Investigation of epigenetic factors TET1-3 and PRC2 in spermatogenesis and mature spermatozoa: role in male infertility

Inaugural Dissertation

submitted to the

Faculty of Medicine

in fulfilment of the requirements

for the PhD-Degree

of the Faculties of Veterinary Medicine and Medicine

of the Justus-Liebig-University Giessen

by

Shiplu Shanjid Ahmed

Giessen, 2024

From the Research Laboratory of Molecular Andrology

(Head: Prof. Dr. rer. nat. U. Schagdarsurengin)

at the Department of Urology, Pediatric Urology and Andrology

(Head: Prof. Dr. med. F. Wagenlehner)

Faculty of Medicine, Justus-Liebig-University Giessen

First Supervisor and Committee Member:

Prof. Dr. rer. nat. U. Schagdarsurengin

Second Supervisor and Committee Member:

Prof. Dr. Christine Wrenzycki

Examination committee members:

1st reviewer: Prof. Dr. Undraga Schagdarsurengin

2nd reviewer: PD Dr. Melanie Freifrau von Brandenstein

Examination vice-chair: Prof. Dr. Daniela Fietz

Examination chair: Prof. Dr. Martin Diener

Date of Doctoral Defence

30 October 2024

Abstract

This study investigated the role of aberrant epigenetic mechanisms in spermatogenesis and male infertility. We examined the expressions and retention of TET1, PRC2 components, and H3K27me3 during human and mouse spermatogenesis and the mRNA and methylation levels of their target genes in mature human sperm from fertile and infertile/subfertile men.

We first studied the impact of *Tet1* knockout on PRC2 and H3K27me3 and its implications for male fertility. In *Tet1*-deficient mice, Eed was absent, and Ezh2 and Suz12 expressions shifted to early and late stages of spermatogenesis respectively. As *Eed* knockout mice are infertile and Eed helps establish H3K27me3 through PRC2, and this mark is passed from sperm to the next generation, these results and altered H3K27me3 expression suggest Tet1's indirect role in male infertility.

Next, we examined the retention of TET1, H3K27me3, and PRC2 components at the protein level in human motile spermatozoa. In patients' spermatozoa, we found aberrant retention of H3K27me3 and a trend of higher enrichment of H3K27me3 in selected genes of developmental relevance (GDR).

Further analysis revealed dysregulation of *TETs*, PRC2 components and selected GDR in subfertile patients. Particularly, *TET1, TET3, SOX1-2, and OCT4* mRNA levels were significantly upregulated in subfertile patients. Additionally, patients exhibited significantly lower mRNA levels of *EED, FOXA2, HERV-w* and dysregulated promoter methylation levels of *EZH2, SUZ12* and *GATA2*.

Our data suggest that high *TET1* and *TET3* mRNA levels and over-retention of H3K27me3 are characteristics of infertile patients. These aberrations were linked with the overexpression and hypomethylation of a few GDR in spermatozoa. As the affected GDR code transcription factors, which are critical for embryonic genome activation (EGA) and preimplantation embryogenesis, sperm-inherited RNA and proteins may play a crucial role in early embryonic development, i.e. their aberrations may impact embryo development, implantation, and cause unsuccessful pregnancy.

In conclusion, our study has established a preliminary link between TETs, PRC2, H3K27me3 and GDR, and suggest that their interplay is important for governing EGA and embryogenesis, and herewith also for male fertility.

Zusammenfassung

In dieser Studie wurde die Rolle abweichender epigenetischer Mechanismen bei der Spermatogenese und männlicher Unfruchtbarkeit untersucht. Wir untersuchten die Expression und Retention von TET1, PRC2-Komponenten und H3K27me3 während der humanen und Maus-Spermatogenese sowie die mRNA- und DNA-Methylierung von deren Target-Genen in reifen humanen Spermien von fertilen und infertilen/subfertilen Männern.

Wir haben zunächst die Auswirkungen des Tet1-Knockouts auf PRC2 und H3K27me3 sowie deren Auswirkungen auf die männliche Fruchtbarkeit untersucht. Bei Tet1-defizienten Mäusen fehlte Eed, und die Expressionen von Ezh2 und Suz12 verschoben sich jeweils in frühe und späte Stadien der Spermatogenese. Da Eed-Knockout-Mäuse unfruchtbar sind und Eed bei der Etablierung von H3K27me3 durch PRC2 hilft, und diese Markierung von Spermien auf die nächste Generation übertragen wird, legen diese Ergebnisse und die veränderte Expression von H3K27me3 eine indirekte Rolle von Tet1 in der männlichen Unfruchtbarkeit nahe.

Als Nächstes untersuchten wir die Retention von TET1, H3K27me3 und PRC2-Komponenten auf Proteinebene in menschlichen motilen Spermatozoen. In den Spermien von Patienten fanden wir eine abweichende Retention von H3K27me3 und einen Trend zu einer höheren Anreicherung von H3K27me3 in ausgewählten Genen, die für die embryonale Entwicklung relevant sind (genes of developmental relevance, GDR).

Weitere Analysen ergaben eine Dysregulation von *TETs*, PRC2-Komponenten und ausgewählten GDR bei subfertilen Patienten. Insbesondere die mRNA-Spiegel von *TET1, TET3, SOX1-2* und *OCT4* waren bei subfertilen Patientinnen signifikant hochreguliert. Darüber hinaus wiesen Patienten signifikant niedrigere mRNA-Spiegel *von EED, FOXA2, HERV-w* und dysregulierte Promotor-Methylierung in *EZH2, SUZ12* und *GATA2* auf.

Unsere Daten legen nahe, dass hohe *TET1*- und *TET3*-mRNA-Spiegel sowie eine Überretention von H3K27me3 in motilen Spermien charakteristisch für unfruchtbare Patienten sind. Diese Abweichungen wurden mit einer Überexpression und Hypomethylierung von einigen GDR in Spermien in Verbindung gebracht. Da die betroffenen GDR Transkriptionsfaktoren kodieren, die für die Aktivierung des embryonalen Genoms (EGA) und die präimplantäre Embryogenese entscheidend sind, könnten spermienvererbte RNA und Proteine eine wichtige Rolle in der frühen embryonalen Entwicklung spielen, und deren epigenetischen Aberrationen sich negativ auf die ganze Embryonalentwicklung, die Implantation bzw. Schwangerschaft auswirken.

Zusammenfassend hat unsere Studie eine vorläufige Verbindung zwischen TETs, PRC2, H3K27me3 und GDR hergestellt und legt nahe, dass ihr Zusammenspiel wichtig ist für die Steuerung von EGA und Embryogenese und damit auch für die männliche Fruchtbarkeit.

Declaration

I hereby declare that the PhD thesis titled "Investigation of Epigenetic Factors TET1-3 and PRC2 in Spermatogenesis and Mature Spermatozoa: Role in Male Infertility" is my work and has been written by me under the guidance of my thesis advisor and by using Grammarly program as supporting tool.

This work has not been submitted for any degree or examination in any other university. The research reported in this thesis, unless otherwise indicated, is my original research. I have acknowledged all the sources of information which have been used in the thesis.

To the best of my knowledge and belief, this thesis does not contain text, graphics, or tables previously published or written by another person, unless specifically acknowledged, and the source is detailed in the thesis and in the References sections.

I have abided by the principles of good scientific conduct laid down in the charter of the Justus Liebig University of Giessen in carrying out the investigations described in the dissertation.

_____

Shanjid Ahmed Shiplu

Date: 12 November 2024

List of Abbreviations

Abbreviation	Description
APS	Ammonium persulfate
ART	Assisted reproductive technologies
ALKBH5	AlkB Homolog 5, RNA Demethylase
bp	Base pair
BSA	Bovine serum albumin
ChIP	Chromatin immunoprecipitation
cm	Centimetre
Ct	Cycle threshold
CFTR	Cystic fibrosis transmembrane conductance regulator
COBRA	Combined-bisulphite-restriction analysis
kDa	Kilo dalton
ddH2O	Distilled water
DEPC	Diethypyrocarbonate
DDR1	Discoidin domain receptor 1
DMRs	Differentially methylated regions
DNA	Desoxyribonucleic acid
DNase I	Deoxyribonuclease I
DNMT	DNA methyltransferase
dNTP	2'-deoxynucleoside-5'-triphosphate
DTT	Dithiothreitol
EZH	enhancer of Zeste 1 or 2
EED	embryonic ectoderm development
EGA	Embryonic genome activation
EDTA	Ethylenediaminetetraacetic acid
EtOH	Ethanol
F	Forward primer
Fig	Figure
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GDR	Genes of developmental relevance
h	Hour
HeLa	Human cervix carcinoma cell line
hESC	Human embryonic stem cells
H2O2	Hydrogen peroxide
hmC	hydroxymethylcytosine
ddH2O	double-distilled water
H3K27me3	histone 3 lysine 27 trimethylation
HCl	Hydrochloric acid
ICSI	Intracytoplasmic sperm injection
ICRs	Imprinting control regions
IGF2	Insulin-like Growth Factor 2
IHC	Immunohistochemistry
IVF	In vitro fertilization
kDa	in Kilodalton
L1	Long interspersed nuclear element 1

L	Litre
mL	Milliliter
M	Molar (mol/L)
MeOH	Methanol
MEST	Mesoderm-specific transcript homologue
MTHFR	Methylenetetrahydrofolate reductase
MI	Male infertility
mg	Milligram
min	Minute
MW	Molecular weight
METTL3	Methyltransferase 3, N6-Adenosine-methyltransferase Complex Catalytic Subunit
n	Number
NaAc	Sodium acetate
NaCl	Sodium chloride
NaOH	Sodium hydroxide
nm	Nanometer
NP-40	Nonyl phenoxypolyethoxylethanol
ns	Not significant
NGS	Next-generation sequencing
PM	Progressive Motility
PRC2	Polycomb repressive complex 2
PBS	Phosphate buffered saline
PFA	Paraformaldehyde
PMSF	Phenylmethylsulfonyl fluoride
PTM	Post-translational modification
PVDF	Polyvinylidene difluoride
R	Reverse primer
RNA	Ribonucleic acid
RNAse	Ribonuclease
RT	Room temperature
RT-qPCR	Real time-quantitative polymerase chain reaction
rpm	Rotation per minute
s	Second
SINE	Short interspersed nuclear element
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
STIs	Sexually transmitted infections
SSCs	Spermatogonial stem cells
SC	Sertoli cell
SCO	Sertoli cell only
SUZ12	Suppressor of Zeste 12
T	Temperature
TEMED	Tetramethylethylenediamine
TET	Ten-eleven translocation
TRIS	Tris-(hydroxymethyl)-amino methane
TBS	Tris-buffered saline
V	Volt

WB	Western blot
WHO	World Health Organization
x	Time(s)

Table of contents

1. Introduction

Fertility refers to the natural ability of an organism to (re)produce offspring. According to the World Health Organization (WHO), infertility is defined as the inability of a sexually active couple to conceive within 12 months of regular, unprotected sexual intercourse [1]. Some clinicians use the term infertility interchangeably with subfertility. However, any form of reduced fertility with a prolonged time of unwanted non-conception is generally described as subfertility. And sterility means permanent infertility. In normal circumstances, having intercourse within the first six cycles in the fertile phase, most of the pregnancies (80%) occur [2,3]. Globally, 8–12% of reproductive-aged couples have infertility problems. Approximately 187 million couples worldwide, which is equivalent to one out of every six couples of reproductive age, are affected by this condition, making it a major global health concern. According to a Global Burden of Disease survey, the age-standardized prevalence of infertility increased annually by 0.370% in women and 0.291% in men between 1990 and 2017 [4,5].

Male fertility refers to the ability of a man to father a child biologically by producing healthy sperm, delivering it into the female reproductive tract during sexual intercourse and fertilizing an egg in the fallopian tube. Males are found to be solely responsible for around 30% of infertility cases but the overall contribution is approximately 50%. Emerging evidence suggests that the health of a man at the time of conception can impact the metabolic health and reproductive potential of their offspring through the transgenerational transmission of epigenetic modifications. Consequently, conditions such as obesity or diabetes may not only contribute to male subfertility but also potentially compromise the health of future generations. A study of 744 infertile men revealed that 15.4% of men with prediabetes symptoms were at an increased risk of hypogonadism, higher sperm DNA fragmentation, and non-obstructive azoospermia. In addition, men with oligozoospermia are more likely to have metabolic syndrome than those with normozoospermia [2,3,5–7]. The data presented underscores the significance of investigating male infertility (MI).

MI may occur due to testicular and post-testicular deficiencies. MI is a multifactorial disease which has been associated with spermatogenic aberrations as well as infertility-related factors for example (epi)genetics, environment, diet etc. MI can be affected by obesity, smoking, heavy alcohol or drug use, exposure to toxins or radiation, and certain medical conditions as presented in Fig. 1. [4,8,9]. About 90% of male infertility (MI) cases are due to low sperm count or poor quality. Causes include hormonal imbalances, infections, genetic disorders, and certain

medications [22,24]. Anatomical problems, hormonal imbalances, and genetic defects are also common causes. Other factors include endocrine-disrupting chemicals and consanguinity [1]. The presence of white blood cells in semen indicates infection. Poor sperm viability and ejaculatory problems also contribute to MI [4,8,9].

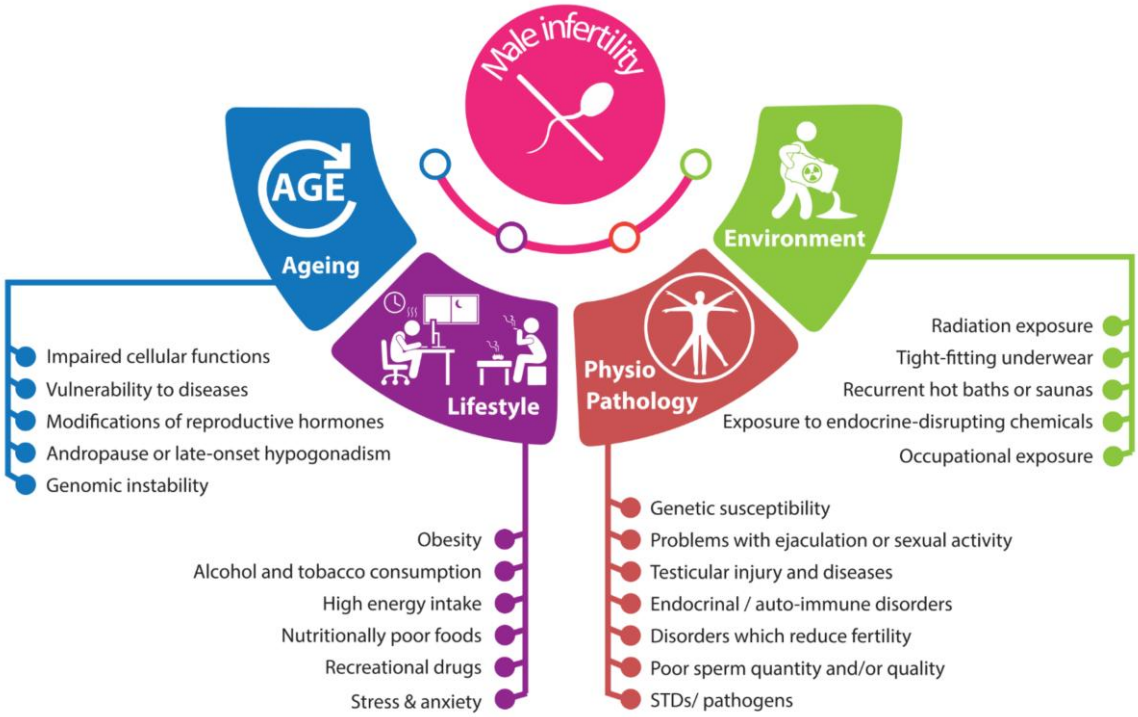


Figure 1: **Aetiologies of male infertility**. Male fertility complications stem from a range of complex factors, which can be classified into four distinct groups. The sperm epigenome is influenced by a variety of factors from all mentioned groups, including environmental exposures, and paternal lifestyle. STDs: sexually transmitted diseases. (The image was adapted [1])

However, in around 30% of MI cases, the aetiology remains unknown and is called idiopathic infertility. Idiopathic infertility might arise from various factors, including impaired spermatogenesis, as well as genetic and epigenetic factors which may be the primary contributors to the condition [2,10]. A nine-year prospective study involving 1,737 patients found that approximately 75% of men with oligozoospermia presented idiopathic MI. Epigenetics may provide insight into this issue [11,91]. Aberration in epigenetic mechanisms like DNA methylation of imprinted genes, especially those related to reproduction, has been observed in various studies [12–16]. This abnormal methylation could help elucidate the causes of idiopathic infertility. In addition, associations between infertility and epigenetic mechanisms such as various histone modifications and short non-coding RNAs (sncRNA) have been also demonstrated. These epigenetic mechanisms are vital for imprinting and reprogramming in early embryos. During fertilization, the sperm transfers its genetic and epigenetic information

to the oocyte [17–19]. Therefore, the mechanisms of sperm epigenetics appear to be a promising area of research for studying idiopathic MI [8].

1.1 Mammalian Spermatogenesis

Spermatogenesis, vital for male fertility, is the process of producing mature sperm from undifferentiated spermatogonial stem cells (SSCs) in the testes [1]. It involves mitotic and meiotic divisions and is studied through "Staging," categorizing germ cells based on spermatids' transformations [20].

1.1.1 Seminiferous epithelial cycle in man

In humans, spermatogenesis starts around the age of 12 and lasts about 74 days on average, but it can vary from 42 to 76 days in healthy men [19,21]. The expected daily sperm production can range from 150 to 275 million spermatozoa per man [22], or 4.4 million per gram of testis tissue [23]. The seminiferous epithelium cycle, divided into mitosis, meiosis, and spermiogenesis stages, represents interactions between Sertoli cells (SCs) and developing germ cells. Spermiogenesis involves transforming round spermatids into motile spermatozoa. SCs support germ cells throughout the cycle [17–19].

The seminiferous epithelium cycle releases new sperm every 16 days, involving cell division and a resting period. This adds new cells to the epithelium periodically. The germinal epithelium of the seminiferous tubules has cells at different developmental stages, indicating simultaneous spermatogenesis waves. Cross-sectional examination of seminiferous tubules reveals six cellular associations that represent various stages of germ cell development [24].

These six stages are represented in Fig. 2 along with eight distinct steps of spermiogenesis: a) Stage I: Early round spermatids with acrosomal vesicles, including Type A and Type B spermatogonia, pachytene primary spermatocytes, round (step 1), and elongating (step 7) spermatids. b) Stage II: Presence of Type A and Type B spermatogonia, pachytene primary spermatocytes, round (step 2), elongated (step 8) spermatids, and the formation of residual bodies within SCs, followed by spermiation. c) Stage III: Nuclear condensation of spermatids begins at step 3, and Type B spermatogonia enters meiosis at the preleptotene stage. d-e) Stages IV and V: Continued nuclear condensation (step 4 and step 5) with differentiation based on the

presence of leptotene primary spermatocytes in stage IV and zygotene primary spermatocytes in stage V, leading to the first meiotic division. f) Stage VI: Presence of secondary spermatocytes that rapidly undergo the second meiotic division after a brief interphase of approximately 6 hours. As a result, stage VI is uncommonly observed [25].

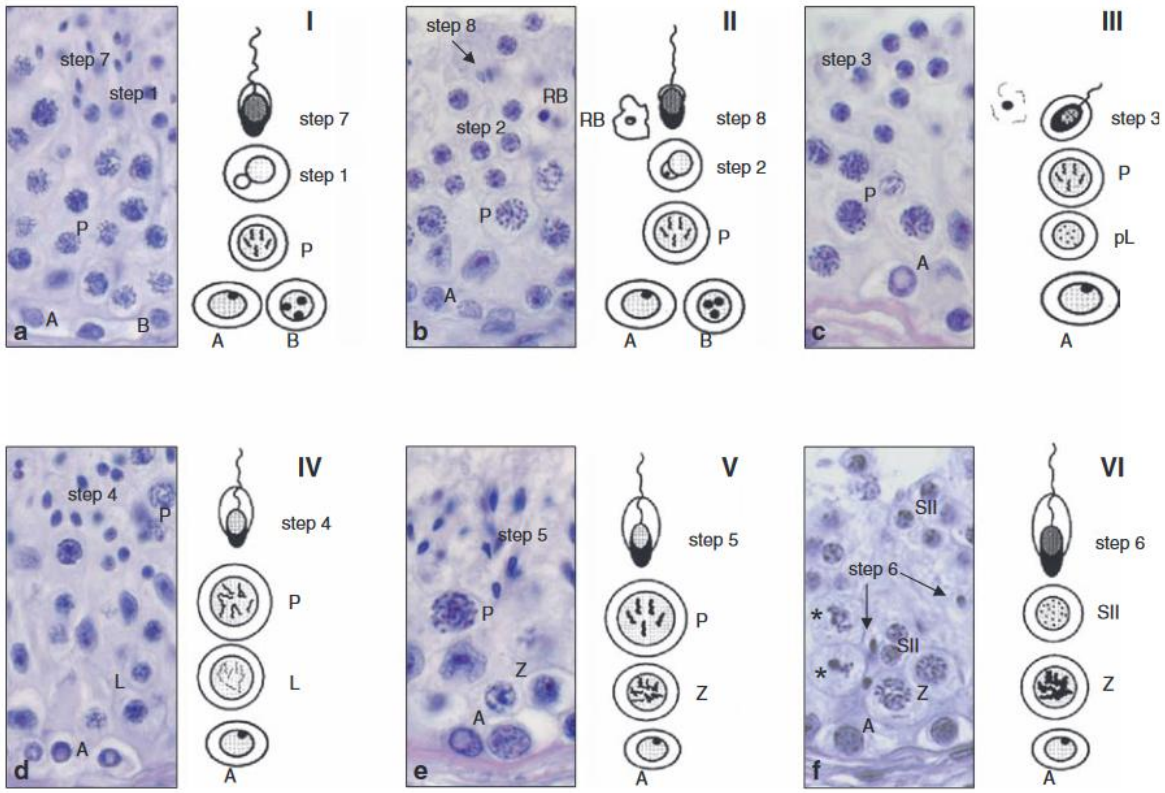


Figure 2: Human normal spermatogenesis with six stages. All stages and steps were visualized using haematoxylin and eosin staining in the Bouin fixed, paraffin-embedded human testis section. The stages of spermatogenesis are designated by Roman numbers from I to VI. Here in the image (a) Stage I, (b) Stage II, (c) Stage III, (d) Stage IV, (e) Stage V, (f) Stage VI. A, B: spermatogonia type A, type B; P: pachytene primary spermatocytes, L: leptotene primary spermatocytes, Z: zygotene primary spermatocytes; SII: secondary spermatocytes; step1–8: different steps of spermiogenesis; RB: residual body; *cells in the first meiotic division. Magnification 40x. (The image was adapted [25]).

1.1.2 Seminiferous epithelial cycle in mouse

The seminiferous epithelial cycle is similar in humans and mice but with differences in timing and duration. In contrast to humans, mice have a short prepubertal period of approximately three weeks. Mice have a shorter prepubertal period and complete one cycle in 8.6 days, with around 35 days for the overall differentiation process [17]. A tubule can contain four to five generations of germ cells at a time. This makes the isolation and molecular characterisation of individual sub-stages exceedingly difficult during spermatogenesis [26]. Gonocytes transition to SSCs or develop into type A1 spermatogonia early in the cycle. Type-B spermatogonia undergo

mitosis to produce preleptotene spermatocytes, which then produce spermatozoa. Meiosis in mice involves two consecutive cell divisions, DNA double-strand break formation, homologous recombination, and chromosomal synapsis, taking around 14 days.

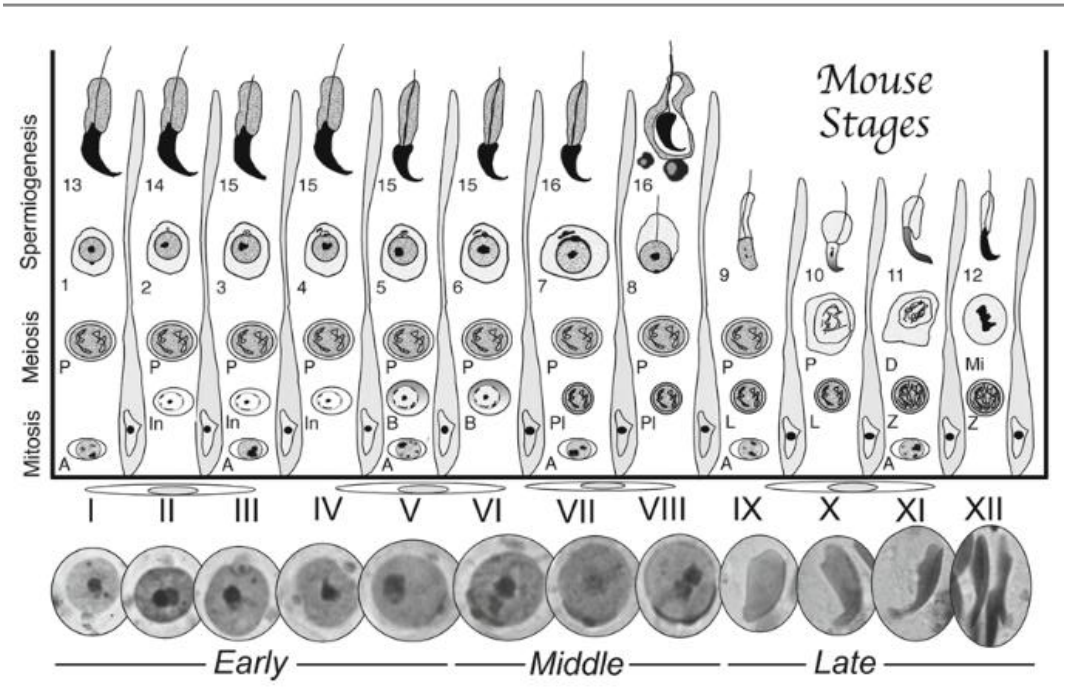


Figure 3: **Schematic representation of mouse seminiferous epithelial cycle**. The figure shows all stages (denoted with Roman numbers) and steps (denoted with alphanumeric numbers) of mouse germ cells during spermatogenesis. Each layer illustrates the cellular associations with SCs separating each stage. At the bottom, image of early, middle, and late spermatid nuclei, stained with the periodic acid-schiff reaction and haematoxylin. Spermatogonia (A, In, B); spermatocytes (Pl: preleptotene, L: leptotene, Z: zygotene, P: pachytene, D: diakinesis, Mi: meiotic division); round spermatids (1-8). (The image was adapted [20]).

Mouse spermatogenesis has twelve stages and spermiogenesis has sixteen steps. As spermatogenesis is a continuous process, therefore, transitional areas may be present between two stages. To overcome this, researchers often rely on the predominant cell types present to identify the stage. Identification of specific stages requires higher-resolution microscopy. However, for most research purposes, it is sufficient to categorize the stages into three groups, making the evaluation process much simpler (Fig. 3). For instance, stages I-V can be grouped as 'early', stages VI-VIII as 'middle', and stages IX-XII as 'late' [20,23,26].

1.1.3 Spermatogenic impairment in man

Spermatogenesis, defined by the stage arrangement of the seminiferous epithelium, can be impaired at various stages. Impairments can occur during spermatogonia to spermatocyte differentiation, meiosis, and spermiogenesis. Lack of growth factors can lead to Sertoli cell only (SCO) syndrome, causing spermatogenic arrest. "Maturation arrest" can occur at later stages, resulting in infertility. Impairments during meiosis can cause azoospermia or oligozoospermia. Hypospermatogenesis is a condition with low numbers of elongated spermatids. Genetic defects can also decrease sperm production. The most complex stage, spermiogenesis, can produce abnormally shaped spermatozoa when impaired. Mixed atrophy of spermatogenesis is a condition where different types of impaired spermatogenesis (e.g., Hyp, SCO, or total atrophy with only a thickened lamina propria) coexist in the same testis. Reduced motility of spermatozoa (asthenozoospermia) is another common feature of infertile men. Table 1 shows the normal semen parameters according to the WHO, based on semen samples from men whose female partners conceived naturally within a year. Table 2 defines various terms for the quality of sperm in semen and with their reference values [27,28].

Table 1: Reference values for normal human semen parameters according to WHO [29]

Semen Parameters	WHO (2010)	WHO (2021)
Semen volume (mL)	1.5 (1.4–1.7)	1.4 (1.3–1.5)
Total sperm number (10^6 per ejaculate)	39 (33–46)	39 (35–40)
Total motility (%)	40 (38–42)	42 (40–43)
Progressive motility (%)	32 (31–34)	30 (29–31)
Non-progressive motility (%)	1	1 (1–1)
Immotile sperm (%)	22	20 (19–20)
Vitality (%)	58 (55–63)	54 (50–56)

WHO: World Health Organization.

Table 2: Description of human semen abnormalities [27,28]

Name of condition	Phenotype
Sertoli cell only	No germ cells
Azoospermia	No spermatozoa in the ejaculate
Oligozoospermia	Less than 15 million/mL sperm
Asthenozoospermia	Fewer than 40% of the sperm are moving and less than 32% are swimming progressively
Teratozoospermia	Less than 4% of spermatozoa have a normal morphology
Dysfunctional spermatozoa	Morphologically normal spermatozoa with dysfunctional molecules

1.2 Epigenetics and sperm epigenome

Epigenetics involves mechanisms that alter gene expression without changing the DNA sequence. These include DNA methylation, histone modifications, and chromatin structure changes. These are specific to cell type, tissue, organ, sex, species, and development stage. Epigenetic modifications can be inherited and sometimes reversed and assisted reproductive technologies have raised concerns about links to genomic imprinting disorders [9,22,30,31]. During spermatogenesis, various precise epigenetic modifications occur to create specialized mature sperm cells, which involves significant reorganization of chromatin structure. As a result, spermatogenesis is highly susceptible to epigenetic changes, and any dysregulation (e.g., DNA methylation, histone modification) in this remodelling process can lead to infertility due to abnormal expression of target genes. Some of the possible aberrations during spermatogenesis are mentioned in Fig. 4 [11,91].

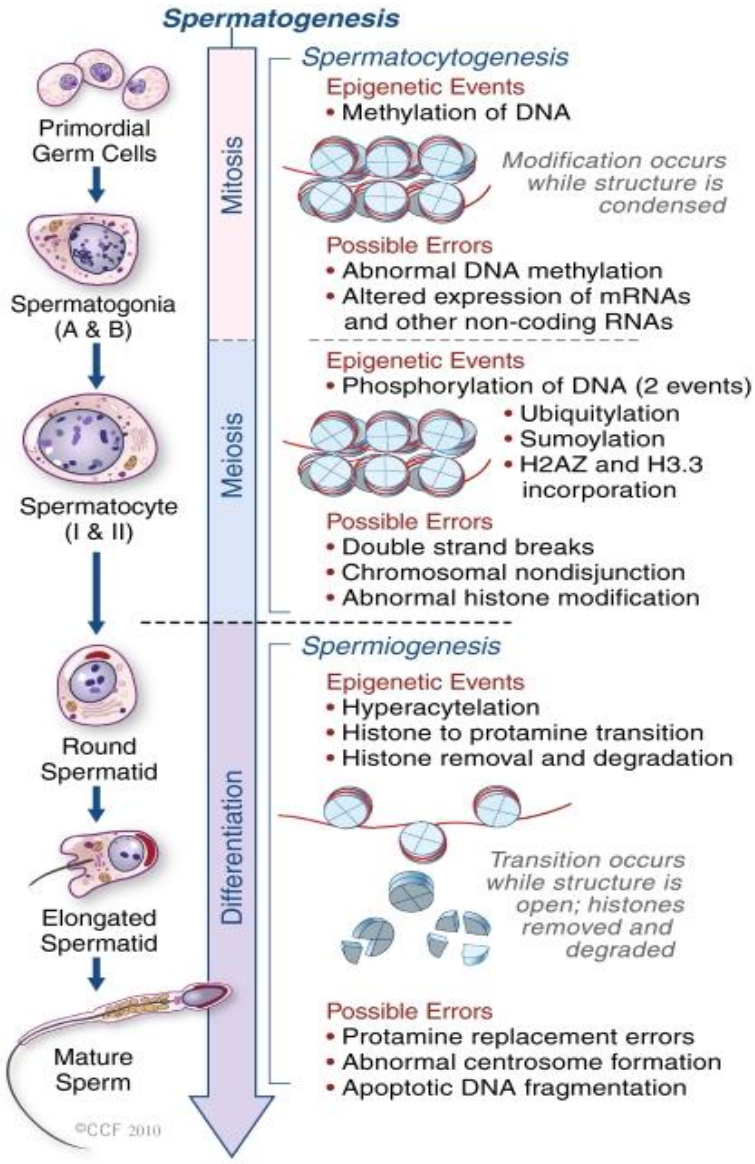


Figure 4: **Major epigenetic events along with possible errors during spermatogenesis**. During spermatogenesis, various epigenetic events take place to ensure the proper development and maturation of sperm cells. In primordial germ cells, DNA methylation occurs during mitosis to establish paternal-specific imprints. In meiotic cells, phosphorylation aids in both recombination and XY body formation. Additionally, XY body formation involves ubiquitylation, sumoylation, and the incorporation of H2AZ and H3.3 variants. Hyperacetylation takes place during spermiogenesis to facilitate the exchange of histones and protamines. However, spermatocytogenesis can also result in chromosome non-disjunction during meiosis I and II, along with double-strand breaks, abnormal histone modification, and changes in the expression of mRNA and other non-coding RNAs. Furthermore, apoptosis can occur following double-strand breaks or abnormal protamination during spermiogenesis, leading to DNA fragmentation. (the image was adapted [32]**).**

1.2.1 Epigenetic mechanisms and the role of epigenetic aberrations in the establishment of male infertility

Impairment of spermatogenesis and male fertility can be caused by aberrant epigenetic mechanisms, which are outlined in Fig. 5. These aberrant mechanisms might be the root cause of the majority of infertility cases that are deemed "unexplained" or "idiopathic" [9,22,30,31]. DNA

methylation, post-translational modifications (PTM) of histone tails and ncRNA (e.g., miRNA) regulation are the most described epigenetic mechanisms in sperm [32,33].

1.2.1.1 DNA-methylation

DNA methylation, a heritable epigenetic mark, involves adding a methyl group to the cytosine ring in DNA. DNA methylation is facilitated by DNA methyltransferase enzymes, while DNA demethylation occurs through ten-eleven translocation (TET) enzymes [34]. Over 98% of DNA methylation occurs at CpG dinucleotides within CpG islands in gene promoters, causing stable, heritable transcriptional silencing. These islands are genomic regions about 500 base pairs long with a high frequency of CpG sites (CG to GC ratio) >55%. These stretches of DNA are located within the promoter region of about 40% of mammalian genes and when methylated, cause stable, heritable transcriptional silencing, as it can prevent transcription factors from binding to the DNA. Both hypomethylation and hypermethylation can occur simultaneously in different genomic regions. Promoters with high CpG content are enriched in nucleosomes, while those with low CpG content lack nucleosomes. This nucleosome enrichment at CpG-rich sequences lacks DNA methylation throughout the genome. Interestingly, these DNA hypomethylated promoters that are present in mature sperm cells are strongly associated with the promoters that are bound by transcription and signalling factors involved in the self-renewal network in human embryonic cells [15,35] and suggesting a significant role of these sperm genomic regions in embryonic development as DNA methylation is heritable mark.

DNA methylation is also crucial for genomic imprinting, where genes are expressed in a parent-of-origin-specific manner. It involves the addition of a methyl group to the DNA and suppressing the expression of genes, which can result in the differential expression of genes inherited from the mother and/or father. Impaired maintenance of imprinted gene methylation in the germline can lead to poor sperm quality, low pregnancy rates, and impaired post-fertilization development. Studies have identified several genes showing differential DNA methylation patterns that are associated with impaired spermatogenesis and reproductive dysfunction [36]. Imprinting takes place in particular DNA sequences known as differentially methylated regions (DMRs) or imprinting control regions (ICRs). An ICR determines whether the imprinted genes should be expressed or not, according to the parental allele [37].

Moreover, sperm cells have incredibly unique DNA methylation patterns which start in the early stages of spermatogenesis and are essential for this process [8,9]. In idiopathic infertile men,

hypomethylation of *IGF2/H19* and hypermethylation of MEST are linked to reduced sperm counts and low sperm motility respectively [38]. DNA methylation defects in these genes, as well as *MTHFR*, are associated with male infertility [36]. Methylenetetrahydrofolate reductase (MTHFR) regulates DNA replication and methylation. Inactivation of *MTHFR* results in male infertility and hyper-homocysteinemia [38]. In non-obstructive azoospermia patients, a significant association between the methylation status of the *MTHFR* gene and infertility has been reported [39]. Aberrant DNA methylation of non-imprinted genes is linked to oligozoospermia and abnormal sperm morphology [40]. Methylation status aberrations were found in the genomic sperm DNA of infertile males. Additionally, the authors also revealed that DNA methylation of spermatozoa was lower in several repetitive sequences (e.g. *LINE-1, Alu*) compared to the somatic cells [12]. In the human genome, there are approximately 600,000 long interspersed nucleotide elements (*LINE1 or L1*) retrotransposons and over 1,000,000 *ALU* (a short interspersed nuclear element, *SINE*) retrotransposons. These non-imprinted elements account for approximately 17% and 11% of the total genomic DNA, respectively. DNA Methylation plays a crucial role in preventing retrotransposon activity and maintaining genome stability, particularly in *ALU* and *L1* elements. The absence of methylation in retrotransposons may facilitate their widespread dissemination across the genome, leading to insertional mutagenesis and, consequently, contributing to various diseases, including male infertility [36]. *L1* expression and retro-transposition events take place in the germline, during embryogenesis, and to a certain extent in somatic tissues. The methylation levels of *L1* in sperm, with a mean of 70% and a range of 20% to 80%, were found to be similar to those observed in somatic cells. Additionally, a study has reported a significant reduction in global *L1* promoter methylation in spermatozoa from males whose partners experienced recurrent pregnancy loss [36,41]. So, the study of retrotransposons' DNA methylation is also important to understand male infertility. These data highlight the crucial role of DNA methylation in spermatogenic aberrations and male infertility.

1.2.1.2 Histone modification

Germline reprogramming involves extensive histone modifications, a type of covalent PTM to histone proteins, which includes methylation, phosphorylation, acetylation, ubiquitylation, and sumoylation on their N- and C-terminal tails [42]. Histone proteins H2A, H2B, H3, and H4 along with DNA constitute octameric nucleosomes, and they are the fundamental DNA packing units. DNA wrapped around histone protein octamer forms a complex known as a nucleosome which is the structural and functional unit of chromatin; moreover, the expression level of genes

depends on chromatin modification [31,43]. Histone modifications can influence gene regulation by affecting the binding of regulatory factors to DNA. Histone methylation, a key modification, involves adding a methyl group to certain amino acids in the histone chain. Methylation at specific sites on histones H3 and H4 can lead to either transcriptional repression or activation. For instance, methylation of H3K9 and H3K27 is linked to gene silencing and a repressed chromatin state. These modifications are commonly found in genomic regions transformed into tightly packed chromatin, known as heterochromatin. The machinery responsible for transcription usually cannot access these regions [43]. However, the effects of histone methylation on gene expression are context-dependent and can vary depending on the specific amino acid being methylated, and the location of the methylation within the gene. For instance, methylation of H3K4 and H3K27 can have both expressive and repressive effects on gene expression [9].

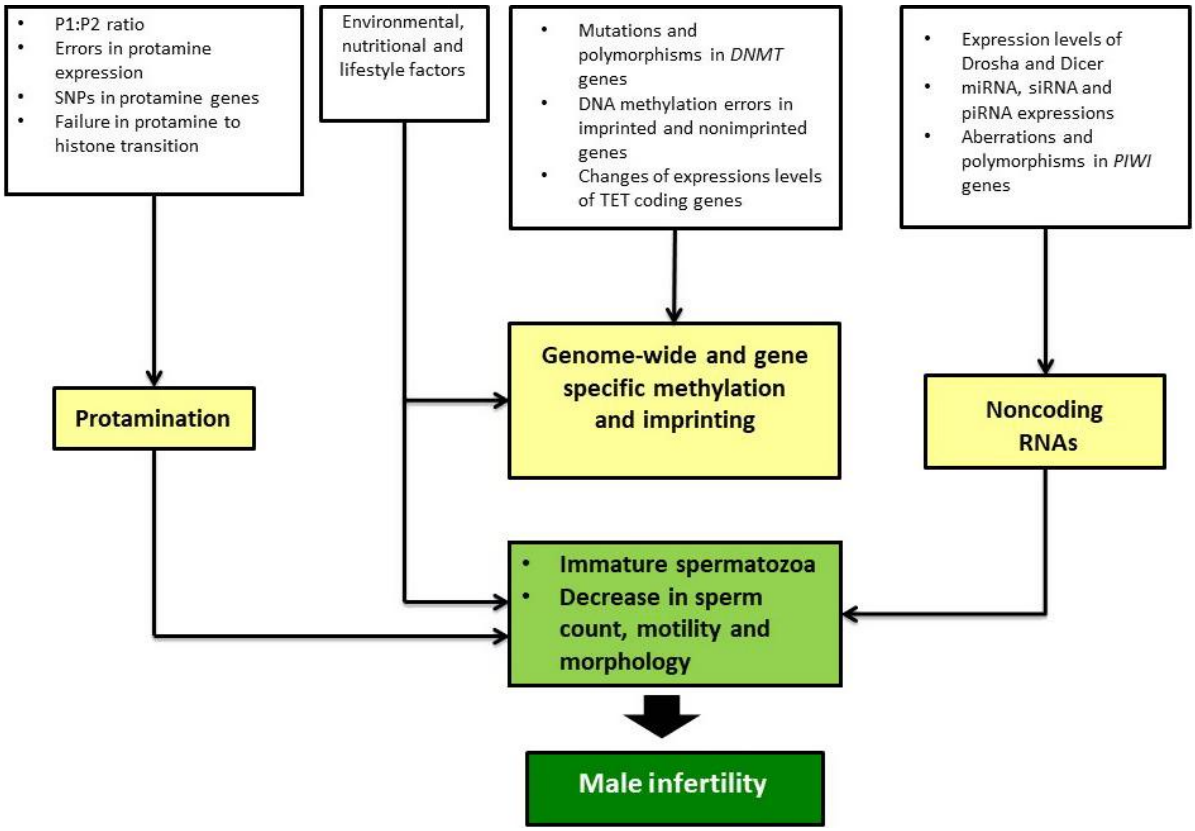


Figure 5: Potential epigenetic abnormalities causing impaired spermatogenesis and male infertility. Protamination, methylation and imprinting, and noncoding RNAs are the three major levels where epigenetic aberrations occur via different factors mentioned above. These aberrations subsequently contribute to male infertility by impacting sperm maturity, count, motility, and morphology. (The image was adapted [9]).

