 500 infertile or subfertile mouse models have been generated by targeted gene disruption (81, 85). These mouse models revealed genes that are essential for spermatogenesis and fertilization (81), thus the corresponding proteins could be candidate targets for male contraceptive design. Disruption of Tex101 gene in mice resulted in absolute sterility,
supporting the idea that TEX101 protein can be a potential drug target. However, this finding cannot be translated into humans, and it is not confirmed yet if disruption of human TEX101 leads to production of incompetent sperm, and results in fertilization deficiency. This prompted us to investigate if a missense SNV in human TEX101 would lead to the translation of truncated or non-functional protein, and furthermore to infertile spermatozoa. Our preliminary results showed that rs35033974 variant leads to degradation of the variant form of TEX101 protein. Thus, rs35033974 homozygotes presented very low levels of TEX101 (<47 ng/mL), and they were considered as human TEX101 knockdowns. However, this study cannot conclude if the presence of this variant and low concentrations of TEX101 result in male subfertility or infertility. Therefore, prospective collection of semen samples accompanied by clinical data for male and female partners, and the outcome of ART is essential to evaluate if TEX101 variant affects fertilization competence of sperm. Identification of additional TEX101 rs35033974 homozygous and heterozygous men and measurement of residual TEX101 will allow the assessment of the TEX101 concentration range and the degradation rate of TEX101 among such individuals. Sperm functional tests, like hemizona assay and sperm penetration assay (303, 304), can be used to assess the fertilization ability of TEX101 knockdown sperm obtained from rs35033974 homozygotes with different concentration of residual TEX101. Such experiments will demonstrate whether the presence of the SNV or the concentration of residual variant TEX101 protein is the actual predictive marker for fertilization success. Once this is proven, measurement of TEX101 protein concentration, and rs35033974 TEX101 genotyping could be implemented in the evaluation of unexplained infertility cases, in which semen analysis is normal.

Considering competent sperm selection and successful ART outcome, it is still not clear if TEX101-positive spermatozoa are superior and of higher quality compared to TEX101-negative. There is evidence that TEX101 is cleaved from the surface of sperm during epididymal maturation, however we demonstrated that a fraction of it remains on the membrane of mature sperm. Thorough examination of TEX101-positive and TEX101-negative sperm, in terms of morphology, motility, interaction and binding with cumulus and ZP in vitro, may assess the maturity of sperm. Thus, TEX101 on sperm may emerge as an inclusion or exclusion criteria for the selection of sperm for ART.
The outstanding performance of TEX101 ELISA test in discriminating between healthy individuals pre- and post-vasectomy could further support the development of a TEX101 point-of-care test which could be implemented as an Operating Room (OR)-, or a home-based test to evaluate the success of vasectomy or vasectomy reversal. A pair of monoclonal antibodies against native TEX101 protein can be employed to develop a lateral flow solid-phase immunoassay that will offer a fast, simple, sensitive and inexpensive qualitative test to detect TEX101 in SP or vas deferens fluid. Other home-based tests are already available, for example SpermCheck Fertility® and SpermCheck Vasectomy®, which are designed to detect sperm in semen by employing a pair of monoclonal antibodies against a spermatid- and spermatozoa-specific protein, acrosomal protein SP-10 (ACRV1) (310, 311).

TEX101 performance in predicting sperm retrieval in NOA patients was moderate. However, this needs to be confirmed in a larger cohort of NOA patients. The lack of sensitive and specific biomarkers to predict sperm retrieval will require thorough examination of testis-specific proteins that are released in SP (DPEP3, ACRV1, PRND, IZUMO1, etc.). Combination of Human Protein Atlas IHC data and SP proteome would be a comprehensive source to identify a protein, or a panel of proteins, with TEX101 included, that may successfully predict sperm retrieval in NOA patients non-invasively.

Overall, human TEX101 is a novel testis-specific protein that has not been fully characterized yet. In the present study we addressed only few aspects of TEX101, mostly associated with its performance as a male infertility biomarker, and its physical and transient interactions during spermatogenesis and sperm maturation. The role of TEX101 glycosylation in its function, the identification of the enzyme(s) that regulate TEX101 processing during sperm maturation, and the involvement of TEX101 in the acrosome reaction are still an open area for future research. New knowledge on TEX101 will improve our understanding of reproductive biology, contribute to rational selection of infertility treatments, and provide new protein targets for the development of non-hormonal male contraceptives.
References


30. Mital, P., Hinton, B. T., and Dufour, J. M. (2011) The blood-testis and blood-epididymis barriers are more than just their tight junctions. *Biology of reproduction* 84, 851-858


149


fertility by affecting sperm migration into the oviduct in mice. *J.Mol.Cell Biol.* 5, 345-347

Comprehensive analysis of reproductive ADAMs: relationship of ADAM4 and ADAM6 with an
ADAM complex required for fertilization in mice. *Biology of reproduction* 80, 1001-1008

165. Kim, T., Oh, J., Woo, J. M., Choi, E., Im, S. H., Yoo, Y. J., Kim, D. H., Nishimura, H., and
reproduction* 74, 744-750


loss of adhesive function in sperm lacking cyritestin or fertilin beta. *Dev.Biol.* 233, 204-213

168. Yamaguchi, R., Muro, Y., Isotani, A., Tokuhiro, K., Takumi, K., Adham, I., Ikawa, M., and
*Biol.Reprod.* 81, 142-146

aggregation in Adam2 and Adam3 null mice. *Fertil.Steril.* 93, 2754-2756

Identification of a binding site in the disintegrin domain of fertilin required for sperm-egg

egg integrin alpha 6 beta 1 functions as a sperm receptor. *Cell* 81, 1095-1104

interactions: analysis of roles of egg integrins and the mouse sperm homologue of PH-30
(fertilin) beta. *J.Cell Sci.* 108 ( Pt 10), 3267-3278

distribution of ADAM3 in sperm from both angiotensin-converting enzyme (Ace)- and
calmegin (Clgn)-deficient mice. *Biology of reproduction* 75, 760-766

A. (2001) Control of mouse cumulus cell-oocyte complex integrity before and after ovulation:
plasminogen activator synthesis and matrix degradation. *Endocrinology* 142, 3033-3040


Development of ELISA for the Measurement of Native TEX101 in Biological Fluids. *Molecular & cellular proteomics : MCP* 14, 1517-1526


Appendices

**Appendix 2.1 Analysis of recombinant human TEX101 protein.** (a) SDS-PAGE analysis of purified recombinant TEX101 of 3, 9, and 12 μg/well. (b) SDS-PAGE analysis of purified recombinant TEX101, before and after treatment with PNGase F. Lanes 1-3, Purified recombinant TEX101 of 2,1 and 5 μg/well; Lane 4, Purified recombinant TEX101 after treatment with PNGase F. (c) Western Blotting analysis of purified recombinant TEX101, before and after treatment with PNGase F. Lane 1, Purified recombinant TEX101 (~90 ng/well) after treatment with PNGase F; Lane 2, Purified recombinant TEX101 (~90 ng/well). Arrows mark the bands of the gel excised for MS/MS analysis. M, Markers.
Appendix 2.2 Calibration curves and linear correlation between endogenous/spike-in TEX101 concentration and dilution in seminal plasma and female serum. (a) Calibration curve for TEX101 immunoassay in seminal plasma. (b) Linear correlation between endogenous TEX101 concentration and dilution in a seminal plasma sample (regression coefficient $\beta_1=0.900$, $P<0.0001$). (c) Calibration curve for TEX101 immunoassay with seminal plasma spiked into female serum. (d) Linear correlation between spike-in TEX101 concentration and dilution in a female serum sample (regression coefficient $\beta_1=1.034$, $P<0.0001$).
Appendix 2.3 Calibration curves for the endogenous TEX101 in seminal plasma (▲) and the deglycosylated recombinant TEX101 diluted in PBS buffer (●).
Appendix 3.1 TEX101 levels measured in SP samples (N=821) by ELISA using GndCl-based protocol (3 M guanidine hydrochloride for 1 hour at RT). Median values for each group are shown as red bars.
Appendix 3.2 Diagnosis, sperm count and TEX101 levels in SP of 17 infertile men with high sperm count ($\geq$7 mln/mL). Notes: *TEX101 protein was detected by additional peptide LMSGILAVGPMFVR, however accurate ratios based on internal standard could not be calculated due to its low levels (<LOQ) in these specific samples.

