

Bangor University

DOCTOR OF PHILOSOPHY

Characterisation of human TDRD12 and LKAAEAR1 as potential oncogenic cancer testis antigen genes with clinical potential

Alsulami, Mishal

Award date:
2019

Awarding institution:
Bangor University

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.



**Characterisation of human *TDRD12* and
LKAAEAR1 as potential oncogenic cancer testis
antigen genes with clinical potential**

Ph.D. Thesis 2018

Mishal O. Alsulami

Abstract

Cancer is a highly complex disease that evolved in response to a wide range of biological and molecular changes that impact disease behaviour, treatment efficacy and clinical outcomes. Studying this diversity in human tumours is essential for gaining insights that will ultimately improve the survival rates of cancer patients. Cancer stem-like cells (CSCs) are believed to be responsible for invasive and metastatic features in tumours and can contribute to chemotherapy resistance and subsequent tumour relapses. There is an increasing need to identify the molecular mechanisms involved in tumour cells, particularly in CSCs. Cancer testis antigens (CTAs) are a subclass of germline proteins normally produced in immune-privileged sites, such as the testis, ovary and placenta of somatic tissues, and the presence of these antigens is increased in a variety of cancers. These characteristics make CTAs highly important immunotherapeutic targets, since they do not harness the immune response in the testes but encode immunogenic proteins that can induce a specific response in cancerous tissues. CTA genes are potentially very importance in clinical applications, including cancer diagnosis, vaccination and immunotherapy. This current study focused on the investigation of two CTAs, *TDRD12* and *LKAAEAR1*, that may have an enhanced presence in cancer and the potential to be immunogenic.

TDRD12 is linked to stemness features and enables the proliferation of germ line tumour cells. It appears to act as a possible transcriptional regulator for germline factors that are essential to cell cycle proliferation, germ cell maintenance and stem marker expression. *TDRD12* may have the potential to drive oncogenesis and CSC targets.

LKAAEAR1 was validated as a CTA at the protein level, showing its production was restricted to germ cells and the central nervous system from normal tissues and showed aberrant production in a wide range of tumours. This protein has been shown to be produced in germ cells undergoing spermatogenesis with strong nuclei staining, suggesting its potential role in this process. *LKAAEAR1* potentially acts as a regulator for transposable elements, thereby increasing its contributions to cancer development. This study demonstrated that *LKAAEAR1* could potentially be used as a cancer biomarker and therapeutic target.

Acknowledgement

I am thrilled to have had the opportunity to complete this work. I would like to thank ALLAH for giving me the power and patience to complete this important step in my educational and professional life.

I owe my deepest gratitude to my supervisor Dr Ramsay McFarlane. Without his continuous optimism concerning this work, and his enthusiasm, guidance and support throughout the research and during the writing of this thesis, this study would probably have never been completed. Nobody could have wished for better supervision. Special thanks and appreciation to Dr Jane Wakeman for sharing her experience, constructive advice and firm support during my study.

I am grateful to all members of the McFarlane laboratory, past and present, who have shared scientific knowledge and technical support as well as good company with me along the way, particularly to, Dr Natalia Gomez-Escobar and Dr Ellen Vernon as well as Dr David Pryce who was always already to offer friendly and generous discussion. Thanks and gratitude extend also to Dr Jana Jezcova, Dr John Sammut, and Dr Jason Williams for their effort in IHC staining and results assessment as well as Dr Ali Hazzazi for his assistance with processing hESCs.

I am deeply grateful to Afrah Alsulami, as a wife and colleague, for her constant love and endless support, encouragement and patience. I wish a bright future and successful journey during her PhD. My love to you and our child, Aziz.

Last but not least, it gives me great pleasure to acknowledge the help and support of my family. I am very grateful to my parents, brothers and sisters for their unwavering belief in me and for supporting my dreams and my goals.

Without these people, I could never have completed this project. I really appreciate their help, most deeply and sincerely.

Thank you all.

Mishal

List of Abbreviations

°C	Degrees Centigrade
3'	Three prime end of DNA
5'	Five prime end of DNA
<i>ACRBP</i>	acrosinbinding protein
AEs	Axial elements
ALDH	Aldehyde dehydrogenase
ALT	alternative lengthening of telomere
ASC	Adult stem cells
ATP	Adenosine-5'-triphosphate
BCA	Bicinchoninic acid
bp	Base pair
BRCA1	Breast cancer susceptibility 1
BRCA2	Breast cancer susceptibility 2
BTB	Blood-testis barrier
CD	Cluster of Differentiation
CDK	Cyclin dependent kinase
cDNA	Complementary DNA
CE	Central element
CNS	central nervous system
CpG	-Cytosine-phosphate-guanine-
CRC	Colorectal Cancer
CSC	Cancer Stem Cell
CT	Cancer-testis
CTA	Cancer testis antigen
<i>CTAG1B</i>	Cancer/testis antigen 1B
C-terminal	Carboxy-terminal domain
CTLs	Cytotoxic T lymphocytes
Da	Dalton
DAPI	4'-6-diamidino-2-phyindole
dH ₂ O	Distilled water
DMC1	Disrupted meiotic cDNA1
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTPs	Four deoxynucleoside triphosphates
DPBS	Dulbecco's phosphate-buffered
DSB	Double-strand break
ECACC	European Collection of Cell Cultures
ECM	extracellular matrix
EGC	Embryonic germ cells
ELDA	Extreme limiting dilution analysis
EMT	Epithelial-mesenchymal transition
ES/ESC	Embryonic stem cells
F	Forward
FACS	Fluorescence-activated cell sorting
FBS	Foetal bovine serum
FFPE	Formalin-fixed, paraffin-embedded
g	Gram
G0	Quiescent phase of the cell cycle

G1	Gap-1 phase
G2	Gap-2 phase
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GBM	Glioblastoma multiforme
GTE _x	Genotype-Tissue Expression
Hela	Henrietta Lacks
HERVs	human endogenous retroviruses
hESC	human embryonic stem cells
HLA	Human leukocyte antigen
HR	Homologous recombination
ICM	inner cell mass
IF	Immunofluorescent staining
IHC	Immunohistochemistry
IP	Immunoprecipitation
iPSC	Induced Pluripotent Stem Cells
IVF	in vitro fertilization
kb	Kilobase
KDa	Kilodalton
L	Litter
LB	Luria bertani
LSC	leukemic stem cells
M	Mitosis phase
mA	Milliamperes
MAGE-A	Melanoma antigen family A
MET	Mesenchymal-epithelial transition
MHC	Major histocompatibility complex
ml	Milliliter
M-PER	Mammalian protein extraction reagent
mRNA	Messenger RNA
MS	Mass spectrometry
MSCs	mesenchymal stem cells
MW	Molecular weight
NCBI	National Centre for Biotechnology Information
ng	Nanogram
nM	Nanomolar
NRT	Non reverse-transcriptase template
NTC	Non-template control
N-terminal	Amino-terminal domain
ORF	Open reading frame
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PFA	Paraformaldehyde
PGC	primordial germ cells
pH	Power of hydrogen
piRNA	piwi-interacting RNA
PVDF	polyvinylidene difluoride
qRT-PCR	Quantitative, real time PCR
R	Reverse
RB	Retinoblastoma-associated gene
RNA	Ribonucleic acid
RNAi	RNA interference

RT	Room temperature
RT-PCR	Reverse transcription PCR
S	Synthesis-phase
SAGE1	Sarcoma antigen 1
SC	Synaptonemal complex
SCCs	Squamous cell carcinomas
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEREX	Serological analysis of recombinant cDNA expression libraries
siRNA	Small interfering RNA
SMC1 β	Structural Maintenance of Chromosomes 1 β
SP	side population
SSC	spermatogonial stem cells
STAG3	Stromal Antigen 3
SW480	Colon adenocarcinoma
SYCP1	Synaptonemal complex protein 1
TAA	Tumour associated antigens
TBE	Tris-borate-EDTA
TBS	Tris buffered saline
TCGA	The Cancer Genomic Atlas
TE	Transposable Element
TFs	Transcription factors
TIC	Tumour/cancer initiating cells
TM	Melting temperature
TME	Tumour microenvironment
Tris	Tris (hydroxymethyl) aminomethane
V	Volts
WB	Western Blot
WCE	Whole cell extract
X-CT	X- chromosome cancer testis genes
μ g	Microgram
μ L	Microliter

List of content

Abstract	I
Acknowledgement	II
Declaration and Consent	III
List of Abbreviations.....	VI
List of content	IX
List of Figures	XIII
List of tables	XVI
1. Introduction	2
1.1 Human Cancer.....	2
1.1.1 Cancer overview.....	2
1.2. The causes of cancer	3
1.3. Cancer Biomarkers	3
1.4. Hallmarks of cancer	4
1.5. Cancer as a stem cell disease.....	6
1.2. Stem cells	7
1.2.1. Stem cell overview	7
1.2.2. Stem cell types and properties	7
1.2.3. Regulation of pluripotency	11
1.2.4. Cancer stem cells (CSC)	12
1.2.5. The germline genes activation is common in embryonic and cancer cells.....	16
1.3. Cancer/Testis antigens.....	19
1.3.1. Overview	19
1.3.2. CTA genes expression in normal cells	20
1.3.3. The role of CTAs in cancer cells.....	24
1.3.4. Genome stability and CTAs	28
1.4. The use of CTAs in in clinical applications.....	32
1.4.1. The use of CTAs in diagnosis and prognosis approaches.....	32
1.4.2. Vaccination.....	32
1.4.3. CTAs in cancer immunotherapeutic.....	33
1.5. CTAs identification	35
1.6. Studied genes.....	37
1.6.1. TDRD12	37

1.6.2. LKAAEAR1 (C20ORF201) gene.....	37
1.7. The aim of project.....	38
1.5.1. TDRD12	38
1.5.2. LKAAEAR1.....	39
2. Materials and Methods.....	41
2.1 Human cancer cell culture	41
2.1.1 Human cancer cell line sources.....	41
2.1.2. Cell culture growth maintenance.....	41
2.1.3. Thawing of stored cancer cell lines.....	41
2.1.4. Cell harvesting and freezing preparation.....	43
2.1.5. Cell counting	43
2.2. siRNA transfection experiment.....	43
2.3. RNA extraction and cDNA synthesis:	44
2.4. Qualitative PCR	45
2.4.1. Qualitative RT-PCR	45
2.4.2. Primers design for RT-PCR:	46
2.4.3. Gel purification of PCR products for sequencing.....	46
2.5. Quantitative Real-time PCR	47
2.6. Protein Extraction:	49
2.6.1. Whole-cell lysate.....	49
2.6.2. Subcellular Protein Fractionation	49
2.6.3. Normal lysates sources	50
2.7. Western Blot	51
2.8. Cell growth and proliferation assay in tissue culture	53
2.8.1. Growth curve analysis.....	53
2.8.2. Cell viability experiment	53
2.9. Cell immunofluorescence staining.....	53
2.10. Immunohistochemistry.....	54
2.11. Differentiation experiment.....	56
2.12. Flow Cytometry (FACS)	56
2.13. Immunoprecipitation	56
3. Functional analysis of human TDRD12 in germ line tumour cells.....	59
3.1. Introduction	59
3.1.1. TDRD Family genes.....	59
3.1.2. TDRD12	60

3.1.3. The aim of the work in this chapter	62
3.2 Results	62
3.2.1 Knockdown of TDRD12 in germ line tumour cells, NTERA2.	62
3.2.2. TDRD12 depletion has an influence on NTERA-2 cell proliferation rate.	66
3.2.3. Cell cycle analysis of TDRD12-depleted NTERA-2 cells.	68
3.2.4. TDRD12 knockdown influences cell cycle regulatory proteins.	71
3.2.7. TDRD12 knockdown does not induce NTERA-2 cell apoptosis	73
3.2.8. TDRD12-depleted NTERA2 cells does not induce senescence.....	75
3.2.9. Double knockdown of <i>CDKN1A</i> and <i>TDRD12</i> rescues proliferation potential in NTERA2 cells.	76
3.2.10. <i>TDRD12</i> is involved in complexes of genes that regulate cancer stem cell stability.	79
3.2.11. <i>TDRD12</i> knockdown regulates transposon element <i>LINE-1</i> in NTERA2 cells.	82
3.3. Discussion.....	83
3.3.1. Characterisation of TDRD12 splice variants and antibody specificity.....	83
3.3.2. TDRD12 regulates cancer stem cell proliferation.	83
3.3.2. TDRD12 regulates other germ genes in NTERA2 cells.	86
3.3.3. <i>TDRD12</i> regulates the transposon element <i>LINE-1</i>	87
3.3.4. Conclusion.....	88
4. The potential roles of TDRD12 gene in human embryonic stem cells.....	90
4.1. Introduction	90
4.2. Results.....	92
4.2.1. Analysis of TDRD12 expression in human embryonic stem cells (hESCs) and induced pluripotent stem cells (iPSCs).	92
4.2.2. The expression pattern of stemness marker genes in <i>TDRD12</i> -depleted NTERA2 cells	94
4.2.3. <i>TDRD12</i> expression during NTERA2 cell differentiation.	95
4.2.4. The expression patterns of <i>TDRD12</i> upon differentiation of human embryonic stem cells (hESCs).....	100
4.2.5. Assessment of <i>EXDL1</i> gene expression in human embryonic stem cells.....	102
4.2.6. Analysis of <i>EXDL1</i> expression profile in human normal tissues and cancer cells.	104
4.2.7. Analysis using qRT-PCR of <i>EXDL1</i> expression in normal and cancerous human tissues and cell lines.....	106
4.3. Discussion.....	109
4.3.1. <i>TDRD12</i> gene is a potentially linked to stemness features of CSCs and hESCs.	109
4.3.2. <i>EXDL1</i> is a potential cancer biomarker.	110
4.3.3. Conclusion.....	111

5. Functional analysis of human LKAAEAR1 gene in normal and cancer tissues.	113
5.1. Introduction	113
5.2. Results.....	116
5.2.1. The characterisation of the LKAAEAR1 protein in normal and cancer tissues	116
5.2.2. Depletion of <i>LKAAEAR1</i> in cancer stem-like cells and breast cancer cells.....	117
5.2.3. <i>LKAAEAR1</i> expression during NTERA2 cell differentiation.	120
5.2.4. The expression patterns of <i>LKAAEAR1</i> upon differentiation of human embryonic stem cells (hESCs).....	125
5.2.5. Assessment of cell growth and proliferation following the depletion of <i>LKAAEAR1</i>	126
5.2.6. <i>LKAAEAR1</i> regulates transposon elements (TEs) in cancer cells.	129
5.2.7. The localisation of LKAAEAR1 proteins in cancer cells.	132
5.2.8. The localization of LKAAEAR1 protein in testis tissue.....	135
5.3. Discussion.....	138
5.3.1. LKAAEAR1 protein is a Testis/CNS CTA	138
5.3.2. The expression of <i>LKAAEAR1</i> in normal hESCs and CSCs.....	138
5.3.3. The <i>LKAAEAR1</i> gene may regulate TEs genes	139
5.3.4. LKAAEAR1 protein localisation in cancer cells	140
5.3.5. Summary	142
6. Analysis of LKAAEAR1 protein in tumour samples	144
6.1. Introduction	144
6.2. Results.....	146
6.2.1. Staining pattern of LKAAEAR1 protein in normal testis and CNS tissues.	146
6.2.2. Staining pattern of LKAAEAR1 in normal tissues sections from the microarray.	148
6.2.3. Staining pattern of LKAAEAR1 protein in tumour tissues sections.....	153
6.2.4. Distribution of LKAAEAR1 protein in epithelial ovarian cancer.	162
6.2.5. Staining pattern of LKAAEAR1 protein colorectal carcinoma progression microarray.....	164
6.3. Discussion.....	166
7. Final discussion and future directions	170
8. References.....	174

List of Figures

Figure 1.1 The Hallmarks of cancer.....	5
Figure 1.2. The postulated manners of stem cell division.....	8
Figure 1.3. A totipotent state during the embryo development.....	9
Figure 1.4. The hypotheses of how CSC was originated.....	14
Figure 1.5. Malignant cells share common features of primordial germ cells (PGC).....	18
Figure 1.6. Cross section of seminiferous tubules.....	22
Figure 1.7 oncogenic functions of CTAs.....	25
Figure 1.8 Examples of CT proteins and CT mRNA that contribute to various pro-tumoral mechanisms.....	27
Figure 1.9. Methods used to identify a new CT gene candidate.....	36
Figure 3.1. The potential roles of TDRD12 in the PET complex.....	61
Figure 3.2. The schematic representation of the four potential transcript variants (TVs) of TDRD12 that are coding protein.....	64
Figure 3.3. Analysis of TDRD12 knockdown efficiency in NTERA2 cells by qRT-PCR and Western blot.....	65
Figure 3.4. Cell growth curve for NTERA2 cells depleted for TDRD12 using three sequences.....	67
Figure 3.5. NTERA2 cell growth curve for six days of transfection with anti-TDRD12 siRNA #2.....	69
Figure 3.6. The influence of TDRD12 depletion on cell cycle profile.....	70
Figure 3.7. Western blot analysis, demonstrating changes in cell cycle protein levels following TDRD12 protein knockdown in NTERA2 cells.....	72
Figure 3.8. Western blot analysis demonstrating the production pattern of apoptotic proteins following TDRD12 protein knockdown in NTERA2 cells.....	74
Figure 3.9. Senescence analysis of <i>TDRD12</i> -depleted NTERA2 cell lines.....	75
Figure 3.10. Cell proliferation curve and microscopic examination for NTERA2 cell proliferation following double knockdown of <i>TDRD12</i> and <i>CDKN1A</i>	77
Figure 3.11. qRT-PCR analysis for mRNA levels of <i>TDRD12</i> and <i>CDKN1A</i> following double knockdown experiment.....	78
Figure 3.12. The expression pattern of <i>TDRD12</i> , <i>EXDL1</i> and <i>PIWIL</i> family genes in NTERA2 cells.....	80
Figure 3.13. Predicted interactions between TDRD12, EXDL1 and PIWIL family proteins in human...81	
Figure 3.14. qRT-PCR analysis of transposon elements (<i>LINE</i> and <i>SINE</i>) following <i>TDRD12</i> depletion in NTERA2 cells.....	82
Figure 4.1. Expression profiles of <i>TDRD12</i> and <i>SOX2</i> in stem cells and iPSCs using qRT-PCR analysis.93	
Figure 4.2. Pattern of stem cell marker gene transcripts in human <i>TDRD12</i> -depleted NTERA2 cells using qRT-PCR analysis.....	94
Figure 4.3. Analysis of the <i>OCT4</i> and <i>TDRD12</i> transcript levels in undifferentiated NTERA2 cells using qRT-PCR.....	96
Figure 4.4. Analysis of the <i>OCT4</i> and <i>TDRD12</i> transcript levels in DMSO treated NTERA2 cells using qRT-PCR.....	97
Figure 4.5. Analysis of the <i>OCT4</i> and <i>TDRD12</i> transcript levels in RA-treated NTERA2 cells using qRT-PCR.....	98

Figure 4.6. Analysis of the <i>OCT4</i> and <i>TDRD12</i> transcript levels in HMBA-treated NTERA2 cells using qRT-PCR.....	99
Figure 4.7 qRT-PCR analysis of <i>TDRD12</i> transcript levels in H9 cells treated with retinoic acid.....	101
Figure 4.8 qRT-PCR analysis of <i>EXDL1</i> transcript levels in H9 cells treated with retinoic acid.....	103
Figure 4.9. RT-PCR analysis of the expression profile of the <i>EXDL1</i> gene in normal human tissues..	104
Figure 4.10. RT-PCR analysis for the expression profile of <i>EXDL1</i> in cancer human tissues and cell lines.....	105
Figure 4.11. qRT-PCR analysis of <i>EXDL1</i> gene expression levels in a range of normal human tissues.....	107
Figure 4.12. qRT-PCR analysis of <i>EXDL1</i> gene expression levels in a range of cancerous human tissues.....	108
Figure 5.1 <i>LKAAEAR1</i> is localised on chromosome 20q13.33 with two predicted splice variants.....	115
Figure 5.2. Analysis of the <i>LKAAEAR1</i> protein in human tissue extracts using Western blot.....	116
Figure 5.3. NTERA2 cells depleted of <i>LKAAEAR1</i> using different siRNAs.....	118
Figure 5.4. MCF7 cells depleted of <i>LKAAEAR1</i> using different siRNA targets.....	119
Figure 5.5. Analysis of the <i>OCT4</i> and <i>LKAAEAR1</i> transcript levels in untreated NTERA2 cells using qRT-PCR.....	121
Figure 5.6. Analysis of the <i>OCT4</i> and <i>LKAAEAR1</i> transcript levels in DMSO treated NTERA2 cells using qRT-PCR.....	122
Figure 5.7. Analysis of the <i>OCT4</i> and <i>LKAAEAR1</i> transcript levels in RA-treated NTERA2 cells using qRT-PCR.....	123
Figure 5.8. Analysis of the <i>OCT4</i> and <i>LKAAEAR1</i> transcript levels in HMBA-treated NTERA2 cells using qRT-PCR.....	124
Figure 5.9 qRT-PCR analysis of <i>LKAAEAR1</i> transcript levels in H9 cells treated with retinoic acid.....	125
Figure 5.10. Analyses of NTERA2 cell proliferation following the depletion of <i>LKAAEAR1</i> transcripts.....	127
Figure 5.11. Analyses of MCF7 cell proliferation following the depletion of <i>LKAAEAR1</i> transcripts.....	128
Figure 5.12. TE mRNA expression levels in MCF7 cells depleted of <i>LKAAEAR1</i> transcripts using qRT-PCR analysis.....	130
Figure 5.13. TE mRNA expression levels in NTERA2 cells depleted of <i>LKAAEAR1</i> transcripts using qRT-PCR analysis.....	131
Figure 5.14. Cellular localisation <i>LKAAEAR1</i> protein in NTERA2 cells.....	133
Figure 5.15. Cellular localisation <i>LKAAEAR1</i> protein in MCF7 cells.....	134
Figure 5.16. IHC staining of <i>LKAAEAR1</i> protein in normal testis section.....	136
Figure 5.17. Negative control staining of testis tissue section by IHC analysis and H&E staining.....	137
Figure 5.18. Sub-nuclear localisation of <i>LKAAEAR1</i> protein in germ cells and breast cancer cells....	141
Figure 6.1. IHC staining of seminiferous tubules and central nervous system selected sections by anti- <i>LKAAEAR1</i> antibody (abcam: #ab108142).....	147
Figure 6.2. IHC staining of <i>LKAAEAR1</i> protein on normal breast, colon, ovarian epithelium, stomach and prostate sections.....	149
Figure 6.3. Example of normal tissues sections from CTHN that stained anti- <i>LKAAEAR1</i> antibody..	150
Figure 6.4. <i>LKAAEAR1</i> staining on normal tissues (control) and normal adjacent to tumour tissues (NAT).....	154
Figure 6.5. <i>LKAAEAR</i> protein staining in prostate tumours.....	157

Figure 6.6. LKAAEAR protein staining in breast tumours.....	158
Figure 6.7. LKAAEAR protein staining in thyroid and bladder tumours.....	159
Figure 6.8. LKAAEAR protein staining in normal and cancerous lung tissues.....	160
Figure 6.9. IHC staining shows LKAAEAR1 protein distribution in normal and different types of epithelial ovarian cancers samples with polyclonal anti-LKAAEAR1 antibody (ab108142).....	163
Figure 6.10 IHC staining of tissue microarray for colorectal cancer tissues (CRC) progression cancer samples by anti-LKAAEAR1 antibody (ab108142).....	164
Figure 6.11. Relative levels of LKAAEAR1 staining detected in CRC progression microarray.....	165

List of tables

Table 1.1. CSC markers for distinct solid tumour types.....	16
Table 2.1. Growth conditions and description of human cell lines used in this study.....	42
Table 2.2. GeneSolution siRNA list.....	44
Table 2.3. primers used in PCR experiments.....	46
Table 2.4. List of commercial primers.....	48
Table 2.5. Designed primers.....	48
Table 2.6. The source of human cell lysates.....	50
Table 2.7. The primary antibodies and its optimized concentrations for Western blot analyses.....	52
Table 2.8. Secondary Antibodies with its optimal concentrations for Western blot analyses.....	52
Table 2.9. Secondary antibodies and optimized dilutions used in IF assay.....	54
Table 2.10. The primary and secondary antibodies that used in IHC method.....	56
Table 3.1. Panel of cell cycle proteins selected to study their responses following TDRD12 depletion in NTERA2 cells.....	71
Table 6.1. Summary of IHC analysis of LKAAEAR1 staining on normal tissues.....	151
Table 6.2. Summary of LKAAEAR1 staining assessment of tumours in the TMA.....	161
Table 6.3. Showing the estimated staining factor for each group cases in CRC progression	165

Chapter 1: Introduction

1. Introduction

1.1 Human Cancer

1.1.1 Cancer overview

Cancer comprises a large group of diseases which account for approximately 14.1 million new diagnosed cases and about 8.2 million deaths worldwide per year. It is responsible for the increasing mortality rates in both economically developed and developing countries (Cavalieri et al., 2002; Siegel et al., 2013). In general, tumours can be classified as benign or malignant. Benign tumours are normally harmless masses with a slow growth rate that usually remain in the original tissue and do not spread to organs. In rare cases, a few types of benign tumours can be life-threatening, depending on their location. Malignant tumours have the potential to migrate and invade local as well as more distant tissues by moving through the lymphatic and circulatory systems in a process called metastasis. Malignant tumours are dangerous and life-threatening, and in contrast to benign tumours, they are usually difficult to treat with surgical intervention and other local treatments (Cooper & Hausman, 2000; Alison et al., 2002). Cancer can be further categorised into four main classes, depending on the cells of origin. These classes include carcinomas, sarcomas, leukaemias and lymphomas (Roy et al., 2017). Carcinomas initiate in the epithelial cells of the respective organ (for example, cancers of the lung, breast, pancreas, prostate and colon), and this type of malignancy comprises about 90 percent of all human cancers. Sarcomas are very rare malignancies that originate in the connective tissues such as bones, cartilage, muscles and fibrous tissues. Leukaemias and lymphomas are those malignancies which arise from blood cells and immune cells, respectively, accounting for about 8 percent of human cancers (Park et al., 2016).

Cancer arises due to genetic/epigenetic alterations and mutations that accumulate in human cells. Over the course of time, these alterations change the cellular characteristics and lead to disruption of control systems within cells which are critical to protecting the cells against uncontrolled proliferation and overgrowth (Loeb et al., 2003; Roy et al., 2017). Cancer cells acquire distinctive capabilities which allow them to divide and self-renew to form tumours through a multistep process called tumorigenesis. Hanahan and Weinberg have described the distinctive features of cancer cells as “hallmarks of cancer,” and they have defined the attainment of distinct biological capabilities during the time of human cell development as being a multistep process (Hanahan & Weinberg, 2000).

1.2. The causes of cancer

To date, many studies suggest that the incidence of cancer is the outcome of environmental (exogenous) and genetic/epigenetic factors (endogenous) (Clapp et al., 2008; Flavahan et al., 2017). The purpose of the concept of environmental factors, which are known as carcinogens, is to elucidate all non-genetic factors such as lifestyle, diet and infections which are responsible for more than 80 percent of cancers. Exposure to carcinogens (for example, air pollution, radiation and chemicals) is implicated in the development of human cancer. A great deal of epidemiological data suggests that approximately 30 percent of cancer-related deaths result from smoking tobacco, and about 35 percent are due to diet-related anomalies, while viral infections account for almost 15 percent. Remarkably, there is evidence that economically developed countries can decrease the rate of cancer to about 50 percent by preventing smoking, limiting alcohol consumption, increasing physical activity and encouraging weight control (Katzke et al., 2015; Weiderpass, 2010; Anand et al., 2008). Cancer-causing genetic factors can not only be induced by environmental factors, but can be inherited or can result from DNA replication errors (Tomasetti et al., 2017). Today, it is well established that cancers arise due to molecular aberrations at the genetic and epigenetic levels that can disrupt gene expression patterns, which can result in elevations/reductions in regulatory protein/RNA molecules. Many cancer-related genes are involved in tumorigenesis as they are vital for cell proliferation, apoptosis and DNA repair in normal cell physiology. Those genes are classified into tumour suppressor genes and proto-oncogenes (Lodish et al., 2000; Roy et al., 2017). The tumour suppressor genes are normally growth inhibitory, and their aberration leads to uncontrolled cell division and initiates tumours. In contrast, the proto-oncogenes are growth-promoting genes, and expression activation may promote tumorigenesis (Lodish et al., 2000; Devi, 2004; Luo et al., 2009). Oncogenic genetic alterations can include mutations, chromosomal rearrangement, aneuploidy, gene amplification and overexpression (Sarasin & Kauffmann, 2008). Epigenetic alterations include histone modifications, such as, phosphorylation, methylation and acetylation, as well as DNA methylation and demethylation (Sharma et al., 2010; Novak, 2004).

1.3. Cancer Biomarkers

Cancer biomarkers are defined as biological molecules detected in blood or other body fluids or tissues that are indicative of tumour conditions, progressions, and treatment responses

(Mishra & Verma, 2010). Many factors can result in the presentation of cancer biomarkers, including transcriptional alterations, somatic or germline mutations, and post-translational modifications. Cancer biomarkers are categorised into different types, including, proteins (enzyme and receptors), nucleic acids (e.g., DNA, RNA and microRNA), antibodies, and peptides (Henry & Hayes, 2012; D'Andrea et al., 2017). The clinical assessment of cancer patients is dependent on these biomarkers because they can be used to estimate the risk of disease, to screen the early incidence of cancer, to differentiate between benign and malignant tumours, to distinguish tumour types from one another, to identify the prognosis and prediction of cancer onset, to detect recurrence, and to demonstrate the progression of the treatment (Burstein et al., 2011).

1.4. Hallmarks of cancer

Hanahan and Weinberg (2000; 2011) have defined the hallmarks of cancer as the acquired functional alterations in normal cell physiology which evolve progressively into the disease. In the vast majority of human cancers, the cancer cells acquire a succession of these functional alterations that enable tumour cells to proliferate, survive and disseminate through various mechanisms and at different times in a multistep process. Hanahan and Weinberg initially proposed six essential hallmarks: sustaining proliferative signalling, enabling replicative potential, evading growth suppressors, resisting cell death, inducing angiogenesis, and tissue invasion and metastasis (Hanahan & Weinberg, 2000). In 2011, the same authors suggested that the acquisition of tumorigenesis is possibly facilitated by two additional enabling characteristics and two additional emerging characteristics which have been added to the group of hallmarks (Figure 1.1). The enabling characteristics include the genomic instability and tumour-promoting inflammatory state of premalignant lesions and malignant cells (Hanahan & Weinberg, 2011). Also, the conceptual progress of research and pathological analyses in the last decade led to the proposal of two additional hallmarks of cancer that are functionally important for the pathogenesis of perhaps all cancers (Colotta et al., 2009; Luo et al., 2009; Negrini et al., 2010). As these new capabilities are still under investigation and validation, they are termed as “emerging hallmarks” and include the deregulation of cellular energy metabolism and avoiding immune destruction (Hanahan & Weinberg, 2000; 2011; Roy et al., 2017). Although the hallmarks of cancer are accepted and can be detected in any carcinogenic process, this simplistic view of cancer has been criticised

by Sonneschein and Soto (2013). These authors criticised all the list of characteristics, for example, they discussed the idea that the cancer is not only a cell-base or genetic disorders that cause uncontrolled cell proliferation, but it is viewed as a tissue-based disease where the interaction among cells through a complicated signalling system and the role of their microenvironment in tumorigenesis (Sonnenschein & Soto, 2013).

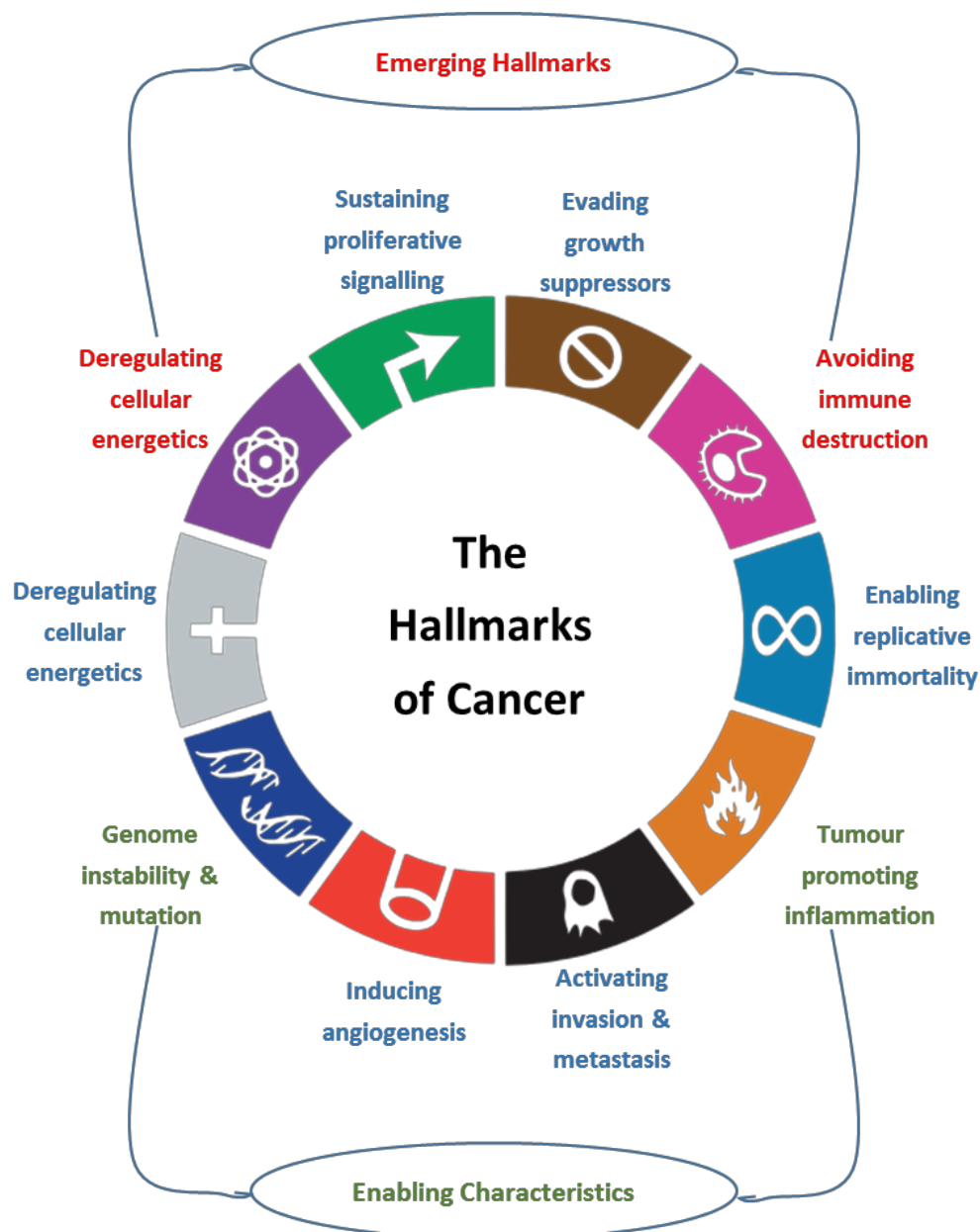


Figure 1.1 The Hallmarks of cancer. The diagram represents the hallmarks of cancer, the picture is adopted from (Hanahan & Weinberg, 2011).

1.5. Cancer as a stem cell disease

Cancer is now known as a stem cell disease. The stem cell characteristics of some cancer cells are considered as key factors for their progression and proliferation (Colak & Medema, 2014; Kaseb et al., 2016; Gedye et al., 2016; Aponte & Caicedo, 2017). Tumours are characterized by an unrestrained progression of malignant cells that are diverse in functions and phenotypes. The heterogeneity of malignant cells within the tumour is explained by two different models. The stochastic model proposes that each single cell, or at least the majority of malignant cells in a tumour, has a similar ability to grow a tumour. According to this model, the tumour initiating cells acquire genomic instability through the sequential somatic mutations or by tumour micro-environmental factors, following subsequent clonal selection. A second model is the cancer stem cell (CSC) model. It hypothesises that a minority of cell populations within the tumour have the potential for self-renewal and the generation of heterogeneous cells, which form the bulk of the tumour. According to this model, the tumour cells are organized in a hierarchy, and the apex of the pyramid is the location of the cancer initiation cells/cancer stem cells. The concept of CSCs depends on the idea that normal stem cells are in normal tissues, and CSCs are present in tumour tissues (Cabrera et al., 2015). Although the ancestry of CSCs is still controversial, the cumulative evidence supports that the CSCs are originally normal stem cells, and that their transformation occurred as the result of mutations or due to the deregulation of genome programs and stability (Welch et al., 2012). The acquisition of genomic instability through mutations allows the stem cells to transform the properties of quiescence and tightly-regulated phenotypes to become overactive stem cells. Indeed, CSCs show some similarities of functional characteristics with normal stem cells as they are able to self-renew to initiate tumour formation in addition to their non-stopped generation of the differentiated progenies that form the heterogeneous tumour bulk (Lau et al., 2017). Moreover, increasing evidence has shown that CSCs contribute to cancer metastasis, chemotherapy resistance and recurrence (Moore & Lyle, 2011; Gedye et al., 2016; Aponte & Caicedo, 2017).

1.2. Stem cells

1.2.1. Stem cell overview

The concept of stem cells first emerged in 1877, when Ernst Haeckel used it to describe the fertilized egg as a single cell, which is also the origin of the multicellular organism. The term “stem cell” was then rapidly adopted to demonstrate the entity of all cell types in the body (Maehle, 2011). Stem cells are characterized by two main features: unlimited self-renewal and the capacity for differentiation into multiple lineages. The self-renewal characteristic of stem cells is essential for maintaining themselves and for regenerating tissues as well. The capacity for differentiation allows them to produce specialised cells that are important for specific organ or tissue function.

There are two types of normal stem cells—embryonic stem cells and adult or somatic stem cells (Reya et al., 2001; Gao, 2008; Verga Falzacappa et al., 2012). Embryonic stem cells (ES) can produce all the somatic and germ cell lines in the organism and are the progenitors of all three of the embryonic germ layers: the ectoderm, endoderm and mesoderm (Lovell-Badge, 2001; Donovan & Gearhart, 2001). The main function of somatic stem cells is to repair the damaged parts of tissues and organs. Thus, the latest studies reveal the potential power of stem cells to be used in the treatment of some diseases through regenerative medicine (O'Brien & Barry, 2009).

1.2.2. Stem cell types and properties

The main characteristic of stem cells is their capacity for infinite self-renewal, by which they can divide symmetrically to produce two identical stem cells, or divide asymmetrically when they produce two daughter cells: one cell is destined to replace the mother stem cell, and the other is destined to undergo the differentiation process (Figure 1.2), (Kapinas et al., 2013).

The differentiation potential is another essential characteristic that is restricted to stem cells (Alison et al., 2002). Generally, stem cells vary in their ability to differentiate or give rise to the descendant cell types. The most primitive stem cells are the fertilized oocyte and the subsequent first two divisions of cells which are known as totipotent stem cells. These cells possess the potential to produce all the undifferentiated cells in the organism, including the embryonic, extra embryonic and trophoblastic cells in the placenta.

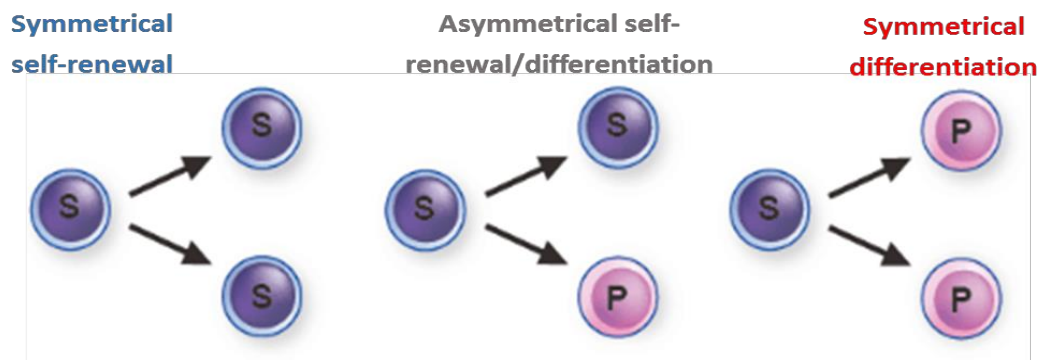


Figure 1.2. The postulated manners of stem cell division, this picture adapted from laboratory of developmental stem cells biology and regenerative medicine, IMSUT.

After the eight-cell stage, these totipotent stem cells begin to differentiate into more specialized cells—the blastocyst and inner cell mass (ICM) from which the foetus develops (Keller et al., 2010; Condic, 2013). The cells of the ICM are considered to be pluripotent stem cells and can give rise to all the types of cells that generate from the three germ layers (the endoderm, mesoderm and ectoderm) but not to new embryos because they lack the potential to give rise to a placenta (Keller et al., 2010). In a continuous process of differentiation, these pluripotent stem cells become more specialized when they can differentiate into multipotent stem cells that have the potential to produce a limited range of functional cell lineages appropriate to the specific tissue or organ. Some examples of multipotent stem cells are central nervous system (CNS) stem cells that possess the trilineage potential to produce neurons, oligodendrocytes and astrocytes, while the stem cells in the small intestine can generate four lineages (Paneth, goblet, absorptive columnar and enteroendocrine cells). Finally, the least potent stem cells are called unipotent stem cells or committed progenitors. These are able to generate at least one type of cells of the tissue in where they are located. For example, the epidermal stem cells that are located in the basal layer can only generate keratinized squamous cells. However, the multipotent and unipotent stem cells are also defined as adult stem cells (Overturf et al., 1997; Alison et al., 2002; Surani & Tischler, 2012). To date, the stem cells can be classified in different categories based on the origin of the cells: embryonic stem cells (ESCs), induced pluripotent stem cells (iPSCs), adult stem cells (ASCs) and mesenchymal stem cells (MSCs) (Gao, 2008; National Institutes of Health, 2009; Park et al., 2014).

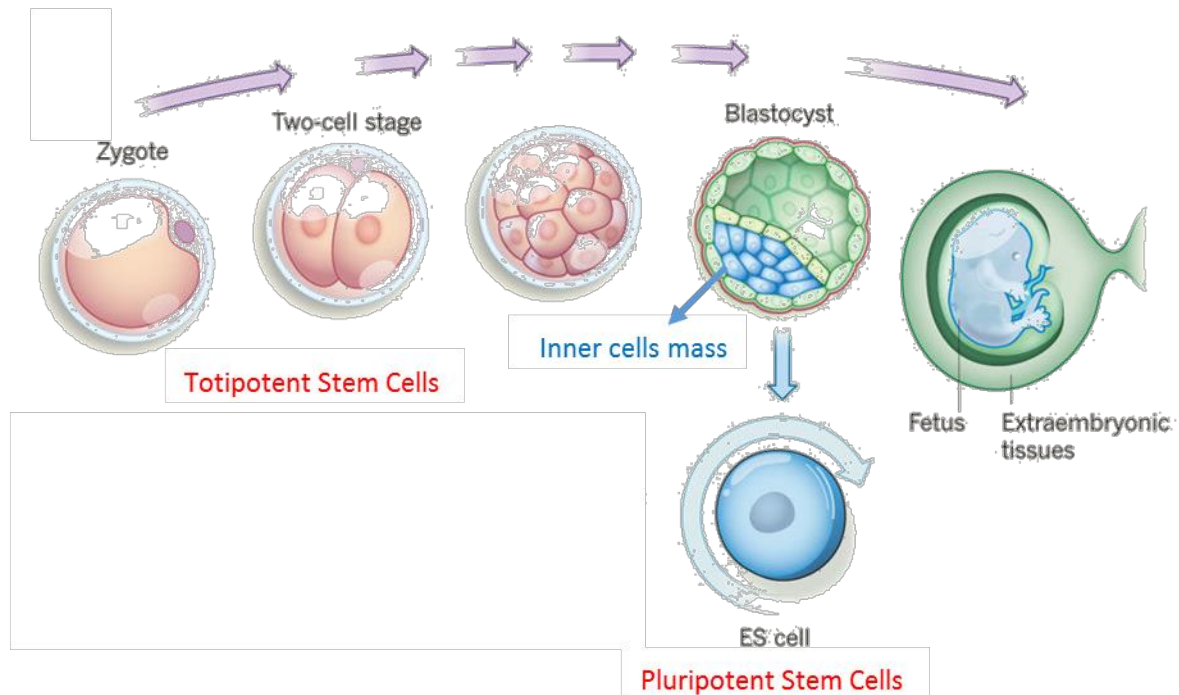


Figure 1.3 A totipotent state during the embryo development: The first two divisions of fertilized egg are totipotent stem cells which differentiate into blastocyst. The ES in the ICM are pluripotent stem cells (Surani & Tischler, 2012).

1.2.2.1. Human embryonic stem cells (hESCs)

Clinical trials have shown that human embryonic stem cells (hESCs) are highly promising for treating complicated diseases and regenerative medicine (Ali et al., 2016; Cao & Zhang, 2017). In 1998, Thomson was the first to produce hESCs from an inner cell mass (ICM) using *in vitro* fertilization (IVF) technology (Thomson et al., 1998). The isolated colonies of hESCs can retain their undifferentiated state as well as their ability to give rise to all three germ layers in a process called gastrulation. Each germ layer (ectoderm, mesoderm and endoderm) can differentiate into specific tissues. For example, ectoderm germ layers can become the external lining of the body, the epidermis and its appendages (hair and nails) and the nervous system. Mesoderm germ cells can develop into cartilage, bone and connective tissues including muscles (heart smooth muscle, kidney, gonads and genital ducts). Endoderm germ layers generate the epithelial internal lining of the gastrointestinal and respiratory tracts, pancreas, liver, thymus, thyroid, urinary bladder and urethra (Damdimopoulou et al., 2016). The hESCs that Thomson isolated in 1998 were grown as colonies on irradiated mouse fibroblasts (MEFs) and have been used in research for decades, as these hESCs are considered

the 'gold standard'. Therefore, efforts have been made to advance the protocols for clinical approaches; however, some considerable caveats must be considered before implementing the clinical use of hESCs, including ethical concerns (Wert & Mummery, 2003), immunogenicity and the increasing possibility of teratocarcinoma formation from undifferentiated colonies (Cao & Zhang, 2017).

1.2.2.2. Induced pluripotent stem cells (iPSCs)

Induced pluripotent stem cells (iPSCs) are derived from somatic adult cells that have been reprogrammed to pluripotent stem cells. This breakthrough in the field of stem cells overcomes the obstacle of immune-rejection using hESCs in clinical trials by using autologous cells. In 2007, the first iPSCs were derived from both mouse and human adult fibroblasts to embryonic stem-like cells through transduction of four transcriptional factors: OCT4, SOX2, KLF4 and c-MYC (Takahashi et al., 2007). The derived iPSCs resemble hESCs in morphology, proliferation and gene expression, especially in stemness markers (OCT4, Nanog & SOX2) and surface antigens (SSEA3, SSEA4, TRA1-60 & TRA 1-81). Moreover, these cells showed their ability to generate cells of all three germ layers. Their self-renewal and differentiation capacity allow iPSCs to provide patient- and disease-specific autologous cells for transplantation (Matsa et al., 2011). Thus, iPSCs constitute excellent *in vitro* models of disease (Matsa et al., 2011), human organ transplantation, tissue repairs, organ synthesis and blood cell formation (Takebe et al., 2013; Park et al., 2014; Singh et al., 2015). Finally, the use of iPSC derivation technology does not require the use of hESCs and thus circumvents the ethical issues involved in the use of hESCs (Cao & Zhang, 2017).

1.2.2.3. Adult stem cells ASCs

ASCs or somatic stem cells are the self-renewing cells in all tissues and organs in the body; ASCs have the ability to generate precursor cells, which develop into specialized cells for general homeostasis in the respective tissue (Slack, 2008; Tweedell, 2017). ASCs are controlled by highly-regulated asymmetric cell divisions to regulate their number in tissues as well as their essential role in organ maintenance function, tissue repair and regeneration (Alison et al., 2002). Many studies reveal that ASCs *in vivo* are usually harboured in a specialized microenvironment called a "stem cells niche" (Rumman et al., 2015). The population of ASCs is a mixture of cells: some exist in a quiescent state for long periods until they receive a re-activation signal, while others are progenitor cells in different stages of

differentiation. The surrounding cells of the niche regulate the stem and progenitor cells, prevent the depletion of ASCs, protect from physical damage and act as supportive tissues (Spradling et al., 2001; Sell, 2004; Scadden, 2006). Although ASCs reside in different niches in various tissue types in the body, they share common properties including self-renewal abilities, undifferentiated state as their capacity of differentiation, genomic repair ability, long cell cycling, micro-environmental protection and support from the niche (Alison et al., 2002; Aponte & Caicedo, 2017). The majority of ASCs are often multipotent cells that generate uni-potent progenitors for terminal differentiated cells. Examples of ASCs are hematopoietic stem cells, mesenchymal stem cells, neural stem cells and intestinal stem cells.

1.2.3. Regulation of pluripotency

The pluripotency of stem cells is governed by many mechanisms which are still not fully understood. Thus far, many studies have found that self-renewal and differentiation are maintained and controlled by a group of key transcription factors: OCT4, SOX2 and NANOG. These transcription factors are used to induce pluripotency in the reprogrammed somatic cells. OCT4 (octamer binding transcription factor-4) is encoded by the *POU5F1* gene and its expression is activated at an early stage (the four- to eight-cells stage) of embryonic development (Zeineddine et al., 2014). OCT4 is a core transcription factor that plays an essential role in the establishment and regulation of ICM pluripotency (Nichols et al., 1998). The down-regulation of OCT4 in human ESCs *in vitro* results in a marked change in morphology, a decrease in growth rate and a clear reduction in the expression of cell surface markers (Matin et al., 2004). Cells depleted for OCT4 also showed a rapid increase in markers associated with differentiation (Niwa et al., 2000). Moreover, RNAi-induced silencing of *POU5F1* is found to change the expression of more than 1000 genes in both positive (pluripotency TFs) and negative (differentiation-associated genes) regulation of many gene sets (Babaie et al., 2007). It has been reported that OCT4 interacts with SOX2, a member of the SOXB1 transcription factor family, in order to regulate the pluripotency of ESCs (Adachi et al., 2010). NANOG is another transcription factor that is present in the ICM of blastocyst stages but not in the early embryonic stages; thus, NANOG plays an important role in enhancing pluripotency. Finally, OCT4, SOX2 and NANOG were demonstrated to regulate the pluripotency of stem cells by transcriptionally activating its genes and/or transcriptionally inactivating the genes that promote differentiation such as *PAX6* and *HOXB1* (Rodda et al.,

2005). Stem cells are also regulated by other transcription factors such as STAT3, GATA2 and FOXD3 (Young, 2011; Krishnakumar et al., 2016; Ye et al., 2016).

1.2.4. Cancer stem cells (CSC)

Malignant cells develop various aspects of stemness, meaning they are unable to retain tissue homeostasis and thus sustain the progression and development of cancer disease (Kaseb et al., 2016). The resemblance of stemness features in SCs and malignant cells provide the cornerstones for the maintenance and survival of cancer cells, starting from the self-renewal and differentiation capacities to the stemness-supporting microenvironment (Visvader, 2011; Gedye et al., 2016; Visvader & Clevers, 2016). So, CSCs are a small group of cells within the tumour that have stemness characteristics that allow the progression of cancer such as self-renewing, unstopped growth, metastasizing and continuous proliferation. CSCs have important, organised features that also allow them to contact neighbouring cells to provide nutrients and contribute in the evasion of the immune system by creating an environment which is suitable for tumour development. The population of CSCs in the microenvironment is composed of heterogeneous cells, often with high plasticity, high resistance to stressful factors such as malnutrition and the ability to be quiescent as a common response to chemotherapy (Moore & Lyle, 2011; Chen et al., 2016). Moreover, while the normal SC population is organised to be protected within a niche, CSCs are also protected and supported by the tumour microenvironment (TME). This TME not only consists of cancer cells but also contains immune cells, mesenchymal cells, endothelial cells and myofibroblasts (Raggi et al., 2016). TME was found to be mainly supported and maintained by the neighbouring mesenchymal stem cells (MSCs). The MSCs also protect the TME during stress and form an immune-privilege barrier to regulate the microenvironment (Cortez et al., 2014; Kfoury & Scadden, 2015; Aponte & Caicedo, 2017).

1.2.4.1. The origin of CSCs

Several studies have been conducted to determine an effective treatment for cancer. These studies focused on the suggestion that only a small subset of cells within the tumour have a key role in tumour initiation and are considered a core origin of tumorigenesis. This subpopulation of cells are termed as CSCs (Kreso & Dick, 2014). The existence of CSCs has been verified by many studies in some types of cancers by detecting molecular markers of CSCs. The first evidence of CSCs was provided in leukaemia when it was confirmed that

CD³⁴⁺CD³⁸⁻ leukemic cells show haematopoietic stem cell features in bone marrow (Bonnet & Dick, 1997; Lapidot et al., 1994). In solid cancers, CSCs were detected in breast cancer (CD⁴⁴⁺CD^{24-/low} Lin⁻); thereafter, the investigation was broadened to confirm the presence of CSCs in many other types of solid tumours including tumours of the pancreas, brain, liver, ovary, colon, skin, prostate and lung (Table 1) (Al-Hajj et al., 2003; Medema, 2013). To date, it is widely accepted that the pathological signs of a deteriorating clinical prognosis are closely linked to the presence of CSCs in tumours. The recurrence of tumours, metastasis and resistance to chemotherapies all relate to the characteristics of CSCs (Bao et al., 2006; Jeon et al., 2011). Other studies also focused on the role of tumour microenvironments (TME) such as hypoxia, perivascular and invasive niches in the development and maintenance of CSCs (Brooks et al., 2015; Sun et al., 2012; Plaks et al., 2015). Recent research has shown the role of TME in tumour plasticity, which is a dynamic transition: epithelial mesenchymal transition (EMT) and mesenchymal epithelial transition (MET) (Friedl & Alexander, 2011; Zabala et al., 2016). The plasticity of tumour has been challenged by work demonstrating that trans-differentiation through cellular fusion process between stem cells and somatic cells may be the derivative of CSCs (Bjerkvig et al., 2005; Zabala et al., 2016). The plasticity is not only responsible for the generation of CSCs but may also be connected to the reconstitution of tumours as well as the capacity of invasion and metastasis (Chen et al., 2014; Eun et al., 2017). There are two representative concepts to explain the origin of CSCs in tumours: one suggested that the CSCs originate from specific, mutated adult stem cell or their progenitors, while another postulates that somatic cells become CSCs through a multistep process (Figure 1.4) (Plaks et al., 2015; Eun et al., 2017).

1.2.4.2. Characterisation, identification, isolation and drug targets for CSCs

It is widely known that CSCs share common stemness signalling with normal stem cells such as Hedgehog, JAK/STAT, Notch, WNT/ β -Catenin and NF κ B (Chen et al., 2013). They are vital for tumour progression and development as well as maintaining stemness characteristics or controlling the differentiation properties during numerous developmental processes (Eun et al., 2017). Recently, many studies have shown that the activation of these signals orchestrates stem cell plasticity in cancerous and normal tissues (Sirko et al., 2013; Yan et al., 2014). Moreover, tumours use the exposure to appropriate stemness signals to establish the cellular hierarchy, while these signals enhance the dedifferentiation mechanisms in normal stem cells (Yan et al., 2014; Eun et al., 2017). More recent papers have also found that the EMT process

is the reprogramming process that occurs the most in physiological conditions and plays an essential role in embryogenesis and other developmental processes; this refers to the ability of mesenchymal type cells to migrate through the extracellular matrix (ECM) (Hay, 1995).

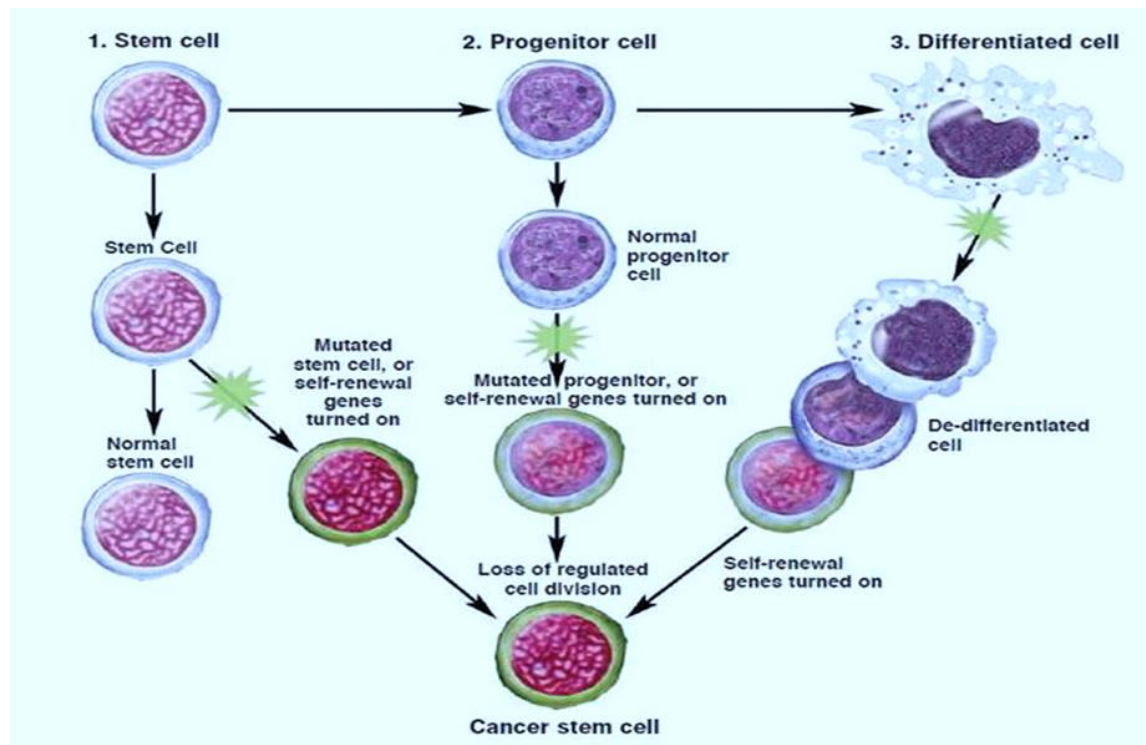


Figure 1.4 The hypotheses of how CSC was originated. It is suggest that they might be generated from mutated stem cells, mutated committed progenitors, or mutation of somatic cells. Taken from (the National Institutes of Health (NIH) resource for SC research, 2015).

Interestingly, this phenomenon takes place in both cancer and normal cells, and cells that undergo the EMT process may acquire many stem cell phenotypes (Mani et al., 2008). Also, the EMT process and CSCs share many transcription factors such as ZEB1/2 HIFs and Twist and have signalling pathways identical to those of Notch, Hedgehog and WNT/ β -catenin (Ouyang et al., 2010). Although CSCs trigger the activation of stemness-signalling pathways, the yield of alteration of gene expression pattern was found to be directly regulated by these TFs (Li et al., 2011; Eun et al., 2017). The transcriptional factors OCT4, NANOG and SOX2 were found to be upregulated in CSCs (Saygin et al., 2016).

It is difficult to identify CSCs from a large bulk of cells within a tumour; however, this identification is important for understanding the role of CSCs in tumour initiation and

progression as well as for effectively diagnosing and treating cancer (Liu & Lathia, 2016; Agliano et al., 2017). The functional identification of CSCs can only be performed depending on the features of self-renewal and differentiation capacities. However, nowadays it has become possible to identify CSCs using specific surface markers and enzymes (Table 1) (Chen et al., 2012; Medema, 2013; Ajani, et al. 2015; Agliano et al., 2017). The CD133, CD24 and CD44 cell surface markers are used to identify and isolate CSCs from different types of tumours (Al-Hajj et al., 2003; Dawood et al., 2014; Liu & Lathia, 2016; Agliano et al., 2017). For example, Al-Hajj showed that the ESA^+ , $CD44^+$ and $CD24^{-/low}$ are surface markers on CSCs extracted from breast cancer (Al-Hajj et al., 2003). The leukemic stem cells (LSC) can be isolated from haematopoietic stem cells by their expression of the $CD34^+$ and $CD38^-$ as cell surface marker phenotypes (Dick, 2008). Other investigations on LSCs demonstrate the expression of other surface markers such as CD^{123} , CD^{47} , CD^{96} and TIM3 (Wang et al., 2017). CSCs from colon cancer (Chen et al., 2017) and breast cancer (Al-Dhfyan et al., 2017) can be identified and isolated based on ALDH enzymatic activity using ALDEFLOUR assay. In addition, isolation of CSCs can be performed using FACS analysis, which uses DNA dye to depict a side population (SP) of low-dye or negative dye. This SP can be easily distinguished from the sorted cells (Boesch et al., 2014). The candidate CSCs can be further analysed to validate their functional stemness properties using sphere forming assay or xenograft assay (Lathia & Liu, 2017). Therefore, a clonogenic assay can be used to identify the drug target for CSCs, and this assay is adapted for conducting high throughput screens on CSCs (Mathews et al., 2012). Many studies on CSCs result in the possibility to be safely used in clinical approach, particularly in developmental agents that eradicate or control the growth of the CSCs population, which could produce more effective therapeutic targets for cancers. For instance, the gene-silencing method was used to detect and control the expression of *STAT3*, which is an important gene in maintaining cancer stemness (Li et al., 2015b). In an *in silico* screen, the inhibition of the STAT3 pathway was used to identify a drug to be utilized in a CSC target. This drug is now under development for use in the clinics (Li et al., 2015b; Lathia & Liu, 2017).

Table 1.1. CSC markers for distinct solid tumour types. This is list of the majority of the molecular markers tested on CSCs but the expression level of scrutiny differ per marker. Some are extensively studied whereas others are tested only on cell lines. The markers are not order according to their importance. Adapted from (Medema, 2013).

Tumour	Most examples of CSC marker
Breast	ALDH1, CD24, CD44, CD90, CD133, Hedgehog-Gli activity and α 6-integrin
Colon	ABCB5, ALDH1, β -catenin activity, CD24, CD26, CD29, CD44, CD133, CD166, and LGR5
Glioma	CD15, CD90, CD133, α 6-integrin and nestin
Liver	CD13, CD24, CD44, CD90, and CD133
Lung	ABCG2 ALDH1 CD90 CD117 CD133
Melanoma	ABCB5, ALDH1, CD20, CD133, and CD271
Ovarian	CD24, CD44, CD117, and CD133
Pancreatic	ABCG2, ALDH1, CD24, CD44, CD133, c-Met, CXCR4, Nestin, and Nodal-Activin
Prostate	ALDH1, CD44, CD133, CD166, α 2 β 1-integrin, α 6-integrin, and Trop2

1.2.5. The germline genes activation is common in embryonic and cancer cells

Immortalization, lack of adhesion, migration, demethylation, downregulation of major histocompatibility complex (MHC) and activation of germline genes are the most characteristics shared by tumour cells and cells undergoing embryogenesis/trophoblast development. Indeed, this has led to the historical proposal that called embryonal theory of cancer. Over a century ago, John Beard suggested the oldest theory of how cancer cells arise which is the “Trophoblast theory of cancer”. In this theory, Beard predicted that cancers result from germ cells that fail to complete their embryonic migration to gonads (Simpson et al., 2005; Akers et al., 2010). Although this prediction was quite far ahead of its time, it has inspired many studies that support an association between germ-cell development and cancer. For instance, one of the first pieces of evidence to support this hypothesis was the observation that a range of human tumours produce the trophoblast hormone, gonadotropin (Simpson et al., 2005; Akers et al., 2010). Today, the release of gonadotropin in somatic cells is a prognostic for epithelial cancer staging (Louhimo et al., 2004). Another piece of supporting

evidence of this theory is the identification of protein antigens group that are produced only in trophoblasts, germ cells and tumours. Thus, the aberrant expression of germline genes in somatic cells leads to the release of trophoblast hormone in tumours (Simpson et al., 2005; Whitehurst, 2014). However, the simplicity of the embryologic theory of cancer has been questioned.

McFarlane and Wakeman discussed the restrictive views of this theory (McFarlane & Wakeman, 2017). In brief, some considerable differences in these core features between cancer cells progression and embryonic cells development. As the embryonic cells during normal physiological development execute these features in highly regulated and controlled pathway as well these cells avoid excessive mutations and maintain proper ploidy. However, these normal constraints that regulate embryonic cells are deviated in cancer cells such as avoiding apoptosis and alternative lengthen of telomere pathway. The new findings of centromeric polarity, altered DNA repair and chromosomal end protection are all potentially regulated by pseudo-meiotic functions that provide solid evidence of distinct oncogenic genome evolution. The cancer cells genomic is reconstituting over time in response of tumour cells requirement in distinct microenvironment. So, the tumour cells are most likely to undergo soma-to-germ transition which can encompass germ-like and embryo-like development but does not mimic the actual embryogenesis in zygote (McFarlane & Wakeman, 2017). In line with this, many recent studies supported that cancer cells may transform to more germ-like state (Janic et al., 2010; Wang et al., 2011; McFarlane et al., 2014; McFarlane & Wakeman, 2017). Also, the findings that the cancer cells possess ability to self-renew and undergo phenotypic changes lead to the postulate that a cellular soma-to-germ transition is a key feature of oncogenesis (McFarlane et al., 2014; Feichtinger et al., 2014; McFarlane et al., 2015; Nassar & Blanpain, 2016; Planells-Palop et al., 2017, McFarlane & Wakeman, 2017). This is supported by the findings that a large group of germline genes are activated during oncogenesis, showing the essential role of these genes in tumour development (Janic et al., 2010; Fagegaltier et al., 2016; Sumiyoshi et al., 2016). Inactivation of many germline genes leads to suppression of tumours (for example, see (Planells-Palop et al., 2017). Additional analyses of gene expression pattern show that the similarity of germline genes activation is apparently observed in both human tumours and germ cells (Figure 1.5) (Feichtinger et al., 2014; Simpson et al., 2005). Finally, all the observations that establish parallels between the development of germ cells and tumour initiation are currently leading

the emergence of new field in cancer biology, the study of cancer testis antigens (CTA) (Figure 1.5). The unique pattern of CT gene expression makes them more attractive targets to be used as cancer biomarkers (in prognosis and diagnosis), and therapeutic targets (Whitehurst, 2014).