In mouse spermatogonia, acetylation of H2A, H2B, H3, and H4 was found to be high. These histones were deacetylated throughout meiosis in round spermatids and re-acetylated in elongating spermatids as illustrated in Fig. 6. Hyperacetylation of H4K has been proven responsible for the histone to protamine exchanges in the elongating spermatids [8,44].

Histone variants H3.3, and H3f3b have functional roles in development, transcriptional memory, and transcriptional reprogramming. A study showed that the knockout of *H3f3b* results in testis atrophy, reduction in germ cell population, and an elevation in abnormal spermatozoa, ultimately resulting in infertility. Moreover, *H3f3b-null* testes showed abnormal chromatin organization with reduced protamine incorporation and higher apoptosis in germ cells. These findings revealed noteworthy changes in histone PTM between abnormal and normal sperm samples, highlighting the crucial role of histone PTM in normal sperm function and fertility [8,9].

In the current study expression pattern of histone 3 lysine 27 trimethylation (H3K27me3) modification in human and mouse testis as well as retention in human motile sperm were investigated. Additionally, the enrichment of this PTM mark in the human sperm genome was also analysed.

1.2.1.3 ncRNAs

ncRNAs are pivotal in male infertility, with dysregulated ncRNAs associated with sperm motility, count, and morphology [45]. ncRNAs can be categorized into two groups by size: short ncRNAs and long ncRNAs (lncRNAs). SncRNAs include miRNAs, siRNAs, and piRNAs, which regulate gene expression post-transcriptionally [45]. miRNAs and siRNAs are abundant in male germ cells during spermatogenesis, while piRNAs are present only during the pachytene stage. SncRNAs are crucial for normal spermatogenesis. Knockout mouse models of Drosha and Dicer, which process miRNAs, revealed infertility due to impaired spermatogenesis (Fig. 5). Human studies identified stable miRNAs like miR-532-5p, miR-374b-5p, and miR-564, and proposed as potential fertility biomarkers in men with normal sperm parameters [8,9]. miRNAs, short single-stranded ncRNAs around 20-23 nucleotides, are crucial in regulating apoptosis, proliferation, and differentiation during spermatogenesis. Disruption of miRNA expression patterns can impact spermatogenesis [9,22,30].

1.2.1.4 Sperm-specific histone to protamine exchange

During spermiogenesis in mammals, male germ cells transform into motile sperm. This involves global nuclear remodelling in haploid round spermatids, where core histones are replaced by transition proteins and then protamines, forming a rigid chromatin structure. Hyperacetylation of histone tails loosens chromatin structure, allowing histone replacement. This protamination enables a highly condensed chromatin structure, aiding sperm motility and protecting the sperm genome. (Fig. 6) [8,9].

Any defects in the replacement or modification of histones during this process may result in conditions such as azoospermia, oligospermia, teratozoospermia and subsequently MI. There are two types of protamines, P1 and P2, which are equally expressed in humans. Faulty protamine transcript processing, leading to the over-production of immature P2 precursors, is associated with subfertility [8,9]. The ratio of different types of protamines also has an impact on fertility, as studies have demonstrated that men with low protamine (P1/P2) ratio have increased DNA fragmentation (Fig. 5). Moreover, the residual DNA content that remains bound to histones is believed to be crucial for proper sperm function and early embryo development [9,22,30,31].

Research has demonstrated that the substitution of histones with protamines in human sperm is not complete, resulting in the persistence of a histone code that is closely linked to crucial genes for development [35]. In human sperm cells, approximately 10% of histone proteins remain bound to chromatin, whereas in mice, this retention is approximately 1%. Analyses from high-throughput sequencing revealed an enrichment of residual nucleosomes at regulatory sequences in both human and mouse sperm cells. Mouse sperm nucleosomes contain both canonical and variant histone H3 proteins. These retained histones may contribute to the inheritance of epigenetic information and gene regulation in offspring. The histone-to-protamine transition provides a model for studying chromatin architecture remodelling by epigenetic regulators [42,46]. In addition, our previous studies in human and bovine sperm revealed that nucleosomes also remain in hypomethylated promoters of genes expressed in preimplantation development (e.g. RNA- and protein processing factors) and in hypomethylated transposable elements like short and long interspersed nuclear elements [47].

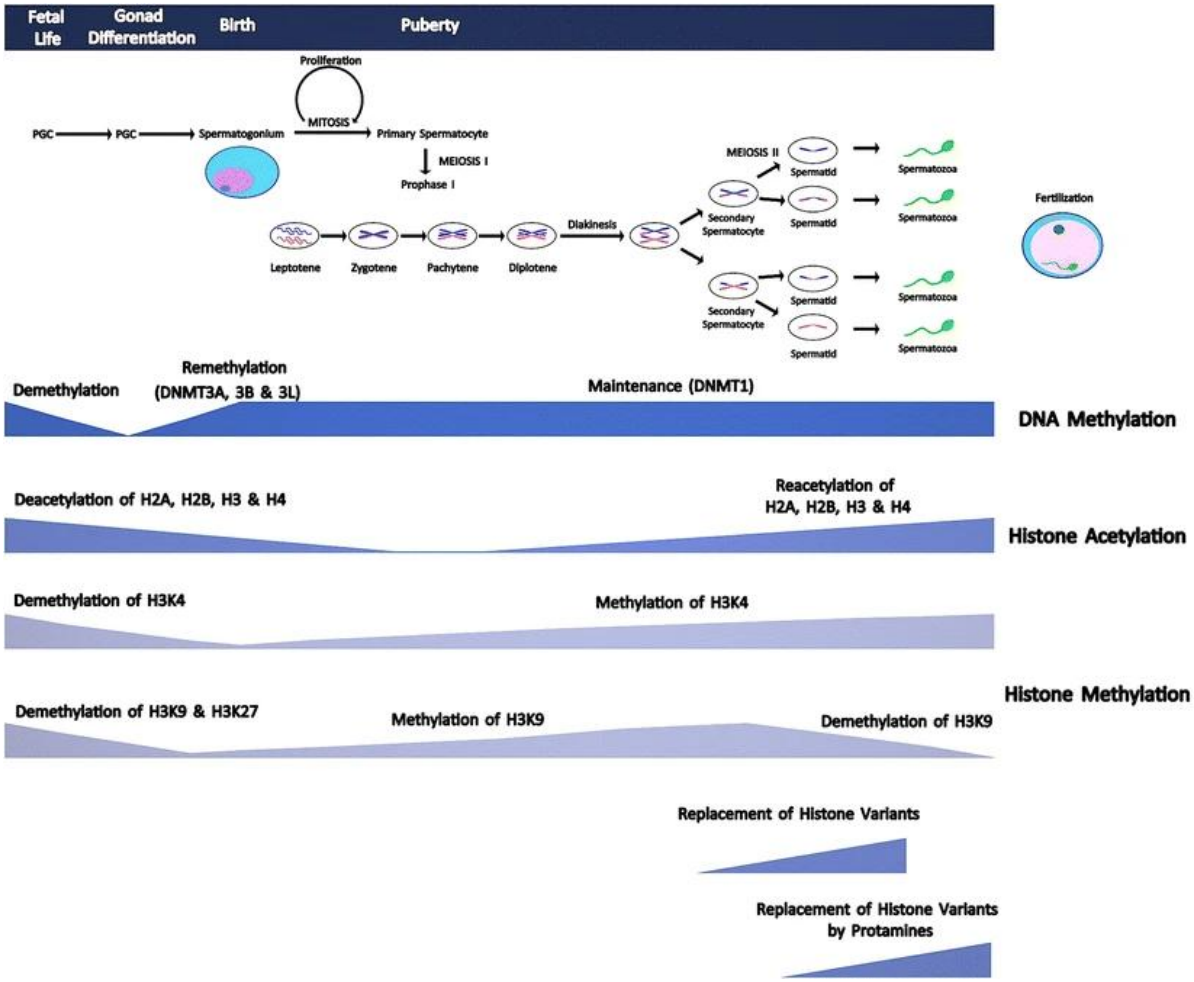


Figure 6: Scheme showing meiotic phases, dynamic changes of DNA methylation patterns, histone modifications, and protamination during spermatogenesis. Meiosis starts with puberty. During embryogenesis, methylation marks of primordial germ cells (PGCs) are erased. After this demethylation process, a specific re-methylation process starts in spermatogonia and primary spermatocytes. High acetylation of H2A, H2B, H3, and H4 is high in spermatogonia. During meiosis, these histones are deacetylated, and round spermatids are re-acetylated in elongating spermatids. (The image was adapted [8]).

1.2.2.1 Role of Ten-eleven translocation enzymes (TET 1-3) in male fertility

It was once considered that DNA methylation is an irreversible epigenetic process, but several studies have since identified a mechanism of active DNA demethylation where oxidation of 5mC to 5-hydroxymethylcytosine (5hmC) occurs by TET proteins, also known as methylcytosine dioxygenase [34]. TETs include TET1, TET2 and TET3, which are a family of 5-mC hydroxylases that regulate gene expression. To facilitate DNA demethylation, TET proteins bind to CpG-rich regions and prevent unwanted DNA methyltransferase activity. TETs enzymatically convert 5-mC (5-methylcytosine) to 5-hmC (5-hydroxymethylcytosine), 5-hmC

to 5-fC (5-formylcytosine), and 5-fC to 5-caC (5-carboxylcytosine) through hydroxylase activity and regulate gene expression [34]. The TET proteins play roles in transcriptional activation and repression (TET1), tumour suppression (TET2), and DNA-methylation reprogramming processes (TET3). TETs also regulate gene repression in an enzyme-independent manner by recruiting other transcription factors to the target genes [48].

The human *TET1* gene is located on chromosome 10q21.3 which has two transcriptional isoforms. These isoforms arise from the utilisation of alternate promoters. *TET2* and *TET3* genes are located on chromosome 4q24 and 2p13.1 respectively. The *TET1/Tet1* gene is an important part of the pluripotency transcriptional network in both human and mouse embryonic stem cells (ESCs). Stemness-specific factors such as Oct3/4, Nanog, and Myc regulate the expression of Tet1 in ESCs. However, *Tet1* is turned off when cells proliferate in adult tissues, and this downregulation is mediated by the Polycomb repressive complex 2 (PRC2) through the deposition of the H3K27me3 histone mark. Genome-wide studies have shown that Tet1 binding is associated with the transcriptional repressor Sin3a, forming a nuclear complex with Tet1 and PRC2. In mouse ESCs, Tet1 interacts with PRC2 and is recruited to the chromatin of bivalent genes to maintain their hypomethylated state (Fig. 7A). The functional interaction between Tet1 and PRC2 in ESCs is either indirect, likely mediated by one of the ESC-specific cofactors or is dependent on post-translational modifications [49,50].

It was reported that when *Tet1* $^{(-/-)}$ was knocked out in ESC, there was an abnormal increase in the expression of bivalent (H3K4me3; H3K27me3) developmental genes, which resulted in differentiation defects. The loss of *Tet1* led to a reduction in the enrichment of PRC2 and Sin3a at bivalent promoters, resulting in decreased H3K27 trimethylation and deacetylation, respectively, without any changes in DNA methylation. In *Tet1*-deficient embryos, *Gata6* was expressed at a higher level, and the developmental progress was delayed. These findings indicate that Tet1's non-catalytic roles in regulating H3K27 modifications to silence developmental genes are more critical for embryonic stem cells and early development than its catalytic functions in DNA demethylation [51].

In *Tet1* knock-out mice, a gradual decline in the number of SSCs was observed which impaired spermatogenesis, leading to an earlier onset of infertility as they aged. As a result, *Tet1* deficiency accelerates the process of premature reproductive ageing in these mice [52].

The impact of active DNA demethylation during sperm maturation could be confirmed by our recent findings in humans showing that TETs are dynamically regulated during the later stages of spermatogenesis and their mRNA levels significantly downregulated in their sperm of that

subfertile men. In particular, a considerable reduction of *TET1* and *TET3* mRNA was found in men undergoing ICSI- procedure (intracytoplasmic sperm injection) who exhibited low fertilization rates and could not achieve pregnancy after ICSI [53]. Apart from this, TET1 governs gene expression and the repression of endogenous retroviruses and regulates H3k27m3 independently of DNA demethylation catalytical activity in mouse ESCs [51,54].

Overall, these data suggest that TET enzymes play an important role in male fertility by regulating DNA methylation and mRNA levels during spermatogenesis. However, further research is needed to fully understand the mechanisms by which TET1 affects male fertility and to develop potential therapeutic strategies to improve male fertility based on this knowledge. In the current thesis, the expression of TET1 during human and mouse spermatogenesis, TET1 preservation and mRNA levels of TETs were examined in mature sperm.

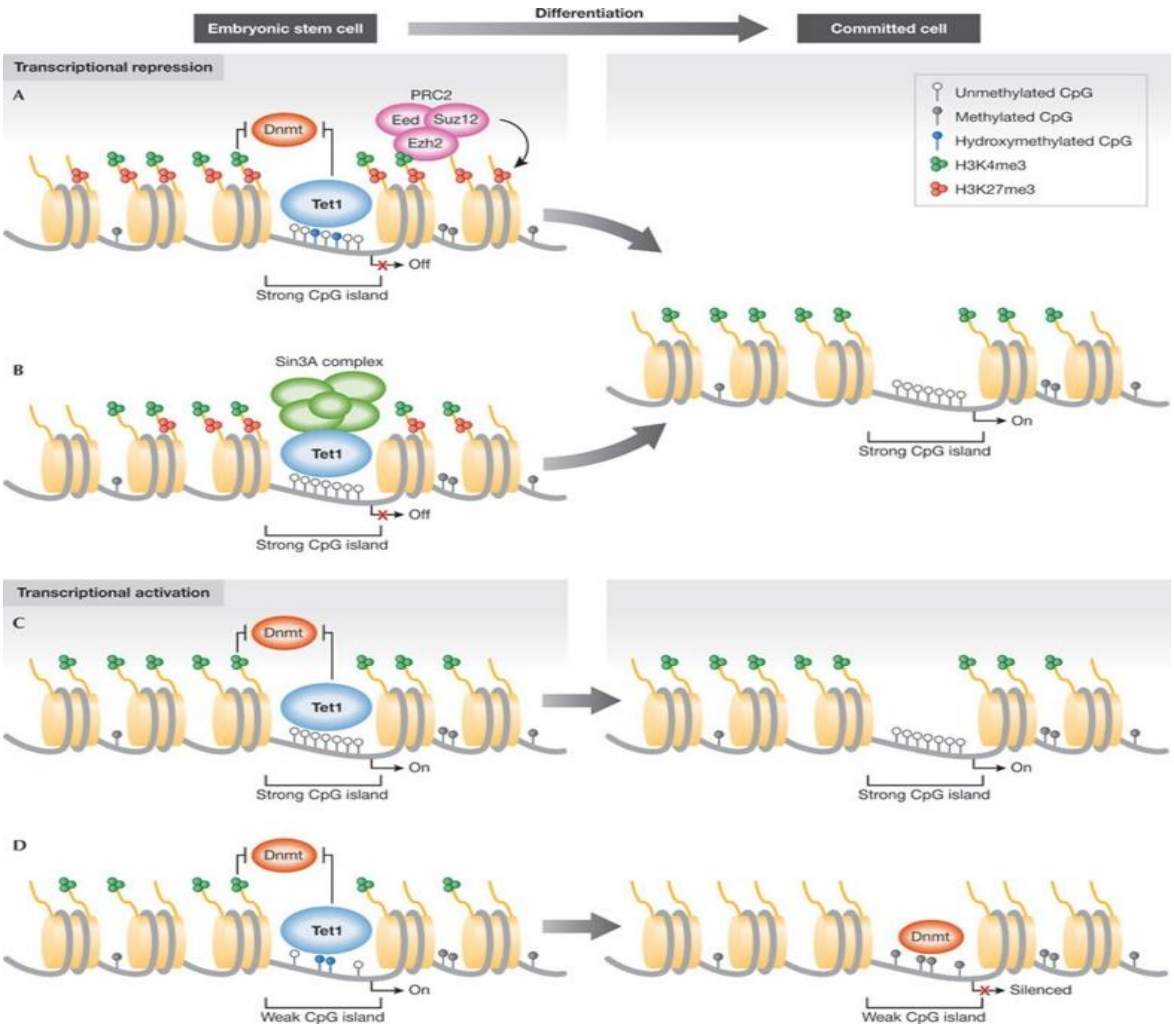


Figure 7: Potential roles of Tet1 in transcriptional activation and repression. A) Tet1 is believed to have a dual role in regulating gene expression. On one hand, it indirectly promotes the binding of the Polycomb Repressive Complex 2 (PRC2) by reducing DNA methylation at PRC2 target genes and gene repression. B) On the other hand, Tet1 can directly recruit

the Sin3A co-repressor complex to some of its target genes, resulting in transcriptional repression. However, In the process of differentiation, Tet1 is typically downregulated, which could enable the activation of previously repressed genes. C) Tet1 may facilitate transcriptional activation by preventing DNA methylation. At strong CpG islands that are infrequently methylated, Tet1 binding could act as a safeguard mechanism to remove any abnormal DNA methylation. D) Tet1 also binds to weak CpG islands that have been observed to become methylated during differentiation. PRC, Polycomb repressive complex; Tet1, ten-eleven translocation 1. (The image was adapted [55]**).**

1.2.2.2 Role of PRC2 complex in male fertility

Polycomb Repressive Complex 2 (PRC2) plays an important role in a variety of cellular processes including chromatin compaction, maintaining cell identity and proliferation, differentiation, and stem-cell plasticity, among others. PRC2 catalyses H3K27me3, which is a repressive histone mark. Studies have demonstrated that the PRC2-dependent H3K27me3 is retained in maturing mouse sperm at specific gene loci, which lack DNA methylation, and are relevant for precise embryo development (e.g. HOX genes) [33]. PRC2 have different subunits including three core components: enhancer of Zeste 1 or 2 (EZH1 and EZH2); suppressor of Zeste 12 (SUZ12); and embryonic ectoderm development (EED), etc. The enzymatic subunits EZH1 and EZH2, participate in the methylation at Lys 27 of histone H3 (H3K27me2/3), where EZH2 is the primary catalytic component of PRC2 [56]. In humans, EZH2 has been identified in the testis, and its expression decreases in testicular germ cell tumours. EED is essential for PRC2 assembly as it plays a critical role in the physical binding of H3K27me3 via five tandemly repeated WD motifs. SUZ12 facilitates the recruitment of PRC2 to the genome. Loss of any one of the three core PRC2 protein subunits severely compromises the functional activity of the complex and results in the loss of H3K27me3 [57,58,59]. Histone modification, particularly histone H3 methylation, which is substantially conserved in mice and humans, is tightly regulated during the process of spermatogenesis. Research has demonstrated that changes in these epigenetic marks are crucial in controlling gene expression in male germ cells both in normal and pathological states. In mice, complete loss of PRC2 function is lethal in embryonic stages. Moreover, mutations resulting in reduced function (hypomorphic), or a tissue-specific deletion, lead to growth and skeletal malformations, heart and immune cell defects, and increased susceptibility to tumorigenesis. Conditional deletion of EED in the male germline results in complete male infertility, demonstrating the essential role of PRC2 in male germline development [57,59]. PRC2-mediated H3K27me3 mark is enriched at the promoter regions of many PRC2 target genes and plays an essential role in cell differentiation. In foetal germ cells, H3K27me3 is enriched in developmental genes. In mature sperm, this histone mark is enriched on nucleosomes that are retained at the promoters of developmental genes indicating that PRC2 may regulate epigenetic information that is transmitted to offspring. Histones and their variants

in germ cells also play an important role in chromatin reorganization during meiosis and later events of spermiogenesis [57,31,58]. Besides, male mice with diminished PRC2 activity fathered progeny that exhibited overexpression of retrotransposons, along with modifications in the rates of preimplantation embryo division and regulation of the cell cycle [59].

Despite these observations, the function of PRC2 in mammalian germ cells remains unknown. In this thesis, the expression of these three core PRC2 components during human and mouse spermatogenesis as well as their retention patterns, and mRNA levels in motile sperm were examined.

1.3 Aims of the study

The establishment of paternal epigenetic information in sperm, its transmission to the offspring and its impact on their development is gaining recognition for its increasing medical significance. Despite this growing interest, how this information is composed and passed down as well as its consequences on embryogenesis and male fertility are not well understood [60].

Based on the known function of TET1-3 and PRC2, we hypothesized that TETs and PRC2 may interact during spermatogenesis and be involved in establishment of epigenetics marks in sperm at genome regions with a developmental relevance, e.g., transcription factors (TF) relevant for embryogenesis and differentiation (Fig. 8). Moreover, we hypothesized that spermatozoa of infertile men may have aberrant DNA methylation and histone modification in promoters of genes encoding those TFs due to dysregulation of TETs and PRC2 (Fig. 8).

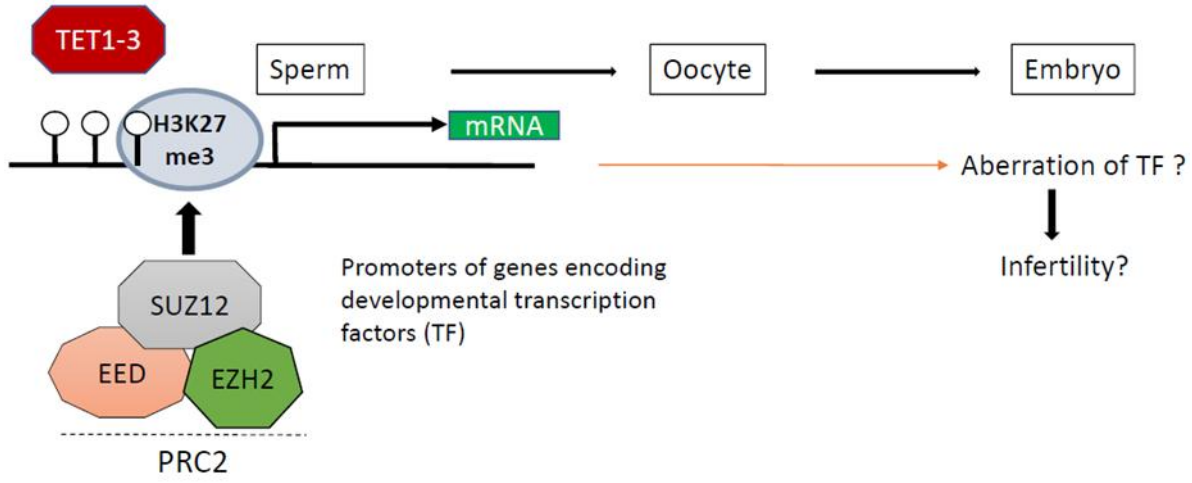


Figure 8: Hypothesized interaction of TET and PRC2 and their potential consequences in infertility. The diagram illustrates a possible interaction between TETs (DNA-demethylases) and PRC2 (H3K27me3 catalyst) at the promoters of genes encoding transcription factors (TFs) in sperm. In infertile patients, this interaction could be dysregulated, leading to aberrations in TF regulation during embryogenesis via mechanisms such as histone modifications (e.g., abnormal retention of H3K27me3) and changes in gene expression patterns due to abnormal DNA methylation. Dysregulation of crucial developmentally relevant TFs, e.g. FOXA2 and SOX2, could result in defective embryogenesis [61–63], which might lead to pregnancy failure and cause infertility.

To clarify the above-mentioned issues, we aimed to perform the following investigations:

- Investigation of the expression of TET1, H3K27me3, and PRC2 components (EED, SUZ12, EZH2) in human and mouse testis tissues using immunohistochemistry (IHC) analysis. For mouse model experiments, *Tet1* knock-out mice (homozygous mutant *Tet1*$^{-/-}$ and heterozygous mutant *Tet1*$^{+/-}$*)* were analysed in comparison to wild-type mice to investigate the impact of *Tet1* absence on the expression of H3K27me3 and PRC2 components.

- As studies have been reporting about the association of aberrant histone retention and DNA methylation to male infertility [64,65], we aimed to determine if TET1, H3K27me3 and PRC2 components (EED, SUZ12, EZH2) are dysregulated at the protein level in spermatozoa of subfertile men (patients) by comparing them to those in fertile men (controls) using quantitative western blot analysis (qWB);

- H3K27me3-binding at gene promoters of TFs with developmental relevance (*SOX1, SOX2, OCT4, NANOG, PAX6, GATA2, FOXA2)* and human endogenous retroviral sequences (*HERV)* were analysed using chromatin immunoprecipitation followed by quantitative PCR (ChIP-qPCR) to detect epigenetic dysregulation of TFs and *HERV* in motile spermatozoa of fertile and subfertile men.

- In previous studies of our group, aberrant levels of *TETs* mRNA in spermatozoa were found to be associated with poor spermiogram parameters [53]. In this study, mRNA levels of *TET1-3, EED, SUZ12, and EZH2* were analysed in motile spermatozoa of fertile and subfertile men using reverse transcription of mRNA into cDNA followed by quantitative PCR (RT-qPCR) to reveal interdependencies.

- CpG-promoter methylation of genes encoding TFs with developmental relevance (*SOX2, PAX6, GATA2, FOXA2*) and PRC2 components (*EED, EZH2, SUZ12*) in spermatozoa of fertile and subfertile men were investigated using bisulphite-treated DNA (BS-DNA) and pyrosequencing to detect DNA methylation aberrations. Correlation analysis between promoter methylation and mRNA levels was also performed.

- CpG-methylation analysis of *LINE1* and *Alu* was performed in motile spermatozoa of fertile and subfertile men using combined-bisulphite-restriction analysis (COBRA) as in our previous study it was found that *LINE1* mRNA was significantly upregulated in the same cohort, and TET enzymes regulate the demethylation of transposable elements.

2. Materials and methods

2.1 Materials

2.1.1 Reagents, buffers, and kits

The list of chemicals, buffers, and kits employed in this thesis is provided in Tables 3 to 7. In Tables 3 and 4, reagents are listed used for nucleic acids, methylation and IHC analysis.

Table 3: Reagents for protein, DNA, RNA, and CpG methylation analysis

Reagents	**Vendor**
NaAc solution 3 mol/l (3 M, pH 5.2)	AppliChem, Darmstadt
Acrylamide solution (30 %); APS;	BioRad, Munich
Agarose	Biozym Scientific, Oldendorf
Citric acid; Orange G; PFA; Tri-sodium citrate	Merck, Darmstadt
PeqGold TriFast, Xylene	VWR Life Science
Acetic acid; chloroform; DTT; Glycine; Glycogen; HCl; Non-fat dry milk; NaCl; NaOH; SDS; Phenol/Chloroform/Isoamyl- alcohol; Proteinase K; Roti®-Phenol; TEMED; Tris; Triton X-100	Roth, Karlsruhe
BSA; DEPC; EDTA; EtOH; H_2O_2; Isopropanol; MeOH; β- mercaptoethanol; PMSF; PBS; Protease inhibitor cocktail Tablets.	Sigma-Aldrich, Steinheim

Table 4: Reagents used for Immunohistochemistry analysis

Reagents/Buffers	**Vendor/Recipe**
10x Tris-HCl Buffer, pH 7.4	1 M Tris-HCl, 1 M NaCl and 1 L ddH2O
1x Tris-HCl-Buffer	1/10 volume 10 x Tris-HCl, 0.1 % Triton in ddH2O
Citrate buffer, pH 6.0	18 mM citric acid monohydrate, 82 mM sodium citrate in ddH2O
2% H_2O_2 / Methanol	63 mL ice-cold MeOH, 7 mL 30 % H2O2
5 % BSA	3.5 g BSA in 70 mL 1x Tris-HCl buffer
AEC+ high sensitivity substrate chromogen, Faramount Mounting Medium	Dako, Glostrup, Denmark
Haematoxylin	Merck, Darmstadt
Vectastain Elite ABC-Peroxidase Staining Kit	Vector, Peterborough, UK

x= times

In the upcoming Tables 5 and 6, chemicals, buffers, and kits used for ChIP and Western blot (WB) analysis are listed below.

Table 5: Reagents used for Western blot analysis

Reagents	Vendor/Recipe
10 % APS	50 mg APS in 500 µL ddH2O
10 x SDS running buffer, pH 8.3	25 mM Tris, 192 mM glycine, 0.1 % SDS in ddH2O
1 x PBS, pH 7.4	4 PBS Tablets in 1 L ddH2O
1x PBST, pH 7.4	1 x PBS, 0.1 % Tween 20 in ddH2O
Coomassie destaining solution	40 % MeOH, 10 % acetic acid in ddH2O
Laemmli buffer (4x); Extra Thick Blot Filter Paper	Bio-Rad, Munich
Coomassie staining solution	0.1 % Coomassie Brilliant blue R250, 50 % MeOH, 10 % acetic acid in ddH2O
Laemmli loading buffer, 1 mL	900 µL Laemmli buffer, 100 µL β-Mercapto ethanol
Odyssey Blocking Buffer TBS/PBS, 1xPBS, REVERT™ Total Protein Stain Kit	LI-COR Biosciences, Bad Homburg
Immobilon-FL PVDF	Merck, Darmstadt
Running (pH 8.8) and Stacking (pH 6.8) gel buffer;	1.5 M Tris in ddH2O
Transfer buffer, pH 8.0-8.5	48 mM Tris, 38 mM glycine, 20 % MeOH, 0.0375 % SDS in ddH2O
Tris-Urea buffer	8 M Urea, 2 % SDS, 50 mM Tris (pH 8.0), 105 mM NaCl, 1 mM PMSF, 10 mM DTT, 1x Protease Inhibitor Cocktail (1 Tablet/ 100 mL)
Coomassie Brilliant Blue, R250	SERVA Electrophoresis, Heidelberg
Page Ruler™ prestained protein ladder 10- 180 kDa	Thermo Fisher Scientific, Darmstadt

Table 6: Reagents used for Chromatin immunoprecipitation analysis

Reagents	Vendor
Micrococcal Nuclease (Cat. No. M0247S)	New England Biolabs
Lysolecithin (1-Palmitoyl-sn-glycero-3-phosphocholine) Cat. no L5254, Protease inhibitor cocktail (Cat. No. S8830)	Sigma
Dynabeads Protein A (Cat. No. 10001D)	Thermo Fisher Scientific

2.1.2 Equipment

In the upcoming Table 7, all the equipment used in the experiments are listed.

Table 7: List of all equipment used in the experiments.

Instruments	Vendor
Gel documentation system, BioDocAnalyze	Biometra, Goettingen
Vibracell 75022, ultrasonic processor, 130W	Bioblock Scientific
Mini-PROTEAN Tetra Cell; Power PAC 200 + 300; Thermocycler T100 + C1000 Touch; Trans-Blot SD Semi-dry transfer cell	Bio-Rad, Munich
Fridge; Freezer Economic GSL 1202 + 3601	BOSCH, Stuttgart
Gourmet cooker	BRAUN, Schwalbach am Taunus
Thermo Mixer; Mastercycler; Pipettes; tubes	Eppendorf, Germany
Mini centrifuge	Fisherbrand, Schwerte
Shaker Reax 2000; Vortex mixer REAX 2000	Heidolph, Schwabach
Incubator Kelvitron T	Heraeus, Hanau
Centrifuge (MIKRO 220R)	Hettich, Frankenberg
Shaker VIBRAX-VXR; Magnetic Stirrer (BLSH0007)	IKA, Staufen
LI-COR Odyssey Fc	LI-COR Biosciences, Bad Homburg
Microwave Untellowave	LG, Ratingen
MilliQ direct water purification system	Merck, Darmstadt
Fine balance Mettler AE 240; pH meter (S20K)	Mettler Toledo, Giessen
Pipette tips	Nerbe plus, Winsen + Biozym,
Microscope BX43	Olympus, Tokyo, Japan
Gel chamber; NanoDrop ND-1000	PeqLab, Erlangen
Stainless steel beads (ball), 5 mm; Pyromark Q24 system and supplies; TissueLyser LT	Qiagen, Hilden
Cover plates; Glass plates superfrost	R. Langenbrinck GmbH
Neubauer-improved counting chamber	Roth, Karlsruhe
Heating plate Präzitherm 28-1	Stoerktronic, Stuttgart
Centrifuge (Heraeus Fresco 21); HERA- freeze™ HFU T Series -86°C Freezer; Invitro- gen Power Ease® 500 Power Supply; Multi- scan Go Photometer; Tube Revolver	Thermo Fisher Scientific, Darmstadt
Heating Block LS 2	VLM, Bielefeld
Magnet, Magna-Grip 8 well 20-400	Millipore

2.1.3 Human semen samples

Motile sperm samples of 48 subfertile men were collected from our collaboration partner, Prof. Dr. med. N. Rogenhofer at the Department of Gynecology and Obstetrics, Clinical Centre, Ludwig-Maximilians-University Munich. These patients underwent ART (assisted reproductive technology), ICSI (intracytoplasmic sperm injection) and/or IVF (*in vitro* fertilization). Semen samples from 113 healthy men (controls) were collected at the Department of Urology, Pediatric Urology and Andrology, JLU Giessen (in collaboration with Prof. Dr. med. H.-C. Schuppe).

Every participant was informed about the study, and written consent was taken from each participant. The collection and molecular-epigenetic analysis of semen was approved by the Ethics Commission of the Medical Faculty, JLU Giessen (approval from 17.12.14 in the frame of the DFG financed Clinical Research Unit KFO181/Period 2/Subproject 6, PI Prof. Dr. rer. nat. U. Schagdarsurengin). The present study particularly dealing with the role of TETs and PRC2 in spermatogenesis and establishment of sperm epigenome was financed by the DFG-GRK1871/Period 2/Subproject 11 (PIs Prof. Dr. rer. nat. U. Schagdarsurengin and Prof. Dr. rer. nat. K. Steger).

2.1.4 Human and mouse testis samples

Human testis tissue samples were obtained from obstructive azoospermia patients and were provided by the Institute for Veterinary Anatomy, Histology and Embryology, JLU Giessen (Prof. Dr. vet. med. D. Fietz). Three types of mouse testes were used in the study, wild type (WT) and *Tet1*-knock out, heterozygous (+/-) and homozygous (-/-). The mouse testes samples were obtained from Prof. Dr. K. P. Koh's lab (KU Leuven, Belgium). Testis samples were subjected to fixation by perfusion in Bouin's fixative, followed by embedding in paraffin using standard techniques. Subsequently, tissue blocks were cut using a Microtome device to obtain sections with a diameter of five micrometres. Then, the tissue sections were carefully placed on slides and kept at a temperature of 37°C overnight, then kept at room temperature for further analysis.

2.2 Methods

2.2.1 Isolation of human motile spermatozoa (swim-up technique)

The semen samples were collected and analysed as per guidelines from WHO (2010). The ejaculated fresh semen samples from controls and patients were obtained by masturbation and collected into sterile tubes. The participants were abstained from sexual intercourse for at least 3 days before semen collection.

For further individual analysis, liquefaction of the semen samples was done at 37 °C for 15-30 minutes and separated into motile and immotile sperm cell fractions using a method called "swim-up". Briefly, liquefied semen was taken on a round tube containing 1.2 mL sperm

wash medium (HTF-Medium/2 % HSA) and was kept at a 45° angle at 37 °C for 60 minutes. During this incubation, motile sperm cells swim up into the wash medium. This medium containing M cells was aspirated, and immotile sperm cells stayed at the bottom of the tube. For both fractions, sperm concentration and motility were determined and recorded. Then sperm samples were centrifuged at 500 g for 10 min at 4 °C and washed twice with PBS. Afterwards, sperm samples were used directly for the experiments or frozen in liquid nitrogen and stored at -80 °C until further processing. Spermiogram parameters, ages of controls, patients, and their female partners, are summarized in Table 8.

Table 8: Spermiogram data with ages of collected controls, patients, and their female ICSI/IVF partners.

Spermiogram parameter	Controls	Patients
Total Participants	113	48
Age (years)	28 (20-48)	39*** (29-50)
Partner Age (years)	-	36 (27-45)
Sperm concentration (Mio/mL)	47*** (0.9-282)	18 (1.2-219)
Total sperm count (Mio)	143** (0-725)	72 (5.12-585)
Progressive motility (%)	83*** (0-100)	39.5 (9-62)
Total motility (%)	90*** (0-100)	61 (25-83)

Median and range values (minimum-maximum) are presented; p-values were calculated by Mann-Whitney U-test: **p=0.0099, ***p <0.0001. Details of patients' and controls' data can be found in the supplement (Tables 20-22).

2.2.2 Protein expression analysis in human and mouse testis (WB, IHC)

2.2.2.1 Immunohistochemistry analysis of TET1, H3K27me3, PRC2 components in human

To analyse the expression of TET1, H3K27me3, and PRC2 components (EED, SUZ12, EZH2) in human testis tissue samples immunohistochemical (IHC) analyses were performed. The testis samples obtained from obstructive azoospermia patients exhibited normal spermatogenesis. Testes sections were fixed in Bouin's solution, embedded in paraffin, and mounted on microscope slides (R. Langenbrinck GmbH). The slides were deparaffinized sequentially in Xylol (3 x 10 min) and Ethanol (2 x 5 min in 100 % EtOH, 2 x 5 min in 96 % EtOH and 2 x 5 min in 70 % EtOH) and boiled for 20 min in citrate buffer (18 mM citric acid monohydrate, 82 mM sodium citrate) or 15 minutes in Tris-EDTA buffer pH 9.0 for antigen retrieval. The slides were cooled down for 30 min at RT and the peroxidase activity was inhibited by blocking with

3 % hydrogen peroxide (H_2O_2) and 100 % ice-cold methanol solution for 30 min at RT. Three times washing was performed for 5 min using Tris-HCl buffer (0.1 M Tris-HCl, 0.1 M NaCl).

Table 9: Antibodies used for the IHC analysis

Antibody	Species	Dilution	MW (kDa)	Vendor/ Cat. No.
Human				
Anti-TET1	Rabbit	1:100-1:800	-	Abcam/ab191698
Anti-TET1	Rabbit	1:400	-	GeneTex/GTX124207
Anti-H3K27me3	Rabbit	1:200	-	Abcam/ab192985
Anti-EZH2	Rabbit	1:100	-	Abcam/ ab191080
Mouse				
Anti-Tet1	Rabbit	1:300	-	GeneTex/GTX124207
Anti-Tet1	Rabbit	1:100-1:800	-	Novus Biologicals/NBP1-78965
Anti-Tet1	Mouse	1:100-1:800	-	Novus Biologicals/NBP2-15135
Anti-Tet1	Rabbit	1:100-1:800	-	Active motif/61444
Anti-Ezh2	Rabbit	1:100-1:400	-	Abcam/ ab191080
Anti-Suz12	Rabbit	1:200	-	Abcam/ ab12073
Anti-Eed	Rabbit	1:100-1:800	-	Abcam/ ab254569
Anti-Eed	Rabbit	1:100-1:800	-	Abcam/ ab236292
Anti-mouse IgG	Goat	1:200	-	Dako/E0433
Anti-rabbit IgG	Goat	1:200	-	Dako/E0432

Abbreviations: MW: Molecular weight in Kilodalton (kDa)

Then blocked for 20 min in 5 % bovine serum albumin (BSA) solution or 60 minutes in 5% milk (in Tris-HCl buffer) to reduce non-specific antibody binding. The blocked slides were incubated with primary antibodies (anti-TET1, GTX124207 Genetex, 1:400-1:500; anti-H3K27me3, ab192985 Abcam, 1:200; anti-EZH2, ab191080 Abcam, 1:100, table 9) overnight at 4 °C in humidity chamber. Slides incubated with blocking buffer without primary antibody were used as a negative control. The slides were washed three times for 5 min in Tris-HCl buffer and incubated with a secondary antibody (goat anti-rabbit IgG, 1:200, Dako, E0432) in a humidity chamber at RT for 1 h. Non-specifically bound antibody was washed away using Tris-HCl buffer (3 x 5 min) and then slides were developed for 2-20 min with ABC-Peroxidase Staining Kit. Slide development was stopped by washing in ddH_2O (3 x 5 min). The nuclei were counterstained for 5 s in Mayer`s haematoxylin (Merck) with final washing in dH_2O (3 x 5 min). Slides were mounted with Dako Faramount aqueous mounting medium and covered with

glass plates for microscopic examination. During slide investigation per microscope, representative images were taken and electronically documented.

2.2.2.2 IHC analysis of Tet1, H3K27me3, and PRC2 components in wild type, *Tet1* $^{(+/-)}$, *Tet1* $^{(-/-)}$ mouse

IHC analyses were performed to analyse Tet1, H3K27me3, PRC2 components (Eed, Suz12, Ezh2) expression in wild type, *Tet1* $^{(+/-)}$, *Tet1* $^{(-/-)}$. mouse testis tissue samples. Testes sections were processed in the same way as described above with little modifications. These modifications are only stated here.

These primary antibodies (anti-Tet1, GTX124207 Genetex, 1:200; anti-H3K27me3, ab192985 Abcam, 1:200; anti-Eed, ab236292 Abcam, 1:800; anti-Eed ab254569, 1:200; anti-Suz12, ab12073 Abcam, 1:200; anti-Ezh2, ab191080 Abcam, 1:200, table 9) were used for overnight incubation at 4 °C in a humidity chamber.

2.2.2.3 Protein analysis using western blot in human motile spermatozoa

2.2.2.3.1 Protein isolation from spermatozoa and WB analysis of TET1 (wet transfer)

For protein isolation, 10-20 million motile sperm cells were taken from frozen aliquots and washed with PBS. Sperm cells were resuspended in Tris-Urea buffer (8 M Urea, 2 % SDS, 50 mM Tris (pH 8.0), 105 mM NaCl, 1 mM PMSF, 10 mM 1,4-Dithiothreitol (DTT), 1x Proteinase Inhibitor Cocktail) and incubated at RT for 30 min. Samples were vortexed frequently. Then sperm cells were sonicated for three cycles, each 10 s (Amplitude 30 %, max. pulse) using ultrasonic sonicator 'Vibra Cell'. Cell debris was removed by centrifugation (10 min, 13.000 rpm, RT) and whole protein extracts were taken into a new tube. Protein concentrations were measured using Nanodrop and extracts were kept at -80 °C. 50 µg protein was mixed with 4x Laemmli loading buffer and run on an SDS-PAGE Gel. For TET1 protein, 7.5 % running gel (4.8 mL ddH_2O, 2.5 mL acrylamide, 2.5 mL running gel buffer, 100 µL 10 % SDS, 100 µL 10 % APS, 4 µL TEMED) and a 4 % stacking gel (2.1 mL ddH2O, 500 µL acrylamide, 380 µL stacking gel buffer, 30 µL 10 % SDS, 30 µL APS, 3 µL TEMED) was used.