<table>
<thead>
<tr>
<th>Sample #</th>
<th>Group</th>
<th>Sperm count (mln/mL)</th>
<th>TEX101 in SP by ELISA (ng/mL)</th>
<th>TEX101 in SP by SRM (ng/mL)</th>
<th>TEX101 in spermatozoa by SRM (ng/µg total protein)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Unexplained infertility</td>
<td>111.1</td>
<td>&lt;0.5</td>
<td>3051.8</td>
<td>0.11</td>
</tr>
<tr>
<td>2</td>
<td>Unexplained infertility</td>
<td>91.1</td>
<td>2.82</td>
<td>742.4</td>
<td>0.16</td>
</tr>
<tr>
<td>3</td>
<td>Unexplained infertility</td>
<td>30</td>
<td>&lt;0.5</td>
<td>420.3</td>
<td>0.02</td>
</tr>
<tr>
<td>4</td>
<td>Unexplained infertility</td>
<td>58.8</td>
<td>&lt;0.5</td>
<td>978.1</td>
<td>0.22</td>
</tr>
<tr>
<td>5</td>
<td>Unexplained infertility</td>
<td>18.5</td>
<td>&lt;0.5</td>
<td>1140.3</td>
<td>0.08</td>
</tr>
<tr>
<td>6</td>
<td>Unexplained infertility</td>
<td>26.1</td>
<td>&lt;0.5</td>
<td>Not Detected</td>
<td>0.13</td>
</tr>
<tr>
<td>7</td>
<td>Unexplained infertility</td>
<td>59</td>
<td>&lt;0.5</td>
<td>2082.2</td>
<td>0.09</td>
</tr>
<tr>
<td>8</td>
<td>Unexplained infertility</td>
<td>104.5</td>
<td>&lt;0.5</td>
<td>1099.7</td>
<td>0.05</td>
</tr>
<tr>
<td>9</td>
<td>Unexplained infertility</td>
<td>123.9</td>
<td>&lt;0.5</td>
<td>389.3</td>
<td>0.15</td>
</tr>
<tr>
<td>10</td>
<td>Unexplained infertility</td>
<td>139.5</td>
<td>&lt;0.5</td>
<td>1289.6</td>
<td>0.81</td>
</tr>
<tr>
<td>11</td>
<td>Oligospermia</td>
<td>6.9</td>
<td>&lt;0.5</td>
<td>Not Detected</td>
<td>Detected*</td>
</tr>
<tr>
<td>12</td>
<td>Oligospermia</td>
<td>10.3</td>
<td>&lt;0.5</td>
<td>104.5</td>
<td>0.05</td>
</tr>
<tr>
<td>13</td>
<td>Oligospermia</td>
<td>7.7</td>
<td>&lt;0.5</td>
<td>758.4</td>
<td>0.18</td>
</tr>
<tr>
<td>14</td>
<td>Oligospermia</td>
<td>8.4</td>
<td>&lt;0.5</td>
<td>265.6</td>
<td>0.06</td>
</tr>
<tr>
<td>15</td>
<td>Oligospermia</td>
<td>9.9</td>
<td>&lt;0.5</td>
<td>Not Detected</td>
<td>0.01</td>
</tr>
<tr>
<td>16</td>
<td>Oligospermia</td>
<td>10.9</td>
<td>&lt;0.5</td>
<td>324.3</td>
<td>0.07</td>
</tr>
<tr>
<td>17</td>
<td>Oligospermia</td>
<td>12.8</td>
<td>&lt;0.5</td>
<td>Detected*</td>
<td>0.05</td>
</tr>
</tbody>
</table>
Appendix 3.3 Relative amounts of TEX101 captured from SP pools by commercial (mPoly) or in-house generated (23ED616.8) antibodies, as measured by SRM. SP samples were treated before capture using the following protocols: (1) untreated; (2) incubated with 3 M guanidine hydrochloride for 1 hour at RT; (3) incubated with 2% sodium deoxycholate for 1 hour at RT; (4) incubated with 3 M guanidine hydrochloride for 1 hour at 63°C; (5) incubated with 2% sodium deoxycholate for 1 hour at 63°C; and (6) incubated for 1 hour at 63°C.
Appendix 3.4 Correlation between sperm count and TEX101 levels in SP. Samples were incubated with 2% sodium deoxycholate for 1 hour at 63oC, prior to TEX101 analysis by ELISA. rs, Spearman correlation coefficient.
Appendix 4.1 Parameters of a multiplex scheduled SRM assay for the validation of TEX101 candidate interacting proteins in human testicular tissue. [cm], carboxyamidomethyl; IS, spike-in internal standard; K*, heavy-isotope-labeled 13C6, 15N2 L-Lysine; R*, heavy-isotope-labeled 13C6, 15N4 L-Arginine