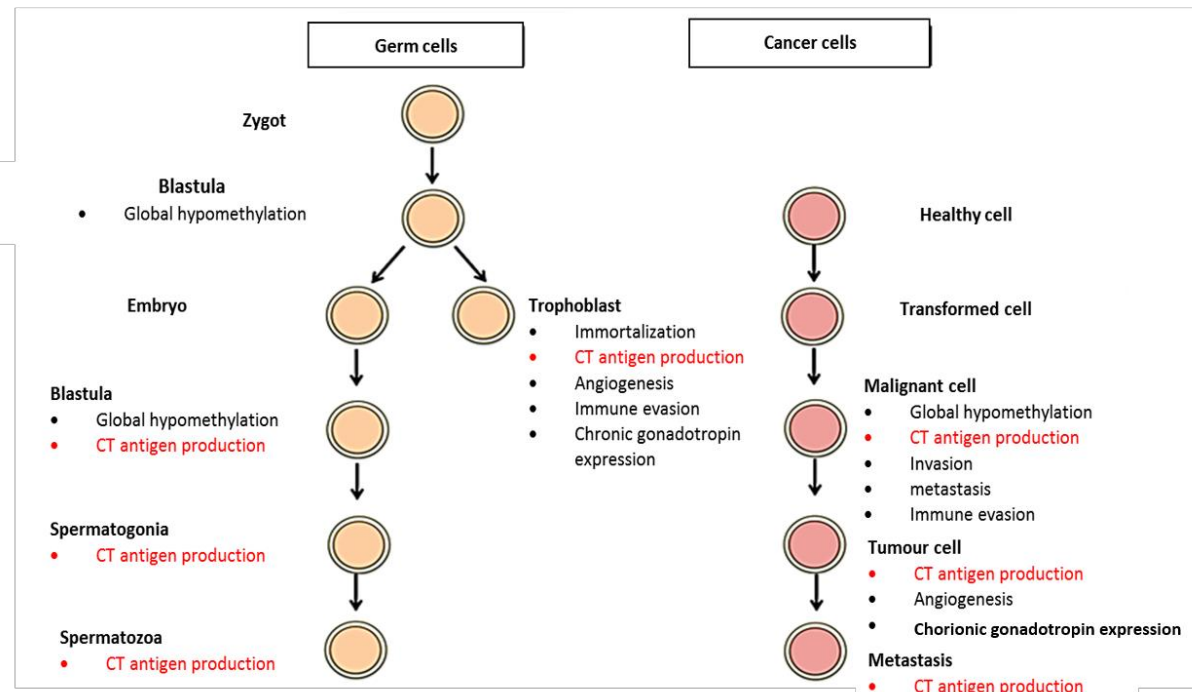


Figure 1.5. Malignant cells share common features of primordial germ cells (PGC). Germ cells (orange) and cancer cells (red) have common phenotypes include CT genes expression, global demethylation, chorionic gonadotropin hormone release, and immune evasion. This figure is adapted from (Simpson et al., 2005).

1.3. Cancer/Testis antigens

1.3.1. Overview

Cancer/Testis (CT) genes are a major group of germline genes that encode cancer testis antigen (CTA) proteins. These antigens were named 'germline proteins' (Chen et al., 1998). CTAs are native germline proteins with production normally restricted to germ cells of testicular and ovarian and/or placental tissues, and are also aberrantly presented in somatic cells in a wide range of human malignancies. In early 1990s, Van der Bruggen and co-workers identified the first CTA gene, called *MAGE-1* (melanoma antigen-1) which evokes a spontaneous cytolytic T lymphocyte response using autologous typing of melanoma cells in a patient (van der Bruggen et al., 1991). Subsequently, *NY-ESO-1* and *SYCP1* genes were also discovered (Chen et al., 1997; Tureci et al., 1998). The limitations of autologous typing technology led to the invention of SEREX (serological analysis of recombinant tumour cDNA expression libraries with autologous serum). The successful cloning of *SSX* and *NY-ESO-1* allowed a series of SEREX studies on several tumours to identify various CTAs including *SCP-1*, *MAGE-C1*, *HOM-TES-85*, *cTAGE-1*, *CAGE* and *OY-TES-1* (Chen et al., 1997; Chen et al., 1998; Li et al., 2004). Collectively, all these antigens are commonly characterised by their activation in tumours (Whitehurst, 2014). The tissue exclusive fashion of CTA production prompted for additional CTA gene identification through mRNA expression profiling. Additionally, the restricted expression of CT genes to testis and cancers, but not other healthy tissues were investigated using different techniques such as bioinformatics analysis, cDNA oligonucleotide array analysis, and representational difference analysis (RDA). To date, over 100 CT genes families with different members and splice variants have been reported using different techniques (Almeida et al., 2008; Pagotto et al., 2013).

The presence of CTAs is commonly identified in several tumours; however, the levels can be variable amongst tumour types. For instance, low levels of CTAs were detected in lymphoma, leukemia, colon, and pancreatic tumours, whereas high levels can be found in melanoma, ovarian, and lung cancers. In addition, certain CTA genes show expressions that is specific to tumour type; for example, 70% of melanomas tumours were found to express at least one gene of the *MAGE* gene family (Brasseur et al., 1995).

CTAs have been originally classified into those that are encoded by genes on the X-chromosome as CT-X genes and those that have their genes distributed throughout the autosomes as non-X-CT genes. Non-X CT genes encode proteins that are found to be

functionally contributing in either spermatogenesis process or in spermatids such as, *ACRBP* (acrosinbinding protein), *ADAM2*, meiotic proteins (*SCP1*, *HORMAD1* and *SYCE*) (Tureci et al., 1998; Ono et al., 2001; Chen et al., 2005). The majority of the non-X genes are activated during or after meiosis but not in the pre-meiotic stage. According to CT gene expression profiles in normal tissues, they can be further categorised into four groups. First, those that are normally expressed in the testis and sometimes in the placenta are named ‘testis-restricted genes’. The second group involves the genes that are also expressed in the central nervous system (CNS) as well as adult male testis, which are called ‘testis/CNS-restricted genes’. The third group includes genes whose expressions is found in the adult male testis with no more than two other selective tissues, other than the CNS; these are known as ‘testis-selective genes’. The fourth group, which is called ‘testis/CNS-selective genes’, includes genes that are expressed in testis, CNS, and no more than two additional tissues. Hofmann and co-workers classified about 153 CT genes into three groups, including testis-restricted, testis-selective and testis/CNS-restricted (Hofmann et al., 2008). The association of CT genes in meiosis and their abnormal expression in cancers have led to the postulate that these proteins may have, to some extent, contributions in the chromosomal instability in cancer cells; for example the occurrence of aneuploidy and trisomy (McFarlane & Wakeman, 2017).

1.3.2. CTA genes expression in normal cells

1.3.2.1. The role of CT genes in spermatogenesis:

CTA genes which, are responsible for controlling and maintaining germline, have many critical and distinct functions in spermatogenesis (Whitehurst, 2014). To date, the role of CT genes in gametogenesis is not fully understood, however, gene expression and knockout studies shed light on the distinct functions of these genes (Whitehurst, 2014). Some CT genes such as *MAGEA1* and *NY-ESO-1* were found to be expressed during early stages of spermatogenesis suggesting that they might contribute in spermatogenesis initiation (Jungbluth et al., 2000; Whitehurst, 2014). Some CT proteins (such as sp17) were detected in late stages of spermatogenesis including spermatocytes to sperm production stages (Cheng et al., 2011). GAGE protein is thought to play a critical function in germ cells as it is found in Oct4+ primordial germ cells (Gjerstorff et al., 2007). Furthermore, gene-targeting studies on a group of CTA genes showed that these CTAs are essential for male fertility. For example, mice that are deficient in a single CTA commonly demonstrate attenuated fertility (de Vries et al., 2005;

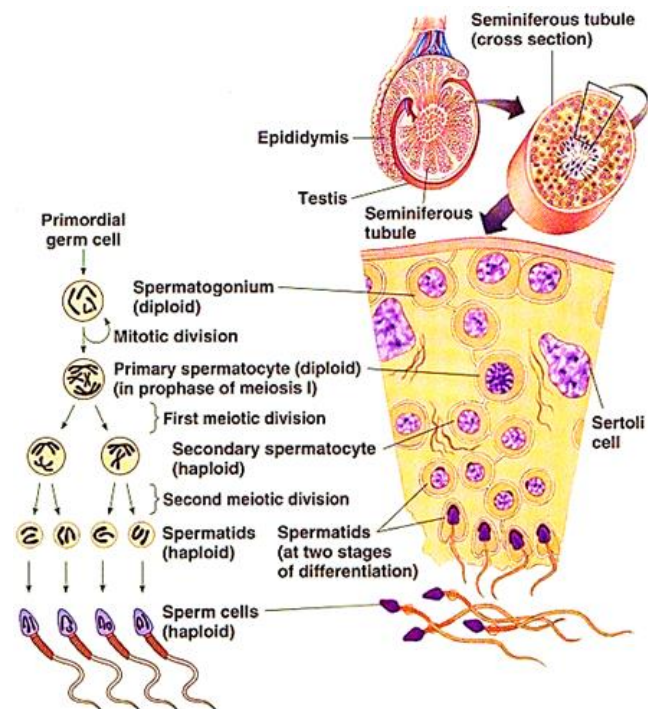
Bolcun-Filas et al., 2009; Schramm et al., 2011).

Spermatogenesis is a physiological process in seminiferous tubules within male testes that produces spermatozoa (Lie et al., 2013). Seminiferous tubules are structurally composed of two compartments: basal and adluminal that are separated by blood/testis barrier (BTB) (Mital et al., 2011). These structures are protected by connective tissues containing Leydig cells, which produce testosterone hormone. Each seminiferous tubule epithelium is composed of germ cells and Sertoli cells. The germ cells are located at the basal membrane and are considered as progenitors for spermatogenesis or spermatogonial stem cells (SSC). These SSCs are thought to undergo self-renewal to maintain the number of SSCs in the microenvironment or differentiate into spermatocytes and then to spermatids (Figure 1.6). The hierarchical structures can be visually seen in the basal compartments when the large spermatogonia are near to the basal membrane while the differentiated spermatids are located close to the lumen. Sertoli cells play critical roles that facilitate the progression and differentiation process in spermatogenesis as well as their physical support to the tissues. Interestingly, Sertoli cells form the BTB barrier between the two compartments of the seminiferous tubules to prevent the invasion of large molecules such as protein, and lymphocytes (Mital et al., 2011). This barrier segregates the spermatogenesis event, including meiosis I and meiosis II, in a special microenvironment away from the host immune system. Thus, BTB confers the immune privilege site in the testis (Figure 1.6) (Lie et al., 2013).

1.3.2.2. The role of CTAs in meiosis

Meiosis is a unique feature of gametogenesis that drives genetic diversity through the production of genetically different haploid cells (Nielsen & Gjerstorff, 2016). Germ cells that commit to spermatogenic differentiation lack proliferative features and enter meiosis, a process highly orchestrated by a range of specific proteins and complexes (Figure 1.6 A). Some CTA proteins have been shown to drive and maintain the physiological molecular events in meiosis such as homologous recombination, synaptonemal complex formation, alignment of sister chromatid, and DNA double-strand break (DSB) formation (McFarlane & Wakeman, 2017). In meiosis, homologous recombination is responsible for the exchange of genetic diversity through the crossing over of chromosome. This process is initiated by DSBs that are induced by a topoisomerase-like protein, SPO11, which is a CTA.

A)



B)

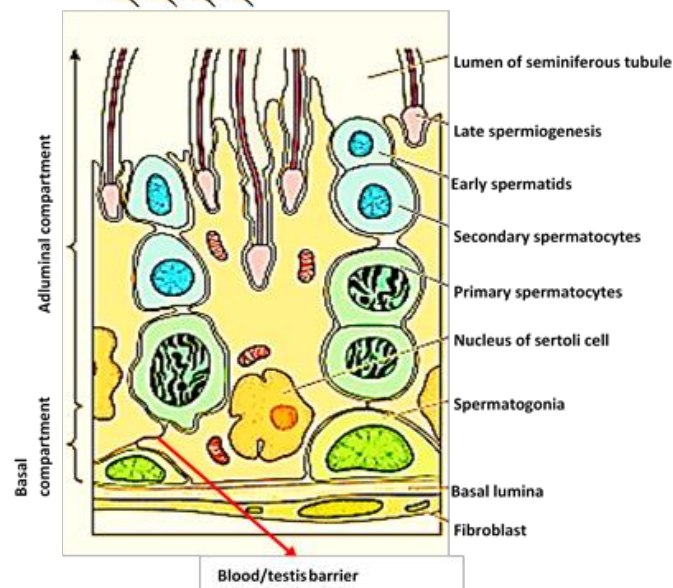


Figure 1.6. cross section of seminiferous tubules

(A) A hierarchical structure spermatogenesis and its central element, meiosis, during the differentiation steps from spermatogonia to sperm production. Diploid cells stop mitotic division and enter meiotic division and finally produce haploid sperms cells.

(B) The Sertoli cells facilitate the differentiation process of spermatogonia stem cells into more specialised cells, physical support to the tissue and provide immune privilege site for SSC (blood/testis barrier)(Holstein et al., 2003; Lui & Cheng, 2013).

Synaptonemal complex (SC) is a meiosis-specific complex that functions to synapse homologous chromosomes during meiosis prophase-I and is required for the formation of meiotic crossover (Cahoon & Hawley, 2016). The SC is a multiprotein structure that built up by two lateral elements formed along the axes of the homologous chromosomes that consist of (SCP2 and SCP3) proteins, the central elements consists of (SYCE1-3 and TEX12), and transverse filaments of SCP1 protein that connect the lateral with central elements (Nielsen & Gjerstorff, 2016). Additionally, the lateral elements of the SC are linked with HORMAD-proteins as well as meiotic cohesins. In human, misregulation of the SC is associated with miscarriage, infertility and birth defects (Schramm et al., 2011; Nagaoka et al., 2012). Importantly, many of the proteins that constitute the SC are aberrantly activated in tumours. The cohesin protein complex has critical roles in tethering sister chromatids during mitosis, but it is also necessary for meiosis. Depletion of the meiotic-specific cohesin component genes such as, *Rec8* and *Rad21* in mice, were found to disturb proper synapses function (Ishiguro & Watanabe, 2007; Llano et al., 2012). The constitution of cohesin complexes involved in mitosis and meiosis differs, and Structural Maintenance of Chromosomes 1 β (SMC1 β), Stromal Antigen 3 (STAG3), REC8 and RAD21L are all subunits unique to meiosis-specific cohesin complexes (Lee et al., 2003; Lee & Hirano, 2011). These genes are also expressed in different types of cancer (Feichtinger et al., 2012; Rosa et al., 2012; Lindsey et al., 2013; Strunnikov, 2013).

Furthermore, HORMAD1 has multiple functions during meiotic recombination such as facilitating formation of the SC, promotion homologue alignment and homologue search by ensuring DSBs are repaired via an inter-homologue pathway. HORMAD1 is aberrantly expressed in breast, lung and colon cancers (Robert et al., 2016; Vrielynck et al., 2016).

Meiosis Specific with OB Domains (MEIOB) was recently identified as a DNA binding exonuclease which is important in meiotic recombination and it is also detected in many tumours, such as, lung cancer (Luo et al., 2013; Souquet et al., 2013; Shiohama et al., 2014; Nielsen & Gjerstorff, 2016). An additional examples of important CTAs in meiosis process are: Testis Expressed 15 (TEX15) and Disrupted Meiotic cDNA1 (DMC1) are proteins that normally involved in DSB repairing formed during meiosis, however, they are abnormally present in melanoma. Recent work found that *Tex19.1* in mouse is normally involved in meiotic recombination-initiation events (Crichton et al., 2017). However its human orthologue *TEX19* is a CT gene which has been detected in many tumours (Planells-Palop et al., 2017).

Additionally PRDM9, the important meiotic recombination hotspot activator is known as meiosis-specific CT gene (Feichtinger et al., 2012).

The above-mentioned proteins/genes facilitate important functions during meiotic recombination. Recombination mechanisms occurs in somatic cells to repair DNA damage, and the factors involved in both types of homologous recombination are overlapping. The aberrant production of one or several meiotic proteins in somatic cells could, in itself or together with endogenous factors, promote the initiation of inappropriate recombination events between homologous and non-homologous chromosomes, thus, may explain the subsequent abnormalities such as insertions, deletions and translocations of chromosomes that frequently present in cancer cells (McFarlane & Wakeman, 2017).

1.3.3. The role of CTAs in cancer cells

Cancers share many common features with germ cells and the activation of germline genes in tumour cells is sought to be the driving force of tumorigenesis (Rousseaux et al., 2013; McFarlane et al., 2014; Whitehurst, 2014; McFarlane et al., 2015). Extensive investigations have been applied in recent years to identify the functions of CT genes in cancer development and initiation and the emerging data are consistent with the idea that re-expression of CT genes in tumours might confer the central characteristics of cancer. Examples of some CT genes and their identified functions are shown in the established diagram in (Figure 1.7).

Meiosis-specific genes are a subclass of CT genes termed (as meiCT genes) that physiologically modulates a reductional segregation of meiosis, a central part of spermatogenesis. Many of meiCT genes are normally silent in healthy somatic cells but their aberrant activation could drive oncogenesis through either the contribution of these genes in soma-to-germline transition or their involvement in oncogenic chromosomes dynamics (McFarlane & Wakeman, 2017). Many examples of meiosis specific genes and their functions in normal meiosis are listed in Section 1.3.2.2. However, the emerging evidence supported the fundamental roles of meiosis-specific CT genes in cancer initiation and maintenance since the finding in *Drosophila* that I(3)mbt brain tumor initiation required an ectopic activation of CT genes, some of which are key meiotic genes (Rousseaux et al., 2013; Rossi et al., 2017).

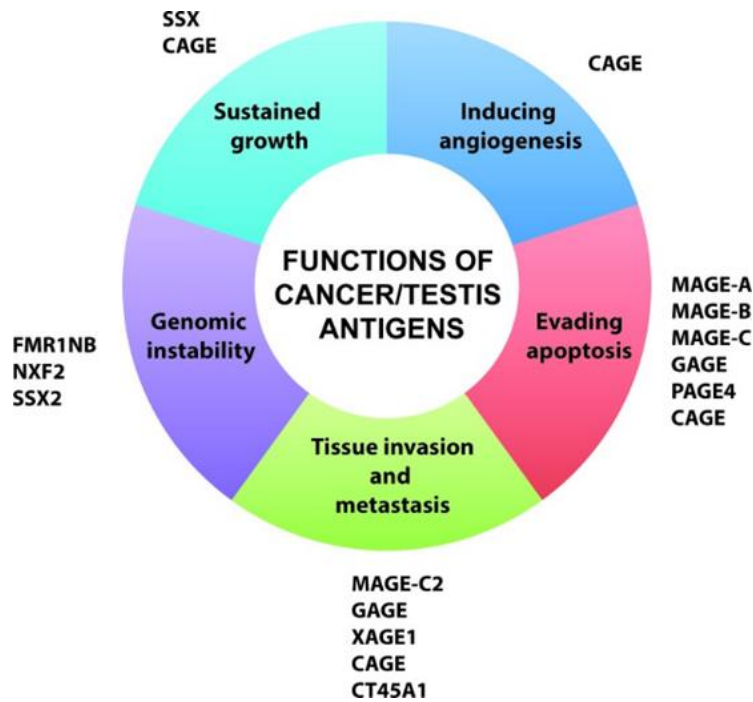


Figure 1.7. Oncogenic functions of CTAs: the diagram shows many examples of CTAs contributions to the central characteristics that lead to tumorigenesis (Gjerstorff et al., 2015).

New emerging evidence indicate important implications of meiCT genes in cancer chromosomal dynamics when Greenberg and co-workers reported that *MND1* and *HOP2*, two meiosis specific genes, that are normally meiotic recombination regulators, play fundamental roles in cancer cells in the absence of normal telomerase, where MND1-HOP2 proteins assist cancerous cells to utilize an alternative lengthening of telomere (ALT) mechanism (Arnoult & Karlseder, 2014; Cho et al., 2014).

Moreover, the expression of meiosis-specific gene *SYCP3* is found to drive ploidy changes and directly influence chromosomal segregation in cancer cells (Hosoya et al., 2011). However, the aberrant presence of its protein SYCP3 in mitosis impairs recombination via disrupting the function of BRCA2, a tumour suppressor recombination regulator (Hosoya et al., 2011; McFarlane & Wakeman, 2017).

Other examples of the role of meiotic recombination regulators activation in cancer cells progression: RAD51 and its orthologue DMC1 is activated in glioblastoma is found to promote proliferative potentials (Rivera et al., 2015). In addition, the human *TEX19* is a CT genes whose expression is found to mediate proliferation in distinct tumour types and also contribute in cancer stem-like cells self-renewal (Planells-Palop et al., 2017).

Sustained growth is an essential feature of all cancer cells that found to result from the deregulation of cell cycle machinery that governs the proliferative and survival signals. Interestingly, many CT genes have been identified to play important roles in cancer cell growth and progression. For instance, study showed that melanoma cancer cells depleted of *SSX2* CT gene expression greatly reduce cellular proliferation (Greve et al., 2015). MAGE proteins are functionally well-characterised CTAs that bind and function to regulate the essential tumour suppressor P53. Knockdown experiments of MAGE-A, MAGE-B and MAGE-C proteins in melanoma cells showed a significant increase in P53 activity and apoptosis (Yang et al., 2007). Furthermore, apoptotic resistance in various types of cancer cells was found to be induced by members of the *GAGE*-like CT gene family (*GAGE* and *PAGE*). Also *GAGE-7* prevents cells undergoing programmed cell death in response to various types of apoptotic agents/stimuli (Cilensek et al., 2002; Kasuga et al., 2008; Kular et al., 2009). However, *PAGE-4* knockdown has been shown to induce and stop tumour growth (Zeng et al., 2011).

Metastasis is a feature of cancer cells which thought to be linked to the EMT. This process allows epithelial cells to lose some features (like cell-cell adhesion and cell polarity) and enhance their potential for motility and invasion. Interestingly, some studies found that CT genes are rarely expressed in benign malignancies (non-metastatic) (Lüftl et al., 2004). Other studies identified that the expression levels of *MAGE-A1* and *MAGE-A4* in primary tumours (early stages of cancer before metastasis) are 20% and 9%, respectively, but their expression levels are remarkably elevated to 51% and 44% respectively in distant metastasis (Barrow et al., 2006). This could be an indication of the functional roles of CT genes in metastasis. Although the direct implication of GAGE proteins in metastasis and invasion remains unclear, the presence of these proteins in migrating germ cells as well as in metastatic melanoma is thought to contribute to their invasions and migrating process (Gjerstorff et al., 2008; Bai et al., 2012). Furthermore, knockdown of GAGE protein in melanoma cell lines significantly reduced their metastatic features (Gjerstorff et al., 2015; Lee et al., 2015).

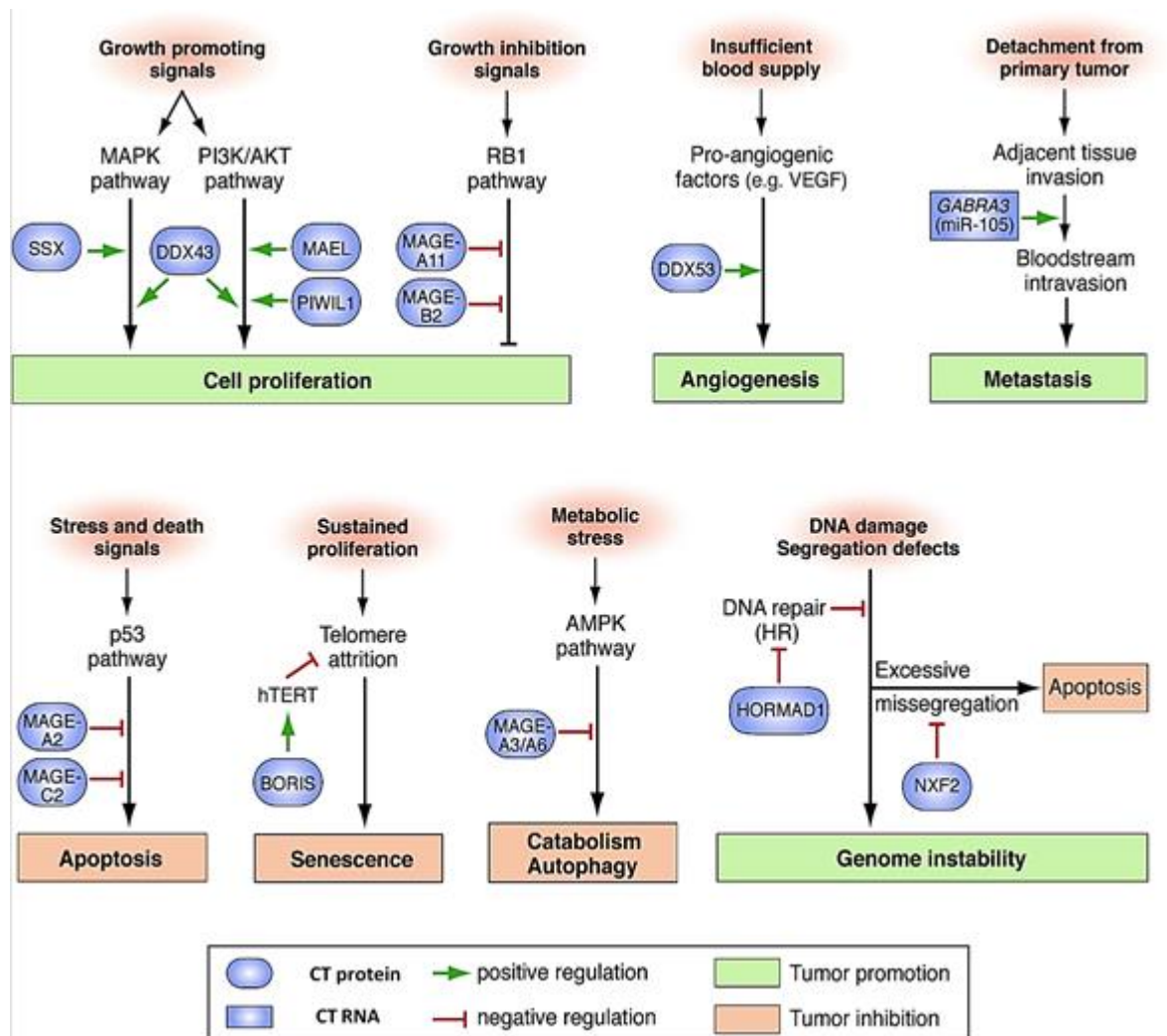


Figure 1.8. Examples of CT proteins and CT mRNA that contribute to various pro-tumoral mechanisms. These examples might function either by stimulating oncogenic functions or by inhibiting tumour suppressor pathways. The figure illustrates examples of CT proteins and miRNA that can regulate many several key processes of cancer development. These observations support the notion that DNA methylation induces the activation of CT genes in tumours (Van Tongelen et al., 2017).

1.3.4. Genome stability and CTAs

There is no doubt that cancer is a highly complex disease and evolved by a wide range of biological and molecular changes showing the behaviour and characteristics of human cancer is extensively variable (GrizziChiriva-Internati, 2006). Tumorigenesis is clearly a nonlinear process whose behaviour is not a clearly predicted or repeated pathway. This evolution does not indicate a lack of order rather than the order is so difficult or impossible to predict and can change in response to distinct tumour environments (Bizzarri, et al. 2011). Some changes such as acquisition of germ-like or embryo-like state may drive tumour cells to undergo a relatively genomic evolution over time to cope with the new surrounding requirements and stress (Flavahan et al., 2017; McFarlane & Wakeman, 2017).

Genome instability is a hallmark of human cancers and is responsible for the genetic abnormalities that drive cancer progression (Hanahan & Weinberg, 2011). Multiple highly regulated mechanisms, such as recombination, are contributing in the maintenance of genomic stability. In the majority of tumours, the alterations in the genome include abnormal chromosomes segregation, mutations, gain and loss of partial or whole chromosome (Janssen & Medema, 2013). Genomic instability in many cancers can result from epigenetic changes including histone modification or DNA hypomethylation of gene promoters that cause ectopic expression of CTAs in cancerous tissues. Furthermore, genomic instability can be generated also by the presence of mobile DNA sequences termed as transposon elements which have ability to change their location in the genome (Downey et al., 2015).

1.3.4.1. Genetic changes associated with cancer:

The advancement of high-throughput techniques in the field of DNA sequencing has allowed large-scale analysis to detect common and rare genomic changes that may drive carcinogenesis. In particular, such these genomic alterations occur in genes encoding proto-oncogenes and tumour suppressor genes. For example, mutations activate proto-oncogenes (e.g., *PRAF*, *KRAS*, *EGFR*) (Davies et al., 2002; Lynch et al., 2004), genomic instability causing fusion oncogenes (e.g., *SYT-SSX*, *BCR-ABL*) (Ladanyi, 2001; Salesse & Verfaillie, 2002), or focal amplification (e.g., *HER2*, *EGFR*) (Slamon et al., 1987; Dacic et al., 2006). Similarly, mutations or deletions of the *TP53* tumour suppressor gene occurs in more than 50% of cancers (Hollstein et al., 1991). Recent studies identified that genomic changes such as mutations in non-coding regions can also lead to carcinogenesis, for instance, mutations in insulators are sufficient for oncogenes activation in several tumours (Hnisz et al., 2016). *BRCA2* is a tumour

suppressor recombination regulator gene and its aberrant expression is reported to increase the risk of cancer incidence especially, breast and ovarian cancers. BRCA2 protein has been reported to interact with the synaptonemal complex protein SYCP3 (Hosoya et al., 2011), however, the over-activation of SYCP3 that was found to impair mitotic recombination by disrupting the function of BRCA2 (Hosoya et al., 2011) is an example of genomic instability in cancer cells.

1.3.4.2. Epigenetics involvement

Epigenetic mechanisms play fundamental roles for the development and maintenance of tissue-specific gene expression profile in healthy tissues. Deregulation of epigenetic processes can lead to a disruptive gene function and therefore, oncogenesis. Global alterations in the epigenetic landscape are a hallmark of cancer (Flavahan et al., 2017). The tumour initiation and progression, traditionally described as a genetic disease, is now realized to involve epigenetic alterations along with genetic abnormalities. Recent improvements in the field of cancer epigenetics have demonstrated extensive reprogramming of every component of the epigenetic mechanisms in cancer including DNA methylation, histone modifications (Sharma et al., 2010; Flavahan et al., 2017).

1.3.4.2.1. DNA methylation

DNA methylation is established during early embryogenesis and plays a role in controlling gene expression in adult cells (Yu et al., 2014). The subsequent process of germline gene activation via DNA hypomethylation in tumours is not understood (De Smet & Lorient, 2013; Van Tongelen et al., 2017).

DNA methylation mostly occurs at C-5 of cytosine in predominantly CpG island sequences in gene promoters and is maintained by DNA methyltransferases (DNMTs) (Jones, 2012). Previous studies identified that the DNA hypomethylation of CT gene promoters was observed in testicular germ cells and tumour cells but was not reported in normal somatic tissues (Kim et al., 2013). The DNA methylation process in gene promoters can inversely affect the gene expression, thereby causing tumourgenesis (Baylin, 2005; Jones & Baylin, 2007; Shen & Laird, 2013). For instance, DNA hypomethylation reportedly increases the expression of CT genes, such as the *MAGE-A1* gene in several tumours and *CAGE* in gastric tumours (Lee et al., 2006; De Smet & Lorient, 2010). In contrast, DNA hypermethylation of the gene promoter of *MAGE* was found to down-regulate its expression (Wischnewski et al., 2007). Many of the CT genes that are regulated by DNA methylation have been shown to be involved in distinct

cancer pathways include growth inhibition and induction, metastasis and invasion, immortality, angiogenesis, and genomic instability (as shown in Figures 1.7 and 1.8).

1.3.4.2.2. Histone modification

Histones are basic proteins that DNA wraps around to form nucleosomes, the basic units of chromatin (Neidhardt et al., 1990). Histone proteins can be post-translationally modified, predominantly in the tail regions and can interact with other proteins to control genome dynamics (Akers et al., 2010). Several histone modifications, including acetylation, methylation, phosphorylation and ubiquitination, take place in special residues in histones that lead to alterations in chromatin structure and transcription, which also has an impact on DNA methylation (Cedar & Bergman, 2009).

Histone deacetylases (HDACs) change chromatin structure by allowing the positively charged lysine residues in the histone tails to interact with negatively charged DNA backbone. This process overall increase the compacting structure of chromatin that makes DNA inaccessible to transcription factors. Inhibition of HDACs by inhibitors (HDACi) activates many genes including CT genes (Karpf, 2006; Ellis et al., 2009).

Histone methylation also plays essential roles in gene expression regulation. Histone methylation can activate or repress genes on the basis of which lysine residue is methylated. For instance, H3K9 or H3K27 (methylation of H3 at lysine 9 or 27 respectively) are repressive marks whereas H3K4 (lysine 4) is an active mark (Lachner et al., 2003). A further direct clue of the function of histone methylation in CT genes regulation was validated by the investigations that mouse ESCs sustaining knockout of H3K9 methyltransferase enzymes, G9a or GLP causes a marked elevation of *MAGE-A* expression (Tachibana et al., 2002; Tachibana et al., 2005). Another study in mice found that the CT gene *Prdm9* mediates histone trimethylation H3K4me3, thus, this process with other possible factors, mark chromatin for recombination (Hayashi et al., 2005). This indicates that *prdm9*, a CT gene, plays a functional role in histone modification.

1.3.4.3. Transposable elements in cancer

Much of our genome is composed of repetitive sequences that include the extensive tandem arrays forming centromeric satellites and telomere ends in addition to other simple repeats distributing within the genome especially in pericentromeric and sub-telomeric regions. They are generated from mobile DNA or transposable elements TEs. Transposable elements are broadly classified to class I retrotransposons that use RNA transcription mechanism to copy

and paste themselves and class II, DNA transposons which excise from donor site to reintegrate elsewhere in the genome through (cut- paste mechanism) (Wicker et al., 2007). This latter class is comprising about 5 percent of human genome but they are no longer active in human. The retrotransposons comprise about 45 percent of human genome and can be categorised into two groups: non-long terminal repeat (non-LTRs) and long terminal repeats (LTRs). Non-LTRs include long interspersed elements (LINE) and short interspersed elements (SINE) whereas the LTRs encompass the human endogenous retroviruses (HERVs). HERVs are similar in structure to infective retroviruses as they are sought to be acquired from the infection of germline cells by ancient retroviruses and later through the evolution they become part of our genome. Therefore, the genetic sequence of HERVs still contain the main retroviral genes gag, pro, pol and env (Kurth & Bannert, 2010; Stoye, 2012; Dewannieux & Heidmann, 2013). The analyses on mRNAs and proteins of active TEs revealed their activation in many different cancers (Downey et al., 2015).

TEs activity is regulated mainly through epigenetic mechanisms. Genomic sequences of TEs are hypermethylated in normal tissues (Schulz et al., 2006); however, majority of tumour samples demonstrate a global hypomethylation state (Alves et al., 1996; Shukla et al., 2013), which may promote the activity of these genes in cancers. On the other hand, there are other layers of cellular protection against the activity of TEs. For example, APOBEC3G is cytidine deaminase with an anti-retroviral role as it inhibits the activity of some HERV K members at a protein level (Lee et al., 2008). A different strategy to control the expression of TEs is through RNA interfering (RNAi). The piwi-interacting RNA (piRNA) pathway is active in germ cells and it uses dicer-independent piRNAs. piRNAs are processed by Piwi or Aubergine proteins that bind to Ago3, which prompts the cleavage of antisense mRNAs. This pathway is crucial to protect the genome against TEs activity (Aravin et al., 2007). In humans, PIWI proteins have been identified to play a similar role. Moreover, PIWI proteins can directly trigger the methylation on TE promoters (Moyano & Stefani, 2015). Additionally, the human PIWI family members are aberrantly expressed in a wide range of cancers, which might result in the dysregulation of TEs in tumours; however, similarly to the effect of activation of TEs, it is not clear whether the abnormal production of PIWI proteins is *per se* tumorigenic or it is a consequence of the disease (Suzuki et al., 2012).

1.4. The use of CTAs in clinical applications

Many immunological approaches have emerged as a potential target for cancer therapy, vaccination and cancer-specific markers (Mellman et al., 2011). Cancer/testis antigens have intriguing characteristics that make them of major interest in this field (Scanlan et al., 2004; Whitehurst, 2014).

1.4.1. The use of CTAs in diagnosis and prognosis approaches

Many strategies have been used to reveal the potential of CTAs in the field of cancer diagnosis and prognosis (Wang et al., 2016). CTAs can be used to differentiate malignant tumour samples from benign tumours. For example, blood sample analysis showed that anti-STAG9 antibody is elevated in lung cancer patients, compared to healthy individuals. Thus, the CTA STAG9 can be used as a candidate biomarker in lung cancer diagnosis (Ren et al., 2016). Furthermore, analyses were conducted on a group of about 200 clinical samples from different stages and grades of colorectal cancer (CRC), this showed that vast majorities of these samples expressed the CT gene *AKAP4* suggesting that *AKAP4* can be used as a diagnostic marker for the early stages of CRC (Jagadish et al., 2016). Currently, PSA protein level in blood samples is used in the investigation and diagnosis of prostate cancers in correlation with other clinical and pathological signs such as age, tumour grade and BORIS protein level. BORIS is a CTA protein whose presence is an indication of aggressive prostate cancer (Cheema et al., 2014). Some CTAs can be used as prognostic markers and predictors for poor clinical outcomes such as MAGEA1 (Zou et al., 2012) and EBI3 (Rousseaux et al., 2013). Analysis of glioblastoma patients reported four different CTAs (CTCF, XAGE3, ACTL8 and OIP5), however, the results also demonstrated that OIP5-positive patients have significantly better overall survival period than CTA-negative patients (Freitas et al., 2013). Finally, John and co-workers found that NY-ESO-1 is a good marker for the response of non-small lung cancer patients to the chemotherapy treatment. The data showed that NY-ESO-1 is linked with a significant survival benefit and downstaging with the given chemotherapy (John et al., 2013).

1.4.2. Vaccination

CTAs possess the potential to be useful as a possible therapeutic vaccines for cancer, either by suppression of cancer development in early stages or at least protecting healthy individuals, especially those with a family history of cancer incidence or genetic predisposition, from the

high risks of cancer. For instance, women with deleterious mutations in *BRCA1* and *BRCA2* genes are more susceptible to develop breast cancer (Adams et al., 2011). CTA SP17 has been shown as a successful vaccine in *in vivo* studies of ovarian cancer in mice (Chiriva et al., 2010).

1.4.3. CTAs in cancer immunotherapeutic

The conventional treatments of cancer, such as chemotherapy and radiotherapy are useful for some patients, however, they are often insufficient alone especially for relapsed or metastatic cases, in addition to their high toxicity and low efficacy potentials. Immunotherapy provide on-target but off-tumour toxicity depending on selection of appropriate tumour antigens. The criteria used for immunological elimination of cancer cells include generation of sufficient numbers of effector T cells with high avidity recognition *in vivo* of tumour antigens. In the past decade, many immunotherapeutic strategies made significant progress as anticancer therapy that initiate and boost an existing immune response against cancers such of these strategies include, cancer vaccine, cell based therapy, immune check point blockade and oncolytic viruses (Mellman et al., 2011).

CTAs can be an ideal tumour antigen family for immunotherapy because of: a) an exclusive expression on cancer cells but not healthy somatic cells, b) stably expression on all/majority of tumour cells C) vital for cancer cells progression and survival d) recognised and targeted by antigen specific cytotoxic T lymphocytes (Simpson et al., 2005; Krishnadas et al., 2013). CTAs are restricted to immune privilege sites of germ cells that are characterised by the presence of the BTB and downregulation of MHC-I class on cells surface. Thus, the germ cells are protected from immune system targeting. CTAs are immunogenic proteins that when processed into peptides and bind to MHC-I presented on cancer cells (Adair & Hogan, 2009). These peptides subsequently can initiate the immune response with other stimulatory factors in immune system to play role in maintaining, enriching, and strengthening the response (Meek & Marcar, 2012).

Adoptive T cell (ATC) therapy depending upon administration of T cells with antitumor activity. The initial approach used extraction of tumour infiltrating lymphocytes (TILs) from patient tumour sample that were then cultured and expanded and re-injected back in the patient (Wu et al., 2012b). For example, expanding *in vitro* of TILs from melanoma patients have been shown to induce immune responses and tumours were regressed in more than 50 per cent of metastatic melanoma patients (Dudley et al., 2008). The autologous T cells can also be genetically engineered with T cell receptors (TCR) that recognise tumour antigens in

order to improve the feasibility of ATC (Klebanoff et al., 2016). An alternative approach depends on using of recombinant or chimeric antigen receptors, which fuse an antigenic-recognition domain with a single-chain fragments that can target common tumour antigens independent of MHC (Klebanoff et al., 2016). This type of therapy makes CTAs among the highest priority target. For example, TCR against epitopes from NY-ESO-1, MAGE-A4 and SSX have been developed and currently under assessment for clinical use (Hiasa et al., 2009; Robbins et al., 2011; Ghafouri-Fard et al., 2015).

Finally, a growing body of work has been demonstrated that epigenetic drugs, particularly DNA methyltransferases inhibitors (DNMTi) can induce immune responses in cancer cells via demethylation of CTAs (Chiappinelli et al., 2016). As CTAs upregulation induce immune response, they represent good candidates for epigenetic mediated immunotherapy. DNMT inhibitors treatment leads to CTA upregulation that can be recognised and targeted by immune system. For example, novel DNMT inhibitor SGI110 which has longer stability than deoxycytidine, has shown promising clinical activity in AML leukemia patients (Taverna et al., 2013). DNMT inhibitor, Decitabine ('5-aza-2'-deoxycytidine) have been reported to improve survival in leukaemia patients along with cancer immunotherapy (Gang et al., 2014).

1.5. CTA identification

A significant issue in the development of prognostic and diagnostic approach for cancer is the identification of new cancer-specific biomarkers (Feichtinger et al., 2012). There is a group of genes in humans that are normally expressed solely in the adult testis; however, the aberrant expression of these genes also signifies a wide range of tumours. Such genes are known as cancer-testis antigens (CTA) genes, which encode CTAs (Whitehurst, 2014). Because of the existence of the blood-testis barriers, the human immune system does not identify the CTAs as 'self-antigens' on cancer cells (Mruk & Cheng, 2010; Li et al., 2012). Thus, CTAs can be used effectively in the clinical field; for example, they can be used as diagnostic markers, vaccination and as immunotherapeutic targets (Fratta et al., 2011; Whitehurst, 2014).

To identify novel human CTA genes, various bioinformatics analytical methods have been employed (e.g., Hofmann et al., 2008). Moreover, Feichtinger et al., (2012; 2014) employed database mining for expressed sequencing tag (EST) and microarray databases to identify more CTA genes. RT-PCR was used to validate candidate CTA genes (Feichtinger et al., 2012; 2014).

The start of the bioinformatic approach was used the GermOnline database to look for mouse meiosis-specific genes, (<http://www.germonline.org/>) (Figure 1.9). Through this method, 744 mouse meiotic spermatocyte-specific genes were obtained. After this, these 744 genes were filtered through the mapping to the human orthologous genes resulting in 408 human genes (Feichtinger *et al.*, 2012). The MitoCheck (<http://www.mitocheck.org>) database was used to exclude the orthologous human genes that are activated during mitosis (Neumann et al., 2010). A total of 375 human potential meiosis-specific genes were obtained following this filtration process. Two different bioinformatics programmes were used to assess these human genes by employing microarray and EST data sets (Feichtinger et al., 2012, Feichtinger et al., 2014). EST and microarray were used to determine the expression of the tissue and cancer specificity of these genes providing a final number of 177 candidates from EST pipeline and 40 candidates from microarray pipeline. Conventional RT-PCR was used to validate the expression profile of these candidates in normal and cancerous human tissues (see, Sammut et al., 2014; Feichtinger et al., 2012; 2014).

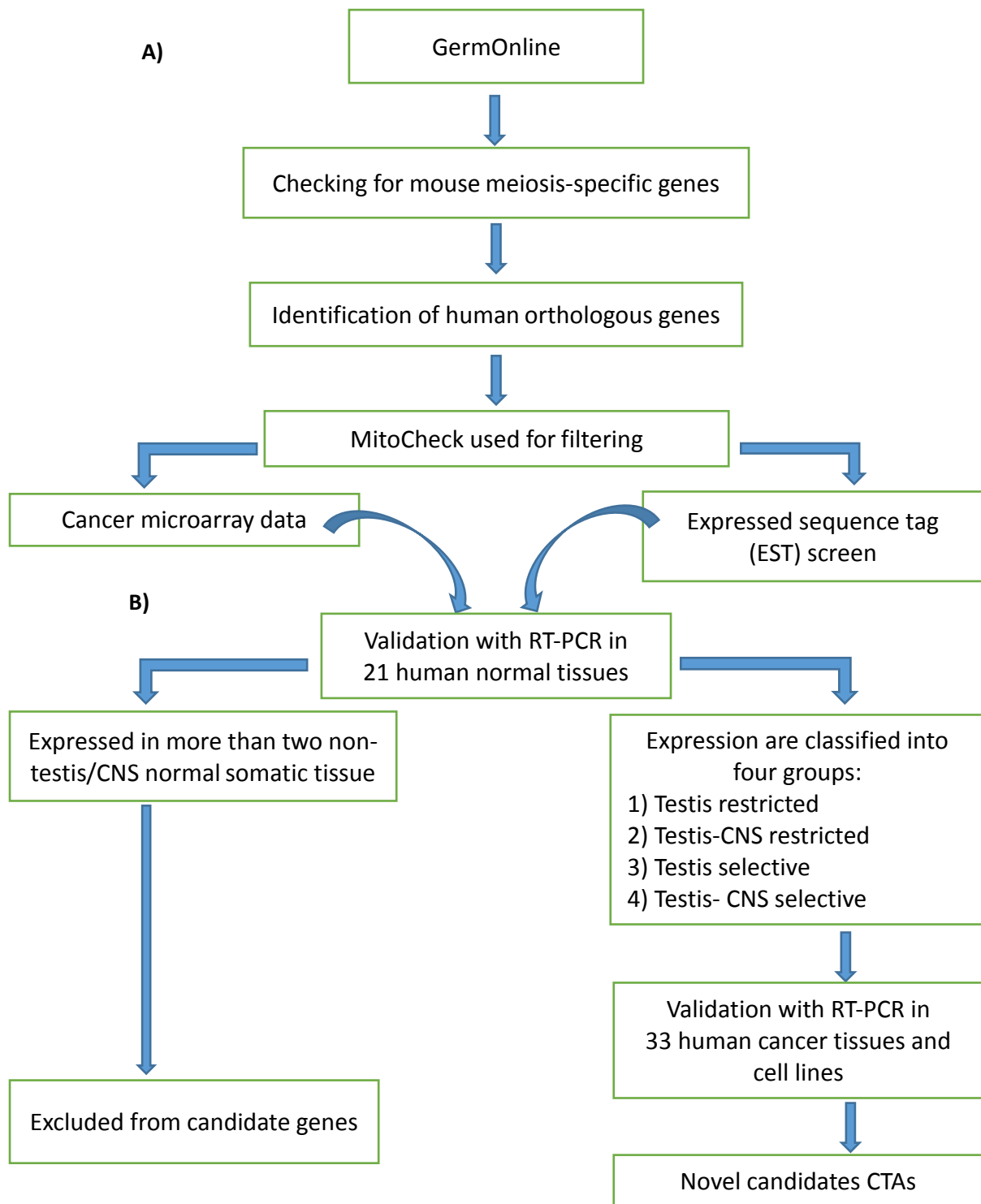


Figure 1.9. Methods used to identify a new CT gene candidate.

(A) Displays the bioinformatics procedure. **(B)** Validation using RT-PCR with 21 human normal tissues. Genes expressed in several normal tissues were excluded from further assessment. Candidate genes show expression as 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 group 1-4 were then examined further 33 human cancer cell lines and tissues to identify potential CT gene candidates.

1.6. Studied genes

1.6.1. *TDRD12* gene

TDRD proteins are a group of mammalian functionally associated methyl arginine binding proteins, which take part in small RNA silencing pathways (Ying & Chen, 2012; Zhou et al., 2014; Iwasaki et al., 2015). Most of these proteins contain several tandem Tudor domain sections, which are repeated. A few of them have domains associated with RNA, such as the RNA helicase domain or the K homology (KH) domain (Cheng et al., 2011; Iwasaki et al., 2015). Numerous TDRD genes are normally expressed in the germ-line as they contribute to p-element induced wimpy testis genes (PIWI)-interacting RNA (piRNA) pathway and gametogenesis process. The TDRD proteins engage with PIWI proteins via symmetrical dimethyl arginines (sDMAs) in different but specific patterns.

TDRD12 is essential for spermatogenesis and its deficiency has led to atrophied testes (Pandey et al., 2013). TDRD12 belongs to Tudor domain family proteins which is containing a central helicase of DEAD box and two flanking Tudor domains, along with CS domain in C-terminal. TDRD12 is conserved gene in eukaryotes and also known as ES cell associated transcript 8 (ECAT8), as its essential roles in germ cells development (Mitsui et al., 2003). In 2012, TDRD12 was described as a good CT gene candidate (Feichtinger et al., 2012). Later, Pandey and colleagues found that TDRD12 is a unique Piwi-interacting RNA biogenesis factor in mice. In addition, they reported that TDRD12 as a component of Piwi protein MILI (PIWIL2 in human), its associated primary pi-RNAs and TDRD1, which all are involved in secondary pi-RNAs biogenesis (Pandey et al., 2013). piRNAs are small noncoding RNAs that act as regulators for genes activation and silencing in stem cells.

Yang and his colleague have recently identified a complex of Piwi-Exd1-Tdrd12 (PET complex), all of which contribute to piRNA biogenesis in addition to transposon element repression in germ cells. Independent loss of either TDRD12 or EXD1 reduces sequences generated by PIWI slicing, impacts biogenesis of piRNAs, and de-represses LINE1 retrotransposons (Pandey et al., 2013; Yang et al., 2016). The regulatory process of piRNA biogenesis is a very critical for maintaining genome integrity in germ cells.

1.6.2. *LKAAEAR1 (C20ORF201)* gene

In 2012, Feichtinger and colleagues were the first who identified *LKAAEAR1* gene as a CT gene through the validation of its expression in normal and cancer tissues followed by meta-analyses of clinical data sets from many types of tumours (Feichtinger et al., 2012). Later,

Kamata et al. also reported the expression of mRNA transcripts of *LKAAEAR1* gene using q-RT-PCR analyses of a range of tumour types (Kamata et al., 2013). Both previous studies confirmed that human *LKAAEAR1* gene transcripts were detected in testis and CNS (cerebellum and foetal brain) in normal tissues and in many cancers including ovarian, breast and prostate. Thus, from the expression profile of *LKAAEAR1* gene, this gene is classified as testis/CNS restricted gene (Feichtinger et al., 2012). Given this, the *LKAAEAR1* gene has potential to be used as cancer biomarker.

Additionally, Kamata and co-workers isolated many antigenic peptides from HLA molecule presented on the prostate cancer cells which could be from *LKAAEAR1* (*C20orf201*). These antigenic peptides are recognised by cytotoxic T lymphocytes (CTL), thus, they are potentially promising to becoming effective vaccines for cancer immunotherapy (Kamata et al., 2013).

1.7. The aim of project

It has been identified that *TDRD12* and *LKAAEAR1* genes as possible cancer biomarkers that may have CSC-specific activity (Feichtinger et al., 2012). This project is to characterise the functional role of germline genes: *TDRD12* and *LKAAEAR1* in cancers, mainly in cancer stem cells (CSCs) to explore their biological functions. These results should help to understand why these genes are active in tumours and how they regulate or interact with different factors in cancer cells. Moreover, these results try to answer if cancer cells strictly require *TDRD12* and/or *LKAAEAR1* to proliferate and promote cancer progression. *TDRD12* and *LKAAEAR1* were investigated for the potentials of therapeutic targets. The main points that were covered in this project as following:

1.5.1. *TDRD12* gene

- 1- We investigated the influence of *TDRD12* protein on NTERA2 cell proliferation and growth.
- 2- We look into the effect of *TDRD12* on cell cycle and further influence on other proteins that are known to play roles in cell cycle.
- 3- Further examinations were carried out to check whether *TDRD12* protein imply on cell apoptosis and senescence.
- 4- The study also examine whether the changes of *TDRD12* transcript levels have a significant effect on other germ genes such as PIWIL family genes.

5- It has been previously identified that mouse Exd-1 gene is a partner for, so these findings were also examined whether the human *EXDL1* expression is linked to *TDRD12* expression. Moreover, this study further examine the expression profile in human stem cells, normal and cancer tissues.

6- The study also explored the possible function of *TDRD12* in stemness features. This point covered the expression of *TDRD12* in induced pluripotent stem cells, human embryonic stem cells and differentiated stem cells. In addition, the connection between *TDRD12* and other stem cell markers were investigated before and after *TDRD12* depletion.

1.5.2. *LKAAEAR1* gene

1- We study the presence of *LKAAEAR1* in many normal and cancer tissues and to validate whether *LKAAEAR1* can be definitely defined as a CTA.

2- The localisation of *LKAAEAR1* was studied on normal testis section and cancer cells.

3- To assess the influence of *LKAAEAR1* transcript level changes on transposon elements.

4- The staining pattern of *LKAAEAR1* protein was investigated on a wide range of cancer tissues that were obtained from diagnosed patients.

Chapter 2:

Materials and Methods

2. Materials and Methods

2.1 Human cancer cell culture

2.1.1 Human cancer cell line sources

The embryonal carcinoma NTERA-2 (clone D1) cell line was originally provided by Prof. P.W. Andrews (University of Sheffield). The ovarian cancer cell line A2780 was kindly gifted by Prof. P. Workman (Cancer Research UK Centre for Cancer Therapeutics, Surrey, UK). The 1321N1, COLO800, COLO857, G-361, HCT116, HT29, LoVo, MCF7, MM127, SW480 and T84 cell lines were purchased from the European Collection of Cell Cultures (ECACC). The two ovarian adenocarcinoma cell lines PEO14 and TO14 were obtained from Cancer Research Technology Ltd. All Cancer cell lines were verified for authenticity once per annum by LGC Standards™ (authentication tracking order: SOJ50865 from April 2017 to March 2018).

2.1.2. Cell culture growth maintenance

Each cancer cell line was cultured in appropriate media as illustrated in (Table 2.1) supplemented with foetal bovine serum (FBS) (Invitrogen; GIBCO 10270). They were then maintained in humidified incubators at 37°C in a 5% CO₂ atmosphere; 10% CO₂ for the NTERA-2 (clone D1) cells. LookOut Mycoplasma PCR Detection Kit (sigma, #MP0035) was used on a regular basis to test the cultures were free from mycoplasma contamination. All media and cell line growth conditions are listed in Table 2.1.

2.1.3. Thawing of stored cancer cell lines

The cell vials were immediately immersed and gently agitated in a 37°C water bath for about 30-60 seconds or until just a small bit of ice was left in the vial. Then, the outside of vial was quickly wiped with 70% ethanol before handling inside the hood where the cells were transferred from cryotube into a 15 ml falcon tube containing roughly 5 ml of pre-warmed medium that was already prepared. The cell suspension was mixed gently and, therefore, spun in a centrifuge at 1500 xg for 5 minutes; after that the medium was aspirated and the cells were resuspended in 10 ml of fresh growth medium. Finally, the cells were mixed and transferred, depending on cells number, into T25 or T75 flasks and incubated at 37°C in a humidified incubator with appropriate CO₂ concentration as shown in Table 2.1. The cells were checked daily until they reached the required confluency.

Table 2.1. Growth conditions and description of human cell lines used in this study.

Cell Line	Description	Medium	CO ₂
NTERA-2 (clone D1)	Human Caucasian pluripotent embryonal carcinoma	Dubeco's modified Eagle's medium (DMEM) + GLUTAMAX™	10%
HEP-G2	Hepatocellular carcinoma	(Invitrogen; 61965) + 10% FBS	5%
SW480	Human colon adenocarcinoma		
LoVo	Human colon adenocarcinoma	Ham's F12 + 2mM Glutamine + 10% Foetal Bovine Serum (FBS).	5%
1321N1	Human brain astrocytoma		
HT29	Human Caucasian colon adenocarcinoma	McCoy's 5A medium + GLUTAMAX™ (Invitrogen; 36600) + 10% FBS	5%
HCT116	Human colon carcinoma		
G-361	Human Caucasian malignant		
A2780	Human ovarian carcinoma		
MCF7	Human Caucasian breast adenocarcinoma	DMEM + GLUTAMAX™ + 10% FBS and 1xNEAA (non-essential amino acids)	5%
HeLa	Cervical cancer		
MM127	Human malignant melanoma	RPMI 1640 + GLUTAMAX™ + 10% FBS and 25 mM HEPES	5%
COLO800	Human melanoma	Roswell Park Memorial Institute 1640 medium (RPMI 1640) + GLUTAMAX™ (Invitrogen; 61870) + 10% FBS	5%
COLO857	Human melanoma		
H460	Large cell lung carcinoma		
MDA-MB-453	Human breast carcinoma	Leibovitz's (L-15) medium + GLUTAMAX™ (Invitrogen; 31415) + 10% FBS	0%
PEO14	Ovarian Adenocarcinoma, peritoneal ascites		
TO14	Ovarian Adenocarcinoma, solid metastasis	RPMI 1640 + GLUTAMAX™ + 10% FBS and 2 mM sodium pyruvate	5%
K562	Leukaemia		
T84	Human colon carcinoma	Ham's F12 + DMEM (1:1) + GLUTAMAX™ (Invitrogen; 31331) + 10% FBS	5%

2.1.4. Cell harvesting and freezing preparation

The cells can be readied for harvesting to be either split or frozen when they reached an 80% confluence. The medium of the culture was aspirated by a sterile pipette. The cells were covered slowly by 5 ml of 1x PBS (Invitrogen; 14190-094) to wash away all traces of FBS. Subsequently, trypsin-EDTA (Sigma-Aldrich; T3924) was carefully added to cover the cells (0.5 – 1 ml) and then re-incubated again for 3 to 5 minutes. After the cells detached from the flask, 5 ml of complete medium was added to stop trypsinization. The number of cells were counted using TC20 cell counter or haemocytometer. To pellet, the cells were centrifuged at 1,500 xg for 5 minutes and the supernatant was aspirated carefully. At this stage, the cells were passaged onto two new flasks (split) or stored in freezing medium. For storing, the pellet was suspended in 1 ml of chilled freezing medium (prepared as 10% of dimethyl sulfoxide (DMSO) [Sigma-Aldrich; D8418] diluted in 90% FBS) and then transferred to a cryotube. The vial was put at -80°C overnight and then stored in liquid nitrogen tanks for long-term storage.

2.1.5. Cell counting

Using haemocytometer: The cell pellet was suspended into 10 ml complete medium and mixed gently by inverting the 15 Falcon tube or by gentle pipetting up and down. The haemocytometer (Sigma-Aldrich; Z169021-1EA) and its coverslip were wiped with 70% IMS, followed by rinsing with ionized water. 10 µl of cell suspension was mixed with 10 µl of 0.4% trypan blue (Invitrogen; 15250-061) in a sterile Eppendorf tube and then a total of 10 µl of the mixture was loaded to each side of the grid of the haemocytometer to count the viable cells. Under the light microscope 10x objectives, the cells on the 25 squares were counted. The mean from both grids was calculated and the total number of the cells can be determined per milliliter using the following calculation:

$\text{Cells/ml} = \text{mean count of live cells} \times 2 \text{ (dilution by Trypan blue)} \times \text{dilution factor} \times 10^4.$

Using TC20™ automated cell counter from Bio-Rad (cat # 145-0102): the cell pellet was suspended in 1 ml PBS and then 10 µl of cells was mixed with the same volume of 0.4% trypan blue. Then, 10 µl of the mixture was injected on to Bio-Rad counting slide (Bio-Rad cat# 145-0011). The TC20 is programmed to calculate the cell number per 1 ml and showed the percentage of viable cells.

2.2. siRNA transfection experiment

The depletion of mRNA transcripts of target genes was performed using siRNA purchased

from Qiagen at a final concentration of 1-5 nM and prepared to obtain 10 nmole of each one as listed in Table 2-2. HiPerFect (Qiagen; #301705) was the transfection reagent used. The procedures were carried out as per the manufacturer's instructions as follows: 150-300 ng of si-RNA with 6 µl of HiPerFect were diluted in 100 µl of serum free media - (use appropriate medium as shown in Table 2-1) and mixed the transfection complex well by vortex. Then, incubate the transfection complex for 20 minutes at room temperature. This recipe is sufficient for 1 – 2 x10⁵ cells which were already seeded in a 6 well plate. After the incubation time, the transfection mixture was pipetted drop-wise into the culture with gently shaking of the plate to distribute the siRNA mixture uniformly. The plates were incubated under normal growth condition, and the first treatment was considered as day 0. In this study, the treatment was repeated daily for three days (hits). On the fourth day, the cells were harvested to extract RNA and protein to validate the efficiency of gene knockdown using quantitative RT-PCR and Western blot, respectively. The negative control siRNA as well as the untreated cell were carried out alongside the target gene treatment.

Table 2.2. GeneSolution siRNA list

Human gene	siRNA name	Target sequence 5' to 3' direct	Cat # *
<i>LKAAEAR1</i>	Hs_LOC198437_2	ACCGCCAAGAGGTGCAGACAA	SI00485135
	Hs_LOC198437_5	CCCGTGGACGCAGTCGCTCGA	SI03186386
	Hs_LOC198437_6	CCAGCCTCCCACATAAAGTTA	SI04258772
	Hs_LOC198437_7	TCCCGCGGTGACGGCGACTGA	SI04319574
<i>TDRD12</i>	Hs_TDRD12_1	ACCCAGGTGGAAGCCAGATTA	SI04713863
	Hs_TDRD12_2	ATCCAGGTTGCTTCTGGGTTA	SI04713870
	Hs_TDRD12_3	ATGGAAGATTCACATGGTGTA	SI04713877
<i>EXDL1</i>	Hs_EXDL1_1	CTAGCTTATGTGGAACACTA	SI04133759
	Hs_EXDL1_2	CAGAAGTTTGGTGCTGCGATA	SI04142446
	Hs_EXDL1_4	CCCTGGTGGATGGTTACCTAA	SI04195058
	Hs_MGC33637_4	CAGAGAATAACTCAGAAAGAA	SI00638764
Negative control	Allstars si-RNA	-----	1027280

*All products source is Qiagen.

2.3. RNA extraction and cDNA synthesis:

The extraction of total RNA was carried out using RNeasy Plus Mini Kit (Qiagen; #74136) as

described in the Company's protocol. Briefly, cells were washed with PBS to inactivate the trypsin and then the PBS was aspirated. An appropriate volume of RLT plus buffer was added to lyse the cells and inactivate RNases. The lysates were transferred to a gDNA Eliminator spin column to get rid of genomic DNA. The flow through was passed to an RNeasy spin column to isolate the RNA using manufacturer's buffers. The final RNA quality and concentration were evaluated by a NanoDrop 2000c (NanoDrop; Thermo Scientific).

First strand cDNA was created using SuperScript III First-Strand Synthesis System (ThermoFisher Scientific; #1808-051). 1 µg of total RNA was transcribed into single-stranded cDNA using an oligo-dT primer. RNase H (supplied in the kit) was added to all samples in order to degrade RNA to obtain pure-single strand cDNA. Control samples were excluded from adding reverse transcriptase enzyme (SSIII, contained in the kit) in all experiments showing no amplified PCR products. The final cDNA in each sample was diluted 1: 8 with DNase/RNase free water and the quality of the cDNA was evaluated using *BACT* primers for qualitative RT-PCR.

2.4. Qualitative PCR

2.4.1. Qualitative RT-PCR

PCR amplifications were performed using MyTaq Red Mix from (Bioline; #BIO25043) in a final products volume of 25 µl as described in the Manufacturer's procedures. Each reaction tube prepared as 1x MyTaq Red Mix, 0.5 µl (10 pmol) of each primers, 1 µl of diluted cDNA (equivalent to 5 ng RNA), and sterile ddH₂O up to 25 µl. PCR amplification was initiated with a pre-denaturation hold 95°C for 1 minute, followed by 35-40 cycles of denaturing temperature at 96°C for 15 seconds; annealing temperature was depending on each set of primers (as shown in Table 2-3). Annealing was for 15 seconds and extension for 30 seconds/kb. Finally, the last cycle of the reaction was followed by the final elongation step at 72°C for 5 minutes. PCR products were loaded on a 0.8 - 1.2% agarose gel (Sigma; #A9539). Gels were stained with peqGreen dye (Peqlab; #37-5000) to be later visualized using a BioRad ChemiDoc XRS+ imaging system. The first lane was loaded with 5 µl DNA markers [100 bp (NEB; #N0467) or 50 bp (NEB; #N0556)].

2.4.2. Primers design for RT-PCR:

Depending on full genes sequences in the database of National Center for Biotechnology Information (NCBI; <http://www.ncbi.nlm.nih.gov/>), primer sets for each gene (forward and reverse) were designed to span at least one interon. Primer3 software was used to design the primers from each gene codon sequence and assess the primers efficiency (from website: <http://www.genome.wi.mit.edu/cgi-bin/primer/primer3www.cgi>). Additionally, Primer3 used to check the sequence length and annealing temperature and also it helps to avoid the possibility of primer dimers. The chosen primers were synthesized by Eurofins MWG operon (<http://www.eurofinsgenomics.eu/>). Finally, the primers were diluted in sterile ddH₂O to a final concentration of 10 pmol prior to use in the PCR reaction.

Table 2.3. Primers used in PCR experiments:

Human gene	Primer name	Primer sequence	Ta*(°C)	Predicted amplicon size
<i>β-ACT</i>	F	5'AGAAAATCTGGCACCACACC3'	58.4	553 bp
	R	5'AGGAAGGAAGGCTGGAAGAG3'		
<i>OCT4</i>	F	5'AGCCCTCATTTACCCAGGCC3'	62.5	872 bp
	R	5'CTCGGACCACATCCTTCTCG3'		
<i>SOX2</i>	F	5'GCAACCAGAAAAACAGCCCG3'	60.5	590 bp
	R	5'CGAGTAGGACATGCTGTAGG3'		
<i>TDRD12</i>	F	5'GGTATTGTCGGTGACCTCAG3'	60.5	355 bp
	R	5'GCTGGAGATCAGAGATTCCG3'		
<i>LKAAEAR1</i>	F	5'ATCTGCTCTTCGGCGACCTG3'	60.0	505 bp
	R	5'ACACTCTCAGTCGCCGTAC3'		
<i>EXDL1</i>	F	5'AAACGGGTGGCTATCTTCCA3'	58.5	472 bp
	R	5'CAGGGTGGTTAGGTCAGACA3'	60.0	

* Annealing Temperature.

2.4.3. Gel purification of PCR products for sequencing

The gel purification was carried out to isolate the desired fragment to be evaluated by sequencing. Fragments were cut out of the gel using a sterile scalpel blade. PCR products were purified from gel samples using the High Pure PCR Products Purification Kit (Roche; #11732676001) as the per Manufacturer's instructions. A NanoDrop was used to measure the concentration of the purified DNA. 75 ng of purified DNA (or 1 µg for plasmid DNA) was adjusted with DNase/RNase free sterile ddH₂O to a final volume of 15 µl for each sequencing reaction. Sequencing primers (forward or reverse) were added at a concentration of 10 µM to each reaction tube. Reactions were sent to Eurofins MWG and the resulting sequences were analyzed using the following website: NCBI nucleotide BLAST search tool, (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>), and aligned with expected products.