Separation of loaded proteins was started at 90V for ~30 min. The run was adjusted to 130 V for ~50 min until the proteins reached the bottom of the gel. The gel was equilibrated for 10 min in a transfer buffer. Wet transfer (Towbin buffer, 25 mM Tris, 192 mM Glycine) was performed for 95 minutes at 250 mA on polyvinylidene-difluoride membranes (PVDF, 0.45 µm pore size, Immobilon-FL, Merck-Millipore). The membrane was dried at 37 °C for 15 min for better protein binding and was re-activated in 100% methanol for 2 min. The whole protein content was analysed after transfer to the membrane using the Revert™ total protein stain Kit. Staining was documented by the Odyssey® Fc imaging system.

Table 10: Antibodies used in Western blot analysis

Antibody	Species	Dilution	MW (kDa)	Vendor/ Cat. No.
Anti-TET1	Rabbit	1:500	235	GeneTex /GTX124207
Anti-H3K27me3	Mouse	1:500	15	Abcam /ab6002
Anti-EZH2	Rabbit	1:1000	98	Cell signaling/D2C9
Anti-SUZ12	Rabbit	1:500	83	Cell signaling/D39F6
Anti-EED	Rabbit	2µg/ml	51	Abcam/ab254569
Anti-EED	Rabbit	1:500	51	Abcam/ab240650
Anti-SUZ12	Rabbit	1:500	83	Abcam/ab175187
Anti-SUZ12	Rabbit	5µg/ml	83	Abcam/ab12073
Anti-GAPDH	Rabbit	1:5000	-	Sigma/G9545
Anti-GAPDH	Mouse	1:1000	-	Abcam/ ab9484
Anti-rabbit IgG 680RD	Goat	1:5000	-	LI-COR/926-68071
Anti-rabbit IgG 800CW	Goat	1:5000	-	LI-COR/926-32211
Anti-mouse IgG 800CW	Goat	1:5000	-	LI-COR/926-32210

Abbreviations: MW: Molecular weight in Kilodalton (kDa)

Blocking was performed using 5% non-fat milk (in TBS) for 1 hour at room temperature. The membranes were incubated overnight at 4°C with anti-TET1 (GTX124207, 1:500) and anti-GAPDH antibodies (ab9484, Abcam, 1:1000) diluted in 1% non-fat dry milk (in TBS) with 0.1% Tween-20. Membranes were washed in TBS-Tween-20 buffer (0.1% Tween-20, TBST) and incubated with secondary antibodies (goat anti-rabbit IgG, 1:5000; conjugated with IRDye 800CW for TET1 and goat anti-mouse IgG, 1:5000; conjugated with IRDye 800CW for

GAPDH detection) for 1 hour at room temperature (dark). Secondary antibodies were diluted in LI-COR Intercept blocking buffer (plus 0.2% Tween-20 and 0.01% SDS). After washing with TBST buffer, the membranes were rinsed again with TBS. The fluorescence signals from the membranes were detected with the Odyssey Fc system and captured images were stored and quantified using the Empiria Studio software (LI-COR Biosciences).

This software offers a wide range of analysis capabilities for WB projects, starting from validation experiments and progressing to Target Analysis Experiments, which include calculations for fold change and statistics. To quantify the blots, Empiria Studio calculates the signal by adding up the intensity values of individual pixels (total) within a bandbox (or lane box) and then subtracting the product of the background value and area. The software automatically assesses the local area surrounding each band box or lane box to determine the background value. The software also calculates the signal-to-noise ratio (SNR) by comparing the signal in a lane with the background noise calculated for that lane. Additionally, the software provides the average SNR for lanes within a replicate group. Furthermore, the software calculates the signal for the target after normalizing it to the internal loading control. It also provides the average of the normalized signal values for target bands in a replicate group[66].

2.2.2.3.2 WB analysis of H3K27me3 and PRC2 components (semi-dry transfer)

For the semidry transfer method, most of the protocol remains the same as the wet transfer method except for the transfer of the protein gel to the PVDF membrane. The modifications are as follows:

The gel with membrane cassette was placed in the semi-dry protein transfer chamber and the transfer was carried out for 30 min at 150 mA. Rest of the procedures were the same. Here (Table 10) primary antibodies (anti-H3K27me3 1:500, ab6002, Abcam; anti-GAPDH1:5000, G9545, Sigma; anti-EZH2 1: 500; 5246, Cell signaling; SUZ12 1:500, ab175187, Abcam) diluted in 1 % NFDM with 0.1 % Tween-20 were added to the membrane and incubated overnight at 4 °C on a rotator (15 rpm). The next morning, the membranes were washed three times with TBST buffer for 5 min (20 rpm) and incubated for 1 h (15 rpm) with secondary antibody (IRDye® 800CW/680LT Goat anti-Rabbit IgG, 1:5000, Goat anti-Mouse IgG, 1:5000) diluted in intercept blocking buffer with 0.1 % Tween-20, and 0.01 % SDS. The membranes were washed, developed, and quantified as stated before.

2.2.3 Chromatin analysis in human motile spermatozoa

2.2.3.1 Native chromatin immunoprecipitation-sequencing of TET1

Native chromatin immunoprecipitation (NChIP) was performed with 5 million aliquots of 10-20 million motile sperm cells from fertile men to analyse DNA loci bound to TET1. The cell pellet was resuspended in 0.1 % Lysolecithin buffer, incubated for 15 min on ice, then centrifuged at 2500 g for 5 min at 4°C, and the supernatant was discarded. The pellets were washed with Phosphate-buffered saline, which was supplemented with 1x protease inhibitor cocktail (PBS-PI). After centrifugation supernatant was discarded. Pellets were resuspended with 20 mM DTT buffer and were incubated at 37°C for 30 min. The washing and centrifugation steps were repeated. The pellets were suspended in a buffer called PBS-PI-CaCl2 (also known as MNase Buffer), with a final concentration of 1 mM CaCl2. This is important because MNase needs calcium for its activity. PBS provides an isotonic and buffered environment, maintaining the stability of the cell or chromatin structure during the experiment. PI (Protease Inhibitor) helps in preventing protein degradation by inhibiting protease activity. Then 30 units of Micrococcal nuclease (Cat No. M0247S, New England Biolabs) was added into each sample and mixed properly by flicking the tube and incubated exactly 5 min at 37°C to digest the linker DNA and producing mono-nucleosomal chromatin fragments. The digestion was stopped by adding EDTA to a final concentration of 6 mM and incubated on ice briefly. Centrifugation was performed at 17000 g for 10 min at 4°C to separate chromatin fragments. All the supernatants (Histone-fraction) were pooled into a new tube and dilution buffer, DB (50 mM Tris-HCl pH 7.5, 5 mM EDTA, 0.1 % Triton X-100) was added. After mixing well, from diluted chromatin 10% was separated to use as the input DNA control, the rest was used for ChIP. Dynabeads protein A (Cat No. 10001D, Thermo Fisher Scientific) was used to prepare the bead-antibody complex. The Dynabeads were resuspended by gentle pipetting and washed once with DB. A magnetic rack was used to separate the beads from the solution. Then beads were mixed with DB and incubated with 5µg anti-TET1 antibody (Cat No. GTX124207, GeneTex) for 20 minutes (15 rpm) at room temperature. This antibody-bead complex was washed once with DB and then incubated with the remaining 90 % soluble chromatin supernatants overnight (10 rpm at 4 °C). In the next morning, the sample was taken from the rotator and chromatin-bound antibody-beads complexes were separated from the buffer using a magnet. The supernatant was taken to clean tubes and DNA was isolated to check the mono-nucleosomal DNA fragment (~147bp) produced after Mnase digestion. The chromatin-bound antibody-beads complexes were washed sequentially using each of the NChIP wash buffers:

buffer A (50 mM Tris-HCl pH 7.5, 10 mM EDTA, 75 mM NaCl), B1 (50 mM Tris-HCl pH 7.5, 10 mM EDTA, 125 mM NaCl, 0.04% Triton X-100), and B2 (50 mM Tris-HCl pH 7.5, 10 mM EDTA, 125 mM NaCl, 0.01% Triton X-100). All the buffers were prepared with protease inhibitors (EDTA-free). The complexes were separated from the buffer on the magnet between each wash and supernatants were discarded. Antibody-bound chromatin was eluted from beads using elution buffer (10 mM Tris/HCl pH 8, 1 mM EDTA, 1% SDS, 100 mM NaHCO3) for 20 min (35 rpm) at RT. The eluate was taken into a new DNA low-bind tube. The elution step was repeated once more, and the pooled eluates and input controls were subjected to RNase A (Thermo Fisher Scientific) treatment at 37°C for 30 minutes and proteinase K (Carl-Roth) treatment at 56°C for 4 hours. DNA was purified by a standard phenol-chloroform procedure and precipitated with 2 volumes absolute ethanol, 1/10 volume NaAc (pH 5.2) and 20 µg glycogen overnight at -20°C. The eluted ChIP-DNA was quantified using the Qubit™ dsDNA HS Assay Kit (Cat No. Q32851, Thermo Fisher Scientific) and selected samples were sent to the company Diagenode for genome-wide sequencing (Illumina platform).

2.2.3.2 H3K27me3-NChIP-qPCR

The purpose of H3K27me3-NChIP-qPCR (quantitative PCR) was to analyze the enrichment of H3K27me3 in sperm of fertile versus sub-fertile men at specific gene promoters known to be involved in early embryogenesis as transcription factors. After performing H3K27me3-NChIP according to the previous protocol, the isolated NChIP-DNA fragments were subjected to qPCR analysis using specific primers designed to amplify the promoters of interest. The quantification of NChIP-DNA using qPCR served to confirm the presence of H3K27me3 at specific genomic loci and to obtain information on relative H3K27me3 enrichment among different samples. NChIP was performed with 10-20 million motile sperm cells from healthy and infertile men with the same protocol used for TET1-NChIP. Here 5µg anti histone H3 tri methylated at K27 antibody (H3K27me3, Cat no. ab6002, Abcam) per ChIP. The eluted NChIP-DNA was quantified and analysed by qPCR for potential H3K27me3-bound developmentally relevant gene promoters (*SOX1, SOX2, OCT4, NANOG, PAX6, GATA2, FOXA2, and HERV*). ChIP-qPCR primers for these genes were either designed or taken from the literature and are presented in Table 11. After qPCR, data were analysed using the per cent input method as described in Allan et al., 2021 [67]. This method normalizes the signals obtained from the ChIP to the signals obtained from an input sample, which represents the amount of chromatin used in the ChIP.

Firstly, the average Ct values for the antibody used were calculated. Then, the Ct value for the input sample was adjusted to represent 100% of the DNA amount. This was done by subtracting the log (base 2) of the dilution factor from the average Ct of the input sample. Next, delta Ct values were calculated by subtracting the Ct values of the antibodies of interest from the adjusted input Ct value. Finally, the per cent input values were calculated based on the delta Ct values. This was done by raising two to the power of the negative delta Ct value and multiplying by 100. The percent input values represent the amount of DNA pulled down by the antibody of interest in the ChIP reaction, relative to the amount of starting material (input sample). These values of individuals were used to represent the ChIP-qPCR results in this study.

Table 11: ChIP qPCR primers with respective product size and annealing temperature

Gene	Primer sequence (5' to 3')	Product size	Annealing T
PAX6	F: CCTCCCACTGGCCACTCTAGT	149 bp	56°C
	R: TAGGGGCTTACCAAGAAACTA		
SOX1	F: CAAGTGGTTTGTGCATCAGG	145 bp	56 °C
	R: GACGGAGAGGAATTCAGACG		
NANOG	F: GATGGGGGAATTCAGCTCAGG	140 bp	60 °C
	R: GTCTCTCTTAATCAGCACAGT		
HERV-K	F: AGAGGAAGGAATGCCTCTTGCAGT	151 bp	56 °C
	R: TTACAAAGCAGTATTGCTGCCCGC		
SOX2	F: GCGCTGATTGGTCGCTAGAA	103 bp	56 °C
	R: TGCCTTGACAACTCCTGATACT		
OCT	F: GAGAAGGCGAAATCCGAAGC	125 bp	56 °C
	R: GCCAGAGGTCAAGGCTAGTG		
GATA2	F: ACCCTGTGCATCCACACTCC	80 bp	56 °C
	R: GACGTGTCCCAAGCTTTCTG		
FOXA2	F: GTCCAAAATTGGGGGCGATG	130 bp	56 °C
	R: TGAGAGTCCTGGGGTTAGGG		

Abbreviations: F: forward, R: reverse, T: temperature, bp: base pair

2.2.4 mRNA analysis in human motile spermatozoa (RT-qPCR)

2.2.4.1 RNA isolation and reverse transcription

Five million motile sperm isolated using a swim-up procedure were thawed on ice and washed with PBS (in DEPC-H_2O). RNA was isolated with the Trifast (cat no. 30-2020, VWR Life Science). In brief, the motile sperms were resuspended in 150 µL Trifast and lysed with

TissueLyser LT for 5 min at 4 °C and 50 osc/sec. Afterwards, 350 µL Trifast (total 500 µL Trifast) and 100 µL chloroform were added and mixed thoroughly by hand. The samples were centrifuged (13000 rpm at 4 °C) for 10 min after incubation for 5 min at RT. The upper aqueous phase containing RNA was put into a new tube and precipitated overnight at -20 °C by the addition of one volume of isopropanol (~400 µL) and 5 µL glycogen (20 mg/mL). The next day the samples were centrifuged (13000 rpm at 4 °C) for 30 min, and the pellet was washed twice with 500 µL Ethanol (75 % in DEPC-H_2O) and air-dried at RT. To avoid DNA contamination during RT-qPCRs, the pelleted RNA was treated with DNase I. RNA was dissolved in 16 µL DEPC-H_2O, 2 µL of DNase I and 2 µL of DNase I 10x- buffer were added and the mixture was incubated for 30 min at 37 °C. 2 µL EDTA (50 mM) were added to the samples and incubated at 65 °C for 10 min to inactivate DNase I. RNA concentration was measured (NanoDrop) and samples were stored at -80 °C or used directly for reverse transcription. Reverse transcription (cDNA synthesis) was performed using M-MLV reverse transcriptase (Promega) or QuantiTect Reverse Transcription Kit (cat no. 205310, Qiagen) according to the manufacturer`s protocol with little modification. Using the Promega kit, 300 ng of RNA were incubated with 200 units of the M-MLV reverse transcriptase provided with 5x M-MLV reaction buffer (cat. No. M3681). Two µL dNTP mix (10 mM, Promega), 2 µL random Hexamere (10 mM), 2 µL Poly dT primer mix (10 mM) and 20 units of RNase inhibitor (0.5 µL of 40 U/µL, cat. no. N2511, Promega) were added with DEPC-H_2O to a final volume of 20 µL. This mixture was incubated with the desired amount of RNA for 10 min/25 °C, 60 min/42 °C and 3 min/ 90 °C in the thermocycler for reverse transcription. Afterwards, the cDNAs were stored at -80 °C until PCR analysis. In the other case, the same amount of RNA (300 ng) was reverse transcribed with QuantiTect Reverse Transcription Kit. In brief, a master mix was prepared using 1 µl of reverse transcriptase and primer mix, 4 µL of RT buffer (5x) and was added to the desired amount of RNA and the final volume was adjusted to 20 µL with DEPC-H_2O. The samples were incubated at 42 °C/15 min, and 95 °C/3 min for reverse transcription reaction. The synthesized cDNAs were also stored at -80 °C or used directly for PCR analysis.

2.2.4.2 RT-qPCR of TET1-3, PRC2 components and H3K27me3-binding genes

Real-time PCRs were performed using RT-qPCR thermocycler (Bio-rad) for *TET1-3*, PRC2 components (*EED, EZH2, and SUZ12*), H3K27me3 binding genes (*SOX1, SOX2, OCT4,*

NANOG, PAX6, GATA2, FOXA2, and *HERV-w*). *GAPDH* was used as reference genes to quantify mRNA levels in motile sperm. All the primes with their corresponding product size and annealing temperature are listed in Table 12.

Table 12: RT-qPCR primers with respective product size and annealing temperature

Gene	Primer sequence (5' to 3')	Amplicon	Annealing T
TET1	F: TCCTGGTGCTATTCCAGTCC	110 bp	60 °C
	R: CAGGAAGGAAGACAGGCAAG		
TET2	F: ACTCACCCATCGCATACCTC	113 bp	60 °C
	R: TCAGCATCATCAGCATCACA		
TET3	F: CCCAGAGCTCCAACTGCTAC	138 bp	60 °C
	R: GAGAATACGAAGAAGTTCATCA		
GAPDH	F: TGGAGAAGGCTGGGGCTCAT	176 bp	60 °C
	R: GACCTTGGCCAGGGGTGCTA		
EED	F: GGGCGATTTGATTACAGCCAG	157 bp	60 °C
	R: GAGTCAGTGTTGTACATTTGGC		
EZH2	F: GAGTTGGTGAATGCCCTTGG	195 bp	60 °C
	R: TGCTGTGCCCTTATCTGGAA		
SUZ12	F: CCGAGCACTGTGGTTGAGTA	152 bp	60 °C
	R: AACTGCATCTGATGGTGGTG		
SOX1	F: AACACTTGAAGCCCAGATGGA	150 bp	60 °C
	R: GCAGGCTGAATTCGGTTCTC		
SOX2	F: GCCGAGTGGAAACTTTTGTCG	154 bp	60 °C
	R: GCAGCGTGTACTTATCCTTCTT		
OCT	F: TCGAGAACCGAGTGAGAGG	125 bp	60 °C
	R: GAACCACACTCGGACCACA		
NANOG	F: ATGCCTCACACGGAGACTGT	103 bp	56 °C
	R: AAGTGGGTTGTTTGCCTTTG		
FOXA2	F: GGGAGCGGTGAAGATGGA	89 bp	60 °C
	R: TCATGTTGCTCACGGAGGAGTA		
PAX6	F: TGGGCAGGTATTACGAGACTG	111 bp	60 °C
	R: ACTCCCGCTTATACTGGGCTA		
GATA2	F: GCGGTCCGCTGAACACC	179	60 °C
	R: CAACGGCCCGAGCGA		
HERV-w	F: TGCCCCATCGTATAGGAGTCT	278	60 °C
	R: CATGTACCCGGGTGAGTTGG		

Abbreviations: F: forward, R: reverse, T: temperature, bp: base pair

The qPCR primer sets were either designed with the Primer Blast online tool (NCBI/NIH) or taken from published articles. 15–45 ng of cDNA was amplified using Quantifast SYBR

Green PCR Kit (Qiagen) and the CFX96 Touch Real-Time PCR Detection System (Biorad) for 45 cycles with the following program. The program included an initial activation of DNA Polymerase at 95°C for 5 minutes, followed by denaturation at 95°C for 30 seconds, primer annealing at 60°C for 30 seconds, and extension at 72°C for 30 seconds. These steps were repeated for 45 cycles, with a final melt curve acquisition for product specificity. All PCR reactions were performed in duplicates and relative values for all mRNA levels were calculated using an internal calibrator, housekeeping gene, *GAPDH*. Differences in mRNA levels between groups were identified using the double delta CT (ΔΔCT) method. PCR product sizes were verified on a 2 % agarose gel.

2.2.5 DNA methylation analysis in human motile spermatozoa (Pyrosequencing)

2.2.5.1 DNA isolation, bisulphite treatment and amplification of BS-DNA

Five million motile sperm isolated using a swim-up procedure were thawed on ice and washed with PBS (in DEPC-H_2O). DNA was also isolated with the Trifast (cat no. 30-2020, VWR Life Science). In brief, the motile sperms were resuspended in 150 µL Trifast and lysed with TissueLyser LT for 5 min at 4 °C and 50 osc/sec. Afterwards, 350 µL Trifast (total 500 µL Trifast) and 100 µL chloroform were added and mixed thoroughly by hand. The samples were centrifuged (13000 rpm at 4 °C) for 10 min after incubation for 5 min at RT. The interphase and lower red phases containing DNA were put into a new tube and precipitated overnight at -20 °C by the addition of one volume of isopropanol (~400 µL) and 5 µL glycogen (20 mg/mL). The next day the samples were centrifuged (13000 rpm at 4 °C) for 30 min, and the pellet was washed twice with 500 µL Ethanol (75 % in DEPC-H_2O) and air-dried at RT. DNA isolation started with Proteinase K digestion, followed by phenol/chloroform extraction and alcohol precipitation. Therefore, a semen sample was resuspended in 120 µL Proteinase K buffer, 80 µL DTT (0.1 M) and lysed on Tissue Lyser LT for 5 min at 4 °C and 50 osc/sec. After short centrifugation (30 s, 13000 rpm, 4 °C) foams were removed and overnight protein digestion at 56 °C was initiated by the addition of 10 µL Proteinase K (20 mg/mL in PBS). Next, the aqueous phase was increased with 300 µL of Tris buffer (10 mM, pH 8.0) to facilitate handling. 500 µL of phenol/chloroform solution was added, well mixed per handshaking, and centrifuged for 10 min at 13000 rpm at RT. The DNA containing the upper aqueous phase was supplemented with 500 µL chloroform, well mixed per handshaking and centrifuged again for 10 min at 13000 rpm at RT. DNA was precipitated overnight at -20 °C through the addition of one volume isopropanol, 1/10 volume

NaAc solution (3 M, pH 5.2) and 5 µL glycogen (20 mg/mL) to the aqueous phase. Centrifugation at 13.000 rpm and 4 °C for 30 min pelletized DNA, which was washed twice with 500 µL Ethanol (75 % in ddH_2O) and air-dried at RT. DNA was resuspended in 30 µL Tris EDTA buffer (10 mM, pH 8.0), concentration was measured (NanoDrop) and samples were stored at -80 °C. Bisulphite treatment was done using the EZ DNA Methylation Kit (Zymo Research) according to the manufacturer`s protocol. 500 ng DNA were treated for 16 hours at 50 °C with 100 µL CT Conversion reagent (reconstituted in 750 µL ddH_2O and 210 µL M-Dilution Buffer). Then samples were cooled for 10 min on ice and transferred on a Zymo- Spin™ IC Column and 400 µL M-Binding-Buffer was added. The column was placed into a provided collection tube. After centrifugation at 13000 rpm for 1 min, the flow-through was discarded. The column was washed with 100 µL M-Wash Buffer and the centrifugation step was repeated and flow-through was discarded. The column was treated with 200 µL M-Desulphonation Buffer at RT for 15-20 min to remove incomplete converted cytosines. Subsequent centrifugation at 13000 rpm for 1 min and washing with 200 µL M-Wash-Buffer (twice) was done. Bisulphite converted DNA was eluted in 20 µL M-elution buffer and samples were stored at -80 °C.

Bisulphite-treated DNA (BS-DNA) was amplified using an EpiTect® Whole Bisulfitome kit. In brief, 150-200 ng BS-DNA was taken to a micro-centrifuge tube and the volume was adjusted to 10 µl using nuclease-free water. EpiTect amplification master mix was prepared using EpiTect WBA reaction buffer and REPLI-g Midi DNA Polymerase by following precise protocol from the vendor and centrifuge briefly. Then 30 µl of the EpiTect amplification master mix was added to 10 µl of bisulphite-converted DNA and the solution was incubated at 28°C for 8 h. REPLI-g Midi DNA Polymerase was inactivated by heating the sample for 5 min at 95°C. Samples were stored at -20 °C.

2.2.5.2 Analysis of *LINE1* and *Alu* methylation using COBRA

DNA methylation of *LINE1* and *Alu* (a type of short interspersed nuclear elements, *SINEs)* repetitive elements was investigated using COBRA which is a molecular biology technique used to sensitively quantify DNA methylation levels at a specific genomic locus within a small sample of genomic DNA. This technique, a modification of bisulphite sequencing, combines bisulphite conversion-based PCR with restriction digestion. In brief, the genomic DNA of interest undergoes treatment with sodium bisulphite, which converts unmethylated cytosine

residues to uracil, while leaving methylated cytosine residues unaffected. The bisulphite-treated DNA is then subjected to PCR amplification. This results in the retention of cytosine residues at positions that were originally methylated, and the replacement of cytosine residues with thymine at positions that were originally unmethylated (which were converted to uracil). During this step, the primers used do not contain CpG sites (the common target of cytosine methylation), ensuring that the amplification process does not discriminate between templates based on methylation status. The resulting PCR products are verified using agarose gel electrophoresis. The next step involves treating the PCR products with a restriction enzyme that cleaves only at sites that were originally methylated, leaving the sites that were originally unmethylated intact. Finally, in the quantification step of COBRA, the DNA methylation levels of the original input sample are determined by comparing and quantifying the number of digested and undigested fragments in the gel. In the current study, bisulphite-treated DNA was amplified using MyTaq^{TM} Mix (Bioline) and specific primer sets (table 13) for *LINE1* and *Alu* sequences were taken from Yang et al, 2004 [68].

Table 13: PCR primers with respective product size and annealing temperature

Gene	Primer sequence (5' to 3')	Amplicon	Annealing T
LINE1	F: TTGAGTTGTGGTGGGTTTTATTTAG	413 bp	52 °C
	R: TCATCTCACTAAAAAATACCAAACA		
ALU	F: GATCTTTTTATTAAAAATATAAAAATTAGT	152 bp	52 °C
	R: GATCCCAAACTAAAATACAATAA		

Abbreviations: F: forward, R: reverse, T: temperature, bp: base pair

PCR was conducted for both genes using the following program: initial denaturation and DNA Polymerase activation at 95°C for 2 min, followed by denaturation at 95°C for 30 seconds, annealing at 52°C for 30 seconds, and extension at 72°C for 30 seconds. These steps were repeated for 40 cycles. Subsequently, melt curve acquisition was performed. For *LINE1* PCR products, CpG-sensitive restriction enzyme *HinfI* was used to digest PCR products. This enzyme recognizes the site (5′ G$^{\downarrow}$ANTC 3′.....3′ CTNA$^{\uparrow}$G 5′) in the PCR products and digests only if the site is methylated. Here in the restriction site, N can be Adenine / Guanine / Cytosine / Thymine. For the *ALU* PCR products, CpG-sensitive restriction enzyme *MboI* was used to digest PCR products and this enzyme recognizes the 5′ $^{\downarrow}$GATC 3′...3′ CTAG$^{\uparrow}$ 5′ site in the PCR product and digests it only if the site is methylated. After agarose gel electrophoresis and GelRed staining were performed, the PCR product bands from the

gel image were quantified through densitometric analysis using Empiria Studio software. In both experiments, the digested bands represented methylated repetitive elements, while the undigested bands at the expected product size represented either unmethylated repetitive elements or repetitive elements with mutated restriction sites, as described in Yang et al, 2004 [68]. This software provided values for each product band. The values of the digested (methylated) product bands and the undigested (unmethylated) bands were used to calculate the percentage of methylation for each sample.

2.2.5.3 Pyrosequencing of H3K27me3-binding genes

The quantification of CpG methylation levels in a gene region of interest was performed using Pyrosequencing. The PyroMark Software (Qiagen) and the pyrosequencing instrument Q24 were employed. The manufacturer's protocol and kit were followed, which included forward and reverse PCR primers. Per the Pyrosequencing reaction, 10-20 ng of BS-DNA was amplified using specific primer sets (table 14). Subsequently, the PCR products were immobilized to streptavidin Sepharose beads (GE Healthcare) and subjected to a denaturation buffer to obtain single-stranded DNA, onto which the provided sequencing primer was annealed. The measurement of light emission, proportional to the accurately incorporated nucleotide, was achieved through sequencing by synthesis. The resulting light signals were displayed as peaks in a pyrogram, indicating the methylation levels of the examined CpGs in percentage. In the pyrosequencing experiment, bisulphite-treated DNA from a prostate cancer cell line (PC3) and commercial control DNA set (cat. no. 59695) from Qiagen were utilized as the (un)methylated control. Table 14 displays the sequencing region, along with the expected product size, the covered CpG sites and the Geneglobe catalogue ID.

2.2.6 Cell culture of human cancer cell lines

HeLa, PC3, and LNCaP cell lines served as positive controls for validating PCR primers and antibodies. These cell lines were cultured following the manufacturer's guidelines. HeLa cells, originating from cervical cancer, were cultured in Dulbecco’s modified Eagle’s medium (DMEM) medium. The PC3 cell line, derived from grade IV prostate cancer, was cultured in the DMEM-F12 medium, while LNCaP cells, isolated from the lymph node of

prostate carcinoma, were grown in the RPMI 1640 medium. The media were supplemented with 10 % fetal calf serum, 1 % penicillin/streptomycin (Life Technologies GmbH) and cultured under a 5% CO_2 atmosphere at 37 °C. Upon reaching 90% confluency, the cells were washed with PBS, trypsinized, and collected. The cell-containing PBS was then centrifuged at 500g for 2 minutes at 4 °C. After discarding the supernatants, the cell pellets underwent Trifast processing for DNA/RNA extraction. In brief, 1000µl of Trifast and 200µl of chloroform were added to the cell pellet and vortexed. The mixture was left at room temperature for 5 minutes and then centrifuged at 13000 rpm for 10 minutes. The aqueous phase was transferred to a new Eppendorf tube for RNA extraction. 500µl of isopropanol and 5µl of glycogen were added and mixed gently by vortexing. The mixture was incubated at -20°C overnight.

Table 14: Pyrosequencing primers, amplicon size, CpG sites with their catalogue ID

Gene	Sequence to Analyse	Amplicon	CpG site	Geneglobe ID
EZH2	CGATGGCGATTGGGCTGCCGCGTTTGGCG CTCGGTCCG	107 bp	7	PM00029806
SUZ12	CGAAAATGGAGCACGTCCAGGCTGACCA CGA	150 bp	3	PM00182602
EED	GCGGGTCGGAGATCGAAGGAACGGGCCA ATTGCGGC	114 bp	5	PM00151627
SOX2	CTGCTGGGGGCGGCGGGTTCCGGCACCTC GGCGCCGGGGA	205 bp	6	PM00016856
PAX6	CGAGCGGCTGCCTGGTCCTGCGCTGTCCC GAGCTCTCTCGGCGGCTTCG	185 bp	7	PM00045514
GATA2	CGCGCGCGCGCGGGCGGGGCGCGA	231 bp	9	PM00109396
FOXA2	CTCGGGTCGGTCCCGCGGGCCTGGCGGCC GACAGGGGGCGC	249 bp	7	PM00196700

Abbreviations: base pair: bp

To the phenol phase and interphase 20µl of Sodium Acetate (NaAC), 5µl of glycogen, and 600µl of 100% Ethanol (EtOH) were added for DNA extraction. The mixture was mixed thoroughly by shaking and incubated at -20°C overnight. The next day, the mixtures were centrifuged at 4°C, 13000 rpm for 20 minutes. For the RNA pellet, the supernatant was discarded, 150µl of 70%-75% EtOH was added, and the mixture was centrifuged at 13000 rpm

for 5 minutes. The supernatant was thrown away, and the pellet was left to dry. For the DNA pellet, the phenol phase was aspirated for protein extraction or discarded. 150µl of 70%-75% EtOH was added, and the mixture was centrifuged at 13000 rpm for 5 minutes. The supernatant was thrown away, and the pellet was left to dry. For both RNA and DNA, the pellets were not allowed to dry too long. To dissolve the pellet, 20µl of DEPC water was added for the RNA and 30µl of ddH2O was added for the DNA. For protein isolation Urea buffer was used (protein extraction protocol described in section 2.2.2.3.1). After the isolation of DNA, RNA and protein, concentrations were measured using NanoDrop and kept frozen at -80 °C if not used directly.

2.3 Statistical analyses

Data compilation, encompassing calculations such as Cycle threshold (Ct) value computation and methylation percentage determination, was executed in Microsoft Excel. Subsequent figure preparation and statistical analyses were conducted utilizing GraphPad Prism 8.2. An initial assessment was performed to determine if the data conformed to a normal distribution. In all instances, including data derived from differential mRNA expression and CpG-methylation analysis, the data did not adhere to a normal distribution. Thus, non-parametric variables were compared employing the Mann-Whitney U-test. The Mann-Whitney U-test essentially ranks all the observations from both groups together and then evaluates whether the ranks for each group differ significantly.

This thesis selectively incorporated significantly correlated data by analysing the mRNA levels and methylation levels of all candidate genes along with clinical data from the Spermiogram. The Pearson correlation analysis was employed to examine the relationship between variables, given its appropriateness for quantitative data. The outcomes are articulated as mean values, with the standard error of the mean (SEM) denoted as mean ± SEM. A P-value of ≤ 0.05 was designated as the threshold for statistical significance. The following thresholds were used to interpret the results: $P<0.0001$ was considered extremely significant; $P<0.001$ was deemed very significant; $P<0.01$ was regarded as significant; $P<0.05$ was viewed as weakly significant; and $P \geq 0.05$ was not considered significant and stated as “ns” in the thesis. All mRNA and methylation analysis data are documented in Table 23-28.

3. Results

3.1 Expression of TET1, H3K27me3 and PRC2 in human spermatogenesis

3.1.1 Detection of TET1 in spermatogonia B, leptotene spermatocytes, round spermatids

The results of IHC analysis in human testis samples showed that TET1 expression started in stage I spermatogonia B and continued until stage IV leptotene spermatocytes. It was then detectable again from stage I to stage III round spermatids. Sertoli, Leydig, and peritubular cells tested positive for TET1, with the staining being weaker in peritubular cells compared to other cell types (Fig. 9A-B). SCs exhibited a stronger TET1 expression pattern compared to germ cells. The images below depict the results of nuclear antibody staining. Deep or light red staining indicates positive results, while blue staining, which is a result of haematoxylin, indicates negative results. To confirm the specificity of the experiment, a negative control was used. Fig. 9C represents negative control (only the secondary antibody, with no primary antibody applied) showing only haematoxylin staining but no positive signal at the same stages (I-II). To visualize the different cell types, a higher magnification (100x) image was taken for the upper section of Fig. 9B, where stage III and IV cell types were visible, and is presented in Fig. 9D. Additionally, a scheme of different stages of spermatogenic cell types was created, where all positive staining was marked with a green rectangle, as shown in Fig. 9E.

3.1.2 Detection of H3K27me3 in spermatogonia B, leptotene spermatocytes, round spermatids

According to the IHC results, the expression of H3K27me3 was detected in stage I spermatogonia B and continued to be present until stage IV leptotene spermatocytes. After that, its expression was detectable again from stage I to stage III round spermatids. The staining intensity was found to be higher in all germ cells that showed a higher expression of this protein. Leydig cells were also found to be positive for H3K27me3, as shown in Fig. 10A. To verify the specificity of the staining, a negative control was utilized in the experiment. This involved the use of only the secondary antibody, with no primary antibody applied, as depicted in Fig. 10B. This section only displayed haematoxylin staining, indicating that there were no unspecific H3K27me3 signals. Additionally, a schematic representation of different stages of

spermatogenic cell types is included in Fig. 10C, with all positively stained cells marked with a green rectangle.

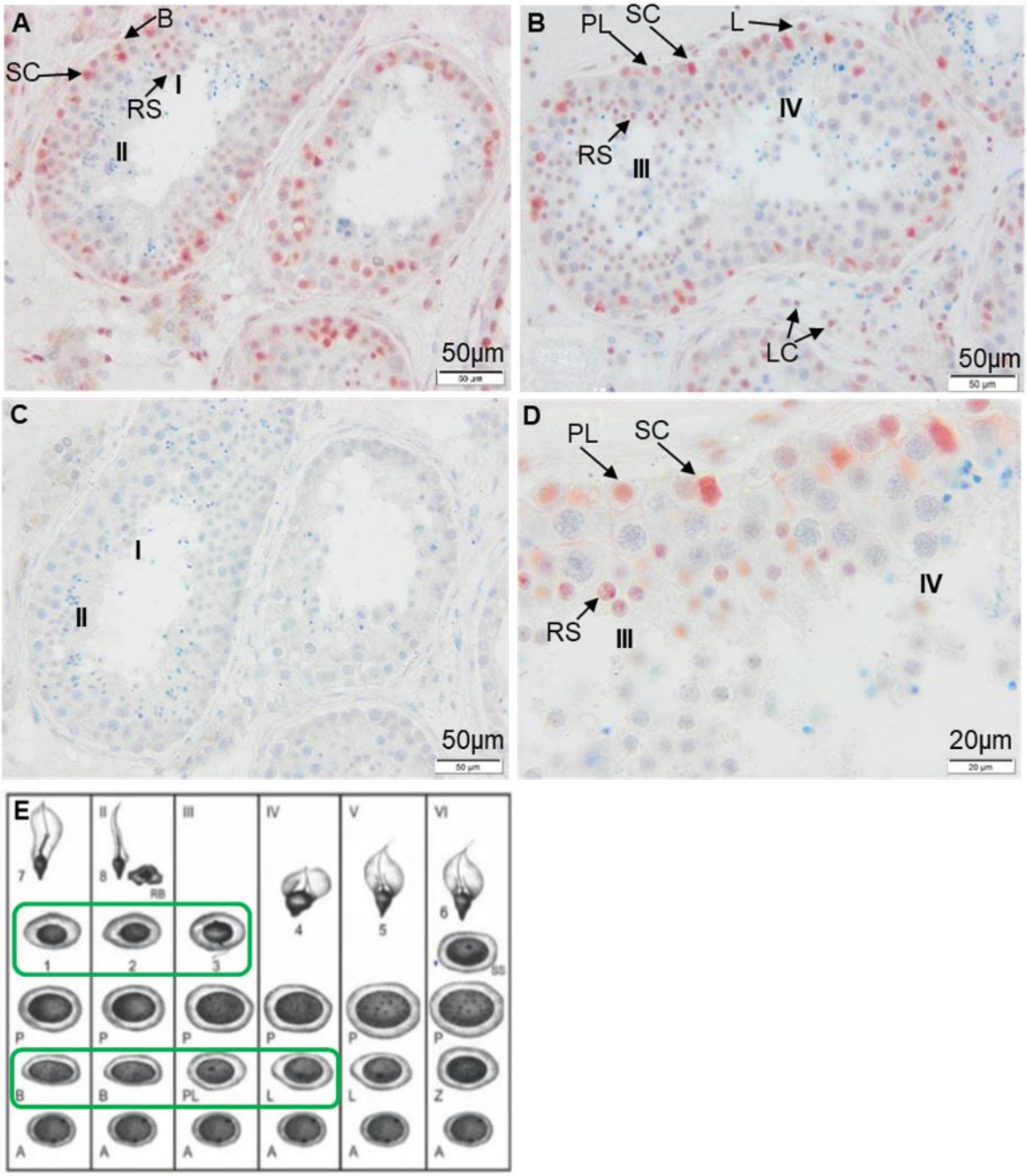


Figure 9: TET1 expression in human testicular tissue sections. A-B) TET1 expression starts in stage I spermatogonia B and persists until stage IV leptotene spermatocytes (L). Stage I-III round spermatids (RS), Sertoli cells (SC), and Leydig cells (LC) were also found positive. C) Negative control of serial section 1A. D) Magnified view of a section of figure B. E) Schematic representation of a human seminiferous epithelial cycle, modified from Clermont, 1963[69]. Roman numbers = various stages of the seminiferous epithelial cycle. Green rectangle= positively stained cells.

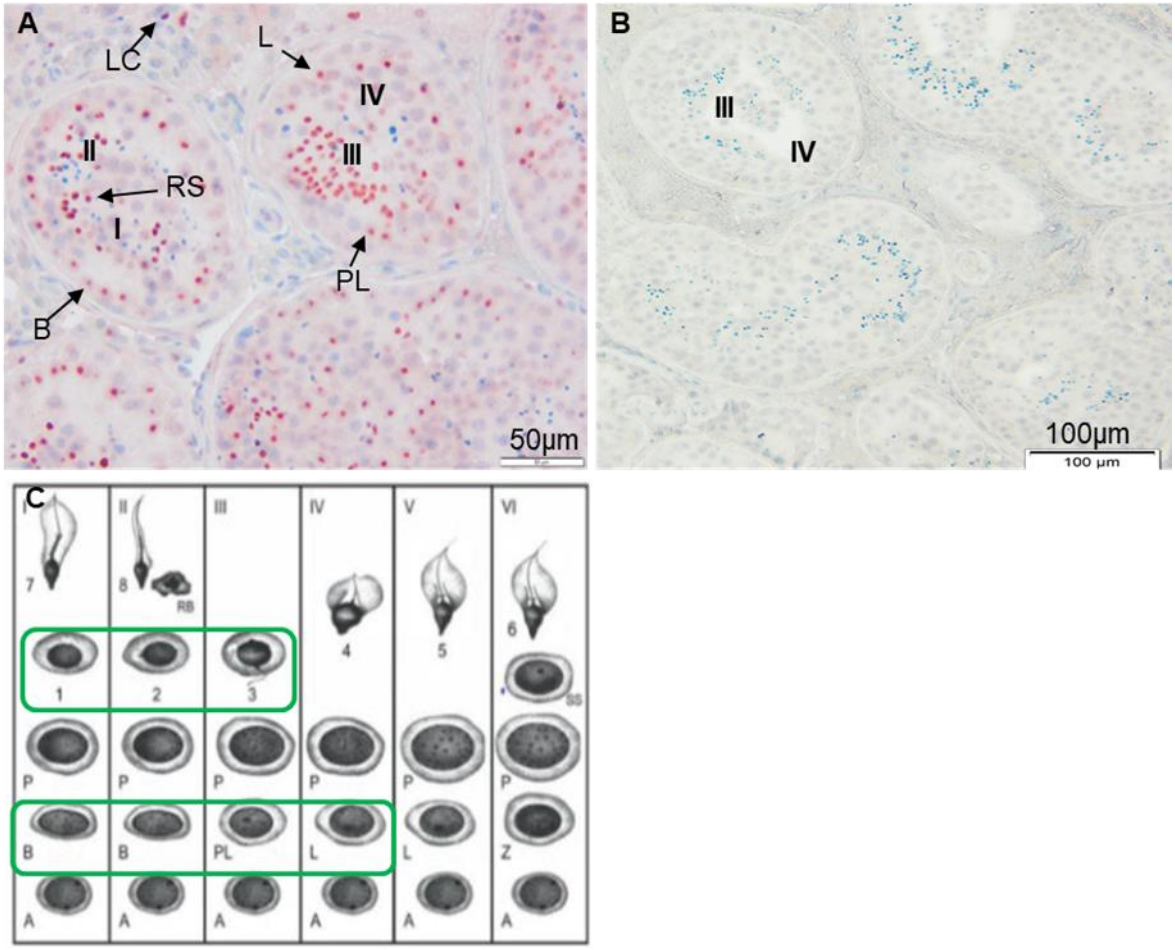


Figure 10: H3K27me3 expression in testicular tissue sections of humans. A) H3K27me3 expression starts in stage I spermatogonia B and persists until stage IV leptotene spermatocytes (L). Stage I-III round spermatids (RS), and Leydig (LC) cells were also found positive. B) Negative control of serial section. C) Schematic representation of a human seminiferous epithelial cycle, modified from Clermont, 1963[69]. Roman numbers = various stages of the seminiferous epithelial cycle. Green rectangle= positively stained cells.