<table>
<thead>
<tr>
<th>Protein (gene symbol)</th>
<th>Peptide</th>
<th>Q1, m/z</th>
<th>Q3, m/z</th>
<th>Collision Energy, V</th>
<th>Dwell time, ms</th>
<th>LC retention time, min</th>
<th>Peptide Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>TEX101</td>
<td>AGTETAILATK</td>
<td>538.3</td>
<td>1004.6</td>
<td>20.6</td>
<td>20</td>
<td>13.4</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>TEX101</td>
<td>AGTETAILATK</td>
<td>538.3</td>
<td>846.5</td>
<td>20.6</td>
<td>20</td>
<td>13.4</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>TEX101</td>
<td>AGTETAILATK</td>
<td>538.3</td>
<td>717.5</td>
<td>20.6</td>
<td>20</td>
<td>13.4</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>TEX101</td>
<td>AGTETAILATK*</td>
<td>542.3</td>
<td>1012.6</td>
<td>20.7</td>
<td>20</td>
<td>13.4</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>TEX101</td>
<td>AGTETAILATK*</td>
<td>542.3</td>
<td>854.5</td>
<td>20.7</td>
<td>20</td>
<td>13.4</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>TEX101</td>
<td>AGTETAILATK*</td>
<td>542.3</td>
<td>725.5</td>
<td>20.7</td>
<td>20</td>
<td>13.4</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>ACE2</td>
<td>AV(C[cm])HTPAWDLGK</td>
<td>452.2</td>
<td>887.5</td>
<td>14.9</td>
<td>20</td>
<td>15</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>ACE2</td>
<td>AV(C[cm])HTPAWDLGK</td>
<td>452.2</td>
<td>317.2</td>
<td>14.9</td>
<td>20</td>
<td>15</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>ACE2</td>
<td>AV(C[cm])HTPAWDLGK</td>
<td>452.2</td>
<td>592.8</td>
<td>14.9</td>
<td>20</td>
<td>15</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>ACE2</td>
<td>AV(C[cm])HTPAWDLGK*</td>
<td>454.9</td>
<td>895.5</td>
<td>14.9</td>
<td>20</td>
<td>15</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>ACE2</td>
<td>AV(C[cm])HTPAWDLGK*</td>
<td>454.9</td>
<td>325.2</td>
<td>14.9</td>
<td>20</td>
<td>15</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>ACE2</td>
<td>AV(C[cm])HTPAWDLGK*</td>
<td>454.9</td>
<td>596.8</td>
<td>14.9</td>
<td>20</td>
<td>15</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>NT5E</td>
<td>VPSYDPLK</td>
<td>459.8</td>
<td>722.4</td>
<td>17.9</td>
<td>20</td>
<td>15.1</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>NT5E</td>
<td>VPSYDPLK</td>
<td>459.8</td>
<td>357.2</td>
<td>17.9</td>
<td>20</td>
<td>15.1</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>NT5E</td>
<td>VPSYDPLK</td>
<td>459.8</td>
<td>410.2</td>
<td>17.9</td>
<td>20</td>
<td>15.1</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>NT5E</td>
<td>VPSYDPLK*</td>
<td>463.8</td>
<td>730.4</td>
<td>18.1</td>
<td>20</td>
<td>15.1</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>NT5E</td>
<td>VPSYDPLK*</td>
<td>463.8</td>
<td>365.3</td>
<td>18.1</td>
<td>20</td>
<td>15.1</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>NT5E</td>
<td>VPSYDPLK*</td>
<td>463.8</td>
<td>414.2</td>
<td>18.1</td>
<td>20</td>
<td>15.1</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>ITGB1</td>
<td>SGEPQTFTLK</td>
<td>554.3</td>
<td>609.4</td>
<td>21.2</td>
<td>20</td>
<td>15</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>ITGB1</td>
<td>SGEPQTFTLK</td>
<td>554.3</td>
<td>361.2</td>
<td>21.2</td>
<td>20</td>
<td>15</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>ITGB1</td>
<td>SGEPQTFTLK</td>
<td>554.3</td>
<td>417.7</td>
<td>21.2</td>
<td>20</td>
<td>15</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>ITGB1</td>
<td>SGEPQTFTLK*</td>
<td>558.3</td>
<td>617.4</td>
<td>21.3</td>
<td>20</td>
<td>15</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>ITGB1</td>
<td>SGEPQTFTLK*</td>
<td>558.3</td>
<td>369.3</td>
<td>21.3</td>
<td>20</td>
<td>15</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>ITGB1</td>
<td>SGEPQTFTLK*</td>
<td>558.3</td>
<td>421.7</td>
<td>21.3</td>
<td>20</td>
<td>15</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>DPEP3</td>
<td>SWSEEELQGVLR</td>
<td>716.9</td>
<td>1159.6</td>
<td>26.7</td>
<td>20</td>
<td>21</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>DPEP3</td>
<td>SWSEEELQGVLR</td>
<td>716.9</td>
<td>814.5</td>
<td>26.7</td>
<td>20</td>
<td>21</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>DPEP3</td>
<td>SWSEEELQGVLR</td>
<td>716.9</td>
<td>444.3</td>
<td>26.7</td>
<td>20</td>
<td>21</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>DPEP3</td>
<td>SWSEEELQGVLR*</td>
<td>721.9</td>
<td>1169.6</td>
<td>26.8</td>
<td>20</td>
<td>21</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>DPEP3</td>
<td>SWSEEELQGVLR*</td>
<td>721.9</td>
<td>824.5</td>
<td>26.8</td>
<td>20</td>
<td>21</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>DPEP3</td>
<td>SWSEEELQGVLR*</td>
<td>721.9</td>
<td>454.3</td>
<td>26.8</td>
<td>20</td>
<td>21</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>ART3</td>
<td>ILEPTQIPAPGPVPGPK</td>
<td>636.4</td>
<td>1112.6</td>
<td>34.7</td>
<td>20</td>
<td>19.3</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>ART3</td>
<td>ILEPTQIPAPGPVPGPK</td>
<td>636.4</td>
<td>398.2</td>
<td>34.7</td>
<td>20</td>
<td>19.3</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>ART3</td>
<td>ILEPTQIPAPGPVPGPK</td>
<td>636.4</td>
<td>301.2</td>
<td>34.7</td>
<td>20</td>
<td>19.3</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>ART3</td>
<td>ILEPTQIPAPGPVPGPK*</td>
<td>639.0</td>
<td>1120.7</td>
<td>34.8</td>
<td>20</td>
<td>19.3</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>Protein</td>
<td>Peptide Sequence</td>
<td>Mz (calc)</td>
<td>Mz (exp)</td>
<td>Intensity</td>
<td>Source</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>-----------------</td>
<td>-----------</td>
<td>----------</td>
<td>-----------</td>
<td>--------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ART3</td>
<td>ILEPTQIPAPGPVPVPGPK*</td>
<td>639.0</td>
<td>406.3</td>
<td>34.8</td>
<td>20</td>
<td>19.3</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>ART3</td>
<td>ILEPTQIPAPGPVPVPGPK*</td>
<td>639.0</td>
<td>309.2</td>
<td>34.8</td>
<td>20</td>
<td>19.3</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>LAMP1</td>
<td>TVESITDIR</td>
<td>517.3</td>
<td>833.4</td>
<td>19.9</td>
<td>20</td>
<td>15.1</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>LAMP1</td>
<td>TVESITDIR</td>
<td>517.3</td>
<td>288.2</td>
<td>19.9</td>
<td>20</td>
<td>15.1</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>LAMP1</td>
<td>TVESITDIR</td>
<td>517.3</td>
<td>201.1</td>
<td>19.9</td>
<td>20</td>
<td>15.1</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>LAMP1</td>
<td>TVESITDIR*</td>
<td>522.3</td>
<td>843.4</td>
<td>20.1</td>
<td>20</td>
<td>15.1</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>LAMP1</td>
<td>TVESITDIR*</td>
<td>522.3</td>
<td>298.2</td>
<td>20.1</td>
<td>20</td>
<td>15.1</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>CD59</td>
<td>FEHC[cm]NFNDVTR</td>
<td>513.9</td>
<td>591.3</td>
<td>16.7</td>
<td>20</td>
<td>13</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>CD59</td>
<td>FEHC[cm]NFNDVTR</td>
<td>513.9</td>
<td>476.3</td>
<td>16.7</td>
<td>20</td>
<td>13</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>CD59</td>
<td>FEHC[cm]NFNDVTR</td>
<td>513.9</td>
<td>377.2</td>
<td>16.7</td>
<td>20</td>
<td>13</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>CD59</td>
<td>FEHC[cm]NFNDVTR*</td>
<td>517.2</td>
<td>601.3</td>
<td>16.8</td>
<td>20</td>
<td>13</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>CD59</td>
<td>FEHC[cm]NFNDVTR*</td>
<td>517.2</td>
<td>486.3</td>
<td>16.8</td>
<td>20</td>
<td>13</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>CD59</td>
<td>FEHC[cm]NFNDVTR*</td>
<td>517.2</td>
<td>387.2</td>
<td>16.8</td>
<td>20</td>
<td>13</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>PRND</td>
<td>LVQELC[cm]SLK</td>
<td>545.3</td>
<td>877.4</td>
<td>20.8</td>
<td>20</td>
<td>15.9</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>PRND</td>
<td>LVQELC[cm]SLK</td>
<td>545.3</td>
<td>749.4</td>
<td>20.8</td>
<td>20</td>
<td>15.9</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>PRND</td>
<td>LVQELC[cm]SLK</td>
<td>545.3</td>
<td>620.3</td>
<td>20.8</td>
<td>20</td>
<td>15.9</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>PRND</td>
<td>LVQELC[cm]SLK*</td>
<td>549.3</td>
<td>885.5</td>
<td>21.0</td>
<td>20</td>
<td>15.9</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>PRND</td>
<td>LVQELC[cm]SLK*</td>
<td>549.3</td>
<td>757.4</td>
<td>21.0</td>
<td>20</td>
<td>15.9</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>PRND</td>
<td>LVQELC[cm]SLK*</td>
<td>549.3</td>
<td>628.4</td>
<td>21.0</td>
<td>20</td>
<td>15.9</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>CD81</td>
<td>NNLC[cm]PSGSNIISNLFK</td>
<td>889.4</td>
<td>1276.7</td>
<td>32.5</td>
<td>20</td>
<td>25.8</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>CD81</td>
<td>NNLC[cm]PSGSNIISNLFK</td>
<td>889.4</td>
<td>608.3</td>
<td>32.5</td>
<td>20</td>
<td>25.8</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>CD81</td>
<td>NNLC[cm]PSGSNIISNLFK</td>
<td>889.4</td>
<td>229.1</td>
<td>32.5</td>
<td>20</td>
<td>25.8</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>CD81</td>
<td>NNLC[cm]PSGSNIISNLFK*</td>
<td>893.5</td>
<td>1284.7</td>
<td>32.6</td>
<td>20</td>
<td>25.8</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>CD81</td>
<td>NNLC[cm]PSGSNIISNLFK*</td>
<td>893.5</td>
<td>616.4</td>
<td>32.6</td>
<td>20</td>
<td>25.8</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>CD81</td>
<td>NNLC[cm]PSGSNIISNLFK*</td>
<td>893.5</td>
<td>229.1</td>
<td>32.6</td>
<td>20</td>
<td>25.8</td>
<td>Heavy IS</td>
</tr>
</tbody>
</table>
Appendix 4.2 Parameters of a multiplex scheduled SRM assay for the validation of TEX101 candidate interacting proteins in human sperm cells from normal fertile individuals. [cm], carboxyamidomethyl; IS, spike-in internal standard; K*, heavy-isotope-labeled 13C6, 15N2 L-Lysine; R*, heavy-isotope-labeled 13C6, 15N4 L-Arginine