2.5. Quantitative Real-time PCR

Quantitative real time PCR (qRT-PCR) reactions were set up using GoTaq qPCR Master Mix from (Promega; #A6001) and carried out in a CFX96 Real-Time Detection System C100 thermal cycler from BioRad. In 96 well plate, each sample was triplicated in a final volume 20 μ l per reaction. Each well contained reaction mixture, which was prepared as 1x qPCR Master Mix, 2 μ l diluted cDNA (equivalent to 10 ng RNA), and primers at a final concentration of 0.2 μ M. The PCR amplification protocol was created as follows: initial denaturing temperature at 95°C for 10 minutes, followed by 40 cycles of denaturation at 95°C for 15 seconds, primer annealing at 60°C for 30 seconds and an elongation step at 72°C for 10 seconds. Finally, melt curve analysis was achieved after a completion of the 40 cycles of reaction. The results were analyzed using BioRad CFX Manager 3.0 software to evaluate the efficiency and specificity of primers, assess threshold cycle value (Ct) and gene expression value. Two types of negative controls; non-reverse transcriptase template (NRT) and non-template control (NTC) were used in each run to check contamination. At least two reference genes were carried out as positive controls to normalize the qPCR results. Both, designed primers and commercial primers (Qiagen; QuantiTech Primer Assay) were used in this study. See Tables 2-4 and 2-5 for the primer list.

Table 2.4. List of commercial primers

Human Gene	Primer name	Source
<i>OCT4</i>	Hs_POU5F1_1_SG	Qiagen; QT00210840
<i>GAPDH</i>	Hs_GAPDH_1_SG	Qiagen; QT00079247
<i>HSP90AB1</i>	Hs_HSP90AB1_2_SG	Qiagen; QT01679790
<i>PIWIL1</i>	Hs_PIWIL1_1_SG	Qiagen; QT00064638
<i>PIWIL2</i>	Hs_PIWIL2_va.1_SG	Qiagen; QT01326990
<i>PIWIL3</i>	Hs_PIWIL3_1_SG	Qiagen; QT00071526
<i>PIWIL4</i>	Hs_PIWIL4_1_SG	Qiagen; QT00011074
<i>SOX2</i>	Hs_SOX2_1_SG	Qiagen; QT00237601
<i>TDRD12</i>	Hs_TDRD12_2_SG	Qiagen; QT1157184
<i>TUBULIN</i>	Hs_TUBA1C_1_SG	Qiagen; QT00062720
<i>β ACT</i>	Hs_ACTB_1_SG	Qiagen; QT00095431
<i>NANOG</i>	HS_NANOG_1_SG	Qiagen; QT01025850
<i>C20ORF201</i>	Hs_LKAAEAR1_1_SG	Qiagen; QT00221816
<i>EXDL1</i>	Hs_EXD1_1_SG	Qiagen; QT00040180

Table 2.5. Non-commercial primers; designed as shown in Section 2.4.2

Gene symbol	Primer name	Primer sequence 5' to 3'	Predicted amplicon size	Ta (°C)	Reference
LINE	F	AAATGGTGCTGGGAAAAGCTG	209	59.9	(Planells-Palop et al., 2017, Oyouni, 2016)
	R	GCCATTGCTTTTGGTGTTTT			
SINE	F	ACGAGGTCAGGAGATCGAGA	173	59.9	
	R	GATCTCGGCTCACTGCAAG			
MM ERV10C	F	CAACCCCTACTGCGGTAAAA	193	59.9	
	R	TTGGAGCATGAGGGTTAAGG		60	
H ERVK107	F	GAGAGCCTCCACAGTTGAG	172	60	
	R	TTTGCCAGAATCTCCCAATC			
H ERVK10	F	GATAACGTGGGGGAGAGGTT	156	60	
	R	TCCATCCATGTGACTGGTGT			
H ERVK Gag	F	CCCAGCATTTGTCTTGGTCT	178	60	
	R	CCTGGGGAATCCTTCTCTTC			
H ERVK Pro	F	TGTTCCCTCAGGGTTTTAGG	153	60.08	
	R	CCCTGGAAGCAGAGAGACTG		60.13	
H ERVK Pol	F	TTGAGCCTTCGTTCTCACCT	216	59.99	
	R	CTGCCAGAGGGATGGTAAAA		60.07	
H ERVK Env	F	AAATTTGGTGCCAGGAAGT	145	59.97	
	R	CCACATTTCCCCCTTTCTT		60.16	
H ERVW Env	F	TAACATTTTGGCAACCACGA	157	59.97	
	R	GGTTCCTTTGGCAGTATCCA		59.93	
H ERVK HML2 Rec	F	ATGAACCCATCGGAGATGCA	160	60	
	R	AACAGAATCTCAAGGCAGAA			

2.6. Protein Extraction:

2.6.1. Whole-cell lysate

The cell pellets were prepared as described in Section 2.1.4 and were transferred in 1x cold PBS to new Eppendorf tube and centrifuged for 3 minutes at 4000 xg at 4°C. The supernatant was discarded to obtain the pellet. The whole-cell lysates were extracted by adding appropriate volume of M-PER Mammalian Protein Extraction Reagent (ThermoFisher Scientific; #78503) based on the pellet weight (15 µl of lysis buffer per 1 µg pellet). Supplements (Halt Protease Inhibitor Cocktail; #87785 and Phosphatase Inhibitor Cocktail; #78420 from ThermoFisher Scientific) were added as per the Manufacturer's recommendations. The cell suspension was incubated (on shaker) at room temperature for 10 minutes then centrifuged at 14,000 xg for 15 minutes in a cold centrifuge at 4°C. The total protein concentration was determined by BCA Protein Assay Kit (ThermoFisher Scientific; #23227) and measured using a NanoDrop.

2.6.2. Subcellular Protein Fractionation

Around 2×10^6 cells were pelleted, washed with cold PBS and dried. The cytoplasmic and nuclear proteins were enriched from the same cells using a Nuclear and Cytoplasmic Extraction Reagents NE-PER™ Kit (ThermoFisher Scientific; #78833) following the Manufacturer's protocol. Briefly, the kit enabled stepwise separation and preparation of cytoplasmic and nuclear extracts from the cells within two hours. Addition of two reagents CER I and CER II to a cell pellet disrupt cell membranes and release cytoplasmic contents. After recovering the intact nuclei from the cytoplasmic extract by centrifugation, the proteins were extracted out of the nuclei with the NER reagent in another pre-chilled tube. The concentrations of isolated proteins were measured using BCA assay for western blot analysis.

2.6.3. Normal lysates sources

Table 2.6 shows the origin of normal human lysates purchases and loaded quantities.

Table 2.6. The source of human cell lysates

whole cell lysates	sources	catalogue number	Amount of protein loaded
Testis	Abcam	AB39257	20 µg
Testis	Novus Biological	NB820-59266	20 µg
Brain Cerebellum	Abcam	AB30069	20 µg
Spinal cord	Abcam	AB29188	20 µg
Brain Whole	Abcam	AB29466	20 µg
Salivary glands	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. Western Blot

A quantity of 30 µg of cell lysate was loaded in each well for western blot analysis. The lysates were mixed with appropriate volume of 10x Reducing agent (Invitrogen; #NP0004) and 4x LDS Sample loading dye 4x (Invitrogen; #NP0007) and adjusted with sterile ddH₂O to total volume of 20 µl. The mixture of each sample was incubated at 70°C for 10 minutes prior to electrophoresis. In the first well, Precision Plus Protein Dual Color Standard from (BioRad; #1610374) was used as protein molecular weight marker and loaded alongside the samples. Electrophoresis was run using NuPAGE Novex 4-12% Bis-Tris Gel (ThermoFisher Scientific; #NP0322) in NuPAGE MOPS SDS buffer (ThermoFisher Scientific; #NP0001) at 100 V for 1 hour. The proteins were transferred onto a PVDF membrane (Millipore; #IPVH00010), which was activated by absolute methanol for 2-3 minutes. Transfer step was carried out using a Trans-blot® Turbo™ RTA Mini PVDF transfer Kit (Bio-Rad cat# 1704272) that was processed in a dry transfer system instrument (Trans-Blot® Turbo™ from Biorad; # 1704150) at 2.5 A for 7 minutes. The transfer buffer was prepared by mixing 200 ml of 5x transfer buffer (BioRad; #10026938), 200 ml 100% ethanol and 600 ml of Nano-pure H₂O.

The PVDF membrane was then incubated for 1 hour at room temperature in fresh blocking buffer [10% dried, fat-free milk + 0.5% Tween-20 (sigma-Aldrich; #P1379) + in 1x PBS]. Primary antibody was diluted in the blocking buffer with appropriate concentration, and incubated with the membrane on a shaker for overnight at 4°C. Triple washes with washing solution (0.5% Tween-20 in 1x PBS) on the shaker at room temperature were performed and followed by incubating the membrane with the corresponding secondary antibody at room temperature for 1 hour with gently shaking. The secondary antibody was diluted in the blocking buffer at appropriate working concentration. After that, the membrane was washed three times each 10 minutes on the shaker at room temperature to remove unbound antibodies. Tables 2.7 and 2.8 show a list of the main antibodies used.

Pierce ECL Plus substrate from (ThermoFisher Scientific; #32132) or Chemiluminescent Peroxidase Substrate-3 from (Sigma-Aldrich; #CPS3100-1KT) were used as substrates to detect western blot signals following the manufacture's procedures. The membrane then was exposed to X-Ray films (ThermoFisher Scientific; #34091) within optimal time of exposure or digital camera detection system (Biorad ChemiDoc XRS+ imaging system; #1708265) to detect signals. ImageLab™ software 4.1 version was used for visualization.

Table 2.7. The primary antibodies and its optimized concentrations for Western blot analyses

Primary Antibody	Cat#	source	Host	Clonality	Application	Dilution *
Anti-LKAAEAR1	AB108142	Abcam	Rabbit	Polyclonal	WB	1/1,000
					IF	1/50
					IHC	1/500
Anti-GAPDH	Sc-365062	Santa Cruz	Mouse	Monoclonal	WB	1/5,000
					IF	1/2000
Anti-Lamin A/C	Sc-7292	Santa Cruz	Mouse	Monoclonal	WB/IF	1/100
Anti-Lamin B	Sc-6217	Santa Cruz	Goat	Polyclonal	WB	1/1,000
Anti-OCT4	ab19857	Abcam	Rabbit	Polyclonal	WB	1/1,000
Anti- α -Tubulin	T6074	Sigma	Mouse	Monoclonal	WB	1/5,000
					IF	1/1000
Anti-Caspase 3	9664	Cell Signaling	Rabbit	Monoclonal	WB	1/1000
Anti-CDK4	2906	Cell Signaling	Mouse	Monoclonal	WB	1/1000
Anti-Cyclin D1	AB16663	Abcam	Rabbit	Monoclonal	WB	1/200
Anti-Cyclin D2	3741	Cell Signaling	Rabbit	Monoclonal	WB	1/1000
Anti-EXDL1	Ab123924	Abcam	mouse	Monoclonal	WB	1/500
Anti-CDKNA1 (P21)	2947	Cell Signaling	Rabbit	Monoclonal	WB	1/1000
Anti-PIWIL1	SAB-4200365	Sigma	Mouse	Monoclonal	WB	1/100
Anti-TDRD1	AB107665	Abcam	Rabbit	Polyclonal	WB	1/100
Anti-TDRD12	Ab182463	Abcam	Rabbit	Polyclonal	WB	1/500

(*) IHC experiments require dilution optimization, attached in Appendix C.

Table 2.8. Secondary Antibodies with its optimal concentrations for Western blot analyses

Secodary Antibody	Cat #	Source	Host	Isotype	Conjugate dilution	dilution
Anti-Goat	AB97051	Abcam	Goat	IgG	HRP	1/10000
Anti-Rabbit	7074	Cell Signalling	Rabbit	IgG	HRP	1/3000
Anti-Mouse	7076	Cell Signalling	Mouse	IgG	HRP	1/3000

2.8. Cell growth and proliferation assay in tissue culture

2.8.1. Growth curve analysis

Cells were cultured in 6-well plates at optimized concentration for each cell line (2×10^5 for NTERA2, 1×10^5 MCF-7 and A2780 cell lines) in 2 ml per well. Every day treatment was carried out according to the procedures in Section 2.2 for up to 6 or 8 days. All experiments were repeated in triplicates and the observations of cells morphology were recorded before harvesting cells. Then the cells were counted in everyday using TC20™ Automated Cell Counter (Biorad; #1450102) and the curve was established.

2.8.2. Cell viability experiment

The NTERA2 cells were plated in 96 well plate, roughly 10^4 per well as described in Section 2.8.1. The transfection complexes of non-interference and interference siRNA were added to specific wells for three sequential days. Each group of cells had four replicates. At day 4, the CellTiter 96® AQueous One Solution Reagent (Promega; # G3582) was added to the cells containing 100 µl medium as 20 µl/well. Then, the plates were incubated for 4 hours in 10% CO₂ humid incubator at 37°C. The absorbance at 490 nm was recorded using an ELISA plate reader [from WALLAC VICTOR² (1420 multilabel counter)] and the background absorbance was measured at zero cells/well.

2.9. Cell immunofluorescence staining

Approximately 10^5 cells were plated upon a sterile coverslip in a well of a 24-well plate and incubated under normal growth conditions until the cell confluency reached about 70-80%. Fixation step was carried out by incubating the cells with 4% Paraformaldehyde (PFA) in PBS at room temperature for 10 minutes and then washed three times with PBS. The permeabilization step was carried out by adding 1 ml of 0.25% Triton X100 diluted in PBS for 15 minutes at room temperature followed by 3 PBS washes. The cells were incubated in blocking buffer (5% of FBS diluted in 1X DPBS) for 1 hour to block non-specific binding of antibodies. The primary antibodies were diluted in blocking buffer at appropriate concentrations as shown in Table 2.7 and added to the cells for overnight incubation at 4°C (or 1 hour at room temperature). After three washes in PBS, the secondary antibodies (Table 2.9) were diluted in blocking buffer and incubated with the cells in the dark at room temperature for about 1 hour. The plates were kept away from light and after the secondary antibody incubation the cells were washed three times with PBS for 10 minutes each wash.

The coverslips were moved onto a clean microscope slides and mounted using Vectashield HardSet Antifade Mounting Medium with DAPI (Vector Laboratories; #H-1500) as per the Manufacturer's instructions. A Zeiss LSM 710 confocal microscope and ZEN software were used to detect the immunofluorescence signals.

Table 2.9. Secondary antibodies and optimized dilutions used in IF assay

Antibody	Dye/Lable	Wavelength of light (nm)	Source & CAT#	Host	Species reactivity	Dilutions
Anti-mouse IgG (H+L)	Alexa flour® 568	578/603	Life Technology A11031	Goat	Mouse	1:500
Anti-rabbit IgG (H+L)	Alexa flour® 488	495/519	Life Technology A11034	Goat	Rabbit	1:500

2.10. Immunohistochemistry

This experiment was carried out on different normal and cancer human tissues in addition to tissue microarrays (TMAs) which were purchased from the Cooperative Human Tissue Network (CHTN) at the University of Virginia (data sheet documents are available at: <https://chtn.sites.virginia.edu/tissue-microarrays>). Written consents were provided from each individual patient and checked via the local research ethics committee (North Wales Research Ethics Committee). All patient consents and ethical approval letters are documented in McFarlane lab. Biopsies were prepared as 4 µm sections fixed in formalin and embedded in paraffin.

The slides were dewaxed at 60°C for 30 minutes in dry oven and allowed to cool at room temperature. The paraffin was removed from the tissues through the deparaffinization step in xylene for 5 minutes. The tissues were then rehydrated through serial concentrations of ethanol (96%, 75% and 50%) each for 5 minutes prior to dH₂O washing. The antigen retrieval mixture was prepared in a large glass jar by adding 450 ml of sterile dH₂O to 50 ml of 10x Dako™ sodium citrate buffer (Dako™: cat# S1700; pH 6.0). The slides were immersed into the 1x buffer to be heated on a heat platform up to 100°C. The buffer containing the slides was maintained at 100°C for 10 minutes and then moved on the bench to cool down at room temperature for approximately 30 minutes. After that, the slides were placed in Dako™

washing buffer for 15 minutes on a gentle rotator. The blocking buffer was prepared for the next step (5% normal serum and 0.3% of 100X triton™ in 1XDPBS) or can be prepared as 995 µl of commercial Dako antibody diluent (Dako™: cat# S3022) mixed with 5 µl of rabbit serum (Dako™: cat# X0902) for at least 10 minutes in room temperature. In the blocking step, volume of 300 µl of blocking buffer was sufficient to cover the tissue area on the slides that were then incubated in a humid chamber in the dark for about 60 minutes. The primary antibody was diluted in Dako™ antibody diluent at appropriate concentrations; 250 µl of the primary antibody mixture was enough to cover the tissue. The primary antibody was then added and incubated overnight at 4°C in a humidified chamber. The primary antibody was then removed by tapping the slides on a clean tissue paper and placed in the rack (every two slides were back-to-back) to be washed twice for 5 minutes each by Dako™ washing buffer. Quenching endogenous peroxidase step was carried out with hydrogen peroxide mixture (prepared as: 10 ml of 3% hydrogen peroxide mixed in 100 ml of methanol) for 7 minutes to get rid of carryover water. The secondary antibody was added to the tissues sections and incubated in a humid chamber for 60 minutes at room temperature followed by triple washes of 5 minutes each by Dako™ washing buffer. The working solution of DAB substrate was prepared by adding one drop (or 20 µl) of DAB chromogen (Dako™ cat# K3467) to 1 ml of substrate buffer and left for at least 30 minutes at room temperature before use. Amount of 200 µl from DAB solution was enough to cover the tissue which was incubated for at least 5 minutes at room temperature. The slides were then rinsed with sterile ddH₂O for 5 minutes in three repeats. The slides counterstained with Hematoxylin (Dako™: cat# CS70030) using a dropper to cover the tissues sections for about 15-30 seconds. Again the slides were placed on a rack and rinsed in a container with running water for 5 minutes. Finally, the slides were dehydrated in absolute ethanol for 5 minutes followed by rinsing in xylene three times before mounting. The coverslips were mounted with DPX mountant for histology (Sigma cat# 06522) and allowed to dry at room temperature for about 30 minutes before imaging. Digital images were obtained using Axio Scan.Z1 scanner (Zeiss). Negative controls were incubated with secondary antibodies only. Primary and secondary antibodies that were used in this experiment are listed in Table 2.10.

Table 2.10. The primary and secondary antibodies that used in IHC method.

Antibody	source	dilutions
Primary: anti-LKAAEAR1	Abcam: ab108142	1:250 1:500
Secondary antibody: goat anti-rabbit IgG H&L (HRP)	Abcam: ab6721	1:1000

2.11. Differentiation experiment

NTERA2 cells were differentiated using two types of induction agents: retinoic acid (RA) which was purchased from (Sigma; cat# R2625) and hexamethylene bisacetamide (HMBA) (Sigma; cat# 224235). The preparation of RA agent as powder when received requires it to be emulsified in dimethyl sulfoxide (DMSO) (Sigma; cat# D8481). So, the DMSO induction agent was used as internal negative control for the RA induction. The cells were differentiated with 3 mg/ml of RA added to the proper medium. The HMBA was prepared at 3 mM concentration as: 3 gm of HMBA (powder) was mixed in 50 ml dH₂O and then 5 ml from this mixture was added to 500 ml prepared media. In DMSO induction: only 0.5 ml was added to 500 ml of complete medium. Untreated cells culture was used as negative control.

2.12. Flow Cytometry (FACS)

NTERA2 cells depleted of TDRD12 using siRNA transfection procedures described in Section 2.2. The cells were harvested and counted and then fixed in 70% ethanol for overnight incubation at -20°C. Cells were incubated in PI stain (50 µg/ml, Sigma; #P4846) with RNase A (0.5 mg/ml, Sigma; #R4642) for 15 minutes at 37°C. Partec CyFlow® Cube 8 flow cytometry instrument was used for GFP fluorescence and DNA content analysis. Software analysis was carried out using the multicycle cell cycle analysis FCS Express 4.

2.13. Immunoprecipitation

Cells were cultured to 80% confluency in T75 flasks, washed with cold PBS, scraped in to 500 µl lysis buffer (25 mM Tris, 7.4 pH, 150 mM NaCl, 1% NP-40, 5% glycerol, 1mM EDTA, and HALT protease inhibitor cocktail); (ThermoFisher Scientific; #87785) and then incubated on ice for 5 minutes with constant agitation in every minute. The cell suspension was centrifuged

for 10 minutes at 13,000 g to isolate total protein. The concentration of protein was determined using BCA assay. Immunoprecipitations were carried out using the Pierce Direct Magnetic Kit (Co-IP kit); (ThermoFisher Scientific; #88828) following the manufacture protocol. Briefly, per IP only a final concentration of 5 µg/100 µl the primary antibody was covalently coupled with 25 µl NHS-activated magnetic beads (Pierce) and the negative control contained only beads. After the beads were washed, the amount of total protein of 500 µl at 1 mg/ml concentration were added to the beads and incubated on a rotator for 2 hours at room temperature. Finally the beads were washed in lysis buffer and the bound proteins were eluted for Western Blot analysis.

Chapter 3:

Functional analysis of human *TDRD12* in germ line tumour cells

3. Functional analysis of human TDRD12 in germ line tumour cells

3.1. Introduction

3.1.1. TDRD Family genes

In humans there are approximately 41 Tudor domain-containing proteins (TDRDs) (Pek et al., 2012; Lu & Wang, 2013). The Tudor domain exists in single or multiple copies, in the presence or absence of other functional domains, such as the Jumonji C (JmjC) domain, the SET domain, the RING-finger domain, the PHD domain and the AT-rich interaction domain (ARID). Tudor domain proteins have been found to contribute to the function of other proteins that associated with RNA metabolism, such as small RNA pathways, alternative RNA splicing and germ cell development (Chen et al., 2011; Pek et al., 2012). These proteins contribute to a wide variety of processes, including the DNA damage response, histone regulation, cell division, differentiation, genome stability and spermatogenesis (Chen et al., 2011; Pek et al., 2012). For example, mTdrd1 (mouse Tudor repeat-1) was detected in spermatogonia as an essential protein for male germ cell development (Wang et al., 2001). Another study determined that Tdrd1 is at high levels in foetal spermatogonia, postnatal primary spermatocytes, chromatid bodies at late stages of spermatocytes, and round spermatids (Chuma et al., 2006; Kim et al., 2016). Tdrd1-deficient mice are infertile due to meiosis being disrupted (Chuma et al., 2006). Tudor domains recognize and bind to methylated lysines and arginines of target substrate proteins, and this is postulated to be the activity of these domains to facilitate the association of protein complexes at different cellular compartments (Liu et al., 2010; Tripsianes et al., 2011). A subclass of Tudor domain proteins interact with methylated lysines, whereas the rest have methyl-arginine-binding capacity. However, the factors that control this specificity remain poorly understood (Chen et al., 2011). Dysregulation of various TDRDs has been reported to play critical roles in the development of several types of tumours (Jiang et al., 2016). For example, over production of a TDRD protein, ARID4B, is predictive of a poor prognosis of breast cancer and metastatic growth (Goldberger et al., 2013); another TDRD, UHRF1, has been suggested to contribute to key cell-cycle inhibitors through an epigenetic mechanism to control cell proliferation in breast cancer (Bronner et al., 2007; Wu et al., 2012a). Moreover, overexpression of *PHF20L1* (TDRD gene family) significantly correlates with low survival rates of breast cancer patients; however,

downregulation of *PHF20L1* stops cell proliferation in breast cancer cell lines (Jiang et al., 2016; Carr et al., 2017).

3.1.2. TDRD12

TDRD12 belongs to the Tudor domain family proteins, which contains a central helicase of DEAD box and two flanking Tudor domains, along with a CS domain in the C-terminal (Figure 3.1 A and Figure 3.2 C). *TDRD12* is a conserved gene in eukaryotes, and is also known as ES cell-associated transcript 8 (*ECAT8*) due to its essential role in germ cell development (Mitsui et al., 2003). Pandey and colleagues found that murine *Tdrd12* is a unique Piwi-interacting RNA (pi-RNA) biogenesis factor. In addition, they reported that murine *Tdrd12* is a component of a Piwi complex associated with primary pi-RNAs and *Tdrd1*, which are all involved in secondary pi-RNAs biogenesis (Pandey et al., 2013). Moreover, the loss of *Tdrd12* function in germ cells leads to impaired pi-RNA biogenesis, defects in spermatogenesis and atrophied testes (Pandey et al., 2013).

The regulatory process of pi-RNA biogenesis is critical for maintaining genome integrity in germ cells. Pi-RNAs are small noncoding RNAs that act as regulators for gene activation and silencing in germ and stem cells (Khurana & Theurkauf, 2010; Senti & Brennecke, 2010; Pek et al., 2012). Yang and colleagues (2016) have recently identified a complex of Piwi-Exd1-*Tdrd12* proteins (PET complex), all of which contribute to pi-RNA biogenesis in addition to transposon element repression in germ cells (Yang et al., 2016) (Figure 3.1 C). Exonuclease domain-containing-1 (Exd1) is characterised as a partner of *Tdrd12* that functions in an RNA-binding role within the PET complex (Yang et al., 2016). Exd1 in mouse contains an Lsm domain that is proposed to provide mutually exclusive interaction between *Tdrd12* and RNA (Yang, 2014) (Figure 3.1 B). It is expected that Exd1 works as an RNA chaperone to deliver ssRNA molecules to *Tdrd12* through the interaction with its helicase domain (Yang, 2014). Independent loss of either *Tdrd12* or Exd1 reduces sequences generated by PIWI slicing, impacts the biogenesis of piRNAs, and de-represses LINE1 retrotransposons (Pandey et al., 2013; Yang et al., 2016).

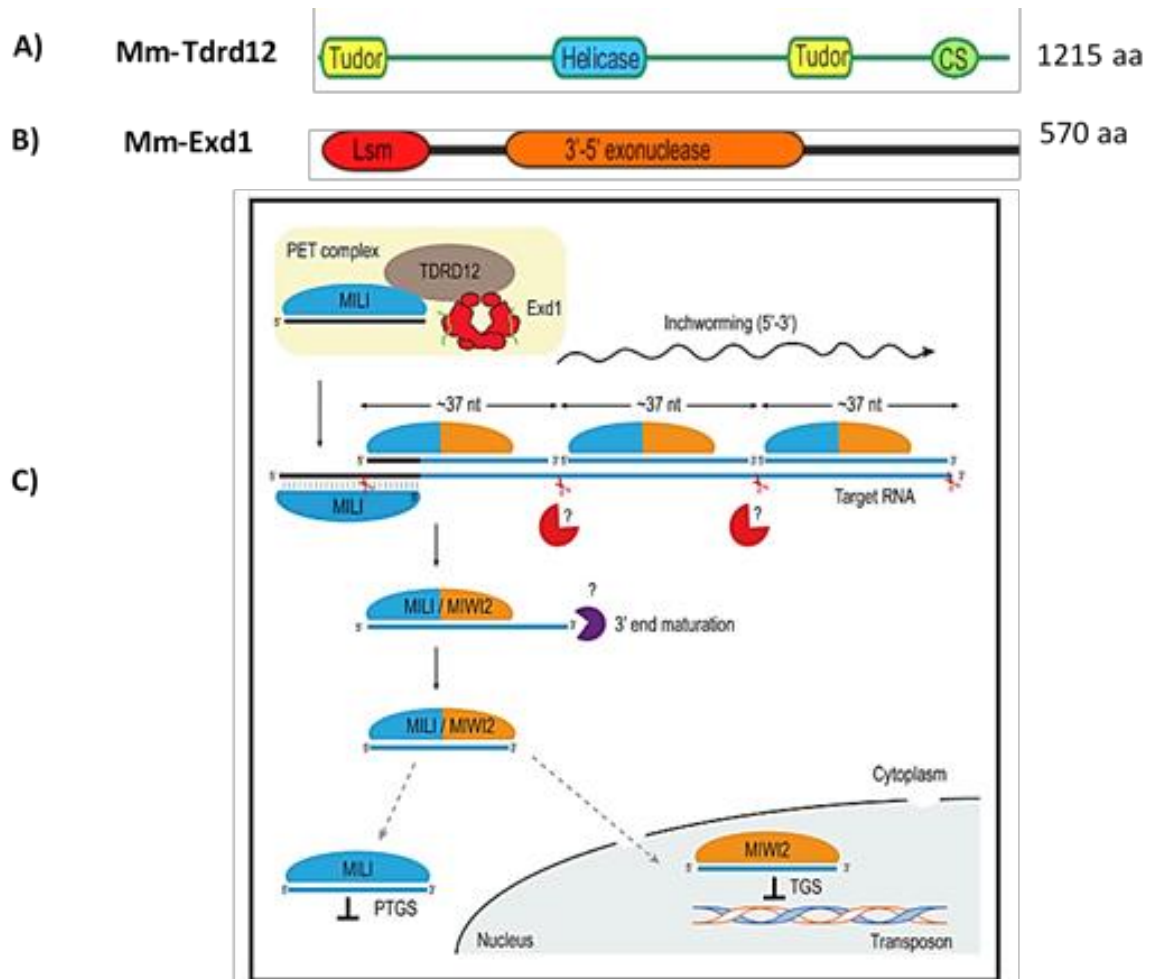


Figure 3.1. The potential roles of TDRD12 in the PET complex. (A) TDRD12 contains a central DEAD box (RNA helicase) domain in-between two Tudor domains and the C-terminal CS-domain (named after CHORD-containing proteins and SGT1). SGT1 is a highly conserved nuclear protein which functions in kinetochore assembly, required for G1/S transition and G2/M transitions and interacts with HSP90. (B) EXD1 (Mm Exd1) is composed of 570 amino acids that encode for an N-terminal Like-Sm (Lsm) domain and a central exonuclease domain, followed by a long C-terminal tail. The Lsm domain is suggested to mediate interaction of EXD1 with TDRD12. (C) A schematic to demonstrate the contribution of the PET complex in pi-RNA biogenesis to repress transposon elements, (adapted from (Pandey et al., 2013; Yang et al., 2016).

3.1.3. The aim of the work in this chapter

Cancer stem cells (CSCs) are cells with self-renewal and differentiation potential that were hypothesised to be responsible for cancer recurrence and chemotherapy resistance (Bao et al., 2006; Jeon et al., 2011). The human *TDRD12* gene is poorly studied, however, work in this chapter aimed to characterise the roles of TDRD12 in human germ tumour cells which have stem cell-like characteristics (NTERA2). The initial depletion of TDRD12 protein in NTERA2 cells showed a significant reduction in cell counts that led to investigation of the possible influence of TDRD12 on cell cycle progression. Furthermore, a possible regulatory mechanism of human *TDRD12* on the PET complex genes and transposable element *LINE-1* is explored.

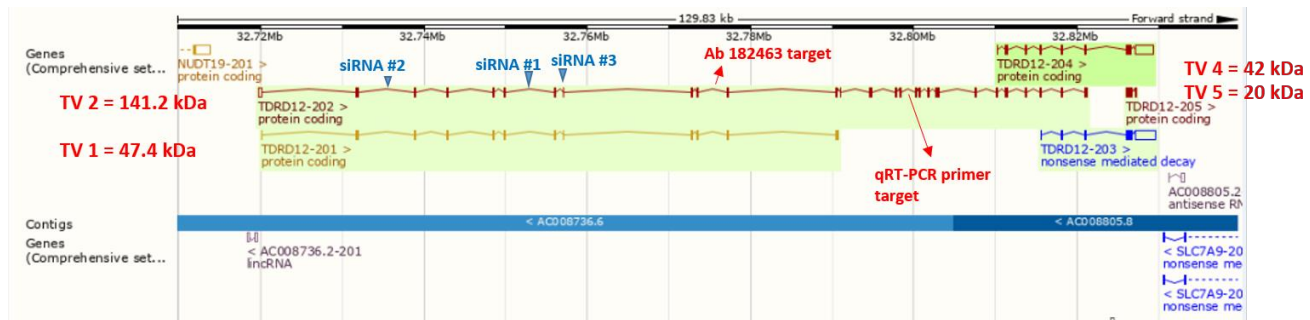
3.2 Results

3.2.1 Knockdown of TDRD12 in germ line tumour cells, NTERA2.

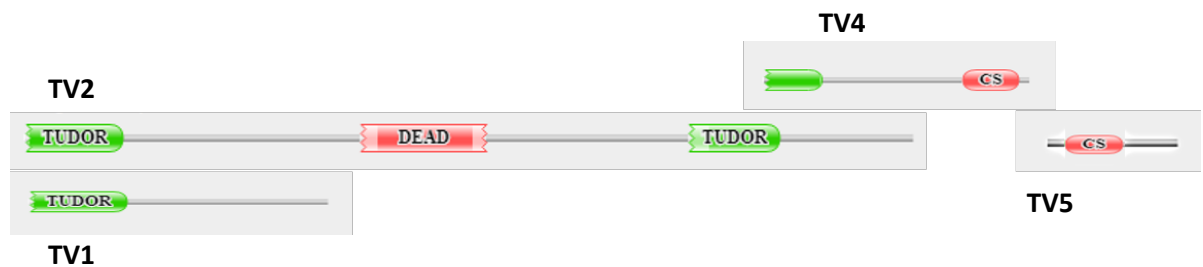
NTERA-2 cells are cancerous stem-like cells that were originally derived from lung metastasis of a primary testicular embryonic carcinoma (Andrews et al., 1984). Small interference RNAs (siRNAs) were employed to deplete *TDRD12* mRNA levels in NTERA2 cells to investigate the effect on cell proliferation. Three distinct siRNAs (#1, #2 and #3) were purchased from Qiagen to target the *TDRD12* mRNA (Figure 3.2 A): siRNA #1 targets *TDRD12* in exon 6; siRNA #2 targets *TDRD12* in exon 2; and siRNA #3 binds to *TDRD12* mRNA at the junction between exons 7 and 8 (Figure 3.2 A). Depletion efficiency was monitored by qRT-PCR analysis; additionally, the polyclonal anti-TDRD12 antibody was used in western blots of whole cell protein extracts (WCEs) following siRNA depletion. The *TDRD12* is predicted to have five splice variants; four variants are potentially protein coding while the last one is predicted to be a non-coding protein [NCBI; gene ID: 91646 and ensemble database: ENSG00000173809] (Figure 3.2 A). The predicted size for each protein variant was calculated (Figure 3.2 A). Additionally, computational analyses of TDRD12 predicted protein variant sequences for pfam matches at: (<https://pfam.xfam.org/>) demonstrate the expected protein domains that allow protein-protein interactions (Figure 3.2 B). The McFarlane group previously demonstrated a new transcript variant 6, which showed to contain matched sequences from TV2, TV4 and TV5 as shown in (Figure 3.2 C; Appendix A-1 ; Oyouni, PhD thesis 2016, Bangor University).

QRT-PCR analysis demonstrated significant reductions of *TDRD12* mRNA transcript levels as compared to negative control siRNA (P value < 0.0001) in all *TDRD12* siRNAs treated NTERA2 cells (siRNA #1, #2, #3 and mixed siRNAs), as shown in (Figure 3.3 A). Western blot analysis showed reductions of a protein migrating at approximately 203 kDa following siRNA #2, siRNA #3 and mixed siRNA with a slight reduction in siRNA #1-treated cells in comparison with the negative control and untreated cells (Figure 3.3 B). The protein signals were detected using Gel Doc™ XR+ Gel Documentation System, therefore ImageLab software analysis was used to determine the size of higher molecular weight band as approximately 203 kDa. This band may be related to transcript variant 6, (TV6) that we predicted at approximately 162 kDa, however, the higher molecular weight given by the instrument might be not accurate or this protein might undergo posttranslational modifications (PTMs) such as ubiquitination. Moreover, in this blot, there is a band detected under 37 kDa which may result from degradation or possibly be produced by a small variant of *TDRD12* that is detected by the antibody but it might be missing exons that are targeted by anti-*TDRD12* siRNAs. The same blot with higher exposure settings is demonstrated in (Figure 3.2 C) showing other bands which are possible of other *TDRD12* variants or might be nonspecific bands that cross-react with other proteins or resulted from *TDRD12* degraded products.

A)



B)



C) TDRD12 TV6 protein, it is predicted to be 162 kDa.

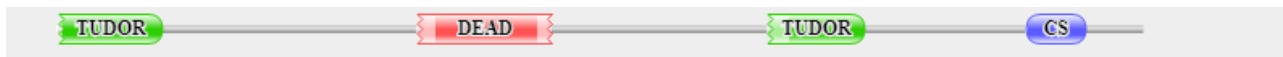


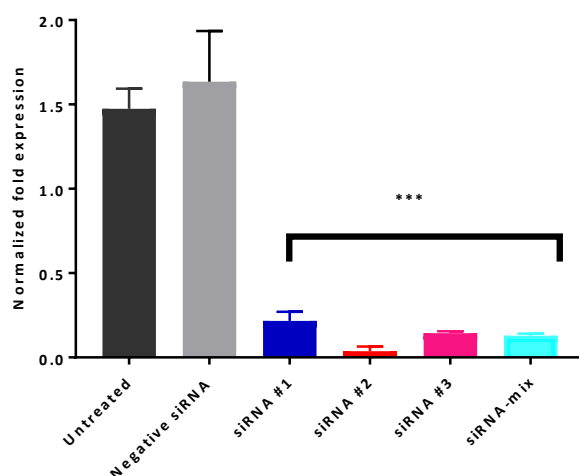
Figure 3.2. The schematic representation of the four potential transcript variants (TVs) of TDRD12 that are coding protein.

(A) Four predicted protein coding isoforms 1, 2, 4 and that defined as (TDRD12-201, TDRD12-202, TDRD12-204 and TDRD12-205) respectively. The schematic also showing that anti-TDRD12 antibody (ab182463) targets isoform 1 and 2 while the used qRT-PCR primer (TDRD12_2_SG) is expected to specifically detect isoform 2. Detected sequences of anti-TDRD12 siRNAs 1, 2, and 3 is shown to target variants 1 and 2 exons. The expected size for each variant was calculated by: (www.calctool.org/CALC/prof/bio/protein_size).

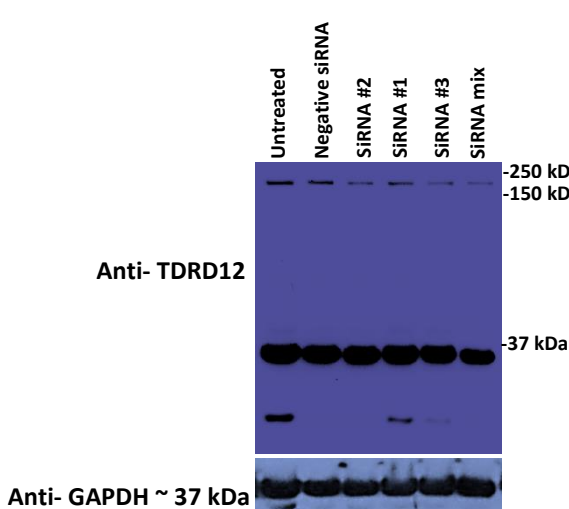
(B) Computational analysis of TDRD12 protein sequence for pfam matches showing the domains of TDRD12 in each isoform using website (<http://pfam.xfam.org/>).

(C) The TDRD12 transcript variant 6 protein pfam domains shown. Tudor domain proteins function as molecular adaptors, binding methylated arginine or lysine residues on their substrates to promote physical interactions and the assembly of macromolecular complexes. DEAD box proteins are highly conserved in nine motifs and are involved in an assortment of metabolic processes such as RNA metabolism. CS domain at C-terminal region of high sequence homology called the Shq1 domain.

TDRD12 expression levels



B) Low exposure



C) High exposure

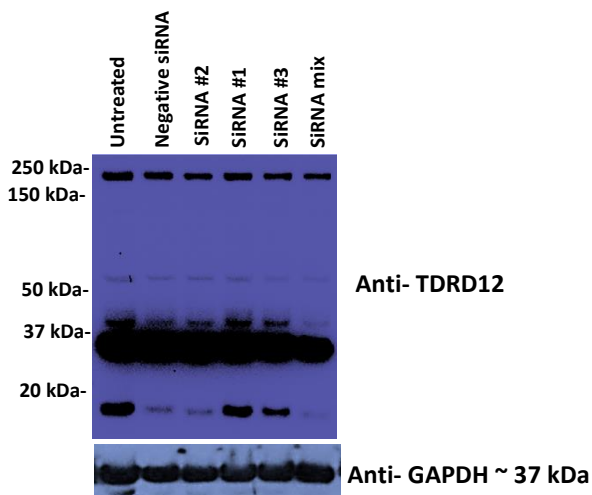


Figure 3.3. Analysis of TDRD12 knockdown efficiency in NTERA2 cells by qRT-PCR and Western blot. (A) Bar charts show the mRNA transcript levels of the *TDRD12* gene by qRT-PCR in NTERA2 cells following depletion. The cells were transfected with a negative control and three types of siRNA targets, and the gene expression was normalised to *GAPDH* and *ACTB*. The fold change was computed using the $\Delta\Delta C_t$ method, and the error bars refer to the standard errors for the mean of three technical repeat. *P* values were calculated in comparison to the control (negative siRNA) treatment (***) *P* value < 0.001) Three biological replicates were performed from RNA extracts. (B) Western blot analysis to show the TDRD12 protein levels after three days of treatment using different siRNAs. (C) The same blot with higher exposure time. GAPDH was used as a loading control. Whole cell protein extracts were carried out three biological repeats (for siRNA2) and two repeats (for siRNA 1 and 3).

3.2.2. TDRD12 depletion has an influence on NTERA-2 cell proliferation rate.

During knockdown experiment, the number of harvested cells in the siRNA treated cultures was lower than the untreated cells and negative controls. Therefore, an experiment was performed to investigate the cell counts following TDRD12 depletion by siRNA. The cell count curve was established depending on cell counts from cultures of siRNAs treated cells along with untreated cells and the non-specific siRNA control for six days of siRNA transfections.

The results showed that the cell count were reduced in positively treated cells with siRNAs #2 and #3 in comparison with negatively treated cells (Figure 3.4). Additionally, the efficiency of protein reduction was assessed by western blot analysis of WCEs after 72 hours of treatment by siRNA #2 and #3 only. In this blot a band was detected at 47 kDa with a significant protein reduction following *TDRD12* specified siRNA treatment. It is at the expected size and might predicted by TDRD12 variant 1. This results suggest that NTERA2 cell proliferation is affected due to TDRD12 protein reduction. The higher band (203 kDa) was not appear in this blot; however, in this assay we used conventional detection system (developer) with normal chemiluminescent substrate which might be unable to detect weak signals or the higher molecular weight species may not have transferred during blotting.

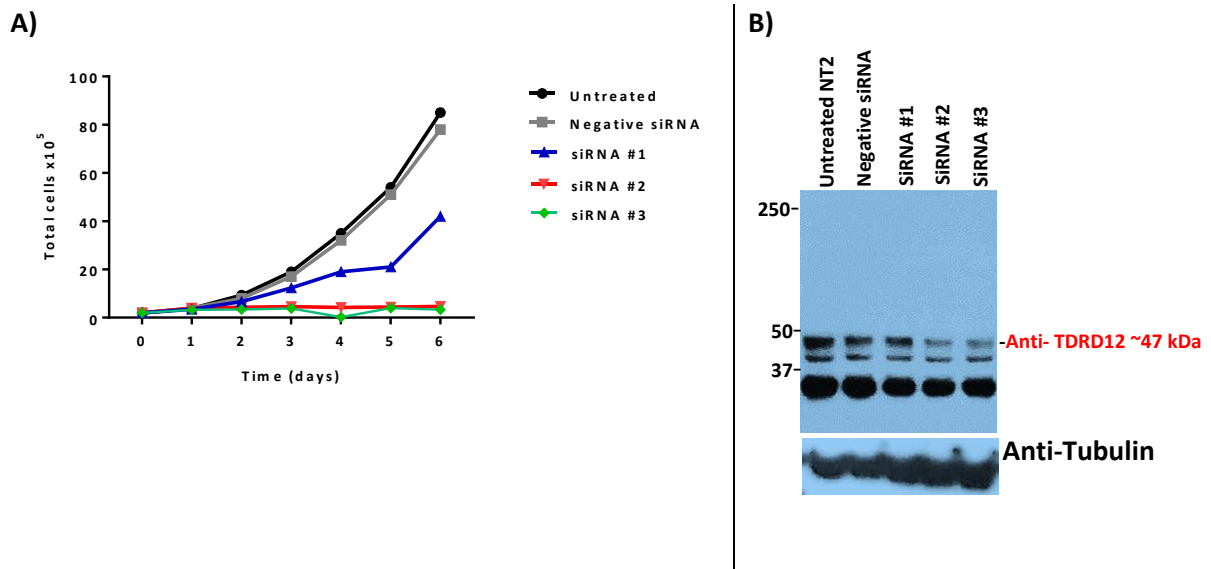


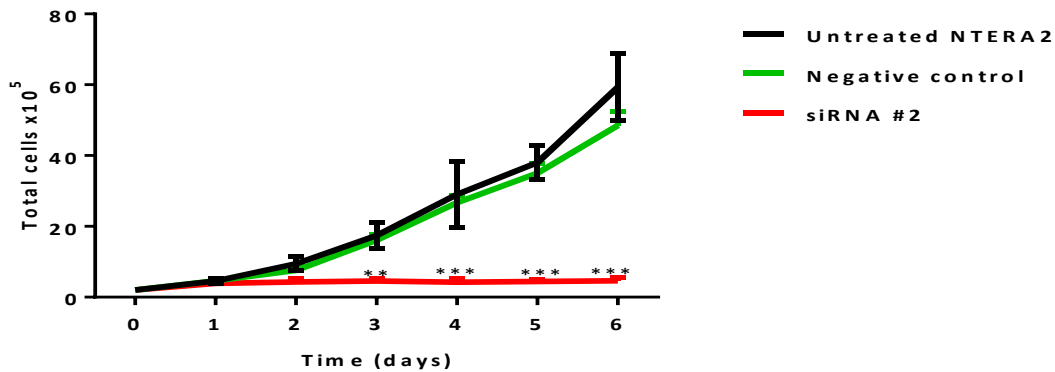
Figure 3.4. Cell growth curve for NTERA2 cells depleted for TDRD12 using three sequences. (A) The cells were counted daily after the depletion of TDRD12 for six days of treatment. Negative control cells were transfected with non-interference RNAi, while the positive treatment used three types of commercial siRNAs (#1, #2 and #3). The results showed a significant reduction of cell counts in siRNAs #2 and #3, as p values < 0.05 from three biological repeats. (B) Western blot to assess the reduced amount of TDRD12 protein in siRNA #2 and #3.

To confirm these observations and to further examine the role of TDRD12 in controlling NTERA2 cell proliferation, TDRD12 depletion was carried out using siRNA #2. The cell counts in siRNA treated cells remained near to the cell counts at the first day post treatment. Cell counts suggested that the cell proliferation rate is markedly reduced in TDRD12-depleted NTERA-2 cells in comparison with negative controls as shown in (Figure 3.5 A). The depletion of TDRD12 efficiency was assessed by western blot analysis as presented in (Figure 3.5 B). A polyclonal anti-TDRD12 antibody (ab182463) from (Abcam) was used for protein detection. Western blot showed many bands (Figure 5.3 B), band (1) of approximately 203 kDa which, is affected by anti-TDRD12 siRNA #2, and may be related to TDRD12 variant 6 protein. Weak signals showing reductions in siRNA #2 treated cells were also detected as shown band (2) at 141 kDa and band (5) at 47 kDa which might be relating to transcript variants 2 and 1 respectively. Additionally, other signals (bands 3 and 4) at approximately 70 and 80 kDa emerged that are not significantly reduced, however, they might be resulted from degradations or another unpredicted variants for TDRD12 protein which are not depleted by siRNA #2. Finally, the strong signal below 37 kDa (band 6) emerged showing a slight reduction due to siRNA #2 transfection and it is expected to be a small variant of TDRD12 protein or a degradation product.

3.2.3. Cell cycle analysis of TDRD12-depleted NTERA-2 cells.

Analysis of NTERA-2 cells transfected with siRNA #2 and #3 showed that cell proliferation was reduced in treated cells, as shown in Figure 3.4 and 3.5. To assess why this might be and what cell cycle stage the cells were arrested in, flow cytometry was carried out on propidium iodide (PI) stained cells following 72 hours of anti-TDRD12 siRNAs depletion. The results showed clear changes in cell cycle distribution, particularly in cells transfected with siRNA #2 and #3 compared to the negative control. The cells appear to be accumulated in the S phase, with a marked reduction of the subpopulation of cells undergoing mitosis at G2/M phase (Figure 3.6).

Growth cell curve



Protein analysis

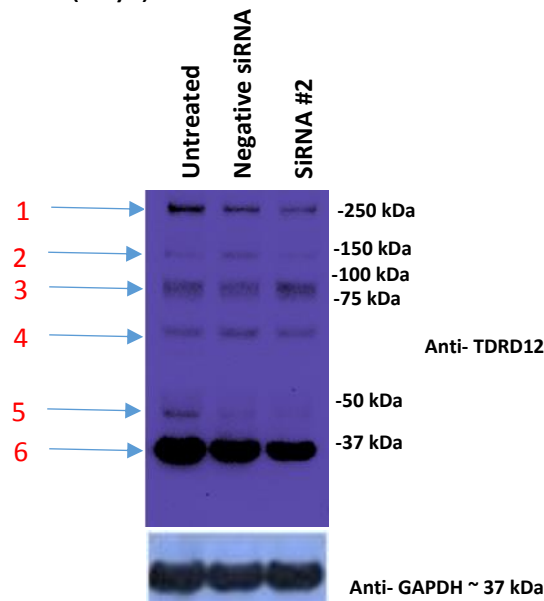


Figure 3.5. NTERA2 cell growth curve for six days of transfection with anti-TDRD12 siRNA #2. (A) The graph shows the daily records of cell numbers for each type of treatment to build the growth curve for six days. There is a significant change in the cell count of positively treated NTERA2 cells compared to the negative control (**: P value < 0.01 and ***: P value < 0.001). P values were obtained from three biological repeats (B) Western blot for TDRD12 protein analysis in NTERA2 cell extracts after six days of treatment from each experiment. The extracts from the cells transfected in the cell growth curve experiment were probed against the anti-TDRD12 antibody. The anti-GAPDH antibody was used as a loading control. Band (1) may be related to TDRD12 variant 6 protein. Bands (2) at 141 kDa and band (5) at 47 kDa which might be relating to transcript variants 2 and 1 respectively. Additionally, other signals (bands 3 and 4) at approximately 70 and 80 kDa emerged that are not significantly reduced, however, they might be resulted from degradations or another unpredicted variants. Finally, the signal below 37 kDa (band 6) emerged showing a slight reduction due to siRNA #2 transfection and it is expected to be a small variant of TDRD12 protein or a degradation product.

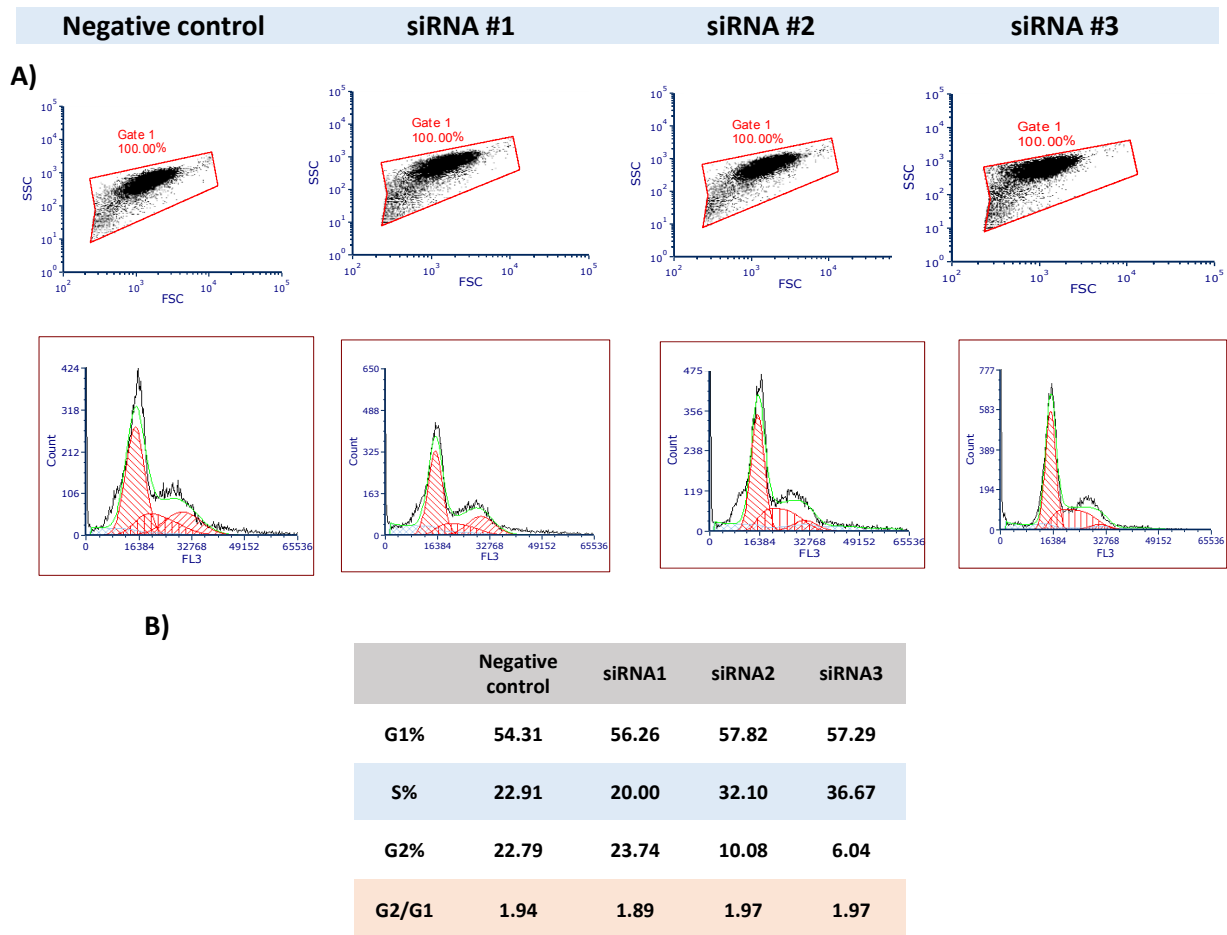


Figure 3.6. The influence of TDRD12 depletion on cell cycle profile.

NTERA-2 cells were treated with either non-interference or anti-TDRD12 siRNAs (#1, #2 and #3). (A) Flow cytometry was carried out to analyse the cell cycle profile in each treated cell by the multicycle cell cycle analysis plug-in for FCS Express 5. Three different TDRD12-silencing siRNAs were used along with the negative control, and the results were categorised in plots for each experiment separately. (B) Percentage of cell cycle stage distribution. This experiment was carried out three replicates from cell culture.

3.2.4. TDRD12 knockdown influences cell cycle regulatory proteins.

The progression through the cell cycle is controlled by the periodic activation of cyclin-dependent kinases (CDKs), which are regulated positively by cyclins and negatively by CDK inhibitors (Serrano et al., 1993; Morgan, 1995; Sherr & Roberts, 1999). D-type cyclins are required for G1 phase progression and reach a peak at the late G1 phase to promote G1/S transition (Won et al., 1992). CDKN1A (also termed p21^{Waf1/Cip1}) participates in many cellular pathways in response to a variety of extracellular and intracellular stimuli. Inhibition of cyclin–CDK activity is the essential function of CDKN1A during the cell cycle. The increase of the CDKN1A level is suggested to mediate G1 and/or G2 cell cycle arrest (Harper et al., 1993). Interestingly, some studies have reported that Cyclin-D–CDK complexes contain the CDKN1A protein, suggesting its role in the initiation of first contact between cyclin D and CDK subunits (Sherr & Roberts, 1999; Perucca et al., 2009).

To further evaluate cell cycle progression, NTERA-2 cells were treated with all three *TDRD12* siRNAs and the negative control. Western blot analysis was carried out using selected antibodies to proteins specific for the cell cycle, as shown in Table 3.1. The results showed an apparent elevation of Cyclin-D type, CDKN1A and CDK4 proteins as shown in (Figure 3.7).

Table 3.1. Panel of cell cycle proteins selected to study their responses following TDRD12 depletion in NTERA2 cells

Protein	Function	Reference
Cyclin D1	progression through G1 phase, activator of CDK4/6 and G1/S transition	(Hunter & Pines, 1994; Sherr & Roberts, 1999)
Cyclin D2	progression through G1 phase, activator of CDK4/6 and M phase regulation	(Hunter & Pines, 1994; Sherr & Roberts, 1999)
CDK4	G1-S phases regulation and progression	(Morgan, 1995)
CDKN1A P21^{Waf1/Cip1}	CDK–cyclin complexes regulation and inhibition	(Harper et al., 1993; Sherr & Roberts, 1999)

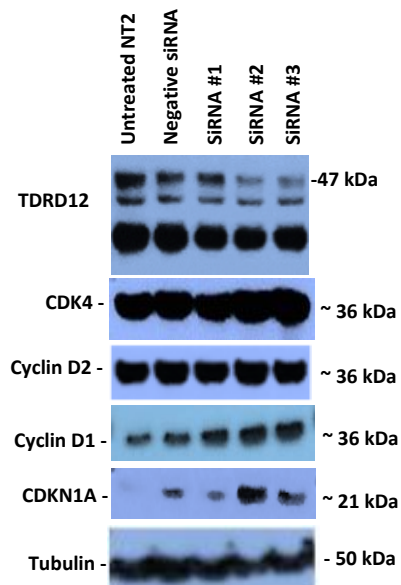


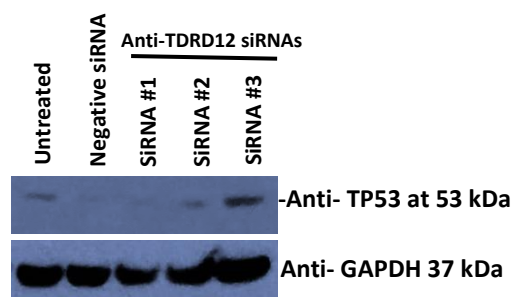
Figure 3.7. Western blot analysis, demonstrating changes in cell cycle protein levels following TDRD12 protein knockdown in NTERA2 cells. The depletion of TDRD12 protein was carried out by the three siRNAs (siRNA #1, #2 and #3). This experiment was performed in three repeats.

3.2.7. TDRD12 knockdown does not induce NTERA-2 cell apoptosis

Apoptosis is a fundamental mechanism whereby cells choose to activate appropriate signals for the death pathway. This process is also induced as a response from physiological and pathophysiological factors to internal or external stimuli, such as aging, irradiation and drugs. Apoptosis is executed by the activation of a series of cysteine proteases called caspases. However, Caspase 3 is the predominant caspase that is involved in both extrinsic and intrinsic apoptotic pathways (Salvesen, 2002; Ghavami et al., 2009). Furthermore, it is well known that the tumour protein 53 (TP53) is reported to mediate apoptosis (Gorospe et al., 1996; Gorospe et al., 1997). It has been reported that TP53 could induce apoptosis by directly regulating anti-apoptotic proteins and pro-apoptotic proteins (Hockenbery et al., 1993; Li et al., 2015a). Additionally, as a transcriptional factor, TP53 activates CDKN1A transcription; this might represent a feedback mechanism to regulate TP53 activity during the apoptotic process (Seoane et al., 2002; Coqueret, 2003).

An experiment was conducted to determine whether the TDRD12-depleted NTERA2 cells induce apoptotic activity. The same lysates that showed TDRD12 protein reductions from previous experiment in Figure 3.7 were used to check TP53 and Caspase 3 proteins. It has been reported that NTERA2 cell has low wild-type TP53 levels (Burger et al., 1998). Our observation showed a possible slight elevation of TP53 protein levels (Figure 3.8 A). However, the assessment of Caspase-3 showed that it is not cleaved to activate and induce apoptosis (Figure 3.8 B).

A)



B)

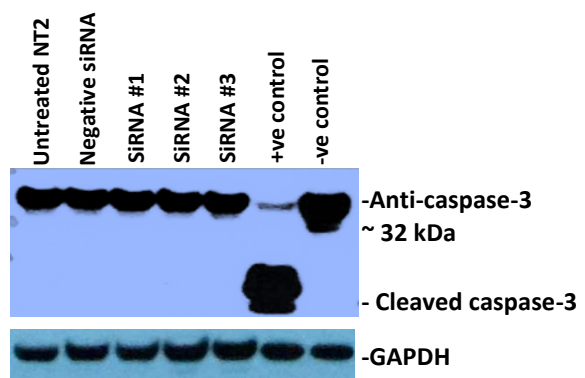


Figure 3.8. Western blot analysis demonstrating the production pattern of apoptotic proteins following TDRD12 protein knockdown in NTERA2 cells. (A) The depletion of TDRD12 protein by many siRNAs showed an elevation of TP53 protein from one experimental repeat. (B) The reduction of TDRD12 protein has no influence on Caspase-3, as cleaved Caspase-3 was non-detectable in positively treated cells by anti-*TDRD12* siRNAs. Positive and negative controls were supplied with antibody. Triplicate repeats was carried for this experiment from whole cell protein extracts.

3.2.8. TDRD12-depleted NTERA2 cells does not induce senescence.

The induction of CDKN1A during cell cycle arrest has been reported by several studies as an indication of common physiological mechanisms, such as senescence and terminal differentiation (Noda et al., 1994; Brown et al., 1997; Zheng et al., 2006; Herbig & Sedivy, 2006). Therefore, NTERA2 cells were also evaluated for the senescence state before and after TDRD12 knockdown. NTERA2 cells were employed for three hits of transfections before staining. The senescence was examined by lysosomal senescence-associated β -galactosidase activity at pH 6. The results suggested that NTERA2 cells depleted of TDRD12 did not induce the senescence state as no staining was detected in positively treated cells (Figure 3.9).

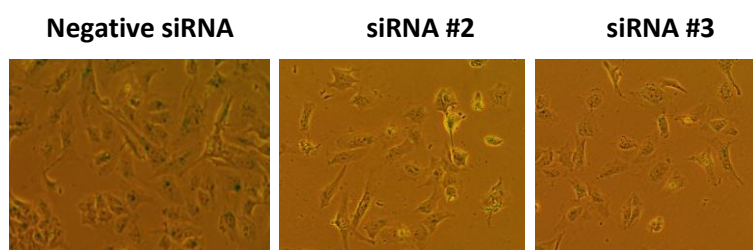


Figure 3.9. Senescence analysis of *TDRD12*-depleted NTERA2 cell lines.

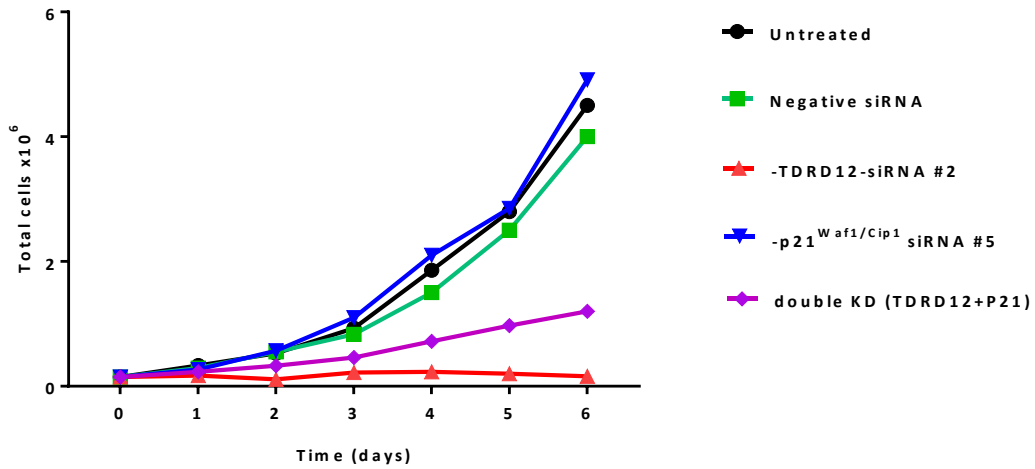
β -galactosidase staining was used to evaluate whether NTERA2 cells that are treated by anti-TDRD12 siRNAs #2 and #3 induce the senescence state compared to untreated cells and the negative control (negative siRNA). The images show that there is no significant changes of β -galactosidase staining in treated and untreated cells. The images were using an ECLIPSE-inverted microscope (5x lens). Senescence experiment was performed twice from cell cultures.

3.2.9. Double knockdown of *CDKN1A* and *TDRD12* rescues proliferation potential in NTERA2 cells.

Based on the observed elevation of *CDKN1A* levels following *TDRD12* knockdown, the NTERA2 cell growth curve was established under conditions of *TDRD12* and *CDKN1A* double knockdown to determine whether simultaneous *CDKN1A* transcript depletion could rescue the inhibition of proliferation that was reported in single *TDRD12* knockdown cells. This experiment was further aimed at determining the link between *TDRD12* and *CDKN1A* during the cell cycle. NTERA2 cells were plated separately in five different cultures depending on treatment: untreated, negative siRNA, anti-*TDRD12* siRNA #2, *CDKN1A* siRNA #5 and double treatment of *CDKN1A* and *TDRD12* siRNAs. The daily records of cell counts were documented to establish the proliferation curve (Figure 3.10 A), the cell morphology and confluency were also depicted (Figure 3.10 B), and the efficiency of knockdown was evaluated by qRT-PCR analysis (Figure 3.11).

The results revealed that cell proliferation inhibition was partially rescued under the condition of *TDRD12* and *CDKN1A* double knockdown. The proliferation rate of NTERA2 cells was observed to improve in double knockdown cells through the assessment of increased cell counts and increased confluency of attached cells in comparison to untreated and non-siRNA (as negative control) and single *TDRD12*-depleted (as positive control) cells in this experiment. The qRT-PCR analyses in (Figure 3.11 A and B) showed that the downregulation of *TDRD12* transcripts in a single anti-*TDRD12* siRNA treated culture has led to an elevation of *CDKN1A* transcript levels, thereby, the proliferation was arrested. However, in the double knockdown culture (anti-*TDRD12* and anti-*CDKN1A* siRNAs) the upregulation of *CDKN1A* transcript levels return back to normal and the cells start proliferation although the levels of *TDRD12* transcripts were still downregulated. Furthermore, the results suggested that the *CDKN1A* might be a downstream target of *TDRD12* gene.