3.1.3 Detection of EZH2 in spermatogonia B, leptotene spermatocytes, round spermatids

The IHC results revealed that EZH2 expression was detectable in stage II spermatogonia B, stage III preleptotene, stage IV-V leptotene spermatocytes, stage I-III round spermatids, as well as in SCs, as shown in Fig. 11A-B. To ensure the specificity of the staining, a negative control (use of only the secondary antibody, with no primary antibody applied) was used in the experiment as shown in Fig. 11C. This section displayed only haematoxylin staining, indicating that there was no positive signal of EZH2 present at the same stages (I-III). Furthermore, a schematic representation of different stages of spermatogenic cell types is included in Fig. 11D, with all positively stained cells marked with a green border.

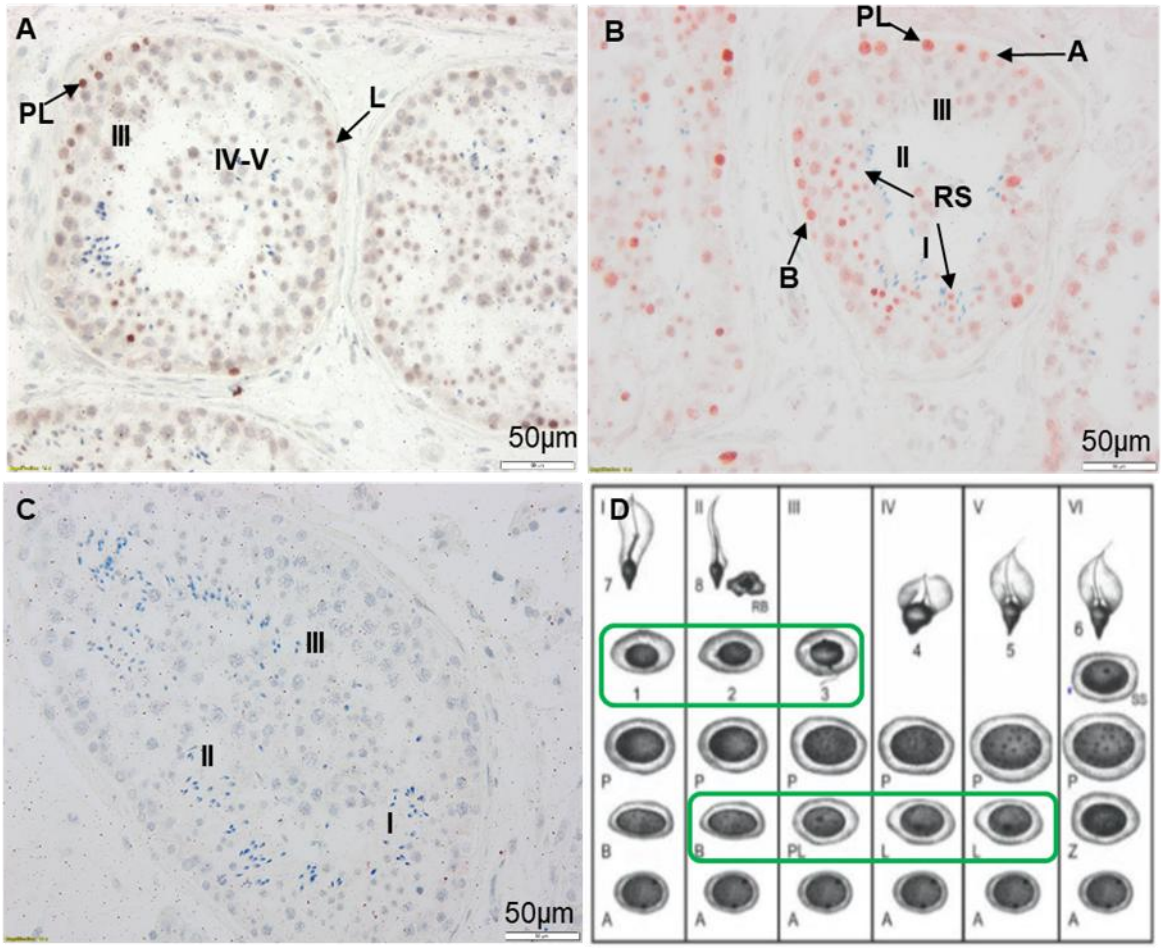


Figure 11: EZH2 expression in testicular tissue sections of humans. A-B) EZH2 starts expression in stage II spermatogonia B and persists until stage V leptotene spermatocytes (L), also expressed in stage I-III round spermatids (RS). C) Negative control of serial section. D) Schematic representation of a human seminiferous epithelial cycle, modified from Clermont, 1963[69]. Roman numbers = various stages of the seminiferous epithelial cycle. Green rectangle= positively stained cells.

3.1.4 EED and SUZ12 analysis did not function with commercial antibodies

To investigate SUZ12 and EED expression, different conditions were tested using two anti-EED and two anti-SUZ12 antibodies. However, specific staining could not be observed, as shown in Fig. 12A-B. The major conditions that were tested included the use of Tris EDTA buffer (pH 9.0) and Citrate buffer (pH 6.0) for antigen retrieval, and the use of 5% bovine serum albumin (BSA) and 5% non-fat dry milk to block non-specific binding sites. For SUZ12, two different anti-SUZ12 antibodies were used from two different vendors, with cat no. ab12073 (two batches, Abcam) and D39F6 (Cell Signaling). As per the vendor's product sheet ab12073 was shown to be expressed in human cancer lines like Hela, and HEP1 and was validated in SUZ12 knockout cell lines. The same antibody was tested in my western blot analysis and was detectable in PC3 cancer cell line protein lysate. For EED, two different anti-EED antibodies

from Abcam were used, with cat no. ab236292 (As per vendor's product sheet, expressed in colon cancer tissues, K562 and Hela cell line) and ab254569 (was tested in our lab and was detectable in Hela cell lysate. As per vendor's product sheet, ab254569 found expressed in human heart and testis tissues). All antibodies were tested with a working dilution ranging from 1:50 to 1:800. To ensure the specificity of the staining, negative controls (use of only the secondary antibody, with no primary antibody applied) were included in the experiment (Fig. 12C-D).

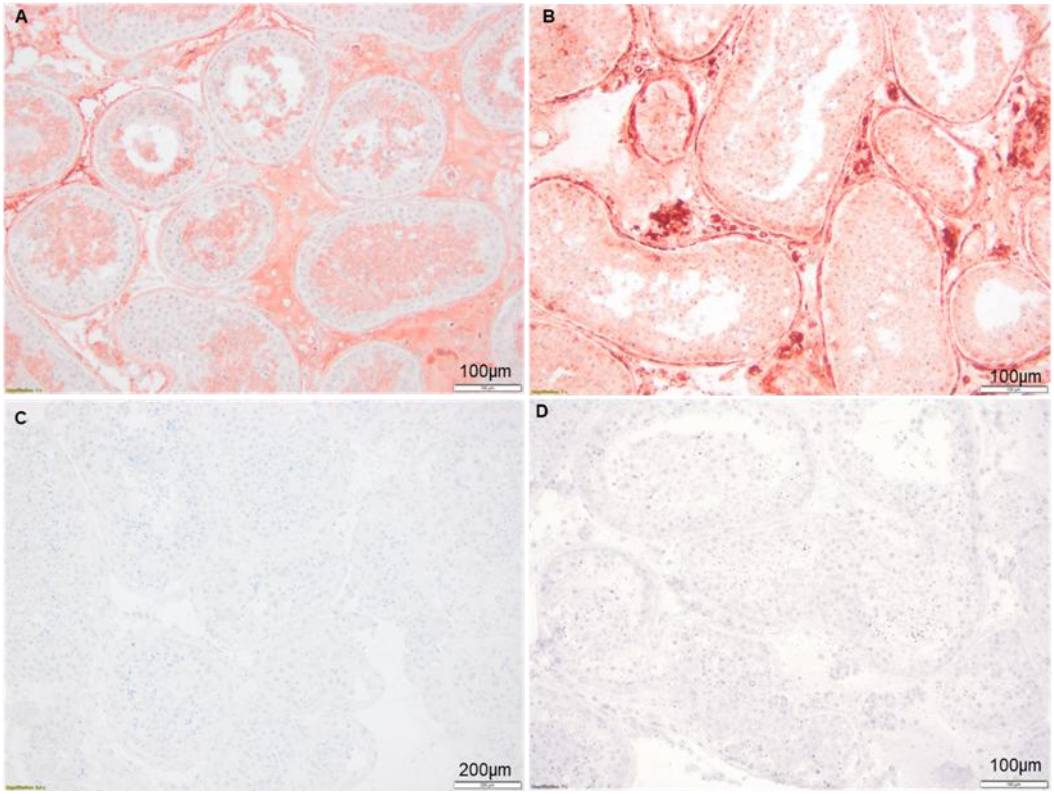


Figure 12: EED and SUZ12 expression in testicular tissue sections of humans. A) Anti-EED antibody showed non-specific staining C) Negative control of serial section for EED. B) Anti-SUZ12 antibody also showed non-specific staining. D) Negative control of serial section for SUZ12.

3.2 Expression of Tet1, H3K27me3 and PRC2 in Tet1 $^{(-/-)}$ mouse spermatogenesis

To assess the expression pattern of Tet1, H3K27me3, and PRC2 components (Eed, Suz12, and Ezh2), IHC was carried out on testis sections obtained from three different mouse models: wild type (WT) mice, Tet1 $^{(+/-)}$ mice, and Tet1 $^{(-/-)}$ mice.

3.2.1 In *Tet1* $^{(-/-)}$ mice, Tet1 was not expressed in germ cells

In the testis of WT mice, Tet1 expression was detected in stage X-XI elongating spermatids, Leydig cells, and SCs (Fig 13A). A serial section of the same testis showing the same stages was used as a negative control (Fig. 13C).

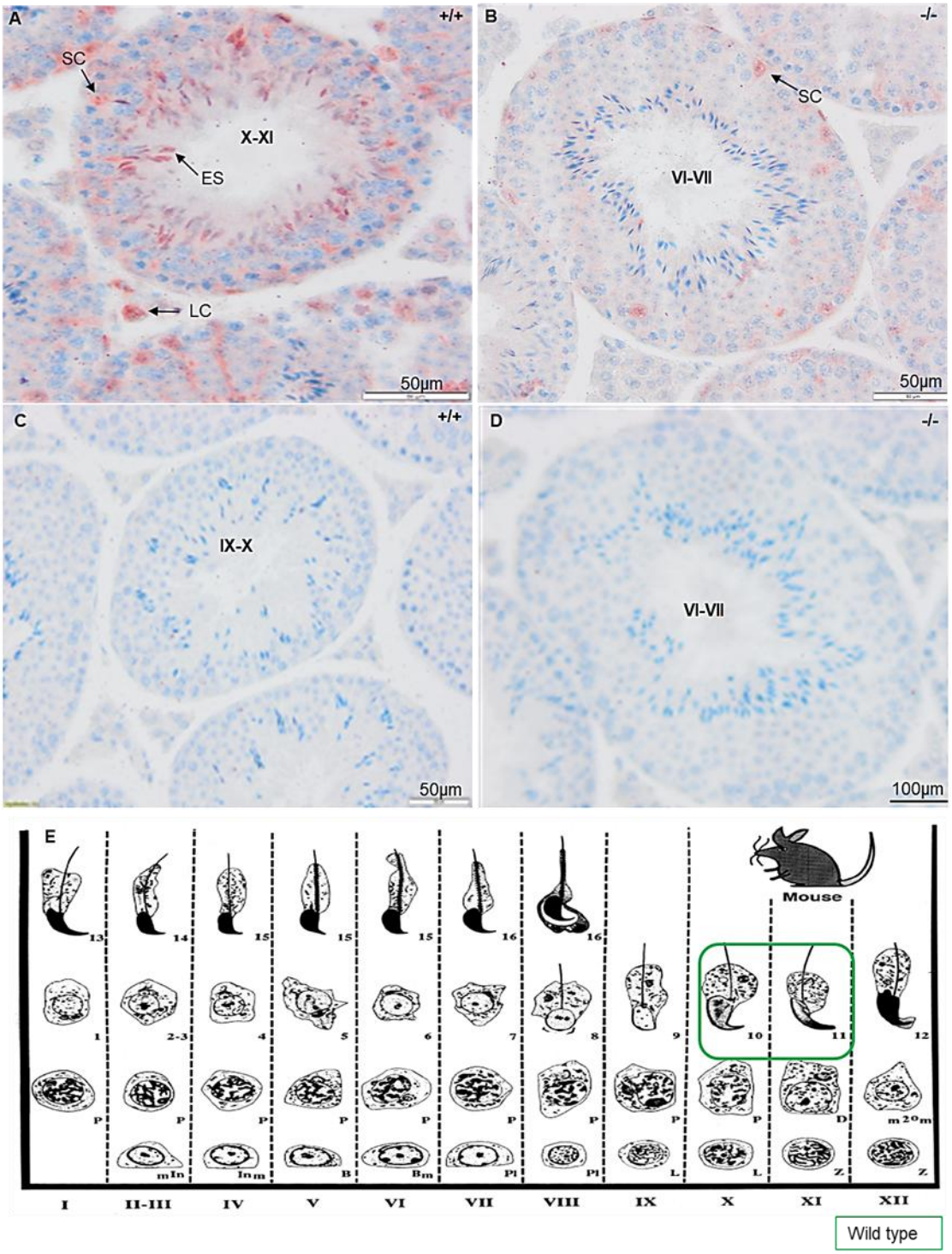


Figure 13: Tet1 expression in testicular tissue sections of mouse. A) In WT mice, Tet1 was expressed in stage X-XI elongating spermatids (ES). Sertoli cells (SC), and Leydig cells (LC) were also positive. B) *Tet1* $^{(-/-)}$ showed no expression

in germ cells, but SCs cells were found positive. C-D) Negative controls of serial sections. E) Schematic representation of mouse seminiferous epithelial cycle, modified from Russell, 1990. Roman numbers = different stages of the seminiferous epithelial cycle. A, C = wildtype (+/+); B, D = *Tet1* KO (-/-). Green rectangle= positively stained cells.

In *Tet1* $^{(-/-)}$ mice testis, Tet1 expression was absent in all germ cells, but SCs showed a positive signal (Fig. 13B). Since *Tet1* knockout was germ cell-specific, somatic cells showed positive signals while germ cells did not. To ensure the specificity of the staining, a serial section of the same testis showing the same stages was used as a negative control (use of only the secondary antibody, with no primary antibody applied) for *Tet1* $^{(-/-)}$ mice (Fig. 13D). A scheme of different stages of spermatogenic cell types is presented in Fig. 13E, where all positively stained cells are marked with a green rectangle.

3.2.2 Shift of H3K27me3 from spermatogonia B to preleptotene spermatocytes

In *Tet1* $^{(-/-)}$ mice, a slight shift in the pattern of H3K27me3 expression was observed compared to WT mice. Specifically, in WT mice, H3K27me3 expression was present in stage V-VI spermatogonia B and stage IX-XI elongating spermatids (as depicted in Fig. 14A). On the other hand, in *Tet1*$^{(-/-)}$ mice, H3K27me3 was expressed in stage VII-VIII pre-leptotene spermatocytes and stage X-XI elongating spermatids (as illustrated in Fig. 14B). To ensure specificity, a serial section of the same testis with the same stages was used as a negative control (use of only the secondary antibody, with no primary antibody applied) for *Tet1*$^{(-/-)}$ mice, respectively (as shown in Fig. 14C and 14D). Leydig cells were also positive in both types of cells. The staining intensity was equivalent in both WT and *Tet1* $^{(-/-)}$ cell types. Fig 14E provides a scheme of various stages of spermatogenic cell types, with all positive staining marked with green (WT) and red (*Tet1* $^{(-/-)}$) borders.

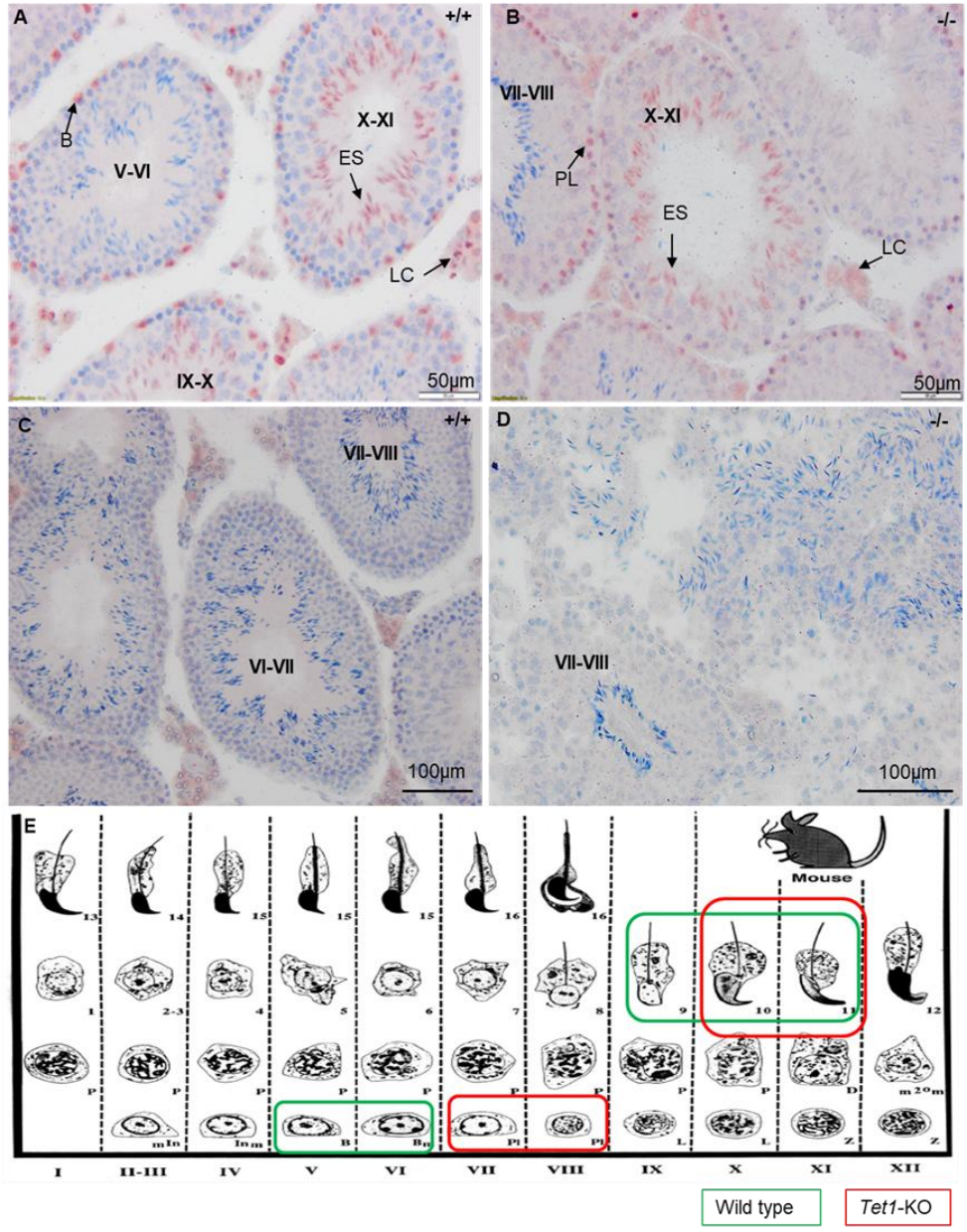


Figure 14: H3K27me3 expression in testicular tissue sections of mouse. A) In WT mice, H3K27me3 expression was detectable in stage V-VI spermatogonia B and stage IX-XI elongating spermatids (ES). B) In *Tet1* $^{(-/-)}$ mice, H3K27me3 was expressed in stage VII-VIII pre-leptotene spermatocytes (PL), stage X-XI elongating spermatids (ES). Leydig cells (LC) were also positive in both cell types. C-D) Negative controls of serial sections. E) Schematic representation of mouse seminiferous epithelial cycle, modified from Russell, 1990: Roman numbers = different stages of seminiferous epithelial cycle. A, C = wildtype (+/+), B, D = *Tet1* KO (-/-). Green/red rectangle= positively stained cells.

3.2.3 Absence of Eed expression in *Tet1* $^{(-/-)}$ mice

The expression pattern of Eed, a component of PRC2, was analysed in WT, *Tet1*$^{(+/-)}$, and *Tet1*$^{(-/-)}$ mice. In WT mice, Eed expression was detected only in stage IX-X elongating spermatids (as shown in Fig. 15A). However, in *Tet1*$^{(-/-)}$ mice, Eed expression was absent (as depicted in Fig. 15B). In *Tet1*$^{(+/-)}$ mice, Eed was expressed not only in elongating spermatids but also in

various cell types, ranging from stage V spermatogonia B to stage XII zygotene spermatocytes, and in stage V-VIII pachytene spermatocytes (as illustrated in Fig. 15C). The positive signals were confirmed by comparing negative control (use of only the secondary antibody, with no primary antibody applied) staining (as shown in Fig. 15D-F, respectively).

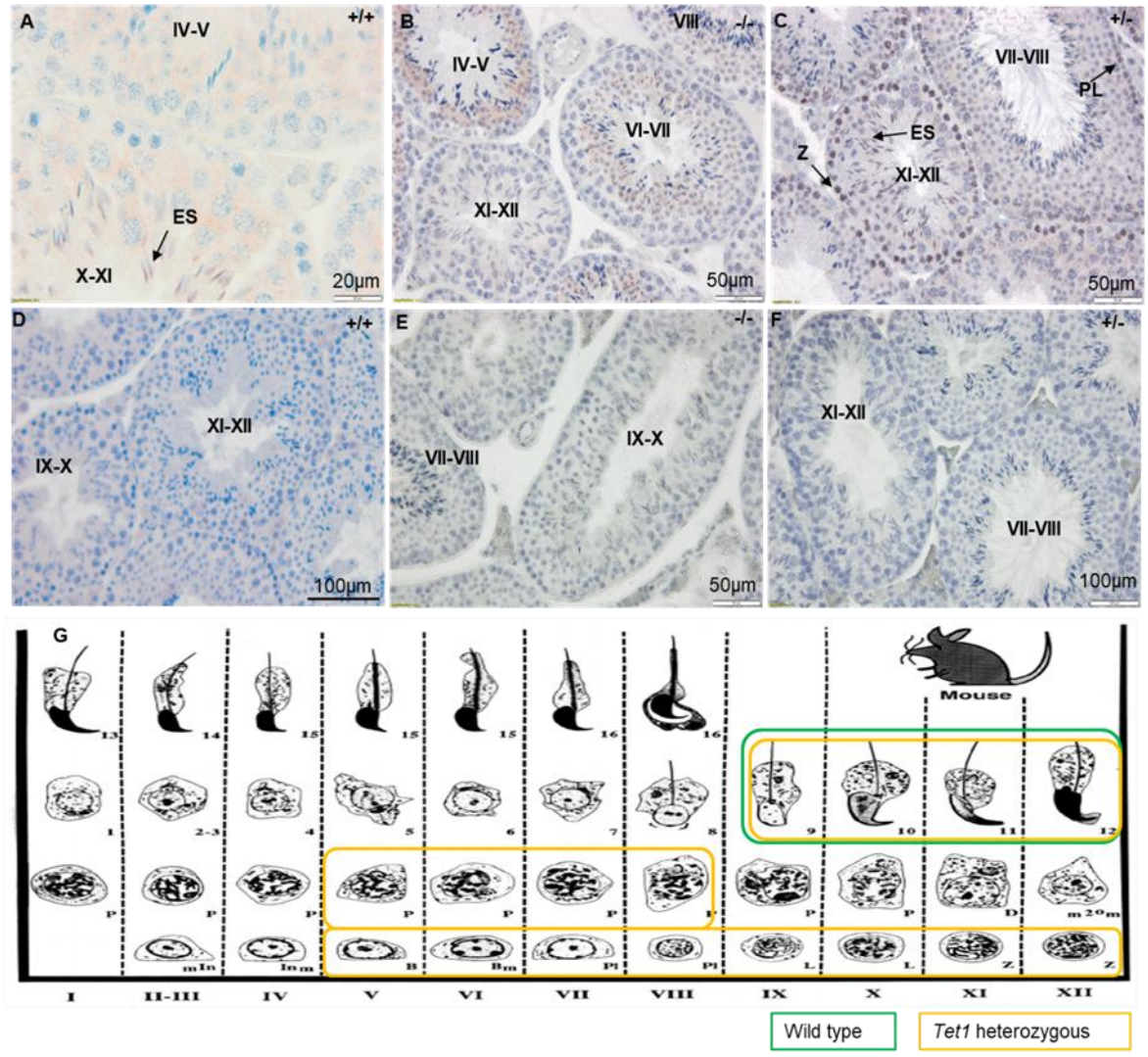


Figure 15: Eed expression in testicular tissue sections of the mouse. A) In WT mice, Eed was detectable in stage IX-XII elongating spermatids (ES). B) In *Tet1* $^{(-/-)}$ mice, Eed expression was not detectable. C) In *Tet1* $^{(+/-)}$ mice, expression starts in stage V spermatogonia B and persists until stage XII zygotene spermatocytes (Z), stage V-VIII pachytene spermatocyte (P), and in stage IX-XII elongating spermatids (ES). D, E, F) Negative controls of serial sections. G) Schematic representation of mouse seminiferous epithelial cycle, modified from Russell, 1990. Roman numbers = different stages of the seminiferous epithelial cycle. A, D = wildtype (+/+), B, E = *Tet1* KO (-/-). C, F = *Tet1* heterozygous (+/-). Green/yellow rectangle= positively stained cells.

3.2.4 Detection of Suz12 in elongating and maturing spermatids

The expression pattern of Suz12, another component of PRC2, was analysed in WT (WT), *Tet1*$^{(+/-)}$, and *Tet1*$^{(-/-)}$ mice. In WT and *Tet1*$^{(+/-)}$ mice, Suz12 expression was detected in

spermatogonia B, preleptotene, and pachytene spermatocytes. Additionally in WT mice, Suz12 expression was also observed in leptotene spermatocytes (as shown in Fig. 16A). Interestingly, *Tet1*$^{(-/-)}$ mice exhibited new expression patterns in elongating and maturing spermatids, as well as in peritubular somatic cells (Fig. 16B), which differed from the expression patterns of WT and *Tet1*$^{(+/-)}$ mice. In *Tet1*$^{(+/-)}$ mice, Suz12 was also expressed in stage V-VIII round spermatids (as shown in Fig. 16C).

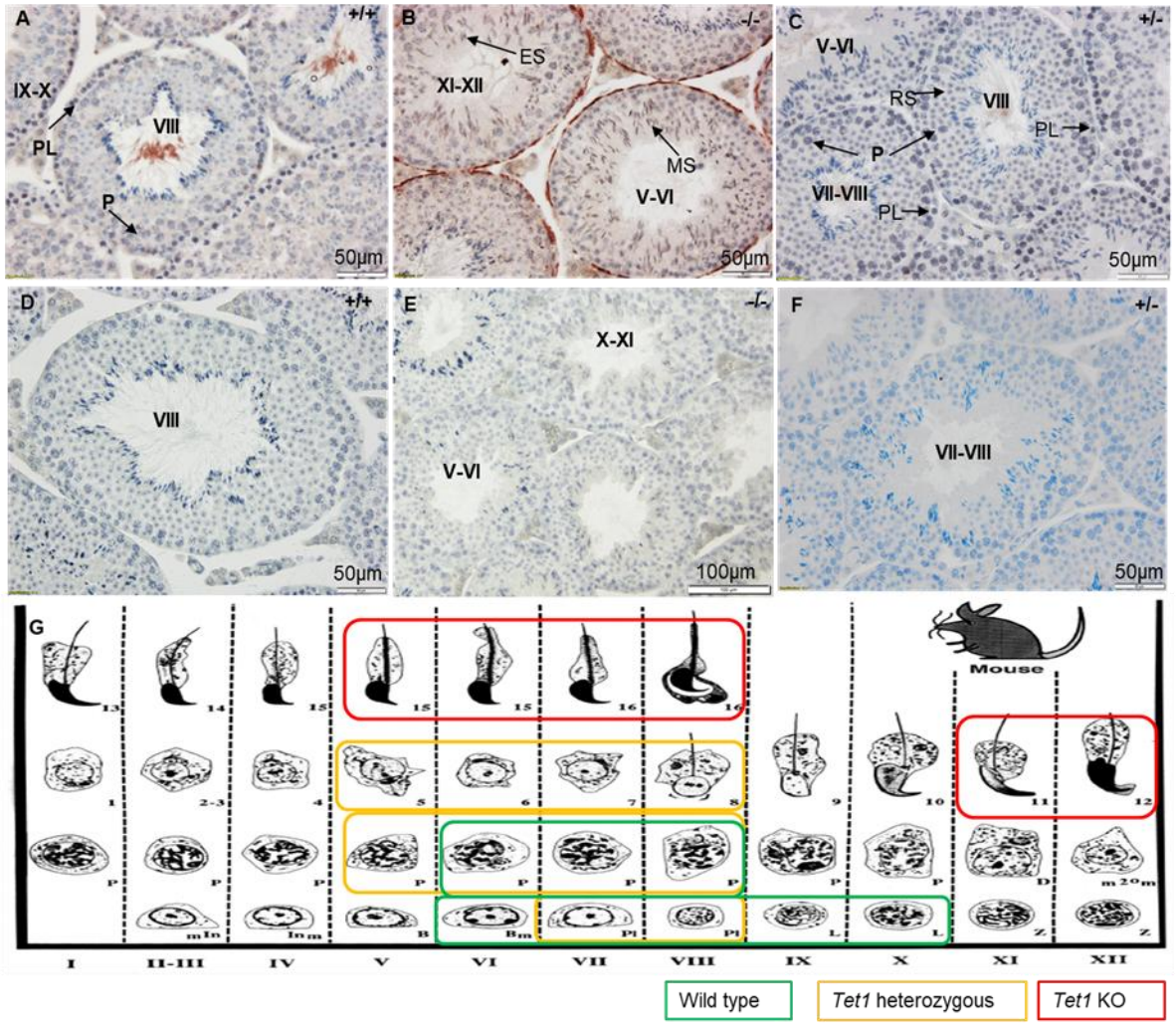


Figure 16: Suz12 expression in testicular tissue sections of mouse. A) In WT mice, Suz12 expression starts in stage VI spermatogonia B and persists until stage X leptotene spermatocytes (L), in stage VI-VIII pachytene spermatocytes (P). B) In *Tet1* $^{(-/-)}$ mice, Suz12 was expressed in stage XI-XII elongating spermatids (ES) and stage V-VIII maturing spermatids (MS). C) In *Tet1* $^{(+/-)}$ mice, Suz12 was expressed in stage VII-VIII pre-leptotene spermatocytes (PL), in stage V-VIII pachytene spermatocytes (P) and round spermatids (RS). D, E, F) Negative controls (use of only the secondary antibody, with no primary antibody) of serial sections. G) Schematic representation of mouse seminiferous epithelial cycle, modified from Russell, 1990. Roman numbers = different stages of the seminiferous epithelial cycle. A, D = wildtype (+/+), B, E = *Tet1* KO (-/-). C, F = *Tet1* heterozygous (+/-). Green/red/yellow rectangle= positively stained cells.

3.2.5 Detection of Ezh2 in pre-leptotene spermatocytes, pachytene spermatocytes, diplotene spermatocytes, and round spermatids

The expression pattern of Ezh2 was also analysed in WT, *Tet1*$^{(+/-)}$, and *Tet1*$^{(-/-)}$ mice. In all types of mice, Ezh2 expression was observed in diplotene spermatocytes and round spermatids (as shown in Fig. 17A-C). However, only in *Tet1*$^{(-/-)}$ mice, Ezh2 expression was also detected in pre-leptotene spermatocytes and pachytene spermatocytes (Fig. 17B). Staining intensities indicated that this epigenetic factor was highly expressed in all types of mouse testes.

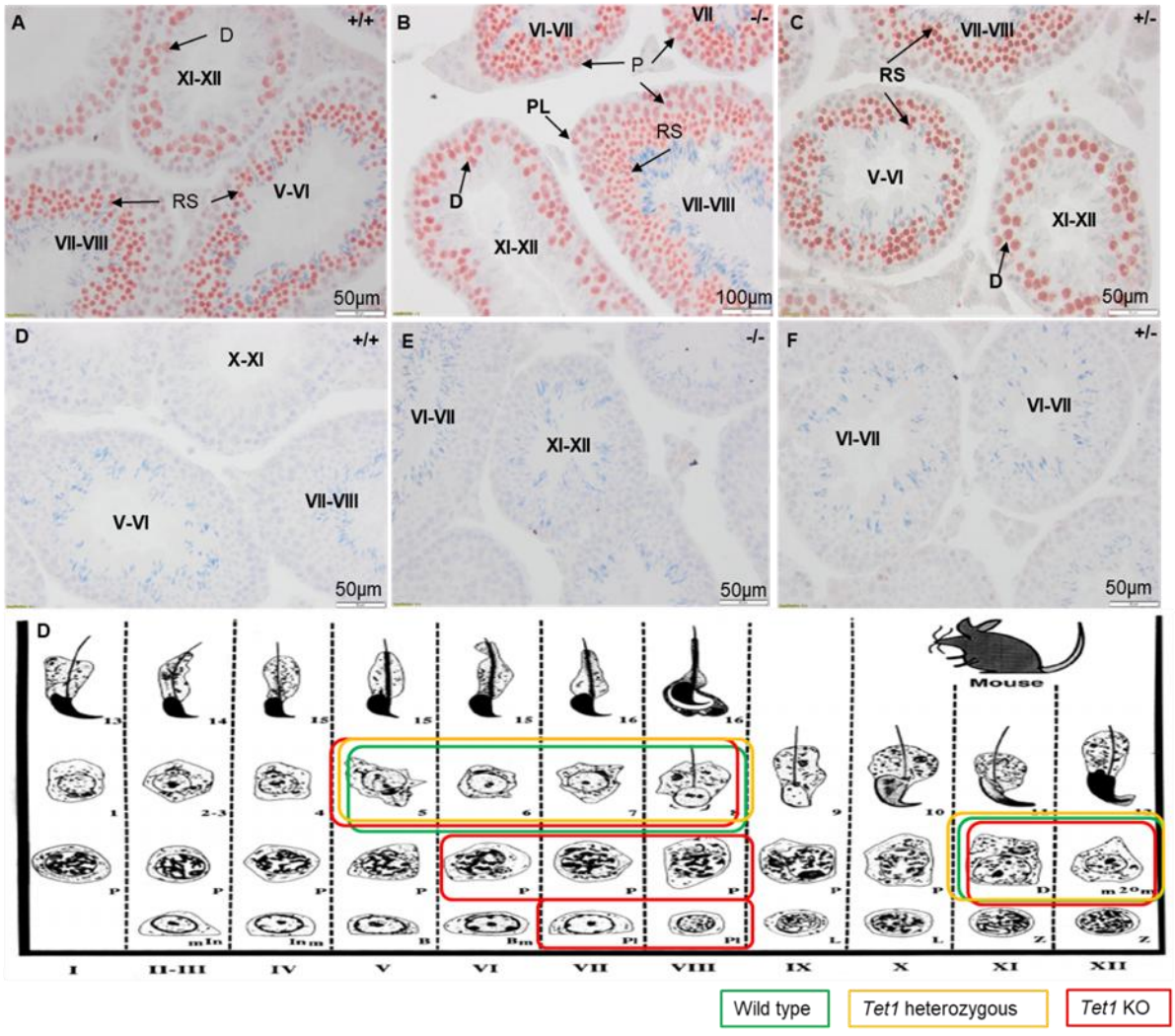


Figure 17: Ezh2 expression in testicular tissue sections of mouse. A) In WT mice, Ezh2 was detectable in stage XI-XII diplotene spermatocyte (D), and stage V-VIII round spermatids (RS). B) In *Tet1* $^{(-/-)}$ mice, Ezh2 was expressed in stage VII-VIII pre-leptotene spermatocytes (PL), stage VI-VIII pachytene spermatocyte (P), stage XI-XII diplotene spermatocyte (D), stage VI-VIII round spermatids (RS). C) In the *Tet1* $^{(+/-)}$ mouse, Ezh2 was expressed in stage XI-XII D, stage V-VIII RS. D, E, F) Negative control (use of only the secondary antibody, with no primary antibody) of serial sections. G) Schematic representation of mouse seminiferous epithelial cycle, modified from Russell, 1990. Roman numbers = different stages of the seminiferous epithelial cycle. A, D = wildtype (+/+), B, E = Tet1 KO (-/-). C, F = *Tet1* heterozygous (+/-). Green/red/yellow rectangle= positively stained cells.

3.3 Investigation of TET1, H3K27me3, EED, SUZ12 and EZH2 in motile sperm of fertile and subfertile men

To investigate the retention of TET1, H3K27me3, EED, SUZ12, and EZH2, western blotting was performed on motile sperm samples collected from subfertile men and fertile controls. However, for mouse sperm samples, it was not possible to collect mature sperm from *Tet1*$^{(-/-)}$ mice as they had a shorter lifespan and typically died by four weeks of age, making the experiment unfeasible.

3.3.1 TET1 was not detectable in the motile sperm of fertile men

Western blotting experiments were conducted to investigate the presence of TET1 protein in the motile sperm of fertile men. After optimizing the experimental conditions, it was found that TET1 protein was not detected in the sperm samples, while it was present in the positive control samples (Fig. 18 A1).

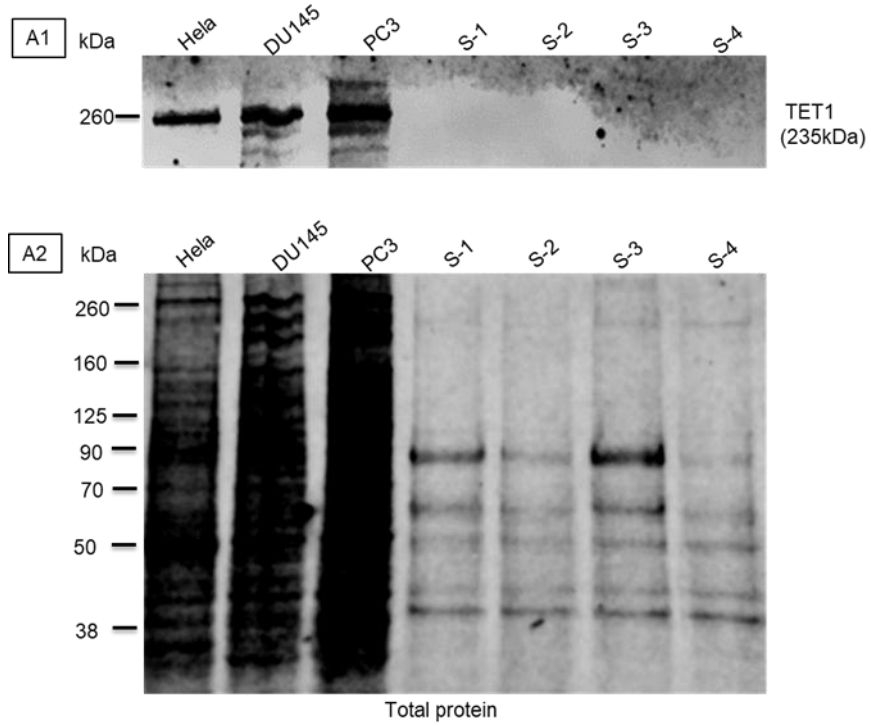
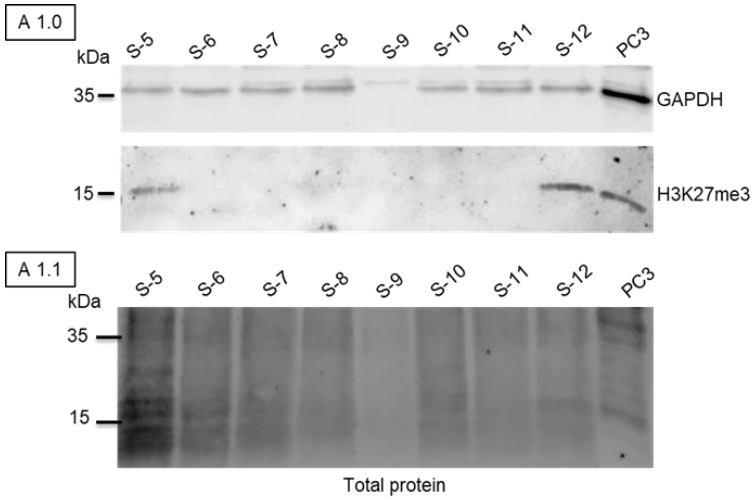


Figure 18: TET1 WB-analysis in motile sperm of fertile men. A1) TET1 (235kDa) was not detectable in the protein lysates of motile sperm of fertile men (S1 to S4) while human cancer cell lines (Hela, DU145, PC3) showed positive signals. A2) Staining of total protein on PVDF membrane after the transfer from the gel.

Total protein staining revealed the presence of various proteins in the sperm protein lysates, including proteins with molecular weights in the expected size range of TET1 protein (250-260 KDa). Fig. 18 A2 shows the total protein staining used as a control for quantification.