<table>
<thead>
<tr>
<th>Protein (gene symbol)</th>
<th>Peptide</th>
<th>Q1, m/z</th>
<th>Q3, m/z</th>
<th>Collision Energy, V</th>
<th>Dwell time, ms</th>
<th>LC retention time, min</th>
<th>Peptide Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>TEX101</td>
<td>AGTETAILATK</td>
<td>538.3</td>
<td>1004.6</td>
<td>21.6</td>
<td>20</td>
<td>13.7</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>TEX101</td>
<td>AGTETAILATK</td>
<td>538.3</td>
<td>846.5</td>
<td>21.6</td>
<td>20</td>
<td>13.7</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>TEX101</td>
<td>AGTETAILATK</td>
<td>538.3</td>
<td>717.5</td>
<td>21.6</td>
<td>20</td>
<td>13.7</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>TEX101</td>
<td>AGTETAILATK*</td>
<td>542.3</td>
<td>1012.6</td>
<td>21.6</td>
<td>20</td>
<td>13.7</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>TEX101</td>
<td>AGTETAILATK*</td>
<td>542.3</td>
<td>854.5</td>
<td>21.6</td>
<td>20</td>
<td>13.7</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>ACE</td>
<td>YQGIC[cm]PPVTR</td>
<td>595.8</td>
<td>899.5</td>
<td>23.6</td>
<td>20</td>
<td>14</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>ACE</td>
<td>YQGIC[cm]PPVTR</td>
<td>595.8</td>
<td>729.4</td>
<td>23.6</td>
<td>20</td>
<td>14</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>ACE</td>
<td>YQGIC[cm]PPVTR</td>
<td>595.8</td>
<td>569.3</td>
<td>23.6</td>
<td>20</td>
<td>14</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>ACE</td>
<td>YQGIC[cm]PPVTR*</td>
<td>600.8</td>
<td>909.5</td>
<td>23.6</td>
<td>20</td>
<td>14</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>ACE</td>
<td>YQGIC[cm]PPVTR*</td>
<td>600.8</td>
<td>739.4</td>
<td>23.6</td>
<td>20</td>
<td>14</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>ACE</td>
<td>YQGIC[cm]PPVTR*</td>
<td>600.8</td>
<td>579.3</td>
<td>23.6</td>
<td>20</td>
<td>14</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>DPEP3</td>
<td>SWSEEELQGVL R</td>
<td>716.9</td>
<td>1159.6</td>
<td>27.7</td>
<td>20</td>
<td>21.2</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>DPEP3</td>
<td>SWSEEELQGVL R</td>
<td>716.9</td>
<td>814.5</td>
<td>27.7</td>
<td>20</td>
<td>21.2</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>DPEP3</td>
<td>SWSEEELQGVL R</td>
<td>716.9</td>
<td>444.3</td>
<td>27.7</td>
<td>20</td>
<td>21.2</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>DPEP3</td>
<td>SWSEEELQGVL R*</td>
<td>721.9</td>
<td>1169.6</td>
<td>27.7</td>
<td>20</td>
<td>21.2</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>DPEP3</td>
<td>SWSEEELQGVL R*</td>
<td>721.9</td>
<td>824.5</td>
<td>27.7</td>
<td>20</td>
<td>21.2</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>DPEP3</td>
<td>SWSEEELQGVL R*</td>
<td>721.9</td>
<td>454.3</td>
<td>27.7</td>
<td>20</td>
<td>21.2</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>LAMP1</td>
<td>TVESITDIR</td>
<td>517.3</td>
<td>833.4</td>
<td>20.9</td>
<td>20</td>
<td>15.3</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>LAMP1</td>
<td>TVESITDIR</td>
<td>517.3</td>
<td>288.2</td>
<td>20.9</td>
<td>20</td>
<td>15.3</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>LAMP1</td>
<td>TVESITDIR</td>
<td>517.3</td>
<td>201.1</td>
<td>20.9</td>
<td>20</td>
<td>15.3</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>LAMP1</td>
<td>TVESITDIR*</td>
<td>522.3</td>
<td>843.4</td>
<td>20.9</td>
<td>20</td>
<td>15.3</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>LAMP1</td>
<td>TVESITDIR*</td>
<td>522.3</td>
<td>298.2</td>
<td>20.9</td>
<td>20</td>
<td>15.3</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>LAMP1</td>
<td>TVESITDIR*</td>
<td>522.3</td>
<td>201.1</td>
<td>20.9</td>
<td>20</td>
<td>15.3</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>CD59</td>
<td>FEHC[cm]NFNDVTTR</td>
<td>513.9</td>
<td>591.3</td>
<td>21.8</td>
<td>20</td>
<td>15.3</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>CD59</td>
<td>FEHC[cm]NFNDVTTR</td>
<td>513.9</td>
<td>476.3</td>
<td>21.8</td>
<td>20</td>
<td>15.3</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>CD59</td>
<td>FEHC[cm]NFNDVTTR</td>
<td>513.9</td>
<td>377.2</td>
<td>21.8</td>
<td>20</td>
<td>15.3</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>CD59</td>
<td>FEHC[cm]NFNDVTTR*</td>
<td>517.2</td>
<td>601.3</td>
<td>21.8</td>
<td>20</td>
<td>15.3</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>CD59</td>
<td>FEHC[cm]NFNDVTTR*</td>
<td>517.2</td>
<td>486.3</td>
<td>21.8</td>
<td>20</td>
<td>15.3</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>CD59</td>
<td>FEHC[cm]NFNDVTTR*</td>
<td>517.2</td>
<td>387.2</td>
<td>21.8</td>
<td>20</td>
<td>15.3</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>ZPB P</td>
<td>IVGSTSFPVK</td>
<td>517.8</td>
<td>822.4</td>
<td>20.9</td>
<td>20</td>
<td>15.1</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>ZPB P</td>
<td>IVGSTSFPVK</td>
<td>517.8</td>
<td>343.2</td>
<td>20.9</td>
<td>20</td>
<td>15.1</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>ZPB P</td>
<td>IVGSTSFPVK</td>
<td>517.8</td>
<td>213.2</td>
<td>20.9</td>
<td>20</td>
<td>15.1</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>Protein</td>
<td>Peptide</td>
<td>MW</td>
<td>pI</td>
<td>AA</td>
<td>Status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>---------------------</td>
<td>-------</td>
<td>--------</td>
<td>---------</td>
<td>--------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZPBP</td>
<td>IVGSTSFPVK*</td>
<td>521.8</td>
<td>8.30</td>
<td>20.9</td>
<td>15.1</td>
<td>Heavy IS</td>
<td></td>
</tr>
<tr>
<td>ZPBP</td>
<td>IVGSTSFPVK*</td>
<td>521.8</td>
<td>8.30</td>
<td>20.9</td>
<td>15.1</td>
<td>Heavy IS</td>
<td></td>
</tr>
<tr>
<td>ZPBP</td>
<td>IVGSTSFPVK*</td>
<td>521.8</td>
<td>8.30</td>
<td>20.9</td>
<td>15.1</td>
<td>Heavy IS</td>
<td></td>
</tr>
<tr>
<td>SPACA1</td>
<td>LAC[cm]IHTSPLNR</td>
<td>427.9</td>
<td>5.84</td>
<td>18.4</td>
<td>20</td>
<td>Light endogenous</td>
<td></td>
</tr>
<tr>
<td>SPACA1</td>
<td>LAC[cm]IHTSPLNR</td>
<td>427.9</td>
<td>5.84</td>
<td>18.4</td>
<td>20</td>
<td>Light endogenous</td>
<td></td>
</tr>
<tr>
<td>SPACA1</td>
<td>LAC[cm]IHTSPLNR*</td>
<td>431.2</td>
<td>5.89</td>
<td>18.4</td>
<td>20</td>
<td>Heavy IS</td>
<td></td>
</tr>
<tr>
<td>SPACA1</td>
<td>LAC[cm]IHTSPLNR*</td>
<td>431.2</td>
<td>5.89</td>
<td>18.4</td>
<td>20</td>
<td>Heavy IS</td>
<td></td>
</tr>
<tr>
<td>SPACA1</td>
<td>LAC[cm]IHTSPLNR*</td>
<td>431.2</td>
<td>5.89</td>
<td>18.4</td>
<td>20</td>
<td>Heavy IS</td>
<td></td>
</tr>
<tr>
<td>SPACA4</td>
<td>ATSC[cm]GLEEPVSYR</td>
<td>734.8</td>
<td>8.79</td>
<td>28.3</td>