A)



B)

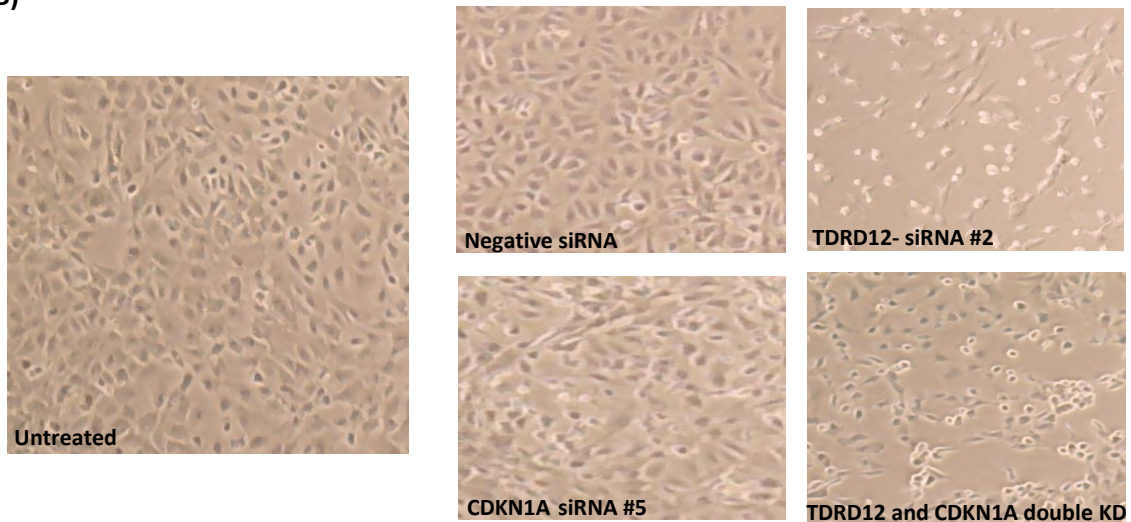
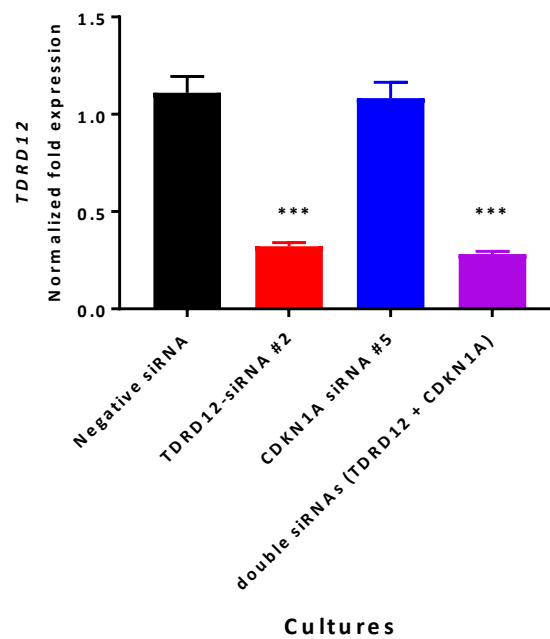


Figure 3.10. Cell proliferation curve and microscopic examination for NTERA2 cell proliferation following double knockdown of *TDRD12* and *CDKN1A*. (A) NTERA2 cells were plated as untreated, negative control, single knockdown *TDRD12*, single knockdown *CDKN1A* and double knockdown *TDRD12* + *CDKN1A* for six days of transfection, and cells were counted. (B) The microscopic examination showing the number of cells and confluency after six days of treatment. Single knockdown *TDRD12* showed a very low number of cells at about 25% confluency; however, cell proliferation was induced and confluency was increased to approximately 50% in *TDRD12* and *CDKN1A* double knockdown. Single *CDKN1A* knockdown showed no change in cell proliferation as well as negative control and untreated cells. The images were depicted before trypsinisation to evaluate the cell density on day 6 using an ECLIPSE-inverted microscope (5x lens). Only one result was represented here from duplicated experiment repeats.

***TDRD12* expression**



***CDKN1A* expression**

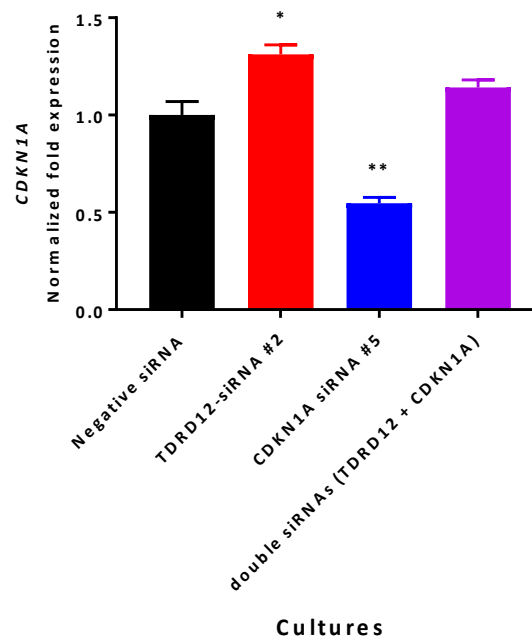


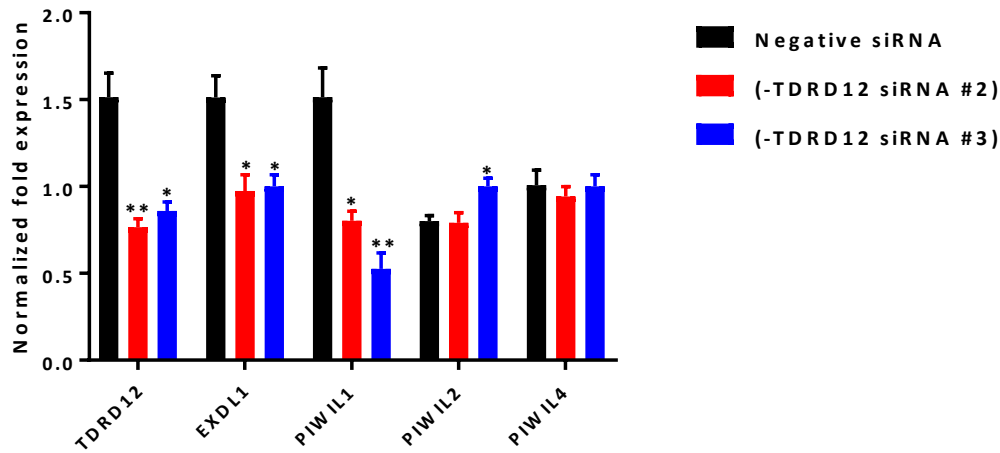
Figure 3.11. qRT-PCR analysis for mRNA levels of *TDRD12* and *CDKN1A* following double knockdown experiment. Bar charts show the levels of mRNA transcripts of the *TDRD12* gene and by qRT-PCR in NTERA2 cells following double gene knockdown. Each gene expression analysis was performed separately and normalised to *GAPDH* and *ACTB*. The fold change was computed using the $\Delta\Delta C_t$ method, and the error bars refer to the standard errors of the mean. Error bars and statistics are derived from three technical repeats not biological repeats (A) The expression levels of *TDRD12* transcripts. (B) The expression levels of *CDKN1A* Transcripts. P values showed significant changes in comparison to the control (negative siRNA) treatment. (***) *P* value < 0.001, (**) *P* value < 0.01 and (*) *P* value < 0.05).

3.2.10. *TDRD12* is involved in complexes of genes that regulate cancer stem cell stability.

The *TDRD12* mRNA level was downregulated in NTERA2 using two siRNAs (anti-*TDRD12* siRNA #2 and #3) prior to the analysis of *TDRD12* expression influence on other genes. Furthermore, human *EXDL1*, which encodes a partner of *TDRD12*, was also depleted in NTERA2 using anti-*EXDL1* siRNAs #2 and #4 to investigate its effect on *TDRD12* transcript levels and other human *PIWIL* gene transcripts. NTERA2 cells express *PIWIL1*, *PIWIL2* and *PIWIL4* from the *PIWIL* gene family, but the expression of *PIWIL3* was not detected in both negatively and positively treated cells, so it was excluded in this study.

The results of qRT-PCR analysis (Figure 3.12 A), demonstrated a significant reduction of *TDRD12* expression in NTERA2 cells treated with anti-*TDRD12* siRNA #2 and #3 (P value < 0.01). The results showed significant reductions in transcripts from *PIWIL1* (P value < 0.01) and *EXDL1* (P value < 0.05) in both anti-*TDRD12* siRNAs-treated cultures. There is no significant changes due to *TDRD12* reduction on the other *PIWIL* genes (*PIWIL2* and *PIWIL4*). The results in (Figure 3.12 B) showed a marked increase in *PIWIL1* transcripts (P value < 0.0001) and *TDRD12* transcript levels (P value < 0.0001) in NTERA2 depleted of *EXDL1* transcripts. Moreover, a significant elevation of *PIWIL2* transcripts (P value < 0.01) was also observed due to *EXDL1* depletion in NTERA2 cells. The results may show that there is relationship between *TDRD12*, *EXDL1* and *PIWIL* genes at the transcription level in NTERA2 cells. Finally, the protein-protein interaction is reported in the PET complex in mice, but no study has been investigated this interaction in human. However, protein interactions databases predict the PET complex (*PIWIL*-*EXDL1*-*TDRD12*) proteins interaction in human as shown in (Figure 3.13).

A)



B)

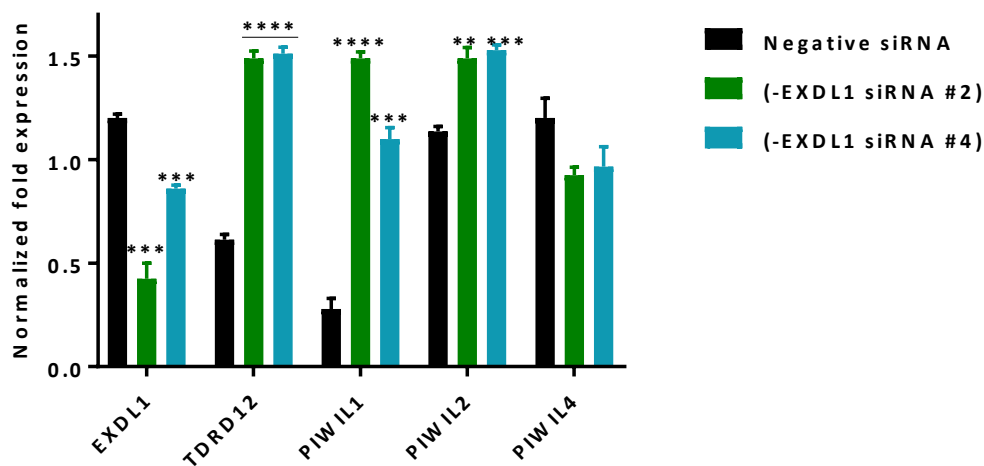


Figure 3.12. The expression pattern of *TDRD12*, *EXDL1* and *PIWIL* family genes in NTERA2 cells. Analyses of qRT-PCR show the expression pattern changes of the *TDRD12*, *EXDL1* and *PIWIL* family genes in two conditions of NTERA2 cell knockdown. The expression of genes was normalised to a combination of two endogenous genes (*ACT-β* and *GAPDH*). The fold change was computed using the $\Delta\Delta C_t$ method, the error bars refer to the standard errors for the mean of three technical repeats and the Y axis scale is linear. These experiments were repeated three times from RNA extracts. (A) The bar chart shows the expression pattern of these genes in NTERA2 depleted of the *TDRD12* gene by two different treatments (siRNA #2 and #3). (B) The bar chart shows the expression pattern of these genes in NTERA2 depleted of *EXDL1* gene by two different treatments (siRNA #2 and #4) (*= P value < 0.05 ** = P value < 0.01, *** = P value < 0.001, **** = P value < 0.0001).

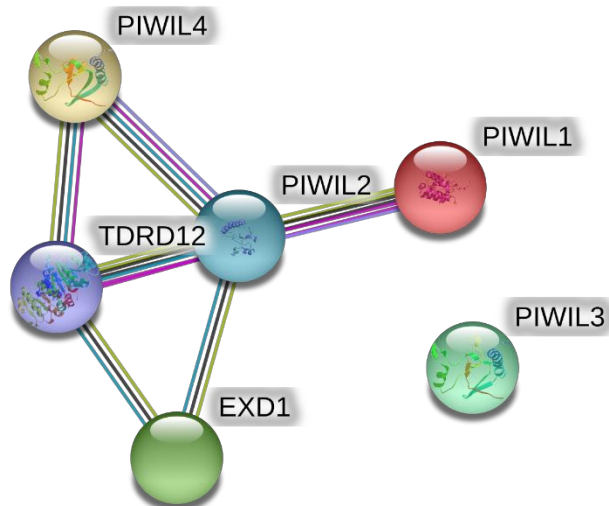


Figure 3.13. Predicted interactions between TDRD12, EXD1 and PIWIL family proteins in human. STRING protein databases predict the interaction between TDRD12, EXD1 and PIWIL proteins. This interaction is predicted as results of neighbouring genes or genes co-expression regulation, see (string-db.org).

Known interaction



Curated databases



Experimentally determined

Predicted interactions



Gene neighbouring



Gene fusions



Gene co-occurrence

Other



Textmining



Co-expression



Protein homology

3.2.11. *TDRD12* knockdown regulates transposon element *LINE-1* in NTERA2 cells.

The level of *TDRD12* mRNA was successfully downregulated using *TDRD12* siRNA #2 in cancer stem cells (NTERA2). To explore the relationship between *TDRD12* and non-LTR transposon elements (*LINE-1* and *SINE*), qRT-PCR analysis was carried out to evaluate the changes in transcript levels for these TEs following *TDRD12* knockdown (Figure 3.14). The initial screening of *LINE-1* transcripts showed a significant downregulation in transcript levels (P value < 0.05) after 72 hours of *TDRD12* siRNA #2 transfection in NTERA2 cells. The level of *SINE* transcripts was not affected, as it showed the same levels as the negative control.

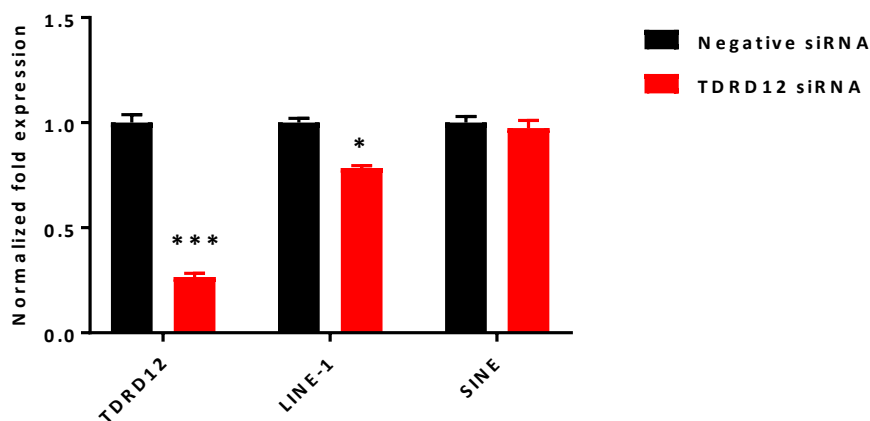


Figure 3.14. qRT-PCR analysis of transposon elements (*LINE* and *SINE*) following *TDRD12* depletion in NTERA2 cells. The bar charts show the levels of mRNA transcripts of *TDRD12*, *LINE* and *SINE* genes by RT-qPCR in NTERA2 cells following *TDRD12* depletion. Gene expression was normalised to *GAPDH* and *ACTB*. The fold change was computed using the $\Delta\Delta C_t$ method, and the error bars refer to the standard errors for the mean of technical triplicates. (***) P value < 0.001, * P value < 0.05).

3.3. Discussion

Previous studies have shown that Tdrd12 is essential for germ cell development and genomic stability in mice (Chuma et al., 2006; Shoji et al., 2009; Yabuta et al., 2011; Pandey et al., 2013; Yang et al., 2016). The roles of Tdrd12 have previously been identified in germinal cell migration, proliferation and maintenance in the early stages of mouse germ cells (Dai et al., 2017), while mutants lacking Tdrd12 resulted in male sterility and atrophied testes (Pandey et al., 2013). The results in this chapter explore the functional roles of human TDRD12 protein on the NTERA-2 cell cycle progression. Additionally, results here also highlight the involvement of the *TDRD12* gene in regulatory mechanisms on other genes in NTERA2 cells.

3.3.1. Characterisation of TDRD12 splice variants and antibody specificity.

TDRD12, is highly conserved protein in eukaryotes, composed of a central DEAD (helicase) domain in between of two Tudor domains along with CS domain in C-terminal (Pandey et al., 2013). Five splice variants are predicted for *TDRD12*, four variants termed as TV1, TV2, TV4 and TV5 are potential coding protein variants. Current study used a polyclonal antibody from Abcam (ab182463), which is specific to TDRD12 protein TV1 and TV2 as the sequence of the antibody was matched with the corresponding TDRD12 transcript sequences as shown in (Figure 3.2) and (Appendix A-2). Antibody and siRNAs confirmed the detection of TDRD12 TV1 at the expected size of 47 kDa and TV2 at 141 kDa. The higher molecular weight band at approximately 203 kDa appeared to be specific because it was reduced due to the knockdown. Based on previous work conducted by the McFarlane group, TV6 transcript was predicted to encode a protein of approximately 183 kDa. However, the identification of exact sizes for high molecular proteins might be not accurate or this protein may undergo posttranslational modifications (PTMs) that led to increase its molecular weight such as ubiquitination.

3.3.2. TDRD12 regulates cancer stem cell proliferation.

In normal cells, the cell cycle process is tightly regulated through a balance of activation and/or inhibition between genes that promote cell proliferation and others that suppress it, to ultimately achieve cell division. Disruption of this balance is one of the hallmarks that initiate cancer development, when cells begin uncontrolled division (Hanahan & Weinberg, 2000; 2011). The downregulation of CT genes, which are essential for cell division and proliferation or which at least contribute to the developmental mechanisms of cancer, was

reported to induce an apoptotic pathway or slow proliferating rates (Assanga & Lujan, 2013; Shange et al., 2014). For instance, knockdown of PHF20L1 stops cell proliferation in breast cancer cell lines (Jiang et al., 2016; Carr et al., 2017). Moreover, the downregulation of *TEX19*, a CT-restricted gene, was reported to inhibit the proliferation of colorectal carcinoma cells (Planells-Palop et al., 2017). The monitoring of cell proliferation and the effects of drugs on the cell cycle can be assessed through the establishment of growth curve assays (Assanga & Lujan, 2013). Therefore, establishing cell growth curves is useful for understanding how cell proliferation occurs during *TDRD12* depletion. An initial screen of cell counts after the depletion of *TDRD12* using three sequences of siRNAs showed that siRNA (#2 and #3)-treated cells were not actively proliferating. In fact, the depletion of NTERA2 cells with anti-*TDRD12* siRNAs #2 and #3 showed significant differences in total cell counts when compared to non-interference (negative control) and untreated cells. The reduction of mRNA transcripts was confirmed by qRT-PCR in all tested siRNAs; however, protein reduction was observed only in independent cultures of siRNA #2- and #3-treated cells. This may suggest that the *TDRD12* protein is essential for NTERA2 cell proliferation. The growth curve analysis clearly demonstrated that the cell cycle is nearly arrested under conditions of *TDRD12* depletion. Flow cytometric analysis confirmed an uneven distribution of treated cells throughout the cell cycle sub-phases when compared to control NTERA2 cells. In *TDRD12* knockdown, NTERA2 cells accumulate in the G1/S phase and a lower proportion of cells undergo mitosis. This implies that *TDRD12* is required for cell cycle progression in NTERA2 cells. *TDRD12* may be able to be conferred to cells in order to decide whether to progress in proliferation or exit the cell cycle through different mechanisms, such as quiescence, senescence or apoptotic pathways. To examine whether *TDRD12* is also part of any of these mechanisms, the influence of *TDRD12* on known cell cycle regulators and quiescence-related proteins was further tested following *TDRD12* knockdown.

First, some cell cycle proteins that are known to control the G1 phase were chosen to be examined. Cyclin-D types function in G1/S transition that activated by CDK4/6 and inhibited by cyclin kinase inhibitors (CKIs), such as; p21^{Waf1/Cip1} (CDKN1A) and/or p27^{Kip1} (CDKN1B) (Harper et al., 1993; Sherr & Roberts, 1999). The results showed that the reduction of *TDRD12* led to an increase of Cyclin-D1, CDK4 and CDKN1A. This confirms cell cycle arrest at a late G1 phase, and an accumulation of Cyclin-D1 protein was reported to promote the transition to the S-phase (Harper et al., 1993). Results showed an activation of TP53 in response to *TDRD12*

depletion. Given that, this is consistent with many studies have suggested that CDKN1A mediates TP53-dependent G1 growth arrest (Deng et al., 1995; Brugarolas et al., 1995; Abbas & Dutta, 2009). Apart from the function of CDKN1A as a cyclin–CDK inhibitor, it was found to be elevated as a response to protect arrested growth cells from apoptosis unless it is degraded by a Caspase-3 protein (Fujiwara et al., 2008; Zhang et al., 1999).

Apoptosis, or programmed cell death, is regulated by many different proteins and pathways as a response to intracellular and extracellular stimuli. However, Caspase-3 is the main core protein in many apoptotic pathways (Salvesen, 2002; Ghavami et al., 2009). Moreover, CDKN1A is transcriptionally regulated by the TP53 protein in response to DNA damage (Harper et al., 1993). Alternatively, CDKN1A also responds to signals independent of TP53 and has been determined to play roles in apoptosis and senescence (Shao et al., 1995; Zengel-Deiry, 1996). Additionally, TP53 has been identified to transcriptionally target CDKN1A, which plays roles in cell growth arrest and senescence (Abbas & Dutta, 2009). In this study, apoptosis and senescence were examined in TDRD12-depleted NTERA2 cells to confirm the status of these cells. The cells do not undergo apoptosis, as determined by Caspase activity showing no cleaved caspase-3 signals. These results, alongside the results from the cell growth curve assay, may confirm that TDRD12-depleted NTERA2 cells are still viable and that the reduction of this protein may not trigger apoptotic pathways. On the other hand, the senescence state was assessed by β -galactosidase activity assay, showing that TDRD12-depleted NTERA2 cells do not induce senescence either. However, the elevation of CDKN1A together with TP53 was observed to establish and maintain the quiescence-like state in human cells (Itahana et al., 2002; Perucca et al., 2009). Therefore, the knockdown of TDRD12 in NTERA2 cells may induce these cells to enter a quiescence-like state (G0), in which the viable cells are still present but do not show signs of progression or division as determined in cell counts assay.

The reduction of *TDRD12* mRNA results in the upregulation of *CDKN1A* mRNA levels, leading to cell growth arrest, but the single knockdown of *CDKN1A* mRNA neither affect the *TDRD12* mRNA levels nor cell proliferation. This may suggest that *CDKN1A* is a downstream target of the *TDRD12* gene. However, double knockdown of *TDRD12* and *CDKN1A* was observed to partially rescue cell proliferation inhibition and increase the confluency percentage of attached cells with normal levels of *CDKN1A* transcripts, which may confirm that *TDRD12*-depleted NTERA2 cells were under quiescent conditions. Finally, we suggested that TDRD12 exhibited an important role for NTERA-2 cells to divide and proliferate, however we do not

know to what extent these cells require the production of TDRD12. To address this, more investigations are required such as knockout system.

3.3.2. TDRD12 regulates other germ genes in NTERA2 cells.

Piwi-interacting (pi-RNAs) RNAs are small, germ-specific RNAs that regulate the activation and silencing of genes in germ cells to maintain genomic stability and provide a defence against transposable elements (TEs). However, the molecular mechanisms and complete set of factors involved in this process remain unclear. Furthermore, many studies reported that Tudor domains interact with methylated lysine and arginine in target protein thereby facilitating the assembly of protein complexes (Liu et al., 2010; Tripsianes et al., 2011). Previous studies on mice and insects found that *Tdrd12* is dispensable for the production of primary pi-RNA, but essential for secondary piRNA biogenesis, and is achieved through the implication of protein–protein interactions in a complex: *Tdrd12*-*Tdrd1*-Piwi and its associated pi-RNAs (Pandey et al., 2013). Another study determined that *Exd1* is a partner to *Tdrd12* and is also essential in secondary piRNA production through involvement in a PET (Piwi-*Exd1*-*Tdrd12*) complex (Yang et al., 2016).

This study was based on the screening of mRNA expression changes in each gene within the PET complex during the depletion of other genes in NTERA2 cell lines. TDRD12-depleted NTERA2 cells show reductions in both *EXDL1* and *PIWIL1*. Moreover, knockdown of *EXDL1* transcripts leads to increased levels of *TDRD12*, *PIWIL1* and *PIWIL2* transcripts. In fact, these results suggested that there is a relationship at transcriptional levels for these genes in NTERA2 cells. It is proposed that these genes are involved in the same biological process that may govern the regulation of these genes although it is not clear to identify which regulatory and regulated gene. Furthermore, many studies reported the co-expression of gene partners under several conditions such as disease state (Hu et al., 2009), tissue types (Pierson et al., 2015), and developmental stages (Xue et al., 2013), because these genes are more likely to be regulators that underlie phenotypic differences. The co-expression regulation of such genes is possibly resulted from other regulatory mechanisms such as protein–protein interactions and/or transcription factors (TFs) and their targets' interactions (Glass et al., 2013; van Dam et al., 2017). However, protein–protein interactions in the PET complex was previously reported in murine (Pandey et al., 2013; Yang et al., 2016), in addition, these interactions are also predicted in human as shown in (Figure 3.12). Given that, the PET (PIWI-

EXD1-TDRD12) complex remains standing in NTERA2 cells and further analyses are required to conclude the PET complex proteins assembly.

3.3.3. *TDRD12* regulates the transposon element *LINE-1*.

Pandey and co-workers reported that the loss of Tdrd12 in mice leads to de-repression of transposon element *LINE-1*. They examined the activation of *LINE-1* through the confirmation of reduced DNA methylation (to approximately 53%) at its promoter and the reduced of pi-RNAs production (Pandey et al., 2013). However, the depletion of human *TDRD12* transcripts in NTERA2 cells showed an unexpectedly reduction of transposable element *LINE-1*, which is inconsistent with the previous study. Despite many conditional variations in the experiments of this study and the previous study, the difference in molecular mechanisms between mice and humans is possible, as the comparison with different species is sometime not applicable. On the other hand, we observed in this chapter that TP53, a tumour suppressor protein, appeared to be elevated following *TDRD12* protein knockdown (Figure 3.8 A). This indication may uncover the functional roles of *TDRD12* in cancer development and maintenance in NTERA-2 cells. It has been demonstrated that the activation of germline genes in cancerous cells play critical roles in their development, maintenance and invasion (Greve et al., 2015; Yang, 2014; Shang et al., 2014; McFarlane et al., 2015; Feichtinger et al., 2014). The regulation of germ genes might not be changed in cancerous cells, however the functional roles may be deviated as a response for new environment; reviewed in (McFarlane & Wakeman, 2017). Results here showed that the levels of *TDRD12* is correlated with *LINE-1* activity, thus, the oncogenic function of *TDRD12* in NTERA-2 cells is proposed. As the NTERA-2 cell is a cancer stem cell model, we cannot conclude that *TDRD12* is required for stemness and/or cancerous development and stability. Finally, there is a significant correlation between *TDRD12* and transposable element *LINE-1* and it is suggested that *TDRD12* plays important roles in *LINE-1* activity in tumour germ cells.

3.3.4. Conclusion

TDRD12 is a germline protein that is produced in cancer stem-like cells NTERA-2. The presence of TDRD12 protein in NTERA-2 is essential for the cell division and proliferation. The interrelation between the activities of *TDRD12* mRNA transcripts influence the activity of other germline genes *EXDL1* and *PIWIL* family genes, suggesting that they are involved in the same biological process. Additionally, TDRD12 protein interaction in the PET (PIWI-EXD1-TDRD12) complex remains standing in NTERA2 cells. The oncogenic function of *TDRD12* is proposed hence its transcriptional activities is affecting the activity of transposable element *LINE-1* and its protein production is observed to have influence on the tumour suppressor protein p53 levels.

Chapter 4:

The potential roles of *TDRD12* gene in human embryonic stem cells.

4. The potential roles of TDRD12 gene in human embryonic stem cells.

4.1. Introduction

The main aim of this chapter was to explore whether human Tudor domain containing 12 (*TDRD12*) expression is correlated with features of stemness. In this study, the expression of the *TDRD12* gene was investigated and compared to a stemness marker in normal stem cells (SC), cancer stem-like cells (CSCs) and induced pluripotent stem cells (iPSCs). SCs possess the capacity for self-renewal and the potential for differentiation to more specialised cell lineages during the embryogenesis process. Moreover, CSCs have stem cell-like features and have been identified as a subpopulation of cells within many tumours. CSCs are postulated to play roles in tumour drug resistance and tumour recurrence (Moore & Lyle, 2011; Gedye et al., 2016; Aponte & Caicedo, 2017). The discovery of CSCs led to improvements in the development and design of oncology treatments and drug efficacies. The identification of new markers for CSCs is required to enable the targeting of these cells. The iPSCs are another type of SCs that can be established *in vitro*. The iPSCs are derived from somatic adult cells that are reprogrammed to pluripotent stem cells (Takahashi et al., 2007). These iPSCs resemble hESCs in morphology, proliferation and gene expression, especially in their stemness markers (*OCT4*, *NANOG* and *SOX2*).

Malignant tumours are composed of a mixture of heterogeneous cells, and the heterogeneity within the tumour bulk might be a result of differentiated and/or dedifferentiated CSCs. The embryonal carcinoma stem cells NTERA2 cells have cancerous and stemness characteristics, and they can be differentiated *in vitro* by exogenous stimuli (Pera et al., 1989; Andrews et al., 1990; Przyborski, 2001). The exposure of NTERA2 cells to compounds such as retinoic acid (RA) results in a heterogeneous culture of differentiating cells with a bias to the neuronal lineages, whereas their response to hexamethylene bisacetamide (HMBA) produces a more homogenous development. RA is a vitamin A metabolite that is normally involved in various processes during embryogenesis, such as proliferation and differentiation. In differentiation, RA directs stem cells to produce neural lineage precursors. RA-induced NTERA-2 cells were shown to upregulate molecules associated with neuroectodermal derivatives (Andrews et al., 1990; Przyborski, 2001; Eisenbarth et al., 1979; Przyborski et al., 2000; Stewart et al., 2003). In contrast, HMBA-induced NTERA-2 cells were observed to lack markers for neuron

maturation in the culture, though they did express antigens associated with the formation of epithelial derivatives. Similarly, human embryonic stem cells (hESCs) can be generated *in vitro* and generally rely on embryonic fibroblast feeder to maintain their undifferentiated state. These cells, however, can be differentiated by RA to produce normal neuronal cells. Several markers can be used to check the undifferentiated phenotype of pluripotent stem cells in hESCs or CSCs. For example, *OCT4* is highly expressed in hESCs and CSCs stem cells but not in differentiated derivatives (Pera et al., 2000; Przyborski, 2001). *OCT4* plays fundamental roles in maintaining pluripotency in stem cells and its downregulation results in dramatic differentiation (Zafarana et al., 2009). Furthermore, it is suggested that *PAX6* is expressed when stem cells are differentiated into neuronal cells (Callaerts et al., 1997).

The *TDRD12* gene has been characterised as a candidate cancer biomarker that may have CSC-specific activity (Feichtinger et al., 2012). In the previous chapter, the expression of *TDRD12* was suggested to influence other germline genes, such as *EXDL1*. There was a correlation at transcription level between *TDRD12* and *EXDL1* genes. The identification of new cancer testis antigens (CTA) is useful in the fields of cancer diagnosis and immunotherapy thereby, the work of this chapter aimed to further explore *EXDL1* and *TDRD12* expression and their association with stemness.

4.2. Results

4.2.1. Analysis of *TDRD12* expression in human embryonic stem cells (hESCs) and induced pluripotent stem cells (iPSCs).

The findings that *TDRD12* is expressed in NTERA2 cells may infer that it could be linked with the stemness characteristics of these cells. So, human *TDRD12* expression pattern was investigated on different stem cells and iPSCs using qRT-PCR analysis. The expression patterns of the *TDRD12* gene were compared to the expression *SOX2* as a stemness marker in this experiment. The SC panel contains NTERA2 cell RNA (as a positive control for the expression of *TDRD12*), commercial calibrator RNA (as a positive control for stem cells), hESC RNA, and iPSC RNA. The iPSCs were generated in the McFarlane lab from human primary fibroblast cells that were reprogrammed by stemness factors. Additionally, fibroblasts were used as a negative control for iPSCs in this experiment as they are negative for stemness marker gene expression. Results in Figure 4.1 show the expression pattern for *TDRD12* in comparison to the expression of *SOX2*. The pattern of *TDRD12* transcript levels match the pattern of *SOX2* transcripts in ESCs. The levels of *TDRD12* transcripts were observed to be relatively low in iPSCs compared to the *SOX2* transcripts, but both genes were clearly undetectable in fibroblasts (negative control). This may suggest that these iPSCs may not possess some stemness properties, or the *TDRD12* might show a weak activation to stemness induction. This result reveals the relationship of the *TDRD12* gene with stemness markers, especially *SOX2*, possibly suggesting that *TDRD12* might play a functional role in stem cells.

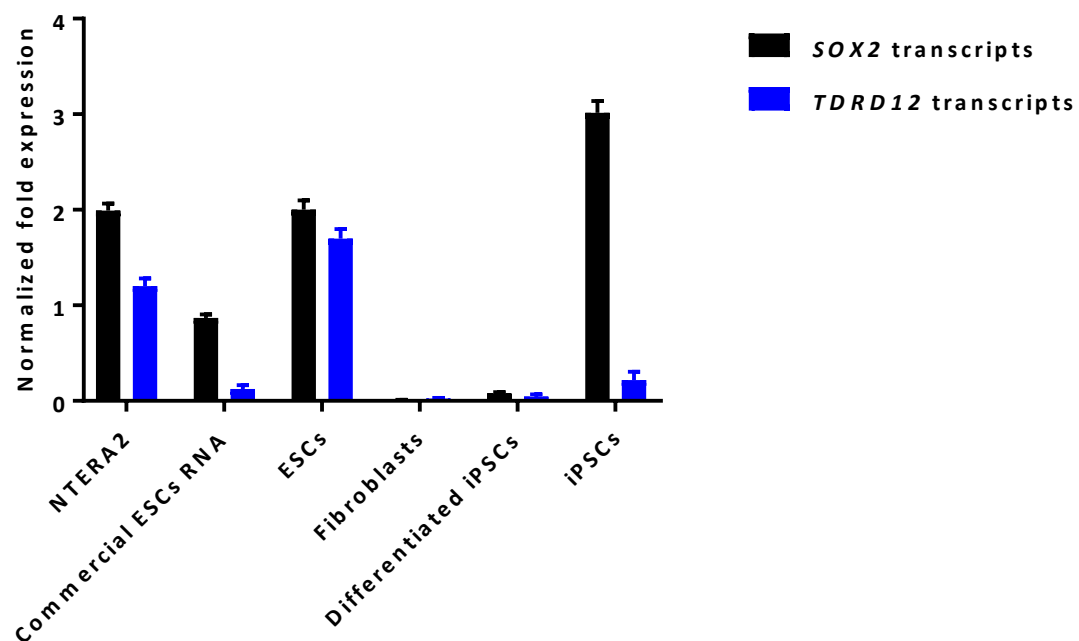


Figure 4.1. Expression profiles of *TDRD12* and *SOX2* in stem cells and iPSCs using qRT-PCR analysis. Bar chart showing expression levels of *TDRD12* in human embryonic stem cells in comparison to the expression of *SOX2*, a stemness marker. The obtained data were normalised to two endogenous reference genes (*HSP90AB1* and *GAPDH*), and the relative fold change was computed by the $\Delta\Delta C_t$ method. Error bars denote the standard error for the mean of three technical repeats. NTERA2 cell RNA was used as positive control for stem cells markers and the studied gene. Fibroblast RNA was used as a negative control.

4.2.2. The expression pattern of stemness marker genes in *TDRD12*-depleted NTERA2 cells

As the previous work suggested that *TDRD12* might be linked to stem-like function, the following experiment was conducted to investigate the influence of *TDRD12* gene on selected stem cell marker genes *OCT4*, *SOX2* and *NANOG*. The levels of *OCT4*, *SOX2* and *NANOG* transcripts were assessed following the depletion of *TDRD12* to determine any regulatory link. NTERA2 cells were depleted of *TDRD12* transcripts using siRNA #2 for three days of transfection. Results in Figure 4.2 showed that *TDRD12* transcript levels were significantly reduced compared to the negative control (non-interference RNA) ($P < 0.0001$) as anticipated. The mRNA levels of *OCT4* were constant showing no changes following *TDRD12* transcripts depletion. However, there is significant reductions in transcript levels of *SOX2* ($p < 0.01$) and transcript levels of *NANOG* ($P < 0.0001$) in *TDRD12*-depleted NTERA2 cells. This may suggest that *SOX2* and *NANOG* genes are functionally linked to *TDRD12*.

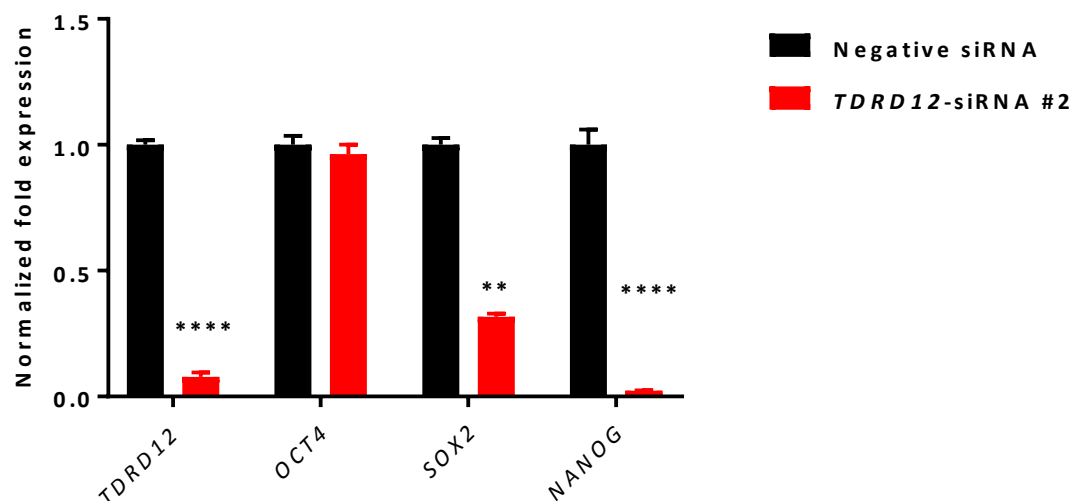


Figure 4.2. Pattern of stem cell marker gene transcripts in human *TDRD12*-depleted NTERA2 cells using qRT-PCR analysis.

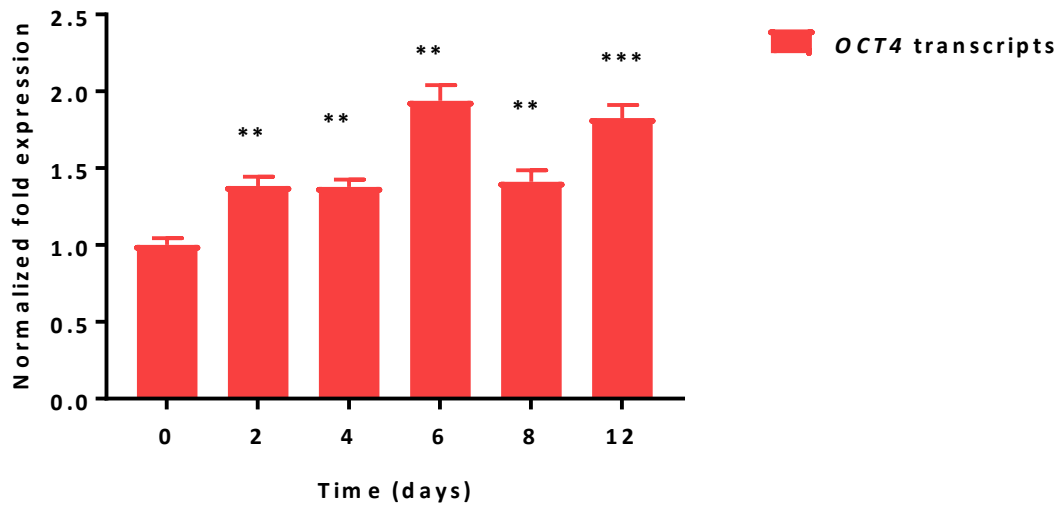
Bar chart showing levels of transcript of stemness marker genes *OCT4*, *SOX2* and *NANOG* following the depletion of *TDRD12* transcripts in NTERA2 cells. The obtained data were normalised to two endogenous reference genes (*ACTB* and *GAPDH*), and the relative fold change was computed by the $\Delta\Delta C_t$ method. Error bars denote the standard errors for the mean of three technical replicates. This experiment was performed in three biological repeats from knockdown. Asterisks above the bars refer to the p-values (**: $p < 0.01$, ****: $p < 0.0001$).

4.2.3. *TDRD12* expression during NTERA2 cell differentiation.

The previous work suggested that the expression of *TDRD12* might be associated with stemness features. This work aimed to monitor the *TDRD12* transcript levels during the differentiation process. NTERA2 cells are embryonal carcinoma stem cells that possess stemness characteristics and can be differentiated *in vitro*. In this experiment, two agents were used to differentiate NTERA2 cells; RA and HMBA; DMSO treatment was used as negative control along with untreated NTERA2 cells. NTERA2 cells were differentiated for 12 days post-treatment and transcript levels were analysed at days 0, 2, 4, 6, 8, and 12. Transcripts of the stemness marker gene *OCT4* were used to monitor the differentiation of NTERA2 cells during the treatment.

The differentiation of NTERA2 cells was successfully achieved. The transcript levels of *OCT4* and *TDRD12* were highly expressed in untreated and DMSO-treated cultures as positive controls (Figure 4.3 and 4.4). Statistical analyses showed increased levels of *OCT4* in untreated cells, but this experiment was to ensure that NTERA2 cells did not undergo differentiation during this experiment in response to other factors and the levels of *OCT4* were still high until the Day 12. In line with this, cells were not differentiated due to DMSO treatment. However, the analyses of *OCT4* levels in RA-treated NTERA2 cells showed significant reductions starting from day 4 post-treatment in comparison to Day 0 in the same experiment, which is not treated cells ($p < 0.0001$) and were undetectable from day 6 post treatment (Figure 4.5 A). This is an indication of the successful differentiation of NTERA2 cells. Additionally, the levels of *TDRD12* mRNA exhibited a significant decline after 12 days of RA treatment ($p < 0.05$) as shown in Figure 4.5 B. NTERA2 cells were successfully differentiated due to HMBA induction as determined through the analysis of *OCT4* mRNA levels in this culture. The significant reductions of *OCT4* levels start from day 4 ($p < 0.01$) while levels were undetectable by day 8 (Figure 4.6 A). Furthermore, *TDRD12* mRNA levels showed significant decreases ($p < 0.05$, $p < 0.001$ and $p < 0.01$) respectively in response to HMBA treatment at days 6, 8 and 12 (Figure 4.6 B). These results suggest that *TDRD12* expression could be correlated with the stem cell marker gene *OCT4*, and might be also required for stemness characteristics.

A)



B)

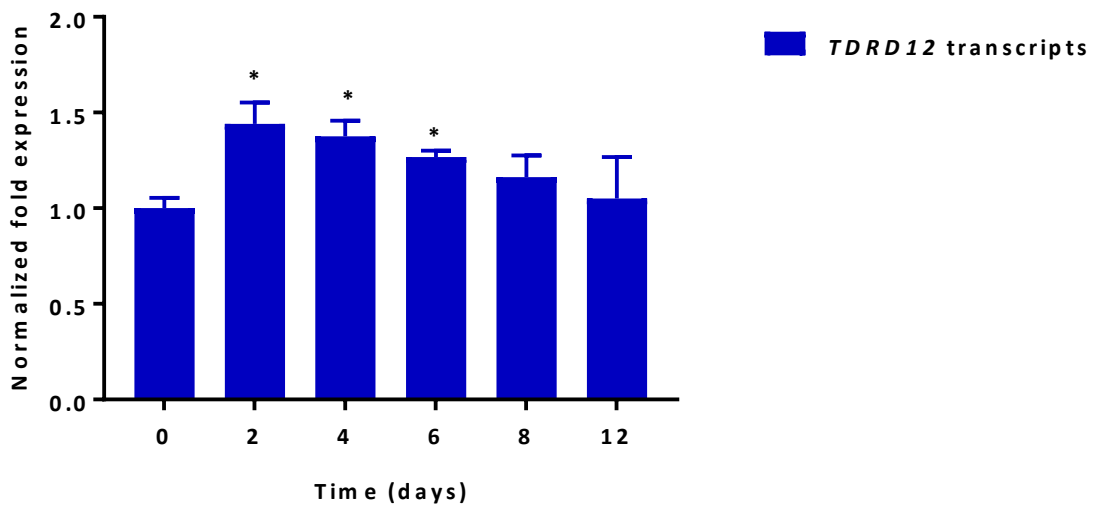
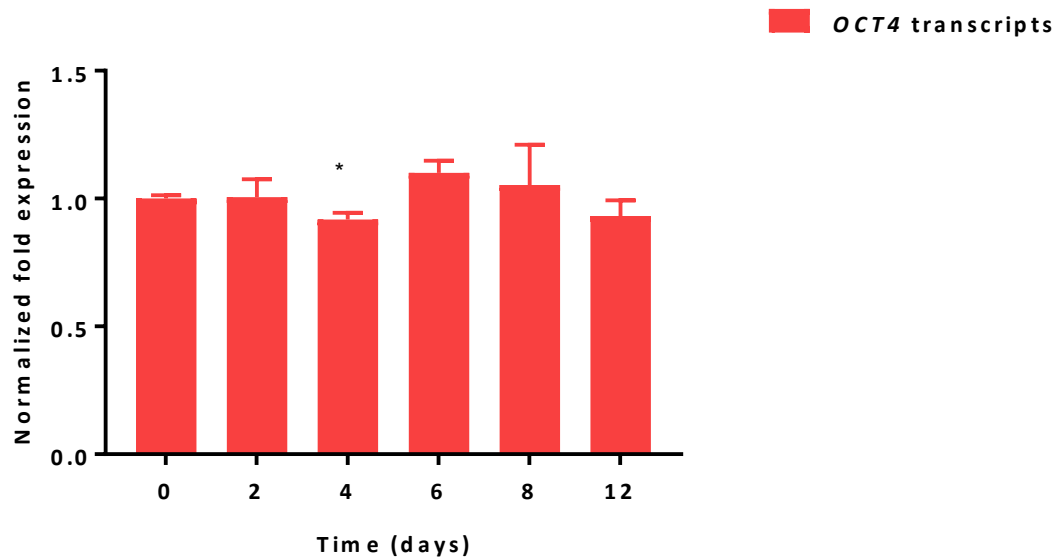


Figure 4.3. Analysis of the *OCT4* and *TDRD12* transcript levels in undifferentiated NTERA2 cells using qRT-PCR.

(A) Bar charts represent the expression levels of *OCT4* transcripts in untreated cultures of NTERA2 cells from different days. (B) Bar charts represent the levels of *TDRD12* transcripts in untreated NTERA2 cells from different days. The data were normalised to two endogenous reference genes (*ACTB* and *GAPDH*), and the relative fold change was computed by the $\Delta\Delta C_t$ method. Error bars denote the standard errors for the mean of technical triplicates and the experiment was performed in two biological repeats. (*: $P < 0.05$, **: $P < 0.01$ and ***: $P < 0.001$).

A)



B)

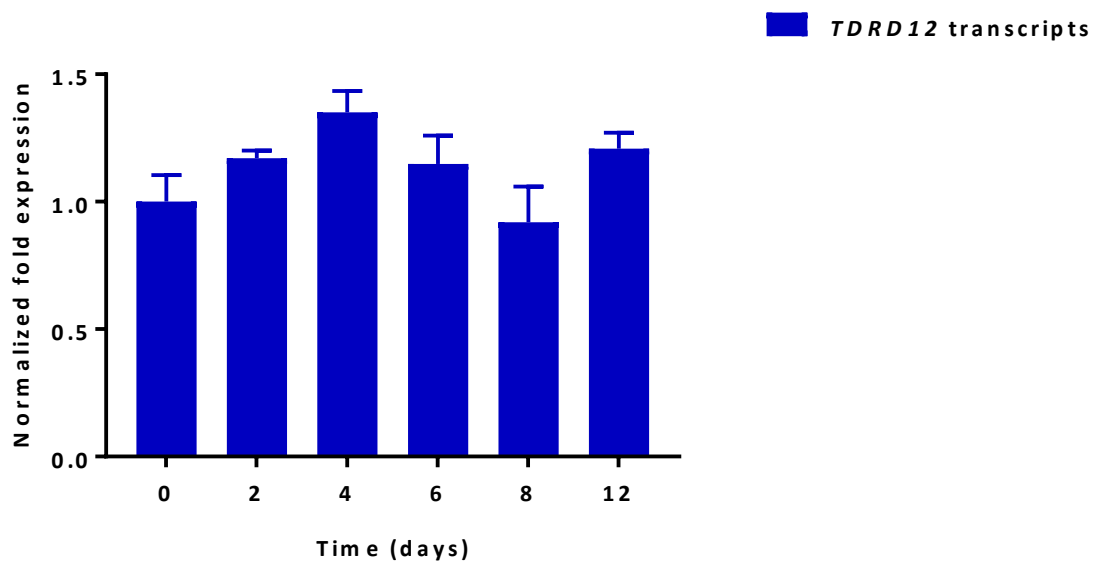
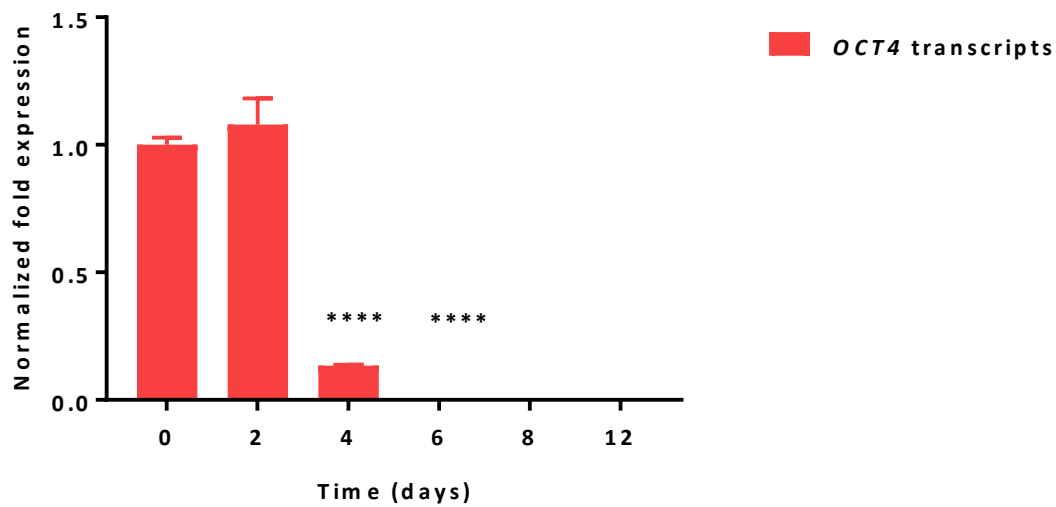


Figure 4.4. Analysis of the *OCT4* and *TDRD12* transcript levels in DMSO treated NTERA2 cells using qRT-PCR.

(A) Bar charts represent the levels of *OCT4* transcripts in DMSO-treated cultures of NTERA2 cells from different days. (B) Bar charts represent the levels of *TDRD12* transcripts in NTERA2 cells treated with DMSO for different days. The data were normalised to two endogenous reference genes (*ACTB* and *GAPDH*), and the relative fold change was computed by the $\Delta\Delta C_t$ method. Error bars denote the standard errors for the mean of technical triplicates and this experiment was performed in two biological repeats. (*: $P < 0.05$).

A)



B)

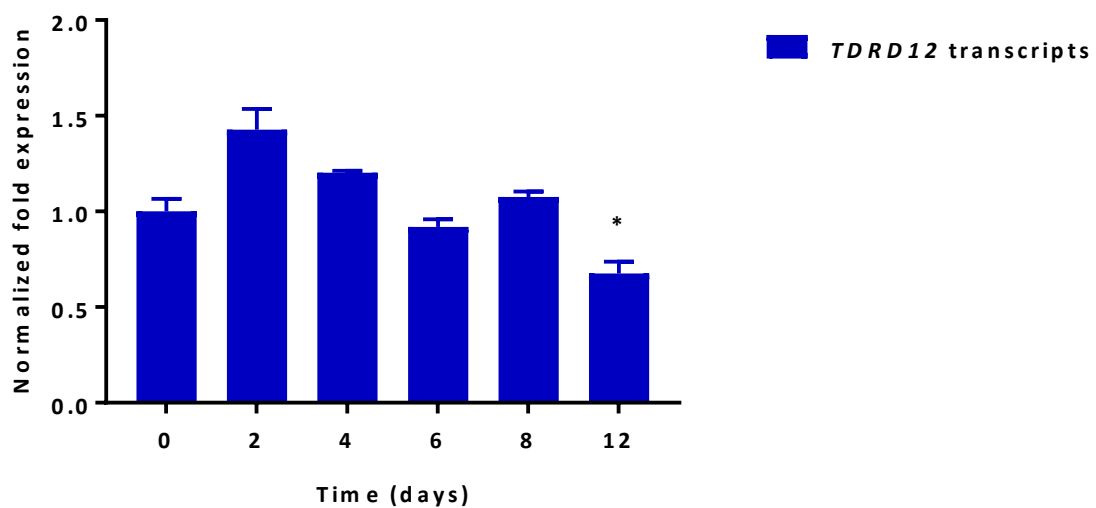
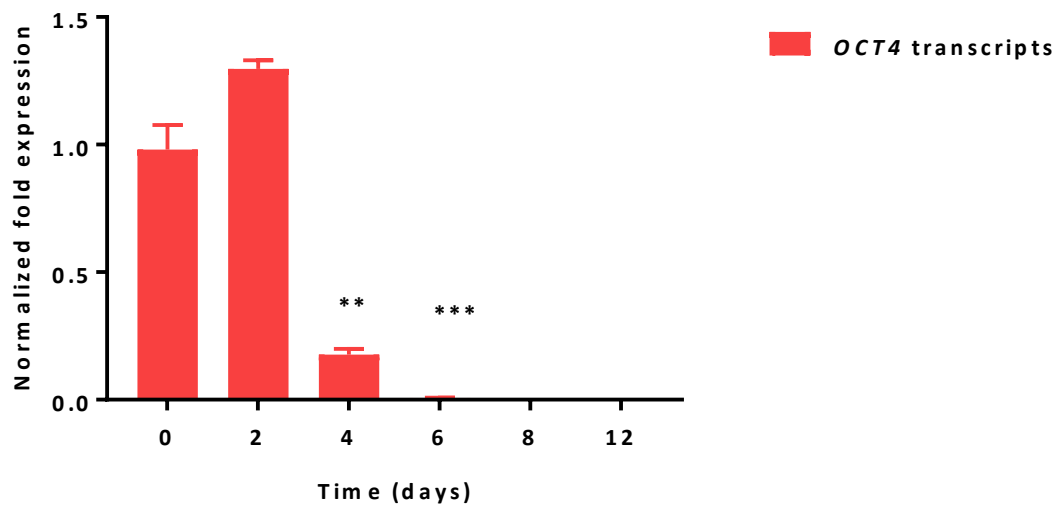


Figure 4.5. Analysis of the *OCT4* and *TDRD12* transcript levels in RA-treated NTERA2 cells using qRT-PCR.

(A) Bar charts represent the levels of *OCT4* transcripts in RA-treated cultures of NTERA2 cells from different days. (B) Bar charts represent the levels of *TDRD12* transcripts in RA-treated NTERA2 cells from different days. The data were normalised to two endogenous reference genes (*ACTB* and *GAPDH*), and the relative fold change was computed by the $\Delta\Delta C_t$ method. Error bars denote the standard errors for the mean of technical triplicates. (*: $P < 0.05$ and ****: $P < 0.0001$). This experiment was carried out three biological repeats.

A)



B)

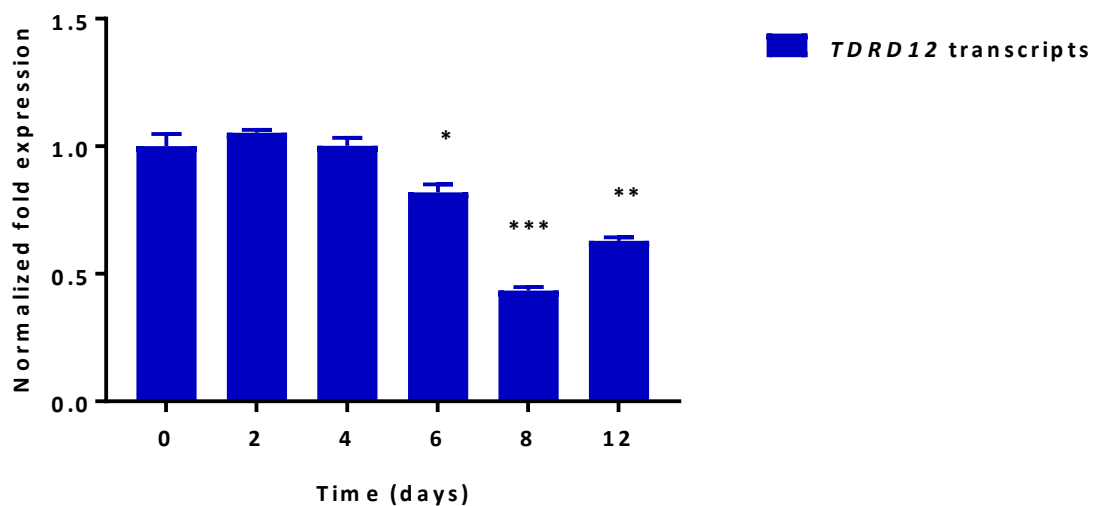


Figure 4.6. Analysis of the *OCT4* and *TDRD12* transcript levels in HMBA-treated NTERA2 cells using qRT-PCR.

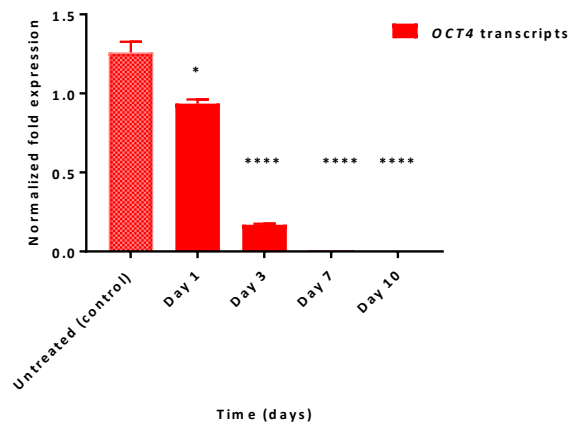
(A) Bar charts represent the levels of *OCT4* transcripts in HMBA-treated cultures of NTERA2 cells from different days. (B) Bar charts represent the levels of *TDRD12* transcripts in HMBA-treated NTERA2 cells from different days. The data were normalised to two endogenous reference genes (*ACTB* and *GAPDH*), and the relative fold change was computed by the $\Delta\Delta C_t$ method. Error bars denote the standard errors for the mean of technical triplicates. Asterisks above the bars refer to the p-values (*: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$ and ****: $p < 0.0001$). This experiment was carried out three biological repeats.

4.2.4. The expression patterns of *TDRD12* upon differentiation of human embryonic stem cells (hESCs)

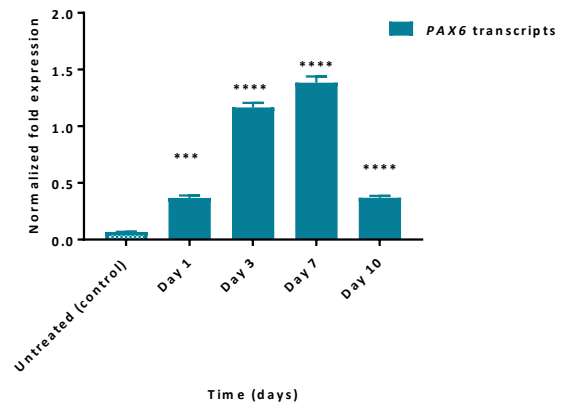
To further exploring possible roles of *TDRD12* in stemness, the levels of *TDRD12* transcripts were investigated in hESC line H9 using qRT-PCR analysis. H9 is a normal karyotype stem cell line, and it was differentiated using RA agent for 10 days. [H9 induction with RA was done in collaboration with A. Hazzazi, Bangor University (PhD thesis, 2017)]. Transcripts were analysed at distinct time points: day 1, day 3, day 7 and day 10; untreated H9 cells were used as a positive control. The level of *OCT4* transcripts as a stemness marker was used to monitor the differentiation process. *PAX6* transcripts were used as a neuronal marker, because the RA agent is known to direct the hESCs into the neuronal lineages.

The expression levels of *OCT4* transcripts showed significant reductions at days 1 and 3 and become undetectable in days 7 and 10 after RA treatment (Figure 4.7 A). This indicates that the H9 cells were successfully differentiated. Additionally, the H9 cells were shown to be differentiated into neuronal lineages as the mRNA levels of *PAX6* are elevated from day 3 of induction, as shown in Figure 4.7 B. The analysis of *TDRD12* transcript levels were significantly reduced in day 3 ($p < 0.01$), day 7 ($p < 0.01$) and day 10 ($p < 0.001$) as shown in Figure 4.7 C. These results may indicate a significant correlation of *TDRD12* expression with the stem marker *OCT4* in hESCs, and also suggested the potential function for this gene in stemness properties.

A)



B)



C)

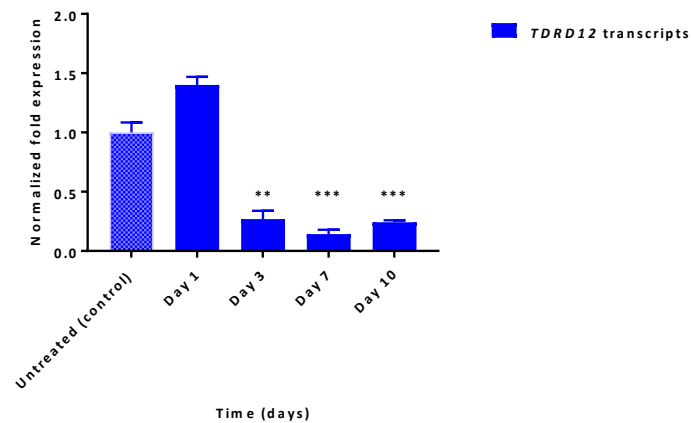


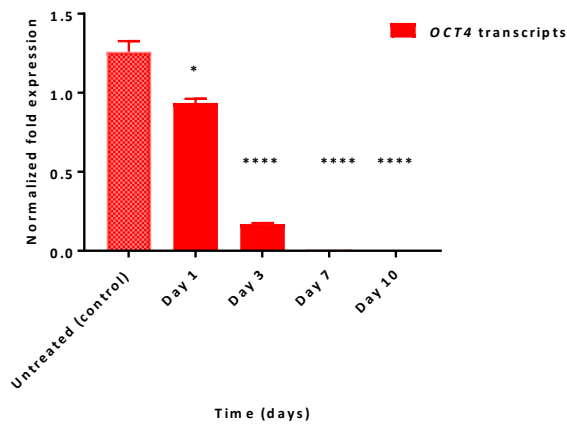
Figure 4.7. qRT-PCR analysis of *TDRD12* transcript levels in H9 cells treated with retinoic acid.

(A) Bar chart representing the transcript levels of *OCT4* in H9 cells treated with RA for different days as a stemness marker. (B) Bar chart representing the transcript levels of *PAX6* gene during the differentiation process as a neuronal marker. (C) Bar chart representing the transcript levels of *TDRD12* during the differentiation process. The transcript levels were normalised to *YWHAZ* and *ACTB* as endogenous reference genes. Error bars represent the standard errors for the mean of three technical repeats. The p-values (*: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$ and ****: $p < 0.0001$). This experiment was repeated two times from cDNA.

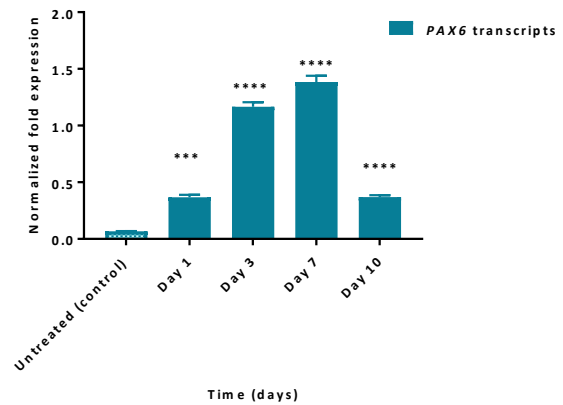
4.2.5. Assessment of *EXDL1* gene expression in human embryonic stem cells.

Previous work from mice reported that mExd1 is a partner to mTdrd12 and contributes in essential roles during germ cells development and genomic integrity (Yang et al., 2016). In addition, the results in the previous chapter suggested that human *EXDL1* expression might be correlated with *TDRD12* expression in NTERA2 cells. This work aimed to investigate the expression pattern of *EXDL1* in hESCs before and after differentiation to assess its relationship with stemness features and *TDRD12* expression. In the previous experiment (Figure 4.7), hESCs were successfully differentiated by RA and the differentiation process was validated through the analyses of *OCT4* and *PAX6* transcript levels. The analysis was also carried out to determine the expression pattern of *EXDL1* during hESCs differentiation. Results in Figure 4.8 showed that *EXDL1* transcript levels were significantly reduced in similar pattern to *TDRD12* transcripts. The significant declines of *EXDL1* mRNA levels start from day 1 ($p < 0.01$). This may suggest that *EXDL1* expression is also correlated to stemness features in hESCs as well as *TDRD12*.

A)



B)



C)

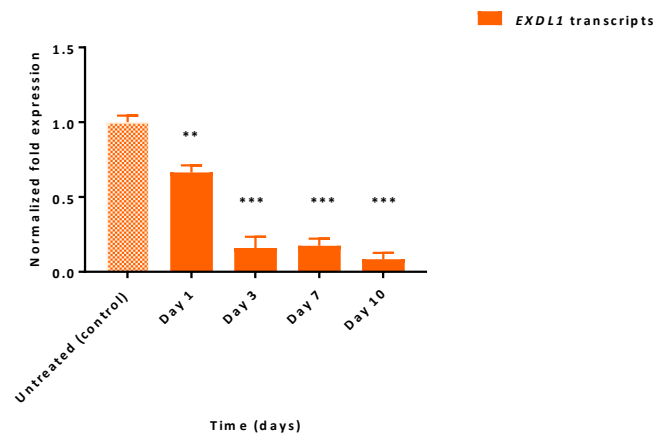


Figure 4.8. qRT-PCR analysis of *EXDL1* transcript levels in H9 cells treated with retinoic acid.

(A) Bar chart representing the transcript levels of *OCT4* in H9 cells treated with RA for different days. (B) Bar chart representing the transcript levels of *PAX6* gene during the differentiation process as a neuronal marker. (C) Bar chart representing the transcript levels of *EXDL1* during the differentiation process. The expression was normalised to *YWHAZ* and *ACTB* as endogenous reference genes. Error bars represent the standard errors for the mean of three technical repeats. The p-values (*: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$ and ****: $p < 0.0001$). Two experimental repeats were performed from cDNA.

4.2.6. Analysis of *EXDL1* expression profile in human normal tissues and cancer cells.

Previous study has described *Exd1* as a partner to *Tdrd12* in mouse (Yang et al., 2016). Furthermore, the results in Chapter 3 suggested that the relationship between *TDRD12* and *EXDL1* in human is still standing. Given that, studying the expression of human *EXDL1* is important to understand this correlation that may detect the function of *TDRD12* as well. To evaluate the expression pattern of *EXDL1* in human normal tissues, RT-PCR analysis was carried out on 20 different normal tissue types. Total RNA was purchased from Clontech and cDNA was synthesised. The quality of synthesised cDNA was validated using *ACTB* as a positive control for cDNA quality. Expression of *EXDL1* was evaluated. Figure 4.9 shows a restricted expression only in male testis for normal tissues. PCR products of *ACTB* and *EXDL1* migrated to expected sizes of 553 bp and 490 bp respectively. Furthermore, sequence analysis of the purified band confirmed that these products belong to human *EXDL1*, see (Appendix A-3).