3.3.2 H3K27me3 retained at higher amounts and more frequently in sperm of subfertile men compared to fertile men

To investigate the presence of H3K27me3 in human sperm of subfertile and fertile men, a western blot analysis was performed. Interestingly, the frequency of H3K27me3 retention was found to be higher in patients (9 out of 16) compared to fertile men (2 out of 9) as shown in Fig. 19 A1.0 and B1.0, B2.0, B3.0. The corresponding blot was also stained for total protein before antibody incubation and is shown in Fig. 19 A1.1 and B1.1, B2.1, and B3.1. This result was further confirmed via the quantification of blots using Empiria Studio (v2.2) software. To visually represent the results, the quantification values were plotted in a scattered plot, which revealed that patients have significantly higher H3K27me3 modification compared to the controls (p = 0.0042) (Fig.19 C).



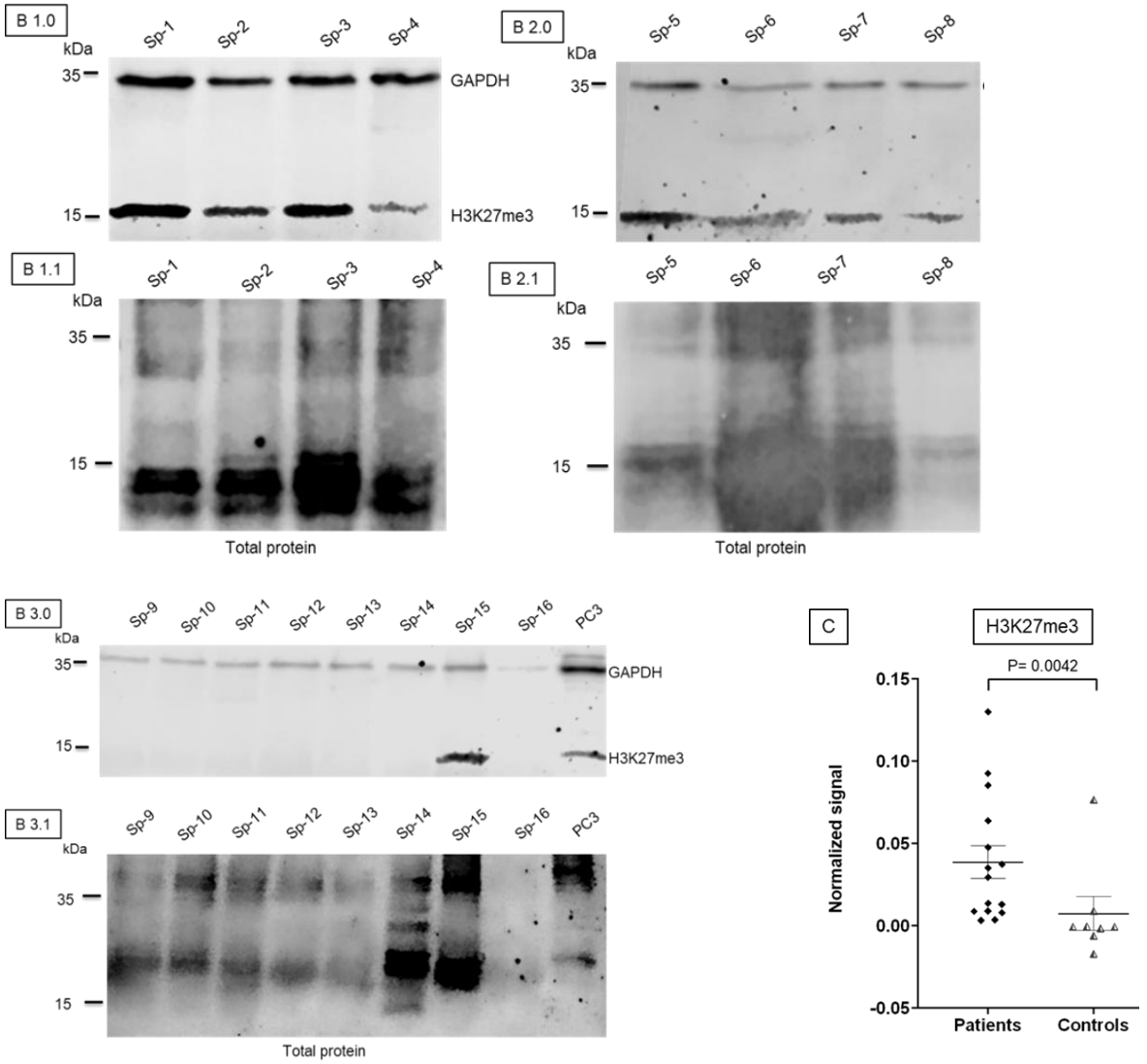


Figure 19: H3K27me3 WB-analysis in motile sperm of fertile men and patients. A1.0) WB results of H3K27me3 (15kDa) with loading control GAPDH (37kDa) in protein lysates of motile sperm of fertile men (S5-12). A1.1) Staining of all proteins after transfer to the same PVDF membrane. B1, B2, B3) WB results of H3K27me3 (15kDa) with loading control GAPDH (37kDa) in protein lysates of motile sperm of subfertile men (Sp1-16). B1.1, B2.1, B3.1) Staining of total proteins after transfer to the same PVDF membrane. S= Sperm protein (fertile); Sp=Sperm protein (patient). C) This figure depicts the normalized signals of H3K27me3 of patients (n=15) and controls (n=8). Results show that patients' sperm retained significantly (p = 0.0042) higher H3K27me3 compared to control men. The results are presented as mean values ± SEM, with P values determined by the Mann–Whitney U-test.

The retention of H3K27me3 was also analysed in immotile sperms (Fig. 20 A1, A2). Results revealed that in immotile sperm, this histone mark retention was significantly higher (p=0.0007) compared to motile sperm. The protein normalization of the target signal against total protein staining and quantification was performed using Empiria Studio 2.2 software from Licor as shown in Fig. 20 A3.

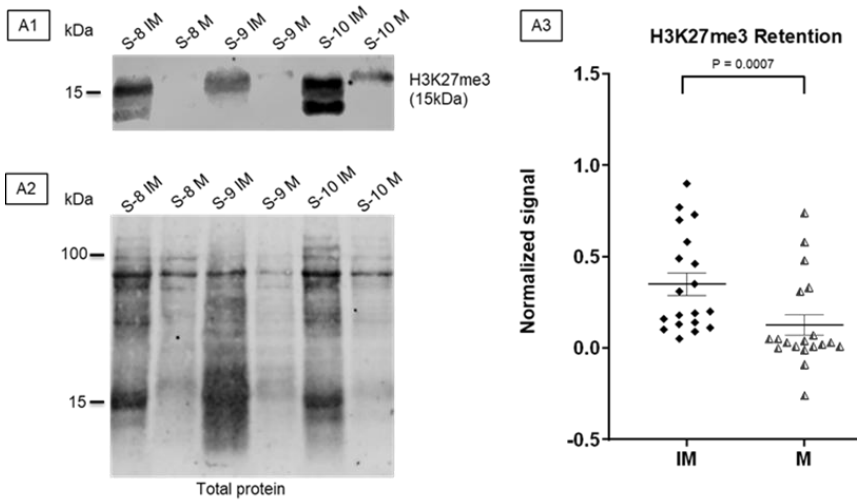


Figure 20: H3K27me3 WB-analysis in motile and immotile sperm of fertile men. The above graphs represent the retention of H3K27me3 in motile (M) and immotile (IM) sperm protein lysates of fertile controls (n=19). A1) WB results of H3K27me3 (15kDa) in protein lysates of sperm of fertile men (S8-10). A2) Staining of all proteins after transfer to the same PVDF membrane while 50µg protein was loaded in each well. A3) Normalized signal showing the immotile group has a significantly higher amount of retention compared to the motile group. The signal intensity of each sample was normalized against the total protein signal of the corresponding sample. The results are presented as mean values ± SEM, with P values determined by the Mann–Whitney U-test.

3.3.3 EED, SUZ12, and EZH2 were not preserved in human motile sperm

The study examined the retention of core PRC2 components in human motile sperm. Results revealed that EED, SUZ12, and EZH2 were not present in the motile sperm of fertile men, but present in the positive controls (Fig. 21 A1, C1). SUZ12 was also not preserved in the motile sperm of subfertile patients (Fig. 21 B1). Unfortunately, the EED and EZH2 antibodies did not function in patients' samples. To validate the absence of PRC2 in human sperm, three different positive controls (PC3, Hela, LnCap cancer cell lines) were used. While total protein staining indicated a low number of proteins in some protein lysates, protein content was better in two (S1-2) motile sperm of fertile men (Fig. 21 A2, C2), where total protein staining demonstrated the presence of proteins, particularly in the expected size of EED and EZH2. To ensure equal loading, GAPDH was used as a loading control (Fig. 21 A1, B1), and it was observed that there were varying amounts of protein presence.

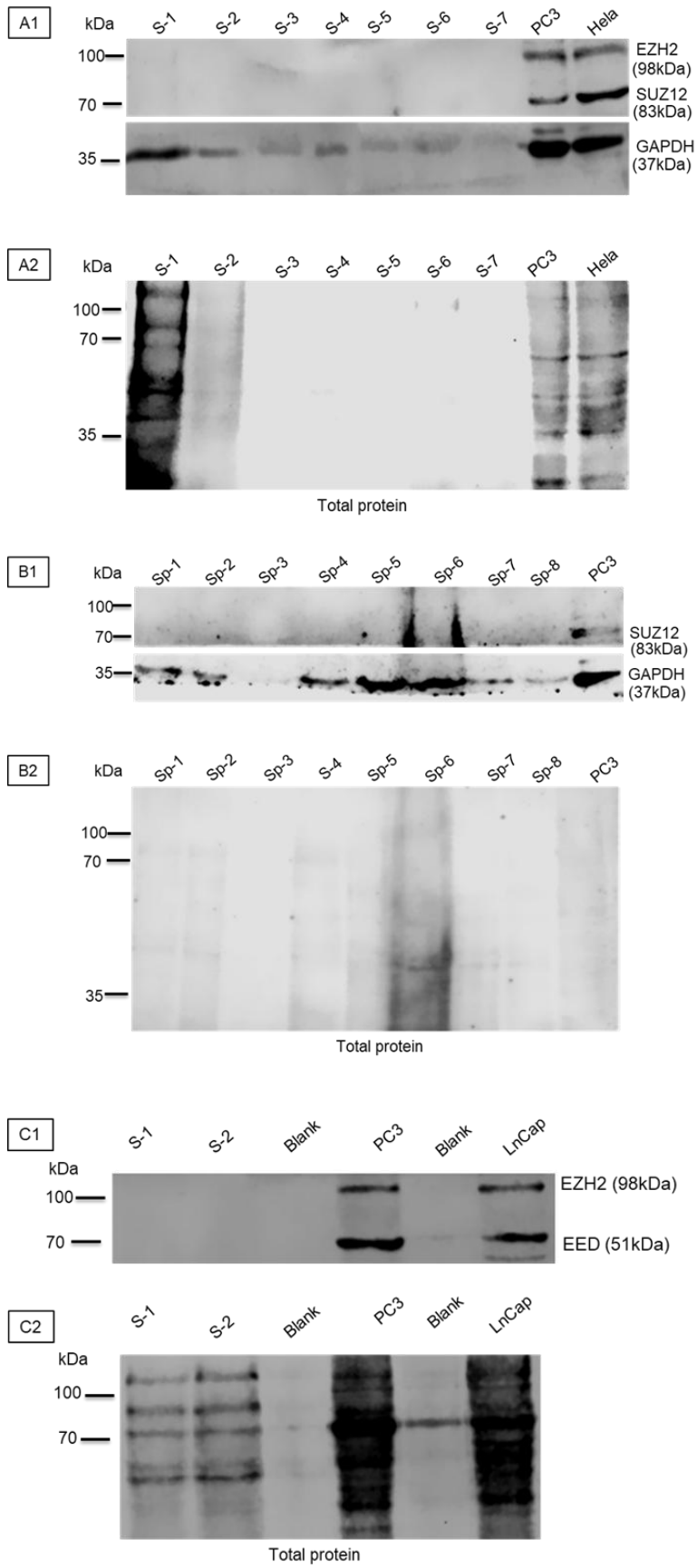


Figure 21: Retention of PRC2 components in human sperm. A1) WB results of SUZ12 (83kDa) and EZH2 (98kDa) with loading control GAPDH (37kDa) in protein lysates of motile sperm of fertile men (S1-7) and positive controls (PC3 and Hela). A2) Staining of all proteins (50µg/well) after transfer to the same PVDF membrane. B1) WB results of SUZ12

(83kDa) with loading control GAPDH (37kDa) in protein lysates of motile sperm of subfertile men (Sp1-8) and positive control (Hela); Sp-6 sample lane showed unspecific staining. B2) Staining of total protein after transfer to the same PVDF membrane. C1) WB results of EED (51kDa) and EZH2 (98kDa) in protein lysates of motile sperm of fertile men (S1-2) and positive controls (PC3 and LnCap). C2) Staining of total proteins after transfer to the same PVDF membrane. S= Sperm protein (fertile); Sp=Sperm protein (patient).

3.4 TET1 and H3K27me3 binding sites in motile sperm of fertile and subfertile men

3.4.1 TET1-NChIP-seq revealed low peak numbers and irregular DNA-binding

The histone fraction was successfully isolated, and the correct isolation of the corresponding nucleosomal fraction was verified following digestion with MNase, DNA preparation, and visualization on an agarose gel (Fig. 22). Subsequently, the soluble chromatin fraction (nucleosomal) was used to perform a ChIP assay with the TET1 antibody. The ChIP-DNA was then isolated and sent to Diagenode for Illumina next-generation sequencing (NGS). The concentrations of the immunoprecipitated and input DNA sent for NGS are detailed in Table 15. After NGS and bioinformatic analysis, it was noted that the TET1 peaks were sparse across all samples, with minimal overlap. The peak calling results showed that when all PCR duplicates and multimapping reads were retained, the highest number of common peaks among samples were three (as shown in Table 16). However, upon removal of PCR duplicates and multimapping reads, no common peak was identified. Due to the failure of the sequencing, Active Motif was requested to isolate sperm chromatin and, if successful, to perform TET1-ChIP. Active Motif carried out TET1 X-ChIP and reported that the chromatin isolated from seventy million motile sperm was insufficient for conducting immunoprecipitation.

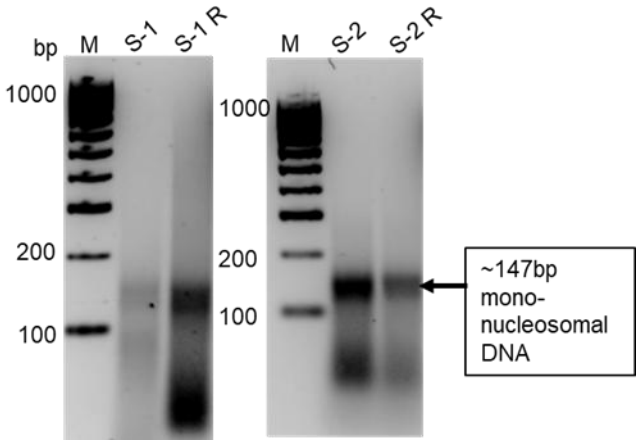


Figure 22: Mono-nucleosomal DNA after Mnase digestion of human motile sperm. A representative agarose gel image of ~147bp mono-nucleosomal DNA fragments produced after Mnase digestion (chromatin fragmentation technique) of fertile men (n=2) sperm chromatin. S= Digested sperm DNA (fertile), R= replicate, M= DNA Marker.

Table 15: Concentrations of input controls and TET1-immunoprecipitated (NChIP) DNA from motile sperm of fertile men

Sample Name	ChIP-DNA (ng/µl)	Input DNA (ng/µl)
S-1 NChIP	1.27	3.96
S-1 R NChIP	1.59	2.1
S-2 NChIP	0.162	2.1
S-2 R NChIP	0.128	0.88

n=2, R=replicates. S= sperm DNA (fertile).

Table 16: Number of detected common peaks among samples in the TET1 ChIP-seq experiment

Sample name	S-1 NChIP	S-1 R NChIP	S-2 NChIP	S-2 R NChIP
S-1 TET1 NChIP	24	12	4	3
S-1 R TET1 NChIP	12	248	9	10
S-2 TET1 NChIP	4	9	227	29
S-2 R TET1 NChIP	3	10	29	131

n=2, R=replicates. S= sperm DNA (fertile).

3.4.2 H3K27me3 DNA-binding sites in sperm of fertile men and subfertile patients

To analyse H3K27me3 enrichment at gene promoters of TFs with developmental relevance in human motile spermatozoa from fertile versus infertile men, H3K27me3-associated chromatin was precipitated using the soluble histone-fraction. qPCR was performed for gene promoters of *SOX2, GATA2, FOXA2, PAX6,* as well as for *HERV* genes. These genes are known to be associated with H3K27me3 in human embryonic stem cells (hESC) [70].

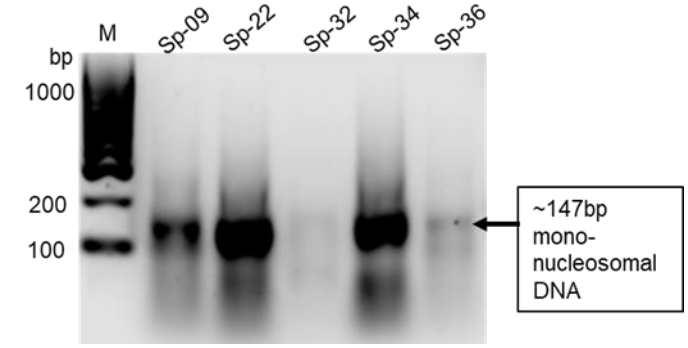


Figure 23: Chromatin fragmentation of human motile sperm and generation of mono-nucleosome. A) A representative agarose gel image of ~147bp mono-nucleosomal DNA fragments produced after Mnase digestion (chromatin fragmentation technique) of subfertile men (n=5) sperm chromatin. Sp= Digested sperm DNA (patient), M= DNA Marker, bp= base pair.

In Fig. 23, a representative image of 147bp mono-nucleosomal DNA band from subfertile patients is shown on agarose gel. In total, H3K27me3-NChIP was performed in five patients and five controls. NChIP- and input-DNA concentrations of patients and controls are presented in Tables 17 and 18 respectively. After qPCR, the relative enrichment was calculated by the per cent input method according to Allan et al., 2021 [67]. The percent input value indicates the proportion of the initial chromatin (input) that was immunoprecipitated in the ChIP experiment. This calculation is derived by comparing the cycle threshold (Ct) values obtained from both the ChIP sample and the input sample for a specific gene. The input sample reflects the quantity of chromatin utilized in the ChIP procedure.

Table 17: Input controls and H3K27me3-immunoprecipitated (NChIP) DNA concentrations from motile sperm of subfertile men

Sample description	Sample	DNA conc. (ng/µl)
P09-Input	human sperm	41.50
P22-Input	human sperm	58.80
P32-Input	human sperm	4.77
P34-Input	human sperm	34.60
P36-Input	human sperm	8.41
P09-H3K27me3-NChIP	human sperm	5.66
P22-H3K27me3-NChIP	human sperm	14.00
P32-H3K27me3-NChIP	human sperm	15.65
P34-H3K27me3-NChIP	human sperm	10.55
P36-H3K27me3-NChIP	human sperm	3.68

n=5, P= patient, Conc= concentration

Table 18: Input controls and H3K27me3-immunoprecipitated (NChIP) DNA concentrations from motile sperm of fertile men

Sample description	Sample	DNA conc. (ng/µl)
D112-Input	human sperm	3.20
D118-Input	human sperm	2.42
D119-Input	human sperm	5.98
D117-Input	human sperm	1.89
D121-Input	human sperm	4.46
D122-Input	human sperm	2.22
D124-Input	human sperm	2.68
D112-H3K27me3-NChIP	human sperm	3.09
D118-H3K27me3-NChIP	human sperm	2.72
D119-H3K27me3-NChIP	human sperm	1.01
D117-H3K27me3-NChIP	human sperm	0.40
D121-H3K27me3-NChIP	human sperm	0.26
D122-H3K27me3-NChIP	human sperm	0.15
D124-H3K27me3-NChIP	human sperm	0.30

n=7, D= fertile men (controls), Conc= concentration.

Fig. 24 displays the relative enrichment of H3K27me3 in the gene promoter region for both patients and fertile controls. The Y-axis shows the relative enrichment normalized to input controls, while the X-axis represents the target protein of interest with its corresponding bead (no antibody) control.

The ChIP-qPCR analysis results showed that the binding of H3K27me3 to the *SOX2* promoter region was lower in subfertile patients compared to control men, although no statistical differences were observed in the Mann-Whitney test (Fig. 24A). Similarly, the *PAX6* promoter exhibited reduced H3K27me3 binding in subfertile patients compared to controls (Fig. 24B), but again, no statistical differences were found. Figures 24C and 24D present the relative enrichment of H3K27me3 in the promoter regions of *GATA2* and *FOXA2* genes, respectively. In both cases, there was a lower binding affinity of H3K27me3 in subfertile patients compared to controls, although statistical differences were not detected by the Mann-Whitney test. Conversely, the retrotransposon *HERV*-K showed an opposite trend, as illustrated in Fig. 24E. However, like the other genes, the Mann-Whitney test did not reveal any statistically significant differences between the patient and control cohort for *HERV*-K.

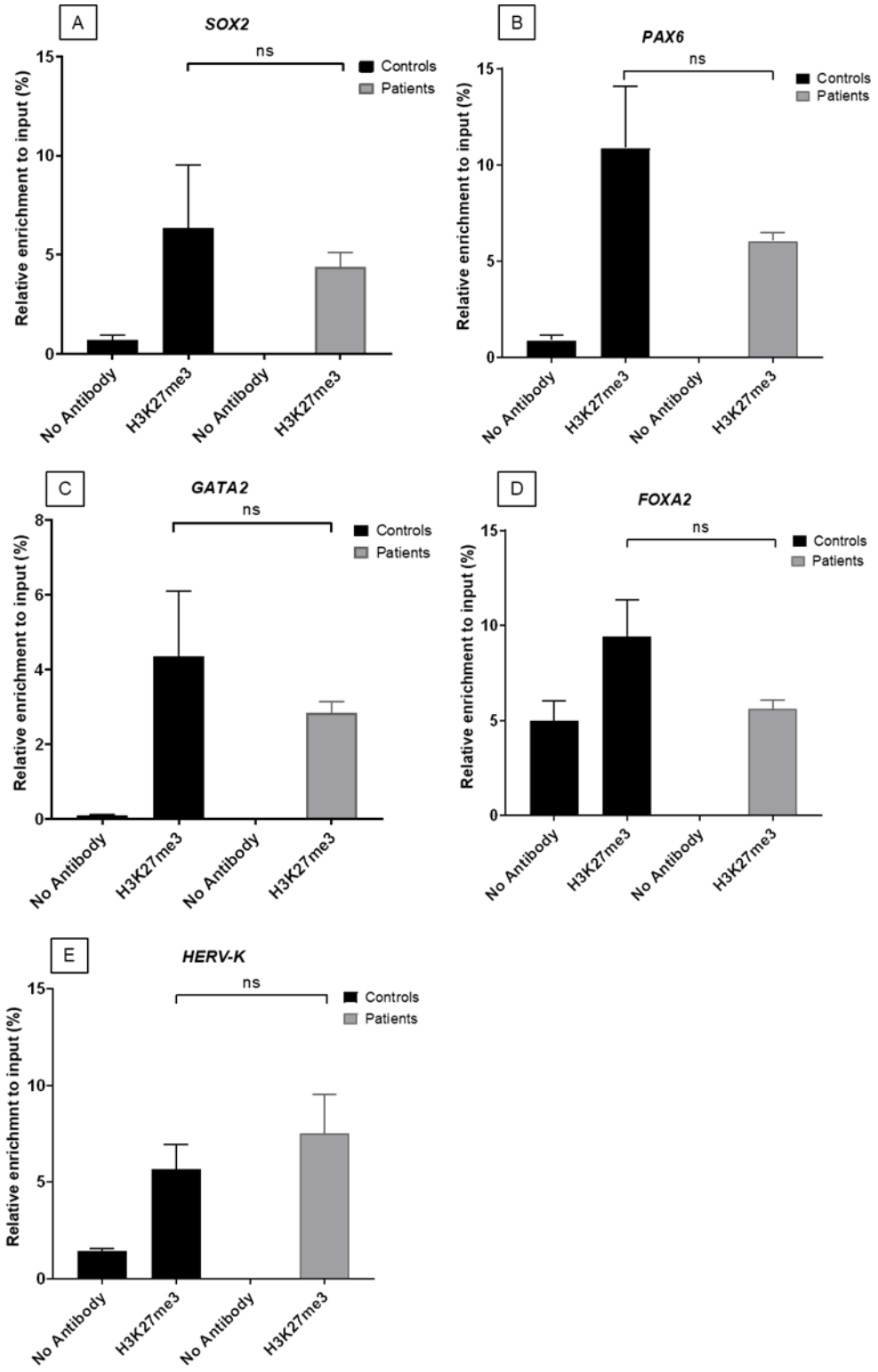


Figure 24: Validation of enriched H3K27me3 bound promoters by ChIP-qPCR in motile sperm of fertile controls and subfertile men (patients); In the X axis, a chipped protein with no antibody control (bead control) from controls and patients is represented in two colours. In the Y axis, per cent input refers to the proportion of the initial chromatin (input) immunoprecipitated in the ChIP experiment. A) Mann Whitney test shows no statistical difference (ns) between controls' (n=6) and patients'(n=5) enrichment for H3K27me3 in the *SOX2* promoter region. B) Enrichment of H3K27me3 in *PAX6* gene promoters in controls (n=5) and patients (n=4) reveals no statistically significant difference (ns) is observed between them. C) Evaluation of enrichment of H3K27me3 in *GATA2* gene promoters in controls (n=6) and patients (n=4) reveals no

statistically significant difference (ns). D) Evaluation of enrichment of H3K27me3 in *FOXA2* gene promoters in controls (n=3) and patients (n=4) reveals no statistically significant difference (ns) between controls and patients is observed. E) Evaluation of enrichment of H3K27me3 in *HERV-K* gene in controls (n=4) and patients (n=4) reveals no statistically significant difference (ns) between them. The results are presented as mean values ± SEM, with P values determined by the Mann–Whitney U-test. ns= not significant (P>0.05).

3.5 3.5 mRNA analysis of *TET1-3, EED, SUZ12, EZH2* and H3K27me3-binding genes in motile sperm of fertile and subfertile men

To measure the mRNA levels of *TETs*, core PRC2 components, and H3K27me3-binding genes, qPCR analysis was performed. The term "mRNA level" is used, as sperm cells are known to be transcriptionally inactive [71]. Following normalization to the GAPDH housekeeping gene mRNA levels, the mRNA levels of corresponding target genes were calculated.

3.5.1 *TET1* and *TET3* were significantly upregulated in the sperm of subfertile patients

The relative mRNA levels of *TET1-3* were measured in motile human spermatozoa from both patients and controls using RT-qPCR with primers from published literature [53]. The resulting data are presented in Fig. 25 as a scattered plot, where the X-axis represents the number of patients and controls, and the Y-axis shows the relative mRNA levels. The findings revealed that the *TET1* mRNA levels were significantly upregulated (P=0.024) in patients' spermatozoa compared to controls (Fig. 25A). *TET2* transcripts, on the other hand, exhibited an opposite trend, while no statistically significant difference was observed between the transcripts of patients and controls (Fig. 25B). For *TET3* transcripts, the upregulation was found to be highly significant (P<0.0001) in patients compared to controls (Fig. 25C).

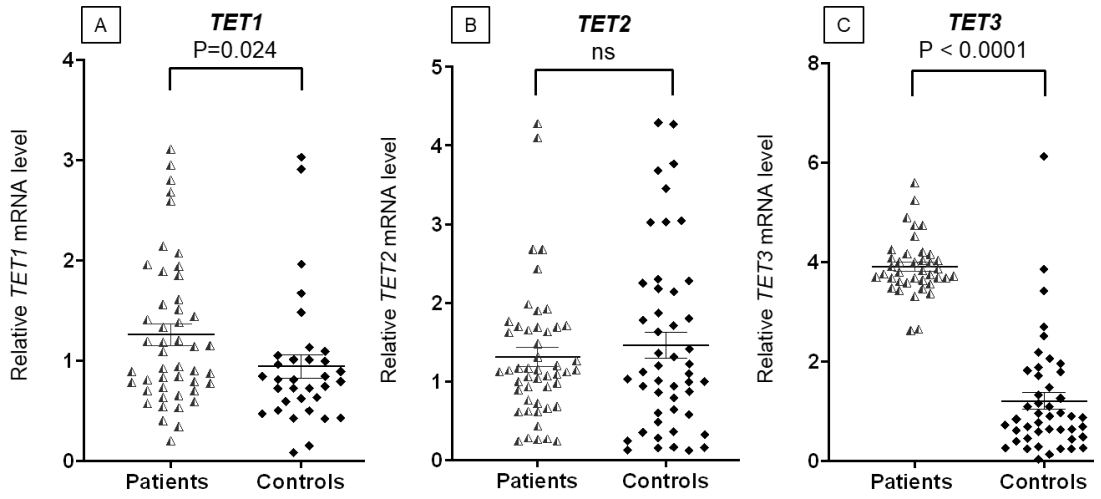


Figure 25: *TET1-3* relative mRNA levels in motile sperm of fertile controls and subfertile men (patients). A) *TET1* transcript levels were significantly (P=0.024) upregulated in patients (n=47) compared to controls (n=33). B) *TET2* transcript levels were not significantly (ns) different between patients (n=49) and controls (n=48) group C) *TET3* relative mRNA levels were significantly (p<0.0001) upregulated in patients (n=39) compared to controls (n=45). The results are presented as mean values ± SEM, with P values determined by the Mann–Whitney U-test. ns= not significant (P>0.05). ΔΔCT=Method of CT values calculation.

3.5.2 *EED, SUZ12* and *EZH2* in human motile sperm

In this study, we aimed to investigate the relative mRNA levels of core PRC2 components, including *EED, EZH2,* and *SUZ12*, in motile human spermatozoa obtained from subfertile patients and fertile controls. RT-qPCR was employed to quantify the mRNA levels, using primers sourced from published literature or designed via primer blast. The analysis of the qPCR results revealed significant differences in the mRNA levels of EED between the patient group and the control group (p=0.0001). Specifically, patients exhibited significantly lower levels of EED mRNA compared to controls, as depicted in Fig. 26A. Conversely, in the case of EZH2, there was no statistically significant difference observed in mRNA levels between patients and controls. However, it is worth noting that controls tended to exhibit slightly higher levels of *EZH2* mRNA, as illustrated in Fig. 26B. In contrast to the observations for *EED* and *EZH2*, the mRNA levels of *SUZ12* demonstrated an opposing trend. Patients displayed higher levels of *SUZ12* mRNA compared to controls, although this difference did not reach statistical significance. This trend is depicted in the corresponding data presented in Fig. 26C.

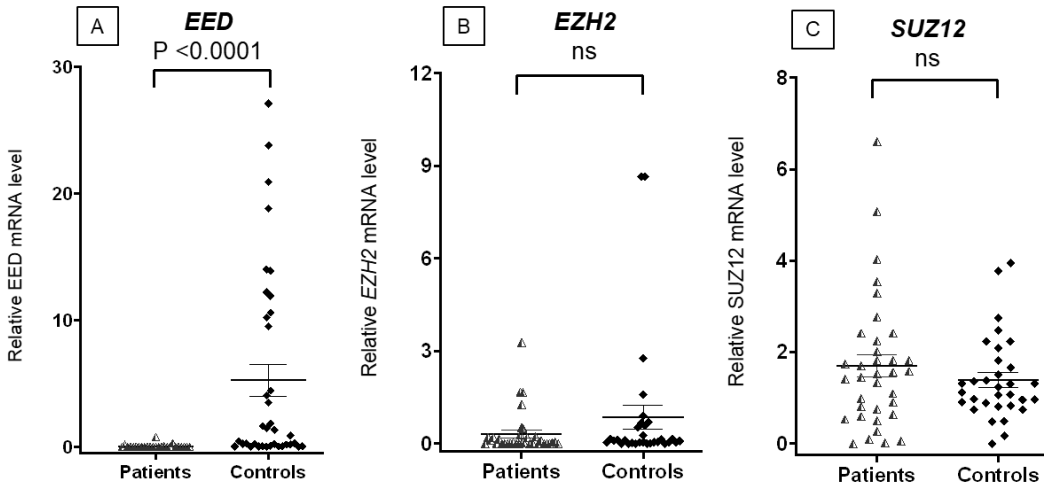


Figure 26: The above graphs represent, PRC2 components *EED, EZH2, and SUZ12* relative mRNA levels in motile sperm of fertile controls and subfertile men (patients). A) *EED* transcript levels were significantly (P<0.0001) upregulated in controls (n=37) compared to patients (n=26). B) *EZH2* relative mRNA levels were not significantly (ns) different between patients (n=34) and controls (n=31) group C) *SUZ12* transcript levels were also not significantly different in patients (n=35) and controls (n=32). The results are presented as mean values ± SEM, with P values determined by the Mann–Whitney U-test. ns= not significant (P>0.05). ΔΔCT=Method of CT values calculation.

3.5.3 *SOX1, SOX2, OCT4, NANOG, PAX6, GATA2, FOXA2,* and *HERV* in human motile sperm

In this study, RT-qPCR was used to measure the mRNA levels of eight H3K27me3 binding genes in motile human spermatozoa from subfertile patients and fertile men. Among the eight genes, *FOXA2*, and *HERV-w* mRNA levels were found to be downregulated in patients, while the mRNA levels of *OCT4, SOX1,* and *SOX2* were upregulated in patients. The difference in mRNA levels between the patient and control groups was statistically significant for most of the genes, except for *GATA2, NANOG,* and *PAX6*. Of particular interest were the results for *SOX1* and *SOX2*, as their transcript levels were significantly upregulated in the spermatozoa of subfertile patients compared to fertile controls (P=0.0126 and P=0.0001, respectively), as shown in Fig. 27A-B.

For both *FOXA2* and *GATA2* genes, the RT-qPCR analysis of human motile sperm revealed minimal transcript levels. However, a significant difference (p = 0.0042) was observed in the *FOXA2* mRNA levels between the fertile control group and subfertile patients, with the former showing higher levels (Fig. 28A). In contrast, the *GATA2* mRNA levels did not show a statistically significant difference between the two groups (Fig. 28B).

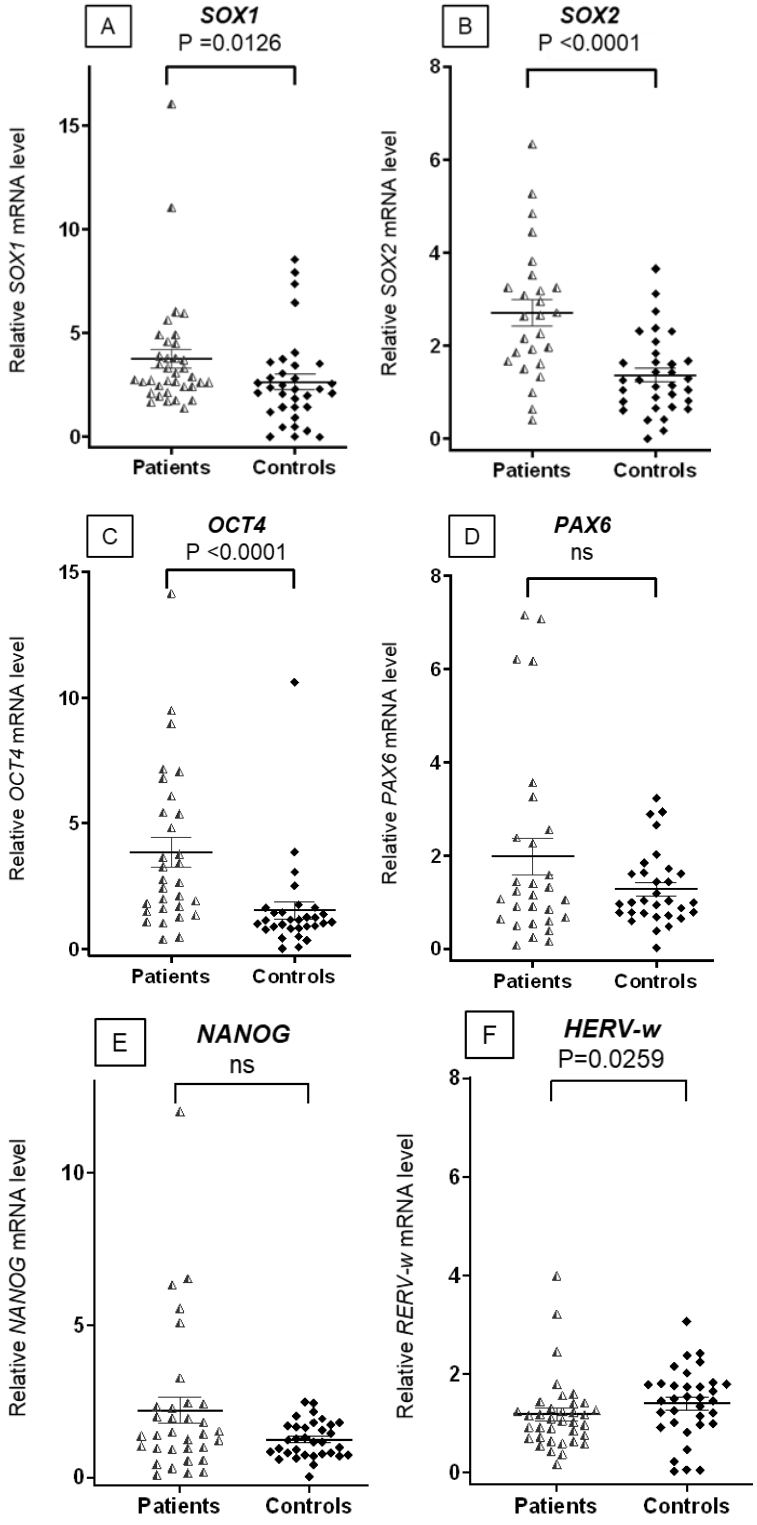


Figure 27: **The above graphs show relative mRNA levels of *SOX1 SOX2, OCT4 and PAX6, NANOG,* and *HERV-w* genes in motile sperm of fertile controls and subfertile men (patients).** A) *SOX1* transcript levels were significantly (P=0.0126) upregulated in patients (n=36) compared to controls (n=34). B) *SOX2* transcript levels were very significantly (P<0.0001) upregulated in patients (n=26) compared to controls (n=33). C) *OCT4* transcript levels were significantly (P<0.0001) upregulated in patients (n=30) compared to controls (n=30). D) *PAX6* mRNA levels were higher in the patient's cohort but no significant (ns) difference was observed between patients (n=29) and controls (n=31). E) *NANOG* transcript levels were not significantly different when patients (n=33) and controls (n=33) were compared. F) *HERV-w* mRNA levels were significantly (P=0.0259) downregulated in patients (n=36) compared to controls (n=34). The results are presented as

mean values ± SEM, with P values determined by the Mann–Whitney U-test. ns= not significant (P>0.05). ΔΔCT=Method of CT values calculation.

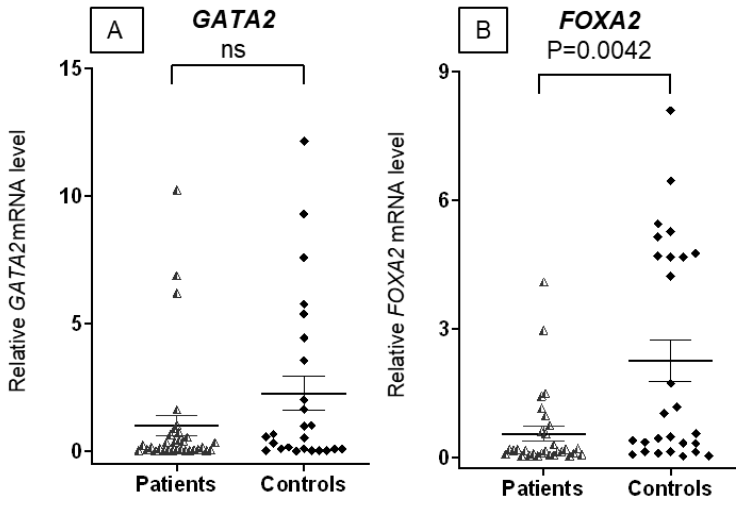


Figure 28: *FOXA2, GATA2* relative mRNA levels in motile sperm. A) *GATA2* relative mRNA levels were not significantly (ns) different between patients (n=32) and controls (n=25). B) *FOXA2* transcript levels were significantly (P=0.0042) decreased in patients (n=30) compared to controls (n=27). The results are presented as mean values ± SEM, with P values determined by the Mann–Whitney U-test. ns= not significant (P>0.05). ΔΔCT=Method of CT values calculation.

3.6 Methylation of *LINE1, Alu, EED, SUZ12, EZH2* and H3K27me3-binding genes in motile sperm of fertile and subfertile men

3.6.1 Methylation levels of *LINE1* and *Alu* in human motile sperm

The methylation status of the *LINE1* repetitive element was investigated in motile sperm of fertile men using COBRA. Firstly, the sperm DNA was treated with bisulphite and PCR was carried out using specific primers as described in the method. The amplified PCR products (413bp) were digested with the *HinfI* restriction enzyme to determine the methylation status of the *LINE1* in each fertile man. *HinfI* enzyme recognizes the site (5´ G$^{\downarrow}$ANTC 3´.....3´ CTNA$^{\uparrow}$G 5´) in the PCR product and digests only if the site is methylated. Here in the restriction site, N can be Adenine / Guanine / Cytosine / Thymine. The agarose gel image in Fig. 29 shows the undigested or digested PCR product in the '+' lane. The digested small fragments indicate methylated DNA, while the undigested fragment at around 418bp could be either unmethylated or mutated restriction sites. The PCR products in the '-' lanes were treated with mock enzyme. Through densitometric analysis of the PCR bands, the amounts of methylation were calculated. The results showed that five out of ten fertile men exhibited heterogeneous methylation levels at the *LINE1* gene, with methylation levels less than 30% (as shown in Table 19).