<td>20</td>
<td>Light endogenous</td>
<td></td>
</tr>
<tr>
<td>SPACA4</td>
<td>ATSC[cm]GLEEPVSYR</td>
<td>734.8</td>
<td>8.79</td>
<td>28.3</td>
<td>20</td>
<td>Light endogenous</td>
<td></td>
</tr>
<tr>
<td>SPACA4</td>
<td>ATSC[cm]GLEEPVSYR*</td>
<td>739.8</td>
<td>8.89</td>
<td>28.3</td>
<td>20</td>
<td>Heavy IS</td>
<td></td>
</tr>
<tr>
<td>SPACA4</td>
<td>ATSC[cm]GLEEPVSYR*</td>
<td>739.8</td>
<td>8.89</td>
<td>28.3</td>
<td>20</td>
<td>Heavy IS</td>
<td></td>
</tr>
<tr>
<td>SPACA4</td>
<td>ATSC[cm]GLEEPVSYR*</td>
<td>739.8</td>
<td>8.89</td>
<td>28.3</td>
<td>20</td>
<td>Heavy IS</td>
<td></td>
</tr>
<tr>
<td>SLC2A14</td>
<td>IAESVEMLILGR</td>
<td>665.9</td>
<td>8.31</td>
<td>26</td>
<td>20</td>
<td>Light endogenous</td>
<td></td>
</tr>
<tr>
<td>SLC2A14</td>
<td>IAESVEMLILGR</td>
<td>665.9</td>
<td>8.31</td>
<td>26</td>
<td>20</td>
<td>Light endogenous</td>
<td></td>
</tr>
<tr>
<td>SLC2A14</td>
<td>IAESVEMLILGR</td>
<td>665.9</td>
<td>8.31</td>
<td>26</td>
<td>20</td>
<td>Light endogenous</td>
<td></td>
</tr>
<tr>
<td>SLC2A14</td>
<td>IAESVEMLILGR*</td>
<td>670.9</td>
<td>8.41</td>
<td>26</td>
<td>20</td>
<td>Heavy IS</td>
<td></td>
</tr>
<tr>
<td>SLC2A14</td>
<td>IAESVEMLILGR*</td>
<td>670.9</td>
<td>8.41</td>
<td>26</td>
<td>20</td>
<td>Heavy IS</td>
<td></td>
</tr>
<tr>
<td>SLC2A14</td>
<td>IAESVEMLILGR*</td>
<td>670.9</td>
<td>8.41</td>
<td>26</td>
<td>20</td>
<td>Heavy IS</td>
<td></td>
</tr>
<tr>
<td>IZUMO4</td>
<td>ATPAFLVSPALR</td>
<td>621.9</td>
<td>7.55</td>
<td>24.5</td>
<td>20</td>
<td>Light endogenous</td>
<td></td>
</tr>
<tr>
<td>IZUMO4</td>
<td>ATPAFLVSPALR</td>
<td>621.9</td>
<td>7.55</td>
<td>24.5</td>
<td>20</td>
<td>Light endogenous</td>
<td></td>
</tr>
<tr>
<td>IZUMO4</td>
<td>ATPAFLVSPALR</td>
<td>621.9</td>
<td>7.55</td>
<td>24.5</td>
<td>20</td>
<td>Light endogenous</td>
<td></td>
</tr>
<tr>
<td>IZUMO4</td>
<td>ATPAFLVSPALR*</td>
<td>626.9</td>
<td>7.65</td>
<td>24.5</td>
<td>20</td>
<td>Heavy IS</td>
<td></td>
</tr>
<tr>
<td>IZUMO4</td>
<td>ATPAFLVSPALR*</td>
<td>626.9</td>
<td>7.65</td>
<td>24.5</td>
<td>20</td>
<td>Heavy IS</td>
<td></td>
</tr>
<tr>
<td>IZUMO4</td>
<td>ATPAFLVSPALR*</td>
<td>626.9</td>
<td>7.65</td>
<td>24.5</td>
<td>20</td>
<td>Heavy IS</td>
<td></td>
</tr>
</tbody>
</table>
Appendix 4.3 Optimization of testicular tissue lysis. After physical disruption, pulverized tissue was incubated in lysis buffers containing various mild non-denaturing non-ionic (NP-40, Triton X-100) and zwitterionic (CHAPS) detergents. Tissue lysis in PBS with sonication (no detergent) was used as a control. Efficiency of tissue lysis and protein solubilisation was assessed by BCA total protein assay and in house TEX101 immunoassay. Highest recovery of TEX101 protein from testicular tissue was achieved by lysis with CHAPS (1% w/v).
Appendix 4.4 Identification of TEX101 interactome by 34ED229 mAb co-immunoprecipitation coupled with mass spectrometry. Volcano plot (-log10(p-value) versus log2(fold change)) revealed proteins co-enriched with target protein TEX101 by anti-TEX101 monoclonal antibody (34ED229) compared to the mouse IgG negative control, from normal testicular tissue with active spermatogenesis. Statistical analysis was performed from 3 biological replicates in Perseus. Fold change and p-value cut-offs are indicated by the hyperbolic selection curve in black (fold change was assessed for the dataset by the variance correction (s0), FDR-adjusted t test p values <0.01). Anti-TEX101 antibody isolated 28% of membrane/secreted proteins among all significantly enriched proteins from testicular tissue, while isotype control mouse IgG antibody isolated 38% of membrane/secreted proteins. Target protein TEX101 is plotted in red, membrane/secreted proteins significantly enriched are plotted in black, and significantly enriched membrane/secreted proteins expressed by testicular germ cells (candidate interactors of TEX101) are plotted in blue.
Appendix 4.5 SDS-PAGE and Western blotting analysis of purified recombinant DPEP3 protein. Arrows indicate the bands that were excised for MS analysis. Majority of DPEP3 protein migrated around 75-85 kDa.
Appendix 4.6 Production of mouse monoclonal antibodies against different epitopes of native DPEP3 protein. (a) Immunocapture-SRM was used for the screening of hybridoma colonies and the selection of mouse monoclonal antibodies against native DPEP3 protein, present in seminal plasma pool. Lane (+) indicates the positive control (serum of immunized mouse) and lane (-) indicates the negative control (sheep anti-mouse IgG- Fcγ fragment-specific antibody only). Asterisks mark the clones that had positive reaction with native DPEP3. (b) The two mouse monoclonal anti-DPEP3 antibodies (40ED139 and 41ED68) were paired in a sandwich immunoassay, targeting two different epitopes on native DPEP3 protein.
Appendix 4.7 Investigation of TEX101 sulfation in testicular tissue, seminal plasma and spermatozoa. (a) Immunoprecipitation of TEX101 protein from testicular tissue lysate, seminal plasma and spermatozoa lysate by 34ED556 mAb, followed by immunoblotting with commercial rabbit polyclonal anti-TEX101 antibody (i) or mouse monoclonal anti-sulfotyrosine antibody (ii). Samples were as follows: (1) input testicular tissue lysate; (2) testicular tissue after IP; (3) input seminal plasma; (4) seminal plasma after IP; (5) input spermatozoa lysate; (6) spermatozoa lysate after IP. (b) Control immunoprecipitation was performed in parallel by using an isotype control mouse IgG antibody. Samples were as follows: (1) testicular tissue lysate after IP; (2) seminal plasma after IP; (3) spermatozoa lysate after IP. Black arrows indicate proteins enriched by 34ED556 TEX101-specific antibody and isotype control mouse IgG, and detected by anti-sulfotyrosine antibody. Protein marker is indicated to the left of each blot.
Appendix 4.8 Western blotting analysis of endogenous DPEP3 protein. A commercial rabbit polyclonal anti-DPEP3 antibody (HPA058607) was used for DPEP3 detection in: (1) testicular tissue lysate; (2) seminal plasma; (3) spermatozoa lysate (~40 ug total protein/well). Protein marker is indicated to the left of the blot. DPEP3 migrates to 75 kDa (monomer) and 155 kDa (homodimer) in testicular tissue and spermatozoa lysate. In seminal plasma, DPEP3 is mostly present as a homodimer.
Appendix 5.1 Parameters of the unscheduled PRM assay for monitoring TEX101 protein in Glu-C digested sperm cell lysate and seminal plasma. [cm], carboxyamidomethyl; [ox], oxidation; IS, spike-in internal standard; V*, heavy-isotope-labeled 13C5, 15N L-Valine

