Bangor University
College of Natural Science
School of Biological Science

Identification of Novel Human Cancer-Testis-Antigen Genes in Cancers

---

Ph.D. Thesis 2014

Mikhlid Hammad Almutairi

Declaration and Consent

Details of the Work

I hereby agree to deposit the following item in the digital repository maintained by Bangor University and/or in any other repository authorized for use by Bangor University.

Author Name: Mikhlid Hammad Almutairi

Title: Identification of novel human cancer-testis-antigen genes in cancers

Supervisor/Department: Dr. Ramsay James McFarlane/ School of Biological Science

Funding body (if any): King Saud University (Riyadh, Saudi Arabia)

Qualification/Degree obtained: 08/09/2014

This item is a product of my own research endeavours and is covered by the agreement below in which the item is referred to as "the Work". It is identical in content to that deposited in the Library, subject to point 4 below.

Non-exclusive Rights

Rights granted to the digital repository through this agreement are entirely non-exclusive. I am free to publish the Work in its present version or future versions elsewhere.

I agree that Bangor University may electronically store, copy or translate the Work to any approved medium or format for the purpose of future preservation and accessibility. Bangor University is not under any obligation to reproduce or display the Work in the same formats or resolutions in which it was originally deposited.

Bangor University Digital Repository

I understand that work deposited in the digital repository will be accessible to a wide variety of people and institutions, including automated agents and search engines via the World Wide Web.

I understand that once the Work is deposited, the item and its metadata may be incorporated into public access catalogues or services, national databases of electronic theses and dissertations such as the British Library's EThOS or any service provided by the National Library of Wales.

I understand that the Work may be made available via the National Library of Wales Online Electronic Theses Service under the declared terms and conditions of use (http://www.llgc.org.uk/index.php?id=4676). I agree that as part of this service the National Library of Wales may electronically store, copy or convert the Work to any approved medium or format for the purpose of future preservation and accessibility. The National Library of Wales is not under any obligation to reproduce or display the Work in the same formats or resolutions in which it was originally deposited.

Statement 1:

This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree unless as agreed by the University for approved dual awards.

Signed **Mikhlid Almutairi** (candidate)

Date **08/09/2014**

Statement 2:

This thesis is the result of my own investigations, except where otherwise stated. Where correction services have been used, the extent and nature of the correction is clearly marked in a footnote(s).

All other sources are acknowledged by footnotes and/or a bibliography.

Signed **Mikhlid Almutairi** (candidate)

Date **08/09/2014**

Statement 3:

I hereby give consent for my thesis, if accepted, to be available for photocopying, for inter-library loan and for electronic storage (subject to any constraints as defined in statement 4), and for the title and summary to be made available to outside organisations.

Signed **Mikhlid Almutairi** (candidate)

Date **08/09/2014**

NB: Candidates on whose behalf a bar on access has been approved by the Academic Registry should use the following version of **Statement 3:**

Statement 3 (bar):

I hereby give consent for my thesis, if accepted, to be available for photocopying, for inter-library loans and for electronic storage (subject to any constraints as defined in statement 4), after expiry of a bar on access.

Signed **Mikhlid Almutairi** (candidate)

Date **08/09/2014**

Statement 4:

Choose **one** of the following options

a)	I agree to deposit an electronic copy of my thesis (the Work) in the Bangor University (BU) Institutional Digital Repository, the British Library ETHOS system, and/or in any other repository authorized for use by Bangor University and where necessary have gained the required permissions for the use of third party material.	√
b)	I agree to deposit an electronic copy of my thesis (the Work) in the Bangor University (BU) Institutional Digital Repository, the British Library ETHOS system, and/or in any other repository authorized for use by Bangor University when the approved **bar on access** has been lifted.	
c)	I agree to submit my thesis (the Work) electronically via Bangor University's e-submission system, however I **opt-out** of the electronic deposit to the Bangor University (BU) Institutional Digital Repository, the British Library ETHOS system, and/or in any other repository authorized for use by Bangor University, due to lack of permissions for use of third party material.	

Options B should only be used if a bar on access has been approved by the University.

In addition to the above I also agree to the following:

1. That I am the author or have the authority of the author(s) to make this agreement and do hereby give Bangor University the right to make available the Work in the way described above.

2. That the electronic copy of the Work deposited in the digital repository and covered by this agreement, is identical in content to the paper copy of the Work deposited in the Bangor University Library, subject to point 4 below.

3. That I have exercised reasonable care to ensure that the Work is original and, to the best of my knowledge, does not breach any laws – including those relating to defamation, libel and copyright.

4. That I have, in instances where the intellectual property of other authors or copyright holders is included in the Work, and where appropriate, gained explicit permission for the inclusion of that material in the Work, and in the electronic form of the Work as accessed through the open access digital repository, *or* that I have identified and removed that material for which adequate and appropriate permission has not been obtained and which will be inaccessible via the digital repository.

5. That Bangor University does not hold any obligation to take legal action on behalf of the Depositor, or other rights holders, in the event of a breach of intellectual property rights, or any other right, in the material deposited.

6. That I will indemnify and keep indemnified Bangor University and the National Library of Wales from and against any loss, liability, claim or damage, including without limitation any related legal fees and court costs (on a full indemnity bases), related to any breach by myself of any term of this agreement.

Signature: **Mikhlid Almutairi** Date: **08/09/2014**

Abstract

Cancer is a large group of diseases that can result in substantially increased mortality. This increased mortality occurs not only due to the limited effectiveness of the available treatments, but also because the diagnosis usually occurs late in the tumour development. Therefore, the identification of new cancer-specific biomarkers is extremely important in order to improve patient diagnosis at an early stage. Humans possess a class of genes that are normally expressed in the testes of adult males, and are also characteristic of several types of cancer cells. These genes are known as cancer-testis (CT) antigen genes and they might be helpful for both diagnosis and immunotherapy drug targeting. For this reason, identifying new CTA genes has significant clinical importance. Meiosis is restricted to germ cells and a number of meiotic proteins have previously been identified as CT antigens; consequently, we postulated that meiosis-specific genes may provide a good source for identifying potential novel CTA genes.

Potential meiotic genes were identified from a bioinformatics analysis. The specificity of these meiosis genes was validated in several types of normal human tissues. Genes that were found to display testis-restricted/selective or testis-CNS-restricted/selective expression profiles in the normal tissues were further screened in a wide range of cancer tissues and cancer cell lines. Out of the 32 genes screened by RT-PCR in the present study, the results identified 11 genes as novel CT genes. From the 11 candidate CT genes, *STRA8* and *C20orf201* were further functionally characterised in cancer cells.

The *STRA8* and *C20orf201* genes are considered to be good CT candidate genes since they are expressed in different types of cancer. Further analysed of *STRA8* and *C20orf201* by developing overexpression cell lines and this work indicates they have CTA potential. Expression of *STRA8* subcloned into Flp-In T-REx-293 cells did not reveal obvious functional activity.

C20orf201 gene was validated in this study at the protein levels in different normal and cancer tissue and cell lines. C20orf201 was found in the normal human testes and central nervous tissues. However, C20orf201 was also found to be present in a range of cancer cells suggesting it could be a CTA; biochemical evidence is also presented to indicate it might have functional activity as it appears to be modified in cancer-specific fashion.

Acknowledgements

The work presented in this thesis was carried out at the college of Natural Science, School of Biological Science, North West Cancer Research Institute at Bangor University from September 2010 to June 2014. First and foremost, I would like to express my sincere gratitude to my supervisor Dr. Ramsay McFarlane, for his continuous encouragement, guidance, patience and invaluable assistance throughout the project research and during writing of this thesis. He is not only a great supervisor, but also a great teacher in my life. Huge thanks also go to Dr. Jane Wakeman for all her help. I am forever thankful to Dr. Natalia Gomez-Escobar for all frequent help and support. I would like to thank all the previous and current lab members for their advice and always being available to answer my questions.

In addition, I would like to take this opportunity to thank Dr. J. Müller (Warwick University) for the Flp-In T-REx-293 and HEP-G2 cancer cell lines and for the pcDNA5/FRT/TO and pOG44 vectors. Furthermore, it is pleasure to thank the Kingdom of Saudi Arabia Government, in particular King Saud University, for funding this study. I would like to give my sincere thanks to all the members of Zoology Department, especially Prof. Faisal Abu-Tarbush for his encouragement and support since I first began studying for a master's degree until now.

Finally, this thesis could not have been completed without the love and support from my family, so special thanks go to my parents, my wife and my children (Khalid, Nada, Sultan, Yazeed and Abdulilah) and my brothers (Khalid, Enad, Fahad, Trad, Bander, Emad, Naif and Nawaf) for supporting my dreams and aspirations, and for making me believe that I can achieve anything I set my mind to.

Without the names listed above, I could never ever have finished this project. I really appreciate their help, deeply and sincerely.

Thank you all.

List of Abbreviations

3'	Three prime end of DNA
5'	Five prime end of DNA
5-azaC	5-aza-2'-deoxycytidine
amp	ampicillin
BCA	Bicinchoninic acid
BGH	Bovine growth hormone
BLAST	Basic Local Alignment Search Tool
bp	Base pair(s)
BSA	Bovine serum albumin
BTB	Blood-testis barrier
°C	Degrees Centigrade
C	Cytoplasm
cDNA	Complementary DNA
CDS	Coding DNA sequence
CE	Central element
Ch	Chromatin
CMW	Cytomegalovirus
CNS	Central nervous system
CO	Crossover
CpG	-cytosine-phosphate-guanine-
Cq	Quantification cycle
CRABP	Cellular RA binding protein
CTAs	Cancer testis antigens
C-terminal	Carboxy-terminal domain
CTLs	Cytotoxic T lymphocytes
DAC	5-aza-2'-deoxycytidine
DAPI	4',6-diamidino-2-phenylindole
DEPC	Diethylpyrocarbonate
dH_2O	Distilled water
dHJ	Double Holliday junction
D-loop	Displacement loop
DMEM	Dulbecco’s Modified Eagle’s Medium
DMSO	Dimethyl sulphoxide

DNA	Deoxyribonucleic acid
DNMT	DNA methyltransferase
DNMTi	DNA methyltransferase inhibitor
DSBR	Double strand break repair
DSBs	Double-strand breaks
dsDNA	Double-strand DNA
DSR	Determinant of selective removal
E. coli	*Escherichia coli*
EC	Embryonal carcinoma
ECACC	European collection of cell cultures
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
EST	Expressed sequence tag
et al.	*et alii* (and others)
F	Forward
FBS	Foetal bovine serum
FSH	Follicle stimulating hormone
g	Gram
G_0	Quiescent phase of the cell cycle
G_1	Gap-1 phase
G_2	Gap-2 phase
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
HATs	Histone acetyltransferases
HBV	Hepatitis B virus
HCC	Hepatocellular carcinoma
HCV	Hepatitis C virus
HDACi	Histone deacetylase inhibitor
HDACs	Histone deacetylases
HEK	Human embryonic kidney
HJ	Holliday junction
HLA	Human leukocyte antigen
HR	Homologous recombination
HRP	Horseradish peroxidase
hrs	Hours
IF	Immunofluorescent staining
IgG	Immunoglobulin G

kb	Kilobase
KDa	Kilodalton
L	Litter
LB	Luria bertani
LEs	Lateral elements
LH	Luteinising hormone
M	Mitosis phase
mA	Milliampere
MCS	Multiple cloning site
mg	Milligram
MHC	Major histocompatibility complex
min	minute
ml	Milliliter
mM	Millimolar
MM	Multiple myeloma
MRN	Mre11-Rad50-Nbs1
mRNA	Messenger RNA
MRX	Mre11-Rad50-Xrs2
MW	Molecular weight
N	Nuclear
NCBI	National Centre for Biotechnology Information
NCO	Non-crossover
ng	Nanogram
NHEJ	Non-homologous end-joining
NL	Non-tumour lung
nM	Nanomolar
NRT	No reverse transcriptase
NTC	No template control
N-terminal	Amino-terminal domain
oligo-dT	Oligodeoxythymidylic acid
ORF	Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PB1	First polar body
PB2	Second polar body
PBS	Phosphate buffered saline
PBST	Phosphate buffered saline-tween-20

PCR	Polymerase chain reaction
PFA	Paraformaldehyde
pH	Power of hydrogen
pmol	picomole
PMSF	Phenylmethylsulfonyl fluoride
PS	Placenta-specific
PSA	Prostate-specific antigen
PTM	Post-translational modification
PVDF	Polyvinylidene difluoride
q-RT-PCR	Quantitative real-time PCR
R	Reverse
r.p.m.	Rotation per minute
RA	Retinoic acid
RARs	Retinoic acid receptors
RBP	Retinol-binding protein
RIPA	Radioimmunoprecipitation assay
RNA	Ribonucleic acid
RNAi	RNA interference
rRNA	Ribosomal RNA
RT	Room temperature
RT-PCR	Reverse transcriptase PCR
RXRs	Retinoid X receptors
s	seconds
SC	Synaptonemal complex
SD	Standard deviation
SDS	Sodium dodecyl sulphate
SDSA	Synthesis-dependent strand annealing
SDS-PAGE	Sodium Dodecyl Sulphate- Polyacrylamide Gel Electrophoresis
SEM	Standard error of the mean
siRNA	Small interfering RNA
SMC	Structural maintenance of chromosomes
SPB	Spindle poles
SSCs	Spermatogonial stem cells
ssDNA	Single strand DNA
Ta	Annealing temperature
TAA	Tumour-associated antigen

TBE	Tris-borate-EDTA
TFs	Transverse filaments
Tm	Melting temperature
T-REx	Tetracycline regulated expression
TS	Testis-specific
Ub	Ubiquitin
UV	Ultra violet
V	Voltage
WB	Western blot
WCE	Whole cell extract
µg	Microgram
µl	Microliter

List of Contents

List of Figures

Chapter 4.0

Chapter 6.0

List of Tables

Chapter 1.0: Introduction

1.1 Cancer in humans

1.1.1 Overview of cancer

Cancer is a major public health problem worldwide; more than 7 million people died in 2008 due to cancer. The number of cancer deaths is expected to increase rapidly to over 11 million per year globally by 2030 (http://www.who.int/cancer/en/). In the United States, for example, one out of four human deaths are caused by cancer and among child deaths, leukaemia and brain cancer are considered to be the second most frequent causes, compared with other types of cancer (Siegel *et al*., 2012). There are also distinctions in cancer deaths in males and female; for example, breast cancer is the second leading cause of cancer related deaths in women in the United Kingdom (Jemal *et al*., 2010).

Cancer is a term for a large group of diseases which are often associated with multiple genetic and epigenetic changes of the cellular genome. Cancer cells proliferate without restriction and changes in cellular adhesion can result in cancer spreading to other locations in the human body and destroying multiple tissues, through a process termed metastasis (Brábek *et al*., 2010; Geutjes *et al*., 2011; Sharma *et al*., 2010). A cancer can arise from a single cell, which is considered the site of the primary cancer. The primary tumour displays uncontrolled growth, invasion and often metastasis through the blood or the lymphatic system, which enables the cancer to spread quickly to different sites. These are the features of malignant cells (Hanahan and Weinberg, 2011; Kopfstein and Christofori, 2006; Shacter and Weitzman, 2002) and distinguish malignant cancers from benign tumours, which grow more slowly and do not invade or metastasise (Klein, 2008; Player *et al*., 2004).

1.1.2 Classification of human cancers

Human cancers are currently classified into various types according to factors such as the origin of the cells, the pathological properties of malignant cells and the phase of tumour progression. The four main types of cancer are carcinoma, sarcoma, leukaemia and lymphoma. (1) Carcinoma is a term which refers to malignant tumours of epithelial tissues, and is found in organs such as the skin, lung and breast. The great majority (approximately 90%) of human cancers are carcinomas; this prevalence may arise because epithelial tissues tend to proliferate during humans life and are most directly exposed to exogenous agents that

assist in promoting tumourigenesis. (2) Sarcomas are cancers that develop in connective tissue, such as muscles, bone and cartilage. (3) Leukaemia and (4) lymphoma arise from haematopoietic cells. Leukaemia is cancer of the white blood cells, while lymphomas begin in B and T lymphocytes. Benign and malignant tumours differ greatly; most benign tumours are adenomas; a type of cancer that generally occurs in glandular tissues, whereas malignant tumours in glandular tissues are named adenocarcinomas (Ruddon, 2007; Weinberg, 2007). Further classification of cancers is available at the National Cancer Institute website (http://www.cancer.gov).

1.1.3 The causes of human cancers

Cancer is caused by exogenous and endogenous agents (Morley and Turner, 1999). The endogenous (internal) agents arise as a result of natural errors in processes occurring during metabolism, DNA duplication, DNA repair and chromosome segregation. Errors in DNA repair or chromosome segregation can lead to alterations in chromosome number and/or structure, thereby causing an increase or decrease in gene copy number. Errors in DNA replication can also result in genetic recombination products that considerably increase or decrease the number of gene copies. Exogenous (external) agents usually enter the human body from the environment. Exogenous agents are classified into three types: physical, chemical and biological (Belpomme *et al.*, 2007; Lutz, 2002; Vogelstein and Kinzler, 2004). Exposure to ultraviolet radiation, a physical agent, can lead to the induction of pyrimidine dimers in DNA, which can cause skin cancer (Douki, 2013). Chemical mutagens, like those inhaled in tobacco smoke, can induce diverse forms of nucleotide damage observed in lung cancer (Gibbons *et al*., 2014). Biological agents can also cause human cancer; for example, hepatitis B or C virus (HBV or HCV) infections are causally associated with hepatocellular carcinoma (Lupinacci *et al*., 2013).

Cancers can be induced by genetic and/or epigenetic changes to tumour suppressor genes and oncogenes. These genetic changes may increase tumour progression, which leads to tumour proliferation, invasion and spreading (Gnoni *et al*., 2013). Nevertheless, recent studies propose that cancer also has a universal foundation in polyclonal epigenetic disruption of stem cells, mediated by tumour stem cells, which can lead to changes in gene expression that commonly contribute to the development and progression of tumours (Choi and Lee, 2013; Feinberg *et al.*, 2006; Vogelstein and Kinzler, 2004).

1.1.4 Genetic alterations and cancer progression

Genetic variations in DNA can result in cancer because failures in chromosome segregation and DNA repair can result in chromosomal aberration, as regularly observed in malignant cells (Ström and Sjögren, 2007). Chromosomal aberrations include amplifications, deletions and chromosomal translocations. However, chromosomal translocations are the main cause of several cancers in humans; for example, leukaemia, lymphoma and some solid tumours (Nambiar *et al.,* 2008). The simplest genetic alteration that can occur in cancer is a point mutation, where a single nucleotide is mutated or deleted. If a mutation is not repaired, the mutated cells can change their behaviour and cancer may initiate (Lengauer *et al*., 1998; Zheng, 2013).

Amplification is a regional copy number increase, and is often associated with overexpression of some of the affected genes. Amplification of a DNA sequence has been observed in various kinds of lymphoma and solid tumours (Santarius *et al*., 2010). Deletion is the loss of genetic material and can be limited in size or involve entire chromosome arms. Abnormal changes in the numbers of chromosome are termed aneuploidy (Albertson *et al*., 2003; Munger, 2002). Chromosomal translocation is a term generally used to describe chromosomal rearrangements, including exchanges of segments between two non-homologous chromosomes. The majority of translocations in humans are related to lymphoid tumours (Küppers, 2005; Nussenzweig and Nussenzweig, 2010). Chromosomal translocations are generally used as diagnostic markers for malignant cells. Chromosomal translocation can lead to the generation of a new gene product by bringing genes close to enhancer or promoter elements, thereby resulting in an alteration in their expression or by the fusion of genes to generate new oncogenic variants (Nambiar *et al*., 2008).

1.1.5 Epigenetic alterations in cancer development

Cancer is considered to be both a genetic and an epigenetic disease. It is now widely recognised that epigenetic events underlie cancer progression. Epigenetic variations result in altered gene expression due to factors other than alteration in the DNA sequence itself. Epigenetic changes can change the phenotype without altering genotype. The main epigenetic factors that may affect global gene expression involve DNA methylation and histone modifications (Hatzimichael and Crook, 2013).

1.1.5.1 DNA methylation

DNA methylation refers to the addition of a methyl group to the 5-carbon position of the cytosine within the dinucleotide CpG (-cytosine–phosphate–guanine-) of a DNA strand, catalysed by DNA methyltransferases (DNMTs). Regions of DNA that are CG rich are called 'CpG islands' and are usually associated with gene promoters (Vinson and Chatterjee, 2012). DNA methylation has a fundamental role in mammalian development (Robertson, 2005; Smith and Meissner, 2013) and an essential function in X-chromosome inactivation (Lee and Bartolomei, 2013) and genomic imprinting (Girardot *et al*., 2013).

The promoter regions of a normal mammalian cell genome comprise approximately 70–80% of all CpG islands being methylated (Ziller *et al*., 2013). On the contrary, DNA hypomethylation in non-promoter regions and hypermethylation in the promoter region of genes is associated with transcription silencing of tumour suppressor and DNA repair genes, and is often identified in different cancerous cells (Akhavan-Niaki and Samadani, 2013). Hypermethylation also leads to genomic instability and can affect apoptosis, DNA repair and cell-cycle control, thereby giving rise to tumourigenesis (Esteller and Herman, 2002; Wu and Bekaii-Saab, 2012). The expression levels and activity of DNMTs are generally greater in distinct human cancers than in normal tissues (Ibrahim *et al*., 2011).

1.1.5.2 Histone modifications

The length of a strand of human DNA is large enough to exceed the borders of the cell nucleus, so it requires packaging in to higher order structures. The nuclei of eukaryotic cells contain both histone and non-histone proteins that organised the genomic DNA into a condensed structure, termed chromatin. Chromatin is typically found in two different functional forms: heterochromatin, a highly condensed and transcriptionally repressive form, and euchromatin, a less condensed form that can be generally permissive for gene transcription. The basic unit of chromatin is termed the nucleosome and consists of 147 base pairs of genomic DNA wrapped twice around a histone octamer, consisting of isoforms of two copies each of the core histones H2A, H2B, H3 and H4. Histones; however, are not only packaging elements, but also critical regulators of gene expression (Clausell *et al*., 2009; Eickbush and Moudrianakis, 1978; Luger *et al*., 1997; Russ *et al*., 2012).

Histone may be subjected many post-translational chemical modifications; for instance, acetylation, methylation, phosphorylation and ubiquitination of the N-terminal regions. This is also limited to H3 and H4 isoforms. These modifications can change the structure of

chromatin and gene expression in human cancers from an open to a closed, condensed form, and vice versa (Figure 1.1). Histone modifications also have an important role in different processes involving transcriptional repression, gene activation and DNA repair (Kouzarides, 2007). For example, the transcriptional activity of euchromatin is often associated with high levels of histone acetylation, while silent heterochromatin is correlated with low levels of histone acetylation (Bannister and Kouzarides, 2011; Nakazawa *et al*., 2012). Histone modifications can alter cellular activity in an oncogenic fashion; for example, acetylation and methylation are the most active in colorectal cancer pathogenesis (Nakazawa *et al*., 2012).

Histone acetylation changes the chromatin structure by altering the interaction between the negatively charged phosphate group of DNA, and the positively charged histone. Most acetylation sites characterised to date are found within the N-terminal tail of histones, which are more accessible for modification (Han *et al*., 2007). Histone acetylation at the histone tails is controlled by histone acetyltransferases (HATs) and histone deacetylases (HDACs) that ordinarily serve as transcriptional co-activators or co-repressors, respectively. HDACs promote the removal of the acetyl group from the amino group of the lysine residue of the histone and are the most characterised proteins among the histone modification enzymes that play important roles in cancer development (Ashktorab *et al*., 2009).

Epigenetic modification can control tissue specific gene regulation; for example, some cancer associated genes are normally expressed in the testis and are controlled by differential epigenetic regulation (see Section 1.2.4).

1.1.6 Antigen markers in cancer

Cancer markers can be defined as factors found in the human body that indicate the presence of cancer (Mäbert *et al*., 2014). Some cancer markers are found in the solid tumour itself or in lymph nodes, in bone marrow, or in other body fluids (urine or blood). These markers can be made by either the body in response to cancer or by the cancerous cells themselves. Many markers in different cancers are proteins, but some of the new types of markers are genes or levels of genetic materials (DNA or RNA) (Mäbert *et al*., 2014). Two distinct types of cancer markers are recognised: specific markers that are associated only with one kind of cancer and unspecific markers that are found in several different cancers (Freidlin and Korn, 2014; Lindblom and Liljegren, 2000).

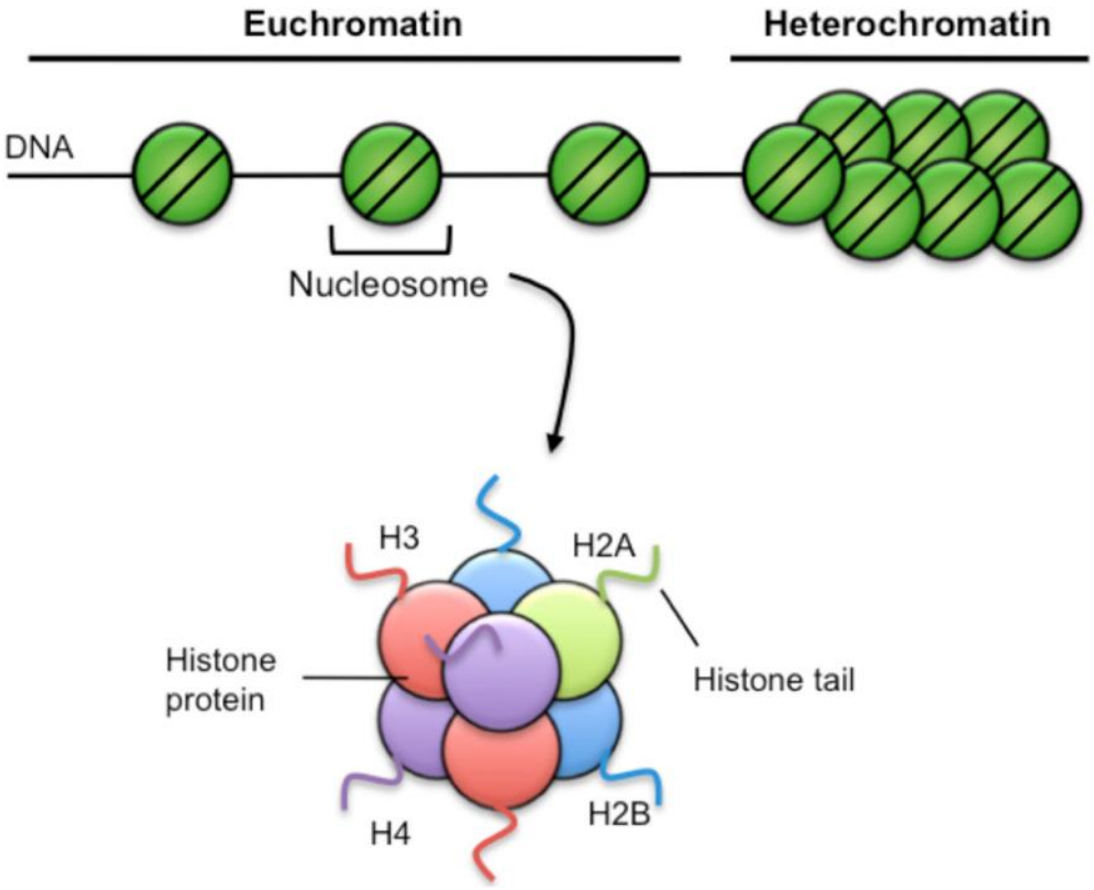


Figure 1.1 Euchromatin and Heterochromatin structure. Euchromatin is an open form of chromatin and is normally associated with active gene transcription. In contrast, heterochromatin is characterised by high nucleosome density, which results in a closed form of chromatin. Therefore, heterochromatin is normally associated with the repression of gene transcription. DNA is wrapped around the histone octamer of the four core histones H2A, H2B, H3 and H4. Each histone has an amino terminal tail that can be covalently changed by specific epigenetic reactions such as methylation, acetylation and phosphorylation (Adapted from Russ *et al.*, 2012).

Cancer markers can be helpful in different ways: (a) Screening and early detection of cancer. The primary goal of a cancer marker is to aid diagnosis of cancer in individuals who have no clinical symptoms of the disease, as early diagnosis of a tumour is important for survival (Duffy, 2001; Thomas and Sweep, 2001). A widely used cancer diagnosis marker is currently the prostate-specific antigen (PSA). The PSA blood test is used to screen men for prostate cancer, as men with prostate cancer usually have high PSA levels (Schröder *et al*., 2009); (b) Diagnosis of a specific type of cancer. Cancer markers in some cases cannot be used alone to diagnose a specific cancer, but can complement biopsy for diagnosis; (c) Determining the prognosis for certain cancers; and (d) Monitoring how well a cancer treatment is working. If a specific type of cancer has a marker, changes in the level of the marker (increased or decreased) can indicate if the treatment is working well (Duffy, 2001; Thomas and Sweep, 2001). Importantly, a number of cancer biomarkers can be used both in diagnosis and in target therapy; for example, CEA and HER2 are biomarkers used for human colon and breast cancer, respectively. However, both markers can be used in passive and active therapy (Even-Desrumeaux *et al*., 2011)

A number of studies have shown that cancer cells can change the expression of genes from up regulated to down regulated or vice versa. Up regulated genes are considered to be cancer markers in different tumours; nevertheless, their use as therapeutic and/or diagnostic targets is restricted for several genes, since the resulting antigens are recognised by the human immune system (Caballero and Chen, 2009; Costa *et al*., 2007).

Several lab studies have focused on a particular class of antigens in order to identify specific biomarkers with potential use in the early detection of malignancy. One such group of antigens have been identified as the cancer-testis antigens (CTAs), and these could play an essential role in clinical studies (Mirandola *et al*., 2011).

1.2 Cancer testis antigens (CTAs)

1.2.1 Overview of CTAs

Van der Bruggen and his colleagues identified the first human cancer testis (CT) genes in 1991 in a patient with malignant melanoma and named the melanoma antigen family (MAGE-1) (van der Bruggen *et a*l., 1991). Further studies showed that the mRNA of *MAGE* was observed in different types of human tumours and also in testis and placental cells of somatic tissues (Fratta *et al*., 2011; Zendman *et al*., 2003). The term cancer testis antigen

(CTA) was first used in 1997 by Old and Chen (Chen *et al*., 1997; Old and Chen, 1998) for this promising class of tumour-associated antigens (TAA), because of their limited presence in somatic tissues (de Carvalho *et al*., 2012). These antigens are also known as cancer germline antigens (Chen *et al*., 1997). Many of the CTAs were discovered from the examination of melanoma samples. The initial discovery of the early CT-antigens, *MAGE-A1*, *PAGE* and *GAGE1,* was reported between 1991 and 1995 (Boël *et al*., 1995; Van den Eynde *et al*., 1995; van der Bruggen *et al.,* 1991), followed by *SSX2* and *NY-ESO-1* in the late 1990s (Chen *et al*., 1997; Tuereci *et al.,* 1998).

CTAs are a heterogeneous group of proteins whose corresponding genes are normally expressed solely in the testes of adult males and in several types of malignant human tumour cells (Whitehurst, 2014). These antigens are encoded by the CT genes (Scanlan *et al.,* 2004; Simpson *et al.,* 2005; Whitehurst, 2014), whose expression levels have been shown to vary in tumours of different histology. For example, these genes display low expression levels in lymphomas, leukaemia, colon and pancreatic cancers, whereas they show high expression in melanoma, ovarian cancer and lung tumours (Caballero and Chen, 2009). In some cases, CT genes are also expressed in the ovaries, the placenta, and the thymus; one example is *NY-ESO-1,* a CT gene originally identified from oesophageal squamous cell carcinoma of a 58-year-old female (Caballero and Chen, 2009).

Due to their unique gene expression profiles, the CTAs are considered attractive targets for use as immunotherapy for tumours, as biomarkers for the early diagnosis of cancer and as strong candidates for cancer vaccines (Simpson *et al.,* 2005; Valmori *et al.,* 2007; van Duin *et al*., 2011). In biological studies, CT genes are considered as good models for understanding complex gene regulation and abnormal gene activation throughout cancer development. Currently, any gene that displays testis-specific expression and expression in tumour cells is defined as a CT gene (Mirandola *et al*., 2011; Scanlan *et al.,* 2004; Whitehurst, 2014).

The majority of CT genes discovered to date are localised on the X chromosome (Kalejs and Erenpreisa, 2005). Examples of CTAs which have been proposed to be used as tumour biomarkers are MAGE-A1/A3 for hepatocellular carcinoma and lung adenocarcinoma (Qiu *et al*., 2006), BORIS for breast cancer (D'Arcy *et al*., 2006) and NY-ESO-1 for lung adenocarcinoma (Gure *et al*., 2005).

Recent study showed that germline genes, including testis-specific/placenta-specific (TS/PS) genes, can be used as prognostic markers in lung cancer patients at all disease stages. These

are good sources of cancer biomarkers because their expressions were not observed in the healthy normal tissues other than the testis or placenta (Rousseaux *et al*., 2013).

1.2.2 Classification and expression of CT genes

The available CTA database has identified more than 100 gene families of potential CT genes, with over 280 members known to date. The CTA members, their gene expression profiles and other characteristics are summarised on a single website supported by the Ludwig Institute (http://www.cta.lncc.br/) (Almeida *et al*., 2009).

CT genes can be classified into two groups according to their chromosomal location. The X-CT group is located on the X chromosome, and those located on the autosomes are called the non-X-CT group. In the normal testis, most of the X-CT genes are expressed in the spermatogonia stage of spermatogenesis, and their function remains poorly understood. The X-CT genes such as *MAGE-A3*, *MAGE-8*, *MAGE-A10*, *XAGE-2* and *XAGE-3* are also expressed in the placenta in addition to the testis (Caballero and Chen, 2009). Some somatic tissues such as the liver, pancreas and spleen show the expression of mRNA of distinct CTAs, but quantitative RT-PCR analysis indicates these mRNA levels are normally less than about 1% of their expression in testicular germ cells (Caballero and Chen, 2009). In contrast, many of the non-X-CT genes are expressed in spermatocytes and perform several tasks during meiotic cell division (Caballero and Chen, 2009; Simpson *et al.*, 2005). It has been proposed that the aberrant expression of CTA genes in malignant cells may result in abnormal chromosome segregation and aneuploidy, thereby giving rise to tumourigenesis (de Carvalho *et al*., 2011; Fratta *et al*., 2011; Whitehurst, 2014).

Members of X-CT genes are found mostly as multi-copy, paralogous gene families, some also express splice variants. Examples of these families are *MAGEA*, *GAGE*, *PAGE*, *XAGE1*, and *SSX* (Figure 1.2). The CT genes located on autosomal chromosomes are mostly single-copy genes. A many as 10% of the genes on the X chromosome have been estimated to be X-CT genes (Ross *et al.*, 2005). Of the CT genes identified, 52% are distributed on the X chromosome, while the remaining (non-X-CT genes) are located on the 22 autosomal chromosomes and the Y chromosome (Rajagopalan *et al*., 2011).

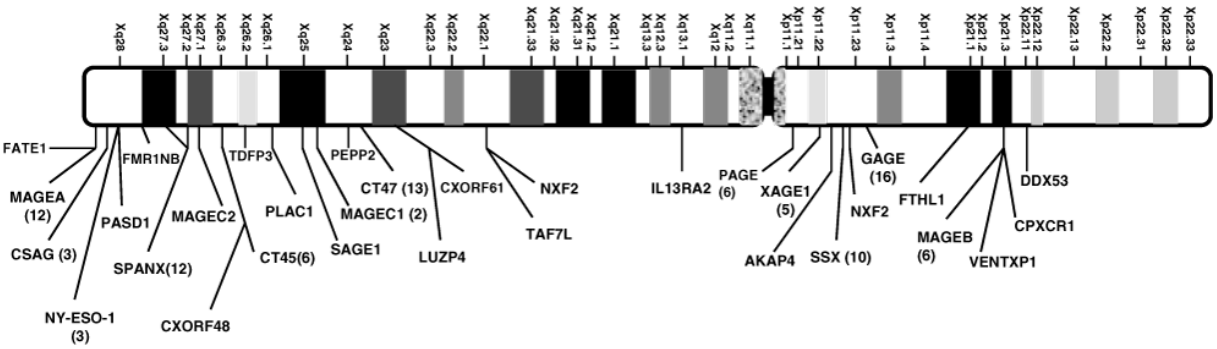


Figure 1.2 Distribution of X-CT genes on the X-chromosome. Different families of cancer testis genes located on the X-chromosome. The number of individual gene families is shown between brackets (Adapted from Caballero and Chen, 2009).

The expression of CT genes in the testis is restricted exclusively to spermatogenic germ cells, which include spermatogonia, primary spermatocytes, secondary spermatocytes, spermatids and spermatozoa. However, the various CT genes are expressed at different stages of sperm development. The expression of some CT genes is controlled by promoter demethylation, indicating as the presence of a testis-specific demethylation mechanism (Fratta *et al*., 2011; Shichijo *et al.,* 1996).

CT antigens have been further classified according to their gene expression profiles into four groups: (a) Testis-restricted (expressed only in the adult testis, placenta and at least one cancer type); (b) Testis-brain-restricted [expressed in the adult testis, central nervous system (CNS) and at least one cancer type]; (c) Testis-selective (expressed in the adult testis and no more than two additional normal tissues at lower expression levels than in the testis, and at least one cancer type) (Hofmann *et al*., 2008); and (d) Testis-brain-selective (expressed in the adult testis, CNS tissue, and no more than two additional normal tissues at lower expression levels than in the testis, and at least one cancer type) (Feichtinger *et al*., 2012a).

According to Hofmann *et al.* (2008), 153 distinct CT genes or isoforms have been described in normal and cancer gene expression libraries. These genes are divided into three categories: testis-restricted, testis-selective and testis-brain-restricted. The chromosomal distribution for each group is shown in the following table:

CT gene group	X-CT	Non-X-CT	mRNA expression in normal tissue
Testis-restricted	35 genes	4 genes	In testis and sometimes in placenta
Testis-selective	36 genes	64 genes	In testis and different normal tissues
Testis-brain-restricted	12 genes	2 genes	In testis and CNS tissue

1.2.3 Functions of cancer testis genes in normal and tumour cells

CTAs have become the most widely studied antigen group in the cancer immunology field. The biological functions of the majority of these genes in both the germ line and tumour cells remain largely unexplored and the connection between gametogenesis and cancer is very poorly understood (Kalejs and Erenpreisa, 2005). However, some CT genes have been identified by their functions in normal cells. For instance, synaptonemal complex protein 1 (SYCP1) functions in chromosome synapsis during meiosis I (Schramm *et al*., 2011; Tuereci *et al.,* 1998). The OY-TES-1 gene plays an essential role in acrosin packaging in the acrosome of sperm heads (Ono *et al.,* 2001). SPO11 acts as a meiotic endonuclease and

generates DNA double-strand breaks to initiate meiotic homologous recombination (Yamada and Ohta, 2013). BORIS has an important role in regulating the promoter methylation process through spermatogenesis by removing imprinting from genes during the final round of mitosis in spermatocytes (Klenova *et al.,* 2002; Loukinov *et al.,* 2002; Martin-Kleiner, 2012).

Although the function of many of these tumour-associated antigens in the disease process is not completely clear, limited evidence infers that they function in many vital cellular processes such as cell growth and transcriptional regulation (Scanlan *et al*., 2002). Some also appear to function as putative proto-oncogenes, such as *Piwil2* that promotes cell proliferation (Cheng *et al*., 2011).

A soma-to-germline transition may result in cancer progression. For example, investigation of the expression patterns of the human orthologues of the *Drosophila melanogaster* germline genes that are normally expressed in brain tumours indicated that most of these genes are generally expressed in germlines and in various human cancer types (Feichtinger *et al*., 2014). Therefore, the expression of these germline genes has a core role in oncogenesis in *D. melanogaster* (Janic *et al*., 2010) and potentially in humans (Feichtinger *et al*., 2014). Inhibition of some of germline genes leads to repression of tumour development, demonstrating that these genes could play a fundamental role in tumourigenesis (Janic *et al*., 2010).

1.2.4 Regulation of CTA gene expression

The expression of some CT genes is regulated by DNA methylation and by modification of a specific amino acid in the N-terminal tail of histone (see Section 1.1.5). DNA methylation mainly occurs in the gene promoter region, and is carried out by DNA methyltransferase enzymes (DNMTs). This process is responsible for CTA gene silencing (for example, see Bai and Zhang, 2010). In somatic tissues, the promoter regions for all X-CT genes contain CpG islands and are often methylated and therefore not expressed (Simpson *et al*., 2005). DNA hypomethylation is the most frequently reported mechanism for gene expression activation in many cancer types (Wang *et al.,* 2004). DNA hypomethylation reportedly promotes the expression of some CT genes; for instance, the melanoma-associated antigen family A (*MAGE-A1*) gene in different cancer cells (De Smet and Loriot, 2010) and the cancer-associated gene (*CAGE*) in gastric cancer (Lee *et al*., 2006). Hypermethylation of the *MAGE* promoter results in repression of its expression (Wischnewski *et al.,* 2007).

DNA methylation is the primary silencing mechanism for CT genes. Inhibition of CT gene transcription is caused either by preventing the binding of specific transcription factors (TFs) to the DNA via methylated recognition sequence, or by the binding of methyl-CpG-binding proteins, which prevent gene expression of methylated genes (Fratta *et al*., 2011). On the other hand, demethylation is necessary for their expression. For instance, a direct correlation exists between demethylated CpG islands in the *MAGE-A1* 5' promoter region and its expression (De Smet *et al.*, 1999). The expression of CT genes such as *SSX* and *NY-ESO-1* can be induced or upregulated in tumour cells by treatment of the cell lines with either DNA methyl-transferase 1 inhibitor (DNMTi), the DNA methyl transferase inhibitor 5-aza-2'-deoxycytidine (5-azaC) (Grunwald *et al.*, 2006; Wang *et al.*, 2004), and/or by histone deacetylase inhibitor (HDACi) (Akers *et al.*, 2010). Indeed, 5-azaC has been shown to induce expression of many of CT genes, especially the X-CT genes (Almstedt *et al*., 2010). Treatment of different human ovarian cancer cell lines with 5-azaC results in increases in the mRNA expression of 11 out of 12 CT genes, which include *MAGE-A1*, *MAGE-A3*, *MAGE-A4*, *MAGE-A6*, *MAGE-A10*, *MAGE-A12*, *NY-ESO1*, *TAG-1*, *TAG-2a*, *TAG-2b* and *TAG-2c* (Adair and Hogan, 2009).

Deacetylation is achieved by histone deacetylases (HDACs); consequently, repression of the HDACs by a specific inhibitor will result in up regulation of many CT genes (Akers *et al*., 2010; Barneda-Zahonero and Parra, 2012). The CTA genes *MAGE-A9* and *MAGE-A11* are silenced by DNA hypermethylation and also by histone deacetylation in breast cancers (Hou *et al*., 2013a).

1.2.5 CTAs as targets for immunotherapy and cancer vaccines

Many studies have been published about the medical use of CTAs as potential targets for cancer bio-markers and immunotherapy (Houghton *et al.*, 2001; Mellman *et al*., 2011; Rosa *et al*., 2012).

Most tumour specific antigens contain small antigenic peptides regions of the protein that can be presented on the surface of cancerous cells by the major histocompatibility complex (MHC) molecule, also known as the human leukocyte antigen (HLA). This MHC peptide complex can be recognised by the cytotoxic T lymphocytes (CTLs), which then kill the tumour cell (Adair and Hogan, 2009). The presence of the blood-testis barrier (BTB) (Li *et al*., 2012) and the absence of HLA class I expression on the surface of testis germ cells (Ghafouri-Fard and Modarressi, 2012a) prevent the human immune system from interacting

with CTA proteins in the testis, so that these antigens are not recognised as self-structures and are immunologically targeted (Kalejs and Erenpreisa, 2005).

Treatment of cancer via chemotherapy and radiotherapy is often characterised by a high level of toxicity and a low level of efficiency. However, immunotherapy can target tumour cells with less toxicity to normal tissues (Aly, 2012). Immunotherapy is considered as a potentially powerful form of cancer therapy (Mellman *et al*., 2011; O'Shea *et al*., 2014). For example, autologous $CD4^+$ T-cells harvested from a patient diagnosed with a metastatic melanoma using an immune reactive CTA, NY-ESO-1 peptide; these isolated T-cells were cultured and expanded *in vitro* for cellular therapy and infused back into the same patient, whose tumour burden was depleted; a CT scan done two months after the infusion showed that the metastasised tumours were completely lost and the patient remained disease free for over two years after receiving the adoptive therapy (Hunder *et al*., 2008). This is just one example of successful adoptive therapeutics using CTAs in immunologically reactive cancers, such as melanomas.

Neither normal testicular germ cells nor placental cells express the human leukocyte antigen class I molecule. Therefore, CTAs are not normally targeted by $CD8^+$ cytotoxic T lymphocytes. For this reason, CTAs are considered attractive targets for antitumour cell vaccines (Jungbluth *et al*., 2005). Vaccination with CT antigens has considerable potential for treating cancer (Renkvist *et al*., 2001; Blanchard *et al*., 2013). The majority of cancer vaccines are antigens expressed by human cancer cells and recognised by the human $CD8^+$ and $CD4^+$ T-cells (Greten and Jaffee, 1999; Blanchard *et al*., 2013). Cancer vaccines based on the CT antigens *MAGE-A1* (melanoma antigen A1) and *NY-ESO-1,* are currently being tested in an attempt to stimulate the T-lymphocyte response against malignancy (Caballero and Chen, 2009) and as an essential target for vaccination against various human tumours (Campos-Perez *et al*., 2013).

1.3 Spermatogenesis

1.3.1 Spermatogenesis overview

Gamete cells act as a link between generations and they transfer genetic information from parents to child. The father contributes genetic information via the sperm. In humans, sperm production begins at puberty and usually continues into old age (Mruk and Cheng, 2010; Rooij and Russell, 2000). Spermatogenesis is a very complex, highly organised and regulated

process that occurs in the seminiferous tubules, the functional unit of the testis that can produce sperm (O'Donnell *et al.*, 2006). The seminiferous tubules are highly coiled loops; they contain Sertoli cells and germ cells and are bordered by a layer of peritubular cells, which provide structural support (Figure 1.3A). The production of spermatozoa in humans is regulated by follicle stimulating hormone (FSH) and luteinising hormone (LH) (O'Shaughnessy, 2014) and through the effects of retinoic acid (RA) (Hogarth and Griswold, 2013). Both hormones are released by the pituitary gland (Holdcraft and Braun, 2004). LH binds to the LH receptors in Leydig cells, which are found in the space between the seminiferous tubules, to produce the male steroid hormone testosterone. Sertoli cells are regulated by FSH and testosterone. Therefore, any mutations or changes in the genes that encode these hormones or their receptors may have effects on normal sperm production (Ge *et al*., 2009; Holdcraft and Braun, 2004; Phillips *et al*., 2010).

1.3.2 Structure of the seminiferous tubule

Spermatogenesis consists of four major phases: (a) Proliferation of progenitor cells via mitosis; (b) Meiotic division of diploid cells to form haploid cells; (c) Cellular differentiation of spermatids to form spermatozoa; and (d) The release of sperm from the seminiferous tubule epithelium into the seminiferous tubule lumen (Hess and de Franca, 2009).

The progenitor cells are normally called spermatogonia, and are divided into undifferentiated and differentiation spermatogonia. The undifferentiated spermatogonia include the stem cells, which are termed spermatogonial stem cells (SSCs). When the SSCs undergo division, they produce two types of cells with different functions: self-renewal (A type) and differentiation (A1 type). Cells destined for self-renewal remain as stem cells, while cells that undergo differentiation continue to undergo the phases of mitotic division before entering meiosis to produce primary spermatocytes (Griswold *et al*., 2012; Mruk and Cheng, 2010). The primary spermatocytes undergo the first phase of meiotic division (see Section 1.4.3) to form secondary spermatocytes. These cells will also divide through meiosis II to produce four haploid cells, the round spermatids, which result from the division of every spermatocyte. Each round spermatid cell then differentiates to form elongated spermatids; this process ends when the cells are released from the seminiferous epithelium into the seminiferous tubule lumen. At this point, the released cells are termed spermatozoa (Figure 1.3B) (Rooij and Russell, 2000). These spermatogenesis cell types are associated with and supported by Sertoli

cells via cell-cell contact with the germ cells. From the testis tubules, spermatozoa flow into the epididymis via the rete testis to become mature (Holstein *et al.,* 2003).

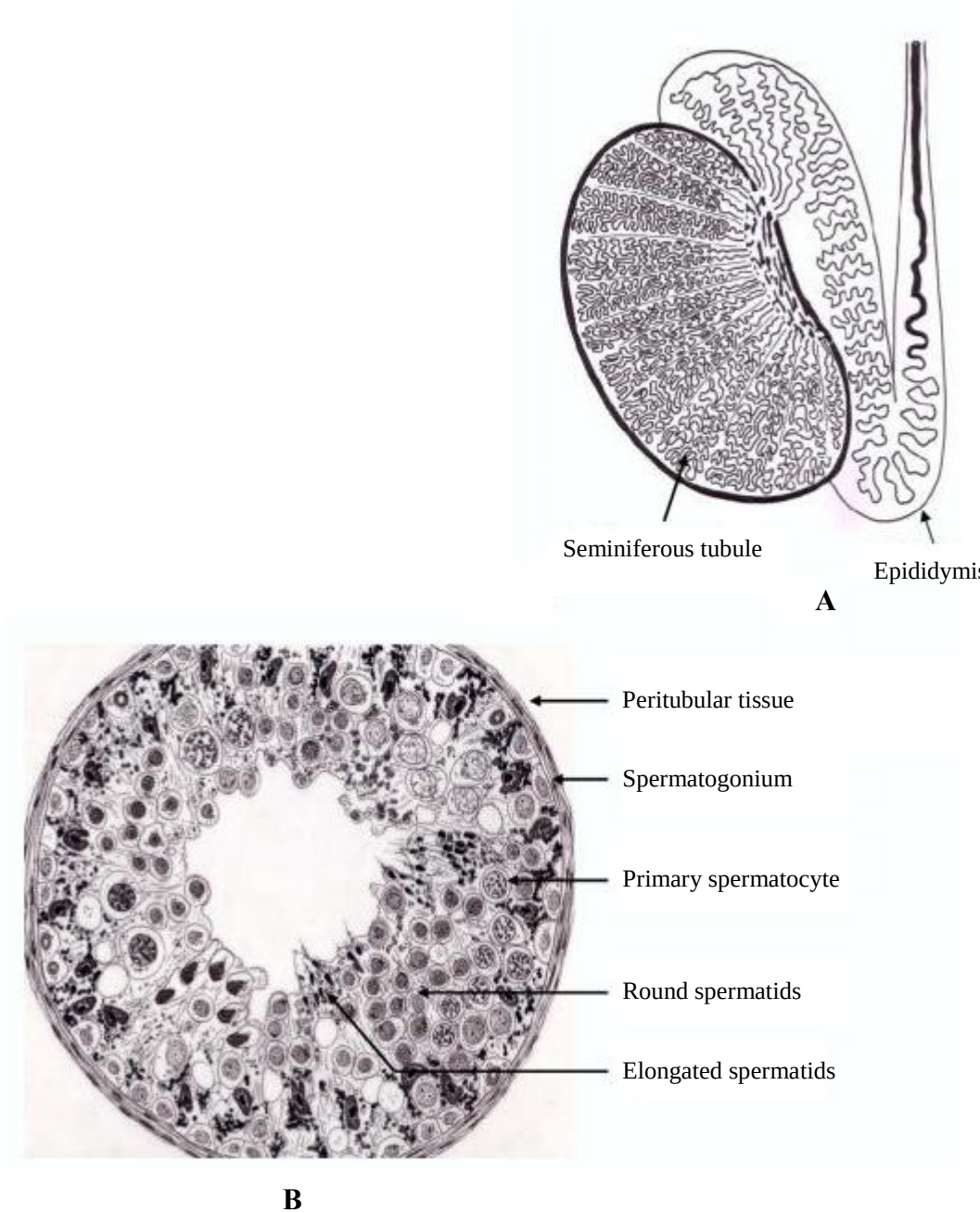


Figure 1.3 The structure of the human seminiferous tubule. (A) The human testis is composed of seminiferous tubules. **(B)** A cross-section of an adult human testis, showing a section of seminiferous tubule at different stages of their development of sperm cell production during spermatogenesis (Adapted from Holstein *et al*., 2003).

1.3.3 Structure and functions of the blood-testis barrier

The testis is a site of immune privilege, protected by the BTB, which is one of the tightest blood-tissue barriers in mammals (Li *et al*., 2012; Lie *et al*., 2013; Mruk and Cheng, 2010). The seminiferous epithelium is anatomically divided into two distinct basal and luminal compartments by the BTB, which is found near the basement membrane in the human testis. The SSC and spermatogonia reside in the basal compartment of the seminiferous epithelium, whereas the remaining germ cells reside in the luminal compartment (Fijak and Meinhardt, 2006; Mruk and Cheng, 2010). The BTB is formed by specialised tight junctions between adjacent Sertoli cells in the seminiferous epithelium (Figure 1.4) (Su *et al*., 2011).

Large molecules, particularly proteins, cannot diffuse from the blood into the lumen of a seminiferous tubule owing to the presence of the BTB. This prevents the production of anti-sperm antibodies, which would result in infertility (Mirandola *et al*., 2011). The presence of autoantigens in the germ cells, outside of the BTB in the basal compartment of the seminiferous epithelium is quite remarkable. Cancer-testis antigens are produced in these germ cells; however, no male human antibodies are developed against those autoantigens within germ cells, except in pathological conditions even through the cells reside outside the BTB (Caballero and Chen, 2009; Simpson *et al*., 2005). This is because the human leukocyte antigen is not expressed in germ cells. Therefore, gene expression does not result in the direct presentation of the antigenic peptides and these normal germ cells should not be the immune responses targets (Ghafouri-Fard and Modarressi, 2012a).

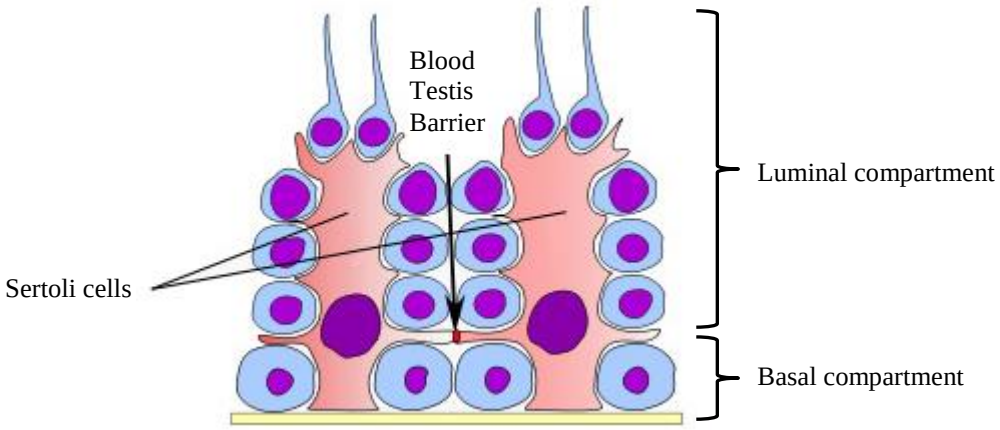


Figure 1.4 The blood-testis barrier. The location of the blood-testis barrier during the seminiferous epithelium cycle of spermatogenesis in the human testes. The main task of the BTB is to protect the developing germ cells from autoimmune attack (Adapted from Modarressi and GhafouriFard, 2011).

1.3.4 The role of retinoic acid in spermatogenesis

Spermatogenesis is a highly regulated process of cellular differentiation that leads to to the initiation of mature sperm from spermatogonial stem cells. One regulatory compound is retinoic acid (RA), a metabolite of Vitamin A (retinol), which is essential for male reproduction. The action of RA in the testis is mediated by activation of its receptors (Livera *et al*., 2002; Wang and Wu, 2013). Vitamin A is normally stored in the liver and Sertoli cells and transported to target tissues. However, spermatids and testicular and epididymal sperm also store retinoids (Ren and Bishop, 1989). Oxidation of Vitamin A to produce RA generally occurs within the target tissues (Paik *et al*., 2004).

RAs have two distinct families of intercellular receptors. The first type is the retinoic acid receptors (RARs), which bind all-*trans* and 9-*cis* RA and the second receptor type is the retinoid X receptors (RXRs), which bind 9-*cis* retinoic acid only (Livera *et al*., 2002). In addition, each family comprises three classes, α, β, and γ, each encoded by different genes. RARα is expressed mainly in the Sertoli cells, RARβ in spermatids and RARγ in spermatogonia (Vernet *et al*., 2006).

These receptors are present in the testis at different developmental stages. The function of these receptors is to regulate gene expression by binding to specific elements in the promoter regions of genes under the control of retinoic acids (Doyle *et al*., 2007). Regulation of the expression of RA receptors is clearly essential for mammalian spermatogenesis and fertility. In the mouse, the spermatogenesis process is arrested at the A spermatogonia stage under conditions of Vitamin A deficiency. However, addition of Vitamin A restores the differentiation of spermatogonia to A1 spermatogonia (Hogarth and Griswold, 2010). Recent studies support the conclusion that retinoic acid is required for adult male spermatogonial differentiation (transition from A to A1) and the entrance into meiosis (Figure 1.5) (Matson *et al*., 2010; Snyder *et al*., 2010).

Sertoli cells have been determined to be the main site of RA synthesis in the testis and are also responsible for delivering RA to germ cells. The Sertoli cells also distribute Vitamin A to germ cells or spermatogonia, which can access Vitamin A and RA directly from serum by means of the vascular system (Livera *et al*., 2002). Inside the cells, such as in spermatogonia, RA is mainly produced from Vitamin A by two biochemical steps. The first step is the oxidation of Vitamin A into retinaldehyde and the second step is conversion of retinaldehyde to RA (Figure 1.6) (Griswold *et al*., 2012).

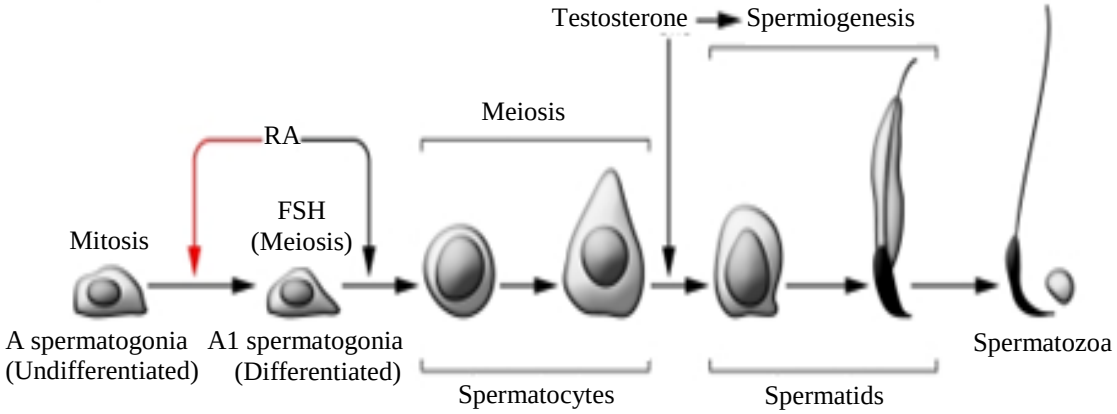


Figure 1.5 Regulation of cellular differentiation in the mammalian testis. The different stages of spermatogenesis in the adult, which takes place in the seminiferous epithelium of the testis tubules. These germ cells are regulated by different regulators. For example, undifferentiated (A type) and differentiated (A1 type) spermatogonia undergo mitotic division, regulated by the FSH hormone. Retinoic acid (RA) also appears to be a requirement for the proliferation and differentiation of spermatogonia A to A1 and for the production of spermatocytes. The initiation of meiotic division also requires RA. Testosterone is needed for the differentiation of secondary spermatocytes and spermiogenesis, which includes round spermatids, elongated spermatids and finally forms functional gametes called spermatozoa or sperm (Adapted from Hogarth and Griswold, 2010).

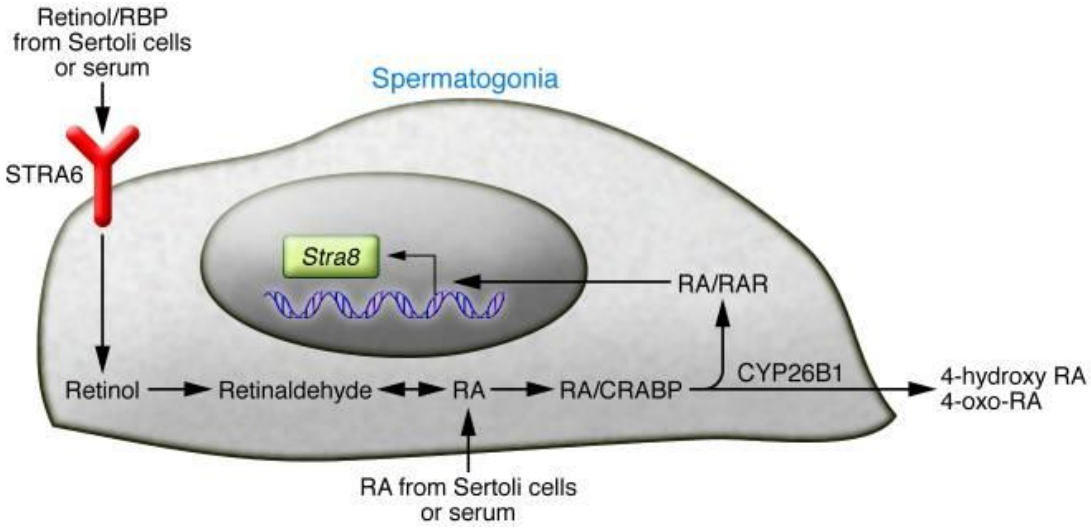


Figure 1.6 Metabolism of Vitamin A in the spermatogonia. The metabolisme of Vitamin A and the cellular mechanism of RA action in the spermatogonia. Vitamin A is provided to the germ cells by the blood, where it circulates and binds to retinol-binding protein (RBP). The RBP carries the majority of retinol in the circulation and binds to the membrane receptor, STRA6 to facilitate the cellular uptake of retinol inside the cells. However, RA is also provided directly inside the spermatogonia by means of Sertoli cells or by serum. Vitamin A can also be converted to RA by a two-step process. The produced RA is then bound to cellular RA binding protein (CRABP), which facilitates its transport to the nucleus. In the nucleus, RA is bound to the RARs, such as RARγ, to form the RA/RAR complex, which allows chromatin opening and stimulates transcriptional activity of a number of genes, such as *STRA8* that has been identified to have an essential role in meiotic progression. Degradation of RA is also an important process to protect cells from excessive RA stimulation. The RA degradation occurs by CYP26B1 enzyme to produce 4-oxo and 4-hydroxy forms, which are then secreted across the cell membrane (Adapted from Hogarth and Griswold, 2010).

1.4 Meiotic division in eukaryotic cells

1.4.1 Meiosis and sexual reproduction

In the sexual reproduction process, two gametes fuse during fertilisation to produce a zygote. However, to prevent the doubling of the genetic material during cell division with every new generation, the number of chromosomes must be reduced by half during the formation of the gametes. This reduction process is performed by a specialised nuclear division termed meiosis (Petronczki *et al*., 2003). Non-reproductive cells comprise two distinct copies of each chromosome, called homologues. One homologous chromosome is inherited from the paternal lineage, and one from the maternal lineage (Lee and Amon, 2001; Villeneuve and Hillers, 2001).

Meiosis takes place solely in the testes (males) and ovaries (females) to produce haploid sex cells or gametes, the sperm and egg, from diploid progenitor cells (Page and Hawley, 2003). Meiotic division consists of two types of cell division: the first is a reductional division (meiosis I) and the second is an equational division (meiosis II) (Gerton and Hawley, 2005).

1.4.2 Phases of cell division

The interval time between each mitotic cell division is termed a cell cycle (Boydston-White *et al*., 2005). The mitotic cell cycle consists of two basic stages: interphase and M phase. Interphase can be subdivided into three ordered stages: G1 (Gap 1), S (DNA synthesis) and G2 (Gap 2) (Figure 1.7). S phase is defined as the stage where the DNA replication occurs and the M phase is where the cell is ready to divide into 2 daughter cells (mitosis). There are two gaps before the S phase and after the M phase, G1 and G2. In the G1 phase, the cell is growing and also preparing for the process of DNA replication, whereas in the second gap, G2, which comes after the S phase, the cell prepares for the process of division (Strahm and Capra, 2005; Thomas and Goodyer, 2003). Non-dividing cells exist outside of the cell cycle in a phase called G0 (O'Farrell, 2011). Most cells spend the majority of their time in interphase; M phase makes up only a relatively short period (Boydston-White *et al*., 2005).

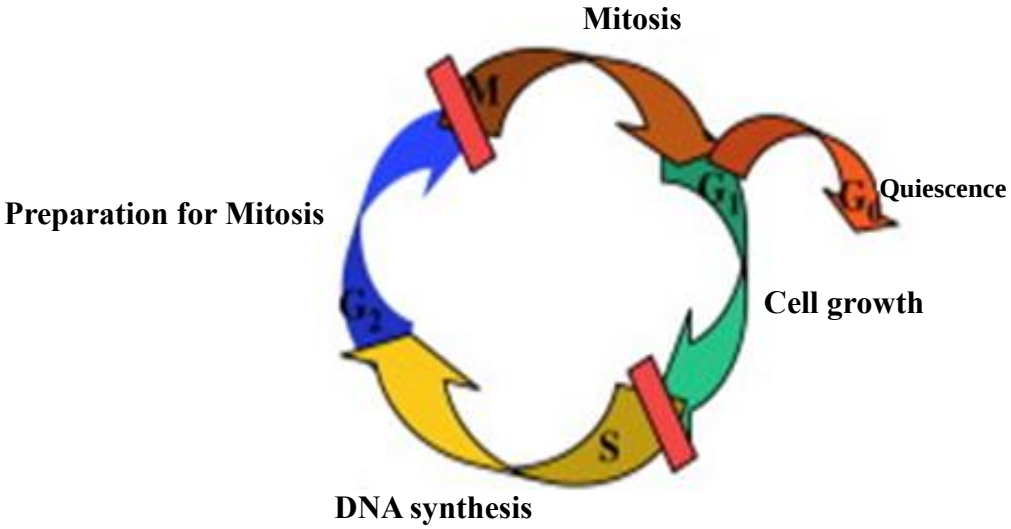


Figure 1.7 The progression of mitotic cell cycle. The phases of the cell cycle are designated G1, S, G2, and M. In the first gap, the cells increase in size. In the S phase, DNA synthesis occurs to produce DNA ready for cell division. In the second gap, the cells grow and produce new proteins. In the M phase, cell growth stops, and each cell undergoes orderly division. In the G0 phase, the cell is in a resting state and does not progress through the cell cycle. (Adapted from Foster, 2008).

1.4.3 The first meiotic division, Meiosis I

Meiosis I is divided into four phases: prophase I, metaphase I, anaphase I, and telophase I. Prior to meiosis, chromosome duplication occurs during the interphase stage in order to double the sister chromatid pairs. The sister chromatids remain associated at the centromeres; consequently, each set of chromosomes is composed of two identical sister chromatids (Lee and Amon, 2001).

Many unique events take place during prophase I, including homologous chromosome pairing, synapsis, DNA double-strand break (DSB) formation and repair, crossover, and chromosome condensation (Kleckner, 1996). Prophase I can be sub-classified into five stages: leptotene, zygotene, pachytene, diplotene and diakinesis, on the basis of the morphology of chromosomes and the association of homologous chromosomes during synapsis (Klutstein and Cooper, 2014; Tsai and McKee, 2011).

In the first sub-stage, leptotene, the duplicated sister chromatids start to condense and coil, and the chromosomes become shorter and thicker. In the zygotene sub-stage, homologous chromosomes draw close to each other to a distance of roughly 100 nm mediated by a special structure celled the synaptonemal complex (SC), which begins to form between paired homologous chromosomes in a process termed synapsis. As a result, the pairs of chromosomes consist of four chromatids, with one chromosome coming from each parent. Each pair of homologous chromosomes is known as a bivalent. By the third sub-stage, pachytene, synapsis is complete and the paired chromosomes are held together tightly with the aid of the SC and structures termed chiasmata (Youds and Boulton, 2011). The chiasma is the physical link between homologous chromosomes during meiosis (Hirose *et al*., 2011). In addition, crossing over between homologous chromosomes occurs and DNA is exchanged between the bivalents in a process called homologous recombination (Bowles and Koopman, 2007; Youds and Boulton, 2011). One consequence of crossing over is the generation of a new combination of genetic material in the gametes. In the fourth sub-stage, diplotene, the homologous chromosomes begin to separate in a process called desynapsis, but remain connected through sister chromatid cohesion and chiasmata until anaphase I. The final subdivision of prophase I is diakinesis. In this sub-stage, the SC has completely dissociated, the chromosomes continue to condense further and spindle formation occurs (Tsai and McKee, 2011).

In metaphase I, spindle fibres are formed and the bivalents are aligned in a double row along the metaphase plate of the cell. The chiasmata promote the monopolar orientation of the sister kinetochores, which leads to the correct segregation of the maternal and paternal chromosomes to opposite poles (Hirose *et al*., 2011). In anaphase I, sister chromatid cohesion along the chromosome arms is resolved, but the sister chromatids remain bound to each other at the centromeres until the beginning of anaphase II. At this stage, the microtubules pull one set of homologous chromosomes toward the opposite poles of the cell (Lee and Amon, 2001; Villeneuve and Hillers, 2001). The microtubule has two ends; one end links to the kinetochore on the chromosome centromere, but the other end is linked to a centrosome or a spindle pole body in human and yeast cells, respectively (Miller *et al*., 2012). In telophase I, a new nuclear membrane forms around the segregating chromosomes to produce two separate nuclei. The resulting nuclei comprise half the number of chromosomes of the original cell and so meiosis I is reductional in nature (Villeneuve and Hillers, 2001).

1.4.4 The Second meiotic division, Meiosis II

Meiosis II, an equational division that does not reduce chromosome number, is a mitosis-like division. When meiosis II begins, each daughter nucleus contains the haploid number of chromosomes, and each chromosome consists of two chromatids associated at the centromere. For this reason, the second meiotic division is necessary to segregate the chromatids from each other (Page and Hawley, 2003). It is shorter than meiosis I, although it consists of four phases: prophase II, metaphase II, anaphase II, and telophase II. It occurs without further DNA replication (Wassmann, 2013).

During prophase II, the chromosomes condense again and the membrane of the nucleus breaks down. In metaphase II, each chromosome contains two chromatids, and these are arrayed on the cell equator by spindle fibres. In anaphase II, the centromere of each chromosome separates, and the spindle fibres pull one chromatid of each chromosome to the opposite end of the cell (Tsai and McKee, 2011). At the end of meiosis II, cytokinesis (the division of the cytoplasm) occurs to divide the original cell into four new haploid cells containing half the original number of chromosomes, each with a diverse genotype. This stage is called telophase II (Page and Hawley, 2004; Zickler and Kleckner, 1998). During meiosis II, the sister kinetochores are attached to microtubules from the different pole (Bio-orientated) (Dudas *et al*., 2011). The key stages of the two divisions of meiosis are shown in Figure 1.8.

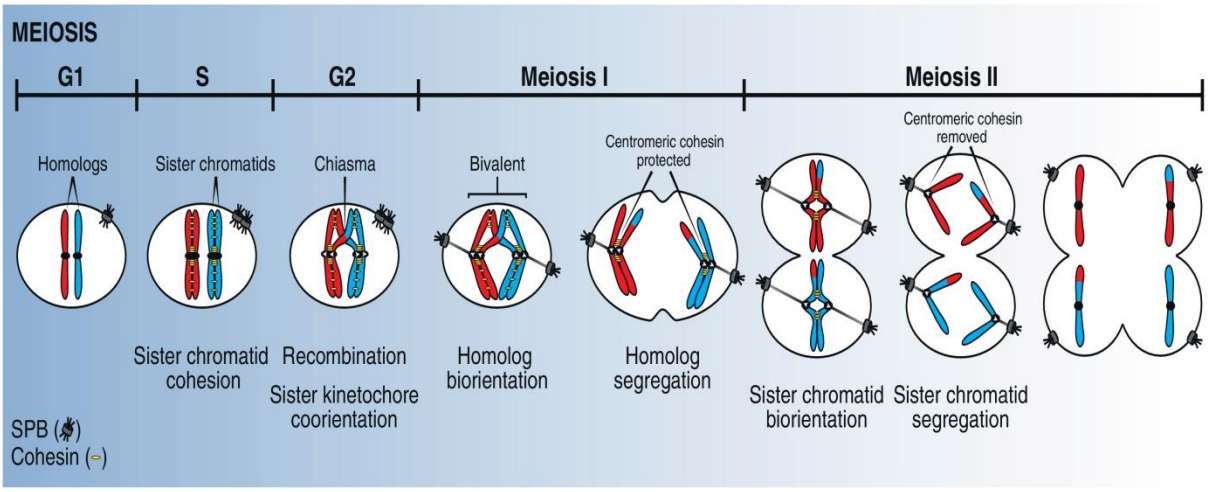


Figure 1.8 The process of meiotic cell division. Meiosis consists of one round of DNA replication that occurs in S phase followed by two nuclear divisions including meiosis I and meiosis II. Each sister chromatid in a non-homologous chromosome, before entering meiosis I, joins by cohesin complexes (yellow). In prophase I, homologous chromosomes connect to each other and recombination takes place to form chiasmata. Homologous centromeres attach to the microtubules from the spindle poles (SPB), which is known as the centrosome in eukaryotes. The paired homologues (bivalents) are aligned on the metaphase plate during metaphase I. This is followed by segregation of homologues to the opposite poles of the cell to initiate anaphase I with the resolution of chiasmata. The generation of two daughter cells occurs at telophase I. In meiosis II, chromosomes align along the metaphase plate during metaphase II and then the sister chromatids separate from each other in anaphase II. The process ends with telophase II producing four haploid cells containing half the original number of chromosomes. Cohesin along chromosome arms in meiosis I is cleaved from the bivalent at anaphase I to facilitate segregation of homologues, but the cohesin at the centromeres (sister kinetochores) is protected, which keeps the two sister chromatids bound to each other. However, at anaphase II, the cohesin remaining (centromeric cohesin) between sister kinetochores is removed, which allows the sister chromatids to orient, segregate and to move to the opposite poles (Adapted from Miller *et al*., 2013).

1.4.5 Differences between meiosis in human males and females

Gametes are haploid, and are produced by meiosis starting from a diploid germ cell. The process of meiosis is extremely important for both males and females; however, there are some essential differences between the sexes. Meiosis begins during foetal development in females and after birth in males. The primary oocytes are analogous to the primary spermatocytes in the male. In females, the primary oocyte undergoes meiosis I up to the diplotene sub-stage of prophase I, and then meiosis is arrested. Once the female reaches puberty, in every menstrual cycle, one of the arrested primary oocytes will proceed to complete meiosis I to form secondary oocyte and a first polar body (PB1). However, the secondary oocyte will enter meiosis II and arrest again in metaphase II and it remains in arrest until fertilisation occurs, at which time the entire meiotic process is completed and a second polar body (PB2) is formed. One oocyte will only produce one functional gamete termed an ovum or egg (Bowles and Koopman, 2007; Holt and Jones, 2009; Hou *et al*., 2013b). In contrast, in males, germ cells are arrested in spermatogonia at G0/G1 stage of the mitotic cell cycle. Therefore, the mitotic division will not complete and meiosis will not initiate until the embryos are born. After the first week of life, meiosis I is completed and secondary spermatocytes are formed, but the division is arrested again until puberty. Once the male reaches puberty, the arrested secondary spermatocytes will continue to complete meiosis II, forming 4 functional gametes termed spermatozoa from one spermatocyte (Figure 1.9) (Clagett-Dame and Knutson, 2011; Hunt and Hassold, 2002).

1.4.6 Generation and repair of meiotic DNA Double-Strand Breaks (DSBs)

Meiotic DSBs are generated by the SPO11 protein (topoisomerase type II-like) (Bergerat *et al.*, 1997; Keeney *et al*., 1997). Most organisms have only one copy of SPO11, and the protein acts as a homodimer to make DSBs (Kenney and Neale, 2006). SPO11 is also essential for the assembly of the SC (Keeney, 2001). SPO11 relies on many accessory factors that support its activation (Keeney and Neale, 2006). Meiotic DSBs are concentrated in narrow regions of chromosomes known as hotspots, which have high rates of meiotic recombination (Baudat *et al*., 2013; de Massy, 2013). After generating DSBs, SPO11 remains covalently bound to the 5' end of each DNA strand and must be removed to generate single-strand 3' ends (Longhese *et al*., 2010).

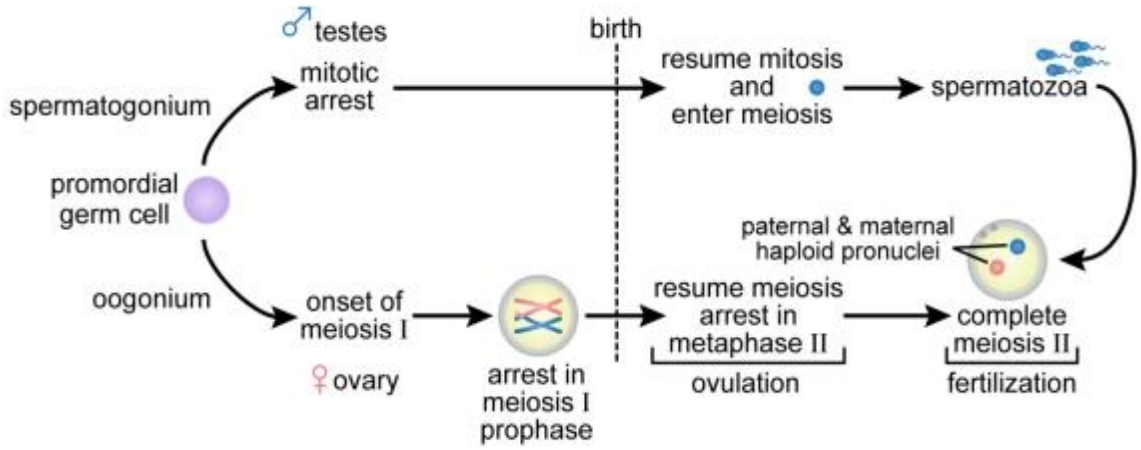


Figure 1.9 Schematic diagrams showing the distinction between meiosis in the testes and ovaries. Meiosis in women initiates in embryogenesis (before birth) and arrests in the prophase of meiosis I. The primary oocytes remain in this state until just before ovulation, at which point they complete of the first division and begin the second division, which is arrested again in metaphase II. After fertilisation of the egg, meiosis II is completed. By contrast, in men, germ cells are arrested in mitotic process in embryogenesis and are completed after birth. Primary spermatocytes initiate both meiotic divisions after birth and complete at puberty (Adapted from Clagett-Dame and Knutson, 2011).

SPO11 is removed by the MRN complex, which also functions in the early steps of meiotic DSB repair in somatic cells (Neale and Keeney, 2006). The MRN complex is an endonuclease, and consists of three proteins, MRE11, RAD50 and NBS1 in mammals and fission yeast and the MRX (Mre11, Rad50 and Xrs2) in budding yeast (Lisby *et al*., 2004). These proteins have an important role in repairing DSBs in somatic and meiotic cells (Borde, 2007; Stracker and Petrini, 2011).

1.4.7 Mechanism of meiotic recombination

Meiotic recombination is essential to meiotic division among most eukaryotes. It may result in both the formation of crossovers (resulting from reciprocal exchange between homologous chromosomes), or non-crossover (nonreciprocal exchange). A minority of the meiotic DSBs result in crossover products in eukaryotic cells (Sung *et al.*, 2003; Youds and Boulton, 2011). In mitotic cells, repair of DSBs occurs mostly via the sister chromatid, and noncrossover recombination is the predominant pathway. In meiotic cells, sister chromatid repairing is repressed, and the invasion of one chromatid of the homologues is stimulated (see Section 1.4.8) (Niu *et al.*, 2005).

The functions of recombination in mitotic cells and meiotic cells are distinct. Mitotic recombination occurs in somatic cells to repair DSBs caused by either a DNA replication problem or exposure to DNA damaging agent, such as ionising radiation and/or genotoxic chemicals. In contrast, meiotic recombination has two functions. First, it is very important in forming a physical connection between homologues and it plays a fundamental role in chromosome alignment and segregation in meiosis I. Second, homologous recombination promotes genetic variation by generating new combinations of paternal and maternal alleles (Strom and Sjogren, 2007).

1.4.8 Induction of crossover (CO) and non-crossover (NCO)

Two main models for the meiotic recombination mechanism have been proposed for both meiotic and mitotic divisions. The first model is DSB repair (DSBR) and the second is the synthesis-dependant strand annealing (SDSA) pathway (Figure 1.10). Meiotic recombination starts when a DSB occurs in one chromatid in a homologue pair. Fragments of DNA around the 5' ends of the break are cut away by MRN complex. The generated 3' ends are covered with proteins called RAD51 and DMC1 in mammals (Rad51 and Dmc1 in *S. cerevisiae*) (San Filippo *et al*., 2008), which promote the invasion of the opposite homologous DNA duplex in

a process called strand invasion, to generate a heteroduplex and a displacement loop (D-loop) that is completed by DNA synthesis (de Massy, 2003; Handel and Schimenti, 2010; Hunter and Kleckner, 2001; Youds and Boulton, 2011). The recombinase, RAD51, is essential for both meiotic and mitotic recombination, but DMC1 is a meiosis-specific protein that is required for normal meiosis (Serrentino and Borde, 2012).

The result of this process is either crossover or non-crossover products. Most DSBs are repaired by SDSA, but about 15% of DSBs are repaired by DSBR. In the DSBR pathway, the second end of the DSB is captured; as a result, a structure called a double Holliday junction (dHJ) is formed by DNA synthesis and ligation. The Holliday junction is then cleaved in the same direction (vertical or horizontal) or the opposite direction (vertical and horizontal) followed by DNA ligation to produce non-crossover (NCO) or crossover (CO), respectively (Cromie and Smith, 2007; de Massy, 2003; Hunter and Kleckner, 2001; Youds and Boulton, 2011). Models have also been proposed involving single Holliday Junctions, but it remains unknown which is correct for human meiosis (Cromie and Smith, 2007).

In the SDSA model, the DSBs occurs without the formation of the dHJs, so the D-loop is released by displacement of the newly synthesised DNA, which anneals to the other ssDNA end of the break. Repair of the break leads to generate either gene conversion or non-crossover (de Massy, 2003; Hunter and Kleckner, 2001). Resolution of dHJ during meiosis in budding yeast can be performed via three different resolvase enzymes (nucleases), Mus81-Mms4, Slx1-Slx4 and GEN1/Yen1, but can also be resolved by the Exo1 exonuclease, Sgs1 and the MutLγ complex (M1h1-M1h3) (Zakharyevich *et al*., 2012).

1.4.9 The cohesin complex

The cohesin complex is a complex protein that binds the sister chromatids together along chromosome arms and at the centromere (Figure 1.11). The cohesin complex associates with chromosomes during the DNA replication (S phase) of both mitotic and meiotic cells (Mehta *et al*., 2013; Nasmyth, 2001). Defects in sister chromatid cohesion may result in the generation of aneuploidy and cancer (Barber *et al*., 2008; Sajesh *et al*., 2013). Cohesin is responsible for sister chromatids pairing and separating, as well as the repair of DNA DSBs during either mitosis or meiosis (Ishiguro *et al*., 2014; Sjögren and Ström, 2010) and it is also important for the generation of tension across centromeres (Ishiguro and Watanabe, 2007).

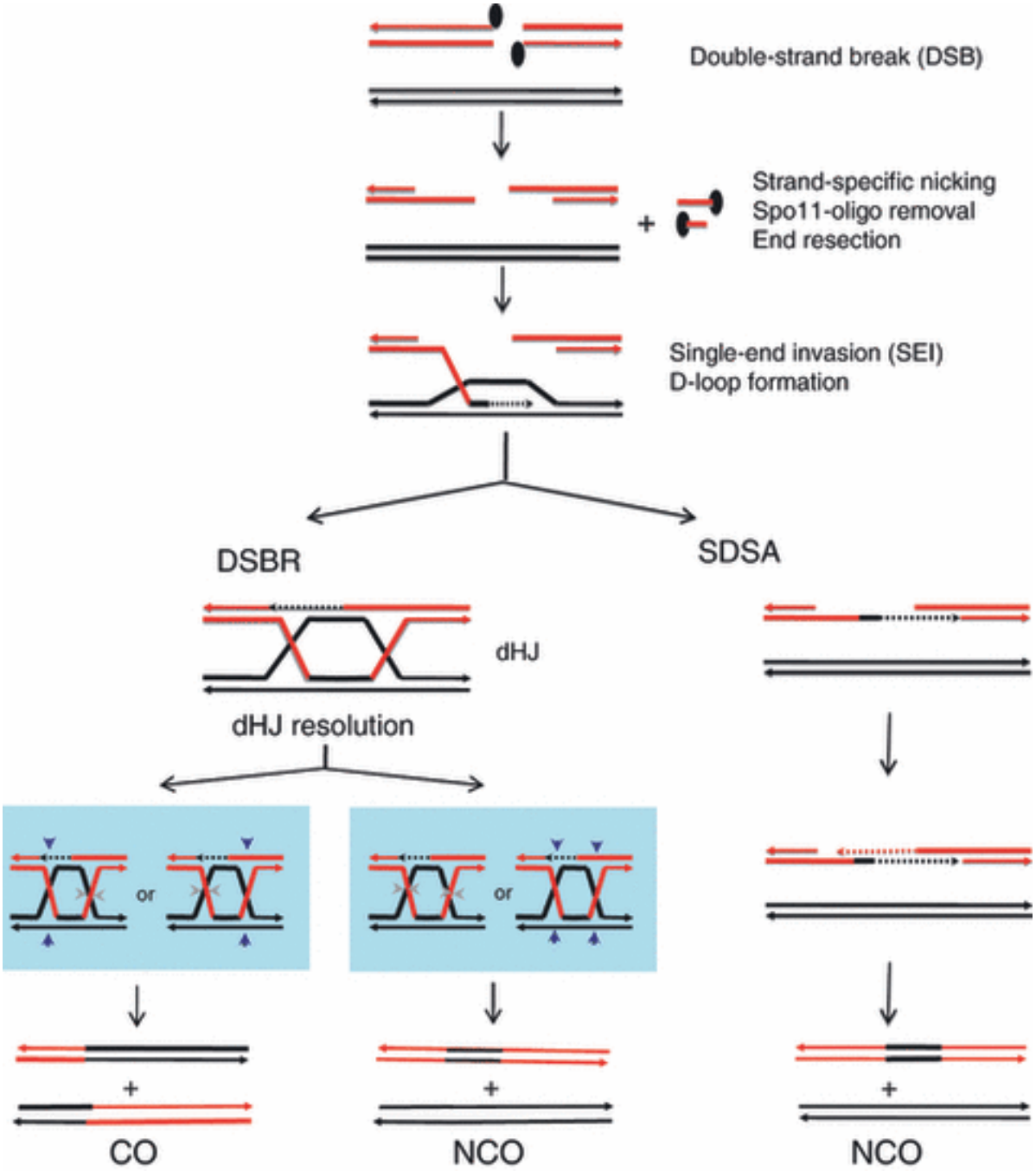


Figure 1.10 The mechanism of DSB repair for meiotic recombination in meiosis I. The DSBs are mainly formed by SPO11 and one SPO11 protein remains covalently bound at each 5' end of the DSB. The 5' end is removed by the MRN complex and single strand (ssDNA) overhangs is generated in both sides of the DSB ends. One of the overhanging 3' ended tail invades the homologous chromosome (single end invasion). However, the repair mechanism can follow two different pathways. First, in the DSBR pathway, the second 3' overhang invades the displaced strand (second end capture) and DNA synthesis produces a double Holliday junction (dHJ). Therefore, the dHJ resolution gives rise to either CO or NCO products. Second, in the SDSA pathway, the DSBs occur without the creation of a dHJ, which results in an NCO product. Meiotic recombination of the majority of DSBs results in either the formation of NCO or gene conversion (Adapted from Szekvoelgyi and Nicolas, 2010).

The cohesin complex is a tripartite ring-shaped structure. In mammalian germ cells, for example, cohesin complex contains: (a) Two members of the structural maintenance of chromosomes (SMC) protein family, named SMC1 and SMC3. Two variants also exist for the SMC1 subunits (SMC1α and SMC1β); (b) Three members of the Kleisin family proteins (RAD21/SCC1; REC8 and RAD21L); and (c) Three accessory subunits, SCC3/SA (STAG1, STAG2 and STAG3) (Nasmyth, 2011; Uhlmann, 2004). Four of these proteins have been identified as being meiosis-specific cohesin subunits, REC8, SMC1β, RAD21L and STAG3 (Nasmyth, 2011). By contrast, the mitotic cohesin complex consists of four subunits including SMC1α, SMC3, RAD21/SCC1 and SCC3/SA (STAG1 and STAG2) (Nasmyth and Haering, 2009).

During mitosis, cohesin is destroyed to allow the separation of sister chromatids, but this is not possible in meiosis, because sister chromatids in the bivalent are connected by homologues arm cohesion. Therefore, the removal of sister chromatid cohesion in meiotic cells is performed in two steps. The first step allows the segregation of bivalent chromosomes by REC8 degradation from the chromosome arms through a specific protease called separase, which is active during anaphase I, while the second step is cleavage of REC8 from the centromeric regions during anaphase II by separase activity, which results in sister chromatid segregation from each other towards opposite poles (Brar *et al*., 2006; Mehta *et al*., 2012).

In the first meiotic division, the sister kinetochores are mono-orientated and the centromeric cohesions are protected (Dudas *et al*., 2011). Chromosome segregation during cell division in mammals is regulated by two member of the Shugoshin (SGO) protein family, SGO1 and SGO2. The function of these proteins is to protect centromeric sister cohesion of sister chromatids and to facilitate bio-orientation attachment of the sister kinetochores (Yao and Dai, 2012). SGO1 plays an essential role in cohesin protection in mitosis and meiosis, but SGO2 functions in meiosis only (Gómez *et al*., 2007; Yao and Dai, 2012). In fission yeast, the Shugoshin proteins attach with protein phosphatase SA (PP2A), which leads to dephosphorylation of REC8 during meiosis I and inhibits separase activity (Xu *et al*., 2009).