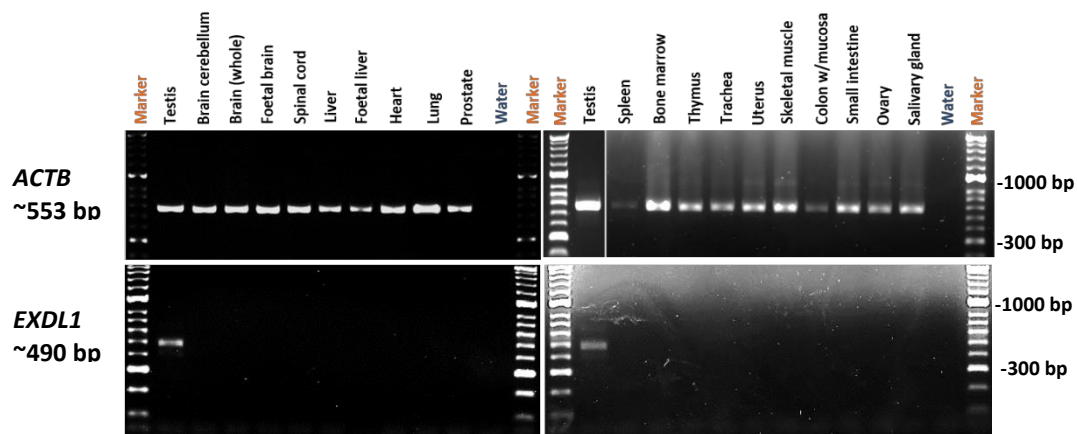


Figure 4.9. RT-PCR analysis of the expression profile of the *EXDL1* gene in normal human tissues. Agarose gels demonstrate the expression of *EXDL1* in normal human tissues. *ACTB* was used as a positive control and quality assessment for synthesised cDNA from normal tissues RNA samples. The expected size of *ACTB* is 553 bp and the expected size for *EXDL1* products was 490 bp. Water samples represent a negative controls for primers.

As the expression of *EXDL1* showed it to be a testis-specific gene, this study was extended for analysis of its expression on multiple cancer tissues and cell lines. The cDNA was created from total RNA and the quality was assessed by *ACTB* gene as a positive control. Non-template control (water) was also run to check that primers are not contaminated. The expression pattern of *EXDL1* on several human cancer tissues and cell lines was carried out using RT-PCR analysis as shown in Figure 4.10. *EXDL1* expression was restricted to testes and no cancer tissues showed expression as measured using this primer pair. These results appear to suggest that *EXDL1* is testis specific and rarely, or not at all expressed in cancer cells.

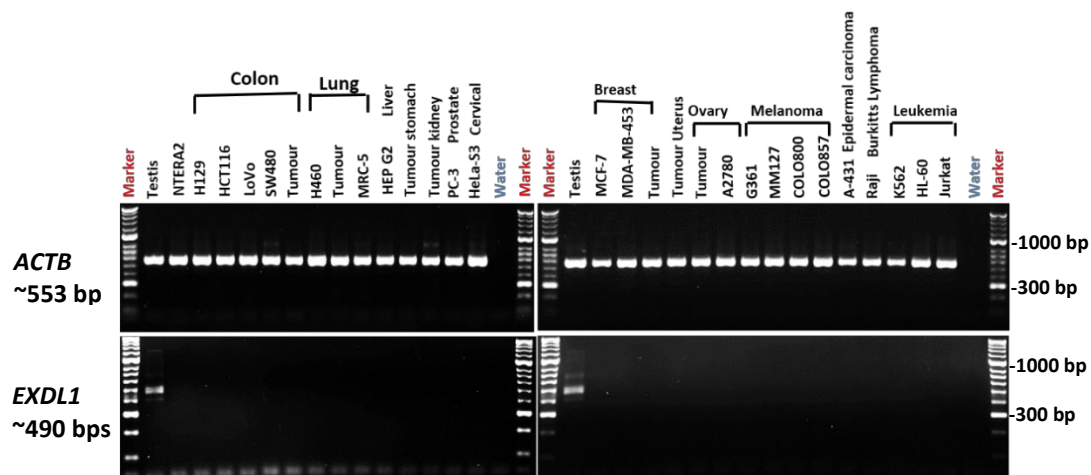


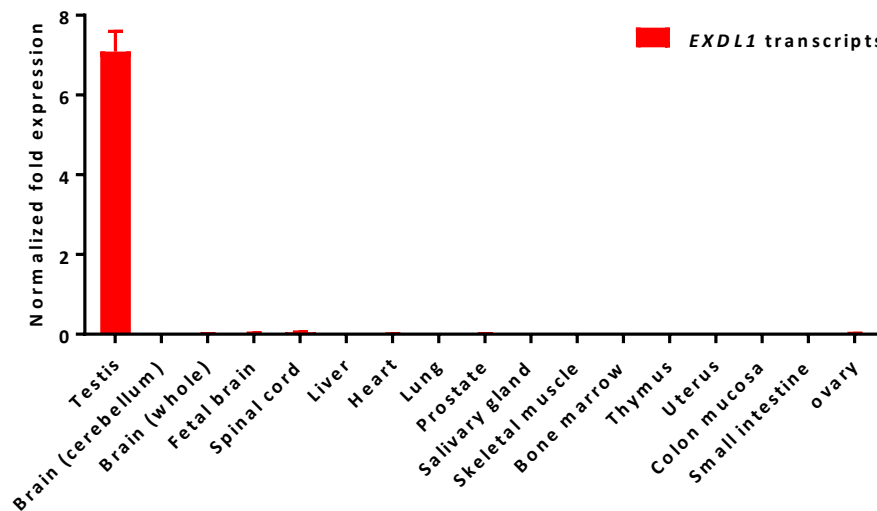
Figure 4.10. RT-PCR analysis for the expression profile of *EXDL1* in cancer human tissues and cell lines. Agarose gels demonstrate the expression of *EXDL1* in cancer samples. *ACTB* was used as a positive control and quality assessment for synthesised cDNA from cancer tissues and cells samples. The expected size of *ACTB* is 553 bp and the expected size for *EXDL1* products was 490 bp. Water samples represent a negative controls for primers.

4.2.7. Analysis using qRT-PCR of *EXDL1* expression in normal and cancerous human tissues and cell lines

Based on the RT-PCR results on normal and cancer tissues, further investigations was carried out using qRT-PCR to confirm the *EXDL1* expression pattern. As the limitations of conventional PCR to detect the weak signals, qRT-PCR was important to either confirm the previous step or quantify the low levels of mRNA transcripts. Using the same cDNA from the last experiment, quantification of levels of *EXDL1* transcripts in normal and cancer human tissues and cells was carried out. Values above Ct = 37 indicate a negative expression as this was given in non-template control samples. *EXDL1* expression profile in normal tissues is shown in Figure 4.11. The results showed a significant expression in the testes with detectable Ct value at 28.01. Very weak expressions were amplified in central nervous system (CNS) tissues as detected by Ct values; for instance, brain cerebellum (Ct = 36.42), brain (whole) (Ct = 36.15), foetal brain (Ct = 34.79) and spinal cord (Ct = 34.25). Additionally, Ct value at 36.18 which is extremely close to the cut-off was observed in normal ovary tissue indicating a potentially weak expression of *EXDL1* in ovary tissues.

The screen of *EXDL1* expression profile in cancerous tissues and cell lines is shown in (Figure 4.12). A significant expression with high level was determined in testis tissues with Ct value of 28.25. Three cell lines: embryonal carcinoma (NTERA2), melanoma (COLO857) and ovarian cancer cell line (A2780) showed weak expressions with detectable Ct values at 34.01, 35.14 and 33.82 respectively. Furthermore, ovarian tumour tissues showed expression with detectable Ct value of 35.72. Amplified products of *EXDL1* were also observed in leukaemia cell lines, HL-60 (Ct = 36.15) and colon cancer cell lines HCT116 (Ct = 36.05) which are very close to the cut-off value.

A)



B)

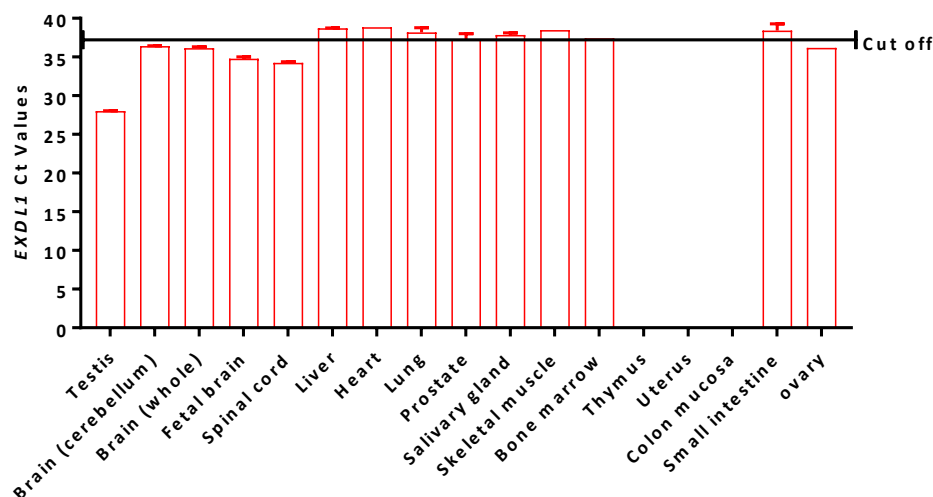
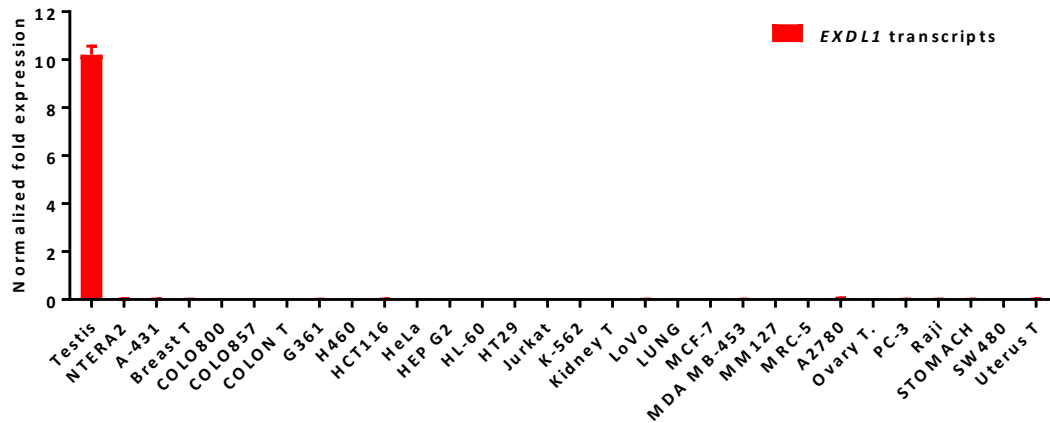


Figure 4.11. qRT-PCR analysis of *EXDL1* gene expression levels in a range of normal human tissues. From different normal human tissues RNA samples, cDNA samples were synthesised to examine the levels of *EXDL1* expression. A) Bar chart demonstrates the level of mRNA transcripts of the *EXDL1* gene by RT-qPCR in testis and multiple normal human tissues. The gene expression was normalised to two endogenous reference genes *GAPDH* and *ACTB*. The fold change was computed using the $\Delta\Delta C_t$ method, and the error bars refer to the standard errors of mean and Y axis scale is linear. B) Cycle threshold (Ct) values obtained following the amplification of *EXDL1* in normal tissues. Black line is the cut off line of Ct value < 37 cycles, indicates the acceptable Ct threshold in this study. Samples above the cut off line is considered as negative expression as no detectable value (Ct= 0). Ct values of reference genes were ranging from 16 to 18 cycles to ensure the reference gene stabilities.

A)



B)

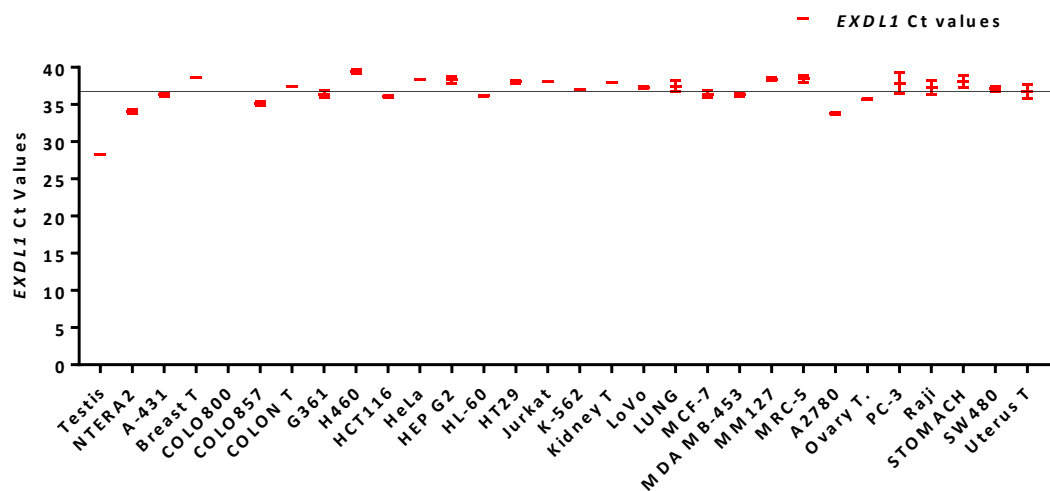


Figure 4.12. qRT-PCR analysis of *EXDL1* gene expression levels in a range of cancerous human tissues. From multiple cancerous human tissues RNA samples, cDNA samples were synthesised to examine the levels of *EXDL1* gene expression. A) Bar chart demonstrates the level of mRNA transcripts of the *EXDL1* gene by qRT-PCR in testis and multiple cancer human tissues. The gene expression was normalised to two endogenous reference genes *GAPDH* and *ACTB*. The fold change was computed using the $\Delta\Delta Ct$ method, and the error bars refer to the standard errors of mean and Y axis scale is linear. B) Cycle threshold (Ct) values obtained following the amplification of *EXDL1* in cancerous tissues. Black line is the cut off line of Ct value < 37 cycles, indicates only the acceptable Ct threshold are below this study. Samples above the cut off line represent negative expression as well as no detectable Ct values (Ct= 0). Ct values of reference genes were ranging from 16 to 18 cycles to ensure the reference gene stabilities.

4.3. Discussion

4.3.1. *TDRD12* gene is a potentially linked to stemness features of CSCs and hESCs.

The human *TDRD12* gene has been identified as a CTA gene candidate that has the potential to be a cancer biomarker and may encode CSC-specific activity (Feichtinger et al., 2012; Almatrafi et al., 2014). In the preliminary screen of *TDRD12* gene expression in multiple cancerous tissues and cells, a restricted expression of *TDRD12* transcripts was seen in embryonic cancer stem-like cell line (NTERA2), which makes *TDRD12* a high-interest target. This led to the hypothesis that *TDRD12* may have a functional role in conferring stemness features in cancer cells.

This chapter investigated the expression of *TDRD12* in hESCs, iPSCs and differentiated cancer stem cells. Firstly, the qRT-PCR analysis demonstrated that *TDRD12* along with a stemness marker, *SOX2*, were expressed in hESCs, CSCs and iPSCs but not in fibroblast somatic cells. Interestingly, *TDRD12* was shown to be activated in iPSCs in response to the transfected stem cell markers during the reprogramming process from precursor fibroblasts. Given this, it was proposed that *TDRD12* may be associated with stemness characteristics and it could be a possible stem marker. Furthermore, the depletion of *TDRD12* in NTERA2 cells reduced the levels of *SOX2* and *NANOG* stemness markers but did not change *OCT4*, suggesting that *TDRD12* could potentially be involved in a transcriptional regulation for some stemness features. These results encouraged further analysis, and NTERA2 cells were differentiated to investigate the expression of *TDRD12*. In HMBA-treated cultures, significant reductions in the *TDRD12* mRNA levels started to be observed upon the inhibition of *OCT4* expression. This may indicate the correlation of *TDRD12* with the stemness marker expression. Moreover, *TDRD12* mRNA levels showed a decline at day 12 post-treatment in RA-treated NTERA2 culture. This observation suggested that the *TDRD12* expression does not behave similarly to the *OCT4* expression in RA-treated NTERA2. However this might be explained by the nature of this cell model that may not represent a normal *TDRD12* profile.

This data justified the extended analysis of *TDRD12* expression in hESCs upon differentiation using RA induction. In this study, RA-treated hESCs were found to inhibit the expression of *OCT4* and activate neuronal marker *PAX6*, indicating a successful differentiation process. The *TDRD12* mRNA level decline started from day three, which paralleled the *OCT4* mRNA levels that showed the same patterns. However, the *TDRD12* gene was not completely silenced upon hESC differentiation. It is possible that these cells did not lose their stemness properties,

and if the time scale of the experiment was increased it may then give different results. Thus, the *TDRD12* gene was suggested to be associated with stemness characteristics in human ES and CSCs, and it could be implicated as a stem cell marker.

4.3.2. *EXDL1* is a potential cancer biomarker.

Previous studies reported an Exd1 protein that was involved in mouse germ cell stability and was described as a Tdrd12 partner (Yang et al., 2016). No studies have characterised the *EXDL1* (human orthologue) protein, though the results in Chapter 3 suggested a significant correlation between *EXDL1* and *TDRD12* genes at a transcriptional level. This led us to investigate *EXDL1* expression in hESCs during differentiation. The transcript levels of *EXDL1* behaved in the same way as *TDRD12* transcript levels in this study, suggesting that *EXDL1* is also correlated to stemness features along with the *TDRD12* gene. Furthermore, this supports the suggested interrelationship between *EXDL1* and *TDRD12* at the transcriptional level in human germ cells in previous chapter.

Identification of new CTAs as potential cancer biomarkers is also useful for developing strategies for cancer diagnosis and immunotherapy. The *TDRD12* gene has been previously characterised as a potential CT gene (Feichtinger et al., 2012), but no previous studies have investigated the expression profile of *EXDL1* in normal and cancer human tissues. This study aimed to identify a novel CT gene and to validate and investigate its relationship to *TDRD12* expression, depending on the tissue type. An RT-PCR analysis showed that *EXDL1* expression was restricted to male testis tissues from normal and cancerous human tissues. These results suggested that *EXDL1* is a testis-specific gene, although its expression is also apparent in hESCs. In contrast to *TDRD12* that was detected in NTERA2 in the previous study, *EXDL1* was not detected in NTERA2 cells using conventional RT-PCR. The expression of *EXDL1* in NTERA2 might be too weak to visualise on agarose gels. For this reason, a further step of analysis was carried out using qRT-PCR to quantify the expression of the *EXDL1* gene. The *EXDL1* mRNA levels were high and were significantly restricted to the testis with weak expression levels in CNS tissues and healthy ovarian tissues. This result suggested that *EXDL1* might be positively expressed in normal CNS tissues in addition to the testis. From cancerous samples, several cancer cells showed positive expression signals for *EXDL1*. They were at low levels in comparison to the expression in testes. Cancer cells that might express *EXDL1* as demonstrated by their Ct values include: NTERA2, COLO857, G361, HCT116, HL-60, A2780 and

ovarian tumour tissue. Based on the criteria of CTA gene characterisation, Hofmann and co-workers suggested that genes with low levels of expression in no more than two healthy tissues compared to the expression detected in the testis are considered to be CTA candidates (Hofmann et al., 2008). Some CTAs may be expressed in other immune privilege organs, such as the brain (CNS), which shows expression levels of less than 1% of those in the testis (Caballero & Chen, 2009; Fratta et al., 2011). Given that, from the expression profile of the *EXDL1* gene, it is a possible CTA candidate. Moreover, these findings may suggest that both *TDRD12* (identified in previous study) and *EXDL1* genes share common features, as they are both possible CTA candidates and are correlated to the stemness features in hESCs.

4.3.3. Conclusion

To summarize, biomarkers that are specific to CSCs represent good targets for generating new drugs and for the early diagnosis of cancer. The work in this chapter revealed a link between *TDRD12* and stemness features, suggesting that this gene is a possible stem cell marker. Furthermore, this study highlighted the potential function of the *TDRD12* gene in CSC and stem cell specificity. Furthermore, the *EXDL1* gene behaves like *TDRD12* in hESCs, suggesting that it might also have a similar potential function in stem cells. Additionally, novel CTA identification is important in cancer diagnosis and treatment. This study validated that the *EXDL1* gene is a possible CTA candidate gene.

Chapter 5:

Functional analysis of human *LKAAEAR1* gene in normal and cancer tissues.

5. Functional analysis of human *LKAAEAR1* gene in normal and cancer tissues.

5.1. Introduction

One of the main features that define CT genes is the exclusive expression in germ cells of the male testis. The testes are an immune-privileged site containing about 250 testicular lobules compartments that are separated by septa (tissue barriers) and each lobule contains from one to three seminiferous tubules. Each seminiferous tubule is formed by two kinds of cells: germ cells that undergo the spermatogenesis process and Sertoli cells that provide nutrition and physical support to the germ cells. The roles of CT genes in spermatogenesis remain comparatively unclear, however, the studies on gene expression analysis and gene knockout revealed many functions of CT genes, for examples, see (Jungbluth et al., 2000; Caballero & Chen, 2009; Whitehurst, 2014). Some CT genes may show expression in other immune-privileged sites, such as the brain, where the blood–brain barrier (BBB) exists, and they are termed testis/CNS-restricted CT genes. Another class of CT genes is found to be expressed in no more than two healthy somatic tissues, although their expression levels always show less than 1% of the expression detected in the testis (Caballero & Chen, 2009; Fratta et al., 2011). The expression pattern in a wide range of cancerous tissues is another key feature of CT genes. CTAs do not harness the immune response in the testes but CTA genes encode immunogenic proteins that have ability to induce a specific response in cancerous tissues (Caballero & Chen, 2009). This is the most attractive characteristic that make CTA proteins potential mediators of tumour cell recognition and targeting by the immune system (Simpson et al., 2005; Caballero & Chen, 2009). Thus, CTAs genes are potentially of a high importance in clinical applications, including cancer vaccination and immunotherapy (Caballero & Chen, 2009; Fratta et al., 2011).

Human *C20orf201* gene is also officially named *LKAAEAR* motif containing 1 (*LKAAEAR1*), and is located on chromosome 20 at q13.33 (GenBank accession number BC036837). Only two splice variants were predicted with the whole open reading frame of 723 bp in variant 1 and 585 bp in variant 2. These variants encode proteins with unknown functions of 240 aa and 194 aa for variants 1 and 2 respectively (Figure 5.1).

In 2012, Feichtinger and colleagues identified *LKAAEAR1* as a CT gene through the validation of its expression in normal and cancer tissues followed by meta-analyses of clinical data sets

from many types of tumours (Feichtinger et al., 2012). Later, Kamata et al. also reported the expression of mRNA transcripts of *LKAAEAR1* using quantitative qRT-PCR analyses of a range of tumour types (Kamata et al., 2013). Both previous studies reported that human *LKAAEAR1* transcripts were detected in testis and CNS (cerebellum and foetal brain) in normal tissues and in many cancers including ovarian, breast and prostate. Thus, from the expression profile of *LKAAEAR1*, this gene is classified as testis/CNS restricted CT gene (Feichtinger et al., 2012). Additionally, Kamata and co-workers isolated many antigenic peptides from HLA molecules presented on the prostate cancer cells which could be encoded by *LKAAEAR1*. These antigenic peptides are recognised by cytotoxic T lymphocytes (CTL), thus, they have promising potential as vaccines for cancer immunotherapy (Kamata et al., 2013). Among isolated peptides from HLA on prostate cancer cells were five- or six- mer peptides of TKLSA and RLRYT. Interestingly in this study, the uncharacterised protein of *LKAAEAR1* was reported to be the possible source of RLRYT peptides (Kamata et al., 2013).

Given these findings, the *LKAAEAR1* gene and *LKAAEAR1* protein have the potential to be used as cancer biomarkers. The aim of the work in this work is to validate the presence of the *LKAAEAR1* protein in normal and cancer tissues and cell lines. Also, this chapter seeks to explore the functional roles of the *LKAAEAR1* gene in stemness and oncogenesis. A further aim is to determine the localisation of *LKAAEAR1* in cancer cells and testis tissues.

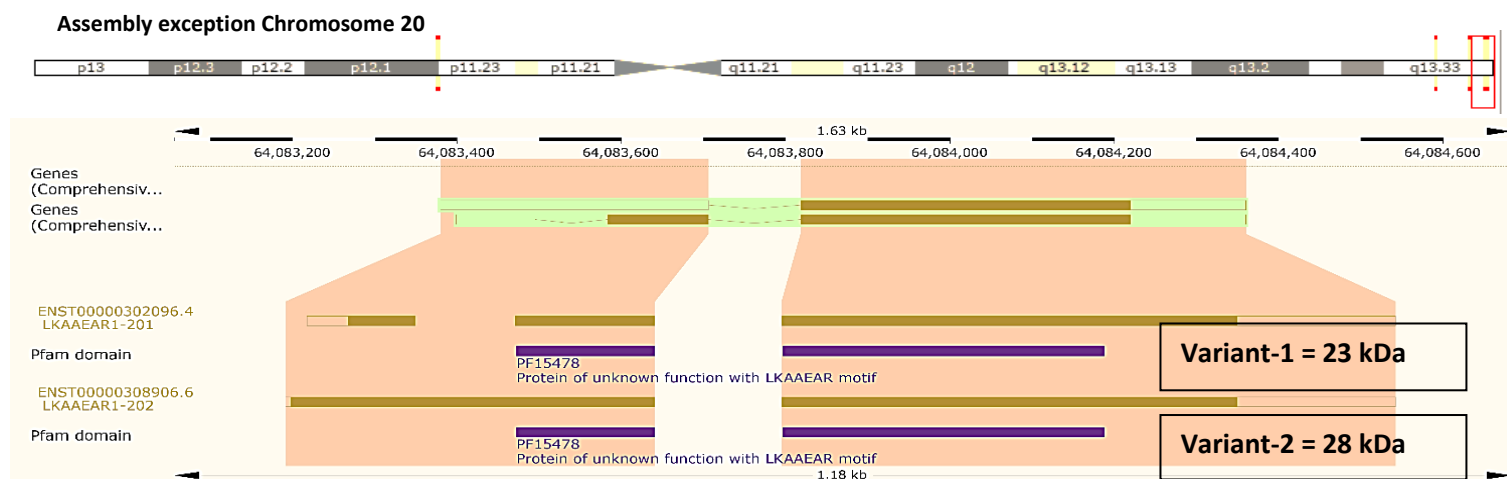


Figure 5.1 *LKAAEAR1* is localised on chromosome 20q13.33 with two predicted splice variants. Splice variant 1 (*LKAAEAR1*-201) which is composed of three exons that are separated by two introns. The splice variant 2 (*LKAAEAR1*-202) which is composed of two exons separated by one intron. From (ensemble.org; Gene: *LKAAEAR1* ENSG00000171695). The expected size of protein is calculated (TV1 at 23 kDa and TV2 at 28 kDa).

5.2. Results

5.2.1. The characterisation of the LKAAEAR1 protein in normal and cancer tissues

Previous studies have characterised the *LKAAEAR1* gene as a testis/CNS-restricted CT gene through the analysis of mRNA by RT-PCR (Feichtinger et al., 2012). This characterisation has subsequently been confirmed through the quantitative qRT-PCR analysis of mRNA transcript levels in many normal and cancerous tissues (Kamata et al., 2013). Here, further characterisation was performed to validate the presence of LKAAEAR1 protein in many normal and cancer tissues and cell line lysates. Western blot was carried out and an anti-LKAAEAR1 polyclonal antibody was purchased from Abcam (ab108142). In normal tissues, the presence of the LKAAEAR1 protein was detected only in testis and brain lysates at the expected size of 28 kDa. However, many cancer tissues and cell lysates have detectable levels of this protein, including those of the lung, liver, breast and ovary, as well as in leukaemia and melanoma. Some cancer cells, such as, NTERA2, HEP, A2780 and K562 cell lines, showed another strong band at approximately 50 kDa. An anti-GAPDH antibody was used as a loading control (Figure 5.2).

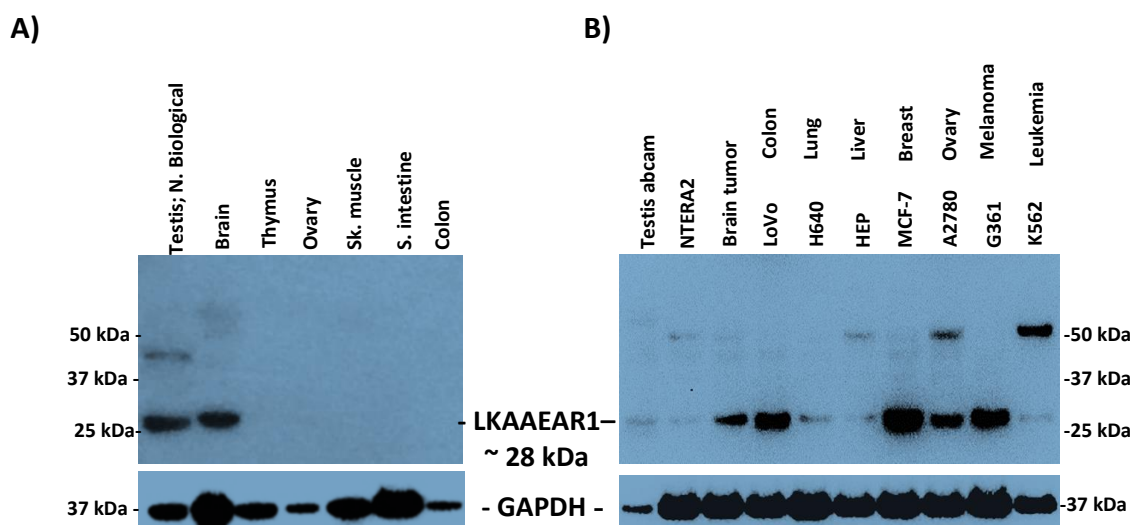


Figure 5.2. Analysis of the LKAAEAR1 protein in human tissue extracts using Western blot. A) Normal human tissue lysates show a strong signal in the testis and brain tissues. B) The LKAAEAR1 protein was detected with variable signal intensities in many cancer tissue extracts. Anti-GAPDH was used as a loading control.

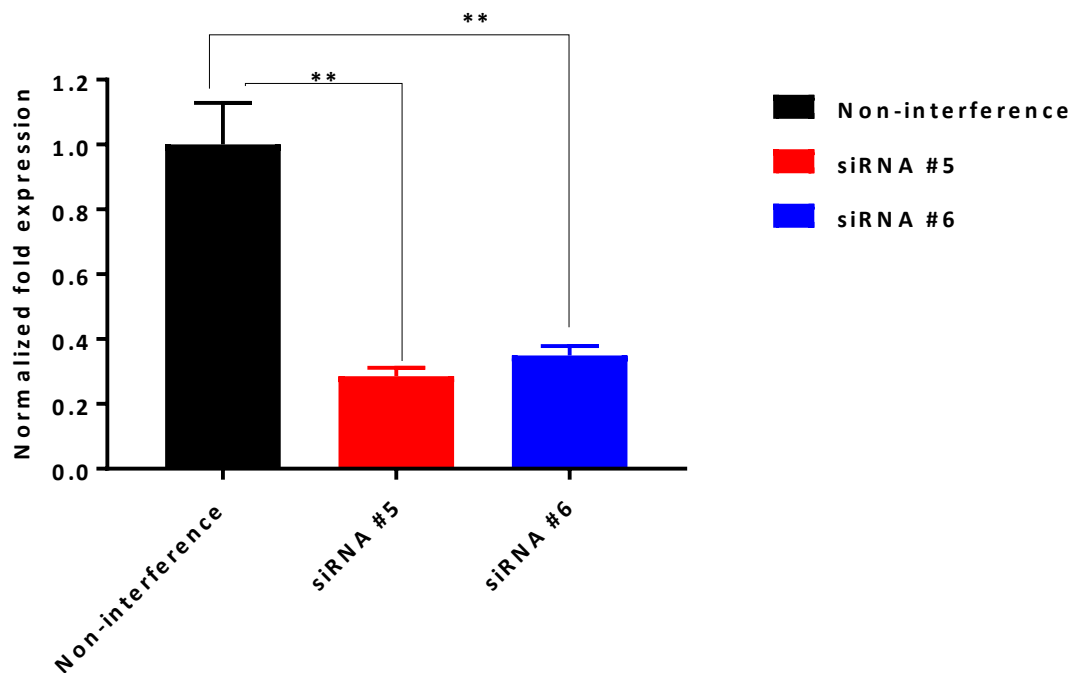
5.2.2. Depletion of *LKAAEAR1* in cancer stem-like cells and breast cancer cells

The aim of this experiment is to target *LKAAEAR1* transcripts in order to validate the antibody and monitor the influence of transcript reductions on cancer cell functions. Numerous studies have relied on silencing techniques, such as siRNA transfection to identify the function of many CTAs. For instance, knockdown of *SSX2* in melanoma cancer cells was reported to effectively slow cell proliferation (Greve et al., 2015). In this study, two cancer cell types were chosen: cancer stem-like cells NTERA2 because of their cancerous and stemness features, and breast cancer cells, MCF7, as these cells showed a strong signal for *LKAAEAR1* protein at expected size of 28 kDa. In previous experiment, the signal of *LKAAEAR1* was weak in NTERA2 cells and this may be because of using a commercial lysate (inaccurate concentration might be recommended), however, lysates of NTERA2 from McFarlane lab were examined and showed strong signals of *LKAAEAR1* protein. Given this, NTERA2 was chosen because of its stemness feature and to investigate whether *LKAAEAR1* is linked to stemness or not. The depletion of *LKAAEAR1* transcripts was carried out using two sequences of siRNAs that both target mRNA transcripts of *LKAAEAR1* in the coding region. Non-interference RNAi was used as a negative control. RNA and protein extracts were obtained after 72 hours of treatment to evaluate the depletion efficiency.

The knockdown of *LKAAEAR1* transcripts in NTERA2 cells was evaluated by qRT-PCR analysis and western blot analysis (Figure 5.3). The results showed that *LKAAEAR1* transcript levels were significantly reduced in positively siRNA-treated cultures ($p < 0.01$) in comparison to the negative control culture. The knockdown of *LKAAEAR1* protein was verified by western blot analysis with a clear reduction in siRNA #6 treated cells at the expected size of 28 kDa. A slight reduction was also observed in 28 kDa protein following the depletion by siRNA #5.

The efficiency of knockdown in MCF7 cells was examined through qRT-PCR analysis of mRNA transcripts and Western blot analysis (Figure 5.4). Significant downregulations of the transcript levels were shown in cells treated by siRNAs (#5 and #6), ($p < 0.001$), compared to the negative control (Figure 5.4 A). Additionally, no changes were observed in the protein at the expected size of 28 kDa. However, two bands of higher molecular weights were observed in a negative control at approximately 50 and 48 kDa. Both bands showed slight reductions in siRNA #6-treated cells (Figure 5.4 B). This suggested that the higher molecular weight bands may belong to *LKAAEAR1* species and the increase of molecular weight may be resulted from a posttranslational modification such as ubiquitination.

A)



B)

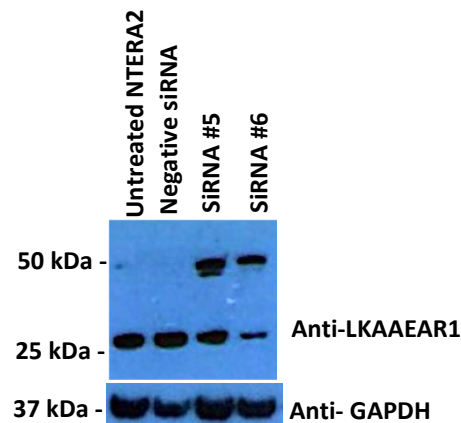
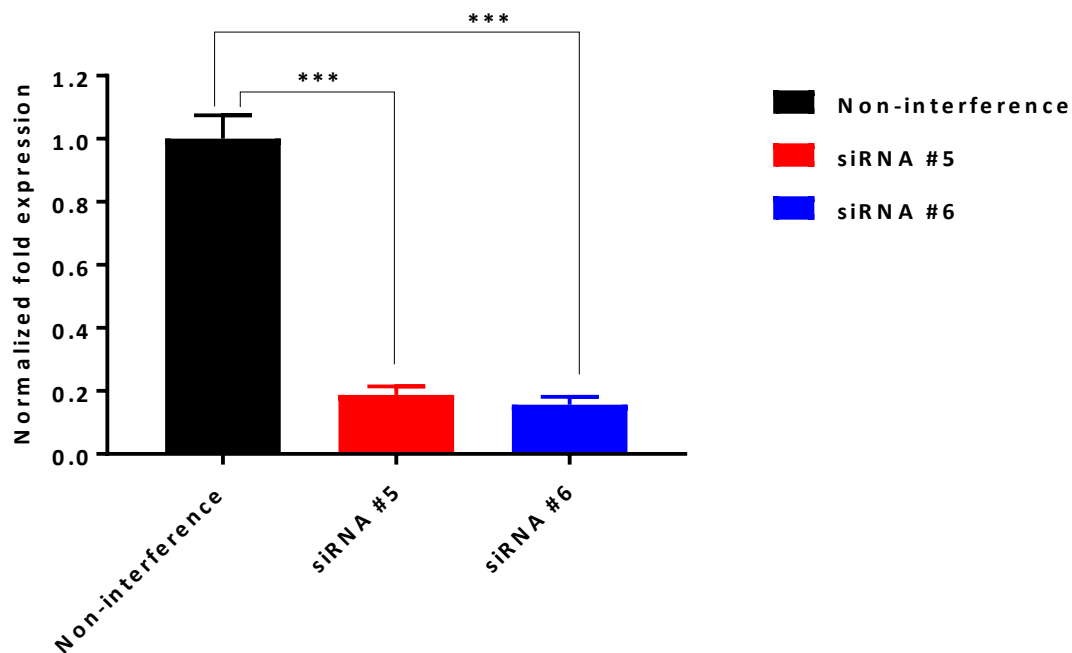


Figure 5.3. NTERA2 cells depleted of *LKAAEAR1* using different siRNAs.

A) Bar charts demonstrate the level of mRNA transcripts of the *LKAAEAR1* gene by qRT-PCR in NTERA2 cells following depletion. The cells were transfected with the negative control and two *LKAAEAR1*- specific siRNAs, and the gene expression was normalised to *GAPDH* and *ACTB*. The fold change was computed using the $\Delta\Delta C_t$ method, and the error bars refer to the standard errors for the mean of three technical repeats. P values were calculated in comparison to the control (non-interference) treatment. (** P value < 0.01). Three biological experiments were performed from knockdown.

B) Western blot analysis to show the *LKAAEAR1* protein by (ab108142) antibody after three days of treatment using different siRNAs. Whole cell protein extracts were performed from three biological replicates, one representative blot is shown. *GAPDH* was used as a loading control.

A)



B)

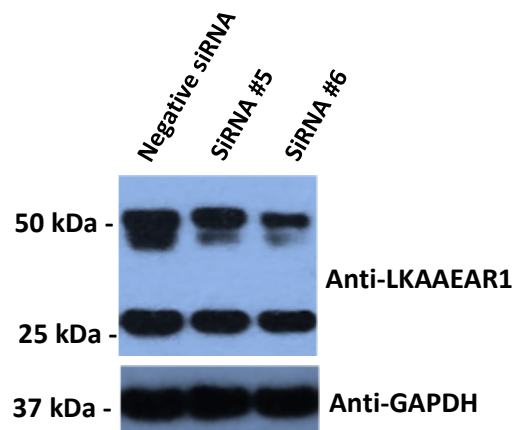


Figure 5.4. MCF7 cells depleted of *LKAAEAR1* using different siRNA targets.

A) Bar charts show the levels of mRNA transcripts of the *LKAAEAR1* gene by qRT-PCR in MCF7 cells following depletion. The cells were transfected with a negative control and two *LKAAEAR1*-specific siRNAs, and the gene expression was normalised to *GAPDH* and *ACTB*. The fold change was computed using the $\Delta\Delta C_t$ method, and the error bars refer to the standard errors of the mean. P values were calculated in comparison to the control (negative siRNA) treatment. Three knockdown repeats reveal consistently reduction in *LKAAEAR1* transcripts (*** P value < 0.001). Untreated cells were used as a positive control for the treatment.

B) Western blot analysis to show the *LKAAEAR1* protein levels after three days of treatment using different siRNAs. . Whole cell protein extracts were performed from three biological replicates, one representative blot is shown. GAPDH was used as a loading control.

5.2.3. *LKAAEAR1* expression during NTERA2 cell differentiation.

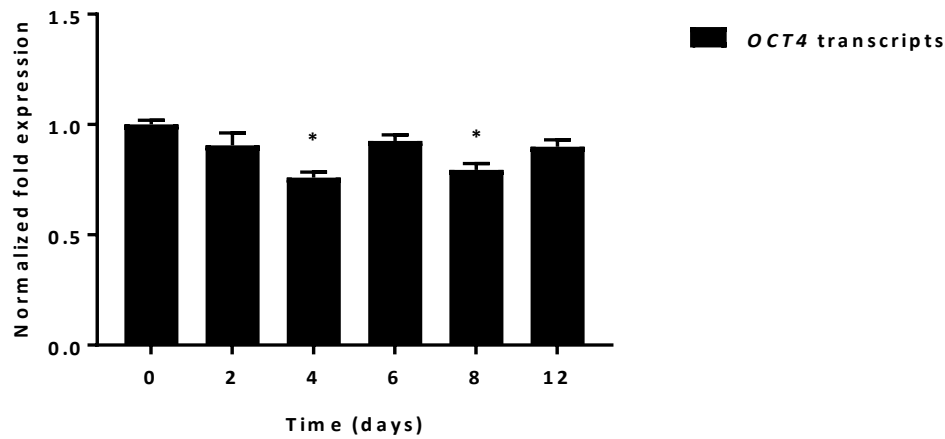
LKAAEAR1 expression in germ cells of testis and embryonal cancer stem-like cells may suggest the potential function of this gene in stemness features. This work aimed to assess the *LKAAEAR1* transcript levels during the differentiation process in NTERA2 cells. In this experiment, two agents were used to differentiate NTERA2 cells, RA and HMBA. Furthermore, DMSO treatment was used as a negative control for the induction treatment and the untreated NTERA2 cultures as a negative control for the experiment. NTERA2 cells were differentiated for 12 days post-treatment and transcript levels were analysed at days 0, 2, 4, 6, 8, and 12. Transcripts of the stemness marker gene *OCT4* were used to monitor the differentiation of NTERA2 cells during the treatment.

The transcript levels of *OCT4* remained generally constant and highly expressed in untreated and DMSO-treated cultures as positive controls (Figure 5.5 A and 5.6 A). No treatment and DMSO-treated cultures were conducted along with induction agent (RA and HMBA) cultures to ensure that the medium or DMSO did not cause cell differentiation. Thus, cells of untreated and DMSO cultures did not show differentiation depending on *OCT4* levels (Figure 5.5 A and 5.6 A). *LKAAEAR1* transcript levels were variable and higher than the levels of day 0 (Figure 5.5 A and 5.6 A). It can be seen that *LKAAEAR1* levels were very high in untreated cells at day 6 (Figure 5.5 B), however, this may depend on the variability during this experiment and this increase was not observed in another repeats.

The differentiation of NTERA2 cells was successfully achieved. The analyses of *OCT4* levels in RA-treated NTERA2 cells showed significant reductions starting from day 4 post-treatment ($p < 0.0001$) and were undetectable from day 6 after treatment (Figure 5.7 A). This is an indication of the successful differentiation of NTERA2 cells. However, *LKAAEAR1* mRNA exhibited variable and high levels with no significant reduction in RA-treated culture as shown in Figure 5.7 B.

NTERA2 cells were successfully differentiated due to HMBA inducing agent as determined through the analysis of *OCT4* mRNA levels in this culture (Figure 5.8 A). The significant reductions of *OCT4* levels start from day 4 ($p < 0.01$) while levels were undetectable by day 8. Furthermore, *LKAAEAR1* mRNA levels were substantially reduced upon HMBA induction at all examined cultures ($p < 0.001$ and $p < 0.01$) (Figure 5.8 B). These results suggest that *LKAAEAR1* expression could be correlated with the stem cell marker gene *OCT4*, and might be also required for stemness characteristics.

A)



B)

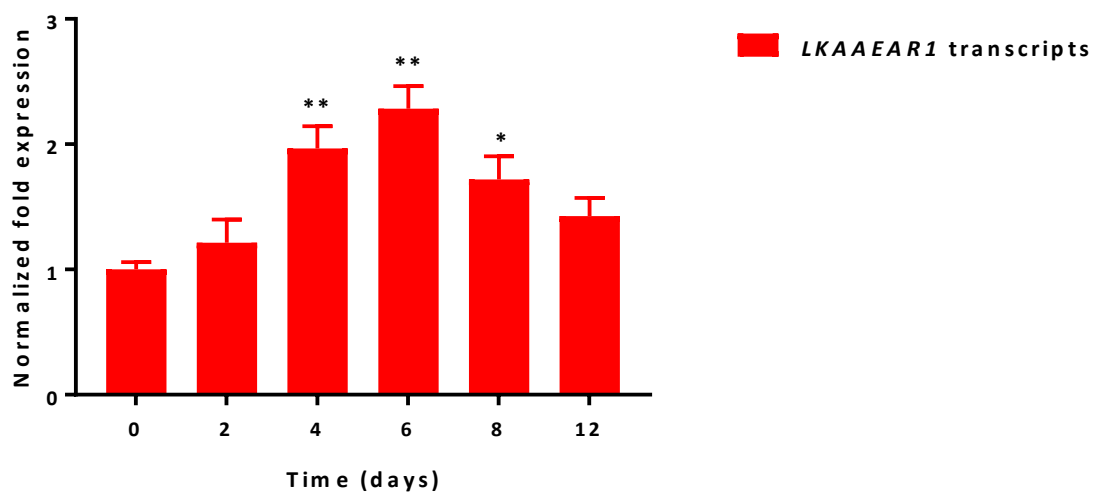
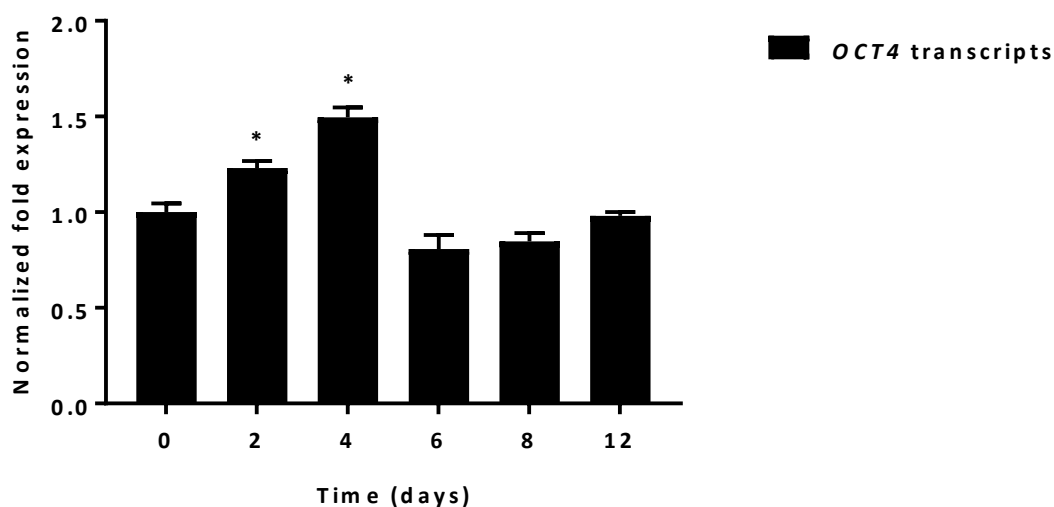


Figure 5.5. Analysis of the *OCT4* and *LKAAEAR1* transcript levels in untreated NTERA2 cells using qRT-PCR.

(A) Bar charts represent the expression levels of *OCT4* transcripts in untreated cultures of NTERA2 cells from different days. (B) Bar charts represent the levels of *LKAAEAR1* transcripts in untreated NTERA2 cells from different days. The data were normalised to two endogenous reference genes (*ACTB* and *GAPDH*), and the relative fold change was computed by the $\Delta\Delta C_t$ method. Error bars denote the standard errors for the mean of technical triplicates.

A)



B)

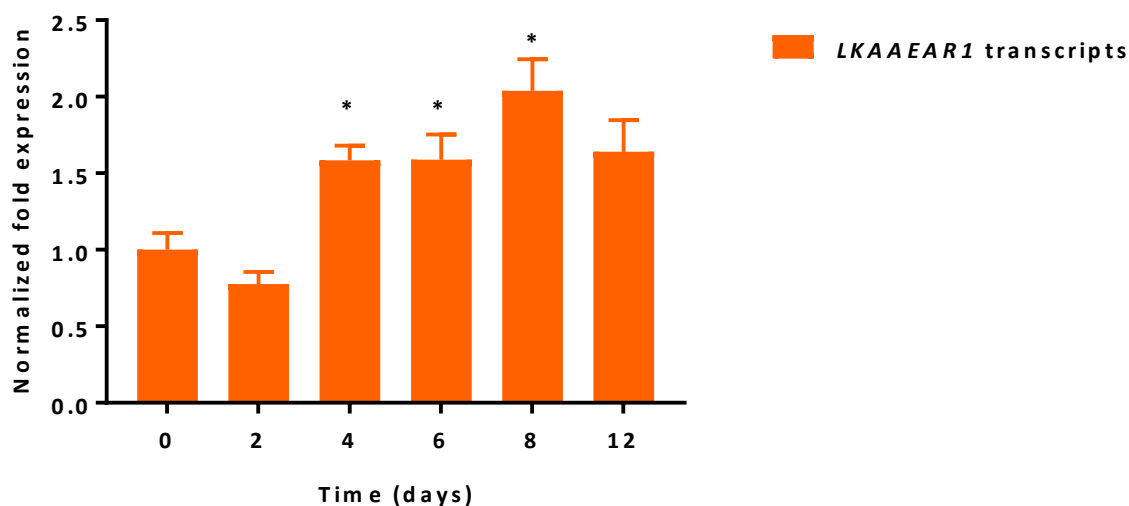
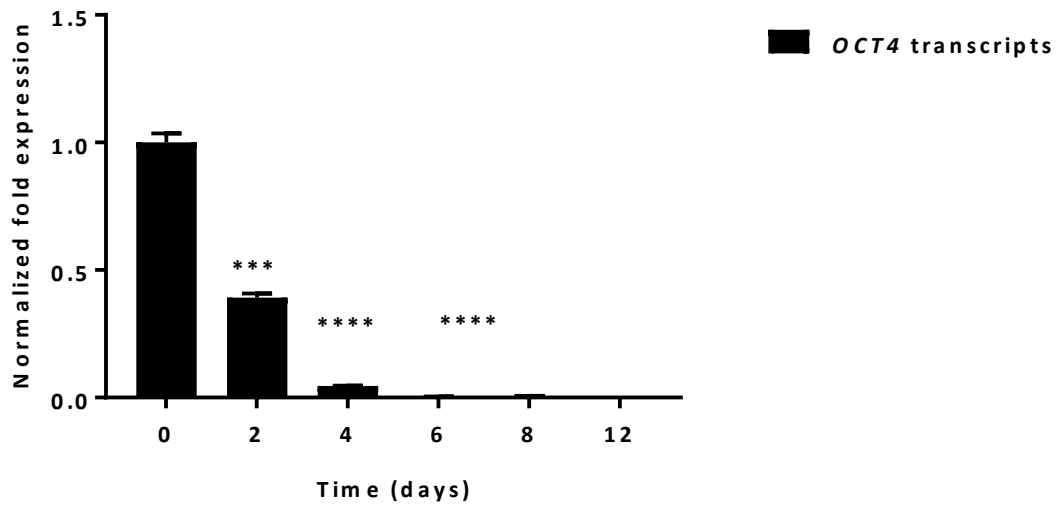


Figure 5.6. Analysis of the *OCT4* and *LKAAEAR1* transcript levels in DMSO treated NTERA2 cells using qRT-PCR.

(A) Bar charts represent the levels of *OCT4* transcripts in DMSO-treated cultures of NTERA2 cells from different days. (B) Bar charts represent the levels of *LKAAEAR1* transcripts in NTERA2 cells treated with DMSO for different days. The data were normalised to two endogenous reference genes (*ACTB* and *GAPDH*), and the relative fold change was computed by the $\Delta\Delta C_t$ method. Error bars denote the standard errors for the mean of technical triplicates.

A)



B)

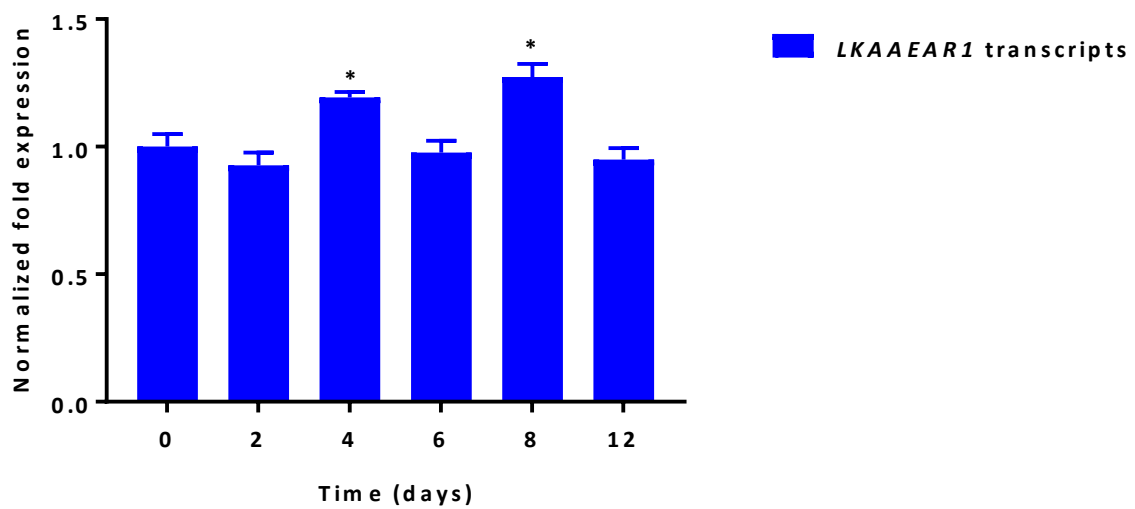
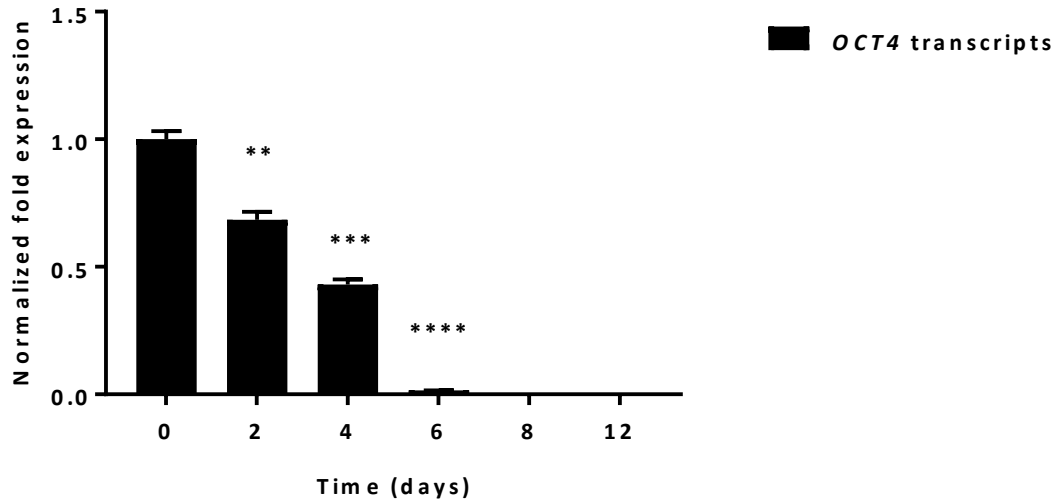


Figure 5.7. Analysis of the *OCT4* and *LKAAEAR1* transcript levels in RA-treated NTERA2 cells using qRT-PCR.

(A) Bar charts represent the levels of *OCT4* transcripts in RA-treated cultures of NTERA2 cells from different days. (B) Bar charts represent the levels of *LKAAEAR1* transcripts in RA-treated NTERA2 cells from different days. The data were normalised to two endogenous reference genes (*ACTB* and *GAPDH*), and the relative fold change was computed by the $\Delta\Delta C_t$ method. Error bars denote the standard errors for the mean of technical triplicates. (*: $P < 0.05$, ***: $P < 0.001$ and ****: $P < 0.0001$). This figure represents results from one of three biological repeats.

A)



B)

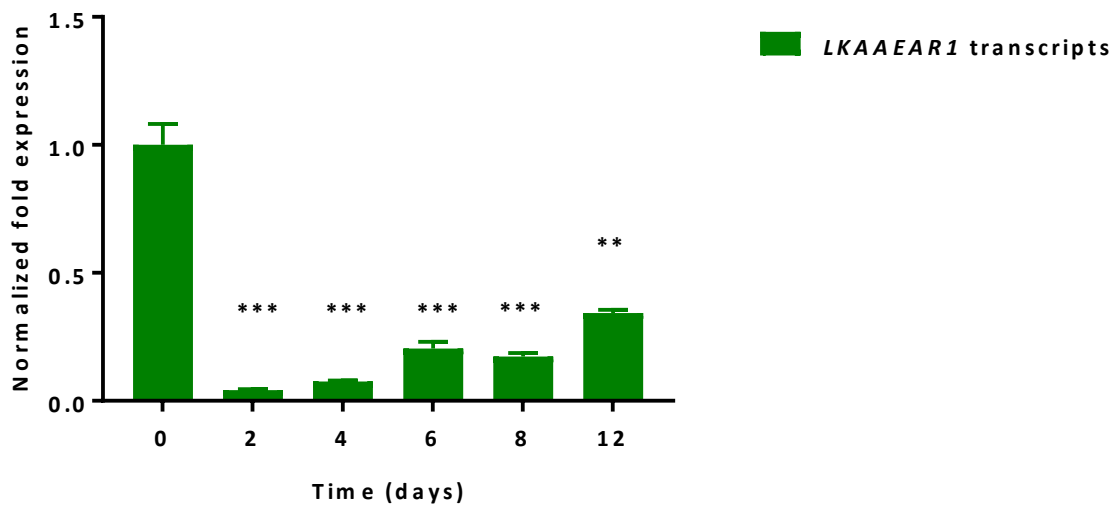


Figure 5.8. Analysis of the *OCT4* and *LKAAEAR1* transcript levels in HMBA-treated NTERA2 cells using qRT-PCR.

(A) Bar charts represent the levels of *OCT4* transcripts in HMBA-treated cultures of NTERA2 cells from different days. (B) Bar charts represent the levels of *LKAAEAR1* transcripts in HMBA-treated NTERA2 cells from different days. The data were normalised to two endogenous reference genes (*ACTB* and *GAPDH*), and the relative fold change was computed by the $\Delta\Delta C_t$ method. Error bars denote the standard errors for the mean of technical triplicates. Asterisks above the bars refer to the p-values (**: $p < 0.01$, and ***: $p < 0.001$). This figure represents results from one of three biological repeats.

5.2.4. The expression patterns of *LKAAEAR1* upon differentiation of human embryonic stem cells (hESCs).

The levels of *LKAAEAR1* transcripts were further assessed in hESC line H9 using qRT-PCR analysis. H9 cell line was differentiated using RA agent for 10 days. The same cDNA was used from Section 4.2.4, to evaluate the expression pattern of *LKAAEAR1* in differentiated hESCs. Results in Figure 5.9 showed that mRNA levels of *LKAAEAR1* were significantly reduced from day 7 with ($p < 0.05$). *LKAAEAR1* transcripts were observed to be expressed in normal neuronal tissues, however, this significant reductions may suggest its role in stemness.

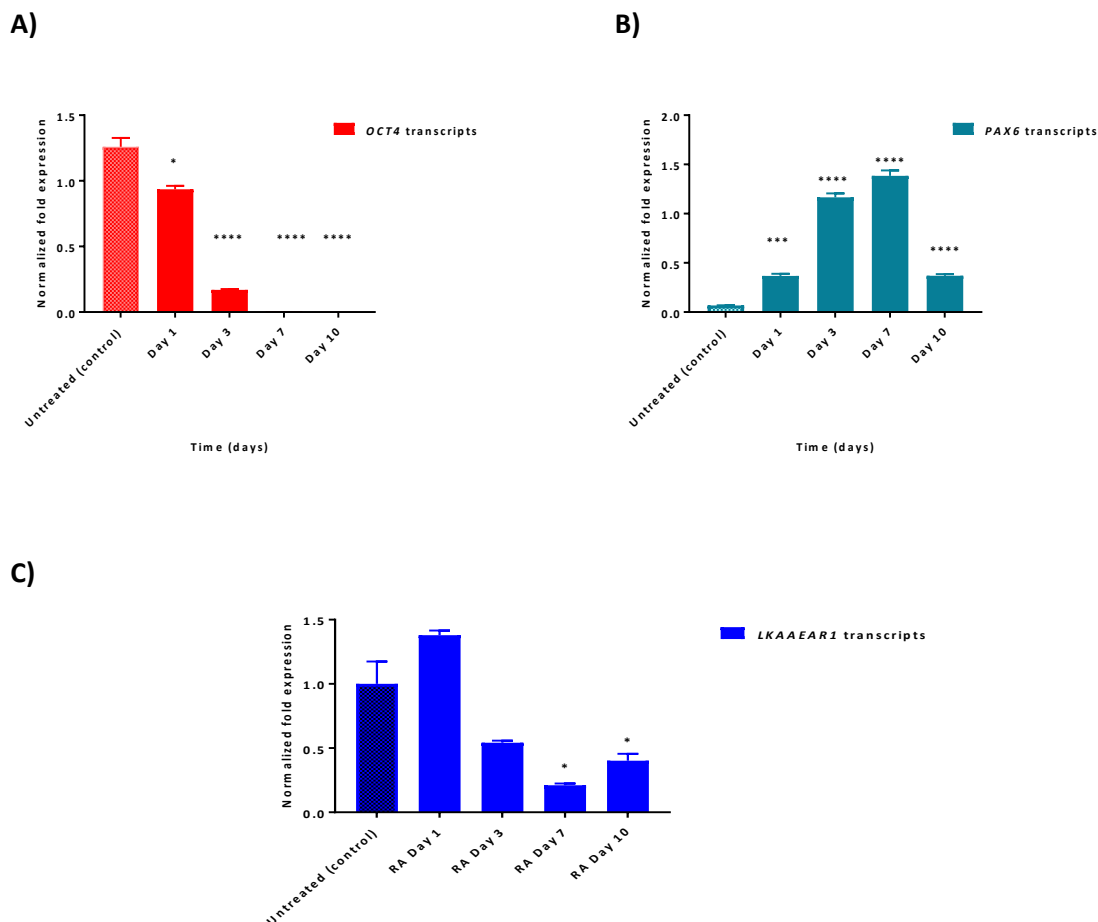


Figure 5.9. qRT-PCR analysis of *LKAAEAR1* transcript levels in H9 cells treated with retinoic acid. (A) Bar chart representing the transcript levels of *OCT4* in H9 cells treated with RA for different days as a stemness marker. (B) Bar chart representing the transcript levels of *PAX6* gene during the differentiation process as a neuronal marker. (C) Bar chart representing the transcript levels of *LKAAEAR1* during the differentiation process. The transcript levels were normalised to *YWHAZ* and *ACTB* as endogenous reference genes. Error bars represent the standard errors for the mean of three technical repeats. The p-values (*: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$ and ****: $p < 0.0001$).

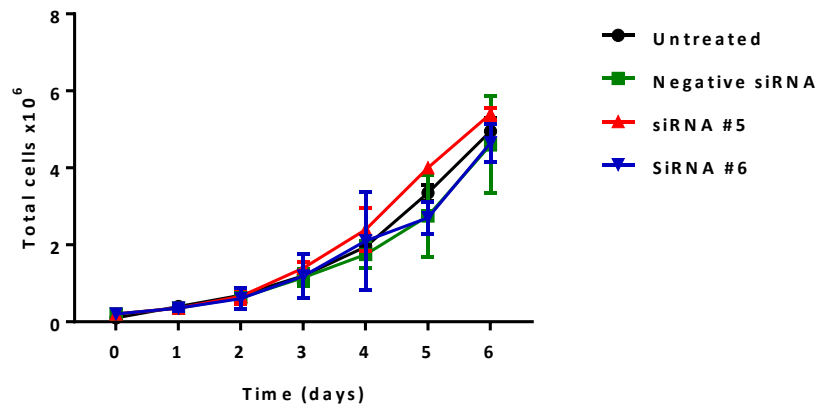
5.2.5. Assessment of cell growth and proliferation following the depletion of *LKAAEAR1*.

Generally, cancer cells are defined by their overgrowth characteristics (Hanahan & Weinberg, 2011). The downregulation of CT genes that contribute to cell developmental mechanisms, were reported to induce an apoptotic pathway or slow their growth rates (Assanga & Lujan, 2013; Shange et al., 2014). In this experiment, two siRNAs, #5 and #6, were employed on independent cultures of MCF7 and NTERA2, along with negative controls, for six days of transfection. Daily records were kept on cell morphology, confluency and cell count number to establish the growth curve.

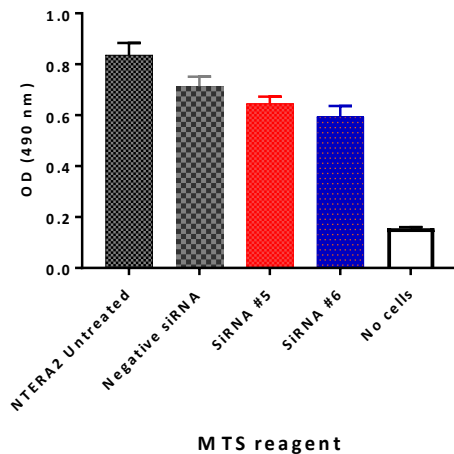
In addition, both cells were seeded on 96 well plate and the transfection experiment were employed, therefore, the viability of cells were examined by Proliferation Colorimetric Assay. This assay is based on the reduction of the MTS tetrazolium compound by viable cells to generate a coloured formazan product that is soluble in cell culture media. The formazan dye produced by viable cells can be quantified by measuring the absorbance at 490–500 nm. In this study, the CellTiter 96 AQueous One Solution Assay was used to measure the viability of MCF7 and NTERA2 cell following anti-*LKAAEAR1* siRNA #5 and #6. The results showed no significant changes of cell viability correlated with the depletion of *LKAAEAR1*. The knockdown efficiency of *LKAAEAR1* transcripts were examined using qRT-PCR on MCF7 and NTERA2 cells.

Results in Figures (5.10 and 5.11) showed no significant influence of *LKAAEAR1* transcript reductions on cell growth and/or proliferation, in comparison to negative controls.