<table>
<thead>
<tr>
<th>Protein (gene symbol)</th>
<th>Peptide</th>
<th>Q1, m/z</th>
<th>Collision Energy, V</th>
<th>LC retention time, min</th>
<th>Peptide Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>TEX101_wt</td>
<td>AITIVQHSSPPGLIVTSYSNYC[cm]E</td>
<td>1268.6</td>
<td>27</td>
<td>8.5</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>TEX101_wt</td>
<td>AITIVQHSSPPGLIV*TSYSNYC[cm]E</td>
<td>1271.6</td>
<td>27</td>
<td>8.5</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>TEX101_G99V</td>
<td>AITIVQHSSPPVLIIVTSYSNYC[cm]E</td>
<td>1289.6</td>
<td>27</td>
<td>9.2</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>TEX101_G99V</td>
<td>AITIVQHSSPPVLIIV*TSYSNYC[cm]E</td>
<td>1292.7</td>
<td>27</td>
<td>9.2</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>TEX101</td>
<td>TAILATKG[C][cm]IPE</td>
<td>637.3</td>
<td>27</td>
<td>6.8</td>
<td>Light endogenous control peptide</td>
</tr>
</tbody>
</table>
Appendix 5.2 Parameters of a multiplex scheduled SRM assay for the quantification of TEX101 and other downregulated proteins in sperm cell samples from TEX101 G99V homozygotes and control fertile individuals. [cm], carboxymidomethyl; IS, spike-in internal standard; K*, heavy-isotope-labeled 13C6, 15N2 L-Lysine; R*, heavy-isotope-labeled 13C6, 15N4 L-Arginine

