The effect of seminal fluid on TBX2 and TBX3 expression and activity in cervical cancer cells

---

George William Cooper

Supervisors:

Professor Arieh A Katz and Professor Sharon Prince

Presented for the degree of

Master of Science in Medicine

Medical Biochemistry

Department of Integrative Biomedical Sciences

Faculty of Health Sciences

University of Cape Town

March 2017

Declaration

I, George William Cooper, hereby declare that the work on which this thesis is based is my original work and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other University. I empower the University of Cape Town to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever.

Signed by candidate

George William Cooper

Date: 12 March 2017

i

Acknowledgements

I would like to thank my supervisors, Professor Arieh A Katz and Professor Sharon Prince; your support and guidance was crucial to the completion of this project and to my development as a scientist.

I would also like to acknowledge the members of the Katz and Prince labs; the culture and environments wouldn't have been the same without any of you. I am particularly grateful to Dr Anthonio Adefuye, Dr Deeya Ballim, Dr Tarryn Willmer, and Dr Aretha Cooper; your patience, diligence and care while teaching me are sincerely appreciated. Furthermore, I would like to specially thank Sandra Jordaan, Graham Christians, Dr Roshan Ebrahim and Xolani Nonzinyana for their continued efforts to maintain a pleasant, safe environment and for the running of the logistical aspects of our laboratories. On a personal note, I would like to thank Jenna Bleloch, Dr Aretha Cooper, Melissa Blumenthal and Sylvia Ujma. I never knew I would find such great friends along this journey. I am eternally grateful to have met all of you, and forever indebted for the support you have given me.

I would also like to thank my parents, Christine and Peter Cooper, as well as my sister, Nina. Your patience and tolerance during this period knew no bounds and your support was overwhelming. I love you all very much.

Marcus Kretschmer, John Laurie Human and Margriet Vollgraaff; life would be extremely lonely without you. Thanks for supporting me through this. Markie, thanks for your patience with my erratic schedule during this period and for your emotional support throughout it.

I gratefully acknowledge the financial support provided by the National Research Foundation in the form of a Scarce Skills Development Fund Scholarship. I also gratefully acknowledge the financial support provided by the Struwig-Germeshuysen Kankernavorsingtrust towards the completion of my studies. Opinions expressed or conclusions drawn in this or any other related publication are mine alone and do not necessarily align with those of the SGKN-Trust.

Table of Contents

List of Figures & Tables

List of Abbreviations

°C	Degrees Celsius
A	
Act D	Actinomycin D
ANOVA	Analysis of Variance
ART	Assisted Reproductive Therapy / Therapies
C	
CCL2	C-C chemokine Ligand 2
CCL20	C-C chemokine Ligand 20
CD4	Cluster of Differentiation 4
cDNA	complementary DNA
CIN	Cervical Intraepithelial Neoplasia
CNRQ	Calibrated Normalised Relative Quantity
CO_2	Carbon dioxide
Cq	Quantification cycle
CSF2	Colony Stimulating Factor 2
Ct	Threshold cycle
D	
DAB	3,3'-diaminobenzidine
dH_2O	Deionised Water
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic Acid
dNTP	deoxyribonucleotide Triphosphate
E	
EDTA	Ethylenediaminetetraacetic Acid
EGFR	Epidermal Growth Factor Receptor
ERK 1/2	Extracellular Signal-Regulated Kinase 1/2
F	
FCS	Foetal Calf Serum
FFPE	Formalin-Fixed, Paraffin-Embedded
G	
-g	gram(s)

H

H_2O	Water
HCl	Hydrochloric Acid
HIV	Human Immunodeficiency Virus
H_2O_2	Hydrogen Peroxide
HPV	Human Papilloma Virus
HREC REF	Human Research Ethics Committee Reference

I

IFN-γ	Interferon γ
IL	Interleukin
IVF	In-vitro Fertilisation

K

kb	Kilobase Pairs
K_2HPO_4	Dipotassium Phosphate
KCl	Potassium Chloride

L

-L	Litre(s)
LCR	Long Control Region
LOF	Loss-Of-Function

M

m-	milli-
-m	metre(s)
-M	Molar
mESC	Mouse Embryonic Stem Cell
$MgSO_4$	Magnesium Sulfate
mRNA	messenger Ribonucleic Acid

N

n-	Nano
Na_2HPO_4	Disodium phosphate
NaCl	Sodium Chloride
$NaHCO_3$	Sodium Bicarbonate
NRQ	Normalised Relative Quantity

P

P/S	Penicillin/Streptomycin

p21	$p21^{CIP1/WAF1}$
PAGE	Polyacrylamide Gel Electrophoresis
PBS	Phosphate Buffered Saline
PBS/T	Phosphate Buffered Saline with 0.1% Tween 20
PCR	Polymerase Chain Reaction
PG	Prostaglandin
PGE_2	Prostaglandin E_2
PI3K	Phosphoinositide 3-Kinase
PTEN	Phosphate and Tensin Homolog
PTGS2	Prostaglandin Endoperoxide Synthase 2

Q

qPCR	Quantitative real time Polymerase Chain Reaction

R

Rb	Retinoblastoma Protein
RNA	Ribonucleic Acid
rpm	revolutions per minute
RT	Room Tempereature
RT-qPCR	Reverse Transcription Quantitative Real Time Polymerase Chain Reaction

S

SDS	Sodium Dodecyl Sulfate
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
SEM	Standard Error of the Mean
SF	Seminal Fluid
shRNA	Short Hairpin Ribonucleic Acid
siRNA	Small Interrupting Ribonucleic Acid
SNP	Single Nuclear Polymorphism
SP	Seminal Plasma

T

TBS	Tris Buffered Saline
TBS/T	Tris Buffered Saline with 0.1% Tween
TBX	T-box Transcription Factor
TEMED	Tetramethylethylenediamine
TGF-β	Transforming Growth Factor β
Tm	Melting Temperature

U

UCT	University of Cape Town

UMS	Ulnar-Mammary Syndrome

V

VEGF	Vascular Endothelial Growth Factor
Veh	Vehicle

µ-	micro-

Abstract

Cervical cancer is one of the most common female cancers in Africa, both in terms of incidence and mortality, and is disproportionately prevalent in developing nations due to a lack of adequate access to healthcare. While new vaccine technologies are rapidly reducing the incidence of Human Papilloma Virus (HPV) infection, the primary causative agent of cervical cancer, new cases continue to accumulate in the developing world. Beyond the role of HPV in the early stages of cancer development, the molecular aetiology of this disease is poorly understood. Frequent exposure to seminal fluid (SF), the liquid component of semen, has been proposed as a potential driver of oncogenesis in cervical cancers and has been shown to exacerbate some aspects of cervical cancers. While some of the cellular signaling pathways responsible for these phenomena have been identified, much remains to be elucidated. We hypothesized that TBX2 and TBX3, two highly homologous transcription factors frequently implicated in other cancers, may be responsible for mediating some of the effects of SF on cervical cancer cells. We established that TBX3 protein is significantly overexpressed in both primary cervical adenocarcinomas and squamous cell carcinomas compared to normal tissue. SF was shown to increase expression of both TBX2 and TBX3 mRNA in HeLa and CaSki, but not C-33 A, cervical cancer cell lines. Furthermore, SF upregulated TBX3 protein expression in both of these cell lines. In contrast, TBX2 protein was undetectable in these cell lines. In addition, our results showed that SF treatment of HeLa cells increases the expression of the known TBX3 target gene, $p21^{CIP1/WAF1}$ (p21), while having no effect on PTEN expression. Transient knockdown of TBX3 resulted in decreased p21 expression in SF-treated cells suggesting that SF upregulation of p21 is dependent on TBX3. This is the first study to investigate TBX3 protein expression in primary cervical tissues and SF regulation of TBX3. However, further research is required in order to elucidate the role of SF-induced TBX3 in cervical cancer development. The identification of the role of TBX3 in cervical cancer development could aid in the development of more effective treatments for cervical cancers and could potentially impact sexual health policy recommendations for women with cervical cancer.

1. Introduction

1.1. Cervical Cancer

Cervical cancer refers to cancer of the uterine cervix, which forms the part of the uterus that protrudes into the vagina in the female reproductive tract. While a number of risk factors have been identified, the necessary causative agent of cervical cancer is infection with the Human Papilloma Virus (HPV). In addition, there is a long period between infection and disease onset during which precancerous changes are easily identified, thus making screening programs very effective preventative interventions. Recently, high-efficacy vaccines targeting the most oncogenic HPV strains have been developed. Given the infectious nature of the disease, nations that have established effective medical screening and vaccination programs are able to largely prevent cervical cancers. Consequently, the major burden of disease falls on the developing world, where cervical cancers will continue to occur with great frequency. The biochemical processes driving cervical cancer formation and progression are not yet fully elucidated despite intensive study and the primary causative agent being known. This has resulted in a lack of effective approaches for treating cervical cancers and therefore while cervical cancer is highly preventable, treatment success rates remain poor. Research focusing on the molecular underpinnings of the disease must be conducted so that targeted therapies can be developed to combat cancers arising in women lacking access to adequate healthcare or who are already infected with HPV.

1.1.1. Epidemiology

Globally, cervical cancer is the fourth most common female cancer and the fourth most common cause of female cancer death (Stewart & Wild, 2014). In 2012, an estimated 528 000 new cases (7.9% of all female cancers) and 266 000 (7.5%) deaths occurred due to cervical cancer (Ferlay *et al.*, 2015). The disease imposes a disproportionate burden on the developing world as it occurs most commonly in

developing nations (Plummer *et al.*, 2016). It is the second most common cancer in developing regions while only being the 11th most common in developed regions. Furthermore, 87% of deaths occur in less developed regions (Ferlay *et al.*, 2015). These disparities are largely due to the prevention and early detection of disease afforded by well-implemented screening procedures and medical care in developed nations or in nations with advanced healthcare infrastructure. In the United States of America, for instance, cervical cancer incidence has decreased by more than 80% between 1930-2012, largely due to improved screening programs (Siegel, Miller & Jemal, 2016). Similarly, the 5-year net survival rate for cervical cancer varies widely around the world. Across 61 countries for which data were available to calculate the age-standardized 5-year net survival of women diagnosed in 2005-2009, ≥70% survival was seen in the top four countries, 60-69% in 31 others, 50-59% in 26, and <50% in the bottom two. Developing countries are disproportionately prevalent in the lower half of the distribution of survival rates (Allemani *et al.*, 2015).

1.1.2. Anatomical origin and major types of cervical cancer

Cervical cancer most often originates at the transformation zone of the cervix, wherein the interface between the columnar epithelium of the endocervix and the squamous epithelium of the ectocervix lies (the squamo-columnar junction). While the cellular origin of cervical cancer remains unclear, a population of cuboidal stem-like cells have been identified in the junction; these may be more prone to malignant transformation following HPV infection relative to other cells (Herfs *et al.*, 2012).

Three major types of cervical cancer exist; squamous cell carcinomas, adenocarcinomas, and other undifferentiated carcinomas. Squamous cell carcinomas are the most common, accounting for 85-90% of all cases. Adenocarcinoma variants, of which the endocervical cell type is most common, constitute 10-15% of cases (Stewart & Wild, 2014).

1.1.3. Human Papilloma Viruses (HPVs) and oncogenesis

HPVs are a family of non-enveloped DNA viruses consisting of relatively small, double-stranded circular genomes (~8000 base pairs) encoding 8-9 open reading frames. While the total gene number is low, the number of encoded proteins is much greater due to transcription from multiple promoter regions as well as complex transcriptional splicing patterns. These genomes are encapsulated in icosahedral capsids consisting of 360 L1 capsid proteins arranged into 72 capsomeres, which assemble with one another and with a variable number of L2 capsid proteins (Doorbar *et al.*, 2016). It is the most common sexually transmitted infection; men and women in the USA have a 90% and 80% lifetime risk of contracting HPV at least once, respectively (Chesson *et al.*, 2014).

HPV infection is associated with (but is not always a necessary or causative agent) a number of other cancers such as penile, prostate, anal, vulvar, vaginal, and oropharyngeal carcinomas (Yang *et al.*, 2015; Plummer *et al.*, 2016). Over 150 different strains of HPV have been described and 13 high-risk subtypes (HPVs 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59 and 68) have been identified as being carcinogenic or possibly carcinogenic (IARC, 2012). Of the 14 million new cancer cases estimated to have been diagnosed globally in 2012, an estimated 2.2 million cases were caused by carcinogenic infections; HPV caused 640 000 of those. HPV is responsible for half of all female cancers that are due to viral infection *(Plummer et al.*, 2016). It is the necessary causative agent of cervical cancer; HPV DNA presence can be detected in essentially all cervical cancers. In the case of cervical cancers, HPV16 and -18 are the most carcinogenic subtypes, collectively causing ~70% of cervical cancer cases (IARC, 2012). However, the distribution of HPV strains detected in cervical carcinomas varies internationally. In the African context, HPV16 contributes slightly less and HPV45 slightly more relative to the rest of the world (Guan *et al.*, 2012).

While the majority of HPV infections are cleared by the immune system within two years, a small fraction of cases persist. Chronic cervical infection with HPV can lead

to the development of precancerous lesions, which in turn can progress to become invasive cervical cancer within two decades (IARC, 2012). During this process, the lesions follow a characteristic and well-defined progression from mild dysplasia to carcinoma (Figure 1.1)(Stewart & Wild, 2014). Briefly, HPV is introduced into the female reproductive tract through sexual intercourse with an infected partner. It is thought to travel through micro-abrasions in the cervical epithelium to reach the basal cell layer where it infects basal epithelial cells. The Viral DNA is then amplified episomatically using the host machinery. High-risk HPV types then rely on their E6 and E7 oncogenes to drive proliferation at the basal and parabasal layers. E7 proteins from high-risk subtypes are able to sequester and degrade Retinoblastoma (Rb) protein family members such as p107, which is responsible for regulating cell cycle entry at the basal epithelium, and p130, which regulates re-entry in the upper epithelial layers. Rb regulates entry from G1 into S phase by inhibiting E2F family members, thus effectively preventing aberrant mitosis (Giacinti & Giordano, 2006). E7 competes with E2F proteins for Rb protein binding and therefore prevents E2F transcription factor family proteins from being inhibited. E2F in turn activates a number of other genes needed for S-phase propagation and entry into the cell cycle. Normally, aberrant expression of these genes results in inactivation of MDM and subsequent p53 activation, which in turn results in cellular senescence or apoptosis. In high-risk HPV-infected cells, however, oncogenic E6 proteins associate with host machinery to ubiquitinate p53, which leads to its degradation and thus prevents it from countering the E7-induced proliferation (reviewed in Doorbar *et al.*, 2016). As the cells proliferate, the normal cervical tissue is replaced by cervical intraepithelial neoplasia (CIN), which is graded into three grades. During this time the viral DNA integrates with the host genome and E6 and E7 expression increases with lesion severity. At its most extreme (CIN3 or carcinoma *in situ*), neoplastic cells eventually disrupt the epithelial basement membrane at which point a diagnosis of invasive carcinoma is made (Stewart & Wild, 2014; Doorbar *et al.*, 2016). Recently, a subset of cervical cancers in which HPV mRNA transcripts were absent despite viral DNA being present in the genome was identified. These cancers are capable of sustaining themselves in the absence of viral E6 and E7 gene expression and exhibit marked differences in global gene expression compared to cancers where viral transcripts

are detected. HPV is therefore necessary for oncogenic initiation, but a small subset of cancers may evolve to become HPV independent over time (Banister *et al.*, 2017).

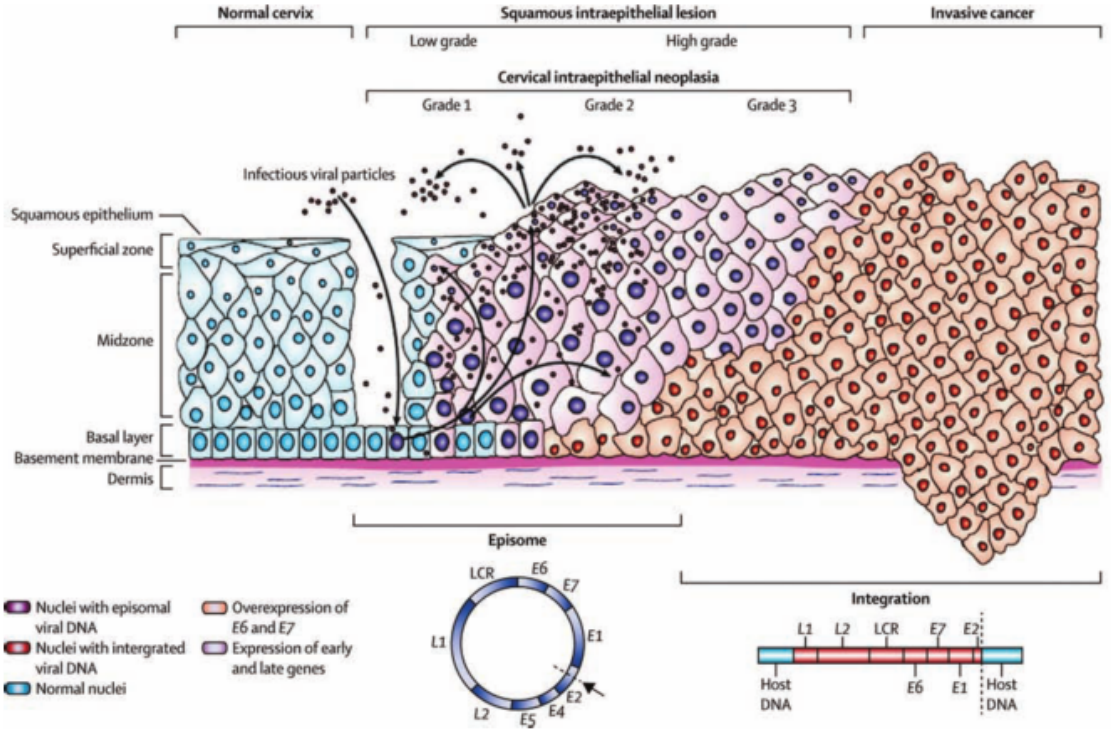

Figure 1.1: Human papillomavirus-mediated progression to cervical cancer.
LCR = Long Control Region. Figure from Stewart & Wild (2014).

1.1.4. Other risk factors

While HPV infection is a necessary cause for cervical carcinogenesis, it is not sufficient. Over 90% of infections spontaneously regress 1.5 years post-infection (Castle *et al.*, 2009). It is unclear why regression fails to occur in the remaining minority of cases, but the probability of clearance appears to be related to the duration of the infection (Rodriguez *et al.*, 2010). It is apparent, however, that other factors are required for carcinogenesis to occur.

A number of factors that increase the risk of cervical cancer have been proposed or identified. Co-infection with the Human Immunodeficiency Virus (HIV) is the greatest risk factor to date. Like most other infection-related cancers, cervical cancer is

significantly more common in HIV+ individuals relative to the general population (Grulich *et al.*, 2007). Immunodeficiency caused by HIV infection seems to be the likely cause. Firstly, because immunosuppressed transplant patients also have a greater risk of developing cervical cancer (Grulich *et al.*, 2007). Secondly, African HIV+ women with cervical cancer exhibit different distributions of high-risk HPV strains relative to HIV- African cervical cancer cases. A meta-analysis by Clifford *et al.* (2016) found that while the rate of HPV16 and -18 detection is similar across the two groups, HIV+ patients exhibit significantly fewer single infections with HPV16, while co-infections of HPV16 and other high risk HPVs is more likely. Conversely, HPV18 single and co-infections are more common in HIV+ women, and all other high-risk types were more common as part of co-infections in HIV+ women. Owing to the fact the HPV16 prevalence is less affected by CD4+ cell counts (Strickler *et al.*, 2003), the authors suggest that HPV16 has adapted more efficient mechanisms to evade the immune response, thus it acquires a lesser benefit from a compromised immune system relative to the other high risk strains that do show an increase in prevalence in immunocompromised women (Clifford *et al.*, 2016).

Some behavioural risk factors have been identified, but their relative impacts are much weaker than the risk of persistent HPV infection or HIV co-infection. Some correlate to increased probability of HPV acquisition, such as number of sexual partners (Munoz *et al.*, 2002) or oral contraceptive use (which implies sexual activity) (Moreno *et al.*, 2002), or the possible duration of infection (age at first intercourse) (Veldhuijzen *et al.*, 2010; Plummer, Peto & Franceschi, 2012; Stewart & Wild, 2014). Smoking has also been very weakly correlated (Schiffman *et al.*, 1987; Moreno *et al.*, 1995).

Limited and contradictory evidence regarding the existence of genetic predispositions to developing cervical cancer exists. Genome-wide association studies on Chinese Han populations have identified genetic risk factors at 3 loci; 4q12, 17q12 and 6p21.32 (Shi *et al.*, 2013). Conversely, while Nordic twin studies have identified relatively excessive familial risk factors for other cancers of the female reproductive tract, no such risk was apparent in cervical cancer (Mucci *et al.*,

2016). Given the paucity of information available, it is unclear whether genetic predispositions play a significant role in all populations.

1.2. T-Box transcription factors

An ancient family of transcription factors all sharing a common DNA-binding domain named the T-box have been identified across a wide spectrum of eukaryotic organisms (Sebé-Pedrós & Ruiz-Trillo, 2017). Called the T-box family, each T-box transcription factor (TBX) binds to a canonical, palindromic core DNA sequence *in vitro* (T(G/C)ACACCT AGGTGTGAAATT), also called the T-element (Kavka & Green, 1997). While it is generally found as a palindrome, functional half sites (TCACACCT) also exist (Fernando *et al.*, 2010). Furthermore, T-box factors may also bind to variants thereof *in vivo* (Kispert & Herrmann, 1993; Sinha *et al.*, 2000; Bruneau *et al.*, 2001; Lingbeek, Jacobs & Van Lohuizen, 2002; Paxton *et al.*, 2002). This sequence variation gives rise to variance in the target genes regulated by each T-box family member (Papaioannou, 2014), possibly in conjunction with context-specific interactions with cofactors (Lu *et al.*, 2010).

More than 20 T-box family members have been discovered in humans. Their expression patterns are spatially and temporally regulated throughout human development (Packham & Brook, 2003). A full review of their developmental roles is beyond the scope of this thesis; for comprehensive reviews the reader is referred to Papaioannaou (2014) and Sheeba & Logan (2017). The developmental roles of the T-box factors of interest to the present study, TBX2 and TBX3, will be briefly discussed in a later section (see section 1.3.2).

1.3. TBX2 and TBX3

1.3.1. Gene and protein structures

TBX2 and TBX3 are part of the Tbx2 subfamily of T-box factors. Comprising *TBX2, TBX3, TBX4,* and *TBX5,* this subfamily is hypothesized to originate from a single

ancestral gene, with unequal crossing over events resulting in two clusters, *Tbx2/3* and *Tbx4/5*, followed by a second event producing the four separate genes. In humans, *TBX2* and *TBX4* are situated on chromosome 17q23, while *TBX3* and *TBX5* reside on 12q24 (Campbell *et al.*, 1995; Agulnik *et al.*, 1996; Bamshad *et al.*, 1997). Despite their relatively distant chromosomal locations, *TBX2* and *TBX3* share the highest degree of homology in humans, with 95% similarity in their DNA-binding domains (Law *et al.*, 1995). Both proteins are highly conserved in vertebrates, with TBX2 and TBX3 being 96% and 98% homologous to their respective mouse proteins (Law et al., 1995; Bamshad *et al.*, 1997).

In humans, the TBX2 coding region spans 7 exons encoding 712 amino acids (Figure 1.2). The T-box DNA binding domain spans amino acids 104-285 in the N-terminal half of the protein. An arginine at position 122 has been shown to be essential for DNA binding (Sinha *et al.*, 2000). TBX2 can bind palindromic and half site T-elements as a monomer and acts predominantly as a transcriptional repressor (Carreira *et al.*, 1998; Sinha *et al.*, 2000). It contains two repression domains as well as a weak activation domain (see Figure 1.2). While gene activation has been observed *in vitro* (Paxton *et al*., 2002; Sakabe *et al*., 2012), to date no direct TBX2-mediated activation activity has been observed *in vivo*.

TBX3 encodes two isoforms, TBX3 and TBX3+2a. The transcripts span seven and eight exons and encode 723 and 743 amino acid proteins, respectively. TBX3+2a contains an additional 60bp sequence in the middle of the T-box that adds 20 amino to the resultant protein (see Figure 1.2)(Bamshad *et al.*, 1997, 1999; Fan *et al.*, 2004). Crystallisation studies performed by Coll *et al.* (2002) have shown that TBX3 binds the consensus sequence as two monomers, each recognising one palindromic sequence. Owing to how it interacts with the DNA in the TBX3-DNA crystal structure, the authors speculate that it can also naturally bind target genes as a monomer through the half-site of the palindromic sequence**.** Like TBX2, it also possesses two repression domains and an activation domain (Carlson *et al.*, 2001). While TBX3 predominantly acts as a repressor of transcription, recent findings have shown functional activity of the activation domain *in vivo* (Willmer, Cooper, *et al.*, 2016).

Whether the two TBX3 isoforms are functionally distinct remains controversial. Fan *et al.* (2004) found that they are differentially expressed in a subset of breast cancer cell lines, and that the ratio between them also differs by cell line. Their work also showed that the two were variably expressed in mature adult human and mouse tissues and that TBX3 can induce immortalisation of mouse embryo fibroblasts, whereas TBX3+2a accelerated senescence and could not bind DNA in their study. Both isoforms are highly expressed in mouse embryonic stem cells (mESCs) and induced pluripotent stem cells, where overexpression of either can induce mESC differentiation. However, while both isoforms repressed Nanog transcription, only TBX3+2a could do so directly (Zhao, Wu & Chen, 2014). Conversely, work by Hoogaars *et al.* (2008) indicates functional similarity in terms of binding the consensus T-element and as repressors of $p21^{CIP1/WAF1}$ (hereafter referred to as p21) and Nppa promoter activity. Furthermore, the authors showed that ectopic expression of either isoform repressed chamber formation in embryonic mouse hearts. Both isoforms behave similarly to promote cancer progression in a model of early breast cancer, and both behave similarly as tumour suppressors in fibrosarcoma cells (Krstic et al., 2016; Willmer, Cooper, *et al.*, 2016). Taken together, these results imply that other context-dependent factors may be required to differentiate the functions of the two isoforms.

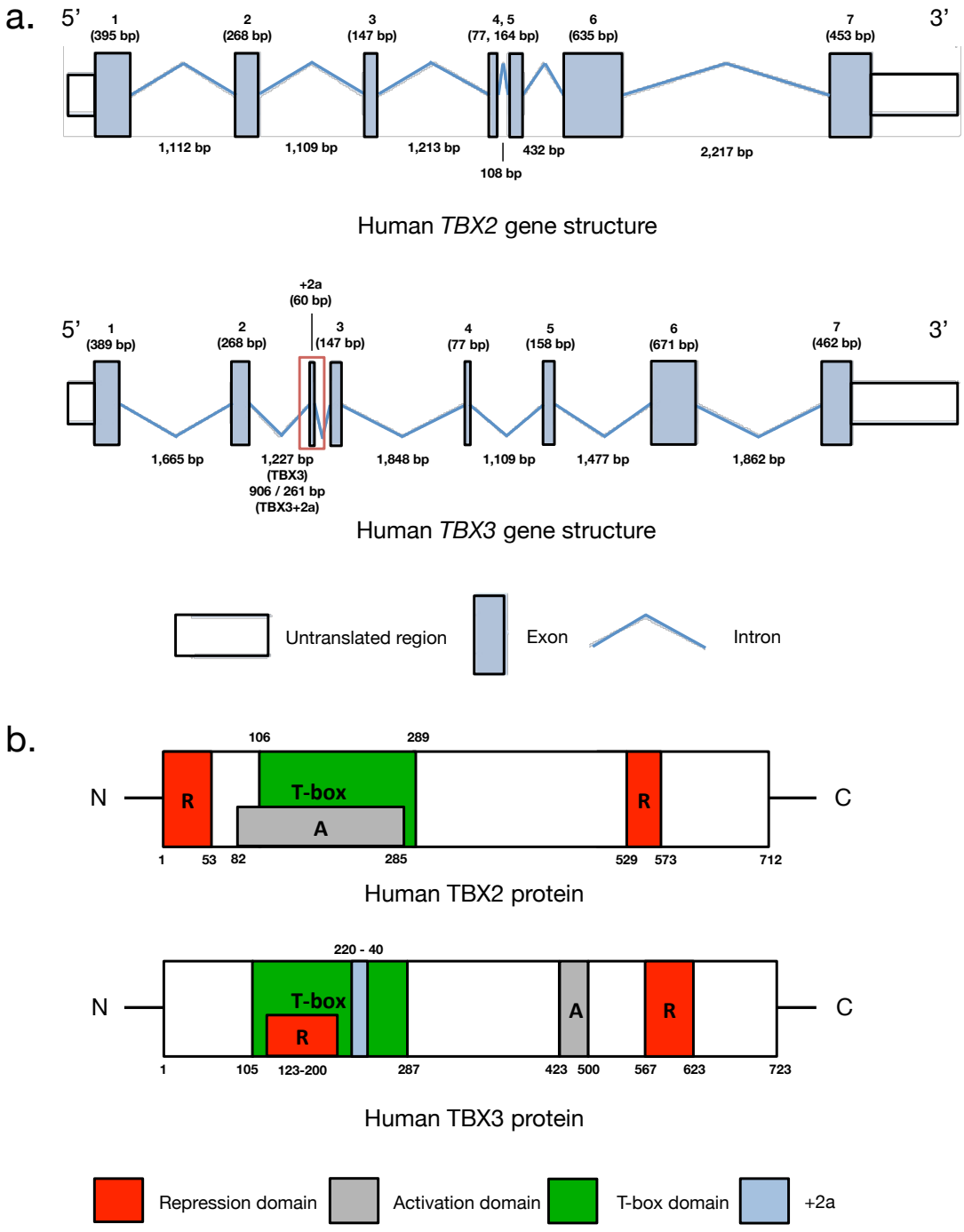

Figure 1.2: Structure of human TBX2 and TBX3 genes and proteins.
a. The human TBX2 and TBX3 mRNA genes. Exon and intron lengths are indicated. The additional 60bp +2a segment found in the TBX3+2a isoform is boxed in red. Information was obtained and figures were adapted from Ensembl.org for transcripts ENST00000240328.3 (TBX2), ENST00000349155.6 (TBX3) and ENST00000257566.7 (TBX3+2a). **b.** Diagrams of the human TBX2 and TBX3 proteins depicting identified domain regions. Amino acid positions are indicated at boundaries between regions. The additional 20 amino acids in TBX3+2a are included in the TBX3 protein at the insert position (220-240) and indicated in blue, but subsequent amino acid position numberings reflect the positions of regions as found in the TBX3 isoform. All regions and domains are drawn to scale. Domain information sources are referenced in-text. Figure design adapted from Willmer (2016).