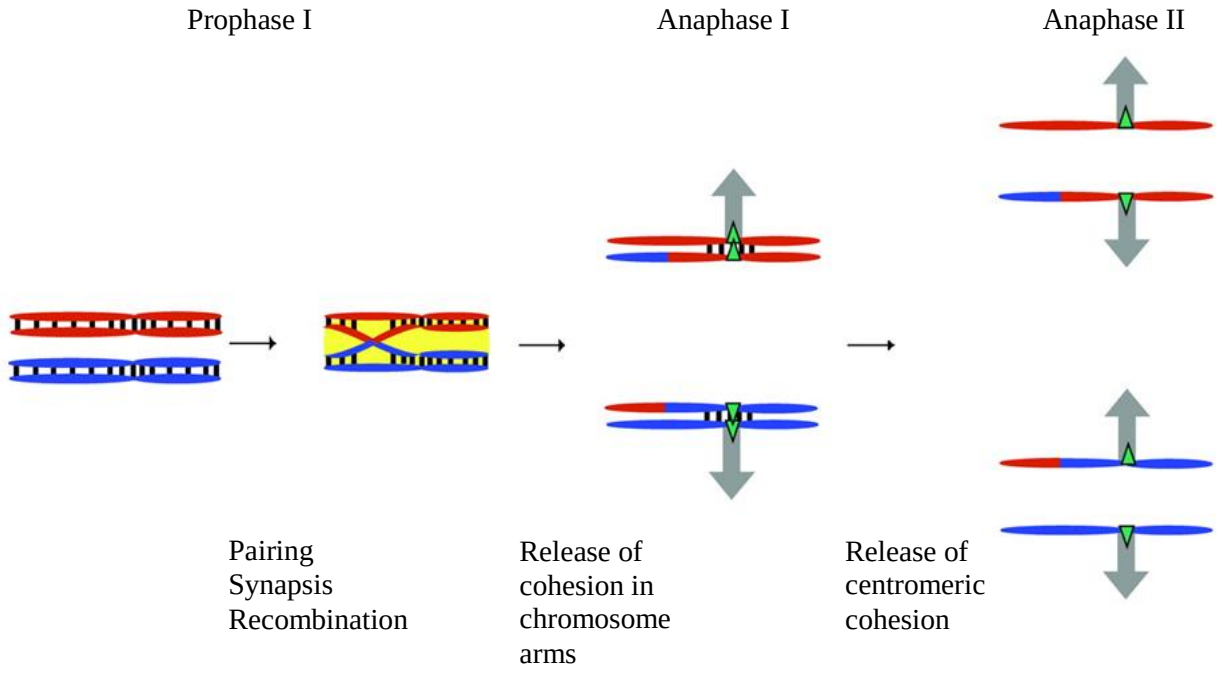


Figure 1.11 Sister chromatid cohesion in mammalian undergoes meiosis. The sister chromatids in each homologue connect to each other via cohesin (black bars). A pair of homologous chromosomes consists of two pairs of sister chromatids and synapsis occurs through the formation of synaptonemal complex (yellow). Meiotic recombination takes place between non-sister chromatids. In anaphase I, cohesin at chromosomal arms is dissolved and the bivalent chromosomes segregate, but cohesion at the centromeric is retained. In anaphase II, centromeric cohesion is destroyed, and the sister chromatids separate to the opposite sides of the cells with the aid of bio-orientation of kinetochores (green arrowheads) (Adapted from Revenkova and Jessberger, 2005).

1.4.10 Homologous chromosome pairing

The homologous chromosomes align and pair during the meiotic prophase in order to recombine and form chiasmata between one sister chromatid from one homologue and another sister chromatid from the other homologue. In several organisms, such as mammals and yeast, homologous chromosome pairing requires the formation and repair of DNA DSBs (Inagaki *et al.*, 2010; Klutstein and Cooper, 2014). During meiotic prophase I, three basic events occur: the pairing of homologous chromosomes, synapsis and meiotic recombination. Chromosome pairing begins in the leptotene stage, progresses in the zygotene stage, and it is completed in the pachytene stage (Klutstein and Cooper, 2014; Page and Hawley, 2004).

1.4.11 Chromosome synapsis

Most of the studies in synapsis and meiotic recombination have been conducted on different organisms, such as *Saccharomyces cerevisiae*, *Drosophila melanogaster* and *Caenorhabditis elegans* (Page and Hawley, 2004). Synapsis takes place along the entire length of the homologous chromosomes and is achieved by the SC formation (Fraune *et al*., 2012). However, synapsis does not occur between the sex chromosomes of male mammalian cells (Handel, 2004).

The formation of the SC, the tripartite protein structure that forms between two homologous chromosomes, leads to the stable association of these homologues during prophase I of meiotic cells (Page *et al*., 2008). The SC consists of three distinct structural parts: two structures known as the lateral (axial) elements (LEs) along the axes of homologous chromosomes, a central connecting element (CE) and transverse filaments (TFs) that connect the lateral and central elements (Figure 1.12). The SC is gradually detached in the diplotene stage (Fung *et al.*, 2004; Page and Hawley, 2004).

Cytological observations indicate that synapsis in humans is initiated from the ends of telomeres of homologous chromosomes (Brown *et al.*, 2005). During the first phase of chromosome pairing, all telomeres on the nuclear membrane are clustered, whilst the arms of homologous chromosomes are moved to the nucleus centre (Harper *et al*., 2004; Scherthan, 2007). Chromosomal movement is very important for the correct pairing of homologous chromosomes (Ding *et al.*, 2006; Klutstein and Cooper, 2014). The induction and repair of the meiotic DSBs are also fundamental to achieving synapsis (Kuznetsov *et al.*, 2007).

The SC in mammals is composed of seven known proteins, which play a fundamental role in its structure during the prophase of meiosis I. These proteins are synaptonemal complex protein 1, 2 and 3 (SYCP1, SYCP2 and SYCP3); synaptonemal complex central element protein 1, 2, and 3 (SYCE1, SYCE2 and SYCE3), as well as the testis expressed 12 (TEX12) (Davies *et al*., 2012).

Synapsis particularly relies on the central element and transverse filament proteins of the SC. Four central element proteins have been determined in mammalian cells including SYCE1, SYCE2, SYCE3 and TEX12 (Yang and Wang, 2009). Mutation in the CE and TF proteins leads to male/female infertility in mice (Schramm *et al*., 2011). In mammals, the transverse filament consists of a single known protein named SYCP1 (Page and Hawley, 2004).

In contrast, the lateral element proteins consist of SYCP2, SYCP3, HORMA-domain proteins (HORMAD1), and the sister chromatid cohesion provided by the meiotic cohesin complex arises due to four proteins REC8, STAG3, SMC1β and SMC3 (Page and Hawley, 2004). These proteins are also essential for the correct pairing of homologous chromosomes, and the formation of crossovers (Kouznetsova *et al.,* 2005; Yang *et al.,* 2006). For instance, deletion of the SYCP3 gene in male mice causes heterologous synapsis, and apoptosis occurs in spermatocytes (Yuan *et al.,* 2000). In female mice; however, meiotic division is not completely blocked by this deletion, but during the first week of postnatal follicle development, many oocytes are lost in the SYCP3 gene knockout line. Those that survive show a high frequency of embryonic aneuploidy (Yuan *et al.,* 2002).

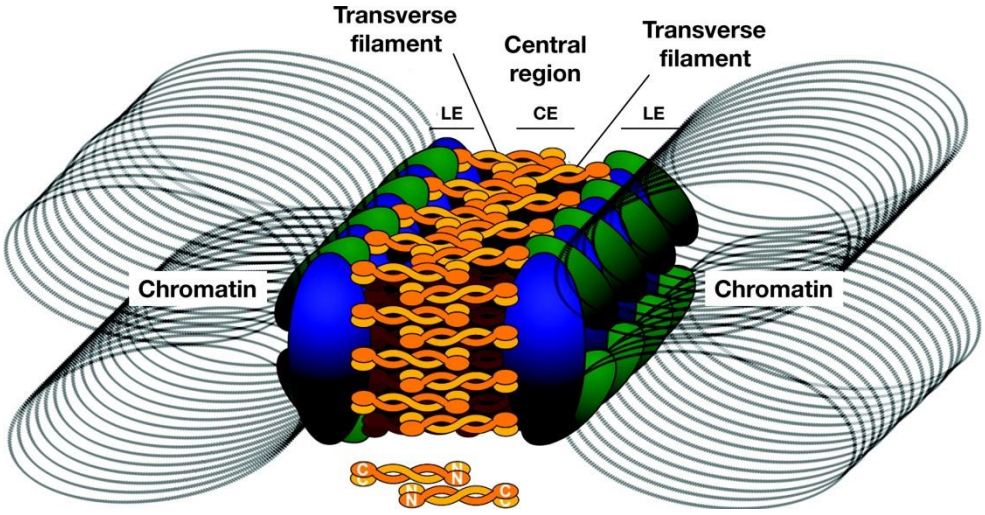


Figure 1.12 The synaptonemal complex structure. The SC is created by lateral elements (LEs), a central element (CE) and proteins known as transverse filaments, which bind the LEs to the CE. In mammals, the LEs contain meiotic cohesins; the structural proteins (SYCP2 and SYCP3) and the HORMA-domain proteins. The CE is composed of four essential proteins including SYCE1; SYCE2; SYCE3 and TEX12. In contrast, the transverse filaments are linked to the LEs by a protein known as the SYCP1 molecule, with the C-terminal embedded in the LEs and interaction of the N-terminal in the middle of the CE. C= carboxyl-terminus; N= amino-terminus (Adapted from Page and Hawley, 2004).

1.5 Project aims

The overall purpose of the present study was to identify candidate genes that could serve as cancer-specific markers. A bioinformatics screening program, which included microarray analysis and an EST analysis pipeline, indicated potential meiotic genes which could serve this purpose. A core aim was to validate these candidate genes, and further explore their potential as oncogenes and/or clinical useful biomarkers

The overarching aims and objectives:

1. Identify new CTA genes via RT-PCR analysis using RNAs isolated from normal human tissues and cancer cells.
2. Functional analysis of two of identified candidate genes (*STRA8* and *C20orf201*) using human cancer cell lines.

Chapter 2.0: Materials and Methods

2.1 Human cell lines and cell culture

2.1.1 Source of human cell lines

The Flp-In T-REx-293 cell line and the liver cancer cell line HEP-G2, were gifted by Dr. J. Müller (University of Warwick). The embryonal carcinoma cell line NTERA-2 clone D1) was gifted by Prof. P.W. Andrews (University of Sheffield) and the A2780 cell line was gifted by Prof. P. Workman (Cancer Research UK Centre for Cancer Therapeutics, Surrey, UK). The following cell lines were obtained from the European Collection of Cell Cultures (ECACC); 1321N1, HT29, HCT116, T84, LoVo, SW480, MCF7, G-361, MM127, COLO800 and COLO857. H460 and MDA-MB-453 cells were purchased from the American Type Culture Collection (ATCC), and human ovarian adenocarcinoma from the cell lines, PEO14 and TO14, were obtained from Cancer Research Technology Ltd.

2.1.2 Culturing the cell lines

All cell lines were maintained at 37°C in a humidified atmosphere containing 5% or 10% CO_2 and passaged based on recommended dilutions and confluences from ATCC. The cell lines were grown in a media containing 10% foetal bovine serum (FBS) (Invitrogen, Catalogue number; 10270, Lot 41Q6208K). The media, cell line description, CO_2 concentration and any other standard condition of each cell line are summarised in Table 2.1. The LookOut® Mycoplasma PCR Detection kit (Sigma Aldrich; MP0035) was routinely used according to the manufacturer's instructions to check for mycoplasma contamination in cell culture.

Table 2.1 Description of the cell standard conditions of different human cancer cell lines

Cell line names	Origin	CO_2	Culture Media conditions
Flp-In T-REx-293	Human embryo kidney carcinoma	5%	Dubeco's modified Eagle's medium (DMEM) + GLATAMAXTM (Invitrogen; 61965) + 10% FBS
HEP-G2	Hepatocellular carcinoma	5%	
NTERA-2 (clone D1)	Human Caucasian pluripotent embryonal carcinoma	10%	
SW480	Human colon adenocarcinoma	5%	
LoVo	Human colon adenocarcinoma	5%	Ham's F12 + 2mM Glutamine + 10% Foetal Bovine Serum (FBS).
1321N1	Human brain astrocytoma	5%	
HT29	Human Caucasian colon adenocarcinoma	5%	McCoy's 5A medium + GLUTAMAXTM (Invitrogen; 36600) + 10% FBS
HCT116	Human colon carcinoma	5%	
G-361	Human Caucasian malignant melanoma		
A2780	Human ovarian carcinoma	5%	DMEM + GLATAMAXTM + 10% FBS and 1xNEAA (non-essential amino acids)
MCF7	Human Caucasian breast adenocarcinoma	5%	
MM127	Human malignant melanoma	5%	RPMI 1640 + GLUTAMAXTM + 10% FBS and 25 mM HEPES
COLO800	Human melanoma	5%	Roswell Park Memorial Institute 1640 medium (RPMI 1640) + GLUTAMAXTM (Invitrogen; 61870) + 10% FBS
COLO857	Human melanoma	5%	
H460	Large cell lung carcinoma	5%	
MDA-MB-453	Human breast carcinoma	0%	Leibovitz's (L-15) medium + GLUTAMAXTM (Invitrogen; 31415) + 10% FBS
PEO14	Ovarian Adenocarcinoma, peritoneal ascites	5%	RPMI 1640 + GLUTAMAXTM + 10% FBS and 2 mM sodium pyruvate
TO14	Ovarian Adenocarcinoma, solid metastasis	5%	
T84	Human colon carcinoma	5%	Ham's F12 + DMEM (1:1) + GLUTAMAXTM (Invitrogen; 31331) + 10% FBS

2.1.3 Preparing of thawing frozen cells and recovery

Vials of frozen cells were removed from the -80°C freezer or the liquid nitrogen tank and immediately immersed in a 37°C water bath. The cells was quickly thawed by gently swirling the vial in the 37°C water bath until there was just a small bit of ice left in the vial. Before opening, the outside of the vial was wiped with 70% ethanol and then the cells transferred from the cryotube to 5 ml of the complete medium in a Falcon tube then collected by centrifugation at 1,500 xg for 5 minutes. The cells were re-suspended in 10 ml of fresh complete media and transferred into T75 cm^2 flask. The cells were then incubated in 37°C incubator at appropriate CO_2 concentration.

2.1.4 Cell harvesting and freeze preparation of cultured cells

Cells were cultured to an approximately 80% confluence prior to freezing. The culture medium was removed and discarded using a sterile pipette. Subsequently, the cells were washed with 5 ml 1xPBS (Invitrogen; 14190-094) to remove all traces of FBS. 1 ml of 1x trypsin-EDTA (Sigma-Aldrich; T3924) was added to the T75 cm^2 flask and incubated for 1 minute. A further 5 ml of complete medium was added to the cell culture to stop trypsinuzation. The cells were counted using a hemocytometer and then centrifuged for 5 minutes at 1,500 xg to obtain a cell pellet. The supernatant was removed from the centrifuged cells and the pellets were re-suspended at a density of at least 3×10^6 cells/ ml in fresh chilled freezing media containing 10% dimethyl sulfoxide (DMSO) (Sigma-Aldrich; D8418) and 90% FBS. The cells were then transferred into a cryotube and stored overnight at -80°C and then the frozen cells were transferred to the liquid nitrogen tank for long-term storage. The viability and recovery of frozen cells were checked 24 hours after storing cryovials in liquid nitrogen.

2.2 Cell counting using hemocytometer

The cells were trypsinized and pipetted up and down for several times to resuspend the cells. Hemocytometer and coverslips were cleaned by rinsing with 70% ethanol followed by deionized water. A 10 µl cell culture was mixed thoroughly with an equal amount of 0.4% Trypan blue (Invitrogen; 15250-061) in a clean Eppendorf tube to determine cell viability. The tube was allowed to stand for 5 minutes at room temperature before loading. The cover slip was placed squarely above the grid of the hemocytometer. 10 µl of the cells and Trypan blue mixture was loaded to each side of the grid (chamber). The average number of cells was

counted on 25 squares for both the girds under the light microscope 10× objectives. The blue cells are dead ones, and the unstained cells are alive. The total number of viable cells was determined per millilitre by using the following calculation:
Cells/ ml = Average count of unstained cells × dilution factor × 10^4.

2.3 Total RNA isolation from cultured cells

Total RNA from the cell lines grown in the lab (Table 2.1) were extracted. Confluent cells at a density of 5×10^6 cells were collected in 1 ml TRIzol reagent (Invitrogen; 15596-026) and incubated for 5 minutes at room temperature. A volume of 200 µl chloroform was added to the TRIzol and the samples were shaken vigorously by vortex for 15 seconds and then incubated for 5 minutes at room temperature. The samples were spun at 12,000 xg for 15 minutes at 4°C. The colourless upper aqueous phase was transferred into a clean 1.5 ml Eppendorf tube. RNA precipitation was carried out by adding 200 µl of 100% isopropanol, followed by incubation for 10 minutes at room temperature. The samples were then centrifuged at 12,000 xg for 20 minutes at 4°C. The RNA pellet was washed using 1 ml of 75% ethanol in DEPC treated water and centrifuged again at 7,500 xg for 5 minutes at 4°C. The supernatant was discarded and the pellet was dried. The total RNA pellet was dissolved in 100 µl DEPC treated water and incubated in heating block at 56°C for 10 minutes. Finally, the RNA sample was treated with 1 µl DNase I (Sigma-Aldrich; D5319), and incubated at 37°C for 10 minutes and inactivated at 75°C for 10 minutes. The RNA concentration and quality of each sample was determined via a NanoDrop (ND_1000) spectrophotometer and agarose gel electrophoresis.

2.4 cDNA synthesis

Total RNA from the normal human tissue panel (Clontech; 636643) and many cell lines and tumour tissues were purchased from Clontech and Ambion. cDNA synthesis was performed according to the instructions of the manual of the SuperScript III First Strand synthesis system for RT-PCR (Invitrogen; 18080-051). 1 µg of the total RNA was transcribed into single-strand cDNA using oligo-dT primer that specifically binds to the poly-A tail of mRNAs. The cDNA was subsequently diluted 1:8 in sterile distilled water and store at -20°C. Primers for human *β ACT* were used in PCR reaction as a control for the quality of the cDNA samples.

2.5 Reverse transcription PCR (RT-PCR) and agarose gel electrophoresis

RT-PCR begins with reverse transcription of RNA into cDNA, and is followed by PCR amplification of the cDNA. A volume of 2 µl diluted cDNA was used in the PCR with 1 µl of each primer (10 pmol) and 25 µl of BioMix Red (BioLine; BIO-25006) and added to a final volume of 50 µl with distilled water. Samples were amplified with a pre-denaturation hold at 96°C for 5 minutes, followed by 40 cycles of denaturing at 96°C for 30 seconds, annealing temperature as described in Table 2.2 for 30 seconds and extension at 72°C for 30 seconds/kb, followed by a final extension step at 72°C for 5 minutes. PCR products were run on 1% agarose gels using 1x TBE buffer and 0.5 µg/ml ethidium bromide (Sigma-Aldrich; 46067) was added to the gel for later visualisation of the DNA. 5 µl of 100 bp DNA marker (NEB; N0467) was loaded for size determination of PCR products.

2.5.1 Primer design for RT-PCR

The specific sequence of all genes was found in the databases of National Center for Biotechnology Information (NCBI; http://www.ncbi.nlm.nih.gov/) and all primers were carefully designed to span at least one intron where possible. The Primer3 software (available from: http://www.genome.wi.mit.edu/cgi-bin/primer/primer3www.cgi) was used for design the primers. All primers for this study were synthesised from Eurofins MWG operon (http://www.eurofinsgenomics.eu/) and diluted with sterile distilled water to final concentration of 10 pmol. The primers used for RT-PCR screening and their expected size for each gene are listed in Table 2.2.

Table 2.2 Primer sequences for RT-PCR and their expected amplicon size

Human gene	Primer name	Primer sequence	Ta*	Expected amplicon size
β ACT	ACTB F1 ACTB R1	5'-AGAAAATCTGGCACCACACC-3' 5'-AGGAAGGAAGGCTGGAAGAG-3'	58.4°C	553 bp
BTG4	BTG4 F1 BTG4 R1	5'-AGTCACTGGCACTCTGATTG-3' 5'-GAATGACACGAGGTTCCTTG-3'	58.4°C	349 bp
CCNA1	CCNA1 F1 CCNA1 R1	5'-CAGCTGGAAAGAAAGCACTC-3' 5'-GAGAAACTGGTTGGTGGTTG-3'	58.4°C	698 bp
C2orf69	C2orf69 F1 C2orf69 R1	5'-GCATCACGTCCTCTATTTCC-3' 5'-CCAAGTATTGCTTCCTCCAG-3'	58.4°C	631 bp
C5orf50	C5orf50 F1 C5orf50 R1	5'-CATTGCTGATCTTGGTGCTG-3' 5'-GCACCTCTTCCTCTAACTTC-3'	58.4°C	219 bp
C11orf65	C11orf65 F1 C11orf65 R1	5'-AGAGCTTCTAGATGCTGCTG-3' 5'-GATGTGTTGACTTAGCCTCC-3'	58.4°C	422 bp
C11orf70	C11orf70 F1 C11orf70 R1	5'-CCAGACCTTTCAGTCCTATCG-3' 5'-GTCTCCCACACCATAACAGTG-3'	58.4°C	643 bp
C20orf201	C20orf201 F1 C20orf201 R1 C20orf201 F2 C20orf201 R2	5'-ATCTGCTCTTCGGCGACCTG-3' 5'-ACACTCTCAGTCGCCGTCAC-3' 5'-ATCTGCTCTTCGGCGACC-3' 5'-GACTTCTGCCGCTGGATG-3'	60°C 58.4°C 	505 bp (variant 1) 256 bp (variant 2)
C20orf195	C20orf195 F1 C20orf195 R1	5'-CCTGGACAAGATGAAGCTG-3' 5'-GTCAAACACGACAGGCATC-3'	58.4°C	613 bp
COX7B2	COX7B2 F1 COX7B2 R1	5'-TCTGCAAAGCATGGCAAGAC-3' 5'-ACAGGGGATAGGTTCCATTC-3'	58.4°C	156 bp
CST8	CST8 F1 CST8 R1	5'-CAATGCCTCAAATGCCAACG-3' 5'-TCCTACCAAAAAGCTGCAGC-3'	58.4°C	259 bp
DAZL	DAZL F1 DAZL R1	5'-ACTGATCGAACTGGTGTGTC-3' 5'-GGTGGAGTAGCTTCATGAAC-3'	58.4°C	497 bp
DMC1	DMC1 F1 DMC1 R1	5'-GAACCAGGATTCTTGACTGC-3' 5'-TGGAGTCGTGACAACATCTG-3'	58.4°C	518 bp
FABP9	FABP9 F1 FABP9 R1	5'-GAACATGGCAGGGTTAGTGA-3' 5'-TGGTGCTCTTTACTTTCCGG-3'	58.4°C	161 bp
FSCN3	FSCN3 F1 FSCN3 R1	5'-GTGCTTTCTACTGCGTTTCC-3' 5'-GTTTAAGCGCTCAGAAGCAG-3'	58.4°C	598 bp
GAGE1	GAGE1 F1 GAGE1 R1	5'-TAGACCAAGGCGCTATGTAC-3' 5'-CATCAGGACCATCTTCACAC-3'	58.4°C	245 bp
HORMAD1	HORMAD1 F1 HORMAD1 R1	5'-CCAGAATGCGCTTATGGAAC-3' 5'-CCATTCGTTCTCTCTCAGTG-3'	58.4°C	556 bp
KCNU1	KCNU1 F1 KCNU1 R1	5'-AGCCAAGACATCCTTAGGAC-3' 5'-GTTAGGAAGGTACACAAGCC-3'	58.4°C	642 bp
LYRM1	LYRM1 F1 LYRM1 R1	5'-AGCATTTTCAGGCTTGCGAG-3' 5'-CTCAGTTTCTCTTGGCTTCG-3'	58.4°C	278 bp
MAGEA1	MAGEA1 F1 MAGEA1 R1	5'-CCCACTACCATCAACTTCAC-3' 5'-CTCTTGCACTGACCTTGATC-3'	58.4°C	676 bp
NOL4	NOL4 F1 NOL4 R1	5'-GATGGGAGACTCCAACAGTG-3' 5'-CCTGCAGGACTTGAGGTAAG-3'	58.4°C	547 bp
NUT	NUT F1 NUT R1	5'-CACCACCAGTTGCTCAACTG-3' 5'-CTCCTTCACAGCTTCTGGTG-3'	60°C	623 bp
ODF3	ODF3 F1 ODF3 R1	5'-CCTGATTCCACCAACAACAG-3' 5'-CAGAGTGTTTGATGCCGAAG-3'	58.4°C	641 bp
PFN3	PFN3 F1 PFN3 R1	5'-AACAGCTGCGTGTGGGCTTC-3' 5'-CGTGCACCGTCTTGTTGAGG-3'	58.4°C	298 bp
PRDM9	PRDM9 F1 PRDM9 R1	5'-CAGGCTCAGAAACCAGTGTC-3' 5'-GTTCCTGGCCGTATTCATCC-3'	60°C	655 bp
PRSS45	PRSS45 F1 PRSS45 R1	5'-GCACCAAAGAGTACTCAGTG-3' 5'-CAACTTCACAAGCCAATGGC-3'	58.4°C	375 bp

Ta* - Annealing temperature

RAD21L	RAD21L F1	5'-TACTTCCTTTGCGGACACAC-3'	58.4°C	655 bp
	RAD21L R1	5'-AGCCAGCTGTTTCTTTAGGAC-3'		
REC8	REC8 F1	5'-GAGAGTTGAAGAGATCCCTC-3'	58.4°C	480 bp
	REC8 R1	5'-CTGGTTTGCAGTTGTTCCTG-3'		
SLC25A41	SLC25A41 F1	5'-GGAAGTGGATAACAAGGAGG-3'	58.4°C	550 bp
	SLC25A41 R1	5'-CTGGAGCATCTCATAGACAG-3'		
SMC1β	SMC1β F1	5'-TCAAGAAATCGAGGCCCACC-3'	60°C	368 bp
	SMC1β R1	5'-CTGGGGCCACACAGTTATAG-3'		
SSX2	SSX2 F1	5'-CAGAGAAGATCCAAAAGGCC-3'	58.4°C	407 bp
	SSX2 R1	5'-CTCGTGAATCTTCTCAGAGG-3'		
SSX5	SSX5 F1	5'-GAGGCCATGACTAAACTAGG-3'	58.4°C	253 bp
	SSX5 R1	5'-GCAGCTGTTTCCCATTGTTC-3'		
STAG3	STAG3 F1	5'-CTCTTCCATCAGGACAAGCAG-3'	60°C	495 bp
	STAG3 R1	5'-CTCTTCTTCCTCGTCCTCTTC-3'		
STRA8	STRA8 F1	5'-TGGCAGGTTCTGAATAAGGC-3'	58.4°C	510 bp
	STRA8 R1	5'-GAAGCTTGCCACATCAAAGG-3'		
	STRA8 F2	5'-GGCAAGAGGAATCACAATCC-3'	58.4°C	306 bp
SYCP1	SYCP1 F1	5'-GGTACAGCAGAAAGCAAGCAAC-3'	60°C	645 bp
	SYCP1 R1	5'-GGCAGATGTCCACAGATAGTC-3'		
SYCP2	SYCP2 F1	5'-CTTGGGAGACCTGGCAAAATG-3'	60°C	354 bp
	SYCP2 R1	5'-GATGAAGCCTCTGTTGTTCGC-3'		
SYCP3	SYCP3 F1	5'-GTCTTCTGCAGGAGTAGTTG-3'	58.4°C	509 bp
	SYCP3 R1	5'-CACTTGCTATCTCTTGCTGC-3'		
TEX19	TEX19 F1	5'-GCTTCAACATGGAGATCAGC-3'	58.4°C	386 bp
	TEX19 R1	5'-GAAGCTCCTCAAATCTCCAG-3'		
TEKT5	TEKT5 F1	5'-CATTGGCTTCTGGAAGTCAG-3'	58.4°C	543 bp
	TEKT5 R1	5'-GACAAGGTCTCAAAGAGGTG-3'		
TEPP	TEPP F1	5'-CTGCTGTCCATAATAAGGGC-3'	58.4°C	422 bp
	TEPP R1	5'-GTGATGTCCAAACACTGCAG-3'		
TMEM225	TMEM225 F1	5'-AGCCAAGATGAACCACAGTC-3'	58.4°C	375 bp
	TMEM225 R1	5'-CTGGTAGACAACTTGCACTC-3'		
TULP2	TULP2 F1	5'-GCAGAATTGGAGGAAGTCTC-3'	58.4°C	680 bp
	TULP2 R1	5'-GTGAACTTGGTGCTGAAGAC-3'		
UBL4B	UBL4B F1	5'-GGATGACAAGCACCTCTCTG-3'	58.4°C	350 bp
	UBL4B R1	5'-CCTTCTCCTCCATGTCACAG-3'		
WDR27	WDR27 F1	5'-GACCAACATCAAGAGTGAGG-3'	58.4°C	579 bp
	WDR27 R1	5'-GCCGATAAACTGGTCATGTC-3'		
ZNF558	ZNF558 F1	5'-CCCAGTTGGAACAAGACAAG-3'	58.4°C	579 bp
	ZNF558 R1	5'-GAATGCTATTGTGCCCAGTG-3'		

Ta* - Annealing temperature

2.5.2 Gel purification of RT-PCR product for sequencing

The gel purification was used to isolate the desired fragment. PCR products were loaded onto a 1% agarose gel prepared in 1x TBE buffer and 0.5 µg/ml ethidium bromide and the desired bands were cut out of from the gel using a sterile scalpel blade. The gel slice were then purified using the High Pure PCR Product Purification Kit (Roche Applied Science; 11732676001), according to the manufacturer's recommendations. Finally, the purified PCR fragment was eluted in elution buffer. The DNA concentration was determined by a NanoDrop (ND_1000).

The DNA amounts of 5 ng/µl in a total volume of 15 µl was put in a clean Eppendorf tube and in other tubes, 2 pmol/µl in a minimum volume 15 µl of the forward and/or reverse primers were transferred. The tubes were then sent at room temperature to Eurofins MWG for DNA sequencing. The result sequencing of each product was subjected to blast analysis and aligned against the expected sequence of PCR product to compare a query sequence with the NCBI databases of sequences.

2.6 Real time quantitative PCR (qRT-PCR)

2.6.1 Total RNA isolation for qRT-PCR

Total RNAs were isolated from confluence of cell cultures by means of the RNeasy Plus Mini Kit (Qiagen; 74134). The isolation process was done according to the manufacturer's recommendation. The concentration and quality of RNA was determined using a NanoDrop (ND_1000) and gel electrophoresis. 1 µg of total RNA was reverse transcribed into cDNA using the SuperScript III First Strand synthesis system for RT-PCR (Invitrogen; 18080-051). The reverse transcription was done according to the manufacturer's instructions.

2.6.2 Primer design for qRT-PCR

For efficient amplification in quantitative RT-PCR, all primers were designed manually at an amplicon size of < 160 bp. The length of each primer was 17-20 nucleotides, containing 40-60% G/C, avoiding predicted internal secondary structure. The forward and reverse primers had no significant complementarity to each other at the 3´ends to avoid primer-dimer forming and had similar melting temperature. Primers were checked with blast search (http://blast.ncbi.nlm.nih.gov/Blast.cgi). All primers for the genes of interest were synthesised by Eurofins MWG (http://www.eurofinsgenomics.eu/) (see Table 2.3), while the primers for

housekeeping genes used in qRT-PCR were made by Qiagen (see Table 2.4). Stock primers were diluted with sterile distilled water to make a final concentration of 10 pmol.

2.6.3 PCR setup for qRT-PCR

To set up the real time PCR reactions, the QuantiTect SYBR Green PCR Kit (Qiagen; 204054) was used, according to the manufacturer's protocol. In a 96 well plate, 1.5 µl (containing ~100 ng/µl cDNA) cDNA and 2.5 µl from each primer were used in a total of 25 µl. Samples were replicated three times and amplified with a pre-denaturation step at 95°C for 5 minutes, followed by 40 cycles of a denaturation step at 95°C for 10 seconds, a primer annealing step at 60°C for 30 seconds and an extension step at 95°C for 10 seconds. Melt curve analysis was achieved after completion of the 40 cycles.

Two negative controls were used to check for contamination. *β ACT* and *LAMIN A* were used as positive controls for normalisation of the qRT-PCR results. Real time PCR was carried out by using a BioRad CFX machine and analysis of result was achieved via the BioRad CFX Manager Software (version 2).

Table 2.3 Primer sequences for real time RT-PCR and their expected amplicon size

Human gene	Primer name	Primer sequence	Ta*	Amplicon size
STRA8	STRA8 F3	5'-GTTCCGAGTACATCTAGCTC-3'	60°C	124 bp
	STRA8 R3	5'-GAGTGTTTGCACAGCTAAGG-3'		
C20orf201	C20orf201 F3	5'-GGCTCCTCGGCGTCCTCAAG-3'	60°C	152 bp
	C20orf201 R3	5'-CGGCAGGAACAGCTCCAGCC-3'		
	C20orf201 F4	5'-ATGCCGCCGCCAGCGAA-3'	60°C	126 bp
	C20orf201 R4	5'-CGGCTTGGGGGGCTCTGT-3'		

Ta* - Annealing temperature

Table 2.4 Positive primer assay for real time RT-PCR

Human gene	Assay name	Source
β ACT	Hs_ACTB_1_SG	Qiagen; QT00095431
LAMIN A	Hs_LMNA_2_SG	Qiagen; QT01678495

2.7 Western blot analysis

2.7.1 Preparation of whole cell lysate from cell culture (Technique A)

Confluent cells were washed twice with 1x cold PBS and harvested with trypsin and then the cells were counted in a complete medium by a haemocytometer to determine the total number of cells. 5×10^5 cells were transferred to a new Eppendorf tube and centrifuged at 5,000 xg for 5 minutes at 4°C. The supernatant was discarded and pellet was washed again with 1x cold PBS. After centrifugation, the cell pellet was resuspend in lysis buffer (50 mM Tris-HCl pH 7.4, 200 mM NaCl, 0.5% Triton X-100, 1 mM AEBSF [4-(2-aminoethyl)-benzenesulfonyl fluoride] (Sigma-Aldrich; A8456) with one complete, mini, EDTA-free protease inhibitor cocktail tablet (Roche Applied Science; 11836170001) per 10 ml of lysis buffer and an equal volume of 2x Laemmli buffer (Sigma-Aldrich; S3401). The protein lysates were boiled at 95°C for 5-10 minutes. Lysate of approximately 5×10^5 cells was loaded per well.

2.7.2 Preparation of whole cell lysate from cell culture (Technique B)

Whole cell protein lysates were isolated from 5×10^5 cells using radio immunoprecipitation assay (RIPA) buffer (Sigma-Aldrich; R0278) with 1 mM AEBSF [4-(2-aminoethyl)-benzenesulfonyl fluoride] and one complete, mini, EDTA-free protease inhibitor cocktail tablet was added to the 10 ml of RIPA buffer. Following, an equal volume of 2x Laemmli buffer was added to the lysates. The lysates were mixed gently by pipetting, and then incubated on ice for 30 minutes. The samples were gently vortexed and boiled at 95°C for 5-10 minutes, followed by transferring the supernatant to a clean Eppendorf tube and discard the cell debris. Lysate of approximately 5×10^5 cells was loaded per well.

2.7.3 Preparation of whole cell lysate from cell culture (Technique C)

Cells were washed with 1x cold PBS, harvested and pelleted by centrifugation at 5,000 xg for 5 minutes at 4°C. The dry pelleted cells were weighted and re-suspended in a volume of 10 µl of M-PER Mammalian Extraction Reagent (Thermo; 78503) for every 1 mg of dry cell pellet. The cell lysate tubes were incubated at room temperature for 10 minutes, followed by centrifugation for 15 minutes at 14,000 xg to pellet cell debris. The protein concentration of each sample was measured using the BCA protein assay kit (Thermo; 23227). Approximately 20 µg protein was used with 2 µl Sample Reducing agent 10x (Invitrogen; NP0004) and 4 µl LDS Sample Buffer (4x) (Invitrogen; NP0007). The samples were then mixed by pipetting and boiled at 70°C for 10 minutes, followed by chilling on ice before gel loading. Approximately 20 µg protein was loaded per well.

2.7.4 Source of human normal tissue lysates

A whole cell lysates from human normal tissues were purchased and detailed in the accompanying table.

Table 2.5 Source of normal lysates used in Western blotting

Whole cell lysates	Source	Catalogue number	Protein loading
Testis	Abcam	AB30257	20 µg
Testis	Novus Biological	NB820-59266	20 µg
Brain: Cerebellum	Abcam	AB30069	20 µg
Spinal cord	Abcam	AB29188	20 µg
Brain	Abcam	AB29466	20 µg
Salivary gland	Abcam	AB29159	20 µg
Skeletal muscle	Abcam	AB29331	20 µg
Thymus	Abcam	AB30146	20 µg
Small intestine	Abcam	AB29276	20 µg
Stomach	Abcam	AB29681	20 µg
Ovary	Abcam	AB30222	20 µg
Liver	Abcam	AB29889	20 µg

2.7.5 Cell lysis and fractionation

Before fractionation, cells were washed twice with cold 1xPBS, tripsinized and pelleted for 5 minutes at 1,500 xg using centrifugation. In the cytoplasmic fraction, the cell pellet was resuspend in hypotonic buffer (50 mM Tris-HCl pH 7.4, 0.1 M sucrose, 1 mM AEBSF with one Roche complete protease inhibitor cocktail tablet/10 ml) and an equal volume of lysis buffer C (1% Triton-X-100, 10 mM $MgCl_2$, 1 mM AEBSF, with one Roche complete protease inhibitor cocktail tablet/10 ml). The suspension were then incubated on ice for 30 minutes and centrifuged for 2 minutes at 6,000 xg at 4°C. The pooled supernatant that contained cytoplasmic protein was transferred to a clean chilled Eppendorf tube.

In contrast, in the nuclear fraction, the cell pellet was then re-suspended in lysis buffer N (50 mM Tris-HCl pH 7.4, 100 mM KAc, 1 mM AEBSF, one Roche complete protease inhibitor cocktail/10 ml). The protein determination for both fractions was carried out using the BCA protein assay according to the recommendations of the supplier. An equal volume of 2x Laemmli buffer was added to cytoplasmic and nuclear tubes and then incubated at 70°C for 10 minutes. 20 µg protein was loaded per well.

2.7.6 Western blotting technique

The protein lysate for each sample was loaded onto a NuPAGE® 4-12% Bis-Tris gel (Invitrogen; NP0322) in 1x MOPS SDS Running Buffer (Invitrogen; NP0001) and Precision plus protein oral standard (BioRad; 161-0377) was also run as molecular weight markers to determine the protein sizes of each sample. The gels were run at 200 V for 35 minutes and then transferred to an Immobilon-P PVDF membrane (Millipore; IPVH00010) which was activated by wetting the PVDF membrane in 100% methanol. The transfer step was carried out at 400 mA for 3 hour using 1x cold Towbin transfer buffer (192 mM Glycine, 25 mM Tris base, 10% methanol). After completing transfer, the membrane was washed in H_2O five times to remove all traces of transfer buffer. Membranes were blocked for one hour at room temperature with blocking solution containing 1xPBST (1xPBS + 0.5% Tween-20, Sigma-Aldrich; P1379) and 5% non-fat dry milk with gentle agitation followed by an overnight incubation on the shaker at 4°C with an appropriate dilution of primary antibody. Membranes were washed three times for 15 minutes using 1xPBST at RT with shaking. The secondary antibody was diluted in blocking solution and incubated with the membranes on the shaker at room temperature for 1 hour. The primary and secondary antibodies are detailed in Tables 2.6 and 2.7, respectively. The membranes were washed again three times for 15 minutes using 1xPBST at RT with shaking. ECL detection using SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific; 34080) or Chemiluminescent Peroxidase Substrate-3 (Sigma-Aldrich; CPS3100-1KT) were used in 1:1 dilution and incubated with membrane for 5 minutes at room temperature to detect the protein. The membranes were then exposed using X-Ray films (Thermo Scientific; 34091) within an appropriate amount of time.

Table 2.6 Primary antibodies used in Western blot (WB) and Immunofluorescence (IF)

Primary antibodies	Catalogue number	Source	Host	Clonality	Application	Optimal dilution
Anti-STRA8	AB130985	Abcam	Rabbit	Polyclonal	WB	1/1,000
Anti-C20orf201	AB108142 AB170783	Abcam	Rabbit	Polyclonal	WB IF	1/1,000 1/15
Anti-6x His tag	AB18184	Abcam	Mouse	Monoclonal	WB	1/1,000
Anti-Lamin A/C (636)	Sc-7292	Santa Cruz	Mouse	Monoclonal	WB	1/100
Anti-Lamin B	Sc-6217	Santa Cruz	Goat	Polyclonal	WB	1/1,000
Anti-α-Tubulin	T6074	Sigma	Mouse	Monoclonal	WB	1/5,000
Anti-GAPDH (G-9)	Sc-365062	Santa Cruz	Mouse	Monoclonal	WB	1/2,000

Table 2.7 Secondary antibodies used in Western blot (WB) and Immunofluorescence (IF)

Secondary antibodies	Catalogue number	Source	Application	Optimal dilution
Peroxidase-conjugated AffiniPure donkey anti-mouse IgG (H+L)	715-035-150	Jackson ImmunoResearch Laboratories Inc.	WB	1/20,000
Peroxidase-conjugated AffiniPure donkey anti-rabbit IgG (H+L)	715-035-152	Jackson ImmunoResearch Laboratories Inc.	WB	1/20,000
Anti-Goat IgG (whole molecule)-Peroxidase produced in rabbit	A5420	Sigma	WB	1/20,000
Alexa Fluor® 568 goat anti-rabbit IgG (H+L)	A11036	Invitrogen	IF	1/250

2.8 siRNA (small interfering RNA) Knockdown

Confluent cells were seeded at 3×10^5 cells per well in 6-well plates which contained fresh complete medium and then the cells were incubated under normal growth conditions. The transfection mixture for each well was prepared in a clean Eppendorf tube containing 100 µl of complete medium serum free, 12 µl of Hiperfect Transfection Reagent (Qiagen; 301705) and 1.2 µl of a 10 µM siRNA of the gene of interest or 1.2 µl of Negative Control siRNA. The siRNAs used in this study are summarised in Table 2.8. The mixtures were mixed and then incubated at room temperature for 20 minutes to allow the formation of transfection complexes. The complexes were added drop-wise onto the cells with gently shaking the plate and then the cells were incubated with the transfection complexes under their normal growth conditions. Untreated cells were used as controls at the same time of transfection to allow measurement of the expression level of the gene of interest.

The day of transfection was considered a day 0 and two 'hits' was used in this study for the siRNA knockdown without changing the complete medium. 48 hours post-transfection the cells were harvested and counted as per Sections 2.7.1 and 2.7.3, and prepared for the Western blotting procedure (refer to Section 2.7.6).

Table 2.8 siRNA used for gene knockdown

Human gene	siRNA name	Target sequence (5'to 3' direction)	Catalogue number	Source
STRA8	Hs_STRA8_1	CTCAAAGTGGCAGGTTCTGAA	Sl04251933	Qiagen
	Hs_STRA8_2	CAGGCTGTGGCAGCTTATAAT	Sl04317614	
	Hs_LOC346673_3	AAGCAGCTTAGAGGAGGTCAA	Sl00505778	
	Hs_LOC346673_5	AGGAGAAGTTTCAGCTCTATA	Sl03145618	
C20orf201	Hs_LOC198437_2	ACCGCCAAGAGGTGCAGACAA	Sl00485135	Qiagen
	Hs_LOC198437_5	CCCGTGGACGCAGTCGCTCGA	Sl03186386	
	Hs_LOC198437_6	CCAGCCTCCCACATAAAGTTA	Sl04258772	
	Hs_LOC198437_7	TCCCGCGGTGACGGCGACTGA	Sl04319574	
NI (non-interfering) siRNA	AllStars Negative Control siRNA	----------	1027280	Qiagen

2.9 Staining of cell culture using immunofluorescence (IF)

Cells were seeded at 5×10^4 cells per well in 24-well plates on glass coverslips coated with poly-D-lysine (VitroCam; 1445-P01) and incubated under normal growth condition until 70-80% confluence. The cells were washed twice with ice cold 1xPBS, then fixed in 4% paraformaldehyde (PFA) in 1xPBS pH 7.4 for 15 minutes at room temperature and washed gently twice with ice cold 1xPBS. The cells were permeabilized by adding 0.25% Triton X-100 diluted in 1xPBS and incubated for 15 minutes at room temperature. The cells were then washed carefully three times with 1xPBS for 15 minutes and incubated with 1% BSA in PBST/ 0.5% Tween-20 for 30 minutes to block unspecific binding of the antibodies. The cells were then incubated with the primary antibody (see Table 2.6) diluted in the blocking buffer for one hour at room temperature or overnight at 4°C, followed by three washes for 15 minutes using 1xPBS. The secondary antibody (see Table 2.7) was diluted in the blocking buffer and incubated in the dark for one hour at room temperature, followed by three washes for 15 minutes in the dark. The coverslips were removed carefully from each well and allowed to semi-air dry for about 10 minutes, before counterstaining and mounting with 7 µl of reagent MD (Cambio; 1124-MD-50) and incubated for 10 minutes at room temperature. The cover slips were placed onto glass slides and sealed with nail polish to prevent drying and movement under microscope. The slides were viewed by using a Zeiss Axioskop 2 fluorescence microscope, and the pictures were performed with an AxioCam digital camera.

2.10 The cloning of *STRA8* and *C20orf201* cDNAs genes

2.10.1 Primers design for cloning

The forward and reverse primers were designed in a 5'---3' orientation and contained a start codon (ATG) in the 5' primer and a stop codon (TTA) in the 3' primer. Restriction enzyme sites were added to the 5' end of the forward and reverse primers before the CDS sequence. A Kozak sequence was added to the 5' end of the forward primer before the initiation codon. A couple of bases were added to the 5' end of the primers to preserve restriction site during PCR. All primers for cloning were synthesised from Eurofins MWG operon and diluted to a concentration of 100 pmol as the manufacturer's instructions and recommended volumes, followed by dilutions, from 100 pmol to 10 pmol.

2.10.2 *STRA8* cloning primers

The primers used for *STRA8* cloning +Kozak sequence are detailed in Table 2.9

Table 2.9 Primer sequences for *STRA8* cloning and their expected amplicon size

Primer name	Primer sequence	Ta*	Amplicon size
STRA8+ Kozak F4	5'-GAGCTCGAGACCATGGGGAAGATTGATGTGG-3'	58.4°C	1212 bp
STRA8 R4	5'-GAGCTCGAGTTACAAATCTTCATCGTCAAAG-3'		
GAG – Extra bases	CTCGAG – *Xho*I recognition site	ACC – Kozak sequence	
ATG – Start codon	TTA – Stop codon		

Ta* - Annealing temperature

2.10.3 *C20orf201* cloning primers

The primers used for *C20orf201* cloning +/- Kozak sequence are detailed in Table 2.10

Table 2.10 Primer sequences for *C20orf201* cloning and their expected amplicon size

Primer name	Primer sequence	Ta*	Amplicon size
C20orf201+ Kozak F5	5'-GTGAAGCTTACCATGCCGCCGCCAGCGAA-3'	60°C	726 bp
C20orf201- Kozak F6	5'-GTGAAGCTTATGCCGCCGCCAGCGAA-3'	60°C	723 bp
C20orf201 R5	5'-GTGAAGCTTTTATAACTTTATGTGGGAGGCTG-3'		
GAG – Extra bases	AAGCTT – *Hin*dIII recognition site	ACC – Kozak sequence	
ATG – Start codon	TTA – Stop codon		

Ta* - Annealing temperature

2.10.4 Cloning primers for different fragments of *C20orf201*

His Taq sequence was added to each reverse primer before the stop codon for cloning of different fragments for *C20orf201* cDNAs. The primers are detailed in Table 2.11.

Table 2.11 Primer sequences for 7 fragments of *C20orf201* and their expected amplicon size

Fragment	Primer name	Primer sequence	Ta*	Amplicon size
	C20orf201 F7	5'-GTGAAGCTTACCATGCCGCCGCCAGCGAA-3'		
1	C20orf201 R7	5'-GTGAAGCTTTTAGTGGTGGTGGTGGTGGTG TAACTTTATGTGGGAGGCTG-3'	58.4°C	744 bp
2	C20orf201 R8	5'-GTGAAGCTTTTAGTGGTGGTGGTGGTGGTG CGCCTCCGCTGCCGG-3'	58.4°C	645 bp
3	C20orf201 R9	5'-GTGAAGCTTTTAGTGGTGGTGGTGGTGGTG TTGGCGGTCCAGGGGGTC-3'	58.4°C	546 bp
4	C20orf201 R10	5'-GTGAAGCTTTTAGTGGTGGTGGTGGTGGTG GAGCAGCGCGATCTC-3'	55°C	447 bp
5	C20orf201 R11	5'-GTGAAGCTTTTAGTGGTGGTGGTGGTGGTG GTTCTGGCGCTCGGC-3'	55°C	348 bp
6	C20orf201 R12	5'-GTGAAGCTTTTAGTGGTGGTGGTGGTGGTG CGCCGCGCCCACGTC-3'	58.4°C	249 bp
7	C20orf201 R13	5'-GTGAAGCTTTTAGTGGTGGTGGTGGTGGTG CGGCTTGGGGGGCTCTGT-3'	58.4°C	150 bp

GAG – Extra bases AAGCTT – *Hin*dIII recognition site ACC – Kozak sequence

ATG – Start codon TTA – Stop codon GTGGTGGTGGTGGTGGTG – His Taq sequence

Ta* - Annealing temperature

2.10.5 PCR Amplification using BioMix Red Master Mix

Specific primers were designed for *STRA8* (refer to Table 2.9) and *C20orf201* cloning (refer to Table 2.10). The PCR reactions were carried out in a total volume of 50 µl by using 25 µl of BioMix red Master Mix, 1 µl of each primer (10 pmol) (refer to Tables 2.9 and 2.10), 2 µl of pCMV6-AC::*C20orf201* vector (5 ng/µl) (Origene; SC321851) or 2 µl of pCMV6-XL5::*STRA8* vector (5 ng/µl) (Origene; SC307401) and 21 µl of water (Sigma-Aldrich; W3500). The PCR reaction complex was firstly predenatured at 96°C for 30 seconds, followed by 40 cycles of denaturation at 96°C for 30 seconds, annealing temperature depending on the primer site for each fragment for 30 seconds and extension at 72°C for 30 seconds. One cycle of final extension was set up at 72°C for 5 minutes.

2.10.6 Loading of PCR product onto agarose gel electrophoresis

After PCR completing, the 50 µl of PCR reactions were loaded onto 1x TBE agarose gel stained with ethidium bromide and run for an hour at 100 V. The right band was cut by a clean blade and purified by using a High Pure PCR purification kit, as per the manufacturer's instructions. 1 µl of purified PCR product with gel loading blue dye (6x) (BioLabs; B7021S) was run on the gel to ensure that the purification has worked. HyperLadder II (BioLine; Bio-33040) was used a marker to determine the size of the fragments.

2.10.7 PCR Amplification using Phusion High Fidelity PCR Master Mix

Specific primers were designed for different fragments of *C20orf201* (refer to Table 2.11) and the PCR amplification were carried out in a total volume of 50 µl by using 25 µl of Phusion® High-Fidelity PCR Master Mix (BioLabs; M0532S), 2.5 µl of each primer (10 pmol) (refer to Table 2.11), 2 µl of pCMV6-AC::*C20orf201* vector (5 ng/µl) and 18 µl of Sigma water. The PCR reaction complex was firstly predenatured at 98°C for 30 seconds, followed by 35 cycles of denaturation at 98°C for 10 seconds, annealing temperature depending on the primer site for each fragment for 30 seconds and extension at 72°C for 30 seconds. One cycle of final extension was set up at 72°C for 10 minutes. Following completion of PCR, the 50 µl of PCR reactions with gel loading dye were run on 1x TBE agarose gel (see Section 2.10.6).

2.10.8 Digestion of PCR inserts by restriction endonucleases

The restriction enzymes used for digestion of *STRA8* or *C20orf201* are listed in Table 2.12. The purified PCR product was digested in a sterile Eppendorf tube by mixing 2 µg of purified DNA, 5 µl of an appropriate 10x restriction enzyme buffer, 1 µl of the desired restriction enzyme (10 u/µl) and Sigma water was added to make up a volume of 50 µl. The mixture was then gently mixed by pipetting and incubated for 2 hours at 37°C.

Table 2.12 Restriction enzymes used for *STRA8* and *C20orf201* digestions

Gene	Restriction enzyme	Enzyme buffer	Recognition site	Catalogue No.	Company
STRA8	*Xho*I	10x Buffer D	5'…C▼TCGAG…3' 3'…GAGCT▲C…5'	R6161	Promega
C20orf201	*Hin*dIII	10x Buffer E	5'…A▼AGCTT…3' 3'…TTCGA▲A…5'	R6041	Promega

2.10.9 Generation of *STRA8* cDNA clone in pCMV6-XL5

The pCMV6-XL5 vector was digested by *Not*I and 10x buffer D was used. The optimal digestion procedure for this is to refer to Section 2.10.8

2.10.10 DNA isolation from agarose gel

All 50 µl was separated on 1x TBE agarose gel by electrophoresis. The DNA bands were cut out under long wave UV light. The DNA purification was carried out using a High Pure PCR purification kit according to the supplier's manual instructions.

2.10.11 Digestion of pcDNA5/FRT/TO plasmid with restriction enzymes

The mammalian expression vector, pcDNA5/FRT/TO was obtained as a gift from Dr. J. Müller (University of Warwick). This digestion was performed by mixing 2 µg of plasmid DNA, 5 µl of an appropriate 10x restriction enzyme buffer, 1 µl of the desired restriction enzyme (the same restriction enzymes and buffers that used in Table 2.12) and Sigma water was added to make up a volume of 50 µl. The mixture was then gently mixed by pipetting and incubated for 2 hours at 37°C. After this, 11 µl of Antartic phosphatase reaction buffer and 1 µl of Antartic phosphatase (BioLabs; M0289S) were added to the digestion mixture and incubated for 15 minutes at 37°C. Thereafter, heat inactivated for five minutes at 65°C. The total volume were split into two well agarose gel and the plasmid DNA was purified (refer to Section 2.10.10). HyperLadder I (BioLine; Bio-33026) was used to determine the size of the migrated fragments.

2.10.12 Calculation the insert and vector molar ratio for ligation

The concentration for insert DNA and plasmid were measured by using a NanoDrop (ND_1000) to determine how much digested insert and plasmid are required for ligation. A molar ratio of pcDNA5/FRT/TO vector was used to insert DNA. The amount of insert and vector needed was calculated using the following formula:

The amount of insert = [Amount of vector in ng × Length of insert in bp × molar ratio / length of vector in bp]

2.10.13 Ligation of digested PCR insert and plasmid

The ligation of the digested insert and vector was performed using the Quick Ligation Kit (BioLabs; M2200S) to anneal the insert of gene of interest into the vector. Around of 50 ng of vector was used in one ligation reaction and combined with a 3-fold molar excess of insert. The volume of DNA ligation was adjusted to 10 µl with Sigma water. 10 µl of 2x Quick Ligation Buffer was added and mixed, followed by the addition of 1 µl of Quick T4 DNA Ligase. The ligation mixtures were mixed, centrifuged briefly and incubated at room temperature for 5 minutes. Finally, the samples were chilled on ice for 1 minute and then transformed into *Escherichia coli*. To do a self-ligation test, a control reaction containing all the components above except the DNA insert was set up.

2.10.14 Preparation of LB and LB agar media for *E. coli* culture

The LB (Luria Bertani) and LB agar media components were detailed in the following table.

Table 2.13 Composition of LB and LB agar media

LB medium (1 L)		**LB agar medium (1 L)**	
Reagent	**Volume**	**Reagent**	**Volume**
Bacto®-Tryptone; 211705	10 g/L	Bacto®-Tryptone	10 g/L
Bacto®-Yeast Extract; 212750	5 g/L	Bacto®-Yeast Extract	5 g/L
NaCl	10 g/L	NaCl	10 g/L
		Bacto®-Agar; 214030	15 g/L

The above reagents (Table 2.13) for each medium was added to a sterile bottle contained 900 ml of distilled water and stirred. Distilled water was added up to one litter, followed by autoclaving for 20 minutes at 120°C on the same day they were made up. The media bottles were cooled and Ampicillin antibiotic (Sigma-Aldrich; A9518) was added to make a final concentration of 100 µg/ml. 25 ml of the LB/ amp/ agar medium was poured onto each labelled Petri dish. The dishes were then left to solidify at room temperature in clean bench. The LB media and plates were stored at 4°C.

2.10.15 Transformation of recombination plasmid into competent *E. coli* cells

The vials of DH5α competent *E. coli* (BioLabs; C2987) were removed from -80°C and placed on ice for 5 minutes to thaw. 5 µl of the ligation mixture (refer to Section 2.10.13) was transferred to a chilled Eppendorf tube. The thawed competent cells were mixed by flicking

tube and 25 µl of the cells were taken to the previous ligation Eppendorf tube. The cells were then mixed with the DNA by pipetting up and down or flicking the tubes 4-5 times. The mixture was immediately placed on ice for 30 minutes, followed by heating shock the cells at 42°C for 30 seconds and then the tubes were immediately placed on ice for 5 minutes. 270 µl of prewarmed (37°C) SOC outgrowth medium (BioLabs; B9020S) was added to each transformation reaction and then the tubes were incubated at 37°C for 60 minutes with vigorous shaking (approximately 250 xg). The Petri dishes plates (refer to Section 2.10.14) were warmed at 37°C for 30 minutes and 50-100 µl of the transformation reactions were spread out onto the required plates using a sterile spreader sealed in a flame. The plates were stored at room temperature until the liquid had been absorbed. The plates were then inverted and incubated overnight (12-16 hours) at 37°C. The tube for testing *E. coli* efficiency was carried out by taking 1 µl from undigested pcDNA5/FRT/TO (20 ng/µl) and then transformation was performed.

2.10.16 PCR colony screening

One colony was picked from the overnight plates by using a micropipette tip and rinsing it into a new Eppendorf tube containing 20 µl of LB medium with Ampicillin. PCR screening was directly carried out using the following reaction: 2 µl was taken from the LB/ colony mixture and transfer to PCR tube containing 12.5 µl of BioMix red, 0.5 µl of each primer and 9.5 µl of distilled water. The primer sequences are listed in Table 2.14. The PCR program is described in Section 2.10.5. 15 µl of PCR mixture was run on 1x TBE agarose gel to determine the presence of the gene of interest.

Table 2.14 Primers used in PCR colony screening and their amplicon size

Human gene	Primer name	Primer sequence	Ta*	Amplicon size
STRA8	STRA8 F1 STRA8 R1	5'-TGGCAGGTTCTGAATAAGGC-3' 5'-GAAGCTTGCCACATCAAAGG-3'	58.4°C	726 bp
C20orf201	C20orf201 F1 C20orf201 R1	5'-ATCTGCCTCTTCGGCGACCTG-3' 5'-ACACTCTCAGTCGCCGTCAC-3'	60°C	505 bp

Ta* - Annealing temperature

2.10.17 Preparation of *E. coli* glycerol stock

As a result of positive PCR screening, the remaining LB/ colony mixture (18 µl) was transferred to a sterile 50 ml Falcon tube containing 10 ml of LB medium containing the appropriate selective antibiotic. The tube was grown overnight at 37°C with vigorous shaking.

Next day, a glycerol stock of plasmid containing *E. coli* was prepared in a new cryovial by taking 1 ml of overnight culture of the bacteria plus 200 µl of 100% a sterile glycerol and mixed well by vortex. The plasmid name, cloned ID, cloning site, date and strain of bacteria were recorded on the vial and stored at -80°C freezer for long-term storage. The miniprep was prepared with 9 ml rest of bacteria culture (refer to Section 2.10.19). The purified DNA was stored at -20°C freezer.

2.10.18 Streaking out from a glycerol stock

To recover bacteria from the glycerol stock, the cryovial was taken from -80°C freezer and opened on a clean bench. A sterile loop was used to scrape off a portion from the top of the frozen glycerol stock. The vial was immediately returned to the -80°C freezer, while the bacteria streak was rinsed onto an LB agar plate containing selective marker (as required) and grown overnight at 37°C. On the next day, one single colony was picked up and incubated into 10 ml of LB medium, which contained appropriate selective antibiotic. It was incubated overnight at 37°C with vigorous shaking.

2.10.19 Purification of plasmid DNA from *E. coli*

The bacterial cells were harvested by centrifugation for 15 minutes at 6,000 xg at 4°C. The supernatant was discarded and the plasmid DNA was isolated from the bacterial pellet using QIAprep Spin Miniprep Kit (Qiagen; 27106) as per the manufacturer's instructions. 1 µl of purified plasmid DNA was run onto 1x TBE agarose gel to check DNA.

2.10.20 Confirming the orientation of *C20orf201*

After cloning with *C20orf201* fragment 1 and 2, the pcDNA5/FRT/TO was further analysed using the *Xho*I restriction enzyme to determine the right orientation of each fragment. 2 µg of plasmid is digested by the required restriction enzyme and then the digestion procedure was described in Section 2.10.8.

2.10.21 Confirming the DNA sequencing of inserts

The plasmid DNA amount of 80 ng/µl in a total volume of 15 µl was put in a clean Eppendorf tube. The tubes were then sent at room temperature to Eurofins MWG for DNA sequencing. The plasmid was sequenced with the CMV forward and BGH reverse primers to confirm that the gene of interest was cloned in the correct orientation for expression. The primer sequences

are listed in Table 2.15. The result of the correct sequencing of inserts was blasted and aligned against the expected sequence.

Table 2.15 Primers used for checking DNA sequencing of inserts

Primer name	Primer sequence	Concentration
CMV forward	5'-CGCAAATGGGCGGTAGGCGTG-3'	10 pmol
BGH reverse	5'-TAGAAGGCACAGTCGAGG-3'	10 pmol

2.11 Flp-In T-REx-293 cell line

2.11.1 Growth and maintenance of Flp-In T-REx-293 cell line

The cell lines were maintained in the complete DMEM, which contained 10% heat inactivated FBS (refer to Section 2.1.3 for cells thawing). The thawed cells were grown in T75 cm^2 and incubated in humidified incubator at 37°C and 5% CO_2 for 16-24 hours, and then the media were aspirated off and replaced with 10 ml of fresh, complete DMEM containing 10 µg/ml of blasticidin S (InvivoGen; ant-bl-1) and 200 µg/ml zeocin (Invitrogen; R25001). The cells were incubated at 37°C and checked daily until the cells were 80-90% confluent. The confluent cells were routinely split 1:2 in T75 cm^2 flasks as follows. After removal of the growth medium, cells were washed once with 5 ml 1xPBS to remove excess medium and serum. 1 ml of trypsin EDTA solution was then added to the flask and placed at 37°C for about 3 minutes. The cells were checked under a microscope to confirm that most of the cells had detached. After the cells were detached, 5 ml pre-warmed culture medium was added to the T75 cm^2 flask, and the suspension was briefly pipetted up and down to break up clumps of cells. The cells then split into new flasks containing zeocin and blasticidin. The cells were frozen as soon as possible to keep stocks of the early passage number (refer to Section 2.1.4 for cells freezing).

2.11.2 Preparation of LyoVec transfection

The cells were transfected with LyoVec (InvivoGen; lyec-12) to generate stable inducible cell lines. The LyoVec was prepared by adding 2 ml of deionized sterile water per vial and homogenized by vortexing gently for 30 seconds. The vial was placed for at least 30 minutes before starting transfection or stored at 4°C until required.

2.11.3 Preparation of transfection and DNA complexes

The LyoVec transfection reagent was incubated at room temperature and gently vortexed to homogenize before used. 1 µg pcDNA5/FRT/TO::*STRA8* or pcDNA5/FRT/TO::*C20orf201* plasmid DNA plus 10 µg of pOG44 at a ratio of 1:9, respectively were taken to a sterile 1.5 ml Eppendorf tube containing 400 µl of LyoVec and then mixed gently. pOG44 plasmid was obtained from Dr. J. Müller (University of Warwick). pcDNA5/FRT/TO construct alone was prepared as a negative control. The LyoVec- DNA complexes were incubated at room temperature for at least 15 minutes to allow the formation of the complex. The LyoVec™-DNA complexes were used immediately for transfection of cells.

2.11.4 Transfection of Flp-In T-REx-293

The Flp-In T-REx-293 cells were seeded at 2×10^6 cells culture medium containing serum in 10 cm dishes and incubated with complete DMEM supplemented with 200 µg/ml zeocin and 10 µg/ml blasticidin medium for 24 hours. Next day, the LyoVec and DNA complexes were prepared (refer to Section 2.11.3). The medium was removed from cells and replaced with fresh complete DMEM containing blasticidin, but without zeocin. The LyoVec -DNA complexes were added directly to the medium and the dish was swirled to distribute the complex. 24 hours after transfection the media were removed and replaced with complete DMEM + blasticidin. 48 hours post-transfection the cells were split directly 1:4 into medium supplemented with 100 µg/ml of hygromycin B (Invitrogen; 10687-010). The cells were split at density of less 25% confluent to increase antibiotic efficiency. The complete DMEM + blasticidin + hygromycin was refreshed every 2-3 days until visible foci appeared.

2.11.5 Isolation of individual colonies of transfected cells

After 7-10 days adding hygromycin to cell culture medium, few hygromycin resistance foci appeared. At this state, the medium was removed from the dish and cells were washed once with 5 ml 1xPBS. The cylinder (Sigma-Aldrich; C1059) was placed over a colony of cells by using a sterile curved forceps to pick up individual colony. The cylinder was pressed gently to seal the space between the cylinder and the dish. The cylinder was filled with the appropriate amount of trypsin and incubated at 37°C for 2 minutes. The cells were then checked for detachment under a microscope and growth media were added to resuspend the cells, and then the cells were transferred to a 6-well plate containing 2 ml of growth media supplemented

with 100 µg/ml hygromycin. The cells from one well of 6-well plates was transferred to 10 cm dish and grown to confluence. Multiple vials were frozen.

2.11.6 Isolation of genomic DNA from individual colony

2×10^6 cells were used for genomic DNA isolation. This was started by removing the medium and washed the cells once with 1xPBS. Cells were trypsinized and the cells were harvested in a new 1.5 ml Eppendorf tube. The cells were then centrifuged at 1,500 xg for 5 minutes to pellet the cells. The supernatant was removed and 500 µl 1xPBS was added to the cells pellet to wash the cells. The cells were spun again as previous, followed by removing the PBS and the pellet was kept. Wizard® SV Genomic DNA Purification System (Promega; A2361) was used to isolate the genomic DNA as per the manufacturer's instructions. The genomic DNA was eluted in 100 µl nuclease free water. 1 µl of RNase (Promega; A797C) was added to the eluted genomic DNA to remove co-purified RNA, followed by incubation at room temperature for 15 minutes. The quality of DNA samples were tested by PCR using *β ACT* primers.

2.11.7 PCR screening of different integration in Flp-In T-REx-293

To confirm that the gene of interest, *STRA8* or *C20orf201,* was integrated correctly into the genome, a PCR was carried out using primers that are listed in Table 2.16. The PCR reactions and program were carried out as the same as in RT-PCR reaction (refer to Section 2.5) but with one difference, 2 µl of genomic DNA was used instead of cDNA.

Table 2.16 Primer sequences for PCR and their expected amplicon size

Primer name	Primer sequence	Ta*	Amplicon size
PSV40 F1	5'-CAGTTAGGGTGTGGAAAGTC-3'	58.4°C	805 bp
Hygromycin R1	5'- GCCATGTAGTGTATTGACCG-3'		
STRA8 F5	5'-CTTTGATGTGGCAAGCTTCC-3'	58.4°C	900 bp
LacZ-Zeocin R1	5'-GCGATTAAGTTGGGTAACGC-3'		
C20orf201 F1	5'-ATCTGCCTCTTCGGCGACCTG-3'	58.4°C	1161 bp
LacZ-Zeocin R1	5'-GCGATTAAGTTGGGTAACGC-3'		
C20orf201 F7	5'-GTGAAGCTTACCATGCCGCCGCCAGCGAA-3'	58.4°C	1400 bp

Ta* - Annealing temperature

2.11.8 *LacZ* staining of Flp-In T-REx-293 cells

Staining of cell lines used in this study was performed in 60 mm plates. Cells were seeded at 3×10^6 cells and incubated under growth condition for 24 hours. On the next day, the growth

medium was removed from the cells and rinsed once with 2 ml 1xPBS. The cells were then fixed and stained using the β-Gal Staining Kit (Invitrogen; K1465-01) according to the instruction of the manufacturer to detect of β galactosidase. The cells were incubated at 37°C for 60 minutes in staining solution and then checked under a microscope for the development of blue colour. The staining solution was replaced with 70% glycerol for long-term storage of stained plates and then kept at 4°C.

2.11.9 Induction of *STRA8* and *C20orf201* genes in Flp-In T-REx-293

The cells were seeded in T75 cm^2 flask at the desired density with complete DMEM containing foetal bovine serum, gold (PAA, Catalogue number; A11-251, Lot A25112-7017) + blasticidin + hygromycin and incubate until attached to the plastic. The cells were refreshed with media containing 2 µg/ml tetracycline (Sigma-Aldrich; T3383) to induce expression and then incubated at 37°C for 24 hours. Two flasks were used as negative controls, untransfected cells and a flask of cells were transfected with the pcDNA5/FRT/TO plasmid. Thereafter, the cells were harvested for RT-PCR (refer to Section 2.5) and Western blot (refer to Section 2.7.3).

The cell lines that were inserted with *STRA8* were also treated with two types of inducer agent, *all-trans* retinoic acid (Sigma-Aldrich; R2625) and *9-cis* retinoic acid (Sigma-Aldrich; R4643) at a final concentration of 5 µM and 10 µM. Thereafter, the cells were harvested for RT-PCR.

Chapter 3.0: mRNA expression profiles of genes identified via a bioinformatics analysis to detect new human CT genes

3.1 Introduction

The identification of new cancer-specific biomarkers will be a major challenge in the development of prognostic and diagnostic interventions for cancer (Feichtinger *et al*., 2012b). Humans possess a class of genes that expression is normally limited to adult testis; however, expression of these same genes is also characteristic of a wide range of cancers: these genes are termed cancer-testis (CT) genes (Whitehurst, 2014). The presence of the blood-testis barrier leads to the products of these genes not being recognised as 'self' by the human immune system when produced by cancer cells (Li *et al*., 2012; Mruk and Cheng, 2010). This feature makes CT antigens very promising targets for clinical purposes, including their use as diagnostic markers, in the development of vaccines and as immunotherapeutic targets (Fratta *et al*., 2011; Lim *et al*., 2012; Mirandola *et al*., 2011; Whitehurst, 2014).

Different bioinformatics analysis methods have been used to identify new human CT antigen candidate genes (for example, see Hofmann *et al*., 2008). Two distinct bioinformatic tools were used by Feichtinger *et al*., (2012b; 2014) to identify potential CTA genes. The first tool was based on expressed sequence tag (EST) analysis data (Feichtinger *et al*., 2014). The second bioinformatic pipeline was based on cancer Microarray data (Feichtinger *et al*., 2012b). Candidate CTA genes were then validated using the RT-PCR technique in the current study.

The bioinformatic approach (Figure 3.1) was initiated by searching for mouse meiosis-specific genes using the GermOnline database (http://www.germonline.org/), which gives 744 mouse genes. These 744 mouse meiotic spermatocyte-specific genes were then mapped to the human orthologous genes and the result of this assignment was 408 human genes (Feichtinger *et al*., 2012a). These 408 orthologous human genes were then filtered using the MitoCheck program (Neumann *et al*., 2010; http://www.mitocheck.org) to eliminate genes that were expressed during mitotic division. The results from the MitoCheck filter left 375 human potential meiosis-specific genes. These human genes were then checked against two distinct bioinformatic programmes using previously published Microarray and EST data sets

(Feichtinger *et al*., 2012a; 2012b; 2014). The rationale for using two bioinformatic pipelines was to check the expression of the tissue and cancer specificity of these genes.

The final number of genes that were potentially of interest from the EST and Microarray analysis were 177 and 40, respectively. The EST analysis produced a much greater number of potential candidate genes when contrasted with that of the Microarray search. The 177 candidate genes identified through the EST pipeline were then divided into four groups based on the expression pattern shown in the EST analysis data. The first group comprises 9 genes, which are predicted to be cancer-testis-restricted genes (class 1); the second group consists of 75 genes, which are predicted to be testis-restricted genes (class 2); the third group comprises 21 genes, which are predicted to be cancer-testis-CNS-restricted genes (class 3) and the fourth group can be described as testis-CNS-restricted genes, and it contains 72 genes (class 4) (Feichtinger *et al*., 2012a; 2014; Sammut *et al*., 2014).

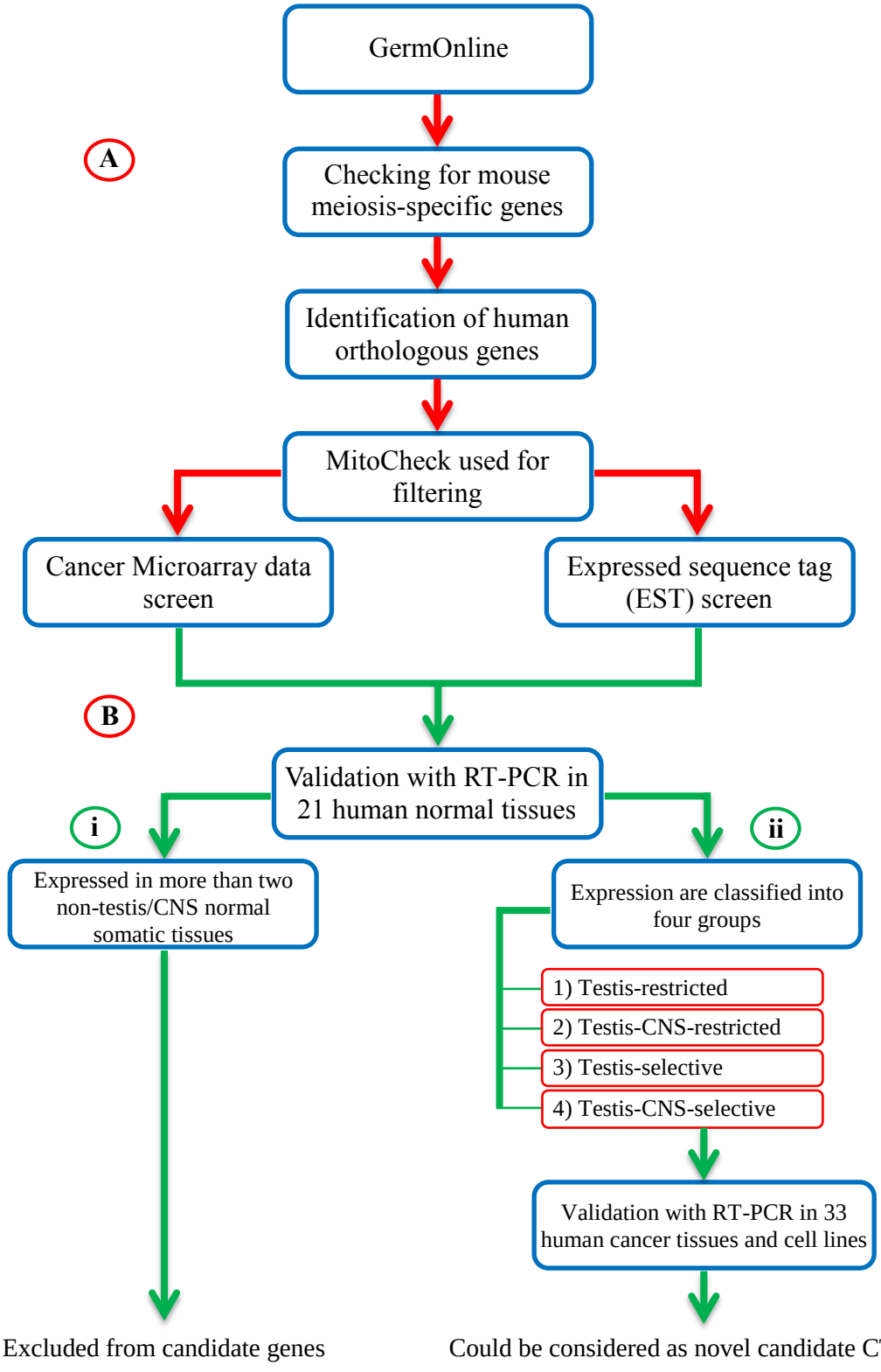


Figure 3.1 Method of identifying a novel CT gene candidate with the use of bioinformatics analyses. (A) Displays the bionformatic procedure that was used to identify CT genes. **(B)** Validation using RT-PCR with 21 human normal tissues. **i)** Genes expressed in several normal tissues were excluded from further assessement. **ii)** Shows the genes expressed only in the following: **1**) in adult testis; **2**) in adult testis and CNS; **3**) in adult testis and selective normal tissues; or **4**) in adult testis, CNS and selective normal tissues (one or two non-testis/CNS normal tissue). Genes from groups 1-4 were then examined further in 33 human cancer cell lines and tissues to identify potential CT gene candidates.

In this chapter, 21 genes were chosen at random from the 105 candidate genes in classes 1-3 identified via the EST analysis. Another 11 genes were also chosen at random from the 40 candidate genes identified by the Microarray analysis. RT-PCR was performed to validate the tissue specificity of these genes. The RT-PCR validation employed RNA from 21 normal tissues, including adult testis. The genes that were expressed only in the testis, or in the testis and CNS tissues, or in the testis and no more than two additional normal tissues, were further examined by RT-PCR using 33 different RNAs isolated from diverse cancer cell lines and tissues.

3.2 Results

3.2.1 Validation of candidate genes identified from Microarray data

11 genes were identified from the gene expression Microarray data analysis pipeline, which predicted they are testis-specific. In the present study, the expression of these genes was analysed using RT-PCR. These predicted genes and their known functions are listed in Table 3.1.

Table 3.1 Meiosis specific-genes identified through the Microarray study and their functions

Gene symbol	Gene name	Function	Reference
CCNA1	Cycline A1	Cell cycle regulator	Weiss *et al*., 2012
C2orf69	Chromosome 2 open reading frame 69	Function unknown	---
C11orf70	Chromosome 11 open reading frame 70	Function unknown	---
C20orf195	Chromosome 20 open reading frame 195	Function unknown	---
FSCN3	Fascin actin-bundling protein 3, testicular	Function unknown	---
GAGE1	G antigen 1	Function unknown	---
HORMAD1	HORMA domain containing 1 (CT46)	Essential for meiotic function in human spermatogenesis, SC formation and homologous alignment	Miyamoto *et al*., 2012 Daniel *et al*., 2011
NOL4	Nucleolar protein 4	Function unknown	---
SSX2	Synovial sarcoma, X breakpoint 2	Associated with transcriptional repressor activity	Chen *et al*., 2012
UBL4B	Ubiquitin-like 4B	Post-translation protein modification	Yang *et al*., 2007
ZNF558	Zinc finger protein 558	Function unknown	---

3.2.1.1 Analysis of expression in normal tissues

In order to investigate their testis-specificity, the genes listed in Table 3.1 were subjected to validation by a RT-PCR analysis with a range of RNAs from 21 human normal tissues. These RNAs were purchased from Clontech and Ambion. RT-PCR intron-spanning primers were designed for every gene (refer to Table 2.2). cDNA was synthesised from the RNA. In the RT-PCR screening, two positive control genes were used; the expression of *β-Actin* was displayed as a positive control for the cDNA quality and *MAGE-A1* was used as a positive control for a known CT gene (testis-specific). Triplicate PCR was performed for each gene.

The RT-PCR screening of these genes on the multiple normal tissues indicated 7 genes (*CCNA1, C2orf69, C11orf70, C20orf195, HORMAD1, NOL4* and *ZNF558*) were expressed in numerous normal tissues, including testis (Figure 3.2); therefore, these genes were all dismissed and were not screened for expression in the human cancer tissues or cell lines, because they did not display the expected expression pattern of either a meiosis-specific gene or a CT gene in normal tissues. The remaining 4 predicted meiosis-specific genes were classified into three sub-classes according to their expression profile in the normal tissues. The first sub-class contained 2 genes, *SSX2* and *UBL4B,* which had expressions restricted to the testis in the normal tissue panel (Figure 3.3). The second sub-class was a testis-selective gene, *GAGE1*, which was expressed in the normal tissues and as a very weak expression in uterus (Figure 3.4). The third sub-class obtained from normal tissue analysis was a testis-CNS-selective gene, *FSCN3*, which had expression limited in the testis, CNS tissues and very weak expression in the prostate, lung and thymus (Figure 3.5). Given the very weak signal in these three somatic tissues, *FSCN3* was classified as ‘selective’ despite the threshold being set as expression in no more than two other somatic tissues.

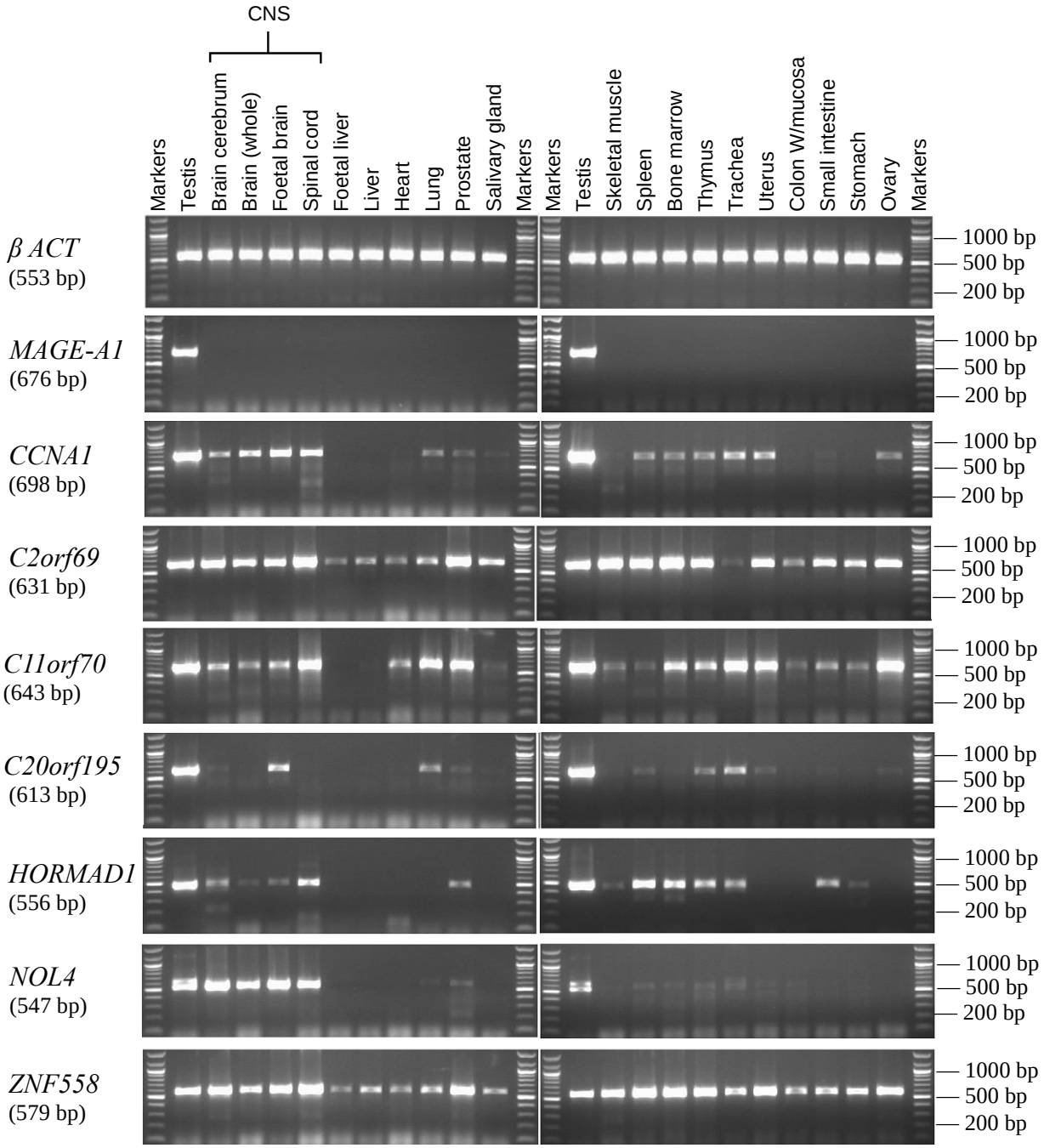


Figure 3.2 RT-PCR analysis of the mRNA from normal human tissues for the genes excluded due to broad somatic expression. Agarose gels show the RT-PCR assays of distinct genes from the Microarray pipeline which were excluded due to their expressions in a range of human normal tissues. These genes are *CCNA1, C2orf69, C11orf70, C20orf195, HORMAD1, NOL4* and *ZNF558*. The expression of all these genes was found in multiple normal cells. *β ACT* expression was used as a positive control for the cDNA samples and the expression of *MAGE-A1* was used as a positive control for the CT antigen genes in normal tissues. The expected amplicon size of each gene is shown on the left between brackets.

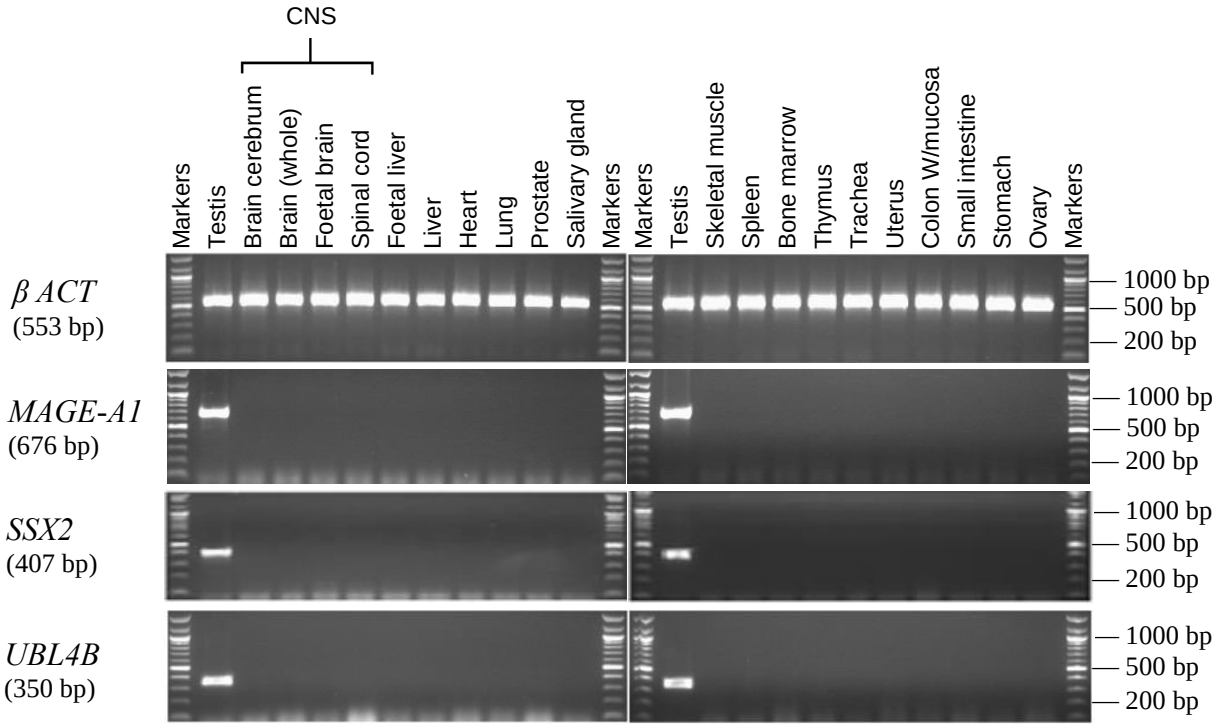


Figure 3.3 RT-PCR analysis of the mRNA from normal human tissues for the testis-restricted genes identified from the Microarray analysis. Agarose gels show the RT-PCR assays for the testis–restricted genes, *SSX2* and *UBL4B* that were identified from the microarray pipeline. cDNAs were isolated from the total RNA from 21 normal tissues. Both genes were expressed only in the normal testis. *β ACT* expression was used as a positive control for the cDNA samples and the expression of *MAGE-A1* was used as a positive control for a CT antigen genes in human normal cells. The expected amplicon size of each gene is shown on the left between brackets.

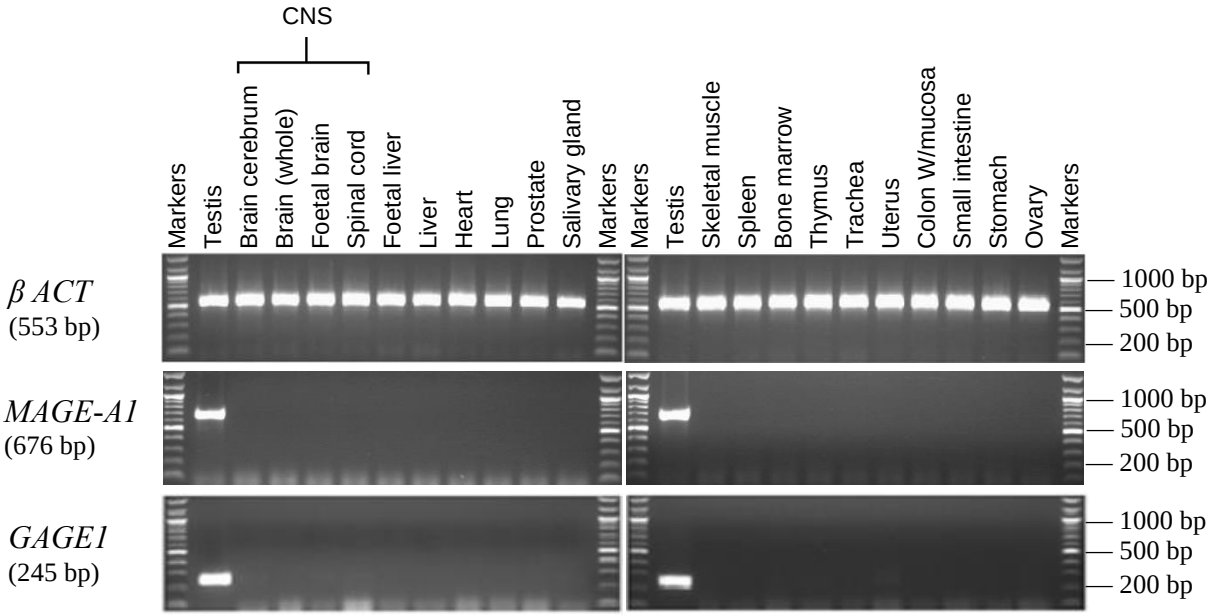


Figure 3.4 RT-PCR analysis of the mRNA from normal human tissues for the testis-selective gene identified from the Microarray analysis. Agarose gels show the RT-PCR assays for the testis–selective gene, *GAGE1* was identified from the microarray pipeline. cDNAs were isolated from the total RNA from 21 normal tissues. The expression of *GAGE1* was in the normal testis and uterus. *β ACT* expression was used as a positive control for the cDNA samples and the expression of *MAGE-A1* was used as a positive control for the CT antigen genes in human normal cells. The expected amplicon size of each gene is shown on the left between brackets.