<table>
<thead>
<tr>
<th>Protein (gene symbol)</th>
<th>Peptide</th>
<th>Q1, m/z</th>
<th>Q3, m/z</th>
<th>Collision Energy, V</th>
<th>Dwell time, ms</th>
<th>LC retention time, min</th>
<th>Peptide Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>TEX101</td>
<td>AGTETAILATK</td>
<td>538.3</td>
<td>1004.6</td>
<td>21.6</td>
<td>20</td>
<td>13.2</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>TEX101</td>
<td>AGTETAILATK</td>
<td>538.3</td>
<td>846.5</td>
<td>21.6</td>
<td>20</td>
<td>13.2</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>TEX101</td>
<td>AGTETAILATK</td>
<td>538.3</td>
<td>717.5</td>
<td>21.6</td>
<td>20</td>
<td>13.2</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>TEX101</td>
<td>AGTETAILATK*</td>
<td>542.3</td>
<td>1012.6</td>
<td>21.6</td>
<td>20</td>
<td>13.2</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>TEX101</td>
<td>AGTETAILATK*</td>
<td>542.3</td>
<td>854.5</td>
<td>21.6</td>
<td>20</td>
<td>13.2</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>TEX101</td>
<td>AGTETAILATK*</td>
<td>542.3</td>
<td>725.5</td>
<td>21.6</td>
<td>20</td>
<td>13.2</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>DPEP3</td>
<td>SWSEEELQGVR</td>
<td>716.9</td>
<td>1159.6</td>
<td>27.7</td>
<td>20</td>
<td>20.8</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>DPEP3</td>
<td>SWSEEELQGVR</td>
<td>716.9</td>
<td>814.5</td>
<td>27.7</td>
<td>20</td>
<td>20.8</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>DPEP3</td>
<td>SWSEEELQGVR</td>
<td>716.9</td>
<td>444.3</td>
<td>27.7</td>
<td>20</td>
<td>20.8</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>DPEP3</td>
<td>SWSEEELQGVR*</td>
<td>721.9</td>
<td>1169.6</td>
<td>27.7</td>
<td>20</td>
<td>20.8</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>DPEP3</td>
<td>SWSEEELQGVR*</td>
<td>721.9</td>
<td>824.5</td>
<td>27.7</td>
<td>20</td>
<td>20.8</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>DPEP3</td>
<td>SWSEEELQGVR*</td>
<td>721.9</td>
<td>454.3</td>
<td>27.7</td>
<td>20</td>
<td>20.8</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>SPAM1</td>
<td>SPLPVFAYTR</td>
<td>575.8</td>
<td>853.5</td>
<td>21.9</td>
<td>20</td>
<td>20.4</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>SPAM1</td>
<td>SPLPVFAYTR</td>
<td>575.8</td>
<td>657.3</td>
<td>21.9</td>
<td>20</td>
<td>20.4</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>SPAM1</td>
<td>SPLPVFAYTR</td>
<td>575.8</td>
<td>298.2</td>
<td>21.9</td>
<td>20</td>
<td>20.4</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>SPAM1</td>
<td>SPLPVFAYTR*</td>
<td>580.8</td>
<td>863.5</td>
<td>22.0</td>
<td>20</td>
<td>20.4</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>SPAM1</td>
<td>SPLPVFAYTR*</td>
<td>580.8</td>
<td>667.3</td>
<td>22.0</td>
<td>20</td>
<td>20.4</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>SPAM1</td>
<td>SPLPVFAYTR*</td>
<td>580.8</td>
<td>298.2</td>
<td>22.0</td>
<td>20</td>
<td>20.4</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>SPACA3</td>
<td>ITQEPQGLGYWEAWR</td>
<td>917.4</td>
<td>1362.7</td>
<td>33.5</td>
<td>20</td>
<td>23.3</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>SPACA3</td>
<td>ITQEPQGLGYWEAWR</td>
<td>917.4</td>
<td>967.4</td>
<td>33.5</td>
<td>20</td>
<td>23.3</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>SPACA3</td>
<td>ITQEPQGLGYWEAWR</td>
<td>917.4</td>
<td>432.2</td>
<td>33.5</td>
<td>20</td>
<td>23.3</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>SPACA3</td>
<td>ITQEPQGLGYWEAWR*</td>
<td>922.5</td>
<td>1372.7</td>
<td>33.6</td>
<td>20</td>
<td>23.3</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>SPACA3</td>
<td>ITQEPQGLGYWEAWR*</td>
<td>922.5</td>
<td>977.5</td>
<td>33.6</td>
<td>20</td>
<td>23.3</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>SPACA3</td>
<td>ITQEPQGLGYWEAWR*</td>
<td>922.5</td>
<td>442.2</td>
<td>33.6</td>
<td>20</td>
<td>23.3</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>ZPBP2</td>
<td>DFLEPLSGLYTC[cm]TLSYK</td>
<td>1004.0</td>
<td>1502.8</td>
<td>36.4</td>
<td>20</td>
<td>26.9</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>ZPBP2</td>
<td>DFLEPLSGLYTC[cm]TLSYK</td>
<td>1004.0</td>
<td>397.2</td>
<td>36.4</td>
<td>20</td>
<td>26.9</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>ZPBP2</td>
<td>DFLEPLSGLYTC[cm]TLSYK</td>
<td>1004.0</td>
<td>376.2</td>
<td>36.4</td>
<td>20</td>
<td>26.9</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>ZPBP2</td>
<td>DFLEPLSGLYTC[cm]TLSYK*</td>
<td>1008.0</td>
<td>1510.8</td>
<td>36.5</td>
<td>20</td>
<td>26.9</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>ZPBP2</td>
<td>DFLEPLSGLYTC[cm]TLSYK*</td>
<td>1008.0</td>
<td>405.2</td>
<td>36.5</td>
<td>20</td>
<td>26.9</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>ZPBP2</td>
<td>DFLEPLSGLYTC[cm]TLSYK*</td>
<td>1008.0</td>
<td>376.2</td>
<td>36.5</td>
<td>20</td>
<td>26.9</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>ACR</td>
<td>VQPTNVC[cm]AGYPVGK</td>
<td>745.4</td>
<td>400.3</td>
<td>27.6</td>
<td>20</td>
<td>13.0</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>ACR</td>
<td>VQPTNVC[cm]AGYPVGK</td>
<td>745.4</td>
<td>631.8</td>
<td>27.6</td>
<td>20</td>
<td>13.0</td>
<td>Light endogenous</td>
</tr>
<tr>
<td></td>
<td>Sequence</td>
<td>MPA 1</td>
<td>MPA 2</td>
<td>PDI 1</td>
<td>PDI 2</td>
<td>Light endogenous</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>--------</td>
<td>--------------------------------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>------------------</td>
<td>----------</td>
</tr>
<tr>
<td>ACR</td>
<td>VQPTNVC[cm]AGYPVGK</td>
<td>745.4</td>
<td>228.1</td>
<td>27.6</td>
<td>20</td>
<td>13.0</td>
<td></td>
</tr>
<tr>
<td>ACR</td>
<td>VQPTNVC[cm]AGYPVGK*</td>
<td>749.4</td>
<td>408.3</td>
<td>27.8</td>
<td>20</td>
<td>13.0</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>ACR</td>
<td>VQPTNVC[cm]AGYPVGK*</td>
<td>749.4</td>
<td>635.8</td>
<td>27.8</td>
<td>20</td>
<td>13.0</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>ACR</td>
<td>VQPTNVC[cm]AGYPVGK*</td>
<td>749.4</td>
<td>228.1</td>
<td>27.8</td>
<td>20</td>
<td>13.0</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>ACRBP</td>
<td>FFALLTPTWK</td>
<td>612.3</td>
<td>531.3</td>
<td>23.1</td>
<td>20</td>
<td>26.3</td>
<td></td>
</tr>
<tr>
<td>ACRBP</td>
<td>FFALLTPTWK</td>
<td>612.3</td>
<td>295.1</td>
<td>23.1</td>
<td>20</td>
<td>26.3</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>ACRBP</td>
<td>FFALLTPTWK</td>
<td>612.3</td>
<td>366.2</td>
<td>23.1</td>
<td>20</td>
<td>26.3</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>ACRBP</td>
<td>FFALLTPTWK*</td>
<td>616.4</td>
<td>539.3</td>
<td>23.3</td>
<td>20</td>
<td>26.3</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>ACRBP</td>
<td>FFALLTPTWK*</td>
<td>616.4</td>
<td>295.1</td>
<td>23.3</td>
<td>20</td>
<td>26.3</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>ACRBP</td>
<td>FFALLTPTWK*</td>
<td>616.4</td>
<td>366.2</td>
<td>23.3</td>
<td>20</td>
<td>26.3</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>IZUMO3</td>
<td>GVFIYQGK</td>
<td>456.3</td>
<td>755.4</td>
<td>17.8</td>
<td>20</td>
<td>15.6</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>IZUMO3</td>
<td>GVFIYQGK</td>
<td>456.3</td>
<td>608.3</td>
<td>17.8</td>
<td>20</td>
<td>15.6</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>IZUMO3</td>
<td>GVFIYQGK</td>
<td>456.3</td>
<td>495.3</td>
<td>17.8</td>
<td>20</td>
<td>15.6</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>IZUMO3</td>
<td>GVFIYQGK*</td>
<td>460.3</td>
<td>763.4</td>
<td>18.0</td>
<td>20</td>
<td>15.6</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>IZUMO3</td>
<td>GVFIYQGK*</td>
<td>460.3</td>
<td>616.4</td>
<td>18.0</td>
<td>20</td>
<td>15.6</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>IZUMO3</td>
<td>GVFIYQGK*</td>
<td>460.3</td>
<td>503.3</td>
<td>18.0</td>
<td>20</td>
<td>15.6</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>ADAM29</td>
<td>IFGAGANTASETC[cm]YK</td>
<td>795.4</td>
<td>858.4</td>
<td>29.3</td>
<td>20</td>
<td>14.1</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>ADAM29</td>
<td>IFGAGANTASETC[cm]YK</td>
<td>795.4</td>
<td>571.3</td>
<td>29.3</td>
<td>20</td>
<td>14.1</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>ADAM29</td>
<td>IFGAGANTASETC[cm]YK</td>
<td>795.4</td>
<td>470.2</td>
<td>29.3</td>
<td>20</td>
<td>14.1</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>ADAM29</td>
<td>IFGAGANTASETC[cm]YK*</td>
<td>799.4</td>
<td>866.4</td>
<td>29.5</td>
<td>20</td>
<td>14.1</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>ADAM29</td>
<td>IFGAGANTASETC[cm]YK*</td>
<td>799.4</td>
<td>579.3</td>
<td>29.5</td>
<td>20</td>
<td>14.1</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>ADAM29</td>
<td>IFGAGANTASETC[cm]YK*</td>
<td>799.4</td>
<td>478.2</td>
<td>29.5</td>
<td>20</td>
<td>14.1</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>LY6K</td>
<td>ENTFEC[cm]QNPR</td>
<td>647.8</td>
<td>803.3</td>
<td>24.3</td>
<td>20</td>
<td>10.3</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>LY6K</td>
<td>ENTFEC[cm]QNPR</td>
<td>647.8</td>
<td>674.3</td>
<td>24.3</td>
<td>20</td>
<td>10.3</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>LY6K</td>
<td>ENTFEC[cm]QNPR</td>
<td>647.8</td>
<td>272.2</td>
<td>24.3</td>
<td>20</td>
<td>10.3</td>
<td>Light endogenous</td>
</tr>
<tr>
<td>LY6K</td>
<td>ENTFEC[cm]QNPR*</td>
<td>652.8</td>
<td>813.4</td>
<td>24.5</td>
<td>20</td>
<td>10.3</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>LY6K</td>
<td>ENTFEC[cm]QNPR*</td>
<td>652.8</td>
<td>684.3</td>
<td>24.5</td>
<td>20</td>
<td>10.3</td>
<td>Heavy IS</td>
</tr>
<tr>
<td>LY6K</td>
<td>ENTFEC[cm]QNPR*</td>
<td>652.8</td>
<td>282.2</td>
<td>24.5</td>
<td>20</td>
<td>10.3</td>
<td>Heavy IS</td>
</tr>
</tbody>
</table>
Appendix 5.3 List of proteins significantly downregulated in sperm cell samples obtained from TEX101 rs35033974 homozygous men when compared with TEX101 wild-type fertile control men. Proteins were ranked according to their Log2 LFQ Fold-change. Testis-specific, membrane-bound and secreted proteins are in bold. (FDR-adjusted t test p-value <0.05 and fold change >1.5).