1.3.2. Role in development

TBX2 plays a significant role in limb, heart, eye, kidney, lung, melanocyte and mammary gland development (Papaioannou, 2014; Sheeba & Logan, 2017). Murine homozygous Tbx2 mutants spontaneously abort mid-gestation due to heart abnormalities while heterozygous mutants exhibit normal phenotypes *(Harrelson et al*., 2004). While no human disorders have been associated with TBX2, studies have associated microdeletions on chromosome 17q23 (where *TBX2* and *TBX4* reside) with an as yet unnamed syndrome that commonly features developmental retardation, microcephaly, heart, hand, foot and limb abnormalities as well as delayed postnatal growth (Ballif *et al.*, 2010; Nimmakayalu *et al.*, 2011; Willmer, 2016). TBX2 may therefore be responsible for this syndrome, either independently or in co-operation with TBX4. TBX2 is widely expressed in adult human tissues but no clear role has been defined (Campbell *et al.*, 1995; Law *et al.*, 1995).

Like TBX2, TBX3 also plays crucial but distinct roles in limb development (Sheeba & Logan, 2017). It is the first T-box gene to be expressed in mammalian development and is important in maintaining pluripotency in mouse embryonic stem cells (Russell *et al*., 2015). It has also been shown to promote neuroepithelial differentiation of human embryonic stem cells, but not endoderm development as it can also do in mice (Esmailpour & Huang, 2012). TBX3 acts during cardiac development, particularly in the development of the atrioventricular canal, the ventricular septum and the cardiac conduction system (Davenport, Jerome-Majewska & Papaioannou, 2003; Hoogaars *et al.*, 2007; Bakker *et al.*, 2012). It also plays a significant role in mammary gland development, as evidenced by the higher incidence of deformities and aplasia observed in murine Tbx3 heterozygous mutants (Davenport, Jerome-Majewska & Papaioannou, 2003). In humans, heterozygous mutation of TBX3 results in an autosomal dominant development disorder called ulnar mammary syndrome (UMS). Sufferers of UMS exhibit abnormalities of the forelimb, apocrine and mammary gland, areola, dental structures and axillary hair development as well as defects of the heart, jaw and genitalia (Wilson & Conlon, 2002; Bamshad *et al.*, 1997; Tada & Smith, 2001; Meneghini *et al.*, 2006). One study found that mutations in the

T-box domain are more likely to cause severe limb defects, indicating the importance of its role as a transcription factor in the aetiology of UMS (Meneghini *et al.*, 2006). Recently, however, TBX3 has been shown to also act as a regulator of alternative splicing of mRNA; TBX3 mutants found in UMS were unable to direct alternative splicing to varying degrees, either due to disruption of the recruitment of RNA-binding proteins and splicing factors or by disrupting direct interaction between TBX3 and mRNA targets (Kumar *et al*., 2014). Disruption of this novel role for TBX3 therefore potentially also contributes to development of the UMS phenotype.

1.3.3. Role in cancer

The roles of TBX2 and TBX3 in cancer have been studied extensively. Both TBX2 and TBX3 have been shown to play distinct, context dependent roles in the development and progression of a number of different cancers. Despite the high degree of structural homology between the two, their roles in oncogenesis often do not overlap and both may play different roles depending on the cancer context. Furthermore, evidence is accumulating to suggest that both TBX2 and TBX3 can act as tumour suppressors in certain contexts. A comprehensive review of all of their oncogenic activities is beyond the scope of this work, and as such only areas of particular relevance are highlighted. For comprehensive review, the reader is directed to Wansleben *et al.* (2014) and Change *et al.* (2016).

TBX2

TBX2 is overexpressed in a wide range of cancers including breast, melanoma, liver, lung, ovarian, cervical, colorectal, endometrial, and pancreatic cancers as well as the soft-tissue cancer, rhabdomyosarcoma (Carreira, Liu & Goding, 2000; Mahlamäki *et al.*, 2002; Vance *et al.*, 2005; Dimova *et al.*, 2009; Liu, Jiang & Zhang, 2010a,b; Wang *et al.*, 2012; Han *et al.*, 2013; Zhu *et al.*, 2014, 2016; Hu *et al.*, 2014; Yu *et al.*, 2015). TBX2 plays multiple roles in cancer; it allows for bypass of senescence and apoptosis, promotes proliferation and survival and induces changes related to epithelial-to-mesenchymal transition (Wansleben *et al.*, 2014). Furthermore, TBX2 overexpression

has been shown to confer resistance to platinum based chemotherapeutics (Davis *et al.*, 2008; Wansleben *et al.*, 2013).

TBX2's significant pro-proliferative / anti-senescence roles are mediated by inactivation of the $p14^{ARF}$–p53–p21 pathway. TBX2 represses $p19^{ARF}$ ($p14^{ARF}$ in humans), a gene encoded by an alternate reading frame in the *CDKN2A* locus (Jacobs *et al.*, 2000). When not repressed, $p14^{ARF}$ inhibits the ubiquitin ligase MDM2, which in turn results in increased p53 levels. p53 in turn directly increases expression of its target, p21, which is a potent tumour suppressor and the primary mediator of p53 tumour suppressor activity. p21 can, however, also be upregulated by a number of p53-independent pathways have been identified in recent years. p21 acts as a cyclin-dependent kinase inhibitor, but is also capable of directly regulating gene expression and its tumour suppressor activity is the result of its ability to induce growth arrest, differentiation or senescence (Abbas & Dutta, 2009). Given that the p53 pathway is often disabled in cancers, the influence of TBX2 on $p14^{ARF}$ expression is often unnecessary or could be bypassed by one of the alternate p21 pathways. However, TBX2 is also capable of directly repressing p21 and can therefore prevent senescence even in cancers with non-functional p53 pathways (Prince *et al.*, 2004; Dobrzycka *et al.*, 2006). Indeed, silencing Tbx2 expression in *CDKN2A*-null B16 melanoma cells induces senescence, which is accompanied by increased expression of p21 (Vance *et al.*, 2005). Similarly, TBX2 mRNA expression was upregulated, and p21 downregulated in an inversely correlative manner, in a study on laryngeal squamous cell carcinomas relative to adjacent normal tissues. Patients with high TBX2 / low p21 expressing cancers had a significantly lower chance of survival (Huang *et al.*, 2014). TBX2 also reduces p21 expression in rhabdomyosarcomas where it functions to promote proliferation (Zhu *et al.*, 2014). p21 repression is enhanced by phosphorylation of TBX2 in response to activation of a stress-induced senescence pathway in MCF-7 breast cancer cells (Abrahams *et al.*, 2008). Furthermore, shRNA knock-down of TBX2 or TBX3 in two vertical growth phase melanoma cell lines showed that TBX2 plays a predominantly pro-proliferative and transformative role in a mechanism involving p21 repression – a finding that was reproducible in MCF7 breast cancer cells (Peres *et al.*, 2010). Despite its clear influence on the $p14^{ARF}$–p53–

p21 pathway, TBX2 is also able to promote anchorage independent growth and increased resistance to apoptotic stimuli in and adrenocortical carcinoma cell line that is deficient in p53 and p21. This indicates that TBX2 is also capable of exerting oncogenic effects beyond this pathway alone (Ismail & Bateman, 2009).

In the context of cervical cancer, TBX2 protein was found to be overexpressed in squamous cell carcinomas relative to adjacent normal tissues, and its expression correlated with HPV16 E7 protein presence and lymph node metastasis, but not with patient age, histological grade or invasiveness (Liu, Jiang & Zhang, 2010b). TBX2 may also be involved in the regulation of the HPV lifecycle; TBX2 (and TBX3) protein interacts with the minor L2 capsid protein of HPV16 to enhance TBX2-mediated repression of the long control region of the integrated viral genome, thus suggesting that it may also contribute to the regulation of oncovirus-mediated carcinogenesis (Schneider *et al.*, 2013).

Interestingly, a protective role for TBX2 has been suggested in recent literature on endometrial carcinomas and bladder cancers. Researchers examined the methylation status of the TBX2 gene in early and late stage endometrial carcinomas and found that early carcinomas expressed relative hypomethylation of TBX2, while late stage carcinomas expressed hypermethylation. Furthermore, patients displaying hypomethylation of the gene had significantly greater 5-year survival rates than those exhibiting hypermethylation (Farkas, Sorbe & Nilsson, 2016). Similarly, TBX2 methylation was also identified as a strong predictor of bladder cancer progression in two independent studies, with low methylation status indicating significantly greater 10-year progression-free survival rates (Kandimalla *et al.*, 2012; Beukers *et al.*, 2015). This suggests that TBX2 may act protectively in these cancers and that silencing its expression is required for them to progress to later stages. A full characterization of the tumour-suppressor activities of TBX2 has not yet been reported; it may be that context-dependent cofactors play a role in determining whether TBX2 acts as an oncogene or a tumour suppressor.

TBX3

TBX3 is overexpressed in a number of primary cancer tissues, including head and neck squamous carcinomas, pancreatic ductal carcinomas and solid-pseudopapillary pancreatic neoplasms, gastric and colorectal cancers, a diverse range of sarcomas, a subset of breast cancers, and Wnt-activated hepatocellular carcinomas and hepatoblastomas (Fan *et al.*, 2004; Renard *et al.*, 2007; Yarosh *et al*., 2008; Cavard *et al.*, 2009; Burgucu *et al.*, 2012; Shan *et al.*, 2015; Wang *et al.*, 2015; Miao *et al.*, 2016; Perkhofer *et al.*, 2016; Willmer, 2016; Willmer, Cooper, *et al.*, 2016). Elevated expression was also seen in plasma obtained from a subset of ovarian and breast cancer patients relative to healthy controls (Lomnytska *et al.*, 2006). TBX3 has multiple roles in cancer and has been shown to influence cell proliferation, migration and metastasis as well as bypass of senescence and apoptosis (Wansleben *et al*., 2014).

Like TBX2, TBX3 also plays a role in melanoma. However, it does not seem to play a pro-proliferative role and instead appears to be necessary for tumour formation and migration in these cells (Peres *et al.*, 2010). Indeed, TBX2 and TBX3 appear to play opposing roles during different growth phases of melanoma. Some of TBX3's effects also occur via p21. When stably-transfected to express TBX3, the non-tumorigenic radial growth phase cell line WM1650 showed decreased proliferative rates, an alteration of cell cycle distribution, decreased levels of p21, TBX2 and E-cadherin, and increased levels of p53 and $p14^{ARF}$. Transfected cells also exhibited increased migratory ability and anchorage independence and tumour forming ability *in vitro* and *in vivo* (Peres & Prince, 2013). Furthermore, both TBX3 isoforms are expressed in 501 mel and YUSIT melanoma cell lines, but not in the melanocyte lines melan-a and –c, and both isoforms were capable of repressing p21 in a functionally equivalent manner (Hoogaars *et al.*, 2008). While the mechanism of this repression was not determined, TBX3 has recently been shown to be able to directly bind and suppress p21 in chondrosarcomas by Willmer *et al*. (2016). However, in this context, TBX3-mediated repression of p21 did result in increased rates of proliferation.

Reasons for the different context-dependent physiological outcomes caused by TBX3-mediated repression of p21 remain to be elucidated.

TBX3's pro-migratory role may be mediated by its ability to repress E-cadherin. TBX3 expression is induced by oncogenic B-RAFV600E activity in melanoma cells to repress E-cadherin expression (Boyd *et al.*, 2013). E-cadherin is also directly repressed by TBX3 in this context and TBX3 depletion led to increased E-cadherin expression and decreased invasiveness in *in vitro* studies. Depletion of TBX3 also led to substantial increases in TBX2 expression, further supporting the concept of opposing roles for TBX2 and TBX3 in melanoma proliferation (Rodriguez *et al.*, 2008). The AKT3 pathway is an important regulator of melanomagenesis and TBX3 has been shown to be a key substrate of this pathway in this context; TBX3 phosphorylation by AKT3 increases its stability, enhances nuclear localization and the repression of E-cadherin expression both *in vitro* and *in vivo,* leading to enhanced cell migration and invasion of melanoma cells (Peres, Mowla & Prince, 2014). Similar effects of TBX3 in proliferation and migration were also observed in MCF-7 breast cancer cells as well as in breast epithelial cells (Peres *et al.*, 2010; Li, Weinberg, *et al.*, 2013). Furthermore, TBX3 may reinforce its own AKT-pathway driven phosphorylation by repressing expression of tumour suppressor phosphatase and tensin homolog (PTEN) (Burgucu *et al.*, 2012). PTEN antagonizes phosphoinositide 3-kinase (PI3K) signaling to inactivate it and the downstream AKT pathway (Leevers, Vanhaesebroeck & Waterfield, 1999). Inactivation of PTEN results in subsequent activation of the PI3K/AKT pathway, leading to increased proliferation, migration and survival (Li & Sun, 1998; Chung & Eng, 2005). It may be that TBX3 acts as the effector of some of these consequences while simultaneously reinforcing them.

TBX3 also acts as a tumour suppressor in some instances. Like TBX2, TBX3 also interacts with the HPV16 L2 capsid protein to enhance viral gene repression (Schneider *et al.*, 2013). However, unlike TBX2, TBX3 expression in cervical cancers appears to be onco-protective. TBX3 mRNA was downregulated in node-positive cervical cancer tumours and this downregulation was statistically associated with metastatic phenotypes. High expression levels were linked to better progression-free

survival and decreased tumour volume (Lyng *et al.*, 2006). Similarly, TBX3 has been shown to act as a tumour suppressor in fibrosarcoma cells both *in vitro* and *in vivo*; both isoforms are capable of inducing p21 expression and overexpression of either results in decreased cancer cell proliferation, migration and anchorage independent growth in these cells (Willmer, Cooper, *et al.*, 2016). Finally, *TBX3* is also highly methylated in glioblastoma, gastric and bladder cancers. Increased methylation negatively correlated with survival in glioblastomas and in bladder cancers was associated with a greater likelihood to progress to muscle-invasive tumours and significantly reduced survival rates (Yamashita *et al.*, 2006; Etcheverry *et al.*, 2010; Kandimalla *et al.*, 2012; Beukers *et al.*, 2015). Interestingly, *in silico* database screens and cell culture assessments found that TBX3 is more likely to be mutated in breast cancers relative to other TBX genes and that frequently-occurring mutants generally exhibit reduced or abrogated protein function (Fischer & Pflugfelder, 2015). Collectively, it is clear that a tumour suppressor role is possible for TBX3 in some cancers; the mechanisms and contexts involved in this activity remain to be determined.

1.4. Other T-box transcription factors in cancer

While TBX2 and TBX3 are most significantly associated with oncogenic processes in the literature, a number of other T-box factors have also been implicated in cancer development.

TBX1 and TBX4 have been implicated as potential tumour suppressors. TBX1 expression is reduced in mouse skin tumours and when overexpressed in mouse spindle carcinoma cells it conveys oncoprotective effects (Trempus *et al.*, 2011). TBX4 has been suggested as a potential biomarker for pancreatic ductal carcinomas where it predicts good prognostic outcomes (Qi *et al.*, 2008; Zong, Meng & Li, 2011). TBX15 and T-bet (TBX21), on the other hand, appear to have oncogenic potential. Thyroid cancer cell lines exhibit altered TBX15 expression relative to normal cells, and it appears to play an anti-apoptotic role in these cancers (Arribas *et al.*, 2015). T-bet acts as an oncogene when expressed in oestrogen-dependent breast cancers and

its presence in primary tumours was suggested as a predictive marker of hormonal therapy-resistant cancer phenotypes (McCune *et al.*, 2010).

Like TBX2 and TBX3, other T-box factors have been implicated as either oncogenes or tumour suppressors depending on the cellular context. Brachyury expression is upregulated in a number of epithelial-derived cancers relative to normal tissues and its expression appears to promote epithelial-to-mesenchymal transition-related cancer phenotypes (Palena *et al.*, 2007; Imajyo *et al.*, 2012; Sarkar *et al.*, 2012; Shimoda *et al.*, 2012). Furthermore, Brachyury expression was correlated to tumour stage in human lung tumour tissues, indicating a biological role in cancer progression (Fernando *et al.*, 2010). Brachyury also promoted resistance to chemotherapy and radiotherapy in lung carcinoma models (Huang *et al.*, 2013). In contrast, Brachyury has also shown potential as a tumour suppressor in primary non-small lung cell carcinomas; preliminary research showed that Brachyury mRNA expression levels were decreased relative to normal tissue levels (Park *et al.*, 2008). TBX5 also seems to play contradictory roles in different cancer contexts. Ectopic expression of TBX5 inhibits colony formation and increases the rate of apoptosis in osteosarcoma and lung carcinoma cell lines (He *et al.*, 2002). Additionally, TBX5 is silenced or downregulated in a number of colon cancer cell lines relative to normal cells, and TBX5 methylation was detected in 68% of primary colon tumours tested and was associated with decreased overall survival (Yu *et al.*, 2010). However, other research has shown that TBX5 is essential in the transformation and survival of β-catenin driven cancer cell lines, including a number of colon cancers (Rosenbluh *et al.*, 2012).

Finally, T-bet and Eomes may indirectly affect cancer progression by influencing immune responses. Colorectal cancers lacking signs of early metastatic invasion exhibit less T-bet mRNA expression relative to tumours that do show such signs (Pagès *et al*., 2005). Given that T-bet is required for naive CD4+ T cell differentiation into type 1 helper T cells in response to Interferon Gamma (IFN-γ) stimulation (Szabo *et al.*, 2003), it seems likely that T-bet may be a marker of an effective immune response targeting these tumours. Reduced Eomes expression was correlated with lymph node metastasis in a study on colorectal cancer tissues and Eomes expression

in CD8+ T cells regulates IFN-γ expression, perforin levels and cytolitic activity of CD8+ T cells. Low-expression of Eomes may therefore result in a poor immune response to colorectal, and possibly other, cancers (Atreya *et al.*, 2007). Eomes and T-bet also collectively decide CD8+ T cell fate and promote and maintain effector and central memory CD8+ T cell phenotypes in mice (Li, Yang, *et al.*, 2013). Taken together, these data suggest that the balance between T-bet and Eomes in immune cells may influence their ability to successfully combat colorectal and possibly other cancers.

1.5. Seminal fluid

Semen can be separated into two distinct components; sperm, the male gamete responsible for transmission of genetic material from the paternal parent into an ovum, and seminal fluid (SF, also called seminal plasma), the liquid within which the sperm are deposited into the female reproductive tract. Once thought to simply act as an ideal environment for sperm survival during transport and fertilization, accumulating evidence indicates that SF also contributes to a successful pregnancy by immunomodulation of the female reproductive tract tissues and immune system in a range of mammals, including humans (Schjenken & Robertson, 2014). At the molecular level, the effects of SF on mammalian pregnancy appear to be largely related to immunomodulation (Bromfield, 2014; Robertson & Sharkey, 2016). Furthermore, evidence exists to indicate that SF could also potentially influence carcinogenic tissues in the female reproductive tract.

1.5.1. Origins and constituents

Several glands in the male reproductive tract contribute specific factors to the production of SF (see Figure 1.3); the resultant fluid is a complex mixture of biomolecules, from simple sugars to complex macromolecules (see Table 1-1). The unique peptides and proteins alone, as identified my mass spectrometry, number in the hundreds to thousands (Fung *et al.*, 2004; Pilch & Mann, 2006; Batruch *et al.*, 2011). Furthermore, the concentrations of various constituents varies from ejaculate

to ejaculate, both intra- and inter-personally (see Table 1-1) (Sharkey *et al.*, 2016). While many constituents certainly play a role in the maintenance and protection of sperm, some also elicit signaling responses in the female reproductive tracts of various mammals, including humans (Schjenken & Robertson, 2014; Robertson & Sharkey, 2016).

Finally, SF also contains numerous types of extracellular vesicles (Machtinger, Laurent & Baccarelli, 2016). While specific roles for these vesicles have not been defined, exosomes within SF have been shown to contain a vast number of small non-coding RNA, including microRNAs, some of which are unique to SF (Weber *et al.*, 2010; Vojtech *et al.*, 2014).

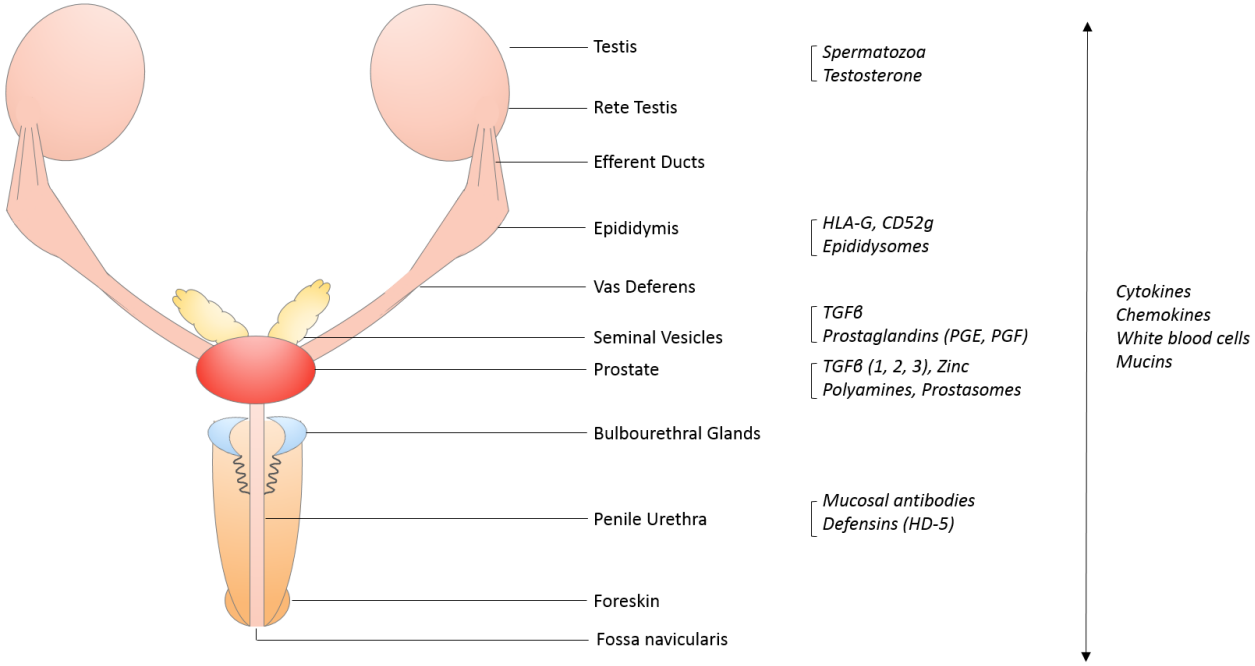

Figure 1.3: A diagrammatic representation of the male reproductive tract.
The origins of notable constituents of seminal fluid are annotated. Figure adapted from Hale *et al.* (2015); redrawn by M Blumenthal.

Table 1-1: Selected constituents of seminal fluid from healthy males.

Descriptive statistics of semen analysis variables of fertile men ($n = 83$)

Variable	Percentiles			Range	Mean ± SE
	25th	50th (Median)	75th		
Semen volume (ml)	2.3	2.9	4.2	0.9–10.5	3.4 ± 0.2
Sperm concentration (10^6/ml)	35.9	48.0	73.6	5.4–281.3	64.4 ± 6.1
Total sperm number (10^6)	99.5	169.4	236.0	13.3–1400.3	198.8 ± 19.9
Motility (%)	41.3	55.0	64.0	18.0–83.0	52.3 + 1.7
Total motile sperm number (10^6)	48.8	80.5	145.5	4.2–420.1	104.3 ± 9.1

Descriptive statistics of cytokines in seminal plasma of fertile men

Variable[a]	*n*	% Positive	Percentiles			Range	Geometric mean (95% CI)[b]
			25th	50th (Median)	75th		
IL-1α	81	68	ND[c]	6.0	16.0	ND-214.0	4.7 (3.2–7.0)
IL-1β	83	54	ND	2.0	8.0	ND-118.0	2.3 (1.6–3.3)
IL-2	77	1	ND	ND	ND	ND-24.0	3.6 (3.4–3.8)
IL-5	19	100	18.0	31.3	80.2	11.0-227.0	39.7 (26.2–60.3)
IL-6	79	99	2.0	6.0	14.0	ND-110.0	6.4 (4.9–8.3)
IL-7	19	100	1480.9	2532.6	3380.7	1109.5–3985.5	2365.8 (1929.5–2901.5)
IL-10	80	9	ND	ND	ND	ND-32.0	2.4 (2.1–2.8)
IL-12	80	13	ND	ND	ND	ND-18.0	3.1 (2.7–3.4)
IL-13	19	53	ND	3.1	5.7	ND-149.3	3.2 (1.6–6.6)
IL-17	19	74	ND	11.6	19.8	ND-84.9	7.1 (3.5–14.6)
TNF-α	59	20	ND	ND	ND	ND-40.3	1.5 (1.2–1.9)
IFN-γ	82	37	ND	ND	10	ND-130.0	3.7 (2.8–5.0)
IFN-α	60	55	ND	3.3	23.3	ND-270.8	15.6 (11.9–20.3)
TGF-β1 (Total)	77	100	68 040.0	85 120.0	112 420.0	37 520.0–192 640.0	85 548.0 (79 063.0–92 595.9
TGF-β1 (Latent)	63	100	65 326.0	80 118.0	109 089.0	36 840.0–190 230.0	81 221.8 (74 309.9–88 787.4
TGF-β1 (Active)	75	100	680.5	1096.0	2072.0	168.0–9214.0	1068.9 (881.8–1296.0)
G-CSF	60	87	25.6	64.7	132.7	ND-5262.6	47.5 (28.8–78.6)
GM-CSF	60	7	ND	ND	ND	ND-1190.6	1.5 (1.0–2.2)

[a]Cytokines in pg/ml; [b]Geometric mean (95% CI); see text for description; [c]Not detectable. TNF, tumor necrosis factor; G-CSF, granulocyte colony-stimulating factor; GM-CSF, granulocyte–macrophage CSF.

Descriptive statistics of chemokines in seminal plasma of fertile men

Variable[a]	*n*	% Positive	Percentiles			Range	Geometric mean (95% CI)[b]
			25th	50th (Median)	75th		
IL-8	82	100	850.0	1305.0	2780.0	384.0–14 712.0	1583.3 (1311.6–1910.3)
MIP-1α	82	15	ND[c]	ND	ND	ND-136.0	6.9 (5.8–8.3)
MIP-1β	83	81	26.5	66.0	157.0	ND-1050.0	54.1 (39.8–73.4)
RANTES	83	98	93.0	126.0	167.5	ND-1480.0	119.9 (97.6–147.4)
SDF-1α	59	98	3870.9	5741.9	8691.5	ND-17 955.4	5087.7 (3928.4–6588.0)
MCP-1	60	100	1362.9	2980.9	6510.9	292.4–81 516.1	3251.5 (2319.3–4555.1)

[a]Chemokines in pg/ml; [b]Geometric mean (95% CI): see text for description; [c]Not detectable MCP, monocyte chemotactic/chemoattractant protein.

Adapted from Politch et al. (2007).