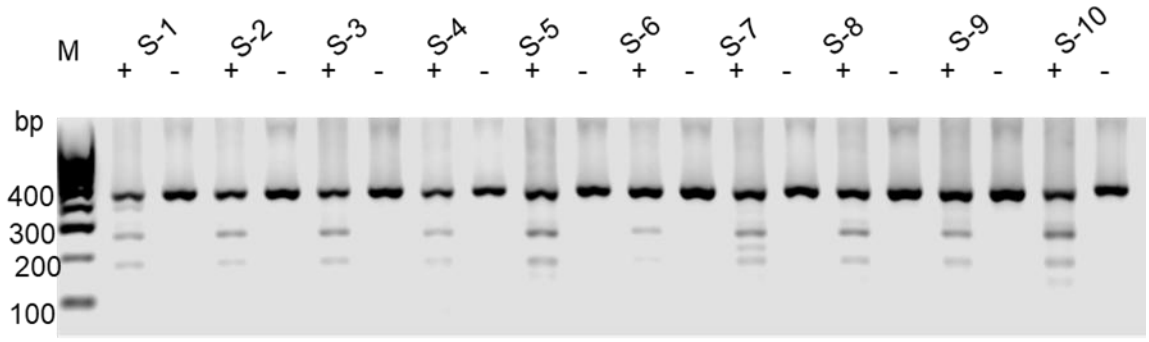


Figure 29: CpG methylation levels of *LINE1* in motile sperm of fertile men. The PCR product (413bp) for COBRA at *LINE1* was treated with a restriction enzyme, *HinfI* that only cuts repetitive elements that were originally methylated. The digested PCR product was separated by agarose gel electrophoresis and stained with GelRed nucleic acid stain. "+" means digestion with restriction enzyme. "–" means mock digestion. M= Marker. S (1-10): numbers of fertile men.

Table 19: Percentage of methylation of *LINE1* gene after quantification of agarose gel image.

Name	Methylation (%)
S1+	51.43
S2+	41.11
S3+	48.77
S4+	21.08
S5+	45.31
S6+	12.89
S7+	28.13
S8+	29.35
S9+	22.93
S10+	65.92

In this study, the methylation status of the *Alu* repetitive element was also analysed using COBRA. To perform the analysis, sperm DNA was subjected to bisulphite treatment, followed by PCR amplification with a specific primer set (mentioned in the method section). The resulting PCR product (152bp) and the '*MboI'* restriction enzyme were used to determine the methylation status of the *Alu* gene for each fertile man. This enzyme recognizes the 5´ ↓GATC 3´...3´ CTAG↑ 5´ site in the PCR product and digests it only if the site is methylated. The agarose gel image in Fig. 30 shows the undigested PCR product in the '+' lane revealing that there was no detectable methylation in these samples or the restriction sites have been mutated. On the other hand, the '-' lanes show the PCR products with mock enzyme treatment of the corresponding samples.

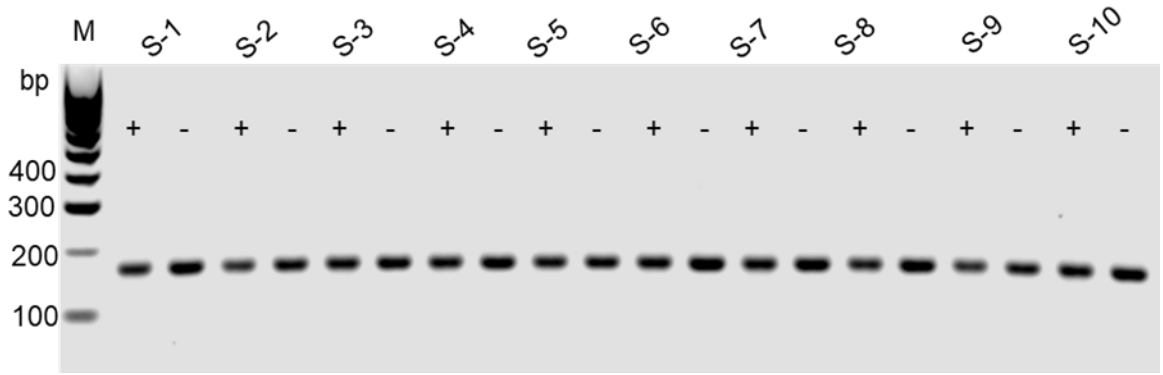


Figure 30: CpG methylation levels of *Alu* in motile sperm of fertile men. The PCR product (152bp) for *Alu* was treated with restriction enzyme *MboI* that only cuts repetitive elements that were originally methylated. The undigested bands represent Alu elements were unmethylated or the restriction site had been mutated in the sperm of fertile men. The treated PCR product was separated by agarose gel electrophoresis and stained with GelRed nucleic acid stain. "+" means digestion with restriction enzyme. "–" means mock digestion. M= Marker. S (1-10): numbers of fertile men.

3.6.2 Methylation levels of *EZH2, SUZ12, and EED* in human motile sperm

The total DNA was extracted from human motile sperm and subjected to bisulphite conversion to analyse gene promoter methylation levels of *EZH2, SUZ12,* and *EED* using Pyrosequencing. For the *EED* gene, the promoter methylation levels were also very low. However, there was no significant difference in methylation levels between patients and control groups (Fig. 31A). Fig. 31B indicates that the DNA methylation levels of the *EZH2* gene promoter were very low in both patients and controls, with an average of around one per cent. Interestingly, the methylation levels in the control group were significantly higher (p=0.0001) than in the patient cohort. For the *SUZ12* gene, the promoter methylation levels were significantly higher (p=0.002) in patients than in control men, with an average methylation level of about one per cent (Fig. 31C).

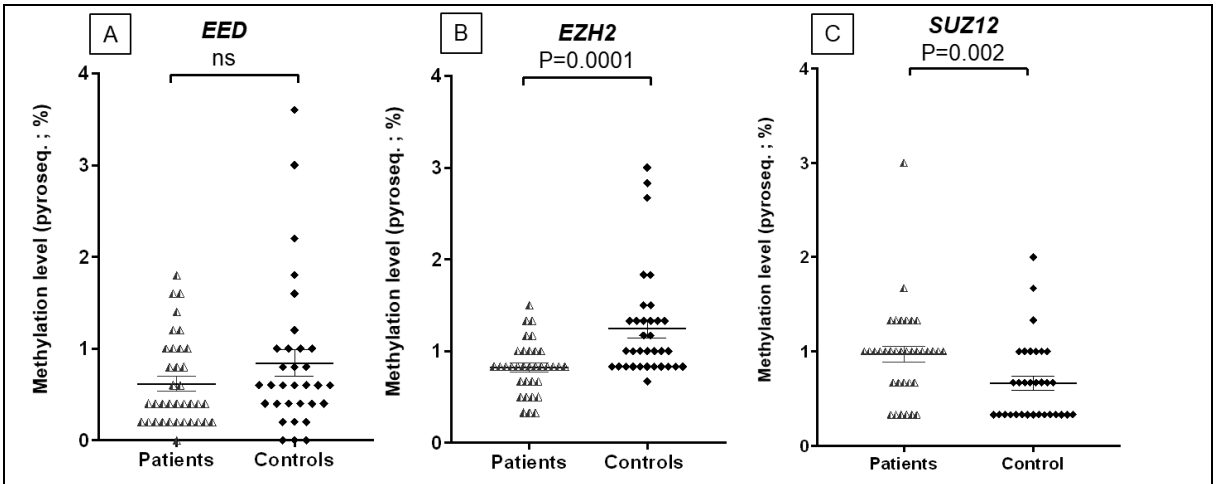


Figure 31: CpG promoter methylation levels of *EED, EZH2, SUZ12* in motile sperm. The figure depicts the CpG promoter methylation levels of PRC2 components *EED, EZH2, and SUZ12* in motile sperm of fertile controls and subfertile men (patients), as analysed by pyrosequencing. The Y-axis denotes the percentage of methylation, while the X-axis indicates the number of patients and controls. A) *EED* methylation levels: Higher in control (n=32) but not significantly different compared to patients (n=36). B) *EZH2* methylation levels: Significantly higher in controls (n=34) compared to patients (n=34) with a P-value of 0.0001. C) *SUZ12* methylation levels: Significantly higher in patients (n=35) compared to controls (n=32) with a P-value of 0.002. The results are presented as mean values ± SEM, with P values determined by the Mann–Whitney U-test. ns= not significant (P>0.05).

3.6.3 Methylation levels of *GATA2, PAX6, FOXA2, and SOX2* in human motile sperm.

This study examined the promoter methylation levels of four H3K27me3 binding genes, *GATA2, PAX6, FOXA2,* and *SOX2,* in motile sperm from subfertile patients and fertile men using pyrosequencing. The result revealed that the *GATA2* gene promoter exhibited significantly higher levels of methylation in the patient group compared to the control group (p=0.0006), despite the low percentage of methylation ranging from 0.12% to 4% (as illustrated in Fig. 32A). The methylation levels of the *PAX6* gene were also relatively low, ranging from 0.3% to 2.4%, and there was no statistically significant difference observed between the patient and control groups (as depicted in Fig. 32B). For *FOXA2* and *SOX2* genes, the study showed that the methylation levels of the promoters of both genes were similar, with an average of approximately 2.5% in both the patient and control groups. The range of methylation levels for these genes fell within 0.1% to 10%. The statistical analysis indicated no significant difference in promoter methylation levels between the two groups for either of these genes, as illustrated in Fig. 32C-D.

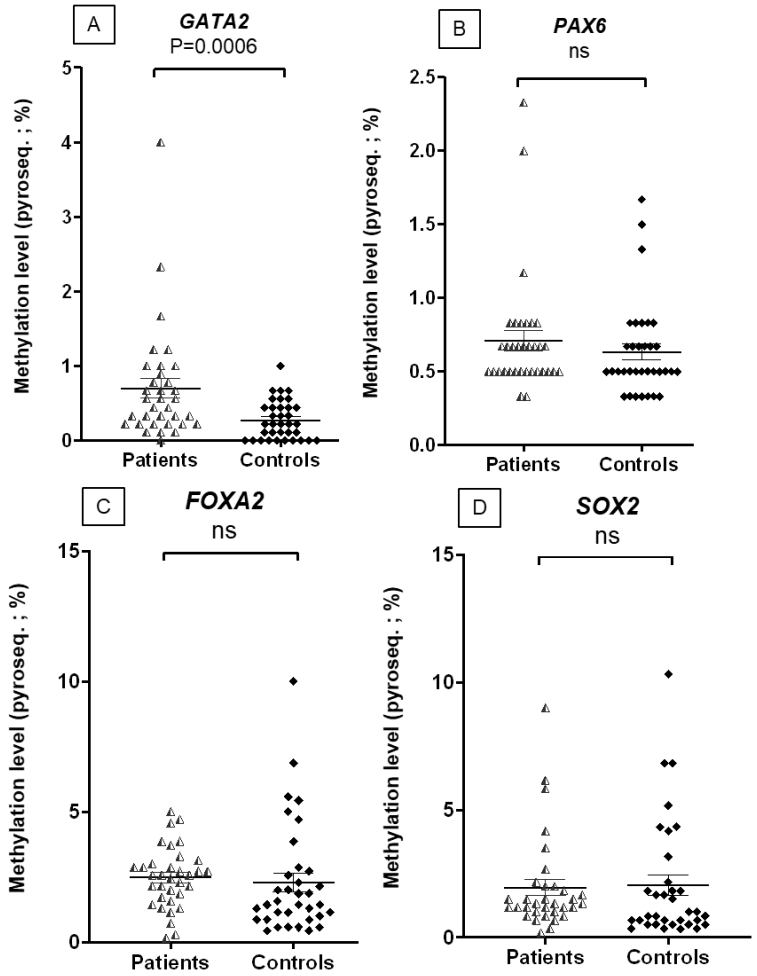


Figure 32: **Methylation levels of *GATA2, PAX6, FOXA2, and SOX2* promoters in motile sperm of fertile and subfertile men (patients).** The pyrosequencing data presented in the graph illustrates the methylation levels of *GATA2* and *PAX6* promoters in motile sperm. A) *GATA2* methylation levels were found to be significantly higher in patients (n=34) compared to controls (n=35), with a P value of 0.0006. B) *PAX6* methylation levels were observed to be higher in patients (n=34) but did not exhibit a statistically significant difference compared to controls (n=34). C) The methylation levels of *FOXA2* in patients (n=34) and controls (n=35) showed no significant difference. D) The methylation levels of *SOX2* in patients (n=33) and controls (n=34) showed no significant difference. The results are expressed as mean values ± SEM, with P values derived from the Mann–Whitney U-test. ns= not significant (P>0.05).

3.6.4 Correlation analysis

This section of the thesis presents significantly correlated most pertinent data, obtained from the correlation analysis of mRNA levels, and methylation levels across all candidate genes, as well as with the clinical data. Results presented in Fig. 33 represent correlation analysis results of *TET1* with *PAX6, and SUZ12* mRNA levels as well as with age. The correlation analysis between *SUZ12* mRNA and *TET1* mRNA revealed a significant positive correlation (r=0.419, P=0.0079). Similarly, the correlation analysis between *PAX6* mRNA and *TET1* mRNA also demonstrated a significant positive correlation (r=0.534, P=0.0008). Data also suggests that *TET1* mRNA increases with the progression of age. A negative correlation (r=-0.385, P=0.016)

was observed between *SOX2* methylation levels and *TET3* mRNA levels. Additionally, investigating the relationship between *SUZ12* mRNA and *OCT4, PAX6, and NANOG* mRNA, we found significant positive correlations (Fig. 33 and S1).

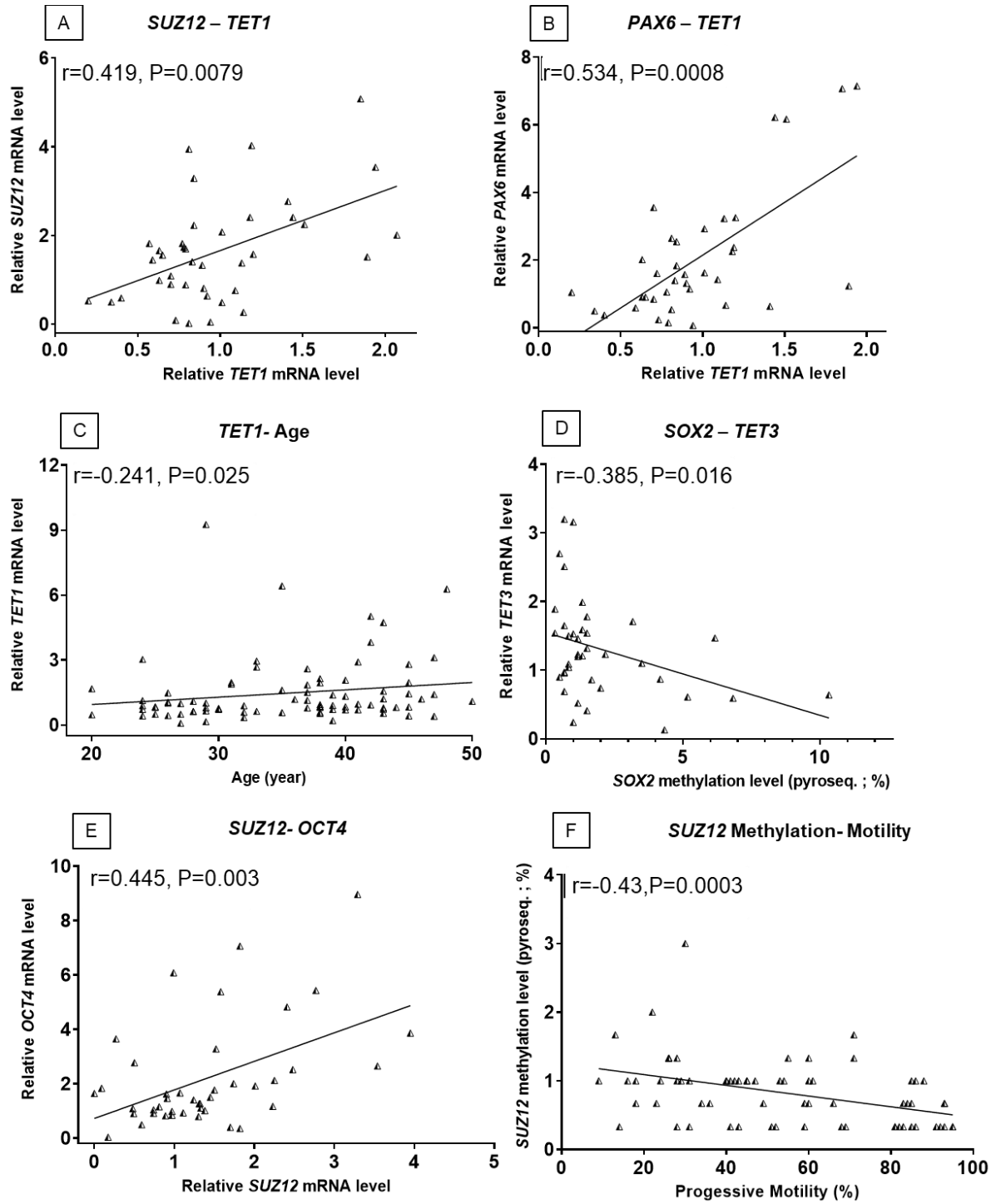


Figure 33: **Correlation analysis revealed multiple correlations**. A) Correlation analysis of *SUZ12* mRNA with *TET1* mRNA shows a significant (P=0.0079) positive correlation (r=0.419). B) Correlation analysis of *PAX6* mRNA with *TET1* mRNA shows a significant (P=0.0008) positive correlation(r=0.534). C) Correlation analysis of *TET1* mRNA with age shows a significant (P=0.025) weak positive correlation (r=0.241) with each other. D) A negative correlation (r=-0.385, P=0.016) was observed between *SOX2* methylation levels and *TET3* mRNA levels. E) Correlation analysis of *SUZ12* mRNA with *OCT4* mRNA shows a significant (P=0.003) positive correlation (r=0.445). F) *SUZ12* promoter methylation levels: Significantly negatively correlated with PM (r = -0.43) with a P-value of 0.0003. The results are presented as mean values ± SEM. The r and P-values were obtained from Pearson correlation analysis.

Fig. 34 represents results from the correlation analysis among mRNA levels and methylation levels of *FOXA2, and TET1* along with their correlation with age. A negative correlation (r = -0.438, P = 0.024) was observed between *FOXA2* promoter methylation levels and *FOXA2* mRNA levels. When methylation levels decrease, mRNA levels tend to increase. A positive correlation (r = 0.380, P = 0.0014) was found between *FOXA2* promoter methylation and age progression. Additionally, there was a significant negative correlation (r = -0.545, P = 0.001) between *FOXA2* mRNA and *TET1* mRNA. *FOXA2* mRNA levels exhibited significant inverse correlations with age (r = -0.343, P = 0.02). Additionally, Fig. 35 shows, a significant inverse correlation between *GATA2* and *TET1, TET3* mRNA levels.

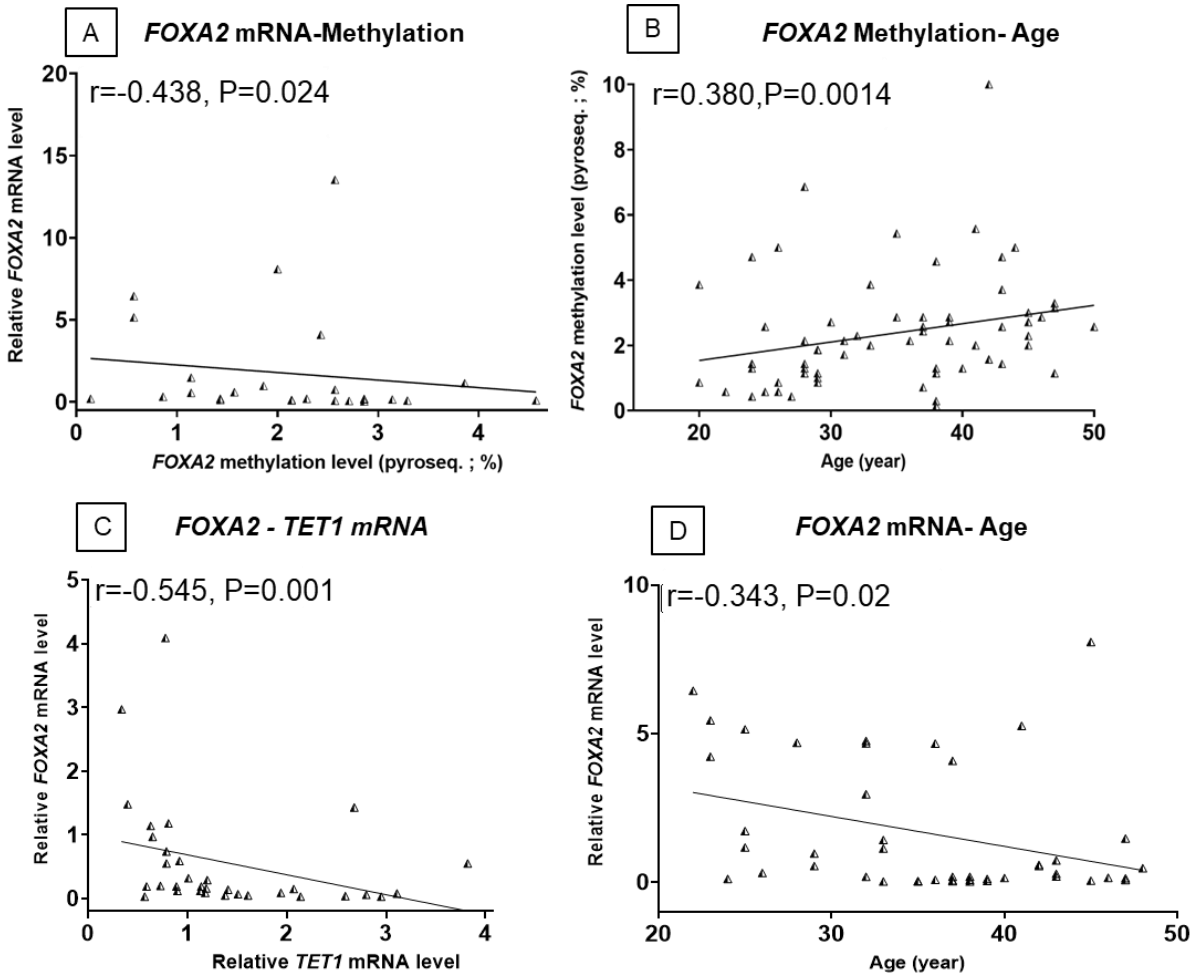


Figure 34: **Correlation analysis among mRNA levels and methylation levels of *FOXA2, TET1*; their correlation with age.** A) A negative correlation (r = - 0.438, P = 0.024) was observed between *FOXA2* promoter methylation levels and *FOXA2* mRNA levels. B) A positive correlation (r = 0.380, P = 0.0014) was found between *FOXA2* promoter methylation and age progression. C) Correlation analysis of *FOXA2* mRNA with *TET1* mRNA shows a significant (P=0.001) negative correlation (r=-0.545). D) Correlation analysis of *FOXA2* mRNA with age shows negative significant correlations (r=-0.343, P=0.02). The results are presented as mean values ± SEM. The r and P-values were obtained from Pearson correlation analysis.

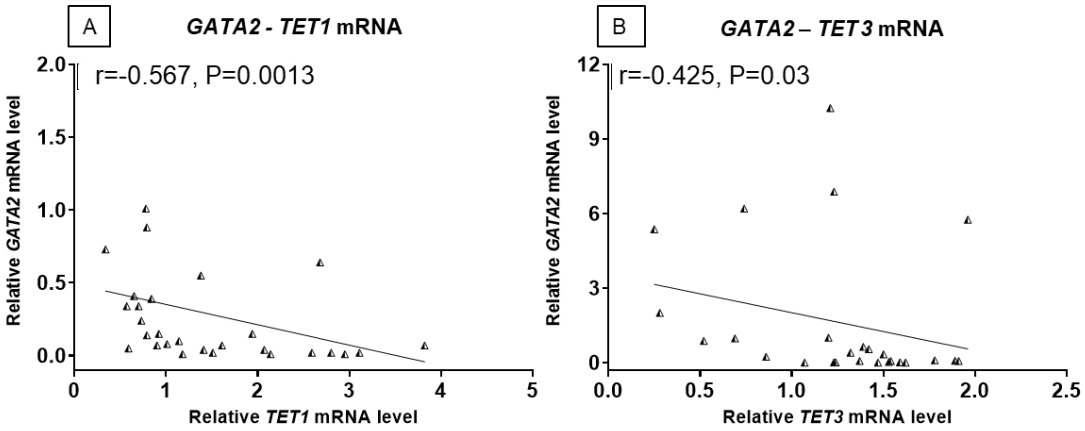


Figure 35: **Correlation analysis between mRNA levels of *GATA2 and TET1, TET3.*** A-B) Correlation analysis of *GATA2* mRNA levels: negatively correlated with *TET1, and TET3* (r = -0.567, -0.425) with P-values of 0.0013, 0.03 respectively. The results are presented as mean values ± SEM. The r and P-values were obtained from Pearson correlation analysis.

4. Discussion

Epigenetic factors TET1-3, PRC2 and H3K27me3, have been identified to have altered regulation in multiple diseases, one of which is infertility [52,59,64,72]. To gain a deeper understanding of the interdependency of these factors during spermatogenesis, first, the expression of these three factors during spermatogenesis was investigated using human and mouse testis samples. For mice, wild type (WT) and mutants (*Tet1* $^{(+/-)}$ and *Tet1* $^{(-/-)}$) samples were used. Furthermore, it was analysed if TET1, H3K27me3 and PRC2 components EED, SUZ12, EZH and H3K27me3 are differentially preserved at the protein level in mature spermatozoa of fertile versus infertile men. This experiment was complemented with ChIP experiments for TET1 and H3K27me3 in human sperm to analyse their differential binding them to the gene promoters of transcription factors (TFs) with developmental relevance. Moreover, comprehensive CpG-promoter methylation analyses and mRNA quantification experiments were performed in big cohorts of fertile versus infertile men for several genes encoding TFs with developmental relevance to evaluate the impact of any plausible epigenetic aberrations on the establishment of male infertility.

Recent study has discovered that RNA molecules present in sperm not only play a pivotal role in embryonic development but are also associated with fertility [30]. Sperm is known to carry various components, including histone proteins, RNAs, and DNA methylation signatures, all of which can be transmitted to the embryo [59]. These transferred elements, known as epigenetic marks, can have a significant impact on the embryo. For instance, the RNAs that are transmitted to the oocyte can be translated by the transcription machinery of embryo, or they can directly regulate gene expression in the embryo at either the transcriptional or post-transcriptional level [73]. Therefore, any aberrations in these elements, such as changes in methylation or mRNA levels in sperm, can potentially affect fertility. The full impact of these aberrations is most evident during the stages of embryo development and pregnancy.

4.1.0 Expression pattern of TET1 during spermatogenesis and confirmation of *Tet1* knock-out in a mouse model

Studies showed that 5-hmC dynamically changes during human spermatogenesis showing the role of TETs during this process [74]. However, the specific role of TET1-mediated DNA demethylation in sperm from idiopathic infertile patients remains less explored.

In the present study, in human testis, weaker TET1 protein expression (assessed based on IHC signal intensities) was observed from spermatogonia B to leptotene spermatocyte whereas a stronger expression signal was found in round spermatids. This result suggests an active DNA demethylation process might be occurring during these stages. In the previous study of our group, Ni et al.[53] showed that TET1 expression was also detectable in round spermatids, this aligns with the current finding. Contrary to the present observations in humans, TET1 expression was also detected in elongating spermatids in their study. Interestingly, the same expression pattern was also noted in the present study in murine-elongated spermatids. The discrepancy in detection could potentially be attributed to the utilization of a different tissue fixation method which might have facilitated enhanced antibody penetration into the densely compacted elongating spermatids. Apart from the germ cells, the TET1 expression signal was found to be very strong in SC and LC. TET1 expression was reported in LCs in the previous study [75], also in SCs where the triangular nuclei of SCs were visible. These findings suggest that TET1 might regulate SC and LC to support smooth spermatogenesis.

In our mouse model study, Tet1 protein expression level was analysed in WT and *Tet1* $^{(-/-)}$ mice. It was observed that Tet1 was highly expressed in elongating spermatids in WT mice. Ni et al. found 5-hmC only in the human elongating spermatids [53] which might be the result of high TET1 activity and indirectly might support our finding in mice. This finding may denote there was an ongoing active DNA demethylation process [53] at this stage or that Tet1 might had the non-catalytical function to regulate other proteins like PRC2. Analyses of *Tet1* $^{(-/-)}$ mice showed that Tet1 is not expressed in germ cells but is expressed in SCs, showing the effectiveness of this knockout technique for studying the role of specific genes in specific cell types. *Tet1* knockout was germ-cell specific which was generated via the Cre-Lox conditional knockout technique. Hence, expression was only visible in somatic cells. SCs support germ cells during spermatogenesis. The strong expression of TET1/Tet1 in SCs in both mice and humans may suggest a conserved role for Tet1 across species. The presence of Tet1 in LCs in WT mice suggests that Tet1 might also play a role in these cells, possibly in testosterone production. TET1/Tet1 might be related to DNA demethylation or non-enzymatic function to regulate gene expression of other genes like retrotransposons [54] in both somatic cell types.

In human testes, TET1 expression has been observed in the early stages of spermatogenesis, while in mice, it appears in the later stages. The timing of this protein expression could be influenced by the variations in the epigenetic landscape between mice and humans [53]. Additionally, differences in regulatory elements like promoters or enhancers of this gene in humans and mice could lead to distinct expression patterns [76]. Extensive optimizations were

carried out to obtain the appropriate signal when performing IHC staining. Moreover, to validate the IHC results, a WB experiment was planned for tissue samples fixed in Bouin's solution. However, previous research conducted by our lab indicated that WB is not feasible with this type of tissue fixation unless it's frozen tissue. Unfortunately, frozen samples were not available, making it impossible to validate IHC results via WB. In future, further analyses like WB with frozen testis samples could further confirm TET1 expression in these mouse samples. In summary, this experiment elucidated the expression pattern of TET1/Tet1 across diverse germ cell populations and substantiated the successful knockout of *Tet1*, validating the efficacy of this methodology. Therefore, the impacts of *Tet1* knockout on the expression levels of other proteins will be more robust and reliable.

4.1.1 Depletion of *Tet1* induces altered expression of core PRC2 components and H3K27me3: link to male infertility

The next research question was whether there were any changes in PRC2 expression when *Tet1* was knocked out. Tet1 aids in the recruitment of PRC2 for H3K27 trimethylation at the promoters of bivalent genes, independently of its catalytic activity [51]. In this study, we present the initial stage-specific expression analysis of core PRC2 components EED/Eed, SUZ12/Suz12, and EZH2/Ezh2 in human and mouse testes during spermatogenesis using IHC.

In WT mice, Eed expression was primarily confined to elongating spermatids. In *Tet1* $^{(+/-)}$ mice, however, Eed expression was detected in a broader range of cell types, including spermatogonia, and spermatocytes, in addition to elongating spermatids. This suggests that the knockout of one allele may trigger a compensatory mechanism leading to expanded Eed expression in various cell types. This type of phenomenon was also observed by Pena et al. where it was shown that *Gata2b* compensated for the effects of *Gata2a* loss [77]. Interestingly, in *Tet1*$^{(-/-)}$ mice, Eed expression was absent.

Thus, our findings suggest that the knockout of the *Tet1* affected the expression patterns of PRC2 components in mouse testes. Specifically, in *Tet1*$^{(-/-)}$ mice, Eed expression was absent. Research suggested conditional deletion of *Eed* in male germ cells leads to complete male infertility [59,72]. As Eed establishes H3K27me3 through PRC2 and this epigenetic mark passes from sperm to next generation [59], therefore, these findings suggest an indirect role of Tet1 in male infertility.

SUZ12 showed expression in later stages of spermatogenesis, in elongating and maturing spermatids in *Tet1* $^{(-/-)}$ mice, compared to WT and *Tet1* $^{(+/-)}$ mice. It was also found expressed in peritubular cells, which was not observed in WT and *Tet1* $^{(+/-)}$ mice, indicating a novel finding; Peritubular myoid cells help support the structure of the tubule and play a role in controlling sperm production and overall testicular function[78]. Ezh2 expression was consistent in diplotene spermatocytes and round spermatids across all testes in all three groups of mice. A similar expression of Ezh2 was expressed in the apical region of the mouse round spermatid nucleus reported by Lambrot et al. [58] which also supports our current finding. In the *Tet1* $^{(-/-)}$ mice, Ezh2 was additionally expressed during the early stages of spermatogenesis which could be another effect of knockout of *Tet1*. Staining intensities showed high expression of Ezh2 in all the mentioned cell types in the mouse testes. Moreover, in *Tet1*$^{(-/-)}$ mice, our results showed Suz12 expression pattern shifted to the late stages of spermatogenesis, and Ezh2 had additional expression in the early stages of spermatogenesis. These types of altered expression could also stem from a compensatory mechanism due to *Tet1* knock-out, also suggested by Stringer et al., this altered expression could change RNA content in the sperm as well which might influence gene expression and embryo development [59]. Studies also indicate that the loss of any of the three essential components of the PRC2 complex significantly disrupts its function, resulting in reduced levels of H3K27me3. Furthermore, the complete depletion of PRC2 culminates in embryonic fatality [59,72].

In addition, this study investigated the expression patterns of PRC2 components in the human testis, aiming to discern their localization during spermatogenesis. Apart from EZH2 components, EED and SUZ12 expression experiments did not function. This could be caused by the non-specificity of the antibodies or improper tissue preparation, such as inappropriate fixation of the testes. EZH2 expression was found in various stages of human spermatogenesis, including spermatogonia, spermatocytes, round spermatids, and SCs. TET1 is also expressed in the same stages of spermatogenesis suggesting both proteins might interact together and facilitate chromatin compaction via H3K27me3. EZH2 expression was very strong in normal testicular tissues indicating an active function in this cell type. Our findings are consistent with previous studies that revealed higher mRNA levels of *EZH2* in normal testicular tissue compared to testicular tumour tissue [79]. Future studies should explore the expression pattern of SUZ12 and EED proteins in human testis using new antibodies and new testis samples preferably with another type of fixation.

Based on the above discussion, our findings hint towards the altered expression of core PRC2 components due to *Tet1* knockout. This altered expression might alter the H3K27me3 level during embryonic development. Since these possible embryonic development issues are related to male germ cells thus this could be a potential cause of male infertility [59].

4.1.2 Altered expression of H3K27me3 was observed in *Tet1*$^{(-/-)}$ *mice*

H3K27me3 mark is well studied in embryo [80] but it is sparsely investigated during human and mouse spermatogenesis, thus further investigations have been done here in the current study. We studied stage-specific expressions of the H3K27me3 mark during human and mouse spermatogenesis.

In our mouse model, the expression of H3K27me3 was compared between the WT and *Tet1*$^{(-/-)}$ mouse models to observe any differences in the expression pattern when *Tet1* is knocked out. In *Tet1*$^{(-/-)}$ mice, a slight shift of H3K27me3 expression was observed compared to WT mice, H3K27me3 expression was shifted from spermatogonia B to preleptotene spermatocytes. In elongating spermatids, expression was just different by one stage which could be due to an observational difference under the light microscope during staging. Also, these changes in expression might be due to a lack of non-enzymatic regulation of Tet1 in that stage, since Tet1 regulates this histone mark non-catalytically [51]. A recent study showed that H3K27me3 modification was present in mouse round spermatids [80] but we did not observe this in our mouse study. They used paraformaldehyde fixed round spermatid immunostaining which might gave better accessibility to the antibody whereas we used Bouin's fixed testis samples which might have resulted in this discrepancy. Moreover, since epigenetic modifications are dynamic; this observed difference could also be due to mouse strain specific. Further approaches, such as cell sorting and conducting WB or immunostaining on specific cell types would be helpful to confirm these findings.

Several studies have demonstrated that Tet1 has a regulatory role in the expression and function of PRC2 and H3K27me3 in various biological contexts, including embryonic stem cell biology, chromatin remodelling, and gene repression [51,54,81,82]. Additionally, *Tet1* knockout could disrupt the landscape of H3K27me3 and interfere with the bivalent state of developmental genes, resulting in abnormal gene activation and impaired differentiation [51,54,82]. Our investigation demonstrated altered expression of PRC2 and H3K27me3 during mouse spermatogenesis, which may also contribute to aberrant gene regulation during this process. These aberrant

epigenetic marks can directly pass to embryo [59] and cause aberrant chromatin remodelling and transcription resulting in defective embryonic genome activation (EGA) and embryogenesis. A precise EGA is very important for proper embryo development, implantation, and fertility [83]. Taken together, these findings suggest Tet1 may have an important role in male subfertility via modulating the expression of PRC2 and H3K27me3 during mouse spermatogenesis.

In humans, we observed that H3K27me3 was expressed during early and mid-stages of spermatogenic cell types, namely, spermatogonia B to leptotene spermatocytes, and in round spermatids. The same expression pattern was observed for the TET1 and EZH2 proteins, which indicates that these three factors might work together [50]. The staining intensities were strong in these cell types, especially in round spermatid intensity was very high depicting a higher expression. This could be because at this stage more H3K27me3 was required for the chromatin compaction and gene repression. This round spermatid is very important to get offspring from nonobstructive azoospermia patients [80]. In line with prior research, we also observed H3K27me3 expression in LCs [84].

Overall, in humans, we noticed that TET1, H3K27me3, and EZH2 are co-expressed in the same types of germ cells during spermatogenesis. This observation suggests that these factors may work in concert. As Tet enzymes have non-catalytical functions in EGA [83], TET1 may recruit EZH2 in these cells, which in turn facilitates the trimethylation of H3K27, thereby gene repression of GDR. In support of this hypothesis, we found negative correlations between the mRNA levels of *TET1* and two genes, *GATA2* and *FOXA2*. Both of these GDRs, bounded by H3K27me3, have been demonstrated to play crucial roles in embryo implantation in mice studies [62,85,86]. Further investigations using testis samples from subfertile patients could provide more insights into these correlations and their potential links to male subfertility.

4.2.0 The absence of TET1 in human sperm could lead to abnormal methylation in paternal DMRs and embryonic defect

This chapter aimed to discuss the presence and potential abnormalities of TET1, EED, EZH2, SUZ12, and H3K27me3 factors in mature sperm from subfertile men compared to fertile controls. Following various optimization procedures, WB analysis revealed the absence of TET1 protein in motile sperm from fertile individuals, despite its presence in other positive controls. TET1, a large protein weighing 235 kDa, presented challenges in effective transfer to the PVDF membrane. Wet transfer was identified as the most suitable method. Protein staining

indicated the presence of various proteins in sperm lysates, including those within the expected size range of 250-260 kDa for TET1. The TET1 antibody from GeneTex (GTX124207) identified a specific band at 260 kDa, possibly due to post-translational modifications or mRNA splice variants as noted by the antibody manufacturer. Despite equal protein loading, cancer cell lines exhibited significantly higher protein abundance compared to sperm lysates. This difference may stem from the diverse protein composition of cell-line lysates versus the limited protein content in sperm lysates due to chromatin compaction, remodelling, and lack of active transcription [87].

The current finding is inconsistent with previous research conducted by Ni et al., who reported the presence of TET1 in human sperm [53]. This inconsistency could be linked to the loss of TET1 during chromatin remodelling in spermiogenesis. We observed TET1 until round spermatids. Moreover, it is essential to acknowledge that the testis samples group and the fertile men group were separate entities. Therefore, the heterogeneity of the samples might have contributed to differing results. Unfortunately, due to constraints in both resources and time, it was not possible to carry out an immunostaining assay using motile sperm to validate this finding. However, this avenue could be explored in future research.

TET1 plays a crucial role in erasing paternal imprints [88]. Sperm from *Tet1*-deficient mice exhibits hypermethylation in paternally expressed DMRs, including *Peg3*; This *Peg3* DMR hypermethylation in *Tet1* knockout placentas showed variable phenotypes, including placental, foetal, and postnatal growth defects, leading to early embryonic lethality [88]. Research on *Tet1*-deficient mice has also shown a steady decrease in SSCs and disrupted spermatogenesis. This eventually results in a faster onset of infertility as the mice age [52]. Indeed, paternal Tet1 appears to play a crucial role in embryo development. These findings, coupled with our current results, indicate that the absence of TET1 protein in human sperm may result in abnormal methylation patterns in paternal DMRs, thereby might potentially affect embryo development and ultimately contributing to male subfertility.

4.2.1 Abnormal sperm H3K27me3 may impact early embryo development and cause male infertility

In mammals, the retention of histones and other TFs are necessary for proper chromatin structure during embryonic development [33,89]. In our investigation, we observed aberrant retention of H3K27me3 in the sperm of subfertile men. These observations were further

validated through the densitometric analysis of the WB data, revealing significantly elevated levels of H3K27me3 modification in the patient group compared to the fertile group. This could be due to chromatin remodelling issues during the histone-to-protamine exchange. Supporting our findings, scholarly inquiry suggests that infertile men tend to exhibit a more random nucleosome retention pattern, in contrast to the localized retention observed in fertile men [15].

The epigenetic marker H3K27me3, is believed to serve as cellular memory, preserving cell-type specific transcriptional states under normal developmental conditions [90]. This marker has been linked to male infertility [91]. Abnormal retention of H3K27me3 can impact the histone-to-protamine ratio, chromatin condensation, and epigenetic inheritance during embryogenesis [35,91,92]. Empirical evidence indicates that changes in the histone-retained sperm epigenome have been linked with idiopathic male infertility and reduced blastocyst growth during *in vitro* fertilization [15,93]. Sperm-inherited H3K27me3 plays an important role in the transcription and development of embryos. Any alterations to the histone retention could result in abnormal reprogramming of the paternal nucleus [15,94]. The retention of histones and DNA demethylation collectively establish a poised state, ensuring transcriptional capability and the activation of developmental regulators in the early embryo [92].

Thus, in our study, sperm-inherited aberrant H3K27me3 from subfertile patients could pose an adverse effect on EGA, and early embryogenesis and this might be one of the causes of idiopathic MI. Additionally, a significantly higher retention of H3K27me3 was observed in immotile sperm compared to motile sperm. This indicates that abnormal retention serves as an indicator of immature sperm, also consistent with findings from prior study [91].

4.2.2 SUZ12 might be linked to TET1 recruitment to sperm epigenome

As abnormal retention of H3K27me3 was identified in our subfertile patient's cohort, it was aimed to explore the retention of core PRC2 components as PRC2 mediates tri-methylation of H3K27. We found that the PRC2 components EZH2, EED, and SUZ12 were absent in the motile sperm of fertile men.

In contrast, in subfertile patients, only the SUZ12 antibody proved effective and demonstrated the absence of this protein. Our findings are consistent with previous studies which showed PRC2 components were not present in mouse sperm [92]. Also, PRC2/SUZ12 recruits TET1 to chromatin at H3K27me3 positive regions of the genome, contributing to epigenetic plasticity

throughout cell differentiation [50,70]. The absence of SUZ12 in sperm might have implications for TET1 recruitment in the sperm genome. However, despite multiple attempts and optimizations, no bands for the rest of the two core PRC2 components were detected in the sperm samples of patients or positive controls. This could be due to non-specific antibodies or inappropriate antigen masking.