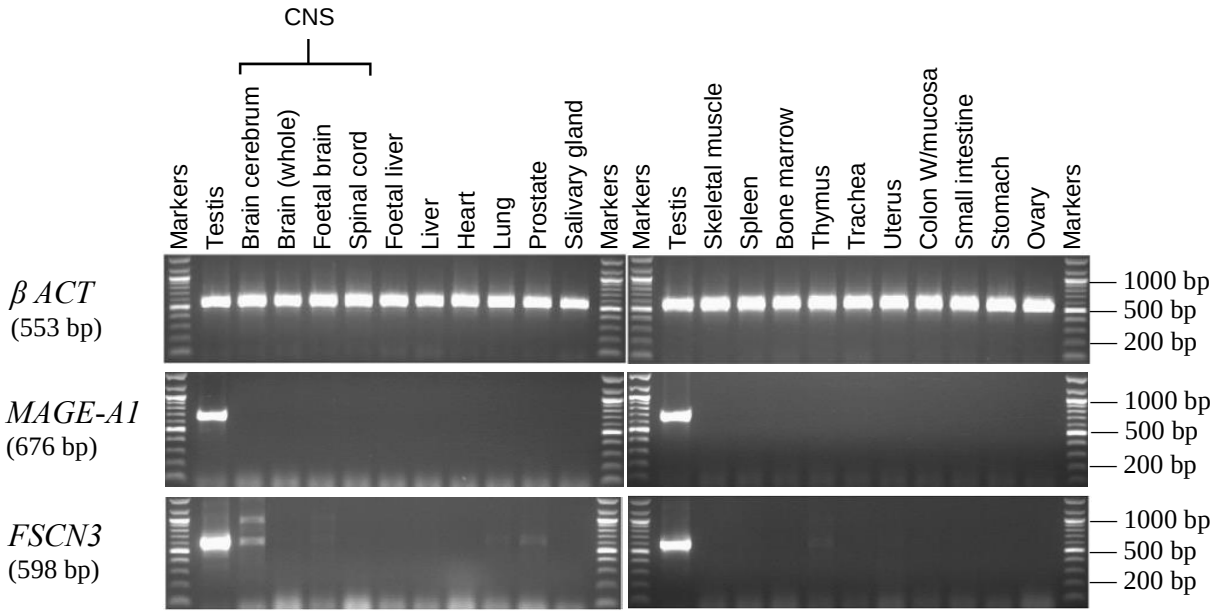


Figure 3.5 RT-PCR analysis of the mRNA from normal human tissues for the testis-CNS-selective gene identified from the Microarray analysis. Agarose gels show the RT-PCR assays for the testis–CNS-selective gene, *FSCN3* was identified from the microarray pipeline. cDNAs were isolated from the total RNA from 21 normal tissues. The expression of *FSCN3* was in the normal testis, prostate, central nervous system, lung and thymus. *β ACT* expression was used as a positive control for the cDNA samples and the expression of *MAGE-A1* was used as a positive control for the CT antigen genes in human normal cells. The expected amplicon size of each gene is shown on the left between brackets.

3.2.1.2 Analysis of candidate gene expression profiles in cancer cells/tissues

Expression of the testis-restricted genes *SSX2* and *UBL4B*, the testis-selective gene *GAGE1* and the testis-CNS-selective gene *FSCN3* were further investigated using cDNA from 33 human cancer cell lines and tissues derived from different organs. The expression of *β-Actin* was displayed as a positive control for the cDNA quality, while *MAGE-A1* was used as a positive control for a known candidate CT genes. Triplicate PCR was performed for each gene.

SSX2 and *GAGE1* were found to display expressions in two or nine cancer cells, respectively (Figure 3.6); this means that both genes are considered to be good candidates for the CT genes and were classified as CT-restricted and CT-selective genes, respectively. *SSX2* displayed expression exclusively in the normal testis and also exhibited expression in one or more cancer types, whereas the expression profile for *GAGE1* was observed in one of the non-testis-CNS normal tissues in addition to the testis, but also displayed expression in one or more cancer samples. The RT-PCR expression profile of *UBL4B* was not observed in any of the cancer samples tested here; therefore, this gene was classified as a testis-specific gene (Figure 3.7). Conversely, this gene may be expressed in cancer cell lines and/or tissues not covered in the current study. For this reason, the *UBL4B* gene was not dismissed as a CT candidate gene.

The *FSCN3* gene was expressed in three cancer cell lines, including HCT116, TO14 and K-562. However, the RT-PCR product for *FSCN3* was different, because it gave the expected band sizes in HCT116 and K-562, but not in the TO14 cell line, when compared to the testis. The RT-PCR product of *FSCN3* in the TO14 cell line was higher than expected (approximately 1000 bp). Therefore, RT-PCR products for *FSCN3* in the testis and TO14 were subjected to DNA sequencing to ensure that the correct target was being amplified (see Table 3.2 for sequencing results). The partial sequence obtained was virtually identical to that of *FSCN3* indicating this higher molecular weight band could be a splice variant. The exact nature of this valiant was not determined. After obtaining the sequencing results, *FSCN3* is considered as a potential CT gene and was classified as a CT-CNS-selective gene due to its expression profile in the normal testis, central nervous system tissues and no more than two additional non-testis-CNS normal tissues, but it also showed expression in at least one cancer tissue and/or cell line.

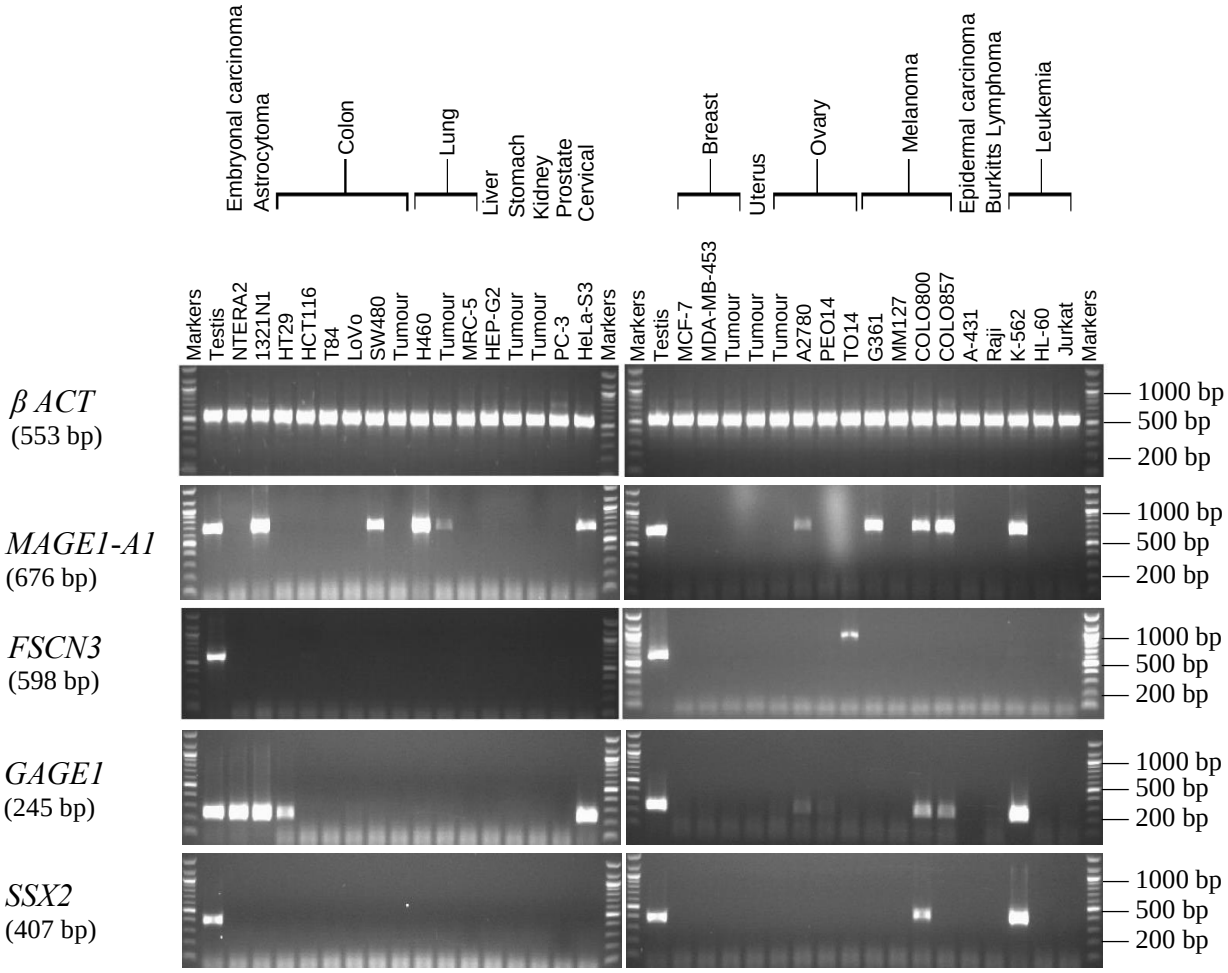


Figure 3.6 RT-PCR analysis of mRNA from human cancer cells for the potential CT antigen genes identified from the Microarray analysis. Agarose gels show the RT-PCR assays for the potential CT antigen genes identified by the Microarray pipeline. The mRNA expressions of these genes were positive in different cancer cell lines and tissues. The expression of *SSX2* was shown in normal testis, COLO800 and K-562 cell lines. The expression of *GAGE1* was observed in 9 cancer cell lines. *FSCN3* expression was shown in three cancer samples, but the product size in TO14 was larger than expected. *β ACT* expression was used as a positive control for the cancer cDNA samples and the expression of *MAGE-A1* was used as a positive control for the CT antigen genes in human cancer cells. The expected amplicon size of each gene is shown on the left between brackets.

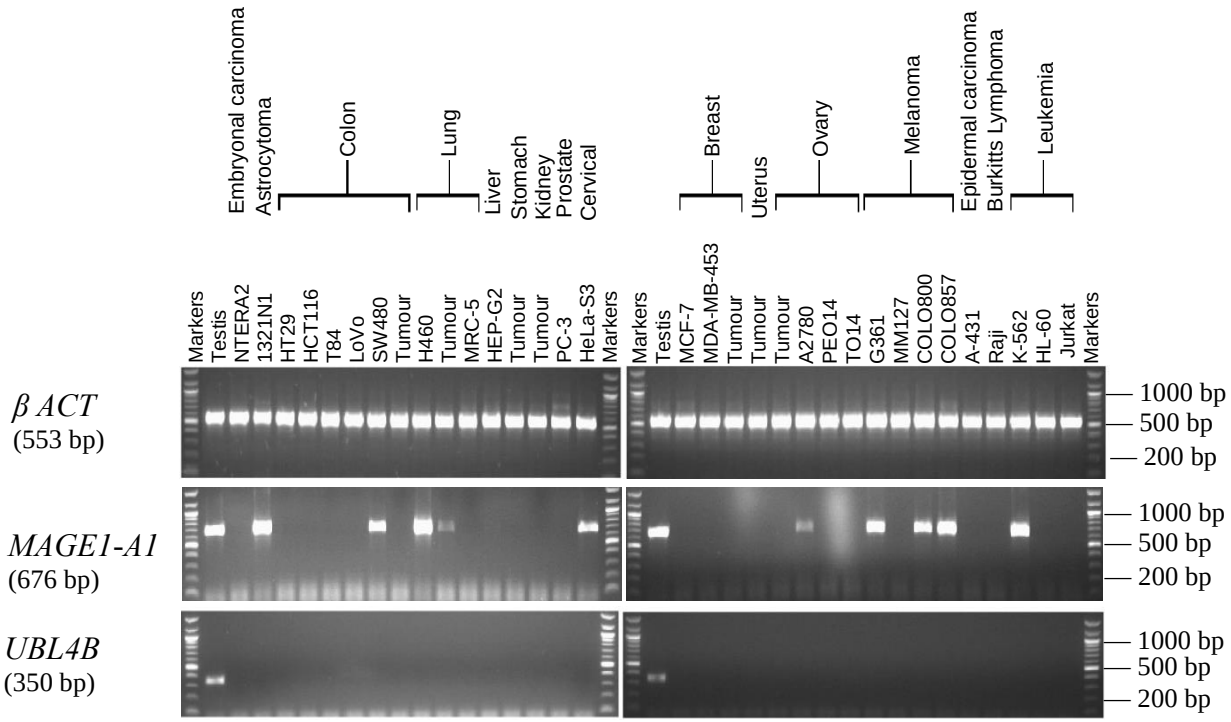


Figure 3.7 RT-PCR analysis of mRNA from human cancer cells for the testis-restricted gene identified from the Microarray analysis. Agarose gels show the RT-PCR assays for the testis–restricted gene, *UBL4B*, from the Microarray pipeline. cDNAs were isolated from the total RNA from 33 cancer tissues and cell lines.*UBL4B* was expressed only in the normal testis tissue and no expression was detected in any cancer samples tested. *β ACT* expression was used as a positive control for the cancer cDNA samples and the expression of *MAGE-A1* was used as a positive control for the CT antigen genes in human cancer cells. The expected amplicon size of each gene is shown on the left between brackets.

3.2.1.3 Summary of the sequencing results for the Microarray identified genes

The PCR product for the desired gene was purified and sequenced in order to amplify the correct sequence. The sequences recovered from the sequencing were aligned using Basic Local Aliment Search Tools (BLAST; NCBI http://blast.ncbi.nlm.nih.gov/Blast.cgi). The sequencing results of the Microarray genes are summarised in Table 3.2.

Table 3.2 Summary of the sequencing results for the RT-PCR screening of the Microarray analysis genes

Gene	Primer	Expected size (bp)	Sequenced in		Sequence size (bp)	Sequence identity %
			Normal tissues	Cancer samples		
FSCN3	F1	598	Testis Prostate Thymus ---	--- --- --- TO14	517 79 268 837	100 100 99 100
GAGE1	F1	245	Uterus	---	79	95
NOL4	F1	547	Testis (upper band) Testis (lower band)	--- ---	154 154	100 99

3.2.2 Validation of candidate genes identified by EST database analysis

177 genes were identified from the EST data analysis pipeline with predicted meiosis-associated genes. In the present study, the mRNA expressions of 21 of these genes from class 1-3 were analysed using the RT-PCR technique. These genes and their known functions are listed in Table 3.3.

Table 3.3 Predicted meiosis associated-genes identified through the EST analysis and their functions

Gene symbol	Gene name	Function	Reference
BTG4	B-cell translocation gene 4	Cell cycle regulator	Auer *et al*., 2005
C5orf50	Chromosome 5 open reading frame 50	Function unknown	---
C11orf65	Chromosome 11 open reading frame 65	Function unknown	---
C20orf201	Chromosome 20 open reading frame 201	Function unknown	---
COX7B2	Cytochrome c oxidase subunit VIIb2	Function unknown	---
CST8	Cystatin 8, also known as *CRES*	A protease inhibitor	Cornwall *et al*., 1999
DAZL	Deleted in azoospermia-like	Encodes a germ-cell-specific RNA-binding protein that is essential in spermatogenesis	Reynolds and Cooke, 2005; Smorag *et al*., 2014
FABP9	Fatty acid binding protein 9	Function unknown	---
KCNU1	Potassium channel, subfamily U, member 1	Function unknown	---
LYRM1	LYR motif containing 1	Function unknown	---
ODF3	Outer dense fibre of sperm tails 3	A component of sperm tail	Ghafouri-Fard and Modarressi, 2012b
PFN3	Profilin 3	A role in spermatogenic cells	Behnen *et al*., 2009
PRSS45	Protease, serine, 45	Function unknown	---
SLC25A41	Solute carrier family 25, member 41	Transport molecule over the mitochondrial membrane	Haitina *et al*., 2006
SSX5	Synovial sarcoma, X breakpoint 5	May act as transcriptional repressor	de Bruijn *et al*., 2002
STRA8	Stimulated by retinoic acid 8	Meiosis initiation in male and female in mice	Anderson *et al*., 2008
TEKT5	Tektin 5	Function unknown	---
TEPP	Testis, prostate and placenta expressed	Function unknown	---
TMEM225	Transmembrane protein 225	Function unknown	---
TULP2	Tubby like protein 2	Function unknown	---
WDR27	WD repeat domain 27	Function unknown	---

3.2.2.1 The results of RT-PCR screening of the EST genes in normal tissues

The genes listed in Table 3.3 were validated by a RT-PCR analysis with a range of RNAs from 21 human normal tissues in order to investigate their potential testis-specificity. These RNAs were purchased from Clontech and Ambion. Intron-spanning primers were designed for every gene (refer to Table 2.2) and were used to perform RT-PCR using cDNA synthesised from the RNA preparation. Two positive control genes were used during the RT-PCR screening; the expression of *β-Actin* was displayed as a positive control for the cDNAs quality and *MAGE-A1* was used for a positive control as a known CT gene. Triplicate PCR was performed for each gene.

The RT-PCR screening of the 21 genes predicted by the EST bioinformatic analysis on the multiple normal tissues demonstrated that 10 genes (*BTG4, C11orf65, COX7B2*, *DAZL, LYRM1, PRSS45, SLC25A41, TEKT5, TULP2* and *WDR27*) were expressed in numerous normal tissues including testis (Figure 3.8); therefore, these genes were all dismissed and were not screened for expression in the human cancer tissues or cell lines because they did not display the expected expression pattern of either a meiosis gene or a CT gene in normal tissues. The remaining 11 genes were categorised into four groups according to their expression profile observed in the normal tissues. The first group contained 6 genes, *C5orf50, CST8, FABP9, ODF3, STRA8* and *TMEM225*, and had expression restricted to the testis in the normal tissue panel (Figure 3.9). For the *STRA8* gene, the expected PCR product size in the testis sample according to design of forward and reverse primers is 510 bp, but the RT-PCR displayed higher band of 723 bp after being confirmed by DNA sequencing (sequencing results are summarised in Table 3.4). The second group was made up of testis-selective genes such as *TEPP*, which was expressed in two or fewer of the somatic normal tissues; for *TEPP* there was a weak expression in the trachea (Figure 3.10). The third group obtained from normal tissue was composed of testis-CNS-selective genes, *PFN3* and *SSX5*, which had expressions limited in the testis and central nervous system tissues and weak expression in uterus and thymus, respectively (Figure 3.11). The last genes were *C20orf201* and *KCNU1*, which had expressions restricted to the testis and central nervous system tissues; therefore, these genes were classified as a testis-CNS- restricted (Figure 3.12).

The expected product size for the *C20orf201* gene is 505 bp; however, it was shown that the PCR product for this gene in the testis cDNA was smaller than 505 bp; therefore, the PCR product was confirmed by sequencing that showed a significant sequence similarity to *C20orf201* (see Table 3.4 for sequencing results). Conversely, *KCNUI* expression in the brain

cerebellum and spinal cord exhibited a band of unexpected size (approximately 280 bp) compared to its expression in the normal testis, which was 642 bp. However, sequencing of this PCR product showed a significant sequence similarity to *KCNU1* (see Table 3.4). The *KCNU1* cDNA containing whole open reading frame (ORF) has a length of 3450 bp and consists of 27 exons. The designed primers included exons from number 8 to 14, but the sequence results obtained cover exons 9. These results could suggest another splice variant for *KCNU1*.

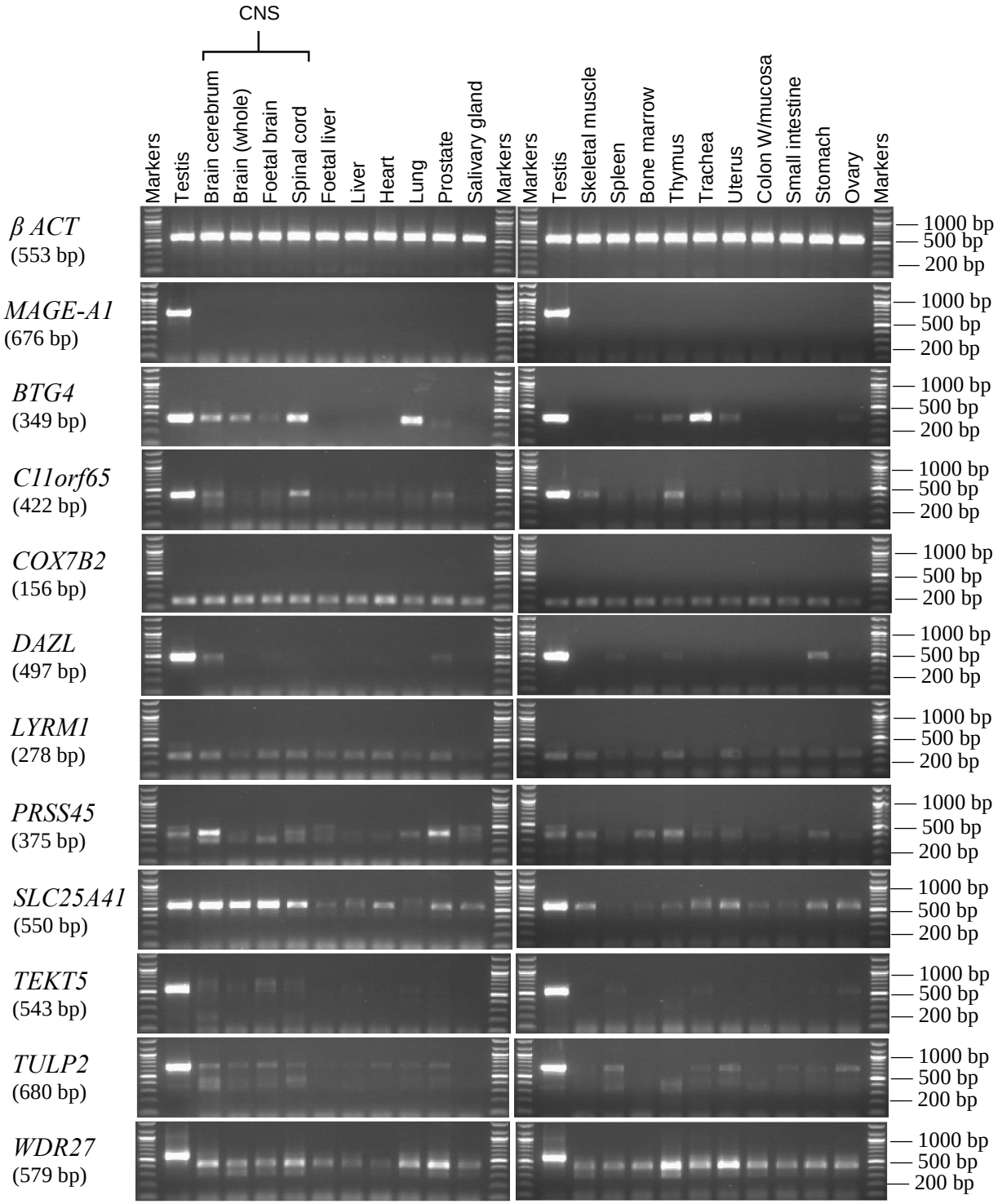


Figure 3.8 RT-PCR analysis of the mRNA for the genes excluded identified from the EST pipeline analysis. Agarose gels show the RT-PCR assays of the genes from the EST pipeline which were excluded because expressions were shown in a several human normal tissues and these genes are *BTG4*, *C11orf65*, *COX7B2*, *DAZL*, *LYRM1*, *PRSS45*, *SLC25A41*, *TEKT5*, *TULP2* and *WDR27*. cDNAs were isolated from the total RNA from 33 cancer tissues and cell lines. The expression of all these genes was found in multiple normal cells. *β ACT* expression was used as a positive control for the cancer cDNA samples and the expression of *MAGE-A1* was used as a positive control for the CT antigen genes in normal tissues. The expected amplicon size of each gene is shown on the left between brackets.

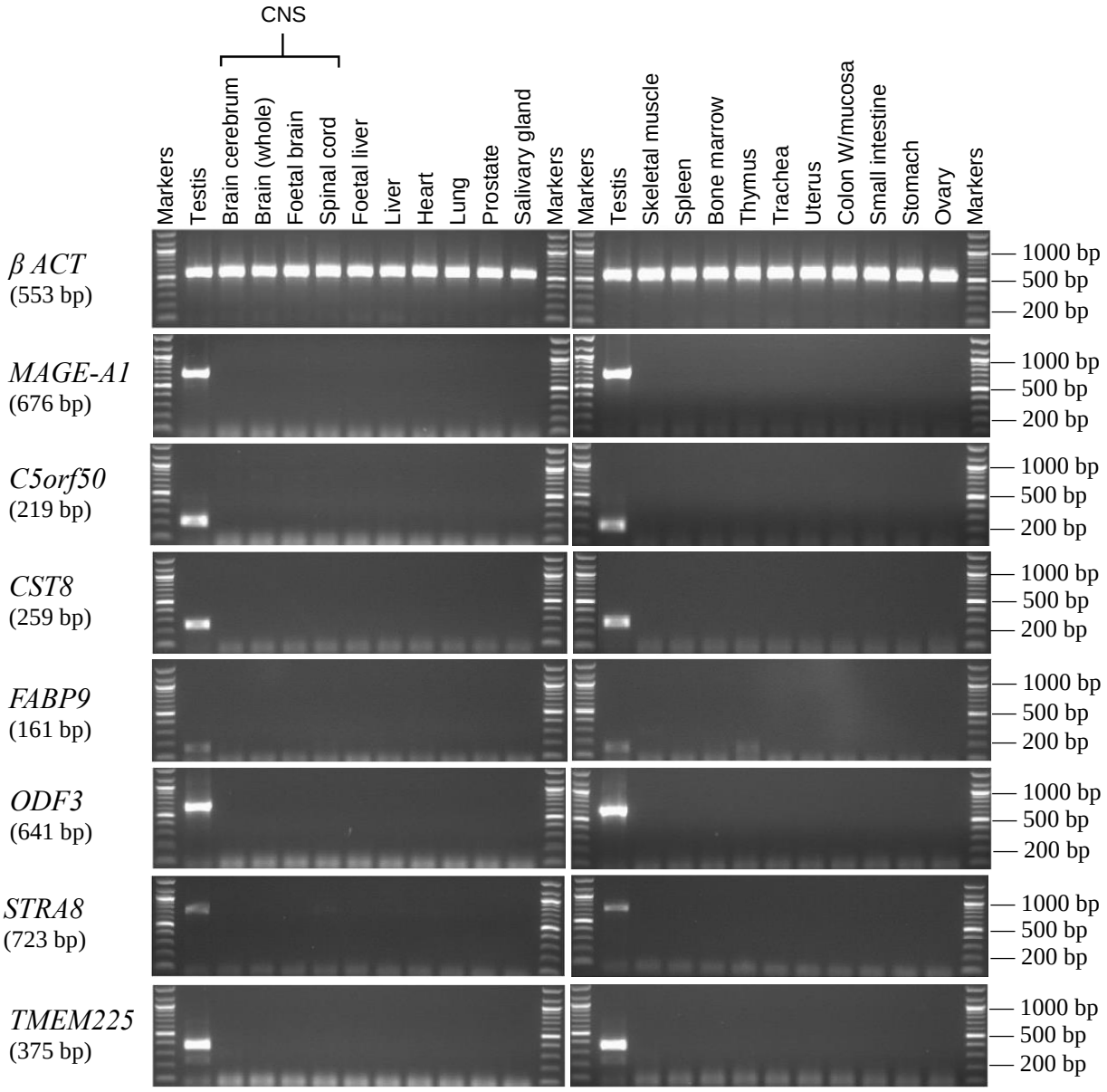


Figure 3.9 RT-PCR analysis of the mRNA for the testis-restricted genes identified from the EST pipeline analysis in human normal tissues. Agarose gels show the RT-PCR assays for the testis–restricted genes; *C5orf50*, *CST8*, *FABP9*, *ODF3*, *STRA8* and *TMEM225* were identified from the EST pipeline. cDNAs were isolated from the total RNA from 21 normal cells. These genes were expressed only in the normal testis. *β ACT* expression was used as a positive control for the cDNA samples and the expression of *MAGE-A1* was used as a positive control for the CT antigen genes in human normal cells. The expected amplicon size of each gene is shown on the left between brackets.

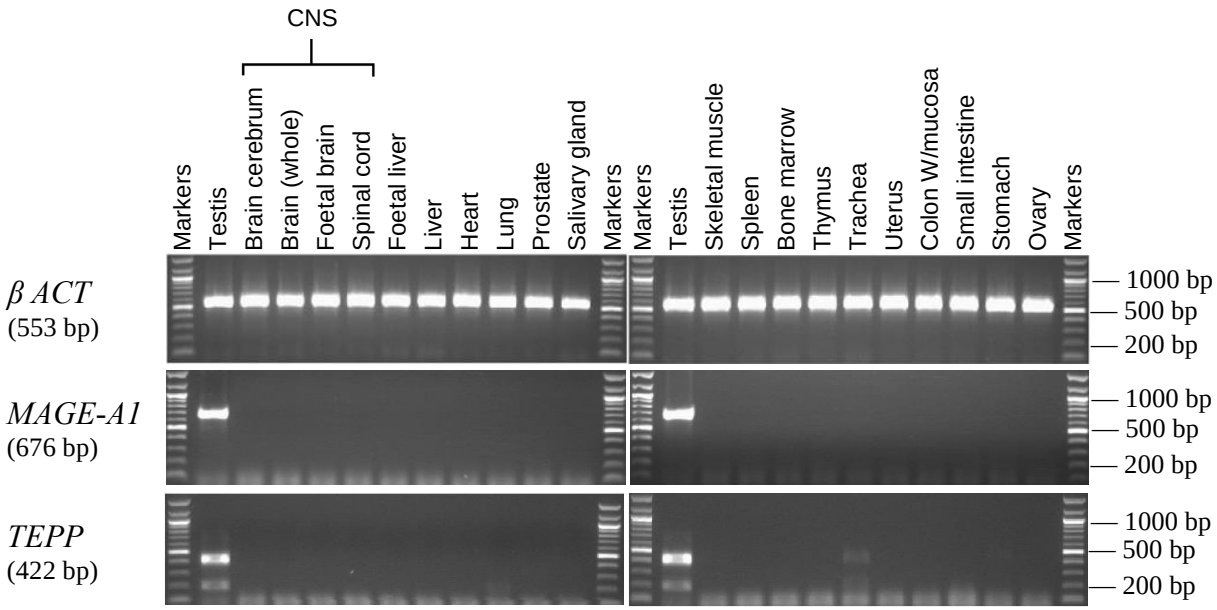


Figure 3.10 RT-PCR analysis of the mRNA for the testis-selective gene identified from the EST pipeline analysis in human normal tissues. Agarose gels show the RT-PCR assays for the testis–selective gene, *TEPP* that was identified from the EST pipeline. cDNAs were isolated from the total RNA from 21 normal cells. The expression of this gene was exclusively in the normal testis and a faint band indicated some expression in the trachea. *β ACT* expression was used as a positive control for the cDNA samples and the expression of *MAGE-A1* was used as a positive control for the CT antigen genes in human normal cells. The expected amplicon size of each gene is shown on the left between brackets.

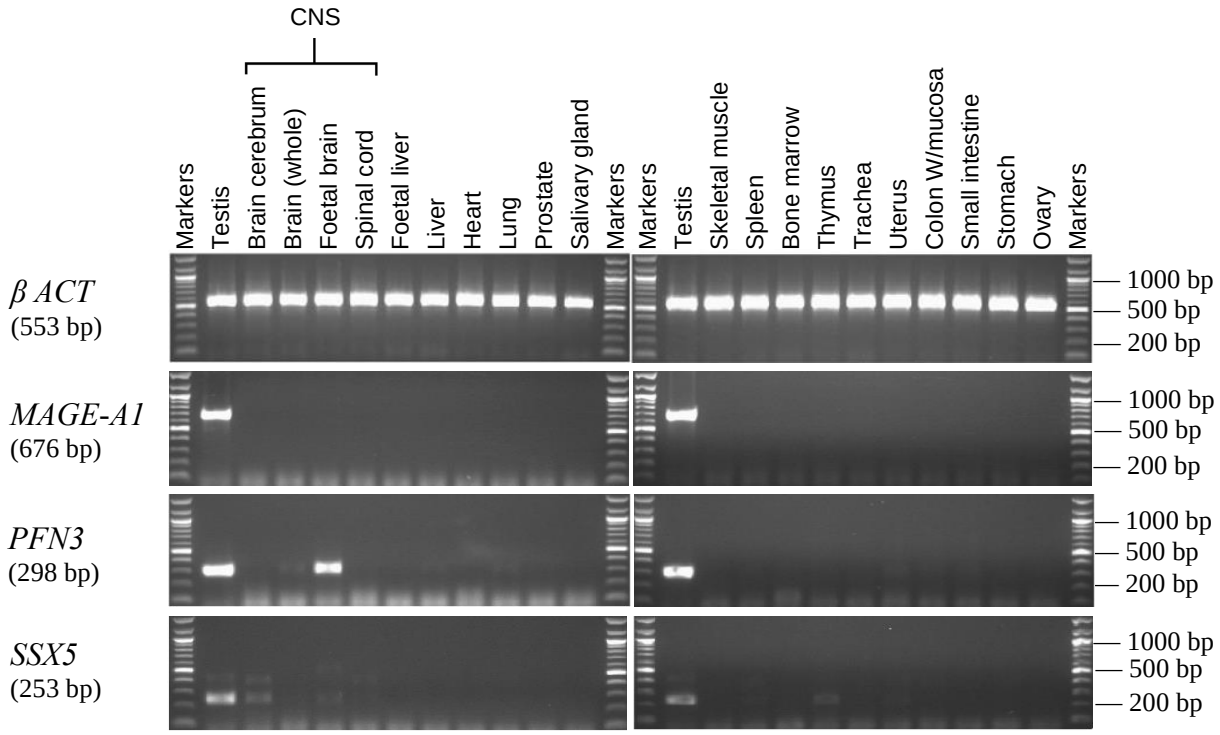


Figure 3.11 RT-PCR analysis of the mRNA for the testis-CNS-selective genes identified from the EST pipeline analysis in human normal tissues. Agarose gels show the RT-PCR assays for the testis–CNS-selective genes, *PFN3* and *SSX5* were identified from the EST pipeline. cDNAs were isolated from the total RNA from 21 normal cells. The expressions of these genes were shown in the testis and central nervous system tissues and also in the number of normal tissues. *β ACT* expression was used as a positive control for the cDNA samples and the expression of *MAGE-A1* was used as a positive control for the CT antigen genes in human normal cells. The expected amplicon size of each gene is shown on the left between brackets.

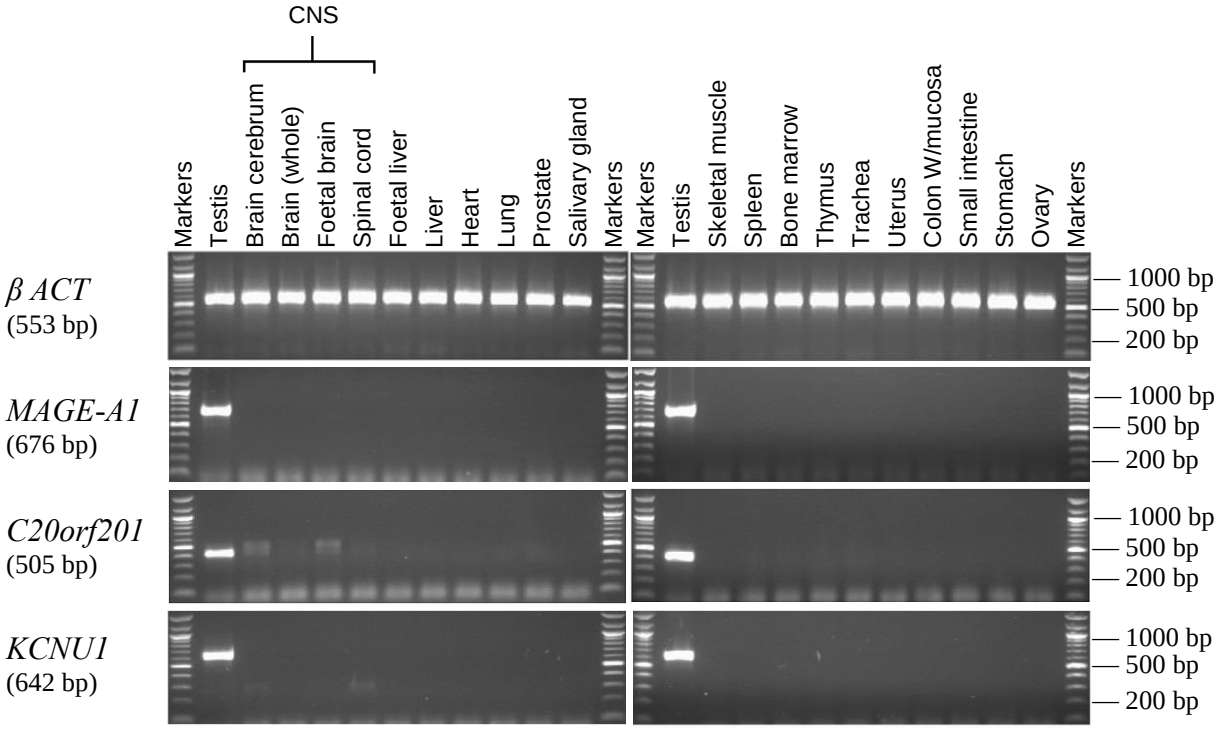


Figure 3.12 RT-PCR analysis of the mRNA for the testis-CNS-restricted genes identified from the EST pipeline analysis in human normal tissues. Agarose gels picture illustrating the RT-PCR assays for the testis-CNS-restricted genes, *C20orf201* and *KCNU1* were identified from the EST pipeline. cDNAs were isolated from the total RNA from 21 normal cells. The expression of these genes was apparent in the testis and in central nervous system tissues. *β ACT* expression was used as a positive control for the cDNA samples and the expression of *MAGE-A1* was used as a positive control for the CT antigen genes in human normal cells. The expected amplicon size of each gene is shown on the left between brackets.

3.2.2.2 The results of RT-PCR screening of the EST genes in cancer tissues and cell lines

The testis-restricted genes *C5orf50, CST8, FABP9, ODF3, STRA8* and *TMEM225*; the testis-selective gene *TEPP*; the testis-CNS-selective genes *PFN3* and *SSX5* and the testis-CNS-restricted genes *C20orf201* and *KCNU1* were further investigated by RT-PCR in a range of 33 human cancer cell lines and tissues resulting from different organs. All cell lines in the current study were grown and RNAs were made for each cell line. The mRNAs were reverse-transcribed to cDNAs and RT-PCR was performed. Two positive control genes were used in the RT-PCR screening; the expression of *β-Actin* was displayed as a positive control for the cDNAs quality, while *MAGE-A1* was used as a positive control for the good candidate CT genes. Triplicate PCR was performed for each gene.

Among the above genes identified by the EST bioinformatic analysis, three genes, *CST8*, *ODF3* and *TMEM225* were not found to exhibit any expression in the cancer samples used in the current study; therefore, they were classified as testis-specific genes, because their expression was restricted to the normal testis only (Figure 3.13). However, these three genes were not dismissed as CT gene candidates and further screening in other cancer cell line and/or tissues not covered in this study may express these genes. In contrast, *C20orf201, C5orf50, FABP9, KCNU1, PFN3, SSX5, STRA8* and *TEPP* were expressed in different types of cancer cells; therefore, they are considered good candidates for CT genes (Figure 3.14). Furthermore, these genes according to their expression were classified into four groups. The first group was CT-restricted genes, which includes the following: *C5orf50, FABP9* and *STRA8*. The second group was CT-selective gene and included only the *TEPP* gene. The third group was CT-CNS-selective and included two genes, *PFN3* and *SSX5*. The last group included two genes, *C20orf201* and *KCNU1* and they were classified as CT-CNS-restricted. The results of DNA sequencing for the genes *C20orf201, C5orf50, FABP9, KCNU1, PFN3, SSX5, STRA8* and *TEPP* were compared with the reference sequence using the NCBI's BLAST software (http://blast.ncbi.nlm.nih.gov) (see Table 3.4).

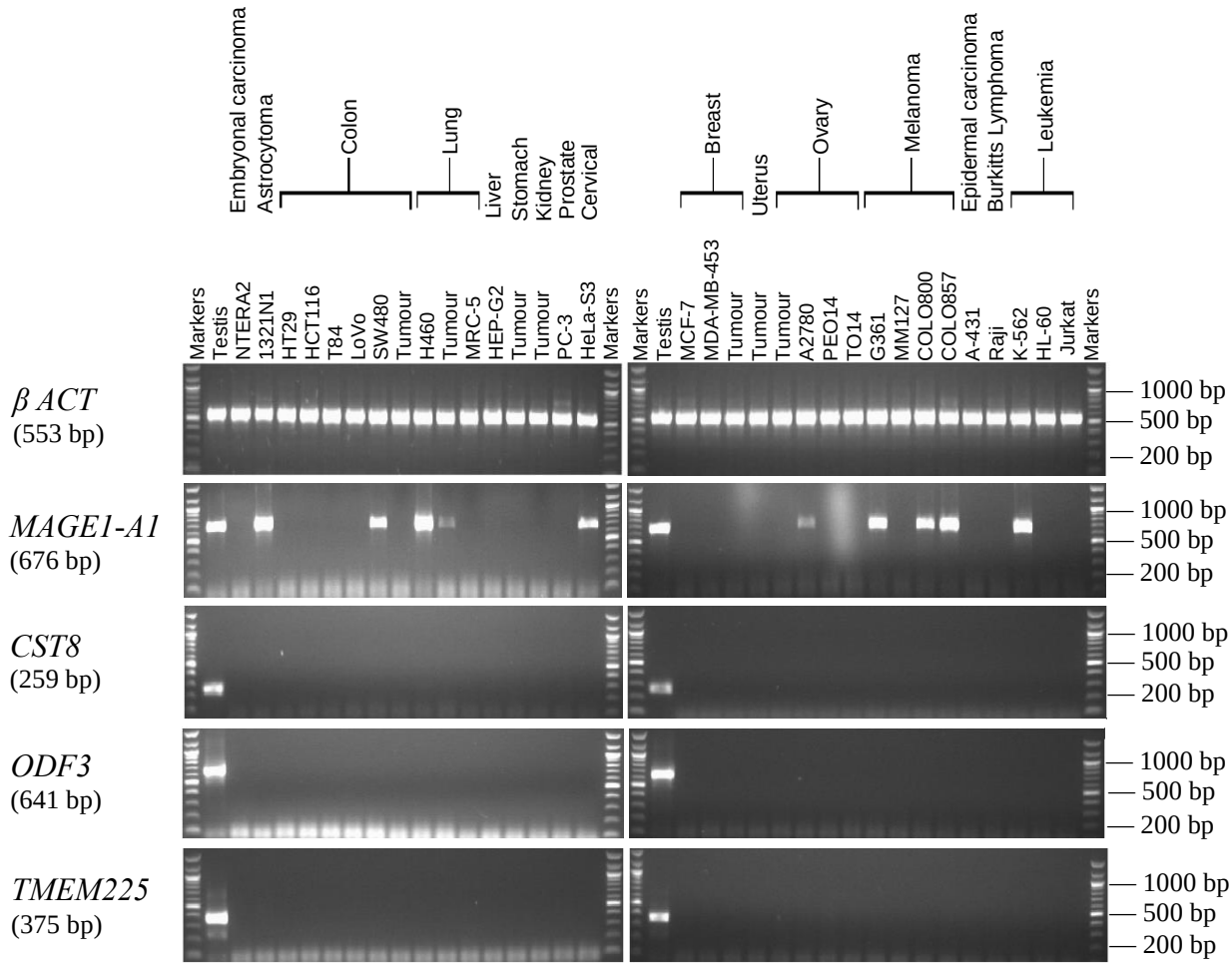


Figure 3.13 RT-PCR analysis of the mRNA for the testis-restricted genes identified from the EST analysis in human cancer cell lines and tumour tissues. Agarose gels show the RT-PCR assays for the testis–restricted genes; *CST8, ODF3* and *TMEM225* from the EST pipeline. cDNAs were isolated from the total RNA from 33 cancer tissues and cell lines. These genes were expressed only in the normal testis tissue and no expression observed in any other cancer samples. *β ACT* expression was used as a positive control for the cancer cDNA samples and the expression of *MAGE-A1* was used as a positive control for the CT antigen genes in human cancer cells. The expected amplicon size of each gene is shown on the left between brackets.

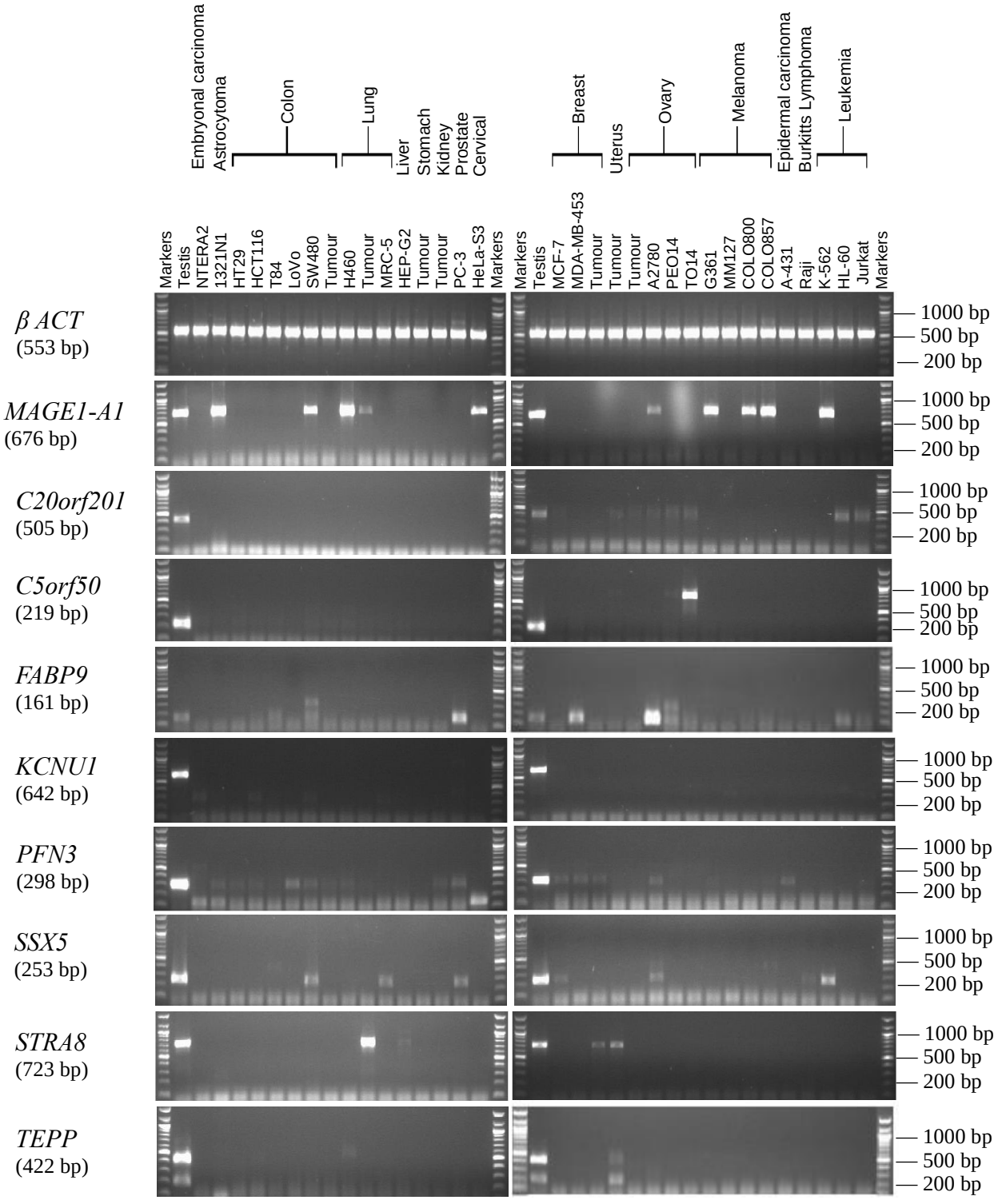


Figure 3.14 RT-PCR analysis of the mRNA for the potential CT antigen genes identified from the EST analysis in human cancer cell lines and tumour tissues. Agarose gels show the RT-PCR assays for the potential CT antigen genes identified by the EST pipeline. The mRNA expression of these genes was positive in different cancer cell lines and tissues. *β ACT* expression was used as a positive control for the cancer cDNA samples and the expression of *MAGE-A1* was used as a positive control for the CT antigen genes in human cancer cells. The expected amplicon size of each gene is shown on the left between brackets.

3.2.2.3 Summary of the sequencing results for the EST identified genes

The PCR product for the desired gene was purified and sequenced in order to amplify the correct sequence. The sequences recovered from the sequencing were blasted using the BLAST program. The sequencing results of the EST genes are summarised in Table 3.4

Table 3.4 Summary of the sequencing results for the RT-PCR screening of the EST analysis genes

Gene	Primer	Expected size (bp)	Sequenced in		Sequence size (bp)	Sequence identity %
			Normal tissues	Cancer samples		
C5orf50	F1	219	Testis ---	--- TO14	111 434	100 100
C20orf201	F1	505	Testis Foetal brain --- --- ---	--- --- A2780 TO14 Jurkat	302 294 456 323 348	100 100 99 100 100
FABP9	F1	161	Testis Thymus --- ---	--- --- PC-3 A2780	114 113 119 60	100 100 100 97
KCNU1	F1	642	Testis Brain cerebellum --- ---	--- --- NT2 HCT116	467 113 96 61	100 98 98 98
PFN3	F1	298	Testis Foetal brain --- --- ---	--- --- LoVo PC-3 A2780	257 94 178 180 166	99 100 100 100 100
SSX5	F1	253	Testis Thymus --- ---	--- --- Uterus tumour G361	215 212 211 204	97 98 98 97
STRA8	F1	510	Testis --- --- --- ---	--- Lung tumour HEP-G2 Breast tumour Uterus tumour	662 660 408 40 246	100 99 98 100 98
TEKT5	F1	543	Testis Spleen	--- ---	382 191	99 99
TEPP	F1	422	Testis (upper band) Testis (lower band) Trachea	--- --- ---	365 158 145	99 98 100
WDR27	F1	579	Testis Uterus (upper band) Uterus (lower band)	--- --- ---	411 136 35	100 99 97

3.3 Discussion

3.3.1 Summary of the RT-PCR expression profiles for the Microarray and the EST genes

Two bioinformatic studies were used to identify new CT gene candidates. In the current study, these genes were tested using a RT-PCR analysis to investigate the expression of these genes in a range of normal tissues or/and cancer samples. According to their expression profiles as assessed using RT-PCR, candidate genes were determined as potential CT genes, because these predicted meiosis-specific genes were observed to show expression in different types of cancer samples.

The Microarray bioinformatic analysis identified 40 predicted meiosis-specific genes and 11 of these genes were chosen at random for RT-PCR screening here (others were screened by other group members). RT-PCR analysis of 7 genes showed expression in multiple normal tissues, thus all 7 genes were dismissed as candidate CT genes. This screen also identified 4 of the 11 genes as potential novel CT genes. These 4 genes were characterised after RT-PCR screening as the following; *SSX2* is a CT-restricted gene; *GAGE1* is a CT-selective gene; *FSCN3* is a CT-CNS-selective gene and *UBL4B* is a testis-restricted gene.

In contrast, 177 predicted meiosis-specific genes were identified via the EST pipeline and 21 of these genes were chosen at random for RT-PCR screening here (all other genes were screened by other members of the group). In this test, 10 genes were excluded from the CT gene candidates due to their expression in different types of non-testis-CNS normal tissues. Furthermore, the RT-PCR screening was also identified 11 of the 21 genes as potential novel CT genes. These 11 genes have been characterised after RT-PCR screening as the following: *CST8, ODF3* and *TMEM225* are testis-restricted genes; *C5orf50, FABP9* and *STRA8* are CT-restricted genes; *PFN3* and *SSX5* are CT-CNS-selective genes; *C20orf201* and *KCNU1* are CT-CNS-restricted genes and finally, *TEPP* was categorised as a CT-selective gene. The RT-PCR expression profiles for the meiosis-specific genes identified by the Microarray and EST analysis pipelines are summarised in Figures 3.15 and 3.16 for the Microarray analysis and Figures 3.17 and 3.18 for the EST pipelines.

According to this study, the expression of the *STRA8* gene was restricted to the testis among normal tissues; with additional expression also being shown in three cancer tumours and one cell line (HEP-G2). However, RT-PCR analysis showed that the expression profile of the *C20orf201* gene was observed in the testis and in the tissues of the CNS (brain cerebellum,

whole brain, foetal brain and spinal cord). Two bands of *C20orf201* were detected in the CNS tissues; the large and small sizes of the bands were related to splice variants 1 and 2, respectively (see Chapter 5 for more details). The expression of *C20orf201* was also found in various types of cancer, including MCF-7, uterus tumour and ovarian cell lines (A2780, PEO14 and TO14). Therefore, both genes are important and seen as good candidates to encode CTAs and promising for being cancer biomarkers in different cancer types. For this reason, these genes were chosen for further, in detail studies, described in Chapters 4 and 5, respectively.

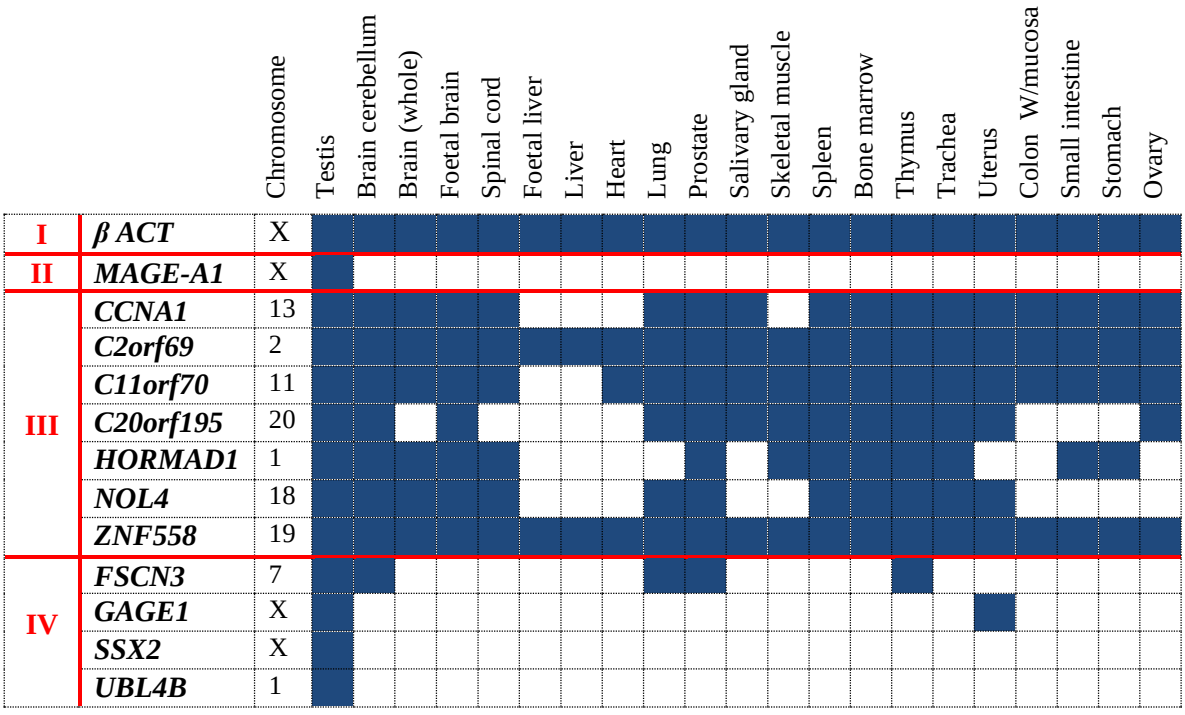


Figure 3.15 Summary of the RT-PCR results for the Microarray analysis genes in different normal tissues. The expression patterns of individual genes are displayed by rows in the grid, with the column corresponding to the normal human tissues used. **I.** The expression profile of the control gene for cDNAs quality. **II.** The expression pattern of the positive control for CT genes. **III.** The expression pattern for seven genes from the Microarray analysis lists that showed expression in several human normal tissues; therefore, these genes were dismissed at this stage. **IV.** The expression pattern for four genes; *FSCN3*, *GAGE1*, *SSX2* and *UBL4B* that might to be good CT gene candidates. The chromosomal location of all genes is shown in a box in front of the gene name. The dark blue boxes indicate the positive expression of the given genes, whereas the empty boxes indicate negative RT-PCR results.

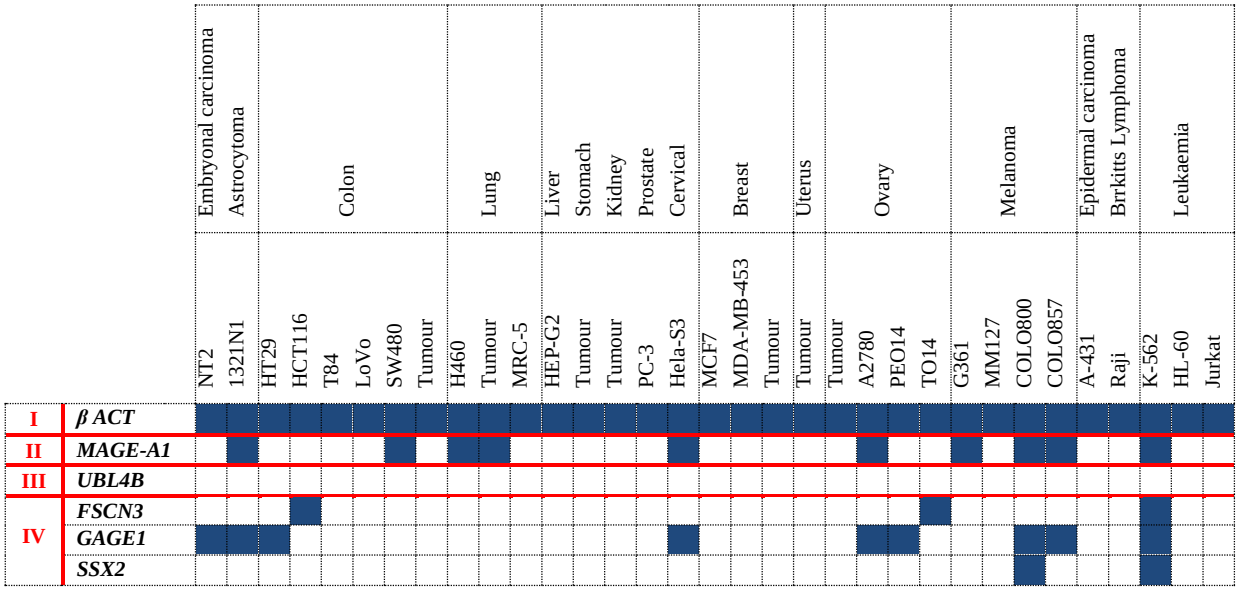


Figure 3.16 Summary of the RT-PCR results for the Microarray analysis genes in different cancer tissues and cell lines. The expression patterns of individual genes are displayed by rows in the grid, with the column corresponding to the cancer human tissues and cell lines used. **I.** The expression profile of the control gene for cDNAs quality. **II.** The expression pattern of the positive control for CT genes. **III.** The expression pattern for the predicted meiosis-specific gene. **IV.** The expression pattern for the potential CT antigen genes, which are the CT–CNS–selective gene, *FSCN3*; the CT–selective gene, *GAGE1* and the CT–restricted gene, *SSX2*. The dark blue boxes indicate the positive expression of the given genes, whereas the empty boxes indicate negative RT-PCR results.

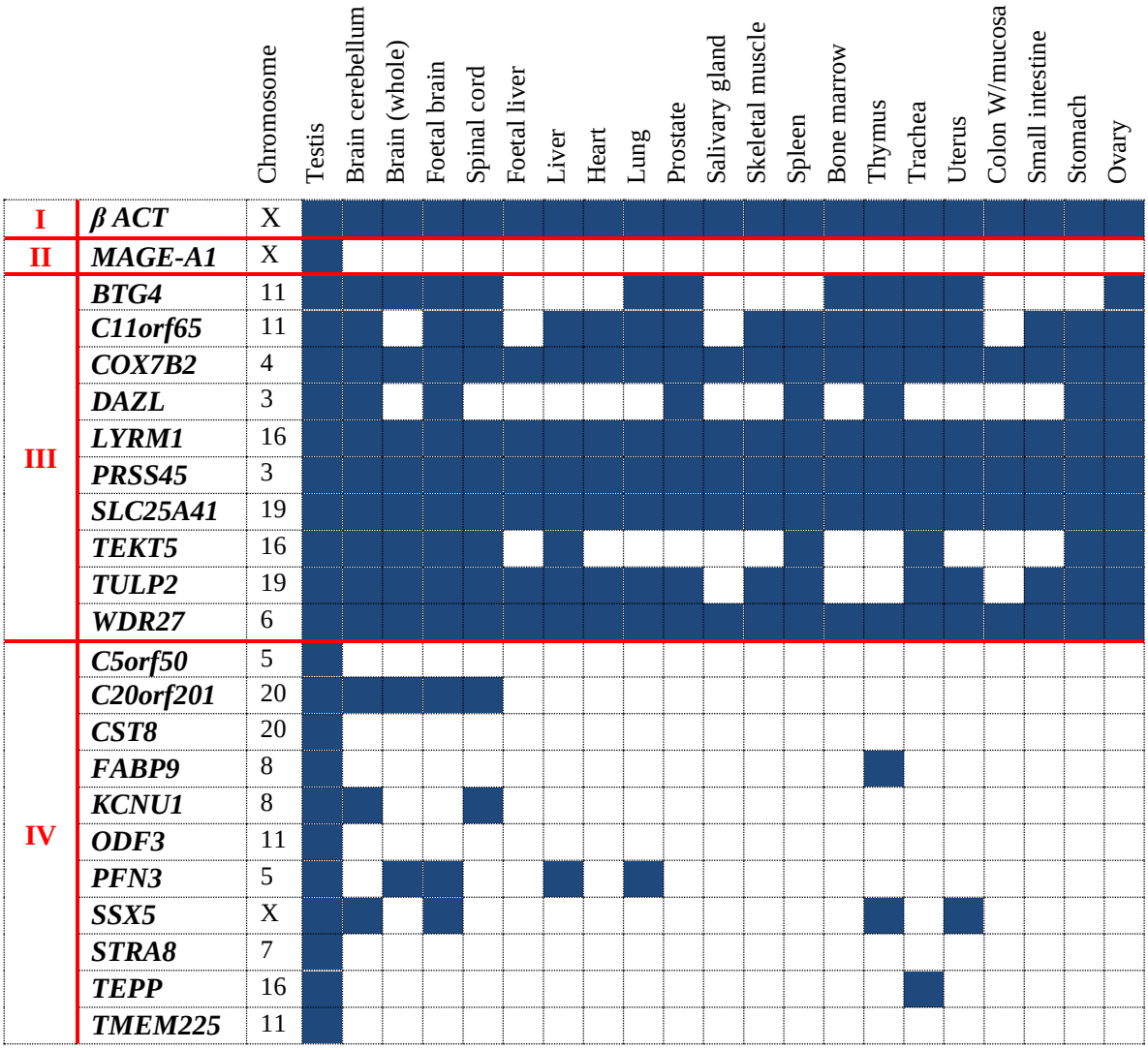


Figure 3.17 Summary of the RT-PCR results for the EST analysis genes in different normal tissues. The expression patterns of individual genes are displayed by rows in the grid, with the column corresponding to the normal human tissues used. **I.** The expression profile of the control gene for cDNAs quality. **II.** The expression pattern of the positive control for CT genes. **III.** The expression pattern for ten genes from the EST analysis lists that showed expression in multiple human normal tissues; therefore, these genes were dismissed at this stage. **IV.** The expression pattern for 11 genes: *C5orf50, C20orf201, CST8, FABP9, KCNU1, ODF3, PFN3, SSX5, STRA8, TEPP* and *TMEM225* that might to be good CT gene candidates. The chromosomal location of all genes is shown in a box in front of the gene name. The dark blue boxes indicate the positive expression of the given genes, whereas the empty boxes indicate negative RT-PCR results.

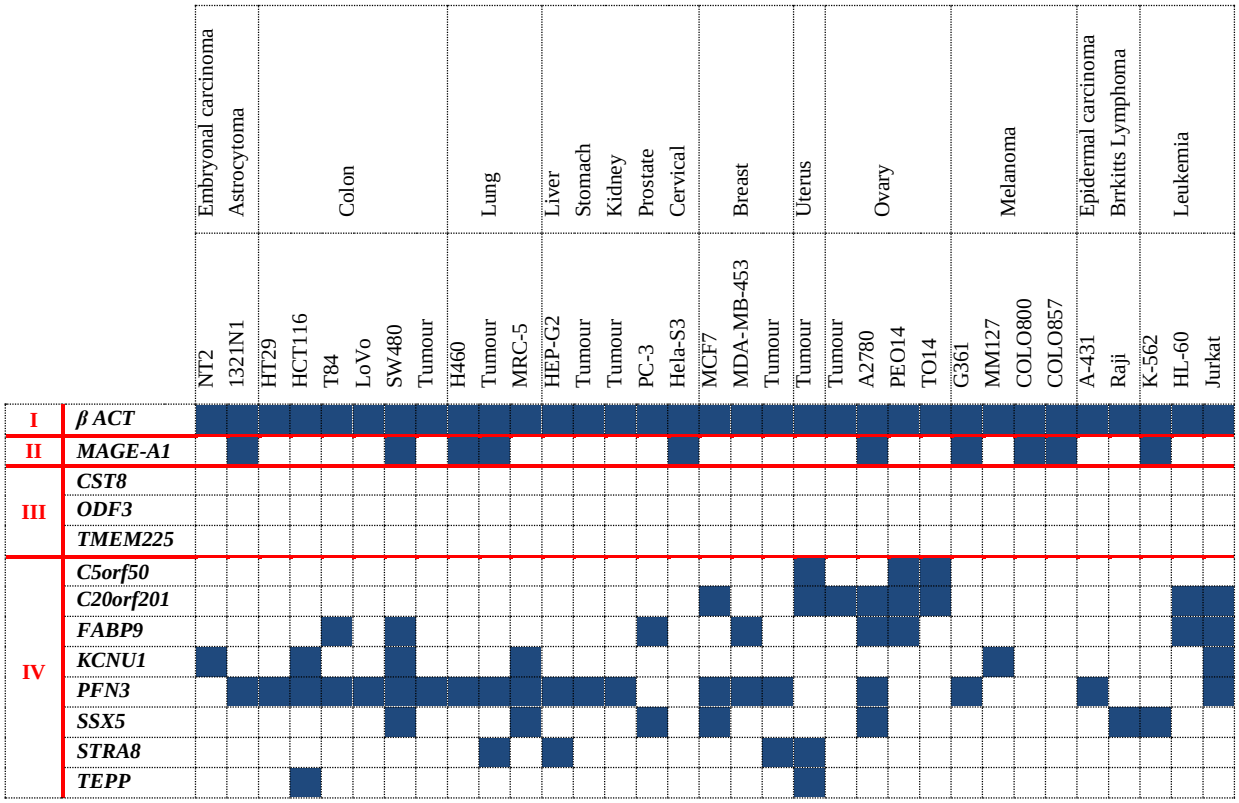


Figure 3.18 Summary of the RT-PCR results for the EST analysis genes in different cancer tissues and cell lines. The expression patterns of individual genes are displayed by rows in the grid, with the column corresponding to the cancer human tissues and cell lines used. **I.** The expression profile of the control gene for cDNAs quality. **II.** The expression pattern of the positive control for CT genes. **III.** The expression pattern for the predicted meiosis-specific genes. **IV.** The expression pattern for the potential CT antigen genes, which are the CT–restricted genes *C5orf50, FABP9* and *STRA8*; the CT–selective gene, *TEPP;* the CT–CNS-restricted, *C20orf201* and *KCNU1* and the CT–CNS-selective, *PFN3* and *SSX5*. The dark blue boxes indicate the positive expression of the given genes, whereas the empty boxes indicate negative RT-PCR results.

3.3.2 Dismissed genes from the EST and Microarray pipelines

10 genes, *BTG4, C11orf65, COX7B2*, *DAZL, LYRM1, PRSS45, SLC25A41, TEKT5, TULP2* and *WDR27* from the EST library and 7 genes, *CCNA1, C2orf69, C11orf70, C20orf195, HORMAD1, NOL4* and *ZNF558* from the Microarray library displayed expressions in various normal tissues, including testis. For this reason, these 17 genes were dismissed from further investigation, because CT genes in our strict criteria are genes whose expression were observed in non-more than two non-testis-CNS normal somatic tissues. Although the expression of the dismissed genes was observed in many normal tissues, qRT-PCR of some of these genes might show testis up-regulation, rather than testis specificity.

In a previous study, *HORMAD1* was identified as a CT gene in a screening that contained 12 normal cells, including cells from the testis and 29 cancer cell lines. Although the RT-PCR expression of *HORMAD1* was observed in different types of normal tissues, including brain, breast, colon, placenta and spleen tissues and it showed a strong expression in the testis, it is still considered to be a CT antigen gene (Chen *et al*., 2005b). However, in the present study, *HORMAD1* is shown to exhibit an expression in 13 out of 21 normal tissues; therefore, it was dismissed as a CT gene candidate. The expression differences between the studies suggests that *HORMAD1* might have been incorrectly identified as a CTA gene or the number of PCR cycles used in those studies was high (40 cycles), which may produce low levels of gene expression.

Another previous study reported that *LYRM1* mRNA expression was observed in distinct normal tissues; for example, heart, brain, placenta, lung, liver, skeletal muscle, kidney, adipose tissue and pancreas; however, among these nine normal tissues, adipose tissue expressed the highest mRNA level of *LYRM1* (Qiu *et al*., 2009). In the present work, RT-PCR expression of *LYRM1* was detected in all of the normal tissues tested, which confirms the study by Qiu *et al*. (2009).

Hanafusa *et al*., (2012) study also identified the mRNA expression of *TEKT5* using qRT-PCR analysis in various normal adult tissues, including testis and in a range of cancer types. In the previous study, *TEKT5* was identified as a novel CT gene candidate, although the mRNA expression was observed in different normal cells, including brain, colon, kidney, liver, lung, ovary, prostate, skeletal muscles, spleen and stomach in addition to testis, which had the highest expression. *TEKT5* mRNA expression was compared to one of the known CT genes, *NY-ESO-1*. *NY-ESO-1* normally demonstrated higher expression levels of gene transcript in

non-testis normal tissues compared to *TEKT5* expression. However, the RT-PCR analysis tested here revealed that the expression of *TEKT5* was found in some normal nongametogenetic tissues and confirmed via DNA sequencing, thus *TEKT5* was excluded from consideration as a novel CT antigen candidate.

The *DAZL* gene was excluded from further study aimed at identifying CT gene candidates due to its expression in the normal tissues, but our findings support that *DAZL* is a testis-CNS-selective gene. However, we followed strict criteria that led to it being excluded. Previous study indicated that DAZL is a germ cell marker in humans (Kerr *et al*., 2008) and the protein levels was observed in the adult human testes, but was absent in the prostate and testis isolated from a patient with Sertoli cells only (Ruggiu *et al*., 2000). A northern blot study showed that *DAZL* expression was observed in the testis, but only 17 normal tissues were tested (Yen *et al*., 1996). Another study suggested that *DAZL* expression was observed in both human male and female germ cells (Brekhman *et al*., 2000).

3.3.3 Testis-restricted genes from the EST and Microarray pipelines

One gene, *UBL4B* from the Microarray, and three genes, *CST8 (CRES)*, *ODF3* and *TMEM225*, from the EST analysis, were shown to have a testis-restricted expression pattern in the normal tissues and demonstrated no evidence of RT-PCR expression in the cancer tissues and/or cell lines. However, these genes were not dismissed from the gene screening, because they might be expressed in other cancer types not used in this work. A northern blot analysis showed that *CST8* was expressed at high levels in human testis, but not in any of the other nine normal tissues examined, including kidney, spleen, brain, cervix, ovary, placenta, lung, prostate and uterus (Wassler *et al*., 2002), which supports the result of this work. A recent study reported by Dianatpour *et al*., (2012) revealed that the testis specific gene, *ODF3* was not expressed in breast cancer patients or breast cancer cell lines (MCF-7 and MDA-231), but was expressed only in the testis. These results therefore provide further support to the results of this study.

A meta-analysis was conducted on the testis-restricted genes, using the CancerMA online tool (http://www.cancerma.org.uk; Feichtinger *et al*., 2012b) to evaluate any potential gene expression for these genes in cancer Microarrays data sets. The meta-analysis results did not show any statistically significant mean upregulation for *CST8*, *ODF3* or *TMEM225* in the cancer array data. However, upregulation was identified for the expression of *UBL4B* in ovarian cancer array data. Therefore, this analysis identified *UBL4B* as a potential CT-

restricted gene. The CTA control gene in this analysis was *MAGE-A1*. The meta-analysis indicated upregulation of the *MAGE-A1* gene in head and neck cancers as well as in lung and brain cancers. The results from this meta-analysis are summarised in Figure 3.19.

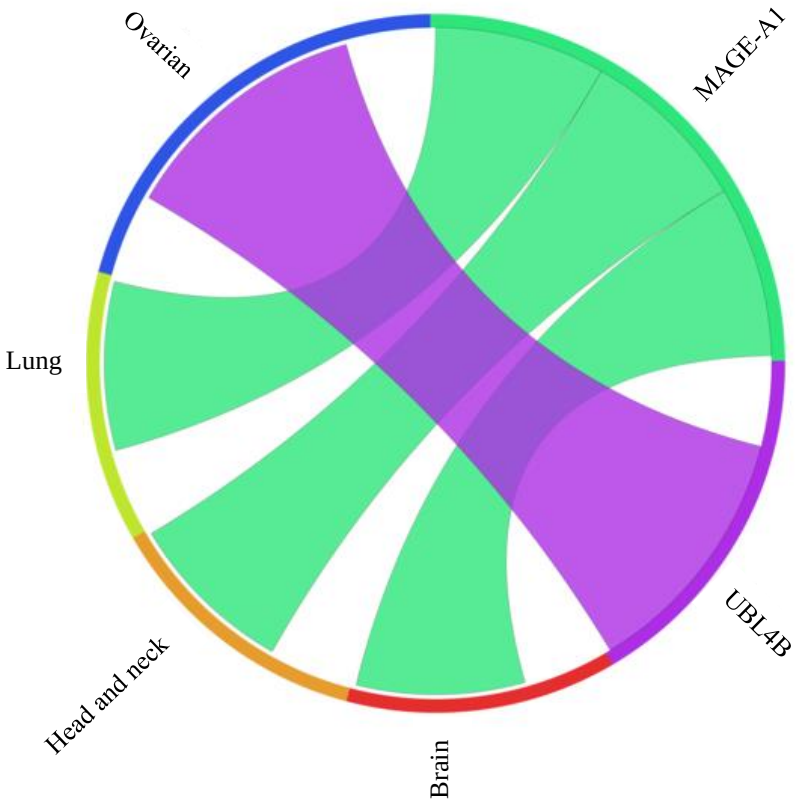


Figure 3.19 The Circos plot for the testis-restricted genes. The Circos plot depicts the meta-change in gene expression in relation to corresponding cancer types for the testis-restricted genes and the control gene. One of the four genes shows a statistically significant upregulation for that type of cancer. The cancerMA tool was used to analyse the expression patterns of these four genes using cancer array data originating from a number of combined array studies for cancer tissues *vs.* normal tissue.

3.3.4 Genes from the EST and Microarray pipelines identified as novel CT gene candidates

CT antigens represent ideal targets for cancer immunotherapy based on their restricted tissue expression. Here, the three Microarray genes *FSCN3*, *GAGE1* and *SSX2*, and the EST analysis techniques of eight genes, *C20orf201*, *C5orf50*, *FABP9*, *KCNU1*, *PFN3*, *SSX5*, *STRA8* and *TEPP*, were found to display possible candidate CT genes according to RT-PCR screening tests. According to Atanackovic *et al*., (2007), the mRNA expression of *SSX2* and *SSX5* were observed in different myeloma cell lines. In addition, it has previously been proposed that *SSX2* and *SSX5* showed a high positive rate of mRNA expression in patients with hepatocellular carcinoma (HCC) and no evidence of expression for both genes in the normal tissue samples adjacent to the tumour tissues (Wu *et al*., 2006). In one recent publication, *GAGE1* and *SSX2* have been determined as positive controls for the known X-CT genes. Furthermore, *STRA8* and *FABP9* were strongly expressed in healthy testis as well as in different malignant cells. It is also reported that the *PFN3* and *TEPP* genes were classified as CT-CNS-selective and CT-selective, respectively (Feichtinger *et al*., 2012a).

Meta-analysis of gene expression profiles was performed on the novel CTA gene candidates (11 genes), using the CancerMA online tool (http://www.cancerma.org.uk; Feichtinger *et al*., 2012b), to assess any potential gene expression for these genes in cancer microarray data sets. The meta-analysis results did not show a statistically significant mean upregulation for *C5orf50*, *FABP9*, *KCNU1*, *PFN3*, *STRA8* and *TEPP* in the cancer array data. Conversely, upregulation was identified for the expression of five of these genes (*C20orf201*, *FSCN3*, *GAGE1*, *SSX2* and *SSX5*). The upregulation of *C20orf201* was observed in ovarian cancer and *FSCN3* was upregulated in ovarian and brain cancers. The upregulation of *GAGE1* was detected in ovarian, lung, breast and brain cancers. The meta-analysis results also identified upregulation in ovarian, lung, brain and adrenal cancers for *SSX2* and in ovarian, leukaemia and brain cancers for *SSX5*. Interestingly, in the present study, the RT-PCR results for *GAGE1* and *SSX2* did not indicate any expression in the three types of lung cancers or in the four types of ovarian cancers tested here, respectively. Therefore, both genes require more studies using more lung cancer types for *GAGE1* and more ovarian cancer samples for *SSX2*. The results from this meta-analysis are summarised in Figure 3.20.

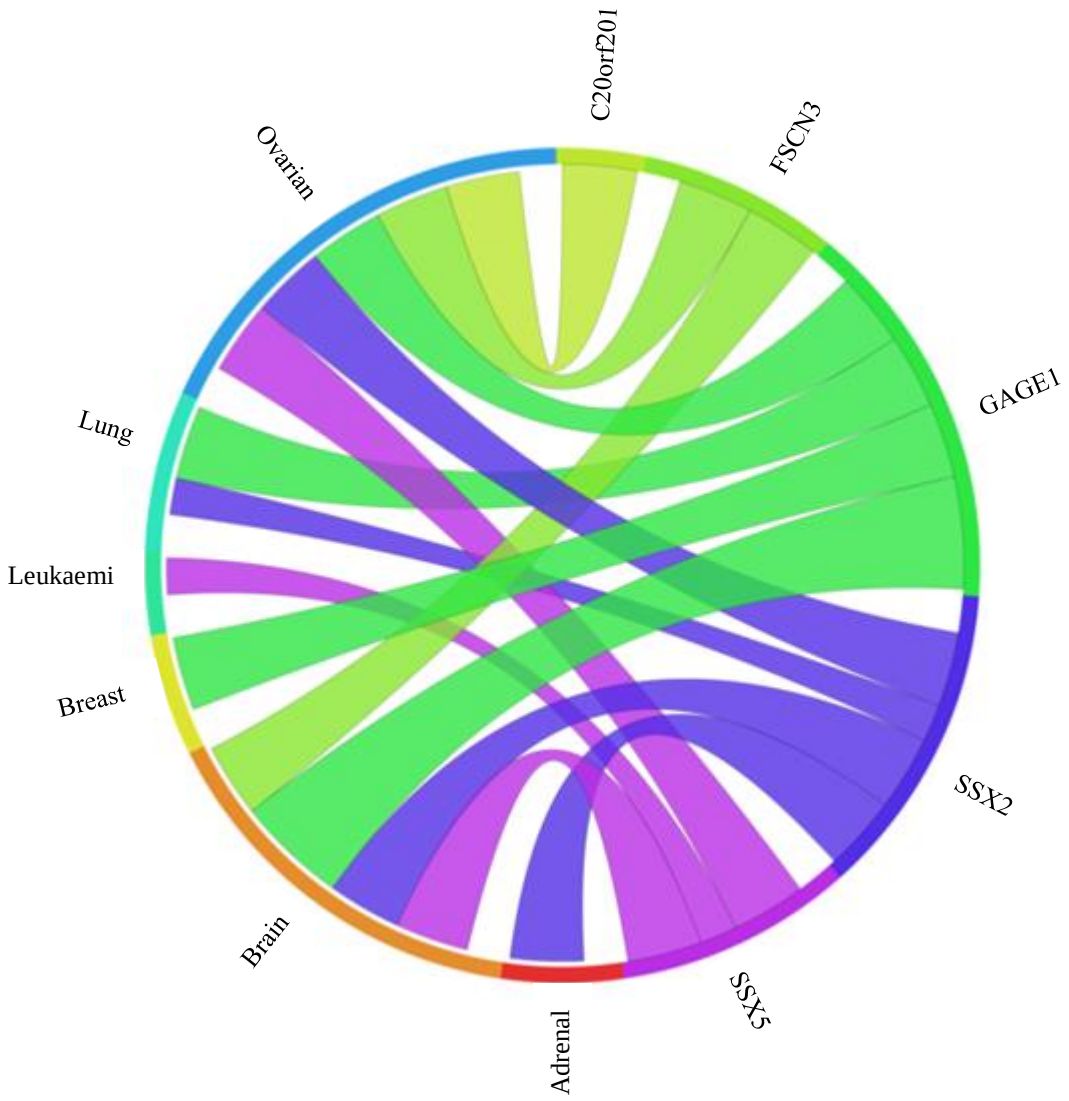


Figure 3.20 The Circos plot for the novel CTA genes. The Circos plot depicts the meta-change in gene expression in relation to corresponding cancer types for the novel CTA genes. Five of the eleven genes show a statistically significant upregulation for that cancer types. The cancerMA tool was used to analyse the expression patterns of these eleven genes using cancer array data originating from a number of combined array studies for cancer tissues *vs.* normal tissue.

The genes identified as novel CT genes could play a role in drug resistance; for example, a recent study revealed that over-expression of the testis specific gene, *CYCLON*, led to poor survival of lymphoma patients treated with a combination of CHOP chemotherapy with the monoclonal antibody Rituximab, but did not affect survival of patients treated with chemotherapy alone. These findings therefore indicated that *CYCLON* expression results in resistance to Rituximab in lymphoma patients (Knapp, 2013).

Germline genes can be used as prognostic markers for cancer; for example, Rousseaux and his co-authors investigated the expression of testis-specific/placenta-specific (TS/PS) genes in 1776 solid human cancers isolated from 14 different cancer types. A total of 506 tissue-restricted genes were identified whose expression profiles were generally restricted to one tissue. These genes were divided into 439 testis/germline genes and 67 placenta-restricted genes. The expressions of hundreds of these 506 genes were aberrantly observed in all cancer types, indicating that these genes represent a promising source of cancer biomarkers. The genes were studied further to determine whether the TS/PS genes represented useful cancer biomarker in a specific type; lung cancer. A group of 293 patients with lung cancer at all stages of the disease were examined, including 152 lung cancer patients with early-stage cancer. The TC/PC genes were expressed to a greater extent in all 293 patients, including the early lung cancer patients, than in non-tumour lung (NL) samples. This study ultimately identified 26 of the TS/PS genes whose expression correlated with lower survival in lung cancer patients. Therefore, these genes are associated with poor prognosis, independent of cancer stage (Rousseaux *et al*., 2013).

3.4 Conclusion

In this work, the validation of the 32 genes using RT-PCR analysis determined that six genes showed a CT-restricted/selective expression pattern and five displayed CT-CNS-restricted/selective expression patterns. These CT candidate genes were expressed in a wide variety of cancer cells and they were limited in the testis, or with no more than two other additional non-testis-CNS normal tissues. Therefore, these CT genes represent promising candidates due to their expression patterns in distinct types of cancer. They may be used as potential targets for the development of cancer vaccines and/or diagnostic markers for several types of cancer. However, further investigation is needed to establish the protein products of these excellent CT candidate genes in a range of normal and cancer tissues.

Chapter 4.0: Analysis of the function of the human *STRA8* gene

4.1 Introduction

Retinoic acid (RA) is required for the switch from mitosis to meiosis. RA stimulates meiosis by activating the RA receptor-dependant (refer to Section 1.3.4) transcription of meiotic inducers, including *Stra8* (Bowels and Koopman, 2007). The *Stra8* gene (Stimulated by Retinoic Acid gene 8) was first identified by a subtractive hybridization cloning strategy that was used to identify retinoic acid-responsive genes in mouse P19 embryonal carcinoma (EC) cells in both the presence and absence of retinoic acid to identify RA-responsive genes in P19 cells. *Stra8* was one of the 50 genes identified in this screen, 40 of which were novel (Bouillet *et al*., 1995). These genes were collectively designated as those stimulated by RA (Stra) (Miyamoto *et al*., 2002; Oulad-Abdelghani *et al*., 1996a). It was shown that *Stra8* expression is restricted to the early premeiotic spermatogonia, and the protein encoded by the gene is mostly cytoplasmic. In addition, the levels of *Stra8* mRNA and protein were upregulated after treatment of the P19 cells by both all-*trans* and 9-*cis* retinoic acids (Oulad-Abdelghani *et al*. 1996a). According to previous studies, *Stra8* gene function is required for germ cells to enter into meiosis (Anderson *et al*., 2008; Baltus *et al*., 2006).

Stra8 encodes a glutamic acid-rich protein with no homology to other known proteins. Two-dimensional gel electrophoresis of RA-stimulated cytosolic cell extracts demonstrated that *Stra8* has 9 different forms of phosphorylation (Oulad-Abdelghani *et al*., 1996b). Studies have suggested a possible Stra8 function in the nucleus, with shuttling of the protein between the nucleus and cytoplasm (Tedesco *et al*., 2009). Furthermore, Stra8 may play a role in binding double-strand DNA (Choi *et al*., 2010; Tedesco *et al*., 2009) and have a transcription activation function *in vitro* (Choi *et al*., 2010). In adult mice, Stra8 protein was restricted in spermatogonia and spermatocytes; the level of Stra8 protein was increased after the addition of RA to the adult spermatogonia (Zhou *et al*., 2008).

Although the molecular function of Stra8 is still unknown, a range of studies have demonstrated that Stra8 is required for successful meiotic initiation in both male and female germ lines (Anderson *et al*., 2008; Mark *et al*., 2008) and in female embryos (Baltus *et al*., 2006). Some of these studies showed that *Stra8* deficiency leads to infertility in mice with no

evidence of mature gametes in both postnatal testes (Anderson *et al*., 2008; Mark *et al*., 2008) and foetal ovaries (Baltus *et al*., 2006). In *Stra8*-deficient foetal ovaries, meiosis was stalled just prior to prophase I entry. However, in *Stra8*-deficient postnatal testes, data were not consistent regarding whether the disruption occurs before or after initiation of meiosis (Baltus *et al*., 2006; Mark *et al*., 2008).

Studies using Vitamin A showed that RA plays a critical role during meiosis in spermatogenesis in postnatal testes (Li *et al*., 2011) and in developing ovaries (Li and Clagett-Dame, 2009). A recent study demonstrated that the RA is critical for spermatogenesis as well as for meiotic entry (Raverdeau *et al*., 2012). It is remarkable that there is a clear connection between the presence of RA and the up-regulation of *Stra8* in germ cells in both adult and foetal meiosis (Feng *et al*., 2014). A previous study indicated that Dmrt1 is also a regulator of meiotic entry. Studies demonstrated that knockout of Dmrt1 in mice leads to the reduction of *Stra8* expression in the foetal ovary and the number of ovarian follicles in the mature ovary is also decreased (Krentz *et al*., 2011). However, knockout of Dmrt1 in postnatal testes demonstrated the opposite effect with increasing *Stra8* expression (Matson *et al*., 2011).

The human *STRA8* gene is localised on chromosome 7 and contains 9 exons separated by 8 introns of different lengths. The translational start of the gene is localised in exon 1 and stop codon TAA in exon 9 (Figure 4.1). The *STRA8* cDNA containing the whole open reading frame (ORF) is 993 bp (GenBank accession number AF513502), encoding a 330 amino acid protein.

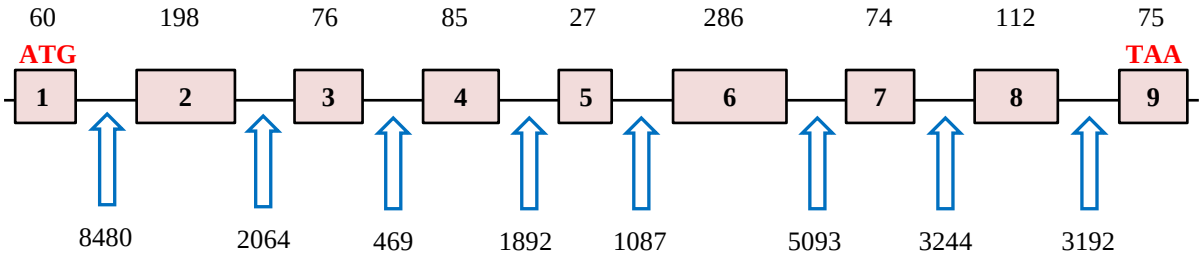


Figure 4.1 Exon and intron structure of the human *STRA8* gene. Schematics show the structure of the *STRA8* gene coding and non-coding regions. The boxes and arrows represent exons and introns, respectively. The lengths of each exon and intron are given above the box in the exons and below the arrow in the introns. Approximate positions of the start codon (ATG) and the stop codon (TAA) are marked in red.