1.5.2. Role in human pregnancy

The degree to which SF contributes to a successful human pregnancy remains somewhat controversial in the literature (Bedford, 2015; Robertson & Sharkey, 2016). Given the countless successful births that have occurred via assisted reproductive therapies (ARTs) in the absence of SF, it is clear that it is not an essential factor in human childbirth. Furthermore, the difficulties of studying these events in women prevents confirmation of their physiological significance or the

mechanisms by which they may occur in humans (Bedford, 2015; Hale *et al.*, 2015; Robertson & Sharkey, 2016). There are, however, a number of lines of evidence to indicate that it may contribute to various aspects of pregnancy, including success in pregnancy, the health of offspring and immunotolerance.

Success in pregnancy

The findings of studies attempting to describe the effects of SF on successful pregnancies are conflicting. One study found that women partaking in coitus with their partners during peri-transfer periods of *in vitro* fertilization (IVF) cycles did not have a significantly greater chance of pregnancy, but were more likely to carry a viable embryo to at least 6-8 weeks of pregnancy (Tremellen, 2000). In contradiction, a more recent meta analysis of 2204 patients suggested that women exposed to SF around ovum pick up or embryo transfer do have significantly higher rates of clinical pregnancy, but show no difference with regards to live birth rate or other clinical parameters (Crawford *et al.*, 2015).

SF-mediated effects may be specific to the partner from whom the semen originates and to their relationship with the woman. Women treated with IVF or intracytoplasmic sperm injection using their partner's sperm and SF exhibit a third of the risk of preeclampsia (a syndrome thought to be caused by failed induction of maternal immunotolerance towards the semi-allogenic fetus) compared to those undergoing the same procedures with sperm that they had not previously been exposed to (Wang *et al.*, 2002). Similarly, women undergoing intrauterine insemination with washed sperm from a donor were more likely to develop preeclampsia than those receiving sperm from their partner. In this instance however, the antigenic factor appears to be associated with the sperm, and not the fluid (Smith *et al.*, 1997). Yet other studies found that cumulatively increased vaginal exposure to paternal semen or duration of relationship prior to conception is significantly inversely correlated with preeclampsia risk (Kho *et al.*, 2009; Saftlas *et al.*, 2014). Yet another study found that intrauterine insemination using donor sperm was found to increase the incidence of preeclampsia and multiple cycles with the

same donor sperm appear to convey a weakly protective effect (Kyrou *et al.*, 2010). Collectively, it seems clear that immunotolerance induced in the female reproductive tract is specific to the partner from whom the SF originates.

Health of offspring

Epidemiological research indicates the possible existence of long-term health effects on babies mediated by SF exposure. Infants conceived via ART have ~2-fold increased risk of perinatal mortality, low birth- and preterm birth weight, are 50% more likely to be small for their gestational age and a 30-35% increase in birth defects is observed relative to naturally conceived infants (Bower & Hansen, 2005). Additionally, a subset of babies born from women who have shorter pre-conception sexual relationships with their partners are small for their gestational ages at birth (Kho *et al.*, 2009). However, other systematic reviews find that short-term outcomes of IVF appear positive, instead cautioning that long term risk factors, such as chronic cardiometabolic disease, may be increased in IVF-conceived children, albeit that the data are only extrapolative at this point given the relatively young age of the technology (Hart & Norman, 2013). While it would be impossible to determine whether the absence of SF exposure significantly contributes to these increased risks, a possible correlation is collectively indicated.

Immunotolerance

Both animal and human studies have shown that SF exposure results in immunological changes in the female reproductive tract. In mice, a soluble form of CD38 released from the seminal vesicles induces the differentiation of tolerogenic dendritic cells and CD4+ Fox3p+ regulatory T cells, thus protecting the semiallogenic fetus from resorption (Kim *et al.*, 2015). In an *in vitro* model, it was shown that human T cells also increase mRNA expression levels of CD25 and Fox3p in response to SF exposure. Furthermore, CD25 protein was detectable on CD4+ T cell membranes, indicating a shift towards an immunotolerant phenotype (Meuleman *et al.*, 2015). In humans, the cervical transition zone appears to be the primary site of

immune induction and cell-mediated immunity in the female reproductive tract (Pudney, 2005). The overall repertoire of leukocytes in the cervix is altered through intercourse. CD4+ T lymphocytes are found in higher concentrations in women who have recently experienced intercourse, whereas those who have not exhibit more macrophages (Prakash *et al.*, 2003). Another study found increased CD45+ cells (mostly comprising CD14+ macrophages and CD1a+ dendritic cells) following intercourse in cervical biopsies compared to biopsies from the same women taken prior to intercourse (Sharkey *et al.*, 2012). It is possible that the immunomodulatory roles of SF vary according to the recipient's ovulatory cycle; work by Kimura *et al.* (2009) showed that CD16+CD56+ bright cells were more common in cervical biopsies of pre-ovulatory women who had recently had intercourse compared to post-ovulatory women, as well as to abstinent women regardless of ovulatory status. Collectively, these responses indicate a shift towards immunotolerance.

1.6. Seminal fluid as a potential oncogenic driver

Inflammation is considered a tumour-enabling characteristic of cancer (Hanahan & Weinberg, 2011). SF contains many pro-inflammatory cytokines and is able to further induce expression of immune-related molecules in the female reproductive tract. It is able to stimulate cytokine production in cervical cells; SF was shown to induce mRNA expression of over 300 genes in primary ectocervical cells as well as in the ectocervical cell line Ect1. Inflammatory cytokine genes were prevalent among these, and Interleukin-6 (IL-6), IL-8, Colony-Stimulating Factor 2 (CSF2, also known as granulocyte-macrophage colony-stimulating factor) and chemokine (c-c) motif ligand 2 (CCL2) protein expression was also induced in Ect1 cells. The endocervical cell line End1 was shown to be less responsive in this respect. Given the biological orientation of the endo- and ectocervix, it is tempting to speculate that this may be physiologically significant, as ectocervical cells undergo greater and primary exposure to SF. While it was clear that SF could induce inflammatory cytokine production, significant inter-donor variation in cytokine-inducing ability was reported (Sharkey *et al.*, 2007). This inflammatory response is also observed in cancerous cervical tissues. IL-1α mRNA is upregulated in both cervical squamous cell

carcinoma and adenocarcinoma tissue explants relative to normal cervical tissue, and SF can further induce its expression in both cancerous and normal tissues as well as in HeLa cells, predominantly via the EP2/EGFR/PI3K signaling pathway (Adefuye, Sales & Katz, 2014). SF is also able to induce expression of the angiogenic factor Vascular Endothelial Growth Factor A (VEGFA), among others, in cervical tissues (Muller *et al.*, 2006). Given the apparent role of SF as an inflammatory agent as well as an inducer of angiogenesis (yet another hallmark of cancer), it may play a role in encouraging the onset and/or progression of cancers of the female reproductive tract (Wang *et al.*, 2010; Hanahan & Weinberg, 2011; Sales & Katz, 2012).

SF induces changes in cervical tissues via a number of pathways. Transforming growth factor-β (TGF-β) elicits changes in Ect1 cell expression of several pro-inflammatory cytokine and chemokine genes, replicating principal aspects of the Ect1 response to SF and induced a similar IL-6 and CSF2 response in primary cervical epithelial cells. TGF-β-neutralizing antibodies, receptor antagonists, and signaling inhibitors ablated seminal plasma induction of CSF2 and IL-6, but did not alter IL-8, CCL2, CCL20, or IL-1α production (Sharkey, Macpherson, *et al.*, 2012). Indeed, the EP2/EGFR/PI3K signaling pathway is the predominant regulator of IL-1α expression in both normal and cancerous cervical tissues (Adefuye, Sales & Katz, 2014). IL-8 and growth-regulated oncogene α expression are increased via transactivation of the epidermal growth factor receptor (EGFR), activation the Extracellular Signal-Regulated Kinase (ERK) pathway and induction of prostaglandin-endoperoxide synthase (PTGS, also known as cyclooxygenase) enzymes (Adefuye & Sales, 2012). SF has also been shown to exert some of its effects on HeLa cells via the EGFR-PTGS1-PGE_2 pathway (Sales *et al.*, 2014). Seminal prostaglandin E_2 (PGE_2) was shown to be a major factor influencing PTGS2 expression (Joseph *et al.*, 2013). SF also exerts PTGS2 and VEGF inducing activity in HeLa cells via the prostaglandin EP-4 receptor, EGFR and ERK1/2 signaling pathways, initiating or exacerbating a potentially already-present positive feedback loop in these cells (Muller *et al.*, 2006). Upregulation of PTGS2 by SF is also observed in vaginal and cervicovaginal explants and SF treatment synergistically potentiates other stimulants in doing the same. Taken together, SF clearly induces multiple cellular changes in a variety of tissues of the female

reproductive tract, including normal and cancerous cervical tissues, via multiple signaling pathways.

SF-treated cells also exhibit altered physiologies. HeLa cell mouse xenograft models treated with intraperitoneal SF injection showed increased inflammatory enzyme, cytokine and VEGFA expression, with treated mice developing larger, more vascularised tumours more rapidly than control mice (Sutherland *et al.*, 2012). This increased rate of xenograft HeLa tumour formation has also been observed elsewhere, where SF was injected together with tumour cells, rather than being intraperitoneally administered post-HeLa cell injection (Liu *et al.*, 2011). Physiological changes are not limited to cervical tissues; treatment of primary endometrial epithelial and stromal cell culture models with SF induces transcription of genes associated with migration, proliferation, viability and inhibition of cell death. Treatment also induced secretion of pro-inflammatory and –chemotactic cytokines as well as the expression of pro-angiogenic and proliferative growth factors by these cells (Chen *et al.*, 2014).

Finally, given its clear effects on immune function in the female reproductive tract (see section 1.5.2), SF exposure may induce a more tolerant environment in which cervical cancers may develop or grow more freely.

1.7. Rationale and Hypothesis

Cervical cancer is primarily caused by HPV infection; however, other factors also contribute to its development. Consequently, regular screening and early intervention significantly reduces morbidity and mortality (American Cancer Society, 2014). Furthermore, vaccines against high-risk HPVs have recently been developed and have markedly reduced the incidence of HPV infection in vaccinated populations (Garland *et al.*, 2015; Herrero, González & Markowitz, 2015; Joura *et al.*, 2015). Despite these great strides in prevention, however, cervical cancer remains a pressing issue. As previously mentioned, the burden of disease disproportionately falls on the developing world, where access to screening and vaccinations remains

limited (see section 1.1.1). These populations will thus continue to exhibit disproportionate morbidity rates. Beyond the necessary HPV infection and viral lifecycle, the molecular mechanisms driving cervical oncogenesis remain poorly understood. Greater insight into these mechanisms is required in order to effectively combat the disease in the aforementioned at-risk populations.

Both TBX2 and TBX3 exhibit altered expression patterns in cancerous cervical tissues *in vivo* and their expression is correlated to prognostic outcomes (Lyng *et al.*, 2006; Liu, Jiang & Zhang, 2010b). Furthermore, both proteins repress viral gene transcription at the LCR of a number of HPV subtypes, and interaction with HPV16 L2 capsid protein augments their repressive effects (Schneider *et al.*, 2013). While no mechanisms have been presented, it is clear that these proteins are in some way involved in the HPV lifecycle and therefore possibly also cervical oncogenesis.

SF influences both normal and malignant cervical tissue physiology and biochemistry. However, the mechanistic underpinnings of these effects are poorly or not yet fully described. It contains a number of signaling molecules that may play a role either individually or synergistically. Some of these, like the TGF-β family, have been shown to mediate their oncogenic effects via TBX2- and TBX3-related pathways in other cancer contexts (Li, Weinberg, *et al.*, 2013; Li *et al.*, 2014).

We hypothesize that TBX2 and TBX3 do play a significant role in cervical carcinomas and that they may act as mediators of the oncogenic effects caused by SF.

1.8. Aims

The aims of this research study were as follows:

1. To determine whether TBX3 protein expression is altered in cervical cancer tissues compared to normal cervical tissues *in vivo*.
2. To determine whether SF treatment influenced TBX2 or TBX3 gene expression in cervical cancer cell culture models.
3. To determine whether SF-mediated changes in TBX2 or TBX3 expression affected TBX2 or TBX3 target gene expression.
4. To determine whether SF-mediated changes in TBX2 or TBX3 gene expression resulted in altered cervical cancer cell physiology.

2. Materials and Methods

2.1. Materials

Standard laboratory reagents were purchased from Sigma-Aldrich unless stated otherwise. HeLa (ATCC® CCL-2™), CaSki (ATCC® CRL-1550™) and C-33 A (ATCC® HTB-31™) cells were obtained from laboratory stocks and were originally purchased from the American Type Culture Collection. Formalin-fixed, paraffin-embedded (FFPE) histological slides with cervical adenocarcinoma, squamous cell carcinoma and normal tissue biopsy samples mounted on them were obtained from the Division of Anatomical Pathology, Faculty of Health Sciences, University of Cape Town (UCT), Cape Town, South Africa. Semen was obtained from healthy male volunteers attending the Andrology Laboratory of the Reproductive Medicine unit at Groote Schuur Hospital, Cape Town, South Africa. Informed consent was obtained from all participants (see appendix I). The Human Research Ethics Committee, Health Sciences Faculty, UCT granted ethical approval for collection of both cervical tissues and SF for use in the research presented herein (HREC REF: 0849/2015).

2.2. Cell culture

2.2.1. Culture conditions

All cell lines were cultured in complete medium (see appendix II). Media were replaced every two or three days and cells were cultured in 174 cm^2 CELLSTAR® cell culture flasks (Greiner Bio-one GmbH) until 80-90% confluent. Thereafter they were washed twice with phosphate-buffered saline (PBS, appendix II), followed by addition of 4 mL Trypsin/EDTA (21 µM / 684 µM in PBS). After a brief incubation period at 37°C, cells were examined by positive phase contrast microscopy with a TMS inverted microscope (Nikon Instruments, Netherlands) to assess the degree of detachment. Once sufficiently detached, 6 mL complete medium was added to inactivate the Trypsin/EDTA and a portion of the suspension (3 mL for CaSki and

HeLa cells, 4 mL for C-33 A) was returned for culturing. The remaining cell suspensions were used in experiments or discarded. Cells were incubated in 37°C incubators (5% CO_2, 65% humidity). All culture flasks, dishes and plates were purchased from Greiner Bio-One GmbH.

2.2.2. Mycoplasma testing

Cells were routinely tested for Mycoplasma infection by culturing cells on a sterile coverslip in antibiotic-free complete medium for three days, after which they were fixed using a 1:3 glacial acetic acid:methanol fixative for 10 seconds. Coverslips were then washed with water and air-dried. Cells were stained with Hoechst 33258 (0.5 µg/ml, Hoechst) for 10 seconds, washed with water, and mounted on a slide using mounting fluid (see appendix II). Slides were examined using an Axioskop 2 MOT fluorescence microscope (Zeiss GmbH) utilizing an HBO 100 mercury short arc lamp and a DAPI filter (excitation: Band Pass 365/12, Beamsplitter: Colour Splitter (FT) 395, Emission: Long Pass 397).

2.3. Immunohistochemistry

2.3.1. Staining

Slides were heated in a 55°C oven for 20 minutes to melt the paraffin in which the tissues were embedded. They were then deparaffinised and rehydrated by two 5 minute washes in xylene (Merck) followed by 5 minute washes in a decreasing alcohol gradient (100%, 100%, 95%, 70%, 50%, Merck) and rinsed in deionized water (dH_2O). Antigen retrieval was performed by immersing slides in 0.01 M sodium citrate (pH 6) inside a pressure cooker (Russell Hobbs, South Africa) for 2 minutes. Thereafter the pressure was released and the cooker was placed in cold water to cool to room temperature. Slides were washed with dH_2O before being incubated in 3% hydrogen peroxide, diluted in methanol, for 30 minutes on a shaker at room temperature (RT) to eliminate endogenous peroxidase activity. They were then rinsed in dH_2O and washed for 5 minutes in tris-buffered saline (TBS, see appendix

II). Slides were then blocked with 5% normal goat serum (Dako, ref no: X0907) diluted in TBS for 30 minutes in a humidifying chamber. Excess goat serum was drained and slides were incubated with primary antibody solution (anti-TBX3 polyclonal antibody (Abcam, AB99302) diluted 1:20 in TBS), or 5% normal goat serum in the case of secondary only controls, in a humidifying chamber at 4°C overnight. Slides were then washed twice for 5 minutes in TBS, after which 1-2 drops of secondary antibody (anti-rabbit Envision+ System HRP-labeled polymer (Dako, ref no: K4002)) was applied to each slide before being returned to the humidifying chamber for 30 minutes at RT. Slides were again washed twice in TBS for 5 minutes and 1-2 drops of chromogenic substrate DAB (3,3'-diaminobenzidine, Dako, ref no: K3467) was added to each slide for 10 minutes, after which they were rinsed in dH_2O and counterstained with Mayer's Hematoxylin and Scott's tap water substitute (see appendix II) for 1 minute each. Sections were then dehydrated by immersion in an increasing alcohol gradient followed by xylene (the reverse of the above, 1 minute each) after which coverslips were attached using Entellan (Merck). Slides were viewed using an Axioskop 2 MOT fluorescence microscope (Zeiss, GmbH) and images were captured using Axiovision v.7.0.0 software (Carl Zeiss Imaging Solutions GmbH).

2.3.2. Quantification of DAB staining

Images of cervical tissues were loaded into Fiji software (http://fiji.sc/Fiji) (Schindelin *et al.*, 2012) and colour-corrected by background subtraction. Thereafter stromal sections were digitally removed (if possible), to leave only normal epithelial cell or cancerous tissue area. Image de-convolution (H-DAB) and automatic thresholding of the resultant hematoxylin (H) and DAB stain monochromes (8-bit) was performed and the area of staining determined by measuring stain area within threshold limits. % DAB staining was calculated as per equation 1:

$$\%\ DAB\ stain = \frac{Area\ DAB\ stained}{Area\ \ DAB\ stained + Area\ H\ stained} \times 100 \qquad [1]$$

Multiple images were taken per slide when sufficient tissue of interest was present.

% DAB staining was averaged for images of cancerous or normal tissue in each slide. These average values were then used for statistical analyses (see section 2.12).

2.4. Preparation of SF

Semen used in this study was produced by men performing voluntary masturbation following 72 hours of sexual abstinence. A trained technician confirmed that all semen was within reference values for normal semen as stipulated by the World Health Organisation (Cooper *et al.*, 2010). Semen was processed within 30 minutes of collection. Collected semen was separated into sperm and SF by centrifugation at 15 000 x g for 20 minutes at 4°C. The supernatant was transferred to sterile eppendorf tubes and the pelleted sperm discarded. 50 µL from each supernatant was incubated overnight in complete medium (see appendix II) without antibiotics to screen for bacterial or fungal infection, with the remaining supernatants stored at -80°C. Incubations were assessed the following day and all infected SF was discarded. Remaining samples were pooled together in groups of 10, aliquotted and stored at -80°C until needed. All experiments used pooled SF (n = 10) at a final concentration of 1:50. This dilution was chosen as it does not influence the viability of HeLa cells (Jeremias *et al.*, 1997), whereas higher concentrations have been shown to reduce the viability of a number of cell lines (Nocera & Chu, 1993; Kim *et al.*, 2010; Chen *et al.*, 2014). Dilutions of 1:50 are routinely used in our laboratory without causing cytotoxic effects (Sutherland *et al.*, 2012; Adefuye, Sales & Katz, 2014).

2.5. SF treatment of cervical cancer cell lines

For reverse transcription-real time quantitative polymerase chain reaction (RT-qPCR) experiments (section 2.6) HeLa (5×10^5 cells / dish), CaSki (6×10^5), or C-33 A (7×10^5) cells were cultured in 3 mL complete medium (appendix II) in 60 mm culture dishes overnight. The following day, cells were washed twice with 2 mL PBS and 3 mL starving medium (appendix II) was added to each well. Following overnight incubation, cells were again washed twice with PBS and either 3 mL starving media only or starving media containing 1:50 SF (final concentration) was added. Cells were

incubated for various amounts of time depending on the particular experiment.

The same procedure was followed for western blot experiments (section 2.7), but in those experiments 2 x 10^5 HeLa or CaSki cells were seeded in 6-well plates instead of 60 mm dishes.

2.6. Reverse transcription-quantitative real time polymerase chain reactions

The relative amounts of mRNA species isolated from cell culture experiments were determined by reverse transcription followed by quantitative real time polymerase chain reactions (RT-qPCR, collectively), as described below.

2.6.1. RNA extraction

High Pure RNA Isolation kits (Roche) were used to extract and isolate RNA. Cells cultured and treated as per section 2.5 were washed twice with PBS and 200 µL PBS was added per dish. Thereafter cells were lifted by scraping with a cell scraper (Nest Scientific) and the resultant suspensions were transferred to sterile 1.5 ml Eppendorf tubes. Thereafter 400 µL Lysis buffer (see appendix II) was added per tube, samples were vortexed briefly and stored at -80°C until needed. Upon thawing, the entirety of the sample was pipetted into High Pure filter assemblies and centrifuged at 8 000 x g for 15 seconds. The eluent was discarded and 10 µL DNAseI suspended in 90 µL DNASeI incubation buffer (appendix II) was added to each assembly and incubated at room temperature for 15 minutes. 500 µL Wash Buffer I (appendix II) was added to each assembly and centrifuged as above. This process was repeated with 500 µL Wash Buffer II (appendix II) and again with 200 µL Wash Buffer II, with the final centrifugation at 13 000 x g for 2 minutes. Finally, samples were eluted in 40 µL Elution Buffer (appendix II) into sterile Eppendorf tubes by centrifugation for 1 minute at 8 000 x g.

2.6.2. RNA quantitation and purity assessment

Purified RNA sample concentrations and purity were determined by spectrophotometry using a NanoDrop 2000 spectrophotometer (Thermo Scientific). Samples with A260/280 or A260/230 ratios less than 1.8 were considered impure.

2.6.3. Synthesis of single strand cDNA

ImProm-II™ Reverse Transcription Systems (Promega) were used to reverse-transcribe isolated RNA to produce complementary DNA (cDNA). Equal amounts of RNA per sample were reverse transcribed within each experiment. Reactions were performed according to the manufacturer's instructions. Briefly: up to 1 µg of RNA was combined with 0.5 g oligo (dT) primer and nuclease-free water was added to make a final volume of 5 µL per reaction. Samples were then incubated at 70°C for 5 minutes to denature the RNA, followed by incubation on ice. Thereafter 15 µL of reverse transcription mix (1X ImProm- II™ Reaction buffer, 3 mM $MgCl_2$, 0.5 mM dNTP mix, 20 units RNasin® ribonuclease inhibitor and 1 µl ImProm-II™ reverse transcriptase) was added to each experimental reaction. A no-reverse-transcription control was also made for each sample, wherein the reverse transcriptase and RNasin® were substituted with nuclease-free water. Reactions were then incubated at 25°C for 5 minutes (primer annealing), then 42°C for 1 hour (reverse transcription) and finally 70°C for 15 minutes (inactivation of reverse transcriptase). The samples were then used immediately in qPCR experiments (section 2.6.4) or stored at 4°C and used the following day.

2.6.4. Real time quantitative PCR

Real time quantitative PCR (qPCR) reactions were performed using a StepOne Plus PCR System (Applied Biosystems, UK). Reactions were performed in MicroAmp optical 96-well plates or 8-well strips (Applied Biosystems). Per well, 2 µL of diluted cDNA was combined with 8 µL of qPCR MasterMix (FG-Power SYBR® Green PCR Master Mix (Applied Biosystems), 10 µM forward primer, 10 µM reverse primer)

containing the appropriate primers for each gene being investigated. Unless otherwise stated, all cDNA samples were diluted 1/5 prior to amplification. Samples were measured in triplicate in each experiment. Matching no-reverse-transcription controls (also diluted 1/5) were included for each sample-target pair and no-template controls (where cDNA is replaced with nuclease-free water) were included for each target gene in all experiments. Continuous melt curves were included in all experiments to confirm the production of single amplicon products and to assess whether genomic contamination had occurred (see appendix III). qPCR thermocycling conditions are detailed in Table 2-1.

Table 2-1: Thermocycling parameters used during RT-qPCR.

	Step	**Temperature (°C)**	**Duration**	**No. Cycles**
1	Polymerase Activation	95	10 min	1
2	Template DNA denaturation	95	15 sec	40
3	Primer-template annealing	55	3 sec	
4	Elongation	60	1 min	
5	Final denaturation	95	15 sec	1
6	Final extension	60	1 min	1
7	Continuous melt curve	60-95, 3.3% gradient	-	-

EEF1A1, HPRT1, and HSPC90AB1 qPCR primers were ordered from the Synthetic DNA laboratory, Department of Molecular Cell Biology, UCT. TBX2 (QT00091266), TBX3 (QT00022484), GUSB (QT00046046), p21 (QT00062090) and PTEN (QT01676969) qPCR primers were ordered from Qiagen. Sequences of primers used in this study are detailed Table 2-2.

2.6.5. Data analysis

Data generated by the StepOne Plus qPCR instrument was exported as a result file, which was subsequently imported into qBase+ version 2.6.1 (Biogazelle). Replicate measurements that differed by more than 0.5 quantification cycles (Cq, also known as threshold cycles, Ct) were excluded from analysis. Negative control thresholds (difference between control Cq values and the highest Cq measured for a specific gene in a specific run) were set at <5 Cq. Failing reactions were excluded from further analysis. Relative quantitation was performed with efficiency correction as described by Hellemans *et al.* (2007). Briefly: primer efficiencies were determined for all primers by creating a standard curve dilution series of a sample, from which the slope and its standard error was calculated by linear regression. Normalised relative quantities (NRQ) of each gene in a given sample were determined as described in equation 2. All expression was normalized to two reference genes, *GUSB* and *HSPC90AB1* (see section 3.2). qBase+ was also used to determine reference gene expression stability via GeNorm (see section 3.2)(Vandesompele *et al*., 2002). Target-specific amplification efficiences for both reference and target genes were applied to all calculations. Reference genes were considered stable within an experiment if GeNorm analysis returned an M-value < 0.5 and a coefficient of variation < 0.2 (Hellemans *et al.*, 2007). Where appropriate, inter-run calibration was performed using untreated HeLa cDNA. Resultant Calibrated-NRQ (CNRQ), or NRQ data when inter-run calibration was not required, was then exported for further analysis (see section 2.12).

$$NRQ = \frac{E_{goi}^{\Delta Cq,goi}}{\sqrt[f]{\prod_0^f E_{ref_0}^{\Delta Ct,ref_0}}} \qquad [2]$$

$Where$:

$\Delta Cq =$ difference of Cq values of a sample of interest and a scaling sample

$goi =$ gene of interest

$ref =$ reference gene

$E_{goi\ or\ ref} =$ The amplification efficiency of the goi or the ref, respectively

Therefore, quantitation occurs by division of the linearized, efficiency-corrected ΔCq of the gene of interest by the geometric means of the linearized, efficiency-corrected ΔCq of the reference genes. Samples may also be scaled to the average expression of a gene across a run.

Table 2-2: qPCR primer sequences, amplicon sizes and melting temperatures.

Gene	Forward primer 5'-3'	Tm (°C)	Reverse primer 5'-3'	Tm (°C)	amplicon size (base pairs)
HPRT1	TGAGGATTTGGAAAGGGTGTT	57.39	CAGAGGGCTACAATGTGATGG	58.43	111
EEF1A1	TGCGGTGGGTGTCATCAAA	59.85	AAGAGTGGGGTGGCAGGTAT	61.12	123
HSPC90AB1	ATGGAAGAGAGCAAGGCAAA	57.41	AATGCAGCAAGGTGAAGACA	58.01	117
GUSB	*Proprietary*	55	*Proprietary*	55	96
P21	*Proprietary*	55	*Proprietary*	55	79
PTEN	*Proprietary*	55	*Proprietary*	55	111
TBX2	*Proprietary*	55	*Proprietary*	55	62
TBX3	*Proprietary*	55	*Proprietary*	55	119

2.7. SDS Poly acrylamide gel electrophoresis and western blotting

Cells cultured as per section 2.5 were placed on ice and washed twice with PBS. Thereafter 70 µL lysis buffer (2 x Boiling Blue, see appendix II) was added to each well and cells were lysed by scraping with the back of a yellow pipette tip. The resultant mixtures were transferred to sterile Eppendorf tubes and heat-treated to denature proteins at 95°C for 7 minutes and immediately returned to ice. Tubes were inverted several times and briefly centrifuged. Thereafter, 20 µL of each sample was inserted into wells of polyacrylamide gels (4% stacking, 8% running) containing sodium dodecyl sulfate (SDS-PAGE, see appendix II) immersed in running buffer (appendix II) and electrophoresed at 100 Volts until adequate separation, as judged by migration of the molecular weight ladder (PageRuler Prestained Protein Ladder, Thermo Fisher Scientific), was achieved. Thereafter, gels were removed from

the apparatus, the stacking gel section was excised and the running section was placed in a transfer cassette assembly and transferred onto Hybond ECL nitrocellulose membrane (Amersham Biosciences) while immersed in Transfer buffer (appendix II) for 90 minutes at a constant current of 0.35 Amperes.