A)



B)



C)

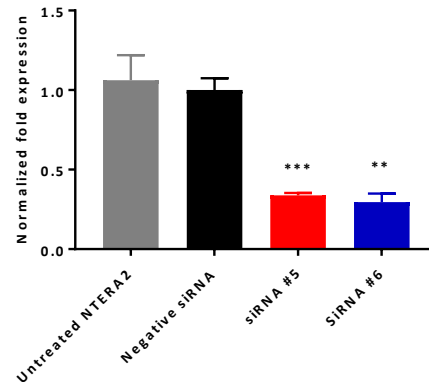
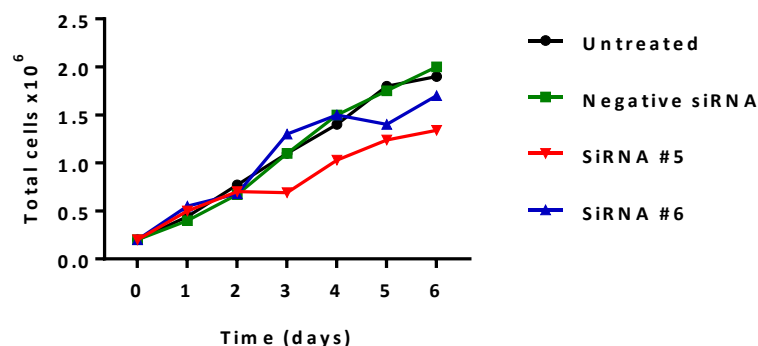


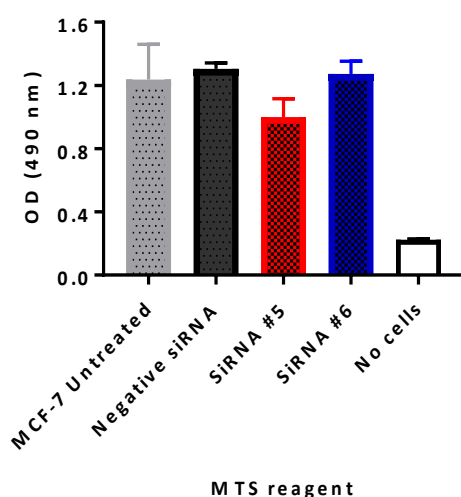
Figure 5.10. Analyses of NTERA2 cell proliferation following the depletion of *LKAAEAR1* transcripts.

(A) The graph shows the daily records of cell numbers for each type of treatment to build the growth curve for six days. There is no significant changes in cell counts of positively treated cells compared to the negative controls from three biological repeats. (B) The cell viability test using MTS Cell Proliferation Colorimetric Assay on NTERA2 cells. Cells treated with negative siRNA were used as a negative control for *LKAAEAR1* knockdown, and untreated cells were utilised as a positive control for the treatment. Three wells were injected with media only (no cells) as a negative control for the experiment. The experiment was conducted in a 96-well plate using the CellTiter 96 AQueous One Solution Assay. Bars show the optical density of NTERA2 viable cells from three biological repeats, there are no significant changes to the negative control. (C) Analysis of *LKAAEAR1* transcript levels using qRT-PCR to check knockdown efficiency, (**: $p < 0.01$ and ***: $p < 0.001$).

A)



B)



C)

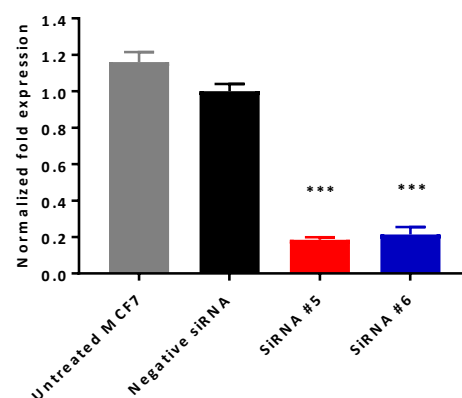


Figure 5.11. Analyses of MCF7 cell proliferation following the depletion of *LKAAEAR1* transcripts.

(A) The graph shows the daily records of cell numbers for each type of treatment to build the growth curve for six days. There is no significant changes in cell counts of positively treated cells compared to the negative controls from one biological repeat on MCF7 cells. (B) The cell viability test using MTS Cell Proliferation Colorimetric Assay on NTERA2 cells. Cells treated with negative siRNA were used as a negative control for LKAAEAR1 knockdown, and untreated cells were utilised as a positive control for the treatment. Three wells were injected with media only (no cells) as a negative control for the experiment. The experiment was conducted in a 96-well plate using the CellTiter 96 AQueous One Solution Assay. Bars show the optical density of MCF7 viable cells from three technical repeat, there are no significant changes to the negative control. (C) Analysis of *LKAAEAR1* transcript levels in MCF7 cells using qRT-PCR to check knockdown efficiency, (***: $p < 0.001$).

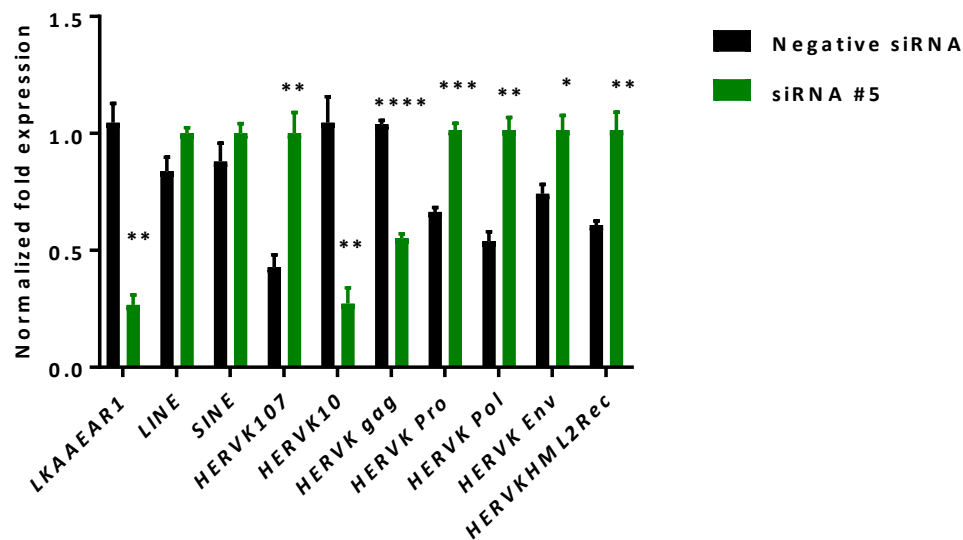
5.2.6. *LKAAEAR1* regulates transposon elements (TEs) in cancer cells.

Approximately half of human genome is composed of mobile DNA sequences, termed transposable elements (TEs). The activity of TEs was reported to be associated with oncogenesis (Goodier, 2014; Chénais, 2015). Furthermore, some CTAs were identified to contribute in tumourigenesis through their influence on TEs. For instance, *Tdrd12* was identified to inhibit the activation of *LINE-1* in murine germ cells (Pandey et al., 2013) and *TEX19* was also reported to regulate many TEs in cancerous cells (Planells-Palop et al., 2017). In this experiment, the levels of *LKAAEAR1* mRNA were downregulated in MCF7 and NTERA2 cells using siRNA #5 and #6 molecules. To inspect the correlation between *LKAAEAR1* and TEs, changes in mRNA transcript levels for multiple TEs in MCF7 and NTERA2 depleted of *LKAAEAR1* were analysed by qRT-PCR. This study examined different groups of TEs, including non-LTR retrotransposons (*LINE-1* and *SINE*) and LTR retrotransposons (*HERVs*). The multiple retrotransposons belonging to the *HERVs* group were *HERV K gag*, *HERV K pro*, *HERV K 107*, *HERV K 10*, *HERV K env*, *HERV K pol*, and *HERV K HML2 rec*. Also, MM ERV 10c, which belongs to the mouse genome, was used as a negative control. Generally, TEs were differentially expressed in distinct cancer types. The results demonstrated that the levels of TE mRNA transcripts were variable depending on the cancer cell types.

The knockdown of the *LKAAEAR1* transcripts in MCF7 cells resulted in the upregulation of transcript levels of most TEs. Transcript levels of *LINE-1* and *SINE* showed significant elevations ($p < 0.01$) following *LKAAEAR1* reduction by siRNA #6 but no changes were observed in siRNA #5-treated cells. From *HERV K* group, the upregulations in transcript levels were observed in *HERV K 107*, *HERV K pro*, *HERV K pol*, and *HERV K env* in both types of treatments. However, the levels of *HERV K 10* transcripts exhibited significant changes but were different dependent on siRNA molecules. Additionally, *HERV K gag* levels showed significant downregulations ($p < 0.0001$ and $p < 0.01$) following *LKAAEAR1* mRNA depletion by siRNA #5 and #6 respectively (Figure 5.12).

The *LKAAEAR1* transcript levels reduction in NTERA2 cells showed significant decreases in *HERV K gag* levels ($p < 0.01$ in siRNA #5- and $p < 0.05$ in siRNA #6- treated cells). Furthermore, NTERA2-treated with siRNA #5 showed reductions in *HERV K pol*, *HERV K pro* and *HERV K env* levels, however, only *HERV K pol* was downregulated due to siRNA #6 treatment (Figure 5.13).

A)



B)

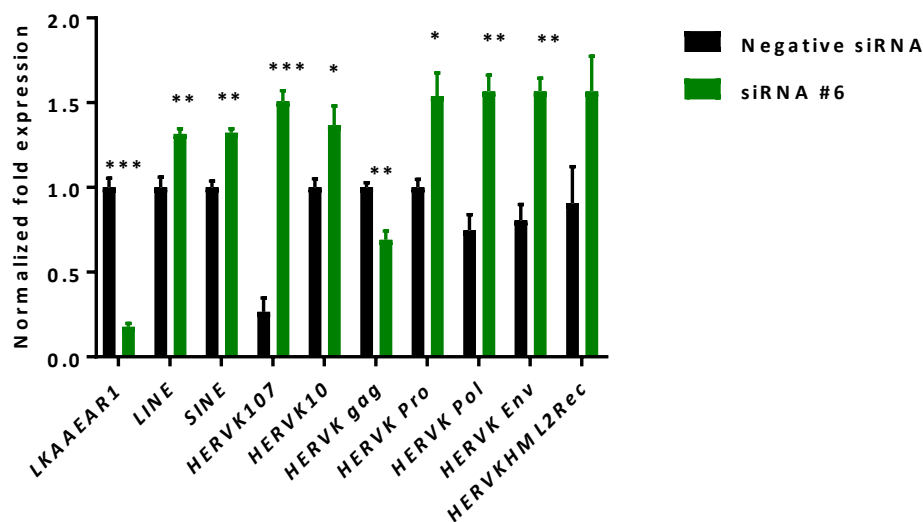
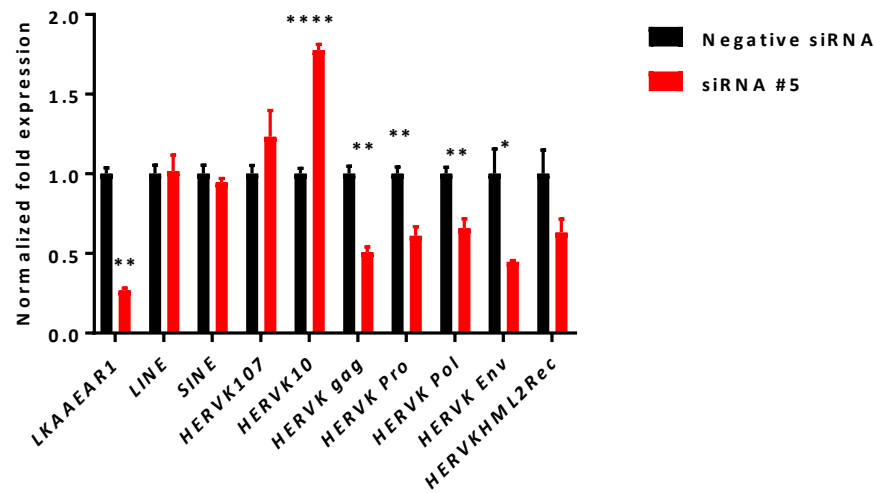


Figure 5.12. TE mRNA expression levels in MCF7 cells depleted of *LKAAEAR1* transcripts using qRT-PCR analysis.

Bar chart shows the changes in mRNA transcript levels of TEs after the depletion of *LKAAEAR1* in MCF7 cells by (A) siRNA #5 and (B) siRNA #6 molecules. *ACTB* and *GAPDH* were the endogenous reference genes used to normalise the data. Error bars refer to the standard error for the mean of three technical repeats. Asterisks above the bars show p-values (*: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$ and ****: $p < 0.0001$).

A)



B)

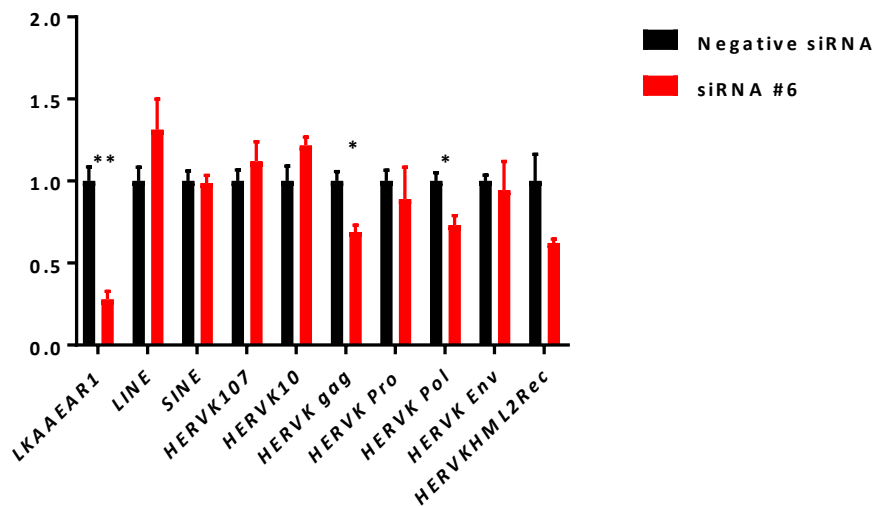


Figure 5.13. TE mRNA expression levels in NTERA2 cells depleted of *LKAAEAR1* transcripts using qRT-PCR analysis.

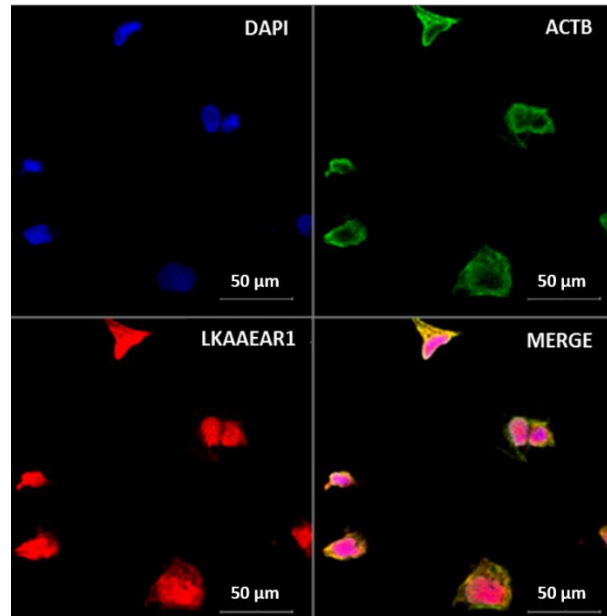
Bar chart shows the changes in mRNA transcript levels of TEs after the depletion of *LKAAEAR1* in NTERA2 cells by (A) siRNA #5 and (B) siRNA #6 molecules. *ACTB* and *GAPDH* were the endogenous reference genes used to normalise the data. Error bars refer to the standard error for the mean of three technical repeats. Asterisks above the bars show p-values (*: $p < 0.05$, and **: $p < 0.01$).

5.2.7. The localisation of LKAAEAR1 proteins in cancer cells.

The presence of the LKAAEAR1 protein in distinct types of cancers suggested possible functions of this protein in oncogenesis. The assessment of cellular localisation may give insight to explore these functions. In this study, two cancer cell lines; NTERA2 and MCF7 cells were employed for indirect immunofluorescent (IF) method to determine the cellular localisation of LKAAEAR1 protein. A polyclonal anti-LKAAEAR1 antibody (Abcam, ab108142) was used to detect the protein. In the IF method, DAPI (blue) was used to stain the nucleus and ACTIN B (green) was used for cytoplasmic staining.

The analysis results of IF images showed that the LKAAEAR1 protein is located in the nucleus and cytoplasm in both examined cells; NTERA2 (Figure 5.14) and MCF7 (Figure 5.15). To further confirm LKAAEAR1 protein localisation, western blot analysis was conducted on nuclear and cytoplasmic extracts of MCF7 cells. The results demonstrated that LKAAEAR1 protein was detected at the expected size of 28 kDa in cytoplasmic extracts, in addition to emerging band at about 50 kDa. However, only one band at approximately 50 kDa was emerged in nuclear extracts of MCF7 cells (Figure 5.15 D).

A)



B)

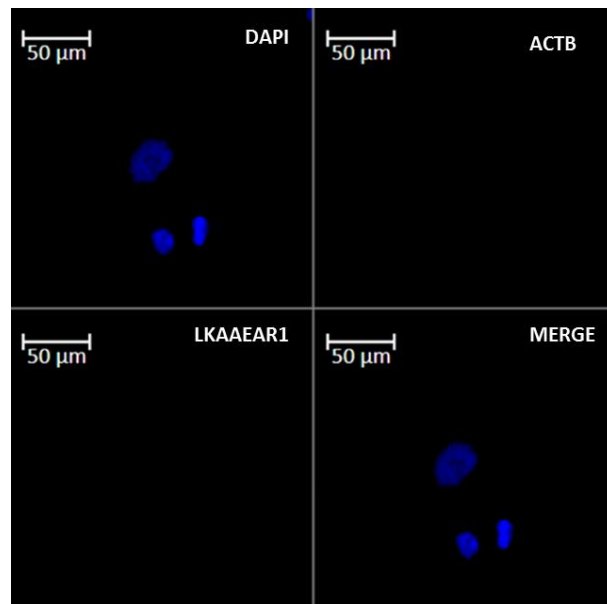


Figure 5.14. Cellular localisation LKAAEAR1 protein in NTERA2 cells. (A) IF staining of fixed NTERA2 cells with ACTIN B antibody (green) detected by mouse monoclonal antibody from (Abcam: ab6277) as a positive control and cytoplasmic marker. LKAAEAR1 protein (red) was detected by rabbit polyclonal antibody from (Abcam: ab108142) in three biological repeats from cell cultures. DAPI blue was used to stain DNA, and images were captured using a Zeiss LSM 710 confocal microscope with 40X oil objective. The LKAAEAR1 protein was shown to be localised in the nucleus and cytoplasm of NTERA2 cells. (B) IF staining of fixed NTERA2 cells with only secondary antibodies: anti-mouse secondary antibody (Invitrogen; #A11029, green) and anti-rabbit secondary antibody (Invitrogen; #A11011, red). This stain is the negative control for the primary antibodies, which is used to fine-tune the setting before viewing the positive samples to avoid background from secondary antibodies.

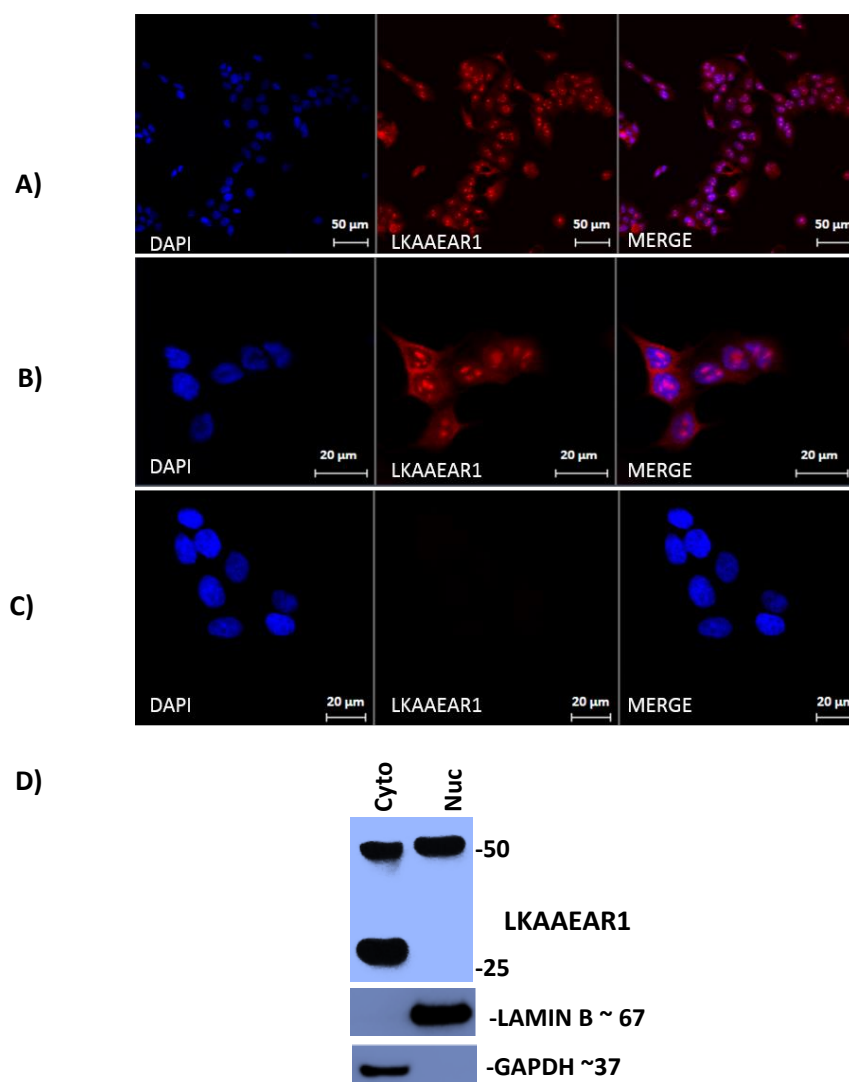


Figure 5.15. Cellular localisation LKAAEAR1 protein in MCF7 cells. IF staining of fixed MCF7 cells shows that the LKAAEAR1 protein (red) was detected by rabbit polyclonal antibody from abcam (ab108142). DAPI blue was used to stain DNA in all images. (A) Image of confluent culture population; (B) image for sub confluent culture; (C) negative control stained with secondary antibody only. The LKAAEAR1 protein is shown to be localised in the nucleus and cytoplasm of MCF7 cells. Images were captured using a Zeiss LSM 710 confocal microscope with 40X oil objective; (D) One western blot analysis is represented from two biological repeats on subcellular protein fractionations from MCF7 cell compartments. The positive controls were anti-GAPDH for cytoplasm and LAMIN B for nuclear proteins.

5.2.8. The localization of LKAAEAR1 protein in testis tissue.

The functional roles of LKAAEAR1 gene is unknown although it is predicted to play a critical role in spermatogenesis as it is strongly expressed in testis. This study utilized IHC staining analysis to identify the presence of LKAAEAR1 protein in normal testis section, to detect the protein localization and also to optimize the antibody concentration and specificity. For this purpose, primary antibody (rabbit polyclonal anti- LKAAEAR1) was purchased from Abcam (ab108142) to detect the distribution of protein. Also, paraffin-embedded testis tissues sections were obtained from Amsbio Company.

The preliminary data demonstrated positive signals in testis tissue sections, particularly in seminiferous tubules. The results clearly show strong positive stain in the basal compartment of seminiferous tubules, which is containing mitotic dividing spermatogonia, meiotic dividing spermatocytes and early spermatid cells. Depending on cell morphology, spermatogonia cells are described as spherical cells with round nuclei containing condensed chromatin, while the primary spermatocytes can be identified by their size alone as they are markedly increased cellular and nuclear size (the largest cells in spermatogenesis) with granular chromatin (Meyer, 1985). Given that, majority of primary spermatocytes were the stained cells with more intense signals in nucleus than in cytoplasm. Some of spermatogonia cells also appeared to be weakly stained and this may indicate that these spermatogonia which are committed to differentiate and this depends on spermatogonia cell types. The stain is obviously seen to be dispersed through the seminiferous tubule cytoplasm which may include Sertoli cells. The spermatid cells (early and elongated sperms) can be clearly seen as not stained. These results show that the production of LKAAEAR1 protein tends to be more tightly limited to meiotic dividing spermatocyte cells within the seminiferous tubules (Figure 5.16). These results were compared to the negative control with nonspecific primary antibody and H&E stain in (Figure 5.17).

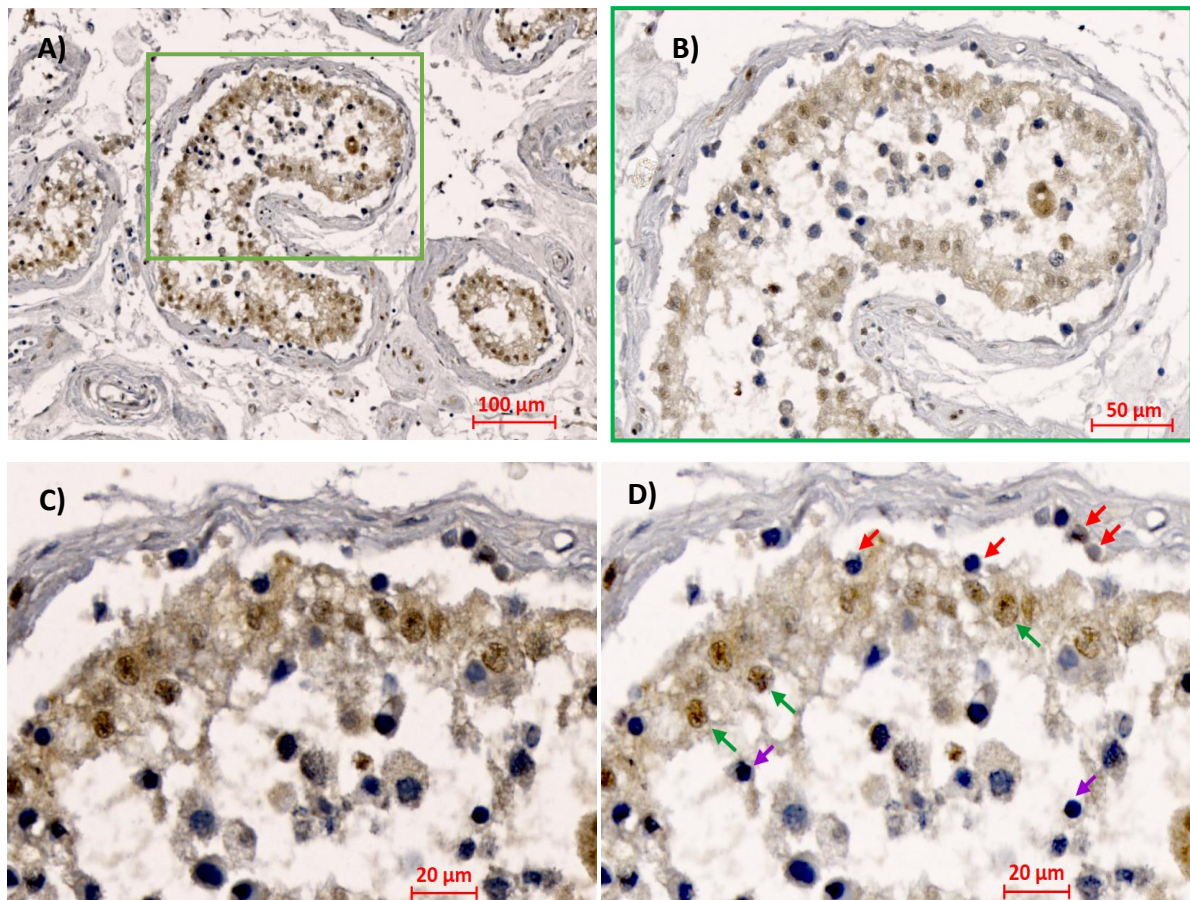
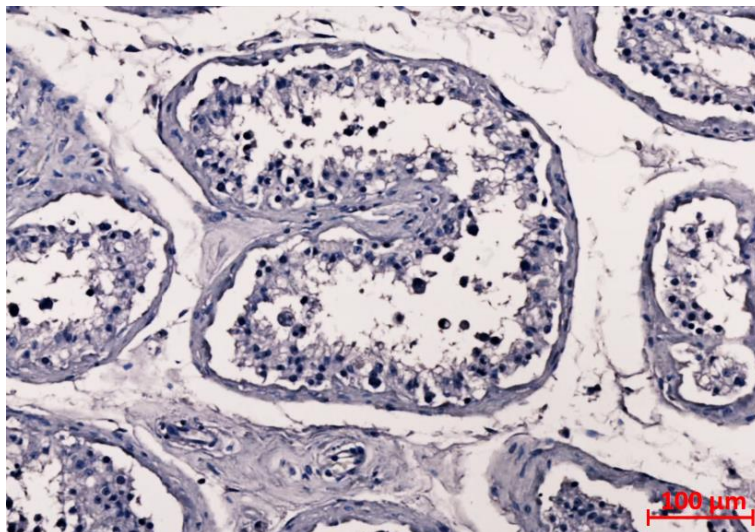


Figure 5.16. IHC staining of LKAAEAR1 protein in normal testis section

IHC was carried out by hand method using the rabbit polyclonal antibody (Abcam: ab108142) at 1/250 ug/ul concentration. (A) The testis section shows positive stain in seminiferous tubules, (B) and (C) shows the periphery of seminiferous tubules is positively stained. D) Image with 20X magnification demonstrating that primary spermatocytes (green arrow) are strongly stained, weak signals of staining in some of spermatogonia cells (red arrow) whereas early spermatid cells (purple arrow) show negative stain.

A)



B)

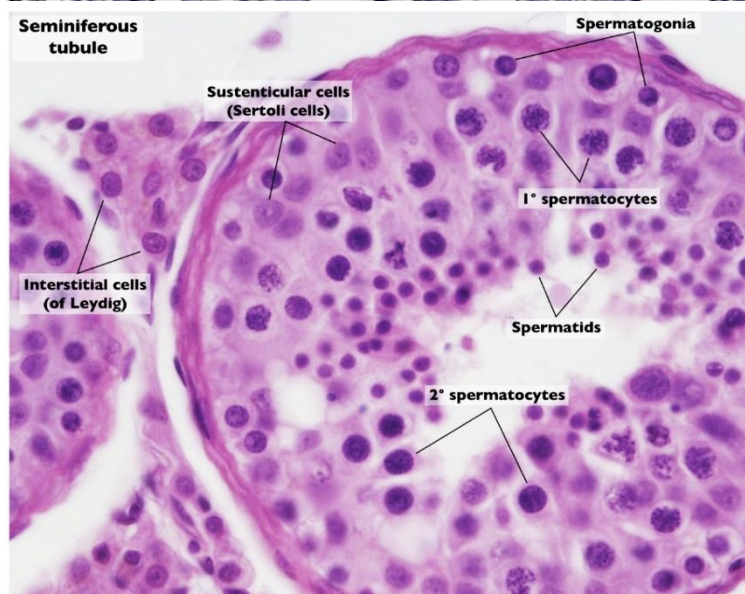


Figure 5.17. Negative control staining of testis tissue section by IHC analysis and H&E staining.

IHC method was performed with only rabbit serum (secondary antibody) as negative control along with positive stain under the same condition.

H&E staining with label obtained from (School of Anatomy and Human Biology - The University of Western Australia).

5.3. Discussion

5.3.1. *LKAAEAR1* protein is a Testis/CNS CTA

In 2012, Feichtinger and co-workers were the first who identified that *LKAAEAR1* is a CTA candidate (Feichtinger et al., 2012). Later, Kamata and colleagues evaluated the levels of mRNA transcripts on multiple normal and cancer human tissues using qRT-PCR and confirmed that *LKAAEAR1* is a good CTA gene candidate (Kamata et al., 2013). This work was the first to validate *LKAAEAR1* protein production in normal and cancerous tissues. The results in the current study demonstrate that *LKAAEAR1* protein was identified in the male testis and brain from human normal tissues, and multiple cancer tissues, including embryonal carcinoma, melanoma, leukaemia, and brain, colon, lung, breast, ovary and liver cancers. These results suggested that the *LKAAEAR1* protein is a good testis/CNS CTA, which is consistent with previous studies. The presence of the *LKAAEAR1* protein in all screened cancers is an indication of its potential as a diagnostic biomarker. The expected size of the protein is recommended for detection at 28 kDa which might belong to transcripts variant 2. However, a band was detected at about 50 kDa in a wide range of examined cancers and another band at approximately 48 kDa was emerged in some cases. In this regard, the downstream experiments of gene knockdown showed a slight reductions of these higher molecular weight bands, which is an initial indication that these bands belong to the *LKAAEAR1* protein species. This is consistent with a previous work was carried out by McFarlane lab and showed that the ~50 kDa band specifically was significantly increased when *LKAAEAR1* was overexpressed in cell lines (Almutairi, PhD thesis 2014, Bangor University). The alteration of protein size is possibly resulted from post-translation modifications. It is suggested that this high molecular *LKAAEAR1* protein might have undergone a polyubiquitination or multi-ubiquitination to be detected at this size.

5.3.2. The expression of *LKAAEAR1* in normal hESCs and CSCs.

It is reported that cancer cells share many commonalities with germ cells suggesting that stemness functions confer many biological features of cancer cells such as metastasis, invasion and proliferation potentials (Fratta et al., 2011; Whitehurst, 2014; Gjerstorff et al., 2015). This has led to the postulate that the activation of germline genes is suggested to drive the germ-like state and soma-to-germline transition in cancer cells (McFarlane & Wakeman, 2017). In addition, the activation of germline genes is also correlated with poor prognosis, for example, expression of germline genes in lung cancer showed to be linked with advanced

stages of aggressive and metastatic tumour (Rousseaux et al., 2013). Furthermore, it was also hypothesised that the existence of CSCs as a minority population within the tumour, with self-renewal potentials maintain cancer proliferation, invasion and metastasis (Bonnet & Dick, 1997; Dick, 2008). So, targeting CSCs within tumours cells is thought to inhibit the aggressive metastatic and recurrence features of these cancers. In this regard, the expression of *LKAAEAR1* in germ cells and CSCs may suggest its function in stemness characteristics. This study showed that *LKAAEAR1* transcript levels were significantly reduced in NTERA2 cells following HMBA induction. This might suggest its potential function and contribution to stemness features. In RA treated NTERA2 cells, the transcript levels of *LKAAEAR1* were constant, however, this is expected as this gene is known to be highly expressed in normal CNS, and RA induces NTERA2 down a neuronal lineage. This was confirmed on hESCs treated with RA showing that *LKAAEAR1* transcripts were downregulated compared to untreated cells. These results proposed the possible link of *LKAAEAR1* gene with stemness.

5.3.3. The *LKAAEAR1* gene may regulate TEs genes

First, depletion of *LKAAEAR1* transcripts showed no influence on cell growth and proliferation. However, this relation of *LKAAEAR* is not clear yet, because these methods were unable to eradicate the protein. Given this, the amount of protein after depletion might be sufficient for cell growth and proliferation. Moreover, *LKAAEAR1* transcripts depletion showed significant changes on important group of genes that are correlated with oncogenesis; transposon elements (TEs) genes. Many studies have shown that the activation of TEs occurs during oncogenesis (Goodier, 2014). The activity of TEs can influence changes of genomic integrity, resulting in genetic disorders and, subsequently, cancer development (Dhivya & Premkumar, 2016). Research focusing on these elements might be of value, as these genes are potentially prognostic and diagnostic biomarkers for some types of cancer, and thus understanding their activation and/or repression may yield various future clinical applications (Downey et al., 2015). In order to identify the function of the *LKAAEAR1* gene in tumours, the behaviour of these elements was covered in this chapter through the preliminary screening of mRNA levels from different TE groups before and after *LKAAEAR1* depletion. Importantly, no previous study has examined the association between *LKAAEAR1* and TEs. In the present study, the screening test was carried out on a range of TEs, including non-LTR (*LINE* and *SINE*), along with *HERVs* members in two cell lines MCF7 and NTERA2. Remarkably, the *HERV gag* gene was found to be a common TE downregulated in both examined cell lines with different

siRNA sequences. This indicates that there might be a strong link between *LKAAEAR1* function and *HERV gag* gene regulation. The association with the remaining studied TEs is also still outstanding, as the findings revealed that *HERV K* members (*Pro*, *pol*, *env*, and *HML2 rec*) were upregulated in MCF7 cells, but their levels were variable in NTERA2 cells; this may depend on the distinct nature of each cell lines. However, the direct connection between *LKAAEAR1* depletion and its influence on the regulation of TEs is not yet clear, and further analyses are required. Importantly, the activation of TEs has been correlated with triggering oncogenic mechanisms that result in different types of cancer development as an important aetiology factor of carcinogenesis (Paterson et al., 2015). These results may highlight that *LKAAEAR1* is possibly involved in oncogenesis through the direct or indirect regulation of TE expression.

5.3.4. LKAAEAR1 protein localisation in cancer cells

IF analysis was conducted to elucidate the cellular localisation of the LKAAEAR1 protein in different cancer cells, NTERA2 and MCF7. IF analyses demonstrated that LKAAEAR1 protein is located in cytoplasm and nucleus of both examined cells. The signals of nucleus protein appeared to be stronger than that of the cytoplasm. The presence of LKAAEAR1 protein in nucleus may suggest its roles in transcription regulation. Western blot analysis confirmed the existence of LKAAEAR1 in the cytoplasmic and nuclear extracts in MCF7 cells. Nuclear extracts showed only the higher molecular weight of LKAAEAR1 protein at (approximately 50 kDa), supporting the idea that this protein might have undergone a posttranslational modification (PTM) during trafficking process to the nucleus. It is expected that a poly/multi-ubiquitination may occur, as ubiquitination is known to play important roles in protein localisation and trafficking and subsequently promote oncogenesis (Glickman & Ciechanover, 2002; Pickart & Eddins, 2004).

Immunohistochemistry (IHC) was also conducted to determine the LKAAEAR1 protein testis section. The results showed positive staining in the periphery (basal compartment) of seminiferous tubules. Spermatocytes were positively stained with LKAAEAR1 antibody showing that the protein is present in both cytoplasm and nucleus with more intense staining in nucleus. We observed that some of spermatogonia cells were also stained but others not. However, this might depend on the spermatogonia cell types. As spermatogonia cells can be classified into type A (dark nuclei) which are usually mitotically inactive to reserve SSCs and type A (pale nuclei) which are mitotic active cells to produce SSCs and type B cells. Type B cells can give rise to primary spermatocytes (Mahla et al., 2012; ChenLiu, 2015). Furthermore,

LKAAEAR1 protein tends to be produced in spermatocyte cells from other spermatogenesis cells, however, specific markers of these cells are required to be co-stained with LKAAEAR1 protein. However, the presence of LKAAEAR1 protein in the majority of spermatocytes but not in spermatids or mature sperm may indicate the fundamental roles of LKAAEAR1 in spermatogenesis or in particular meiosis, as the spermatocytes are meiotic dividing cells. This study confirmed the potential function of LKAAEAR1 protein during spermatogenesis, although it remains unclear.

Data on the Human Protein Atlas (HPA) is available on www.proteinatlas.org, suggests that LKAAEAR1 protein is strongly staining nucleus of testis tissues and human cell lines. Although the results here showed strong nucleus staining, perinuclear and cytoplasmic region appear to be faintly stained. We also observed that LKAAEAR1 protein was accumulating in sub-nuclear regions of testis tissues and MCF7 cells as shown in Figure 5.18. The functional relevance of protein foci within the nucleus is unknown, however, it is a noteworthy observation. It might be that LKAAEAR1 is located at nucleoli, therefore, this protein might be involved in nucleolus functions; such as, cell cycle, cell differentiation, aging, ribosome biogenesis, etc. Furthermore, this localisation is possibly associated to chromatin. To obtain better conclusion, further analyses are required, for example, LKAAEAR1 would be co-stained with nucleolar markers to inspect its localisation in nucleoli. Further analysis is needed to examine LKAAEAR1-chromatin association.

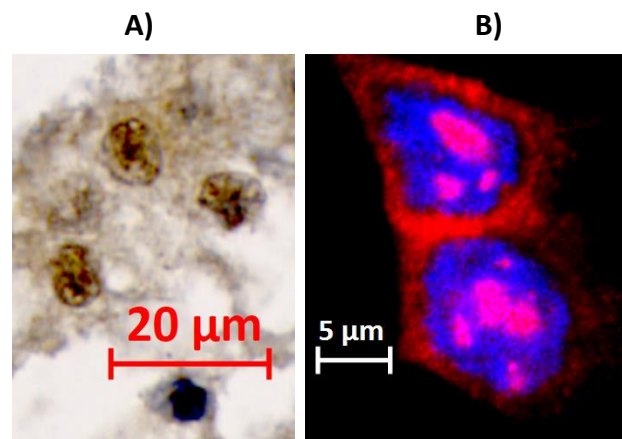


Figure 5.18. Sub-nuclear localisation of LKAAEAR1 protein in germ cells and breast cancer cells. A) Image of IHC analysis on seminiferous tubule section. B) Image of IF analysis on MCF7 cells. Both images to illustrate the dark staining inside the nucleus compartments.

5.3.5. Summary

In conclusion, the work in this chapter identified that LKAAEAR1 protein is a good CTA candidate, confirming previous studies on gene expression levels. Transcripts of *LKAAEAR1* appeared to be correlated to the stemness features in hESCs and CSCs. The downregulation of *LKAAEAR1* mRNA levels was successfully achieved, showing that this gene might contribute to oncogenesis through the influence on TE genes. It was clearly observed that LKAAEAR1 protein is located in cytoplasm and nucleus compartments. High molecular weight of LKAAEAR1 protein is detected in nuclear extracts of cancerous cells, suggesting that this protein may undergo PTMs, such as, ubiquitination. Since ubiquitination affects so many biochemical processes within the cell, it is not surprising that modifications in this system play a vital role in oncogenesis (Dwane et al., 2017). To ensure if LKAAEAR1 is undergoing ubiquitination process, more investigations are recommended to be carried out, for example, immunoprecipitation of LKAAEAR1 protein and the isolates then should be probed against ubiquitin protein and reversal experiment as well to investigate whether the ubiquitin protein isolates contain LKAAEAR1 protein or not.

Finally, the basal compartments of seminiferous tubules produce LKAAEAR1 protein, highlighting its fundamental roles in spermatogenesis. It is a noteworthy observation that the LKAAEAR1 protein was appeared to accumulate in nuclear substructure regions, however, its functional roles are not clear yet.

Chapter 6:

Analysis of LKAAEAR1 protein in tumour samples

6. Analysis of LKAAEAR1 protein in tumour samples

6.1. Introduction

Gene expression analyses are very useful to classify cancers into clinically relevant categories (Sadanandam et al., 2014). Many germline genes, particularly CTAs, are attractive because they may encode antigens that are highly specific for cancer cells but not healthy cells, except of the immune-privilege sites in the body including testis, ovary and placenta (Whitehurst, 2014). However, in terms of treatment, there is importance for those proteins that can be produced by cancer cells and they can be recognised by immune system as non-self antigens which initiates immunological responses to destroy the cancer cells that present these CTAs. Moreover, the known functional roles of these CTAs in addition to their immunogenicity increase the therapeutic potentials of these antigens. It has been shown that the presence of protein is correlated with the high levels of transcripts expression, for example, CT45 protein is detectable at high levels in breast, ovarian and lung cancers that have > 10% of transcripts level higher than in testicular expression. However, normal tissues with less than 1% of mRNA level compared to testicular expression often have no detectable protein. This may be a consequence of the gene expression regulations such as post-transcriptional regulatory mechanisms (Shang et al., 2014).

Immunohistochemistry (IHC), also defined as protein profiling, is an important method for studying the distribution and localisation of proteins. IHC enables the detection of antigens within the tissue sections through the use of a specialised stain which binds to specific targeted antigens. In cancer diagnosis, IHC can give prognostic and predictive data that are valuable for treatment decision making. The IHC analysis of CTA distribution has proven that CTAs are not only promising for diagnostic markers for cancers but also have ability to differentiate tumour cells from normal cells in tumour samples and detect the stages of cancers. For example, some cases show that CTAs homogenously stain almost all cells in the tumour tissue, suggesting that they were originally produced from a single cell that followed clonal expansion. In some cases, the CTA was unevenly present in the tumour, a group of cells that stained in compare to others in the bulk of tumour cells, this might be explained by the presence of cancer stem cells within the tumour or may be referred to a specific mutation or changes in DNA methylation in these CTA-positive cells. Interestingly, some CTAs were reported to be produced at higher levels in different stages of cancer, for example, TEX19 can

be detected as more frequently in early stages of colorectal carcinoma (CRC) than other stages (Planells-Palop et al., 2017).

LKAAEAR1 was defined as a meiCT gene that shows testis/CNS restricted expression pattern in normal tissues and highly expressed in breast and ovarian cancers and leukemia (Feichtinger et al., 2012). This study also carried out the meta-analyses of clinical data sets from many types of tumours. The results detected in Chapter 5 confirm the presence of its protein in many tumours but no protein was produced in normal tissues except in testis and CNS. Finally, a more recent study has investigated the differential expression of long non-coding RNAs (lnc RNAs) and mRNA between normal ovarian tissues and different types of epithelial ovarian tissues. *LKAAEAR1* mRNAs of one of many transcripts observed to be differentially expressed in epithelial ovarian cancers by 3.2 fold-change upregulation than in adjacent normal ovary epithelium (Ding et al., 2017).

The aim of work in this chapter is to inspect the protein intensity and distribution in normal and cancer biopsies that were obtained from diagnosed patients. The differential expression has been reported at mRNA levels in normal and cancerous ovarian epithelium, however, the investigation of the protein level has not previously been carried out. Thus, this study aimed to examine the staining intensity for *LKAAEAR1* in cancerous ovarian tissues compared to adjacent normal tissues. Further analysis was also carried out for first time on the progression of colorectal cancer arrays to inspect the suggestion that support the correlation of *LKAAEAR1* in cancer development.

6.2. Results

6.2.1. Staining pattern of LKAAEAR1 protein in normal testis and CNS tissues.

The distribution of LKAAEAR1 protein has been studied in human tissues biopsies that are recorded as clinically normal tissues. The tissue microarray (TMA) of formalin fixed paraffin embedded (FFPE) samples slide contains several human tissue types from 15 organs in 1.0 mm spot sizes. Most samples are of normal, non-neoplastic adult tissue, obtained from surgical resection specimens, excepting parathyroid from a hyperplastic parathyroid gland and central nervous system tissue (cerebral cortex and white matter) from autopsy (data sheet information in Appendix B-1). This TMA was purchased from the Cooperative Human Tissue Network at University of Virginia (CHTN- NORM2 slide) showing some clinical information such as tissue anatomy site, age and sex of the patients (the details in Appendix B-2). Rabbit polyclonal anti-LKAAEAR1 antibody (abcam: ab108142) was used as an initial screening of LKAAEAR1 protein presence by IHC method. The manual optimization of staining was examined to choose the best concentration at 1/500 of the antibody.

The results from Chapter 5 on the normal lysates confirmed that LKAAEAR1 protein was detected in testis and central nervous system. The tissue microarray screen analysis shows the presence of LKAAEAR1 protein in testis section that also used as a positive control for any CTA protein. This study showed that LKAAEAR1 protein is present in testis and CNS sections as shown in (Figure 6.1). In the testis, the stain of LKAAEAR1 protein is dispersed mostly through the cytoplasm and some nuclei of basal compartments in seminiferous tubules. Multiple tissues sections from CNS showed strong positive staining for LKAAEAR1 protein. The staining pattern of LKAAEAR1 protein is clearly distributed in the cytoplasm of the cerebral cortex and peripheral nerve tissues. However, the white matter (subcortical) section showed positive staining in cytoplasm and about 40% of nuclei (Figure 6.1 and Table 6.1).

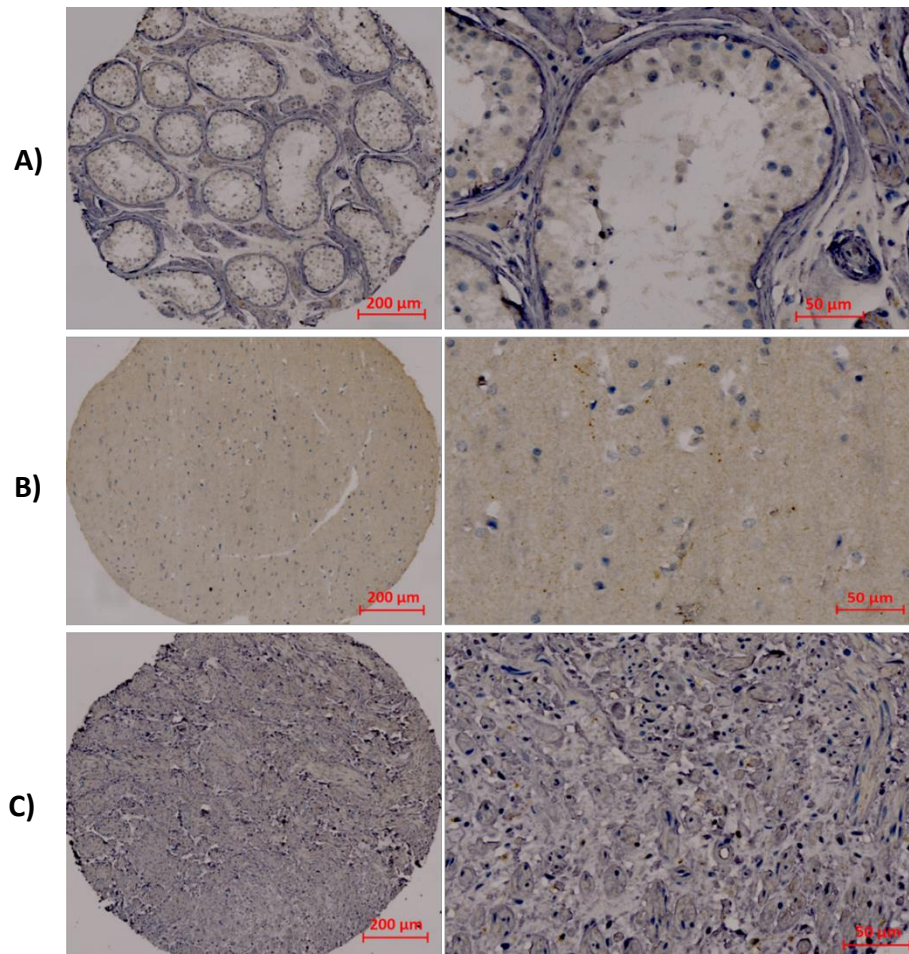


Figure 6.1. IHC staining of seminiferous tubules and central nervous system selected sections by anti-LKAAEAR1 antibody (abcam: #ab108142). A) Normal seminiferous tubules section from male testis. B) Cerebral cortex section shows mostly cytoplasmic staining. C) Normal peripheral nerve section with weak cytoplasmic staining. Images were captured by a ZEISS AXIO scan.Z1 digital scanner and each figure with a different power magnification.

6.2.2. Staining pattern of LKAAEAR1 in normal tissues sections from the microarray.

The TMA slides contain a variety of normal tissues were used to assess the staining pattern of LKAAEAR1 protein in normal tissues. Results in Figure 6.2 showed that sections from breast, prostate and stomach tissues are clearly negative. However, a very weak staining was observed in the cytoplasm of normal ovarian epithelium tissue. This staining is consistent with a previous study that identified *LKAAEAR1* mRNAs in normal ovarian and placenta tissues (Feichtinger et al., 2012). Furthermore, the cytoplasm of colonic crypt cells showed negative to a very weak signal of LKAAEAR1 staining (Figure 6.2).

Further analysis is demonstrated in Figure 6.3 to show the LKAAEAR1 staining on many normal tissues. This results demonstrated that normal tissues sections of fallopian tube, salivary gland, and heart myocardium smooth muscle tissues are clearly negative. Staining on bronchus epithelium tissues appeared to be very weak which is considered negative in comparison to testis tissues as a positive control. The placenta section showed a very weak stain in cytoplasm with scant nuclear staining. The central part of bladder tissue section is clearly negative but there is a weak staining at the edge of the biopsy and this is suggested to be resulted from the poor quality, folded edge and/or atrophied sample (often referred to as 'edge effect' in IHC). Sections from uterine, endometrium tissue, showed positive staining at cytoplasm of crypt cells. Moreover, there is a weak staining in the cytoplasmic parts of pancreas, liver, kidney, small intestine and thymus tissues. Table 6.1 demonstrated a brief description for the staining pattern of LKAAEAR1 protein on all examined normal tissues in this microarray slide. The positive staining in the normal tissues does not correlate to gene expression analyses (Feichtinger et al., 2012).

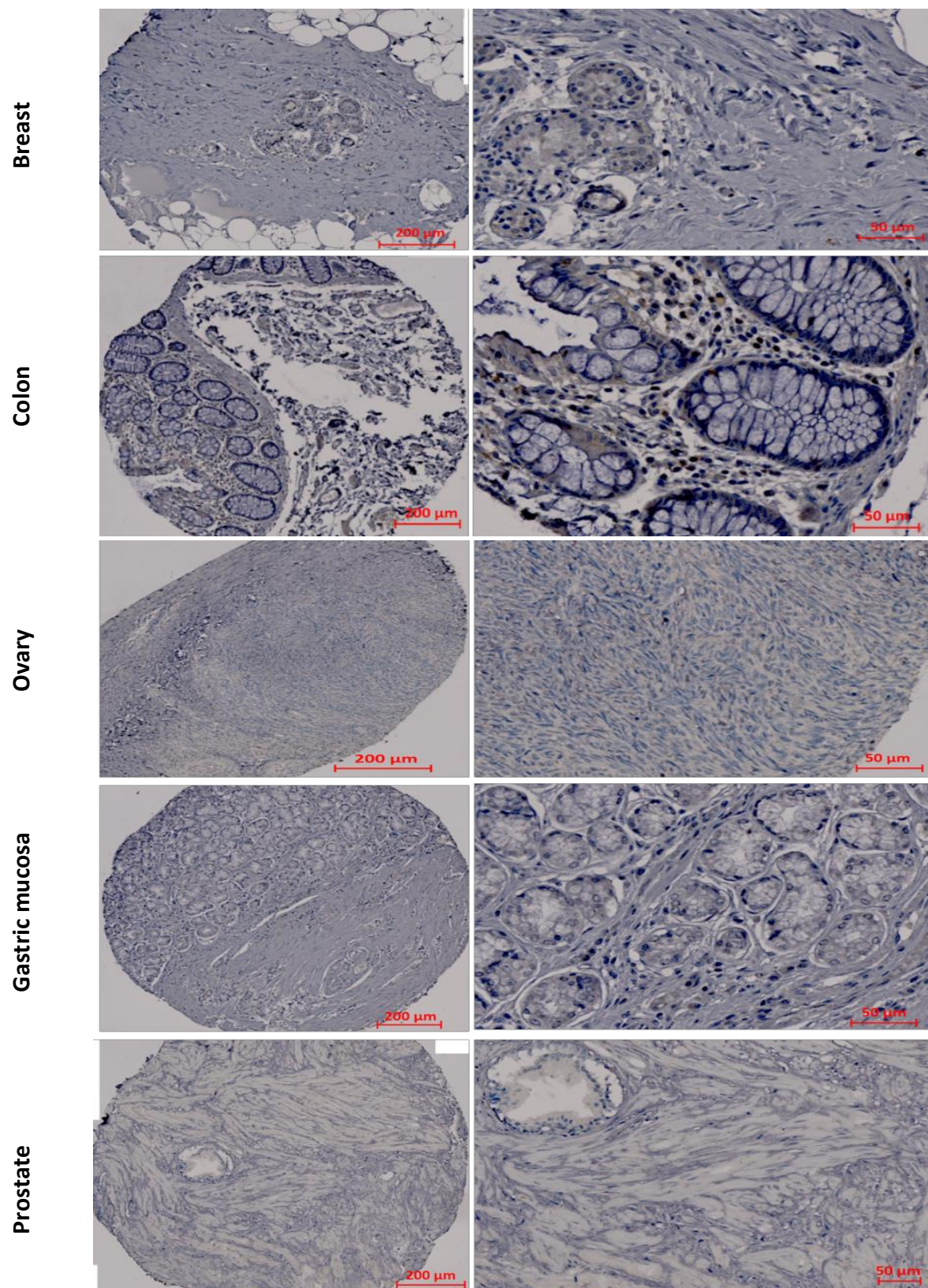


Figure 6.2. IHC staining of LKAAEAR1 protein on normal breast, colon, ovarian epithelium, stomach and prostate sections. Staining with the rabbit polyclonal anti-LKAAEAR1 antibody from abcam (ab108142) to show the protein distribution in many normal tissues. Images were captured by a ZEISS AXIO scan.Z1 digital scanner and each figure with different power magnification.

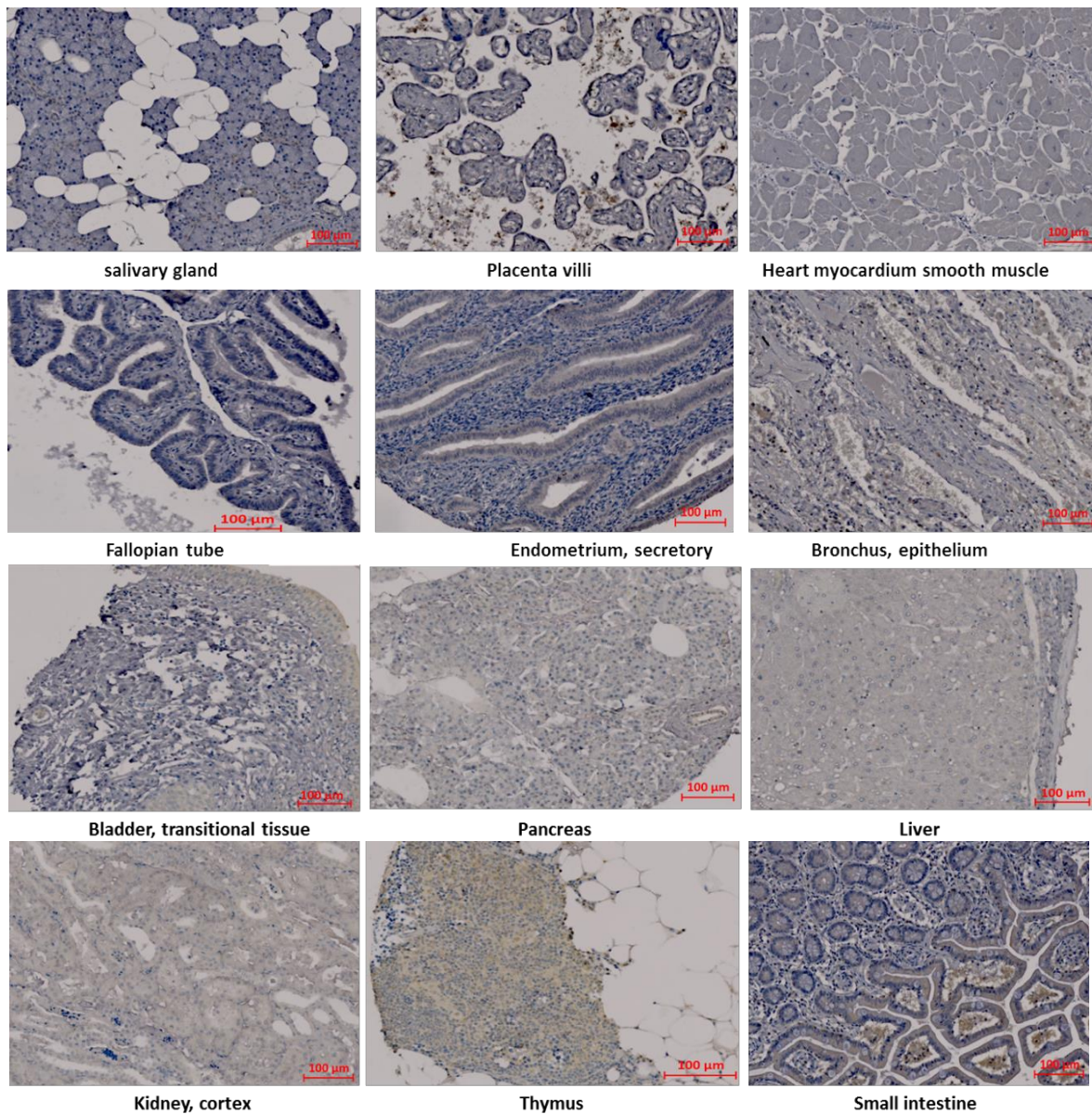


Figure 6.3. Example of normal tissues sections from CTHN that stained anti-LKAAEAR1 antibody. IHC staining was carried out on many normal tissues showing different distributions of LKAAEAR1 staining pattern. Images was captured by a ZEISS AXIO scan.Z1 digital scanner.

Table 6.1. Summary of IHC analysis of LKAAEAR1 staining on normal tissues.

Tissue	Staining pattern
Breast, epithelium	Negative
Aorta, smooth muscle	Negative
Heart, myocardium	Negative
Adrenal gland, cortex	Weak cytoplasmic
Adrenal gland, medulla	Weak cytoplasmic and ~ 10% nuclear staining
Parathyroid (hyperplastic)	Moderate staining (cytoplasmic)
Ectocervix	Negative
Endocervix	Negative
Endometrium, secretory	Negative with faint cytoplasmic staining at crypts
Fallopian tube	Negative
Ovary, corpus luteum	Moderate staining in cytoplasm with 40-50 % nuclei
Ovary, epithelium	Faint cytoplasmic staining
anus, mucosa	Negative
Colon, mucosa	Faint cytoplasmic staining at crypts (5-10% nuclei)
Esophagus, squamous mucosa	Negative (some part showed a very weak staining at cytoplasm)
Gastric mucosa,	Negative
Small intestine, mucosa	Negative with moderate to strong staining at crypts
Epididymis	Moderate positive
Prostate	Weak cytoplasmic staining
Seminiferous tubules	Positive
Seminal vessicle	Positive
Gallbladder	Positive (cytoplasmic mostly)
Pancreas, islet cell	Weak to moderate staining
Liver	Weak cytoplasmic
Pancreas	Moderate staining at cytoplasmic with 30% of nuclei faintly stained
Lymph node	Negative to very weak cytoplasmic staining
Thymus	Weak staining (cytoplasmic with about 60 % of nuclei)
Mesothelium	Negative
Cerebral cortex	Positive cytoplasmic staining (20 % nuclei staining)

White matter (subcortical)	Strong cytoplasmic staining and 50 % of nuclei are faint
Peripheral nerve	Weak staining at cytoplasm
Salivary gland (parotid)	Negative
Tonsil, squamous epithelium	Negative (some parts of the section with weak positive cytoplasmic staining)
Amniotic membrane	Negative
Placenta, villi	Negative
Alveoli	Negative
Bronchus, epithelium	Very weak positive staining at cytoplasm with 30 % nuclei moderate staining
Skin, squamous epithelium	Negative
Cartilage, articular	Negative
Skeletal muscle	Negative
Smooth muscle, intestine	Negative
Smooth muscle, uterus	Negative
Bladder, transitional epithelium	Negative with some parts weakly positive in the section
Kidney, cortex	Moderate positive
Kidney, medulla	Moderate positive

6.2.3. Staining pattern of LKAAEAR1 protein in tumour tissues sections.

Based on the observations from previous studies that identified LKAAEAR1 antigens were highly expressed in patients with prostate cancer (Kamata et al., 2013), in addition to the differential expression of LKAAEAR1 mRNAs between normal and cancerous ovarian epithelium tissues (Ding et al., 2017), we carried out IHC analysis on tumour microarray samples to inspect the staining pattern of this protein. Microarray containing multiple cancer tissue sections were stained to study the presence of LKAAEAR1 protein and its distribution. TMA slide employed in the following experiments was purchased from the US BIOMAX Inc. and the list containing all different tissue types with its related information such as age, sex, anatomical site of each section and the confirmed diagnosis of each case with relative cancer stage (Appendix B-3). The TMA slide (FDA808-c2) contains about 72 tissues sections that include roughly 54 cancer cases and about 18 sections are from normal adjacent cancer tissues. Some cores of these biopsies are of poor quality which are not clear to be assessed here. Generally, LKAAEAR1 protein was present in many cancer cases although the staining pattern vary from case to case and this may depend on the cancer stage. The data show that the diagnosis of cancer is included with international grading and staging system. There are two types of cancer staging: TNM staging and number staging systems (Edge et al., 2010). TNM staging refers to (Tumour, Nodes and Metastasis) where T describes the tumour size and can be included with number from 1 to 4 (e.g., T1 indicates the tumour is small whereas 4 means large tumour. Lymph node (N) refers to whether the tumour has spread to lymph nodes or not and may be attached with numbers (between N0- no cancer cells invade lymph nodes and N3 stage means a lot of lymph nodes are containing cancer cells). Metastasis (M stage) refers to whether this cancer has invade or metastasised to distant organs and can be expressed by M0 which means no metastasis or M1 indicates that cancer has spread to other parts of the body. Examples on this system are: T2 N1 M0, this diagnosis means that a small cancer has spread to lymph nodes but not to anywhere else, T4 N3 M1 means more advanced cancer case with aggressive invasion to other organs. Number staging system categorises cancer to four different stages expressed by Roman numerals. Briefly, stage I means the cancer is small and still contained within its start organ. Stage II refers to its increasing and becomes larger but has not spread to the surrounding tissues. Stage III, means that cancer is larger and may penetrate the nearby tissues and close lymph nodes in the area. Stage IV usually means aggressive cancer that so called secondary or metastatic cancer that spread to

distant organs. Some diagnostic criteria include (stage 0) for reporting small group of cells that are growing faster or with different morphology comparing with neighbouring cells that may develop into cancerous cells in future and they are termed as carcinoma or neoplasm *in situ* (Edge et al., 2010).

6.2.3.1. Staining pattern of LKAAEAR1 on control samples (normal tissues in tumour array).

The FDA808c-2 slide contains some normal sections and normal adjacent to cancer tissues that were used as a control for the staining. *LKAAEAR1* is CT/CNS restricted gene and its protein production should not be produced in normal tissues except of CNS tissues. Figure 6.4 showed some examples of normal tissues that are included in the slide to inspect the LKAAEAR1 protein staining pattern. Normal lingual gland and mesothelial pericardium tissues are diagnosed as normal sections from healthy individuals and they were clearly negative for LKAAEAR1 protein. Normal adjacent cancer (NAT) skin showed different pattern as the top of section is clearly negative, the middle part is weakly stained and the bottom edge was strongly stained by anti-LKAAEAR1 protein which may detect the cancerous part of this section. Peripheral nerve tissue section is expected to be positively stained as the presence of LKAAEAR1 protein is detectable in CNS tissues.