4.2 Results

4.2.1 Cloning of *STRA8* cDNA into the pcDNA5/FRT/TO vector for over expression

The purpose of cloning *STRA8* cDNA was to address the protein function of this gene in cancer cells. The mammalian expression vector, pcDNA5/FRT/TO (Figure 4.2), was used for the cloning of the *STRA8* ORF with and without adding the Kozak consensus sequence. The Kozak sequence is added to *STRA8* cDNA sequences to stimulate the translation process (Kozak, 1991). Constructs with and without Kozak sequences were generated to potentially produce cell lines producing different amounts of STRA8 protein. *Not*I and *Xho*I restriction enzymes were used for cloning of *STRA8*–Kozak and *STRA8*+Kozak, respectively. Both enzymes were chosen because they are available in the cloning plasmid at the multiple cloning site (MCS) and do not appear in *STRA8* cDNA sequences.

For *STRA8* cloning without the Kozak sequence, the cDNA coding clone was cut from the pCMV6-XL5::*STRA8* vector with the *Not*I restriction enzyme. To evaluate vector digestion, gel electrophoresis was performed for both the undigested and digested pCMV6-XL5::*STRA8* vectors, which gave three molecules (Figure 4.3A left) and two bands (Figure 4.3A right), respectively. The two bands indicate a pCMV6-XL5 vector of about 4482 bp and a *STRA8* cDNA fragment of about 1209 bp. The insert fragment, verified by DNA sequencing, was subcloned into the *Not*I site of mammalian expression vector, pcDNA5/FRT/TO, to construct the plasmid termed pcDNA5/FRT/TO::*STRA8*–Kozak. Four molecules were observed in the undigested pcDNA5/FRT/TO vector (Figure 4.3B) and one band of approximately 5137 bp was observed, as expected, in the digested vector (Figure 4.3C left). After purification, the gel also displayed a *STRA8*–Kozak cDNA fragment within the expected size of 1209 bp (Figure 4.3C right).

In contrast, the cloning of *STRA8*+Kozak was initiated by amplification of *STRA8* cDNA from the pCMV6-XL5::*STRA8* vector using STRA8 F4+Kozak and R4 primers. The amplified fragment was 1212 bp, based on DNA sequencing, which confirms the cDNA sequences of the gene of interest. The purified PCR insert was subcloned into the pcDNA5/FRT/TO vector after digestion with the same restriction enzyme, *Xho*I. The vector was analysed on agarose gel before and after *Xho*I digestion in order to test the digestion efficiency of the vector, which yielded four molecules in the undigested vector (Figure 4.4A) and a single fragment of about 5137 bp in the digested vector (Figure 4.4B left). The insert

was also separated on the gel after digestion and purification to show a single band of about 1212 bp, based on DNA sequencing (Figure 4.4B right).

The purified *STRA8* inserts (with and without Kozak) were ligated into the purified digested pcDNA5/FRT/TO plasmid using quick T4 DNA ligase. The recombinant plasmids were transformed into competent *E. coli* DH5α. For analysing the clones, PCR screening was performed on 17 randomly selected colonies. Primers internally positioned within the *STRA8* gene, *STRA8* F1 and R1, were used in the PCR screening. The PCR results indicated that all the *E. coli* colonies had the gene of interest (*STRA8*) without Kozak (Figures 4.5) or with Kozak (Figure 4.6). The PCR primers were tested using testis cDNA. pcDNA5/FRT/TO::*STRA8*+Kozak and pcDNA5/FRT/TO::*STRA8*–Kozak plasmids were isolated from the *E. coli*. The presence of *STRA8* inserts (+Kozak or –Kozak) was verified by restriction enzyme digestion, which cut once in the vector and once in the insert. Therefore, 6 samples of purified DNA from the 17 pcDNA5/FRT/TO::*STRA8*–Kozak plasmids were digested with *Not*I, while 2 of the 17 pcDNA5/FRT/TO::*STRA8*+Kozak plasmids were digested with *Xho*I. Two fragments were observed in each digested plasmid. The upper fragment indicated the expected vector size of 5137 bp, while the lower bands indicated fragments of 1209 bp and 1212 bp for the *STRA8*–Kozak (Figure 4.7) and *STRA8*+Kozak, respectively (Figure 4.8).

Two plasmids were sent for DNA sequencing using CMV forward and BGH reverse primers to confirm that the inserts were subcloned in the correct orientation for expression and contained an ATG initiation codon and a stop codon. The sequences obtained were compared to the original sequence through the NCB1 (BLAST). They were 100% identical to *STRA8* without mutation for the whole length of the *STRA8* gene. As soon as the sequences were confirmed, *E. coli* glycerol stocks were made and stored in the -80°C freezer.

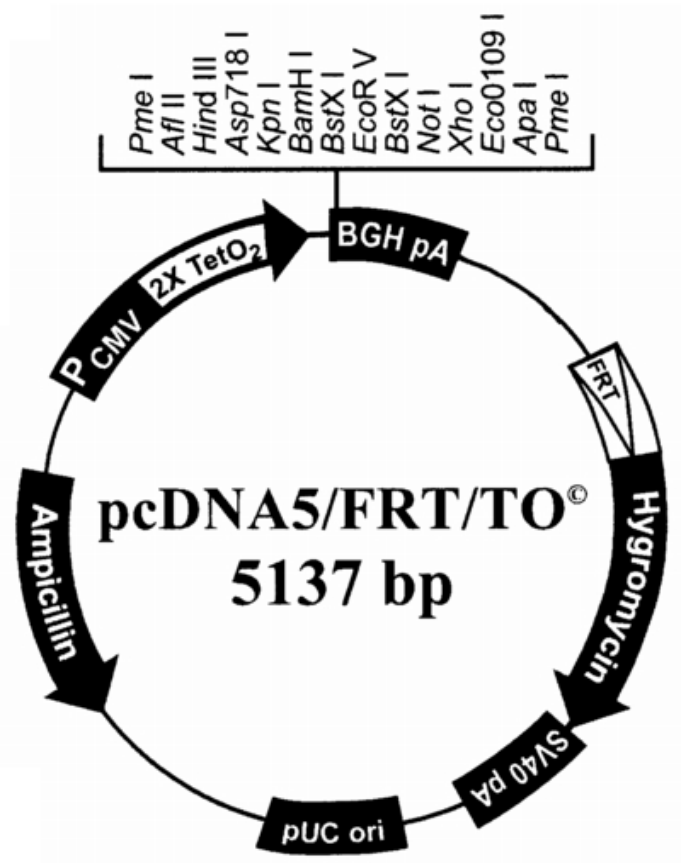


Figure 4.2 Map of pcDNA5/FRT/TO plasmid. pcDNA5/FRT/TO is a 5.1 kb inducible expression plasmid designed for use with the Flp-In T-REx system and contains the following components: A hybrid human cytomegalovirus (CMV)/TetO2 promoter for high-level, tetracycline-regulated expression of *STRA8* in a Flp-In T-REx-293 cell line; multiple cloning site (MCS) that contains *Xho*I and *Not*I restriction sites to mediate cloning of the *STRA8* gene with and without Kozak, respectively; an FLP Recombination Target (FRT) site for Flp recombinase to facilitate integration of the vector into the Flp-In T-REx-293; and a hygromycin gene to select stable Flp-In T-REx-293 cell lines (Adapted from Invitrogen; www.invitrogen.com).

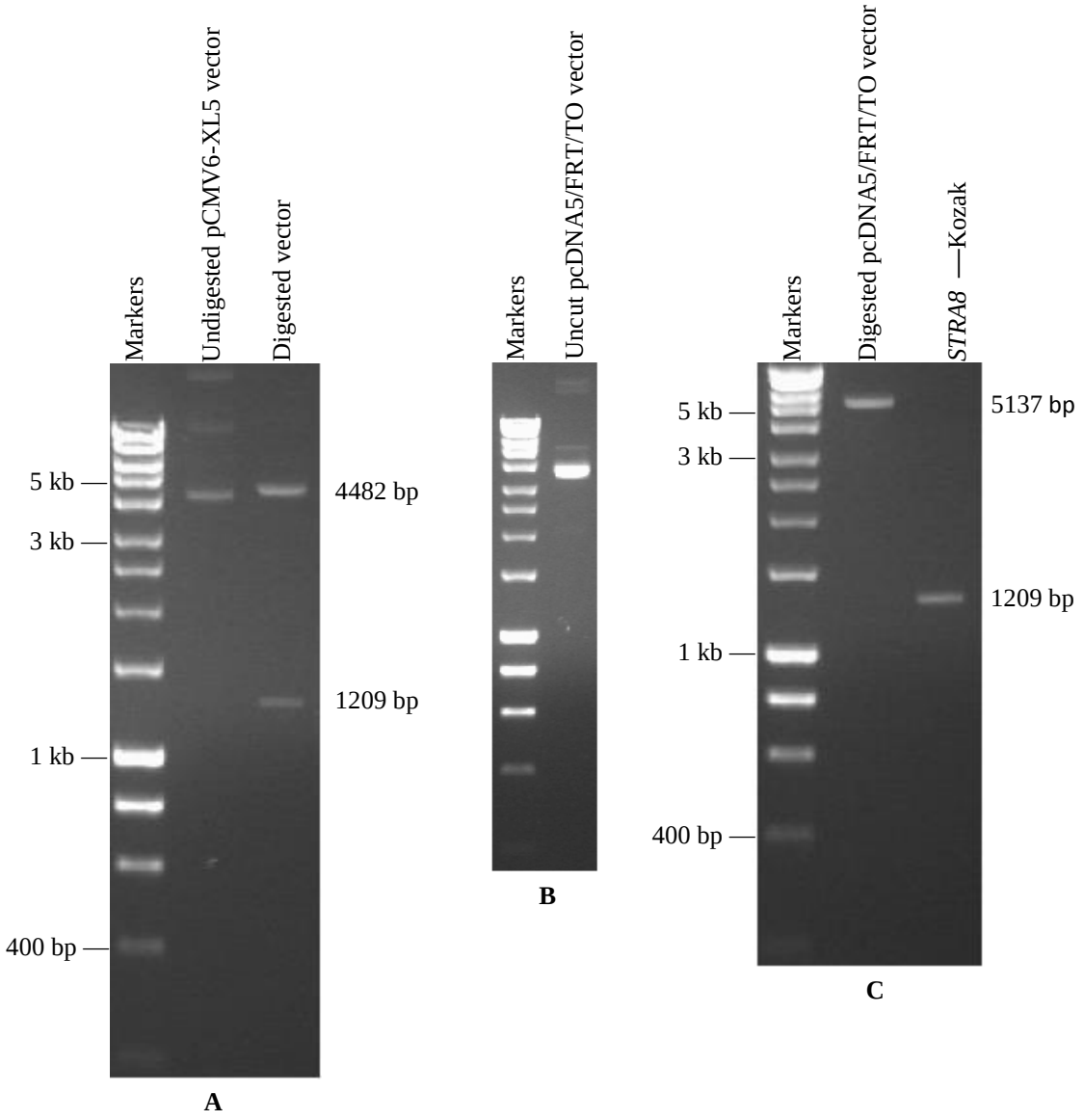


Figure 4.3 Generation of a *STRA8* cDNA clone in the pCMV6-XL5::*STRA8* vector. Agarose gel picture (panel **(A)**) showing undigested pCMV6-XL5::*STRA8* mammalian expression vector (left), as a control to compare to the digestion vector (right), which was digested by *Not*I restriction enzyme to extract *STRA8* cDNA of 1209 bp. Panel **(B)** shows undigested pcDNA5/FRT/TO, a mammalian expression vector. Panel **(C)** (left) shows the pcDNA5/FRT/TO vector after digestion by the *Not*I restriction enzyme and purification. The enzyme linearizes the 5137 bp plasmid into one single fragment. The right side of the gel **(C)** displays the *STRA8* sequence after isolation from the pCMV6-XL5::*STRA8* vector and purification.

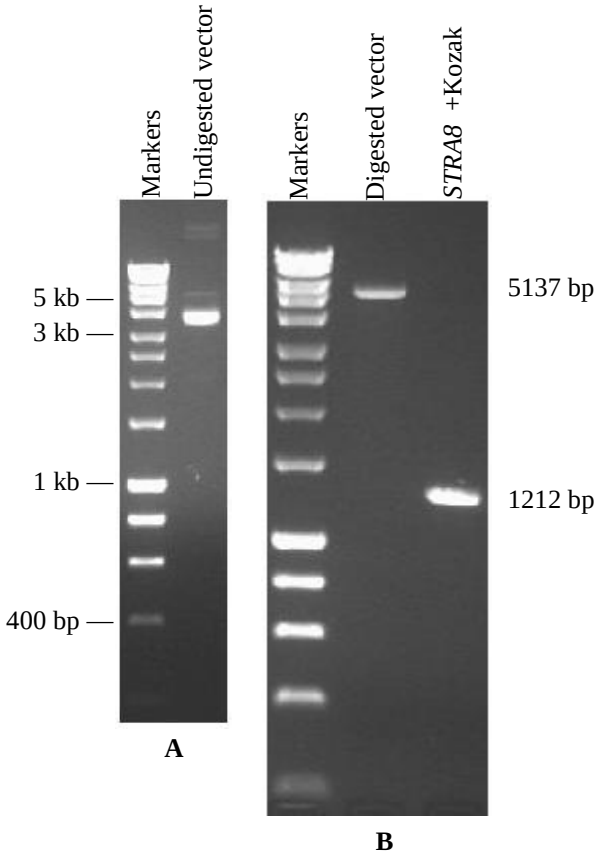


Figure 4.4 Amplification of *STRA8*+Kozak from the recombinant pCMV6-XL5::*STRA8* vector. Agarose gel picture showing undigested pcDNA5/FRT/TO, mammalian expression vector, as a control to compare to the digestion vector (panel **(A)**). Panel **(B)** (left) shows the pcDNA5/FRT/TO vector after digestion by *Xho*I restriction enzyme and purification. The enzyme linearizes the 5137 bp plasmid into one single fragment. The right side of gel B displays the amplification of the *STRA8* cDNAs +Kozak sequence from the pCMV6-XL5::*STRA8* vector, which was cloned with the full length of the open reading frame of *STRA8*. This band was digested with the same restriction enzyme, *Xho*I, and then underwent purification.

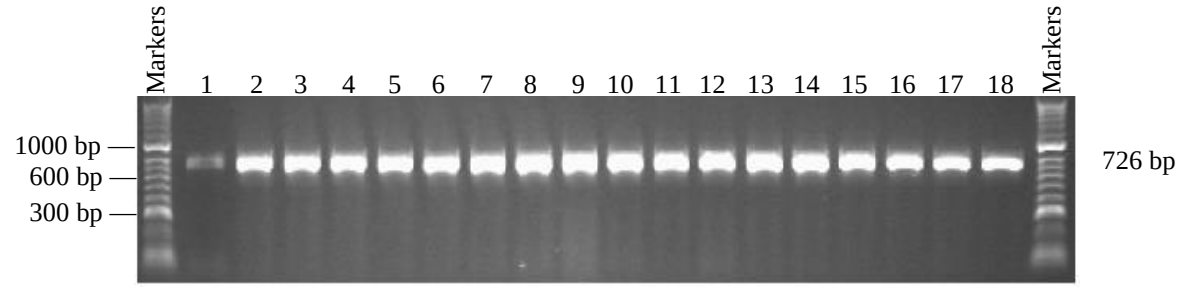


Figure 4.5 PCR screening of colonies for the cloning of *STRA8*. Agarose gel picture showing PCR screening of *E. coli* colonies for the cloning of *STRA8* into the pcDNA5/FRT/TO using an internal primer for the *STRA8* gene (F1, R1). Lane 1 shows the presence of *STRA8* in the testis to confirm that the primers work. The results show that all the samples in lanes 2-18 contain a *STRA8* insert of the expected size, 726 bp.

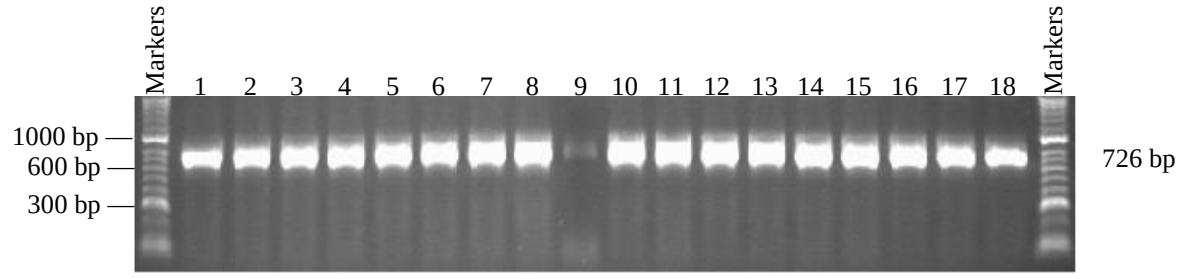


Figure 4.6 PCR screening of colonies for the cloning of *STRA8*. Agarose gel picture showing PCR screening of *E. coli* colonies for the cloning of *STRA8*+Kozak into the pcDNA5/FRT/TO using an internal primer for the *STRA8* gene (F1, R1). Lane 1 shows the presence of *STRA8* in pCMV6-XL5::*STRA8* to confirm that the primers work. The results show that all the samples in lanes 2-18 have a *STRA8* insert.

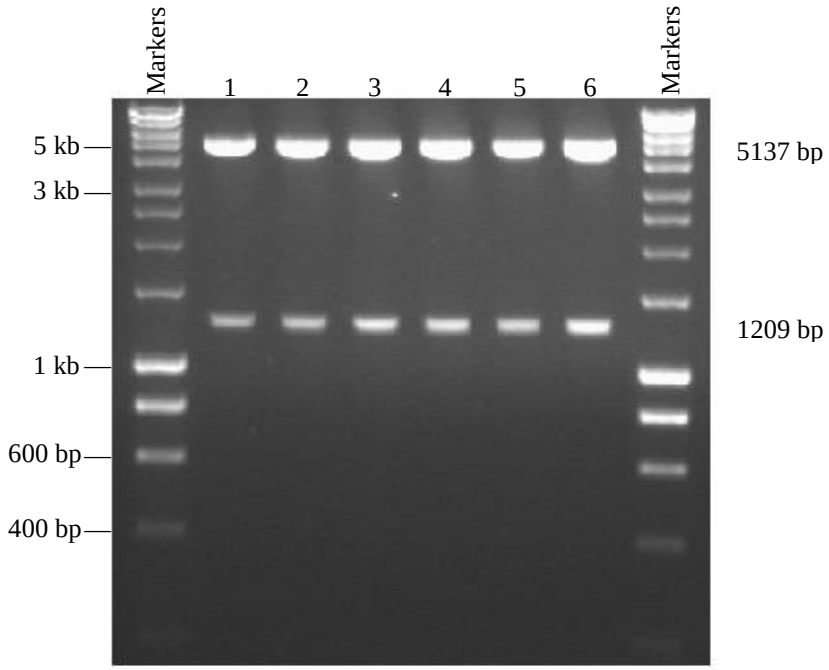


Figure 4.7 Digestion of recombinant plasmids. Agarose gel picture showing the digestion of the recombinant purified pcDNA5/FRT/TO::*STRA8*-Kozak vectors by the *Not*I restriction enzyme. Two bands are formed for all the recombinants. The upper bands show a pcDNA5/FRT/TO vector of 5137 bp, while the lower bands show the full open reading frame of the *STRA8* gene of 1209 bp in size. All the minipreps in lanes 1-6 show successful cloning of *STRA8* cDNA.

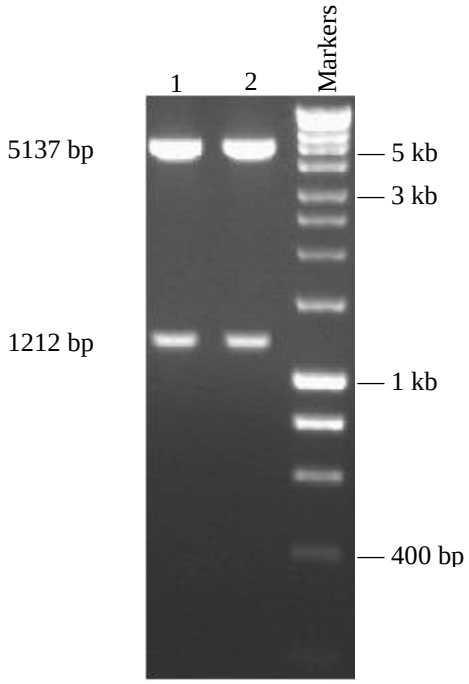


Figure 4.8 Digestion of recombinant plasmids. Agarose gel picture showing the digestion of purified recombinant plasmids (pcDNA5/FRT/TO::*STRA8*+Kozak) by the *Xho*I restriction enzyme. Two bands are formed for all recombinants. The upper bands show a pcDNA5/FRT/TO vector of approximately 5137 bp in size, while the lower bands show the full open reading frame of *STRA8* gene +Kozak at the size of 1212 bp. The plasmids in lanes 1-2 contain bands commensurate with the expected size for the cloning of *STRA8*+Kozak.

4.2.2 Integration of pcDNA5/FRT/TO containing *STRA8* cDNA into the Flp-In T-REx-293 cell line

4.2.2.1 Generation of a stable cell line

The Flp-In T-REx-293 cell line is derived from the HEK 293 cell (human embryonic kidney) and has already been transfected with two different plasmids (Figure 4.9). The first plasmid is pFRT/*LacZ*eo, designed for use with the Flp-In system. This plasmid contains a single integrated FRT site. The FRT has been inserted just downstream of the ATG initiation codon of the *LacZ-Zeocin* fusion gene. Cell lines containing this plasmid stably express the *LacZ-Zeocin* fusion gene under the control of a P_{SV40} promoter. Consequently, this inserted gene was used as a resistant marker for selection. The second plasmid, pcDNA6/TR, is designed for use with the T-REx system for tetracycline-regulated expression of the gene of interest. This plasmid expresses high levels of tetracycline (Tet) repressor under the control of human cytomegalovirus (CMV) promoter. The vector also contains the blasticidin resistance gene. Based on the insertion of the two plasmids above, the non-transfected Flp-In T-REx-293 has three major features: it expresses β-galactosidase activity, it confers resistance to zeocin and blasticidin, and it is sensitive to hygromycin.

Insertion of *STRA8* into the Flp-In T-REx-293 cell line also required the pOG44 vector (Figure 4.10), which encodes the *FLP* gene. The *FLP* gene expresses Flp recombinase which is required for integration of the pcDNA5/FRT/TO containing *STRA8* with and without Kozak via the Flp recombination target (FRT) into the genome of the Flp-In T-REx-293 cell line (Buchholz *et al*., 1996). Therefore, in this study, the cell line was co-transfected with a DNA mixture containing *STRA8* cDNA (+Kozak or –Kozak) in pcDNA5/FRT/TO and pOG44 at a ratio of 9:1 (pOG44: pcDNA5/FRT/TO). A negative control was set up via transfection of the Flp-In T-REx-293 cell line with pOG44 and pcDNA5/FRT/TO alone without the gene of interest. In brief, a non-transfected Flp-In T-REx-293 cell line was grown in medium supplemented with 200 µg/ml zeocin and 10 µg/ml blasticidin. 48 hours following transfection, the cells were refreshed with media containing 100 µg/ml of hygromycin instead of zeocin. Indeed, if the gene of interest, *STRA8*, integrated successfully at the FRT site of the cell line genome, the cells should be resistant to hygromycin, sensitive to zeocin, and lack any β-galactosidase activity. The majority of the cells did not survive due to the presence of hygromycin, which was used for the selection of a stable cell line. After pre-screening with a light microscope, single hygromycin resistant clones were selected for each construct (Figure 4.11).

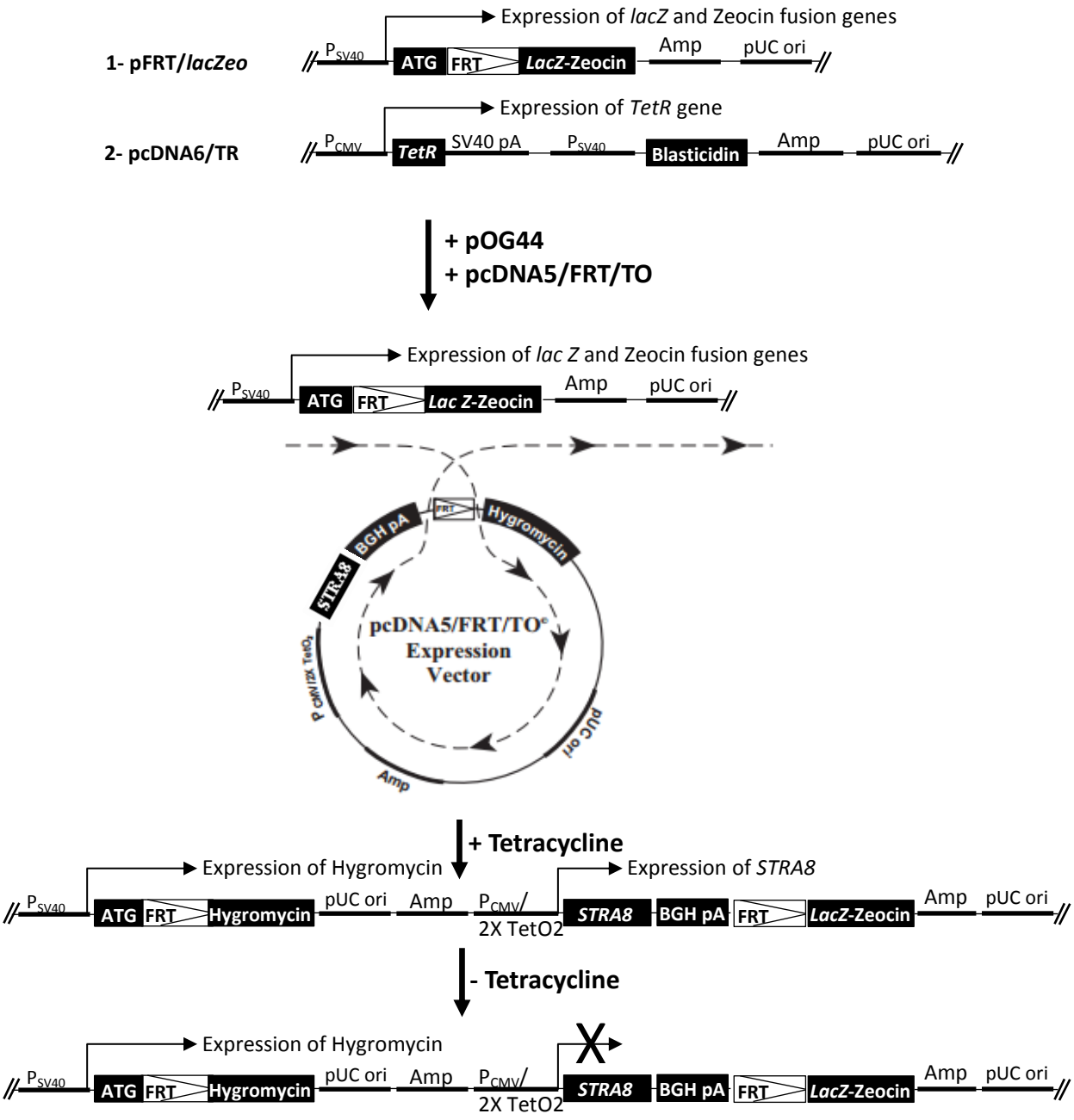


Figure 4.9 Diagram of the Flp-In T-REx-293 system. The figure above shows the major components of the system to generate a stable, inducible mammalian expression cell line by Flp recombinase-mediated DNA recombination at the FRT site. Expression of the *STRA8* gene from a specific genomic location is induced by the addition of tetracycline (Adapted from Invitrogen with some modification).

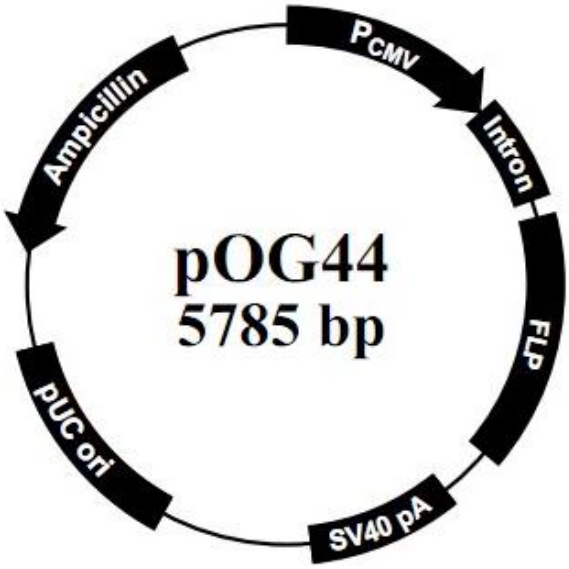


Figure 4.10 Map of the pOG44 plasmid. pOG44 is a 5.8 kb Flp recombinase expression vector designed for use with the Flp-In T-REx system and contains the following components: an *FLP* gene encoding Flp recombinase to facilitate integration of the pcDNA5/FRT/TO plasmid containing *STRA8* gene into the genome of the Flp-In T-Rex-293 cell line; a synthetic intron to enhance expression of the *FLP* gene; and a human cytomegalovirus (CMV) immediate-early enhancer/promoter for high-level expression of Flp recombinase (Adapted from Invitrogen; www.invitrogen.com**)**.

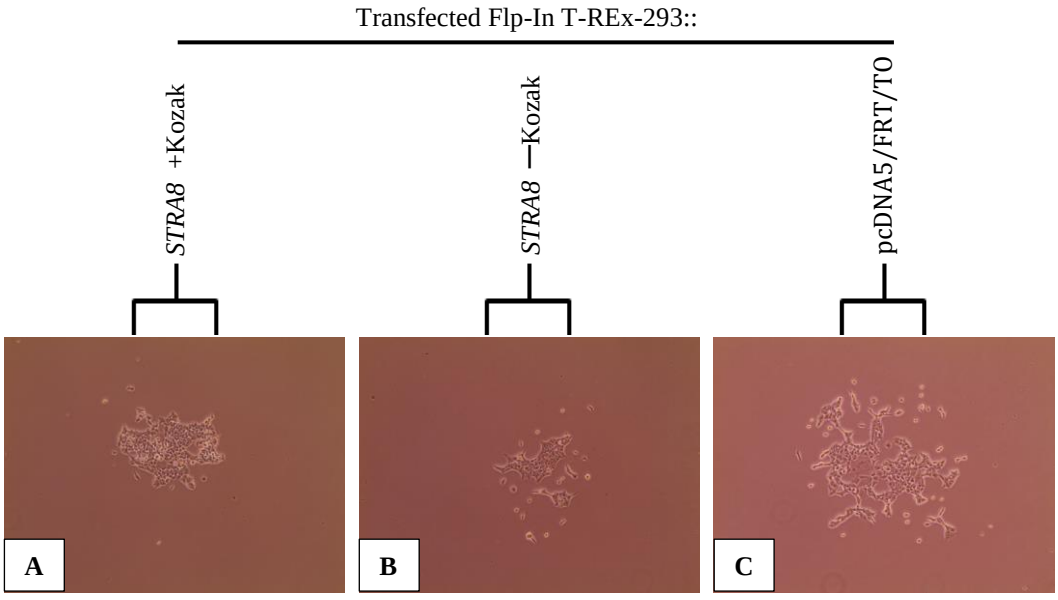


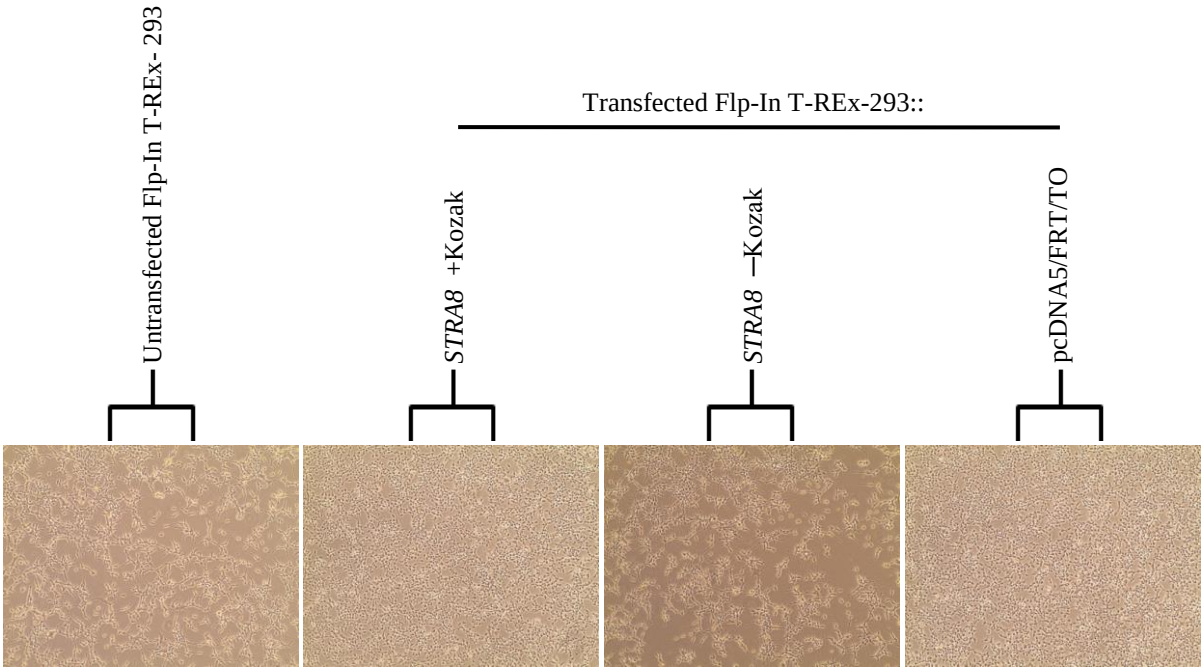
Figure 4.11 Examples of individual hygromycin resistant colonies. Panels **(A** and **B)** show a single colony still growing after hygromycin treatment; these cells were integrated with *STRA8*+Kozak and with *STRA8*–Kozak, respectively. Panel **(C)** shows cells integrated with pcDNA5/FRT/TO alone, without the target gene; these cells served as a negative control.

4.2.2.2 Evaluation of different integrant Flp-In T-REx-293 cell lines

Two different methods were used to analyse the successful integration of pcDNA5/FRT/TO::*STRA8* into the genome of the Flp-In T-REx-293 cell line at the FRT site. The first method was screening for β-galactosidase activity. Confluent cells were analysed using β-Gal staining kit in media containing hygromycin or zeocin for the transfected and non-transfected cells, respectively. The *lacZ* gene product, β-galactosidase, can be visualised by β-Gal staining of the cell. Such activity was determined in triplicate for each cell line and light microscope images were obtained. In the positive control for *lacZ*-staining (untransfected Flp-In T-REx-293 cell line), all cells were blue, indicating active *LacZ* expression. In contrast, none of the cells in the clonal cell lines showed blue cells, indicating that the insertion of the *STRA8* constructs were successful (Figure 4.12).

The second method used was PCR screening of the integrant Flp-In T-REx-293. Genomic DNA was isolated from the non-transfected and transfected Flp-In T-REx-293 cells. *β*-actin was used as a positive control for the genomic DNA samples. PCR products were found in all of the transfected cells that integrated with the pcDNA5/FRT/TO::*STRA8*+Kozak, the pcDNA5/FRT/TO::*STRA8*–Kozak or the empty pcDNA5/FRT/TO; however, the non-transfected cells did not display any PCR products when using primers located in the PSV40 as a forward primer and in the hygromycin gene as a reverse primer. Primers located in the *STRA8* gene (forward) and in the *LacZ-Zeocin* gene (reverse) were also used. The PCR results using STRA8 and LacZ-Zeocin primers showed products only for the Flp-In T-REx-293 that had been integrated with pcDNA5/FRT/TO::*STRA8*+Kozak or with pcDNA5/FRT/TO::*STRA8*–Kozak (Figure 4.13).

The results of both methods (β-galactosidase activity and PCR screening) suggest that the *STRA8* gene was correctly inserted into the genomic DNA of the Flp-In T-REx-293 cells.



A- Before staining

B- After staining

Figure 4.12 Assay for β-galactosidase activity. These images show the untransfected Flp-In T-REx-293 alone exhibiting β-galactosidase (panel **(B)**); the transfected cells that integrated with different inserts (*STRA8*+Kozak or *STRA8*–Kozak or pcDNA5/FRT/TO) are not stained.

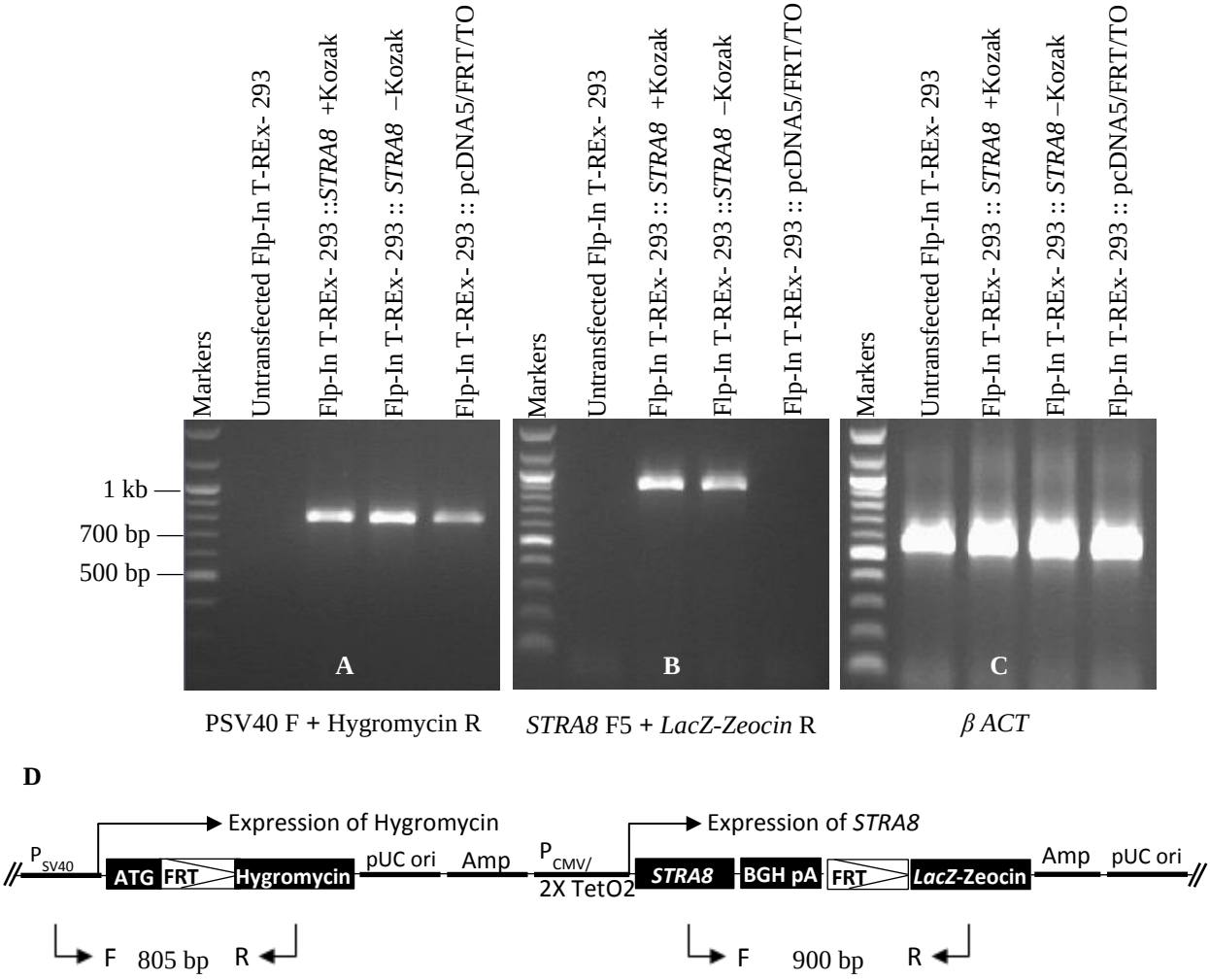


Figure 4.13 PCR screening of different *STRA8* integrant Flp-In T-REx-293 cell lines. Panel **(A)** shows the presence of the *STRA8* gene in the transfected cells when using primers located in the PSV40 (forward) and in the hygromycin gene (reverse). Panel **(B)** shows the PCR of Flp-In T-REx-293 with the *STRA8* genes when using primers in the *STRA8* gene (forward) and in the *LacZ-Zeocin* gene (reverse). PCR using primers within the *β*-actin gene are shown as a positive control for the genomic DNA samples (panel (**C)**). The diagram in panel **(D)** shows the localisation of *STRA8* cDNAs in the genome of the Flp-In T-REx-293 cell line.

4.2.2.3 Evaluation of the *STRA8* expression level in the Flp-In T-REx-293 cell line

Analysis of expression of the *STRA8* gene was achieved using both RT-PCR and qRT- PCR analyses. *STRA8* expression was induced by the addition of tetracycline to the culture cells (2 µg/ml). Upon addition of tetracycline to the cell medium, tetracycline binds to the Tet repressor, which is expressed from the pcDNA6/TR in the Flp-In T-REx-293 cells. Subsequently, the *STRA8* promoter initiates the transcription (Yao *et al*., 1998). However, expression of *STRA8* is repressed in the absence of tetracycline (Yao *et al*., 1998), because the Tet repressor binds to the Tet operator 2 (TetO2) sequences in the pcDNA5/FRT/TO vector (CMV)/TetO2), which inhibit the transcription process for the *STRA8* gene (refer to Figure 4.9).

cDNAs were synthesised from the mRNAs of these cell lines. Two negative controls were used: untransfected cells and cells transfected with empty pcDNA5/FRT/TO. *β-ACT* was used as a positive control for the cDNA samples. RT-PCR results showed that the expression of the *STRA8* gene (using F1 and R1 primers) was observed in the cell lines that had been integrated with either pcDNA5/FRT/TO::*STRA8*+Kozak or pcDNA5/FRT/TO::*STRA8*–Kozak (Figure 4.14). However, the mRNA levels of *STRA8* cDNA PCR products were higher in the cells treated with tetracycline compared to the same cells with no treatment. The mRNA levels of the *STRA8* gene were not detectable in either negative control. The results in Figure 4.14 also show two distinct bands (18S and 28S ribosomal RNAs), which suggest good quality RNAs for all Flp-In T-REx-293 cell lines.

qRT-PCR was carried out for the *STRA8* gene in different Flp-In T-REx-293 cell lines using STRA8 F3 and R3 primers, while commercial positive primers were used. The STRA8 qRT-PCR primers (F3 and R3) were tested by RT-PCR to confirm their performance using different cDNAs from the different Flp-In T-REx-293 cells before and after tetracycline treatment. The results showed that the expression of *STRA8* was found in cell lines that had been inserted with *STRA8* at the expected size of 124 bp (Figure 4.15), which suggests that the performance of the primers was good. Introns-intron primers and intron flanking primers were designed for *STRA8* to prevent the amplification of potential contaminating genomic DNA. RT-PCR of the two sets of primer showed that the expression was observed in the cell lines with the integrated *STRA8* gene when using intron flanking primers (STRA8 F1 and STRA8 R1). However, *STRA8* expression was not observed in any of the cell lines when using intron-intron primers (STRA8 F2 and STRA8 R1) (Figure 4.16), which suggests that there was no genomic DNA contamination. The performance of the intron-intron primers was

confirmed using a serial dilution of genomic DNA isolated from Flp-In T-REx-293::*STRA8*+Kozak. The expression was indicated as 306 bp in DNA amounts of 100, 10 and 1 ng (Figure 4.17).

In the *STRA8* qRT-PCR results, the NRT (no reverse transcriptase) and NTC (no template control) negative controls did not give a Cq (quantification cycle) reading, which suggests that there was no non-specific background and no genomic DNA contamination (Table 4.1). The qRT-PCR results show a higher level of *STRA8* expression in the Flp-In T-REx-293::*STRA8*+Kozak and Flp-In T-REx-293::*STRA8*–Kozak cells compared to the Flp-In T-REx-293::pcDNA5/FRT/TO cells after treatment with tetracycline (Figure 4.18). The RT-PCR results shown in Figure 4.14 suggest that the Flp-In T-REx-293::pcDNA5/FRT/TO cell line was negative for *STRA8* expression. The intensity of the *STRA8* PCR band was stronger in cell lines that had been integrated with *STRA8* after treatment with tetracycline than that observed in other cell lines with no treatment. Therefore, these qRT-PCR results correspond to the RT-PCR results previously observed (Figure 4.14). The *STRA8* qRT-PCR results were normalised to the qRT-PCR results for *Lamin A* and *β ACT*; the results are shown in Table 4.1

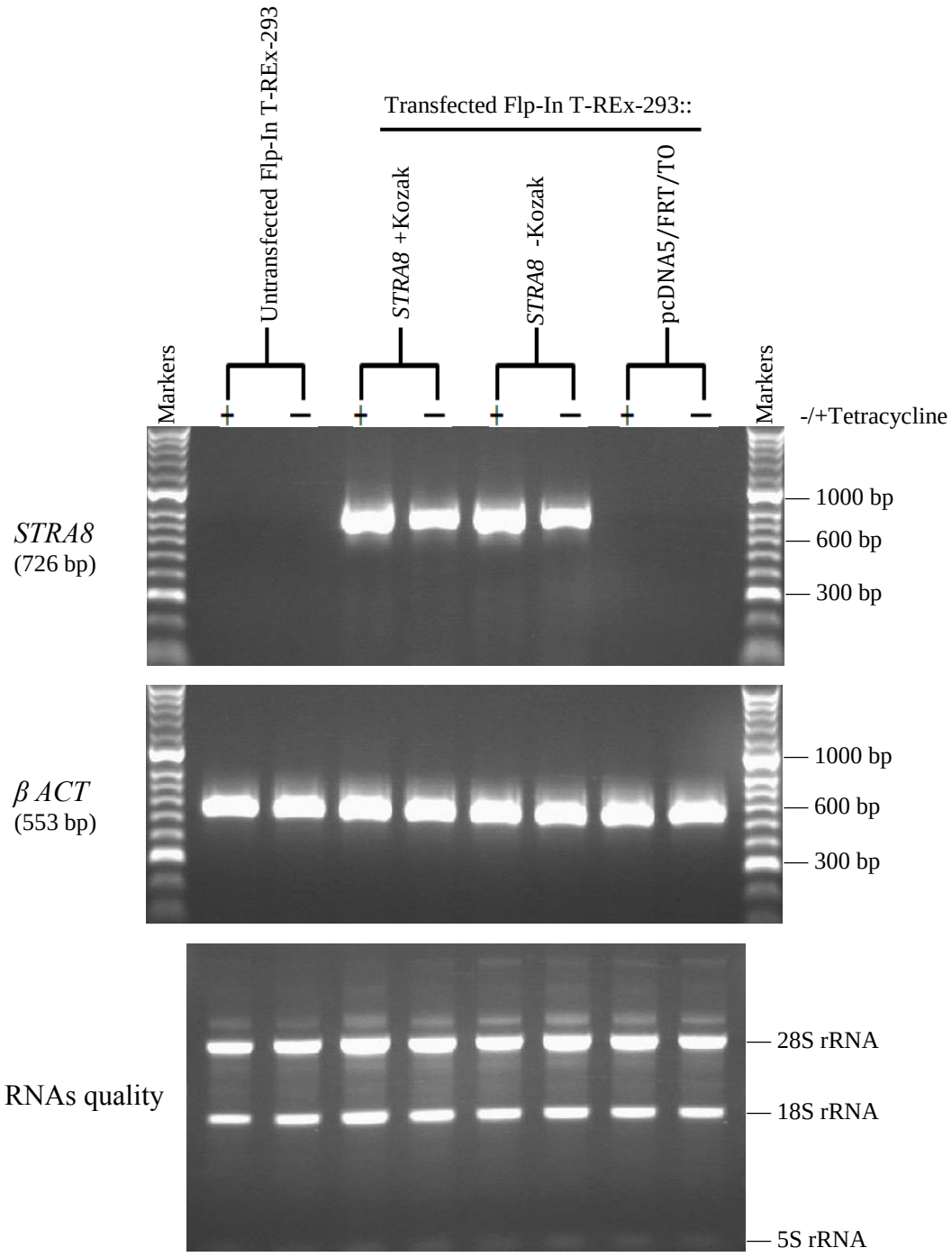


Figure 4.14 RT-PCR expression analysis of *STRA8* and *β ACT* in Flp-In T-REx-293 cell lines. The top agarose gel shows the expression of *STRA8* in Flp-In T-REx-293 integrated with three different insertions (*STRA8*+Kozak or *STRA8*–Kozak or pcDNA5/FRT/TO) with and without tetracycline treatment. The middle agarose gel displays the expression profile for *β*-actin as a positive control for the cDNA samples. The bottom agarose gel shows the quality of the total RNAs for the different Flp-In T-REx-293 cell lines. Three bands of ribosomal RNAs appear in each sample. The 28S and 18S ribosomal RNA bands are clearly visible in the RNA samples at the ratio of approximately 2:1. The RNAs were isolated using the RNeasy Plus Mini Kit.

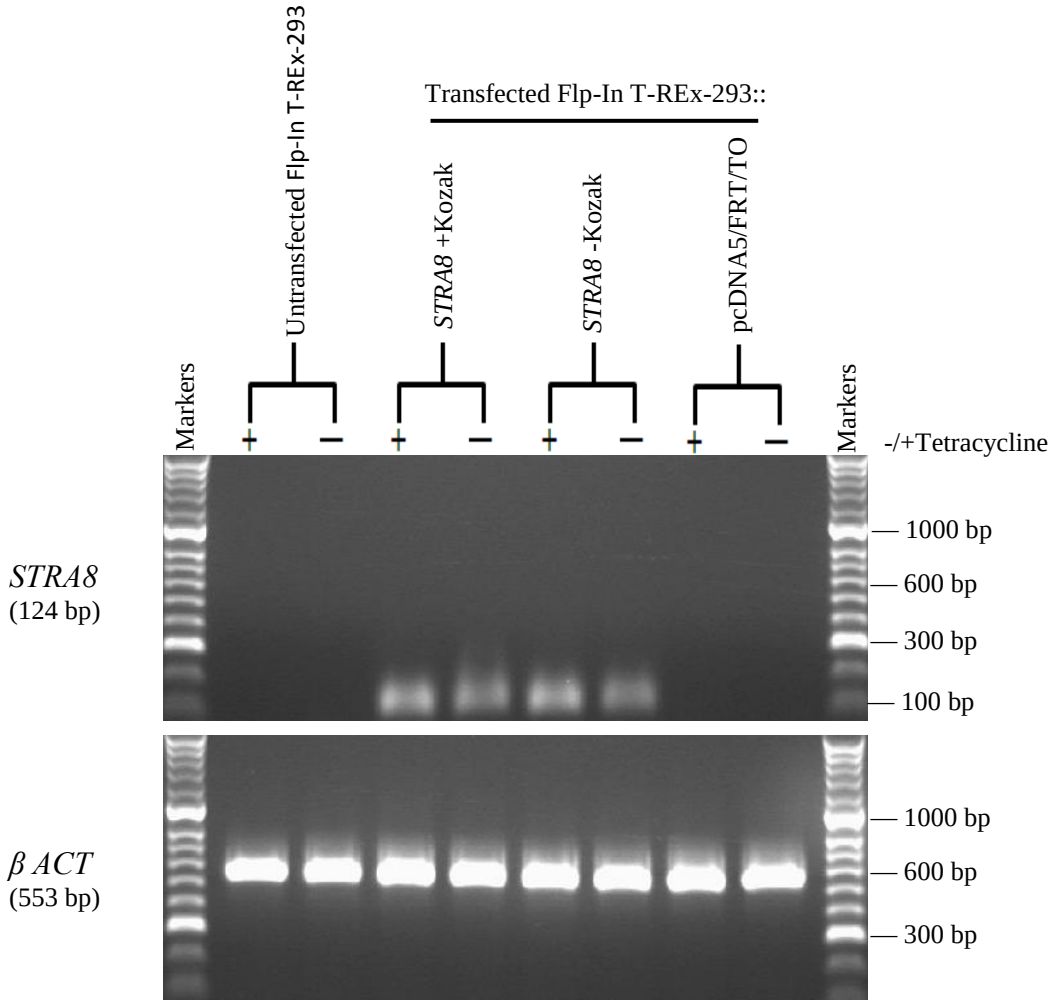


Figure 4.15 RT-PCR analysis of *STRA8* expression using qRT-PCR primers in different Flp-In T-REx-293 cell lines. The top agarose gel shows the expression of *STRA8* in Flp-In T-REx-293 integrated with three different insertions (*STRA8*+Kozak or *STRA8*–Kozak or pcDNA5/FRT/TO) with and without tetracycline treatment. The bottom agarose gel displays the expression profile for *β*-actin as a positive control for the cDNA samples.

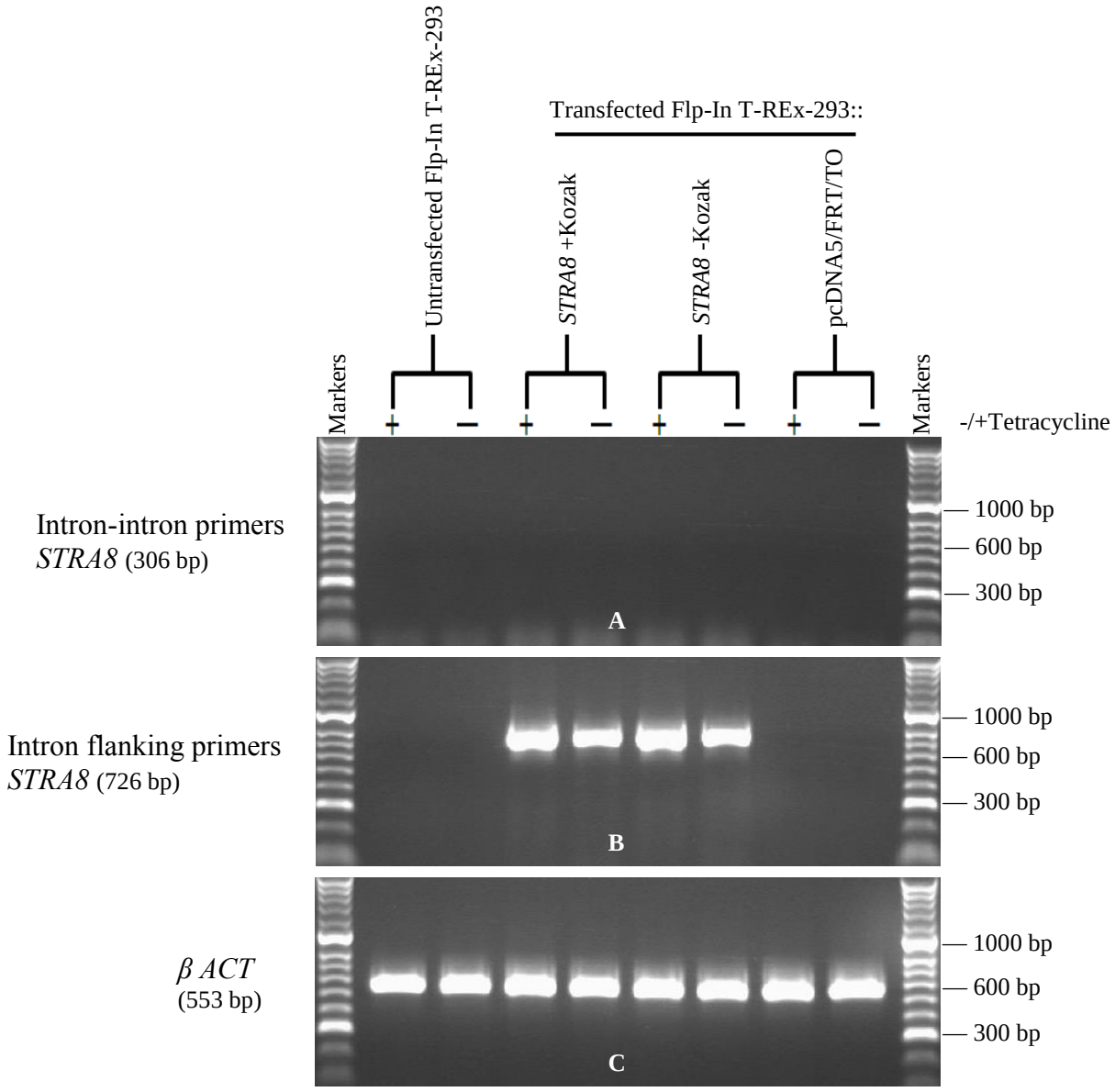


Figure 4.16 RT-PCR analysis of *STRA8* expression in different Flp-In T-REx-293 cells using intron-intron and intron flanking primers. The agarose gel on panels **(A** and **B)** show the expression of *STRA8* in Flp-In T-REx-293 integrated with three different insertions (*STRA8*+Kozak or *STRA8*–Kozak or pcDNA5/FRT/TO) with and without tetracycline treatment. In panel **(A)**, no amplification is apparent in any of the cells, whereas in gel **(B),** there are strong expressions in the cells that had been integrated with *STRA8*, with or without Kozak. The bottom agarose gel on panel **(C)** displays the expression profile for *β*-actin as a positive control for the cDNA samples.

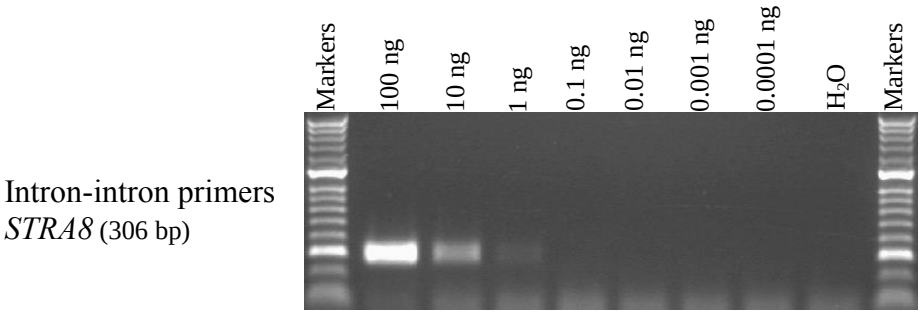


Figure 4.17 Serial dilution of genomic DNA from Flp-In T-REx-293:: *STRA8*+Kozak cells. The agarose gel shows the PCR analysis of *STRA8* in different dilutions of genomic DNA of Flp-In T-REx-293::*STRA8*+Kozak to confirm the performance of the intron-intron primers.

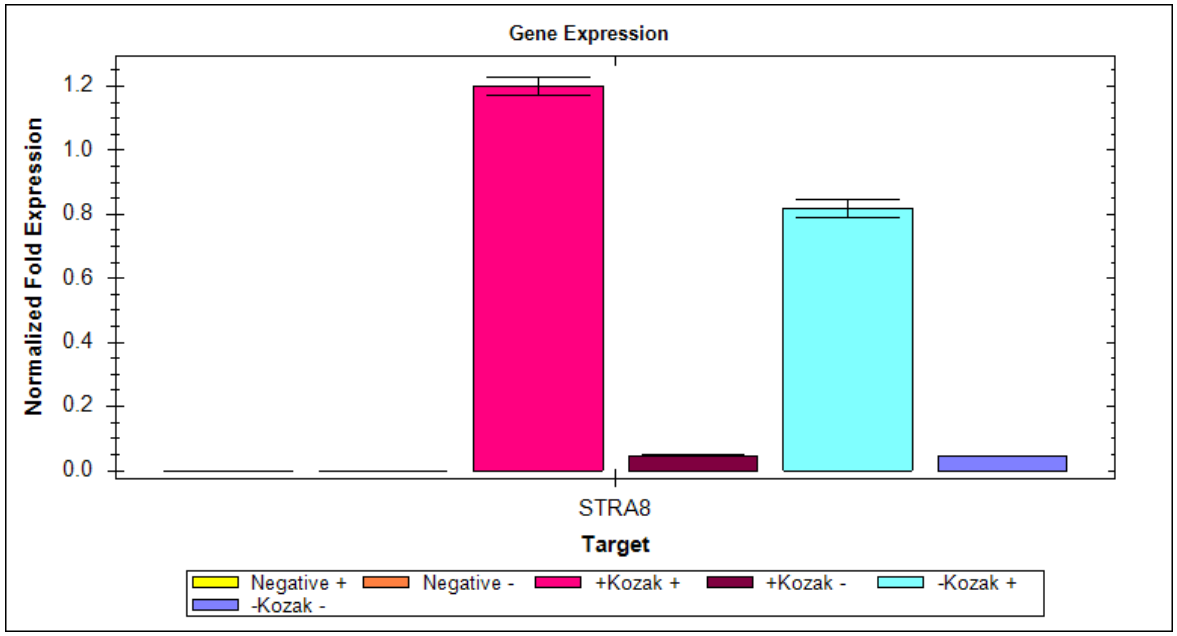


Figure 4.18 qRT-PCR analysis of *STRA8* expression in Flp-In T-REx-293 cells. The bar chart displays the gene expression results for *STRA8* with and without Kozak consensus in Flp-In T-REx-293 cell lines before (-) and after (+) adding tetracycline to the cell culture, compared to the cells with the integrated pcDNA5/FRT/TO vector alone (negative). Expression data were normalised to the *β ACT* and *Lamin A* reference genes. The error bars represent the standard error of the mean (SEM) for three repeats.

Table 4.1 Real-time qRT-PCR analysis for *STRA8* expression in Flp-In T-REx-293 cells. The table displays the mean of three repeats of the quantification cycle (Cq) and the standard deviation (SD) for *STRA8* expression with and without Kozak consensus in Flp-In T-REx-293 cell lines before (-) and after (+) tetracycline treatment. These results were normalised by *β ACT* and *Lamin A* expression. Also shown in the table are the Cq readings for the no reverse transcriptase (NRT) and no template control (NTC).

Target name	Sample name	Wells	Cq Mean	Cq SD
STRA8	Flp-In T-Rex-293: pcDNA5/FRT/TO+	3	31.26	0.079
β ACT	Flp-In T-Rex-293: pcDNA5/FRT/TO+	3	16.65	0.024
Lamin A	Flp-In T-Rex-293: pcDNA5/FRT/TO+	3	21.99	0.127
STRA8	Flp-In T-Rex-293: pcDNA5/FRT/TO-	3	31.02	0.125
β ACT	Flp-In T-Rex-293: pcDNA5/FRT/TO-	3	16.74	0.068
Lamin A	Flp-In T-Rex-293: pcDNA5/FRT/TO-	3	22.00	0.121
STRA8	Flp-In T-Rex-293: STRA8+Kozak+	3	18.09	0.022
β ACT	Flp-In T-Rex-293: STRA8+Kozak+	3	16.95	0.110
Lamin A	Flp-In T-Rex-293: STRA8+Kozak+	3	22.17	0.042
STRA8	Flp-In T-Rex-293: STRA8+Kozak-	3	22.64	0.073
β ACT	Flp-In T-Rex-293: STRA8+Kozak-	3	16.89	0.052
Lamin A	Flp-In T-Rex-293: STRA8+Kozak-	3	21.95	0.090
STRA8	Flp-In T-Rex-293: STRA8-Kozak+	3	18.68	0.074
β ACT	Flp-In T-Rex-293: STRA8-Kozak+	3	17.00	0.057
Lamin A	Flp-In T-Rex-293: STRA8-Kozak+	3	22.18	0.047
STRA8	Flp-In T-Rex-293: STRA8-Kozak-	3	22.75	0.036
β ACT	Flp-In T-Rex-293: STRA8-Kozak-	3	16.95	0.046
Lamin A	Flp-In T-Rex-293: STRA8-Kozak-	3	22.05	0.046
STRA8	Flp-In T-Rex-293: pcDNA5/FRT/TO+ (NRT)	1	0.00	0.000
β ACT	Flp-In T-Rex-293: pcDNA5/FRT/TO+ (NRT)	1	0.00	0.000
Lamin A	Flp-In T-Rex-293: pcDNA5/FRT/TO+ (NRT)	1	0.00	0.000
STRA8	Flp-In T-Rex-293: pcDNA5/FRT/TO- (NRT)	1	0.00	0.000
β ACT	Flp-In T-Rex-293: pcDNA5/FRT/TO- (NRT)	1	0.00	0.000
Lamin A	Flp-In T-Rex-293: pcDNA5/FRT/TO- (NRT)	1	0.00	0.000
STRA8	Flp-In T-Rex-293: STRA8+Kozak+ (NRT)	1	0.00	0.000
β ACT	Flp-In T-Rex-293: STRA8+Kozak+ (NRT)	1	0.00	0.000
Lamin A	Flp-In T-Rex-293: STRA8+Kozak+ (NRT)	1	0.00	0.000
STRA8	Flp-In T-Rex-293: STRA8+Kozak- (NRT)	1	0.00	0.000
β ACT	Flp-In T-Rex-293: STRA8+Kozak- (NRT)	1	0.00	0.000
Lamin A	Flp-In T-Rex-293: STRA8+Kozak- (NRT)	1	0.00	0.000
STRA8	Flp-In T-Rex-293: STRA8-Kozak+ (NRT)	1	0.00	0.000
β ACT	Flp-In T-Rex-293: STRA8-Kozak+ (NRT)	1	0.00	0.000
Lamin A	Flp-In T-Rex-293: STRA8-Kozak+ (NRT)	1	0.00	0.000
STRA8	Flp-In T-Rex-293: STRA8-Kozak- (NRT)	1	0.00	0.000
β ACT	Flp-In T-Rex-293: STRA8-Kozak- (NRT)	1	0.00	0.000
Lamin A	Flp-In T-Rex-293: STRA8-Kozak- (NRT)	1	0.00	0.000
STRA8	NTC	1	0.00	0.000
β ACT	NTC	1	0.00	0.000
Lamin A	NTC	1	0.00	0.000

4.2.3 siRNA knockdown of STRA8 in HEP-G2 and NT2

Knockdown of STRA8 levels was attempted using four types of *STRA8*-specific siRNAs in both HEP-G2 and NT2 cell lines to determine any effect that knockdown of this protein may have on these cells and to test the specificity of the anti-STRA8 antibody. Whole cell extracts were collected from the cells after siRNA treatment and analysed using Western blotting.

The predicted size of the *STRA8* cDNA according to the NCBI, is 993 bp, which gives a molecular weight of about 37 KDa. However, the DNA sequencing for RT-PCR products confirms that the entire open reading frame of *STRA8* is 1209 bp (see Section 4.2.5). Therefore, the expected molecular weight of STRA8 is approximately 44 KDa. However, the band observed in HEP-G2 and NTERA2 producing a molecular weight of about 48 KDa. The band thought to correspond to STRA8 appears fainter in HEP-G2 after 48 hours treatment (2 "hits") with all siRNAs, especially siRNAs 3, 5 and a combination of four STRA8-specific siRNAs compared to the negative controls (untreated cells and cells treated with non-interfering siRNA). α-Tubulin levels were shown as a loading control. Therefore, the Western blot results for the two negative controls and α-Tubulin suggest that there was significant STRA8 knockdown in the HEP-G2 cells (Figure 4.19).

Although NT2 did not show any detectable expression of the *STRA8* gene (refer to Figure 3.14), Western blot analysis showed significant protein levels at a molecular weight of ~48 KDa. The intensity of the band corresponding to STRA8 in the NT2 cells treated for 48 hours (2 "hits") with 5 nM siRNA appears to be reduced compared to the untreated cells and the cells treated with non-interfering siRNA. The α-Tubulin signal appears to be relatively uniform, which suggests that this difference is due to successful STRA8 knockdown, especially in cells transfected with siRNA number 5 (Figure 4.19).

4.2.4 Evaluation of the protein levels of STRA8 in Flp-In T-REx-293 cell lines

STRA8 protein levels were assessed in different integrants of Flp-In T-REx-293 cell lines before and after tetracycline treatment. The molecular weight for STRA8 protein is ~48 KDa, according to Western blotting (Figure 4.19). Western blots using the anti-STRA8 antibody showed different bands, especially in cell lines treated with tetracycline. This may suggest that tetracycline treatment could give a nonspecific background. However, the Flp-In T-REx-293 cell line inserted with *STRA8*+Kozak and treated with tetracycline showed a strong band at a molecular weight of ~48 KDa. This suggests that the band is related to STRA8 protein (Figure 4.20). The protein loading for each clone was indicated by α-Tubulin.

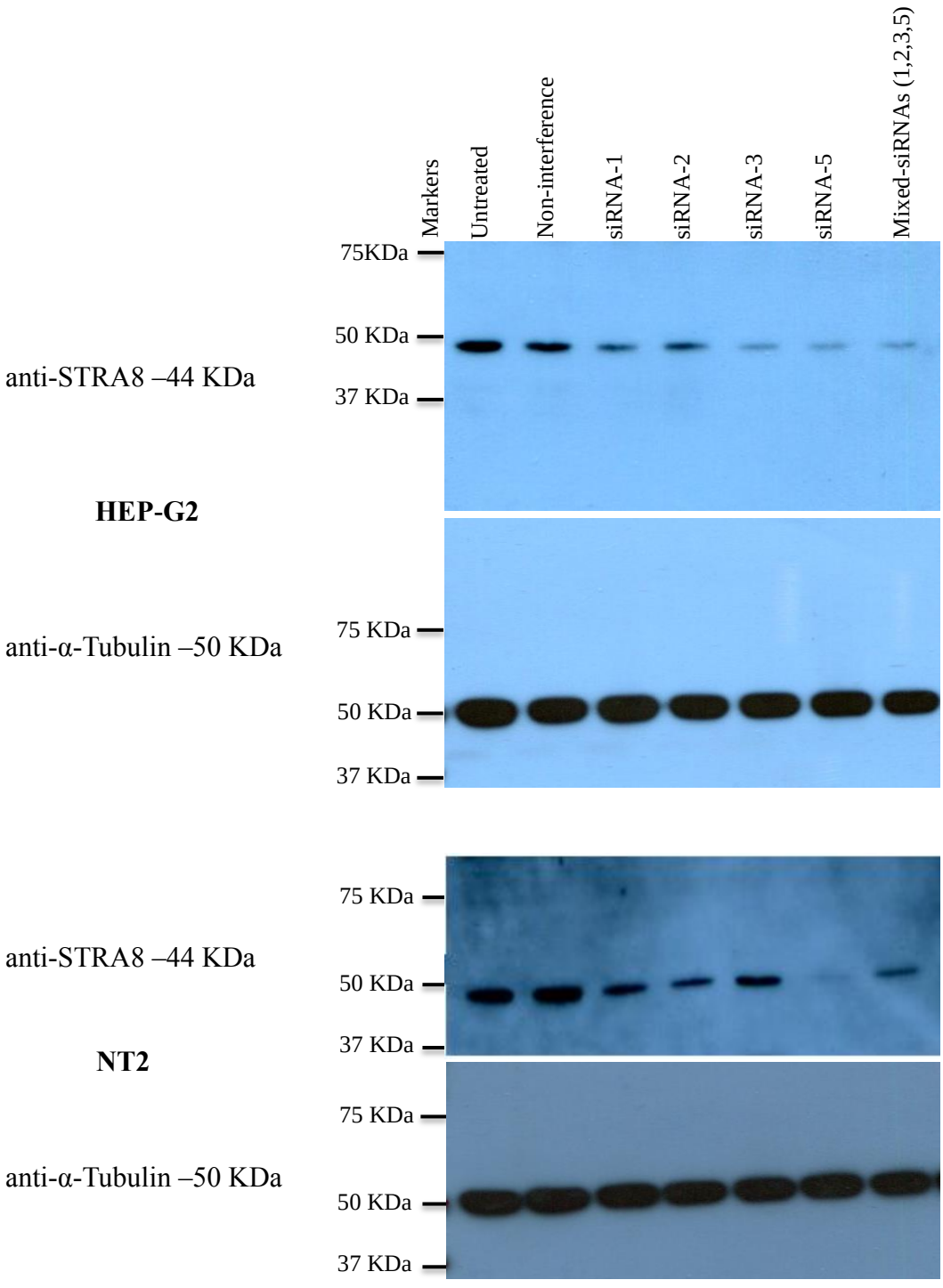


Figure 4.19 Western blot analysis for siRNA knockdown of STRA8 in HEP-G2 and NT2 cancer cells. Two negative controls for the siRNA knockdown were used: cells with no transfection (untreated), and cells treated with non-interfering siRNA. Cells were transfected with 5 nM concentrations of four independent *STRA8*-specific siRNAs, a combination of the four siRNAs (siRNAs 1+2+3+5) or with 5 nM of non-interfering siRNA. The transfection reagent and siRNAs were added to the cell cultures at seeding. The cells were collected 48 hours after siRNA transfection. α-Tubulin protein levels were used as controls for protein loading.

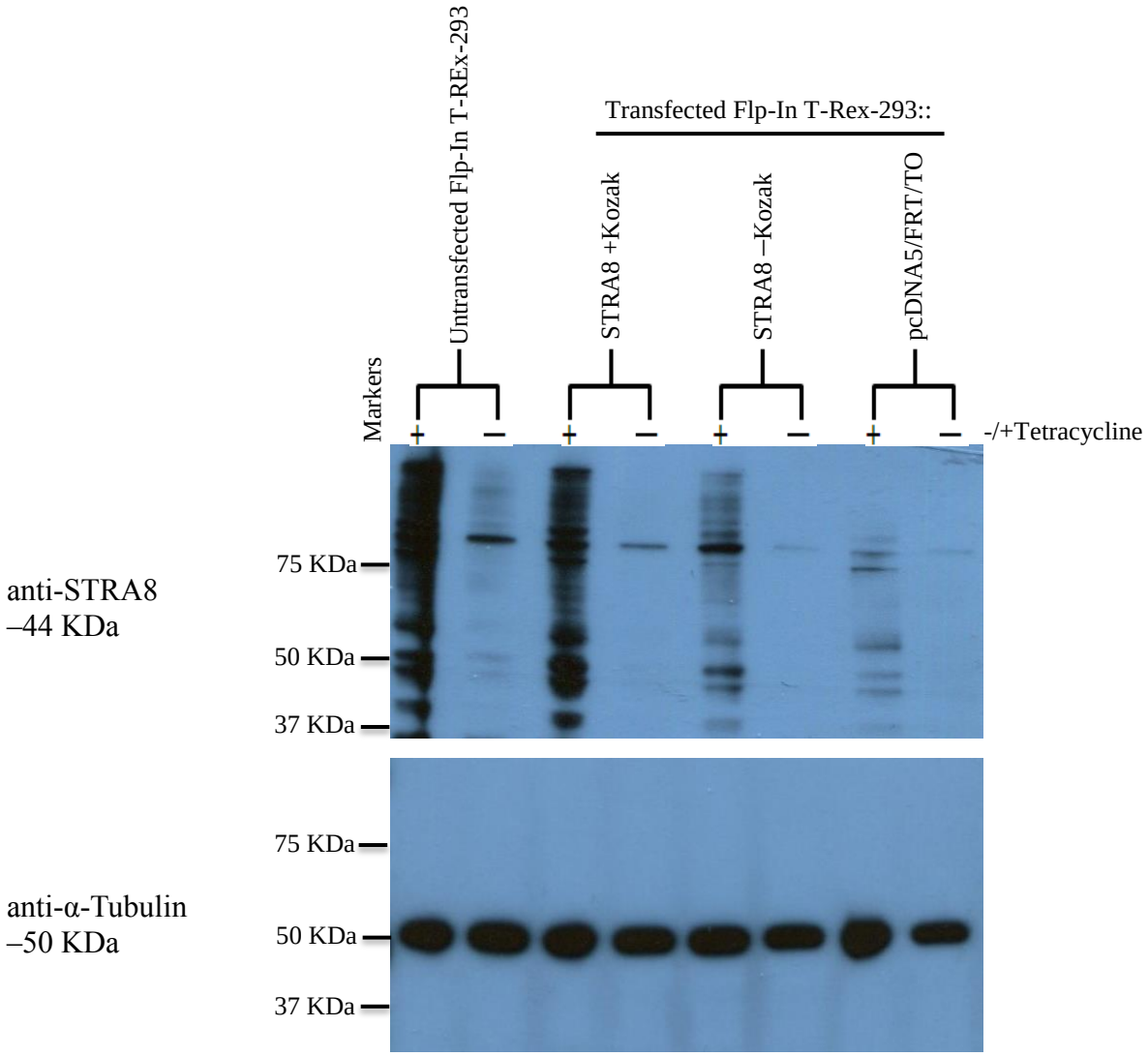


Figure 4.20 Western blot analysis showing STRA8 induction using tetracycline in the Flp-In T-REx-293 cell lines. Four of these cultures were subjected to 2 µg/ml of tetracycline and four were not. Positive signals are seen in cells containing the *STRA8* gene with Kozak sequencing after treatment with tetracycline. The cells were harvested 24 hours after induction. The lysates were separated on a 10% SDS-PAGE gel. α-Tubulin protein levels were used as controls for protein loading.

4.2.5 Comparison between human *STRA8* cDNA and amino acid sequences from the GenBank and those obtained from RT-PCR

The RT-PCR analysis of the *STRA8* gene in the testis cDNA gave a band 216 bp, larger than expected, when we used internal primers; the forward primer located in exon 3 and the reverse primer located in exon 7. These extra base pairs were confirmed by DNA sequencing. Therefore, other primers were designed for cloning to cover the whole open reading frame of *STRA8* and to confirm the final results. A forward primer was designed and located in exon 1 starting at a start codon (ATG) and a reverse primer was located in exon 9, which starts by sequences complementary to a stop codon (TAA). RT-PCR results for the *STRA8* gene using both primers and testis as a template for amplification was 1209 bp, according to the DNA sequence; although the whole cDNA length identified in the GenBank is 993 bp. Indeed, in this study, 216 bp were detected by RT-PCR analysis and the majority of these sequences were translated to glutamic acid (Figure 4.21). Previous studies showed that the *STRA8* human protein does not have a glutamic acid-rich domain. However, Stra8 mouse protein is rich in glutamic acid, which confers a high acidity to the Stra8 protein (Miyamoto *et al*., 2002; Oulad-Abdelghani *et al*., 1996b). Some proteins of the cytoskeleton, such as neurofilaments are composed of glutamic acid (Levy *et al*., 1987). Figure 4.21 shows the differences between both sequences.

atgggaagattgatgtggacaagatcctctttttcaatcaagaaatcaggctgtggcag
M G K I D V D K I L F F N Q E I R L W Q
cttataatggcaacccctgaagaaaacagcaatccccatgacagagcaacaccccagctg
L I M A T P E E N S N P H D R A T P Q L
ccagcacagctgcaggagcttgagcatcgggtggcccggagacggctgtcccaggcccgc
P A Q L Q E L E H R V A R R R L S Q A R
caccgagccaccctggcagcgctcttcaacaacctcaggaagacagtgtactctcagtct
H R A T L A A L F N N L R K T V Y S Q S
gatctcatagcctcaaagtggcaggttctgaataaggcaaagagtcatattccagaactg
D L I A S K W Q V L N K A K S H I P E L
gagcaaaccctggataatttgctgaagctgaaagcatccttcaacctggaagatgggcat
E Q T L D N L L K L K A S F N L E D G H
gcaagcagcttagaggaggtcaagaaagaatatgccagcatgtattctggaaatgacagc
A S S L E E V K K E Y A S M Y S G N D S
ctgctttcaaacagttttcctcagaatggttcctccccttggtgcccaactgaggcagtc
L L S N S F P Q N G S S P W C P T E A V
aggaaggatgctgaggaggaggaagatgaggaagaggaagatcaagaagaagaggaggag
R K D A E E E E E D E E E E D Q E E E E E
gaagaagaagaggaggaggaggaggaagaggaggaagaggaagaggaggaggaggaagag
E E E E E E E E E E E E E E E E E E E E E
gagaaaaaagtgatcttatactccccaggaactttgtcgcctgacctcatggaatttgaa
E K K V I L Y S P G T L S P D L M E F E
cggtatctcaacttttacaaacagacgatggaccttctgactggcagcgggatcattacc
R Y L N F Y K Q T M D L L T G S G I I T
ccgcaggaggcggcgctgcccatcgtctccgcggccatctcccacctgtggcagaacctc
P Q E A A L P I V S A A I S H L W Q N L
tcggaggagaggaaggccagcctccggcaggcctgggcgcagaagcaccgcggccctgcg
S E E R K A S L R Q A W A Q K H R G P A
accctggcggaggcctgccgagagccggcctgtgccgagggcagcgtgaaggacagcggc
T L A E A C R E P A C A E G S V K D S G
gtggacagccaggggccagctgctcgctggtctccacgcccgaggagatccttttgag
V D S Q G A S C S L V S T P E E I L F E
gatgcctttgatgtggcaagcttcctggacaaaagtgaggttccgagtacatctagctcc
D A F D V A S F L D K S E V P S T S S S
agttcagtgcttgccagctgcaacccagaaaacccagaggagaagtttcagctctatatg
S S V L A S C N P E N P E E K F Q L Y M
cagatcatcaacttttttaaaggccttagctgtgcaaacactcaagtaaagcaggaagca
Q I I N F F K G L S C A N T Q V K Q E A
tcctttcccgttgatgaagagatgatcatgttgcagtgcacagagacctttgacgatgaa
S F P V D E E M I M L Q C T E T F D D E
gatttgtaa
D L -

Figure 4.21 Nucleotide and deduced amino acid sequence of human *STRA8* cDNA. The ATG codon and the TAA terminating codon are bold and coloured in red. The *STRA8* cDNA sequence was detected using testis cDNA, forward and reverse primers. The positions of the two primers used for expression analysis are underlined. The nucleotides marked in black letters indicate identical sequences between the *STRA8* cDNA in the GenBank and those obtained via RT-PCR analysis. The extra bases identified by RT-PCR are marked in green. The *STRA8* nucleotide sequence was translated to amino acid using the Expasy program; http://www.expasy.org/tools/.

4.2.6 Analysis of meiotic gene expressions in different Flp-In T-REx-293 cell lines treated with retinoic acids

It has been demonstrated that RA induces *STRA8* gene expression in germ cells (Feng *et al*., 2014). Therefore, our hypothesis asks: Does the *STRA8* gene lead to the induction of any other meiosis-specific gene when treating cell lines that have the *STRA8* gene with RAs? The Flp-In T-REx-293 cell lines were treated with two types of RAs, *all-trans* and *9-cis* retinoic acids, for 24 hours. These RAs were used at a final concentration of 5 µM and 10 µM. 12 meiosis-specific/CTA genes were randomly selected to perform this experiment: *GAGE1, MAGE-A1, NUT, PRDM9, RAD21L, REC8, SMC1β, STAG3, SYCP1, SYCP2, SYCP3* and *TEX19*.

The experiment was initiated by treating the Flp-In T-REx-293 cells inserted with *STRA8*+Kozak, *STRA8*–Kozak or empty pcDNA5/FRT/TO with two different concentrations of two types of RAs, 5 or 10 µM *all-trans* and 5 or 10 µM *9-cis* retinoic acids. cDNAs were synthesised from the mRNA for these cells. Two negative controls were used to detect the absence of the *STRA8* gene in the cells, untransfected Flp-In T-REx-293 cells and cells transfected with pcDNA5/FRT/TO plasmid only. Testis cDNA was used to validate the primer for each gene. *β ACT* gene served as a positive control to validate the quality of the cDNA, and the *STRA8* gene was used as a positive control for analysis of the induction of the 12 meiosis-specific genes (Figure 4.22). RAs were prepared in DMSO. Therefore, cells were subjected to DMSO, which itself may induce cell differentiation; therefore, a control culture was made.

RT-PCR expression was performed for 9 meiosis-specific genes including *GAGE1, MAGE-A1, NUT, PRDM9, RAD21L, SMC1β, SYCP1, SYCP3* and *TEX19*. The aim of this analysis was to establish which of these genes become overexpressing when the *STRA8* gene was activated in the cells. RT-PCR showed that there was no induction of mRNA expression for these genes (Figures 4.23 and 4.24), except for *MAGE*-A1 (Figure 4.23), but a strong band appeared in cells that had been integrated with plasmid (pcDNA5/FRT/TO) alone. This could be due to the effect of DMSO. The expressions of these genes were compared to negative and positive controls.

After obtaining the results of the previous RT-PCR, another experiment was designed using 10 µM *all-trans* RA and 15 different Flp-In T-REx-293 clones: 5 different clones from Flp-In T-REx-293::*STRA8*+Kozak; 5 different clones from Flp-In T-REx-293::*STRA8*–Kozak; and 5

different clones from Flp-In T-REx-293::pcDNA5/FRT/TO. Cells containing *STRA8*+Kozak or *STRA8*–Kozak were used as the positive controls for the RT-PCR of the *STRA8* gene. In contrast, untransfected Flp-In T-REx-293 cells and cells transfected with pcDNA5/FRT/TO were used as the negative controls for the RT-PCR. The RT-PCR for the *STRA8* gene did not show any expression, as expected, in the untransfected cells and in all 5 clones of Flp-In T-REx-293::pcDNA5/FRT/TO (Figure 4.25). The results of the RT-PCR expression in Flp-In T-REx-293::pcDNA5/FRT/TO for *GAGE1*, *MAGE-A1*, *NUT*, *PRDM9*, *RAD21L* and *REC8* are shown in Figure 4.25. The *SMC1β, STAG3, SYCP1, SYCP2, SYCP3* and *TEX19* results are shown in Figure 4.26.

The RT-PCR for the *STRA8* gene showed strong expression in all 5 clones of Flp-In T-REx-293::*STRA8*+Kozak (Figure 4.27), but there was no significant overexpression of the other meiosis genes tested here. The results of the RT-PCR expression in Flp-In T-REx-293::*STRA8*+Kozak for *GAGE1, MAGE-A1*, *NUT, PRDM9, RAD21L* and *REC8* are shown in Figure 4.27. The results for *SMC1β, STAG3, SYCP1, SYCP2, SYCP3* and *TEX19* are shown in Figure 4.28. In contrast, The RT-PCR results for the 12 selected meiosis genes in the 5 different clones from Flp-In T-REx-293::*STRA8*–Kozak showed results identical to those obtained from Flp-In T-REx-293::*STRA8*+Kozak. The RT-PCR expression results for the selected meiosis genes in Flp-In T-REx-293::*STRA8*–Kozak are shown in Figures 4.29 and 4.30.

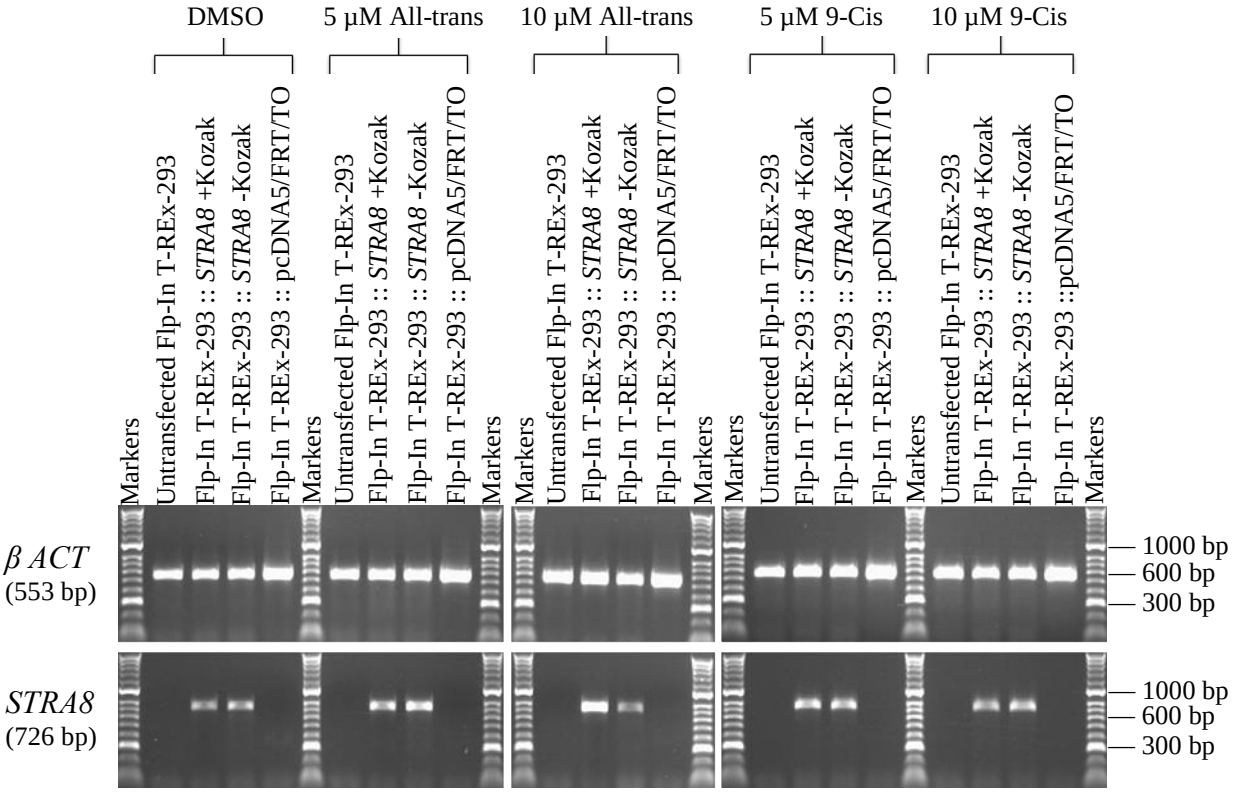


Figure 4.22 Analysis of the induction of *STRA8* gene expression in different Flp-In T-REx-293 cell lines treated with retinoic acid. RT-PCR analysis of *STRA8* mRNA expression in different Flp-In T-REx-293 cells. Prior to RNA extraction, the cells were treated with *all-trans* or *9-cis* retinoic acids at different concentrations, or with the DMSO vehicle alone for 24 hours. *β ACT* expression was used as a positive control for the cDNA samples. The expected amplicon size of each gene is shown on the left between brackets.