Following transfer, membranes were blocked for 1 hour at RT in PBS containing 5% non-fat dry milk and 0.1% tween-20 (PBS/T+5% milk), after which they were probed with the appropriate primary antibodies overnight on a shaker at 4°C. Membranes were then washed with PBS containing 0.1% Tween-20 (PBS/T) and incubated with the appropriate peroxide-conjugated secondary antibody for 1 hour at RT. Thereafter they were washed again with PBS/T and protein bands were visualised by enhanced chemiluminescence (SuperSignal® West Pico Chemiluminescent Substrate #34080, Thermo Scientific) and captured by exposing x-ray film to luminescent membranes. Developed films were digitized by scanning and densitometry analysis was performed using ImageJ v.10.2 software (Abràmoff, Magalhães & Ram, 2004).

Primary antibodies used in this study were rabbit polyclonal anti-TBX3 (1:1000)(AB99302, Abcam), goat polyclonal anti-TBX2 (1:500)(sc-17880, Santa Cruz Biotechnology), rabbit polyclonal anti-p21 (1:500)(sc-756, Santa Cruz Biotechnology), rabbit monoclonal anti-GAPDH (1:20 000) (14C10, Cell Signaling), all diluted in PBS/T+5% milk, and rabbit polyclonal anti-p38 (1:10 000)(9219, Cell Signaling) diluted in PBS/T. Secondary antibodies used were peroxide-conjugated anti-goat (Santa Cruz Biotechnology) or anti-rabbit antibody (BioRad), both diluted 1:5000 in PBS/T+5% milk.

2.8. Actinomycin D treatment

HeLa cells were cultured for RT-qPCR experiments as described in section 2.5. However, prior to incubation with SF, cells were washed twice with PBS, starving medium (appendix II) was added and cells were pre-treated with either Actinomycin D (Act D, 5 µg/mL), or vehicle (DMSO) for 1 hour. SF (1:50 final concentration) was then added drop-wise to treatment dishes and all dishes were incubated for a

further 4 hours. All incubations occurred in the dark. Thereafter, cells were lysed and gene expression was quantified via RT-qPCR as described in section 2.6.

2.9. TBX3 transient knock-down experiments

HeLa cells were seeded in 60 mm dishes (3 x 10^5 cells / dish) and cultured overnight in complete medium (appendix II) to 50% confluence. Cells were then transfected with 50 nM siTBX3 (SI00083503, Qiagen, USA) or siCtrl (1027310, Qiagen, USA) siRNA, as appropriate, using HiPerfect® transfection reagent (Qiagen, USA) according to the manufacturer's instructions and cultured overnight. The following day, cells were washed twice with PBS and starving medium (appendix II) with or without SF (1:50 final concentration) was added to each well as required. Cells were treated for 4 hours before being lysed and processed for gene expression analyses via RT-qPCR as per section 2.6.

2.10. Wound healing assays

HeLa cells were seeded in 24-well plates (6 x 10^4 cells / well) in complete medium (appendix II) and cultured overnight. The following day, all wells were examined to confirm cell confluence and wounds were induced into the cell monolayer of each well by scratching the surface with a sterile pipette tip. Each well was then washed twice with complete media before addition of 500µL complete media supplemented with 10 µg/mL mitomycin C to inhibit cellular proliferation throughout the wound-healing assay. Treatment-well medium was also supplemented with SF (1:50 final concentration). Immediately after medium addition, each well was photographed using a digital camera attached to an Olympus-CKX41 inverted light microscope (Olympus Life Sciences, South Africa) in three unique locations along the wound, as demarcated by lines drawn on the bottom of the plate prior to cell seeding. These images served as the 0 hour time point for each experimental condition. Image acquisition was repeated 3, 6, 9, 12 and 24 hours after the wound induction.

Once all images were taken, ImageJ v.10.2 software (Abràmoff, Magalhães & Ram, 2004) was used to determine the area of the scratch in each image. The 0 hour image in each condition served as the reference, and cell migration distance was determined by subtraction of the remaining area in later images from that of the 0 hour image. Data from the three locations in each well were averaged to produce mean migration rates for each well. Area manipulation calculations were performed in Microsoft Excel v.14.7.1 (Microsoft Corporation) before exporting the results from all independent experiments into GraphPad Prism® v.6.0c (GraphPad Software, USA) for statistical analysis.

2.11. Proliferation assays

HeLa cells were seeded in two 60 mm dishes (4×10^5 cells / dish) and cultured overnight. Thereafter, one dish was treated with 1:50 SF (final concentration) and the other left untreated for 24 hours. Cells were then lifted by trypsinisation and 5×10^4 cells / well were seeded in 24-well plates and allowed to proliferate over 7 days. Cells were lifted by trypsinisation and counted at 1, 3, 5 and 7 days to assess population growth. All populations were resuspended in the same volume of media. On day 2 and 5 cells that were not being counted that day were lifted and transferred to larger culture vessels to avoid contact inhibition of proliferation. Complete medium (appendix II) was used throughout and media were replaced daily. Time points were assessed in triplicate within each independent experiment.

2.12. Statistical analyses

GraphPad Prism® v.6.0c (GraphPad Software, USA) was used to produce all graphs and to conduct all statistical analyses presented herein, unless specifically stated otherwise. Statistical significance was defined as $p < 0.05$ in all analyses. The statistical analyses performed on each dataset are specified in the relevant figure legends.

3. Results

3.1. TBX3 protein is overexpressed in cervical cancer tissues

While the expression status of TBX2 protein in cervical cancers relative to normal tissues is known (Liu, Jiang & Zhang, 2010b), similar comparisons have not been reported for TBX3. TBX3 mRNA is expression levels have only been compared across different cervical cancers, where low levels of expression are associated with lymph node metastasis and poor prognosis (Lyng *et al.*, 2006). No research has investigated levels of TBX3 mRNA or protein expression relative to normal cervical tissues. To determine whether different TBX3 protein expression levels exist in normal and cancerous cervical tissues, we investigated TBX3 protein expression in FFPE cervical carcinomas and normal tissue via immunohistochemistry. The results, shown in **Figure 3.1, -3.2** and **-3.3**, indicate that TBX3 is indeed significantly overexpressed in both cervical squamous cell carcinoma (~4,35 fold, $p < 0.001$; $n = 3$) and adenocarcinomas (~4,45 fold, $p < 0.001$; $n = 4$) relative to normal cervical tissues from patients who did not have cervical cancer ($n = 3$)**(Figure 3.1 b)**. Additionally, significant differences were seen between adenocarcinoma tissues and matched adjacent normal cervical epithelia present in the same biopsy sections. Compared to matched normal epithelia, adenocarcinoma tissues exhibited ~1,83 fold greater TBX3 expression ($p<0.05$; $n = 4$; **Figure 3.2 a, c**). While a similar trend appears evident in the squamous carcinomas tested in this study **(Figure 3.2 a, b)**, only two sections contained significant regions of normal epithelium, thus precluding statistical interrogation.

Staining indicated that TBX3 was detectable in the nuclei and cytoplasm of cancerous tissues **(Figure 3.3)**. Minimal staining was observed in the nuclei of normal cervical epithelial cells – the majority of quantified staining of these tissues appears to be nonspecific staining of the epithelial extracellular matrix **(Figure 3.3)**.

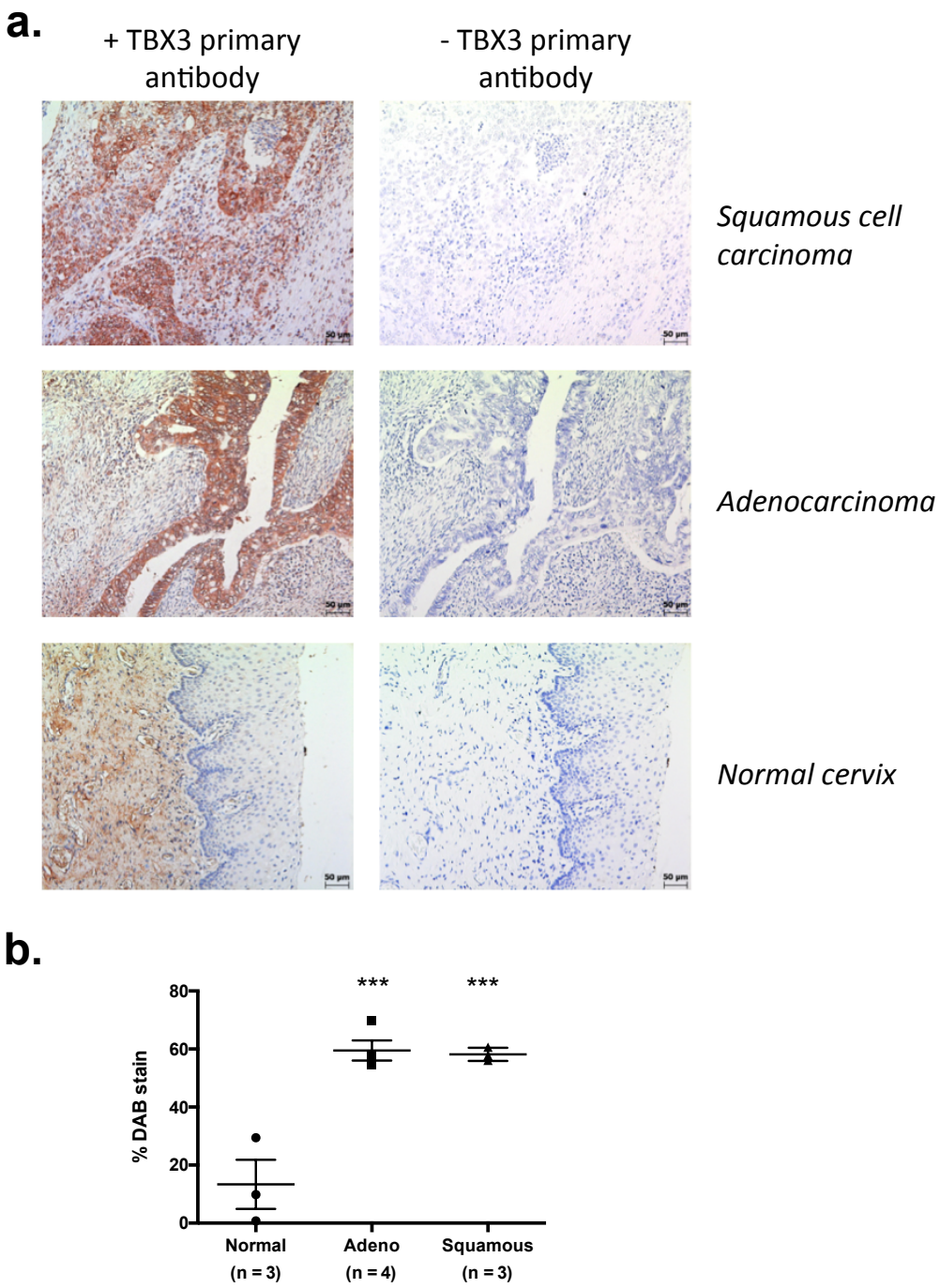

Figure 3.1: TBX3 protein is overexpressed in cervical cancer tissues compared to normal tissues. Formalin-fixed, paraffin-embedded cervical tissues were probed for TBX3 expression via immunohistochemistry. **a**, Staining of squamous carcinoma (top), adenocarcinoma (middle), and normal cervix (bottom). Staining with primary antibody directed against TBX3 is displayed in a panel on the left (+ TBX3 primary antibody) and matching secondary antibody-only control slides are pictured on the right (- TBX3 primary antibody). **b**, Area of staining of normal epithelial tissues and cancerous tissues were quantified as a percentage of total staining. For quantification 2-3 images were taken of each section, depending on the distribution of cancerous tissue. Resultant values were averaged and pooled by tissue type. Statistical significance was determined by ordinary one-way ANOVA with Dunnet's multiple comparisons test applied to comparisons between normal and cancerous tissues. Asterisks indicate significant differences compared to normal tissue. Means ± SEM are displayed. *** = $p < 0.001$.

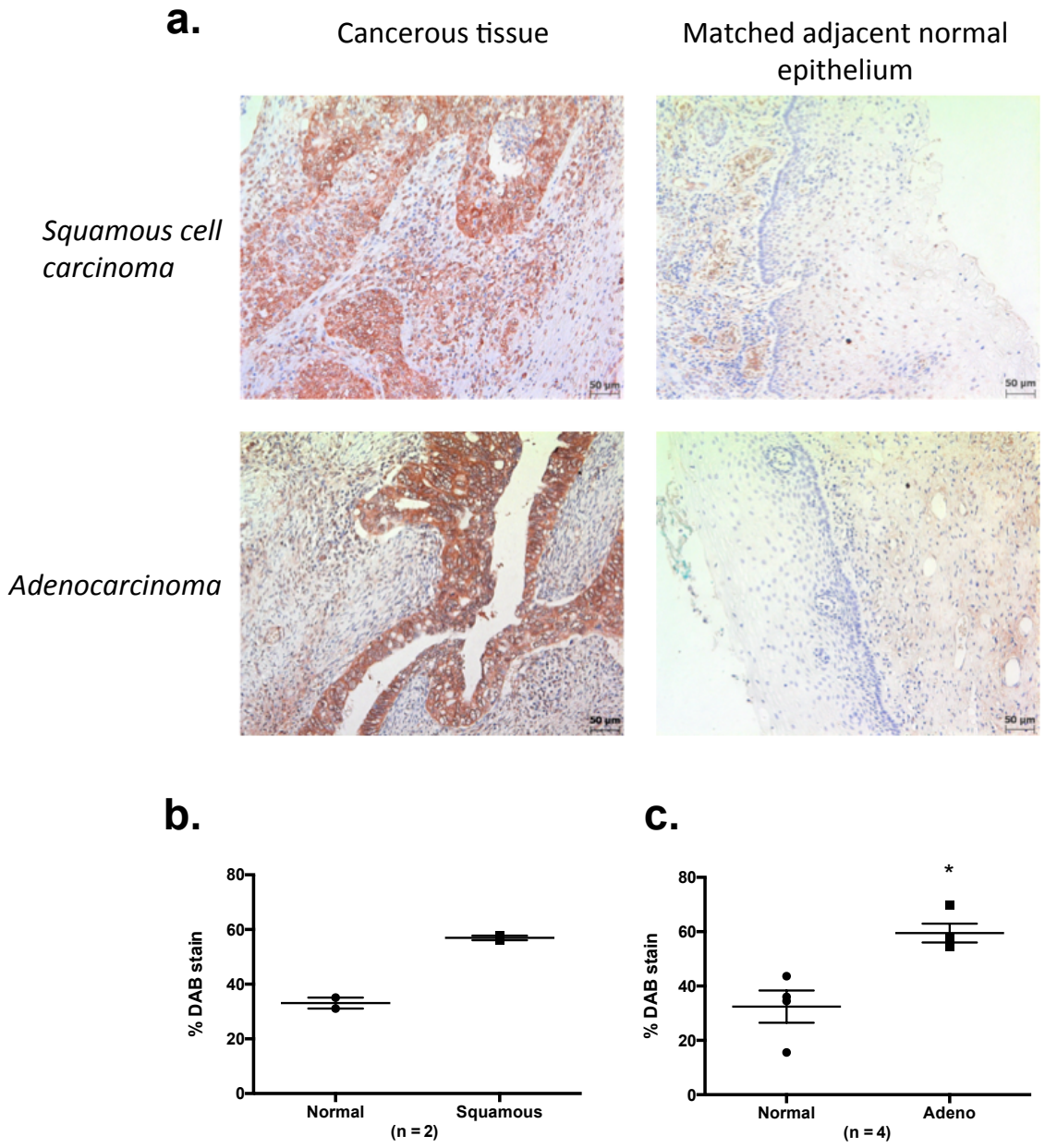

Figure 3.2: TBX3 protein is overexpressed in cancerous tissues compared to matched adjacent normal tissues.

Formalin-fixed, paraffin-embedded cervical tissues were immunostained with antibody directed against TBX3. **a**, Examples of squamous cell carcinoma (top left) and adenocarcinoma (bottom left) tissues compared to matched adjacent normal cervical epithelium (right) from the same tissue section. The area of staining of normal epidethelial tissues and squamous cell carcinoma (**b**) or adenocarcinoma (**c**) tissues were quantified as a percentage of total staining. 2-3 images of each section were taken depending on the distribution of cancerous tissue throughout the section. Resultant values were averaged and the means pooled by tissue type (squamous carcinoma n = 2, adenocarcinoma n = 4). Statistical significance was determined by paired t-test. Means ± SEM are displayed. * = $p < 0.05$.

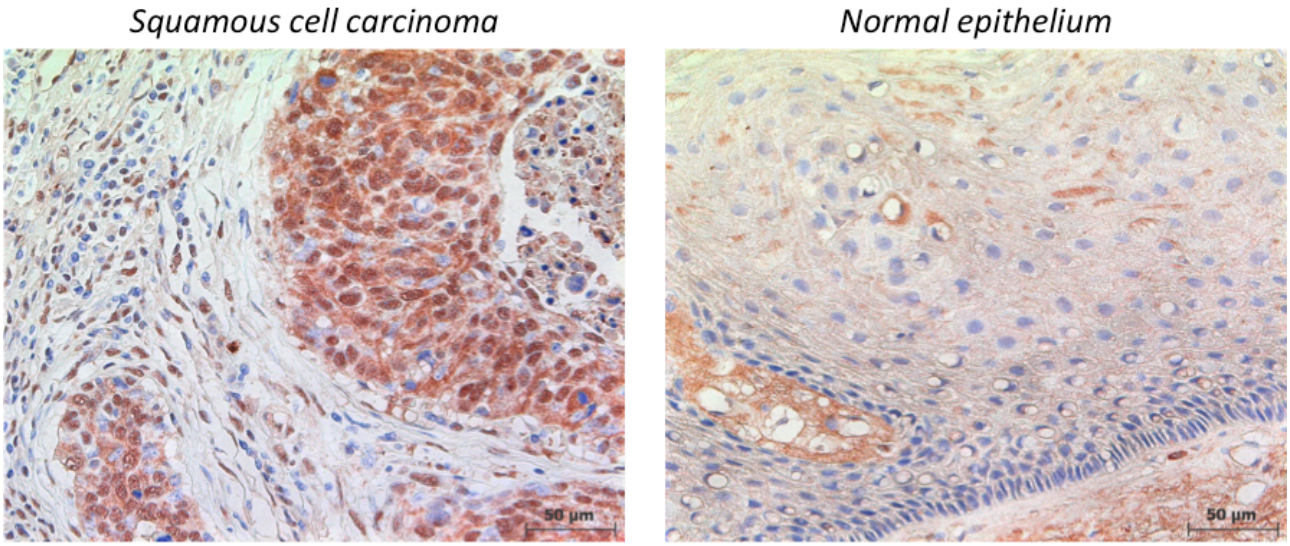

Figure 3.3: TBX3 protein is overexpressed in the nuclei and cytoplasm of cancerous cervical tissues compared to normal epithelium.
Formalin-fixed, paraffin-embedded cervical tissues were immunostained with antibody directed against TBX3. Representative images of staining patterns in squamous cell carcinoma tissue (left) and normal epithelium (right) are displayed.

3.2. *GUSB and HSPC90AB1* are suitable reference genes for use in RT-qPCR experiments investigating cervical cancer cell lines

We hypothesized that TBX2 and TBX3 expression may be regulated by SF treatment. In order to test this, we sought to examine the relative levels of TBX2 and TBX3 mRNA expression in SF-treated and untreated cervical cancer cell lines. Therefore suitable reference genes (also called house keeping genes) in cervical cancer cells had to be identified in order to examine the relative levels of mRNA expression of these (and other) gene targets in RT-qPCR experiments. Multiple reference genes are required for normalization of gene expression in order to accurately determine subtle differences in target gene expression levels between samples (Vandesompele *et al.*, 2002). As per the MIQE[1] guidelines, reference genes must be constitutively expressed in all samples and their expression levels must remain stable under all experimental conditions in order to provide valid normalization (Bustin *et al.*, 2009).

[1] Minimum Information for publication of Quantitative real-time PCR Experiments

Based on work done in our laboratory as well as published literature examining gene expression stability in cervical cancer, the following candidate genes were selected; *GUSB, HSPC90AB1, HPRT1, EEF1A1,* and *PPIA* (Steinau, Rajeevan & Unger, 2006; Krainova *et al.*, 2013). Primer pairs that cross exon boundaries or which produce products that span two or more exons were selected where possible as a means of protecting against PCR products due to genomic contamination. Primer information for all candidate genes is detailed in Table S1 (appendix III).

Primer sequences were first analysed *in silico*; each pair was subjected to a primer BLAST search (ncbi.nlm.nih.gov) across the human genome database to assess target specificity. Primers were considered to be target-specific if the BLAST search only returned PCR products of a single size that aligned to validated sequences of the intended gene. Thereafter, primer pairs were submitted to the RT Primer Database (rtprimerdb.org) to determine whether their binding regions contained known single nucleotide polymorphisms (SNPs) that may influence binding. From these analyses the primers for *HSPC90AB1*, and *HPRT1* were confirmed as suitable for further investigation. The *EEF1A1* reverse primer had a mismatch on its ‘3 terminal nucleotide; removal thereof did not significantly influence its 50% melting temperature (Tm) and resulted in a perfect *in silico* alignment to the target gene sequence. The modified primer pair was therefore also included for further investigation. *GUSB* primers were purchased from Qiagen and have proprietary sequences, thus rendering *in silico* analysis impossible. However, they are guaranteed to produce a specific product under the cycling conditions used in this study and have been used extensively in our laboratory in the past and therefore were also included for further analysis. The *PPIA* primers were unable to distinguish between the two mRNA transcript species produced by this gene, thus producing two unique amplicons of different sizes (116 / 273 base pairs, respectively; data not shown). This lack of specificity resulted in exclusion from further investigation.

Remaining candidate reference genes were analysed in a pilot HeLa cell experiment containing both SF-treated and untreated cell populations. Reference gene stability was assessed using the GeNorm functionality in qBase+. Gene expression was

considered very stable for *HRPT1*, *HSPC90AB1* and *GUSB* (GeNorm M ≤ 0.2). *GUSB* and *HSPC90AB1* were determined to be the best combination of genes with which to normalize HeLa cell gene expression (GeNorm V ≤ 0.15). While it is recommended to use at least three reference genes, the results of this analysis indicated that the addition of more reference genes did not increase the strength of the normalisation. Reference gene primer amplification efficiencies were determined **(Table 3-1)**; efficiencies ranging between 85-110% were considered acceptable. *GUSB* and *HSPC90AB1* both showed high efficiencies and were therefore accepted as suitable reference genes for use in this study. The amplification efficiencies of target gene primers were also determined; all primer pairs were confirmed to have acceptable amplification efficiencies, thus allowing for accurate efficiency-corrected relative quantitation (Vandesompele *et al.*, 2002).

Table 3-1: qPCR primer gene amplification efficiencies as determined by target-specific standard curves.

Target Gene	Efficiency (%)	Std. Error	r^2	Slope	Slope error	Intercept	Intercept error
Reference genes							
EEF1A1	102,8	0,051	0,989	-3,257	0,116	9,364	0,227
HPRT1	106,2	0,083	0,976	-3,182	0,178	19,962	0,347
HSPC90AB1	106,5	0,055	0,991	-3,175	0,116	14,119	0,204
GUSB	109,7	0,022	0,998	-3,110	0,044	20,535	0,082
Genes of interest							
TBX2	103,0	0,022	0,998	-3,252	0,050	27,154	0,084
TBX3	98,7	0,034	0,995	-3,354	0,083	24,459	0,141
p21	86,7	0,023	0,996	-3,689	0,071	19,759	0,134
PTEN	94,6	0,021	0,998	-3,459	0,057	17,719	0,110

3.3. SF regulates TBX2 and TBX3 mRNA expression in a context-dependent manner

Cell line RT-qPCR studies were conducted to investigate the effects of SF treatment on TBX2 and TBX3 gene expression. Three different cervical cancer cell lines, namely HeLa, C-33 A and CaSki cells, were treated with SF at a final concentration of 1:50 and incubated for various periods spanning 1-24 hours. Untreated controls were included at each time point as well as at 0 hours. SF treatment significantly increased TBX2 and TBX3 gene expression in HeLa cells **(Figure 3.4 a)**. SF-Treated HeLa cells exhibited significantly increased expression TBX2 mRNA at 4 (4.5-fold, $p < 0.001$) and 8 (2.65-fold, $p < 0.05$) hours relative to corresponding untreated cells. SF treatment also significantly upregulated TBX3 mRNA expression at 4 (2.8-fold, $p < 0.001$) and 8 hours (2.7-fold, $p < 0.05$) in these cells. Both TBX2 and TBX3 mRNA remained elevated at 24 hours, albeit not significantly. In CaSki cells, SF treatment elicited a similar response with regards to TBX3 regulation, with significantly increased expression observed at 2 and 4 hours post treatment (5.4-fold, $p < 0.01$; 4.8-fold, $p < 0.0001$), but with expression returning to basal levels by 24 hours. TBX2 mRNA was not detectable above no-template and no-reverse-transcription control values in treated or untreated CaSki cells **(Figure 3.4 b)**. In contrast to these findings, while C-33 A cells were shown to express both mRNA species, SF treatment did not regulate the expression of either TBX2 or TBX3 mRNA at any of the time points investigated **(Figure 3.4 c)**.

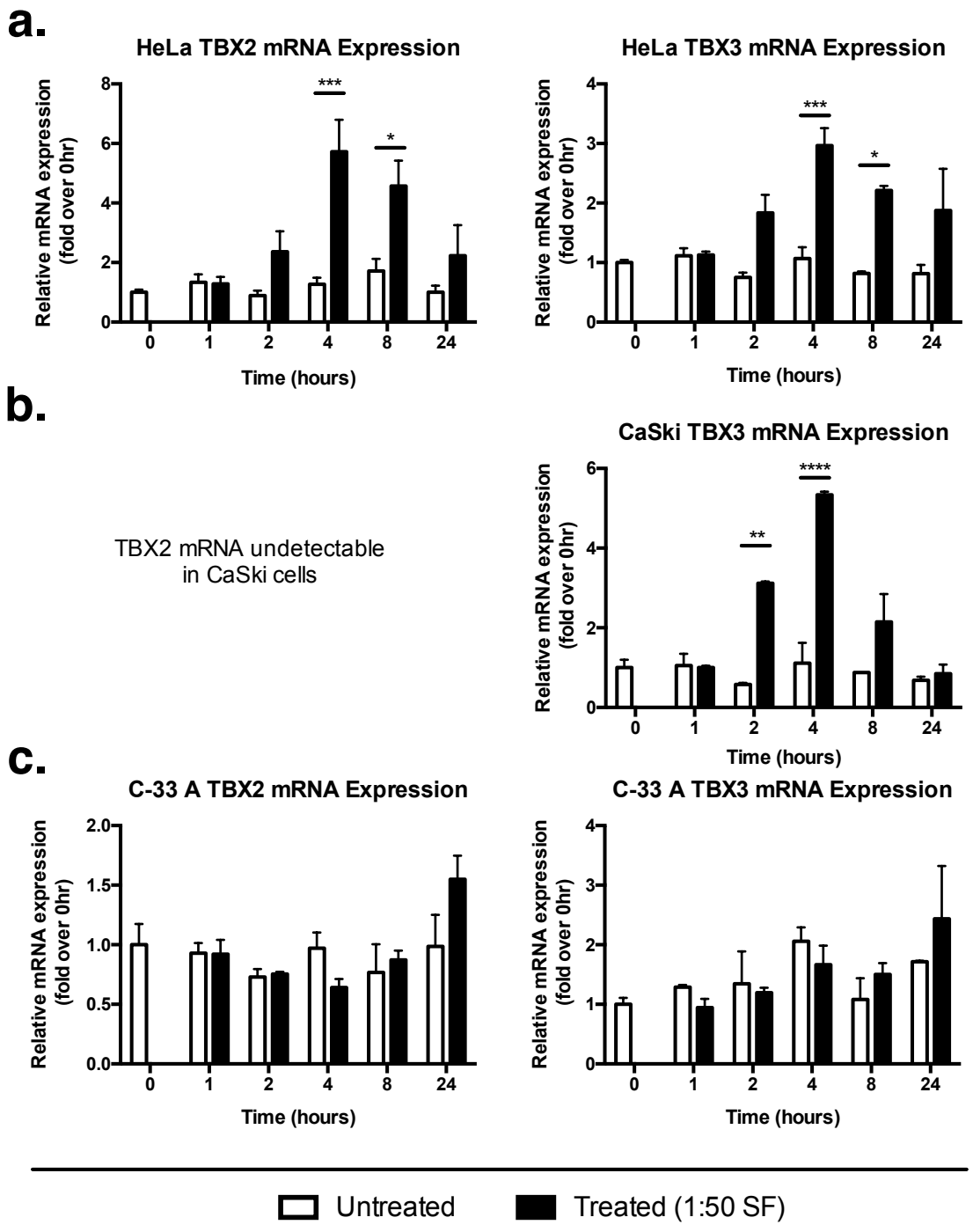

Figure 3.4: SF treatment increases TBX2 and TBX3 expression in cervical cancer cell lines in a cell line-dependent manner.