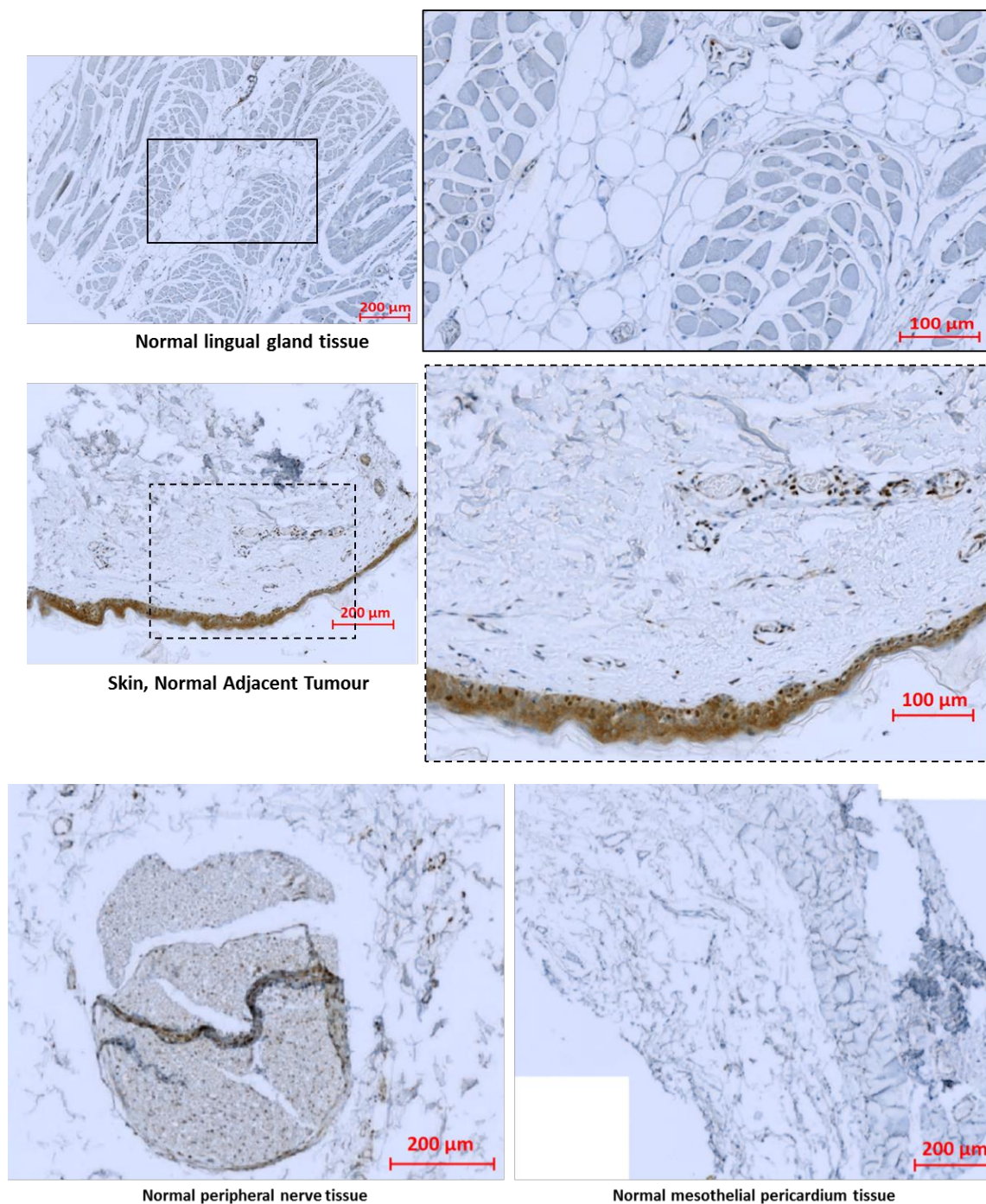


Figure 6.4. LKAAEAR1 staining on normal tissues (control) and normal adjacent to tumour tissues (NAT). These normal tissues are selected from TMA (FDA808-c2) slide to test the specificity of a polyclonal anti-LKAAEAR1 antibody from Abcam (ab108142). Images was captured by a ZEISS AXIO scan.Z1 digital scanner and each figure with different power magnification.

6.2.3.2. Staining pattern of multiple types of tumours of different stages.

The TMA contains many types of cancers and some biopsies are resected from different tissues in cancerous organ. All cancerous tissues show relatively strong staining with anti-LKAAEAR1 antibody with different staining pattern which could be dependent on the onset stage of cancer in each tissues. Figure 6.5, 6.6, 6.7 and 6.8 represent many examples of screened sections from multiple tumour types. For example, there was a clear difference in the stain intensity in prostate adenocarcinoma SII (weak) and prostate adenocarcinoma SIV (strong) as shown in Figure 6.5. This differential staining pattern is consistent with the previous study that has identified higher levels of LKAAEAR1 antigens on prostate cancer cells (Kamata et al., 2013). Another example on LKAAEAR1 differential staining pattern is shown in Figure 6.6. for two cases of breast cancer, intraductal carcinoma from two different patients, the first case from a patient at age 48 years diagnosed with early stage of breast cancer, (Tis N0 M0) which means cancer in situ (few abnormal cells that may develop in cancer in near future), the staining pattern for this tissue showed few cells were positive staining but in general the staining is at very weak signals. However, in the second case from 58 years old woman diagnosed with T3 N1 M0 which means more advanced cancer case has spread to lymph node, this case showed very strong staining pattern. This indicates that strong staining is clearly seen in more advanced cancer stages but the early stages of cancers show very weak staining as no many cancerous cells. Further analysis of LKAAEAR1 staining pattern showed stronger staining pattern in advanced cancer stages than early stages, for example, thyroid medullary carcinomas SI and thyroid papillary carcinomas SII and bladder carcinomas SI and SIII (Figure 6.7). Additionally, the comparison of staining was extended to show that few nuclei were stained in normal lung tissues although the quality of this biopsy was at poor level, however, the staining signals become stronger in lung tumours with SI and very strong in SII case (Figure 6.8). Finally, the summary of all cancer cases are presented in Table 6.2 with the staining pattern for each case. The level of staining was evaluated in all examined tissues and the assessment used a range of values (0 – 4) for each tissue. The values were assigned for staining levels as (0: negative, 1: very weak staining, 2: weak staining, 3: strong and 4: very strong staining).

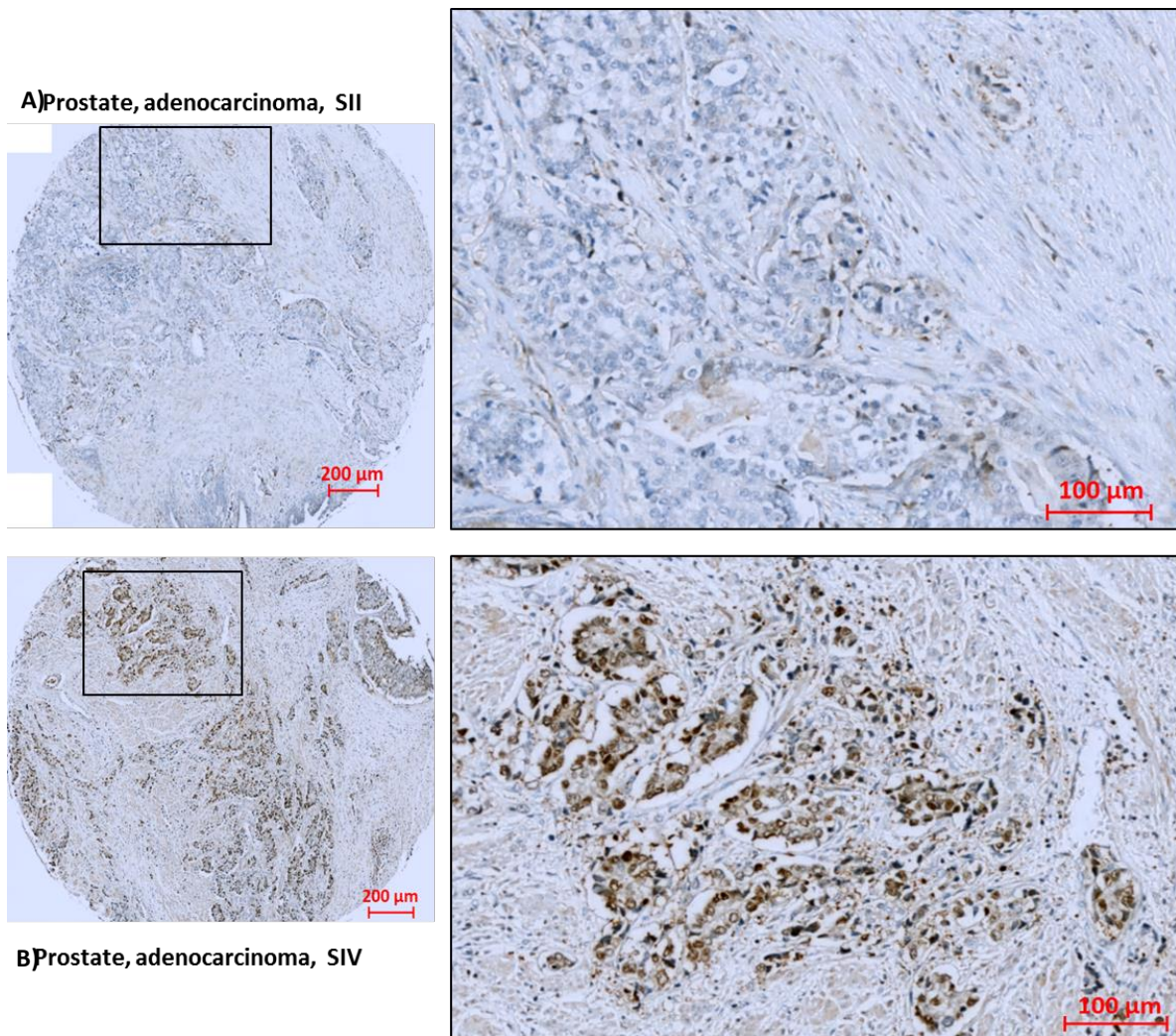
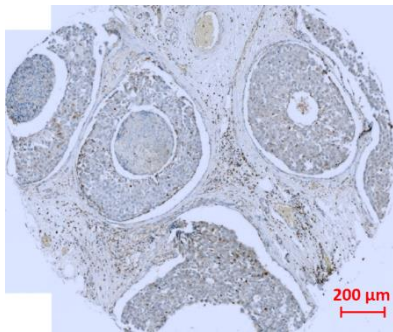
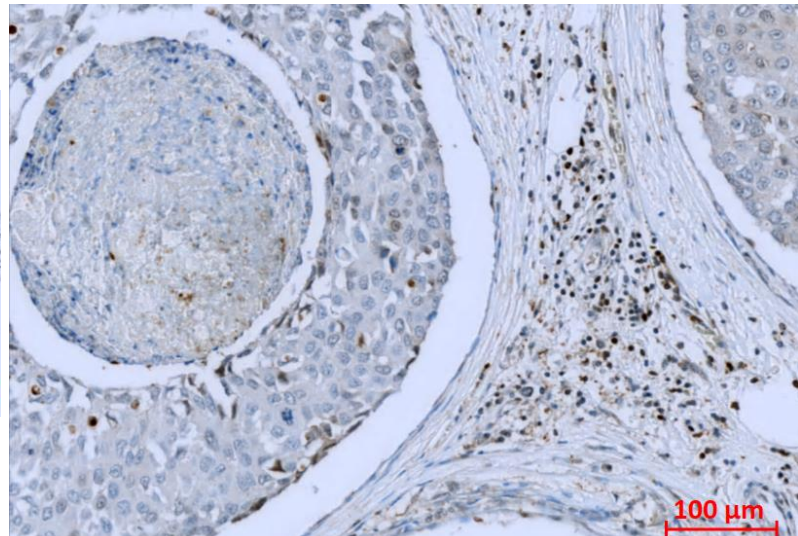


Figure 6.5. LKAAEAR protein staining in prostate tumours. (A) Prostate tumour diagnosed as adenocarcinoma SII, showing negative or very weak nuclei staining. (B) Prostate adenocarcinoma tumour in stage IV with a staining pattern is stronger and predominantly in nuclei. The tissues were scanned by Zeiss scanner and captured by Lite 2 software. Each figure is labelled with the pathology diagnosis and scaling bar is also included.

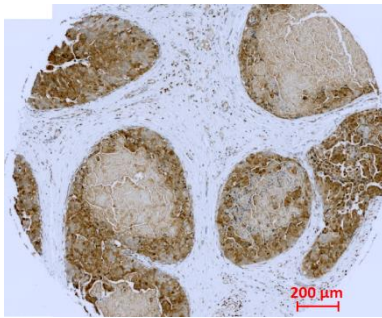
A)



Breast, carcinoma, S0



B)



Breast, carcinoma, SII

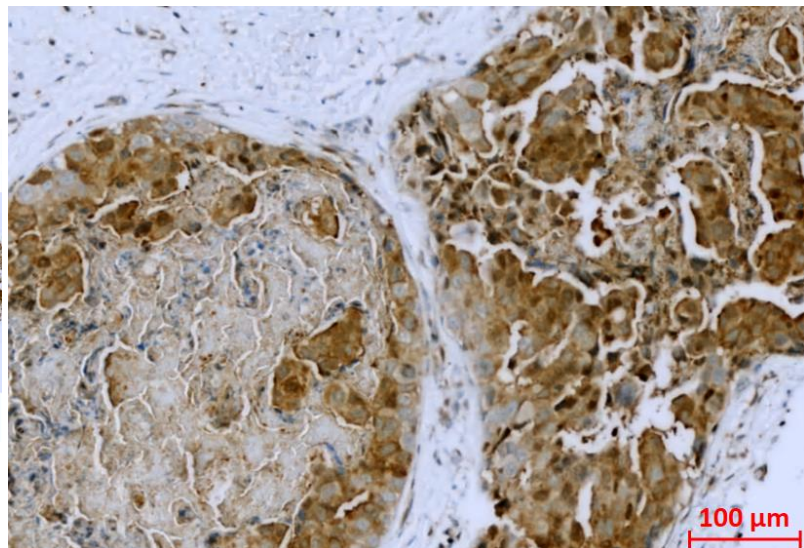
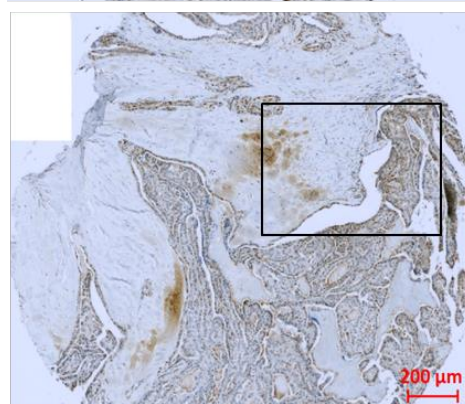
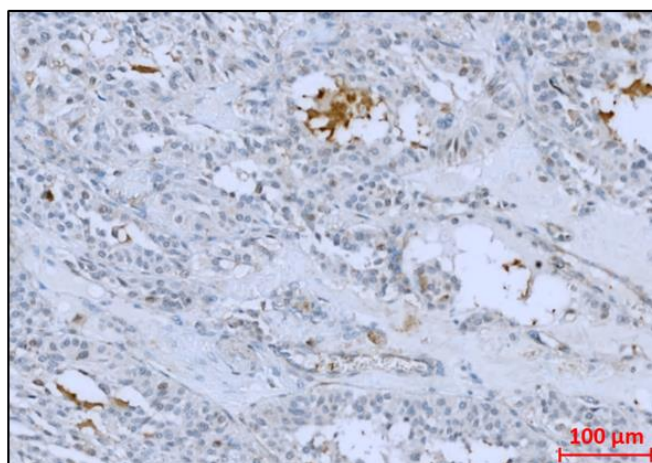
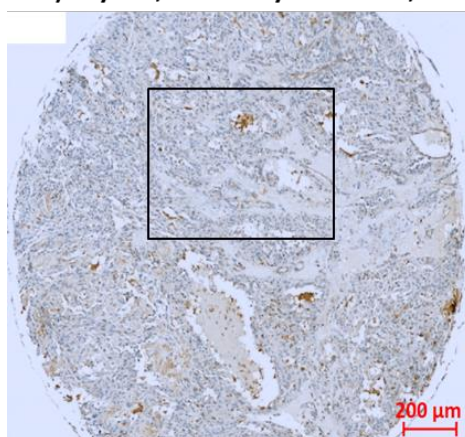
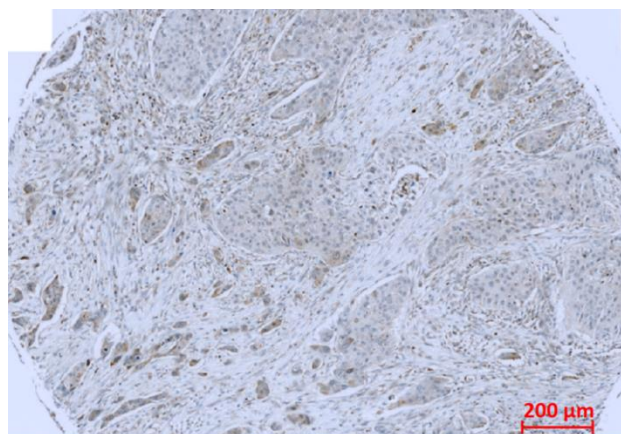
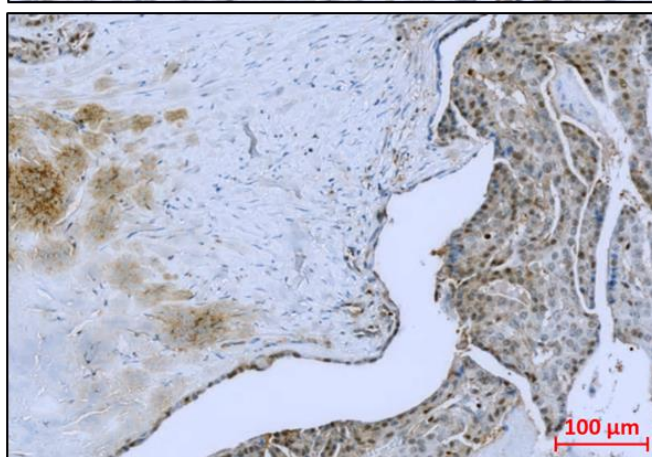


Figure 6.6. LKAEAR protein staining in breast tumours. (A) Section from breast tumour diagnosed as breast carcinoma S0, showing negative or very weak staining. (B) Section from breast tumour in stage II with a staining pattern is stronger in SII in comparison to early stage of tumour. The tissues were scanned by Zeiss scanner and captured by Lite 2 software. Each figure is labelled with the pathology diagnosis and scaling bar is also included.

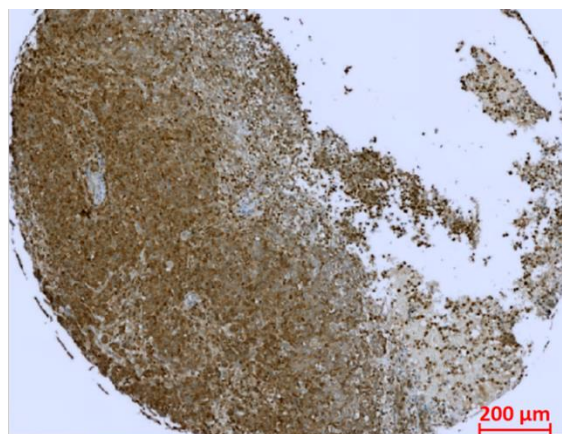
A) Thyroid, medullary carcinoma, SII



B) Thyroid, papillary carcinoma, SIII



C) Bladder, low grade malignant leiomyosarcoma, SI



D) Bladder, transitional cell carcinoma, SIII

Figure 6.7. LKAAEAR protein staining in thyroid and bladder tumours. (A) Thyroid tumour diagnosed as medullary carcinoma SII, showing weak and patchy staining. (B) Thyroid papillary carcinoma in stage IV with a staining pattern is stronger and predominantly in nuclei. (C) Bladder carcinoma SI with weak signals in different parts of the tissue and (D) Bladder carcinoma SIII with very strong staining pattern. The tissues were scanned by Zeiss scanner and captured by Lite 2 software. Each figure is labelled with the pathology diagnosis and scaling bar is also included.

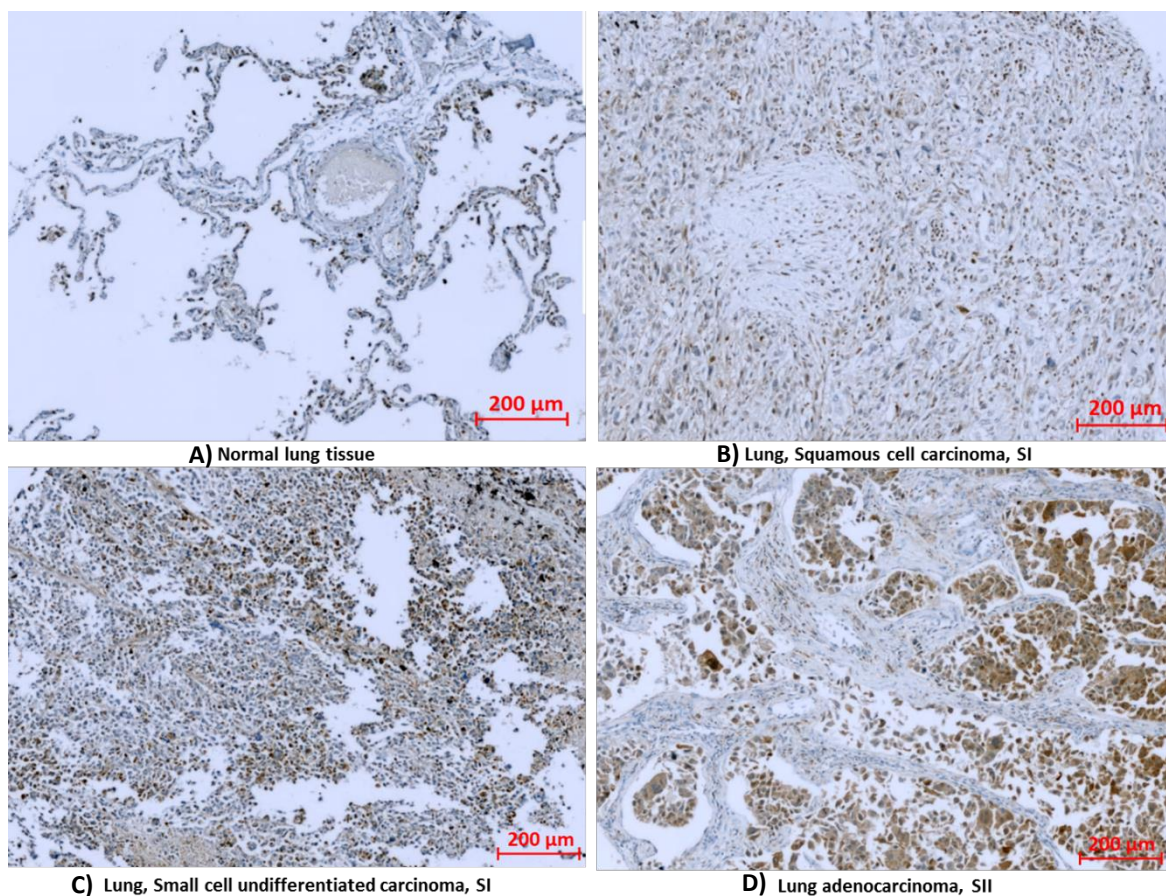


Figure 6.8. LKAAEAR protein staining in normal and cancerous lung tissues. (A) Normal lung tissue showed a very weak staining in few nuclei. (B) Lung squamous carcinoma (SI) with a weak staining pattern which is predominantly staining large number of nuclei. (C) Lung undifferentiated carcinoma SI with moderate stain signals in nuclei and faint in cytoplasm (D) Lung adenocarcinoma SII with very strong staining pattern. The tissues were scanned by Zeiss scanner and captured by Lite 2 software. Each figure is labelled with the pathology diagnosis and scaling bar is also included.

Table 6.2. Summary of LKAAEAR1 staining assessment of tumours in the TMA.

Organ/Anatomic Site	Pathology diagnosis	TNM	Stage	Type	Score*
Lingual gland	Normal lingual gland tissue	-	-	normal	0
Skin	Cancer adjacent normal skin tissue	-	-	NAT	0 and 3
Skin	Normal skin tissue of foot	-	-	normal	0 and 3
Skin	Normal skin tissue of head	-	-	normal	0 and 3
Nerve	Normal peripheral nerve tissue	-	-	normal	2
Nerve	Normal peripheral nerve tissue	-	-	normal	2
Nerve	Normal peripheral nerve tissue	-	-	normal	1
Lung	Normal mesothelial tissue	-	-	normal	2
Cardiac pericardium	Normal mesothelial tissue	-	-	normal	0
Lung	Normal lung and mesothelium tissue	-	-	normal	1
Cerebrum	Glioblastoma	-	-	malignant	2
Cerebrum	Atypical meningioma	-	-	malignant	3
Cerebrum	Malignant ependymoma	-	-	malignant	3
Cerebrum	Oligodendroglioma	-	-	malignant	1
Ovary	Serous adenocarcinoma	T2N0M0	II	malignant	4
Ovary	Adenocarcinoma	T3N0M0	III	malignant	3
Pancreas	Islet cell tumor	-	-	malignant	2
Pancreas	Adenocarcinoma	T3N0M0	II	malignant	2
Testis	Seminoma	T1N0M0	I	malignant	1
Testis	Embryonal carcinoma	T2N0M0	I	malignant	1
Thyroid	Medullary carcinoma	T3N0M0	II	malignant	2
Thyroid	Papillary carcinoma	T4N0M0	III	malignant	3
Breast	Invasive ductal carcinoma	T2N1M0	IIA	malignant	4
Breast	Intraductal carcinoma	TisN0M0	0	malignant	1
Breast	Invasive ductal carcinoma	T2N1M0	IIB	malignant	2
Spleen	Diffuse B-cell lymphoma	-	-	malignant	4
Lung	Small cell undifferentiated carcinoma	T2N0M0	I	malignant	2
Lung	Squamous cell carcinoma	T2N0M0	I	malignant	1
Lung	Adenocarcinoma	T2N0M0	IB	malignant	3
Esophagus	Neuroendocrine carcinoma	T2N1M0	IIA	malignant	4
Esophagus	Adenocarcinoma	T3N0M0	IIA	malignant	1 to 2
Stomach	Signet-ring cell carcinoma	T2N1M0	II	malignant	1
Intestine	Adenocarcinoma	T4N0M0	II	malignant	3
Intestine	Stromal sarcoma	T2N0M0	IIB	malignant	4
Colon	Adenocarcinoma	T4N0M0	IIB	malignant	2 to 3
Colon	Interstitialoma	T2N0M0	IIB	malignant	1
Rectum	Adenocarcinoma	T2N0M0	I	malignant	3
Rectum	Moderate malignant interstitialoma	T2N0M0	IIB	malignant	3 to 4
Liver	Hepatocellular carcinoma	T3N0M0	IIIA	malignant	3
Liver	Hepatoblastoma	-	-	malignant	1
Kidney	Clear cell carcinoma	T2N0M0	II	malignant	1
Prostate	Adenocarcinoma	T4N1M1c	IV	malignant	3
Prostate	Adenocarcinoma	T2N0M0	II	malignant	1
Uterus	Leiomyoma	T2N0M0	IB	malignant	1
Uterus	Adenocarcinoma	T1bN0M0	IB	malignant	2
Uterus	Clear cell carcinoma of endometrium	T2bN0M0	IIB	malignant	2
Uterine cervix	Squamous cell carcinoma	T1bN0M0	IB	malignant	2
Uterine cervix	Squamous cell carcinoma	T2N0M0	II	malignant	4
Striated muscle	Embryonal rhabdomyosarcom of left leg	T1aN0M0	IA	malignant	1
Rectum	Malignant melanoma	T4N0M0	IIB	malignant	2 to 3
Skin	Basal cell carcinoma of left face	T2N0M0	II	malignant	1
Skin	Squamous cell carcinoma of chest wall	T3N0M0	II	malignant	2

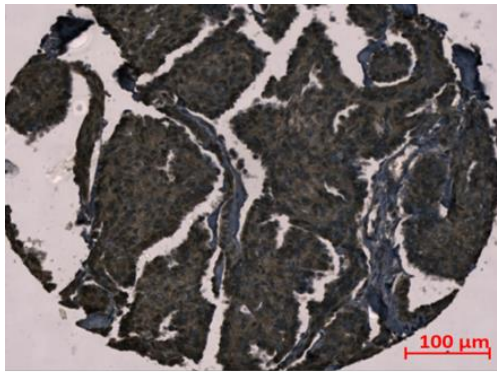
Back	Neurofibroma	-	-	malignant	1
Retroperitoneum	Neuroblastoma	T2bN0M0	IIIB	malignant	2
Abdominal cavity	Malignant mesothelioma	T2N0M0	IIB	malignant	3
Mediastinum	Diffuse B-cell lymphoma of lymph node	-	-	malignant	3
Lymph node	Diffuse B cell lymphoma of right thigh	-	-	malignant	4
Pelvic cavity	Anaplastic large cell lymphoma	-	-	malignant	4
Bladder	Transitional cell carcinoma	T3aN0M0	III	malignant	2
Bladder	Low grade malignant leiomyosarcoma	T2bN0M0	IB G1	malignant	4
Bone	Osteosarcoma of right femur	T2N0M0	IIIB	malignant	3
Retroperitoneum	Spindle cell rhabdomyosarcoma	T2N0M0	IIB	malignant	1
Smooth muscle	Moderate malignant leiomyosarcoma of left buttock	T2bN0M0	IIIB	malignant	1

(*) 0: negative, 1: very weak staining, 2: weak staining, 3: strong and 4: very strong staining, this staining pattern was categorised in collaboration with Dr. S. J. Sammut and Dr. J. Jezkova.

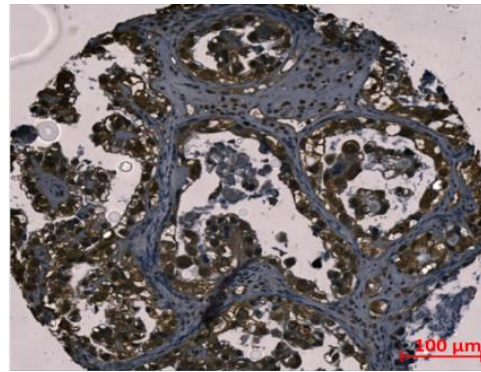
6.2.4. Distribution of LKAAEAR1 protein in epithelial ovarian cancer.

It has been recently published that *LKAAEAR1* transcript levels were detected in normal ovarian tissues at low rates, however, its levels were observed to markedly upregulated in different types of epithelial ovarian tumours (Ding et al., 2017). IHC analysis was conducted in this study to determine the LKAAEAR1 protein in normal and cancerous ovarian sections. Normal ovary section were also examined from different slide but with the same condition for experiment procedures and antibody concentration. A polyclonal rabbit antibody from Abcam (ab108142) with concentration 1/500 was used. Furthermore, benign ovarian tissues from serous cystadenomas and mucinous cystadenomas cases were used as negative controls from the same slide of diagnosed ovarian cancer sections. The results showed that the different types of epithelial ovarian tumours produce strong staining for LKAAEAR1 protein in comparison to the negative controls (Figure 6.9). This stain was also compared to normal epithelial ovarian tissue from different slide (Figure 6.9 C). The findings here suggest that LKAAEAR1 protein may have roles in epithelial ovarian cancer progression and development. Furthermore, LKAAEAR1 protein is also suggested to be a useful diagnostic marker for ovarian tumours.

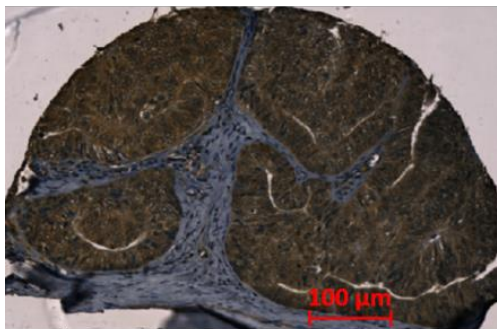
A)



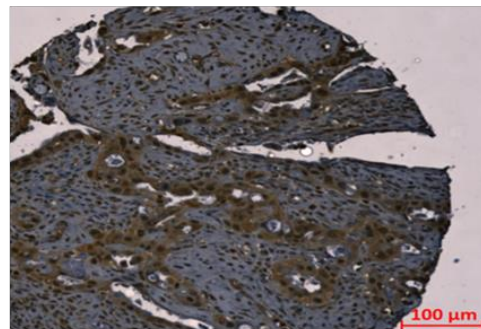
Serous papillary carcinoma (T3)



Clear cell carcinoma (T1)

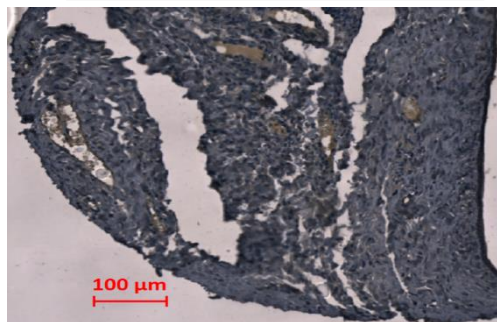


Endometrioid adenocarcinoma (T1)

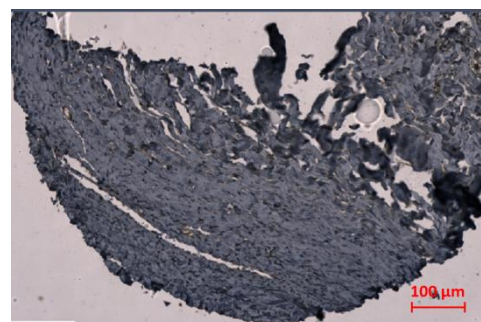


Mucinous adenocarcinoma (T3)

B)



Serous cystadenomas ovary



Mucinous cystadenomas, Ovary

C) Normal ovary epithelium

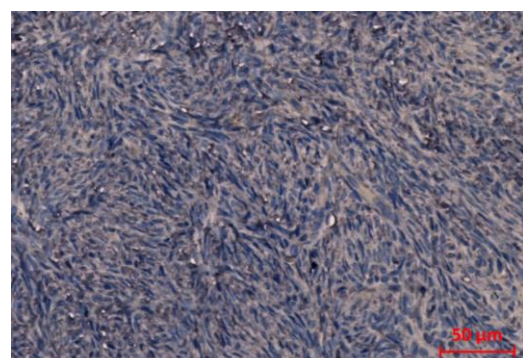
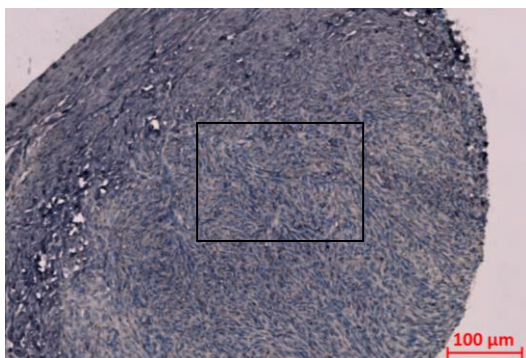


Figure 6.9. IHC staining shows LKAAEAR1 protein distribution in normal and different types of epithelial ovarian cancers samples with polyclonal anti-LKAAEAR1 antibody (ab108142). (A) The results showed a strong staining pattern in four different types of epithelial ovarian tumours (T refers to tumour stage). (B) Two cases of benign ovarian tumour from the same slide. (C) Normal ovarian epithelium from different slide showing a faint staining for LKAAEAR1 protein.

6.2.5. Staining pattern of LKAAEAR1 protein colorectal carcinoma progression microarray.

IHC staining was carried out on tissue microarrays containing multiple replicated samples from colorectal cancer (CRC) patients at different stages of disease progression. The array included 7 different progression group (series) (each case with replicates). Example of staining pattern is illustrated in Figure 6.10. LKAAEAR1 concentration was observed to be increased in adenocarcinomas and cancer stages and at lower levels in metastatic cases. Table 6.3 showed the estimated strength of staining for each case. The average for each CRC case was also demonstrated on bar chart in Figure 6.11.

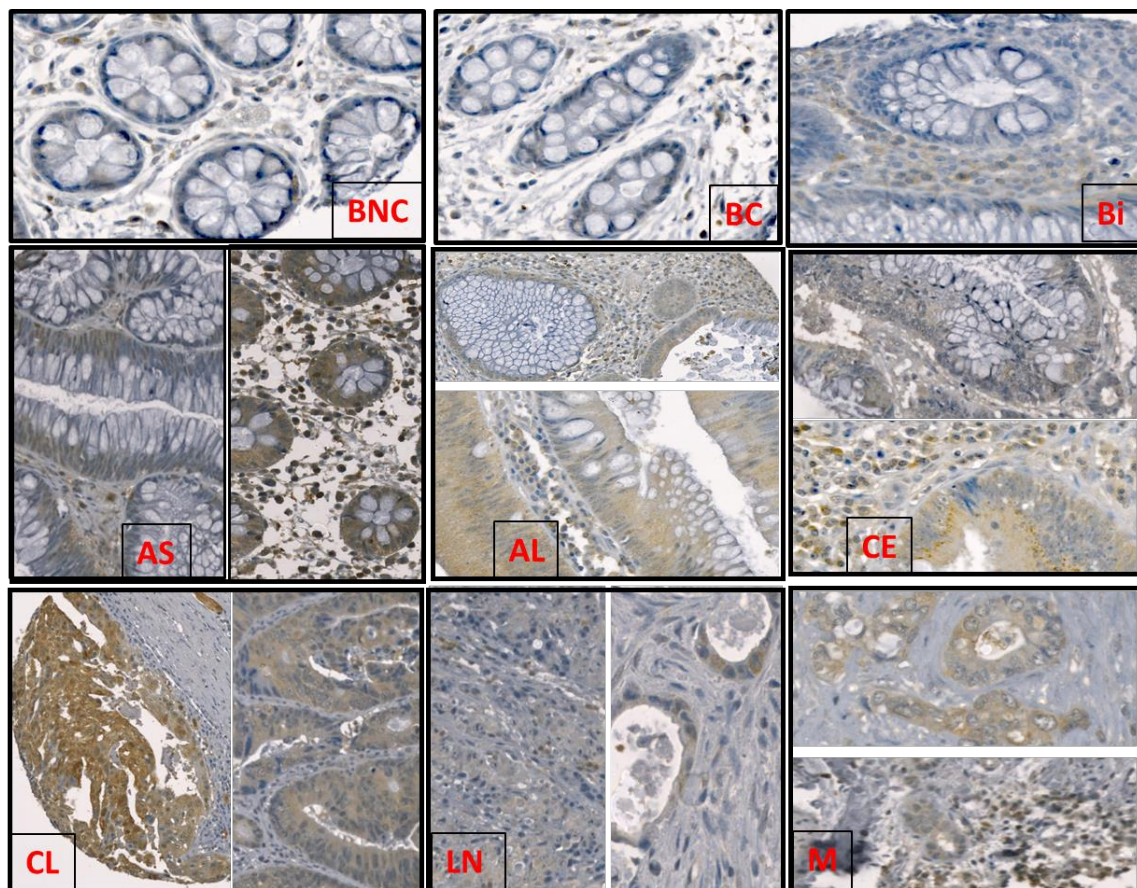


Figure 6.10. IHC staining of tissue microarray for colorectal cancer tissues (CRC) progression cancer samples by anti-LKAAEAR1 antibody (ab108142). Results show the staining pattern at different stages of CRC development. BNC: non-neoplasm colonic mucosa (pre cancer stage), BC: non-neoplasm colonic mucosa (cancer case), Bi: inflamed non-neoplastic mucosa (ulcerative colitis), AS: adenoma (< 2 cm), AL: adenoma (> 2 cm), CE: invasive adenocarcinoma (T1 or T2), CL: invasive adenocarcinoma (T3 or T4), LN: colorectal adenocarcinoma metastatic to lymph nodes and M: colorectal adenocarcinoma metastatic to distant sites. The figures magnification is 10X.

To evaluate the presence and distribution of LKAAEAR1 protein in CRC samples at different stages of developing, values (from 0-4) were assigned to each section depending on the levels of staining. The non-detectable stain was given number 0, very weak staining which also include the edges staining 1 and the strong staining was expressed by number 4. The distribution and production of LKAAEAR1 was constantly stronger in large adenomas and cancer late stages, although different levels of LKAAEAR1 were detected among the remaining samples but at lower levels as shown in Figure 6.11. A list containing the information of all samples/ patients are included in (Appendix B-5).

Table 6.3. Showing the estimated staining factor for each group cases in CRC progression microarray

	Group/series							
CRC	G1	G2	G3	G4	G5	G6	G7	Average
BNC	1	1	2	1	2	2	1	1.4
BC	2	2	2	1	1	3	4	2.1
Bi	2	3	3	2	1	3	3	2.4
AS	2	4	4	3	2	2	3	2.9
AL	3	3	3	4	4	4	3	3.4
CE	4	4	3	4	3	4	2.3	3.5
CL	4	4	4	3	3	3	3	3.4
LN	1	2	0	4	2	1	2	1.7
M	4	2	2	1	3	3	3	2.6

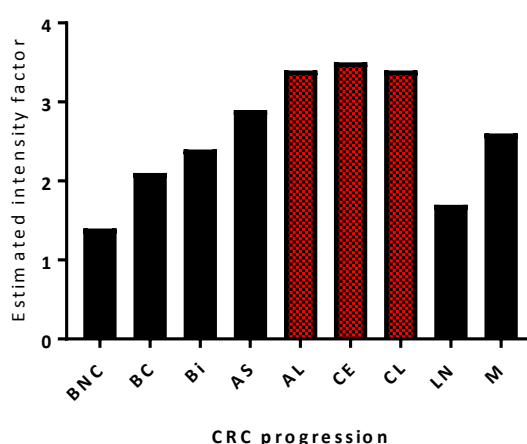


Figure 6.11. Relative levels of LKAAEAR1 staining detected in CRC progression microarray.

Values were assigned as follow: 0 = no stain, 1= very weak staining, 2 weak staining, 3= strong staining and 4 = very strong staining. BNC: non-neoplasm colonic mucosa (pre cancer stage), BC: non-neoplasm colonic mucosa (cancer case), Bi: inflamed non-neoplastic mucosa (ulcerative colitis), AS: adenoma (< 2 cm), AL: adenoma (> 2 cm), CE: invasive adenocarcinoma (T1 or T2), CL: invasive adenocarcinoma (T3 or T4), LN: colorectal adenocarcinoma metastatic to lymph nodes and M: colorectal adenocarcinoma metastatic to distant sites.

6.3. Discussion.

The improvement of personalised therapeutic strategies in recent years has been facilitated by the identification of novel biomarkers. These biomarkers are required for accurate stratification of clinical samples during diagnostic and therapeutic approaches. The work presented here outlines the initial steps towards a clinical use at the antigenic level for a novel and poorly characterised germline gene. The *LKAAEAR1* gene is defined as promising CT gene at the transcription level and these results were validated at the level of protein in Chapter 5. The aim of this chapter was to inspect the presence of LKAAEAR1 protein on clinical samples that were obtained from normal tissues and individuals diagnosed with cancer.

CTAs are tumour antigens that have potential targets for cancer immunotherapy and vaccination, because they are produced on several types of malignant cells but not presented on healthy cells, except of the immune-privilege sites in the body including testis, ovary and placenta (Whitehurst, 2014). IHC analysis of LKAAEAR1 on tissue section from testis showed positive staining in the periphery (basal compartment) of seminiferous tubules (Figure 6.1). The classification of LKAAEAR1 as a CTA would require its presence in germinal cells of testis tissues which is a positive control sample for any CTA. Although its function is unknown, it was predicted that LKAAEAR1 protein might have a fundamental role in spermatogenesis and this is supported by the evidence of strong expression in testes. Analysis on CNS tissues demonstrated positive staining that determined the protein to be in cerebral cortex, subcortical brain and peripheral nerve tissues. This is expected as LKAAEAR1 is classified as a CT-CNS restricted protein (Feichtinger et al., 2012). Figures 6.2 and 6.3 presented many examples of normal tissue sections of negative staining pattern such as breast, stomach, heart muscle, fallopian tube, and salivary gland tissues. A very weak staining was observed in normal ovarian tissues, however, this is consistent with previous studies that identified a weak expression of *LKAAEAR1* transcripts in ovarian samples (Feichtinger et al., 2012). Bladder samples show negative staining at the central part of the section, however the edge of sample exhibited weak staining and this may related to the folded-edge or atrophy or dried edge of this sample. The crypts of colon and uterine tissues showed weak staining signal, which is not expected, as the surrounding tissues appeared to be negative of LKAAEAR1 protein. In such cases, CTAs could be of potential targets by immunotherapeutic approach if it is present in normal tissues at low levels in comparison to a far higher levels in cancerous cells. However, some normal tissues showed faint staining with anti-LKAAEAR1 antibody as

can be seen in kidney, liver, small intestine and thymus tissues. The presence of LKAAEAR1 in these tissues is not expected because the analysis of this protein on tissue lysates in Chapter 5 did not reveal any signal, in addition, the previous studies on mRNA analysis also did not detect transcripts in these tissues. Alternatively, gene expression in some normal tissues might be at very low levels that are not detected by the used methods in the previous study. So, the gene may not have CTA characteristics. Another reason behind this positive staining in some normal tissues is relating to a lack specificity of the polyclonal antibody. This is unsurprising as differences of antibody batch specificity was experienced during the lab work. Some batches of this antibody were discarded as they produced strong staining in all normal tissues with low quality and background showing that these batches may cross react with other proteins. The antibody validation and the related issues have been discussed in (Pozner-Moulis et al., 2007; Bordeaux et al., 2010; Baker, 2015). To address this, monoclonal antibody must be developed/used to give a greater clarity.

Kamata and his colleague provided evidence that LKAAEAR1 protein is highly produced in patient with prostate cancer (Kamata et al., 2013). Furthermore, another study reported that LKAAEAR1 transcripts were differentially expressed at higher levels in ovarian epithelium tumours in comparison to adjacent normal tissues (Ding et al., 2017). Given that, the analysis was conducted on tumour microarray tissues to investigate the staining pattern of LKAAEAR1 protein on cancers. The results showed a significant difference for LKAAEAR1 staining pattern in many cancers. For example, there was a clear difference of staining on prostate adenocarcinoma stage II and stage IV, weak staining was observed in early tumour stages compared to strong staining in late stages of prostate cancer. This observation was confirmed on other tumours with different stages such as, breast, thyroid, bladder and lung tumours. These results suggested that LKAAEAR1 may play roles in cancer progression and development. The analyses were expanded on different types of ovarian epithelial tumours to be compared with normal ovarian tissues. Ovarian epithelial tumours showed a clear difference when the staining was very strong in many types of cancerous tissues but was weak in normal ovarian tissues. This is consistent with Ding and co-workers findings of mRNA differential expression between normal and cancer ovarian tissues. Additionally, IHC analysis was carried out on colon progression microarray showed that LKAAEAR1 protein is not only produced at different stages of colon cancer but also an interesting higher staining signals for this protein were detected at the early stages and specifically to colon adenocarcinomas.

These findings suggested that LKAAEAR1 protein may play important roles in cancer development. Up to this step of analyses, this protein may have potentials to be used as a diagnostic marker for many types of cancer and also might have ability to differentiate early stages from late stages of cancer to help in treatment planning.

To summarize, IHC analysis supported the findings that LKAAEAR1 is a CTA protein. LKAAEAR1 protein is produced in testis and NCS tissues from normal tissues, however, its presence is highly increased in several tumour tissues. Given that, LKAAEAR1 has potential to be a good cancer biomarker and immunotherapeutic target. Finally, our conclusion for these results is that, the presence of weak staining in some normal tissues is not expected and using of more specific antibody may show more reliable results. LKAAEAR1 is still of high interesting CTA that is potentially promising for cancer diagnosis and therapeutics approaches. The function of this protein is not clear, however, the strong staining in cancer tissues may suggest its fundamental roles in oncogenesis.

Chapter 7:

Final discussion and future directions

7. Final discussion and future directions

Cancer is a leading cause of mortality around the world, making studies that gain insights into oncogenesis increasingly important. The capability of tumour cells to self-renew and undergo phenotypic changes has led to the idea that some cancer cells have common features with stem cells (McFarlane et al., 2014; Nassar & Blanpain, 2016). In this context, it has been suggested that the transformation of cancer cells might be driven by a soma-to-germline transition (Feichtinger et al., 2014; McFarlane et al., 2015; Nassar & Blanpain, 2016; Nielsen & Gjerstorff, 2016). This is supported by the observation that a large group of germline genes critical for tumour development were activated during oncogenesis in *Drosophila melanogaster* (Janic et al., 2010; Fagegaltier et al., 2016; Sumiyoshi et al., 2016). Some studies have proposed that not all cancer cells have the ability to self-renew, or have limitless proliferation and invasive potentials, though a small population of cancer cells do have this capacity and are referred to as cancer stem-like cells (CSCs) (Moore & Lyle, 2011; Gedye et al., 2016; Aponte & Caicedo, 2017). The CSC model is supported by the existence of heterogeneity and plasticity within the tumour bulk. Taken together, gene expression analysis indicated that germline genes were activated in human tumour cells in a similar pattern to stem cells, suggesting a possible functional requirement (Feichtinger et al., 2014). Importantly, CSCs are promising targets for new human cancer therapies development of drugs and the early diagnosis of cancer. Therefore, more research is needed into the function and role of CSCs.

Cancer-testis genes are a major group of germline genes that encode CTAs, whose expression is restricted to germ cells and a wide range of tumours. This makes CTAs highly interesting as diagnostic biomarkers and immunotherapy targets. The known functional roles of these proteins along with their immunogenicity increase their therapeutic potential. There is also growing evidence for a functional role of CT genes in cancer biology. New emerging evidence indicates that meiosis-specific genes (a subclass of CT genes) have an important influence on cancer progression and development. For example, Greenberg and colleagues found that two meiosis specific genes, *MND1* and *HOP2*, normally play a role in meiotic recombination, though they also promote telomere length maintenance via the oncological alternative elongation of telomere (ALT) mechanism (Cho et al., 2014; McFarlane & Wakeman, 2017).

Activation of germline genes in a wide range of cancers is suggested to be a result of genome and epigenetic evolution in response to the immediate requirements and pressures of tumour cells, which supports the transformation of tumour cells to a germ-like state (Flavahan et al., 2017; Venkatesan et al., 2017; McGranahan & Swanton, 2017). Two novel germline genes, *TDRD12* and *LKAAEAR1*, have been shown herein to be potential immunotherapeutic targets for cancer treatment, identified as CTA gene candidates that may encode CSC-specific activity (Feichtinger et al., 2012; Almatrafi et al., 2014). Our results demonstrated that *TDRD12* and *LKAAEAR1* gene expression correlates to stemness features, serving as potential stemness markers for CSCs, and these genes might be required for pluripotency. The findings hint at the possible functions of these genes in the tumour cell transformation to the germ-like state and oncogenesis maintenance.

Previous studies have shown that *Tdrd12* is essential for germ cell development and genomic stability in mice (Chuma et al., 2006; Shoji et al., 2009; Yabuta et al., 2011; Pandey et al., 2013; Yang et al., 2016). It has been previously identified as playing a role in germ cell migration, proliferation and maintenance in the early stages of mice germ cell development (Dai et al., 2017), while mutants lacking *Tdrd12* resulted in male sterility and atrophied testes (Pandey et al., 2013). Our analysis showed that *TDRD12* enabled proliferation in human cancer stem-like cell line, NTERA2. Moreover, in *TDRD12* knockdown experiments, NTERA2 cells accumulate in the S phase, and a lower proportion of cells undergo mitosis, causing the cells to enter a quiescent-like state. This is the first study to report the influence of *TDRD12* on CSC proliferation, suggesting that *TDRD12* may be a potential tumour immunotherapy target. Aberrant expression of *TDRD12* was observed to inhibit CDKN1A and TP53 proteins in NTERA2 cells, which might cause the cell proliferation and overgrowth, highlighting a possible functional role of *TDRD12* in oncogenesis. Furthermore, the downregulation of *TDRD12* was observed to change the levels of stemness markers, particularly NANOG and SOX2. This indicated that *TDRD12* might also act as a transcriptional regulator for some stemness markers, supporting its possible relationship to stemness features. However, more research is needed to clarify the mechanism by which *TDRD12* influences these markers, for example, *TDRD12* cloning in a protein expressing system.

Computational databases and bioinformatics tools predicted that *TDRD12* may interact with other germline proteins. Studies in mice identified *Tdrd12* as essential in pi-RNA production to maintain the development of germ cells, and this was observed to be achieved through

TDRD12 protein interactions with the Piwi family and Exd1 proteins (Pandey et al., 2013; Yang et al., 2016). Furthermore, Exd1 had been reported to interact with Tdrd12 in mice (Yang et al., 2016). However, analysis in this study explore the correlation of *TDRD12* and *EXDL1* genes at transcriptional levels, therefore, the potential interactions at protein levels need to be confirmed. For this purpose, co-immunoprecipitation should be conducted using antibodies against TDRD12-interacting partners. Moreover, TDRD12 cloning into HaloTag system could be established for characterisation of protein interactions and analysing other possible potentials and functions.

In 2012, Feichtinger and co-workers first identified that *LKAAEAR1* was a CTA candidate using conventional RT-PCR (Feichtinger et al., 2012). Recently these results were verified by quantitative RT-PCR done by Kamata and co-workers (2013). The study presented herein further analysed the *LKAAEAR1* protein in normal and cancerous tissues, confirming that *LKAAEAR1* is a testis/CNS-restricted candidate CTA. However, a higher molecular weight was observed for the *LKAAEAR1* protein in all examined cancer lysates. This higher molecular protein was determined in nucleus extracts of cancer cells, suggesting that this protein might undergo posttranslational modification. Further analysis is required, such as co-immunoprecipitation with an anti-*LKAAEAR1* antibody followed by WB probing against ubiquitin, as it is expected to be covalently attached to this protein. The analyses were expanded from normal tissue lysates to clinical biopsies, confirming that the distribution of *LKAAEAR1* is restricted to testis and CNS normal tissues. However, these results need further investigation with a monoclonal antibody.

The activation of transposable elements (TEs) occurs during oncogenesis (Goodier, 2014; Chénais, 2015), though the influence of *LKAAEAR1* on TE expression has not been investigated. Our results showed differences in the levels of TE transcription following the depletion of *LKAAEAR1* mRNAs. A specific *HERV K gag* was found to be a common element downregulated in all examined cancer cells, indicating a possible interface with *LKAAEAR1*.

The presence of the *LKAAEAR1* protein in the basal compartment of seminiferous tubules suggested its fundamental role in spermatogenesis. Interestingly, majority of spermatocytes showed a strong nuclear staining of this protein. We also observed nuclear foci in testis germ cells, which were also clearly seen in the IF staining of the breast cancer cell line, MCF7. This phenomenon is a noteworthy observation that should be investigated further. To gain insight into the *LKAAEAR1* protein function in spermatogenesis process, immunogold labelling or

immunogold staining is recommended for further analysis of the LKAAEAR1 protein location during meiosis, to assess whether it associate with specific meiotic chromosome structures. Finally, the functional roles of many germline and CT genes remain poorly understood despite the interest in CTAs as targets for immunotherapy, some of which have already been targeted. Characterisation of functions for some genes, such as *LKAAEAR1*, should be a priority because it has appeared to be widely expressed in human cancer. Two novel putative CTAs (TDRD12 and LKAAEAR1) with immunogenic potentials have been investigated here. A different approach to examine the immunogenic potential, for instance, a T-cell response *in vitro* to targeted protein, including the analysis of proliferation and secretion of cytokines and the expression of cell surface markers, would allow us to understand if TDRD12 or LKAAEAR1 antigens are immunogenic and therefore can be potentially used in immunotherapy.

8. References

- Abbas, T. & Dutta, A. 2009. p21 in cancer: intricate networks and multiple activities. *Nature Reviews Cancer*. 9 (6). pp. 400.
- Adachi, K., Suemori, H., Yasuda, S., Nakatsuji, N. & Kawase, E. 2010. Role of SOX2 in maintaining pluripotency of human embryonic stem cells. *Genes to Cells*. 15 (5). pp. 455-470.
- Adair, S.J. & 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.
- Adams, S., Greeder, L., Reich, E., Shao, Y., Fosina, D., Hanson, N., Tassello, J., Singh, B., Spagnoli, G.C. & Demaria, S. 2011. Expression of cancer testis antigens in human BRCA-associated breast cancers: potential targets for immunoprevention? *Cancer immunology, immunotherapy*. 60 (7). pp. 999.
- Agliano, A., Calvo, A. & Box, C. 2017. The challenge of targeting cancer stem cells to halt metastasis. *Seminars in Cancer Biology*. Elsevier: .
- Ajani, J.A., Song, S., Hochster, H.S. & Steinberg, I.B. 2015. Cancer stem cells: the promise and the potential. *Seminars in oncology*. Elsevier: pp. S3.
- Akers, S.N., Odunsi, K. & Karpf, A.R. 2010. Regulation of cancer germline antigen gene expression: implications for cancer immunotherapy. *Future oncology*. 6 (5). pp. 717-732.
- Al-Dhfyhan, A., Alhoshani, A. & Korashy, H.M. 2017. Aryl hydrocarbon receptor/cytochrome P450 1A1 pathway mediates breast cancer stem cells expansion through PTEN inhibition and β -Catenin and Akt activation. *Molecular cancer*. 16 (1). pp. 14.
- Alexander Przyborski, S., Smith, S. & Wood, A. 2003. Transcriptional profiling of neuronal differentiation by human embryonal carcinoma stem cells in vitro. *Stem cells*. 21 (4). pp. 459-471.
- Al-Hajj, M., Wicha, M.S., Benito-Hernandez, A., Morrison, S.J. & Clarke, M.F. 2003. Prospective identification of tumorigenic breast cancer cells. *Proceedings of the National Academy of Sciences of the United States of America*. 100 (7). pp. 3983-3988.
- Ali, F., Tareh, S., Al-Nuzaily, M., Mok, P.L., Ismail, A. & Ahmad, S. 2016. Stem cells differentiation and probing their therapeutic applications in hematological disorders: a critical review. *European review for medical and pharmacological sciences*. 20 (20). pp. 4390-4400.
- Alison, M.R., Poulosom, R., Forbes, S. & Wright, N.A. 2002. An introduction to stem cells. *The Journal of pathology*. 197 (4). pp. 419-423.
- Almatrafi, A., Feichtinger, J., Vernon, E.G., Escobar, N.G., Wakeman, J.A., Larcombe, L.D. & McFarlane, R.J. 2014. Identification of a class of human cancer germline genes with transcriptional silencing refractory to the hypomethylating drug 5-aza-2'-deoxycytidine. *Oncoscience*. 1 (11). pp. 745-750.

- Almeida, L.G., Sakabe, N.J., Deoliveira, A.R., Silva, M.C.C., Mundstein, A.S., Cohen, T., Chen, Y., Chua, R., Gurung, S. & Gnjjatic, S. 2008. CTdatabase: a knowledge-base of high-throughput and curated data on cancer-testis antigens. *Nucleic acids research*. 37 (suppl_1). pp. D816-D819.
- Alves, G., Tatro, A. & Fanning, T. 1996. Differential methylation of human LINE-1 retrotransposons in malignant cells. *Gene*. 176 (1). pp. 39-44.
- Anand, P., Kunnumakara, A.B., Sundaram, C., Harikumar, K.B., Tharakan, S.T., Lai, O.S., Sung, B. & Aggarwal, B.B. 2008. Cancer is a preventable disease that requires major lifestyle changes. *Pharmaceutical research*. 25 (9). pp. 2097-2116.
- Andrews, P.W., Nudelman, E., Hakomori, S. & Fenderson, B.A. 1990. Different patterns of glycolipid antigens are expressed following differentiation of TERA-2 human embryonal carcinoma cells induced by retinoic acid, hexamethylene bisacetamide (HMBA) or bromodeoxyuridine (BUdR). *Differentiation*. 43 (2). pp. 131-138.
- Andrews, P.W., Damjanov, I., Simon, D., Banting, G.S., Carlin, C., Dracopoli, N.C. & Fogh, J. 1984. Pluripotent embryonal carcinoma clones derived from the human teratocarcinoma cell line Tera-2. Differentiation in vivo and in vitro. *Laboratory investigation; a journal of technical methods and pathology*. 50 (2). pp. 147-162.
- Aponte, P.M. & Caicedo, A. 2017. Stemness in Cancer: Stem Cells, Cancer Stem Cells, and Their Microenvironment. *Stem cells international*. 2017.
- Aravin, A.A., Hannon, G.J. & Brennecke, J. 2007. The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. *Science (New York, N.Y.)*. 318 (5851). pp. 761-764.
- Arnoult, N. & Karlseder, J. 2014. ALT telomeres borrow from meiosis to get moving. *Cell*. 159 (1). pp. 11-12.
- Assanga, I. & Lujan, L. 2013. Cell growth curves for different cell lines and their relationship with biological activities. *International Journal of Biotechnology and Molecular Biology Research*. 4 (4). pp. 60-70.
- Babaie, Y., Herwig, R., Greber, B., Brink, T.C., Wruck, W., Groth, D., Lehrach, H., Burdon, T. & Adjaye, J. 2007. Analysis of Oct4-dependent transcriptional networks regulating self-renewal and pluripotency in human embryonic stem cells. *Stem cells*. 25 (2). pp. 500-510.
- Bai, Q., Assou, S., Haouzi, D., Ramirez, J., Monzo, C., Becker, F., Gerbal-Chaloin, S., Hamamah, S. & De Vos, J. 2012. Dissecting the first transcriptional divergence during human embryonic development. *Stem Cell Reviews and Reports*. 8 (1). pp. 150-162.
- Baker, M. 2015. Blame it on the antibodies. *Nature*. 521 (7552). pp. 274.
- Bao, S., Wu, Q., McLendon, R.E., Hao, Y., Shi, Q., Hjelmeland, A.B., Dewhirst, M.W., Bigner, D.D. & Rich, J.N. 2006. Glioma stem cells promote radioresistance by preferential activation of the DNA damage response. *Nature*. 444 (7120). pp. 756.
- Barrow, C., Browning, J., MacGregor, D., Davis, I.D., Sturrock, S., Jungbluth, A.A. & Cebon, J. 2006. Tumor antigen expression in melanoma varies according to antigen and stage. *Clinical cancer*

research : an official journal of the American Association for Cancer Research. 12 (3 Pt 1). pp. 764-771.

Baylin, S.B. 2005. DNA methylation and gene silencing in cancer. *Nature clinical practice Oncology.* 2 pp. S4-S11.

Bizzarri, M., Giuliani, A., Cucina, A., D'Anselmi, F., Soto, A. & Sonnenschein, C. 2011. Fractal analysis in a systems biology approach to cancer. *Seminars in cancer biology.* Elsevier: pp. 175.

Bjerkvig, R., Tysnes, B.B., Aboody, K.S., Najbauer, J. & Terzis, A. 2005. Opinion: the origin of the cancer stem cell: current controversies and new insights. *Nature reviews.Cancer.* 5 (11). pp. 899.

Boesch, M., Zeimet, A.G., Reimer, D., Schmidt, S., Gastl, G., Parson, W., Spoeck, F., Hatina, J., Wolf, D. & Sopper, S. 2014. The side population of ovarian cancer cells defines a heterogeneous compartment exhibiting stem cell characteristics. *Oncotarget.* 5 (16). pp. 7027-7039.

Bolcun-Filas, E., Speed, R., Taggart, M., Grey, C., de Massy, B., Benavente, R. & Cooke, H.J. 2009. Mutation of the mouse Syce1 gene disrupts synapsis and suggests a link between synaptonemal complex structural components and DNA repair. *PLoS genetics.* 5 (2). pp. e1000393.

Bonnet, D. & Dick, J.E. 1997. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nature medicine.* 3 (7). pp. 730-737.

Bordeaux, J., Welsh, A., Agarwal, S., Killiam, E., Baquero, M., Hanna, J., Anagnostou, V. & Rimm, D. 2010. Antibody validation. *BioTechniques.* 48 (3). pp. 197-209.

Brasseur, F., Rimoldi, D., Liénard, D., Lethé, B., Carrel, S., Arienti, F., Suter, L., Vanwijck, R., Boursol, A. & Humblet, Y. 1995. Expression of MAGE genes in primary and metastatic cutaneous melanoma. *International journal of cancer.* 63 (3). pp. 375-380.

Bronner, C., Achour, M., Arima, Y., Chataigneau, T., Saya, H. & Schini-Kerth, V.B. 2007. The UHRF family: oncogenes that are drugable targets for cancer therapy in the near future? *Pharmacology & therapeutics.* 115 (3). pp. 419-434.

Brooks, M.D., Burness, M.L. & Wicha, M.S. 2015. Therapeutic implications of cellular heterogeneity and plasticity in breast cancer. *Cell stem cell.* 17 (3). pp. 260-271.

Brown, J.P., Wei, W. & Sedivy, J.M. 1997. Bypass of senescence after disruption of p21CIP1/WAF1 gene in normal diploid human fibroblasts. *Science (New York, N.Y.).* 277 (5327). pp. 831-834.

Brugarolas, J., Chandrasekaran, C., Gordon, J.I., Beach, D., Jacks, T. & Hannon, G.J. 1995. Radiation-induced cell cycle arrest compromised by p21 deficiency. *Nature.* 377 (6549). pp. 552.

Burger, H., Nooter, K., Boersma, A., Kortland, C. & Stoter, G. 1998. Expression of p53, Bcl-2 and Bax in cisplatin-induced apoptosis in testicular germ cell tumour cell lines. *British journal of cancer.* 77 (10). pp. 1562.

Burstein, H.J., Mangu, P.B., Somerfield, M.R., Schrag, D., Samson, D., Holt, L., Zelman, D. & Ajani, J.A. 2011. American Society of Clinical Oncology clinical practice guideline update on the use of

- chemotherapy sensitivity and resistance assays. *Journal of Clinical Oncology*. 29 (24). pp. 3328-3330.
- Caballero, O.L. & Chen, Y. 2009. Cancer/testis (CT) antigens: potential targets for immunotherapy. *Cancer science*. 100 (11). pp. 2014-2021.
- Cabrera, M.C., Hollingsworth, R.E. & Hurt, E.M. 2015. Cancer stem cell plasticity and tumor hierarchy. *World journal of stem cells*. 7 (1). pp. 27-36.
- Cahoon, C.K. & Hawley, R.S. 2016. Regulating the construction and demolition of the synaptonemal complex. *Nature structural & molecular biology*. 23 (5). pp. 369-377.
- Callaerts, P., Halder, G. & Gehring, W.J. 1997. PAX-6 in development and evolution. *Annual Review of Neuroscience*. 20 (1). pp. 483-532.
- Cao, Y. & Zhang, P. 2017. Regenerative medicine in cardiovascular diseases—an update. *European review for medical and pharmacological sciences*. 21 (6). pp. 1335-1340.
- Carr, S.M., Munro, S., Sagum, C.A., Fedorov, O., Bedford, M.T. & La Thangue, N.B. 2017. Tudor-domain protein PHF20L1 reads lysine methylated retinoblastoma tumour suppressor protein. *Cell death and differentiation*. 24 (12). pp. 2139.
- Cavalieri, E., Rogan, E. & Chakravarti, D. 2002. Initiation of cancer and other diseases by catechol ortho-quinones: a unifying mechanism. *Cellular and Molecular Life Sciences*. 59 (4). pp. 665-681.
- Cedar, H. & Bergman, Y. 2009. Linking DNA methylation and histone modification: patterns and paradigms. *Nature Reviews Genetics*. 10 (5). pp. 295-304.
- Cheema, Z., Hari-Gupta, Y., Kita, G., Farrar, D., Seddon, I., Corr, J. & Klenova, E. 2014. Expression of the cancer-testis antigen BORIS correlates with prostate cancer. *The Prostate*. 74 (2). pp. 164-176.
- Chen, C., Nott, T.J., Jin, J. & Pawson, T. 2011. Deciphering arginine methylation: Tudor tells the tale. *Nature reviews Molecular cell biology*. 12 (10). pp. 629-642.
- Chen, S., Fisher, R.C., Signs, S., Molina, L.A., Shenoy, A.K., Lopez, M., Baker, H.V., Koomen, J.M., Chen, Y. & Gittleman, H. 2017. Inhibition of PI3K/Akt/mTOR signaling in PI3KR2-overexpressing colon cancer stem cells reduces tumor growth due to apoptosis. *Oncotarget*. 8 (31). pp. 50476.
- Chen, W., Dong, J., Haiech, J., Kilhoffer, M. & Zeniou, M. 2016. Cancer stem cell quiescence and plasticity as major challenges in cancer therapy. *Stem cells international*. 2016.
- Chen, Y., Venditti, C.A., Theiler, G., Stevenson, B.J., Iseli, C., Gure, A.O., Jongeneel, C.V., Old, L.J. & Simpson, A.J. 2005. Identification of CT46/HORMAD1, an immunogenic cancer/testis antigen encoding a putative meiosis-related protein. *Cancer Immunity Archive*. 5 (1). pp. 9.
- Chen, J., Li, Y., Yu, T.S., McKay, R.M., Burns, D.K., Kernie, S.G. & Parada, L.F. 2012. A restricted cell population propagates glioblastoma growth after chemotherapy. *Nature*. 488 (7412). pp. 522-526.