<table>
<thead>
<tr>
<th>Gene Names</th>
<th>Protein Names</th>
<th>Log2 LFQ Fold-change (TEX101 WT/TEX101 G99V)</th>
<th>-Log10 FDR-adjusted t-test p-value</th>
<th>Peptide Sequence coverage (%)</th>
<th>Razor + Unique peptides</th>
</tr>
</thead>
<tbody>
<tr>
<td>PMFBP1</td>
<td>Polyamine-modulated factor 1-binding protein 1</td>
<td>-7.0</td>
<td>33</td>
<td>37.2</td>
<td></td>
</tr>
<tr>
<td>CCDC168</td>
<td>Coiled-coil domain-containing protein 168</td>
<td>-6.8</td>
<td>31</td>
<td>20.5</td>
<td></td>
</tr>
<tr>
<td>FAM205C</td>
<td>Protein FAM205C</td>
<td>-6.1</td>
<td>23</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>C5orf60</td>
<td>Putative uncharacterized protein C5orf60</td>
<td>-5.9</td>
<td>16</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>DYNLL1</td>
<td>Dynen light chain 1</td>
<td>-5.8</td>
<td>19</td>
<td>75.3</td>
<td></td>
</tr>
<tr>
<td>CROCC2</td>
<td>Putative cilary rootlet coiled-coil protein 2</td>
<td>-5.7</td>
<td>2.2</td>
<td>4</td>
<td>2.9</td>
</tr>
<tr>
<td>C2orf70</td>
<td>UPF0573 protein C2orf70</td>
<td>-5.6</td>
<td>1.9</td>
<td>4</td>
<td>32.3</td>
</tr>
<tr>
<td>LY6K</td>
<td>Lymphocyte antigen 6K</td>
<td>-5.5</td>
<td>1.7</td>
<td>9</td>
<td>42.4</td>
</tr>
<tr>
<td>FAM71C</td>
<td>Protein FAM71C</td>
<td>-5.5</td>
<td>2.4</td>
<td>8</td>
<td>34.9</td>
</tr>
<tr>
<td>IQCF5</td>
<td>IQ domain-containing protein F5</td>
<td>-5.4</td>
<td>1.8</td>
<td>18</td>
<td>18.9</td>
</tr>
<tr>
<td>IQCF3</td>
<td>IQ domain-containing protein F3</td>
<td>-5.3</td>
<td>2.1</td>
<td>4</td>
<td>23.4</td>
</tr>
<tr>
<td>SLC44A5</td>
<td>Choline transporter-like protein 5</td>
<td>-5.3</td>
<td>2.2</td>
<td>13</td>
<td>22.6</td>
</tr>
<tr>
<td>SMRP1</td>
<td>Spermatid-specific manchette-related protein 1</td>
<td>-5.3</td>
<td>3.4</td>
<td>3</td>
<td>62.2</td>
</tr>
<tr>
<td>PLCZ1</td>
<td>1-phosphatidylinositol 4,5-bisphosphate phospholipase C-zeta-1</td>
<td>-5.2</td>
<td>1.9</td>
<td>16</td>
<td>27.5</td>
</tr>
<tr>
<td>TEX101</td>
<td>Testis-expressed sequence 101 protein</td>
<td>-5.0</td>
<td>2.3</td>
<td>6</td>
<td>31.1</td>
</tr>
<tr>
<td>HBZ</td>
<td>Hemoglobin subunit zeta</td>
<td>-4.6</td>
<td>1.9</td>
<td>2</td>
<td>21.8</td>
</tr>
<tr>
<td>CAPZA3</td>
<td>F-actin-capping protein subunit alpha-3</td>
<td>-4.5</td>
<td>2.0</td>
<td>18</td>
<td>64.5</td>
</tr>
<tr>
<td>APCS</td>
<td>Serum amyloid P-component</td>
<td>-4.5</td>
<td>2.0</td>
<td>10</td>
<td>37.7</td>
</tr>
<tr>
<td>IQCF6</td>
<td>IQ domain-containing protein F6</td>
<td>-4.5</td>
<td>2.1</td>
<td>4</td>
<td>36.4</td>
</tr>
<tr>
<td>DYNLL2</td>
<td>Dynen light chain 2</td>
<td>-4.2</td>
<td>1.7</td>
<td>5</td>
<td>74.2</td>
</tr>
<tr>
<td>DEPD1C</td>
<td>DEP domain-containing protein 1A</td>
<td>-4.1</td>
<td>1.9</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>PRR14L</td>
<td>Protein PRR14L</td>
<td>-4.0</td>
<td>2.8</td>
<td>2</td>
<td>2.3</td>
</tr>
<tr>
<td>CEP70</td>
<td>Centrosomal protein of 70 kDa</td>
<td>-4.0</td>
<td>2.0</td>
<td>8</td>
<td>19.6</td>
</tr>
<tr>
<td>C1orf234</td>
<td>Uncharacterized protein C1orf234</td>
<td>-4.0</td>
<td>4.0</td>
<td>4</td>
<td>34.3</td>
</tr>
<tr>
<td>GAPDH2</td>
<td>Glyceraldehyde-3-phosphate dehydrogenase</td>
<td>-3.9</td>
<td>3.3</td>
<td>33</td>
<td>76.2</td>
</tr>
<tr>
<td>TTC16</td>
<td>Tetratricopeptide repeat protein 16</td>
<td>-3.9</td>
<td>1.9</td>
<td>15</td>
<td>19.1</td>
</tr>
<tr>
<td>GOLGA8F</td>
<td>Putative golgin subfamily A member 8D</td>
<td>-3.9</td>
<td>3.1</td>
<td>1</td>
<td>3.3</td>
</tr>
<tr>
<td>TSPAN1</td>
<td>Tetraspanin-1</td>
<td>-3.8</td>
<td>1.9</td>
<td>3</td>
<td>14.5</td>
</tr>
<tr>
<td>FAM71B</td>
<td>Protein FAM71B</td>
<td>-3.8</td>
<td>1.8</td>
<td>23</td>
<td>53.7</td>
</tr>
<tr>
<td>NDUFB2</td>
<td>NADH dehydrogenase [ubiquitone] 1 beta subcomplex subunit 2</td>
<td>-3.8</td>
<td>3.1</td>
<td>1</td>
<td>8.6</td>
</tr>
<tr>
<td>TMCO2</td>
<td>Transmembrane and coiled-coil domain-containing protein 12</td>
<td>-3.7</td>
<td>2.0</td>
<td>8</td>
<td>42.9</td>
</tr>
<tr>
<td>ARL4A</td>
<td>ADP-ribosylation factor-like protein 4A</td>
<td>-3.7</td>
<td>2.0</td>
<td>3</td>
<td>21</td>
</tr>
<tr>
<td>SPATA31A3</td>
<td>Spermatogenesis-associated protein 31A3</td>
<td>-3.6</td>
<td>2.2</td>
<td>18</td>
<td>25.1</td>
</tr>
<tr>
<td>LRRQ3</td>
<td>Leucine-rich repeat and IQ domain-containing protein 3</td>
<td>-3.6</td>
<td>1.9</td>
<td>10</td>
<td>19.9</td>
</tr>
<tr>
<td>FAM209A</td>
<td>Protein FAM209A</td>
<td>-3.6</td>
<td>2.4</td>
<td>10</td>
<td>59.6</td>
</tr>
<tr>
<td>C2orf16</td>
<td>Uncharacterized protein C2orf16</td>
<td>-3.6</td>
<td>2.3</td>
<td>57</td>
<td>38.3</td>
</tr>
<tr>
<td>GHTM</td>
<td>Growth hormone-inducible transmembrane protein</td>
<td>-3.5</td>
<td>1.8</td>
<td>6</td>
<td>21.2</td>
</tr>
<tr>
<td>SLC2A14</td>
<td>Solute carrier family 2, facilitated glucose transporter member 14</td>
<td>-3.4</td>
<td>2.4</td>
<td>12</td>
<td>25.1</td>
</tr>
<tr>
<td>UQCR10</td>
<td>Cytochrome b-c1 complex subunit 9</td>
<td>-3.3</td>
<td>2.1</td>
<td>4</td>
<td>52.4</td>
</tr>
<tr>
<td>SYT5</td>
<td>Synaptotagmin-5</td>
<td>-3.3</td>
<td>2.1</td>
<td>4</td>
<td>22</td>
</tr>
<tr>
<td>TMEM190</td>
<td>Transmembrane protein 190</td>
<td>-3.3</td>
<td>2.0</td>
<td>5</td>
<td>35.6</td>
</tr>
<tr>
<td>C9orf117</td>
<td>Uncharacterized protein C9orf117</td>
<td>-3.3</td>
<td>2.3</td>
<td>17</td>
<td>38.3</td>
</tr>
<tr>
<td>Gene</td>
<td>Description</td>
<td>Log2 Ratio</td>
<td>Fold Change</td>
<td>Expression Level</td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>--------------------------------------------------</td>
<td>------------</td>
<td>-------------</td>
<td>-----------------</td>
<td></td>
</tr>
<tr>
<td>CCIN</td>
<td>Calcin</td>
<td>-3.3</td>
<td>2.1</td>
<td>36</td>
<td></td>
</tr>
<tr>
<td>SDR39U1</td>
<td>Epimerase family protein SDR39U1</td>
<td>-3.2</td>
<td>2.2</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>TMEM89</td>
<td>Transmembrane protein 89</td>
<td>-3.2</td>
<td>2.1</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>IQCK</td>
<td>IQ domain-containing protein K</td>
<td>-3.2</td>
<td>2.3</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>KNSTRN</td>
<td>Small kinetochore-associated protein</td>
<td>-3.2</td>
<td>2.4</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>YIF1A</td>
<td>Protein YIF1A</td>
<td>-3.2</td>
<td>2.1</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>MORN3</td>
<td>MORN repeat-containing protein 3</td>
<td>-3.2</td>
<td>2.1</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>WDR88</td>
<td>WD repeat-containing protein 88</td>
<td>-3.2</td>
<td>2.5</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>TATDN3</td>
<td>Putative deoxyribonuclease TATDN3</td>
<td>-3.1</td>
<td>2.0</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>TSGA10</td>
<td>Testis-specific gene 10 protein</td>
<td>-3.1</td>
<td>1.9</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>GSG1</td>
<td>Germ cell-specific gene 1 protein</td>
<td>-3.1</td>
<td>2.4</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>CYLC2</td>
<td>Cylicin-2</td>
<td>-3.1</td>
<td>1.9</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>FLJ43738</td>
<td>Uncharacterized protein FLJ43738</td>
<td>-3.1</td>
<td>3.0</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>CYCS</td>
<td>Cytochrome c</td>
<td>-3.0</td>
<td>1.9</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>CCDC159</td>
<td>Coiled-coil domain-containing protein 159</td>
<td>-3.0</td>
<td>4.2</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>GLB1L</td>
<td>Beta-galactosidase-1-like protein</td>
<td>-3.0</td>
<td>3.0</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>RND2</td>
<td>Rho-related GTP-binding protein RhoN</td>
<td>-2.9</td>
<td>1.9</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>H2AFB1</td>
<td>Histone H2A-Bbd type 2/3</td>
<td>-2.9</td>
<td>2.0</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>FNDC8</td>
<td>Fibrinectin type III domain-containing protein 8</td>
<td>-2.9</td>
<td>1.9</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>SUN5</td>
<td>SUN domain-containing protein 5</td>
<td>-2.9</td>
<td>1.9</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>ITPR1PL</td>
<td>Inositol 1,4,5-triphosphate receptor-interacting protein-like 1</td>
<td>-2.9</td>
<td>2.5</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>ANTXRL</td>
<td>Anthrax toxin receptor-like</td>
<td>-2.9</td>
<td>2.1</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>ACR</td>
<td>Acrosin</td>
<td>-2.9</td>
<td>2.9</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>CPA5</td>
<td>Carboxypeptidase A5</td>
<td>-2.9</td>
<td>2.1</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>ZPB2</td>
<td>Zona pellucida-binding protein 2</td>
<td>-2.9</td>
<td>2.4</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>CPTP</td>
<td>Ceramide-1-phosphate transfer protein</td>
<td>-2.8</td>
<td>2.2</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>CPVL</td>
<td>Probable serine carboxypeptidase CPVL</td>
<td>-2.8</td>
<td>3.4</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>TSSK2</td>
<td>Testis-specific serine/threonine-protein kinase 2</td>
<td>-2.8</td>
<td>2.4</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>CD177</td>
<td>CD177 antigen</td>
<td>-2.8</td>
<td>2.5</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>ABCB7</td>
<td>ATP-binding cassette sub-family B member 7</td>
<td>-2.8</td>
<td>2.3</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>EFCAB3</td>
<td>EF-hand calcium-binding domain-containing protein 3</td>
<td>-2.7</td>
<td>1.9</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>TCP11</td>
<td>T-complex protein 11 homolog</td>
<td>-2.7</td>
<td>2.1</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>SSEM8M1</td>
<td>Serine-rich single-pass membrane protein 1</td>
<td>-2.6</td>
<td>2.0</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>SPACA3</td>
<td>Sperm acrosome membrane-associated protein 3</td>
<td>-2.6</td>
<td>2.4</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>LYZL4</td>
<td>Lysozyme-like protein 4</td>
<td>-2.6</td>
<td>2.3</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>IZUMO3</td>
<td>Izumo sperm-egg fusion protein 3</td>
<td>-2.6</td>
<td>2.4</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>FSD1</td>
<td>Fibronectin type III and SPRY domain-containing protein 1</td>
<td>-2.6</td>
<td>2.2</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>SPATA31A1</td>
<td>Spermatogenesis-associated protein 31A1</td>
<td>-2.6</td>
<td>2.2</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>SPAM1</td>
<td>Hyaluronidase PH-20</td>
<td>-2.6</td>
<td>2.0</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>ACRBP</td>
<td>Acrosin-binding protein</td>
<td>-2.6</td>
<td>3.4</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>ADCY10</td>
<td>Adenylyl cyclase type 10</td>
<td>-2.6</td>
<td>2.3</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>GLTID1</td>
<td>Glycorytransferase 1 domain-containing protein 1</td>
<td>-2.5</td>
<td>2.9</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>B4GALNT3</td>
<td>Beta-1,4-N-acetylgalactosaminyltransferase 3</td>
<td>-2.5</td>
<td>2.0</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>CBLN1</td>
<td>Cerebellin-1</td>
<td>-2.5</td>
<td>2.1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>STK16</td>
<td>Serine/threonine-protein kinase 16</td>
<td>-2.5</td>
<td>2.2</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>RAB2A</td>
<td>Ras-related protein Rab-2A</td>
<td>-2.4</td>
<td>2.9</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>TOMM7</td>
<td>Mitochondrial import receptor subunit TOM7 homolog</td>
<td>-2.4</td>
<td>2.2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>CALM1</td>
<td>Calmodulin</td>
<td>-2.4</td>
<td>2.1</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>IL4I1</td>
<td>L-amino-acid oxidase</td>
<td>-2.3</td>
<td>3.2</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>LMMRD1</td>
<td>Probable lysosomal cobalamin transporter</td>
<td>-2.3</td>
<td>2.4</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>XKR3</td>
<td>XK-related protein 3</td>
<td>-2.2</td>
<td>2.5</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>RPL23</td>
<td>60S ribosomal protein L23</td>
<td>-2.2</td>
<td>2.2</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>NDRG2</td>
<td>Protein NDRG2</td>
<td>-2.1</td>
<td>2.5</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>TMEM167A</td>
<td>Protein kish-A</td>
<td>-2.1</td>
<td>2.3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>EMD</td>
<td>Emerin</td>
<td>-2.0</td>
<td>2.4</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>FAM78A</td>
<td>Protein FAM78A</td>
<td>-2.0</td>
<td>2.3</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>ARL8B</td>
<td>ADP-ribosylation factor-like protein 8B</td>
<td>-1.8</td>
<td>2.6</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>ARL1</td>
<td>ADP-ribosylation factor-like protein 1</td>
<td>-1.7</td>
<td>3.3</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>ENPP1</td>
<td>Ectonucleotide pyrophosphatase/phosphodiesterase family member 1</td>
<td>-1.5</td>
<td>4.6</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>
Appendix 5.4 ADAM proteins identified. Three ADAM family proteins (ADAM18, ADAM29, ADAM32), with adhesion activity, were identified in our sperm proteome, however, they were not selected as candidate according to the applied criteria. Only ADAM29 was closer to our criteria, therefore it was further verified along with the other candidates. We speculate that ADAM29 may be the ortholog of one of the murine ADAM3-6 proteins, which are affected in Tex101 knockout mice.