HeLa (**a**), CaSki (**b**) and C-33 A (**c**) cervical cancer cell lines were cultured in 60 mm dishes for 24 hours after seeding. Thereafter they were washed in PBS and serum starved overnight. The cells were then washed again and fresh starving media was added to each dish. Experimental dishes received seminal fluid (SF) in the media (1:50 final concentration) while control dishes were left untreated. Cells were harvested by scraping and RNA was isolated, reverse transcribed and TBX2 and TBX3 mRNA was quantified by relative quantitation. Each graph represents the pooled mean data from 3 independent experiments. Time points were assessed in triplicate within each experiment. Means ± SEM are displayed. Statistical significance was determined by normal Two-way ANOVA and Šídák's multiple comparison post-hoc test. Significant differences compared to time-matched untreated controls are displayed. * = $p < 0.05$; ** = $p < 0.01$; *** = $p < 0.001$; **** = $p < 0.0001$.

3.4. SF treatment upregulates TBX3 mRNA transcription

To confirm that SF-mediated TBX2 and TBX3 mRNA upregulation occurred at the transcriptional level, HeLa cells were pre-incubated with Act D, a potent repressor of *de novo* mRNA transcription, prior to treatment with SF. SF was then added drop-wise to treatment dishes without replacement of the Act D-containing media. If SF treatment does induce increased transcription of TBX3 in HeLa cells, no response would be seen in Act D + SF-treated cells. Conversely, if the mRNA products were being post-transcriptionally regulated by a component of the SF or as a result of SF treatment, then differences in mRNA levels would persist. From the results obtained **(Figure 3.5)** it is clear that Act D pretreatment abrogated the effects of SF treatment on HeLa cells, indicating that SF treatment transcriptionally regulates TBX3 expression. TBX2 expression was not induced by SF treatment in both vehicle and Act D-treated cells (data not shown), making it impossible to determine whether SF upregulation of TBX2 mRNA also occurs at the transcriptional level. The potential reasons for the discrepancy between these results and those reported in section 3.3 are discussed in section 4.

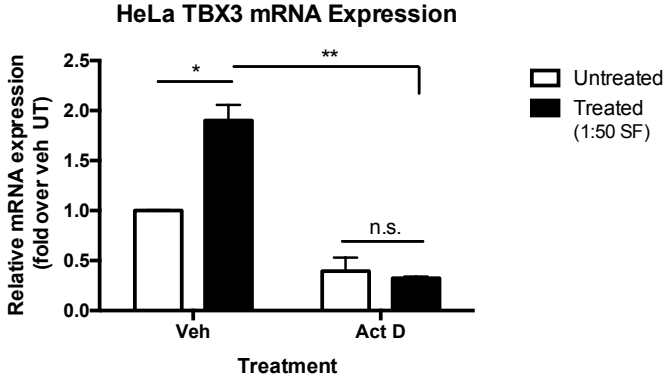

Figure 3.5: Actinomycin D treatment abrogates SF-induced TBX3 mRNA upregulation.
HeLa cells were seeded in 60 mm dishes and cultured for 24 hours. Cells were then washed with PBS and serum starved overnight. Thereafter they were washed again and fresh starving media was added. Dishes were pre-treated with Actinomycin D (Act D, 5 µg / mL) or with vehicle (Veh, DMSO) for 1 hour. One dish in each group was further treated with SF (1:50 final concentration) and the other left otherwise untreated for 4 hours. Thereafter cells were harvested by scraping and RNA was isolated, reverse transcribed and TBX3 mRNA was quantified by relative quantitation. Data represents the pooled means from 3 independent experiments. Treatment groups were assessed in triplicate within each experiment. Means ± SEM are displayed. Statistical significance was determined by normal Two-way ANOVA and Šídák's multiple comparison post-hoc test. * = $p < 0.05$; ** = $p < 0.01$; n.s. = not statistically significant.

3.5. TBX3 protein expression is increased in SF-treated HeLa and CaSki cells

While significant changes in both TBX2 and TBX3 mRNA were observed in SF-treated HeLa and CaSki cells relative to their respective untreated controls **(Figure 3.4)**, it remained to be established whether these changes persisted to the level of translation. To this end, we used western blotting to probe protein lysates from SF-treated and untreated cells with antibodies against TBX2 and TBX3. p38 was used as a loading control because although its phosphorylation state is variable, total levels of p38 protein do not increase in response to various treatments **(Figure 3.6)**.

TBX3 protein was detectable in both HeLa and CaSki cell lines and expression levels were upregulated by SF treatment. In CaSki cells, TBX3 protein expression was upregulated relative to untreated controls at 2 and 24 hours post-treatment **(Figure 3.6 a)**. In HeLa cells, TBX3 expression was initially undetectable in untreated cells and was only detectable in SF-treated cells 24 hours post-treatment **(Figure 3.6 a)**. In a subsequent experiment that included additional time points it was revealed that TBX3 protein is indeed detectable in untreated HeLa cells. Furthermore, SF treatment markedly increased protein expression at 2, 4, 8, and 24 hours post-treatment **(Figure 3.6 b)**.

Conversely, TBX2 protein was undetectable in both HeLa and CaSki cells at 2 and 24 hours in treated and untreated cells in all experiments **(Figure 3.6 c)**. Endogenously expressed TBX2 in MCF-7 cell lysates was used as a positive control and was reliably detectable, thus indicating the functionality of our assays. The lack of detectable TBX2 expression in HeLa and CaSki cells precluded the determination of an effect of SF treatment on TBX2 protein expression in HeLa or CaSki cells.

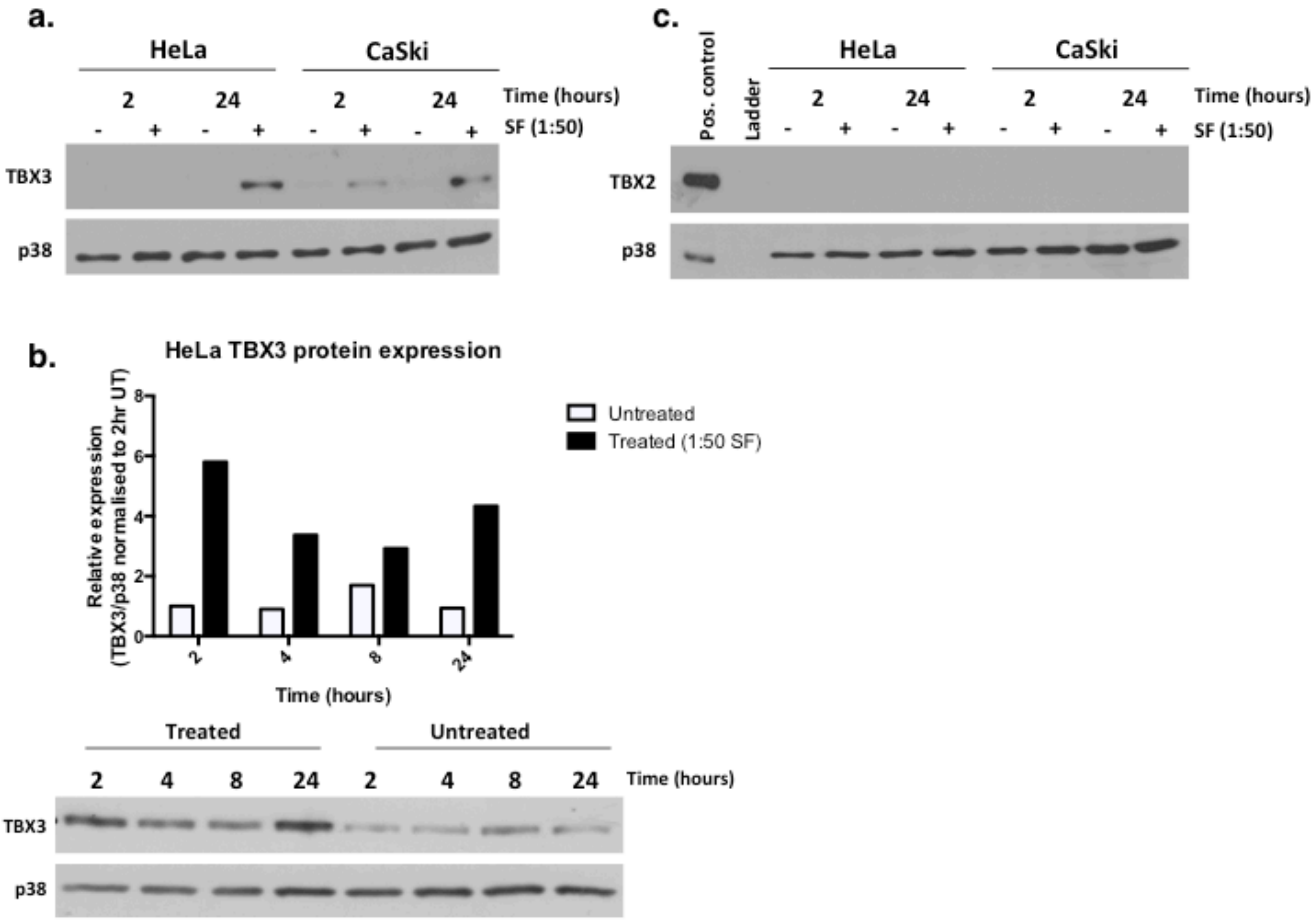

Figure 3.6: SF increases TBX3 protein expression in HeLa and CaSki cells.
HeLa and CaSki cells were cultured in 6-well plates for 24 hours. They were then washed with PBS and serum starved overnight. The following day, cells were washed again and fresh starving media was added to each well. Experimental wells were treated with SF (1:50 final concentration) while control wells were left untreated. Protein was then isolated at set time points and seperated by SDS-PAGE, transferred to nitrocellulose membranes and probed with antibodies against TBX3 (**a, b**) and TBX2 (**c**). MCF-7 protein lysates were used as TBX2 positive controls. TBX3 expression in **b** was quantified via relative densitometry by normalising TBX3 expression to to p38 and expressed as fold-change relative to the 2 hour untreated control. p38 was used as a loading control in all experiments. **b** is representative of two independent experiments, in which TBX2 remained undetectable (not shown).

3.6. SF treatment selectively increases mRNA and protein expression of the TBX3 target gene p21.

Owing to its primary role as a transcription factor, TBX3 upregulation is likely to result in differential gene transcription of downstream target genes. Therefore, well-established cancer-related TBX3 target genes were selected for investigation, namely PTEN and p21. Cells were treated with SF for 4 hours, as this was the time of maximal induction of TBX3 mRNA expression observed in earlier experiments **(Figure 3.4)**, and TBX3 protein levels were shown to be increased as early as 2 hours post-treatment **(figure 3.6)**.

PTEN mRNA expression was detectable in HeLa cells but expression levels were not affected by 4 hour SF incubation **(Figure 3.7 a)**. Similar results were observed in CaSki cells **(Figure 3.7 b).** These results demonstrate that SF does not regulate PTEN mRNA in either cell line.

p21 mRNA expression was significantly increased in HeLa cells treated with SF compared to untreated controls **(Figure 3.7 a)**; a statistically significant increase of ~3.5-fold was observed 4 hours after SF addition ($p < 0.01$). Furthermore, a general trend of increased mRNA expression was also observed in SF-treated CaSki cells **(figure 3.7 b)**, thus indicating that this effect does not exclusively occur in HeLa cells. Additionally, SF treatment also increased HeLa cell p21 protein expression relative to untreated controls at 2 and 4 hours in a western blot experiment **(Figure 3.7 c)**. Collectively, these results show that while SF does not regulate PTEN, it does upregulate p21 mRNA and protein expression.

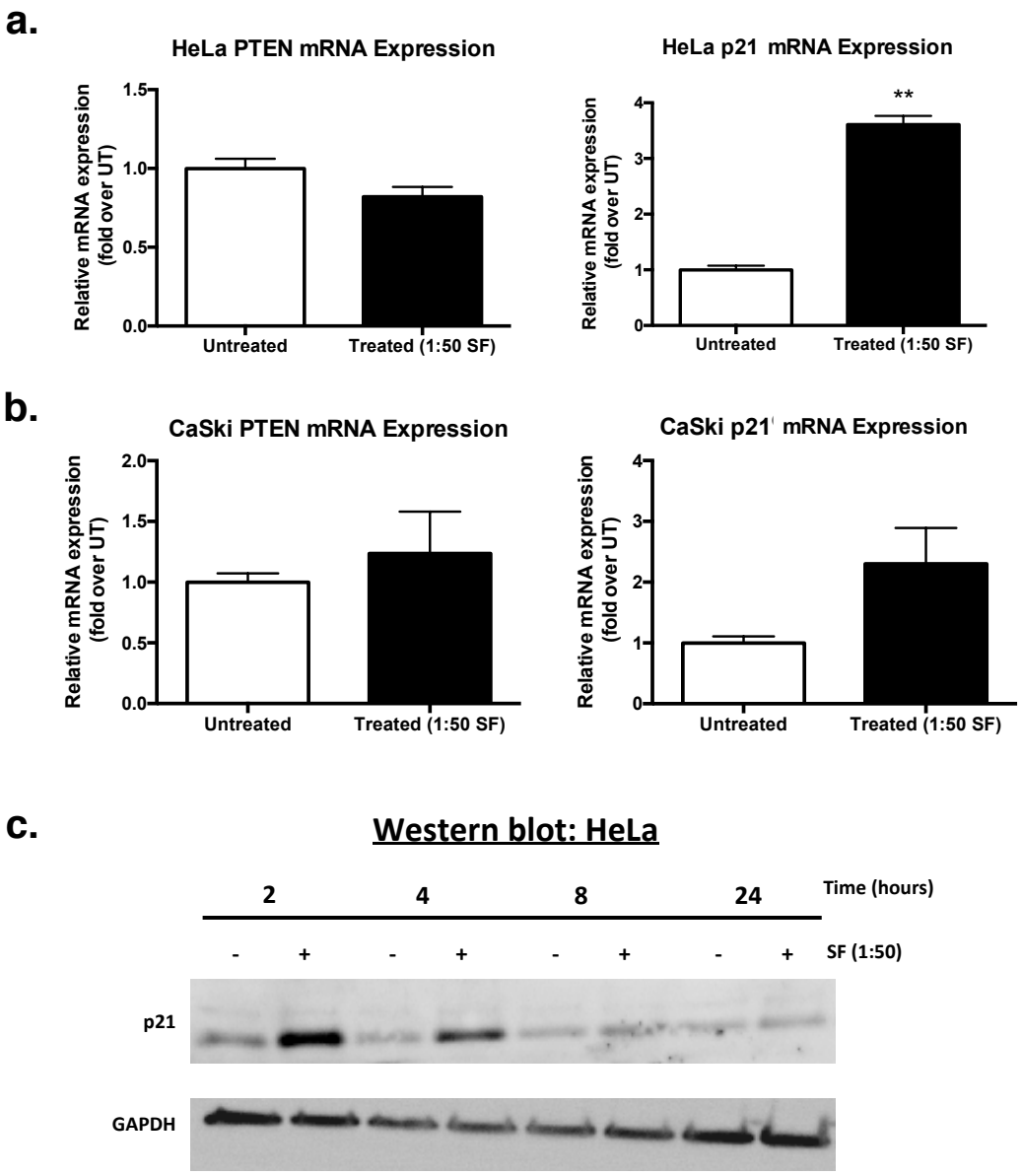

Figure 3.7: SF treatment selectively induces p21 mRNA and protein expression.
HeLa (**a)** and CaSki (**b**) cervical cancer cell lines were cultured in 60 mm dishes for 24 hours after seeding. Thereafter they were washed in PBS and serum starved overnight. Cells were then washed again and fresh starving media was added to each dish. Experimental dishes received seminal fluid (SF) in the media (1:50 final concentration) while control dishes were left untreated. Thereafter cells were harvested by scraping and RNA was isolated, reverse transcribed and PTEN and p21 mRNA was quantified by relative quantitation. A preliminary western blot of HeLa protein lysates was performed to detect p21 expression in SF-treated and untreated cells (**c**). GAPDH was used as a loading control. Each graph represents the pooled means from 3 independent experiments. Time points were assessed in triplicate within each experiment. Means ± SEM are displayed. Normal Two-way ANOVA and Šídák's multiple comparison post-hoc tests were used to determine statistical significance. ** = $p < 0.01$.

3.7. TBX3 knock down reduces p21 mRNA expression in SF-treated HeLa cells

TBX3 was transiently knocked down in HeLa cells to determine whether SF-mediated upregulation of p21 mRNA expression in HeLa cells occurred via a TBX3-related pathway. Cells were transiently transfected with siRNA targeting TBX3 (siTBX3) or with scrambled siRNA sequences as a control (siCtrl) and then treated with SF (1:50 final concentration) for 4 hours. siTBX3 HeLa cells treated with SF expressed significantly less TBX3 mRNA than SF-treated siCtrl cells (~30% of siCtrl, $p < 0.05$), indicating successful knockdown of TBX3 **(Figure 3.8)**. The result of TBX3 knockdown on p21 expression, also shown in **Figure 3.8**, indicates that SF-treated siTBX3 cells express significantly less p21 mRNA expression relative to SF-treated siCtrl cells (~50%, $p < 0.01$). Therefore TBX3 mRNA is not upregulated by SF treatment in siTBX3 cells as it is in siCtrls. This suggests that the SF-regulated induction of p21 is mediated by TBX3.

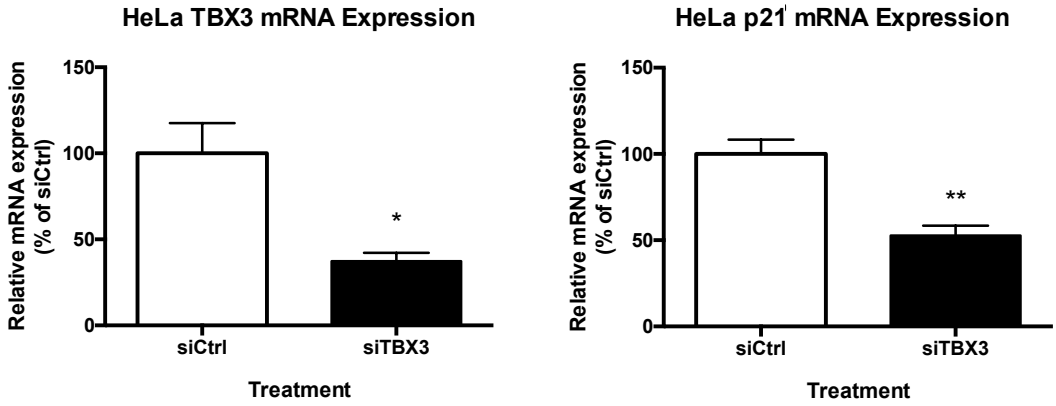

Figure 3.8: Knockdown of TBX3 reduces p21 expression in SF-treated HeLa cells.
HeLa cells were cultured in 60 mm dishes for 24 hours after seeding. Cells were then transfected with either an siRNA targeting TBX3 (siTBX3) or a scrambled siRNA sequence (siCtrl) and cultured overnight. They were then washed in PBS and serum starved overnight. Cells were washed again and fresh starving media was added to each dish. Experimental dishes received seminal fluid (SF) in the media (1:50 final concentration) while control dishes were left untreated. Cells were treated for 4 hours, after which they were harvested by scraping and RNA was isolated, reverse transcribed and TBX3 and p21 mRNA was quantified by relative quantitation. This figure represents the pooled mean data from 3 independent experiments. Gene expression in each treatment group was assessed in triplicate within each experiment. Statistical significance was determined by unpaired, two-tailed t-test. Means ± SEM are displayed. * = $p < 0.5$, ** = $p < 0.01$.

3.8. Examining the role of TBX3 in mediating SF-stimulation of HeLa cell proliferation

Collectively, our results thus far have shown that SF treatment upregulates TBX3 expression and through it p21 expression in HeLa cells. In other contexts, TBX3-mediated regulation of p21 alters cellular proliferation (Willmer, Cooper, *et al.*, 2016). Given that SF has previously been shown to increase the rate of HeLa cell proliferation (Sutherland *et al.*, 2012), we sought to investigate whether TBX3 plays a role in mediating this pro-proliferative regulation. SF-treated and untreated HeLa cells were cultured overnight and used as seeding stock for proliferation assays wherein cells were allowed to proliferate for set periods (1-7 days) of time before being counted. The results, displayed in **Figure 3.9**, indicated that while a slight trend of an increased growth rate may be evident in SF-treated cells, no statistically significant differences were observed between proliferation rates of treated or untreated cells at the investigated time points. Owing to the unexpected result that our SF does not induce increased proliferation, we could not test whether TBX3 plays a role in this process.

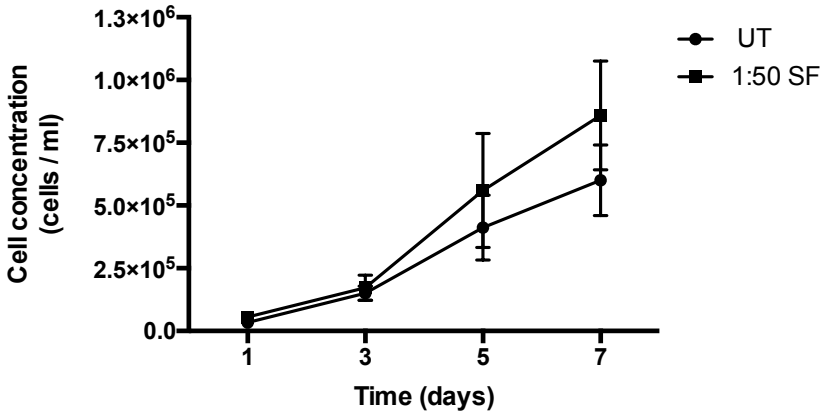

Figure 3.9: SF treatment does not significantly influence HeLa cell proliferation.
HeLa cells were treated with 1:50 SF (final concentration) (■) or left untreated (●), cultured overnight and subsequently used as seeding stock in proliferation assays spanning 7 days. Cells were lifted by trypsinisation and counted at 1, 3, 5 and 7 days to assess population growth. All populations were resuspended in the same volume of media. This figure represents the pooled means from 3 independent experiments. Time points were assessed in triplicate within each experiment. Means ± SEM are displayed. UT = untreated, SF = seminal fluid.

3.9. SF treatment does not influence HeLa cell migration

Because SF treatment did not significantly affect HeLa cell proliferation despite significant induction of TBX3-mediated p21 expression, we turned our attention to other aspects of cell physiology. We hypothesized that SF-mediated upregulation of TBX3 might result in decreased cell migratory ability as TBX3 overexpression does in fibrosarcoma cells (Willmer, Cooper, *et al.*, 2016). If high TBX3 levels do reduce cervical cancer cell migration rates, this effect might contribute to the mechanism underlying the observation that high levels of TBX3 mRNA expression are associated with decreased lymph node metastasis in cervical carcinomas (Lyng *et al.*, 2006). Wound healing assays were conducted using HeLa cells to determine whether SF treatment alters their rate of migration **(Figure 3.10)**. Cells were treated with 10 μg / mL mitomycin C, an antibiotic that causes DNA cross-linkage formation, to prevent proliferation which can confound observations of migration. Our results indicate that while migration was observed in both treatment conditions, SF-treated cells did not migrate at a significantly different rate relative to untreated controls at any time up to 24 hours post-wound creation. The absence of a treatment effect in our assays precluded investigation into the role of TBX3 in HeLa cell migration. Therefore, the explicit role of TBX3 upregulation by SF on cervical cell physiology and cervical cancer progression remains to be determined.

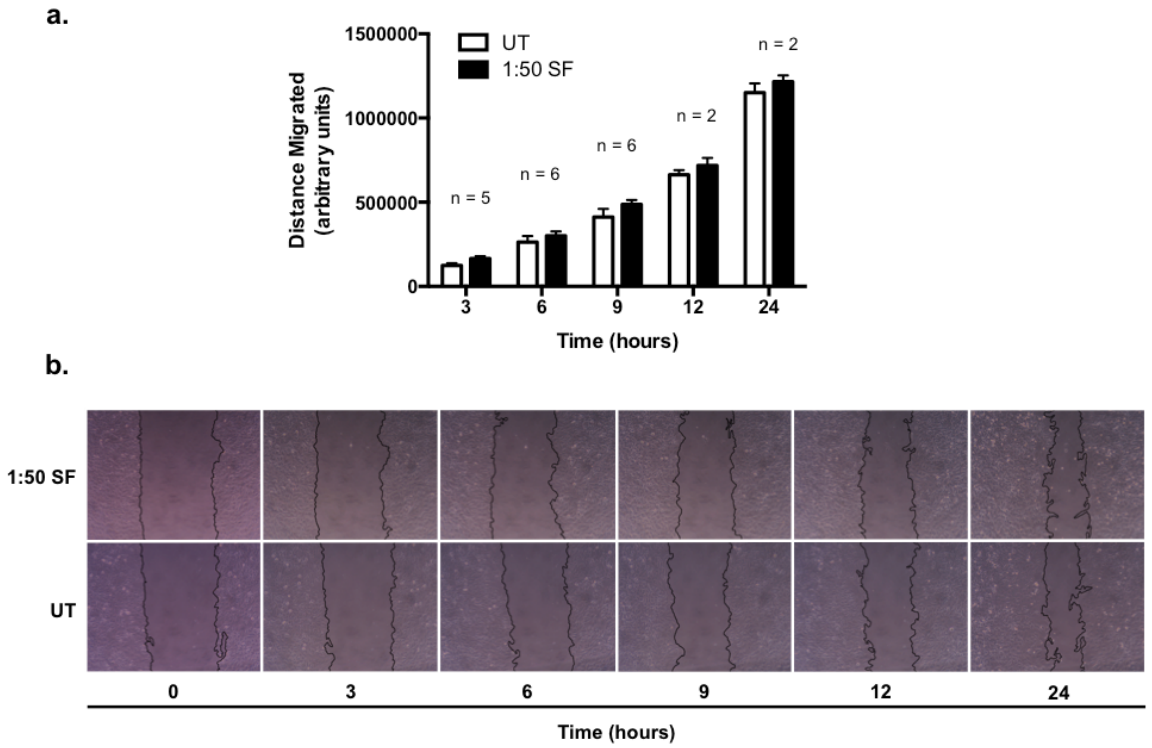

Figure 3.10: SF treatment does not influence HeLa cell migration.
HeLa cells were grown until confluent in 24-well plates after which a wound was induced in each well by scratching with a p20 pipette tip. Wells were washed with PBS to clear lifted cellular debris, and the remaining cells were incubated in media containing mitomycin C (10 µg/mL). Cells were further treated with 1:50 SF (final concentration) (black bars) or left otherwise untreated (white bars). Each well was photographed via light microscopy at three different locations along the wound at the specified times. ImageJ software was used to determine the area of each wound, after which distance migrated was calculated by subtracting the wound area in the 0 hour image of each location from the areas determined from later time points. **a** Represents the pooled means of independent experiments, with the number of replicates specified above each time point. **b** Displays a representative series of paired images from one location in one experiment, with wound margins emphasized for visibility. Means ± SEM are displayed. UT = untreated, SF = seminal fluid.

4. Discussion

This study sought to determine whether TBX2 and TBX3 play a role in mediating the effects of SF on cancerous cervical tissue. Given that neither TBX2 nor TBX3 are universally expressed in all adult tissues, we initially investigated the expression statuses of these proteins in cervical tissues. TBX2 protein is overexpressed in primary cervical squamous cell carcinomas relative to normal tissues (Liu, Jiang & Zhang, 2010b). While TBX3 mRNA is expressed in cervical carcinomas and was shown to be reduced in cancers with metastatic potential (Lyng *et al.*, 2006), the expression status of TBX3 in cervical cancers relative to normal cervical tissue is unknown. Our immunohistochemical analyses (Figure 3.1) showed that TBX3 protein is indeed overexpressed in both cervical adenocarcinoma and squamous cell carcinomas compared to normal cervical tissue samples, which exhibited little to no staining. Furthermore, adenocarcinoma and squamous cell carcinoma tissues also exhibited significant overexpression relative to matched adjacent normal cervical tissues from the same sections (Figure 3.2). These results indicate that TBX3 expression is overexpressed in cancerous cervical tissues. Lyng *et al.* (2006) showed that decreased TBX3 mRNA expression is an independent prognostic predictor of lymph node metastasis. High TBX3 expression formed part of a gene expression cluster associated with significantly improved progression-free survival. This cluster was associated with increased tumour volume and the composition thereof implied a pro-proliferative role. Collectively, these results imply that TBX3 expression is activated in the oncogenic transformation of normal cervical tissues to potentially play a pro-proliferative role, but that further differentiation is required wherein TBX3 expression is downregulated in order for the cancerous cells to metastasize beyond the cervix. In Melanomas, TBX3 plays a pro-migratory role and is required for tumour formation (Peres *et al*., 2010). It is tempting to speculate that it plays a similar role in the context of cervical cancers; TBX3 may be required for migration into the cervical stroma and subsequent tumour formation (i.e. progression from CIN3 to cancer).