Future studies could use immunostaining and immunofluorescence microscopy to clarify the retention of PRC2 components in patients' sperm samples. A mass spectrometric analysis could provide a comprehensive overview of the protein composition within sperm cells, potentially revealing the presence of TETs and PRC2 components.

4.3.0 Genome-wide binding analysis of TET1 faced limitations due to the compact chromatin structure and transient binding nature of TET1

TET1 regulate active DNA demethylation and transcription via chromatin remodelling [34,54]. In our study, TET1 native chromatin immunoprecipitation (TET1-NChIP) was performed to explore the genome-wide binding sites of TET1 in histone-associated fractions isolated from motile sperm of fertile men to analyse the binding pattern of TET1 to developmentally relevant TFs. The protamine-associated fraction was discarded. The immunoprecipitated DNA, along with the input DNA, was subjected to NGS and subsequent bioinformatic analysis. Surprisingly, it was observed a limited number of TET1 peaks across all samples, and the overlap between these peaks was also very low. These low number of peaks may be attributed to the transient nature of TET1's binding to the genome, suggesting that its interactions are brief and fleeting. Similar trends were reported that TET1 binds to weak CpG islands with weak enrichment [55,95]. An NChIP experiment was conducted without formaldehyde cross-linking, which might have affected TET1 binding and resulted in a low peak count. Thus, we performed TET1 X-ChIP with cross-linked human sperm in collaboration with Active Motif. The experiment was unsuccessful due to insufficient chromatin from seventy million motile sperm. The robust cell membrane and compact chromatin of sperm cells posed challenges for ChIP, leading to unsatisfactory results.

Research showed that PRC2 recruits TET1 to chromatin at H3K27me3 positive regions [50]. We found core PRC2 components and TET1 were absent in the mature sperm of fertile men. This evidence suggests that a low number of peaks resulted from either faulty TET1 recruitment in the sperm genome or due to its absence.

Future studies should explore sophisticated methods for isolating more chromatin from this specialized cell type. The CUT&Tag method requires fewer cells, may provide valuable insights, and clarify whether these challenges stem from the sample or methodology.

4.3.1 Investigating epigenetic regulation of H3K27me3/TET1 binding genes in the absence of TET1-ChIP-Seq

In this study, as TET1-ChIP-seq was unsuccessful, H3K27me3 binding at certain genes was assessed using NChIP followed by qPCR. The focus was on genes known to be bound by both H3K27me3 and TET1 based on a previous study done by Verma et al. [70]. The aim was to understand the epigenetic regulation of these genes without direct TET1 ChIP-seq data. By regulating H3K27 modifications, Tet1 plays a crucial role in ESCs and early development, primarily through its non-catalytic functions, which leads to the suppression of developmental genes [51].

Analysis of H3K27me3 chipped DNA concentration revealed higher levels in patient samples. This may be indicative of higher retention of H3K27me3 which potentially makes a stronger association with target genes. However, relative enrichment of H3K27me3 at the promoter regions of *SOX2, FOXA2, GATA2,* and *PAX6* was lower in subfertile patients, though not statistically significant. In contrast, *HERV*-k showed a trend of higher enrichment in patients, but the difference was not significant. The small sample size (4-6 individuals' sperm samples per group) limits data interpretation.

Except *HERV-k,* these genes are translated in TFs which have critical roles in embryo implantation, stemness in early embryogenesis, embryonic development and maintaining pluripotency of stem cells [96–99]. *HERVs* are of particular interest as *HERV-k* transcription has been observed during the process of normal human embryogenesis, starting with EGA at the eight-cell stage and continuing until the emergence of epiblast cells in preimplantation blastocysts [100]. However, research showed an observed upregulation of these transcripts and retroviral protein in certain cancers [98]. Since the sperm epigenome passes to the next generation from sperm [59], thus, any aberration in TFs regions may impact embryo development, such as abnormal expression of *HERV-k* due to H3K27me3 aberrant enrichment, which may potentially impact embryo development, pregnancy loss and leading to male subfertility. These preliminary findings need further research with a larger cohort for a comprehensive understanding.

4.4.0 Aberrant promoter methylation levels of *EZH2, SUZ12* and aberrant *EED* mRNA levels might be linked with an embryonic defect

Investigations propose that epigenetic modifications, including alterations in DNA methylation status at promoters of genes, are crucial for imprinting and development [15,92]**.**

In our patient cohort, it was detected a significant retention of H3K27me3. This prompted us to explore the potential variations in the methylation and mRNA levels of PRC2 components when PRC2 is recruited to chromatin by Tet1 for H3K27 trimethylation [51]. Using pyrosequencing, we found significant differences in the methylation levels of *EZH2* and *SUZ12* gene promoters between the patient and control groups. But both groups displayed notably low levels of methylation in the promoter regions of these three genes, with an average of approximately one per cent. Hypomethylated promoters are found in GDR and genes encoding signalling factors [35,101] which also align with our findings. This hypomethylation plays a crucial role in the processes of spermatogenesis, fertilization, and early embryonic development [101].

SUZ12 gene promoter methylation level was significantly higher in patients as compared to the control group of fertile men. As Suz12 is crucial for H3K27 trimethylation and Tet1 recruitment [50], this observed hypermethylation in patients could result in the aberration of H3K27 trimethylation and TET1 recruitment. Consequently, this aberration might impact various GDR due to its enrichment, as previously demonstrated by Erkek et. al. [33] . In contrast, the mRNA levels of *SUZ12* in patients were higher compared to controls, albeit without statistical significance.

EED is widely recognized as a regulator of gene silencing, playing a crucial role in maintaining the pluripotency of ESCs and promoting cell proliferation [102]. In the present study, the *EED* gene promoter displayed consistently low methylation levels, showing no significant difference between patients and controls. However, a notable contrast emerged in *EED* mRNA levels, where patients exhibited significantly lower *EED* mRNA levels than the fertile group. This aberrant observation might disrupt the *EED*-mediated gene regulation during embryo development, potentially impacting fertility. Because in mice, conditional deletion of *Eed* in the male germline resulted in complete male infertility and the deletion of *Eed* in developing oocytes led to a notable overgrowth phenotype in offspring. Interestingly, a similar overgrowth phenomenon has been observed in humans with Cohen-Gibson syndrome and Weaver syndrome, these syndromes arise from de novo germline mutations in *EED* or *EZH2*, respectively. These observations highlight the significance of both factors in regulating normal

growth and development, both in mice and humans [57,102]. However, in the current study, the mRNA levels of *EZH2* did not show statistically significant differences between patients and controls. However, for *EZH2* promoters, the patient group exhibited significantly lower methylation levels compared to the fertile men. Previous research indicated that *EZH2* may play a regulatory role in both oncogenes and tumour suppressor genes [103] and PRC2 regulates self-renewal and differentiation of developmental TFs [92]. Thus, in sperm, aberrations in this *EZH2* regulation could potentially result in developmental defects such as oncogenes expression, embryonic lethality suggested by Xu et al. [104], and consequently male subfertility. However, abnormal promoter methylation levels of *EZH2* and *SUZ12* didn't correlate with the mRNA level of corresponding genes. This could be due to a very low level of methylation which might not influence mRNA levels or post-transcriptional regulations e.g. mRNA stability which can be independent of methylation [36] in the spermatozoa.

As mentioned earlier, Sperm carries RNAs that affect embryo development [73]. The impact of observed methylation or mRNA differences on fertility can only be fully understood during pregnancy. Future studies may consider integrating functional assays, such as *in vitro* fertilization or embryo development experiments, to elucidate the functional consequences of altered PRC2 gene regulation in sperm cells.

4.4.1 Aberrant *TET1* and *TET3* mRNA levels may represent a potential risk factor for male infertility

Despite being transcriptionally inactive, mature sperm cells contain RNA, primarily composed of fragments from longer transcripts, including ribosomal RNA and specific mRNAs related to testes and spermatogenesis [71]. *TET* mRNA levels were found aberrant in different diseases including male subfertility, type 2 diabetes and breast cancer [53,105,106].

In the present research, the absence of TET1 protein in human spermatozoa prompted us to explore the levels of *TETs* mRNA in this cell. We demonstrated a significant increase in *TET1*, and *TET3* mRNA levels in spermatozoa from subfertile patients compared to fertile men. Impairments of *TET1* and *TET3* mRNA levels in sperm cells may represent a potential risk factor for male subfertility as suggested by a previous study [36]. Our findings in line with the work of Nettersheim et al., demonstrated increased *TET1* mRNA levels in pathological conditions [74]. Another study reported that the *Tet1, Tet2,* and *Tet3* mRNA levels were lower in haploid cells compared to diploid cells during murine spermatogenesis [107]. In line with this

study, we also found that haploid human spermatozoa have very low levels of *TETs* mRNA. This low level of mRNA could be attributed to the remnant of the previous transcription process during spermatogenesis as there is no active transcription occurs in mature sperm [87]. This small amount of RNA might function after fertilization in the EGA stage, as research showed sperm transcripts have potential as fertility markers [71].

The current findings contrasted with a previous study by Ni et al., who showed significantly elevated *TET1* and *TET3* mRNA levels in controls compared to subfertile patients [53]. It is important to note that the current study utilized a distinct cohort from the previous study, potentially accounting for the observed discrepancy in findings. However, for our *TET2* transcripts, no statistically significant difference was observed between the two groups. *TET2* is often found mutated in blood cancers and associated with DNA hypomethylation [55] which also might have happened in our samples.

4.4.2 Aberrant TF promoter methylation and mRNA level hint at a possible link to failures in embryo implantation and cause male infertility

Research showed that H3K27me3 histone mark enriched at genomic loci relevant for embryo development and these regions lack DNA methylation [33,57]. Thus, studying the DNA methylation status of GDR bound by H3K27me3 could shed light on embryo development issues and male infertility.

Here in this study, the promoter methylation levels of H3K27me3 binding genes (*SOX2, GATA2, PAX6,* and *FOXA2*), were analysed in motile sperm from subfertile patients and fertile men using pyrosequencing. For the rest of the H3K27me3 binding genes, pyrosequencing was not feasible as commercial primers were not available. Our result revealed that the overall promoter methylation levels of all GDR were very low (around 0.5 – 2.5 %). Similar findings were reported before in a gene ontology study [92]. Interestingly, Hammoud et al. showed that genomic regions responsible for regulating gene expression during development have reduced levels of promoter methylation in infertile men [15] which is in concordance with our findings. Notably, hypomethylated promoters overlapped with developmental TF promoters bound by self-renewal network TFs like OCT4, SOX2, NANOG, and FOXD3 found in hESC. These TFs promote self-renewal and collaborate with PRC2 to silence many developmental regulators [92]. In the present study, observed promoter hypomethylation could also result from the combined function of PRC2 and self-renewal network TFs to silence GDR and might play an important

role in establishing and maintaining the poised DNA methylation state in sperm. This low DNA methylation could prepare the paternal genome for early embryogenesis. Because Carrel et al. suggested that epigenetic modifications remain poised in sperm [92].

Research also showed knockout of the *Gata2* gene leads to hematopoietic failure and embryonic lethality [86]. Gata2 (GATA binding protein 2), a gene involved in progesterone signalling during early pregnancy [85], was found to have higher methylation levels in the promoter in subfertile patients than in fertile men. This could negatively affect embryo implantation and be a cause of male infertility. Because progesterone stimulates the uterus to facilitate embryo implantation [85]. The implication of higher methylation levels in the patient group was also reflected in *GATA2* mRNA levels as the patient's cohort showed a very low level of mRNA compared to the control group.

McSwiggin et al. reported associations between sperm-borne RNA molecules and fertility, showing that they have key roles in subsequent embryonic development [30]. In GDR, the H3K27me3 mark was found enriched reported by Erkek et al. [33]. The study also found that the retention of the H3K27me3 mark was aberrant in patients. This led to an investigation into the consequences of this aberration on the GDR. H3K27me3 binding eight TF genes, which are important for embryo development, were selected for this purpose.

Among the eight genes examined, mRNA levels of *HERV*-w exhibited significant downregulation in subfertile patients compared to fertile controls. This might be linked with aberrant retention of the H3K27m3 mark on these genes causing a repressed chromatin state. *HERV-w env* (also known as syncytin-1) transcript was reported in early embryogenesis at the eight-cell stage [100]. It is a vital membrane glycoprotein that has a significant role in the development of the placenta, fusion of syncytiotrophoblasts, fertilized egg fusion and embryo implantation [96,97]. Therefore, downregulation of the *HERV*-w gene in patients might lead to implantation defect of embryo and be one of the causes of male infertility.

In scientific literature, FOXA2 is a key TF involved in multiple aspects of development and reproduction e.g. embryo development, organ development, implantation, and fertility [61,62,108]. Basseres et al. reported that *Foxa2* expression was downregulated in lung cancer caused due to promoter hypermethylation [61]. In line with this research, in our study, *FOXA2* mRNA levels were very low and significantly downregulated in subfertile patients compared to fertile men. This is further underscored by studies on *Foxa2*-deficient mice, which demonstrated complete infertility due to complications with blastocyst implantation and stromal cell decidualization, thereby highlighting the critical role of *Foxa2* in processes related to embryogenesis and

fertility [62]. In summary, an aberration in the *FOXA2* mRNA level might play a role in the advancement of male infertility by disrupting implantation and early embryogenesis.

Interestingly, despite the downregulation of *FOXA2* mRNA, the methylation levels of the *FOXA2* gene promoters were low (around 2.5%), and there was no significant difference between the subfertile patients and fertile controls. This relationship between mRNA and methylation levels of *FOXA2* may be ascribed to the marginal disparities in methylation, approximately 2.5 %. These negligible fluctuations may not exert a substantial biological influence on transcription reported by Song et al. [109]. Also, these mRNAs could be the remnants of the transcription process and stored for functioning during fertilization [87].

In the present study, a significant upregulation of *SOX2* mRNA levels was observed in the spermatozoa of subfertile patients compared to fertile controls. However, there was no difference in the methylation levels of the *SOX2* gene between these two groups. Keramari et al. demonstrated that embryos lacking SRY-related HMG-box gene 2 (*Sox2*$^{-/-}$), do not survive beyond implantation, indicating the essentiality of Sox2 during implantation. Specifically, Sox2 is vital for the formation of the trophectoderm, an important component of the preimplantation embryo [63]. Thus aberrant paternal *SOX2* transcript might potentially impact embryo implantation and contribute to male infertility as sperm-inherited RNA influences fertility [30]. In addition to *SOX2*, it was also found an upregulation in the mRNA levels of *OCT4* and *SOX1* in subfertile patients. SOX1 has important functions in various stages of embryogenesis, including neurogenesis and stem cell protection and proliferation [110]. During preimplantation mammalian development, Sox2 and Oct4 form a vital complex, essential for maintaining pluripotency in embryonic stem cells within the inner cell mass (ICM) [63]. Our findings indicate that aberrations in *OCT4* and *SOX1* genes could impact embryogenesis and male infertility.

Moreover, despite low DNA methylation in the *PAX6* gene promoter, mRNA levels were also low. A low level of methylation, around 2.5%, might not significantly impact transcription, as noted in another study [109]. Isolation of sperm RNAs remains methodologically challenging due to their very low abundance and affinity of certain RNAs to chromatin [111]. This could be another reason of very low level of mRNA levels were observed for most of the genes.

In conclusion, given the critical function of TFs and the observed significant aberrations of *HERV-w, FOXA2, OCT4, SOX1, SOX2, and GATA2* genes in patients, hint at a possible connection to failures in embryo implantation and embryogenesis, resulting in miscarriage and male infertility.

4.5 Heterogeneous *LINE1* methylation pattern was observed in fertile men

Research demonstrated that Tet governs gene expression and repression of the endogenous retroviruses specifically *L1* in mouse ESCs, independent of DNA demethylation [54,112]. In our previous study, it was found that *L1* mRNA was significantly upregulated in the current cohort. Thus, in the present study methylation status of the *L1* and *Alu* (a type of short interspersed nuclear elements, SINEs) retrotransposon was investigated in motile sperm of fertile men using COBRA. Due to materials and time constraints, it was not possible to experiment with subfertile patients' samples.

The findings revealed that fertile men exhibited heterogeneity in methylation levels at the *L1* gene. Specifically, these individuals had methylation levels less than 30%. A previous study reported lower levels of *L1* methylation (ranging from 20% to 50%) could potentially pose a risk to genome stability [41]. Therefore, the present study showed that among the fertile men there was heterogeneity in *L1* methylation level and further investigation involving samples from patients could give us more information. For *Alu* elements, no detectable restriction sites had been observed in agarose gel, indicating no methylation or it is possible that restriction sites of restriction enzyme *Mbol* have been mutated. A previous study demonstrated that most of the CpG restriction sites for the enzyme *Mbol* were mutated [68] which supports our finding.

4.6 Correlation between promoter methylation, mRNA level, clinical parameters, and their link to male subfertility

Numerous genes exhibit altered methylation and mRNA levels associated with impaired spermatogenesis and reproductive dysfunction. Rotondo et al. showed inverse correlations between DNA methylation and mRNA levels of genes involved in various pathways, including hormone response and reproduction, in sperm from severe oligozoospermic and azoospermic patients [36].

Chrysanthou et al. showed Tet1's crucial roles in ESCs and early development involves non-catalytic regulation of H3K27 modifications to silence developmental genes [51]. Our results indicate that overexpression of *TET1* and *TET3*, along with over-retention of H3K27me3, are characteristics of infertile patients. These aberrations are linked with the overexpression and hypomethylation of a few GDR of spermatozoa. Since these sperm inherited factors have important functions during early embryonic development; increased levels of *TET1, TET3*

mRNA and over-retention of H3K27me3 may be involved in abnormal suppression of genes that are supposed to be expressed, such as *OCT4*.

We found a positive correlation between the mRNA levels of *TET1* and those of *SUZ12, PAX6,* and age. A study suggests that a deficiency in *TET1* impairs the ability of human embryonic stem cells to differentiate into neuroectoderm by decreasing the expression of PAX6, a key regulator in the development of human neuroectoderm [113]. Interestingly, the ectoderm marker Pax6 was significantly downregulated exclusively in *Tet1*$^{(-/-)}$ embryonic stem cells [51]. Given the abnormal *TET1* mRNA levels in the patient's sperm and its association with *PAX6* mRNA level and the role of PAX6 during embryogenesis, there might be a link to disrupted gene regulation, leading to defects in neurogenesis during embryo development and ultimately contributing to male infertility.

Additionally, our data showed that the methylation level of *SUZ12* was significantly higher in patients and that an increase in methylation correlates with a decrease in sperm motility. Since sperm motility is crucial for fertility, elevated methylation levels could potentially affect fertility. Moreover, both *TET1* and *OCT4* mRNA levels were upregulated in patients and positively correlated with *SUZ12* mRNA levels. We also observed a robust positive correlation between the mRNA levels of *SUZ12* and those of *NANOG* and *PAX6*. Both OCT4 and NANOG are essential for maintaining pluripotency in the ICM, and their lack of expression can lead to ICM failure, causing embryo development to halt at the blastocyst stage [114]. Therefore, our data suggests that the abnormalities observed in these factors in sperm may impact the regulation of genes associated with embryo development, considering the significant functions attributed to these remnants of sperm RNAs.

Additionally, research has shown that SOX2 is a crucial regulator in embryonic development and preimplantation processes [63]. We observed a significant inverse correlation between the mRNA level of *TET3* and the methylation status of the *SOX2* promoter. We also observed that both *SOX1* and *SOX2* mRNA levels were upregulated in patients. This suggests that TET3 might play a role in the demethylation of the *SOX2* promoter region. Therefore, we suggest that the upregulated *TET3* mRNA level in subfertile patients might facilitate the overexpression of *SOX2*, which in turn might trigger embryo implantation defects and cause male infertility.

In contrast to these findings, our data also suggest that more *FOXA2* and *GATA2* mRNA seem to be good for fertility, and both are upregulated in fertile men and inversely correlated with *TET1* mRNA. *GATA2* mRNA is also inversely correlated with *TET3* mRNA levels.

Interestingly, the *FOXA2* mRNA level was significantly lower which also corresponds with a relatively high *FOXA2* methylation trend, and they are negatively correlated to each other. The global *FOXA2* promoter methylation levels of sperm exhibited a positive correlation with age. This trend is supported by a recent study demonstrating a significant increase in sperm ribosomal DNA promoter methylation levels with advancing age [115].

Overall, these correlation analyses have established a preliminary link between GDR, PRC2 and TETs as a key component in the regulatory machinery governing the EGA, embryogenesis and male infertility. This is significant because TET enzymes, including TET1, play a crucial role in EGA and the precise regulation of transposable elements [83].

5. Limitation of the study and perspectives

For protein expression analysis

Despite several interesting findings, the present study has several drawbacks which should be noted. Limitations of the study include the unavailability of testis samples from patients undergoing assisted reproductive technology (ART), preventing a comparison of TET1 expression in spermatogenesis between fertile and infertile men. Future studies could explore TET expression in patient testis samples to address this gap. Additionally, validating the IHC results through WB was not possible due to Bouin's solution fixed paraffin embedded tissues, which is incompatible with WB proven by previous experiments unless the tissue is frozen. Unfortunately, frozen testis samples were unavailable, precluding validation of IHC results via WB. Future analyses using WB with frozen testis samples could confirm TET1, PRC2 and H3K27me3 expression in these samples.

Further research is needed to determine the expression pattern of SUZ12 and EED proteins in human testis using new antibodies and alternative tissue fixation methods, as the attempted IHC for SUZ12 and EED in humans did not yield conclusive results. Future investigations, such as IHC or cell sorting, coupled with WB or immunostaining on specific cell types, could help confirm this stage-specific characterization of these analysed factors.

For protein retention analysis

WB experiments of Tet1, H3K27me3, and PRC2 components were planned for mature mouse sperm. However, it was not feasible to collect mature sperm from *Tet1* $^{(-/-)}$ mice due to their shorter lifespan, as they typically died at four weeks of age. They reach sexual maturity at the age of five to six weeks. In future, other strains of mice could be used to overcome the limitation. An alternative approach would be to perform a *Tet1* knockdown experiment and investigate its effects on TET1, H3K27me3, and PRC2 retention in mouse sperm.

Conducting a TET1 WB experiment with motile sperm from subfertile ICSI patients was not possible due to a lack of materials. This avenue could also be explored in future studies to observe any possible differences between control and patient samples.

For EED and EZH2 Western blots, no bands were detected in patient sperm samples or positive controls. This lack of detection may be due to nonspecific antibodies, as other conditions yielded successful results for other target proteins with the same samples and conditions. Future studies could utilize immunofluorescence microscopy to visualize PRC2 components in patient sperm samples and conduct mass spectrometric analysis to comprehensively assess sperm protein composition, potentially identifying PRC2 components.

For TET1 and H3K27 binding analysis

For ChIP experiment, sperm sample heterogenicity, lots of optimizations were needed to generate mono-nucleosomal chromatin fragments, spending lot of time and resources. Sperm cells itself a challenging material due to their robust cell membrane and highly compact chromatin structure, which led to the failure of the TET1-N-ChIP experiment. So, to design future experiments, these aspects should be considered. Future studies should explore sophisticated methods for isolating more chromatin from this specialized cell type.

Moreover, each ChIP experiment required around 20 million motile spermatozoa which was challenging to obtain from each fertile men and subfertile patients which made the experiment more challenging for TET1 and H3K27me3 ChIP. In future, researchers might consider the CUT&Tag method which requires fewer cells and may provide valuable insights and TET1 and H3K27me3 binding analysis.

During the handling of sperm cells and isolated nucleic acids and proteins, measures were implemented to ensure proper storage. However, given the inherent susceptibility of biological samples to degradation during freeze-thaw cycles, the handling process may have inadvertently led to a reduction in sample quality.

Given the critical function of TFs and the observed significant aberrations of *HERV-w, FOXA2, OCT4, SOX1, SOX2, and GATA2* genes in subfertile patients, these observations highlight their potential role in embryo development and the progression of male subfertility. In future, working with knock-out human early-stage embryonic stem cell models of these selected genes and their *in vitro* development might give more insight into complex mechanisms of embryogenesis and their implications for infertility.

7. References

1. Assidi, M. Infertility in Men: Advances towards a Comprehensive and Integrative Strategy for Precision Theranostics. *Cells* **11**, 1711 (2022).

2. Vander Borght, M. & Wyns, C. Fertility and infertility: Definition and epidemiology. *Clin Biochem* **62**, 2–10 (2018).

3. Gnoth, C. *et al.* Definition and prevalence of subfertility and infertility. *Human Reproduction* **20**, 1144–1147 (2005).

4. Agarwal, A., Majzoub, A., Parekh, N. & Henkel, R. A Schematic Overview of the Current Status of Male Infertility Practice. *World J Mens Health* **38**, 308–322 (2020).

5. Agarwal, A. *et al.* Male infertility. *Lancet* **397**, 319–333 (2021).

6. Ding, G.-L. *et al.* The effects of diabetes on male fertility and epigenetic regulation during spermatogenesis. *Asian J Androl* **17**, 948 (2015).

7. Craig, J. R., Jenkins, T. G., Carrell, D. T. & Hotaling, J. M. Obesity, male infertility, and the sperm epigenome. *Fertil Steril* **107**, 848–859 (2017).

8. Gunes, S., Arslan, M. A., Hekim, G. N. T. & Asci, R. The role of epigenetics in idiopathic male infertility. *J Assist Reprod Genet* **33**, 553–569 (2016).

9. Gunes, S. & Esteves, S. C. Role of genetics and epigenetics in male infertility. *Andrologia* **53**, e13586 (2021).

10. Duca, Y., Calogero, A. E., Cannarella, R., Condorelli, R. A. & La Vignera, S. Current and emerging medical therapeutic agents for idiopathic male infertility. *Expert Opin Pharmacother* **20**, 55–67 (2019).

11. Punab, M. *et al.* Causes of male infertility: a 9-year prospective monocentre study on 1737 patients with reduced total sperm counts. *Human reproduction (Oxford)* **32**, 18–31 (2017).

12. Urdinguio, R. G. *et al.* Aberrant DNA methylation patterns of spermatozoa in men with unexplained infertility. *Hum Reprod* **30**, 1014–1028 (2015).

13. Ferfouri, F. *et al.* A genome-wide DNA methylation study in azoospermia. *Andrology* **1**, 815–821 (2013).

14. Houshdaran, S. *et al.* Widespread Epigenetic Abnormalities Suggest a Broad DNA Methylation Erasure Defect in Abnormal Human Sperm. *PLOS ONE* **2**, e1289 (2007).

15. Hammoud, S. S. *et al.* Genome-wide analysis identifies changes in histone retention and epigenetic modifications at developmental and imprinted gene loci in the sperm of infertile men. *Hum Reprod* **26**, 2558–2569 (2011).

16. Schütte, B. *et al.* Broad DNA methylation changes of spermatogenesis, inflammation and immune response-related genes in a subgroup of sperm samples for assisted reproduction. *Andrology* **1**, 822–829 (2013).

17. La, H. M. & Hobbs, R. M. Mechanisms regulating mammalian spermatogenesis and fertility recovery following germ cell depletion. *Cell Mol Life Sci* **76**, 4071–4102 (2019).

18. Salas-Huetos, A. *et al.* New insights into the expression profile and function of micro-ribonucleic acid in human spermatozoa. *Fertil Steril* **102**, 213-222.e4 (2014).

19. Durairajanayagam, D., Rengan, A. K., Sharma, R. K. & Agarwal, A. Sperm Biology from Production to Ejaculation. in *Unexplained Infertility: Pathophysiology, Evaluation and Treatment* (eds. Schattman, G. L., Esteves, S. C. & Agarwal, A.) 29–42 (Springer, New York, NY, 2015). doi:10.1007/978-1-4939-2140-9_5.

20. Hess, R. A. & de Franca, L. R. Spermatogenesis and Cycle of the Seminiferous Epithelium. in *Molecular Mechanisms in Spermatogenesis* (ed. Cheng, C. Y.) vol. 636 1–15 (Springer New York, New York, NY, 2009).

21. Misell, L. M. *et al.* A stable isotope-mass spectrometric method for measuring human spermatogenesis kinetics in vivo. *J Urol* **175**, 242–246; discussion 246 (2006).

22. Neto, F. T. L., Bach, P. V., Najari, B. B., Li, P. S. & Goldstein, M. Spermatogenesis in humans and its affecting factors. *Seminars in Cell & Developmental Biology* **59**, 10–26 (2016).

23. Fayomi, A. P. & Orwig, K. E. Spermatogonial stem cells and spermatogenesis in mice, monkeys and men. *Stem Cell Research* **29**, 207–214 (2018).

24. Cameron, D. F. & Hudson, J. C. Testicular Function. in *Reference Module in Biomedical Sciences* (Elsevier, 2014). doi:10.1016/B978-0-12-801238-3.00266-X.

25. Bergmann, M. & Kliesch, S. Testicular Biopsy and Histology. in *Andrology* (eds. Nieschlag, E., Behre, H. M. & Nieschlag, S.) 155–167 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2010). doi:10.1007/978-3-540-78355-8_11.

26. Ernst, C., Eling, N., Martinez-Jimenez, C. P., Marioni, J. C. & Odom, D. T. Staged developmental mapping and X chromosome transcriptional dynamics during mouse spermatogenesis. *Nat Commun* **10**, 1251 (2019).

27. Toshimori, K. *et al.* Impairment of spermatogenesis leading to infertility. *Anato Sci Int* **79**, 101–111 (2004).

28. Tüttelmann, F., Ruckert, C. & Röpke, A. Disorders of spermatogenesis: Perspectives for novel genetic diagnostics after 20 years of unchanged routine. *Medizinische Genetik* **30**, 12–20 (2018).

29. Boitrelle, F. *et al.* The Sixth Edition of the WHO Manual for Human Semen Analysis: A Critical Review and SWOT Analysis. *Life (Basel)* **11**, 1368 (2021).

30. McSwiggin, H. M. & O'Doherty, A. M. Epigenetic reprogramming during spermatogenesis and male factor infertility. *Reproduction* **156**, R9–R21 (2018).

31. Wang, T., Gao, H., Li, W. & Liu, C. Essential Role of Histone Replacement and Modifications in Male Fertility. *Frontiers in Genetics* **10**, (2019).

32. Dada, R. *et al.* Epigenetics and its role in male infertility. *J Assist Reprod Genet* **29**, 213–223 (2012).

33. Erkek, S. *et al.* Molecular determinants of nucleosome retention at CpG-rich sequences in mouse spermatozoa. *Nat Struct Mol Biol* **20**, 868–875 (2013).

34. Tahiliani, M. *et al.* Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1. *Science* **324**, 930–935 (2009).

35. Hammoud, S. S. *et al.* Distinctive chromatin in human sperm packages genes for embryo development. *Nature* **460**, 473–478 (2009).

36. Rotondo, J. C., Lanzillotti, C., Mazziotta, C., Tognon, M. & Martini, F. Epigenetics of Male Infertility: The Role of DNA Methylation. *Front. Cell Dev. Biol.* **9**, 689624 (2021).

37. Higashimoto, K. *et al.* Loss of CpG Methylation Is Strongly Correlated with Loss of Histone H3 Lysine 9 Methylation at DMR-LIT1 in Patients with Beckwith-Wiedemann Syndrome. *Am J Hum Genet* **73**, 948–956 (2003).

38. Poplinski, A., Tüttelmann, F., Kanber, D., Horsthemke, B. & Gromoll, J. Idiopathic male infertility is strongly associated with aberrant methylation of MEST and IGF2/H19 ICR1. *Int J Androl* **33**, 642–649 (2010).

39. Khazamipour, N., Noruzinia, M., Fatehmanesh, P., Keyhanee, M. & Pujol, P. MTHFR promoter hypermethylation in testicular biopsies of patients with non-obstructive azoospermia: the role of epigenetics in male infertility. *Hum Reprod* **24**, 2361–2364 (2009).

40. Marques, C. J. *et al.* Abnormal methylation of imprinted genes in human sperm is associated with oligozoospermia. *Mol Hum Reprod* **14**, 67–74 (2008).

41. Potabattula, R. *et al.* Increasing methylation of sperm rDNA and other repetitive elements in the aging male mammalian germline. *Aging Cell* **19**, e13181 (2020).

42. Bao, J. & Bedford, M. T. Epigenetic regulation of the histone-to-protamine transition during spermiogenesis. *Reproduction* **151**, R55–R70 (2016).

43. Campos, E. I. & Reinberg, D. Histones: annotating chromatin. *Annu Rev Genet* **43**, 559–599 (2009).

44. Hazzouri, M. *et al.* Regulated hyperacetylation of core histones during mouse spermatogenesis: involvement of histone deacetylases. *Eur J Cell Biol* **79**, 950–960 (2000).

45. Kyrgiafini, M.-A., Sarafidou, T. & Mamuris, Z. The Role of Long Noncoding RNAs on Male Infertility: A Systematic Review and In Silico Analysis. *Biology* **11**, 1510 (2022).

46. Hisano, M. *et al.* Genome-wide chromatin analysis in mature mouse and human spermatozoa. *Nat Protoc* **8**, 2449–2470 (2013).

47. Samans, B. *et al.* Uniformity of Nucleosome Preservation Pattern in Mammalian Sperm and Its Connection to Repetitive DNA Elements. *Developmental Cell* **30**, 23–35 (2014).

48. Joshi, K., Liu, S., Breslin S J, P. & Zhang, J. Mechanisms that regulate the activities of TET proteins. *Cell Mol Life Sci* **79**, 363 (2022).

49. Neri, F. *et al.* TET1 is controlled by pluripotency-associated factors in ESCs and downmodulated by PRC2 in differentiated cells and tissues. *Nucleic Acids Res* **43**, 6814–6826 (2015).

50. Neri, F. *et al.* Genome-wide analysis identifies a functional association of Tet1 and Polycomb repressive complex 2 in mouse embryonic stem cells. *Genome Biol* **14**, R91 (2013).

51. Chrysanthou, S. *et al.* The DNA dioxygenase Tet1 regulates H3K27 modification and embryonic stem cell biology independent of its catalytic activity. *Nucleic Acids Research* **50**, 3169–3189 (2022).

52. Huang, G. *et al.* Tet1 Deficiency Leads to Premature Reproductive Aging by Reducing Spermatogonia Stem Cells and Germ Cell Differentiation. *iScience* **23**, 100908 (2020).

53. Ni, K. *et al.* TET enzymes are successively expressed during human spermatogenesis and their expression level is pivotal for male fertility. *Hum. Reprod.* **31**, 1411–1424 (2016).

54. Stolz, P. *et al.* TET1 regulates gene expression and repression of endogenous retroviruses independent of DNA demethylation. *Nucleic Acids Research* **50**, 8491–8511 (2022).

55. Williams, K., Christensen, J. & Helin, K. DNA methylation: TET proteins—guardians of CpG islands? *EMBO reports* **13**, 28–35 (2012).

56. Margueron, R. & Reinberg, D. The Polycomb complex PRC2 and its mark in life. *Nature* **469**, 343–349 (2011).

57. Prokopuk, L., Stringer, J. M., Hogg, K., Elgass, K. D. & Western, P. S. PRC2 is required for extensive reorganization of H3K27me3 during epigenetic reprogramming in mouse fetal germ cells. *Epigenetics & Chromatin* **10**, 7 (2017).

58. Lambrot, R., Jones, S., Saint-Phar, S. & Kimmins, S. Specialized Distribution of the Histone Methyltransferase Ezh2 in the Nuclear Apical Region of Round Spermatids and Its Interaction With the Histone Variant H1t2. *Journal of Andrology* **33**, 1058–1066 (2012).

59. Stringer, J. M. *et al.* Reduced PRC2 function alters male germline epigenetic programming and paternal inheritance. *BMC Biol* **16**, 104 (2018).

60. Kaneshiro, K. R., Rechtsteiner, A. & Strome, S. Sperm-inherited H3K27me3 impacts offspring transcription and development in C. elegans. *Nat Commun* **10**, 1271 (2019).

61. Basseres, D. S. *et al.* Frequent downregulation of the transcription factor Foxa2 in lung cancer through epigenetic silencing. *Lung Cancer* **77**, 31–37 (2012).

62. Kelleher, A. M. *et al.* Forkhead box a2 (FOXA2) is essential for uterine function and fertility. *Proceedings of the National Academy of Sciences* **114**, E1018–E1026 (2017).

63. Keramari, M. *et al.* Sox2 Is Essential for Formation of Trophectoderm in the Preimplantation Embryo. *PLoS One* **5**, e13952 (2010).

64. Yu, B. *et al.* Epigenetic Alterations in Density Selected Human Spermatozoa for Assisted Reproduction. *PLOS ONE* **10**, e0145585 (2015).

65. Rahiminia, T. *et al.* Sperm chromatin and DNA integrity, methyltransferase mRNA levels, and global DNA methylation in oligoasthenoteratozoospermia. *Clin Exp Reprod Med* **45**, 17–24 (2018).

66. Blot Quantification. https://odyssey.licor.com/bio/support/contents/software/empiria-studio/blot-quantification.html.

67. Solomon, E. R., Caldwell, K. K. & Allan, A. M. A novel method for the normalization of ChIP-qPCR data. *MethodsX* **8**, 101504 (2021).

68. Yang, A. S. A simple method for estimating global DNA methylation using bisulfite PCR of repetitive DNA elements. *Nucleic Acids Research* **32**, 38e–338 (2004).

69. Heller, C. G. & Clermont, Y. Spermatogenesis in man: an estimate of its duration. *Science* **140**, 184–186 (1963).

70. Verma, N. *et al.* TET proteins safeguard bivalent promoters from de novo methylation in human embryonic stem cells. *Nat Genet* **50**, 83–95 (2018).

71. Casas, E. & Vavouri, T. Sperm epigenomics: challenges and opportunities. *Front Genet* **5**, 330 (2014).

72. Prokopuk, L., Stringer, J. M., Hogg, K., Elgass, K. D. & Western, P. S. PRC2 is required for extensive reorganization of H3K27me3 during epigenetic reprogramming in mouse fetal germ cells. *Epigenetics & Chromatin* **10**, 7 (2017).

73. Jodar, M. Sperm and seminal plasma RNAs: what roles do they play beyond fertilization? *Reproduction* **158**, R113–R123 (2019).

74. Nettersheim, D. *et al.* Analysis of TET expression/activity and 5mC oxidation during normal and malignant germ cell development. *PLoS One* **8**, e82881 (2013).

75. Zhou, S. *et al.* TET1 involved in bisphenol A induced TM3 Leydig cell toxicity by regulating Cav3.3 hydroxymethylation. *Chemosphere* **312**, 137171 (2023).

76. da Cruz, I. *et al.* Transcriptome analysis of highly purified mouse spermatogenic cell populations: gene expression signatures switch from meiotic-to postmeiotic-related processes at pachytene stage. *BMC Genomics* **17**, 294 (2016).

77. Peña, O. A. *et al.* Differential Requirement of Gata2a and Gata2b for Primitive and Definitive Myeloid Development in Zebrafish. *Front Cell Dev Biol* **9**, 708113 (2021).

78. Chen, L.-Y., Brown, P. R., Willis, W. B. & Eddy, E. M. Peritubular Myoid Cells Participate in Male Mouse Spermatogonial Stem Cell Maintenance. *Endocrinology* **155**, 4964–4974 (2014).

79. Hinz, S. *et al.* Deregulation of EZH2 expression in human spermatogenic disorders and testicular germ cell tumors. *World J Urol* **28**, 631–635 (2010).

80. Sakamoto, M. *et al.* Paternally inherited H3K27me3 affects chromatin accessibility in mouse embryos produced by round spermatid injection. *Development* **149**, dev200696 (2022).

81. Hagihara, Y., Asada, S., Maeda, T., Nakano, T. & Yamaguchi, S. Tet1 regulates epigenetic remodeling of the pericentromeric heterochromatin and chromocenter organization in DNA hypomethylated cells. *PLOS Genetics* **17**, e1009646 (2021).

82. Chrysanthou, S., Flores, J. C. & Dawlaty, M. M. Tet1 Suppresses p21 to Ensure Proper Cell Cycle Progression in Embryonic Stem Cells. *Cells* **11**, 1366 (2022).

83. Arand, J. *et al.* Tet enzymes are essential for early embryogenesis and completion of embryonic genome activation. *EMBO reports* **23**, e53968 (2022).

84. Kilcoyne, K. R. *et al.* Fetal programming of adult Leydig cell function by androgenic effects on stem/progenitor cells. *Proc Natl Acad Sci U S A* **111**, E1924-1932 (2014).

85. Rubel, C. A. *et al.* A Gata2-Dependent Transcription Network Regulates Uterine Progesterone Responsiveness and Endometrial Function. *Cell Reports* **17**, 1414–1425 (2016).

86. Fu, Y.-K. *et al.* Gata2-L359V impairs primitive and definitive hematopoiesis and blocks cell differentiation in murine chronic myelogenous leukemia model. *Cell Death Dis* **12**, 1–17 (2021).

87. Carrell, D. T. Epigenetics of the male gamete. *Fertility and Sterility* **97**, 267–274 (2012).

88. Vasconcelos, S., Caniçais, C., Chuva de Sousa Lopes, S. M., Marques, C. J. & Dória, S. The role of DNA hydroxymethylation and TET enzymes in placental development and pregnancy outcome. *Clinical Epigenetics* **15**, 66 (2023).

89. Torres-Flores, U. & Hernández-Hernández, A. The Interplay Between Replacement and Retention of Histones in the Sperm Genome. *Front Genet* **11**, 780 (2020).

90. Fukushima, H. S., Takeda, H. & Nakamura, R. Targeted in vivo epigenome editing of H3K27me3. *Epigenetics & Chromatin* **12**, 17 (2019).

91. Kutchy, N. A. *et al.* Acetylation and methylation of sperm histone 3 lysine 27 (H3K27ac and H3K27me3) are associated with bull fertility. *Andrologia* **50**, e12915 (2018).

92. Carrell, D. T. & Hammoud, S. S. The human sperm epigenome and its potential role in embryonic development. *Mol Hum Reprod* **16**, 37–47 (2010).

93. Denomme, M. M., McCallie, B. R., Parks, J. C., Schoolcraft, W. B. & Katz-Jaffe, M. G. Alterations in the sperm histone-retained epigenome are associated with unexplained male factor infertility and poor blastocyst development in donor oocyte IVF cycles. *Human Reproduction* **32**, 2443–2455 (2017).