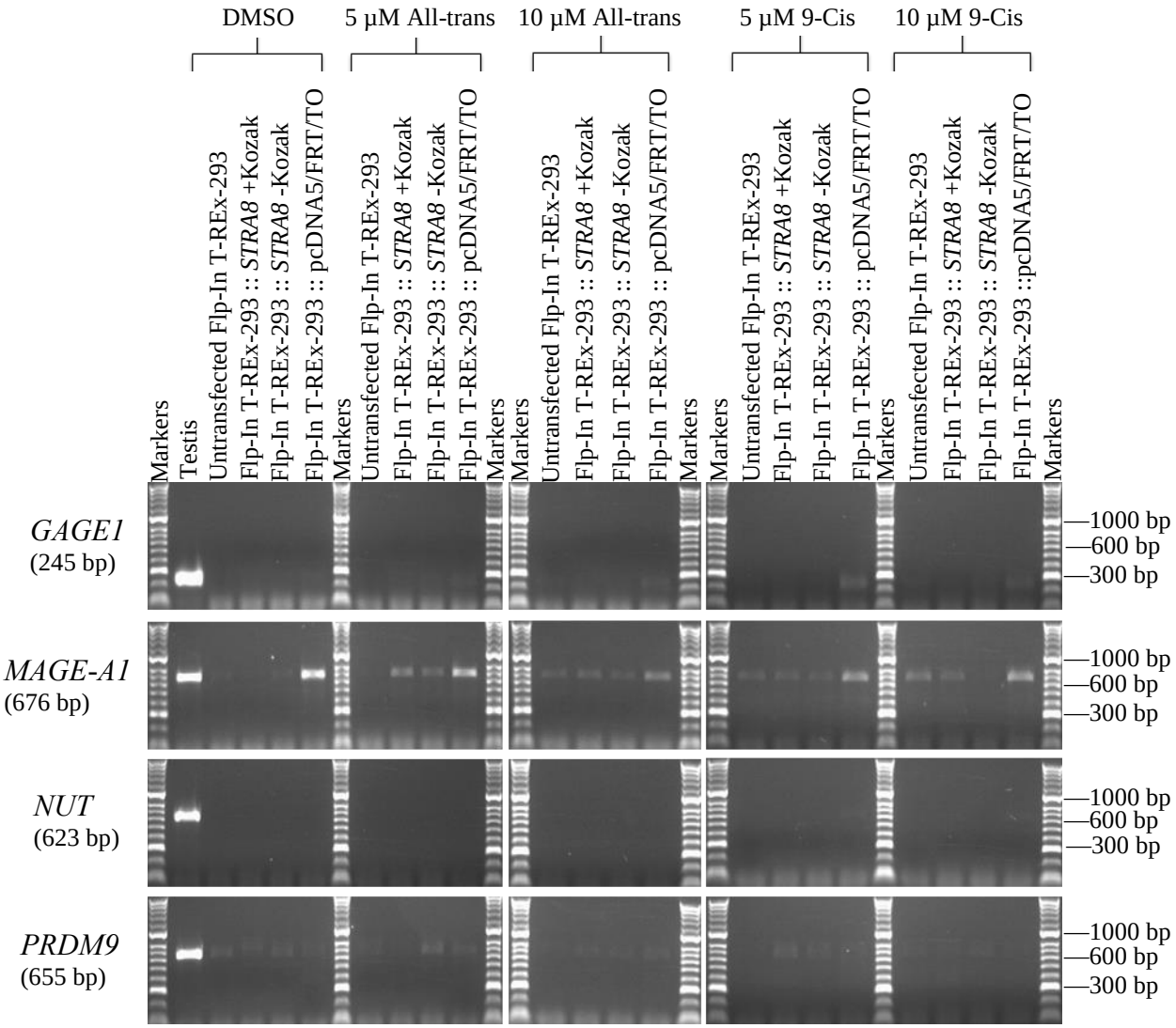


Figure 4.23 Analysis of induction of meiotic genes expression in different Flp-In T-REx-293 cell lines treated with retinoic acid. RT-PCR analysis of *GAGE1, MAGE-A1, NUT* and *PRDM9* mRNA expression in different Flp-In T-REx-293 cells. Prior to RNA extraction, the cells were treated with *all-trans* or *9-cis* retinoic acids at different concentrations or with the DMSO vehicle alone for 24 hours. The expected amplicon size of each gene is shown on the left between brackets.

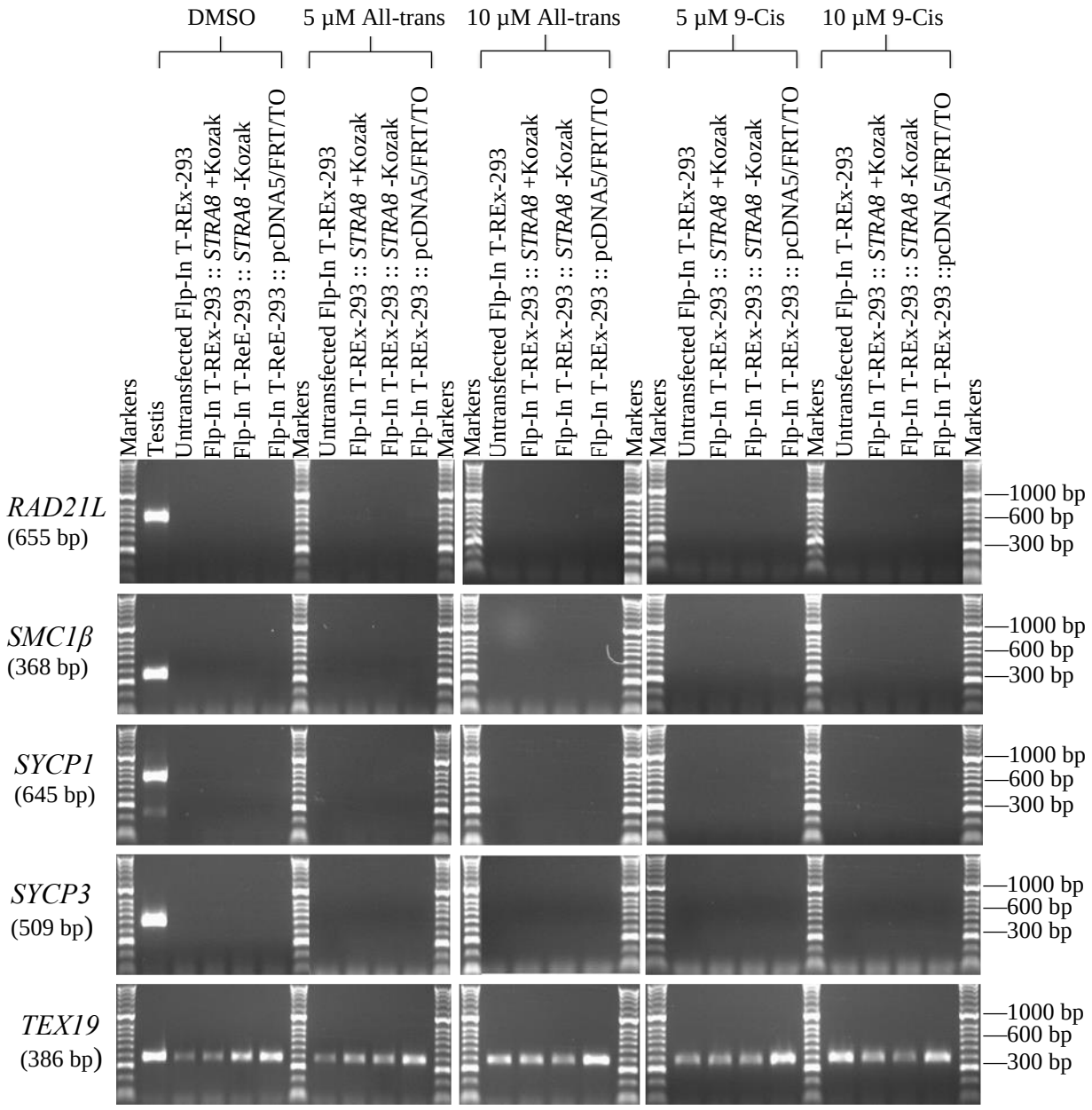


Figure 4.24 Analysis of meiotic genes expression in different Flp-In T-REx-293 cell lines treated with retinoic acid. RT-PCR analysis of *RAD21L*, *SMC1β*, *SYCP1*, *SYCP3* and *TEX19* mRNA expression in different Flp-In T-REx-293 cells. Prior to RNA extraction, the cells were treated with *all-trans* or *9-cis* retinoic acids at different concentrations or with the DMSO vehicle alone for 24 hours. The expected amplicon size of each gene is shown on the left between brackets.

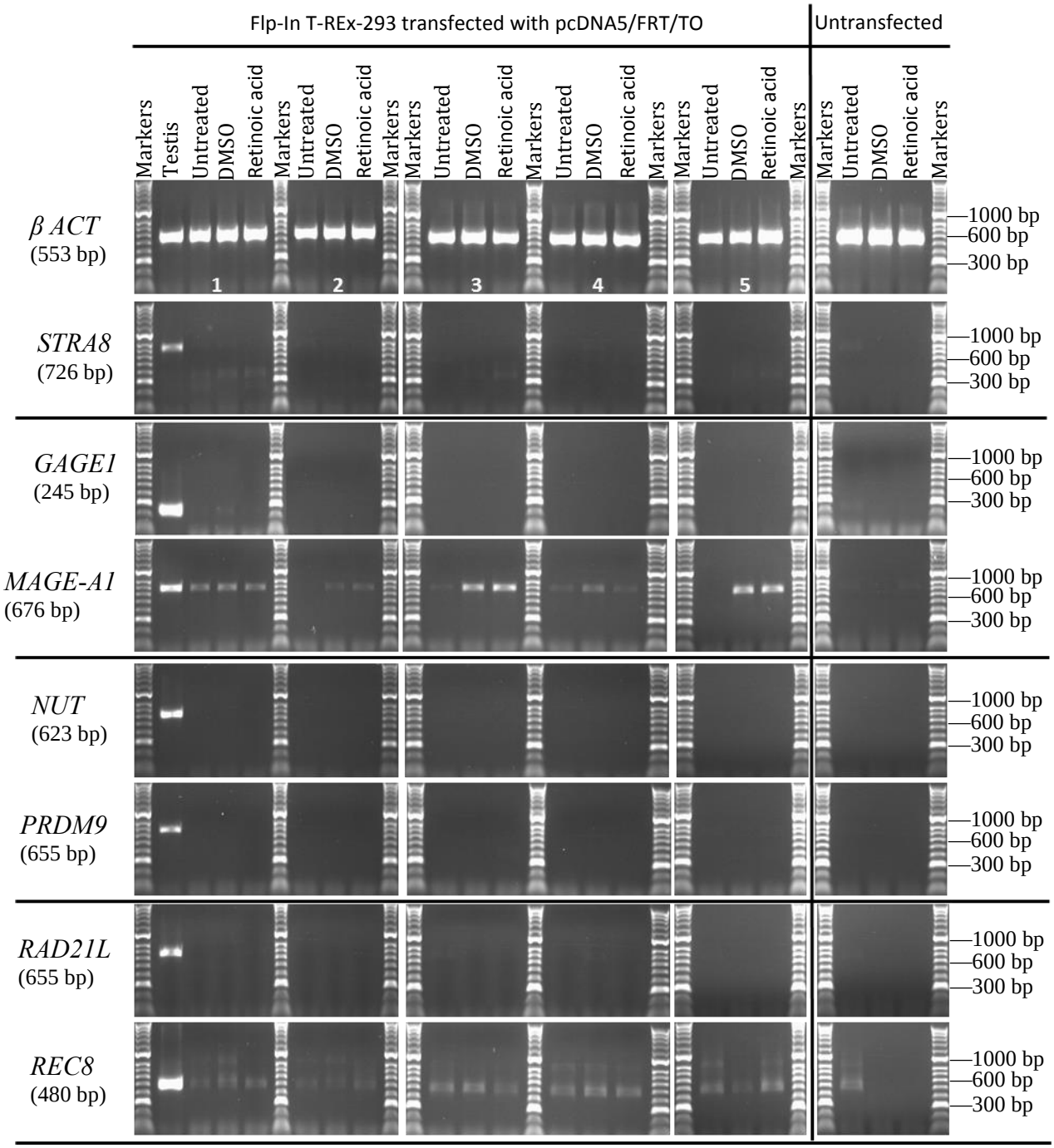


Figure 4.25 Analysis of meiotic genes expression in different Flp-In T-REx-293:: pcDNA5/FRT/TO cell lines treated with *all-trans* retinoic acid. RT-PCR analysis for *STRA8, GAGE1, MAGE-A1, NUT, PRDM9, RAD21L* and *REC8* mRNA expression in five different colonies (1-5) of Flp-In T-REx-293 cells that had been transfected with the pcDNA5/FRT/TO vector alone. Prior to RNA extraction, the cells were treated with 10 µM *all-trans* retinoic acids or with the DMSO vehicle alone for 24 hours. *β ACT* expression was used as a positive control for the cDNA samples. The expected amplicon size of each gene is shown on the left between brackets.

.

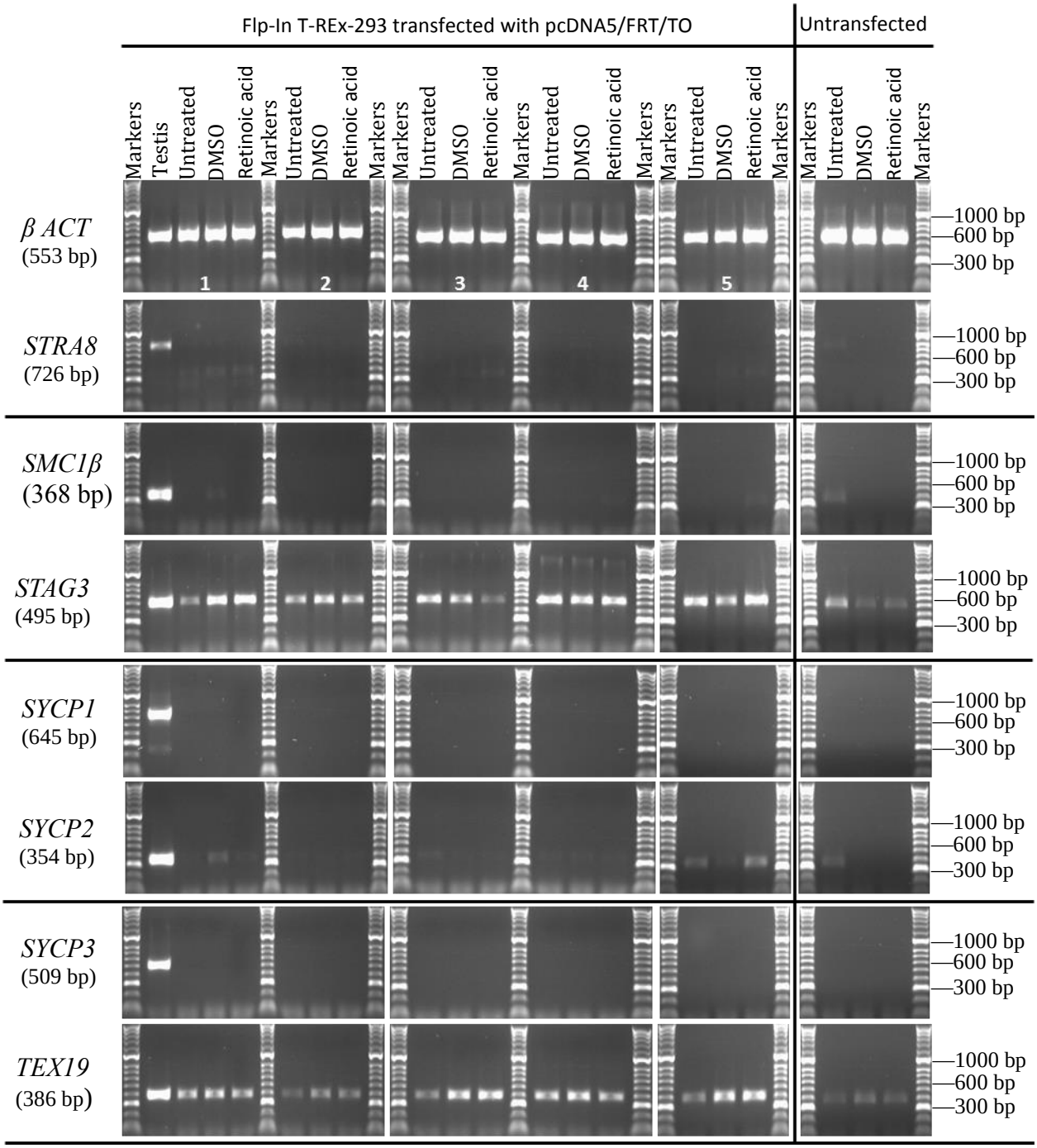


Figure 4.26 Analysis of meiotic genes expression in different Flp-In T-REx-293:: pcDNA5/FRT/TO cell lines treated with *all-trans* retinoic acid. RT-PCR analysis for *STRA8, SMC1β, STAG3, SYCP1, SYCP2, SYCP3* and *TEX19* mRNA expression in five different colonies (1-5) of Flp-In T-REx-293 cells that had been transfected with the pcDNA5/FRT/TO vector alone. Prior to RNA extraction, the cells were treated with 10 µM *all-trans* retinoic acids or with the DMSO vehicle alone for 24 hours. *β ACT* expression was used as a positive control for the cDNA samples. The expected amplicon size of each gene is shown on the left between brackets.

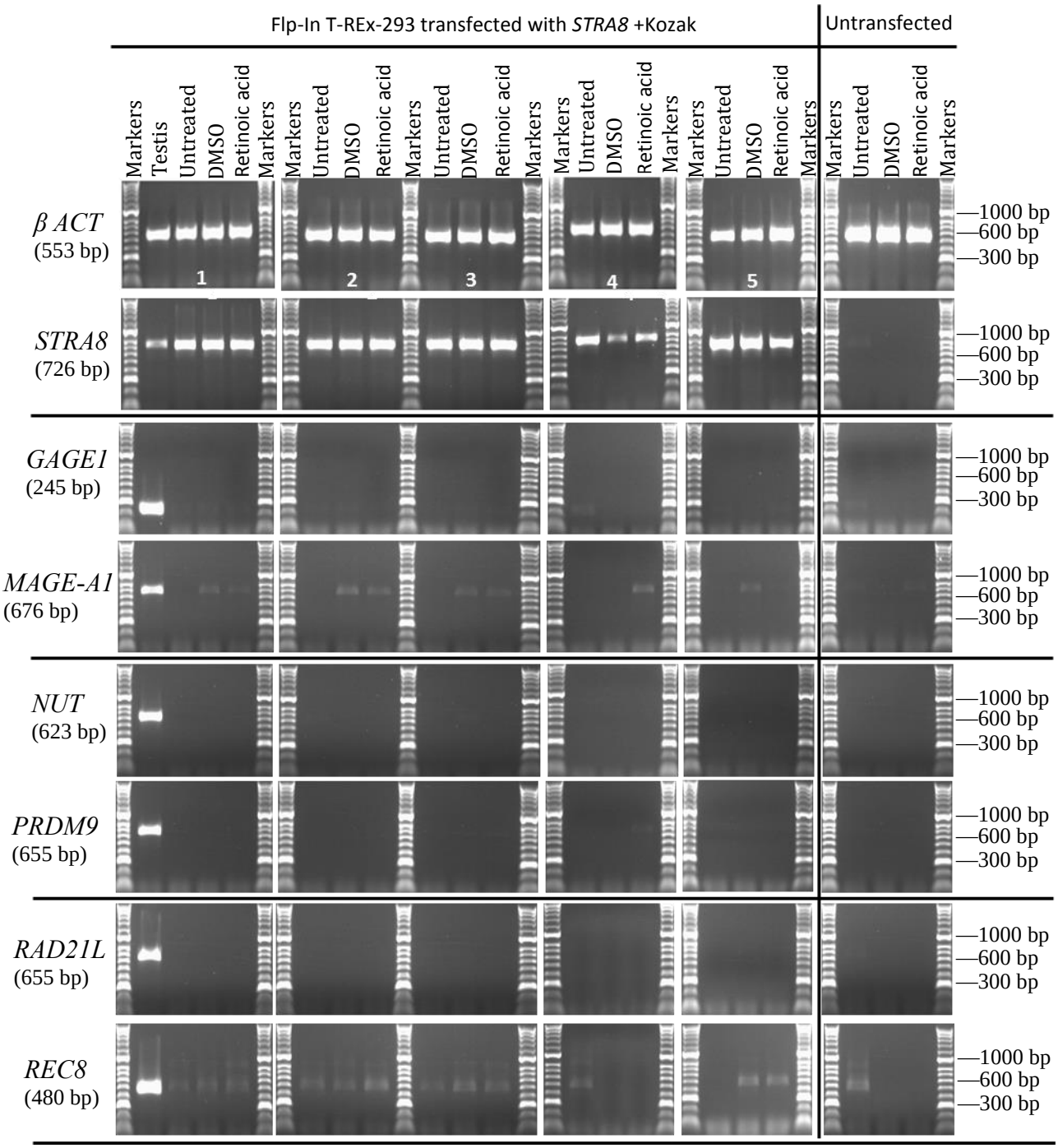


Figure 4.27 Analysis of meiotic genes expression in different Flp-In T-REx-293::*STRA8*+Kozak cell lines treated with *all-trans* retinoic acid. RT-PCR analysis for *GAGE1, MAGE-A1, NUT, PRDM9, RAD21L* and *REC8* mRNA expression in five different colonies (1-5) of Flp-In T-REx-293 cells that had been transfected with *STRA8* cDNA +Kozak. Prior to RNA extraction, the cells were treated with 10 µM *all-trans* retinoic acids or with the DMSO vehicle alone for 24 hours. *STRA8* expression was used as a positive control for the transfected cells with *STRA8*+Kozak. *β ACT* expression was used as a positive control for the cDNA samples. The expected amplicon size of each gene is shown on the left between brackets.

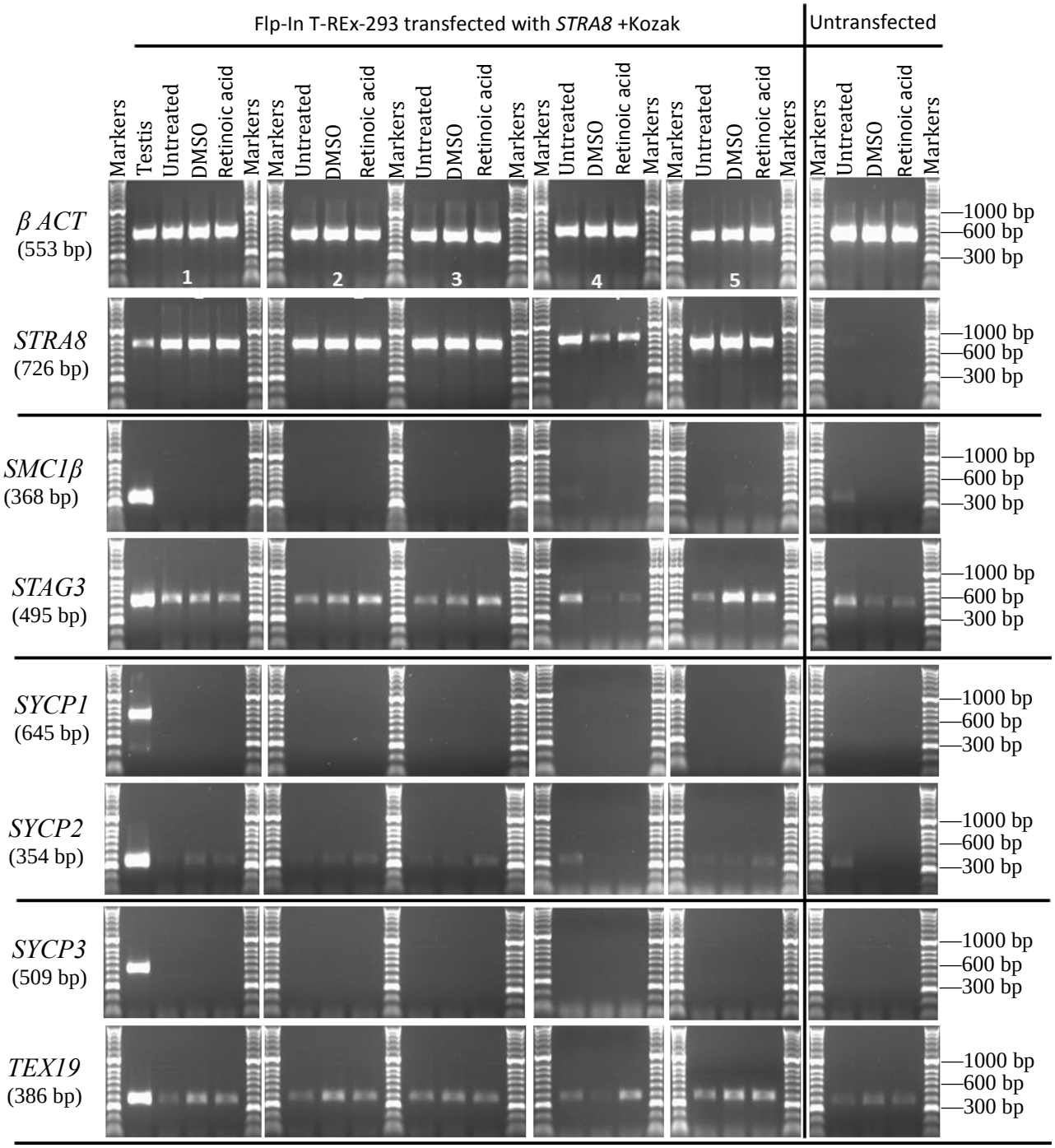


Figure 4.28 Analysis of meiotic genes expression in different Flp-In T-REx-293::*STRA8*+Kozak cell lines treated with *all-trans* retinoic acid. RT-PCR analysis for *SMC1β, STAG3, SYCP1, SYCP2, SYCP3* and *TEX19* mRNA expression in five different colonies (1-5) of Flp-In T-REx-293 cells that had been transfected with *STRA8* cDNA +Kozak. Prior to RNA extraction, the cells were treated with 10 µM *all-trans* retinoic acids or with the DMSO vehicle alone for 24 hours. *STRA8* expression was used as a positive control for the transfected cells with *STRA8*+Kozak. *β ACT* expression was used as a positive control for the cDNA samples. The expected amplicon size of each gene is shown on the left between brackets.

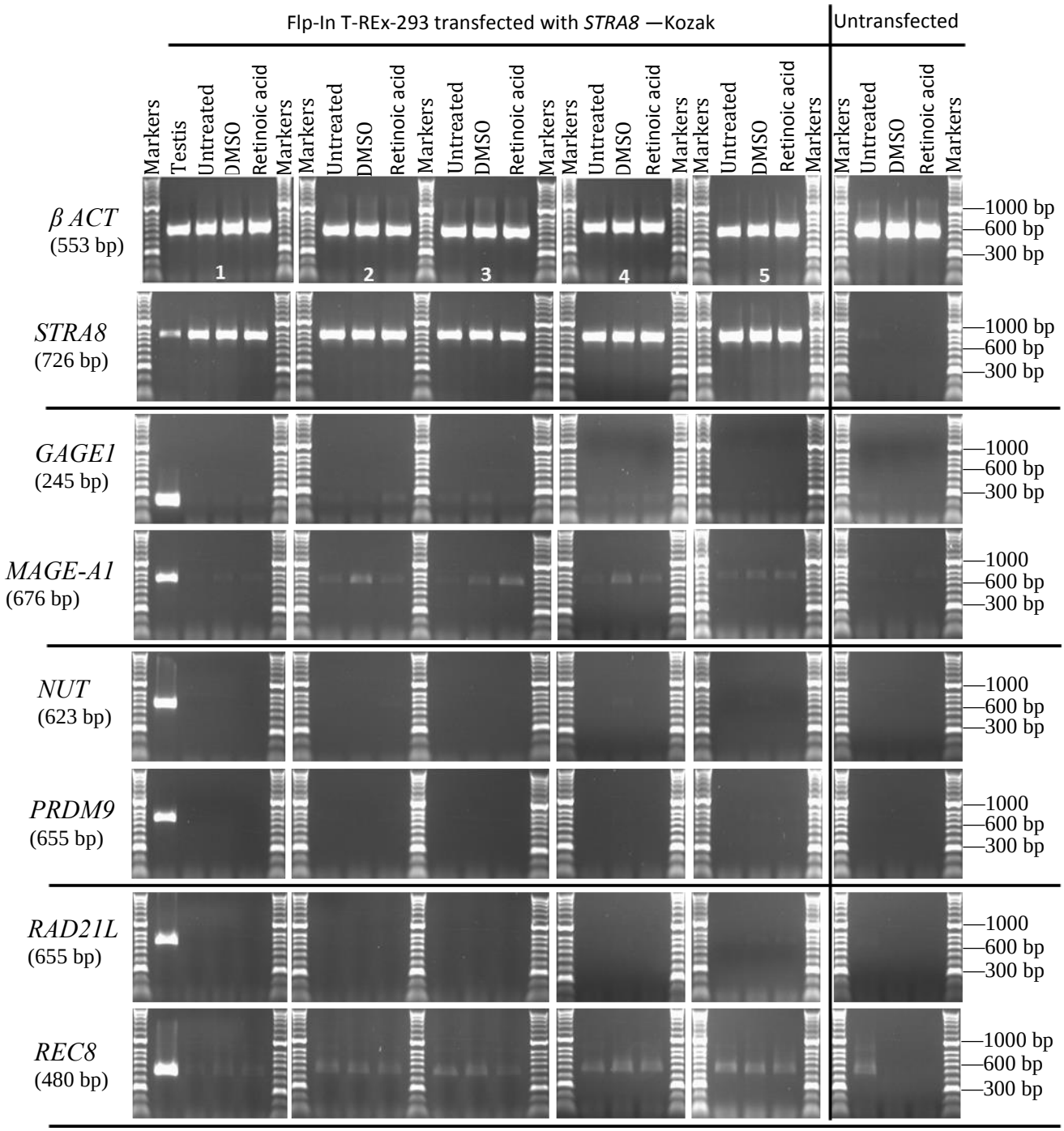


Figure 4.29 Analysis of meiotic genes expression in different Flp-In T-REx-293::*STRA8*–Kozak cell lines treated with *all-trans* retinoic acid. RT-PCR analysis for *GAGE1, MAGE-A1, NUT, PRDM9, RAD21L* and *REC8* mRNA expression in five different colonies (1-5) of Flp-In T-REx-293 cells that had been transfected with *STRA8* cDNA –Kozak. Prior to RNA extraction, the cells were treated with 10 µM *all-trans* retinoic acids or with the DMSO vehicle alone for 24 hours. *STRA8* expression was used as a positive control for the transfected cells with *STRA8*–Kozak. *β ACT* expression was used as a positive control for the cDNA samples. The expected amplicon size of each gene is shown on the left between brackets.

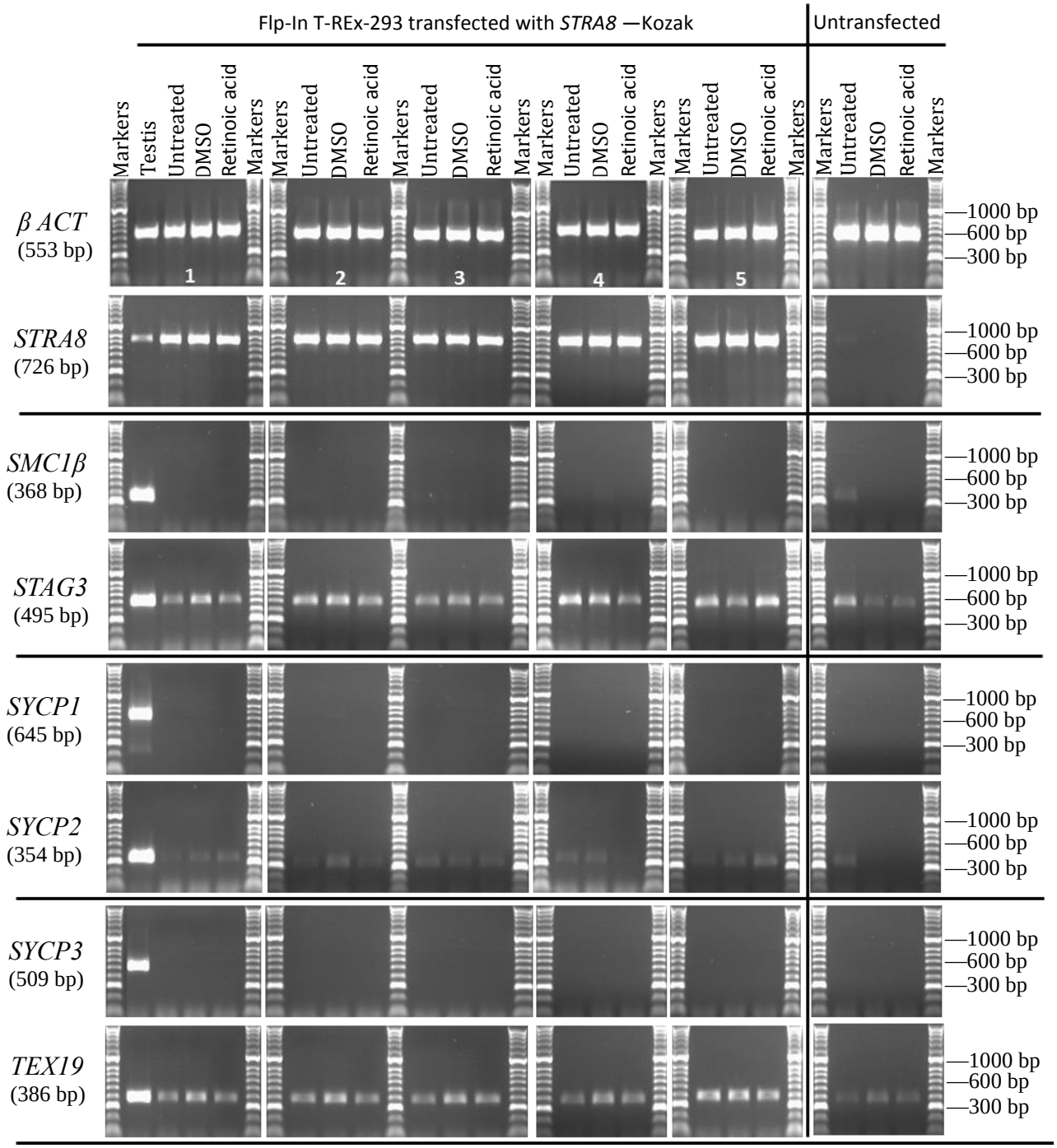


Figure 4.30 Analysis of meiotic genes expression in different Flp-In T-REx-293::*STRA8*–Kozak cell lines treated with *all-trans* retinoic acid. RT-PCR analysis for *SMC1β, STAG3, SYCP1, SYCP2, SYCP3* and *TEX19* mRNA expression in five different colonies (1-5) of Flp-In T-REx-293 cells that had been transfected with *STRA8* cDNA –Kozak. Prior to RNA extraction, the cells were treated with 10 µM *all-trans* retinoic acids or with the DMSO vehicle alone for 24 hours. *STRA8* expression was used as a positive control for the transfected cells with *STRA8*–Kozak. *β ACT* expression was used as a positive control for the cDNA samples. The expected amplicon size of each gene is shown on the left between brackets.

4.3 Discussion

4.3.1 Integration of the human *STRA8* gene into Flp-In T-REx-293 cells

The *STRA8* gene is a mammalian regulator of meiotic progression in both females and males (Anderson *et al*., 2008). STRA8 is required for transcriptional activation of REC8 and DMC1 (Anderson *et al*., 2008). To study the function of the human STRA8, we considered the integration of *STRA8* in a cell line that does not already express it to determine if the overexpression of *STRA8* in a human cell line leads to the induction of *REC8* or *DMC1* or the upregulation of other meiotic genes. The *STRA8*+Kozak and *STRA8*–Kozak cDNA inserts were subcloned into the mammalian expression vector, pcDNA5/FRT/TO. The reason for choosing this vector was because it comprises an important element for integration of the *STRA8* gene into the Flp-In T-REx-293 cell line. This element is hybrid human cytomegalovirus (CMV)/TetO2 promoter for high level tetracycline-regulated expression of any gene of interest in most mammalian cells (Andersson *et al*., 1989; Hillen and Berens, 1994).

All the cells that were transfected with pcDNA5/FRT/TO containing *STRA8*+Kozak or *STRA8*–Kozak were grown in medium containing hygromycin and β-galactosidase was examined in the cells before and after inserting pcDNA5/FRT/TO. The cells expressed β-galactosidase before inserting the pcDNA5/FRT/TO, whereas they did not express it after pcDNA5/FRT/TO insertion, which suggests successful *STRA8* integration.

4.3.2 Induction of *STRA8* expression by retinoic acid

The expression of *Stra8* in mice is related to the availability of RA in both sexes (Zhou *et al*., 2008). In the present study, the Flp-In T-REx-293 cell lines that had been integrated with the *STRA8* gene were treated with 5 µM or 10 µM RA to induce other genes that could induce the expression of *STRA8*. *GAGE1, MAGE-A1, NUT, PRDM9, RAD21L, REC8, SMC1β, STAG3, SYCP1, SYCP2, SYCP3* and *TEX19* meiosis genes were investigated in these cells that were treated with *all-trans* and *9-cis* retinoic acids. The purpose of the investigation was to determine which of these meiosis genes, if any, were upregulated. The RT-PCR results did not indicate any overexpression among these genes. This result could be due to the concentration of RA, to the duration of treatment or to the possibility that STRA8 could play other roles in Flp-In T-REx-293 cells. It might also indicate that *STRA8* expression alone is not sufficient to induce meiotic genes and could demonstrate that other factors are required.

A previous study investigated the expression of *Stra8* in adult mice testes after subcutaneous injection with 100 µl of 7.5 mg/ml *all-trans* retinoic acid for 24 hours; the *Stra8* transcript levels increased three-fold (Zhou *et al*., 2008). Another study implied that mRNA expression of *STRA8* was induced in human foetal testes subjected to 1 µM *all-trans* retinoic acid for 24 hours, but based on qRT-PCR, expression in other meiosis-associated genes, *SYCP3* and *DMC1*, there was no subsequent downstream gene activation (Childs *et al*., 2011). This could indicate that the role of *STRA8* in meiotic gene activation is not identical in humans and mouse. Indeed, recent work showing *REC8* expression in human non-testis tissue supports this (Feichtinger *et al*., 2012a).

In *Stra8*-deficient spermatogenic cells, the meiotic specific cohesin (Rec8) and synaptonemal complex protein (Sycp3) are not properly localised to the chromosomes when examined by immunohistochemical staining, compared to Rec8 and Sycp3 proteins in wild-type spermatogenic cells (Anderson *et al*., 2008; Baltus *et al*., 2006). This indicates that Stra8 may be required in the formation of meiotic cohesin and synaptonemal complexes in mouse. Neither up-regulation of Spo11 and Dmc1 nor the proteins were detected in Stra8-deficient testes (Anderson *et al*., 2008). Furthermore, DNA DSBs are marked by γ-H2AX and were not observed by immunostaining in Stra8-deficient testes; whereas, γ-H2AX staining was detected in several cells of wild-type testes showing that DNA DSBs are formed. Results for Spo11, Dmc1 and γ-H2AX strongly indicate that DSBs are not formed and there is a lack of meiotic recombination in spermatogenic cells of Stra8-deficient testes (Anderson *et al*., 2008). Lack of Stra8 in females gives rise to the repression of many events in meiosis, including DNA replication, sister chromosome cohesion, synapsis and meiotic recombination (Baltus *et al*., 2006).

Overexpression of 12 meiosis genes tested here was not observed by RT-PCR analysis using F1 and R1 primers and Flp-In T-REx-293 cell lines, which contained *STRA8* cDNA with and without Kozak sequences. These results could potentially indicate that the RT-PCR analysis used was not adequate for studying whether *STRA8* can induce a meiosis gene. In this chapter, the 12 meiosis genes were investigated in cell lines inserted with the *STRA8* cDNA and treated with retinoic acids; these genes were chosen at random. Many genes inside the cells can play fundamental roles in meiosis. Consequently, the *STRA8* gene could induce other meiosis genes that were not tested here or expression of *STRA8* alone was not sufficient to initiate a meiotic gene expression programme. Further analysis can be performed to study whether *STRA8* can induce any other meiosis genes; for example, *STRA8* gene was expressed

at low levels in the HEP-G2 cell line and we thought that treatment of these cells with retinoic acid might induce a *STRA8* transcript that could be detected using different techniques, such as RT-PCR, qRT-PCR and Western blotting. Furthermore, this experiment can be carried out using genes other than the 12 meiosis genes, including other CTA genes, like *NY-ESO-1* (Whitehurst, 2014).

In total, 7 of the 36 CTA genes identified by Feichtinger *et al*. (2012a) were expressed in HEP-G2 cells. These genes are: *C19orf67*, *C22orf33*, *ODF4*, *STRA8*, *TEX19*, *CXorf27* and *ZCCHC13*. We speculate that knockdown of the *STRA8* gene with siRNA might downregulate these genes in the HEP-G2 cell line after siRNAs treatment. This remains to be tested thoroughly.

No detectable levels of *STRA8* mRNA were observed in the NTERA2 cell line by RT-PCR analysis; however, Western blotting results indicated a clear band for the STRA8 protein product in this cell. The NTERA2 cell has the ability to differentiate into neurons as well as other cell types following exposure to retinoic acid (Andrews, 1984). Therefore, both NTERA2 and the *STRA8* gene are responsive to retinoic acid. The RT-PCR results indicated that the undifferentiated NTERA2 cells did not show mRNA expression for *STRA8*. Therefore, treatment of the NTERA2 cell line with retinoic acid followed by analysis of *STRA8* expression/protein levels would be informative.

4.4 Conclusion

Based on the RT-PCR screening performed in Chapter 3, the *STRA8* gene is considered to be a good CT gene, because its expression profile is clearly restricted to the normal testis among the normal tissues and in different type of cancers. The STRA8 protein plays a role in the progression of meiosis. In this chapter, *STRA8* was inserted into the human cell line to establish a stable cell line expressing the *STRA8* gene. The target of this insertion was to investigate the role of *STRA8* in a stable cell line. We demonstrated that the *STRA8* gene can be overexpressed in the human Flp-In T-REx-293 cell line, but the expression of other meiosis-specific genes did not increase after stimulation by retinoic acids. Therefore, further investigation is needed to study the function of this gene in the normal human testis and in cancer cells.

Chapter 5.0: Functional analysis of the human *C20orf201* gene

5.1 Introduction

Cancer is a serious disease that can lead to substantially increased mortality. This occurs not only due to the limited effectiveness of available treatments, but also because diagnosis usually takes place late in the tumour development (Abramovitz and Leyland-Jones, 2006; Miturski *et al*., 2002). For this reason, the identification of new cancer-specific biomarkers is essential to improve patient diagnosis (Baulande *et al*., 2014). The detection of cancer biomarkers will play a fundamental role in the early detection/screening of diseases, but also in other aspects such as the classification and staging of cancer development, prognosis assessment and assessment of patient responses to treatment. Diagnosis at the early and asymptomatic stage is considered to be the key to effective cancer treatment (Yang and Sweedler, 2014).

A recent study has suggested that cancer-testis (CT) genes might be useful biomarkers of cancer cells (van Duin *et al*., 2011). The profile of CTA gene expression has significance for the early detection and immunotherapy of cancer (Ghafouri-Fard and Modarressi, 2009). One CTA gene that shows potential as a good candidate cancer biomarker is the human *C20orf201* gene, based on the findings presented in Chapter 3 (see Figures 3.12 and 3.14); however, its actual function is not known.

Cancer cells naturally present antigenic peptides termed human leukocyte antigen (HLA) molecules on their surfaces; these peptides are normally recognised by specific cytotoxic T-lymphocytes. These cancer-associated antigens are potentially promising tools for the production of peptide vaccines against cancer (Kamata *et al*., 2013). For example, Kamata and his colleagues identified HLA class-I-binding peptides that originated from cancer-associated proteins expressed in human prostate cancer cells. This identification was performed by immunoproteomics and indicated that these isolated peptides could possibly be used for the generation of vaccines against prostate cancer. The 5- or 6-mer peptides (TKLSA and RLRYT) were selected among these peptides isolated from the HLA molecule located on the prostate cancer cells. However, an uncharacterised protein *C20orf201* was observed to be the source protein of the RLRYT peptide. A qRT-PCR analysis of *C20orf201* gene expression

revealed that this gene was significantly expressed in different cancer cell types, whereas no expression was found in non-cancerous tissues, except for the testis and brain (whole or cerebellum) (Kamata *et al*., 2013).

The human *C20orf201* gene is localised on chromosome 20 and has two different splice variants. The first splice variant contains 2 exons separated by one intron of 112 bp in length. However, the second splice variant comprises 3 exons separated by 2 introns of different lengths (Figure 5.1). The translational start of the gene is localised in exon 1 and stop codon TAA is in exon 2 or 3 for splice variant 1 or 2, respectively. The *C20orf201* cDNA containing the whole open reading frame (ORF) is 723 bp for splice variant 1 or 585 bp for splice variant 2 (GenBank accession number BC036837). The splice variants 1 and 2 encode 240 and 194 amino acid proteins, respectively.

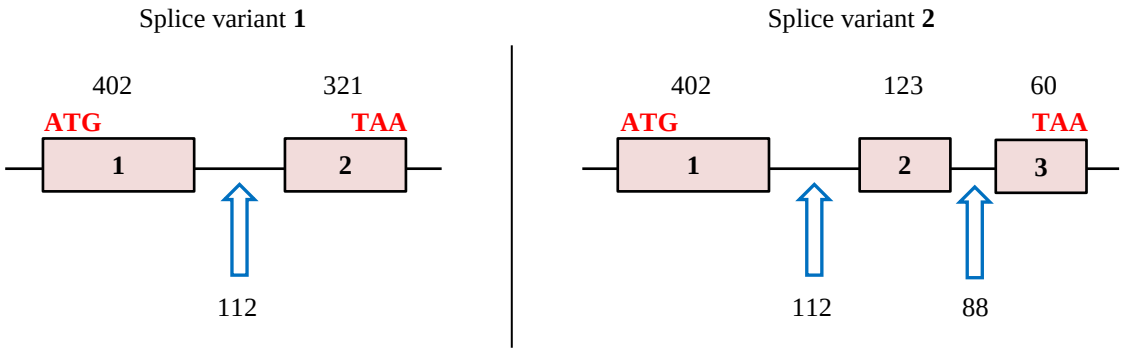


Figure 5.1 Exon and intron structure of the human *C20orf201* gene. Schematics show the structure of the *C20orf201* gene coding and non-coding regions for two different splice variants. The boxes and arrows represent exons and introns, respectively. The lengths of each exon and intron are given above the box in the exons and below the arrow in the introns. Approximate positions of start codon (ATG) and stop codon (TAA) are marked with red.

5.2 Results

5.2.1 Analysis of expression of the *C20orf201* gene, variant 2, in normal and cancer tissues

RT-PCR analysis was performed for *C20orf201* gene splice variant number 2 in a range of RNAs obtained from 21 human normal tissues and 33 cancer tissues and cell lines. RT-PCR intron-spanning primers were designed for *C20orf201* gene splice variant 2 (refer to Table 2.2). The cDNAs were synthesised from the RNAs and *β-Actin* was used in the RT-PCR screening as a positive control for the cDNA quality.

Two transcript variants have been identified for *C20orf201* gene. The RT-PCR screening of the splice variant 1 was previously described in Chapter 3 (see Figures 3.12 and 3.14). The expression of *C20orf201*, variant 2, is analysed in this chapter. The expression of *C20orf201* was observed in normal testis, CNS tissues and trachea (Figure 5.2), which suggests that this variant is testis-CNS-selective, based on its expression pattern. This variant was investigated further in multiple cancer tissues/cell lines, and importantly, it was detailed in 13 different cancers (Figure 5.3).

5.2.2 Cloning of *C20orf201* cDNA into the mammalian pcDNA5/FRT/TO vector for over expression

The aim of cloning *C20orf201* cDNA is to address the function of this gene in cancer cells. The mammalian expression vector, pcDNA5/FRT/TO (refer to Figure 4.2), was used for the cloning of the *C20orf201* variant 1 with and without addition of the Kozak consensus sequence. The Kozak sequence is added to the *C20orf201* cDNA sequences to stimulate the translation process (Kozak, 1991). Constructs with and without Kozak sequences were generated to potentially produce cell lines that differed in the amounts of C20orf201 protein produced. The *Hin*dIII restriction enzyme was used for cloning of *C20orf201*–Kozak and *C20orf201*+Kozak, since it does not appear in *C20orf201* cDNA sequences and is found at the multiple cloning site (MCS) of the pcDNA5/FRT/TO vector.

The cloning of *C20orf201*+Kozak was initiated by amplification of *C20orf201* cDNA from the pCMV6-AC::*C20orf201* vector using C20orf201 F5+Kozak and R5 primers. The amplified fragment was 726 bp, based on DNA sequencing, which confirmed the cDNA sequences of the gene. The purified PCR insert was subcloned into the pcDNA5/FRT/TO vector after digestion with the same restriction enzyme, *Hin*dIII. The vector was analysed on

agarose gel before and after *Hin*dIII digestion in order to confirm the digestion efficiency of the vector, which yielded four molecules with the undigested vector (Figure 5.4A) and a single fragment of about 5137 bp with the digested vector (Figure 5.4B left). The insert was separated on a gel after digestion and purified to confirm the presence of a single band of about 726 bp, as predicted by DNA sequencing (Figure 5.4B right).

In contrast, *C20orf201* cloning without the Kozak sequence was initiated by amplification of *C20orf201* cDNA from the pCMV6-AC::*C20orf201* vector using C20orf201 F6 and R5 primers. The amplified fragment was subcloned into the pcDNA5/FRT/TO vector after digestion with *Hin*dIII restriction enzyme and purification. The vector was analysed on agarose gel before and after *Hin*dIII digestion in order to confirm the digestion efficiency of the vector, which produced four molecules with the undigested vector (Figure 5.5A) and one fragment of about 5137 bp with the digested vector (Figure 5.5B left). Conversely, the *C20orf201* insert was loaded onto an agarose gel following digestion and purification and displayed a single band of about 723 bp, as predicted by DNA sequencing (Figure 5.5B right).

The purified *C20orf201* inserts (with and without Kozak) were ligated into the purified digested pcDNA5/FRT/TO plasmid using quick T4 DNA ligase. The recombinant plasmids were transformed into competent *E. coli* DH5α. The clones were analysed by PCR screening performed on 20 randomly selected colonies. Primers internally positioned within the *C20orf201* gene, namely C20orf201 F1 and R1, were used in the PCR screening. The PCR results indicated that 8 of the *E. coli* colonies had the gene of interest (*C20orf201*) with Kozak (Figure 5.6) and 6 colonies without Kozak (Figure 5.7). The PCR primers were tested using testis cDNA, which gave a band lower than the expected size, with about 88 bp missing, which was confirmed by DNA sequencing.

The pcDNA5/FRT/TO::*C20orf201*+Kozak and pcDNA5/FRT/TO::*C20orf201*–Kozak plasmids were isolated from *E. coli*. The presence of *C20orf201* inserts (+Kozak or –Kozak) was verified by restriction enzyme digestion, which cut once in the vector and once in the insert. Therefore, 8 samples of purified DNA from the pcDNA5/FRT/TO::*C20orf201*+Kozak plasmids and 6 samples of purified DNA from the pcDNA5/FRT/TO::*C20orf201*–Kozak plasmids were digested with *Hin*dIII. Two fragments were observed for each digested plasmid. The upper fragments gave the expected vector size of 5137 bp, while the lower bands gave fragments of 726 bp and 723 bp for the *C20orf201*+Kozak (Figure 5.8) and *C20orf201*–Kozak, respectively (Figure 5.9).

Two purified plasmids (pcDNA5/FRT/TO::*C20orf201*+Kozak and pcDNA5/FRT/TO::*C20orf201*–Kozak) were sent for DNA sequencing using two sets of primers, CMV (forward) and BGH (reverse) to confirm that the *C20orf201* inserts were subcloned in the correct orientation for expression and contained an ATG start codon and a TAA stop codon. The sequences obtained from the DNA sequencing were compared with the original sequence via the NCB1 (BLAST). They were 100% identical to *C20orf201,* without any mutation evident in the whole length of the *C20orf201* gene.

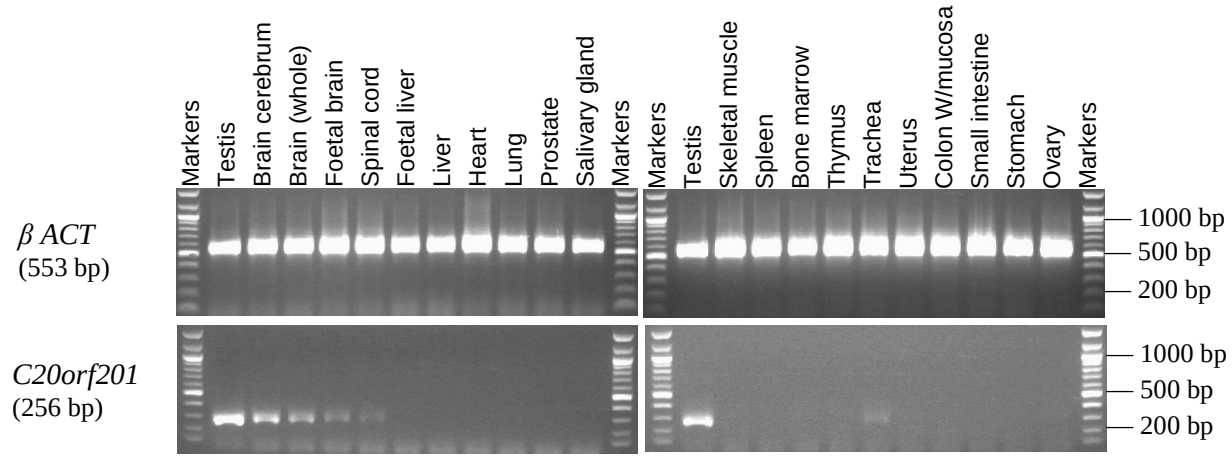


Figure 5.2 RT-PCR analysis of the messenger RNA for *C20orf201* gene splice variant 2 with normal human tissues. Agarose gel image illustrating the RT-PCR assays for the *C20orf201* gene. cDNAs were isolated from the total RNA from 21 normal tissues. The expression of *C20orf201* was observed in the normal testis, central nervous system tissues and trachea. *β ACT* expression was used as a positive control for the cDNA samples. The expected amplicon size of each PCR product is shown on the left between brackets.

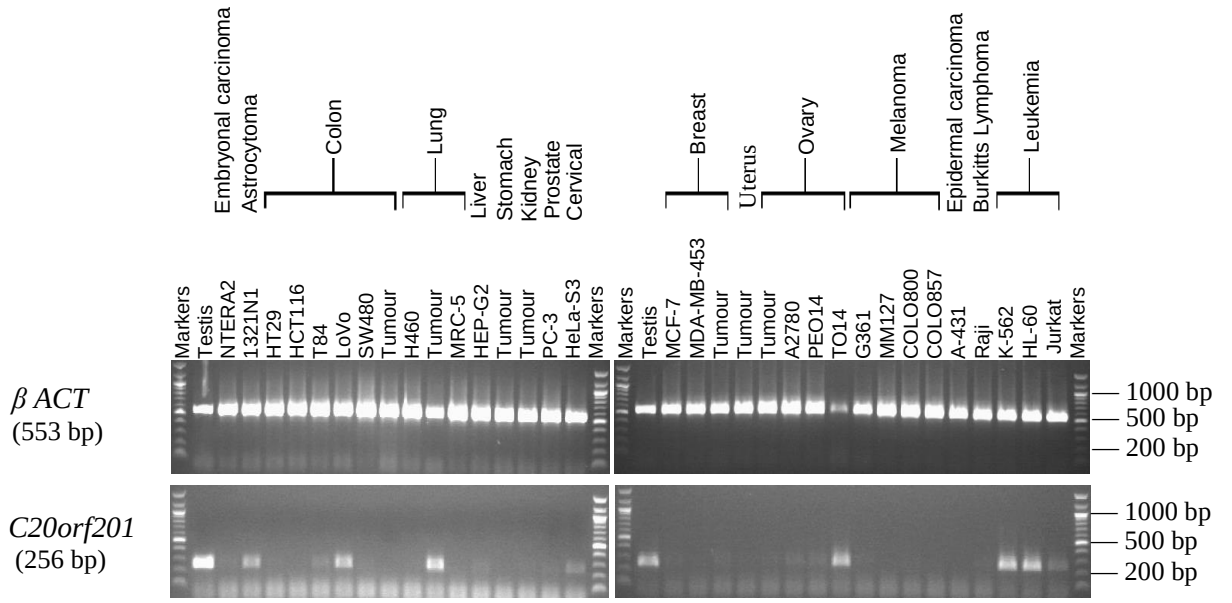


Figure 5.3 RT-PCR analysis of the messenger RNA for *C20orf201* gene splice variant 2 with human cancer cell lines and tumour tissues. Agarose gel image illustrating the RT-PCR assays for the *C20orf201*. cDNAs were isolated from the total RNA from 33 cancer tissues and cell lines. This gene was expressed in the normal testis tissue and in different cancer samples. *β ACT* expression was used as a positive control for the cancer cDNA samples. The expected amplicon size of PCR product is shown on the left between brackets.

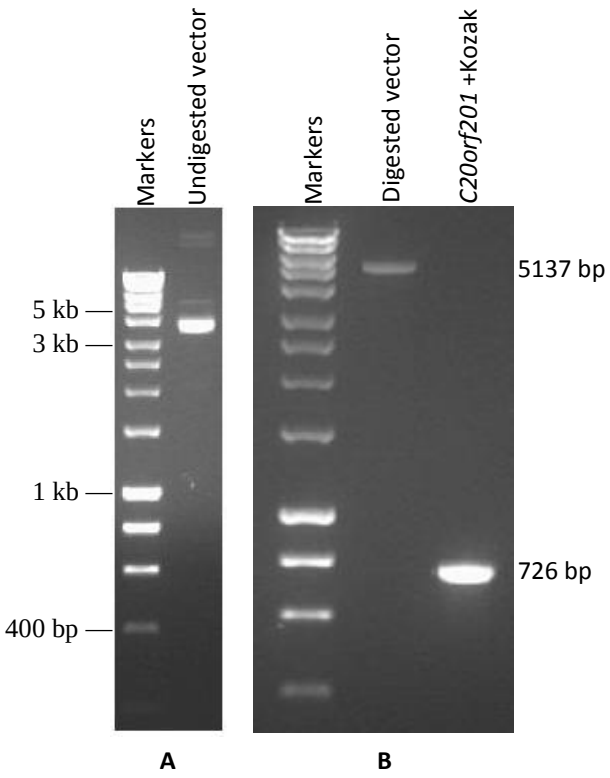


Figure 5.4 Amplification of *C20orf201*+Kozak from recombinant pCMV6-AC::*C20orf201* vector. Agarose gel image shown in panel **(A)** is undigested pcDNA5/FRT/TO, mammalian expression vector, used as a control to compare to the digestion vector. In contrast, the left side of panel **(B)** shows the pcDNA5/FRT/TO vector after digestion by *Hin*dIII restriction enzyme and purification. The enzyme linearizes the 5137 bp plasmid into one single fragment. The right side of gel B displays the amplification of *C20orf201* cDNAs +Kozak sequence from pCMV6-AC::*C20orf201* vector, which was already cloned with the full length open reading frame of *C20orf201* by using F5 and R5 primers. This band was digested with the same restriction enzyme, *Hin*dIII, and then underwent purification.

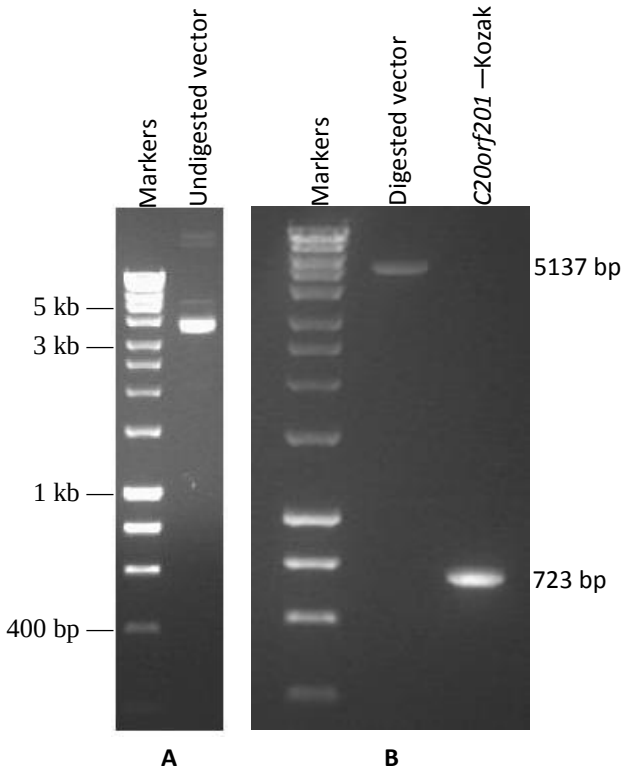


Figure 5.5 Amplification of *C20orf201*–Kozak from recombinant pCMV6-AC::*C20orf201* vector. Agarose gel image shown in panel **(A)** is undigested pcDNA5/FRT/TO, mammalian expression vector used as a control to compare to the digestion vector. In contrast, the left side of the panel **(B)** shows the pcDNA5/FRT/TO vector after digestion by *Hin*dIII restriction enzyme and purification. The enzyme linearizes the 5137 bp plasmid into one single fragment. The right side of gel B displays the amplification of *C20orf201* cDNAs –Kozak sequence from pCMV6-AC::*C20orf201* vector, which was already cloned with the full length open reading frame of *C20orf201* by using F6 and R5 primers. This band was also digested with the same restriction enzyme, *Hin*dIII, and then underwent purification.

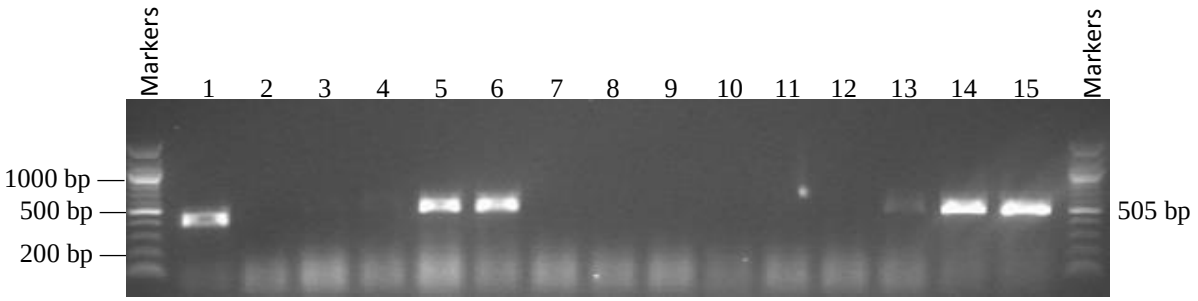


Figure 5.6 PCR screening of colonies for cloning of *C20orf201*. Agarose gel image showing PCR screening of *E. coli* colonies for cloning of *C20orf201*+Kozak into pcDNA5/FRT/TO using an internal primer for the *C20orf201* gene (F1,R1). The first band (lane 1) displays the presence of *C20orf201* in the testis cDNA to confirm that the primers worked. The samples in lanes 4-6 and 13-15 have an insert of the expected size of 505 bp for *C20orf201*.

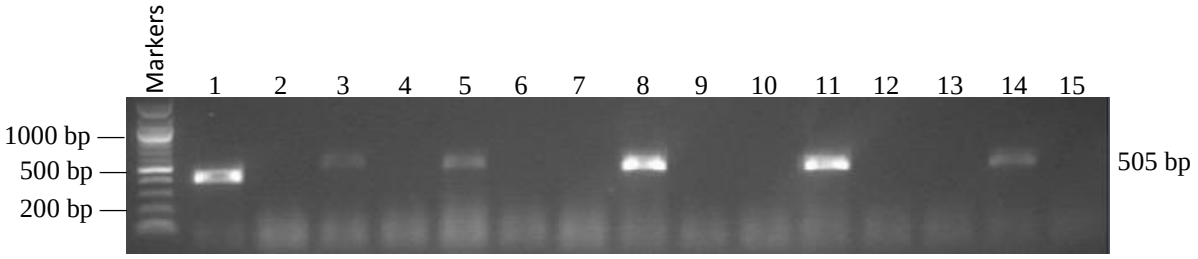


Figure 5.7 PCR screening of colonies for cloning of *C20orf201*. Agarose gel image showing PCR screening of *E. coli* colonies for cloning of *C20orf201*–Kozak into pcDNA5/FRT/TO using an internal primer of *C20orf201* gene (F1,R1). The first band (lane 1) displays the presence of *C20orf201* in the testis cDNA to confirm that the primers worked. The samples in lanes 3, 5, 8, 11 and 14 have an insert of the expected size of 505 bp for *C20orf201*.

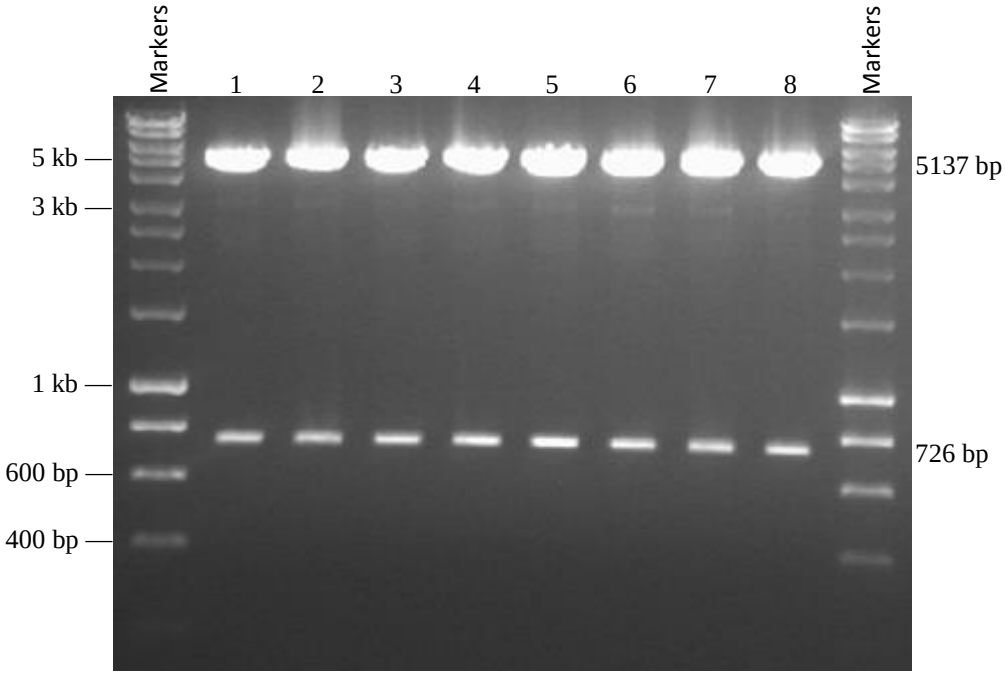


Figure 5.8 Digestion of recombinant plasmids. Agarose gel image showing digestion of the recombinant purified pcDNA5/FRT/TO vectors by the *Hin*dIII restriction enzyme. Two bands are formed for all recombinants. The upper bands demonstrate a pcDNA5/FRT/TO vector of 5137 bp, while the lower bands demonstrate the full open reading frame of *C20orf201* gene +Kozak with a size of 726 bp. All the minipreps in lanes 1-8 show successful cloning of *C20orf201* cDNA.

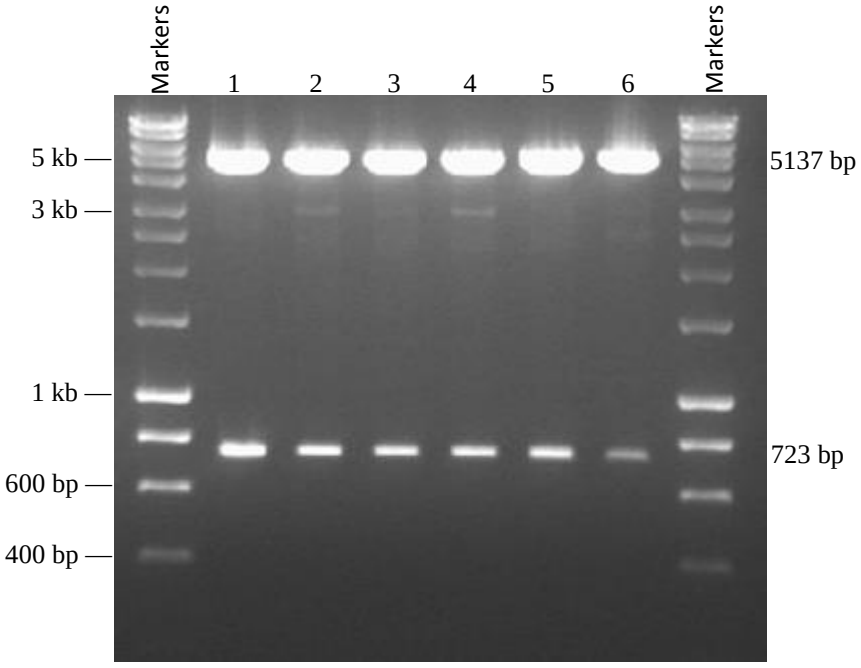


Figure 5.9 Digestion of recombinant plasmids. Agarose gel image showing the digestion of recombinant purified pcDNA5/FRT/TO vectors by the *Hin*dIII restriction enzyme. Two bands are formed for all recombinants. The upper bands demonstrate a pcDNA5/FRT/TO vector of 5137 bp, while the lower bands demonstrate the full open reading frame of *C20orf201* gene –Kozak with a size of 723 bp. The minipreps in lanes 1-6 are all successful clones of *C20orf201* cDNA.

5.2.3 Integration of pcDNA5/FRT/TO containing *C20orf201* cDNA into the Flp-In T-REx-293 cell line

5.2.3.1 Generation of a stable cell line

Insertion of *C20orf201* into the genome of the Flp-In T-REx-293 requires pOG44 vector (refer to Figure 4.10). In this study, the Flp-In T-REx-293 was co-transfected with a DNA mixture containing *C20orf201* cDNA (+Kozak or –Kozak) in pcDNA5/FRT/TO and pOG44 at a ratio of 9:1 (pOG44: pcDNA5/FRT/TO). A negative control was set up via transfection of the Flp-In T-REx-293 cell line with pOG44 and pcDNA5/FRT/TO alone without the gene of interest. In brief, a non-transfected Flp-In T-REx-293 cell line was grown in medium supplemented with 200 µg/ml zeocin and 10 µg/ml blasticidin. 48 hours following transfection, the cells were refreshed with media containing 100 µg/ml of hygromycin instead of zeocin.

Insertion of pcDNA5/FRT/TO containing the *C20orf201* gene with or without the Kozak sequence or the vector alone into the genome of the Flp-In T-REx-293 cell line at the FRT site brings the P_{SV40} promoter and the start codon (ATG) (from pFRT/*lacZeo* plasmid) frame with the hygromycin gene. Therefore, the hygromycin gene becomes active and the *lacZ-Zeocin* gene is inactivated (see Figure 4.9). Stable cell lines can be selected for hygromycin resistance, zeocin sensitivity and lack of β-galactosidase activity. Following integration, the *C20orf201* gene is expressed under the control of the (CMV)/TetO2 promoter. The majority of the cells did not survive due to the presence of hygromycin, which was used for the selection of a stable cell line. After pre-screening by viewing with a light microscope, single hygromycin resistant clones were selected for each construct (Figure 5.10).

5.2.3.2 Evaluation of different integrant Flp-In T-REx-293 cell lines

Two different methods were used to analyse the successful integration of pcDNA5/FRT/TO::*C20orf201* into the genome of the Flp-In T-REx-293 cell line at the FRT site. The first method involved screening for β-galactosidase activity. Confluent cells were analysed using a β-Gal staining kit in media containing hygromycin or zeocin for the transfected and non-transfected cells, respectively. The *lacZ* gene product, β-galactosidase, can be visualized by β-Gal staining of the cell. This activity was determined in triplicate for each cell line and light microscope images were obtained. The positive control for *lacZ*-staining (untransfected Flp-In T-REx-293 cell line) contained only blue cells, indicating active *LacZ* expression. In contrast, none of the cells in the clonal cell lines were blue,

indicating that the insertion of the *C20orf201* constructs was successful (Figure 5.11).

The second method used was PCR screening of the integrant Flp-In T-REx-293. Genomic DNA was isolated from the non-transfected and transfected Flp-In T-REx-293 cells. *β*-actin was used as a positive control for the genomic DNA samples. PCR products were found in all transfected cells that integrated with the pcDNA5/FRT/TO::*C20orf201*+Kozak, the pcDNA5/FRT/TO::*C20orf201*–Kozak or the empty pcDNA5/FRT/TO; however, the non-transfected cells did not display any PCR products when using primers located in the P_{SV40} as a forward primer and the hygromycin gene as a reverse primer. Primers located in the *C20orf201* gene (forward) and in the *LacZ-Zeocin* gene (reverse) were also used. The PCR results using C20orf201 and LacZ-Zeocin primers showed products only for the Flp-In T-REx-293 that had been integrated with pcDNA5/FRT/TO::*C20orf201*+Kozak or with pcDNA5/FRT/TO::*C20orf201*–Kozak (Figure 5.12).

The results of both methods (β-galactosidase activity and PCR screening) suggest that the *C20orf201* gene was correctly inserted into the genomic DNA of the Flp-In T-REx-293 cells.

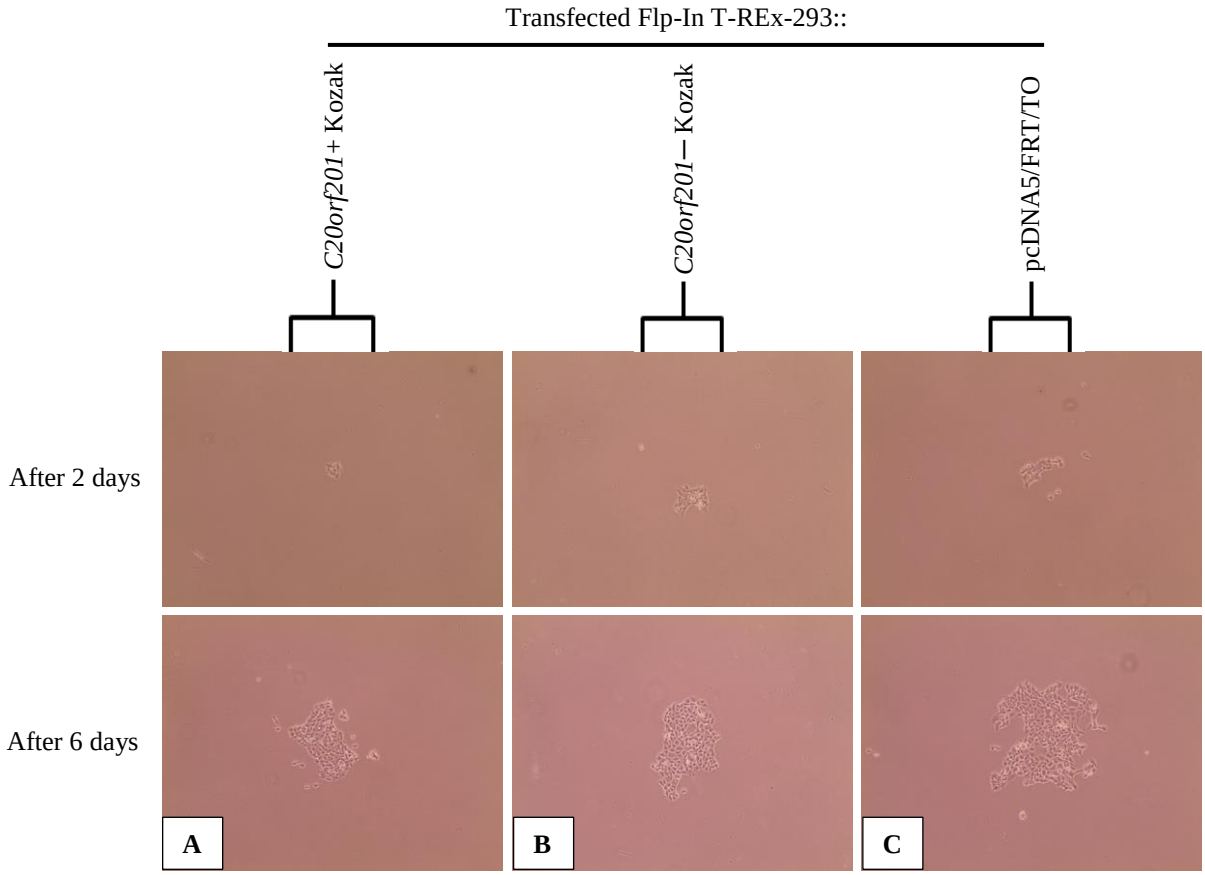


Figure 5.10 Examples of individual hygromycin resistance colonies. These images (panels **(A)** and **(B)**) show examples of a single colony still growing after hygromycin treatment, suggesting these cells have integrated with *C20orf201*+Kozak and *C20orf201*−Kozak, respectively. Panel **(C)** shows an example of cells that have integrated with pcDNA5/FRT/TO alone, without the target gene, and these cells served as a negative control.

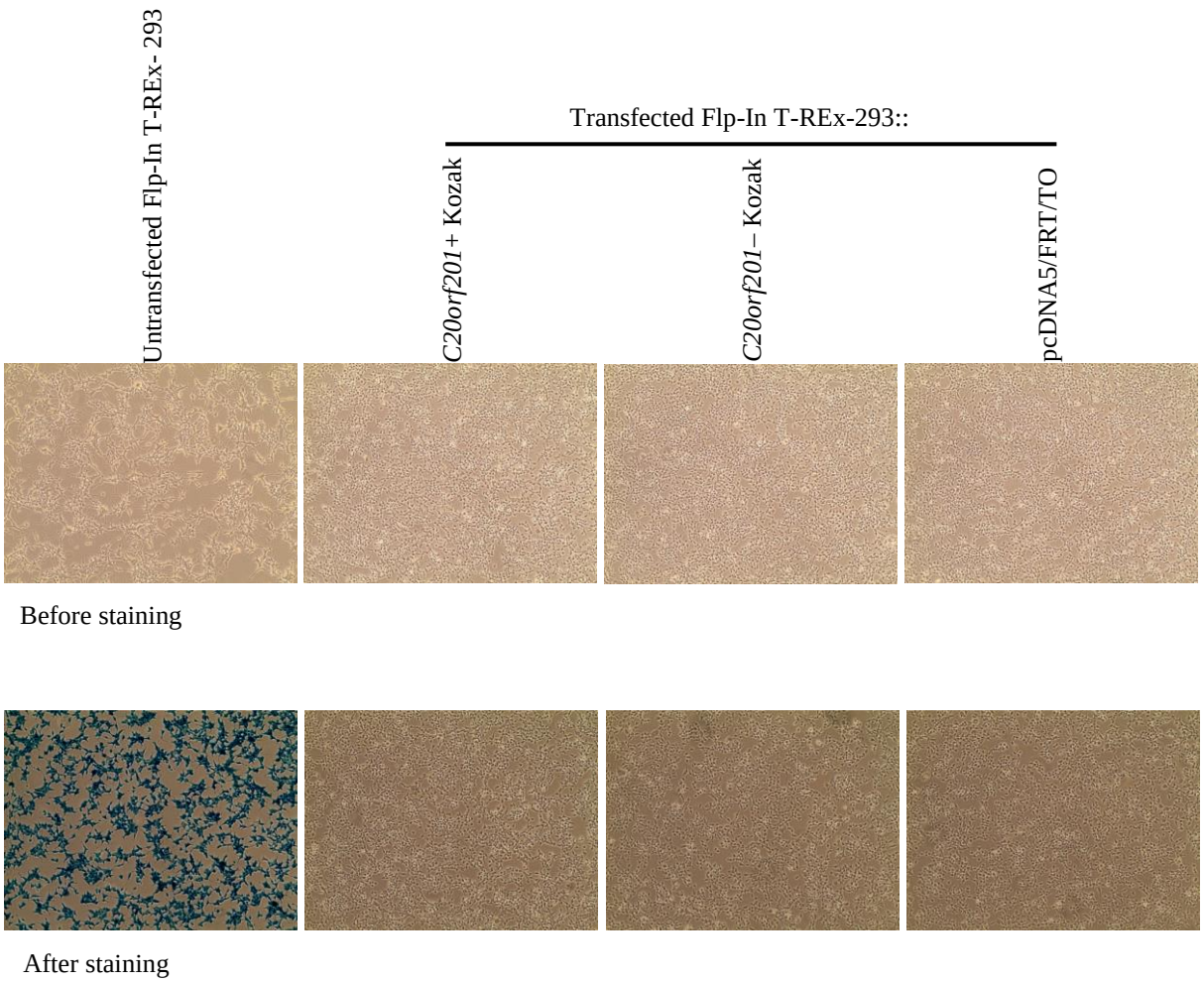


Figure 5.11 Assay for β-galactosidase Activity. These images show the untransfected cells of Flp-In T-REx-293 cells only produce β-galactosidase, but the transfected cells that integrated with different inserts of *C20orf201*+Kozak or *C20orf201*–Kozak or pcDNA5/FRT/TO do not produce β-galactosidase, as assessed by the colour of the cells following staining.

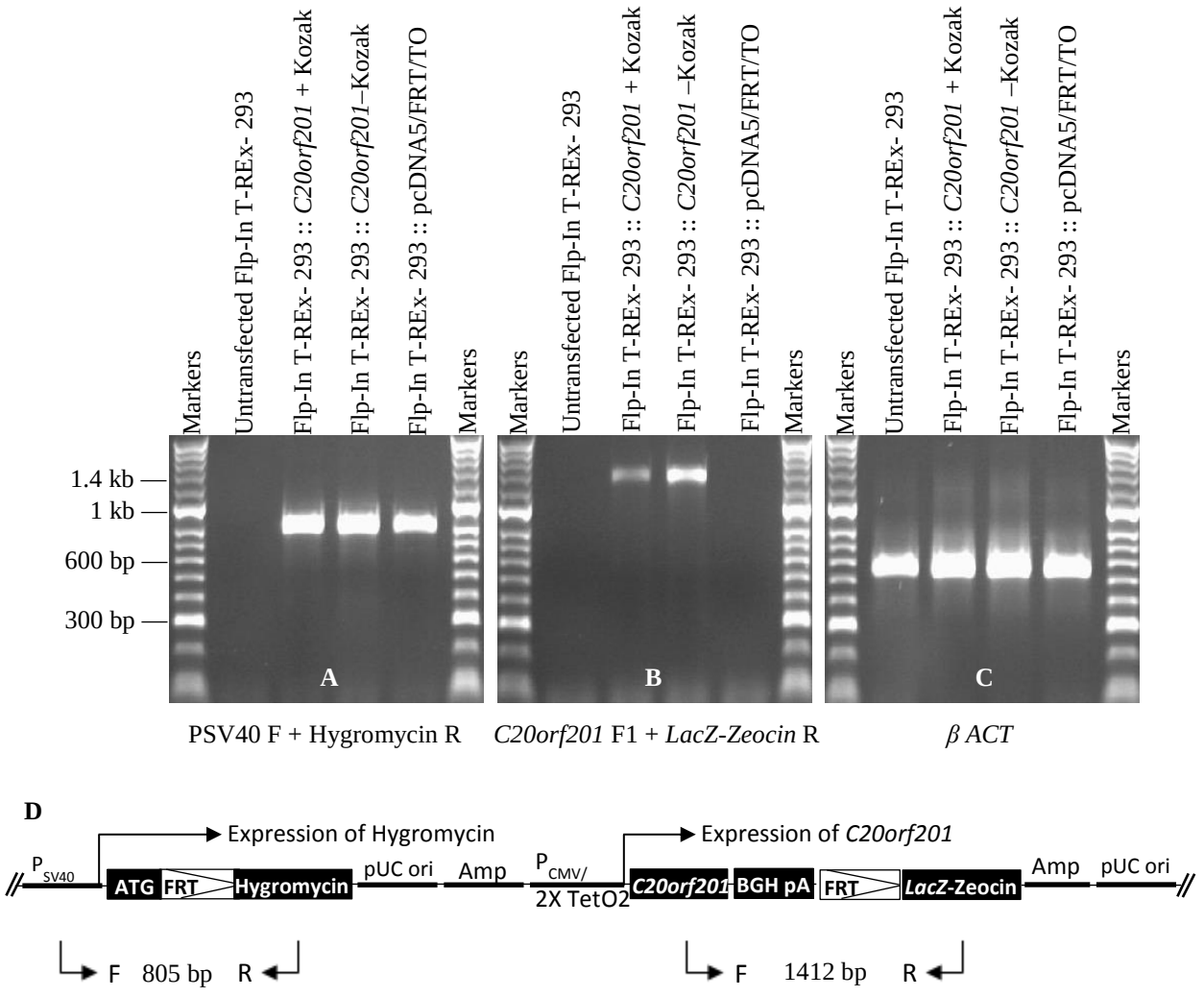


Figure 5.12 PCR screening of different *C20orf201* integrations into Flp-In T-REx-293 cell lines. Panel **(A)** shows the presence of the *C20orf201* gene in the transfected cells when using PSV40 as a forward primer and hygromycin as a reverse primer. Panel **(B)** shows PCR products of Flp-In T-REx-293 with the *C20orf201* genes when using primers in the *C20orf201* gene (forward) and in the *LacZ-Zeocin* gene (reverse). PCR using primers within the *β*-actin gene are shown as a positive control for the genomic DNA samples (panel **(C)**). The diagram in panel **(D)** shows the localisation of *C20orf201* cDNAs in the genome of Flp-In T-REx-293 cell line.

5.2.3.3 Evaluation of the *C20orf201* expression level in the Flp-In T-REx-293 cell line

The expression of the *C20orf201* gene was analysed using both RT-PCR and qRT-PCR. *C20orf201* expression was induced by the addition of tetracycline to the culture cells (2 µg/ml) and cDNAs were synthesised from the mRNAs of these cell lines. Two negative controls were used: untransfected cells and cells transfected with empty pcDNA5/FRT/TO. *β*-actin was used as a positive control for the cDNA samples. RT-PCR results confirmed the expression of the *C20orf201* gene (using F1 and R1 primers) in the cell lines that had been integrated with either pcDNA5/FRT/TO::*C20orf201*+Kozak or pcDNA5/FRT/TO::*C20orf201*–Kozak (Figure 5.13). However, higher levels of PCR products were observed for *C20orf201* in the cells treated with tetracycline relative to the same cells with no treatment. The mRNA levels of the *C20orf201* gene were detectable, indicating weak expression, in both negative controls using the same primers. The results in Figure 5.13 also show two distinct bands (18S and 28S ribosomal RNAs), which suggest good quality RNAs for all Flp-In T-REx-293 cell lines.

qRT-PCR was carried out for the *C20orf201* gene in different Flp-In T-REx-293 cell lines using C20orf201 F3 and R3 primers, while commercial positive primers were used. The NRT (no reverse transcriptase) and NTC (no template control) negative controls did not give a Cq (quantification cycle) reading, which suggests no non-specific background and no genomic DNA contamination (Table 5.1). The qRT-PCR results show a higher level of *C20orf201* expression in the Flp-In T-REx-293::*C20orf201*+Kozak and Flp-In T-REx-293::*C20orf201*–Kozak cells with tetracycline than in the Flp-In T-REx-293::pcDNA5/FRT/TO cells before treatment with tetracycline (Figure 5.14). In contrast, the mRNA levels were distinguished between cells containing *C20orf201*+Kozak or *C20orf201*–Kozak in the absence of tetracycline and the negative control with and without addition of tetracycline (Figure 5.15). This demonstrated that the expression was higher in cells containing the *C20orf201* gene, even though no tetracycline was added, compared to negative control (Figure 5.15). The RT-PCR results shown in Figure 5.13 suggest that the Flp-In T-REx-293::pcDNA5/FRT/TO cell line was negative for *C20orf201* expression. The intensity of the *C20orf201* PCR band was stronger in cell lines that had been integrated with *C20orf201* after treatment with tetracycline than in the other cell lines with no treatment. Therefore, these qRT-PCR results corresponded to the RT-PCR results previously observed (Figure 5.13). The *C20orf201* qRT-PCR results were normalised to the qRT-PCR results for *Lamin A* and *β ACT*; the results are shown in Table 5.1.