How TBX3 would mediate an effect in cervical cancer is unclear. Our analyses identified positive staining in cancer cell nuclei and cytoplasm, as well as cases where cytoplasmic staining was high and nuclear staining was absent (Figure 3.3). TBX3 is known to localize in both cellular compartments and its distribution varies by cell cycle stage in SW1353 chondrosarcoma cells, where its presence is important for progression from S phase, when it is predominantly located in the nucleus, to G2 (Willmer *et al.*, 2015). A similar cell cycle-regulated localization may be present in cervical cancer cells, wherein TBX3 only exerts its effects in certain phases. Given its primary role as a transcription factor, it is tempting to conclude that its influence on gene expression is the sole mechanism whereby it contributes to oncogenesis. However, TBX3 has been shown to regulate mRNA splicing *in vivo* (Kumar *et al.*, 2014). Altered splicing patterns significantly contribute to oncogenic phenotypes in a number of cancers and therefore may also be influencing oncogenesis by altering the composition of protein isoforms in cervical cancer cells (He *et al.*, 2009; Singh & Eyras, 2016).

Our findings and those of Liu *et al*. (2010b) clearly indicate that both TBX2 and TBX3 are overexpressed in primary cancerous cervical tissues. However, their apparent functions in these tissues remained to be elucidated. Given their frequent roles as terminal effectors of signaling pathways (see section 1.3.3), we hypothesized they may mediate some of the effects exerted by SF on cervical tissues. To determine if this was the case, we treated three cervical cancer cell lines with SF and assessed the expression patterns of TBX2 and TBX3 mRNA and protein (Figure 3.4, Figure 3.6). Two cell lines, HeLa and CaSki, showed responses to treatment in this regard. TBX3 mRNA and protein expression were both reliably upregulated in both cell lines following treatment and expression was also detected in untreated cells of both lines. Furthermore, we showed that this increased expression is transcriptionally regulated in HeLa cells (Figure 3.5). While basal TBX3 protein expression has been reported for both cell lines in existing literature (Mowla *et al.*, 2011; Burgucu *et al.*, 2012), our finding that SF increases TBX3 expression in HeLa and CaSki cell lines is novel. Given that HeLa cells are derived from an adenocarcinoma whereas CaSki cells originate from a squamous cell carcinoma, TBX3 expression does not appear to be

limited to a particular cervical cancer subtype. This finding is in agreement with our *in vivo* investigations, which showed that both squamous cell carcinoma and adenocarcinoma tissues significantly overexpress TBX3 protein compared to normal tissue, as discussed above. Furthermore, our results indicate that SF is capable of inducing increased expression in both subtypes.

TBX2 mRNA was detectable in HeLa, but not CaSki, cells (Figure 3.4). While TBX2 mRNA expression was increased in response to SF treatment in HeLa cells, TBX2 protein was undetectable at any of the time points tested in SF-treated or untreated cells. In agreement with our mRNA findings, TBX2 protein was also undetectable in CaSki cells (Figure 3.6). The lack of TBX2 protein expression in untreated HeLa cells agrees with published findings, while no research on CaSki cells was found (Carreira *et al.*, 1998; Schneider *et al.*, 2013). To the best of our knowledge, no prior research has investigated TBX2 mRNA expression or the effects of SF treatment thereon in HeLa or CaSki cells. The contradiction in our observations that TBX2 mRNA is present and regulated by SF in HeLa cells while TBX2 protein remains undetectable in treated or untreated cells is puzzling. Despite being reliably detectable in multiple independent time point experiments (Figure 3.4), TBX2 mRNA expression was not increased by SF in Act D treatment experiments in either vehicle or Act D-treated cells (data not shown). The lack of TBX2 induction in these experiments was surprising and precluded assessment of the impact of Act D treatment on TBX2 mRNA expression. We can therefore offer no explanation for the apparent disconnect between TBX2 mRNA and protein expression observed in our experiments. The duration of storage of SF might explain the lack of TBX2 mRNA induction in Act D experiments. It may be that the component(s) of SF responsible for TBX2 induction degrade more rapidly than those influencing TBX3 (which continued to be reliably inducible in all experiments), even when SF is stored at -80°C. Given that a second, recently collected pool of SF was also able to induce increased TBX2 mRNA expression (data not shown), the relative age of SF seems to be an important factor. Nonetheless, this would not account for the contradiction between TBX2 protein and mRNA expression, as western blot experiments occurred at roughly the same time.

We also report the novel finding that C-33 A cells express both TBX2 and TBX3 mRNA (Figure 3.4). However, expression of either was unaffected by SF treatment, thus implying that context dependent factors are required for SF-mediated upregulation of these genes. Interestingly, C-33 A cells have tested negative for both HPV DNA and RNA in tests conducted by the ATCC, whereas both HeLa and CaSki cells test positive. Research by Schneider *et al.* (2013) has shown that TBX2 and TBX3 are capable of interacting with the HPV16 L2 capsid protein. Furthermore, TBX2 and TBX3 were shown to repress the HPV promoter region known as the long control region of a number of HPV strains. This repression was significantly enhanced by the presence of L2. TBX2 and HPV L2 were shown to colocalise in primary CIN 1 and -2 tissue samples. Other research has also shown that TBX2 protein expression in squamous cell carcinomas is significantly associated with HPV16 E7 detection (Liu, Jiang & Zhang, 2010b). Clearly, interactions between these TBX proteins and HPV proteins and genomes exist; while further research would be required to confirm any direct association, it is possible that SF-mediated TBX2 and TBX3 upregulation may be in some way dependent on the presence of HPV.

While a clear effect of SF treatment on TBX3 gene and protein expression was observed, the functional significance of this effect remained to be established. Therefore, we turned our attention to known downstream targets of TBX3's transcription factor activity, namely PTEN and p21. TBX3 has been inversely correlated with PTEN expression in previous work on head and neck squamous cell carcinomas (Burgucu *et al.*, 2012). Given that these cancers are also commonly associated with HPV16 (as well as other oncogenic HPV subtypes) presence it is tempting to speculate that similar oncogenic processes may be found in cervical carcinomas (Agalliu *et al.*, 2016). Indeed, Burgucu *et al.* (2012) showed that overexpression of TBX3 in HeLa cells resulted in repression of both basal and induced endogenous PTEN expression. However, PTEN mRNA was unaffected by SF treatment in either CaSki or HeLa cells in our study (Figure 3.7). It may be that the time point investigated (4 hours) was either too long or short to detect an effect or, alternatively, SF-mediated increases in TBX3 may not result in PTEN repression due

to other effects of SF. SF treatment did, however, successfully induce p21 mRNA expression in HeLa cells at 4 hours and induced protein expression at 2 and 4 hours, but not at later time points (Figure 3.7). Furthermore, transient knockdown of TBX3 resulted in decreased p21 mRNA expression in SF-treated HeLa cells, indicating that SF-mediated p21 upregulation is at least partially mediated via a TBX3-related pathway (Figure 3.8). In the majority of cases, TBX3 has been shown to repress p21 either directly or indirectly (Peres & Prince, 2013; Willmer, Hare, *et al.*, 2016). Yet, recent work from our laboratory has shown that TBX3 is capable of inducing p21 expression in HT1080 fibrosarcoma cells (Willmer, Cooper, *et al.*, 2016). The role of TBX3 in this context is oncoprotective, whereas the majority of SF studies using HeLa cells have shown that treatment induces oncogenic changes (see section 1.6). However, it cannot be assumed that TBX3-mediated increases in p21 will always have protective consequences. Indeed, overexpression of p21 is correlated with poor prognostic outcomes in a number of cancers, including cervical cancers, where its expression levels correlate with tumour stage (Bae *et al.*, 2001; Cheung *et al.*, 2001; reviewd in Abbas & Dutta, 2009). The importance of this SF-induced, TBX3-mediated upregulation of p21 expression therefore remains to be elucidated.

TBX3 and its interactions with p21 frequently result in altered rates of cell proliferation. TBX3-mediated repression of p21 is required for increased rates of cell proliferation in chondrosarcoma cell lines, and TBX3's ability to promote proliferation in mammary epithelial and hepatic progenitor cells is inversely correlated with p21 expression (Platonova *et al.*, 2007; Suzuki *et al.*, 2008; Willmer, Cooper, *et al.*, 2016; Willmer, Hare, *et al.*, 2016). Furthermore, when acting oncoprotectively in HT1080 fibrosarcomas, TBX3's induction of p21 expression decreases cell proliferation relative to cells in which TBX3 has been knocked down (Willmer, Cooper, *et al.*, 2016). We therefore set out to investigate whether SF-mediated TBX3 upregulation and the consequent increase in p21 expression influences HeLa cell proliferation. This would be investigated by comparing the proliferative rate of SF-treated HeLa cells with that of SF-treated HeLa cells in which TBX3 was knocked down. Our results, however, indicated no significant difference in the rate of proliferation between cells pre-treated with SF and untreated cells

(Figure 3.9), thus precluding investigation into a possible role for TBX3 or p21 in SF-mediated changes in HeLa cell proliferation. Previous investigations have shown that SF induces increased rates of HeLa cell proliferation, both *in vitro* and in mouse xenograft experiments (Liu *et al.*, 2011; Sutherland *et al.*, 2012). The discrepancy between our results and those in existing literature could be accounted for by several factors. Given the heterogeneous nature of SF, the pool of samples used in this study may lack the pro-proliferative induction ability observed in other studies. Alternatively, methodological differences may be involved. The *in vitro* findings by Sutherland *et al.* (2012) were established by ki67 expression assays, wherein expression levels were measured after 6 hours of incubation with SF. Our study assessed growth by physical counting of cells over much longer periods and assessed pre-treatment, rather than constant incubation. Nonetheless, owing to the lack of increased proliferation in SF-treated HeLa cells in our study, we could not determine a role for TBX3 or p21 in SF-mediated HeLa cell proliferation.

Besides influencing proliferation, TBX3 expression is also associated with increased migratory ability in melanoma, chondrosarcoma, breast, head and neck and bladder cancer cell lines. Furthermore, the expression status of p21 is also altered by TBX3 in many of these cell lines (Peres *et al*., 2010; Li, Weinberg, *et al.*, 2013; Du *et al.*, 2014; Willmer, Cooper, *et al.*, 2016). We therefore also examined the effects of SF on HeLa cell motility in order to determine whether SF-induced TBX3 and p21 expression may play a role therein. SF treatment failed to induce changes in the rate of cellular migration in our experiments (Figure 3.10). However, given the nature of our assays, we can only conclude that SF treatment does not increase the overall motility of HeLa cells. We cannot discount whether it may increase directional migration, which might be observed in transwell assays where only one compartment contains SF, nor can we discount whether SF would improve HeLa cell invasiveness. To our knowledge, no reports regarding the effects of SF on HeLa cell migration have been published. Future work will establish whether SF universally does not influence HeLa cell motility or whether our particular assays were unable to detect changes that are induced.

This study has several limitations. Firstly, the sample size of our immunohistochemical analysis was relatively small and requires replication in a larger study. As a consequence of the small size, no stratification by tumour stage, patient age, or other clinical variables could be made. Furthermore, some nonspecific staining was also observed in the stromal extracellular matrix, especially in low-staining normal tissues (Figure 3.3), or in highly necrotic cancerous tissues (not shown), indicating that current protocols could be optimized for these tissue types. Secondly, while we established that SF treatment does indeed influence TBX3 expression in two cell lines as well as TBX2 mRNA expression in HeLa cells, the discrepancy between this finding and the lack of TBX2 protein as well as mRNA expression in later experiments remains to be explained. Thirdly, we have yet to identify the mechanism whereby SF mediates this upregulation beyond confirming that it is transcriptional. Finally, we could not establish the functional significance of these altered expression states beyond target gene regulation, nor have we determined whether any changes are reproducible in primary cervical tissues.

5. Conclusion

This study has presented the novel finding that TBX3 protein is overexpressed in both cervical squamous cell carcinomas and adenocarcinomas relative to normal cervical epithelium. In an *in vitro* context, we found that SF is able to increase the expression of TBX2 mRNA, but not protein, only in HeLa cells and was able to induce TBX3 mRNA and protein expression in HeLa and CaSki cells but not in C-33 A cells. We also established that this effect was mediated at the transcriptional level. Furthermore, the resultant increase in TBX3 expression selectively influenced the expression of known target genes, and we report the second instance wherein TBX3 is shown to increase p21 expression rather than repress it. While we were able to demonstrate that SF-mediated upregulation of p21 was at least partially dependent on TBX3, we were unable to determine the functional significance of our findings. Future work focusing on this protein and its functions in cervical cancer progression could yield valuable insights into the underlying oncogenic mechanisms driving this disease.

6. References

Abbas, T. & Dutta, A. 2009. P21 in Cancer: Intricate Networks and Multiple Activities. *Nature Reviews Cancer*. 9(6):400–414. DOI: 10.1038/nrc2657.

Abrahams, A., Mowla, S., Parker, M.I., Goding, C.R. & Prince, S. 2008. UV-mediated regulation of the anti-senescence factor Tbx2. *Journal of Biological Chemistry*. 283(4):2223–2230. DOI: 10.1074/jbc.M705651200.

Abràmoff, M.D., Magalhães, P.J. & Ram, S.J. 2004. Image processing with imageJ. *Biophotonics International*. 11(7):36–42.

Adefuye, A. & Sales, K. 2012. Regulation of Inflammatory Pathways in Cancer and Infectious Disease of the Cervix. *Scientifica*. 2012:1–8. DOI: 10.6064/2012/548150.

Adefuye, A.O., Sales, K.J. & Katz, A. A. 2014. Seminal plasma induces the expression of IL-1α in normal and neoplastic cervical cells via EP2/EGFR/ PI3K/AKT pathway. *Journal of Molecular Signaling*. 9(1):8. DOI: 10.1186/1750-2187-9-8.

Agalliu, I., Gapstur, S., Chen, Z., Wang, T., Anderson, R.L., Teras, L., Kreimer, A.R., Hayes, R.B., et al. 2016. Associations of Oral α-, β-, and γ-Human Papillomavirus Types With Risk of Incident Head and Neck Cancer. *JAMA Oncology*. 10461:1–8. DOI: 10.1001/jamaoncol.2015.5504.

Agulnik, S.I., Garvey, N., Hancock, S., Ruvinsky, I., Chapman, D.L., Agulnik, I., Bollag, R., Papaioannou, V., et al. 1996. Evolution of mouse T-box genes by tandem duplication and cluster dispersion. *Genetics*. 144(1):249–254.

Allemani, C., Weir, H.K., Carreira, H., Harewood, R., Spika, D., Wang, X.S., Bannon, F., Ahn, J. V., et al. 2015. Global surveillance of cancer survival 1995-2009: Analysis of individual data for 25 676 887 patients from 279 population-based registries in 67 countries (CONCORD-2). *The Lancet*. 385(9972):977–1010. DOI: 10.1016/S0140-6736(14)62038-9.

American Cancer Society. 2014. Cancer Facts & Figures. *Cancer Facts and Figures*. Available: http://www.cancer.org/acs/groups/content/@research/documents/webcontent/acspc-042151.pdf.

Arribas, J., Giménez, E., Marcos, R. & Velázquez, A. 2015. Novel antiapoptotic effect of TBX15: overexpression of TBX15 reduces apoptosis in cancer cells. *Apoptosis*. 20(10):1338–1346. DOI: 10.1007/s10495-015-1155-8.

Atreya, I., Schimanski, C.C., Becker, C., Wirtz, S., Dornhoff, H., Schnurer, E., Berger, M.R., Galle, P.R., et al. 2007. The T-box transcription factor eomesodermin controls CD8 T cell activity and lymph node metastasis in human colorectal cancer. *Gut*. 56(11):1572–1578.

DOI: 10.1136/gut.2006.117812.

Bae, D.-S., Cho, S.-B., Kim, Y.-J., Whang, J.-D., Song, S.-Y., Park, C.-S., Kim, D.-S. & Lee, J.-H. 2001. Aberrant Expression of Cyclin D1 Is Associated with Poor Prognosis in Early Stage Cervical Cancer of the Uterus. *Gynecologic Oncology*. 81(3):341–347. DOI: 10.1006/gyno.2001.6196.

Bakker, M.L., Boink, G.J.J., Boukens, B.J., Verkerk, A.O., Van Den Boogaard, M., Den Haan, A. D., Hoogaars, W.M.H., Buermans, H.P., et al. 2012. T-box transcription factor TBX3 reprogrammes mature cardiac myocytes into pacemaker-like cells. *Cardiovascular Research*. 94(3):439–449. DOI: 10.1093/cvr/cvs120.

Ballif, B.C., Theisen, A., Rosenfeld, J.A., Traylor, R.N., Gastier-Foster, J., Thrush, D.L., Astbury, C., Bartholomew, D., et al. 2010. Identification of a Recurrent Microdeletion at 17q23.1q23.2 Flanked by Segmental Duplications Associated with Heart Defects and Limb Abnormalities. *American Journal of Human Genetics*. 86(3):454–461. DOI: 10.1016/j.ajhg.2010.01.038.

Bamshad, M., Lin, R.C., Law, D.J., Watkins, W.C., Krakowiak, P. A, Moore, M.E., Franceschini, P., Lala, R., et al. 1997. Mutations in human TBX3 alter limb, apocrine and genital development in ulnar-mammary syndrome. *Nature genetics*. 16:311–315. DOI: 10.1038/ng0797-311.

Bamshad, M., Le, T., Watkins, W.S., Dixon, M.E., Kramer, B.E., Roeder, A.D., Carey, J.C., Root, S., et al. 1999. The spectrum of mutations in TBX3: Genotype/Phenotype relationship in ulnar-mammary syndrome. *American journal of human genetics*. 64(6):1550–1562. DOI: 10.1086/302417.

Banister, C.E., Liu, C., Pirisi, L., Creek, K.E. & Buckhaults, P.J. 2017. Identification and characterization of HPV-independent cervical cancers. *Oncotarget*. (January, 5). DOI: 10.18632/oncotarget.14533.

Batruch, I., Lecker, I., Kagedan, D., Smith, C.R., Mullen, B.J., Grober, E., Lo, K.C., Diamandis, E.P., et al. 2011. Proteomic analysis of seminal plasma from normal volunteers and post-vasectomy patients identifies over 2000 proteins and candidate biomarkers of the urogenital system. *Journal of Proteome Research*. 10(3):941–953. DOI: 10.1021/pr100745u.

Bedford, J.M. 2015. The Functions--or Not--of Seminal Plasma? *Biology of Reproduction*. 92(1):18–18. DOI: 10.1095/biolreprod.114.126045.

Beukers, W., Kandimalla, R., Masius, R.G., Vermeij, M., Kranse, R., van Leenders, G.J. & Zwarthoff, E.C. 2015. Stratification based on methylation of TBX2 and TBX3 into three

molecular grades predicts progression in patients with pTa-bladder cancer. *Modern Pathology*. 28(4):515–522. DOI: 10.1038/modpathol.2014.145.

Bower, C. & Hansen, M. 2005. Assisted reproductive technologies and birth outcomes: overview of recent systematic reviews. *Reprod Fertil Dev*. 17(3):329–333. DOI: RD04095.

Boyd, S.C., Mijatov, B., Pupo, G.M., Tran, S.L., Gowrishankar, K., Shaw, H.M., Goding, C.R., Scolyer, R. a, et al. 2013. Oncogenic B-RAFV600E Signaling Induces the T-Box3 Transcriptional Repressor to Repress E-Cadherin and Enhance Melanoma Cell Invasion. *Journal of Investigative Dermatology*. 133(5):1269–1277. DOI: 10.1038/jid.2012.421.

Bromfield, J.J. 2014. Seminal fluid and reproduction: Much more than previously thought. *Journal of Assisted Reproduction and Genetics*. 31(6):627–636. DOI: 10.1007/s10815-014-0243-y.

Bruneau, B.G., Nemer, G., Schmitt, J.P., Charron, F., Robitaille, L., Caron, S., Conner, D. A., Gessler, M., et al. 2001. A murine model of Holt-Oram syndrome defines roles of the T-Box transcription factor Tbx5 in cardiogenesis and disease. *Cell*. 106(6):709–721. DOI: 10.1016/S0092-8674(01)00493-7.

Burgucu, D., Guney, K., Sahinturk, D., Ozbudak, I.H., Ozel, D., Ozbilim, G. & Yavuzer, U. 2012. Tbx3 represses PTEN and is over-expressed in head and neck squamous cell carcinoma. *BMC Cancer*. 12(1):481. DOI: 10.1186/1471-2407-12-481.

Bustin, S. a, Benes, V., Garson, J. A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., et al. 2009. The MIQE guidelines:Minimum Information for publication of quantitative real-time PCR experiments. *Clinical Chemistry*. 55(4):611–622. DOI: 10.1373/clinchem.2008.112797.

Campbell, C., Goodrich, K., Casey, G. & Beatty, B. 1995. DOI: 10.1006/geno.1995.1139.

Carlson, H., Ota, S., Campbell, C.E. & Hurlin, P.J. 2001. A dominant repression domain in Tbx3 mediates transcriptional repression and cell immortalization: relevance to mutations in Tbx3 that cause ulnar-mammary syndrome. *Human molecular genetics*. 10(21):2403–2413. DOI: 10.1093/hmg/10.21.2403.

Carreira, S., Dexter, T.J., Yavuzer, U., Easty, D.J. & Goding, C.R. 1998. Brachyury-related transcription factor Tbx2 and repression of the melanocyte-specific TRP-1 promoter. *Molecular and cellular biology*. 18(9):5099–5108.

Carreira, S., Liu, B. & Goding, C.R. 2000. The gene encoding the T-box factor Tbx2 is a target for the microphthalmia-associated transcription factor in melanocytes. *Journal of Biological Chemistry*. 275(29):21920–21927. DOI: 10.1074/jbc.M000035200.

Castle, P.E., Rodriguez, A.C., Burk, R.D., Herrero, R., Wacholder, S., Alfaro, M., Morales, J.,

Guillen, D., et al. 2009. Short term persistence of human papillomavirus and risk of cervical precancer and cancer: population based cohort study. *Bmj*. 339:b2569. DOI: 10.1136/bmj.b2569.

Cavard, C., Audebourg, A., Letourneur, F., Audard, V., Beuvon, F., Cagnard, N., Radenen, B., Varlet, P., et al. 2009. Gene expression profiling provides insights into the pathways involved in solid pseudopapillary neoplasm of the pancreas. *The Journal of Pathology*. 218:201–209. DOI: 10.1002/path.

Chang, F., Xing, P., Song, F., Du, X., Wang, G., Chen, K. & Yang, J. 2016. The role of T-box genes in the tumorigenesis and progression of cancer (Review). *Oncology Letters*. 1–7. DOI: 10.3892/ol.2016.5296.

Chen, J.C., Johnson, B. A., Erikson, D.W., Piltonen, T.T., Barragan, F., Chu, S., Kohgadai, N., Irwin, J.C., et al. 2014. Seminal plasma induces global transcriptomic changes associated with cell migration, proliferation and viability in endometrial epithelial cells and stromal fibroblasts. *Human Reproduction*. 29(6):1255–1270. DOI: 10.1093/humrep/deu047.

Chesson, H.W., Dunne, E.F., Hariri, S. & Markowitz, L.E. 2014. The estimated lifetime probability of acquiring human papillomavirus in the United States. *Sexually transmitted diseases*. 41(11):660–664. DOI: 10.1097/OLQ.0000000000000193.

Cheung, T.H., Lo, K.W.K., Yu, M.M.Y., Yim, S.F., Poon, C.S., Chung, T.K.H. & Wong, Y.F. 2001. Aberrant expression of p21WAF1/CIP1 and p27KIP1 in cervical carcinoma. *Cancer Letters*. 172(1):93–98. DOI: 10.1016/S0304-3835(01)00624-3.

Chung, J.H. & Eng, C. 2005. Nuclear-cytoplasmic partitioning of phosphatase and tensin homologue deleted on chromosome 10 (PTEN) differentially regulates the cell cycle and apoptosis. *Cancer Research*. 65(18):8096–8100. DOI: 10.1158/0008-5472.CAN-05-1888.

Clifford, G.M., de Vuyst, H., Tenet, V., Plummer, M., Tully, S. & Franceschi, S. 2016. Effect of HIV infection on human papillomavirus types causing invasive cervical cancer in Africa. *Journal of acquired immune deficiency syndromes (1999)*. DOI: 10.1097/QAI.0000000000001113.

Coll, M., Seidman, J.G. & Müller, C.W. 2002. Structure of the DNA-bound T-box domain of human TBX3, a transcription factor responsible for ulnar-mammary syndrome. *Structure*. 10(3):343–356. DOI: 10.1016/S0969-2126(02)00722-0.

Cooper, T.G., Noonan, E., von Eckardstein, S., Auger, J., Baker, H.W.G., Behre, H.M., Haugen, T.B., Kruger, T., et al. 2010. World Health Organization reference values for human semen characteristics. *Human Reproduction Update*. 16(3):231–245. DOI: 10.1093/humupd/dmp048.

Crawford, G., Ray, A., Gudi, A., Shah, A. & Homburg, R. 2015. The role of seminal plasma for improved outcomes during in vitro fertilization treatment: review of the literature and meta-analysis. *Human reproduction update*. 21(2):275–284. DOI: 10.1093/humupd/dmu052.

Davenport, T.G., Jerome-Majewska, L.A. & Papaioannou, V.E. 2003. Mammary gland, limb and yolk sac defects in mice lacking Tbx3, the gene mutated in human ulnar mammary syndrome. *Development (Cambridge, England)*. 130(10):2263–2273. DOI: 10.1242/dev.00431.

Davis, E., Teng, H., Bilican, B., Parker, M.I., Liu, B., Carriera, S., Goding, C.R. & Prince, S. 2008. Ectopic Tbx2 expression results in polyploidy and cisplatin resistance. *Oncogene*. 27(7):976–984. DOI: 10.1038/sj.onc.1210701.

Dimova, I., Orsetti, B., Negre, V., Rouge, C., Ursule, L., Lasorsa, L., Dimitrov, R., Doganov, N., et al. 2009. Genomic markers for ovarian cancer at chromosomes 1, 8 and 17 revealed by array CGH analysis. *Agro Food Industry Hi-Tech*. 20(3 SUPPL.):357–366.

Dobrzycka, K.M., Kang, K., Jiang, S., Meyer, R., Rao, P.H., Lee, A. V. & Oesterreich, S. 2006. Disruption of scaffold attachment factor B1 leads to TBX2 up-regulation, lack of p19ARF induction, lack of senescence, and cell immortalization. *Cancer Research*. 66(16):7859–7863. DOI: 10.1158/0008-5472.CAN-06-1381.

Doorbar, J., Egawa, N., Griffin, H., Kranjec, C. & Murakami, I. 2016. Human papillomavirus molecular biology and disease association. *Reviews in Medical Virology*. 25(1):2–23. DOI: 10.1002/rmv.1822.

Du, H.F., Ou, L.P., Yang, X., Song, X.D., Fan, Y.R., Tan, B., Luo, C.L. & Wu, X.H. 2014. A new PKCα/β/TBX3/E-cadherin pathway is involved in PLCε-regulated invasion and migration in human bladder cancer cells. *Cellular Signalling*. 26(3):580–593. DOI: 10.1016/j.cellsig.2013.11.015.

Esmailpour, T. & Huang, T. 2012. TBX3 promotes human embryonic stem cell proliferation and neuroepithelial differentiation in a differentiation stage-dependent manner. *Stem Cells*. 30(10):2152–2163. DOI: 10.1002/stem.1187.

Etcheverry, A., Aubry, M., de Tayrac, M., Vauleon, E., Boniface, R., Guenot, F., Saikali, S., Hamlat, A., et al. 2010. DNA methylation in glioblastoma: impact on gene expression and clinical outcome. *BMC genomics*. 11(1):701. DOI: 10.1186/1471-2164-11-701.

Fan, W., Huang, X., Chen, C., Gray, J. & Huang, T. 2004. TBX3 and Its Isoform TBX3 + 2a Are Functionally Distinctive in Inhibition of Senescence and Are Overexpressed in a Subset of Breast Cancer Cell Lines. *Cancer Research*. 64:5132–5139. DOI: 10.1158/0008-5472.CAN-

04-0615.

Farkas, S.A., Sorbe, B.G. & Nilsson, T.K. 2016. Epigenetic changes as prognostic predictors in endometrial carcinomas. *Epigenetics*. 00–00. DOI: 10.1080/15592294.2016.1252891.

Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., Parkin, D.M., Forman, D., et al. 2015. Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *International Journal of Cancer*. 136(5):E359–E386. DOI: 10.1002/ijc.29210.