94. Oliva, R. & Luís Ballescà, J. Altered histone retention and epigenetic modifications in the sperm of infertile men. *Asian J Androl* **14**, 239–240 (2012).

95. Zhang, W. *et al.* Isoform Switch of TET1 Regulates DNA Demethylation and Mouse Development. *Molecular Cell* **64**, 1062–1073 (2016).

96. Noorali, S. *et al.* Role of HERV-W syncytin-1 in placentation and maintenance of human pregnancy. *Appl Immunohistochem Mol Morphol* **17**, 319–328 (2009).

97. Wang, Q. *et al.* Molecular mechanisms of syncytin-1 in tumors and placental development related diseases. *Discov Onc* **14**, 104 (2023).

98. Shah, A. H. *et al.* Human endogenous retrovirus K contributes to a stem cell niche in glioblastoma. *J Clin Invest* **133**, (2023).

99. Verma, N. *et al.* TET proteins safeguard bivalent promoters from de novo methylation in human embryonic stem cells. *Nat Genet* **50**, 83–95 (2018).

100. Mao, J., Zhang, Q. & Cong, Y.-S. Human endogenous retroviruses in development and disease. *Computational and Structural Biotechnology Journal* **19**, 5978–5986 (2021).

101. Sujit, K. M. *et al.* Genome-wide differential methylation analyses identifies methylation signatures of male infertility. *Human Reproduction* **33**, 2256–2267 (2018).

102. Prokopuk, L. *et al.* Loss of maternal EED results in postnatal overgrowth. *Clinical Epigenetics* **10**, 95 (2018).

103. Kumari, K., Das, B., Adhya, A. K., Rath, A. K. & Mishra, S. K. Genome-wide expression analysis reveals six contravened targets of EZH2 associated with breast cancer patient survival. *Sci Rep* **9**, 1974 (2019).

104. Xu, R., Li, C., Liu, X. & Gao, S. Insights into epigenetic patterns in mammalian early embryos. *Protein & Cell* **12**, 7–28 (2021).

105. Yang, L., Yu, S.-J., Hong, Q., Yang, Y. & Shao, Z.-M. Reduced Expression of TET1, TET2, TET3 and TDG mRNAs Are Associated with Poor Prognosis of Patients with Early Breast Cancer. *PLOS ONE* **10**, e0133896 (2015).

106. Zhao, S. *et al.* Reduced mRNA and Protein Expression Levels of Tet Methylcytosine Dioxygenase 3 in Endothelial Progenitor Cells of Patients of Type 2 Diabetes With Peripheral Artery Disease. *Frontiers in Immunology* **9**, (2018).

107. Gan, H. *et al.* Dynamics of 5-hydroxymethylcytosine during mouse spermatogenesis. *Nat Commun* **4**, 1995 (2013).

108. Li, J. *et al.* TET1 dioxygenase is required for FOXA2-associated chromatin remodeling in pancreatic beta-cell differentiation. *Nat Commun* **13**, 3907 (2022).

109. Song, J. *et al.* Aberrant DNA methylation and expression of SPDEF and FOXA2 in airway epithelium of patients with COPD. *Clinical Epigenetics* **9**, 42 (2017).

110. Kanwore, K. *et al.* SOX1 Is a Backup Gene for Brain Neurons and Glioma Stem Cell Protection and Proliferation. *Mol Neurobiol* **58**, 2634–2642 (2021).

111. Goodrich, R. J., Anton, E. & Krawetz, S. A. Isolating mRNA and small noncoding RNAs from human sperm. *Methods Mol Biol* **927**, 385–396 (2013).

112. de la Rica, L. *et al.* TET-dependent regulation of retrotransposable elements in mouse embryonic stem cells. *Genome Biol* **17**, 234 (2016).

113. Li, H. *et al.* TET1 Deficiency Impairs Morphogen-free Differentiation of Human Embryonic Stem Cells to Neuroectoderm. *Sci Rep* **10**, 10343 (2020).

114. Tulay, P. Control of Embryonic Gene Expression and Epigenetics. in *Embryo Cleavage* (IntechOpen, 2017). doi:10.5772/67851.

115. Li, L. *et al.* Sperm Ribosomal DNA Promoter Methylation Levels Are Correlated With Paternal Aging and May Relate With in vitro Fertilization Outcomes. *Frontiers in Genetics* **11**, (2020).

8. Supplements

Table 20: Clinical data of fertile men denoted as controls

ID	Age	PM (%)	TM (%)	SC(Mio/ml)	TSC (Mio)	Mor (%)
DFG-US-01	40	78	100	21	63	2
DFG-US-02	37	50	79	155	930	3
DFG-US-03	45	71	100	11.8	23.6	5
DFG-US-04	45	42	83	53.4	138.84	5
DFG-US-05	43	25	75	61.8	123.6	16
DFG-US-06	39	93	99	195	351	17
DFG-US-07	42	81	95	79.8	518.7	2
DFG-US-08	49	89	93	159.5	701.8	4
DFG-US-09	41	69	84	137.5	440	4
DFG-US-10	42	18	30	29.3	102.55	3
DFG-US-11	38	83	95	84	378	?
DFG-US-12	26	71	90	22.9	91.6	8
DFG-US-13	48	97	100	74.9	247.17	3
DFG-US-14	35	71	79	96.5	289.5	15
DFG-US-15	43	44	75	107.5	215	4
DFG-US-16	29	88	97	20	60	9
DFG-US-17	24	0	50	25.2	108.36	8
DFG-US-18	24	14	14	63.4	177.52	9
DFG-US-19	30	81	94	44.8	134.4	3
DFG-US-20	30	20	30	19.5	42.9	5
DFG-US-21	20	35	41	35.5	120.7	8
DFG-US-22	20	100	100	25.3	161.92	3
DFG-US-23	41	43	93	105	262.5	5
DFG-US-24	29	95	98	53	174.9	9
DFG-US-25	28	66	80	100	300	16
DFG-US-26	24	45	73	127	254	9
DFG-US-27	38	7	19	11.3	88.14	5
DFG-US-28	25	52	79	83	290.5	7
DFG-US-29	26	68	95	13.8	45.54	6
DFG-US-30	24	22	22	31.2	99.84	78
DFG-US-31	28	61	67	88.6	177.2	?
DFG-US-32	23	76	84	28	28	?
DFG-US-33	45	85	87.5	79	237	?
DFG-US-34	28	84	87	138	621	?
DFG-US-35	26	85	89	50	210	22
DFG-US-36	24	91	97	25.5	140.25	9
DFG-US-37	20	92	95	164.8	824	20
DFG-US-38	32	92	95	73.5	235.2	?

DFG-US-39	45	81	85	151.3	347.99	?
DFG-US-40	25	57	63	43.6	139.52	11
DFG-US-41	20	49	54	173.4	173.4	24
DFG-US-42	28	90	95	63.8	191.4	13
DFG-US-43	24	79	83	48.2	409.7	12
DFG-US-44	29	68	90	0.9	2.88	2
DFG-US-45	26	86	94	157.8	568.08	8
DFG-US-46	27	86	88	64.5	258	11
DFG-US-47	29	78	82	74.2	222.6	8
DFG-US-48	28	95	96	65.4	228.9	6
DFG-US-49	33	28	39	18.5	33.3	11
DFG-US-50	20	82	86	22.6	67.8	2
DFG-US-51	25	83	88	60.2	210.7	9
DFG-US-52	22	93	99	44.8	116.48	12
DFG-US-53	27	85	86	91	109.2	11
DFG-US-54	26	92	93	151.6	606.4	12
DFG-US-55	32	93	97	55	302.5	10
DFG-US-56	35	65	74	74.2	289.38	7
DFG-US-57	27	92	96	209.5	314.25	27
DFG-US-58	28	93	98	30	153	9
DFG-US-59	24	72	79	30.8	30.8	5
DFG-US-60	29	88	98	96.8	532.4	1
DFG-US-61	31	91	93	282	705	14
DFG-US-62	25	80	88	89.8	107.76	12
DFG-US-63	26	89	89	189.5	568.5	13
DFG-US-64	39	78	96	105.4	316.2	9
DFG-US-65	27	92	95	35.4	123.9	12
DFG-US-66	25	89	94	173	519	1
DFG-US-67	39	0	0	29.4	47.04	0
DFG-US-68	30	84	90	57.6	230.4	18
DFG-US-69	26	94	95	37.2	119.04	13
DFG-US-70	32	86	90	108.6	325.8	13
DFG-US-71	20	90	92	116	406	31
DFG-US-72	25	74	79	14	70	19
DFG-US-73	27	70	74	12.7	49.53	5
DFG-US-74	27	68	75	12	42	11
DFG-US-75	26	85	94	44.6	133.8	17
DFG-US-76	23	89	93	148	725.2	10
DFG-US-77	23	92	94	86	275.2	2
DFG-US-78	36	91	92	34.2	71.82	
DFG-US-79	32	76	82	27.6	85.56	6
DFG-US-80	30	0	0	41.8	104.5	0
DFG-US-81	34	89	94	45	144	10
DFG-US-82	31	90	97	44.4	173.16	4

DFG-US-83	41	96	97	57	142.5	17
DFG-US-84	36	94	97	25.4	124.46	14
DFG-US-85	31	89	93	35.2	88	15
DFG-US-86	39	83	89	21.2	78.44	14
DFG-US-87	48	78	81	15.4	46.2	14
DFG-US-88	30	88	93	45.8	178.62	8
DFG-US-89	31	20	64	4.3	13.33	5
DFG-US-90	29	92	94	69.8	195.44	0
DFG-US-91	29	87	91	31.2	74.88	14

SC=Sperm concentration (Mio/ml); TSC=Total sperm concentration (Mio); PM= Progressive motility; TM= Total motility; Mor=Morphology; ID = Identification Number

Table 21: Clinical data of subfertile men denoted as patients

ID	Age	SC(Mio/ml)	TSC (Mio)	PM (%)	TM (%)	Mor (%)
P01	42	13	20,8	28	48	0
P02	47	4	14,8	18	39	0
P03	39	11	64,9	29	55	0
P04	37	18,4	31,28	23	50	0
P05	50	36	136,8	28	46	0
P06	41	3,2	18,24	41	59	0
P07	32	2,4	9,12	12	32	0
P08	29	46,6	219,2	40	71	4
P09	33	48	110,4	43	67	4
P10	43	3,2	10,24	14	25	0
P11	38	10,2	55,08	13	30	0
P12	43	6,5	22,75	54	66	0
P13	38	2,5	5,75	31	52	0
P14	38	8	5,12	16	52	0
P15	45	219	394,2	60	82	5
P16	37	126	491,4	60	82	6
P17	47	4,3	22,36	40	61	0
P18	32	33	135,3	26	53	0
P19	47	57	68,4	26	46	0
P20	37	10,8	30,24	24	35	1
P21	45	17,7	157,3	53	76	0
P22	31	4,8	15,84	26	51	0
P23	39	1,2	6,84	9	30	0
P24	30	84	310,8	41	63	1
P25	38	31	213,9	55	82	0
P26	37	34	268,6	30	60	0
P27	35	23	85,1	59	81	0
P28	46	25	87,5	51	75	0

P29		11,6	85,84	43	58	0
P30	36	78	163,8	31	61	2
P31	40	58	63,8	59	77	0
P32	44	155	294,5	36	70	0
P33	40	10,1	36,36	34	56	0
P34	43	11	71,5	43	66	0
P35	43	45	585	60	83	0
P36	45	68,4	177,84	34	64	0
P37	40	63,9	166,14	54	80	0
P38	39	14	51,8	60	79	0
P39	38	81	186,3	41	72	0
P40	33	21	52,5	31	61	2
P41	40	17,7	72,57	38	53	0
P42	35	15	130,5	62	83	0
P43	38	19	72,2	42	74	0
P44	33	11,6	27,84	46	62	0
P45	42	3,4	8,16	26	51	0
P46	31	44	233,2	47	72	0
P47	43	49	176,4	45	70	0
P48	37	16	75,2	39	59	0

SC=Sperm concentration (Mio/ml); TSC=Total sperm concentration (Mio); PM= Progressive motility; TM= Total motility; Mor=Morphology; ID = Identification Number.

Table 22: Clinical data of female partners of patients

ID	Age	BMI	FSH	Mt. P.	AMH	Vit. D	Vit. B12	Folate	Estradiol	FR
P01	31	22	175	normal	2.9	29.9	939	19	2899	60
P02	36	22	187	normo	1,2	32	319		1734	67
P03	36	24	137	normo	nd	nd	nd	nd	5085	60
P04	33		100		nd	nd	nd	nd	1368	70
P05	40	22	225	hetero	1.5	29	1026	20	5154	69
P06	28	22	187	normo	3.0	26.6	491	0.8	4745	50
P07	31	20	175	hetero	2.3	34,8	460	19,5	2675	78
P08	31	22	150	normo	1.2	nd	656	20	3562	89
P09	30	22	18	normo	1.5	33	491	20	2235	100
P10	45	21	225	hetero	0,9	32	983	20	1276	60
P11	35	23	200	hetero	nd	nd	nd	nd	3240	50
P12	40	25	225	normo	0.2	32	500	20	1823	67
P13	38	23	183	normo	nd	nd	nd	nd	2889	76
P14	35	24		hetero	5,4	17,5	593	4,7	3203	60
P15	28	20	132	hetero	7.3	34.2	872	20	2090	80
P16	38	22	150		2.5	36.8	489	17.8	2893	67

P17	41	23	225	hetero	0,7	21.2	800	20	2152	50
P18	30	22	175	hetero	2,6	19.9	679	20	2902	86
P19	33	21	225	hetero	0.5	nd	794	15.4	2142	80
P20	36	23	75	normo	2.1	nd	583	13	3006	71
P21	43	21	150	normo	3.2	42	583	17.9	6387	58
P22	31	22	125	hetero	3.4	23.9	600	20	220.8	
P23	36	25	200	normo					1875	22
P24	33	26	225	normo	0.5	nd	252	11	1434	100
P25	34	35	225		0,5	nd	nd	nd	1772	100
P26	35	23	225	hetero	nd	29	1112	16	1566	83
P27	37	22	225	normo	0.9	21	451	20	5850	75
P28	34		150	hetero	4.7	49.6	426	11.1	3241	58
P29	33	26	132	homo	7.4	23.5	410	13.0	1757	57
P30	33	22	175	hetero	2.0	nd	1376	20	2710	56
P31	38	24	100	normo	0,7	36,3	753	20		
P32	37	22	225	hetero	1.3	19.7	476	8.4	4101	38
P33	39	22	225	normo		21	416	7.9	1572	
P34	37	24	200	hetero	0,5	19	523		1060	38
P35	40	20	225	hetero	0.6	24.4	377	20	1433	56
P36	43	24	175	normo	3.0	22.7	407	20	3345	79
P37	37	22	225	hetero	0,6	nd	nd	nd	2394	57
P38	37	21	225	hetero	1.7	60	397	20	4884	69
P39	35	23	132	normo	3,8	25.2	786	20	1380	50
P40	37	22		hetero	3.7	28	479	16.2	1629	100
P41	33	22	225	hetero	nd	nd	nd	nd	5198	67
P42	37	26,6	150	normo	4.0	21,2	485	16,8	2244	55
P43	35	21	150	normo	0.7	13	413	19	3609	80
P44	32	21	150	homo	2.8	nd	271		1602	77
P45	36	20	225	hetero	1,0	26,4	426	16,4		0
P46	28	24	50	normo	nd	nd	nd	nd	3970	64
P47	44	24	200	hetero	0,8	17,2	615	20	3895	79
P48	27	20	125	hetero	3.1	34	403	17	1506	53

Nd = not detectable, ID = Identification Number. Mt. P. =MTHFR677 polymorphism; AMH=Anti- Müllerian hormone; Vit=Vitamin; Folate; Estradiol; FR=Fertilization rate

Table 23: mRNA analysis data of *TET1, TET2, TET3, SUZ12, EZH2, EED* in fertile men

ID	***TET1***	***TET2***	***TET3***	***SUZ12***	***EZH2***	***EED***
DFG-US-03		4.27	40.69			
DFG-US-06		1.80	0.90			
DFG-US-09	2.91				0.12	4.44
DFG-US-10	5.02		2.70			
DFG-US-11	0.81		0.97			
DFG-US-12	1.48		2.51			1.46
DFG-US-13	6.28		3.42			

DFG-US-14	6.42				0.15	
DFG-US-18	3.03		1.26		8.66	9.50
DFG-US-19	0.72		8.93	21.36		
DFG-US-20	0.74					20.94
DFG-US-21						10.19
DFG-US-22	1.67		0.90			
DFG-US-23	0.96					
DFG-US-24	1.01		1.09	2.08		0.90
DFG-US-25	0.63	1.63		1.66	0.03	0.05
DFG-US-26	1.13			1.38		
DFG-US-27	0.84			2.23	0.03	0.01
DFG-US-28	0.50				35.06	0.28
DFG-US-29	0.43					
DFG-US-31	0.62	0.13			0.90	1.64
DFG-US-33	1.96	0.94			0.04	13.89
DFG-US-34		1.01	0.24			
DFG-US-35	1.01	2.30	1.89	0.49	0.00	0.06
DFG-US-36	0.42				0.03	65.02
DFG-US-37						
DFG-US-38				1.07		
DFG-US-39	0.42				0.04	53.17
DFG-US-40	0.84	2.18	0.63			
DFG-US-41		2.25	0.96	0.83		
DFG-US-42		3.04	2.06		0.00	0.00
DFG-US-43		1.12	1.96	0.00	0.00	
DFG-US-44	0.72	0.86	0.61			
DFG-US-45	0.15					
DFG-US-46		0.94	0.64			
DFG-US-47	0.08					
DFG-US-48		0.87	1.71			
DFG-US-50		0.48	0.69			3.48
DFG-US-51	0.47					
DFG-US-52				1.24		14.04
DFG-US-53				0.48		18.77
DFG-US-54		1.71	0.77			11.89
DFG-US-55		0.58	0.59			
DFG-US-56		3.45	2.18	1.82		10.61
DFG-US-57		3.77	0.87			
DFG-US-58	0.50					
DFG-US-59	1.09					
DFG-US-60	0.89					
DFG-US-61	0.79			0.89	0.00	0.47
DFG-US-62		0.99	0.13			
DFG-US-63	0.81			3.95	0.01	27.06
DFG-US-64	1.05					
DFG-US-65		1.87	0.44			23.81
DFG-US-66	0.99	1.36	0.45			
DFG-US-68	0.72					

DFG-US-69		2.14	0.26			
DFG-US-70		1.03	0.48			
DFG-US-71	0.59					
DFG-US-72		3.03	0.25			
DFG-US-73		0.16	0.03			
DFG-US-74		2.28	12.12			
DFG-US-76		0.36	0.26			
DFG-US-77			2.07			
DFG-US-78		3.68	0.64			
DFG-US-79		1.78	6.13			
DFG-US-80		1.31	0.39			
DFG-US-81		0.64	1.09			
DFG-US-82		0.32	0.84			
DFG-US-83		0.60	1.82	3.78		
DFG-US-84		0.24	0.72			
DFG-US-85		0.28	1.79			
DFG-US-86		0.16	0.28			
DFG-US-87		1.00	1.16			
DFG-US-88		1.41	3.86			
DFG-US-89		0.15	0.24	0.17	52.88	12.25
DFG-US-90		1.22	1.48			
DFG-US-91		0.12	0.69			
DFG-US-95				0.97	0.53	1.36
DFG-US-97				0.96	0.00	
DFG-US-108				2.48	0.70	0.23
DFG-US-109				0.97	0.09	0.19
DFG-US-110				1.11	0.28	0.19
DFG-US-113				0.91	0.15	1.84
DFG-US-114				0.81	0.04	0.09
DFG-US-115				0.74	0.05	0.04
DFG-US-116				1.36		4.07
DFG-US-118				1.32	0.67	0.25
DFG-US-119				2.23	0.59	0.25
DFG-US-120				0.74	0.09	0.04
DFG-US-121				1.30	2.78	0.07
DFG-US-123				1.31	1.60	0.03
DFG-US-125				1.50	0.14	0.05

ID = Identification Number.

Table 24: mRNA analysis data of *TET1, TET2, TET3, SUZ12, EZH2, EED* in patients

ID	***TET1***	***TET2***	***TET3***	***SUZ12***	***EZH2***	***EED***
P01	0.92	0.72	3.16	0.64	0.02	0.06
P02	0.40	0.62	1.21	0.59		
P03	0.90	1.04	1.54	0.81	0.29	0.04
P04	1.14	1.12	1.78	0.27	0.19	0.02
P05	1.09	1.48	1.23	0.76	0.04	0.02
P06	0.70	2.68	1.50	0.90	0.22	0.30

P07	0.34	0.43		0.50	0.01	0.77
P08	0.65	0.93	1.32	1.56	0.21	0.00
P09	0.63	0.98	1.04	0.99	0.09	0.02
P10	0.73	0.76	0.86	0.09	0.49	0.02
P11	1.94	1.31		3.54	1.67	0.02
P12	0.79	0.68	0.52	1.70	3.28	0.02
P13	0.59	0.93		1.45	0.00	0.02
P14	0.53	1.17	1.10		0.52	0.04
P15	1.44	1.14	0.41	2.41		
P16	1.85	1.15		5.08		
P17	1.41	0.28		2.77	0.00	0.00
P18	0.89	1.17	0.74	1.33	0.00	
P19	3.11	4.28	1.23		1.65	0.02
P20	1.51	1.09		2.25	1.27	0.01
P21	2.80	1.20	1.59		0.03	0.01
P22	1.89	1.71	1.53	1.52	0.00	0.00
P23	0.20	0.24		0.53		0.20
P24	0.77	1.14		1.82		
P25	0.87	0.66	0.66		0.01	0.00
P26	0.78	1.98	1.20	1.74	0.15	0.00
P27	0.57	0.24	3.20	1.82	0.00	
P28	1.19	0.89		4.03	0.00	
P29	1.18	1.90	1.47	2.41	0.00	
P30	0.70	1.07	53.99	1.09		
P31	0.81	1.00		0.02		
P32	0.54	0.27			0.20	0.03
P33	0.83	0.61	1.99	1.41	0.11	0.01
P34	1.33	2.43	1.11		0.01	0.00
P35	0.94	0.61	1.65	0.05		0.01
P36	1.96	1.69	1.54		0.06	0.00
P37	1.56	1.65	1.46			
P38	1.15	2.68				
P39	0.84	0.26		3.29	0.00	
P40	1.20	1.12	1.67	1.58	0.00	
P41	1.38	1.26	1.42			
P42	2.95	1.69	1.62			
P43	2.07	1.70	1.53	2.01	0.00	
P44	1.61	1.62	1.91		0.00	
P45	2.14	1.76	1.07			
P46	2.68	1.92	1.39		0.00	
P47	3.82	4.10	1.37		0.00	
P48	2.59	1.63	1.24			

ID = Identification Number.

Table 25: Pyrosequencing methylation analysis data from fertile men

ID	*EED*	*SUZ12*	*EZH2*	*PAX6*	*SOX2*	*GATA2*	*FOXA2*
DFG-US-03	1	1.33	1.83	0.67	0.67	0.33	
DFG-US-04	1.2	1.00	1.33	0.50	1.50	0.00	2.00
DFG-US-06	0.6	0.67	1.33	0.83	0.50	0.00	2.86
DFG-US-07	0.8	0.33	2.83	0.33	0.33	0.00	10.00
DFG-US-08							
DFG-US-09	0.8	0.33	1.33	0.67	1.67	0.67	5.57
DFG-US-10	0.2	0.67	0.83	0.67	0.50	0.33	1.57
DFG-US-11	1	0.33	1.83	0.33	0.67	0.00	1.14
DFG-US-12	0.4	1.67	1.33	0.83	0.67	0.56	0.57
DFG-US-16	1	1.00	3	1.5	0.33	0.00	1.86
DFG-US-24	0.6	0.33	1.50	0.67	0.83	0.44	0.86
DFG-US-25	0.6	0.67	1.17	0.50	0.33	0.56	2.14
DFG-US-26	0.8	1.00	1.00	0.50	0.50	0.11	1.43
DFG-US-27							
DFG-US-28	0.6	0.33	0.83	0.50	2.17	0.00	2.57
DFG-US-30	0.6	2.00	1.00	0.50	1.67	0.22	1.43
DFG-US-31	0.6	1.00	0.83	0.50	0.83	0.00	1.43
DFG-US-33	0.6	0.33	1.33	0.50	0.50	0.11	2.71
DFG-US-34	0.4	0.67	1.00	0.50	1.00	0.22	1.29
DFG-US-35	0.4	0.67	0.83	1.67	0.33	0.44	0.86
DFG-US-36	0.2	0.33	1.17	0.33	1.00	0.00	4.71
DFG-US-39	0.4	0.33	0.83	0.50	0.83	1.00	2.29
DFG-US-41	1	0.67	1.00	0.50	0.67	0.00	0.86
DFG-US-44	0	0.33	1.00	0.33	5.17	0.44	1.29
DFG-US-45	0.2	1.00		0.83	4.33	0.44	1.86
DFG-US-46	0.4	0.33	1.50	0.83	10.33	0.11	5.00
DFG-US-48	0.4	0.33	2.67	0.67	3.17	0.44	1.00
DFG-US-49	0	0.33	1.00	1.33	1.83	0.22	1.14
DFG-US-50	3.6	0.33	1.33	0.33	0.67	0.67	2.00
DFG-US-51	3	0.67	0.83	0.50	0.50	0.11	3.86
DFG-US-52	2.2	0.67	0.83	0.83	1.83	0.67	0.57

DFG-US-53	1.6	1.00	0.83	0.67	6.83	0.22	0.57
DFG-US-55	1.8	0.33	0.67	0.33	6.83	0.22	0.57
DFG-US-57	0	0.33	0.83	0.33	4.17	0.56	5.43
DFG-US-58			1.00	0.50		0.11	0.43
DFG-US-59			0.83	0.50		0.33	6.86
DFG-US-60			0.83			0.00	0.43
DFG-US-61					1.83		1.14
DFG-US-62					4.33		

ID = Identification Number.

Table 26: Pyrosequencing methylation analysis data from patients

ID	***EED***	***SUZ12***	***EZH2***	***PAX6***	***SOX2***	***GATA2***	***FOXA2***
P01	0.4	1.33	1.50	0.50	1.00	0.67	1.57
P02	0.6	1.00	1.00	0.83	1.33	0.11	1.14
P03	0.4	1.00	0.67	0.67	1.50	0.22	2.14
P04	0.2	0.67	0.50	0.50	1.50	1.00	2.86
P05	0.2	1.00	0.83	2.00	1.17	0.22	2.57
P06	0.4		0.33	0.50	0.83	0.67	2.00
P07							
P08	0.2	1.00	0.83	0.67	1.50	0.22	1.86
P09	0.2	1.00	0.83	0.50	0.83	0.44	3.86
P10	0.2	0.33	0.83	0.50	1.67	1.00	1.43
P11	0.2	1.67	0.83	0.50	1.83	0.33	4.57
P12	0.4	1.00	0.50	2.33	1.17	0.22	2.57
P13	0	1.00	0.83	0.50	1.17	0.00	0.14
P14	0.8	1.00	0.67	0.83	3.50	0.78	
P15	0.2	1.33	0.50	0.50	1.50	0.33	
P16	0.2	1.00	0.50	0.50	0.83	0.11	0.71
P17	0.4	1.00	0.33	0.83	2.67	0.56	3.14
P18	0.6	1.33	1.17	0.83	2.00	0.56	2.29
P19	1	1.33	1.00	0.67	2.17	0.89	3.29
P20	0.4	1.00	0.83	0.50	1.17	0.44	2.57
P21	0.2	1.00	0.83	0.50	1.33	0.33	2.71
P22	0.8	1.33	1.33	0.67	1.00	1.00	1.71
P23	0.4	1.00	0.33	1.17	2.00	0.22	2.71
P24	0.4	1.00	0.83	0.83		1.67	2.71
P25	1.2	1.33	1.00	0.50		4.00	0.29
P26	1.6	3.00	1.33	0.33	1.17	0.56	2.43
P27	1	0.67	1.17	0.83	0.67	1.22	2.86
P28	1.2	0.33	0.83	0.67	0.17	0.67	2.86
P29	0.4	0.33	0.83	0.67	6.17	0.11	2.14
P30	0.2	0.33	1.00	0.67	5.83	0.22	1.29
P31	0.2	0.67	0.67	0.33	9.00	0.33	5.00
P32	0.8	0.67	1.00	0.50	4.17	0.78	4.71

P33	1.6	0.67	0.83	0.67	1.33	0.33	3.00
P34							
P35	1	0.33	0.83	0.50	0.67	2.33	1.29
P36	1.4	1.00		0.67	0.33	1.22	2.14
P37	1	1.00	0.67		1.17		3.71
P38	1.8	0.33					3.86

ID = Identification Number.

Table 27: mRNA analysis of H3K27me3 binding genes from fertile men

ID	*NANOG*	*PAX6*	*HERV- w*	*SOX1*	*SOX2*	*GATA2*	*OCT 4*	*FOXA*
DFG-US-04			1.45	2.38	1.67			8.09
DFG-US-19	2.44	1.61						
DFG-US-22							3.07	36.46
DFG-US-23	0.70							
DFG-US-24	1.72	1.63	1.15	8.55	1.65			
DFG-US-25	1.92	2.02						
DFG-US-26	0.79	3.23					1.01	0.12
DFG-US-27	1.80	1.84	2.01	7.37	2.31			
DFG-US-35	1.64	2.93	1.52	2.61	1.04	0.08	0.90	0.32
DFG-US-38	1.27	1.71					1.66	4.76
DFG-US-41	0.58	0.68						
DFG-US-43	0.41	1.43	1.76	0.47	1.29	5.76	1.64	4.70
DFG-US-52	0.92	0.78					1.40	5.15
DFG-US-53	1.78	1.60					1.08	6.45
DFG-US-56	1.69	2.89	0.21	2.08	1.63		0.35	4.68
DFG-US-61	0.76		1.78	2.30	2.74	0.14	0.82	0.55
DFG-US-63	1.63	2.65					3.86	1.18
DFG-US-72					1.60	5.37		
DFG-US-73							0.08	61.08
DFG-US-74							1.44	1.73
DFG-US-78							10.61	5.45
DFG-US-79							0.50	4.23
DFG-US-80							1.18	4.67
DFG-US-83	0.79	0.99						
DFG-US-85							0.44	5.27
DFG-US-86						2.01		
DFG-US-89	0.01	0.02				94.18	0.03	0.48
DFG-US-90								
DFG-US-91			1.74	2.11	1.83	0.98		
DFG-US-95	0.69	0.99	0.46	2.81	8.56		0.84	1.03
DFG-US-96			1.54	6.46	3.65	0.31		
DFG-US-97	0.74							
DFG-US-98			2.15	4.06	0.80	0.66		
DFG-US-101			1.79	3.59	2.31	37.79		
DFG-US-102			1.82	2.86	1.42	34.74		
DFG-US-105			2.42	3.07	0.66	1.63		
DFG-US-106			1.25	2.49		4.43		

DFG-US-108	2.48	1.43	1.34	3.54	2.08	7.58	2.52	0.35
DFG-US-109	1.21	0.76	1.80	1.43	0.68	0.09	0.98	0.44
DFG-US-110	0.98	0.59	1.74	1.43	1.26	0.09	0.93	0.32
DFG-US-113	1.17	0.77	0.91	0.00	0.61	66.52	1.46	0.39
DFG-US-114	0.94	0.78	0.05	0.00	0.17	0.07	1.15	
DFG-US-115	1.44	0.65	0.04	1.85	0.00	0.02	1.04	
DFG-US-116	0.62	0.47	0.02	0.01				
DFG-US-117								
DFG-US-118	2.01	1.02	0.96	1.19	1.12	0.03	1.27	0.13
DFG-US-119	2.14	0.94	1.49	2.30	1.25	0.01	1.17	0.09
DFG-US-120	1.17	0.71	1.44	1.45	1.14	1.00	0.91	0.02
DFG-US-121	1.30	0.96	1.64	2.13	0.89	0.52	0.79	0.04
DFG-US-122								
DFG-US-123	0.73	0.87	3.06	3.44	3.12	0.03	1.26	0.13
DFG-US-124								
DFG-US-125	1.56	1.03	2.37	2.56	2.37	0.02	1.77	0.06
DFG-US-126								
DFG-US-127			1.01	0.93	0.81	9.29		
DFG-US-184			0.81	7.93		12.15		
DFG-US-301			1.21	0.28	0.41			
DFG-US-302			2.24	1.99	1.43	0.55		
DFG-US-369			0.98	0.48	0.63	36.44		
DFG-US-601			1.22	3.75	0.40	3.55		
P5G61	0.84	1.18						
P5G123					1.05			
P5G125					0.95			

ID = Identification Number.

Table 28: mRNA analysis data of H3K27me3 binding genes from patients

ID	*NANOG*	*PAX6*	*HERV- w*	*SOX1*	*SOX2*	*GATA2*	*OCT 4*	*FOXA*
P01	1.43	1.15	0.75	3.49	5.27	0.15		0.59
P02	0.30	0.38	0.72	2.45		10.24	0.49	1.48
P03	1.37	1.32	1.41	4.51	0.63	0.07		0.12
P04	1.36	0.67	1.05	2.75		0.10	3.65	0.19
P05	1.80	1.43	1.17	2.91		6.88	138.77	13.53
P06	2.44	3.56	0.59	1.76		0.34	1.62	
P07	0.55	0.50	0.54	5.65	2.71	0.73	2.77	2.97
P08	0.98	0.91	1.39	2.63	1.67	0.41	29.26	0.97
P09	0.17	0.91	0.57	11.07	2.26		6.08	1.14
P10	0.14	0.24	0.37	1.67	3.52	0.24	1.83	0.20
P11	6.52	7.15	0.63	2.66	1.50	0.15	2.65	0.09
P12	1.48	0.15	0.16	5.96		0.88	0.40	0.74
P13	0.95	0.59	1.15	4.61	1.97	0.05	1.50	0.19
P14			1.21	6.04	2.96			
P15	6.31	6.22						
P16	12.00	7.07						
P17	1.52	0.64	1.23	1.70	6.34	0.04	5.43	0.14
P18	2.00	1.58				6.20	1.10	0.19

P19			0.96	4.94	2.64	0.02	3.43	0.08
P20	5.55	6.17	1.17	3.81	3.25	0.02	2.12	0.07
P21			0.69	2.46	4.84	0.02	7.15	0.06
P22	1.24	1.24					3.28	
P23	1.92	1.05	1.25	2.73				
P24	5.07							
P25			1.56	1.75	0.40			
P26	0.96	1.06	0.42	1.98		1.01	2.00	4.09
P27						0.34	7.06	0.03
P28	2.26	2.38				1.62	14.14	0.16
P29	1.94	2.26				0.01	4.82	0.09
P30	1.22	0.85						
P31	0.43	0.54						
P32			0.89	1.39	1.93			
P33	2.41	1.40	0.63	2.12	3.08			
P34			0.92	3.72	2.65			
P35	0.09	0.07	0.84	3.33				
P36								
P37			1.02	2.13	4.45			
P38			2.45	3.34	2.16			
P39	2.31	2.55	1.79	2.42		0.39	8.96	21.20
P40	1.02	3.26					5.38	0.29
P41						0.55	1.04	0.05
P42			1.30	2.64	3.18	0.01	3.77	0.03
P43	0.91		3.21	3.07	1.33	0.04	1.92	0.15
P44			1.43	2.63	1.61	0.07	1.28	0.05
P45			1.09	4.92	3.82	0.01	2.44	0.03
P46			1.23	2.69	3.25	0.64	9.49	1.43
P47			1.59	3.90	1.85	0.07	6.80	0.55
P48			0.91	3.79	0.99	0.02	1.71	0.04

ID = Identification Number.

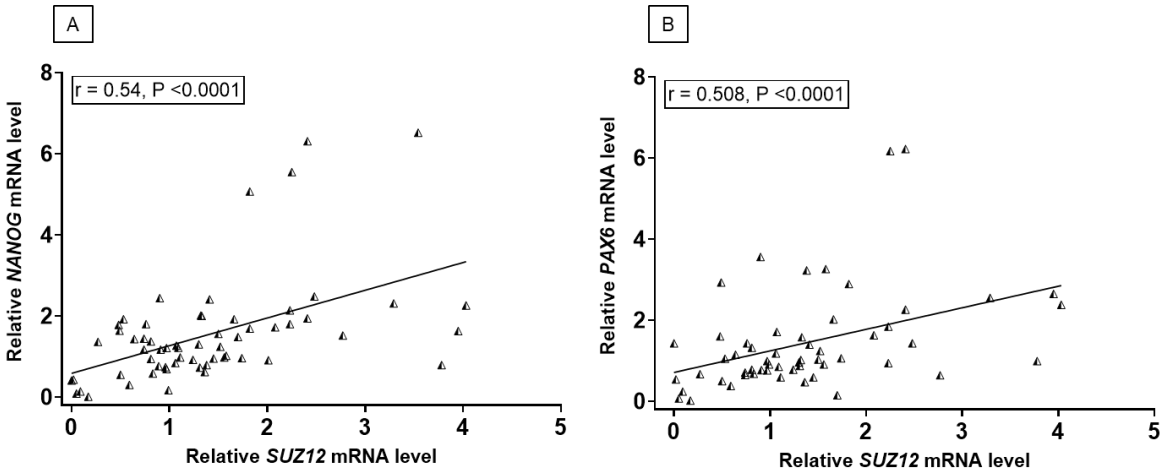


Figure S1: Correlation analysis between *SUZ12* and *NANOG, PAX6* mRNA levels. A) Correlation analysis of *SUZ12* and *NANOG* mRNA levels: strongly positively correlated each other (r = 0.54) with a P-value < 0.0001 B) Correlation analysis of *SUZ12* and *PAX6* mRNA strongly positively correlated each other (r = 0.508) with a P-value < 0.0001.

Publications and Presentations

Article

- Ozturk N, Dansranjavin T, Gies S, Calay D, Shiplu S, Creppe C, Hendrickx J, Schagdarsurengin U. H4K20me3 marks distal intergenic and repetitive regions in human mature spermatozoa. Development. 2021 Aug 1;148(15): dev196477

Conference abstracts, GGL and IRTG (GRK1871) presentations

- Oral Presentation at IRTG Milestone Meeting in Melbourne (December 2018): "Impacts of epigenetic interactions between TET1 and PRC2 in spermatogenesis and fertility."
- Oral Presentation at the 12th meeting of the Network of Young Researchers in Andrology (September 2019). "Impacts of epigenetic interactions between TET1 and PRC2 in spermatogenesis and fertility."
- Poster Presentation at 12th GGL Annual Conference 2019 (September 2019): "Impacts of epigenetic interactions between TET1 and PRC2 in spermatogenesis and fertility."
- Poster Presentation at 6th International Workshop Molecular Andrology (September 2019): "Impacts of epigenetic interactions between TET1 and PRC2 in spermatogenesis and fertility."
- Oral Presentation at GGL Retreat (May 2019): "Impacts of epigenetic interactions between TET1 and PRC2 in spermatogenesis and fertility."
- Poster Presentation at Congress of Andrology 2020 (December 2020): "Role of epigenetic factors TETs, and PRC2 in male infertility."
- Poster Presentation at First Digital Annual GGL Conference (September 2020): "TET1 and PRC2: Roles in spermatogenesis and fertility."
- Oral Presentation at IRTG Progress Meeting (July 2020): "TET1 and PRC2: Roles in spermatogenesis and male fertility."
- Oral Presentation at GGL Retreat and Mid-candidature Review (July 2021): "TET1 and PRC2: Roles in spermatogenesis and male fertility."
- Poster Presentation at AUF Symposium: Urologische Forschung der Deutschen Gesellschaft für Urologie (November 2021): "Epigenetic enzymes TETs, DNMTs and PRC2: Links to male infertility."
- Poster Presentation at 14th GGL Annual Conference 2021 (September 2021): "TETs, DNMTs, and PRC2: Links to male infertility."

- Oral Presentation at IRTG Progress Meeting (July 2021): "TETs, DNMTs, and PRC2: Links to male infertility."
- Oral Presentation at IRTG Progress Meeting (December 2021): "TETs, DNMTs, and PRC2: Links to male infertility."
- Oral Presentation at IRTG Milestone Meeting / Retreat (June 2022): "Investigation of the role of epigenetic factors TET1-3 and PRC2 in the occurrence of male infertility."

Acknowledgement

I would like to express my deepest appreciation to a multitude of individuals who have provided me with their unwavering support throughout my PhD journey. First and foremost, I am profoundly grateful to the Almighty for His countless blessings and grace that have allowed me to complete my work. My heartfelt gratitude goes to my supervisor, Prof. Dr. rer. nat. Undraga Schagdarsurengin. Your constant encouragement and guidance, especially during the thesis writing process, have been invaluable and have greatly facilitated my progress over the past few years. I would also like to extend my special thanks to Prof. Dr. rer. nat. Klaus Steger for his guidance throughout the immunohistochemistry work and thesis writing. Thanks to Associate Prof. Patrick Western for his guidance during the early phase of the project.

I am thankful to Prof. Dr. med. Hans-Christian Schuppe and all the technicians from our department, particularly Tania and Kerstin, for their efforts in collecting and performing an andrological analysis of the semen samples from healthy donors. I am also thankful to Prof. Dr. med. N. Rogenhofer, Prof. Dr. vet. med. D. Fietz, Prof. Dr. K. P. Koh's lab for providing testis and semen samples.

I would also like to express my gratitude to all the members of the Molecular Andrology and Urology Laboratory, both present and former colleagues: Nihan, Nils, Sabrina, Alex, Andreas, Othmane, and the late Babara. Your companionship, encouragement, and support have made the long hours in the laboratory more bearable. I am grateful to the IRTG – GRK1871 "Molecular Pathogenesis of Male Reproductive Disorders", DFG, Department of Urology for funding my employment and my project. I would also like to thank the IRTG, members of IRTG and Giessen Graduate Centre for the Life Sciences, JLU for providing exciting seminars, workshops, and conferences.

I am also grateful to my friend, Dr. S. Jamali, for his motivational, encouraging words and support during the challenging periods of my PhD. I would like to extend my thanks to Dr. Jamali, and Dr. A. Romero for their advice during the thesis writing process.

Finally, I would like to express my deepest gratitude and special thanks to all my family members, especially my respected parents and beloved wife, for their unconditional love, patience, and support, which have been crucial during my PhD journey. Your efforts have made all this possible.