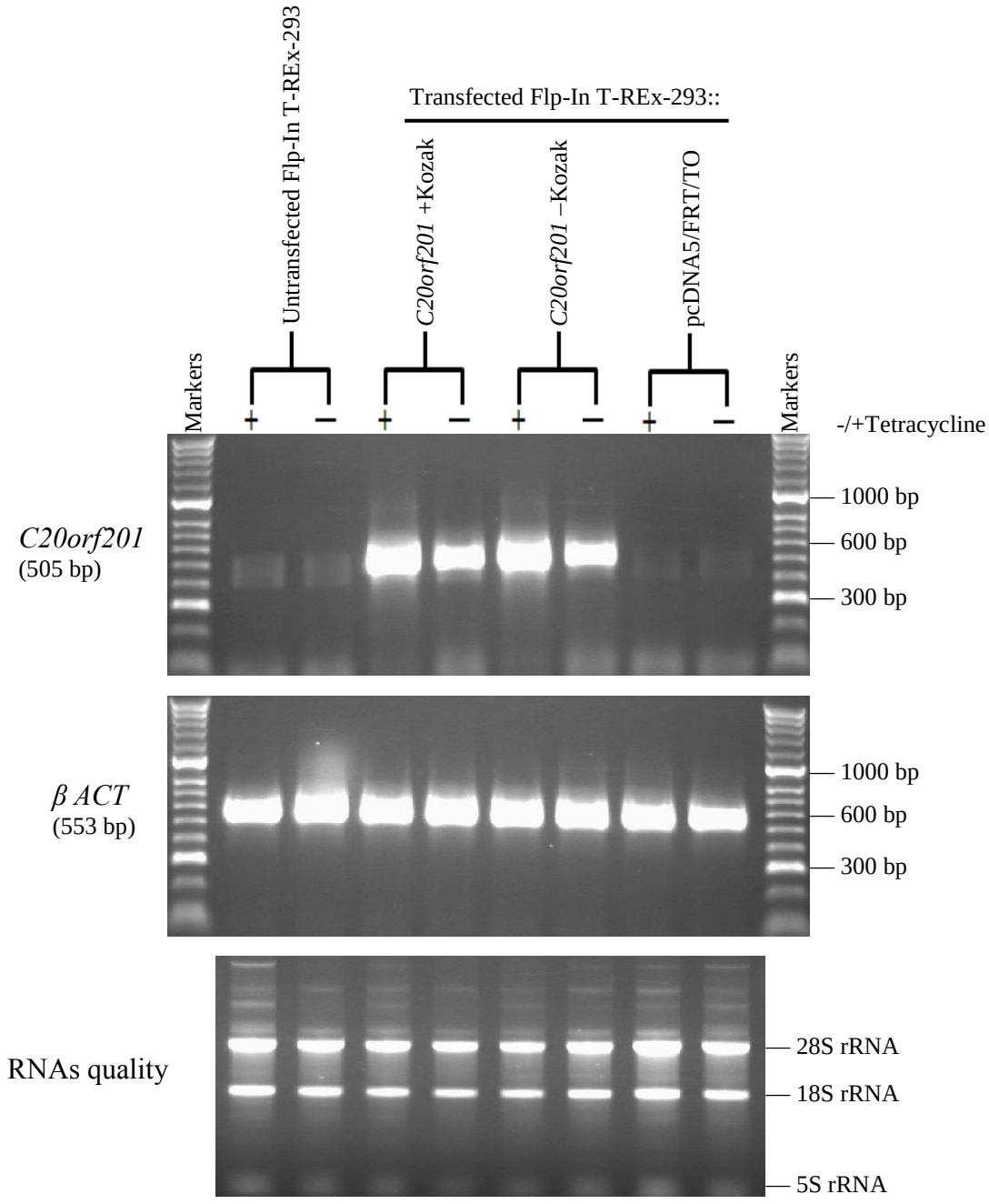


Figure 5.13 RT-PCR expression of *C20orf201* and *β ACT* in Flp-In T-REx-293 cell lines. The top agarose gel shows the expression of *C20orf201* in cells with three different insertions (*C20orf201* +Kozak or *C20orf201*–Kozak or pcDNA5/FRT/TO) with and without tetracycline treatment. The middle agarose gel displays the expression profile for *β*-actin as a positive control for the cDNA samples. The bottom agarose gel shows the quality of the total RNAs for different cell lines. Three bands of ribosomal RNAs appear in each sample. The 28S and 18S ribosomal RNA bands are clearly visible in the RNA samples at the ratio of approximately 2:1.

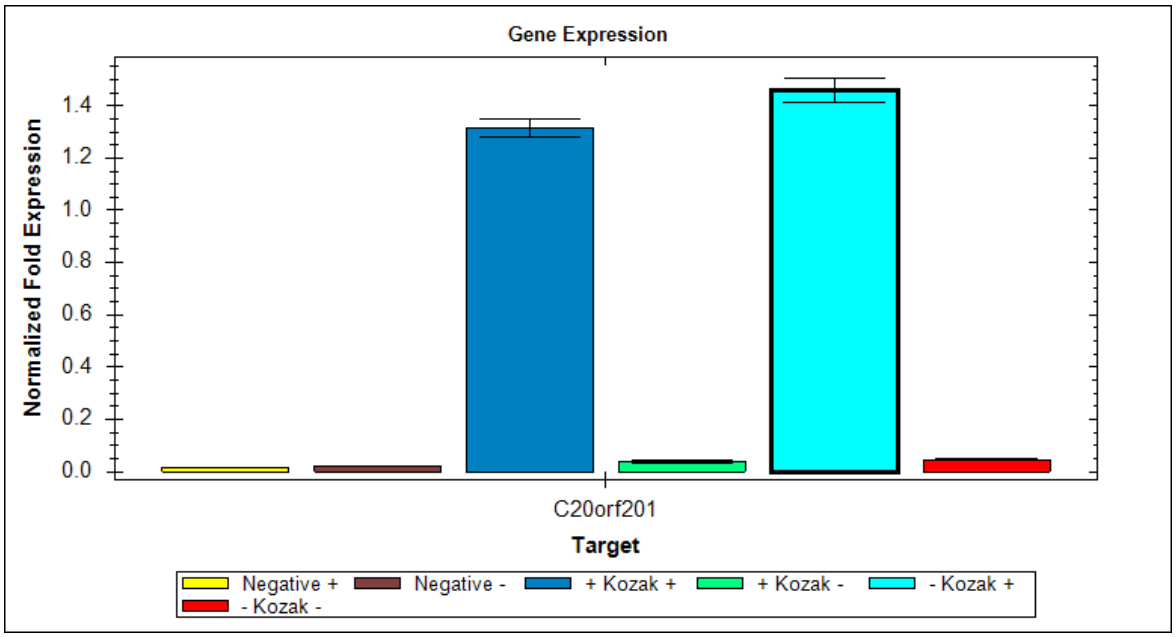


Figure 5.14 Real-time quantitative RT-PCR analysis for *C20orf201* expression in Flp-InT-REx-293 cells. The bar chart displays the gene expression results for *C20orf201* with and without Kozak consensus in Flp-In T-REx-293 cell lines before (-) or after (+) adding tetracycline to the cell culture, compared to the cells which contain the pcDNA5/FRT/TO vector alone (negative). Expression data were normalised to the *β-ACT* and *Lamin A* reference genes. The error bars represent the standard error of the mean (SEM) for three repeats.

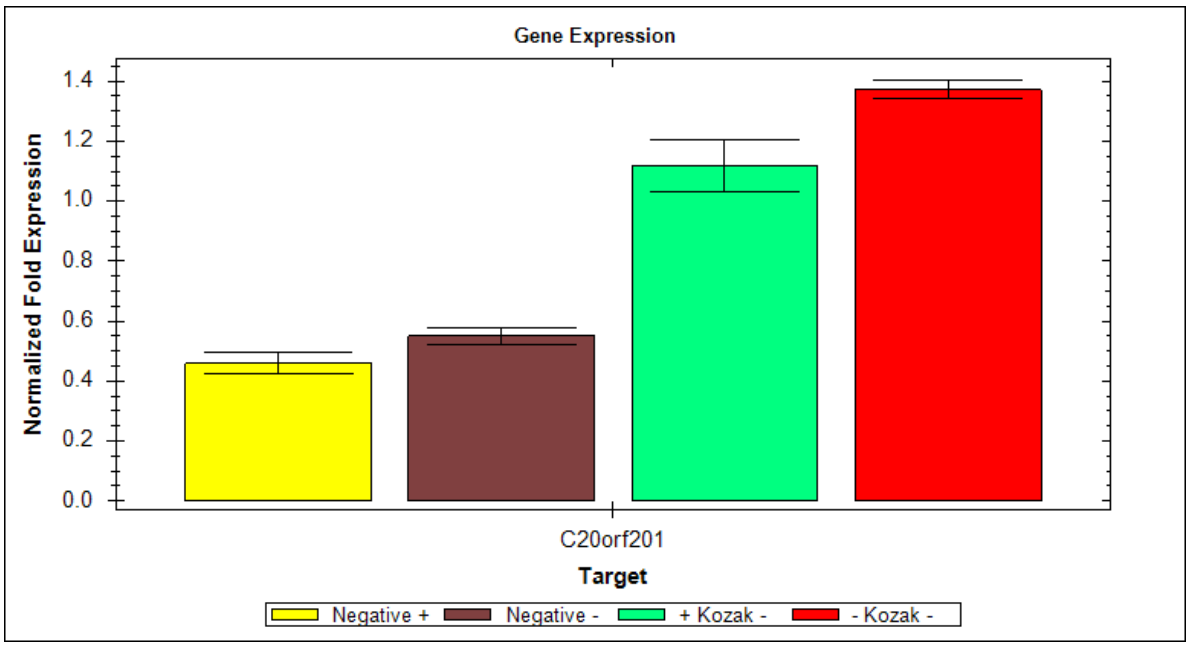


Figure 5.15 Real-time quantitative RT-PCR analysis for *C20orf201* expression in Flp-In T-REx-293 cells. The bar chart displays the gene expression results for *C20orf201* with and without Kozak sequence in Flp-In T-REx-293 cell lines before (-) adding tetracycline to the cell culture, compared to the cells which contain the pcDNA5/FRT/TO vector alone, with and without tetracycline treatment (negative). Expression data were normalised to the *β-ACT* and *Lamin A* reference genes. The error bars represent the standard error of the mean (SEM) for three repeats.

Table 5.1 Real-time qRT-PCR analysis for *C20orf201* expression in Flp-In T-REx-293 cells. The table displays the mean of three repeats of quantification cycle (Cq) and standard deviation (SD) for analysis of *C20orf201* expression with and without Kozak consensus in Flp-In T-REx-293 cell lines before (-) or after (+) adding tetracycline treatment. These results were normalised against *β-ACT* and *Lamin A* expression. The Cq readings for the no reverse transcriptase (NRT) and no template controls (NTC) are also shown.

Target name	Sample name	Wells	Cq Mean	Cq SD
C20orf201	Flp-In T- REx-293: pcDNA5/FRT/TO+	3	23.62	0.184
β ACT	Flp-In T- REx-293: pcDNA5/FRT/TO+	3	16.70	0.035
Lamin A	Flp-In T- REx-293: pcDNA5/FRT/TO+	3	21.83	0.137
C20orf201	Flp-In T- REx-293: pcDNA5/FRT/TO-	3	23.47	0.090
β ACT	Flp-In T- REx-293: pcDNA5/FRT/TO-	3	16.83	0.104
Lamin A	Flp-In T- REx-293: pcDNA5/FRT/TO-	3	21.94	0.164
C20orf201	Flp-In T- REx-293: *C20orf201*+Kozak+	3	18.07	0.053
β ACT	Flp-In T- REx-293: *C20orf201*+Kozak+	3	17.53	0.063
Lamin A	Flp-In T- REx-293: *C20orf201*+Kozak+	3	22.70	0.047
C20orf201	Flp-In T- REx-293: *C20orf201*+Kozak-	3	22.58	0.162
β ACT	Flp-In T- REx-293: *C20orf201*+Kozak-	3	16.88	0.040
Lamin A	Flp-In T- REx-293: *C20orf201*+Kozak-	3	22.13	0.062
C20orf201	Flp-In T- REx-293: *C20orf201*-Kozak+	3	17.61	0.048
β ACT	Flp-In T- REx-293: *C20orf201*-Kozak+	3	17.18	0.071
Lamin A	Flp-In T- REx-293: *C20orf201*-Kozak+	3	22.44	0.111
C20orf201	Flp-In T- REx-293: *C20orf201*-Kozak-	3	22.48	0.056
β ACT	Flp-In T- REx-293: *C20orf201*-Kozak-	3	17.13	0.026
Lamin A	Flp-In T- REx-293: *C20orf201*-Kozak-	3	22.25	0.022
C20orf201	Flp-In T- REx-293: pcDNA5/FRT/TO+ (NRT)	1	0.00	0.000
β ACT	Flp-In T- REx-293: pcDNA5/FRT/TO+ (NRT)	1	0.00	0.000
Lamin A	Flp-In T- REx-293: pcDNA5/FRT/TO+ (NRT)	1	0.00	0.000
C20orf201	Flp-In T- REx-293: pcDNA5/FRT/TO- (NRT)	1	0.00	0.000
β ACT	Flp-In T- REx-293: pcDNA5/FRT/TO- (NRT)	1	0.00	0.000
Lamin A	Flp-In T- REx-293: pcDNA5/FRT/TO- (NRT)	1	0.00	0.000
C20orf201	Flp-In T- REx-293: *C20orf201*+Kozak+ (NRT)	1	0.00	0.000
β ACT	Flp-In T- REx-293: *C20orf201*+Kozak+ (NRT)	1	0.00	0.000
Lamin A	Flp-In T- REx-293: *C20orf201*+Kozak+ (NRT)	1	0.00	0.000
C20orf201	Flp-In T- REx-293: *C20orf201*+Kozak- (NRT)	1	0.00	0.000
β ACT	Flp-In T- REx-293: *C20orf201*+Kozak- (NRT)	1	0.00	0.000
Lamin A	Flp-In T- REx-293: *C20orf201*+Kozak- (NRT)	1	0.00	0.000
C20orf201	Flp-In T- REx-293: *C20orf201*-Kozak+ (NRT)	1	0.00	0.000
β ACT	Flp-In T- REx-293: *C20orf201*-Kozak+ (NRT)	1	0.00	0.000
Lamin A	Flp-In T- REx-293: *C20orf201*-Kozak+ (NRT)	1	0.00	0.000
C20orf201	Flp-In T- REx-293: *C20orf201*-Kozak- (NRT)	1	0.00	0.000
β ACT	Flp-In T- REx-293: *C20orf201*-Kozak- (NRT)	1	0.00	0.000
Lamin A	Flp-In T- REx-293: *C20orf201*-Kozak- (NRT)	1	0.00	0.000
C20orf201	NTC	1	0.00	0.000
β ACT	NTC	1	0.00	0.000
Lamin A	NTC	1	0.00	0.000

5.2.4 Evaluation of the C20orf201 protein levels in the Flp-In T-REx-293 cell lines

The C20orf201 protein levels were assessed in different integrants of Flp-In T-REx-293 cell lines before and after tetracycline treatment using anti-C20orf201 antibody (Abcam; AB108142). The molecular weight of C20orf201 protein is ~26 KDa, according to Western blotting (Figure 5.16). Western blots indicated that the levels of C20orf201 protein were clearly strong in the Flp-In T-REx-293 cell lines inserted with either *C20orf201*+Kozak or *C20orf201*–Kozak following treatment with tetracycline, when compared to the two negative controls (untransfected cells and cells transfected with pcDNA5/FRT/TO only) (Figure 5.16). The mRNA of *C20orf201* gene was expressed in cells with *C20orf201*+Kozak or *C20orf201*–Kozak without tetracycline treatment (refer to Figure 5.13); however, no proteins were detectable in either sample. Therefore, the Western blotting results suggest that some mechanism might operate that inhibits translation in the absence of tetracycline (high mRNA levels). Appropriate protein loading was confirmed for each clone by the GAPDH control.

5.2.5 Protein analysis of the C20orf201 gene in the normal and cancer tissues

Western blotting was performed to investigate protein levels of C20orf201 in 11 different normal tissues and 6 different cancer cell lines. Two types of testis lysate were used, obtained from different companies. Normal tissues showed C20orf201 protein levels in both testis lysates and central nervous system tissues (brain, brain-cerebellum and spinal cord) (Figure 5.17), in agreement with the RT-PCR results obtained for *C20orf201* gene in non-cancer cells (refer to Figure 3.12). In contrast, the presence of the C20orf201 protein was validated in 4 cell lines that showed previously RT-PCR expression in Chapter 3, while NTERA2 and HEP-G2 cell lysates were used as negative controls for the C20orf201, because no PCR product was detected. Interestingly, The C20orf201 protein was detected in breast cancer (MCF-7) at a high level and in ovarian cancer (TO14) at low levels (Figure 5.18). Neither negative control displayed any levels of protein for C20orf201, which corresponded with the RT-PCR results (refer to Figure 3.14).

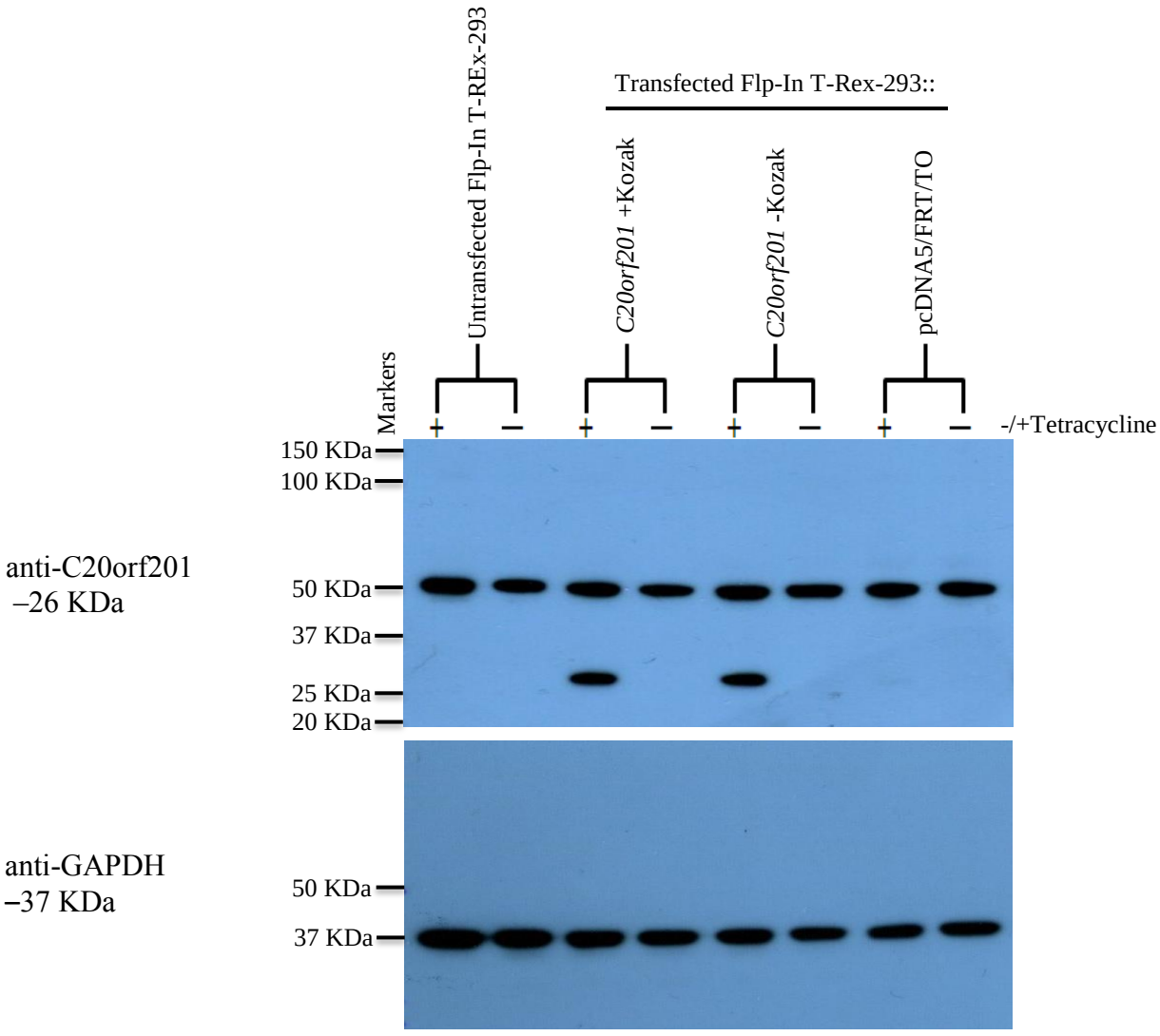


Figure 5.16 Western blot analysis showing C20orf201 induction using tetracycline in the Flp-In T-REx-293 cell lines. Four cultures were subjected to 2 µg/ml of tetracycline and four were not. Positive signals are seen in cells containing the *C20orf201* gene with and without Kozak sequencing after treatment with tetracycline. The cells were harvested 24 hours after induction. The lysates were separated on a 10% SDS-PAGE gel. GAPDH (Glyceraldehyde 3-phosphate dehydrogenase**)** was used as the loading control.

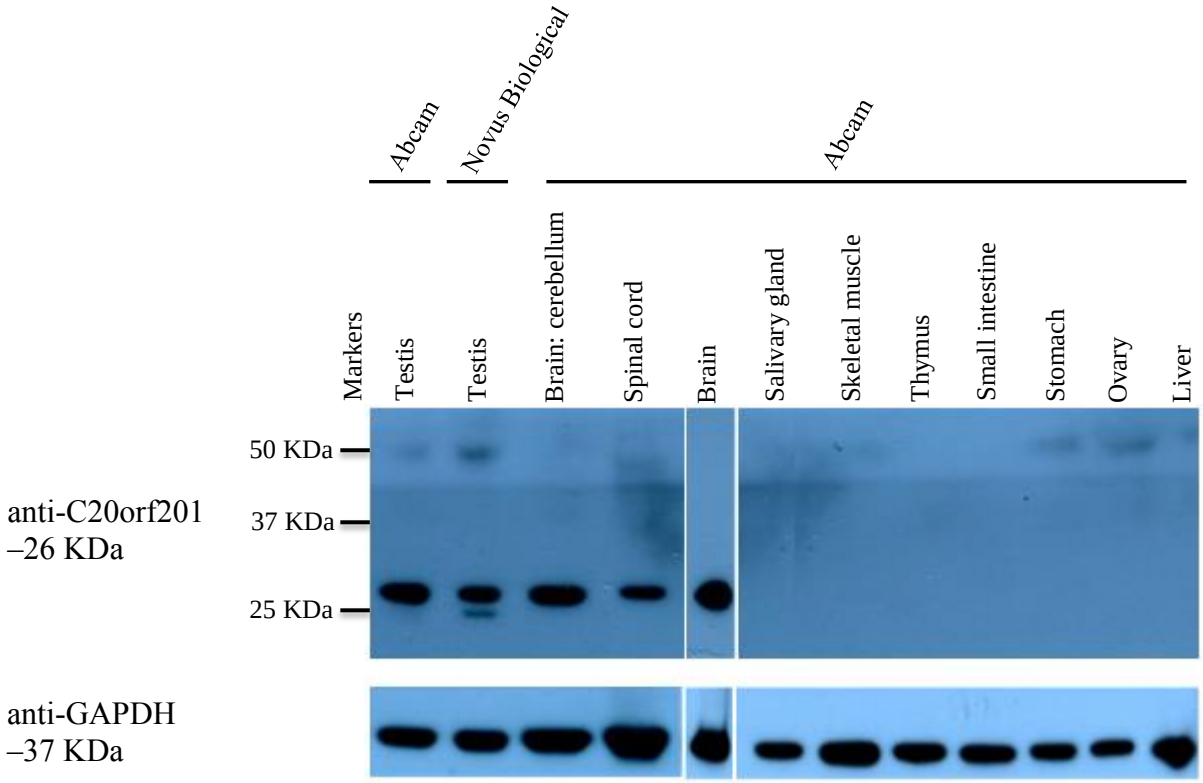


Figure 5.17 Western blot analysis showing C20orf201 protein in distinct normal lysates. Positive signals are seen in testes and central nervous tissues (top). The lysates were separated on a 10% SDS-PAGE gel. GAPDH was used as the loading control (bottom).

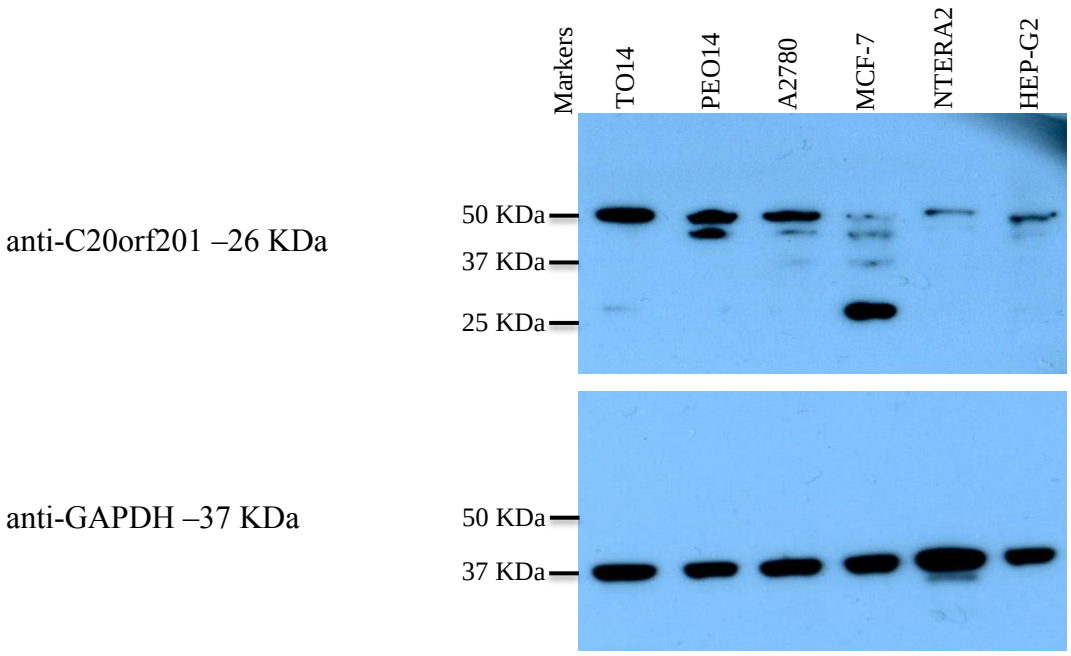


Figure 5.18 Western blot analysis for C20orf201 protein levels in distinct cancer cell line lysates. Western blot analysis showing a strong signal in the MCF-7 cell line compared to other cell lines (top). The NTERA2 and HEP-G2 lysates were used as negative controls as RT-PCR of these cell lines showed no expression of *C20orf201*. The lysates were separated on a 10% SDS-PAGE gel. GAPDH was used as the loading control (bottom).

5.2.6 RT-PCR and protein analyses of the C20orf201 gene in cancer cell lines

MCF-7, A2780, PEO14 and TO14 cell lines were grown; RNAs and whole lysates were extracted from them for RT-PCR and Western blotting analyses, respectively. The cDNAs were also made from these cells, with and without adding reverse transcriptase. Expression of *C20orf201* was observed in all cell lines tested here, but with some differences in mRNAs levels. The intensity of the *C20orf201* PCR band was significantly fainter in the breast cancer (MCF-7) cell line than in the other cell lines tested (Figure 5.19, top left). Samples containing no reverse transcriptase were used as negative controls and *C20orf201* expression was not detectable in any of the cell lines under this condition, indicating no genomic DNA contamination (Figure 5.19, top right). Testis cDNA was used in RT-PCR to confirm the performance of primers. *β ACT* gene was used as well to assess the cDNA quality synthesised from these cell lines.

Post-transcriptional mechanisms may affect the association between gene expression and protein levels (Tan *et al*., 2009). Consequently, *C20orf201* expression in four cancer cell lines does not guarantee that the C20orf201 protein is produced. Therefore, Western blotting analysis was used to investigate the protein abundance for the C20orf201 protein in one breast cancer and three ovarian cancer cell lines. The GAPDH results suggest that the gel loading was relatively equal, so differences in protein level intensities are likely to be due to variations in protein abundance in the four cell lines. Western blot analysis using anti-C20orf201 identified two bands at approximately 50 KDa (upper band) and at 26 KDa (lower band) in all cell lines. These Western blot analysis results showed a fainter band for C20orf201 in the A2780 cells compared to the other cells. Importantly, the strongest band for C20orf201 was found in the MCF-7 (Figure 5.20), although the gene expression shown in Figure 5.19 was extremely faint.

qRT-PCR was carried out to compare the mRNA expression for the *C20orf201* gene in the MCF-7, A2780, PEO14 and TO14 cell lines (Figure 5.21). The RT-PCR (Figure 5.19) and the qRT-PCR (Figure 5.21) results indicated a lower level of *C20orf201* gene expression in the MCF-7 cells, but the protein level was high in this cell line compared to the three ovarian cancer cells (Figure 5.20). Therefore, these Western blot results did not appear to correspond with the gene expression results. The C20orf201 qRT-PCR results were normalised to the qRT-PCR results for *Lamin A* and *β ACT*. The qRT-PCR results are shown in Figure 5.21.

RT-PCR and Western blotting results for the C20orf201 gene in 11 normal tissues and 4

cancer cell lines confirmed that the *C20orf201* gene is potentially a good CTA gene candidate.

The band that appeared in the testis protein lysate through Western blot analysis using anti-C20orf201 antibody was confirmed to have the same molecular weight in the MCF-7, Flp-In T-REx-293::*C20orf201*+Kozak and Flp-In T-REx-293::*C20orf201*–Kozak by running these cell lysates again, side by side, to identify any differences in the molecular weight of C20orf201 protein among them. The Western blotting results for the C20orf201 gene are shown in Figure 5.22. These results indicated that the molecular weight for C20orf201 protein is approximately 26 KDa.

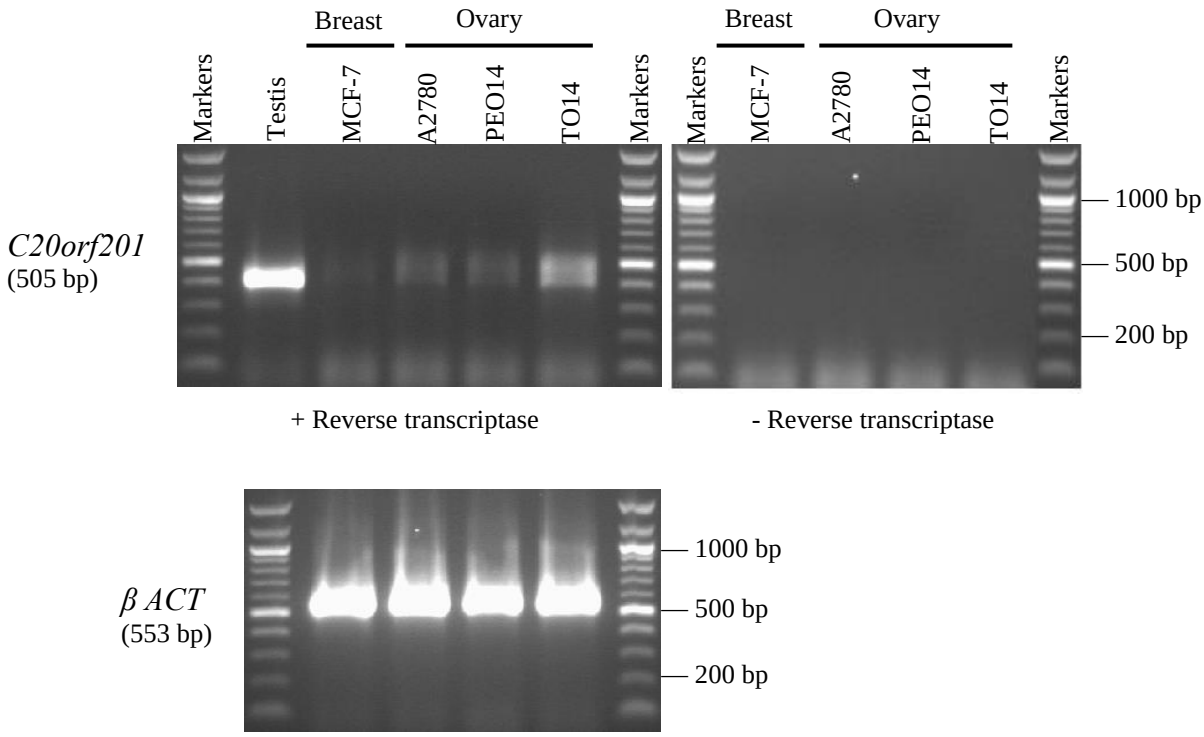


Figure 5.19 RT-PCR analysis for *C20orf201* gene expression in different human cancer cells. Agarose gel image illustrating the RT-PCR analysis for *C20orf201* expression in normal testis, three ovarian cancer cell lines (A2780, PEO14 and TO14) and a breast cancer cell line (MCF-7). cDNAs were isolated for these cells with (+) and without (-) adding reverse transcriptase. The *C20orf201* gene was expressed in all cells, but with some differences. *β ACT* expression was used as a positive control for the cDNA samples. The expected amplicon size of each gene is shown on the left between brackets.

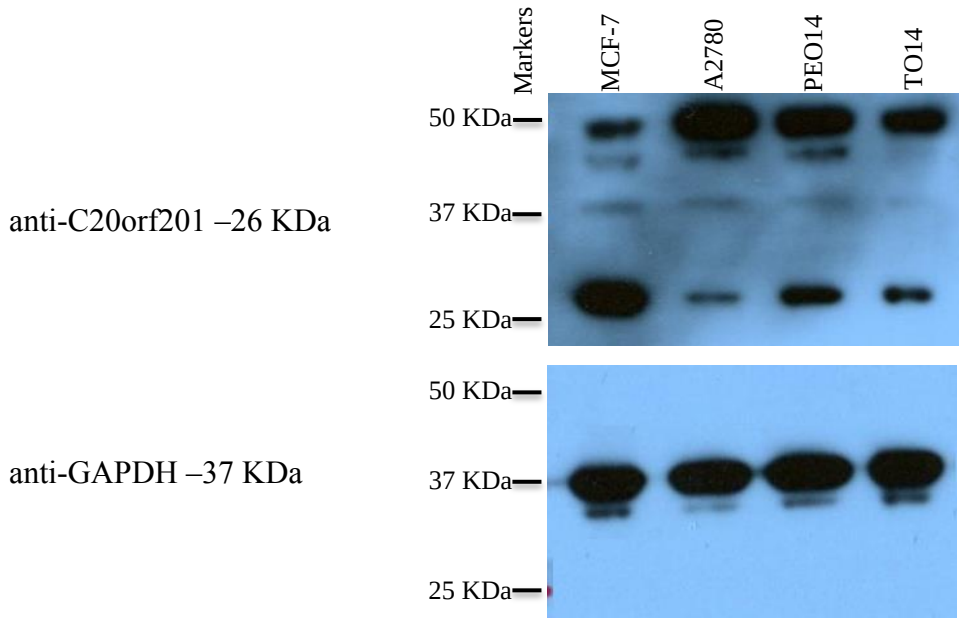
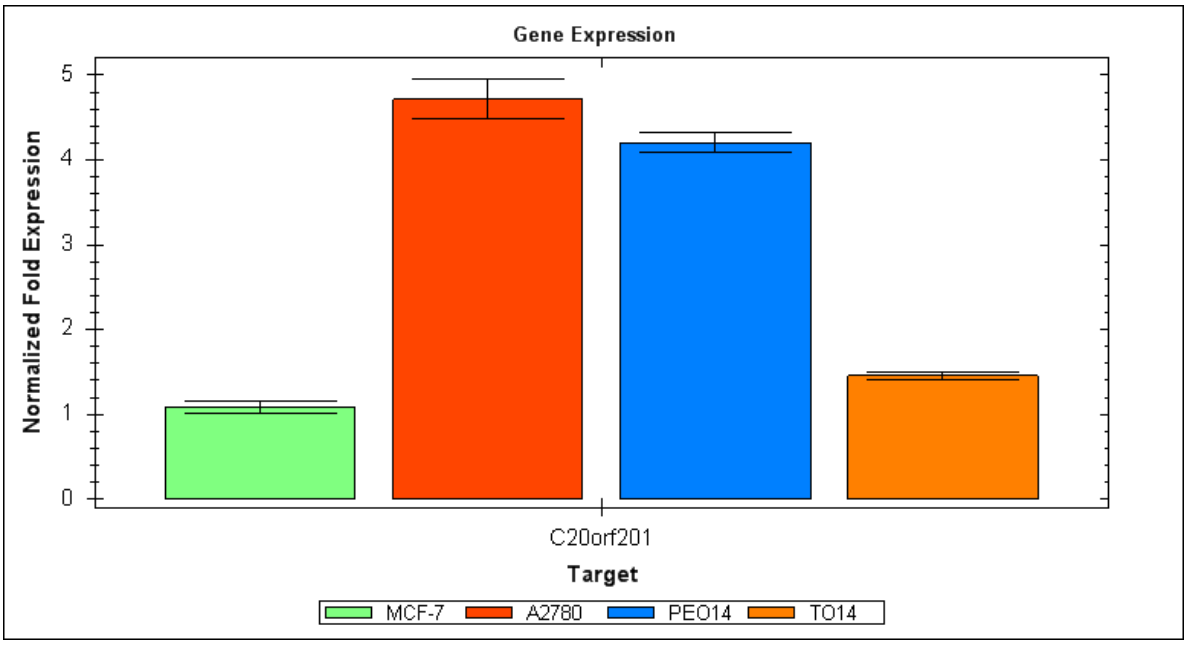


Figure 5.20 Western blot analysis for C20orf201 protein levels in distinct cancer cell line lysates. Western blot analysis showed a strong signal in the MCF-7 cell line compared to other cell lines (top). The lysates were separated on a 10% SDS-PAGE gel. GAPDH was used as the loading control (bottom).



A

Target name	Sample name	Wells	Cq Mean	Cq SD
C20orf201	MCF7	3	23.49	0.053
β ACT	MCF7	3	14.41	0.315
Lamin A	MCF7	3	18.89	0.081
C20orf201	A2780	3	23.63	0.101
β ACT	A2780	3	15.91	0.033
Lamin A	A2780	3	21.90	0.138
C20orf201	PEO14	3	24.39	0.045
β ACT	PEO14	3	17.91	0.078
Lamin A	PEO14	3	21.10	0.076
C20orf201	TO14	3	23.60	0.073
β ACT	TO14	3	15.72	0.061
Lamin A	TO14	3	18.65	0.033
C20orf201	MCF7 (NRT)	1	0.00	0.000
β ACT	MCF7 (NRT)	1	0.00	0.000
Lamin A	MCF7 (NRT)	1	0.00	0.000
C20orf201	A2780 (NRT)	1	0.00	0.000
β ACT	A2780 (NRT)	1	0.00	0.000
Lamin A	A2780 (NRT)	1	0.00	0.000
C20orf201	PEO14 (NRT)	1	0.00	0.000
β ACT	PEO14 (NRT)	1	0.00	0.000
Lamin A	PEO14 (NRT)	1	0.00	0.000
C20orf201	TO14 (NRT)	1	0.00	0.000
β ACT	TO14 (NRT)	1	0.00	0.000
Lamin A	TO14 (NRT)	1	0.00	0.000
C20orf201	NTC	1	0.00	0.000
β ACT	NTC	1	0.00	0.000
Lamin A	NTC	1	0.00	0.000

B

Figure 5.21 Real-time quantitative RT-PCR analysis for *C20orf201* expression in different cells. The bar chart in panel **(A)** displays the gene expression results for *C20orf201* in MCF-7, A2780, PEO14 and TO14 cancer cell lines. The error bars represent the standard error of the mean (SEM) for 3 repeats. The table in panel **(B)** displays the mean of three repeats of quantification cycle (Cq) and standard deviation (SD) for *C20orf201*. Expression data were normalised to the *β ACT* and *Lamin A* reference genes. Also shown in the table are the Cq readings for the no reverse transcriptase (NRT) and no template controls (NTC).

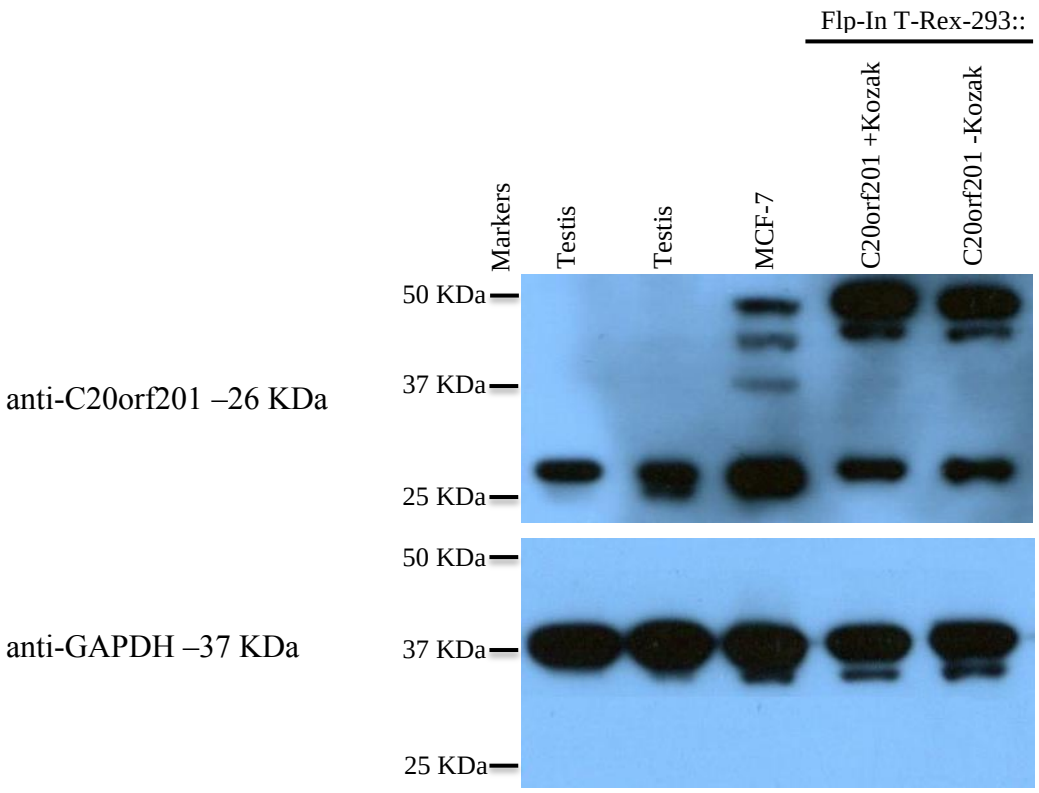


Figure 5.22 Western blot analyses for C20orf201 protein levels in normal and cancer cell lysates. Western blot analysis showed strong signals in all cells for a protein with molecular weight of about 26 KDa (top). The lysates were separated on a 10% SDS-PAGE gel. GAPDH was used as the loading control (bottom).

5.2.7 Cellular localisation of C20orf201

The cellular location of C20orf201 was analysed in MCF-7 (a breast cancer cell line) cells using Western blot analysis and Immunofluorescent staining of fixed cells.

5.2.7.1 Western blot analysis

The nuclear and cytoplasmic protein fractions were prepared alongside whole cell extracts for the MCF-7 cell line, to identify the cellular location of C20orf201 gene using anti-C20orf201 antibody (Abcam; AB108142). Nuclear, cytoplasmic and whole protein lysates were prepared for MCF-7 to localise the protein for the C20orf201 gene. These lysates were then analysed using Western blot analysis, as shown in Figure 5.23. Antibodies against GAPDH and lamin B were used as the positive controls for the cytoplasmic and nuclear fractions, respectively. The predicted band size for C20orf201 using anti-C20orf201 is approximately 26 KDa. Interestingly, the C20orf201 protein was detected in the MCF-7 cells as a stronger band in both the cytoplasm at 26 KDa and the nucleus at 50 KDa.

5.2.7.2 Immunofluorescent staining of fixed cells

Immunofluorescent staining was also performed for C20orf201 in fixed MCF-7 cells using anti-C20orf201 antibody (Abcam; AB170783). Two images are shown for anti-C20orf201 staining of MCF-7 cells; the upper and lower rows show cells magnified with a 60x objective lens. Both stained images suggest that C20orf201 is found in the nucleus and the cytoplasm of the MCF-7 cells (Figure 5.24). Therefore, the immunostaining results appear to correspond with the Western blotting results.

The anti-mouse secondary antibody was used as a negative control and to check the level of non-specific background in the MCF-7 cells. However, the non-specific background produced through the secondary antibody was not significant compared with the staining observed with the primary antibody. Therefore, this indicates that the staining detected for anti-C20orf201 resulted from staining of the primary antibody.

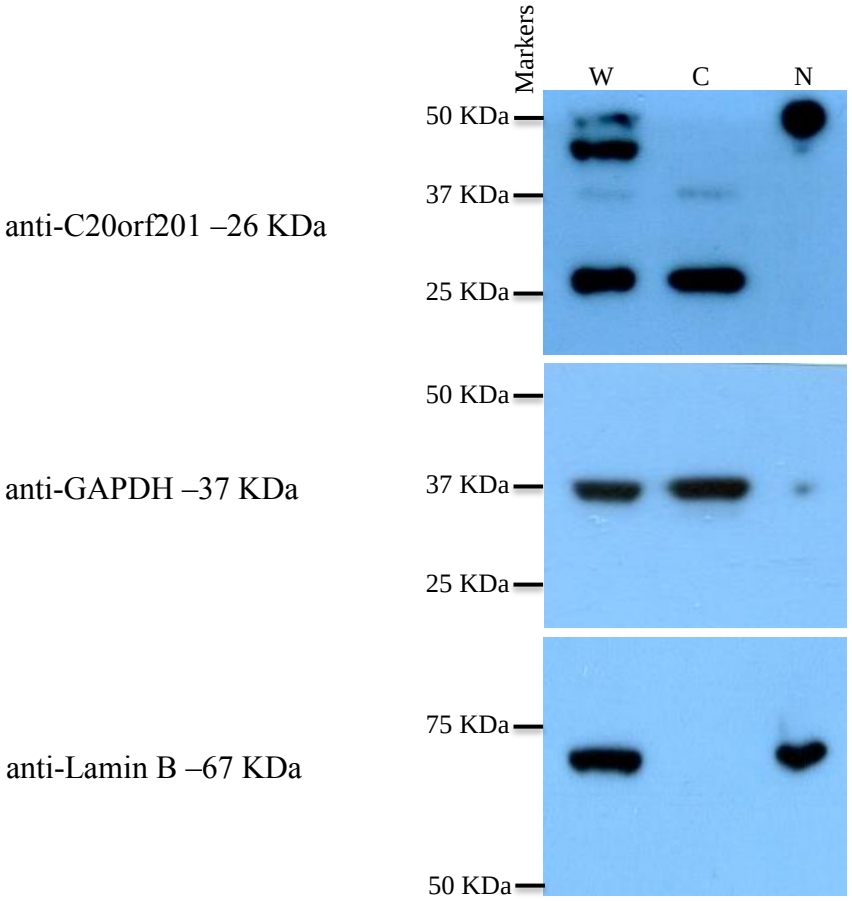


Figure 5.23 Western blot analysis of C20orf201 sub-cellular localisation in MCF-7 cancer cell line. Western blot analysis showing the cellular localisation of C20orf201 protein in the MCF-7 cell line using a whole cell extract (W), a cytoplasmic lysate fraction (C) and a nuclear lysate fraction (N). Antibodies against lamin B (Nuclear) and GAPDH (Cytoplasmic) were used as the positive controls, to confirm the fractionation efficiency and the loading of the gel.

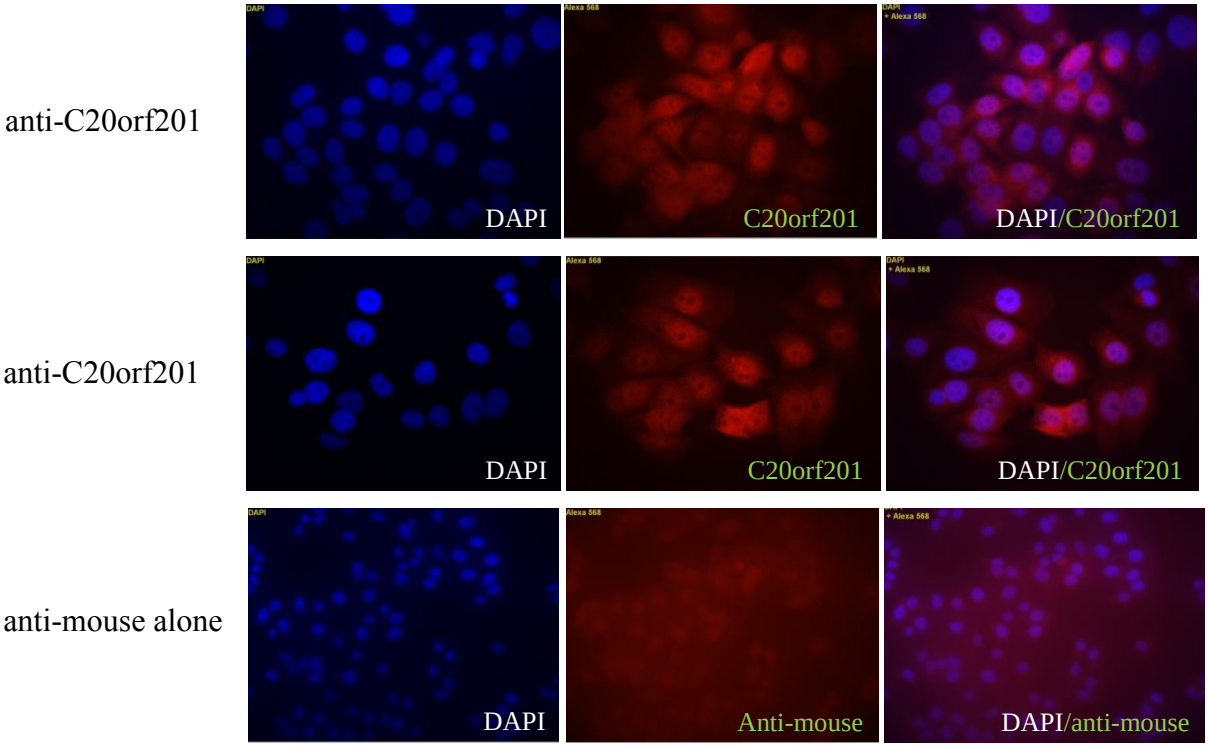


Figure 5.24 Immunofluorescent staining of fixed MCF-7 cells with antibody against C20orf201. Images for immunofluorescent staining for C20orf201 (red) in MCF-7 cells, viewed using a Zeiss Axioskop 2 fluorescence microscope (60x lens). Staining with the secondary antibodies (anti-mouse) only was used as the negative control. Nuclei are stained blue with DAPI.

5.2.8 siRNA knockdown of C20orf201 in MCF-7 cell line

Knockdown of C20orf201 levels was attempted using four types of *C20orf201*-specific siRNAs in the MCF-7 cell line to determine any effect that knockdown of this protein may have on these cells and to test the specificity of the anti-C20orf201 antibody. The predicted molecular weight of the C20orf201 gene is approximately 26 KDa. Whole cell extracts were collected from the cells following siRNA treatment and analysed using Western blot and qRT-PCR analyses. The number of siRNA treatments ('hits') was increased to optimise any knockdown in C20orf201 protein levels or gene expression.

The band thought to correspond to C20orf201 showed only a slight decrease after 48 hours treatment (2 'hits') with siRNA 7 and a combination of four C20orf201-specific siRNAs compared to the negative control (cells treated with non-interfering siRNA). Furthermore, faint but clear protein bands were also observed at approximately 50 KDa in cells treated with siRNA 7 and mixed siRNAs (2, 5, 6 and 7) compared to transfected cells (Figure 5.25). However, this band was also observed as well as a faint band in a few protein lysates in

normal tissues (refer to Figure 5.17) and in the NTERA2 and HEP-G2 cell lines that were negative for RT-PCR expression of the C20orf201 gene (refer to Figure 5.18). These results suggest that this band may be specific to C20orf201. GAPDH levels were used as a loading control.

Gene expression was also examined using qRT-PCR to determine whether expression of *C20orf201* decreases after siRNA knockdown after 48 hours (two 'hits') treatment with different types of siRNAs. Importantly, the qRT-PCR results show significant *C20orf201* expression knockdown in the MCF-7 cells. The knockdown was observed in siRNAs 5, 6, 7 and a combination of four C20orf201-specific siRNAs compared to the non-interfering siRNA control (Figure 5.26). The *C20orf201* qRT-PCR results were normalised to the qRT-PCR for *Lamin A* and *β ACT*; the results are shown in Table 5.2. Neither negative control, NRT and NTC, gave a Cq reading, which indicated the absence of genomic contamination and non-specific background (Table 5.2).

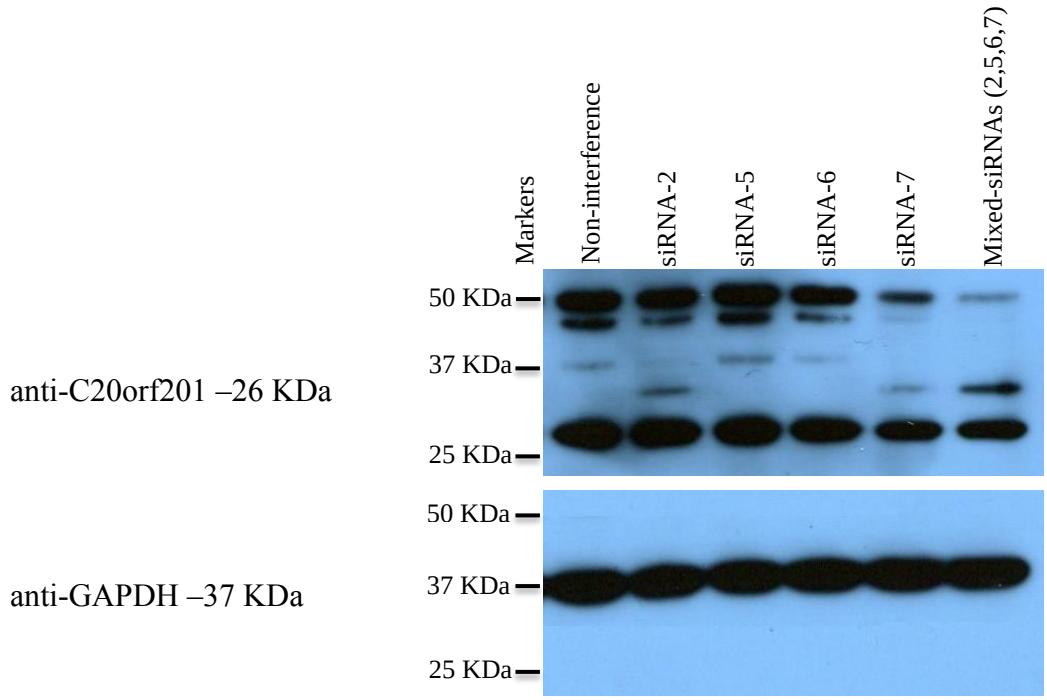


Figure 5.25 Western blot analysis for two hits of siRNA depletion of C20orf201 in the MCF-7 cell line. Negative control for the siRNA knockdown was used: cell treated with non-interfering siRNA. Cells were transfected with 5 nM concentrations of four independent C20orf201 siRNAs and also a combination of the four siRNAs (siRNAs 2+5+6+7) or with 5 nM of non-interference siRNA. The transfection reagent and siRNAs were added to the cells culture upon seeding. The cells were collected at 48 hours after siRNA transfection. GAPDH protein levels were used as a control for protein loading.

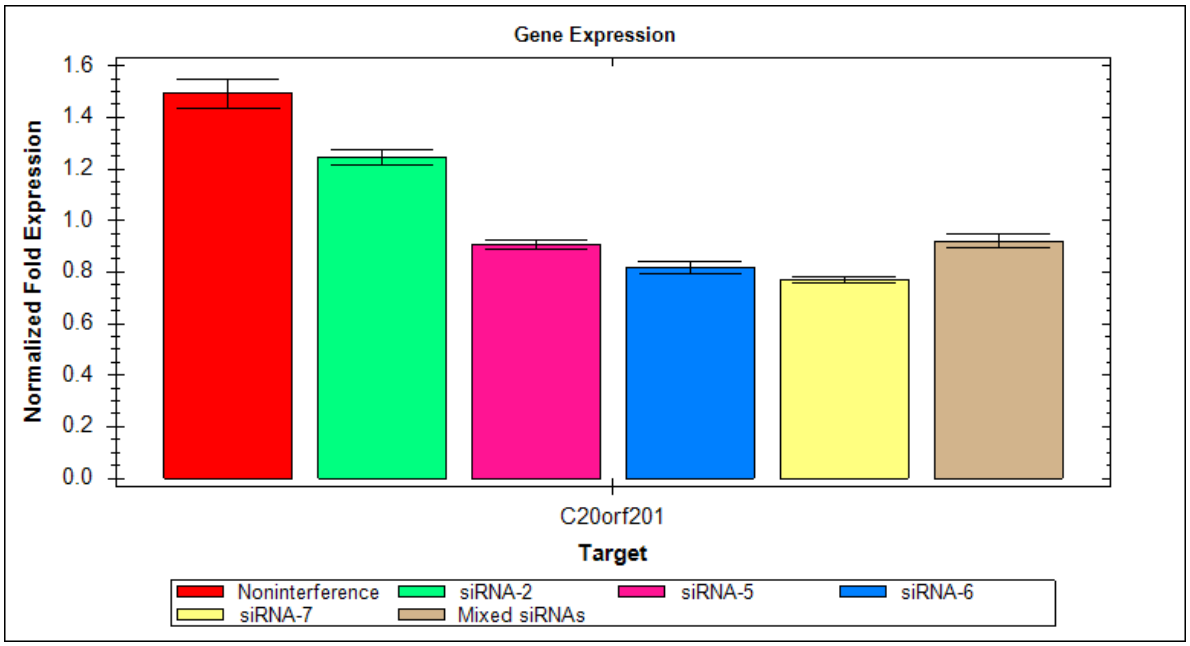


Figure 5.26 Real-time quantitative RT-PCR analysis for two hits of siRNA knockdown of *C20orf201* expression in MCF-7 cells. The bar chart displays the gene expression results for *C20orf201* in the presence of four types of siRNAs for *C20orf201* in the MCF-7 cell line compared to the cells treated with non-interfering siRNA (negative control). Expression data were normalised to the *β-ACT* and *Lamin A* reference genes. The error bars represent the standard error of the mean (SEM) for three repeats.

Table 5.2 Real-time qRT-PCR analysis for siRNA knockdown (2 'hits') of *C20orf201* in MCF-7 cells. The table displays the mean of three repeats of quantification cycle (Cq) and standard deviation (SD) for *C20orf201*. These results were normalised by *β ACT* and *Lamin A* expression. Also shown in the table are the Cq readings for the no reverse transcriptase (NRT) and no template controls (NTC).

Target name	Sample name	Wells	Cq Mean	Cq SD
C20orf201	Non-interference	3	29.41	0.059
β ACT	Non-interference	3	16.67	0.006
Lamin A	Non-interference	3	21.40	0.093
C20orf201	siRNA-2	3	27.62	0.062
β ACT	siRNA-2	3	16.60	0.080
Lamin A	siRNA-2	3	21.45	0.048
C20orf201	siRNA-5	3	23.20	0.018
β ACT	siRNA-5	3	16.46	0.023
Lamin A	siRNA-5	3	21.23	0.044
C20orf201	siRNA-6	3	30.11	0.051
β ACT	siRNA-6	3	16.25	0.038
Lamin A	siRNA-6	3	21.10	0.043
C20orf201	siRNA-7	3	29.86	0.177
β ACT	siRNA-7	3	16.19	0.055
Lamin A	siRNA-7	3	21.06	0.036
C20orf201	Mixed siRNAs	3	29.95	0.202
β ACT	Mixed siRNAs	3	16.48	0.083
Lamin A	Mixed siRNAs	3	21.43	0.024
C20orf201	Non-interference (NRT)	1	0.00	0.000
β ACT	Non-interference (NRT)	1	0.00	0.000
Lamin A	Non-interference (NRT)	1	0.00	0.000
C20orf201	siRNA-2 (NRT)	1	0.00	0.000
β ACT	siRNA-2 (NRT)	1	0.00	0.000
Lamin A	siRNA-2 (NRT)	1	0.00	0.000
C20orf201	siRNA-5 (NRT)	1	0.00	0.000
β ACT	siRNA-5 (NRT)	1	0.00	0.000
Lamin A	siRNA-5 (NRT)	1	0.00	0.000
C20orf201	siRNA-6 (NRT)	1	0.00	0.000
β ACT	siRNA-6 (NRT)	1	0.00	0.000
Lamin A	siRNA-6 (NRT)	1	0.00	0.000
C20orf201	siRNA-7 (NRT)	1	0.00	0.000
β ACT	siRNA-7 (NRT)	1	0.00	0.000
Lamin A	siRNA-7 (NRT)	1	0.00	0.000
C20orf201	Mixed siRNAs (NRT)	1	0.00	0.000
β ACT	Mixed siRNAs (NRT)	1	0.00	0.000
Lamin A	Mixed siRNAs (NRT)	1	0.00	0.000
C20orf201	NTC	1	0.00	0.000
β ACT	NTC	1	0.00	0.000
Lamin A	NTC	1	0.00	0.000

5.3 Discussion

5.3.1 Expression of the human *C20orf201* gene in normal and cancer tissues

The *C20orf201* gene splice variant 1 was expressed in different normal tissues, including testis, brain (whole or cerebellum), foetal brain and spinal cord, while splice variant 2 was expressed in the same tissues as well as in the trachea. During the course of our work another study indicated that *C20orf201* was expressed only in the testis and brain (whole or cerebellum) (Kamata *et al*., 2013). Both studies confirm the results presented in Chapter 3 for *C20orf201* splice variant 1 and in Chapter 5 for splice variant 2.

RT-PCR for the *C20orf201* gene in various cancer tissues and cell lines indicated expression of variants 1 and 2 in 8 and 13 cancer cell lines, respectively. Feichtinger *et al*. (2012a) reported detectable *C20orf201* expression in 8 cancer cell lines, including MCF-7, A2780, PEO14, TO14, HL-60, Jurkat, and uterus and ovarian tumours. Another previous study using qRT-PCR analysis revealed that the expression of *C20orf201* was detected in 13 cancer types derived from different organs (Kamata *et al*., 2013). The results obtained in this chapter by RT-PCR and Western blotting and the previous studies described above confirm that *C20orf201* is a good example of a cancer testis antigen and may be useful as a cancer biomarker. Subsequent work within the group has now demonstrated immunohistochemistry staining of C20orf201 in ovarian cancer tissue, but not healthy tissues.

5.3.2 Regulation of meiosis-specific transcripts

The Western blot and RT-PCR analyses using anti-C20orf201 antibody and primers (F1, R1), respectively, showed strong bands for C20orf201 protein and expression in the presence of tetracycline in both Flp-In T-REx-293::*C20orf201*+Kozak and Flp-In T-REx-293::*C20orf201*–Kozak cell lines. Conversely, in the absence of tetracycline, the *C20orf201* gene was expressed in the same cell lines, but no protein at the expected size was detected. Therefore, the Western blotting results obtained in the presence and absence of tetracycline indicated that there might be some mechanism that represses translation of the C20orf201 mRNA in these cell lines in the absence of tetracycline treatment (such as system exists in fission yeast).

Meiosis leads to a reduction in the genetic material of the cell. The meiotic program among eukaryotes is tightly regulated, although much of the mechanism regulating meiosis remains unknown (Holm and Thon, 2012). In fission yeast, *Saccharomyces pombe* meiosis takes place

in zygote diploid cells that are heterozygous for the mating-type genes and are exposed to nitrogen starvation. These cells arrest in the G1 phase of the mitotic cell cycle, initiate DNA synthesis, and then perform two different consecutive nuclear divisions, the first and second meiotic divisions (Borgen *et al*., 2002). The fission yeast utilises a RNA degradation system that selectively eliminates meiosis-specific mRNA during the mitotic cell cycle. Mmi1, a novel RNA binding protein of the YTH-family, is critical for this process (Yamamoto, 2010). When fission yeast cells undergo meiosis, the expression of hundreds genes is newly induced or significantly up-regulated by the function of specific transcription factors, such as Ste11, which are activated by nitrogen starvation (Yamamoto, 2010).

The Mmi1 RNA elimination system is a type of RNA interference (RNAi) pathway found in fission yeast and used for degradation of meiotic transcripts (Harigaya *et al*., 2006). In this pathway, the RNA binds a specific region on the mRNA, termed the DSR (determinant of selective removal), with the aid of multiple factors (Sugiyama and Sugioka- Sugiyama, 2011). During meiotic progression, Mmi1 is sequestered by the master regulator of meiosis, Mei2 nuclear dot, in meiotic prophase, which results in the failure of the mRNAs to be destroyed by Mmi1. As a result, meiosis-specific mRNAs are stably expressed (Holm and Thon, 2012). However, in mitosis, repression of meiotic-specific transcripts occurs through Mmi1. Mmi1 functions by binding to the DSR sequences in a meiotic mRNAs. Therefore, the function of Mmi1 is to mediate elimination of meiotic mRNAs during mitosis (Harigaya *et al*., 2006).

5.3.3 A possible ubiquitinated variant?

Ubiquitination is an important post-translational modification (PTM) and plays a fundamental role in the regulating of protein functions and it involves the covalent attachment of the ubiquitin (Ub) peptide to lysine residues on substrate proteins, which control all aspects of cellular functioning (Xu and Jaffrey, 2013). The binding between the Ub and substrate protein can lead to protein monoubiquitination or polyubiquitination (Sadowski *et al*., 2012). Protein may be ubiquitinated on one or multiple lysines, resulting in a modification called either monoubiquitination or multiubiquitination, respectively (Petroski and Deshaies, 2005). Ubiquitination can change protein-protein interactions or alter protein localisation (Mukhopadhyay and Riezman, 2007). Many human diseases, such as cancer, are caused by the alteration of protein ubiquitination (Xu and Jaffrey, 2013). Protein ubiquitination is catalysed by Ub-conjugating and Ub-ligase enzymes, which are recognised a specific type of Ub modification on protein substrates; however, how these enzymes operate is still poorly

understood (Sadowski *et al*., 2012). In order to understand the biological function of ubiquitination and its role in cancer, it is important to identify protein ubiquitination sites (Hoeller and Dikic, 2009). However, the identification of ubiquitination sites is a major challenge for two reasons. First, the number of ubiquitinated proteins in cells is very low under normal physiological conditions (Mann and Jensen, 2003). Second, only one or a few lysine residues are modified in an ubiquitinated protein (Jadhav and Wooten, 2009).

In this chapter, we do not know if the C20orf201 protein is ubiquitinated in the MCF-7 cell line; however, we noted that the patterns of three high-molecular-weight bands are consistent with ubiquitination patterns the largest being a 50 KDa species. Interestingly, it is 50 KDa species which shows a nuclear localisation suggesting this has functional relevance (refer to Figures 5.20 and 5.22). One way to prove this notion is the immunoprecipitation and purification of C20orf201 protein before running the pure C20orf201 protein on the gel. We can probe the Western blot with an antibody to ubiquitin to determine if the C20orf201 is ubiquitinated. To confirm such a result, another experiment should be run, but in reverse, using an antibody against ubiquitin and purifying all ubiquitinated proteins. This should be followed by running them on the gel and probing the Western blot with an antibody against C20orf201. If the results are still the same, we would then have strong evidence that the protein is ubiquitinated.

5.3.4 C20orf201 function

Only one paper has been published about the human *C20orf201* gene. In this study, *C20orf201* ORF was integrated into Flp-In T-REx-293 to ask the question "Does the overexpression of *C20orf201* in the Flp-In T-REx-293 cell lines lead to changes in the morphology of the cells or affect the cells ability to survive? These cells were grown in a medium containing tetracycline to induce the gene expression of *C20orf201*; however, no abnormal observations were observed and cells apparent to grow normally, without any obvious phenotypic changes, although to date this analysis was limited due to time constraints.

5.4 Conclusion

The *C20orf201* gene variant 2 can be considered a good CT candidate gene because its expression patterns are clearly restricted to the normal testis and in central nervous tissues. Importantly, it is also expressed in different types of cancer tissues and cell lines, which could

present cancer biomarkers. The results from the Western blots correspond to the RT-PCR results in both the normal and cancer tissues tested here. To study the *C20orf201* function, it was inserted into a human cell line to establish a stable cell line expressing the *C20orf201* gene. We originally thought that translation of *C20orf201* mRNA was inhibited in the Flp-In T-REx-293::*C20orf201* (+Kozak or –Kozak) cells that were not treated with tetracycline. However, the subsequent proposal that the 50 KDa species could in fact be a modified (ubiquitinate) form of C20orf201 lead us to moderate this hypothesis, although the question of a translational inhibition mechanism is addressed in Chapter 6. Therefore, further studies are required to study this protein. As a first step, the hypothesis that C20orf201 is ubiquitinated in cancer cells should be tested. Importantly, Western blotting of C20orf201 on normal testis detected a faint band of 50 KDa species, but in a breast cancer was a strong band in the nucleus. This could be a cancer-specific mechanism.

Chapter 6.0: Functional analysis of different fragments of the human *C20orf201* gene

6.1 Introduction

The Western blots presented in Chapter 5 showed no detectable protein levels of C20orf201 at 26 KDa in Flp-In T-REx-293 cells, which had been inserted with *C20orf201* cDNA with and without the Kozak sequences, unless the cells were supplemented with tetracycline (refer to Figure 5.16). Nevertheless, RT-PCR results showed evidence of significant gene expression in all cells, regardless of the presence or absence of tetracycline treatment (refer to Figure 5.13). This indicated the possible presence of a mechanism that allows transcription but inhibits translation of the C20orf201 gene in these cell lines in the absence of tetracycline. The question addressed in this chapter is to determine whether *C20orf201* mRNA sequences serve to inhibit translation. The fact that further induction of gene expression by addition of tetracycline permits translation might suggest an inhibitory mechanism becomes saturated by excessive mRNA thus triggering translation. Such a translational inhibition mechanism exists in yeast for meiotic mRNAs that are produced during the mitotic cell cycle. This mechanism involves the Mmi1 protein which is a YTH family protein. Members of this family of proteins are also found in humans indicating such a mechanism could exist. To test this hypothesis that a similar mechanism is operating here, we divided the entire open reading frame of the *C20orf201* gene into 7 fragments to determine which fragment/s might act as a binding site to stop the translation of the *C20orf201* mRNA.

6.2 Results

6.2.1 Cloning of *C20orf201* fragments 1 to 7 cDNAs into the pcDNA5/FRT/TO vector

A mammalian expression plasmid (pcDNA5/FRT/TO) was used in the cloning application to clone each fragment of *C20orf201* into individual vector. The cloning of the *C20orf201* fragments was performed by adding the Kozak consensus sequence to the 5' end of each fragment. The reason for this addition was because this sequence plays a fundamental role in stimulating protein translation (Kozak, 1991). Each *C20orf201* cDNA fragment differs from the others by 99 bp (33 amino acids). The full length size of C20orf201 amino acids is 240. The anti-C20orf201 antibody used in Chapter 5 recognises 51 amino acids from the entire

C20orf201 amino acids, so a His Tag sequence (CACCACCACCACCACCAC) was added to the 3' end of *C20orf201* of each fragment before the stop codon to ensure Western blot monitoring of translation using anti-His monoclonal antibody. The *Hin*dIII restriction enzyme was used in cloning each of the 7 fragments of the *C20orf201* gene. A *Hin*dIII recognition sequence is present at the multiple cloning site (MCS) of the pcDNA5/FRT/TO vector but does not exist in the *C20orf201* cDNA sequences; for this reason, the *Hin*dIII was chosen from the 10 unique restriction enzymes appearing at the MCS of the vector.

The cloning of individual *C20orf201* fragments was initiated by amplification of *C20orf201* cDNA from the pCMV6-AC::*C20orf201* vector using C20orf201 F7 and R7 primers for fragment 1; C20orf201 F7 and R8 primers for fragment 2; C20orf201 F7 and R9 primers for fragment 3; C20orf201 F7 and R10 primers for fragment 4; C20orf201 F7 and R11 primers for fragment 5; C20orf201 F7 and R12 primers for fragment 6; or C20orf201 F7 and R13 primers for fragment 7. The amplification of *C20orf201* fragments 1 and 2 were performed using the Phusion High Fidelity PCR Master Mix (BioLabs; M0532S), while other *C20orf201* fragments were amplified using the BioMix Red Master Mix (BioLine; BIO-25006). The amplified PCR products for fragments 1, 2, 3, 4, 5, 6 and 7 were 744 bp, 645 bp, 546 bp, 447 bp, 348 bp, 249 bp and 150 bp, respectively. The sequence of each fragment was confirmed by sequencing.

Each fragment of *C20orf201* was cloned into the pcDNA5/FRT/TO plasmid using *Hin*dIII. The vector was analysed on agarose gel before and after *Hin*dIII digestion to verify the digestion efficiency of the vector, which yielded three molecules in the undigested vector (Figure 6.1A left) and a single fragment of about 5137 bp in the digested vector (Figure 6.1A right). Each fragment (PCR insert) was separated on a gel after digestion, purification and showed the expected single bands, based on DNA sequencing of all fragments (Figure 6.1B).

Each purified *C20orf201* inserts for every fragment was ligated into the purified digested pcDNA5/FRT/TO plasmid. The recombinant plasmids were transformed into competent *E. coli* DH5α. The clones for each fragment were analysed by PCR screening of 14, 29, and 17 randomly selected colonies for fragments 1, 2 and 3, respectively. A random selection of 15 colonies was screened by PCR for fragments 4, 5, 6 and 7. The set of primers used in the PCR amplification for each fragment was also used in the PCR screening. The PCR analysis results indicated that: 12 of the 14 *E. coli* colonies had *C20orf201* fragment 1 (Figure 6.2); 13 of the 29 *E. coli* colonies had *C20orf201* fragment 2 (Figure 6.3); 16 of the 17 *E. coli* colonies had

C20orf201 fragment 3 (Figure 6.4); all of the *E. coli* colonies had *C20orf201* fragment 4 (Figure 6.5); 2 of the 15 *E. coli* colonies had *C20orf201* fragment 5 (Figure 6.6); 14 of the 15 *E. coli* colonies had *C20orf201* fragment 6 (Figure 6.7) and 10 of the 15 *E. coli* colonies had *C20orf201* fragment 7 (Figure 6.8). All the PCR primers for each *C20orf201* fragment were validated using pCMV6-AC::*C20orf201* cDNA.

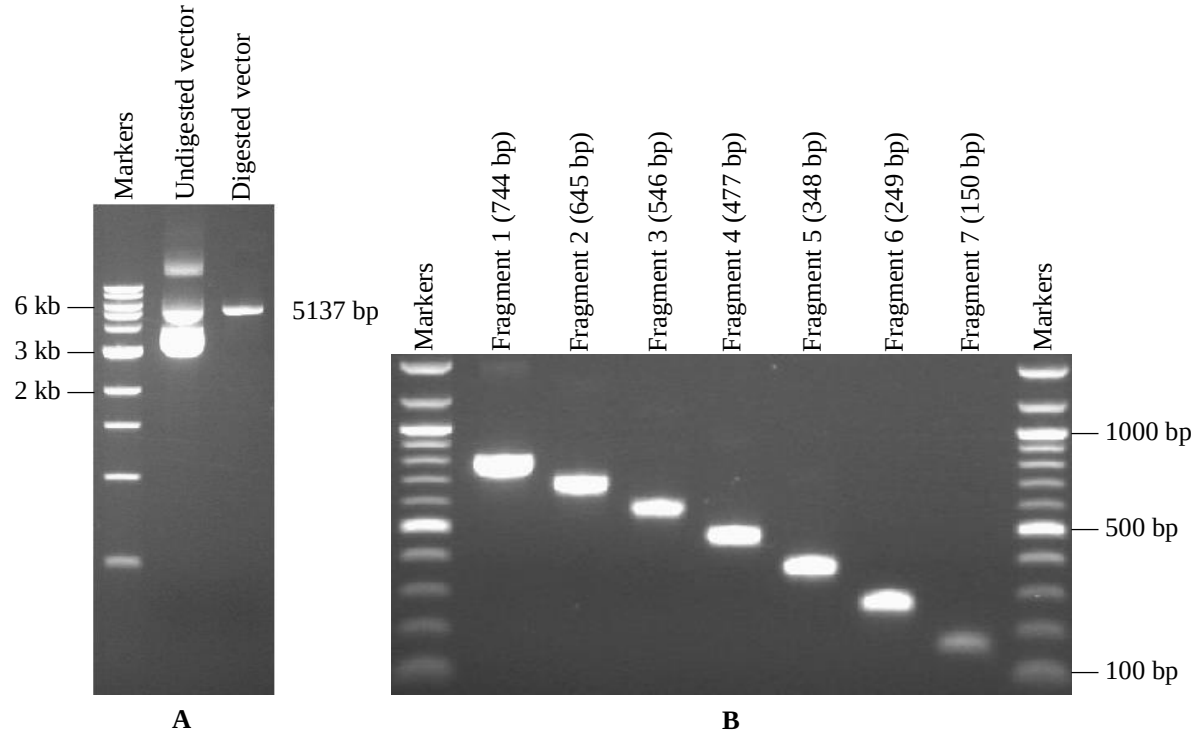


Figure 6.1 Amplification of *C20orf201* cDNA fragments from recombinant pCMV6-AC::*C20orf201* vector. Agarose gel showing on panel (**A**) the pcDNA5/FRT/TO vector before (left) and after (right) digestion with *Hin*dIII restriction enzyme and purification. The enzyme linearizes the 5137 bp plasmid into one single fragment. Panel **(B)** shows the PCR amplification for seven fragments for *C20orf201* from pCMV6- AC::*C20orf201* vector using the same forward and different reverse primers. These fragments were digested with *Hin*dIII and then purified.

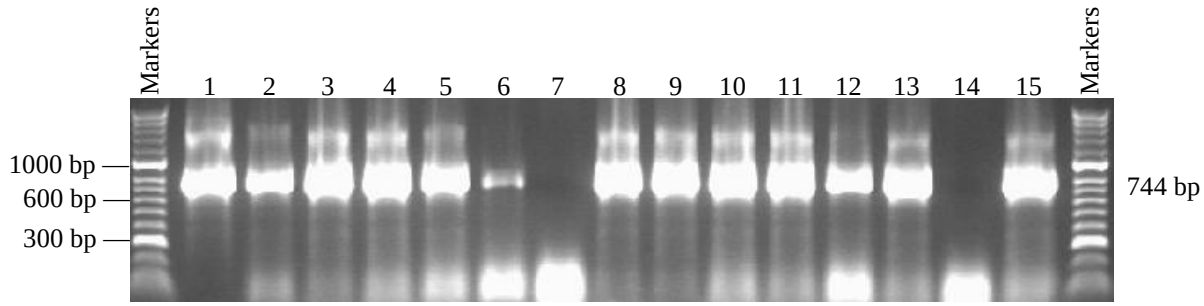


Figure 6.2 PCR screening of colonies for cloning of *C20orf201* fragment 1. Agarose gel showing the PCR expression of different *E. coli* colonies using cloning primers of the *C20orf201* fragment 1 (F7, R7). The first band on the left (lane 1) shows the presence of *C20orf201* in the pCMV6-AC::*C20orf201* vector to validate that the primers were working. The samples in lanes 2-6, 8-13 and 15 have an insert of the expected size of 744 bp.

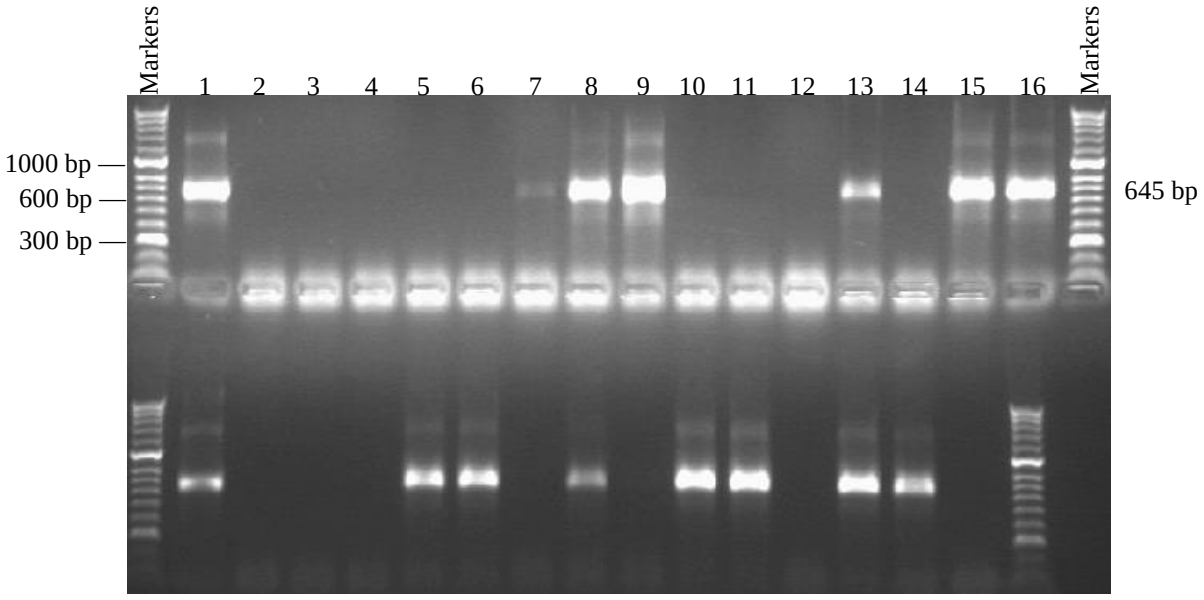


Figure 6.3 PCR screening of colonies for cloning of *C20orf201* fragment 2. Agarose gel showing the PCR expression of different *E. coli* colonies by using cloning primers of *C20orf201* fragment 2 (F7, R8). The first band on the left (lane 1) shows the presence of *C20orf201* in the pCMV6-AC::*C20orf201* vector to validate the primers. The samples in lanes 7-9, 13 and 15-16 have an insert of the expected size of 645 bp.

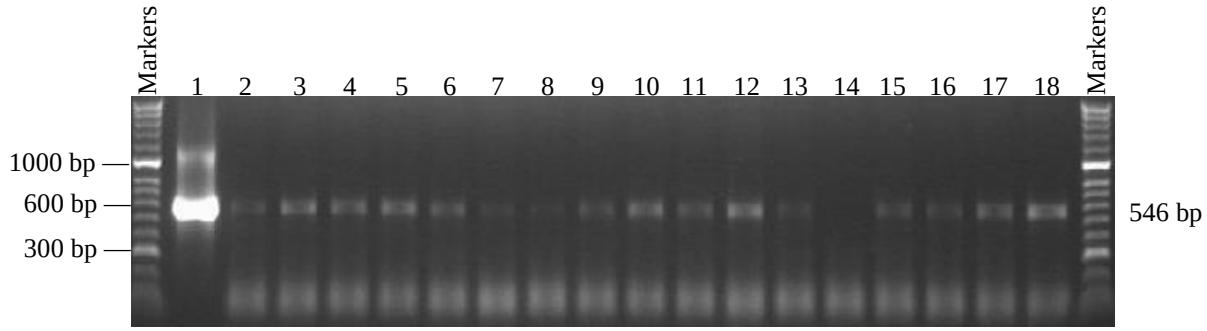


Figure 6.4 PCR screening of colonies for cloning of *C20orf201* fragment 3. Agarose gel showing the PCR products from different *E. coli* colonies using cloning primers of *C20orf201* fragment 3 (F7, R9). The first band on the left (lane 1) shows the presence of *C20orf201* in the pCMV6-AC::*C20orf201* vector to validate the primers. The samples in lanes 2-18, except lane 14, have an insert of the expected size of 546 bp.

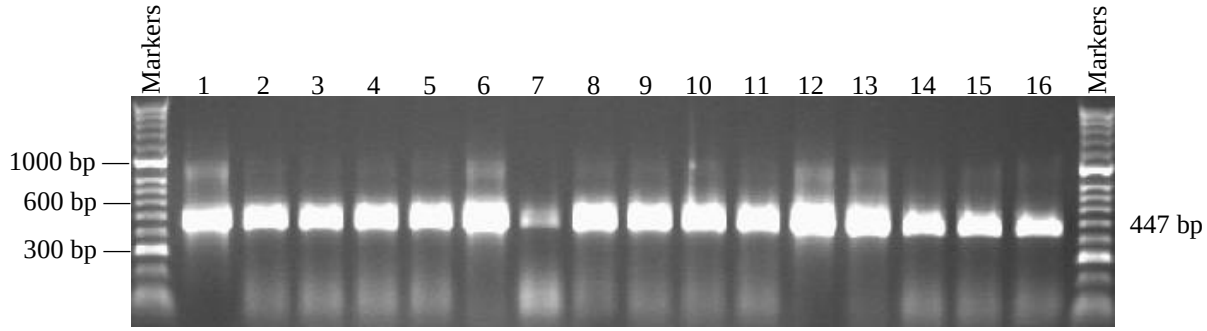


Figure 6.5 PCR screening of colonies for cloning of *C20orf201* fragment 4. Agarose gel showing the PCR products from different *E. coli* colonies using cloning primers of *C20orf201* fragment 4 (F7, R10). The first band on the left (lane 1) shows the presence of *C20orf201* in the pCMV6-AC::*C20orf201* vector to validate the primers. All samples in lanes 2-16 have an insert of the expected size of 447 bp.

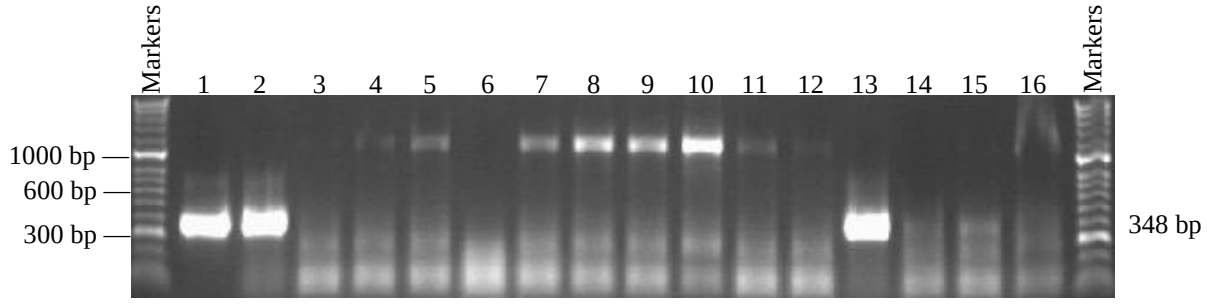


Figure 6.6 PCR screening of colonies for cloning of *C20orf201* fragment 5. Agarose gel showing the PCR products from different *E. coli* colonies using cloning primers of *C20orf201* fragment 5 (F7, R11). The first band on the left (lane 1) shows the presence of *C20orf201* in the pCMV6-AC::*C20orf201* vector to validate the primers. The samples in lanes 2 and 13 have an insert of the expected size of 348 bp.

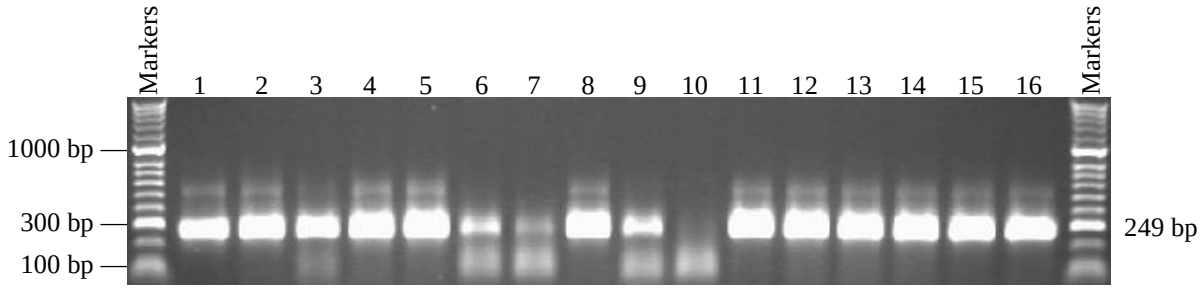


Figure 6.7 PCR screening of colonies for cloning of *C20orf201* fragment 6. Agarose gel showing the PCR products from different *E. coli* colonies using cloning primers of *C20orf201* fragment 6 (F7, R12). The first band on the left (lane 1) shows the expression of *C20orf201* in the pCMV6-AC::*C20orf201* vector to validate the primers. All samples in lanes 2-16, except lane 10, have an insert of the expected size of 249 bp.

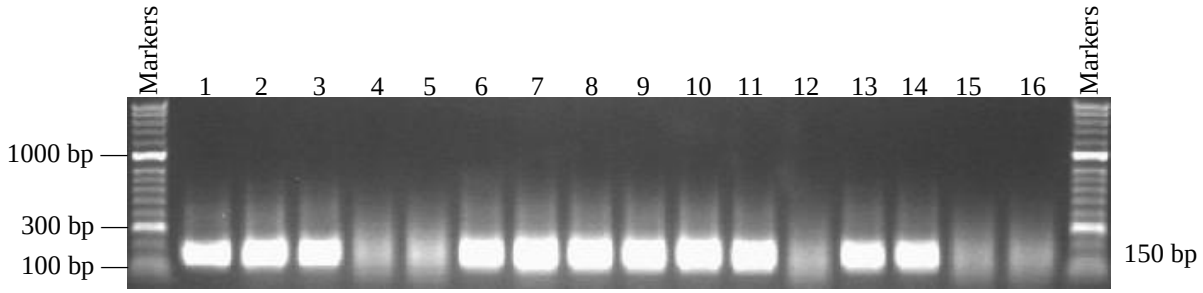


Figure 6.8 PCR screening of colonies for cloning of *C20orf201* fragment 7. Agarose gel showing the PCR products from different *E. coli* colonies using cloning primers of *C20orf201* fragment 7 (F7, R13). The first band on the left (lane 1) shows the presence of *C20orf201* in the pCMV6-AC::*C20orf201* vector to validate the primers. The samples in lanes 2-3, 6-11 and 13-14 have an insert of the expected size of 150 bp.

6.2.2 Screening of successfully cloned inserts by restriction enzyme digestion

A total of 7 pcDNA5/FRT/TO::*C20orf201* plasmids containing different *C20orf201* fragments were isolated from the *E. coli* cells. The presence of each *C20orf201* fragment was verified by a *Hin*dIII digestion. This digestion confirmed that all the digested recombinant plasmids for both fragments 1 and 2 contained the desired fragments (Figures 6.9 and 6.10, respectively). Figures 6.11 to 6.15 show the number of successful cloning for fragment 3 to 7, respectively. The successful cloning of each fragment of *C20orf201* revealed two fragments (bands) in each digested plasmid. The larger fragments showed the expected vector size of 5137 bp, whereas the smaller bands showed *C20orf201* fragments of different sizes, according to the insert length.

6.2.3 Confirming the orientation of *C20orf201* fragments 1 and 2 by restriction enzyme digestion

Two different methods were used to validate the insert (*C20orf201* fragments) orientations: restriction enzyme digestion for *C20orf201* fragments 1 and 2 and DNA sequencing for other fragments. Insertion of distinct fragments of *C20orf201* cDNA into the pcDNA5/FRT/TO vector using a single restriction enzyme leads to either correct or incorrect orientations. For this reason, digestion of another restriction enzyme is also required to compare between the correct and the opposite orientations of these inserts. *Xho*I restriction enzyme was chosen to check the insert orientation subcloned into the pcDNA5/FRT/TO. This is because *Xho*I can cut only once within the insert, but significantly off centre (close the 5'end), and also cuts nearby in the pcDNA5/FRT/TO backbone. Therefore, *Xho*I gave different size fragments depending on the orientation of the fragment.