Fernando, R., Litzinger, M., Trono, P., Hamilton, D.H., Schlom, J. & Palena, C. 2010. The T-box transcription factor Brachyury promotes epithelial-mesenchymal transition in human tumor cells. *The Journal of clinical …*. 120(2):533–544. DOI: 10.1172/JCI38379.(10).

Fischer, K. & Pflugfelder, G.O. 2015. Putative Breast Cancer Driver Mutations in TBX3 Cause Impaired Transcriptional Repression. *Frontiers in Oncology*. 5:244. DOI: 10.3389/fonc.2015.00244.

Fung, K.Y.C., Glode, L.M., Green, S. & Duncan, M.W. 2004. A comprehensive characterization of the peptide and protein constituents of human seminal fluid. *The Prostate*. 61(2):171–181. DOI: 10.1002/pros.20089.

Garland, S.M., Cheung, T.H., McNeill, S., Petersen, L.K., Romaguera, J., Vazquez-Narvaez, J., Bautista, O., Shields, C., et al. 2015. Safety and immunogenicity of a 9-valent HPV vaccine in females 12-26 years of age who previously received the quadrivalent HPV vaccine. *Vaccine*. 33(48):6855–6864. DOI: 10.1016/j.vaccine.2015.08.059.

Giacinti, C. & Giordano, A. 2006. RB and cell cycle progression. *Oncogene*. 25(38):5220–7. DOI: 10.1038/sj.onc.1209615.

Grulich, A.E., van Leeuwen, M.T., Falster, M.O. & Vajdic, C.M. 2007. Incidence of cancers in people with HIV/AIDS compared with immunosuppressed transplant recipients: a meta-analysis. *The Lancet*. 370(9447):59–67. DOI: 10.1016/S0140-6736(07)61050-2.

Guan, P., Howell-Jones, R., Li, N., Bruni, L., De Sanjosé, S., Franceschi, S. & Clifford, G.M. 2012. Human papillomavirus types in 115,789 HPV-positive women: A meta-analysis from cervical infection to cancer. *International Journal of Cancer*. 131(10):2349–2359. DOI: 10.1002/ijc.27485.

Hale, B.J., Keating, A.F., Yang, C.-X. & Ross, J.W. 2015. The Male Role in Pregnancy Loss and Embryo Implantation Failure. *Advances in experimental medicine and biology*. 868:49–79. DOI: 10.1007/978-3-319-18881-2.

Han, Y., Tu, W.-W., Wen, Y.-G., Yan, D.-W., Qiu, G.-Q., Peng, Z.-H. & Zhou, C.-Z. 2013. Increased expression of TBX2 is a novel independent prognostic biomarker of a worse

outcome in colorectal cancer patients after curative surgery and a potential therapeutic target. *Medical Oncology*. 30:688. DOI: 10.1007/s12032-013-0688-3.

Hanahan, D. & Weinberg, R. A. 2011. Hallmarks of cancer: The next generation. *Cell*. 144(5):646–674. DOI: 10.1016/j.cell.2011.02.013.

Harrelson, Z., Kelly, R.G., Goldin, S.N., Gibson-Brown, J.J., Bollag, R.J., Silver, L.M. & Papaioannou, V.E. 2004. Tbx2 is essential for patterning the atrioventricular canal and for morphogenesis of the outflow tract during heart development. *Development*. 131(20):5041–5052. DOI: 10.1242/dev.01378.

Hart, R. & Norman, R.J. 2013. The longer-term health outcomes for children born as a result of IVF treatment: Part I-general health outcomes. *Human Reproduction Update*. 19(3):232–243. DOI: 10.1093/humupd/dms062.

He, C., Zhou, F., Zuo, Z., Cheng, H. & Zhou, R. 2009. A global view of cancer-specific transcript variants by subtractive transcriptome-wide analysis. *PLoS ONE*. 4(3). DOI: 10.1371/journal.pone.0004732.

He, M.-L., Chen, Y., Peng, Y., Jin, D., Du, D., Wu, J., Lu, P., Lin, M.C., et al. 2002. Induction of apoptosis and inhibition of cell growth by developmental regulator hTBX5. *Biochemical and biophysical research communications*. 297(2):185–192. DOI: 10.1016/S0006-291X(02)02142-3.

Hellemans, J., Mortier, G., De Paepe, A., Speleman, F. & Vandesompele, J. 2007. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome biology*. 8(2):R19. DOI: 10.1186/gb-2007-8-2-r19.

Herfs, M., Yamamoto, Y., Laury, A., Wang, X., Nucci, M.R., McLaughlin-Drubin, M.E., Münger, K., Feldman, S., et al. 2012. A discrete population of squamocolumnar junction cells implicated in the pathogenesis of cervical cancer. *Proceedings of the National Academy of Sciences of the United States of America*. 109(26):10516–21. DOI: 10.1073/pnas.1202684109.

Herrero, R., González, P. & Markowitz, L.E. 2015. DOI: 10.1016/S1470-2045(14)70481-4.

Hoogaars, W.M.H., Engel, A., Brons, J.F., Verkerk, A.O., De Lange, F.J., Wong, L.Y.E., Bakker, M.L., Clout, D.E., et al. 2007. Tbx3 controls the sinoatrial node gene program and imposes pacemaker function on the atria. *Genes and Development*. 21(9):1098–1112. DOI: 10.1101/gad.416007.

Hoogaars, W.M.H., Barnett, P., Rodriguez, M., Clout, D.E., Moorman, A.F.M., Goding, C.R. & Christoffels, V.M. 2008. TBX3 and its splice variant TBX3 + exon 2a are functionally similar.

Pigment Cell and Melanoma Research. 21(3):379–387. DOI: 10.1111/j.1755-148X.2008.00461.x.

Hu, B., Mu, H.-P., Zhang, Y.-Q., Su, C.-Y., Song, J.-T., Meng, C. & Liu, D.-X. 2014. Prognostic significance of TBX2 expression in non-small cell lung cancer. *Journal of molecular histology*. DOI: 10.1007/s10735-014-9569-0.

Huang, B., Cohen, J.R., Fernando, R.I., Hamilton, D.H., Litzinger, M.T., Hodge, J.W. & Palena, C. 2013. The embryonic transcription factor Brachyury blocks cell cycle progression and mediates tumor resistance to conventional antitumor therapies. *Cell Death & Disease*. 4(6):e682. DOI: 10.1038/cddis.2013.208.

Huang, Y., Li, Z., Zhong, Q., Li, G., Zhang, Y. & Huang, Z. 2014. Association of TBX2 and P21 expression with clinicopathological features and survival of laryngeal squamous cell carcinoma. *Int J Clin Exp Med*. 7(12):5394–5402.

IARC. 2012. Biological Agents. *IARC Monogr Eval Carcinog Risks Hum*. 100B:1–443.

Imajyo, I., Sugiura, T., Kobayashi, Y., Shimoda, M., Ishii, K., Akimoto, N., Yoshihama, N., Kobayashi, I., et al. 2012. T-box transcription factor Brachyury expression is correlated with epithelial-mesenchymal transition and lymph node metastasis in oral squamous cell carcinoma. *International Journal of Oncology*. 41(6):1985–1995. DOI: 10.3892/ijo.2012.1673.

Ismail, A. & Bateman, A. 2009. Expression of TBX2 promotes anchorage-independent growth and survival in the p53-negative SW13 adrenocortical carcinoma. *Cancer Letters*. 278(2):230–240. DOI: 10.1016/j.canlet.2009.01.006.

Jacobs, J.J., Keblusek, P., Robanus-Maandag, E., Kristel, P., Lingbeek, M., Nederlof, P.M., van Welsem, T., van de Vijver, M.J., et al. 2000. Senescence bypass screen identifies TBX2, which represses Cdkn2a (p19(ARF)) and is amplified in a subset of human breast cancers. *Nature genetics*. 26(3):291–299. DOI: 10.1038/81583.

Jeremias, J., David, S.S., Toth, M. & Witkin, S.S. 1997. Induction of messenger RNA for the 70 kDa heat shock protein in HeLa cells and the human endocervix following exposure to semen: Implications for antisperm antibody production and susceptibility to sexually transmitted infections. *Human Reproduction*. 12(9):1915–1919. DOI: 10.1093/humrep/12.9.1915.

Joseph, T., Zalenskaya, I. A, Sawyer, L.C., Chandra, N. & Doncel, G.F. 2013. Seminal plasma induces prostaglandin-endoperoxide synthase (PTGS) 2 expression in immortalized human vaginal cells: involvement of semen prostaglandin E2 in PTGS2 upregulation. *Biology of reproduction*. 88(1):13. DOI: 10.1095/biolreprod.112.101956.

Joura, E. A., Giuliano, A.R., Iversen, O.-E., Bouchard, C., Mao, C., Mehlsen, J., Moreira, E.D., Ngan, Y., et al. 2015. A 9-valent HPV vaccine against infection and intraepithelial neoplasia in women. *The New England journal of medicine*. 372(8). DOI: 10.1056/NEJMoa1405044.

Kandimalla, R., Van Tilborg, A.A.G., Kompier, L.C., Stumpel, D.J.P.M., Stam, R.W., Bangma, C.H. & Zwarthoff, E.C. 2012. Genome-wide analysis of CpG Island methylation in bladder cancer identified TBX2, TBX3, GATA2, and ZIC4 as pTa-specific prognostic markers. *European Urology*. 61(6):1245–1256. DOI: 10.1016/j.eururo.2012.01.011.

Kavka, A.I. & Green, J.B.A. 1997. Tales of tails: Brachyury and the T-box genes. *Biochimica et Biophysica Acta - Reviews on Cancer*. 1333(2):73–84. DOI: 10.1016/S0304-419X(97)00016-4.

Kho, E.M., McCowan, L.M.E., North, R.A., Roberts, C.T., Chan, E., Black, M.A., Taylor, R.S. & Dekker, G.A. 2009. Duration of sexual relationship and its effect on preeclampsia and small for gestational age perinatal outcome. *Journal of Reproductive Immunology*. 82(1):66–73. DOI: 10.1016/j.jri.2009.04.011.

Kim, B., Choi, Y., Rah, S., Park, D.-R., Park, S., Chung, Y., Park, S.-M., Park, J.K., et al. 2015. Seminal CD38 is a pivotal regulator for fetomaternal tolerance. *Proceedings of the National Academy of Sciences*. 112(5):1559–1564. DOI: 10.1073/pnas.1413493112.

Kim, K.-A., Yolamanova, M., Zirafi, O., Roan, N.R., Staendker, L., Forssmann, W.-G., Burgener, A., Dejucq-Rainsford, N., et al. 2010. Semen-mediated enhancement of HIV infection is donor-dependent and correlates with the levels of SEVI. *Retrovirology*. 7:55. DOI: 10.1186/1742-4690-7-55.

Kimura, H., Fukui, A., Fujii, S., Yamaguchi, E., Kasai, G. & Mizunuma, H. 2009. Timed sexual intercourse facilitates the recruitment of uterine CD56 bright natural killer cells in women with infertility. *American Journal of Reproductive Immunology*. 62(2):118–124. DOI: 10.1111/j.1600-0897.2009.00720.x.

Kispert, A. & Herrmann, B.G. 1993. The Brachyury gene encodes a novel DNA binding protein. *The EMBO journal*. 12(8):3211–3220.

Krainova, N. A., Khaustova, N. A., Makeeva, D.S., Fedotov, N.N., Gudim, E. A., Ryabenko, E. A., Shkurnikov, M.U., Galatenko, V. V., et al. 2013. Evaluation of potential reference genes for qRT-PCR data normalization in HeLa cells. *Applied Biochemistry and Microbiology*. 49(9):743–749. DOI: 10.1134/S0003683813090032.

Krstic, M., Macmillan, C.D., Leong, H.S., Clifford, A.G., Souter, L.H., Dales, D.W., Postenka, C.O., Chambers, A.F., et al. 2016. The transcriptional regulator TBX3 promotes progression from non-invasive to invasive breast cancer. *BMC Cancer*. 16(1):671. DOI: 10.1186/s12885-

016-2697-z.

Kumar, P., Franklin, S., Emechebe, U., Hu, H., Moore, B., Lehman, C., Yandell, M. & Moon, A.M. 2014. TBX3 Regulates Splicing In Vivo: A Novel Molecular Mechanism for Ulnar-Mammary Syndrome. *PLoS Genetics*. 10(3). DOI: 10.1371/journal.pgen.1004247.

Kyrou, D., Kolibianakis, E.M., Devroey, P. & Fatemi, H.M. 2010. Is the use of donor sperm associated with a higher incidence of preeclampsia in women who achieve pregnancy after intrauterine insemination? *Fertility and Sterility*. 93(4):1124–1127. DOI: 10.1016/j.fertnstert.2008.12.021.

Law, D.J., Gebuhr, T., Garvey, N., Agulnik, S.I. & Silver, L.M. 1995. Identification, characterization, and localization to Chromosome 17q21-22 of the human TBX2 homolog, member of a conserved developmental gene family. *Mammalian Genome*. 6(11):793–797. DOI: 10.1007/BF00539006.

Leevers, S.J., Vanhaesebroeck, B. & Waterfield, M.D. 1999. Signalling through phosphoinositide 3-kinases: The lipids take centre stage. *Current Opinion in Cell Biology*. 11(2):219–225. DOI: 10.1016/S0955-0674(99)80029-5.

Li, D.M. & Sun, H. 1998. PTEN/MMAC1/TEP1 suppresses the tumorigenicity and induces G1 cell cycle arrest in human glioblastoma cells. *Proceedings of the National Academy of Sciences of the United States of America*. 95(26):15406–15411. DOI: 10.1073/pnas.95.26.15406.

Li, G., Yang, Q., Zhu, Y., Wang, H.R., Chen, X., Zhang, X. & Lu, B. 2013. T-Bet and Eomes Regulate the Balance between the Effector/Central Memory T Cells versus Memory Stem Like T Cells. *PLoS ONE*. 8(6):1–10. DOI: 10.1371/journal.pone.0067401.

Li, J., Weinberg, M.S., Zerbini, L. & Prince, S. 2013. The oncogenic TBX3 is a downstream target and mediator of the TGF-β1 signaling pathway. *Molecular biology of the cell*. 24(22):3569–76. DOI: 10.1091/mbc.E13-05-0273.

Li, J., Ballim, D., Rodriguez, M., Cui, R., Goding, C.R., Teng, H. & Prince, S. 2014. The Anti-proliferative Function of the TGF-β1 Signaling Pathway Involves the Repression of the Oncogenic TBX2 by Its Homologue TBX3. *Journal of Biological Chemistry*. 289(51):35633–35643. DOI: 10.1074/jbc.M114.596411.

Lingbeek, M.E., Jacobs, J.J.L. & Van Lohuizen, M. 2002. The T-box repressors TBX2 and TBX3 specifically regulate the tumor suppressor gene p14ARF via a variant T-site in the initiator. *Journal of Biological Chemistry*. 277(29):26120–26127. DOI: 10.1074/jbc.M200403200.

Liu, L., Liu, C., Lou, F., Zhang, G., Wang, X., Fan, Y., Yan, K., Wang, K., et al. 2011. Activation of telomerase by seminal plasma in malignant and normal cervical epithelial cells. *Journal of*

Pathology. 225(2):203–211. DOI: 10.1002/path.2914.

Liu, W.K., Jiang, X.Y. & Zhang, Z.X. 2010a. Expression of PSCA, PIWIL1, and TBX2 in endometrial adenocarcinoma. *Onkologie*. 33(5):241–245. DOI: 10.1159/000305098.

Liu, W.K., Jiang, X.Y. & Zhang, Z.X. 2010b. Expression of PSCA, PIWIL1 and TBX2 and its correlation with HPV16 infection in formalin-fixed, paraffin-embedded cervical squamous cell carcinoma specimens. *Archives of Virology*. 155(5):657–663. DOI: 10.1007/s00705-010-0635-y.

Lomnytska, M., Dubrovska, A., Hellman, U., Volodko, N. & Souchelnytskyi, S. 2006. Increased expression of cSHMT, Tbx3 and utrophin in plasma of ovarian and breast cancer patients. *International Journal of Cancer*. 118(2):412–421. DOI: 10.1002/ijc.21332.

Lu, J., Li, X.P., Dong, Q., Kung, H.F. & He, M.L. 2010. TBX2 and TBX3: The special value for anticancer drug targets. *Biochimica et Biophysica Acta - Reviews on Cancer*. 1806(2):268–274. DOI: 10.1016/j.bbcan.2010.07.001.

Lyng, H., Brøvig, R.S., Svendsrud, D.H., Holm, R., Kaalhus, O., Knutstad, K., Oksefjell, H., Sundfør, K., et al. 2006. Gene expressions and copy numbers associated with metastatic phenotypes of uterine cervical cancer. *BMC genomics*. 7:268. DOI: 10.1186/1471-2164-7-268.

Machtinger, R., Laurent, L.C. & Baccarelli, A.A. 2016. Extracellular vesicles: Roles in gamete maturation, fertilization and embryo implantation. *Human Reproduction Update*. 22(2):182–193. DOI: 10.1093/humupd/dmv055.

Mahlamäki, E.H., Bärlund, M., Tanner, M., Gorunova, L., Höglund, M., Karhu, R. & Kallioniemi, A. 2002. Frequent amplification of 8q24, 11q, 17q, and 20q-specific genes in pancreatic cancer. *Genes, Chromosomes and Cancer*. 35(4):353–358. DOI: 10.1002/gcc.10122.

McCune, K., Bhat-Nakshatri, P., Thorat, M.A., Nephew, K.P., Badve, S. & Nakshatri, H. 2010. Prognosis of hormone-dependent breast cancers: Implications of the presence of dysfunctional transcriptional networks activated by insulin via the immune transcription factor T-bet. *Cancer Research*. 70(2):685–696. DOI: 10.1158/0008-5472.CAN-09-1530.

Meneghini, V., Odent, S., Platonova, N., Egeo, A. & Merlo, G.R. 2006. Novel TBX3 mutation data in families with Ulnar-Mammary syndrome indicate a genotype-phenotype relationship: Mutations that do not disrupt the T-domain are associated with less severe limb defects. *European Journal of Medical Genetics*. 49(2):151–158. DOI: 10.1016/j.ejmg.2005.04.021.

Meuleman, T., Snaterse, G., Beelen, E. Van, Anholts, J.D.H. & Pilgram, G.S.K. 2015. The

immunomodulating effect of seminal plasma on T cells. *Journal of Reproductive Immunology*. 110:1–8. DOI: 10.1016/j.jri.2015.01.012.

Miao, Z.F., Liu, X.Y., Xu, H.M., Wang, Z.N., Zhao, T.T., Song, Y.X., Xing, Y.N., Huang, J.Y., et al. 2016. Tbx3 overexpression in human gastric cancer is correlated with advanced tumor stage and nodal status and promotes cancer cell growth and invasion. *Virchows Archiv*. 469(5):505–513. DOI: 10.1007/s00428-016-2007-9.

Moreno, V., Munoz, N., Bosch, F.X., de Sanjose, S., Gonzalez, L.C., Tafur, L., Gili, M., Izarzugaza, I., et al. 1995. Risk factors for progression of cervical intraepithelial neoplasm grade III to invasive cervical cancer. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*. 4(5):459–467.

Moreno, V., Bosch, F.X., Muñoz, N., Meijer, C.J.L.M., Shah, K. V., Walboomers, J.M.M., Herrero, R. & Franceschi, S. 2002. Effect of oral contraceptives on risk of cervical cancer in women with human papillomavirus infection: The IARC multicentric case-control study. *Lancet*. 359(9312):1085–1092. DOI: 10.1016/S0140-6736(02)08150-3.

Mowla, S., Pinnock, R., Leaner, V.D., Goding, C.R. & Prince, S. 2011. PMA-induced up-regulation of TBX3 is mediated by AP-1 and contributes to breast cancer cell migration. *The Biochemical journal*. 433(1):145–153. DOI: 10.1042/BJ20100886.

Mucci, L.A., Hjelmborg, J.B., Harris, J.R., Czene, K., Havelick, D.J., Scheike, T., Graff, R.E., Holst, K., et al. 2016. Familial Risk and Heritability of Cancer Among Twins in Nordic Countries Lorelei. *JAMA : the journal of the American Medical Association*. 2115(1):68–76. DOI: 10.1001/jama.2015.17703.

Muller, M., Sales, K.J., Katz, A. A. & Jabbour, H.N. 2006. Seminal plasma promotes the expression of tumorigenic and angiogenic genes in cervical adenocarcinoma cells via the E-series prostanoid 4 receptor. *Endocrinology*. 147(7):3356–3365. DOI: 10.1210/en.2005-1429.

Munoz, N., Franceschi, S., Bosetti, C., Moreno, V., Herrero, R., Smith, J.S., Shah, K., Meijer, C.J.L.M., et al. 2002. Role of parity and human papillomavirus in cervical cancer: the IARC multicentric case-control study. TT - [Rôle de la parité et du papillomavirus humain dans le cancer du col de l'utérus : étude multi-centrée de contrôle de cas de l'IARC]; [Papel de la. *Lancet*. 359(9312):1093–1101. DOI: 10.1016/S0140-6736(02)08151-5.

Nimmakayalu, M., Major, H., Sheffield, V., Solomon, D.H., Smith, R.J., Patil, S.R. & Shchelochkov, O.A. 2011. Microdeletion of 17q22q23.2 encompassing TBX2 and TBX4 in a patient with congenital microcephaly, thyroid duct cyst, sensorineural hearing loss, and

pulmonary hypertension. *American Journal of Medical Genetics, Part A*. 155(2):418–423. DOI: 10.1002/ajmg.a.33827.

Nocera, M. & Chu, T. 1993. Transforming growth factor β as an immunosuppressive protein in human seminal plasma. *American Journal of Reproductive Immunology*. 30(1):1–8.

Packham, E.A. & Brook, J.D. 2003. T-box genes in human disorders. *Human molecular genetics*. 12 Spec No(1):R37–R44. DOI: 10.1093/hmg/ddg077.

Pagès, F., Berger, A., Camus, M., Sanchez-Cabo, F., Costes, A., Molidor, R., Mlecnik, B., Kirilovsky, A., et al. 2005. Effector Memory T Cells, Early Metastasis, and Survival in Colorectal Cancer. *New England Journal of Medicine*. 353(25):2654–2666. DOI: 10.1056/NEJMoa051424.

Palena, C., Polev, D.E., Tsang, K.Y., Fernando, R.I., Litzinger, M., Krukovskaya, L.L., Baranova, A.V., Kozlov, A.P., et al. 2007. The human T-box mesodermal transcription factor Brachyury is a candidate target for T-cell - Mediated cancer immunotherapy. *Clinical Cancer Research*. 13(8):2471–2478. DOI: 10.1158/1078-0432.CCR-06-2353.

Papaioannou, V.E. 2014. The T-box gene family: emerging roles in development, stem cells and cancer. *Development*. 141(20):3819–3833. DOI: 10.1242/dev.104471.

Park, J.C., Chae, Y.K., Son, C.H., Kim, M.S., Lee, J., Ostrow, K., Sidransky, D., Hoque, M.O., et al. 2008. Epigenetic silencing of human T (brachyury homologue) gene in non-small-cell lung cancer. *Biochemical and Biophysical Research Communications*. 365(2):221–226. DOI: 10.1016/j.bbrc.2007.10.144.

Paxton, C., Zhao, H., Chin, Y., Langner, K. & Reecy, J. 2002. Murine Tbx2 contains domains that activate and repress gene transcription. *Gene*. 283(1–2):117–124. DOI: 10.1016/S0378-1119(01)00878-2.

Peres, J. & Prince, S. 2013. The T-box transcription factor, TBX3, is sufficient to promote melanoma formation and invasion. *Molecular cancer*. 12(1):117. DOI: 10.1186/1476-4598-12-117.

Peres, J., Davis, E., Mowla, S., Bennett, D.C., Li, J.A., Wansleben, S. & Prince, S. 2010. The Highly Homologous T-Box Transcription Factors, TBX2 and TBX3, Have Distinct Roles in the Oncogenic Process. *Genes & cancer*. 1(3):272–282. DOI: 10.1177/1947601910365160.

Peres, J., Mowla, S. & Prince, S. 2014. The T-box transcription factor, TBX3, is a key substrate of AKT3 in melanomagenesis. *Oncotarget*. 6(3):1821–1833.

Perkhofer, L., Walter, K., Costa, I.G., Carrasco, M.C.R., Eiseler, T., Hafner, S., Genze, F., Zenke, M., et al. 2016. Tbx3 fosters pancreatic cancer growth by increased angiogenesis and Activin/Nodal-dependent induction of stemness. *Stem Cell Research*. 17(2):367–378. DOI:

10.1016/j.scr.2016.08.007.

Pilch, B. & Mann, M. 2006. Large-scale and high-confidence proteomic analysis of human seminal plasma. *Genome biology*. 7(5):R40. DOI: 10.1186/gb-2006-7-5-r40.

Platonova, N., Scotti, M., Babich, P., Bertoli, G., Mento, E., Meneghini, V., Egeo, A., Zucchi, I., et al. 2007. TBX3, the gene mutated in ulnar-mammary syndrome, promotes growth of mammary epithelial cells via repression of p19ARF, independently of p53. *Cell and Tissue Research*. 328:301–316. DOI: 10.1007/s00441-006-0364-4.

Plummer, M., Peto, J. & Franceschi, S. 2012. Time since first sexual intercourse and the risk of cervical cancer. *International Journal of Cancer*. 130(11):2638–2644. DOI: 10.1002/ijc.26250.

Plummer, M., De Martel, C., Vignat, J., Bray, F. & Franceschi, S. 2016. Global burden of cancers attributable to infections in 2012: A review and synthetic analysis. *Lancet Glob Health*. 4(9):e609–e616. DOI: http://dx.doi.org/10.1016/ S2214-109X(16)30143-7.

Politch, J.A., Tucker, L., Bowman, F.P. & Anderson, D.J. 2007. Concentrations and significance of cytokines and other immunologic factors in semen of healthy fertile men. *Human Reproduction*. 22(11):2928–2935. DOI: 10.1093/humrep/dem281.

Prakash, M., Patterson, S., Gotch, F. & Kapembwa, M.S. 2003. Recruitment of CD4+ T lymphocytes and macrophages into the cervical epithelium of women after coitus. *American Journal of Obstetrics and Gynecology*. 188(2):376–381. DOI: 10.1067/mob.2003.16.

Prince, S., Carreira, S., Vance, K.W., Abrahams, A. & Goding, C.R. 2004. Tbx2 Directly Represses the Expression of the p21 WAF1 Cyclin-Dependent Kinase Inhibitor Kinase Inhibitor. *Cancer research*. 64(0):1669–1674. DOI: 10.1158/0008-5472.CAN-03-3286.

Pudney, J. 2005. Immunological Microenvironments in the Human Vagina and Cervix: Mediators of Cellular Immunity Are Concentrated in the Cervical Transformation Zone. *Biology of Reproduction*. 73(6):1253–1263. DOI: 10.1095/biolreprod.105.043133.

Qi, T., Han, J., Cui, Y., Zong, M., Liu, X. & Zhu, B. 2008. Comparative proteomic analysis for the detection of biomarkers in pancreatic ductal adenocarcinomas. *Journal of clinical pathology*. 61(1):49–58. DOI: 10.1136/jcp.2006.044735.

Renard, C.A., Labalette, C., Armengol, C., Cougot, D., Wei, Y., Cairo, S., Pineau, P., Neuveut, C., et al. 2007. Tbx3 is a downstream target of the Wnt/β-catenin pathway and a critical mediator of β-catenin survival functions in liver cancer. *Cancer Research*. 67(3):901–910. DOI: 10.1158/0008-5472.CAN-06-2344.

Robertson, S.A. & Sharkey, D.J. 2016. Seminal fluid and fertility in women. *Fertility and*

sterility. 106(3):511–519. DOI: 10.1016/j.fertnstert.2016.07.1101.

Rodriguez, A.C., Schiffman, M., Herrero, R., Hildesheim, A., Bratti, C., Sherman, M.E., Solomon, D., Guillén, D., et al. 2010. Longitudinal study of human papillomavirus persistence and cervical intraepithelial neoplasia grade 2/3: Critical role of duration of infection. *Journal of the National Cancer Institute*. 102(5):315–324. DOI: 10.1093/jnci/djq001.

Rodriguez, M., Aladowicz, E., Lanfrancone, L. & Goding, C.R. 2008. Tbx3 represses E-cadherin expression and enhances melanoma invasiveness. *Cancer Research*. 68(19):7872–7881. DOI: 10.1158/0008-5472.CAN-08-0301.

Rosenbluh, J., Nijhawan, D., Cox, A.G., Li, X., Neal, J.T., Schafer, E.J., Zack, T.I., Wang, X., et al. 2012. β-Catenin-driven cancers require a YAP1 transcriptional complex for survival and tumorigenesis. *Cell*. 151(7):1457–1473. DOI: 10.1016/j.cell.2012.11.026.

Russell, R., Ilg, M., Lin, Q., Wu, G., Lechel, A., Bergmann, W., Eiseler, T., Linta, L., et al. 2015. A Dynamic Role of TBX3 in the Pluripotency Circuitry. *Stem Cell Reports*. 5(6):1155–1170. DOI: 10.1016/j.stemcr.2015.11.003.

Saftlas, A.F., Rubenstein, L., Prater, K., Harland, K.K., Field, E. & Triche, E.W. 2014. Cumulative exposure to paternal seminal fluid prior to conception and subsequent risk of preeclampsia. *Journal of Reproductive Immunology*. 101–102(1):104–110. DOI: 10.1016/j.jri.2013.07.006.

Sakabe, M., Kokubo, H., Nakajima, Y. & Saga, Y. 2012. Ectopic retinoic acid signaling affects outflow tract cushion development through suppression of the myocardial Tbx2-Tgf 2 pathway. *Development*. 139:385–395. DOI: 10.1242/dev.067058.

Sales, K.J. & Katz, A. A. 2012. Inflammatory pathways in cervical cancer – the University of Cape Town ' s contribution. *S Afr Med J*. 102(6):493–496.

Sales, K.J., Adefuye, A., Nicholson, L. & Katz, A. A. 2014. CCR5 expression is elevated in cervical cancer cells and is up-regulated by seminal plasma. *Molecular Human Reproduction*. 20(11):1144–1157. DOI: 10.1093/molehr/gau063.

Sarkar, D., Shields, B., Davies, M.L., Müller, J. & Wakeman, J.A. 2012. BRACHYURY confers cancer stem cell characteristics on colorectal cancer cells. *International Journal of Cancer*. 130(2):328–337. DOI: 10.1002/ijc.26029.

Schiffman, M.H., Haley, N.J., Felton, J.S., Andrews, A.W., Kaslow, R.A., Lancaster, W.D., Kurman, R.J., Brinton, L.A., et al. 1987. Biochemical epidemiology of cervical neoplasia: measuring cigarette smoke constituents in the cervix. *Cancer research*. 47(14):3886–8. Available: http://www.ncbi.nlm.nih.gov/pubmed/3594446.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., et al. 2012. Fiji: an open-source platform for biological-image analysis. *Nature Methods*. 9(7):676–682. DOI: 10.1038/nmeth.2019.

Schjenken, J. & Robertson, S. 2014. Seminal Fluid and Immune Adaptation for Pregnancy - Comparative Biology in Mammalian Species. *Reproduction in Domestic Animals*. 49:27–36. DOI: 10.1111/rda.12383.

Schneider, M.A., Scheffer, K.D., Bund, T., Boukhallouk, F., Lambert, C., Cotarelo, C., Pflugfelder, G.O., Florin, L., et al. 2013. The transcription factors TBX2 and TBX3 interact with human papillomavirus 16 (HPV16) L2 and repress the long control region of HPVs. *Journal of virology*. 87(8):4461–74. DOI: 10.1128/JVI.01803-12.

Sebé-Pedrós, A. & Ruiz-Trillo, I. 2017. Evolution and Classification of the T-Box Transcription Factor Family. In *Current Topics in Developmental Biology*. 1–26. DOI: 10.1016/bs.ctdb.2016.06.004.

Shan, Z., Yan, X.-B., Yan, L.-L., Tian, Y., Meng, Q.-C., Qiu, W.-W., Zhang, Z. & Jin, Z.-M. 2015. Overexpression of Tbx3 is correlated with Epithelial-Mesenchymal Transition phenotype and predicts poor prognosis of colorectal cancer. *AM J Cancer Res*. 5(1):344–353.

Sharkey, D.J., Macpherson, A.M., Tremellen, K.P. & Robertson, S.A. 2007. Seminal plasma differentially regulates inflammatory cytokine gene expression in human cervical and vaginal epithelial cells. *Molecular Human Reproduction*. 13(7):491–501. DOI: 10.1093/molehr/gam028.

Sharkey, D.J., Tremellen, K.P., Jasper, M.J., Gemzell-Danielsson, K. & Robertson, S. A. 2012. Seminal fluid induces leukocyte recruitment and cytokine and chemokine mRNA expression in the human cervix after coitus. *Journal of immunology (Baltimore, Md. : 1950)*. 188(5):2445–54. DOI: 10.4049/jimmunol.1102736.

Sharkey, D.J., Macpherson, A.M., Tremellen, K.P., Mottershead, D.G., Gilchrist, R.B. & Robertson, S. A. 2012. TGF-β Mediates Proinflammatory Seminal Fluid Signaling in Human Cervical Epithelial Cells. *The Journal of Immunology*. 189(2):1024–1035. DOI: 10.4049/jimmunol.1200005.

Sharkey, D.J., Tremellen, K.P., Briggs, N.E., Dekker, G.A. & Robertson, S.A. 2016. Seminal plasma transforming growth factor-β, activin A and follistatin fluctuate within men over time. *Human Reproduction*. 31(10):2183–2191. DOI: 10.1093/humrep/dew185.

Sheeba, C.J. & Logan, M.P.O. 2017. The Roles of T-Box Genes in Vertebrate Limb Development. In *Current Topics in Developmental Biology*. 1st ed. V. 122. Elsevier Inc. 355–381. DOI: 10.1016/bs.ctdb.2016.08.009.

Shi, Y., Li, L., Hu, Z., Li, S., Wang, S., Liu, J., Wu, C., He, L., et al. 2013. A genome-wide association study identifies two new cervical cancer susceptibility loci at 4q12 and 17q12. *Nature genetics*. 45(8):918–22. DOI: 10.1038/ng.2687.

Shimoda, M., Sugiura, T., Imajyo, I., Ishii, K., Chigita, S., Seki, K., Kobayashi, Y. & Shirasuna, K. 2012. The T-box transcription factor Brachyury regulates epithelial-mesenchymal transition in association with cancer stem-like cells in adenoid cystic carcinoma cells. *BMC Cancer*. 12(1):377. DOI: 10.1186/1471-2407-12-377.

Siegel, R.L., Miller, K.D. & Jemal, A. 2016. Cancer statistics, 2016. *CA Cancer J Clin*. 66(1):7–30. DOI: 10.3322/caac.21332.

Singh, B. & Eyras, E. 2016. The role of alternative splicing in cancer. *Transcription*. 0. DOI: 10.1080/21541264.2016.1268245.

Sinha, S., Abraham, S., Gronostajski, R.M. & Campbell, C.E. 2000. Differential DNA binding and transcription modulation by three T-box proteins, T, TBX1 and TBX2. *Gene*. 258(1–2):15–29. DOI: 10.1016/S0378-1119(00)00417-0.

Smith, G.N., Walker, M., Tessier, J.L. & Millar, K.G. 1997. Increased incidence of preeclampsia in women conceiving by intrauterine insemination with donor versus partner sperm for treatment of primary infertility. *American Journal of Obstetrics and Gynecology*. 177(2):455–458. DOI: 10.1016/S0002-9378(97)70215-1.

Steinau, M., Rajeevan, M.S. & Unger, E.R. 2006. DNA and RNA references for qRT-PCR assays in exfoliated cervical cells. *The Journal of molecular diagnostics : JMD*. 8(1):113–118. DOI: 10.2353/jmoldx.2006.050088.

Stewart, B.W. & Wild, C. 2014. World cancer report 2014. *Geneva: WHO*.

Strickler, H.D., Palefsky, J.M., Shah, K. V, Anastos, K., Klein, R.S., Minkoff, H., Duerr, A., Massad, L.S., et al. 2003. Human papillomavirus type 16 and immune status in human immunodeficiency virus-seropositive women. *Journal of the National Cancer Institute*. 95(14):1062–1071.

Sutherland, J.R., Sales, K.J., Jabbour, H.N. & Katz, A. A. 2012. Seminal plasma enhances cervical adenocarcinoma cell proliferation and tumour growth in vivo. *PLoS ONE*. 7(3):1–12. DOI: 10.1371/journal.pone.0033848.

Suzuki, A., Sekiya, S., Büscher, D., Izpisúa Belmonte, J.C. & Taniguchi, H. 2008. Tbx3 controls the fate of hepatic progenitor cells in liver development by suppressing p19ARF expression. *Development (Cambridge, England)*. 135(9):1589–1595. DOI: 10.1242/dev.016634.

Szabo, S.J., Sullivan, B.M., Peng, S.L. & Glimcher, L.H. 2003. Molecular Mechanisms

Regulating Th1 Immune Responses. *Annual Review of Immunology*. 21(1):713–758. DOI: 10.1146/annurev.immunol.21.120601.140942.

Tada, M. & Smith, J.C. 2001. T-targets: Clues to understanding the functions of T-box proteins. *Development Growth and Differentiation*. 43(1):1–11. DOI: 10.1046/j.1440-169X.2001.00556.x.

Tremellen, K.P. 2000. The effect of intercourse on pregnancy rates during assisted human reproduction. *Human Reproduction*. 15(12):2653–2658. DOI: 10.1093/humrep/15.12.2653.

Trempus, C.S., Wei, S., Humble, M.M., Dang, H., Carl, D., Sifre, M.I., Kissling, G.E., Sunman, J.A., et al. 2011. A novel role for the T-box transcription factor Tbx1 as a negative regulator of tumor cell growth in mice. *Molecular carcinogenesis*. 50:981–91. DOI: 10.1002/mc.20768.A.

Vance, K.W., Carreira, S., Brosch, G. & Goding, C.R. 2005. Tbx2 Is Overexpressed and Plays an Important Role in Maintaining Proliferation and Suppression of Senescence in Melanomas. *Cancer Research*. 65(6):2260–2268. DOI: 10.1158/0008-5472.CAN-04-3045.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A. & Speleman, F. 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome biology*. 3(7):RESEARCH0034. DOI: 10.1186/gb-2002-3-7-research0034.

Veldhuijzen, N.J., Snijders, P.J.F., Reiss, P., Meijer, C.J.L.M. & van de Wijgert, J.H.H.M. 2010. Factors affecting transmission of mucosal human papillomavirus. *The Lancet Infectious Diseases*. 10(12):862–874. DOI: 10.1016/S1473-3099(10)70190-0.

Vojtech, L., Woo, S., Hughes, S., Levy, C., Ballweber, L., Sauteraud, R.P., Strobl, J., Westerberg, K., et al. 2014. Exosomes in human semen carry a distinctive repertoire of small non-coding RNAs with potential regulatory functions. *Nucleic Acids Research*. 42(11):7290–7304. DOI: 10.1093/nar/gku347.

Wang, B., Lindley, L.E., Fernandez-Vega, V., Rieger, M.E., Sims, A.H. & Briegel, K.J. 2012. The T box transcription factor TBX2 promotes epithelial-mesenchymal transition and invasion of normal and malignant breast epithelial cells. *PLoS ONE*. 7(7). DOI: 10.1371/journal.pone.0041355.

Wang, H., Meng, Q., Shan, Z., Yuan, Z. & Huang, X. 2015. Overexpression of Tbx3 Predicts Poor Prognosis of Patients with Resectable Pancreatic Carcinoma. *Asian Pacific Journal of Cancer Prevention*. 16(4):1397–1401. DOI: 10.7314/APJCP.2015.16.4.1397.

Wang, J.X., Knottnerus, A.M., Schuit, G., Norman, R.J., Chan, A. & Dekker, G.A. 2002. Surgically obtained sperm, and risk of gestational hypertension and pre-eclampsia. *Lancet*.

359(9307):673–674. DOI: 10.1016/S0140-6736(02)07804-2.

Wang, X., Zhuang, J., Wu, K., Xu, R., Li, M. & Lu, Y. 2010. Human semen: The biological basis of sexual behaviour to promote human papillomavirus infection and cervical cancer. *Medical Hypotheses*. 74(6):1015–1016. DOI: 10.1016/j.mehy.2010.01.009.

Wansleben, S., Davis, E., Peres, J. & Prince, S. 2013. A novel role for the anti-senescence factor TBX2 in DNA repair and cisplatin resistance. *Cell death & disease*. 4(10):e846. DOI: 10.1038/cddis.2013.365.

Wansleben, S., Peres, J., Hare, S., Goding, C.R. & Prince, S. 2014. T-Box Transcription Factors in Cancer Biology. *BBA - Reviews on Cancer*. 1846(2):380–391. DOI: 10.1016/j.bbcan.2014.08.004.

Weber, J.A., Baxter, D.H., Zhang, S., Huang, D.Y., Huang, K.H., Lee, M.J., Galas, D.J. & Wang, K. 2010. The microRNA spectrum in 12 body fluids. *Clinical Chemistry*. 56(11):1733–1741. DOI: 10.1373/clinchem.2010.147405.

Willmer, T. 2016. The role and regulation of the T-box transcription factor 3 in soft tissue and bone sarcomas.

Willmer, T., Peres, J., Mowla, S., Abrahams, A. & Prince, S. 2015. The T-Box factor TBX3 is important in S-phase and is regulated by c-Myc and cyclin A-CDK2. *Cell Cycle*. 14(19):3173–3183. DOI: 10.1080/15384101.2015.1080398.

Willmer, T., Cooper, A., Sims, D., Govender, D. & Prince, S. 2016. The T-box transcription factor 3 is a promising biomarker and a key regulator of the oncogenic phenotype of a diverse range of sarcoma subtypes. *Oncogenesis*. 5(2):e199. DOI: 10.1038/oncsis.2016.11.

Willmer, T., Hare, S., Peres, J. & Prince, S. 2016. The T-box transcription factor TBX3 drives proliferation by direct repression of the p21WAF1 cyclin-dependent kinase inhibitor. *Cell Division*. 11(1):6. DOI: 10.1186/s13008-016-0019-0.

Wilson, V. & Conlon, F.L. 2002. The T-box family. *Genome biology*. 3(6):REVIEWS3008. DOI: 10.1186/gb-2002-3-6-reviews3008.

Yamashita, S., Tsujino, Y., Moriguchi, K., Tatematsu, M. & Ushijima, T. 2006. Chemical genomic screening for methylation-silenced genes in gastric cancer cell lines using 5-aza-2'-deoxycytidine treatment and oligonucleotide microarray. *Cancer Science*. 97(1):64–71. DOI: 10.1111/j.1349-7006.2006.00136.x.

Yang, L., Xie, S., Feng, X., Chen, Y., Zheng, T., Dai, M., Ke Zhou, C., Hu, Z., et al. 2015. Worldwide Prevalence of Human Papillomavirus and Relative Risk of Prostate Cancer: A Meta-analysis. *Scientific Reports*. 5:14667. DOI: 10.1038/srep14667.

Yarosh, W., Barrientos, T., Esmailpour, T., Lin, L., Carpenter, P.M., Osann, K., Anton-Culver,

H. & Huang, T. 2008. TBX3 is overexpressed in breast cancer and represses p14ARF by interacting with histone deacetylases. *Cancer Research*. 68(3):693–699. DOI: 10.1158/0008-5472.CAN-07-5012.

Yu, H., Liu, B., Liu, A., Li, K. & Zhao, H. 2015. T-box 2 expression predicts poor prognosis in gastric cancer. *Oncology Letters*. 10(3):1689–1693. DOI: 10.3892/ol.2015.3428.

Yu, J., Ma, X., Cheung, K.F., Li, X., Tian, L., Wang, S., Wu, C.W., Wu, W.K.K., et al. 2010. Epigenetic inactivation of T-box transcription factor 5, a novel tumor suppressor gene, is associated with colon cancer. *Oncogene*. 29(49):6464–6474. DOI: 10.1038/onc.2010.370.

Zhao, D., Wu, Y. & Chen, K. 2014. Tbx3 isoforms are involved in pluripotency maintaining through distinct regulation of Nanog transcriptional activity. *Biochemical and Biophysical Research Communications*. 444(3):411–414. DOI: 10.1016/j.bbrc.2014.01.093.

Zhu, B., Zhang, M., Byrum, S.D., Tackett, A.J. & Davie, J.K. 2014. TBX2 blocks myogenesis and promotes proliferation in rhabdomyosarcoma cells. *International Journal of Cancer*. 135:785–797. DOI: 10.1002/ijc.28721.

Zhu, B., Zhang, M., Williams, E.M., Keller, C., Mansoor, A. & Davie, J.K. 2016. TBX2 represses PTEN in rhabdomyosarcoma and skeletal muscle. *Oncogene*. 35(32):4212–4224. DOI: 10.1038/onc.2015.486.

Zong, M., Meng, M. & Li, L. 2011. Low expression of TBX4 predicts poor prognosis in patients with stage II pancreatic ductal adenocarcinoma. *International Journal of Molecular Sciences*. 12(8):4953–4963. DOI: 10.3390/ijms12084953.

7. Appendices

7.1. Appendix I – SF collection consent form

INSTITUTE OF INFECTIOUS DISEASE AND MOLECULAR MEDICINE
Wernher and Beit North
Faculty of Health Sciences

Anzio Road, Observatory
Tel: (+27) 21 406 6268
Fax: (+27) 21 406 6061
E-Mail: arieh.katz@uct.ac.za

Patients Information Sheet and
Seminal Fluid Collection Consent Form

Role of seminal fluid and inflammatory pathways/agents on cervical cancer progression and HIV infection, UCT HREC REF: 0849/2015

Why is this study being done?

Cervical cancer is one of the leading causes of cancer-related deaths in women. The prevalence of this disease is particularly high in South Africa. This research study is aimed at investigating the biological process responsible for the development and progression of cervical cancer.

HIV/Aids is also a big health problem in Africa and the world. We still do not know everything about how HIV infects people.

There is evidence implicating seminal fluid (the fluid of a mans ejaculate, without the sperm) may play a role in how infective HIV is in women. There is also evidence that it may play a role in cervical cancer formation or progression.

The study will investigate role of seminal fluid on HIV infection and cervical cancer.

Why are you being asked to take part?

As a patient of the Groote Schuur Fertility clinic, you are already providing a semen sample today as part of your medical treatment. Once the clinic has

analysed your sample, we would like to use the what's left over for our study. You don't need to contribute anything more than you already would have done today.

How many people will take part in the study?

A total of 100 men will take part in this study.

How long will the study last?

The study will last until we have used all of the seminal fluid, or for one year.

What do we do to decide if you are eligible to be taking part?

To be eligible, you must have healthy semen. A technician at the Groote Schuur Hospital Fertility Clinic will confirm that your semen is healthy according to World Health Organisation guidelines.

What will happen if you decide to take part in the study?

You will be asked to sign an informed consent form. Once the fertility clinic is done with your semen sample that was provided today, we will receive it from them. We receive them in numbered tubes. There is no way of knowing whose is whose once we collect them.

What are the risks and discomforts of this study?

There are no risks or discomforts to you if you participate in this study.

Are there any benefits to you for being in the study?

Other than contributing to the advancement of scientific research, there are no benefits to participating in this study.

What other choices do you have?

You may choose to not take part in our study. Your choice to participate or not will have no influence on your future treatment at this or any other hospital.

What will happen when the study is over?

The research will be analysed and presented as part of one or more theses for degree purposes. The results may be used in academic publications.

Will your test results be shared with you?

We will not produce results that are relevant to you as a patient. No results will be shared with you directly.

Will the results of the research be shared with you?

We will not share the results of our research with you directly. However, university theses and journal publications resulting from the research will be available to the public.

Will any of your blood, tissue or other samples be stored and used for research in the future?

Your semen sample will be stored until it is used, or for one year after collection.

Will you receive any reward (money or food vouchers) for taking part in this study?

There are no rewards for participating in this study.

Who will see the information that is collected about you during the study?

We will not collect personal information about you during this study.

Who do I speak to (or contact) if I have any questions about the study?

You are encouraged to discuss any worries or issues you may have with the investigators of this project.

If you have any questions regarding this study, please do not hesitate to contact us at any time at the numbers listed below:

Professor Arieh Katz, Division of Medical Biochemistry, Faculty of Health Sciences, UCT, Tel: 021-4066268, email: arieh.katz@uct.ac.za

In addition, in case you have any ethical concerns or questions about your rights or welfare as a participant on this research study, UCT's Faculty of Health Sciences "Human Research Ethics Committee" can be contacted on 021 406 6338, email: shuretta.thomas@uct.ac.za

CONSENT FORM

I am consenting to participate in the study entitled: **Role of seminal fluid and inflammatory pathways/agents on cervical cancer progression and HIV infection (UCT HREC REF: 0849/2015)**.

I have read (or have read to me) the consent form. I understand, that I will be included in this study, the effect of which has been explained to me by _____

I am aware that the remainder of the semen sample that I provide today will be will be taken for purposes of this study. I voluntarily agree to participate in the study without any coercion. I also know that should I change my mind about participating in the study, I may do so without compromising my future prospects of treatment or care.

Name of patient: _____

Signature of patient: _____

The above signed this consent form in my presence and appears to understand it.

Signature of witness _____Name _____ Date_____

7.2. Appendix II – Solutions and Gels

1. Complete Media

Dulbecco's Modified Eagle Medium (DMEM), 10% fetal calf serum (FCS) (Highveld Biological, South Africa), 2 mM L-glutamine, penicillin (921 U/mL) and streptomycin (781 U/mL) (Highveld Biological, South Africa).

2. Starving Media

Dulbecco's Modified Eagle Medium (DMEM), 2 mM L-glutamine, penicillin (921 U/mL) and streptomycin (781 U/mL) (Highveld Biological, South Africa).

3. Mycoplasma test mounting fluid

20 mM Citric acid, 55 mM $Na_2HPO_4 \bullet 2H_2O$, 50% glycerol, pH adjusted to 5.5 with NaOH.

4. Phosphate-buffered saline (PBS)

138 mM NaCl, 2.67 mM KCl, 8.06 mM Na_2HPO_4, 1.47 mM K_2HPO_4, pH adjusted to 7.4 with NaOH or HCl, as appropriate.

5. PBS-Tween (PBS/T)

As for PBS; add 0.1% Tween20.

6. Mayer's Hematoxylin

50 g Potassium aluminum sulfate (Alum), 1 g hematoxylin, 0.2 g sodium iodate, 20 ml glacial acetic acid, all dissolved in 1 L dH_2O and boiled.

7. Scott's Tap Water Substitute

30 g $MgSO_4$, 2 g $NaHCO_3$ dissolved in 3 L dH_2O.

8. 2 X Boiling Blue

1.25 mL 1M Tris-HCl (pH 6.8), 4 mL 10% sodium dodecyl sulfate (SDS), 1mL β-mercaptoethanol, 2 mL glycerol, 1.75 mL dH_2O. Add a pinch of bromophenol blue.

9. Running Buffer

1 g SDS, 3.03 g Tris, 14.41 g glycine dissolved in 1 L dH_2O.

10. Transfer Buffer

3.03 g Tris, 14.4 g glycine, 200 mL methanol, made up to 1 L with dH_2O. Used ice-cold.

11. Tris-Buffered Saline (TBS)

50 mM Tris-HCl, 150 mM NaCl, pH adjusted to 7.4 with NaOH or HCl, as appropriate.

12. TBS/T

As for TBS; add 0.1% Tween20.

13. Sodium Dodecyl Sulphate (SDS)-polyacrylamide gels

a. Resolving gel:

8% Acryl-bisacryl-amide mix (30:08), 0.375 M Tris (pH 8.8), 0.1% SDS, 0.01% tetramethylethylenediamine (TEMED), 0.1% Ammonium persulphate.

b. Stacking gel:

4% Acryl-bisacryl-amide mix (30:08), 0.192 M Tris (pH6.8), 0.1% SDS, 0.01% TEMED, 0.1% Ammonium persulphate.

14. Roche High Pure RNA Isolation Kit buffer compositions

a. Lysis buffer:

4.5 M guanidine-HCl, 50 mM Tris-HCl, 30% Triton X-100 (w/v), pH 6.6.

b. DNAse incubation buffer:

1 M NaCl, 20 mM Tris-HCl, 20 mM $MnCl_2$, pH 7.0.

c. Wash Buffer I:

5 M guanidine-HCl, 20 mM Tris-HCl, in ethanol, pH 6.6.

d. Wash Buffer II:

20 mM NaCl, 2 mM Tris-HCl, in ethanol, pH 7.5.

a. Elution buffer:

PCR grade water (nuclease, nucleotide free).

7.3. Appendix III – qPCR primer sequences and melt curve exemplars

Table S-1: Candidate reference gene primer information.

Gene	Forward primer 5'-3'	Tm (°C)	Reverse primer 5'-3'	Tm (°C)	amplicon size (base pairs)
PPIA	**C**C-CACCGTGTTCTTCGA**C**AT[1]	60.32	CCAGTGCTCAGAGCACGAAA	60.66	116/273[2]
HPRT1	TGAGGATTTGGAAAGGGTGTT	57.39	CAGAGGGCTACAATGTGATGG	58.43	111
EEF1A1	TGCGGTGGGTGTCATCAAA	59.85	AAGAGTGGGGTGGCAGGTATT[3]	61.12	123
HSPC90AB1	ATGGAAGAGAGCAAGGCAAA	57.41	AATGCAGCAAGGTGAAGACA	58.01	117
GUSB	*Proprietary*	55	*Proprietary*	55	96

[1] Bolded letters and '–' (insertion of an extra C) indicate the position of SNPs identified via *in silico* analysis.
[2] Cannot discriminate between PPIA mRNA transcript variants 1 and 2, thus producing two amplicons
[3] The bolded T at the end of the sequence was removed owing to a sequence mismatch. Original sequence displayed as identified in existing literature.

Figure S1: Example of successful qPCR amplifications.

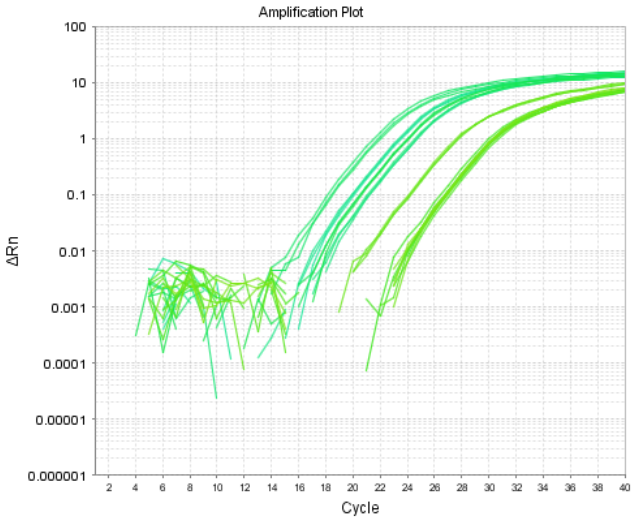

Figure S2: Example of a melt curve from a successful amplification.

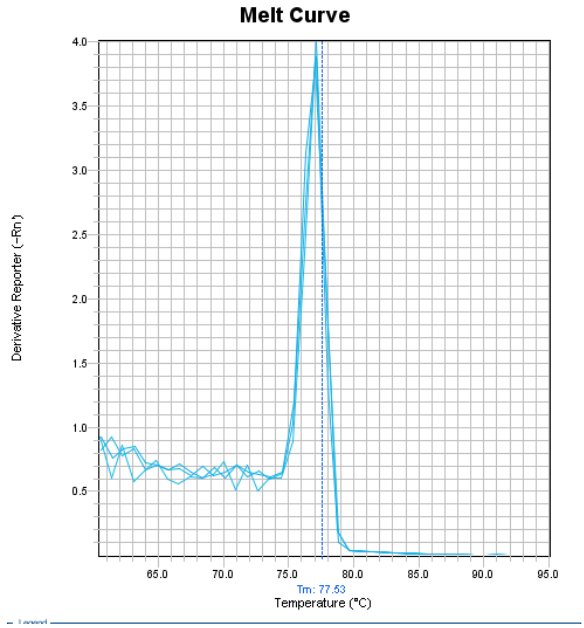

Figure S3: Example of primer dimers present in a melt curve from a failed amplification.

Primer dimers in experimental reactions are indicated by the blue arrow.

Nonspecific amplification in no-template controls is indicated by the purple arrow.

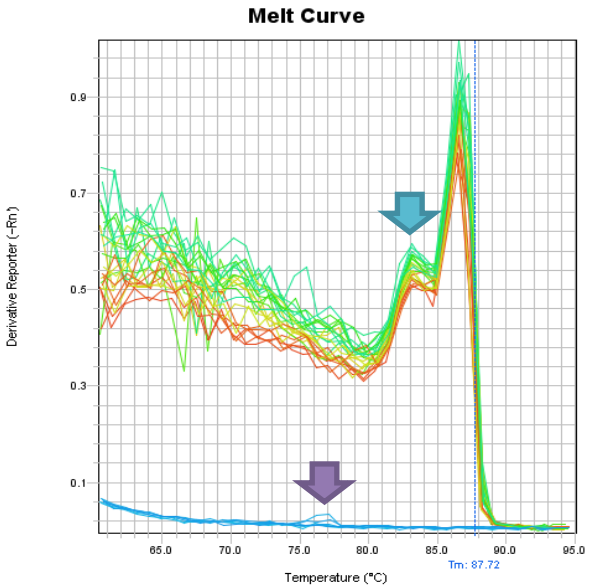

Figure S4: Example of genomic contamination present in a melt curve from a failed amplification.

Contamination indicated by blue arrow.