Digestion of pcDNA5/FRT/TO::*C20orf201* fragment 1 plasmids with *Xho*I gave two bands; the 524 bp insert and the 5357 bp backbone for the correct orientation or the 386 bp insert and the 5495 bp backbone for the incorrect orientation of *C20orf201* fragment 1. A total of 5 of the 11 purified plasmids contained only the right orientation of fragment 1 (Figure 6.16). By contrast, digestion of pcDNA5/FRT/TO::*C20orf201* fragment 2 plasmids with *Xho*I restriction enzyme gave two bands, but of different sizes. These bands are the 425 bp insert and the 5357 bp backbone for the right direction or the 386 bp insert and the 5396 bp backbone for the opposite direction for *C20orf201* fragment 2. A total of 4 of the 13 purified plasmids contained only the correct orientation of fragment 2 (Figure 6.17). One plasmid that

contained the correct orientation for each fragment was sent for DNA sequencing to check the entire cDNA fragment.

To confirm the orientation of other *C20orf201* inserts (fragments 3-7), two plasmids of each *C20orf201* fragment were sent for DNA sequencing using two sets of uniform primers, including CMV forward and BGH reverse primers to verify that the fragments were subcloned in the correct orientation for expression and contained an ATG initiation codon and a stop codon. The sequences obtained were compared to the original sequence through the NCB1 (BLAST). These sequences were 100% identical to *C20orf201* without any mutation for the entire length of each *C20orf201* fragment.

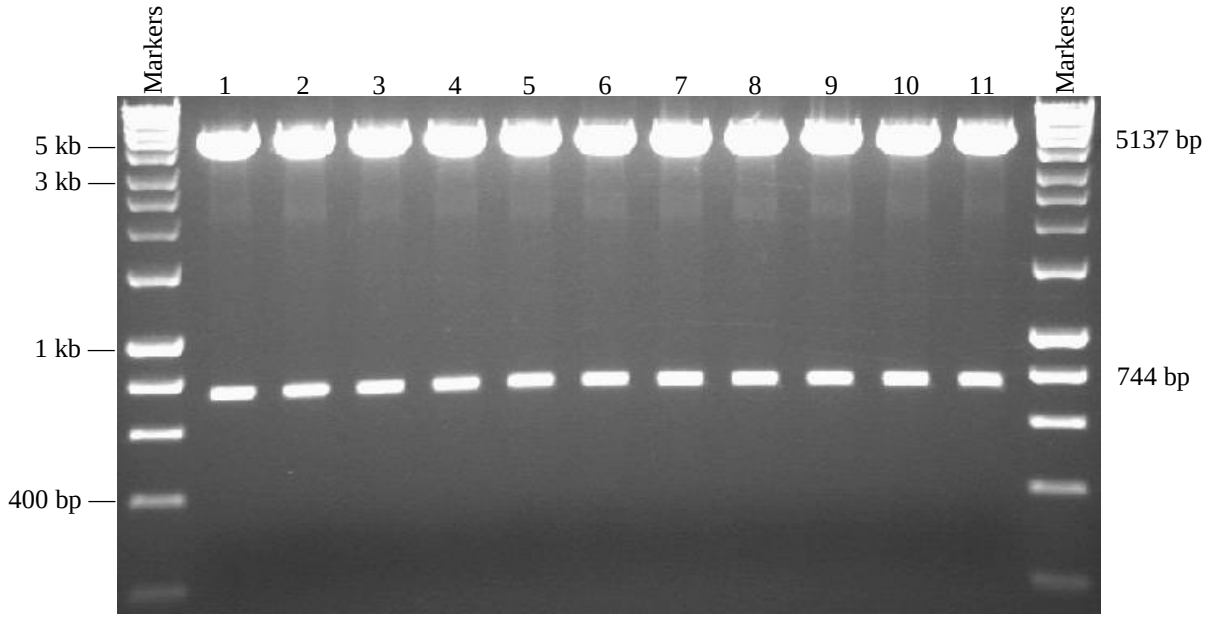


Figure 6.9 Digestion of recombinant plasmids. Agarose gel showing recombinant plasmids containing *C20orf201* fragment 1 digested by *Hin*dIII. Two bands are formed for all the recombinants. The upper bands show the pcDNA5/FRT/TO vector of 5137 bp in size, while the lower bands are fragment 1 of the *C20orf201* gene with a size of 744 bp. All minipreps in lanes 1-11 show successful cloning of the *C20orf201* fragment 1 cDNA.

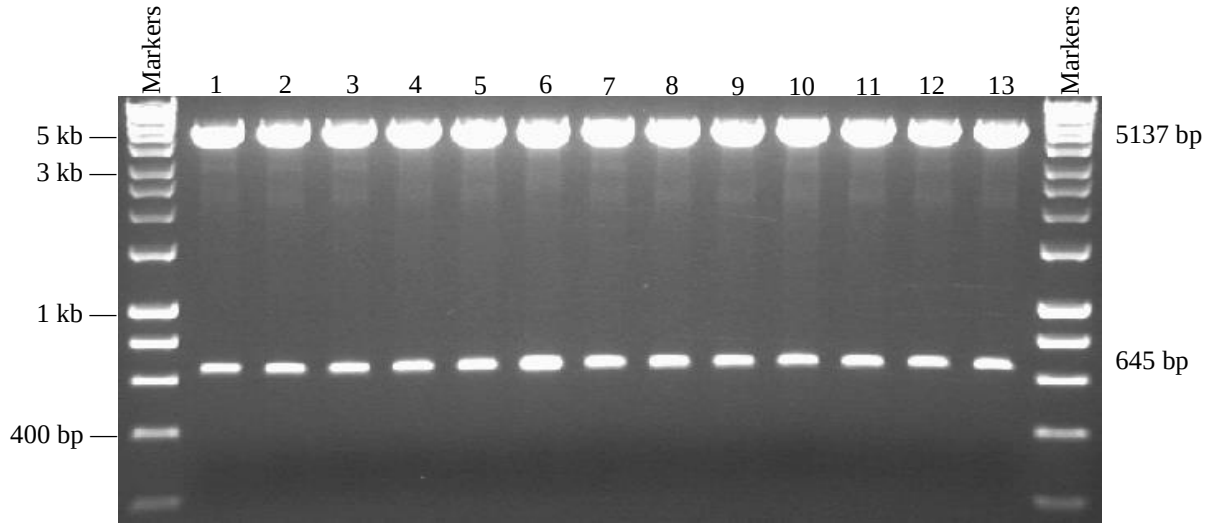


Figure 6.10 Digestion of recombinant plasmids. Agarose gel showing recombinant plasmids containing *C20orf201* fragment 2 digested by *Hin*dIII. Two bands are formed for all the recombinants. The upper bands demonstrate the pcDNA5/FRT/TO vector of 5137 bp in size, while the lower bands demonstrate fragment 2 of *C20orf201* gene with a size of 645 bp. All minipreps in lanes 1-13 have the *C20orf201* fragment 2 cDNA cloned.

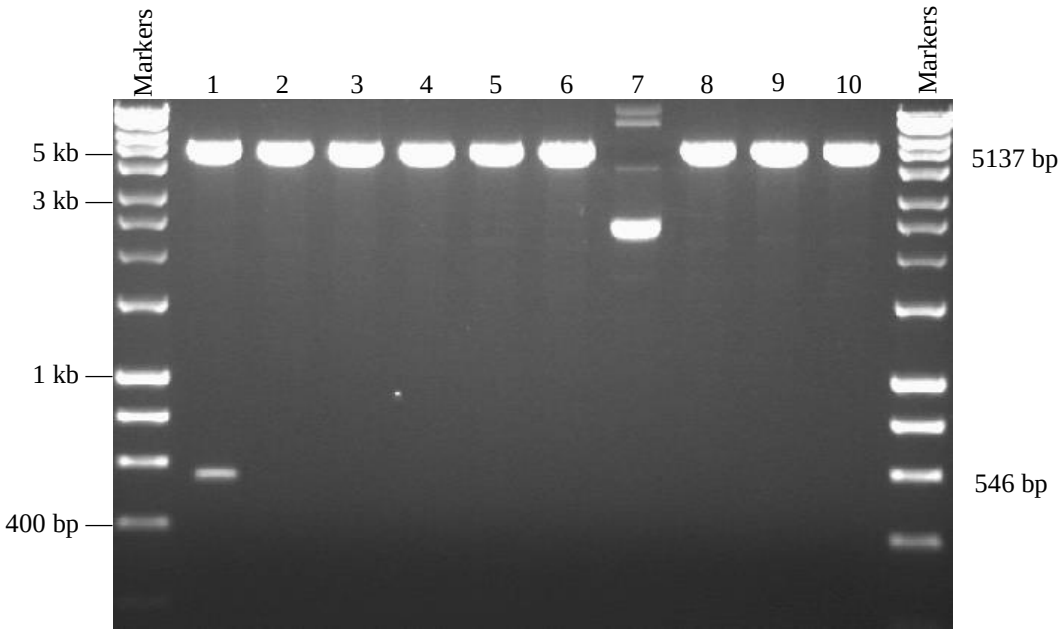


Figure 6.11 Digestion of recombinant plasmids. Agarose gel showing recombinant plasmids containing *C20orf201* fragment 3 digested by *Hin*dIII. Two bands are formed for only one recombinant. The upper band demonstrates the pcDNA5/FRT/TO vector of 5137 bp in size, while the lower band demonstrates fragment 3 of the *C20orf201* gene with a size of 546 bp. The miniprep in lane 1 has the only successful cloning of the *C20orf201* fragment 3 cDNA.

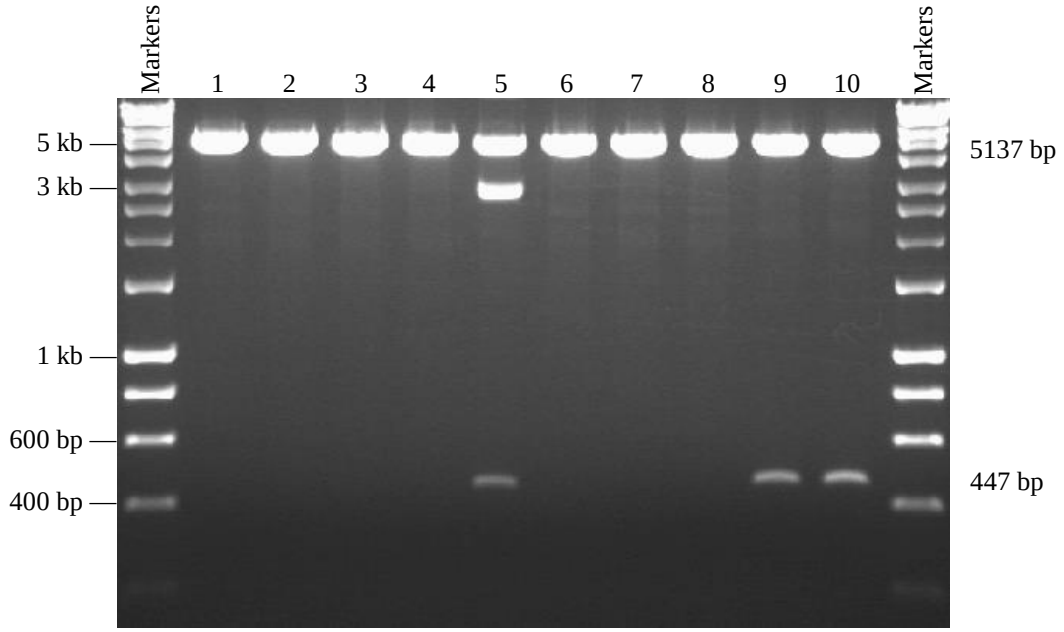


Figure 6.12 Digestion of recombinant plasmids. Agarose gel showing recombinant plasmids containing *C20orf201* fragment 4 digested by *Hin*dIII. Two bands are formed for three recombinants. The upper bands demonstrate the pcDNA5/FRT/TO vector of 5137 bp in size, while the lower bands demonstrate fragment 4 of the *C20orf201* gene with a size of 447 bp. The minipreps in lanes 5, 9 and 10 show successful cloning of the *C20orf201* fragment 4 cDNA.

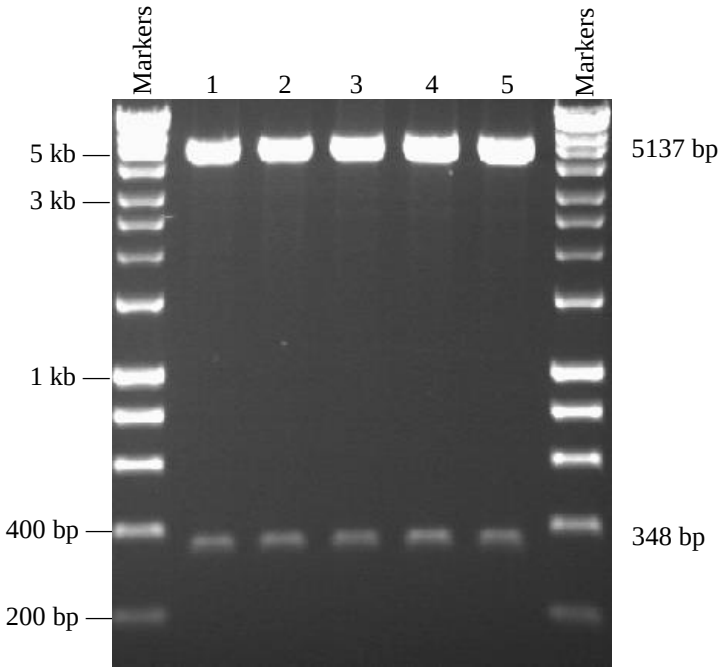


Figure 6.13 Digestion of recombinant plasmids. Agarose gel showing recombinant plasmids containing *C20orf201* fragment 5 digested by *Hin*dIII. Two bands are formed for all the recombinants. The upper bands demonstrate the pcDNA5/FRT/TO vector of 5137 bp in size, while the lower bands demonstrate fragment 5 of the *C20orf201* gene with a size of 348 bp. All minipreps in lanes 1-5 have successful cloning of the *C20orf201* fragment 5 cDNA.

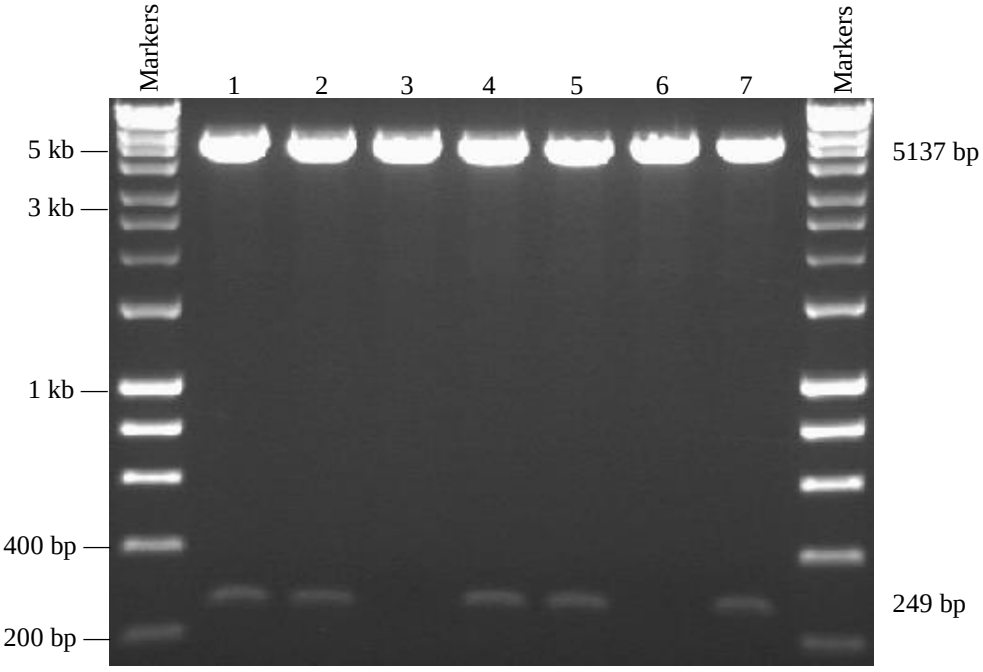


Figure 6.14 Digestion of recombinant plasmids. Agarose gel showing recombinant plasmids containing *C20orf201* fragment 6 digested by *Hin*dIII. Two bands are formed for 5 recombinants. The upper bands demonstrate the pcDNA5/FRT/TO vector of 5137 bp in size, while the lower bands demonstrate fragment 6 of the *C20orf201* gene with a size of 249 bp. Minipreps in lanes 1, 2, 4, 5 and 7 show successful cloning of the *C20orf201* fragment 6 cDNA.

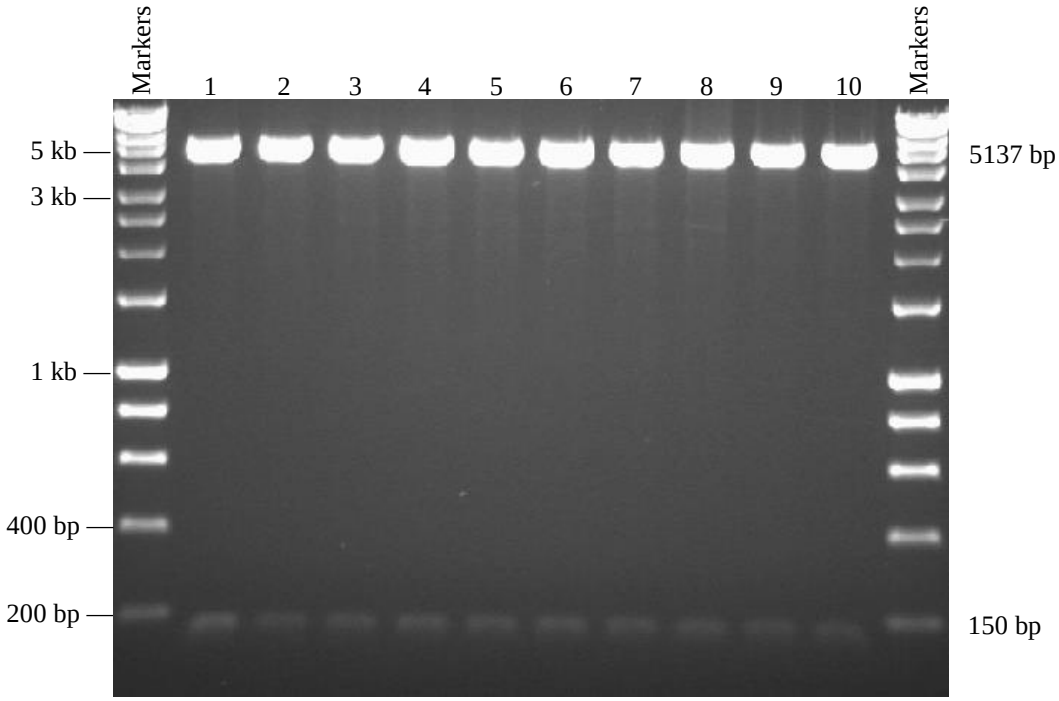


Figure 6.15 Digestion of recombinant plasmids. Agarose gel showing recombinant plasmids containing *C20orf201* fragment 7 digested by *Hin*dIII. Two bands are formed for all the recombinants. The upper bands demonstrate the pcDNA5/FRT/TO vector of 5137 bp in size, while the lower bands demonstrate fragment 7 of the *C20orf201* gene with a size of 150 bp. All the minipreps in lanes 1-10 show successful cloning of the *C20orf201* fragment 7 cDNA.

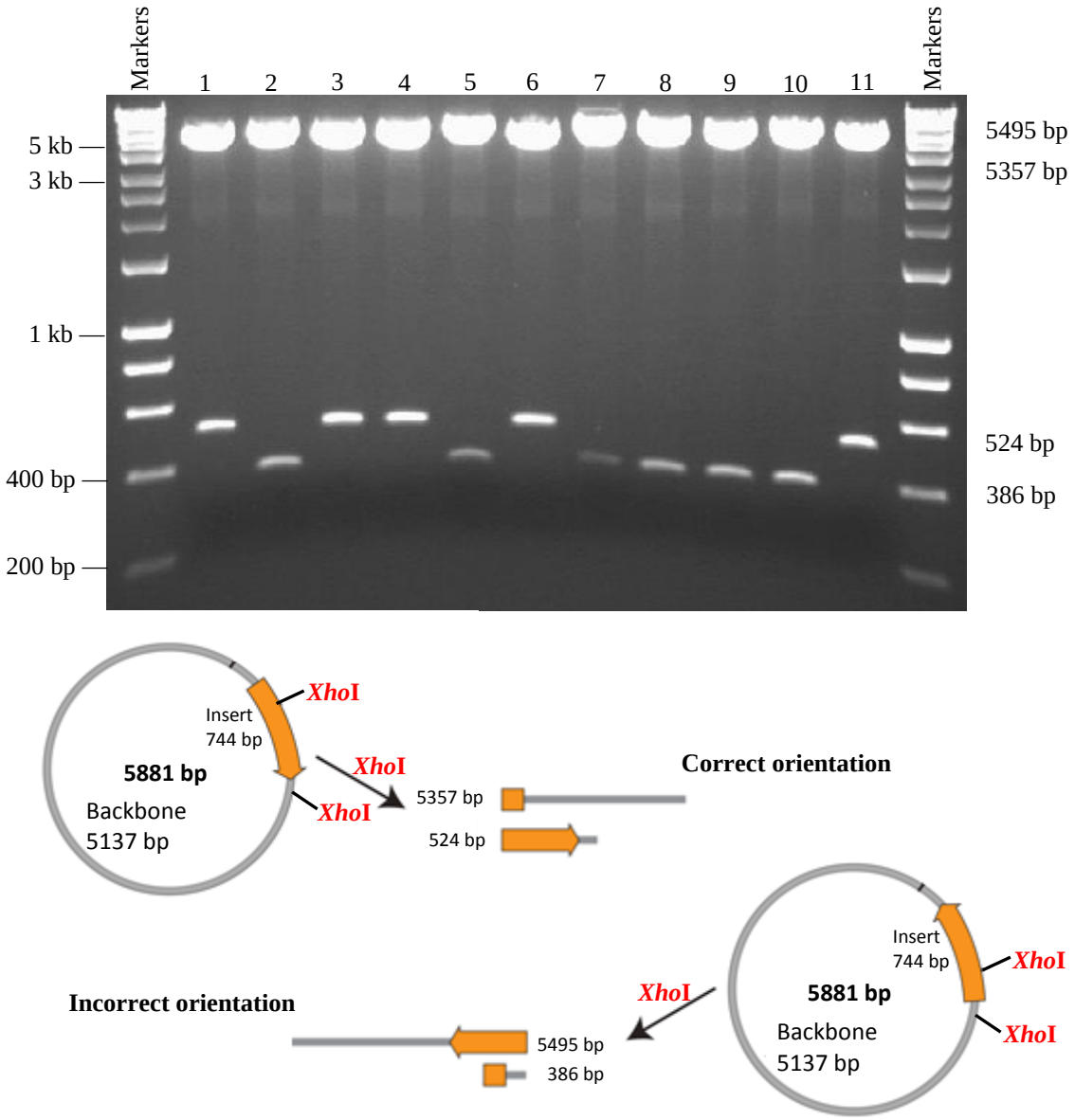


Figure 6.16 Verifying fragment 1 orientation by restriction digest. Agarose gel showing digestion by *Xho*I of recombinant plasmids containing fragment 1 to determine the correct orientation of the inserts. Depending on which way the *C20orf201* fragment 1 is inserted, the *Xho*I enzyme will produce fragments of 524 + 5357 bp, or 386 + 5495 bp; this is because the inserts are oriented in two different directions in the vector. Minipreps 1, 3, 4, 6 and 11 all contain inserts with the right direction, while minipreps 2, 5, 7, 8, 9 and 10 have the fragment inserted in the opposite direction.

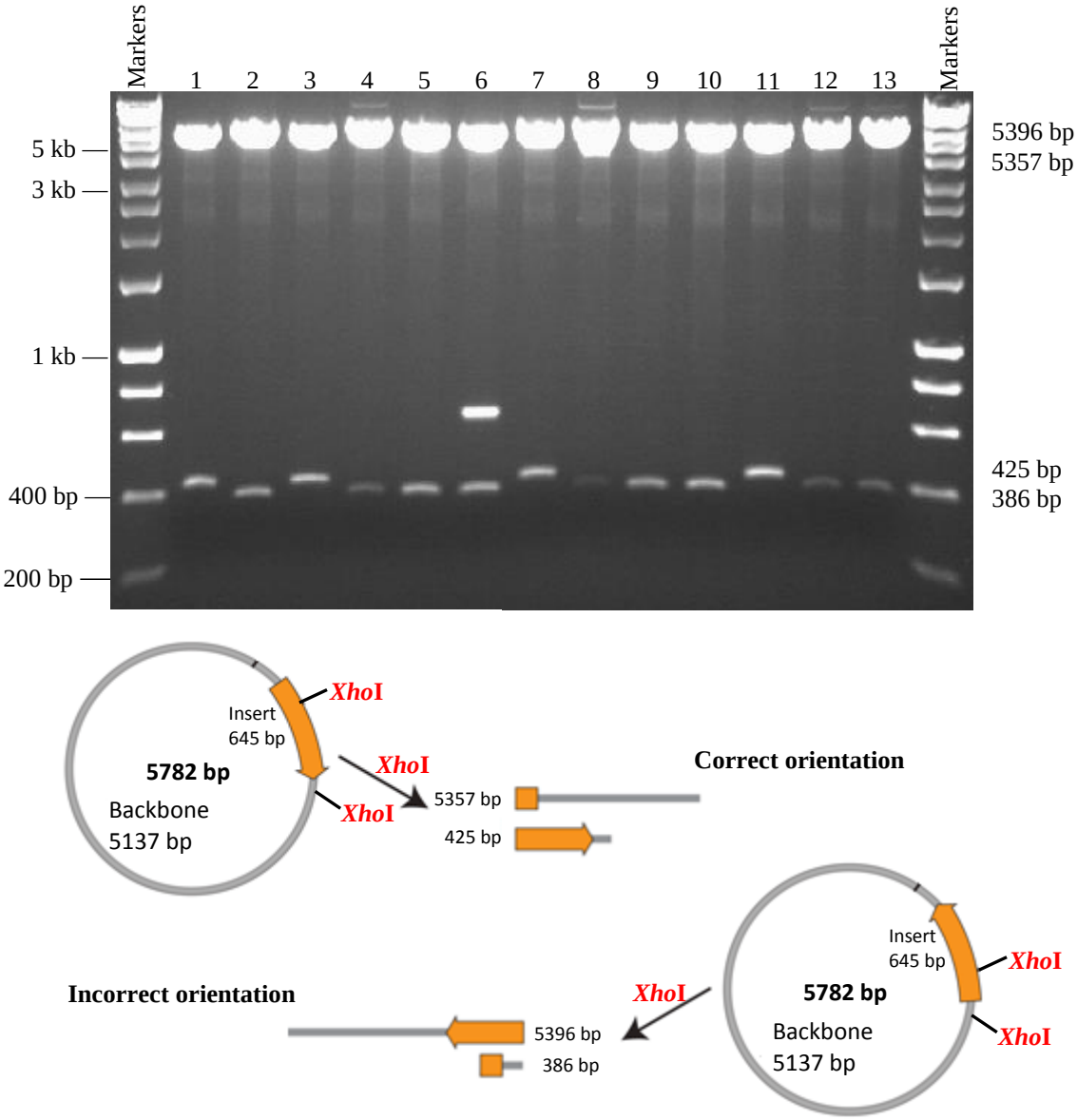


Figure 6.17 Verifying fragment 2 orientation by restriction digest. Agarose gel showing digestion by *Xho*I of recombinant plasmids containing *C20orf201* fragment 2 to determine the correct orientation of the inserts. Depending on which way the *C20orf201* fragment 2 is inserted, the *Xho*I enzyme will produce fragments of 425 + 5357 bp, or 386 + 5396 bp; this is because the inserts are oriented in two different directions in the vector. Minipreps 1, 3, 7 and 11 all contain inserts with the right direction, while minipreps 2, 4, 5, 6, 8, 9, 10, 12 and 13 have the fragment inserted in the opposite direction.

6.2.4 Insertion of pcDNA5/FRT/TO containing different fragments of *C20orf201* cDNA into the human Flp-In T-REx-293 cell line

6.2.4.1 Generation and selection of a stable expression cell line

Since different fragments were inserted into different pcDNA5/FRT/TO plasmids, the different recombinant plasmids required digestion to make sure the plasmid that contained the desired *C20orf201* fragment was chosen correctly. The digestion was carried out prior to insertion of the fragment of interest into the Flp-In T-REx-293 cells. The results of this digestion gave two bands for each recombinant plasmid; the larger bands showed the size of the plasmid DNA, while the smaller bands were related to the different sizes of inserts, which depended on the fragment size (Figure 6.18)

Integration of different fragments of *C20orf201* cDNA into the entire genome of the Flp-In T-REx-293 cell lines was dependent on the pOG44 vector (refer to Figure 4.10). The insertion occurred by co-transfection of the Flp-In T-REx-293 with a mixed DNA containing the following: pcDNA5/FRT/TO::*C20orf201* cDNA (fragments 1-7) and pOG44 vector at a ratio of 1:9 (pcDNA5/FRT/TO: pOG44). The Flp-In T-REx-293 cell line transfected with pOG44 and pcDNA5/FRT/TO alone (without the fragment of interest) was used as a negative control. In brief, an untransfected Flp-In T-REx-293 cell line was grown in media containing 200 µg/ml zeocin and 10 µg/ml blasticidin. 48 hours after transfection; the media of cells were supplemented with 100 µg/ml of hygromycin instead of zeocin antibiotic to select of stable cell lines.

Stable cell lines can be selected for hygromycin resistance, zeocin sensitivity, lack of *β*-galactosidase activity and tetracycline-regulated gene expression. The expression of the *C20orf201* gene following insertion is controlled by the (CMV)/TetO2 promoter. When the cell was refreshed with medium containing hygromycin, the majority of the cells did not survive. After hygromycin selection and pre-screening under a light microscope, single hygromycin resistant clones were isolated and expanded for each construct to generate individual clonal cell lines (Figure 6.19A). Examples of picked cells after growing in 6 well plates for 24 hours and selecting for hygromycin are shown in Figure 6.19B.

6.2.4.2 Analysis of the successful integrant Flp-In T-REx-293 cell lines

The correct integration of the pcDNA5/FRT/TO construct containing *C20orf201* into the FRT site of the Flp-In T-REx-293 cell line genome was evaluated by PCR screening performed using genomic DNA as a template. The genomic DNA was extracted from the untransfected

and transfected Flp-In T-REx-293 cells. PCR products were found in all transfected cells that were integrated with the pcDNA5/FRT/TO::*C20orf201* (fragments 1-7), or the empty pcDNA5/FRT/TO when using primers located in the P_{SV40} (forward) and located in the hygromycin gene (reverse); however, the untransfected cells did not show any PCR products when using the same set of primers (Figure 6.20A). By contrast, PCR results using primers located in the *C20orf201* gene (forward (F7)) and in the *LacZ-Zeocin* gene (reverse) gave products only in the Flp-In T-REx-293 that had been integrated with pcDNA5/FRT/TO::*C20orf201* fragments 1-7 (Figure 6.20B). *β*-actin was used in PCR analysis to test the genomic DNA quality in each sample (Figure 6.20C). Therefore, the results of PCR screening suggest that the different fragments of the *C20orf201* gene were correctly integrated into the genomic DNA of the Flp-In T-REx-293 cells.

6.2.4.3 PCR amplification of different integrated *C20orf201* fragments for DNA sequencing

The analysis described in Section 6.2.4.2 suggested that all pcDNA5/FRT/TO constructs containing distinct fragments of *C20orf201* cDNA were inserted successfully into the genome. However, the integration of these fragments could lead to deletions in any of these fragments. Therefore, PCR analysis was conducted using genomic DNA isolated from these cell lines and tetracycline operator 2 (TetO2), and BGH as forward and reverse primers, respectively. The PCR results indicated that every cell line contained different fragments of *C20orf201* that gave the expected size of the PCR product when using TetO2 and BGH primers (Figure 6.21A). *β*-actin was used to validate the genomic DNA quality (Figure 6.21B). Genomic DNA for each sample was sent for DNA sequencing using TetO2 and BGH primers to confirm that the fragments integrated into the Flp-In T-REx-293 did not have mutations along the entire length of each *C20orf201* fragment. The sequences obtained were compared to the original sequence through the NCB1 (BLAST). These results showed that the sequences were 100% identical to each fragment of *C20orf201,* with no mutations found.

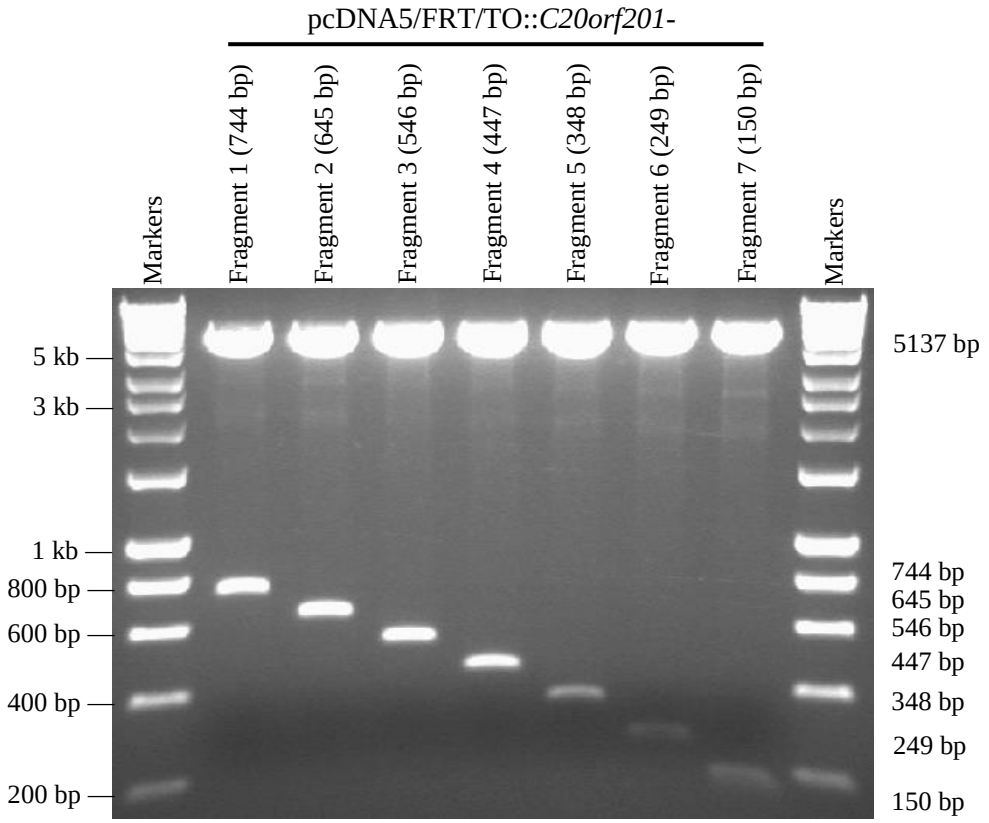


Figure 6.18 Confirmation of the digestion of different recombinant plasmids. Agarose gel showing digestion by *Hin*dIII of recombinant plasmids containing different fragments of *C20orf201* cDNAs (1, 2, 3, 4, 5, 6 and 7). Two bands are formed for all recombinants. The upper bands show the pcDNA5/FRT/TO vector of approximately 5137 bp in size, while the lower bands show different fragments of *C20orf201* cDNAs with different sizes of base pairs. The expected sizes for fragments 1, 2, 3, 4, 5, 6 and 7 are 744 bp, 645 bp, 546 bp, 447 bp, 348 bp, 249 bp and 150 bp, respectively.

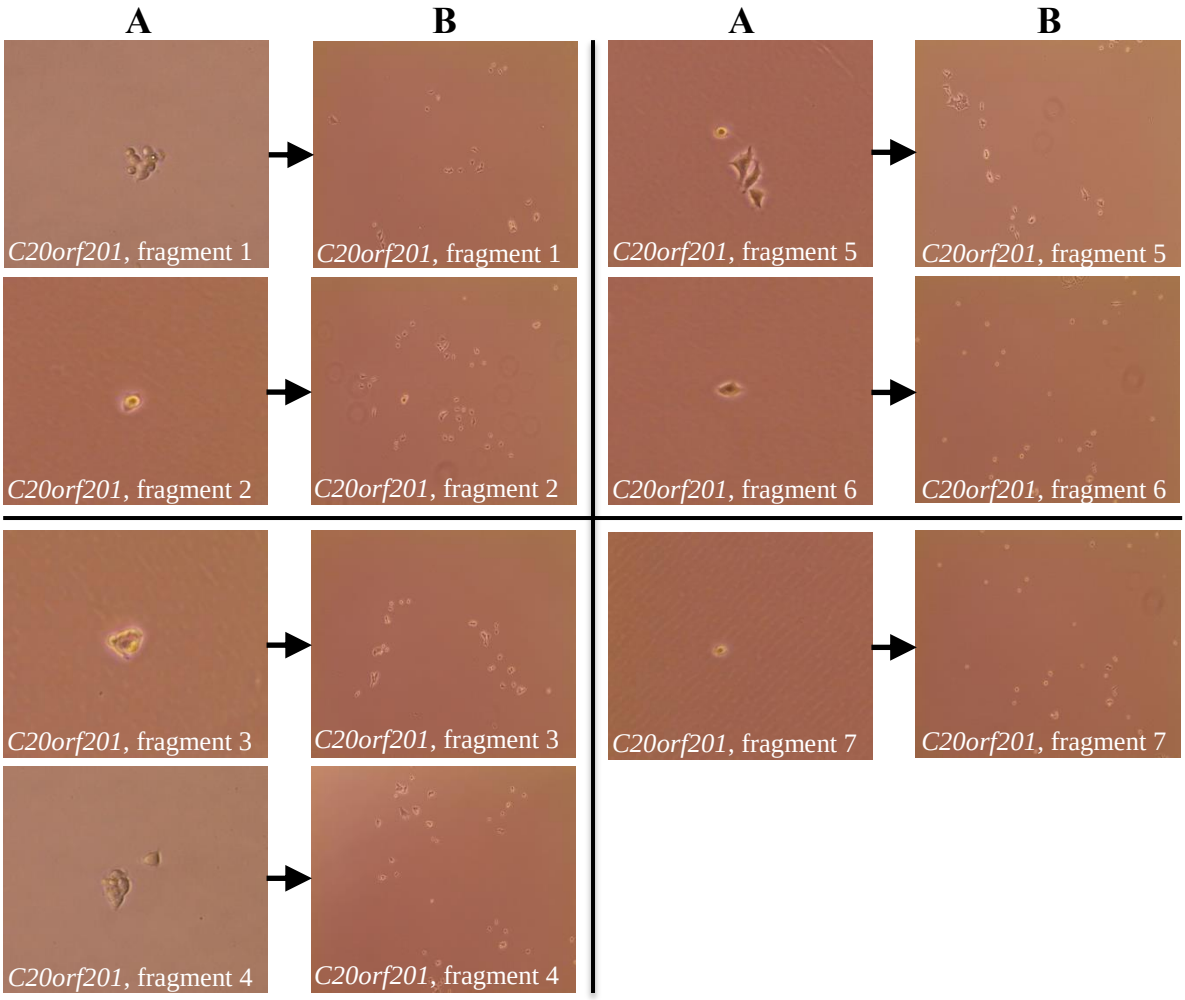


Figure 6.19 Examples of individual hygromycin resistance colonies. Panel **A** shows different resistance Flp-In T-REx-293:*C20orf201* colonies 2 days after adding an antibiotic marker, hygromycin (fragments from 1-7). Panel **B** shows the colonies 24 hours after picking and transfer to 6 well plates.

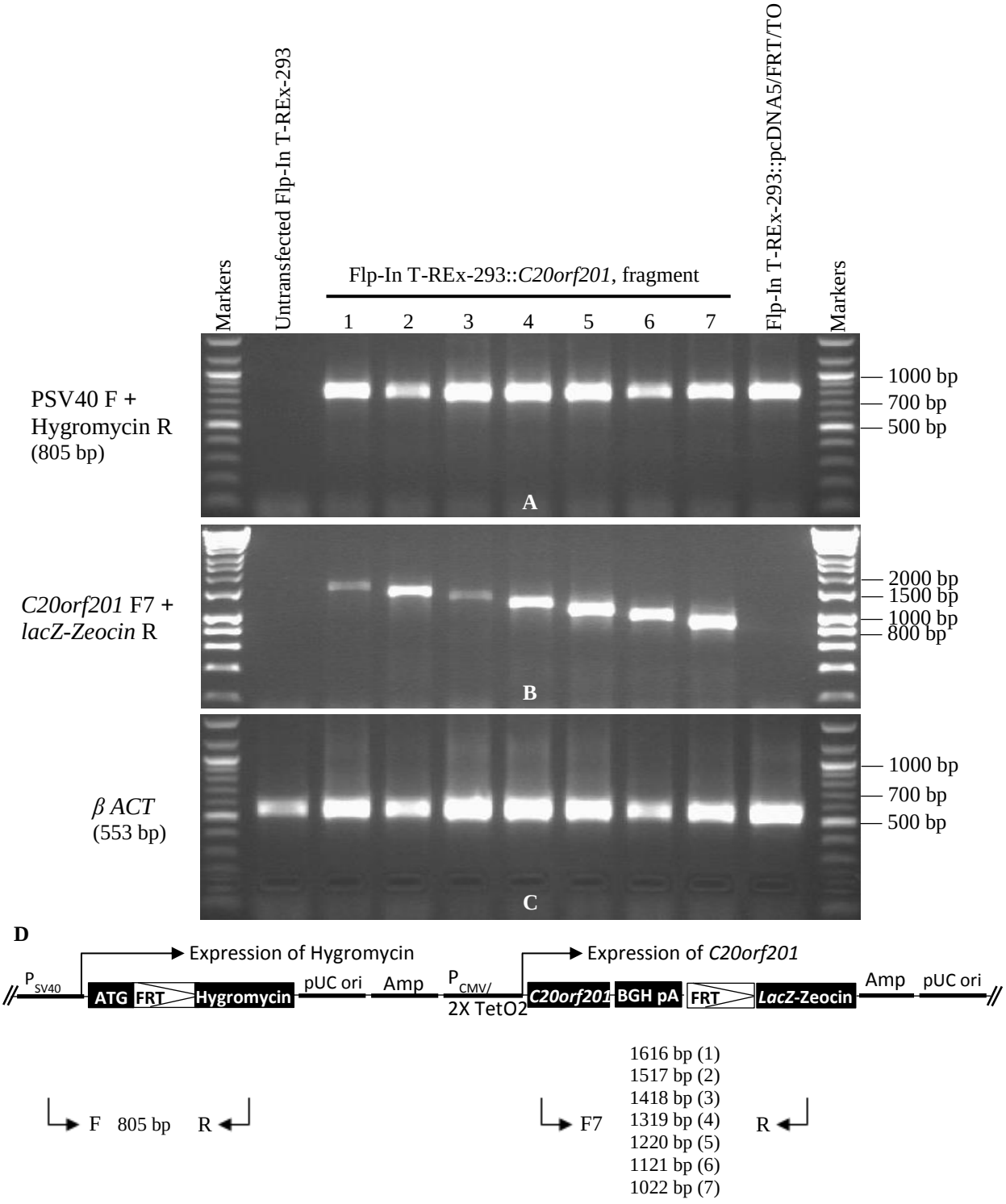


Figure 6.20 PCR screening of different *C20orf201* integrations in Flp-In T-REx-293 cell lines. Panel **(A)** shows the PCR analysis of the vector in the transfected cells with distinct fragments of *C20orf201* cDNA (lanes 1-7) or with a pcDNA5/FRT/TO vector alone (penultimate lane on the right) when using PSV40 as a forward primer and hygromycin as a reverse primer. Panel **(B)** shows the PCR products for *C20orf201* fragments when using the C20orf201 as forward primer and LacZ-Zeocin gene as a reverse primer. The expression profile for *Beta-actin* is shown as a positive control for the genomic DNA samples (Panel **(C)**). Panel **(D)** shows the localisation of *C20orf201* cDNAs in the genome of Flp-In T-REx-293 cell line.

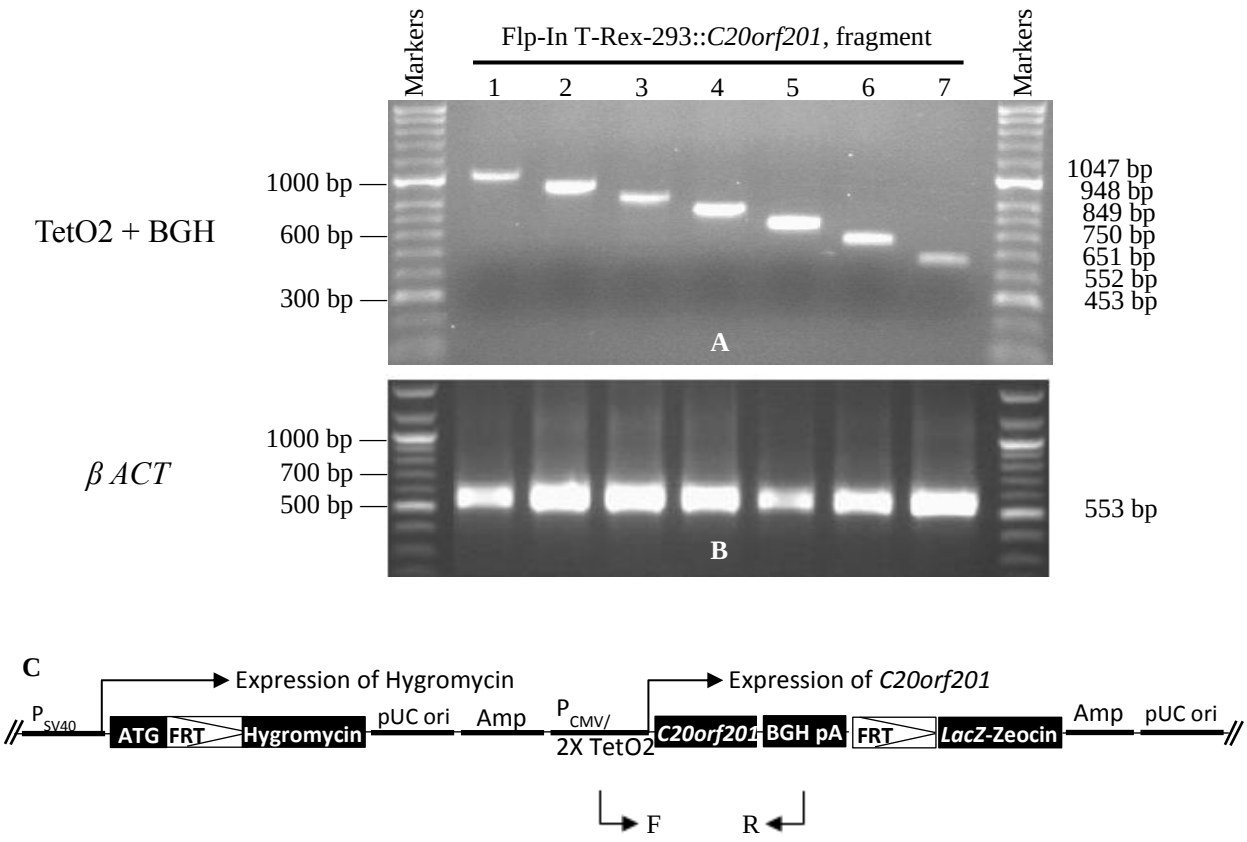


Figure 6.21 Amplification of different fragments of *C20orf201* for genomic DNA sequencing. Panel **(A)** shows the PCR products of distinct fragments of *C20orf201* in Flp-In T-Rex-293 cell lines (lanes 1-7) using tetracycline operator 2 as a forward primer and BGH gene as a reverse primer. The PCR profile for *β ACT* is shown as a positive control for the genomic DNA samples (Panel (**B)**). Panel **(C)** shows the localisation of *C20orf201* cDNAs in the genome of Flp-In T-Rex-293 cell line.

6.2.4.4 Investigation of the expression level of each *C20orf201* fragment in the human Flp-In T-REx-293 cell line

The expression of the *C20orf201* fragments 1 to 7 cDNA was investigated through RT-PCR and qRT- PCR analyses. Tetracycline was used at a final concentration of 2 µg/ml to induce transcription of the *C20orf201* from the pcDNA5/FRT/TO expression vector. The cDNAs were synthesised from the mRNAs from these cell lines. Cells transfected with empty pcDNA5/FRT/TO was used as a negative control. *β*-actin was used as a positive control for the cDNA samples. RT-PCR results showed that the expression of each *C20orf201* fragment was strongly observed in cell lines treated with or without tetracycline except fragment two in the absence of tetracycline. The primers used in RT-PCR analysis were the same forward and reverse primers used in cloning (refer to Section 6.2.1). However, neither untransfected cells nor transfected cells with empty pcDNA5/FRT/TO produced any RT-PCR products (Figure 6.22).

Another analysis used here to validate the *C20orf201* expression of each fragment gene was the qRT-PCR study. C20orf201 F4 and R4 primers were used in this investigation for *C20orf201* fragments 1 to 7 in different Flp-In T-REx-293 cell lines. In contrast, *β ACT* and *Lamin A* commercial positive primers were used. The NRT (no reverse transcriptase) and NTC (no template control) negative controls did not give Cq (quantification cycle) readings, which suggested no non-specific background and no genomic DNA contamination (Tables 6.1, 6.2 and 6.3). In the absence of tetracycline, the qRT-PCR results show a higher level of expression in the Flp-In T-REx-293::*C20orf201*, fragment 3 compared to other fragments and no expression was detected in the negative control tested here (Figure 6.23). However, in the presence of tetracycline, the qRT-PCR results displayed a higher expression levels in all *C20orf201* fragments compared to the cells not treated with tetracycline. Among the treated cells, the expression of *C20orf201* gene was highest for fragment 6, followed by 3, 5, 4, 7, 1 and 2 (Figures 6.24 and 6.25). The *C20orf201* qRT-PCR results were normalised to the qRT-PCR results for *Lamin A* and *β ACT*; the results are shown in Tables 6.1, 6.2 and 6.3.

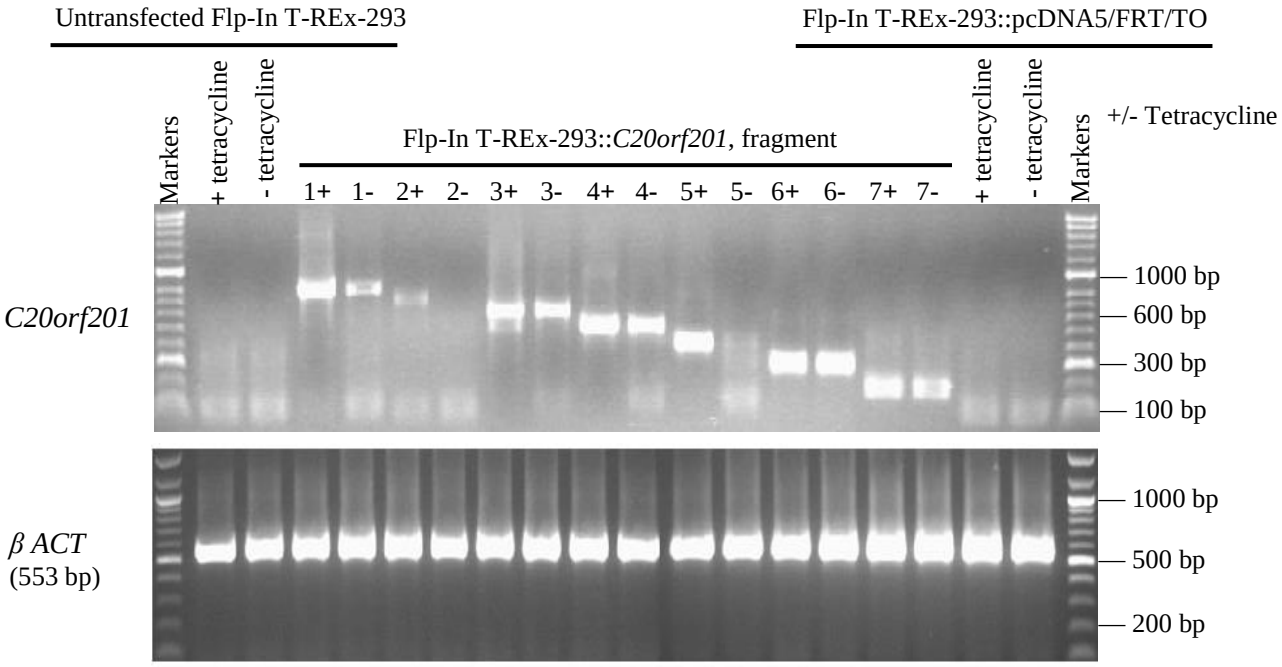


Figure 6.22 RT-PCR analysis of expression of *C20orf201* and *β ACT* in integrated Flp-In T-REx-293 cell lines. The top agarose gel shows the expression of different fragments of *C20orf201* in Flp-In T-REx-293 cell lines with and without tetracycline treatment for each fragment (lanes 1-7). The upper gel also shows the analysis of expression of *C20orf201* in the untransfected cells (the first two lanes on the left after the markers) and cells transfected with the vector alone (the last two wells on the right before the markers) in Flp-In T-REx-293 cell lines. The bottom agarose gel displays the expression profile for *β*-actin as a positive control for the cDNA samples.

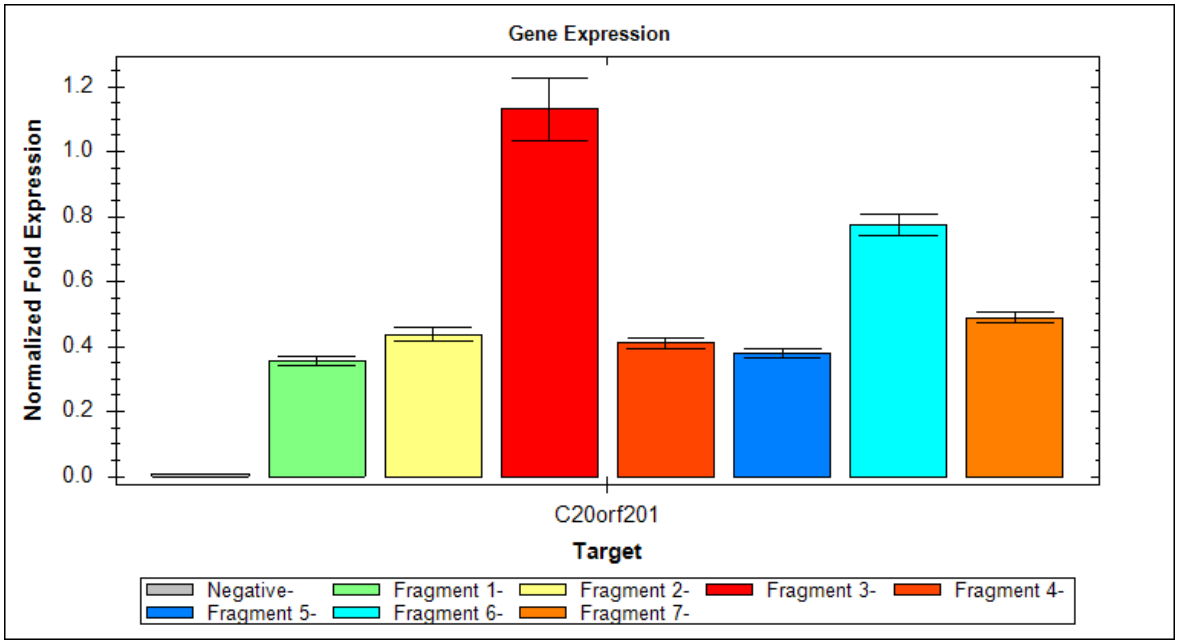


Figure 6.23 Real-time quantitative RT-PCR analysis of distinct fragments of *C20orf201* in Flp-In T-REx-293 cells. The bar chart displays the gene expression results for different fragments (1-7) of *C20orf201* in the Flp-In T-REx-293 cell lines without (-) adding tetracycline to the cell culture, compared to the cells cloned with the pcDNA5/FRT/TO vector alone (negative). Expression data were normalised to the *β ACT* and *Lamin A* reference genes. The error bars represent the standard error of the mean (SEM) for 3 repeats.

Table 6.1 Real-time qRT-PCR analysis of distinct fragments of *C20orf201* in Flp-In T-REx-293 cells. The table displays the mean of three repeats of the quantification cycle (Cq) and standard deviation (SD) for seven fragments (1-7) of *C20orf201* before (-) adding tetracycline. These results were normalised by *β ACT* and *Lamin A* expressions. The table also shows the Cq reading for the no reverse transcriptase (NRT) and no template control (NTC).

Target name	Sample name	Wells	Cq Mean	Cq SD
C20orf201	Flp-In T-REx-293: pcDNA5/FRT/TO-	3	30.09	0.134
β ACT	Flp-In T-REx-293: pcDNA5/FRT/TO-	3	17.76	0.009
Lamin A	Flp-In T-REx-293: pcDNA5/FRT/TO-	3	22.87	0.176
C20orf201	Flp-In T-REx-293: *C20orf201* F1-	3	25.41	0.056
β ACT	Flp-In T-REx-293: *C20orf201* F1-	3	17.34	0.054
Lamin A	Flp-In T-REx-293: *C20orf201* F1-	3	22.86	0.143
C20orf201	Flp-In T-REx-293: *C20orf201* F2-	3	24.96	0.108
β ACT	Flp-In T-REx-293: *C20orf201* F2-	3	17.25	0.062
Lamin A	Flp-In T-REx-293: *C20orf201* F2-	3	22.66	0.073
C20orf201	Flp-In T-REx-293: *C20orf201* F3-	3	23.56	0.210
β ACT	Flp-In T-REx-293: *C20orf201* F3-	3	17.31	0.039
Lamin A	Flp-In T-REx-293: *C20orf201* F3-	3	22.55	0.109
C20orf201	Flp-In T-REx-293: *C20orf201* F4-	3	25.18	0.081
β ACT	Flp-In T-REx-293: *C20orf201* F4-	3	17.42	0.075
Lamin A	Flp-In T-REx-293: *C20orf201* F4-	3	22.75	0.096
C20orf201	Flp-In T-REx-293: *C20orf201* F5-	3	25.06	0.083
β ACT	Flp-In T-REx-293: *C20orf201* F5-	3	17.26	0.055
Lamin A	Flp-In T-REx-293: *C20orf201* F5-	3	22.44	0.079
C20orf201	Flp-In T-REx-293: *C20orf201* F6-	3	24.27	0.058
β ACT	Flp-In T-REx-293: *C20orf201* F6-	3	17.35	0.147
Lamin A	Flp-In T-REx-293: *C20orf201* F6-	3	22.84	0.083
C20orf201	Flp-In T-REx-293: *C20orf201* F7-	3	24.61	0.025
β ACT	Flp-In T-REx-293: *C20orf201* F7-	3	17.07	0.053
Lamin A	Flp-In T-REx-293: *C20orf201* F7-	3	22.47	0.142
C20orf201	Flp-In T-REx-293: pcDNA5/FRT/TO- (NRT)	1	0.00	0.000
β ACT	Flp-In T-REx-293: pcDNA5/FRT/TO- (NRT)	1	0.00	0.000
Lamin A	Flp-In T-REx-293: pcDNA5/FRT/TO- (NRT)	1	0.00	0.000
C20orf201	Flp-In T-REx-293: *C20orf201* F1- (NRT)	1	0.00	0.000
β ACT	Flp-In T-REx-293: *C20orf201* F1- (NRT)	1	0.00	0.000
Lamin A	Flp-In T-REx-293: *C20orf201* F1- (NRT)	1	0.00	0.000
C20orf201	Flp-In T-REx-293: *C20orf201* F2- (NRT)	1	0.00	0.000
β ACT	Flp-In T-REx-293: *C20orf201* F2- (NRT)	1	0.00	0.000
Lamin A	Flp-In T-REx-293: *C20orf201* F2- (NRT)	1	0.00	0.000
C20orf201	Flp-In T-REx-293: *C20orf201* F3- (NRT)	1	0.00	0.000
β ACT	Flp-In T-REx-293: *C20orf201* F3- (NRT)	1	0.00	0.000
Lamin A	Flp-In T-REx-293: *C20orf201* F3- (NRT)	1	0.00	0.000
C20orf201	Flp-In T-REx-293: *C20orf201* F4-(NRT)	1	0.00	0.000
β ACT	Flp-In T-REx-293: *C20orf201* F4- (NRT)	1	0.00	0.000
Lamin A	Flp-In T-REx-293: *C20orf201* F4- (NRT)	1	0.00	0.000
C20orf201	Flp-In T-REx-293: *C20orf201* F5- (NRT)	1	0.00	0.000
β ACT	Flp-In T-REx-293: *C20orf201* F5- (NRT)	1	0.00	0.000
Lamin A	Flp-In T-REx-293: *C20orf201* F5- (NRT)	1	0.00	0.000
C20orf201	Flp-In T-REx-293: *C20orf201* F6- (NRT)	1	0.00	0.000
β ACT	Flp-In T-REx-293: *C20orf201* F6- (NRT)	1	0.00	0.000
Lamin A	Flp-In T-REx-293: *C20orf201* F6- (NRT)	1	0.00	0.000
C20orf201	Flp-In T-REx-293: *C20orf201* F7- (NRT)	1	0.00	0.000
β ACT	Flp-In T-REx-293: *C20orf201* F7- (NRT)	1	0.00	0.000
Lamin A	Flp-In T-REx-293: *C20orf201* F7- (NRT)	1	0.00	0.000
C20orf201	NTC	1	0.00	0.000
β ACT	NTC	1	0.00	0.000
Lamin A	NTC	1	0.00	0.000

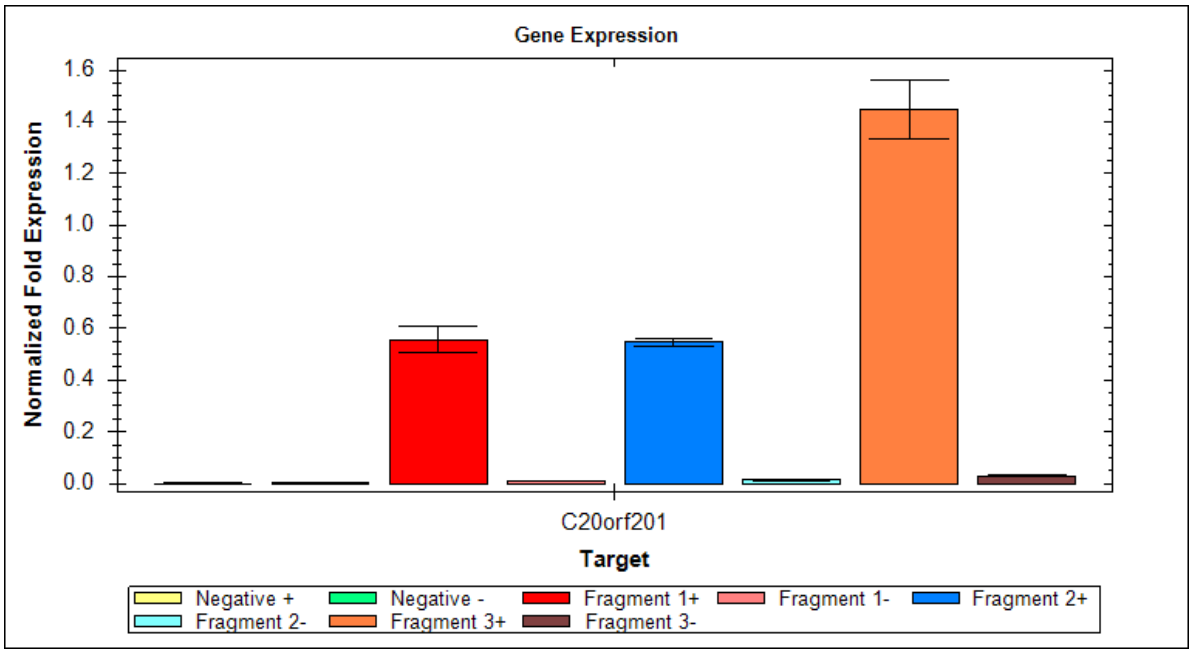


Figure 6.24 Real-time quantitative RT-PCR analysis of distinct fragments of *C20orf201* in Flp-In T-REx-293 cells. The bar chart displays the gene expression results for three fragments (1-3) of *C20orf201* in the Flp-In T-REx-293 cell lines with (+) and without (-) adding tetracycline to the cell culture, compared to the cells cloned with the pcDNA5/FRT/TO vector alone (negative). Expression data were normalised to the *β ACT* and *Lamin A* reference genes. The error bars represent the standard error of the mean (SEM) for 3 repeats.

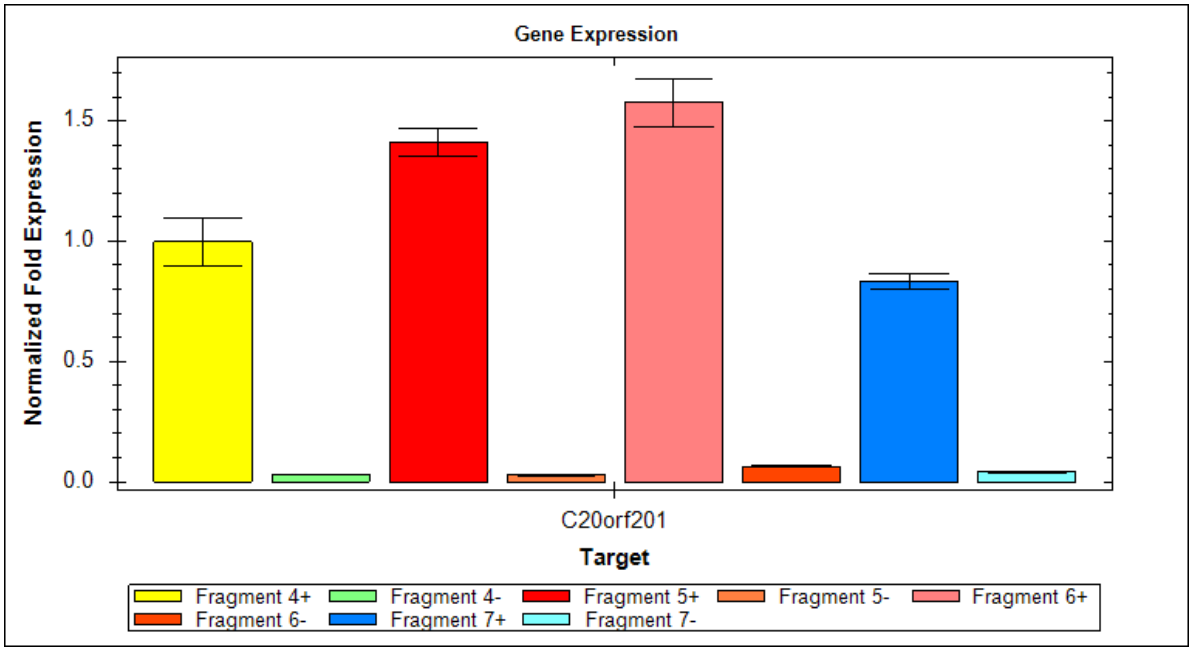


Figure 6.25 Real-time quantitative RT-PCR analysis of distinct fragments of *C20orf201* in Flp-In T-REx-293 cells. The bar chart displays the gene expression results for four fragments (4-7) of *C20orf201* in the Flp-In T-REx-293 cell lines with (+) and without (-) adding tetracycline to the cell culture, compared to the cells integrated with the pcDNA5/FRT/TO vector alone (negative). Expression data were normalised to the *β ACT* and *Lamin A* reference genes. The error bars represent the standard error of the mean (SEM) for 3 repeats.

Table 6.2 Real-time qRT-PCR analysis of distinct fragments of *C20orf201* in Flp-In T-Rex-293 cells. The table displays the mean of three repeats of the quantification cycle (Cq) and standard deviation (SD) for three fragments (1-3) of *C20orf201* before (-) and after (+) adding tetracycline. These results were normalised by *β ACT* and *Lamin A* expressions. The table also shows the Cq reading for the no reverse transcriptase (NRT) and no template control (NTC).

Target name	Sample name	Wells	Cq Mean	Cq SD
C20orf201	Flp-In T-REx-293: pcDNA5/FRT/TO+	3	30.41	0.106
β ACT	Flp-In T-REx-293: pcDNA5/FRT/TO+	3	16.93	0.077
Lamin A	Flp-In T-REx-293: pcDNA5/FRT/TO+	3	22.41	0.070
C20orf201	Flp-In T-REx-293: pcDNA5/FRT/TO-	3	30.59	0.103
β ACT	Flp-In T-REx-293: pcDNA5/FRT/TO-	3	17.66	0.019
Lamin A	Flp-In T-REx-293: pcDNA5/FRT/TO-	3	22.94	0.143
C20orf201	Flp-In T-REx-293: *C20orf201* F1+	3	19.97	0.230
β ACT	Flp-In T-REx-293: *C20orf201* F1+	3	17.40	0.074
Lamin A	Flp-In T-REx-293: *C20orf201* F1+	3	22.90	0.104
C20orf201	Flp-In T-REx-293: *C20orf201* F1-	3	26.09	0.205
β ACT	Flp-In T-REx-293: *C20orf201* F1-	3	17.33	0.047
Lamin A	Flp-In T-REx-293: *C20orf201* F1-	3	23.00	0.054
C20orf201	Flp-In T-REx-293: *C20orf201* F2+	3	19.38	0.051
β ACT	Flp-In T-REx-293: *C20orf201* F2+	3	16.69	0.066
Lamin A	Flp-In T-REx-293: *C20orf201* F2+	3	22.39	0.089
C20orf201	Flp-In T-REx-293: *C20orf201* F2-	3	24.99	0.060
β ACT	Flp-In T-REx-293: *C20orf201* F2-	3	16.89	0.060
Lamin A	Flp-In T-REx-293: *C20orf201* F2-	3	22.51	0.102
C20orf201	Flp-In T-REx-293: *C20orf201* F3+	3	18.49	0.184
β ACT	Flp-In T-REx-293: *C20orf201* F3+	3	17.34	0.136
Lamin A	Flp-In T-REx-293: *C20orf201* F3+	3	22.75	0.068
C20orf201	Flp-In T-REx-293: *C20orf201* F3-	3	23.68	0.119
β ACT	Flp-In T-REx-293: *C20orf201* F3-	3	16.87	0.055
Lamin A	Flp-In T-REx-293: *C20orf201* F3-	3	22.34	0.047
C20orf201	Flp-In T-REx-293: pcDNA5/FRT/TO+ (NRT)	1	0.00	0.000
β ACT	Flp-In T-REx-293: pcDNA5/FRT/TO+ (NRT)	1	0.00	0.000
Lamin A	Flp-In T-REx-293: pcDNA5/FRT/TO+ (NRT)	1	0.00	0.000
C20orf201	Flp-In T-REx-293: pcDNA5/FRT/TO- (NRT)	1	0.00	0.000
β ACT	Flp-In T-REx-293: pcDNA5/FRT/TO- (NRT)	1	0.00	0.000
Lamin A	Flp-In T-REx-293: pcDNA5/FRT/TO- (NRT)	1	0.00	0.000
C20orf201	Flp-In T-REx-293: *C20orf201* F1+ (NRT)	1	0.00	0.000
β ACT	Flp-In T-REx-293: *C20orf201* F1+ (NRT)	1	0.00	0.000
Lamin A	Flp-In T-REx-293: *C20orf201* F1+ (NRT)	1	0.00	0.000
C20orf201	Flp-In T-REx-293: *C20orf201* F1- (NRT)	1	0.00	0.000
β ACT	Flp-In T-REx-293: *C20orf201* F1- (NRT)	1	0.00	0.000
Lamin A	Flp-In T-REx-293: *C20orf201* F1- (NRT)	1	0.00	0.000
C20orf201	Flp-In T-REx-293: *C20orf201* F2+ (NRT)	1	0.00	0.000
β ACT	Flp-In T-REx-293: *C20orf201* F2+ (NRT)	1	0.00	0.000
Lamin A	Flp-In T-REx-293: *C20orf201* F2+ (NRT)	1	0.00	0.000
C20orf201	Flp-In T-REx-293: *C20orf201* F2- (NRT)	1	0.00	0.000
β ACT	Flp-In T-REx-293: *C20orf201* F2- (NRT)	1	0.00	0.000
Lamin A	Flp-In T-REx-293: *C20orf201* F2- (NRT)	1	0.00	0.000
C20orf201	Flp-In T-REx-293: *C20orf201* F3+ (NRT)	1	0.00	0.000
β ACT	Flp-In T-REx-293: *C20orf201* F3+ (NRT)	1	0.00	0.000
Lamin A	Flp-In T-REx-293: *C20orf201* F3+ (NRT)	1	0.00	0.000
C20orf201	Flp-In T-REx-293: *C20orf201* F3- (NRT)	1	0.00	0.000
β ACT	Flp-In T-REx-293: *C20orf201* F3- (NRT)	1	0.00	0.000
Lamin A	Flp-In T-REx-293: *C20orf201* F3- (NRT)	1	0.00	0.000
C20orf201	NTC	1	0.00	0.000
β ACT	NTC	1	0.00	0.000
Lamin A	NTC	1	0.00	0.000

Table 6.3 Real-time qRT-PCR analysis of distinct fragments of *C20orf201* in Flp-In T-Rex-293 cells. The table displays the mean of three repeats of the quantification cycle (Cq) and standard deviation (SD) for four fragments (4-7) of *C20orf201* before (-) and after (+) adding tetracycline. These results were normalised by *β ACT* and *Lamin A* expressions. The table also shows the Cq reading for the no reverse transcriptase (NRT) and no template control (NTC).

Target name	Sample name	Wells	Cq Mean	Cq SD
C20orf201	Flp-In T-REx-293: *C20orf201* F4+	3	20.83	0.250
β ACT	Flp-In T-REx-293: *C20orf201* F4+	3	17.70	0.044
Lamin A	Flp-In T-REx-293: *C20orf201* F4+	3	23.23	0.107
C20orf201	Flp-In T-REx-293: *C20orf201* F4-	3	25.91	0.097
β ACT	Flp-In T-REx-293: *C20orf201* F4-	3	17.64	0.048
Lamin A	Flp-In T-REx-293: *C20orf201* F4-	3	23.21	0.092
C20orf201	Flp-In T-REx-293: *C20orf201* F5+	3	20.41	0.088
β ACT	Flp-In T-REx-293: *C20orf201* F5+	3	17.89	0.096
Lamin A	Flp-In T-REx-293: *C20orf201* F5+	3	23.20	0.041
C20orf201	Flp-In T-REx-293: *C20orf201* F5-	3	26.00	0.103
β ACT	Flp-In T-REx-293: *C20orf201* F5-	3	17.72	0.041
Lamin A	Flp-In T-REx-293: *C20orf201* F5-	3	23.15	0.066
C20orf201	Flp-In T-REx-293: *C20orf201* F6+	3	21.14	0.152
β ACT	Flp-In T-REx-293: *C20orf201* F6+	3	18.74	0.082
Lamin A	Flp-In T-REx-293: *C20orf201* F6+	3	24.13	0.067
C20orf201	Flp-In T-REx-293: *C20orf201* F6-	3	25.09	0.084
β ACT	Flp-In T-REx-293: *C20orf201* F6-	3	18.01	0.072
Lamin A	Flp-In T-REx-293: *C20orf201* F6-	3	23.50	0.070
C20orf201	Flp-In T-REx-293: *C20orf201* F7+	3	20.59	0.079
β ACT	Flp-In T-REx-293: *C20orf201* F7+	3	17.45	0.123
Lamin A	Flp-In T-REx-293: *C20orf201* F7+	3	23.19	0.028
C20orf201	Flp-In T-REx-293: *C20orf201* F7-	3	25.06	0.039
β ACT	Flp-In T-REx-293: *C20orf201* F7-	3	17.26	0.054
Lamin A	Flp-In T-REx-293: *C20orf201* F7-	3	22.84	0.128
C20orf201	Flp-In T-REx-293: *C20orf201* F4+ (NRT)	1	0.00	0.000
β ACT	Flp-In T-REx-293: *C20orf201* F4+ (NRT)	1	0.00	0.000
Lamin A	Flp-In T-REx-293: *C20orf201* F4+ (NRT)	1	0.00	0.000
C20orf201	Flp-In T-REx-293: *C20orf201* F4- (NRT)	1	0.00	0.000
β ACT	Flp-In T-REx-293: *C20orf201* F4- (NRT)	1	0.00	0.000
Lamin A	Flp-In T-REx-293: *C20orf201* F4- (NRT)	1	0.00	0.000
C20orf201	Flp-In T-REx-293: *C20orf201* F5+ (NRT)	1	0.00	0.000
β ACT	Flp-In T-REx-293: *C20orf201* F5+ (NRT)	1	0.00	0.000
Lamin A	Flp-In T-REx-293: *C20orf201* F5+ (NRT)	1	0.00	0.000
C20orf201	Flp-In T-REx-293: *C20orf201* F5- (NRT)	1	0.00	0.000
β ACT	Flp-In T-REx-293: *C20orf201* F5- (NRT)	1	0.00	0.000
Lamin A	Flp-In T-REx-293: *C20orf201* F5- (NRT)	1	0.00	0.000
C20orf201	Flp-In T-REx-293: *C20orf201* F6+ (NRT)	1	0.00	0.000
β ACT	Flp-In T-REx-293: *C20orf201* F6+ (NRT)	1	0.00	0.000
Lamin A	Flp-In T-REx-293: *C20orf201* F6+ (NRT)	1	0.00	0.000
C20orf201	Flp-In T-REx-293: *C20orf201* F6- (NRT)	1	0.00	0.000
β ACT	Flp-In T-REx-293: *C20orf201* F6- (NRT)	1	0.00	0.000
Lamin A	Flp-In T-REx-293: *C20orf201* F6- (NRT)	1	0.00	0.000
C20orf201	Flp-In T-REx-293: *C20orf201* F7+ (NRT)	1	0.00	0.000
β ACT	Flp-In T-REx-293: *C20orf201* F7+ (NRT)	1	0.00	0.000
Lamin A	Flp-In T-REx-293: *C20orf201* F7+ (NRT)	1	0.00	0.000
C20orf201	Flp-In T-REx-293: *C20orf201* F7- (NRT)	1	0.00	0.000
β ACT	Flp-In T-REx-293: *C20orf201* F7- (NRT)	1	0.00	0.000
Lamin A	Flp-In T-REx-293: *C20orf201* F7- (NRT)	1	0.00	0.000
C20orf201	NTC	1	0.00	0.000
β ACT	NTC	1	0.00	0.000
Lamin A	NTC	1	0.00	0.000

6.2.4.5 Investigation of the presence of C20orf201 protein in the human Flp-In T-REx-293 cell lines

A monoclonal antibody against the 6X His tag (Abcam; AB18184) was used to study the protein levels of C20orf201 for distinct integrants of Flp-In T-REx-293 cell lines containing different fragments of *C20orf201* cDNA. The protein for each *C20orf201* fragment was evaluated in cultured cells in the presence and absence of tetracycline treatment. Western blots indicated that C20orf201 protein fragments were clearly observed in the Flp-In T-REx-293 cell line inserted with only *C20orf201* fragments 1, 2 or 3 when the membrane was exposed for 10 minutes using the X-ray film and Chemiluminescent Peroxidase Substrate-3 (Sigma-Aldrich; CPS3100-1KT) (Figure 6.26 upper film); whereas no protein was detected for fragment 3 when the membrane was exposed for 2 minutes (Figure 6.26 lower film). These proteins for the fragments 1 to 3 were detected after treating the cells with tetracycline. Appropriate protein loading was confirmed for each clone by the GAPDH control. The molecular weights for each fragment of C20orf201 gene are shown in Table 6.4.

Table 6.4 Calculation of the molecular weights (MW) of each C20orf201 fragment

Fragment	Size (bp)	Number of amino acids	C20orf201 MW (KDa)	Histidine MW (KDa)	Expected MW (KDa)
1	720	720/3= 240	240×110/1000= 26.4	0.66	27.06
2	621	621/3= 207	207×110/1000= 22.77	0.66	23.43
3	522	522/3= 174	174×110/1000= 19.14	0.66	19.8
4	423	423/3= 141	141×110/1000= 15.51	0.66	16.17
5	324	324/3= 108	108×110/1000= 11.88	0.66	12.54
6	225	225/3= 75	225×110/1000= 8.25	0.66	8.91
7	126	126/3= 42	126×110/1000= 4.62	0.66	5.28

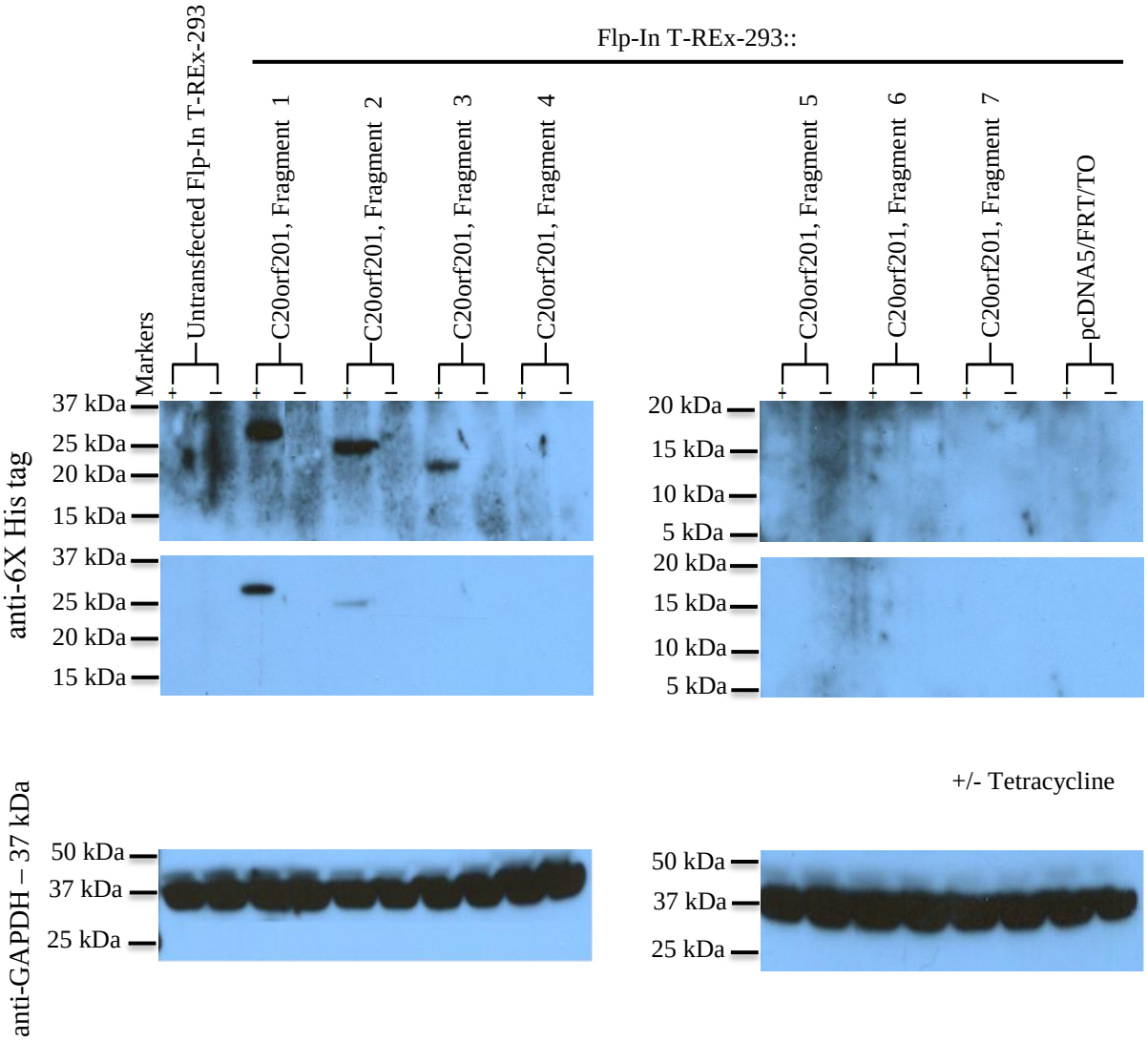


Figure 6.26 Western blot analysis showing different fragments of C20orf201 induced by tetracycline in the Flp-In T-REx-293 system. One group of 7 cell lines with distinct fragments of the *C20orf201* gene inserted into the Flp-In T-REx-293 cells were subjected to 2 µg/ml of tetracycline and another group of 7 were not. Positive signals are seen in cells containing the *C20orf201* fragment 1, 2 and 3 after treatment with tetracycline. The cells were harvested 24 hours after induction. The lysates were separated on a 10% SDS-PAGE gel. GAPDH (Glyceraldehyde 3-phosphate dehydrogenase**)** was used as the loading control (bottom). The upper membrane for anti-6X His tag was exposed for 10 minutes, while the lower membrane was exposed for 2 minutes.

6.3 Discussion

We postulate that a translation inhibition mechanism was operating with similarities to that in the fission yeast for the inhibition of translation of meiotic gene transcripts during the mitotic cell cycle (Harigaya *et al*., 2006; Yamamoto, 2010). In the fission yeast pathway, a small sequence within the mRNA known as the DSR serves as a target site for the inhibitory proteins. We postulated that if such a target site mechanism were working in humans, then deletion of the region encompassing this site would result in translation of the mRNA. To test this deletion constructs of the *C20orf201* gene were made which encoded an epitope tag which would enable us to employ a monoclonal antibody specifically to the products of these constructs. It was hoped that these staggered deletions would enable us to identify a DSR-like region within the human *C20orf201* gene.

RT-PCR analysis showed strong bands in all *C20orf201* fragments either in the presence and absence of tetracycline treatment (refer to Figure 6.22). However, the protein levels were detectable by Western blots only for fragments 1, 2 and 3. No C20orf201 protein fragment was detected in Flp-In T-REx-293 cells for the other fragments. We thought that the absence of the proteins for fragments 4-7 might be due to insufficient exposure time for detecting the protein bands. Therefore, the membrane was also exposed for 30 minutes, but the results were still the same: no proteins were detected for fragments 4-7 and a background appeared on the membrane. We also noticed that the apparent protein levels for fragments 1, 2 and 3 decreased gradually and we thought this was because the small molecular weight protein could have failed to transfer to the membrane fully. Therefore, we used a specific membrane for small molecular weight proteins, but the Western blot results using this membrane remained the same.

Chapter 5 indicated the possible presence of a mechanism that permits transcription but inhibits the translation of the *C20orf201* gene in cell lines that were not treated with tetracycline. The fact that the addition of tetracycline allows the translation of *C20orf201* mRNA in these cells might suggest a translational inhibitory mechanism that could lead to mRNA degradation, but it can only cope with low levels of a specific RNA. Such pathway could be related to the Mmi1 pathway in *S. pombe*. This mechanism involves the Mmi1 protein, which is a YTH family protein; members of this family of proteins are also found in humans. The question here is whether this protein operates in humans. According to Western blot results, this objective was not achieved. Therefore, we can do another experiment, such as a knockdown of the human YTH genes, and ask if that knockdown could change the

translation profile of *C20orf201* in cells. The cells translating the *C20orf201* protein could confirm that the mechanism is similar to the Mmi1 pathway. However, if the mRNA un-translated with no protein production, that may suggest that another mechanism must be found. The finding that a possible ubiquitinated form of C20orf201 resides at 50 KDa confuses the issue, as this species is present with and without tetracycline, although the levels do not change.

6.4 Conclusion

C20orf201 ORF was divided into seven fragments to study the mRNA expression of each fragment after insertion into the whole genome of the Flp-In T-REx-293 cells. The RT-PCR results showed strong expression of each fragment; the investigation into the translation of these mRNAs using an antibody against the 6X His tag showed no protein levels for fragments 4 to 7. Further analysis is required to address whether a translation inhibition mechanism is in operation.

Chapter 7.0: General Discussion

7.1 Screening of meiotic genes in a wide range of normal and cancer tissues

Among numerous tumour antigens, CTAs have been selected as potential targets for promising immunotherapy including cancer vaccination, because they are present in various malignant tumours, but are not present in normal adult tissues other than testis, ovary and placenta (Whitehurst, 2014). Consequently, we suggested that meiotic genes may provide an excellent source for the identification of new CTA genes. CTA gene expression profiling can also be used as a novel tool for cancer biomarkers for the early diagnosis of cancer (for example, see Rousseaux *et al*., 2013). The CTAs are also promising candidates as potential drug targets (for example, see Whitehurst *et al*., 2007). Therefore, identification of new CTA genes should improve and increase therapeutic strategies.

In the current study, genes identified by a bioinformatics pipeline were used to identify a number of novel potential CTA genes. These genes were then tested using RT-PCR analysis using a range of normal tissues and the promising candidates were additionally explored in distinct types of cancer tissues and cell lines. The RT-PCR results presented in Chapter 3.0 identified 11 potential CTA genes: 6 CT-restricted/selective genes and 5 CT-CNS-restricted/selective genes. Four genes from the 32 genes screened here were also found expressed only in the normal testis, but not in other normal or cancer tissues and cell lines. However, these testis-specific genes might be expressed in other cancers not included in the present study. For example, a meta-analysis of Microarray data was performed for the four testis-specific genes (*UBL4B*, *CST8*, *ODF3* and *TMEM225*) and one gene (*UBL4B*) was found to display expression in an ovarian cancer sample (www.cancerma.org.uk; Feichtinger *et al*., 2012b). Conversely, the results from RT-PCR analysis for *UBL4B* did not show any expression in the ovarian tumour tissue or in the three different human ovarian cancer cell lines (A2780, PEO14 and TO14) that were screened in the present study.

One unexpected finding in this study was that 17 of the 32 genes screened showed expression in most normal somatic tissues, although two of these were thought to be meiosis-specific genes. This expression in normal tissues might be explained in two different ways. First, these genes have characterised functions during meiotic division, but might have other functions in mitotic cells. Next, some of the normal tissues might have had undiagnosed neoplasia

(abnormal growth), which could ultimately lead to tumour progression. For instance, a previous study for identification of CTA genes showed that differences in gene expression were found in a distinct panel of normal human tissues (Chen *et al*., 2005a), inferring some contain neoplastic cells.

Importantly, 11 genes were identified in the present project as being novel CT gene candidates because they displayed expression that was either restricted to the normal testis or was found in the normal testis and fewer than two other somatic normal tissues. However, these candidate genes displayed expression in many types of tumours. A meta-analysis of Microarray was performed for the 11 promising CT genes and the results were upregulated in five genes (*C20orf201, FSCN3, GAGE1, SSX2* and *SSX5*). The *SSX2* gene is a good example of the candidate CTA genes identified by RT-PCR. This gene was expressed in COLO800 (melanoma cancer) and K-562 (leukaemia cancer). However, the meta-analysis indicates that this gene was up-regulated in four different types of cancer, such as ovarian, lung, brain and adrenal cancers. Therefore, this study suggests that further investigation of the *SSX2* gene is required using different cancer samples.

The *SSX* gene family of cancer-testis antigens consists of ten genes (*SSX1-10*) (Güre *et al*., 2002; Smith *et al*., 2011). The presence of the SSX family of cancer-testis antigens has been correlated with more advanced stages of disease and worse patient prognosis in a number of human cancers (Güre *et al*., 2002). *SSX* expression has been investigated in a wide range of cancers (Güre *et al*., 2005), while reports on protein levels are limited such as, malignant melanoma (dos Santos *et al*., 2000), prostate cancer (Smith *et al*., 2011), hepatocellular carcinoma (Liang *et al*., 2013), and oesophageal cancer (Akcakanat *et al*., 2006). The SSX proteins are a highly homologous group of CT-X antigens with displayed immunogenicity in patients diagnosed with cancer. They have a number of features that make them attractive targets for cancer therapy. For example, a recent study reported that SSX2, in particular, may be a high priority target for tumour therapy, based upon certain predefined criteria for prioritisation of tumour antigens, such as tumour specificity, oncogenicity, expression level and number of identified epitopes (Cheever *et al*., 2009).

A recent study examined the expression of *SSX1, SSX2, SSX4* and *MAGE-C1/CT7* in multiple myeloma (MM) cell lines using RT-PCR analysis. This study also validated these CTA genes in bone marrow (BM) cells isolated from 25 MM patients and 18 healthy volunteers. The RT-PCR results showed that four CTA genes were expressed in MM cell lines and the expression

rate of each CTA gene, *SSX1*, *SSX2*, *SSX4* and *MAGE-C1/CT7*, in the BM cells of 25 MM patients was 80% (20/25), 40% (10/25), 68% (17/25) and 28% (7/25), respectively. However, BM cells of 18 healthy volunteers did not show any expression of any member of the CTA genes. The conclusion drawn from this work was that these four CTA genes were co-expressed in MM cell lines and may serve as useful targets for cancer immunotherapy in MM in the future (He *et al*., 2014).

The SSX2 gene in the present study and other SSX family members reported in previous studies can be used as important indicators for early diagnosis and prognostic determination of different cancer types.

7.2 Functional analysis of the human *STRA8* gene

The RT-PCR analysis of the human *STRA8* gene using cDNAs from normal adult testis, HEP-G2, lung tumour, breast tumours and uterus tumour gave a band 216 bp larger than expected. The sequence of this was therefore confirmed initially using PCR product derived from testis cDNA and *STRA8* internal primers F1 and R1. Subsequent RT-PCR analysis of the *STRA8* gene using different sets of forward and reverse primers and different cDNA templates confirmed the observation that the whole ORF of *STRA8* identified by RT-PCR analysis is 1209 bp and not 993bp as recorded in GenBank (AF513502).

The majority of the 'additional' base pairs potentially translate into glutamic acid residues. If the smaller cDNA GenBank (AF513502) sequences is expressed as protein the molecular weight of STRA8 should be at 37 kDa. Western blot results from HEP-G2 cell extracts detected a high level of STRA8 protein at a molecular weight that would include the additional glutamic acid residue sequence. We did not observe the predicted 37kDa protein. This data provides strong evidence that the 'larger' sequence identified via RT-PCR is translated into protein and could possibly be a splice variant of STRA8 with a specific function in cancer cells. Further study is required to confirm these extra sequences in other cell lines using RT-PCR and Western blot analyses.

In the present study, the *STRA8* gene was identified as a meiosis specific gene since its expression was observed only in the normal adult testis and no other normal human tissues. However, four different cancer cell lines expressed *STRA8*; therefore, this gene was identified as a good CTA candidate gene. Although the function of the human STRA8 is not known, a previous study revealed that Stra8 is required for meiotic progression in mice in both sexes (Anderson *et al*., 2008). Retinoic acid (RA) plays fundamental roles in both spermatogenesis

and meiotic entry (Raverdeau *et al*., 2012) and the *Stra8* gene is up-regulated by RA in germ cells during both adult and foetal meiosis (Feng *et al*., 2014).

In addition, in the present study, Flp-In T-REx-293 cell lines were inserted with the *STRA8* ORF with and without the addition of the Kozak sequences in order to study the function of this gene in the integrated cell lines. Overexpression of *GAGE1, MAGE-A1, NUT, PRDM9, RAD21L, REC8, SMC1β, STAG3, SYCP1, SYCP2, SYCP3* and *TEX19* meiosis genes, chosen randomly, was not detected in Flp-In T-REx-293 cell lines by RT-PCR analysis. It is possible that the RT-PCR analysis used in this study perhaps was not sufficient for studying whether *STRA8* can promote overexpression of any meiosis-specific gene. Therefore, we thought that qRT-PCR analysis is also very important to validate the mRNA expression of the 12 meiosis genes in the cell lines integrated with the *STRA8* gene; however, the qRT-PCR analysis was not performed in this study due to time constraints.

We also considered that treating cell lines integrated with *STRA8* with retinoic acid could induce meiosis selected genes; however, no overexpression was observed for any of the genes tested here. This may suggest that *STRA8* could possibly induce other genes that were not examined here or that expression of *STRA8* in these cell lines was not sufficient for meiosis progression. A previous study indicated that Depleted in azoospermia-like (DAZL), an RNA-binding protein and key intrinsic factor, enables spermatogonia in mouse to express *Stra8* and initiate differentiation and meiosis in response to RA (Lin *et al*., 2008). The correlation between DAZL and STRA8 can be investigated by comparing RT-PCR, qRT-PCR and Western blot analyses in the cell lines that had been integrated with *STRA8* ORF to these cell lines without the *STRA8* ORF or in cells integrated with the vector alone (negative control).

Cytochrome P450, family 26, subfamily B, polypeptide 1 (CYP26B1) is one of the RA-degrading enzymes and is expressed in both foetal and adult testis (Kasimanickam and Kasimanickam, 2013). CYP26B1 prevents the RA signalling and shields the RA-induced entry of meiosis in testis; in contrast, the absence of CYP26B1 leads to the RA-induced entry of meiosis in ovary and the induction of STRA8 (Bowles *et al*., 2006). The cell lines treated with retinoic acid did not show overexpression of the meiotic genes tested here. It is possible that this could be due to activity of the CYP26B1 gene, which led to RA degradation. Treating these cell lines with a CYP26B1 inhibitor could test this hypothesis.

7.3 Functional analysis of the human *C20orf201* gene

The RT-PCR results presented here showed that *C20orf201* gene expression in normal human tissues was restricted to the testis and CNS tissues for splice variant 1 or to the testis, CNS tissues and trachea for splice variant 2. This observation for the *C20orf201* variant 1 was confirmed in a recent study by Kamata *et al*., (2013), who indicated that the *C20orf201* mRNA was expressed in the same normal tissues as described here. However, the expression of this gene was found in various types of cancer.

The results obtained by RT-PCR for *C20orf201* variant 1 were confirmed by means of Western blot analysis using normal and cancer tissue lysates, which gave the same results from RT-PCR analysis; therefore, this study suggests that *C20orf201* could be considered as a good candidate for a CTA gene and may be a good choice for a cancer-specific marker in the future. Indeed, subsequent work by another group member has demonstrated C20orf201 could be a good immunohistochemistry marker in ovarian cancer. This work is ongoing. The results from Western blot analysis and Immunofluorescent staining suggested that the C20orf201 protein displayed nuclear and cytoplasmic localisations of the cell line tested here (MCF-7; breast cancer). One of the most important observations in this study was the mRNA expression of *C20orf201* in MCF-7 cells. The results from RT-PCR and qRT-PCR analyses confirmed that this expression was extremely faint compared to other ovarian cell lines tested here (A2780, PEO14 and TO14); whereas, the results from Western blot analysis demonstrated that a higher level of the 26 KDa C20orf201 protein in MCF-7 cells. This observation may suggest that some mechanism could be destroying mRNA levels in these cell lines that were expressing high levels of *C20orf201* mRNA.

Moreover, in this study, the *C20orf201* ORF was inserted with and without the addition of the Kozak sequences into the whole genome of the Flp-In T-REx-293 cell lines to address the function of this gene in these cell lines. Interestingly, the results from RT-PCR analysis indicated that expression of *C20orf201* was observed in Flp-In T-REx-293 cell lines with or without tetracycline treatment. However, Western blot results revealed that the 26 KDa C20orf201 protein was found only in the presence of tetracycline. The absence of the 26 KDa protein in the cell line not treated with tetracycline may suggest that some mechanism exists that could lead to inhibition of C20orf201 translation in the absence of tetracycline. This potential inhibition of C20orf201 translation may represent a mechanism similar to the one that exists in fission yeast, termed the Mmi1 RNA elimination system. To test this possibility,

seven different fragments were amplified from the *C20orf201* ORF and then each fragment was inserted individually into the Flp-In T-REx-293 cell line genome to determine which of these fragments could be recognised by the human YTH family (members of the Mmi1 system). All fragments were expressed in these cell lines with and without tetracycline treatment; however, only three fragments showed C20orf201 protein in the presence of tetracycline (1-3). Fragment 3 was expressed more highly than other two fragments, but the protein levels were too low compared to fragments 1 and 2. A brief reiteration of the 50 KDa band and the fact it could be a Ub species.

In conclusion, RT-PCR screening of meiotic genes, using a wide range of normal and cancer tissues and cell lines, identified 11 novel CT genes, and one of the four validated testis-specific genes was up-regulated in ovarian cancer, which may suggest further study of this gene in more ovarian cancer samples would be worthwhile. These results therefore confirm our hypothesis that meiotic genes may provide an excellent source for identifying potential novel CTA genes. Further studies are also required to establish the protein products of the good CTA genes identified in the current study in cancer cells. These proteins, if any, may serve as useful tumour-specific targets.

References

ABRAMOVITZ, M. and LEYLAND-JONES, B., 2006. A systems approach to clinical oncology: focus on breast cancer. *Proteome Science,* **4**(1), pp. 5-20.

ADAIR, S.J. and HOGAN, K.T., 2009. Treatment of ovarian cancer cell lines with 5-aza-2′-deoxycytidine upregulates the expression of cancer-testis antigens and class I major histocompatibility complex-encoded molecules. *Cancer Immunology, Immunotherapy,* **58**(4), pp. 589-601.

AKCAKANAT, A., KANDA, T., TANABE, T., KOMUKAI, S., YAJIMA, K., NAKAGAWA, S., OHASHI, M. and HATAKEYAMA, K., 2006. Heterogeneous expression of GAGE, NY-ESO-1, MAGE-A and SSX proteins in esophageal cancer: Implications for immunotherapy. *International Journal of Cancer,* **118**(1), pp. 123-128.

AKERS, S.N., ODUNSI, K. and KARPF, A.R., 2010. Regulation of cancer germline antigen gene expression: implications for cancer immunotherapy. *Future Oncology,* **6**(5), pp. 717-732.

AKHAVAN-NIAKI, H. and SAMADANI, A.A., 2013. DNA methylation and cancer development: molecular mechanism. *Cell Biochemistry and Biophysics,* **67**(2), pp. 501-513.

ALBERTSON, D.G., COLLINS, C., MCCORMICK, F. and GRAY, J.W., 2003. Chromosome aberrations in solid tumors. *Nature Genetics,* **34**(4), pp. 369-376.

ALMEIDA, L.G., SAKABE, N.J., SILVA, M.C.C., MUNDSTEIN, A.S., COHEN, T., CHEN, Y., CHUA, R., GURUNG, S., GNJATIC, S. and JUNGBLUTH, A.A., 2009. CTdatabase: a knowledge-base of high-throughput and curated data on cancer-testis antigens. *Nucleic Acids Research,* **37**(suppl 1), pp. D816-D819.

ALMSTEDT, M., BLAGITKO-DORFS, N., DUQUE-AFONSO, J., KARBACH, J., PFEIFER, D., JÄGER, E. and LÜBBERT, M., 2010. The DNA demethylating agent 5-aza-2′-deoxycytidine induces expression of NY-ESO-1 and other cancer/testis antigens in myeloid leukemia cells. *Leukemia Research,* **34**(7), pp. 899-905.

ALY, H.A., 2012. Cancer therapy and vaccination. *Journal of Immunological Methods,* **382**(1), pp. 1-23.

ANDERSON, E.L., BALTUS, A.E., ROEPERS-GAJADIEN, H.L., HASSOLD, T.J., DE ROOIJ, D.G., VAN PELT, A.M. and PAGE, D.C., 2008. Stra8 and its inducer, retinoic acid, regulate meiotic initiation in both spermatogenesis and oogenesis in mice. *Proceedings of the National Academy of Sciences of the United States of America,* **105**(39), pp. 14976-14980.

ANDERSSON, S., DAVIS, D.L., DAHLBACK, H., JORNVALL, H. and RUSSELL, D.W., 1989. Cloning, structure, and expression of the mitochondrial cytochrome P-450 sterol 26-hydroxylase, a bile acid biosynthetic enzyme. *The Journal of Biological Chemistry,* **264**(14), pp. 8222-8229.

ANDREWS, P.W., 1984. Retinoic acid induces neuronal differentiation of a cloned human embryonal carcinoma cell line *in vitro*. *Developmental Biology,* **103**(2), pp. 285-293.

ASHKTORAB, H., BELGRAVE, K., HOSSEINKHAH, F., BRIM, H., NOURAIE, M., TAKKIKTO, M., HEWITT, S., LEE, E.L., DASHWOOD, R. and SMOOT, D., 2009. Global histone H4 acetylation

and HDAC2 expression in colon adenoma and carcinoma. *Digestive Diseases and Sciences,* **54**(10), pp. 2109-2117.

ATANACKOVIC, D., ARFSTEN, J., CAO, Y., GNJATIC, S., SCHNIEDERS, F., BARTELS, K., SCHILLING, G., FALTZ, C., WOLSCHKE, C., DIERLAMM, J., RITTER, G., EIERMANN, T., HOSSFELD, D.K., ZANDER, A.R., JUNGBLUTH, A.A., OLD, L.J., BOKEMEYER, C. and KROGER, N., 2007. Cancer-testis antigens are commonly expressed in multiple myeloma and induce systemic immunity following allogeneic stem cell transplantation. *Blood,* **109**(3), pp. 1103-1112.

AUER, R.L., STARCZYNSKI, J., MCELWAINE, S., BERTONI, F., NEWLAND, A.C., FEGAN, C.D. and COTTER, F.E., 2005. Identification of a potential role for POU2AF1 and BTG4 in the deletion of 11q23 in chronic lymphocytic leukemia. *Genes, Chromosomes and Cancer,* **43**(1), pp. 1-10.

BAI, G. and ZHANG, H., 2010. Promoter demethylation mediates the expression of ZNF645, a novel cancer/testis gene. *Biochemistry and Molecular Biology Reports,* **43**(6), pp. 400-406.

BALTUS, A.E., MENKE, D.B., HU, Y., GOODHEART, M.L., CARPENTER, A.E., DE ROOIJ, D.G. and PAGE, D.C., 2006. In germ cells of mouse embryonic ovaries, the decision to enter meiosis precedes premeiotic DNA replication. *Nature Genetics,* **38**(12), pp. 1430-1434.

BANNISTER, A.J. and KOUZARIDES, T., 2011. Regulation of chromatin by histone modifications. *Cell Research,* **21**(3), pp. 381-395.

BARBER, T.D., MCMANUS, K., YUEN, K.W., REIS, M., PARMIGIANI, G., SHEN, D., BARRETT, I., NOUHI, Y., SPENCER, F., MARKOWITZ, S., VELCULESCU, V.E., KINZLER, K.W., VOGELSTEIN, B., LENGAUER, C. and HIETER, P., 2008. Chromatid cohesion defects may underlie chromosome instability in human colorectal cancers. *Proceedings of the National Academy of Sciences of the United States of America,* **105**(9), pp. 3443-3448.

BARNEDA-ZAHONERO, B. and PARRA, M., 2012. Histone deacetylases and cancer. *Molecular Oncology,* **6**(2012), pp. 579-589.

BAUDAT, F., IMAI, Y. and DE MASSY, B., 2013. Meiotic recombination in mammals: localization and regulation. *Nature Reviews Genetics,* **14**(11), pp. 794-806.

BAULANDE, S., CRIQUI, A. and DUTHIEUW, M., 2014. Circulating miRNAs as a new class of biomedical markers. *Medicine Sciences : M/S,* **30**(3), pp. 289-296.

BEHNEN, M., MURK, K., KURSULA, P., CAPPALLO-OBERMANN, H., ROTHKEGEL, M., KIERSZENBAUM, A.L. and KIRCHHOFF, C., 2009. Testis-expressed profilins 3 and 4 show distinct functional characteristics and localize in the acroplaxome-manchette complex in spermatids. *BMC Cell Biology,* **10**(1), pp. 34-49.

BELPOMME, D., IRIGARAY, P., HARDELL, L., CLAPP, R., MONTAGNIER, L., EPSTEIN, S. and SASCO, A.J., 2007. The multitude and diversity of environmental carcinogens. *Environmental Research,* **105**(3), pp. 414-429.

BERGERAT, A., DE MASSY, B., GADELLE, D., VAROUTAS, P., NICOLAS, A. and FORTERRE, P., 1997. An atypical topoisomerase II from Archaea with implications for meiotic recombination. *Nature,* **386**(27), pp. 414-417.

BLANCHARD, T., SRIVASTAVA, P.K. and DUAN, F., 2013. Vaccines against advanced melanoma. *Clinics in Dermatology*, **31**(2), pp. 179-190.

BOËL, P., WILDMANN, C., SENSI, M.L., BRASSEUR, R., RENAULD, J., COULIE, P., BOON, T. and VAN DER BRUGGEN, P., 1995. *BAGE*: a new gene encoding an antigen recognized on human melanomas by cytolytic T lymphocytes. *Immunity*, **2**(2), pp. 167-175.

BORDE, V., 2007. The multiple roles of the Mre11 complex for meiotic recombination. *Chromosome Research*, **15**(5), pp. 551-563.

BORGNE, A., MURAKAMI, H., AYTE, J. and NURSE, P., 2002. The G1/S cyclin Cig2p during meiosis in fission yeast. *Molecular Biology of the Cell*, **13**(6), pp. 2080-2090.

BOUILLET, P., OULAD-ABDELGHANI, M., VICAIRE, S., GARNIER, J., SCHUHBAUR, B., DOLLÉ, P. and CHAMBON, P., 1995. Efficient cloning of cDNAs of retinoic acid-responsive genes in P19 embryonal carcinoma cells and characterization of a novel mouse gene, Stra1 (mouse LERK-2/Eplg2). *Developmental Biology*, **170**(2), pp. 420-433.

BOWLES, J. and KOOPMAN, P., 2007. Retinoic acid, meiosis and germ cell fate in mammals. *Development (Cambridge, England)*, **134**(19), pp. 3401-3411.

BOWLES, J., KNIGHT, D., SMITH, C., WILHELM, D., RICHMAN, J., MAMIYA, S., YASHIRO, K., CHAWENGSAKSOPHAK, K., WILSON, M.J., ROSSANT, J., HAMADA, H. and KOOPMAN, P., 2006. Retinoid signaling determines germ cell fate in mice. *Science (New York, N.Y.)*, **312**(5773), pp. 596-600.

BOYDSTON-WHITE, S., CHERNENKO, T., REGINA, A., MILJKOVIĆ, M., MATTHÄUS, C. and DIEM, M., 2005. Microspectroscopy of single proliferating HeLa cells. *Vibrational Spectroscopy*, **38**(1), pp. 169-177.

BRÁBEK, J., MIERKE, C.T., RÖSEL, D., VESELÝ, P. and FABRY, B., 2010. The role of the tissue microenvironment in the regulation of cancer cell motility and invasion. *Cell Commun Signal*, **8**, pp. 22-29.

BRAR, G.A., KIBURZ, B.M., ZHANG, Y., KIM, J., WHITE, F. and AMON, A., 2006. Rec8 phosphorylation and recombination promote the step-wise loss of cohesins in meiosis. *Nature*, **441**(7092), pp. 532-536.

BREKHMAN, V., ITSKOVITZ-ELDOR, J., YODKO, E., DEUTSCH, M. and SELIGMAN, J., 2000. The DAZL1 gene is expressed in human male and female embryonic gonads before meiosis. *Molecular Human Reproduction*, **6**(5), pp. 465-468.

BROWN, P.W.P., JUDIS, L.L., CHAN, E.R.E., SCHWARTZ, S.S., SEFTEL, A.A., THOMAS, A.A. and HASSOLD, T.J.T., 2005. Meiotic synapsis proceeds from a limited number of subtelomeric sites in the human male. *American Journal of Human Genetics*, **77**(4), pp. 556-566.

BUCHHOLZ, F., RINGROSE, L., ANGRAND, P.O., ROSSI, F. and STEWART, A.F., 1996. Different thermostabilities of FLP and Cre recombinases: implications for applied site-specific recombination. *Nucleic Acids Research*, **24**(21), pp. 4256-4262.

CABALLERO, O.L.O. and CHEN, Y.Y., 2009. Cancer/testis (CT) antigens: potential targets for immunotherapy. *Cancer Science*, **100**(11), pp. 2014-2021.

CAMPOS-PEREZ, J., RICE, J., ESCORS, D., COLLINS, M., PATERSON, A., SAVELYEVA, N. and STEVENSON, F.K., 2013. DNA fusion vaccine designs to induce tumor-lytic CD8 T-cell attack via the immunodominant cysteine-containing epitope of NY-ESO 1. *International Journal of Cancer,* **133**(6), pp. 1400-1407.

CHEEVER, M.A., ALLISON, J.P., FERRIS, A.S., FINN, O.J., HASTINGS, B.M., HECHT, T.T., MELLMAN, I., PRINDIVILLE, S.A., VINER, J.L., WEINER, L.M. and MATRISIAN, L.M., 2009. The prioritization of cancer antigens: a national cancer institute pilot project for the acceleration of translational research. *Clinical Cancer Research: an official journal of the American Association for Cancer Research,* **15**(17), pp. 5323-5337.

CHEN, L., ZHOU, W., ZHAO, Y., LIU, X., DING, Q., ZHA, X. and WANG, S., 2012. Cancer/testis antigen SSX2 enhances invasiveness in MCF-7 cells by repressing ERα signaling. *International Journal of Oncology,* **40**(6), pp. 1986-1994.

CHEN, Y., SCANLAN, M.J., SAHIN, U., TÜRECI, Ö., GURE, A.O., TSANG, S., WILLIAMSON, B., STOCKERT, E., PFREUNDSCHUH, M. and OLD, L.J., 1997. A testicular antigen aberrantly expressed in human cancers detected by autologous antibody screening. *Proceedings of the National Academy of Sciences,* **94**(5), pp. 1914-1918.

CHEN, Y., VENDITTI, C.A., THEILER, G., STEVENSON, B.J., ISELI, C., GURE, A.O., JONGENEEL, C.V., OLD, L.J. and SIMPSON, A., 2005a. Identification of CT46/HORMAD1, an immunogenic cancer/testis antigen encoding a putative meiosis-related protein. *Cancer Immun,* **5**, pp. 9-16.

CHEN, Y.T., SCANLAN, M.J., VENDITTI, C.A., CHUA, R., THEILER, G., STEVENSON, B.J., ISELI, C., GURE, A.O., VASICEK, T., STRAUSBERG, R.L., JONGENEEL, C.V., OLD, L.J. and SIMPSON, A.J., 2005b. Identification of cancer/testis-antigen genes by massively parallel signature sequencing. *Proceedings of the National Academy of Sciences of the United States of America,* **102**(22), pp. 7940-7945.

CHENG, Y., WONG, E.W. and CHENG, C.Y., 2011. Cancer/testis (CT) antigens, carcinogenesis and spermatogenesis. *Spermatogenesis,* **1**(3), pp. 209-220.

CHILDS, A.J., COWAN, G., KINNELL, H.L., ANDERSON, R.A. and SAUNDERS, P.T., 2011. Retinoic acid signalling and the control of meiotic entry in the human fetal gonad. *PloS One,* **6**(6), pp. e20249.

CHOI, J.D. and LEE, J., 2013. Interplay between Epigenetics and Genetics in Cancer. *Genomics & Informatics,* **11**(4), pp. 164-173.

CHOI, Y., YOON, J., PYO, C., KIM, J., BAE, S. and PARK, S., 2010. A possible role of STRA8 as a transcriptional factor. *Genes & Genomics,* **32**(6), pp. 521-526.

CLAGETT-DAME, M. and KNUTSON, D., 2011. Vitamin A in reproduction and development. *Nutrients,* **3**(4), pp. 385-428.

CLAUSELL, J., HAPPEL, N., HALE, T.K., DOENECKE, D. and BEATO, M., 2009. Histone H1 subtypes differentially modulate chromatin condensation without preventing ATP-dependent remodeling by SWI/SNF or NURF. *PLoS One,* **4**(10), pp. e0007243.

CORNWALL, G., HSIA, N. and SUTTON, H., 1999. Structure, alternative splicing and chromosomal localization of the cystatin-related epididymal spermatogenic gene. *Biochem.J,* **340**, pp. 85-93.

COSTA, F.F., LE BLANC, K. and BRODIN, B., 2007. Concise review: cancer/testis antigens, stem cells, and cancer. *Stem Cells*, **25**(3), pp. 707-711.

CROMIE, G.A. and SMITH, G.R., 2007. Branching out: meiotic recombination and its regulation. *Trends in Cell Biology*, **17**(9), pp. 448-455.

DANIEL, K., LANGE, J., HACHED, K., FU, J., ANASTASSIADIS, K., ROIG, I., COOKE, H.J., STEWART, A.F., WASSMANN, K. and JASIN, M., 2011. Meiotic homologue alignment and its quality surveillance are controlled by mouse HORMAD1. *Nature Cell Biology*, **13**(5), pp. 599-610.

D'ARCY, V., ABDULLAEV, Z.K., PORE, N., DOCQUIER, F., TORRANO, V., CHERNUKHIN, I., SMART, M., FARRAR, D., METODIEV, M., FERNANDEZ, N., RICHARD, C., DELGADO, M.D., LOBANENKOV, V. and KLENOVA, E., 2006. The potential of BORIS detected in the leukocytes of breast cancer patients as an early marker of tumorigenesis. *Clinical Cancer Research: an Official Journal of the American Association for Cancer Research*, **12**(20 Pt 1), pp. 5978-5986.

DAVIES, O.R., MAMAN, J.D. and PELLEGRINI, L., 2012. Structural analysis of the human SYCE2-TEX12 complex provides molecular insights into synaptonemal complex assembly. *Open Biology*, **2**(7), pp. 120099.

DE BRUIJN, D.R., DOS SANTOS, N.R., KATER-BAATS, E., THIJSSEN, J., VAN DEN BERK, L., STAP, J., BALEMANS, M., SCHEPENS, M., MERKX, G. and GEURTS VAN KESSEL, A., 2002. The cancer-related protein SSX2 interacts with the human homologue of a Ras-like GTPase interactor, RAB3IP, and a novel nuclear protein, SSX2IP. *Genes, Chromosomes and Cancer*, **34**(3), pp. 285-298.

DE CARVALHO, F., COSTA, E.T., CAMARGO, A.A., GREGORIO, J.C., MASOTTI, C., ANDRADE, V.C., STRAUSS, B.E., CABALLERO, O.L., ATANACKOVIC, D. and COLLEONI, G.W., 2011. Targeting MAGE-C1/CT7 expression increases cell sensitivity to the proteasome inhibitor bortezomib in multiple myeloma cell lines. *PloS One*, **6**(11), pp. e27707.

DE CARVALHO, F., VETTORE, A.L. and COLLEONI, G.W., 2012. Cancer/Testis Antigen MAGE-C1/CT7: new target for multiple myeloma therapy. *Journal of Immunology Research*, **2012**, pp. 257695.

DE MASSY, B., 2003. Distribution of meiotic recombination sites. *Trends in Genetics*, **19**(9), pp. 514-522.

DE MASSY, B., 2013. Initiation of Meiotic Recombination: How and Where? Conservation and Specificities Among Eukaryotes. *Annual Review of Genetics*, **47**, pp. 563-599.

DE SMET, C. and LORIOT, A., 2010. DNA hypomethylation in cancer. *Epigenetics*, **5**(3), pp. 206-213.

DE SMET, C., LURQUIN, C., LETHE, B., MARTELANGE, V. and BOON, T., 1999. DNA methylation is the primary silencing mechanism for a set of germ line- and tumor-specific genes with a CpG-rich promoter. *Molecular and Cellular Biology*, **19**(11), pp. 7327-7335.

DIANATPOUR, M., MEHDIPOUR, P., NAYERNIA, K., MOBASHERI, M., GHAFOURI-FARD, S., SAVAD, S. and MODARRESSI, M.H., 2012. Expression of testis specific genes TSGA10, TEX101 and ODF3 in breast cancer. *Iranian Red Crescent Medical Journal*, **14**(11), pp. 722-726.

DING, D.Q., SAKURAI, N., KATOU, Y., ITOH, T., SHIRAHIGE, K., HARAGUCHI, T. and HIRAOKA, Y., 2006. Meiotic cohesins modulate chromosome compaction during meiotic prophase in fission yeast. *The Journal of Cell Biology,* **174**(4), pp. 499-508.

DOS SANTOS, N.R., TORENSMA, R., DE VRIES, T.J., SCHREURS, M.W., DE BRUIJN, D.R., KATER-BAATS, E., RUITER, D.J., ADEMA, G.J., VAN MUIJEN, G.N. and VAN KESSEL, A.G., 2000. Heterogeneous expression of the SSX cancer/testis antigens in human melanoma lesions and cell lines. *Cancer Research,* **60**(6), pp. 1654-1662.

DOUKI, T., 2013. The variety of UV-induced pyrimidine dimeric photoproducts in DNA as shown by chromatographic quantification methods. *Photochemical & Photobiological Sciences,* **12**(8), pp. 1286-1302.

DOYLE, T.J., BRAUN, K.W., MCLEAN, D.J., WRIGHT, R.W., GRISWOLD, M.D. and KIM, K.H., 2007. Potential functions of retinoic acid receptor A in Sertoli cells and germ cells during spermatogenesis. *Annals of the New York Academy of Sciences,* **1120**(1), pp. 114-130.

DUDAS, A., AHMAD, S. and GREGAN, J., 2011. Sgo1 is required for co-segregation of sister chromatids during achiasmate meiosis I. *Cell Cycle (Georgetown, Tex.),* **10**(6), pp. 951-955.

DUFFY, M., 2001. Clinical uses of tumor markers: a critical review. *Critical Reviews in Clinical Laboratory sciences,* **38**(3), pp. 225-262.

EICKBUSH, T.H. and MOUDRIANAKIS, E.N., 1978. The histone core complex: an octamer assembled by two sets of protein-protein interactions. *Biochemistry,* **17**(23), pp. 4955-4964.

ESTELLER, M. and HERMAN, J.G., 2002. Cancer as an epigenetic disease: DNA methylation and chromatin alterations in human tumours. *The Journal of Pathology,* **196**(1), pp. 1-7.

EVEN-DESRUMEAUX, K., BATY, D. and CHAMES, P., 2011. State of the art in tumor antigen and biomarker discovery. *Cancers,* **3**(2), pp. 2554-2596.

FEICHTINGER, J., ALDEAILEJ, I., ANDERSON, R., ALMUTAIRI, M., ALMATRAFI, A., ALSIWIEHRI, N., GRIFFITHS, K., STUART, N., WAKEMAN, J.A., LARCOMBE, L. and MCFARLANE, R.J., 2012a. Meta-analysis of clinical data using human meiotic genes identifies a novel cohort of highly restricted cancer-specific marker genes. *Oncotarget,* **3**(8), pp. 843-853.

FEICHTINGER, J., LARCOMBE, L. and MCFARLANE, R.J., 2014. Meta-analysis of expression of l (3) mbt tumor-associated germline genes supports the model that a soma-to-germline transition is a hallmark of human cancers. *International Journal of Cancer,* **134**, pp. 2359-2365.

FEICHTINGER, J., MCFARLANE, R.J. and LARCOMBE, L.D., 2012b. CancerMA: a web-based tool for automatic meta-analysis of public cancer microarray data. *Database : The Journal of Biological Databases and Curation,* **2012**, pp. bas055.

FEINBERG, A.P.A., OHLSSON, R.R. and HENIKOFF, S.S., 2006. The epigenetic progenitor origin of human cancer. *Nature Reviews Genetics,* **7**(1), pp. 21-33.

FENG, C., BOWLES, J. and KOOPMAN, P., 2014. Control of mammalian germ cell entry into meiosis. *Molecular and Cellular Endocrinology,* **382**(1), pp. 488-497.

FIJAK, M. and MEINHARDT, A., 2006. The testis in immune privilege. *Immunological Reviews,* **213**(1), pp. 66-81.

FOSTER, I., 2008. Cancer: A cell cycle defect. *Radiography,* **14**(2), pp. 144-149.

FRATTA, E., CORAL, S., COVRE, A., PARISI, G., COLIZZI, F., DANIELLI, R., MARIE NICOLAY, H.J., SIGALOTTI, L. and MAIO, M., 2011. The biology of cancer testis antigens: putative function, regulation and therapeutic potential. *Molecular Oncology,* **5**(2), pp. 164-182.

FRAUNE, J., SCHRAMM, S., ALSHEIMER, M. and BENAVENTE, R., 2012. The mammalian synaptonemal complex: protein components, assembly and role in meiotic recombination. *Experimental Cell Research,* **318**(12), pp. 1340-1346.

FREIDLIN, B. and KORN, E.L., 2014. Biomarker enrichment strategies: matching trial design to biomarker credentials. *Nature Reviews Clinical Oncology,* **11**(2), pp. 81-90.

FUNG, J.C., ROCKMILL, B., ODELL, M. and ROEDER, G.S., 2004. Imposition of Crossover Interference through the Nonrandom Distribution of Synapsis Initiation Complexes. *Cell,* **116**(6), pp. 795-802.

GE, R., CHEN, G. and HARDY, M.P., 2009. The role of the Leydig cell in spermatogenic function. *Molecular Mechanisms in Spermatogenesis.* Springer, pp. 255-269.

GERTON, J.L. and HAWLEY, R.S., 2005. Homologous chromosome interactions in meiosis: diversity amidst conservation. *Nature Reviews Genetics,* **6**(6), pp. 477-487.

GEUTJES, E., BAJPE, P. and BERNARDS, R., 2011. Targeting the epigenome for treatment of cancer. *Oncogene,* **31**(34), pp. 3827-3844.

GHAFOURI-FARD, S. and MODARRESSI, M., 2009. Cancer-testis antigens: potential targets for cancer immunotherapy. *Archives of Iranian Medicine (AIM),* **12**(4), pp. 395-404.

GHAFOURI-FARD, S. and MODARRESSI, M., 2012a. Expression of cancer-testis genes in brain tumors: implications for cancer immunotherapy. *Immunotherapy,* **4**(1), pp. 59-75.

GHAFOURI-FARD, S. and MODARRESSI, M., 2012b. Expression of splice variants of cancer-testis genes ODF3 and ODF4 in the testis of a prostate cancer patient. *Genetics and Molecular Research,* **11**(4), pp. 3642-3648.

GIBBONS, D.L., BYERS, L.A. and KURIE, J.M., 2014. Smoking, p53 Mutation, and Lung Cancer. *Molecular Cancer Research,* **12**(1), pp. 3-13.

GIRARDOT, M., FEIL, R. and LLÈRES, D., 2013. Epigenetic deregulation of genomic imprinting in humans: causal mechanisms and clinical implications. *Epigenomics,* **5**(6), pp. 715-728.

GNONI, A., LICCHETTA, A., SCARPA, A., AZZARITI, A., BRUNETTI, A.E., SIMONE, G., NARDULLI, P., SANTINI, D., AIETA, M. and DELCURATOLO, S., 2013. Carcinogenesis of Pancreatic Adenocarcinoma: Precursor Lesions. *International Journal of Molecular Sciences,* **14**(10), pp. 19731-19762.

GÓMEZ, R., VALDEOLMILLOS, A., PARRA, M.T., VIERA, A., CARREIRO, C., RONCAL, F., RUFAS, J.S., BARBERO, J.L. and SUJA, J.A., 2007. Mammalian SGO2 appears at the inner centromere domain and redistributes depending on tension across centromeres during meiosis II and mitosis. *EMBO Reports,* **8**(2), pp. 173-180.

GRETEN, T.F. and JAFFEE, E.M., 1999. Cancer vaccines. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology,* **17**(3), pp. 1047-1060.

GRISWOLD, M.D., HOGARTH, C.A., BOWLES, J. and KOOPMAN, P., 2012. Initiating meiosis: the case for retinoic acid. *Biology of Reproduction,* **86**(2), pp. 35-41.

GRUNWALD, C., KOSLOWSKI, M., ARSIRAY, T., DHAENE, K., PRAET, M., VICTOR, A., MORRESI-HAUF, A., LINDNER, M., PASSLICK, B. and LEHR, H., 2006. Expression of multiple epigenetically regulated cancer/germline genes in nonsmall cell lung cancer. *International Journal of Cancer,* **118**(10), pp. 2522-2528.

GÜRE, A.O., CHUA, R., WILLIAMSON, B., GONEN, M., FERRERA, C.A., GNJATIC, S., RITTER, G., SIMPSON, A.J., CHEN, Y.T., OLD, L.J. and ALTORKI, N.K., 2005. Cancer-testis genes are coordinately expressed and are markers of poor outcome in non-small cell lung cancer. *Clinical Cancer Research: an Official Journal of the American Association for Cancer Research,* **11**(22), pp. 8055-8062.

GÜRE, A.O., WEI, I.J., OLD, L.J. and CHEN, Y., 2002. The SSX gene family: characterization of 9 complete genes. *International Journal of Cancer,* **101**(5), pp. 448-453.

HAITINA, T., LINDBLOM, J., RENSTRÖM, T. and FREDRIKSSON, R., 2006. Fourteen novel human members of mitochondrial solute carrier family 25 (SLC25) widely expressed in the central nervous system. *Genomics,* **88**(6), pp. 779-790.

HAN, J., ZHOU, H., HORAZDOVSKY, B., ZHANG, K., XU, R. and ZHANG, Z., 2007. Rtt109 acetylates histone H3 lysine 56 and functions in DNA replication. *Science,* **315**(5812), pp. 653-655.

HANAFUSA, T., MOHAMED, A.E., DOMAE, S., NAKAYAMA, E. and ONO, T., 2012. Serological identification of Tektin5 as a cancer/testis antigen and its immunogenicity. *BMC Cancer,* **12**, pp. 520-2407-12-520.

HANAHAN, D. and WEINBERG, R.A., 2011. Hallmarks of cancer: the next generation. *Cell,* **144**(5), pp. 646-674.

HANDEL, M.A. and SCHIMENTI, J.C., 2010. Genetics of mammalian meiosis: regulation, dynamics and impact on fertility. *Nature Reviews Genetics,* **11**(2), pp. 124-136.

HANDEL, M.A., 2004. The XY body: a specialized meiotic chromatin domain. *Experimental Cell Research,* **296**(1), pp. 57-63.

HARIGAYA, Y., TANAKA, H., YAMANAKA, S., TANAKA, K., WATANABE, Y., TSUTSUMI, C., CHIKASHIGE, Y., HIRAOKA, Y., YAMASHITA, A. and YAMAMOTO, M., 2006. Selective elimination of messenger RNA prevents an incidence of untimely meiosis. *Nature,* **442**(7098), pp. 45-50.

HARPER, L., GOLUBOVSKAYA, I. and CANDE, W.Z., 2004. A bouquet of chromosomes. *Journal of Cell Science,* **117**(Pt 18), pp. 4025-4032.

HATZIMICHAEL, E. and CROOK, T., 2013. Cancer Epigenetics: New Therapies and New Challenges. *Journal of Drug Delivery,* **2013**, pp. 1-9.

HE, L., JI, J., LIU, S., XUE, E., LIANG, Q. and MA, Z., 2014. Expression of cancer-testis antigen in multiple myeloma. *Journal of Huazhong University of Science and Technology [Medical Sciences]*, **34**, pp. 181-185.

HESS, R.A. and DE FRANCA, L.R., 2009. Spermatogenesis and cycle of the seminiferous epithelium. *Molecular Mechanisms in Spermatogenesis.* Springer, pp. 1-15.

HILLEN, W. and BERENS, C., 1994. Mechanisms underlying expression of Tn10 encoded tetracycline resistance. *Annual Reviews in Microbiology,* **48**(1), pp. 345-369.

HIROSE, Y., SUZUKI, R., OHBA, T., HINOHARA, Y., MATSUHARA, H., YOSHIDA, M., ITABASHI, Y., MURAKAMI, H. and YAMAMOTO, A., 2011. Chiasmata promote monopolar attachment of sister chromatids and their co-segregation toward the proper pole during meiosis I. *PLoS Genetics,* **7**(3), pp. e1001329.

HOELLER, D. and DIKIC, I., 2009. Targeting the ubiquitin system in cancer therapy. *Nature,* **458**(7237), pp. 438-444.

HOFMANN, O., CABALLERO, O.L., STEVENSON, B.J., CHEN, Y., COHEN, T., CHUA, R., MAHER, C.A., PANJI, S., SCHAEFER, U., KRUGER, A., LEHVASLAIHO, M., CARNINCI, P., HAYASHIZAKI, Y., JONGENEEL, C., SIMPSON, A.J., OLD, L.J. and HIDE, W., 2008. Genome-wide analysis of cancer/testis gene expression. *Proceedings of the National Academy of Sciences, USA,* **105**(51), pp. 20422-20427.

HOGARTH, C.A. and GRISWOLD, M.D., 2010. The key role of vitamin A in spermatogenesis. *The Journal of Clinical Investigation,* **120**(4), pp. 956-962.

HOGARTH, C.A. and GRISWOLD, M.D., 2013. Retinoic acid regulation of male meiosis. *Current Opinion in Endocrinology, Diabetes, and Obesity,* **20**(3), pp. 217-223.

HOLDCRAFT, R.W. and BRAUN, R.E., 2004. Androgen receptor function is required in Sertoli cells for the terminal differentiation of haploid spermatids. *Development (Cambridge, England),* **131**(2), pp. 459-467.

HOLM, L.R. and THON, G., 2012. New romance between RNA degradation pathways: Mmi1 and RNAi meet on heterochromatic islands. *The EMBO Journal,* **31**(10), pp. 2242-2243.

HOLSTEIN, A., SCHULZE, W. and DAVIDOFF, M., 2003. Understanding spermatogenesis is a prerequisite for treatment. *Reprod Biol Endocrinol,* **1**, pp. 107-122.

HOLT, J.E. and JONES, K.T., 2009. Control of homologous chromosome division in the mammalian oocyte. *Molecular Human Reproduction,* **15**(3), pp. 139-147.

HOU, S., SANG, M., GENG, C., LIU, W., LÜ, W., XU, Y. and SHAN, B., 2013a. Expressions of MAGE-A9 and MAGE-A11 in Breast Cancer and Their Mechanism of Expression. *Archives of Medical Research,* **45**(2014), pp. 44-51.

HOU, Y., FAN, W., YAN, L., LI, R., LIAN, Y., HUANG, J., LI, J., XU, L., TANG, F. and XIE, X.S., 2013b. Genome analyses of single human oocytes. *Cell,* **155**(7), pp. 1492-1506.

HOUGHTON, A.N., GOLD, J.S. and BLACHERE, N.E., 2001. Immunity against cancer: lessons learned from melanoma. *Current Opinion in Immunology,* **13**(2), pp. 134-140.

HUNDER, N.N., WALLEN, H., CAO, J., HENDRICKS, D.W., REILLY, J.Z., RODMYRE, R., JUNGBLUTH, A., GNJATIC, S., THOMPSON, J.A. and YEE, C., 2008. Treatment of metastatic melanoma with autologous CD4 T cells against NY-ESO-1. *New England Journal of Medicine,* **358**(25), pp. 2698-2703.

HUNT, P.A. and HASSOLD, T.J., 2002. Sex matters in meiosis. *Science (New York, N.Y.),* **296**(5576), pp. 2181-2183.

HUNTER, N. and KLECKNER, N., 2001. The Single-End Invasion: An Asymmetric Intermediate at the Double-Strand Break to Double-Holliday Junction Transition of Meiotic Recombination. *Cell,* **106**(1), pp. 59-70.

IBRAHIM, A.E., ARENDS, M.J., SILVA, A., WYLLIE, A.H., GREGER, L., ITO, Y., VOWLER, S.L., HUANG, T.H., TAVARÉ, S. and MURRELL, A., 2011. Sequential DNA methylation changes are associated with DNMT3B overexpression in colorectal neoplastic progression. *Gut,* **60**(4), pp. 499-508.

INAGAKI, A.A., SCHOENMAKERS, S.S. and BAARENDS, W.M.W., 2010. DNA double strand break repair, chromosome synapsis and transcriptional silencing in meiosis. *Epigenetics : official journal of the DNA Methylation Society,* **5**(4), pp. 255-266.

ISHIGURO, K. and WATANABE, Y., 2007. Chromosome cohesion in mitosis and meiosis. *Journal of Cell Science,* **120**(Pt 3), pp. 367-369.

ISHIGURO, K., KIM, J., SHIBUYA, H., HERNANDEZ-HERNANDEZ, A., SUZUKI, A., FUKAGAWA, T., SHIOI, G., KIYONARI, H., LI, X.C., SCHIMENTI, J., HOOG, C. and WATANABE, Y., 2014. Meiosis-specific cohesin mediates homolog recognition in mouse spermatocytes. *Genes & Development,* **28**(6), pp. 594-607.

JADHAV, T. and WOOTEN, M.W., 2009. Defining an Embedded Code for Protein Ubiquitination. *Journal of Proteomics & Bioinformatics,* **2**, pp. 316-340.

JANIC, A., MENDIZABAL, L., LLAMAZARES, S., ROSSELL, D. and GONZALEZ, C., 2010. Ectopic expression of germline genes drives malignant brain tumor growth in Drosophila. *Science,* **330**(6012), pp. 1824-1827.

JEMAL, A., SIEGEL, R., XU, J. and WARD, E., 2010. Cancer statistics, 2010. *CA: A Cancer Journal for Clinicians,* **60**(5), pp. 277-300.

JUNGBLUTH, A.A., ELY, S., DILIBERTO, M., NIESVIZKY, R., WILLIAMSON, B., FROSINA, D., CHEN, Y.T., BHARDWAJ, N., CHEN-KIANG, S., OLD, L.J. and CHO, H.J., 2005. The cancer-testis antigens CT7 (MAGE-C1) and MAGE-A3/6 are commonly expressed in multiple myeloma and correlate with plasma-cell proliferation. *Blood,* **106**(1), pp. 167-174.

KALEJS, M. and ERENPREISA, J., 2005. Cancer/testis antigens and gametogenesis: a review and "brain-storming" session. *Cancer Cell International,* **5**, pp 4-14.

KAMATA, Y., KUHARA, A., IWAMOTO, T., HAYASHI, K., KOIDO, S., KIMURA, T., EGAWA, S. and HOMMA, S., 2013. Identification of HLA class I-binding peptides derived from unique cancer-associated proteins by mass spectrometric analysis. *Anticancer Research,* **33**(5), pp. 1853-1859.

KASIMANICKAM, V. and KASIMANICKAM, R., 2013. Expression of CYP26b1 and Related Retinoic Acid Signalling Molecules in Young, Peripubertal and Adult Dog Testis. *Reproduction in Domestic Animals*, **48**(1), pp. 171-176.

KEENEY, S. and NEALE, M., 2006. Initiation of meiotic recombination by formation of DNA double-strand breaks: mechanism and regulation. *Biochemical Society Transactions*, **34**(4), pp. 523-525.

KEENEY, S., GIROUX, C.N. and KLECKNER, N., 1997. Meiosis-specific DNA double-strand breaks are catalyzed by Spo11, a member of a widely conserved protein family. *Cell*, **88**(3), pp. 375-384.

KEENEY, S.S., 2001. Mechanism and control of meiotic recombination initiation. *Current Topics in Developmental Biology*, **52**, pp. 1-53.

KERR, C.L., HILL, C.M., BLUMENTHAL, P.D. and GEARHART, J.D., 2008. Expression of pluripotent stem cell markers in the human fetal testis. *Stem Cells*, **26**(2), pp. 412-421.

KLECKNER, N., 1996. Meiosis: how could it work? *Proceedings of the National Academy of Sciences*, **93**(16), pp. 8167-8174.

KLEIN, C.A., 2008. The metastasis cascade. *Science*, **321**(5897), pp. 1785-1787.

KLENOVA, E.M., MORSE, H.C., OHLSSON, R. and LOBANENKOV, V.V., 2002. The novel *BORIS + CTCF* gene family is uniquely involved in the epigenetics of normal biology and cancer, *Seminars in Cancer Biology*, **12**, pp. 399-414.

KLUTSTEIN, M. and COOPER, J.P., 2014. The Chromosomal Courtship Dance—homolog pairing in early meiosis. *Current Opinion in Cell Biology*, **26**, pp. 123-131.

KNAPP, S., 2013. Testis specific gene expression drives disease progression and Rituximab resistance in lymphoma. *EMBO Molecular Medicine*, **5**(8), pp. 1149-1150.

KOPFSTEIN, L. and CHRISTOFORI, G., 2006. Metastasis: cell-autonomous mechanisms versus contributions by the tumor microenvironment. *Cellular and Molecular Life Sciences CMLS*, **63**(4), pp. 449-468.

KOUZARIDES, T., 2007. Chromatin modifications and their function. *Cell*, **128**(4), pp. 693-705.

KOUZNETSOVA, A., NOVAK, I., JESSBERGER, R. and HOEOEG, C., 2005. SYCP2 and SYCP3 are required for cohesin core integrity at diplotene but not for centromere cohesion at the first meiotic division. *Journal of Cell Science*, **118**(10), pp. 2271-2278.

KOZAK, M., 1991. An analysis of vertebrate mRNA sequences: intimations of translational control. *The Journal of Cell Biology*, **115**(4), pp. 887-903.

KRENTZ, A.D., MURPHY, M.W., SARVER, A.L., GRISWOLD, M.D., BARDWELL, V.J. and ZARKOWER, D., 2011. DMRT1 promotes oogenesis by transcriptional activation of *Stra8* in the mammalian fetal ovary. *Developmental Biology*, **356**(1), pp. 63-70.

KÜPPERS, R., 2005. Mechanisms of B-cell lymphoma pathogenesis. *Nature Reviews Cancer*, **5**(4), pp. 251-262.

KUZNETSOV, S., PELLEGRINI, M., SHUDA, K., FERNANDEZ-CAPETILLO, O., LIU, Y., MARTIN, B.K., BURKETT, S., SOUTHON, E., PATI, D., TESSAROLLO, L., WEST, S.C., DONOVAN, P.J., NUSSENZWEIG, A. and SHARAN, S.K., 2007. RAD51C deficiency in mice results in early prophase I arrest in males and sister chromatid separation at metaphase II in females. *Journal of Cell Biology,* **176**(5), pp. 581-592.

LEE, B.B. and AMON, A.A., 2001. Meiosis: how to create a specialized cell cycle. *Current Opinion in Cell Biology,* **13**(6), pp. 770-777.

LEE, J.T. and BARTOLOMEI, M.S., 2013. X-inactivation, imprinting, and long noncoding RNAs in health and disease. *Cell,* **152**(6), pp. 1308-1323.

LEE, T.S., KIM, J.W., KANG, G.H., PARK, N.H., SONG, Y.S., KANG, S.B. and LEE, H.P., 2006. DNA hypomethylation of CAGE promotors in squamous cell carcinoma of uterine cervix. *Annals of the New York Academy of Sciences,* **1091**(1), pp. 218-224.

LENGAUER, C., KINZLER, K.W. and VOGELSTEIN, B., 1998. Genetic instabilities in human cancers. *Nature,* **396**(6712), pp. 643-649.

LEVY, E., D'EUSTACHIO, P., COWAN, N.J. and LIEM, R.K., 1987. Structure and evolutionary origin of the gene encoding mouse NF-M, the middle-molecular-mass neurofilament protein. *European Journal of Biochemistry,* **166**(1), pp. 71-77.

LI, H. and CLAGETT-DAME, M., 2009. Vitamin A deficiency blocks the initiation of meiosis of germ cells in the developing rat ovary *in vivo*. *Biology of Reproduction,* **81**(5), pp. 996-1001.

LI, H., PALCZEWSKI, K., BAEHR, W. and CLAGETT-DAME, M., 2011. Vitamin A deficiency results in meiotic failure and accumulation of undifferentiated spermatogonia in prepubertal mouse testis. *Biology of Reproduction,* **84**(2), pp. 336-341.

LI, N., WANG, T. and HAN, D., 2012. Structural, cellular and molecular aspects of immune privilege in the testis. *Frontiers in Immunology,* **3**(152), pp. 1-12.

LIANG, J., DING, T., GUO, Z., YU, X., HU, Y., ZHENG, L. and XU, J., 2013. Expression pattern of tumour-associated antigens in hepatocellular carcinoma: association with immune infiltration and disease progression. *British Journal of Cancer,* **109**(4), pp. 1031-1039.

LIE, P.P., CHENG, C.Y. and MRUK, D.D., 2013. Signalling pathways regulating the blood–testis barrier. *The International Journal of Biochemistry & Cell Biology,* **45**(3), pp. 621-625.

LIM, S.H., ZHANG, Y. and ZHANG, J., 2012. Cancer-testis antigens: the current status on antigen regulation and potential clinical use. *American Journal of Blood Research,* **2**(1), pp. 29-35.

LIN, Y., GILL, M.E., KOUBOVA, J. and PAGE, D.C., 2008. Germ cell-intrinsic and -extrinsic factors govern meiotic initiation in mouse embryos. *Science (New York, N.Y.),* **322**(5908), pp. 1685-1687.

LINDBLOM, A. and LILJEGREN, A., 2000. Regular review: tumour markers in malignancies. *BMJ: British Medical Journal,* **320**(7232), pp. 424-427.

LISBY, M., BARLOW, J.H., BURGESS, R.C. and ROTHSTEIN, R., 2004. Choreography of the DNA damage response: spatiotemporal relationships among checkpoint and repair proteins. *Cell,* **118**(6), pp. 699-713.

LIVERA, G., ROUILLER-FABRE, V., PAIRAULT, C., LEVACHER, C. and HABERT, R., 2002. Regulation and perturbation of testicular functions by vitamin A. *Reproduction (Cambridge, England),* **124**(2), pp. 173-180.

LONGHESE, M.P., BONETTI, D., MANFRINI, N. and CLERICI, M., 2010. Mechanisms and regulation of DNA end resection. *The EMBO Journal,* **29**(17), pp. 2864-2874.

LOUKINOV, D.I., PUGACHEVA, E., VATOLIN, S., PACK, S.D., MOON, H., CHERNUKHIN, I., MANNAN, P., LARSSON, E., KANDURI, C. and VOSTROV, A.A., 2002. BORIS, a novel male germ-line-specific protein associated with epigenetic reprogramming events, shares the same 11-zinc-finger domain with CTCF, the insulator protein involved in reading imprinting marks in the soma. *Proceedings of the National Academy of Sciences,* **99**(10), pp. 6806-6811.

LUGER, K., MÄDER, A.W., RICHMOND, R.K., SARGENT, D.F. and RICHMOND, T.J., 1997. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature,* **389**(6648), pp. 251-260.

LUPINACCI, R.M., MENEGAUX, F. and TRESALLET, C., 2013. Hepatocellular carcinoma, from monitoring to treatment. *Soins; La Revue De Reference Infirmiere,* **780**, pp. 35-37.

LUTZ, W.K., 2002. Differences in individual susceptibility to toxic effects of chemicals determine the dose–response relationship and consequences of setting exposure standards. *Toxicology Letters,* **126**(3), pp. 155-158.

MÄBERT, K., COJOC, M., PEITZSCH, C., KURTH, I., SOUCHELNYTSKYI, S. and DUBROVSKA, A., 2014. Cancer biomarker discovery: current status and future perspectives. *International Journal of Radiation Biology,* **12**, pp. 1-19.

MANN, M. and JENSEN, O.N., 2003. Proteomic analysis of post-translational modifications. *Nature Biotechnology,* **21**(3), pp. 255-261.

MARK, M., JACOBS, H., OULAD-ABDELGHANI, M., DENNEFELD, C., FERET, B., VERNET, N., CODREANU, C.A., CHAMBON, P. and GHYSELINCK, N.B., 2008. STRA8-deficient spermatocytes initiate, but fail to complete, meiosis and undergo premature chromosome condensation. *Journal of Cell Science,* **121**(Pt 19), pp. 3233-3242.

MARTIN-KLEINER, I., 2012. BORIS in human cancers–A review. *European Journal of Cancer,* **48**(6), pp. 929-935.

MATA, J., LYNE, R., BURNS, G. and BÄHLER, J., 2002. The transcriptional program of meiosis and sporulation in fission yeast. *Nature Genetics,* **32**(1), pp. 143-147.

MATSON, C.K., MURPHY, M.W., GRISWOLD, M.D., YOSHIDA, S., BARDWELL, V.J. and ZARKOWER, D., 2010. The mammalian doublesex homolog DMRT1 is a transcriptional gatekeeper that controls the mitosis versus meiosis decision in male germ cells. *Developmental Cell,* **19**(4), pp. 612-624.

MATSON, C.K., MURPHY, M.W., SARVER, A.L., GRISWOLD, M.D., BARDWELL, V.J. and ZARKOWER, D., 2011. DMRT1 prevents female reprogramming in the postnatal mammalian testis. *Nature,* **476**(7358), pp. 101-104.

MEHTA, G.D., KUMAR, R., SRIVASTAVA, S. and GHOSH, S.K., 2013. Cohesin: Functions beyond sister chromatid cohesion. *FEBS Letters,* **587**(15), pp. 2299-2312.

MEHTA, G.D., RIZVI, S.M.A. and GHOSH, S.K., 2012. Cohesin: a guardian of genome integrity. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, **1823**(8), pp. 1324-1342.

MELLMAN, I., COUKOS, G. and DRANOFF, G., 2011. Cancer immunotherapy comes of age. *Nature*, **480**(7378), pp. 480-489.

MILLER, M.P., AMON, A. and ÜNAL, E., 2013. Meiosis I: when chromosomes undergo extreme makeover. *Current Opinion in Cell Biology*, **25**(6), pp. 687-696.

MILLER, M.P., ÜNAL, E., BRAR, G.A. and AMON, A., 2012. Meiosis I chromosome segregation is established through regulation of microtubule–kinetochore interactions. *Elife*, **1**, pp. 1-30.

MIRANDOLA, L., J. CANNON, M., COBOS, E., BERNARDINI, G., JENKINS, M.R., KAST, W.M. and CHIRIVA-INTERNATI, M., 2011. Cancer testis antigens: novel biomarkers and targetable proteins for ovarian cancer. *International Reviews of Immunology*, **30**(2-3), pp. 127-137.

MITURSKI, R., BOGUSIEWICZ, M., CIOTTA, C., BIGNAMI, M., GOGACZ, M. and BURNOUF, D., 2002. Mismatch repair genes and microsatellite instability as molecular markers for gynecological cancer detection. *Experimental Biology and Medicine (Maywood, N.J.)*, **227**(8), pp. 579-586.

MIYAMOTO, T., SENGOKU, K., TAKUMA, N., HASUIKE, S., HAYASHI, H., YAMAUCHI, T., YAMASHITA, T. and ISHIKAWA, M., 2002. Isolation and expression analysis of the testis-specific gene, STRA8, stimulated by retinoic acid gene 8. *Journal of Assisted Reproduction and Genetics*, **19**(11), pp. 531-535.

MIYAMOTO, T., TSUJIMURA, A., MIYAGAWA, Y., KOH, E., NAMIKI, M., HORIKAWA, M., SAIJO, Y. and SENGOKU, K., 2012. Single-nucleotide polymorphisms in HORMAD1 may be a risk factor for azoospermia caused by meiotic arrest in Japanese patients. *Asian Journal of Andrology*, **14**(4), pp. 580-583.

MODARRESSI, M.H. and GHAFOURIFARD, S., 2011. Potential of cancer-testis antigens as targets for cancer immunotherapy. *Bridging Cell Biology and Genetics to the Cancer Clinic, Kerala, India*, **2**, pp. 83-98.

MORLEY, A.A. and TURNER, 1999. The contribution of exogenous and endogenous mutagens to *in vivo* mutations. *Mutation Research-Fundamental and Molecular Mechanisms of Mutagenesis*, **428**(1-2), pp. 11-15.

MRUK, D.D. and CHENG, C.Y., 2010. Tight junctions in the testis: new perspectives. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **365**(1546), pp. 1621-1635.

MUKHOPADHYAY, D. and RIEZMAN, H., 2007. Proteasome-independent functions of ubiquitin in endocytosis and signalling. *Science (New York, N.Y.)*, **315**(5809), pp. 201-205.

MUNGER, K., 2002. The role of human papillomaviruses in human cancers. *Front Biosci*, **7**(3), pp. 641-649.

NAKAZAWA, T., KONDO, T., MA, D., NIU, D., MOCHIZUKI, K., KAWASAKI, T., YAMANE, T., IINO, H., FUJII, H. and KATOH, R., 2012. Global histone modification of histone H3 in colorectal cancer and its precursor lesions. *Human Pathology*, **43**(6), pp. 834-842.

NAMBIAR, S.S., MIRMOHAMMADSADEGH, A.A. and HENGGE, U.R.U., 2008. Cutaneous melanoma: fishing with chips. *Current Molecular Medicine*, **8**(3), pp. 235-243.

NASMYTH, K. and HAERING, C.H., 2009. Cohesin: its roles and mechanisms. *Annual Review of Genetics,* **43**, pp. 525-558.

NASMYTH, K., 2001. Disseminating the genome: joining, resolving, and separating sister chromatids during mitosis and meiosis. *Annual Review of Genetics,* **35**(1), pp. 673-745.

NASMYTH, K., 2011. Cohesin: a catenase with separate entry and exit gates? *Nature Cell Biology,* **13**(10), pp. 1170-1177.

NEALE, M.J. and KEENEY, S., 2006. Clarifying the mechanics of DNA strand exchange in meiotic recombination. *Nature,* **442**(7099), pp. 153-158.

NEUMANN, B., WALTER, T., HÉRICHÉ, J., BULKESCHER, J., ERFLE, H., CONRAD, C., ROGERS, P., POSER, I., HELD, M. and LIEBEL, U., 2010. Phenotypic profiling of the human genome by time-lapse microscopy reveals cell division genes. *Nature,* **464**(7289), pp. 721-727.

NIU, H., WAN, L., BAUMGARTNER, B., SCHAEFER, D., LOIDL, J. and HOLLINGSWORTH, N.M., 2005. Partner Choice during Meiosis Is Regulated by Hop1-promoted Dimerization of Mek1. *Molecular Biology of the Cell,* **16**(12), pp. 5804-5818.

NUSSENZWEIG, A. and NUSSENZWEIG, M.C., 2010. Origin of chromosomal translocations in lymphoid cancer. *Cell,* **141**(1), pp. 27-38.

O'SHEA, J.J., KANNO, Y. and CHAN, A.C., 2014. In Search of Magic Bullets: The Golden Age of Immunotherapeutics. *Cell,* **157**(1), pp. 227-240.

O'DONNELL, L., MEACHEM, S.J., STANTON, P.G. and MCLACHLAN, R.I., 2006. Chapter 21 - Endocrine Regulation of Spermatogenesis. In: J.D. NEILL, T.M. PLANT, D.W. PFAFF, J.R.G. CHALLIS, D.M.D. KRETSER, J.S. RICHARDS and P.M. WASSARMAN, eds, *Knobil and Neill's Physiology of Reproduction (Third Edition).* St Louis: Academic Press, pp. 1017-1069.

O'FARRELL, P.H., 2011. Quiescence: early evolutionary origins and universality do not imply uniformity. *Philosophical Transactions of the Royal Society of London.Series B, Biological Sciences,* **366**(1584), pp. 3498-3507.

OLD, L.J. and CHEN, Y.T., 1998. New paths in human cancer serology. *The Journal of Experimental Medicine,* **187**(8), pp. 1163-1167.

ONO, T., KURASHIGE, T., HARADA, N., NOGUCHI, Y., SAIKA, T., NIIKAWA, N., AOE, M., NAKAMURA, S., HIGASHI, T., HIRAKI, A., WADA, H., KUMON, H., OLD, L.J. and NAKAYAMA, E., 2001. Identification of proacrosin binding protein sp32 precursor as a human cancer/testis antigen. *Proceedings of the National Academy of Sciences, USA,* **98**(6), pp. 3282-3287.

O'SHAUGHNESSY, P.J., 2014. Hormonal control of germ cell development and spermatogenesis, *Seminars in Cell & Developmental Biology*, **29**(2014), pp. 55-65.

OULAD-ABDELGHANI, M., BOUILLET, P., CHAZAUD, C., DOLLÉ, P. and CHAMBON, P., 1996a. AP-2.2: a novel AP-2-related transcription factor induced by retinoic acid during differentiation of P19 embryonal carcinoma cells. *Experimental Cell Research,* **225**(2), pp. 338-347.

OULAD-ABDELGHANI, M., BOUILLET, P., DECIMO, D., GANSMULLER, A., HEYBERGER, S., DOLLE, P., BRONNER, S., LUTZ, Y. and CHAMBON, P., 1996b. Characterization of a

premeiotic germ cell-specific cytoplasmic protein encoded by Stra8, a novel retinoic acid-responsive gene. *The Journal of Cell Biology,* **135**(2), pp. 469-477.

PAGE, S.L. and HAWLEY, R.S., 2004. The genetics and molecular biology of the synaptonemal complex. *Annual Review of Cell and Developmental Biology,* **20**, pp. 525-558.

PAGE, S.L., KHETANI, R.S., LAKE, C.M., NIELSEN, R.J., JEFFRESS, J.K., WARREN, W.D., BICKEL, S.E. and HAWLEY, R.S., 2008. Corona is required for higher-order assembly of transverse filaments into full-length synaptonemal complex in Drosophila oocytes. *PLoS Genetics,* **4**(9), pp. e1000194.

PAGE, S.L.S. and HAWLEY, R.S.R., 2003. Chromosome choreography: the meiotic ballet. *Science,* **301**(5634), pp. 785-789.

PAIK, J., VOGEL, S., QUADRO, L., PIANTEDOSI, R., GOTTESMAN, M., LAI, K., HAMBERGER, L., VIEIRA MDE, M. and BLANER, W.S., 2004. Vitamin A: overlapping delivery pathways to tissues from the circulation. *The Journal of Nutrition,* **134**(1), pp. 276S-280S.

PETRONCZKI, M., SIOMOS, M.F. and NASMYTH, K., 2003. Un menage a quatre: the molecular biology of chromosome segregation in meiosis. *Cell,* **112**(4), pp. 423-440.

PETROSKI, M.D. and DESHAIES, R.J., 2005. Function and regulation of cullin–RING ubiquitin ligases. *Nature Reviews Molecular Cell Biology,* **6**(1), pp. 9-20.

PHILLIPS, B.T., GASSEI, K. and ORWIG, K.E., 2010. Spermatogonial stem cell regulation and spermatogenesis. *Philosophical Transactions of the Royal Society of London.Series B, Biological Sciences,* **365**(1546), pp. 1663-1678.

PLAYER, A., BARRETT, J.C. and KAWASAKI, E.S., 2004. Laser capture microdissection, microarrays and the precise definition of a cancer cell. *Expert Review of Molecular Diagnostics,* **4**(6), pp. 831-840.

QIU, G., FANG, J. and HE, Y., 2006. 5′ CpG island methylation analysis identifies the MAGE-A1 and MAGE-A3 genes as potential markers of HCC. *Clinical Biochemistry,* **39**(3), pp. 259-266.

QIU, J., GAO, C.L., ZHANG, M., CHEN, R.H., CHI, X., LIU, F., ZHANG, C.M., JI, C.B., CHEN, X.H., ZHAO, Y.P., LI, X.N., TONG, M.L., NI, Y.H. and GUO, X.R., 2009. LYRM1, a novel gene promotes proliferation and inhibits apoptosis of preadipocytes. *European Journal of Endocrinology / European Federation of Endocrine Societies,* **160**(2), pp. 177-184.

RAJAGOPALAN, K., MOONEY, S.M., PAREKH, N., GETZENBERG, R.H. and KULKARNI, P., 2011. A majority of the cancer/testis antigens are intrinsically disordered proteins. *Journal of Cellular Biochemistry,* **112**(11), pp. 3256-3267.

RAVERDEAU, M., GELY-PERNOT, A., FERET, B., DENNEFELD, C., BENOIT, G., DAVIDSON, I., CHAMBON, P., MARK, M. and GHYSELINCK, N.B., 2012. Retinoic acid induces Sertoli cell paracrine signals for spermatogonia differentiation but cell autonomously drives spermatocyte meiosis. *Proceedings of the National Academy of Sciences of the United States of America,* **109**(41), pp. 16582-16587.

REN, P. and BISHOP, P.D., 1989. Uptake and metabolism of retinol in isolated cells of germinal epithelium *in vitro*. *Journal of Tongji Medical University,* **9**(1), pp. 36-43.

RENKVIST, N.N., CASTELLI, C.C., ROBBINS, P.F.P. and PARMIANI, G.G., 2001. A listing of human tumor antigens recognized by T cells. *Cancer Immunology, Immunotherapy : CII,* **50**(1), pp. 3-15.

REVENKOVA, E. and JESSBERGER, R., 2005. Keeping sister chromatids together: cohesins in meiosis. *Reproduction (Cambridge, England),* **130**(6), pp. 783-790.

REYNOLDS, N. and COOKE, H.J., 2005. Role of the *DAZ* genes in male fertility. *Reproductive Biomedicine Online,* **10**(1), pp. 72-80.

ROBERTSON, K.D., 2005. DNA methylation and human disease. *Nature Reviews Genetics,* **6**(8), pp. 597-610.

ROOIJ, D.G. and RUSSELL, L.D., 2000. All you wanted to know about spermatogonia but were afraid to ask. *Journal of Andrology,* **21**(6), pp. 776-798.

ROSA, A.M., DABAS, N., BYRNES, D.M., ELLER, M.S. and GRICHNIK, J.M., 2012. Germ cell proteins in melanoma: prognosis, diagnosis, treatment, and theories on expression. *Journal of Skin Cancer,* **2012**, 621968.

ROSS, M.T., GRAFHAM, D.V., COFFEY, A.J., SCHERER, S., MCLAY, K., MUZNY, D., PLATZER, M., HOWELL, G.R., BURROWS, C., BIRD, C.P., FRANKISH, A., LOVELL, F.L., HOWE, K.L., ASHURST, J.L., FULTON, R.S., SUDBRAK, R., WEN, G., JONES, M.C., HURLES, M.E., ANDREWS, T., SCOTT, C.E., SEARLE, S., RAMSER, J., WHITTAKER, A., DEADMAN, R., CARTER, N.P., HUNT, S.E., CHEN, R., CREE, A., GUNARATNE, P., HAVLAK, P., HODGSON, A., METZKER, M.L. and RICHARDS, S., 2005. The DNA sequence of the human X chromosome. *Nature,* **434**(7030), pp. 325-337.

ROUSSEAUX, S., DEBERNARDI, A., JACQUIAU, B., VITTE, A.L., VESIN, A., NAGY-MIGNOTTE, H., MORO-SIBILOT, D., BRICHON, P.Y., LANTUEJOUL, S., HAINAUT, P., LAFFAIRE, J., DE REYNIES, A., BEER, D.G., TIMSIT, J.F., BRAMBILLA, C., BRAMBILLA, E. and KHOCHBIN, S., 2013. Ectopic activation of germline and placental genes identifies aggressive metastasis-prone lung cancers. *Science Translational Medicine,* **5**(186), pp. 186ra66.

RUDDON, R.W., 2007. *Cancer biology.* Oxford University Press, USA.

RUGGIU, M., SAUNDERS, P.T. and COOKE, H.J., 2000. Dynamic subcellular distribution of the DAZL protein is confined to primate male germ cells. *Journal of Andrology,* **21**(3), pp. 470-477.

RUSS, B.E., DENTON, A.E., HATTON, L., CROOM, H., OLSON, M.R. and TURNER, S.J., 2012. Defining the molecular blueprint that drives CD8 T cell differentiation in response to infection. *Frontiers in Immunology,* **3**(371), pp. 1-11.

SADOWSKI, M., SURYADINATA, R., TAN, A.R., ROESLEY, S.N.A. and SARCEVIC, B., 2012. Protein monoubiquitination and polyubiquitination generate structural diversity to control distinct biological processes. *IUBMB Life,* **64**(2), pp. 136-142.

SAJESH, B.V., LICHTENSZTEJN, Z. and MCMANUS, K.J., 2013. Sister chromatid cohesion defects are associated with chromosome instability in Hodgkin lymphoma cells. *BMC Cancer,* **13**(1), pp. 391-400.

SAMMUT, S.J., FEICHTINGER, J., STUART, N., WAKEMAN, J.A., LARCOMBE, L. and MCFARLANE, R.J., 2014. A novel cohort of cancer-testis biomarker genes revealed through meta-analysis of clinical data sets. *Oncoscience,* **1**(5), pp. 349-359.

SAN FILIPPO, J., SUNG, P. and KLEIN, H., 2008. Mechanism of eukaryotic homologous recombination. *Annu.Rev.Biochem.,* **77**, pp. 229-257.

SANTARIUS, T., SHIPLEY, J., BREWER, D., STRATTON, M.R. and COOPER, C.S., 2010. A census of amplified and overexpressed human cancer genes. *Nature Reviews Cancer,* **10**(1), pp. 59-64.

SCANLAN, M.J., GURE, A.O., JUNGBLUTH, A.A., OLD, L.J. and CHEN, Y., 2002. Cancer/testis antigens: an expanding family of targets for cancer immunotherapy. *Immunological Reviews,* **188**, pp. 22-32.

SCANLAN, M.J.M., SIMPSON, A.J.G.A. and OLD, L.J.L., 2004. The cancer/testis genes: review, standardization, and commentary. *Cancer Immunity,* **4**, pp. 1-15.

SCHERTHAN, H., 2007. Telomeres and meiosis in health and disease: Telomere attachment and clustering during meiosis. *Cellular and Molecular Life Sciences CMLS,* **64**(2), pp. 117-124.

SCHRAMM, S., FRAUNE, J., NAUMANN, R., HERNANDEZ-HERNANDEZ, A., HÖÖG, C., COOKE, H.J., ALSHEIMER, M. and BENAVENTE, R., 2011. A novel mouse synaptonemal complex protein is essential for loading of central element proteins, recombination, and fertility. *PLoS Genetics,* **7**(5), pp. e1002088.

SCHRÖDER, F.H., HUGOSSON, J., ROOBOL, M.J., TAMMELA, T.L., CIATTO, S., NELEN, V., KWIATKOWSKI, M., LUJAN, M., LILJA, H. and ZAPPA, M., 2009. Screening and prostate-cancer mortality in a randomized European study. *New England Journal of Medicine,* **360**(13), pp. 1320-1328.

SERRENTINO, M. and BORDE, V., 2012. The spatial regulation of meiotic recombination hotspots: Are all DSB hotspots crossover hotspots? *Experimental Cell Research,* **318**(12), pp. 1347-1352.

SHACTER, E. and WEITZMAN, S.A., 2002. Chronic inflammation and cancer. *Oncology,* **16**(2), pp. 217-230.

SHARMA, S., KELLY, T.K. and JONES, P.A., 2010. Epigenetics in cancer. *Carcinogenesis,* **31**(1), pp. 27-36.

SHICHIJO, S.S., YAMADA, A.A., SAGAWA, K.K., IWAMOTO, O.O., SAKATA, M.M., NAGAI, K.K. and ITOH, K.K., 1996. Induction of MAGE genes in lymphoid cells by the demethylating agent 5-aza-2'-deoxycytidine. *Japanese Journal of Cancer Research: Gann,* **87**(7), pp. 751-756.

SIEGEL, R., NAISHADHAM, D. and JEMAL, A., 2012. Cancer statistics, 2012. *CA: A Cancer Journal for Clinicians,* **62**(1), pp. 10-29.

SIMPSON, A.J.G.A., CABALLERO, O.L.O., JUNGBLUTH, A.A., CHEN, Y.Y. and OLD, L.J.L., 2005. Cancer/testis antigens, gametogenesis and cancer. *Nature Reviews Cancer,* **5**(8), pp. 615-625.

SJÖGREN, C. and STRÖM, L., 2010. S-phase and DNA damage activated establishment of sister chromatid cohesion—importance for DNA repair. *Experimental Cell Research,* **316**(9), pp. 1445-1453.

SMITH, H.A., CRONK, R.J., LANG, J.M. and MCNEEL, D.G., 2011. Expression and immunotherapeutic targeting of the SSX family of cancer-testis antigens in prostate cancer. *Cancer Research,* **71**(21), pp. 6785-6795.

SMITH, Z.D. and MEISSNER, A., 2013. DNA methylation: roles in mammalian development. *Nature Reviews Genetics,* **14**(3), pp. 204-220.

SMORAG, L., XU, X., ENGEL, W. and PANTAKANI, D., 2014. The roles of DAZL in RNA biology and development. *Wiley Interdisciplinary Reviews,* doi:10.1002/wrna.1228.

SNYDER, E.M., SMALL, C. and GRISWOLD, M.D., 2010. Retinoic acid availability drives the asynchronous initiation of spermatogonial differentiation in the mouse. *Biology of Reproduction,* **83**(5), pp. 783-790.

STRACKER, T.H. and PETRINI, J.H., 2011. The MRE11 complex: starting from the ends. *Nature Reviews Molecular Cell Biology,* **12**(2), pp. 90-103.

STRAHM, B. and CAPRA, M., 2005. Insights into the molecular basis of cancer development. *Current Paediatrics,* **15**(4), pp. 333-338.

STROM, L. and SJOGREN, C., 2007. Chromosome segregation and double-strand break repair - a complex connection. *Current Opinion in Cell Biology,* **19**(3), pp. 344-349.

SU, L., MRUK, D.D. and CHENG, C.Y., 2011. Drug transporters, the blood-testis barrier, and spermatogenesis. *The Journal of Endocrinology,* **208**(3), pp. 207-223.

SUGIYAMA, T. and SUGIOKA-SUGIYAMA, R., 2011. Red1 promotes the elimination of meiosis-specific mRNAs in vegetatively growing fission yeast. *The EMBO Journal,* **30**(6), pp. 1027-1039.

SUNG, P.P., KREJCI, L.L., VAN KOMEN, S.S. and SEHORN, M.G.M., 2003. Rad51 recombinase and recombination mediators. *The Journal of Biological Chemistry,* **278**(44), pp. 42729-42732.

SZEKVOELGYI, L. and NICOLAS, A., 2010. From meiosis to postmeiotic events: Homologous recombination is obligatory but flexible. *FEBS Journal,* **277**(3), pp. 571-589.

TAN, C.S., SALIM, A., PLONER, A., LEHTIÖ, J., CHIA, K.S. and PAWITAN, Y., 2009. Correlating gene and protein expression data using Correlated Factor Analysis. *BMC Bioinformatics,* **10**(1), pp. 272-284.

TEDESCO, M., LA SALA, G., BARBAGALLO, F., DE FELICI, M. and FARINI, D., 2009. STRA8 shuttles between nucleus and cytoplasm and displays transcriptional activity. *The Journal of Biological Chemistry,* **284**(51), pp. 35781-35793.

THOMAS, C.M. and SWEEP, C.G., 2001. Serum tumor markers: past, state of the art, and future. *The International Journal of Biological Markers,* **16**(2), pp. 73-86.

THOMAS, N. and GOODYER, I., 2003. Stealth sensors: real-time monitoring of the cell cycle. *Targets,* **2**(1), pp. 26-33.

TSAI, J.H. and MCKEE, B.D., 2011. Homologous pairing and the role of pairing centers in meiosis. *Journal of Cell Science,* **124**(Pt 12), pp. 1955-1963.

TUERECI, O., SAHIN, U., ZWICK, C., KOSLOWSKI, M., SEITZ, G. and PFREUNDSCHUH, M., 1998. Identification of a meiosis-specific protein as a member of the class of cancer/testis antigens. *Proceedings of the National Academy of Sciences, USA*, **95**(9), pp. 5211-5216.

UHLMANN, F., 2004. The mechanism of sister chromatid cohesion. *Experimental Cell Research*, **296**(1), pp. 80-85.

VALMORI, D., SOULEIMANIAN, N.E., TOSELLO, V., BHARDWAJ, N., ADAMS, S., O'NEILL, D., PAVLICK, A., ESCALON, J.B., CRUZ, C.M., ANGIULLI, A., ANGIULLI, F., MEARS, G., VOGEL, S.M., PAN, L., JUNGBLUTH, A.A., HOFFMANN, E.W., VENHAUS, R., RITTER, G., OLD, L.J. and AYYOUB, M., 2007. Vaccination with NY-ESO-1 protein and CpG in Montanide induces integrated antibody/Th1 responses and CD8 T cells through cross-priming. *Proceedings of the National Academy of Sciences, USA*, **104**(21), pp. 8947-8952.

VAN DEN EYNDE, B., PEETERS, O., DE BACKER, O., GAUGLER, B., LUCAS, S. and BOON, T., 1995. A new family of genes coding for an antigen recognized by autologous cytolytic T lymphocytes on a human melanoma. *The Journal of Experimental Medicine*, **182**(3), pp. 689-698.

VAN DER BRUGGEN, P., TRAVERSARI, C., CHOMEZ, P., LURQUIN, C., DE PLAEN, E., VAN DEN EYNDE, B., KNUTH, A. and BOON, T., 1991. A gene encoding an antigen recognized by cytolytic T lymphocytes on a human melanoma. *Science*, **254**(5038), pp. 1643-1647.

VAN DUIN, M., BROYL, A., DE KNEGT, Y., GOLDSCHMIDT, H., RICHARDSON, P.G., HOP, W.C., VAN DER HOLT, B., JOSEPH-PIETRAS, D., MULLIGAN, G., NEUWIRTH, R., SAHOTA, S.S. and SONNEVELD, P., 2011. Cancer testis antigens in newly diagnosed and relapse multiple myeloma: prognostic markers and potential targets for immunotherapy. *Haematologica*, **96**(11), pp. 1662-1669.

VERNET, N., DENNEFELD, C., ROCHETTE-EGLY, C., OULAD-ABDELGHANI, M., CHAMBON, P., GHYSELINCK, N.B. and MARK, M., 2006. Retinoic acid metabolism and signaling pathways in the adult and developing mouse testis. *Endocrinology*, **147**(1), pp. 96-110.

VILLENEUVE, A.M. and HILLERS, K.J., 2001. Whence meiosis? *Cell*, **106**(6), pp. 647-650.

VINSON, C. and CHATTERJEE, R., 2012. CG methylation. *Epigenomics*, **4**(6), pp. 655-663.

VOGELSTEIN, B. and KINZLER, K.W., 2004. Cancer genes and the pathways they control. *Nature Medicine*, **10**(8), pp. 789-799.

WANG, K. and WU, Y.J., 2013. Signaling pathway of meiosis induced by retinoic acid during spermatogenesis. *Zhonghua nan ke xue = National Journal of Andrology*, **19**(2), pp. 173-177.

WANG, Z., ZHANG, Y., RAMSAHOYE, B., BOWEN, D. and LIM, S., 2004. Sp17 gene expression in myeloma cells is regulated by promoter methylation. *British Journal of Cancer*, **91**(8), pp. 1597-1603.

WASSLER, M., SYNTIN, P., SUTTON-WALSH, H.G., HSIA, N., HARDY, D.M. and CORNWALL, G.A., 2002. Identification and characterization of cystatin-related epididymal spermatogenic protein in human spermatozoa: localization in the equatorial segment. *Biology of Reproduction*, **67**(3), pp. 795-803.

WASSMANN, K., 2013. Sister chromatid segregation in meiosis II: Deprotection through phosphorylation. *Cell Cycle*, **12**(9), pp. 1352-1359.

WEINBERG, R.A., 2007. *The biology of cancer.* Garland Science New York.

WEISS, D., KOOPMANN, M., BASEL, T. and RUDACK, C., 2012. Cyclin A1 shows age-related expression in benign tonsils, HPV16-dependent overexpression in HNSCC and predicts lower recurrence rate in HNSCC independently of HPV16. *BMC Cancer,* **12**(1), pp. 259-268.

WHITEHURST, A.W., 2014. Cause and Consequence of Cancer/Testis Antigen Activation in Cancer. *Annual Review of Pharmacology and Toxicology,* **54**, pp. 251-272.

WISCHNEWSKI, F., FRIESE, O., PANTEL, K. and SCHWARZENBACH, H., 2007. Methyl-CpG binding domain proteins and their involvement in the regulation of the MAGE-A1, MAGE-A2, MAGE-A3, and MAGE-A12 gene promoters. *Molecular Cancer Research : MCR,* **5**(7), pp. 749-759.

WU, C. and BEKAII-SAAB, T., 2012. CpG island methylation, microsatellite instability, and BRAF mutations and their clinical application in the treatment of colon cancer. *Chemotherapy Research and Practice,* **2012**, 359041.

WU, L., LU, Y., WANG, X., LV, Z., ZHANG, B. and YANG, J., 2006. Expression of cancer-testis antigen (CTA) in tumor tissues and peripheral blood of Chinese patients with hepatocellular carcinoma. *Life Sciences,* **79**(8), pp. 744-748.

XU, G. and JAFFREY, S.R., 2013. Proteomic identification of protein ubiquitination events. *Biotechnology and Genetic Engineering Reviews,* **29**(1), pp. 73-109.

XU, Z., CETIN, B., ANGER, M., CHO, U.S., HELMHART, W., NASMYTH, K. and XU, W., 2009. Structure and function of the PP2A-shugoshin interaction. *Molecular Cell,* **35**(4), pp. 426-441.

YAMADA, T. and OHTA, K., 2013. Initiation of meiotic recombination in chromatin structure. *Journal of Biochemistry,* **154**(2), pp. 107-114.

YAMAMOTO, M., 2010. The selective elimination of messenger RNA underlies the mitosis-meiosis switch in fission yeast. *Proceedings of the Japan Academy.Series B, Physical and Biological Sciences,* **86**(8), pp. 788-797.

YANG, F., DE LA FUENTE, R., LEU, N., BAUMANN, C., MCLAUGHLIN, K. and WANG, P., 2006. Mouse SYCP2 is required for synaptonemal complex assembly and chromosomal synapsis during male meiosis. *Journal of Cell Biology,* **173**(4), pp. 497-507.

YANG, F., SKALETSKY, H. and WANG, P.J., 2007. *Ubl4b,* an X-derived retrogene, is specifically expressed in post-meiotic germ cells in mammals. *Gene Expression Patterns,* **7**(1), pp. 131-136.

YANG, F.F. and WANG, P.J.P., 2009. The Mammalian synaptonemal complex: a scaffold and beyond. *Genome Dynamics,* **5**, pp. 69-80.

YANG, Z. and SWEEDLER, J.V., 2014. Application of capillary electrophoresis for the early diagnosis of cancer. *Analytical and Bioanalytical Chemistry,* **10**, pp. 1-19.

YAO, F., SVENSJÖ, T., WINKLER, T., LU, M., ERIKSSON, C. and ERIKSSON, E., 1998. Tetracycline repressor, tetR, rather than the tetR-mammalian cell transcription factor fusion derivatives, regulates inducible gene expression in mammalian cells. *Human Gene Therapy,* **9**(13), pp. 1939-1950.

YAO, Y. and DAI, W., 2012. Shugoshins function as a guardian for chromosomal stability in nuclear division. *Cell Cycle,* **11**(14), pp. 2631-2642.

YEN, P.H., CHAI, N.N. and SALIDO, E.C., 1996. The human autosomal gene DAZLA: testis specificity and a candidate for male infertility. *Human Molecular Genetics,* **5**(12), pp. 2013-2017.

YOUDS, J.L. and BOULTON, S.J., 2011. The choice in meiosis - defining the factors that influence crossover or non-crossover formation. *Journal of Cell Science,* **124**(Pt 4), pp. 501-513.

YUAN, L., LIU, J., HOJA, M., WILBERTZ, J., NORDQVIST, K. and HOEOEG, C., 2002. Female Germ Cell Aneuploidy and Embryo Death in Mice Lacking the Meiosis-Specific Protein SCP3. *Science (Washington),* **296**(5570), pp. 1115-1118.

YUAN, L.L., LIU, J.G.J., ZHAO, J.J., BRUNDELL, E.E., DANEHOLT, B.B. and HÖÖG, C.C., 2000. The murine SCP3 gene is required for synaptonemal complex assembly, chromosome synapsis, and male fertility. *Molecular Cell,* **5**(1), pp. 73-83.

ZAKHARYEVICH, K., TANG, S., MA, Y. and HUNTER, N., 2012. Delineation of joint molecule resolution pathways in meiosis identifies a crossover-specific resolvase. *Cell,* **149**(2), pp. 334-347.

ZENDMAN, A.J., RUITER, D.J. and VAN MUIJEN, G.N., 2003. Cancer/testis-associated genes: Identification, expression profile, and putative function. *Journal of Cellular Physiology,* **194**(3), pp. 272-288.

ZHENG, J., 2013. Oncogenic chromosomal translocations and human cancer (Review). *Oncology Reports,* **30**(5), pp. 2011-2019.

ZHOU, Q., LI, Y., NIE, R., FRIEL, P., MITCHELL, D., EVANOFF, R.M., POUCHNIK, D., BANASIK, B., MCCARREY, J.R., SMALL, C. and GRISWOLD, M.D., 2008. Expression of stimulated by retinoic acid gene 8 (Stra8) and maturation of murine gonocytes and spermatogonia induced by retinoic acid *in vitro*. *Biology of Reproduction,* **78**(3), pp. 537-545.

ZICKLER, D. and KLECKNER, N., 1998. The leptotene-zygotene transition of meiosis. *Annual Review of Genetics,* **32**(1), pp. 619-697.

ZILLER, M.J., GU, H., MÜLLER, F., DONAGHEY, J., TSAI, L.T., KOHLBACHER, O., DE JAGER, P.L., ROSEN, E.D., BENNETT, D.A. and BERNSTEIN, B.E., 2013. Charting a dynamic DNA methylation landscape of the human genome. *Nature,* **500**(7463), pp. 477-481.