- Chen, K., Huang, Y.H. & Chen, J.L. 2013. Understanding and targeting cancer stem cells: therapeutic implications and challenges. *Acta Pharmacologica Sinica*. 34 (6). pp. 732-740.
- Chen, L., Fan, J., Chen, H., Meng, Z., Chen, Z., Wang, P. & Liu, L. 2014. The IL-8/CXCR1 axis is associated with cancer stem cell-like properties and correlates with clinical prognosis in human pancreatic cancer cases. *Scientific reports*. 4 pp. 5911.
- Chen, S.R. & Liu, Y.X. 2015. Regulation of spermatogonial stem cell self-renewal and spermatocyte meiosis by Sertoli cell signaling. *Reproduction (Cambridge, England)*. 149 (4). pp. R159-67.
- Chen, Y.T., Gure, A.O., Tsang, S., Stockert, E., Jager, E., Knuth, A. & Old, L.J. 1998. Identification of multiple cancer/testis antigens by allogeneic antibody screening of a melanoma cell line library. *Proceedings of the National Academy of Sciences of the United States of America*. 95 (12). pp. 6919-6923.
- Chen, Y.T., Scanlan, M.J., Sahin, U., Tureci, O., Gure, A.O., Tsang, S., Williamson, B., Stockert, E., Pfreundschuh, M. & Old, L.J. 1997. A testicular antigen aberrantly expressed in human cancers detected by autologous antibody screening. *Proceedings of the National Academy of Sciences of the United States of America*. 94 (5). pp. 1914-1918.
- Chénais, B. 2015. Transposable elements in cancer and other human diseases. *Current cancer drug targets*. 15 (3). pp. 227-242.
- Cheng, Y., Wong, E.W. & Cheng, C.Y. 2011. Cancer/testis (CT) antigens, carcinogenesis and spermatogenesis. *Spermatogenesis*. 1 (3). pp. 209-220.
- Chiappinelli, K.B., Zahnow, C.A., Ahuja, N. & Baylin, S.B. 2016. Combining Epigenetic and Immunotherapy to Combat Cancer. *Cancer research*. 76 (7). pp. 1683-1689.
- Chiriva, M., Yu, Y., Mirandola, L., Jenkins, M., Chapman, C., Cannon, M., Cobos, E. & Kast, W.M. 2010. *Effective prevention and therapy of ovarian cancer with sperm protein 17 vaccination (95.7)*.
- Cho, N.W., Dilley, R.L., Lampson, M.A. & Greenberg, R.A. 2014. Interchromosomal homology searches drive directional ALT telomere movement and synapsis. *Cell*. 159 (1). pp. 108-121.
- Chuma, S., Hosokawa, M., Kitamura, K., Kasai, S., Fujioka, M., Hiyoshi, M., Takamune, K., Noce, T. & Nakatsuji, N. 2006. Tdrd1/Mtr-1, a tudor-related gene, is essential for male germ-cell differentiation and nuage/germinal granule formation in mice. *Proceedings of the National Academy of Sciences of the United States of America*. 103 (43). pp. 15894-15899.
- Cilensek, Z.M., Yehiely, F., Kular, R.K. & Deiss, L.P. 2002. A Member of the GAGE Family of Tumor Antigens is an Anti-Apoptotic Gene that Confers Resistance to Fas/CD95/APO-1, Interferon- γ , Taxol and γ -irradiation. *Cancer biology & therapy*. 1 (4). pp. 379-386.
- Clapp, R.W., Jacobs, M.M. & Loechler, E.L. 2008. Environmental and occupational causes of cancer: new evidence 2005-2007. *Reviews on environmental health*. 23 (1). pp. 1-38.
- Colak, S. & Medema, J.P. 2014. Cancer stem cells—important players in tumor therapy resistance. *The FEBS journal*. 281 (21). pp. 4779-4791.

- Colotta, F., Allavena, P., Sica, A., Garlanda, C. & Mantovani, A. 2009. Cancer-related inflammation, the seventh hallmark of cancer: links to genetic instability. *Carcinogenesis*. 30 (7). pp. 1073-1081.
- Condic, M.L. 2013. Totipotency: what it is and what it is not. *Stem cells and development*. 23 (8). pp. 796-812.
- Cooper, G.M. & Hausman, R.E. 2000. *The cell*. Sinauer Associates Sunderland: .
- Coqueret, O. 2003. New roles for p21 and p27 cell-cycle inhibitors: a function for each cell compartment? *Trends in cell biology*. 13 (2). pp. 65-70.
- Cortez, E., Roswall, P. & Pietras, K. 2014. Functional subsets of mesenchymal cell types in the tumor microenvironment. *Seminars in cancer biology*. Elsevier: pp. 3.
- Crichton, J.H., Playfoot, C.J., MacLennan, M., Read, D., Cooke, H.J. & Adams, I.R. 2017. Tex19. 1 promotes Spo11-dependent meiotic recombination in mouse spermatocytes. *PLoS genetics*. 13 (7). pp. e1006904.
- D'Andrea, V., Panarese, A., Tonda, M., Biffoni, M. & Monti, M. 2017. Cancer stem cells as functional biomarkers. *Cancer Biomarkers*. (Preprint). pp. 1-4.
- Dacic, S., Flanagan, M., Cieply, K., Ramalingam, S., Luketich, J., Belani, C. & Yousem, S.A. 2006. Significance of EGFR protein expression and gene amplification in non-small cell lung carcinoma. *American Journal of Clinical Pathology*. 125 (6). pp. 860-865.
- Dai, X., Shu, Y., Lou, Q., Tian, Q., Zhai, G., Song, J., Lu, S., Yu, H., He, J. & Yin, Z. 2017. Tdrd12 Is Essential for Germ Cell Development and Maintenance in Zebrafish. *International journal of molecular sciences*. 18 (6). pp. 1127.
- Damdimopoulou, P., Rodin, S., Stenfelt, S., Antonsson, L., Tryggvason, K. & Hovatta, O. 2016. Human embryonic stem cells. *Best Practice & Research Clinical Obstetrics & Gynaecology*. 31 pp. 2-12.
- Davies, H., Bignell, G.R., Cox, C., Stephens, P., Edkins, S., Clegg, S., Teague, J., Woffendin, H., Garnett, M.J. & Bottomley, W. 2002. Mutations of the BRAF gene in human cancer. *Nature*. 417 (6892). pp. 949-954.
- Dawood, S., Austin, L. & Cristofanilli, M. 2014. Cancer stem cells: implications for cancer therapy. *Oncology*. 28 (12). pp. 1101-1107.
- De Smet, C. & Lorient, A. 2013. DNA hypomethylation and activation of germline-specific genes in cancer. In: *Epigenetic Alterations in Oncogenesis*. Springer: pp. 149-166.
- De Smet, C. & Lorient, A. 2010. DNA hypomethylation in cancer: epigenetic scars of a neoplastic journey. *Epigenetics*. 5 (3). pp. 206-213.
- de Vries, F.A., de Boer, E., van den Bosch, M., Baarends, W.M., Ooms, M., Yuan, L., Liu, J.G., van Zeeland, A.A., Heyting, C. & Pastink, A. 2005. Mouse Sycp1 functions in synaptonemal complex assembly, meiotic recombination, and XY body formation. *Genes & development*. 19 (11). pp. 1376-1389.

- Deng, C., Zhang, P., Harper, J.W., Elledge, S.J. & Leder, P. 1995. Mice lacking p21CIP1/WAF1 undergo normal development, but are defective in G1 checkpoint control. *Cell*. 82 (4). pp. 675-684.
- Devi, P.U. 2004. Basics of carcinogenesis. *Health Adm.* 17 (1). pp. 16-24.
- Dewannieux, M. & Heidmann, T. 2013. Endogenous retroviruses: acquisition, amplification and taming of genome invaders. *Current opinion in virology*. 3 (6). pp. 646-656.
- Dhivya, S. & Premkumar, K. 2016. Nomadic genetic elements contribute to oncogenic translocations: Implications in carcinogenesis. *Critical reviews in oncology/hematology*. 98 pp. 81-93.
- Dick, J.E. 2008. Stem cell concepts renew cancer research. *Blood*. 112 (13). pp. 4793-4807.
- Ding, Y., Yang, D., Zhai, Y., Xue, K., Xu, F., Gu, X. & Wang, S. 2017. Microarray expression profiling of long non-coding RNAs in epithelial ovarian cancer. *Oncology letters*. 14 (2). pp. 2523-2530.
- Donovan, P.J. & Gearhart, J. 2001. The end of the beginning for pluripotent stem cells. *Nature*. 414 (6859). pp. 92.
- Downey, R.F., Sullivan, F.J., Wang-Johanning, F., Ambs, S., Giles, F.J. & Glynn, S.A. 2015. Human endogenous retrovirus K and cancer: innocent bystander or tumorigenic accomplice? *International journal of cancer*. 137 (6). pp. 1249-1257.
- Dudley, M.E., Yang, J.C., Sherry, R., Hughes, M.S., Royal, R., Kammula, U., Robbins, P.F., Huang, J., Citrin, D.E. & Leitman, S.F. 2008. Adoptive cell therapy for patients with metastatic melanoma: evaluation of intensive myeloablative chemoradiation preparative regimens. *Journal of Clinical Oncology*. 26 (32). pp. 5233-5239.
- Dwane, L., Gallagher, W.M., Ni Chonghaile, T. & O'Connor, D.P. 2017. The Emerging Role of Non-traditional Ubiquitination in Oncogenic Pathways. *The Journal of biological chemistry*. 292 (9). pp. 3543-3551.
- Edge, S.B., Byrd, D.R., Compton, C., Fritz, A., Greene, F. & Trotti, A. 2010. American joint committee on cancer staging manual. 7.
- Eisenbarth, G.S., Walsh, F.S. & Nirenberg, M. 1979. Monoclonal antibody to a plasma membrane antigen of neurons. *Proceedings of the National Academy of Sciences of the United States of America*. 76 (10). pp. 4913-4917.
- Ellis, L., Atadja, P.W. & Johnstone, R.W. 2009. Epigenetics in cancer: targeting chromatin modifications. *Molecular cancer therapeutics*. 8 (6). pp. 1409-1420.
- Eun, K., Ham, S.W. & Kim, H. 2017. Cancer stem cell heterogeneity: origin and new perspectives on CSC targeting. *BMB reports*. 50 (3). pp. 117-125.
- Fagegaltier, D., Falcatori, I., Czech, B., Castel, S., Perrimon, N., Simcox, A. & Hannon, G.J. 2016. Oncogenic transformation of Drosophila somatic cells induces a functional piRNA pathway. *Genes & development*. 30 (14). pp. 1623-1635.

- Feichtinger, J., Larcombe, L. & McFarlane, R.J. 2014. Meta-analysis of expression of 1 (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 (10). pp. 2359-2365.
- Feichtinger, J., Aldeaij, I., Anderson, R., Almutairi, M., Almatrafi, A., Alsiwiehri, N., Griffiths, K., Stuart, N., Wakeman, J.A., Larcombe, L. & McFarlane, R.J. 2012. 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.
- Flavahan, W.A., Gaskell, E. & Bernstein, B.E. 2017. Epigenetic plasticity and the hallmarks of cancer. *Science*. 357 (6348). pp. eaal2380.
- Fratta, E., Coral, S., Covre, A., Parisi, G., Colizzi, F., Danielli, R., Nicolay, M., Jean, H., Sigalotti, L. & Maio, M. 2011. The biology of cancer testis antigens: putative function, regulation and therapeutic potential. *Molecular oncology*. 5 (2). pp. 164-182.
- Freitas, M., Malheiros, S., Stavale, J.N., Biassi, T.P., Zamuner, F.T., de Souza Begnami, M., Soares, F.A. & Vettore, A.L. 2013. Expression of cancer/testis antigens is correlated with improved survival in glioblastoma. *Oncotarget*. 4 (4). pp. 636-646.
- Friedl, P. & Alexander, S. 2011. Cancer invasion and the microenvironment: plasticity and reciprocity. *Cell*. 147 (5). pp. 992-1009.
- Fujiwara, K., Daido, S., Yamamoto, A., Kobayashi, R., Yokoyama, T., Aoki, H., Iwado, E., Shinojima, N., Kondo, Y. & Kondo, S. 2008. Pivotal role of the cyclin-dependent kinase inhibitor p21WAF1/CIP1 in apoptosis and autophagy. *The Journal of biological chemistry*. 283 (1). pp. 388-397.
- Gang, A., Frøsig, T., Brimnes, M., Lyngaa, R., Treppendahl, M., Grønbæk, K., Dufva, I., thor Straten, P. & Hadrup, S. 2014. 5-Azacytidine treatment sensitizes tumor cells to T-cell mediated cytotoxicity and modulates NK cells in patients with myeloid malignancies. *Blood cancer journal*. 4 (3). pp. e197.
- Gao, J. 2008. Cancer stem cells: the lessons from pre-cancerous stem cells. *Journal of Cellular and Molecular Medicine*. 12 (1). pp. 67-96.
- Gedye, C., Sirskyj, D., Lobo, N.C., Meens, J., Hyatt, E., Robinette, M., Fleshner, N., Hamilton, R.J., Kulkarni, G. & Zlotta, A. 2016. Cancer stem cells are underestimated by standard experimental methods in clear cell renal cell carcinoma. *Scientific reports*. 6 pp. 25220.
- Ghafouri-Fard, S., Seifi-Alan, M., Shamsi, R. & Esfandiary, A. 2015. Immunotherapy in multiple myeloma using cancer-testis antigens. *Iranian journal of cancer prevention*. 8 (5).
- Ghavami, S., Hashemi, M., Ande, S.R., Yeganeh, B., Xiao, W., Eshraghi, M., Bus, C.J., Kadkhoda, K., Wiechec, E., Halayko, A.J. & Los, M. 2009. Apoptosis and cancer: mutations within caspase genes. *Journal of medical genetics*. 46 (8). pp. 497-510.
- Gjerstorff, M.F., Harkness, L., Kassem, M., Frandsen, U., Nielsen, O., Lutterodt, M., Møllgård, K. & Ditzel, H.J. 2008. Distinct GAGE and MAGE-A expression during early human development indicate specific roles in lineage differentiation. *Human reproduction*. 23 (10). pp. 2194-2201.

- Gjerstorff, M.F., Kock, K., Nielsen, O. & Ditzel, H.J. 2007. MAGE-A1, GAGE and NY-ESO-1 cancer/testis antigen expression during human gonadal development. *Human Reproduction*. 22 (4). pp. 953-960.
- Gjerstorff, M.F., Andersen, M.H. & Ditzel, H.J. 2015. Oncogenic cancer/testis antigens: prime candidates for immunotherapy. *Oncotarget*. 6 (18). pp. 15772-15787.
- Glass, K., Huttenhower, C., Quackenbush, J. & Yuan, G. 2013. Passing messages between biological networks to refine predicted interactions. *PloS one*. 8 (5). pp. e64832.
- Glickman, M.H. & Ciechanover, A. 2002. The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction. *Physiological Reviews*. 82 (2). pp. 373-428.
- Goldberger, N., Walker, R.C., Kim, C.H., Winter, S. & Hunter, K.W. 2013. Inherited variation in miR-290 expression suppresses breast cancer progression by targeting the metastasis susceptibility gene *Arid4b*. *Cancer research*. 73 (8). pp. 2671-2681.
- Goodier, J.L. 2014. Retrotransposition in tumors and brains. *Mobile DNA*. 5 (1). pp. 11.
- Gorospe, M., Cirielli, C., Wang, X., Seth, P., Capogrossi, M.C. & Holbrook, N.J. 1997. p21 Waf1/Cip1 protects against p53-mediated apoptosis of human melanoma cells. *Oncogene*. 14 (8). pp. 929.
- Gorospe, M., Wang, X., Guyton, K.Z. & Holbrook, N.J. 1996. Protective role of p21(Waf1/Cip1) against prostaglandin A2-mediated apoptosis of human colorectal carcinoma cells. *Molecular and cellular biology*. 16 (12). pp. 6654-6660.
- Greve, K.B., Lindgreen, J.N., Terp, M.G., Pedersen, C.B., Schmidt, S., Mollenhauer, J., Kristensen, S.B., Andersen, R.S., Relster, M.M. & Ditzel, H.J. 2015. Ectopic expression of cancer/testis antigen SSX2 induces DNA damage and promotes genomic instability. *Molecular oncology*. 9 (2). pp. 437-449.
- Grizzi, F. & Chiriva-Internati, M. 2006. Cancer: looking for simplicity and finding complexity. *Cancer Cell International*. 6 (1). pp. 4.
- Hanahan, D. & Weinberg, R.A. 2011. Hallmarks of cancer: the next generation. *Cell*. 144 (5). pp. 646-674.
- Hanahan, D. & Weinberg, R.A. 2000. The hallmarks of cancer. *Cell*. 100 (1). pp. 57-70.
- Harper, J.W., Adami, G.R., Wei, N., Keyomarsi, K. & Elledge, S.J. 1993. The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases. *Cell*. 75 (4). pp. 805-816.
- Hay, E. 1995. An overview of epithelio-mesenchymal transformation. *Cells Tissues Organs*. 154 (1). pp. 8-20.
- Hayashi, K., Yoshida, K. & Matsui, Y. 2005. A histone H3 methyltransferase controls epigenetic events required for meiotic prophase. *Nature*. 438 (7066). pp. 374-378.
- Henry, N.L. & Hayes, D.F. 2012. Cancer biomarkers. *Molecular oncology*. 6 (2). pp. 140-146.

- Herbig, U. & Sedivy, J.M. 2006. Regulation of growth arrest in senescence: telomere damage is not the end of the story. *Mechanisms of ageing and development*. 127 (1). pp. 16-24.
- Hiasa, A., Nishikawa, H., Hirayama, M., Kitano, S., Okamoto, S., Chono, H., Yu, S., Mineno, J., Tanaka, Y. & Minato, N. 2009. Rapid $\alpha\beta$ TCR-mediated responses in $\gamma\delta$ T cells transduced with cancer-specific TCR genes. *Gene therapy*. 16 (5). pp. 620-628.
- Hnisz, D., Weintraub, A.S., Day, D.S., Valton, A.L., Bak, R.O., Li, C.H., Goldmann, J., Lajoie, B.R., Fan, Z.P., Sigova, A.A., Reddy, J., Borges-Rivera, D., Lee, T.I., Jaenisch, R., Porteus, M.H., Dekker, J. & Young, R.A. 2016. Activation of proto-oncogenes by disruption of chromosome neighborhoods. *Science (New York, N.Y.)*. 351 (6280). pp. 1454-1458.
- Hockenbery, D.M., Oltvai, Z.N., Yin, X., Millman, C.L. & Korsmeyer, S.J. 1993. Bcl-2 functions in an antioxidant pathway to prevent apoptosis. *Cell*. 75 (2). pp. 241-251.
- Hofmann, O., Caballero, O.L., Stevenson, B.J., Chen, Y.T., Cohen, T., Chua, R., Maher, C.A., Panji, S., Schaefer, U., Kruger, A., Lehvaslaiho, M., Carninci, P., Hayashizaki, Y., Jongeneel, C.V., Simpson, A.J., Old, L.J. & Hide, W. 2008. Genome-wide analysis of cancer/testis gene expression. *Proceedings of the National Academy of Sciences of the United States of America*. 105 (51). pp. 20422-20427.
- Hollstein, M., Sidransky, D., Vogelstein, B. & Harris, C.C. 1991. P53 Mutations in Human Cancers. *Science (New York, N.Y.)*. 253 (5015). pp. 49-53.
- Holstein, A., Davidoff, M. & Schulze, W. 2003. Understanding spermatogenesis is a prerequisite for treatment. *Reproductive Biology and Endocrinology*. 1 (1). pp. 107.
- Hosoya, N., Okajima, M., Kinomura, A., Fujii, Y., Hiyama, T., Sun, J., Tashiro, S. & Miyagawa, K. 2011. Synaptonemal complex protein SYCP3 impairs mitotic recombination by interfering with BRCA2. *EMBO reports*. 13 (1). pp. 44-51.
- Hu, R., Qiu, X., Glazko, G., Klebanov, L. & Yakovlev, A. 2009. Detecting intergene correlation changes in microarray analysis: a new approach to gene selection. *BMC bioinformatics*. 10 (1). pp. 20.
- Hunter, T. & Pines, J. 1994. Cyclins and cancer II: cyclin D and CDK inhibitors come of age. *Cell*. 79 (4). pp. 573-582.
- Ishiguro, K. & Watanabe, Y. 2007. Chromosome cohesion in mitosis and meiosis. *Journal of cell science*. 120 (Pt 3). pp. 367-369.
- Itahana, K., Dimri, G.P., Hara, E., Itahana, Y., Zou, Y., Desprez, P.Y. & Campisi, J. 2002. A role for p53 in maintaining and establishing the quiescence growth arrest in human cells. *The Journal of biological chemistry*. 277 (20). pp. 18206-18214.
- Iwasaki, Y.W., Siomi, M.C. & Siomi, H. 2015. PIWI-interacting RNA: its biogenesis and functions. *Annual Review of Biochemistry*. 84.
- Jagadish, N., Parashar, D., Gupta, N., Agarwal, S., Sharma, A., Fatima, R., Suri, V., Kumar, R., Gupta, A. & Lohiya, N.K. 2016. A novel cancer testis antigen target a-kinase anchor protein (AKAP4) for the early diagnosis and immunotherapy of colon cancer. *Oncoimmunology*. 5 (2). pp. e1078965.

- Janic, A., Mendizabal, L., Llamazares, S., Rossell, D. & Gonzalez, C. 2010. Ectopic expression of germline genes drives malignant brain tumor growth in *Drosophila*. *Science (New York, N.Y.)*. 330 (6012). pp. 1824-1827.
- Janssen, A. & Medema, R. 2013. Genetic instability: tipping the balance. *Oncogene*. 32 (38). pp. 4459-4470.
- Jeon, H.M., Sohn, Y.W., Oh, S.Y., Kim, S.H., Beck, S., Kim, S. & Kim, H. 2011. ID4 imparts chemoresistance and cancer stemness to glioma cells by derepressing miR-9*-mediated suppression of SOX2. *Cancer research*. 71 (9). pp. 3410-3421.
- Jiang, Y., Liu, L., Shan, W. & Yang, Z. 2016. An integrated genomic analysis of Tudor domain-containing proteins identifies PHD finger protein 20-like 1 (PHF20L1) as a candidate oncogene in breast cancer. *Molecular oncology*. 10 (2). pp. 292-302.
- John, T., Starmans, M.H., Chen, Y., Russell, P.A., Barnett, S.A., White, S.C., Mitchell, P.L., Walkiewicz, M., Azad, A. & Lambin, P. 2013. The role of Cancer-Testis antigens as predictive and prognostic markers in non-small cell lung cancer. *PloS one*. 8 (7). pp. e67876.
- Jones, P.A. 2012. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nature Reviews Genetics*. 13 (7). pp. 484-492.
- Jones, P.A. & Baylin, S.B. 2007. The epigenomics of cancer. *Cell*. 128 (4). pp. 683-692.
- Jungbluth, A., Stockert, E., Chen, Y., Kolb, D., Iversen, K., Coplan, K., Williamson, B., Altorki, N., Busam, K. & Old, L. 2000. Monoclonal antibody MA454 reveals a heterogeneous expression pattern of MAGE-1 antigen in formalin-fixed paraffin embedded lung tumours. *British journal of cancer*. 83 (4). pp. 493-497.
- Kamata, Y., Kuhara, A., Iwamoto, T., Hayashi, K., Koido, S., Kimura, T., Egawa, S. & 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.
- Kapinas, K., Grandy, R., Ghule, P., Medina, R., Becker, K., Pardee, A., Zaidi, S.K., Lian, J., Stein, J. & van Wijnen, A. 2013. The abbreviated pluripotent cell cycle. *Journal of cellular physiology*. 228 (1). pp. 9-20.
- Karpf, A.R. 2006. A potential role for epigenetic modulatory drugs in the enhancement of cancer/germ-line antigen vaccine efficacy. *Epigenetics*. 1 (3). pp. 116-120.
- Kaseb, H.O., Fohrer-Ting, H., Lewis, D.W., Lagasse, E. & Gollin, S.M. 2016. Identification, expansion and characterization of cancer cells with stem cell properties from head and neck squamous cell carcinomas. *Experimental cell research*. 348 (1). pp. 75-86.
- Kasuga, C., Nakahara, Y., Ueda, S., Hawkins, C., Taylor, M.D., Smith, C.A. & Rutka, J.T. 2008. Expression of MAGE and GAGE genes in medulloblastoma and modulation of resistance to chemotherapy.
- Katzke, V.A., Kaaks, R. & Kuhn, T. 2015. Lifestyle and cancer risk. *Cancer journal (Sudbury, Mass.)*. 21 (2). pp. 104-110.

- Keller, G.M., Kouskoff, V., Kubo, A. & Fehling, H.J. 2010. *Mesoderm and definitive endoderm cell populations*.
- Kfoury, Y. & Scadden, D.T. 2015. Mesenchymal cell contributions to the stem cell niche. *Cell stem cell*. 16 (3). pp. 239-253.
- Khurana, J.S. & Theurkauf, W. 2010. piRNAs, transposon silencing, and Drosophila germline development. *The Journal of cell biology*. 191 (5). pp. 905-913.
- Kim, R., Kulkarni, P. & Hannenhalli, S. 2013. Derepression of Cancer/testis antigens in cancer is associated with distinct patterns of DNA hypomethylation. *BMC cancer*. 13 (1). pp. 144.
- Kim, M., Ki, B.S., Hong, K., Park, S.P., Ko, J.J. & Choi, Y. 2016. Tudor Domain Containing Protein TDRD12 Expresses at the Acrosome of Spermatids in Mouse Testis. *Asian-Australasian journal of animal sciences*. 29 (7). pp. 944-951.
- Klebanoff, C.A., Rosenberg, S.A. & Restifo, N.P. 2016. Prospects for gene-engineered T cell immunotherapy for solid cancers. *Nature medicine*. 22 (1). pp. 26-36.
- Kreso, A. & Dick, J.E. 2014. Evolution of the cancer stem cell model. *Cell stem cell*. 14 (3). pp. 275-291.
- Krishnadas, D.K., Bai, F. & Lucas, K.G. 2013. Cancer testis antigen and immunotherapy. *ImmunoTargets and therapy*. 2 pp. 11.
- Krishnakumar, R., Chen, A.F., Pantovich, M.G., Danial, M., Parchem, R.J., Labosky, P.A. & Blelloch, R. 2016. FOXD3 regulates pluripotent stem cell potential by simultaneously initiating and repressing enhancer activity. *Cell stem cell*. 18 (1). pp. 104-117.
- Kular, R.K., Yehiely, F., Kotlo, K.U., Cilensek, Z.M., Bedi, R. & Deiss, L.P. 2009. GAGE, an antiapoptotic protein binds and modulates the expression of nucleophosmin/B23 and interferon regulatory factor 1. *Journal of Interferon & Cytokine Research*. 29 (10). pp. 645-656.
- Kurth, R. & Bannert, N. 2010. Beneficial and detrimental effects of human endogenous retroviruses. *International journal of cancer*. 126 (2). pp. 306-314.
- Lachner, M., O'Sullivan, R.J. & Jenuwein, T. 2003. An epigenetic road map for histone lysine methylation. *Journal of cell science*. 116 (Pt 11). pp. 2117-2124.
- Ladanyi, M. 2001. Fusions of the SYT and SSX genes in synovial sarcoma. *Oncogene*. 20 (40). pp. 5755-5762.
- Lapidot, T., Sirard, C., Vormoor, J., Murdoch, B., Hoang, T., Caceres-Cortes, J., Minden, M., Paterson, B., Caligiuri, M.A. & Dick, J.E. 1994. A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. *Nature*. 367 (6464). pp. 645-648.
- Lathia, J.D. & Liu, H. 2017. Overview of Cancer Stem Cells and Stemness for Community Oncologists. *Targeted Oncology*. pp. 1-13.
- Lau, E.Y., Ho, N.P. & Lee, T.K. 2017. Cancer Stem Cells and Their Microenvironment: Biology and Therapeutic Implications. *Stem cells international*. 2017.

- Lee, E.K., Song, K., Chae, J., Kim, K., Kim, S. & Kang, M. 2015. GAGE12 mediates human gastric carcinoma growth and metastasis. *International journal of cancer*. 136 (10). pp. 2284-2292.
- Lee, T.S., Kim, J.W., Kang, G.H., Park, N.H., Song, Y.S., Kang, S.B. & Lee, H.P. 2006. DNA hypomethylation of CAGE promoters in squamous cell carcinoma of uterine cervix. *Annals of the New York Academy of Sciences*. 1091 (1). pp. 218-224.
- Lee, J. & Hirano, T. 2011. RAD21L, a novel cohesin subunit implicated in linking homologous chromosomes in mammalian meiosis. *The Journal of cell biology*. 192 (2). pp. 263-276.
- Lee, J., Iwai, T., Yokota, T. & Yamashita, M. 2003. Temporally and spatially selective loss of Rec8 protein from meiotic chromosomes during mammalian meiosis. *Journal of cell science*. 116 (Pt 13). pp. 2781-2790.
- Lee, Y.N., Malim, M.H. & Bieniasz, P.D. 2008. Hypermutation of an ancient human retrovirus by APOBEC3G. *Journal of virology*. 82 (17). pp. 8762-8770.
- Li, G., Miles, A., Line, A. & Rees, R.C. 2004. Identification of tumour antigens by serological analysis of cDNA expression cloning. *Cancer Immunology, Immunotherapy*. 53 (3). pp. 139-143.
- Li, N., Wang, T. & Han, D. 2012. Structural, cellular and molecular aspects of immune privilege in the testis. *Frontiers in immunology*. 3 pp. 152.
- Li, Y., Qi, H., Li, X., Hou, X., Lu, X. & Xiao, X. 2015a. A novel dithiocarbamate derivative induces cell apoptosis through p53-dependent intrinsic pathway and suppresses the expression of the E6 oncogene of human papillomavirus 18 in HeLa cells. *Apoptosis*. 20 (6). pp. 787-795.
- Li, Y., Li, A., Glas, M., Lal, B., Ying, M., Sang, Y., Xia, S., Trageser, D., Guerrero-Cazares, H., Eberhart, C.G., Quinones-Hinojosa, A., Scheffler, B. & Laterra, J. 2011. c-Met signaling induces a reprogramming network and supports the glioblastoma stem-like phenotype. *Proceedings of the National Academy of Sciences of the United States of America*. 108 (24). pp. 9951-9956.
- Li, Y., Rogoff, H.A., Keates, S., Gao, Y., Murikipudi, S., Mikule, K., Leggett, D., Li, W., Pardee, A.B. & Li, C.J. 2015b. Suppression of cancer relapse and metastasis by inhibiting cancer stemness. *Proceedings of the National Academy of Sciences of the United States of America*. 112 (6). pp. 1839-1844.
- Lie, P.P., Cheng, C.Y. & Mruk, D.D. 2013. Signalling pathways regulating the blood–testis barrier. *The international journal of biochemistry & cell biology*. 45 (3). pp. 621-625.
- Lindsey, S.F., Byrnes, D.M., Eller, M.S., Rosa, A.M., Dabas, N., Escandon, J. & Grichnik, J.M. 2013. Potential role of meiosis proteins in melanoma chromosomal instability. *Journal of skin cancer*. 2013 pp. 190109.
- Liu, H. & Lathia, J.D. 2016. Cancer Stem Cells: Targeting the Roots of Cancer, Seeds of Metastasis, and Sources of Therapy Resistance.
- Liu, K., Chen, C., Guo, Y., Lam, R., Bian, C., Xu, C., Zhao, D.Y., Jin, J., MacKenzie, F., Pawson, T. & Min, J. 2010. Structural basis for recognition of arginine methylated Piwi proteins by the extended Tudor domain. *Proceedings of the National Academy of Sciences of the United States of America*. 107 (43). pp. 18398-18403.

- Llano, E., Herran, Y., Garcia-Tunon, I., Gutierrez-Caballero, C., de Alava, E., Barbero, J.L., Schimenti, J., de Rooij, D.G., Sanchez-Martin, M. & Pendas, A.M. 2012. Meiotic cohesin complexes are essential for the formation of the axial element in mice. *The Journal of cell biology*. 197 (7). pp. 877-885.
- Lodish, H., Berk, A., Zipursky, S., Matsudaira, P., Baltimore, D. & Darnell, J. 2000. Proto-oncogenes and tumor-suppressor genes. *Molecular cell biology*. 4.
- Loeb, L.A., Loeb, K.R. & Anderson, J.P. 2003. Multiple mutations and cancer. *Proceedings of the National Academy of Sciences of the United States of America*. 100 (3). pp. 776-781.
- Louhimo, J., Alfthan, H., Stenman, U.H. & Haglund, C. 2004. Serum HCG beta and CA 72-4 are stronger prognostic factors than CEA, CA 19-9 and CA 242 in pancreatic cancer. *Oncology*. 66 (2). pp. 126-131.
- Lovell-Badge, R. 2001. The future for stem cell research. *Nature*. 414 (6859). pp. 88.
- Lüftl, M., Schuler, G. & Jungbluth, A. 2004. Melanoma or not? Cancer testis antigens may help. *British Journal of Dermatology*. 151 (6). pp. 1213-1218.
- Lui, W. & Cheng, C.Y. 2013. Transcriptional regulation of cell adhesion at the blood-testis barrier and spermatogenesis in the testis. In: *Biology and Regulation of Blood-Tissue Barriers*. Springer: pp. 281-294.
- Luo, J., Solimini, N.L. & Elledge, S.J. 2009. Principles of cancer therapy: oncogene and non-oncogene addiction. *Cell*. 136 (5). pp. 823-837.
- Luo, M., Yang, F., Leu, N.A., Landaiche, J., Handel, M.A., Benavente, R., La Salle, S. & Wang, P.J. 2013. MEIOB exhibits single-stranded DNA-binding and exonuclease activities and is essential for meiotic recombination. *Nature communications*. 4 pp. 2788.
- Lynch, T.J., Bell, D.W., Sordella, R., Gurubhagavatula, S., Okimoto, R.A., Brannigan, B.W., Harris, P.L., Haserlat, S.M., Supko, J.G. & Haluska, F.G. 2004. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *New England Journal of Medicine*. 350 (21). pp. 2129-2139.
- Maehle, A.H. 2011. Ambiguous cells: the emergence of the stem cell concept in the nineteenth and twentieth centuries. *Notes and records of the Royal Society of London*. 65 (4). pp. 359-378.
- Mahla, R.S., Reddy, N. & Goel, S. 2012. Spermatogonial stem cells (SSCs) in buffalo (*Bubalus bubalis*) testis. *PLoS One*. 7 (4). pp. e36020.
- Mani, S.A., Guo, W., Liao, M., Eaton, E.N., Ayyanan, A., Zhou, A.Y., Brooks, M., Reinhard, F., Zhang, C.C. & Shipitsin, M. 2008. The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell*. 133 (4). pp. 704-715.
- Mathews, L.A., Keller, J.M., Goodwin, B.L., Guha, R., Shinn, P., Mull, R., Thomas, C.J., de Kluyver, R.L., Sayers, T.J. & Ferrer, M. 2012. A 1536-well quantitative high-throughput screen to identify compounds targeting cancer stem cells. *Journal of biomolecular screening*. 17 (9). pp. 1231-1242.

- Matin, M.M., Walsh, J.R., Gokhale, P.J., Draper, J.S., Bahrami, A.R., Morton, I., Moore, H.D. & Andrews, P.W. 2004. Specific Knockdown of Oct4 and β 2-microglobulin Expression by RNA Interference in Human Embryonic Stem Cells and Embryonic Carcinoma Cells. *Stem cells*. 22 (5). pp. 659-668.
- Matsa, E., Rajamohan, D., Dick, E., Young, L., Mellor, I., Staniforth, A. & Denning, C. 2011. Drug evaluation in cardiomyocytes derived from human induced pluripotent stem cells carrying a long QT syndrome type 2 mutation. *European heart journal*. 32 (8). pp. 952-962.
- McFarlane, R.J., Feichtinger, J. & Larcombe, L. 2014. *Cancer germline gene activation: friend or foe?*.
- McFarlane, R.J., Feichtinger, J. & Larcombe, L. 2015. Germline/meiotic genes in cancer: new dimensions. *Cell cycle (Georgetown, Tex.)*. 14 (6). pp. 791-792.
- McFarlane, R.J. & Wakeman, J.A. 2017. Meiosis-like Functions in Oncogenesis: A New View of Cancer. *Cancer research*. 77 (21). pp. 5712-5716.
- McGranahan, N. & Swanton, C. 2017. Clonal heterogeneity and tumor evolution: past, present, and the future. *Cell*. 168 (4). pp. 613-628.
- Medema, J.P. 2013. Cancer stem cells: the challenges ahead. *Nature cell biology*. 15 (4). pp. 338.
- Meek, D.W. & Marcar, L. 2012. MAGE-A antigens as targets in tumour therapy. *Cancer letters*. 324 (2). pp. 126-132.
- Mellman, I., Coukos, G. & Dranoff, G. 2011. Cancer immunotherapy comes of age. *Nature*. 480 (7378). pp. 480-489.
- Meyer, D.B. 1985. *Laboratory Guide for Human Histology*. Wayne State University Press: .
- Mishra, A. & Verma, M. 2010. Cancer biomarkers: are we ready for the prime time? *Cancers*. 2 (1). pp. 190-208.
- Mital, P., Hinton, B.T. & Dufour, J.M. 2011. The blood-testis and blood-epididymis barriers are more than just their tight junctions. *Biology of reproduction*. 84 (5). pp. 851-858.
- Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M. & Yamanaka, S. 2003. The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell*. 113 (5). pp. 631-642.
- Moore, N. & Lyle, S. 2011. Quiescent, slow-cycling stem cell populations in cancer: a review of the evidence and discussion of significance. *Journal of oncology*. 2011 pp. 10.1155/2011/396076. Epub 2010 Sep 29.
- Morgan, D.O. 1995. Principles of CDK regulation. *Nature*. 374 (6518). pp. 131.
- Moyano, M. & Stefani, G. 2015. piRNA involvement in genome stability and human cancer. *Journal of hematology & oncology*. 8 (1). pp. 38.

- Mruk, D.D. & 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.
- Nagaoka, S.I., Hassold, T.J. & Hunt, P.A. 2012. Human aneuploidy: mechanisms and new insights into an age-old problem. *Nature Reviews Genetics*. 13 (7). pp. 493-504.
- Nassar, D. & Blanpain, C. 2016. Cancer stem cells: basic concepts and therapeutic implications. *Annual Review of Pathology: Mechanisms of Disease*. 11 pp. 47-76.
- National Institutes of Health 2009. Stem cell basics. *Stem Cell Information*, available at <http://stemcells.nih.gov/info/basics/basics6.asp>.
- Negrini, S., Gorgoulis, V.G. & Halazonetis, T.D. 2010. Genomic instability—an evolving hallmark of cancer. *Nature reviews Molecular cell biology*. 11 (3). pp. 220-228.
- Neidhardt, F.C., Ingraham, J.L. & Schaechter, M. 1990. *Physiology of the bacterial cell: a molecular approach*. Sinauer Sunderland: .
- Neumann, B., Walter, T., Hériché, J., Bulkescher, J., Erfle, H., Conrad, C., Rogers, P., Poser, I., Held, M. & Liebel, U. 2010. Phenotypic profiling of the human genome by time-lapse microscopy reveals cell division genes. *Nature*. 464 (7289). pp. 721.
- Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Schöler, H. & Smith, A. 1998. Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell*. 95 (3). pp. 379-391.
- Nielsen, A.Y. & Gjerstorff, M.F. 2016. Ectopic expression of testis germ cell proteins in cancer and its potential role in genomic instability. *International journal of molecular sciences*. 17 (6). pp. 890.
- Niwa, H., Miyazaki, J. & Smith, A.G. 2000. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nature genetics*. 24 (4). pp. 372.
- Noda, A., Ning, Y., Venable, S.F., Pereira-Smith, O.M. & Smith, J.R. 1994. Cloning of senescent cell-derived inhibitors of DNA synthesis using an expression screen. *Experimental cell research*. 211 (1). pp. 90-98.
- Novak, K. 2004. Epigenetics changes in cancer cells. *MedGenMed : Medscape general medicine*. 6 (4). pp. 17.
- O'Brien, T. & Barry, F.P. 2009. Stem cell therapy and regenerative medicine. *Mayo Clinic proceedings*. 84 (10). pp. 859-861.
- 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. & Nakayama, E. 2001. Identification of proacrosin binding protein sp32 precursor as a human cancer/testis antigen. *Proceedings of the National Academy of Sciences of the United States of America*. 98 (6). pp. 3282-3287.
- Ouyang, G., Wang, Z., Fang, X., Liu, J. & Yang, C.J. 2010. Molecular signaling of the epithelial to mesenchymal transition in generating and maintaining cancer stem cells. *Cellular and Molecular Life Sciences*. 67 (15). pp. 2605-2618.

- Overturf, K., al-Dhalimy, M., Ou, C.N., Finegold, M. & Grompe, M. 1997. Serial transplantation reveals the stem-cell-like regenerative potential of adult mouse hepatocytes. *The American journal of pathology*. 151 (5). pp. 1273-1280.
- Oyouni, A.A.A. 2016. *The medical biomarker and oncogenic potential of the human cancer-and stem/germ cell-specific gene TDRD12*.
- Pagotto, A., Caballero, O.L., Volkmar, N., Devalle, S., Simpson, A.J., Lu, X. & Christianson, J.C. 2013. Centrosomal localisation of the cancer/testis (CT) antigens NY-ESO-1 and MAGE-C1 is regulated by proteasome activity in tumour cells. *PLoS one*. 8 (12). pp. e83212.
- Pandey, R.R., Tokuzawa, Y., Yang, Z., Hayashi, E., Ichisaka, T., Kajita, S., Asano, Y., Kunieda, T., Sachidanandam, R., Chuma, S., Yamanaka, S. & Pillai, R.S. 2013. Tudor domain containing 12 (TDRD12) is essential for secondary PIWI interacting RNA biogenesis in mice. *Proceedings of the National Academy of Sciences of the United States of America*. 110 (41). pp. 16492-16497.
- Park, C.S., Shen, Y., Suppipat, K., Tomolonis, J., Puppi, M., Mistretta, T., Ma, L., Green, M. & Lacorazza, D. 2016. *KLF4 promotes self-renewal by repressing DYRK2-mediated degradation of c-Myc in leukemic stem cells: development of targeted therapy*.
- Park, T.S., Bhutto, I., Zimmerlin, L., Huo, J.S., Nagaria, P., Miller, D., Rufaihah, A.J., Talbot, C., Aguilar, J., Grebe, R., Merges, C., Reijo-Pera, R., Feldman, R.A., Rassool, F., Cooke, J., Luty, G. & Zambidis, E.T. 2014. Vascular progenitors from cord blood-derived induced pluripotent stem cells possess augmented capacity for regenerating ischemic retinal vasculature. *Circulation*. 129 (3). pp. 359-372.
- Paterson, A.L., Weaver, J.M., Eldridge, M.D., Tavaré, S., Fitzgerald, R.C. & Edwards, P.A. 2015. Mobile element insertions are frequent in oesophageal adenocarcinomas and can mislead paired-end sequencing analysis. *BMC genomics*. 16 (1). pp. 473.
- Pek, J.W., Anand, A. & Kai, T. 2012. Tudor domain proteins in development. *Development (Cambridge, England)*. 139 (13). pp. 2255-2266.
- Pera, M.F., Cooper, S., Mills, J. & Parrington, J.M. 1989. Isolation and characterization of a multipotent clone of human embryonal carcinoma cells. *Differentiation*. 42 (1). pp. 10-23.
- Pera, M.F., Reubinoff, B. & Trounson, A. 2000. Human embryonic stem cells. *Journal of cell science*. 113 (Pt 1) (Pt 1). pp. 5-10.
- Perucca, P., Cazzalini, O., Madine, M., Savio, M., Laskey, R.A., Vannini, V., Prosperi, E. & Stivala, L.A. 2009. Loss of p21CDKN1A impairs entry to quiescence and activates a DNA damage response in normal fibroblasts induced to quiescence. *Cell cycle*. 8 (1). pp. 105-114.
- Pickart, C.M. & Eddins, M.J. 2004. Ubiquitin: structures, functions, mechanisms. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*. 1695 (1). pp. 55-72.
- Pierson, E., Koller, D., Battle, A., Mostafavi, S. & GTEx Consortium 2015. Sharing and specificity of co-expression networks across 35 human tissues. *PLoS computational biology*. 11 (5). pp. e1004220.

- Plaks, V., Kong, N. & Werb, Z. 2015. The cancer stem cell niche: how essential is the niche in regulating stemness of tumor cells? *Cell stem cell*. 16 (3). pp. 225-238.
- Planells-Palop, V., Hazazi, A., Feichtinger, J., Jezkova, J., Thallinger, G., Alsiwiehri, N.O., Almutairi, M., Parry, L., Wakeman, J.A. & McFarlane, R.J. 2017. Human germ/stem cell-specific gene TEX19 influences cancer cell proliferation and cancer prognosis. *Molecular cancer*. 16 (1). pp. 84.
- Pozner-Moulis, S., Cregger, M., Camp, R.L. & Rimm, D.L. 2007. Antibody validation by quantitative analysis of protein expression using expression of Met in breast cancer as a model. *Laboratory investigation*. 87 (3). pp. 251-260.
- Przyborski, S., Christie, V., Hayman, M., Stewart, R. & Horrocks, G. 2004. Human embryonal carcinoma stem cells: models of embryonic development in humans. *Stem cells and development*. 13 (4). pp. 400-408.
- Przyborski, S., Morton, I., Wood, A. & Andrews, P. 2000. Developmental regulation of neurogenesis in the pluripotent human embryonal carcinoma cell line NTERA-2. *European Journal of Neuroscience*. 12 (10). pp. 3521-3528.
- Przyborski, S.A. 2001. Isolation of human embryonal carcinoma stem cells by immunomagnetic sorting. *Stem cells*. 19 (6). pp. 500-504.
- Raggi, C., Mousa, H., Correnti, M., Sica, A. & Invernizzi, P. 2016. Cancer stem cells and tumor-associated macrophages: a roadmap for multitargeting strategies. *Oncogene*. 35 (6). pp. 671-682.
- Ren, B., Wei, X., Zou, G., He, J., Xu, G., Xu, F., Huang, Y., Zhu, H., Li, Y. & Ma, G. 2016. Cancer testis antigen SPAG9 is a promising marker for the diagnosis and treatment of lung cancer. *Oncology reports*. 35 (5). pp. 2599-2605.
- Reya, T., Morrison, S.J., Clarke, M.F. & Weissman, I.L. 2001. Stem cells, cancer, and cancer stem cells. *Nature*. 414 (6859). pp. 105.
- Rivera, M., Wu, Q., Hamerlik, P., Hjelmeland, A., Bao, S. & Rich, J. 2015. Acquisition of meiotic DNA repair regulators maintain genome stability in glioblastoma. *Cell death & disease*. 6 (4). pp. e1732.
- Robbins, P.F., Morgan, R.A., Feldman, S.A., Yang, J.C., Sherry, R.M., Dudley, M.E., Wunderlich, J.R., Nahvi, A.V., Helman, L.J. & Mackall, C.L. 2011. Tumor regression in patients with metastatic synovial cell sarcoma and melanoma using genetically engineered lymphocytes reactive with NY-ESO-1. *Journal of Clinical Oncology*. 29 (7). pp. 917-924.
- Robert, T., Nore, A., Brun, C., Maffre, C., Crimi, B., Bourbon, H.M. & de Massy, B. 2016. The TopoVIB-Like protein family is required for meiotic DNA double-strand break formation. *Science (New York, N.Y.)*. 351 (6276). pp. 943-949.
- Rodda, D.J., Chew, J.L., Lim, L.H., Loh, Y.H., Wang, B., Ng, H.H. & Robson, P. 2005. Transcriptional regulation of nanog by OCT4 and SOX2. *The Journal of biological chemistry*. 280 (26). pp. 24731-24737.

- Rosa, A.M., Dabas, N., Byrnes, D.M., Eller, M.S. & Grichnik, J.M. 2012. Germ cell proteins in melanoma: prognosis, diagnosis, treatment, and theories on expression. *Journal of skin cancer*. 2012 pp. 621968.
- Rossi, F., Molnar, C., Hashiyama, K., Heinen, J.P., Pampalona, J., Llamazares, S., Reina, J., Hashiyama, T., Rai, M., Pollarolo, G., Fernandez-Hernandez, I. & Gonzalez, C. 2017. An in vivo genetic screen in *Drosophila* identifies the orthologue of human cancer/testis gene SPO11 among a network of targets to inhibit lethal(3)malignant brain tumour growth. *Open biology*. 7 (8). pp. 10.1098/rsob.170156.
- 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. & Khochbin, S. 2013. Ectopic activation of germline and placental genes identifies aggressive metastasis-prone lung cancers. *Science translational medicine*. 5 (186). pp. 186ra66.
- Roy, N.K., Bordoloi, D., Monisha, J., Anip, A., Padmavathi, G. & Kunnumakkara, A.B. 2017. Cancer—An Overview and Molecular Alterations in Cancer. In: *Fusion Genes and Cancer*. pp. 1-15.
- Rumman, M., Dhawan, J. & Kassem, M. 2015. Concise review: quiescence in adult stem cells: biological significance and relevance to tissue regeneration. *Stem cells*. 33 (10). pp. 2903-2912.
- Sadanandam, A., Wang, X., de Sousa E Melo, Felipe, Gray, J.W., Vermeulen, L., Hanahan, D. & Medema, J.P. 2014. Reconciliation of classification systems defining molecular subtypes of colorectal cancer: interrelationships and clinical implications. *Cell cycle*. 13 (3). pp. 353-357.
- Sallesse, S. & Verfaillie, C.M. 2002. BCR/ABL: from molecular mechanisms of leukemia induction to treatment of chronic myelogenous leukemia. *Oncogene*. 21 (56). pp. 8547-8559.
- Salvesen, G.S. 2002. *Caspases: opening the boxes and interpreting the arrows*.
- Sammut, S.J., Feichtinger, J., Stuart, N., Wakeman, J.A., Larcombe, L. & 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.
- Sarasin, A. & Kauffmann, A. 2008. Overexpression of DNA repair genes is associated with metastasis: a new hypothesis. *Mutation Research/Reviews in Mutation Research*. 659 (1). pp. 49-55.
- Saygin, C., Samour, M., Chumakova, A., Jarrar, A., Lathia, J.D. & Reizes, O. 2016. Reporter Systems to Study Cancer Stem Cells. *Stem Cell Heterogeneity: Methods and Protocols*. pp. 319-333.
- Scadden, D.T. 2006. The stem-cell niche as an entity of action. *Nature*. 441 (7097). pp. 1075.
- Scanlan, M.J., Simpson, A.J. & Old, L.J. 2004. The cancer/testis genes: review, standardization, and commentary. *Cancer Immunity Archive*. 4 (1). pp. 1.
- Schramm, S., Fraune, J., Naumann, R., Hernandez-Hernandez, A., Höög, C., Cooke, H.J., Alsheimer, M. & 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.

- Schulz, W., Steinhoff, C. & Florl, A. 2006. Methylation of endogenous human retroelements in health and disease. In: *DNA Methylation: Development, Genetic Disease and Cancer*. Springer: pp. 211-250.
- Sell, S. 2004. Stem cell origin of cancer and differentiation therapy. *Critical reviews in oncology/hematology*. 51 (1). pp. 1-28.
- Senti, K. & Brennecke, J. 2010. The piRNA pathway: a fly's perspective on the guardian of the genome. *Trends in Genetics*. 26 (12). pp. 499-509.
- Seoane, J., Le, H. & Massagué, J. 2002. Myc suppression of the p21 Cip1 Cdk inhibitor influences the outcome of the p53 response to DNA damage. *Nature*. 419 (6908). pp. 729.
- Serrano, M., Hannon, G.J. & Beach, D. 1993. A new regulatory motif in cell-cycle control causing specific inhibition of cyclin D/CDK4. *Nature*. 366 (6456). pp. 704.
- Shang, B., Gao, A., Pan, Y., Zhang, G., Tu, J., Zhou, Y., Yang, P., Cao, Z., Wei, Q., Ding, Y., Zhang, J., Zhao, Y. & Zhou, Q. 2014. CT45A1 acts as a new proto-oncogene to trigger tumorigenesis and cancer metastasis. *Cell death & disease*. 5 pp. e1285.
- Shao, Z.M., Dawson, M.I., Li, X.S., Rishi, A.K., Sheikh, M.S., Han, Q.X., Ordonez, J.V., Shroot, B. & Fontana, J.A. 1995. p53 independent G0/G1 arrest and apoptosis induced by a novel retinoid in human breast cancer cells. *Oncogene*. 11 (3). pp. 493-504.
- Sharma, S., Kelly, T.K. & Jones, P.A. 2010. Epigenetics in cancer. *Carcinogenesis*. 31 (1). pp. 27-36.
- Shen, H. & Laird, P.W. 2013. Interplay between the cancer genome and epigenome. *Cell*. 153 (1). pp. 38-55.
- Sherr, C.J. & Roberts, J.M. 1999. CDK inhibitors: positive and negative regulators of G1-phase progression. *Genes & development*. 13 (12). pp. 1501-1512.
- Shiohama, Y., Ohtake, J., Ohkuri, T., Noguchi, D., Togashi, Y., Kitamura, H. & Nishimura, T. 2014. Identification of a meiosis-specific protein, MEIOB, as a novel cancer/testis antigen and its augmented expression in demethylated cancer cells. *Immunology letters*. 158 (1). pp. 175-182.
- Shoji, M., Tanaka, T., Hosokawa, M., Reuter, M., Stark, A., Kato, Y., Kondoh, G., Okawa, K., Chujo, T. & Suzuki, T. 2009. The TDRD9-MIWI2 complex is essential for piRNA-mediated retrotransposon silencing in the mouse male germline. *Developmental cell*. 17 (6). pp. 775-787.
- Shukla, R., Upton, K.R., Muñoz-Lopez, M., Gerhardt, D.J., Fisher, M.E., Nguyen, T., Brennan, P.M., Baillie, J.K., Collino, A. & Ghisletti, S. 2013. Endogenous retrotransposition activates oncogenic pathways in hepatocellular carcinoma. *Cell*. 153 (1). pp. 101-111.
- Siegel, R., Naishadham, D. & Jemal, A. 2013. Cancer statistics, 2013. *CA: a cancer journal for clinicians*. 63 (1). pp. 11-30.
- Simpson, A.J., Caballero, O.L., Jungbluth, A., Yao-Tseng, C. & Old, L.J. 2005. Cancer/testis antigens, gametogenesis and cancer. *Nature reviews.Cancer*. 5 (8). pp. 615.

- Singh, A., Singh, D.K. & Bhorla, U. 2015. Induced pluripotent stem cells: An update. *International Journal of Blood transfusion and Immunohematology*. 5 pp. 6-13.
- Siomi, M.C., Mannen, T. & Siomi, H. 2010. How does the royal family of Tudor rule the PIWI-interacting RNA pathway? *Genes & development*. 24 (7). pp. 636-646.
- Sirko, S., Behrendt, G., Johansson, P.A., Tripathi, P., Costa, M.R., Bek, S., Heinrich, C., Tiedt, S., Colak, D. & Dichgans, M. 2013. Reactive glia in the injured brain acquire stem cell properties in response to sonic hedgehog. *Cell stem cell*. 12 (4). pp. 426-439.
- Slack, J.M. 2008. Origin of stem cells in organogenesis. *Science (New York, N.Y.)*. 322 (5907). pp. 1498-1501.
- Slamon, D., Clark, G., Wong, S., Levin, W., Ullrich, A. & McGuire, W. 1987. Human breast cancer: correlation of relapse and. *Science*. 3798106 (177). pp. 235.
- Sonnenschein, C. & Soto, A.M. 2013. The aging of the 2000 and 2011 Hallmarks of Cancer reviews: a critique. *Journal of Biosciences*. 38 (3). pp. 651-663.
- Souquet, B., Abby, E., Hervé, R., Finsterbusch, F., Tourpin, S., Le Bouffant, R., Duquenne, C., Messiaen, S., Martini, E. & Bernardino-Sgherri, J. 2013. MEIOB targets single-strand DNA and is necessary for meiotic recombination. *PLoS genetics*. 9 (9). pp. e1003784.
- Spradling, A., Drummond-Barbosa, D. & Kai, T. 2001. Stem cells find their niche. *Nature*. 414 (6859). pp. 98.
- Stewart, R., Christie, V.B. & Przyborski, S.A. 2003. Manipulation of human pluripotent embryonal carcinoma stem cells and the development of neural subtypes. *Stem cells*. 21 (3). pp. 248-256.
- Stoye, J.P. 2012. Studies of endogenous retroviruses reveal a continuing evolutionary saga. *Nature Reviews Microbiology*. 10 (6). pp. 395-406.
- Strunnikov, A. 2013. Cohesin complexes with a potential to link mammalian meiosis to cancer. *Cell Regeneration*. 2 (1). pp. 4.
- Sumiyoshi, T., Sato, K., Yamamoto, H., Iwasaki, Y.W., Siomi, H. & Siomi, M.C. 2016. Loss of l(3)mbt leads to acquisition of the ping-pong cycle in Drosophila ovarian somatic cells. *Genes & development*. 30 (14). pp. 1617-1622.
- Sun, Y., Campisi, J., Higano, C., Beer, T.M., Porter, P., Coleman, I., True, L. & Nelson, P.S. 2012. Treatment-induced damage to the tumor microenvironment promotes prostate cancer therapy resistance through WNT16B. *Nature medicine*. 18 (9). pp. 1359-1368.
- Surani, A. & Tischler, J. 2012. Stem cells: a sporadic super state. *Nature*. 487 (7405). pp. 43-45.
- Suzuki, R., Honda, S. & Kirino, Y. 2012. PIWI expression and function in cancer. *Frontiers in genetics*. 3 pp. 204.
- Tachibana, M., Sugimoto, K., Nozaki, M., Ueda, J., Ohta, T., Ohki, M., Fukuda, M., Takeda, N., Niida, H., Kato, H. & Shinkai, Y. 2002. G9a histone methyltransferase plays a dominant role in

- euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. *Genes & development*. 16 (14). pp. 1779-1791.
- Tachibana, M., Ueda, J., Fukuda, M., Takeda, N., Ohta, T., Iwanari, H., Sakihama, T., Kodama, T., Hamakubo, T. & Shinkai, Y. 2005. Histone methyltransferases G9a and GLP form heteromeric complexes and are both crucial for methylation of euchromatin at H3-K9. *Genes & development*. 19 (7). pp. 815-826.
- Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K. & Yamanaka, S. 2007. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell*. 131 (5). pp. 861-872.
- Takebe, T., Sekine, K., Enomura, M., Koike, H., Kimura, M., Ogaeri, T., Zhang, R., Ueno, Y., Zheng, Y. & Koike, N. 2013. Vascularized and functional human liver from an iPSC-derived organ bud transplant. *Nature*. 499 (7459). pp. 481-484.
- Taverna, P., Lyons, J.F., Hao, Y., Azab, M., Kantarjian, H.M., Kropf, P. & Issa, J. 2013. *Determinants of demethylation and clinical response in AML patients treated with SGI-110, a novel subcutaneous (SQ) hypomethylating agent (HMA) in a phase 1 study.*
- Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S. & Jones, J.M. 1998. Embryonic stem cell lines derived from human blastocysts. *Science (New York, N.Y.)*. 282 (5391). pp. 1145-1147.
- Tomasetti, C., Li, L. & Vogelstein, B. 2017. Stem cell divisions, somatic mutations, cancer etiology, and cancer prevention. *Science (New York, N.Y.)*. 355 (6331). pp. 1330-1334.
- Tripsianes, K., Madl, T., Machyna, M., Fessas, D., Englbrecht, C., Fischer, U., Neugebauer, K.M. & Sattler, M. 2011. Structural basis for dimethylarginine recognition by the Tudor domains of human SMN and SPF30 proteins. *Nature Structural and Molecular Biology*. 18 (12). pp. 1414.
- Tureci, O., Sahin, U., Zwick, C., Koslowski, M., Seitz, G. & 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 of the United States of America*. 95 (9). pp. 5211-5216.
- Tweedell, K.S. 2017. The Adaptability of Somatic Stem Cells: A Review. *Journal of stem cells & regenerative medicine*. 13 (1). pp. 3.
- van Dam, S., Vösa, U., van der Graaf, A., Franke, L. & de Magalhães, J.P. 2017. Gene co-expression analysis for functional classification and gene–disease predictions. *Briefings in bioinformatics*. pp. bbw139.
- van der Bruggen, P., Traversari, C., Chomez, P., Lurquin, C., De Plaen, E., Van den Eynde, B., Knuth, A. & Boon, T. 1991. A gene encoding an antigen recognized by cytolytic T lymphocytes on a human melanoma. *Science (New York, N.Y.)*. 254 (5038). pp. 1643-1647.
- Van Tongelen, A., Lorient, A. & De Smet, C. 2017. Oncogenic roles of DNA hypomethylation through the activation of cancer-germline genes. *Cancer letters*. 396 pp. 130-137.
- Venkatesan, S., Birkbak, N.J. & Swanton, C. 2017. Constraints in cancer evolution. *Biochemical Society transactions*. 45 (1). pp. 1-13.

- Verga Falzacappa, M.V., Ronchini, C., Reavie, L.B. & Pelicci, P.G. 2012. Regulation of self-renewal in normal and cancer stem cells. *The FEBS journal*. 279 (19). pp. 3559-3572.
- Visvader, J.E. 2011. Cells of origin in cancer. *Nature*. 469 (7330). pp. 314.
- Visvader, J.E. & Clevers, H. 2016. Tissue-specific designs of stem cell hierarchies. *Nature cell biology*. 18 (4). pp. 349-355.
- Vrielynck, N., Chambon, A., Vezon, D., Pereira, L., Chelysheva, L., De Muyt, A., Mezard, C., Mayer, C. & Grelon, M. 2016. A DNA topoisomerase VI-like complex initiates meiotic recombination. *Science (New York, N.Y.)*. 351 (6276). pp. 939-943.
- Wang, J., Emadali, A., Le Bescont, A., Callanan, M., Rousseaux, S. & Khochbin, S. 2011. Induced malignant genome reprogramming in somatic cells by testis-specific factors. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms*. 1809 (4). pp. 221-225.
- Wang, P.J., McCarrey, J.R., Yang, F. & Page, D.C. 2001. An abundance of X-linked genes expressed in spermatogonia. *Nature genetics*. 27 (4). pp. 422-426.
- Wang, X., Huang, S. & Chen, J. 2017. Understanding of leukemic stem cells and their clinical implications. *Molecular cancer*. 16 (1). pp. 2.
- Wang, B.B., Li, Z.J., Zhang, F.F., Hou, H.T., Yu, J.K. & Li, F. 2016. Clinical significance of stem cell marker CD133 expression in colorectal cancer. *Histology and histopathology*. 31 (3). pp. 299-306.
- Weiderpass, E. 2010. Lifestyle and cancer risk. *J Prev Med Public Health*. 43 (6). pp. 459-471.
- Welch, J.S., Ley, T.J., Link, D.C., Miller, C.A., Larson, D.E., Koboldt, D.C., Wartman, L.D., Lamprecht, T.L., Liu, F. & Xia, J. 2012. The origin and evolution of mutations in acute myeloid leukemia. *Cell*. 150 (2). pp. 264-278.
- Wert, G.d. & Mummery, C. 2003. Human embryonic stem cells: research, ethics and policy. *Human reproduction*. 18 (4). pp. 672-682.
- Whitehurst, A.W. 2014. Cause and consequence of cancer/testis antigen activation in cancer. *Annual Review of Pharmacology and Toxicology*. 54 pp. 251-272.
- Wicker, T., Sabot, F., Hua-Van, A., Bennetzen, J.L., Capy, P., Chalhoub, B., Flavell, A., Leroy, P., Morgante, M. & Panaud, O. 2007. A unified classification system for eukaryotic transposable elements. *Nature Reviews Genetics*. 8 (12). pp. 973-982.
- Wischnewski, F., Friese, O., Pantel, K. & 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.
- Won, K.A., Xiong, Y., Beach, D. & Gilman, M.Z. 1992. Growth-regulated expression of D-type cyclin genes in human diploid fibroblasts. *Proceedings of the National Academy of Sciences of the United States of America*. 89 (20). pp. 9910-9914.

- Wu, J., Liu, S., Liu, G., Dombkowski, A., Abrams, J., Martin-Trevino, R., Wicha, M., Ethier, S. & Yang, Z. 2012a. Identification and functional analysis of 9p24 amplified genes in human breast cancer. *Oncogene*. 31 (3). pp. 333.
- Wu, R., Forget, M.A., Chacon, J., Bernatchez, C., Haymaker, C., Chen, J.Q., Hwu, P. & Radvanyi, L.G. 2012b. Adoptive T-cell therapy using autologous tumor-infiltrating lymphocytes for metastatic melanoma: current status and future outlook. *Cancer journal (Sudbury, Mass.)*. 18 (2). pp. 160-175.
- Xue, Z., Huang, K., Cai, C., Cai, L., Jiang, C., Feng, Y., Liu, Z., Zeng, Q., Cheng, L. & Sun, Y.E. 2013. Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing. *Nature*. 500 (7464). pp. 593.
- Yabuta, Y., Ohta, H., Abe, T., Kurimoto, K., Chuma, S. & Saitou, M. 2011. TDRD5 is required for retrotransposon silencing, chromatoid body assembly, and spermiogenesis in mice. *The Journal of cell biology*. 192 (5). pp. 781-795.
- Yan, G., Yang, L., Lv, Y., Shi, Y., Shen, L., Yao, X., Guo, Q., Zhang, P., Cui, Y. & Zhang, X. 2014. Endothelial cells promote stem-like phenotype of glioma cells through activating the Hedgehog pathway. *The Journal of pathology*. 234 (1). pp. 11-22.
- Yang, Z. 2014. *Functional studies of the Tdrd12-Exd1 complex in the piRNA pathway*.
- Yang, Z., Chen, K., Pandey, R.R., Homolka, D., Reuter, M., Janeiro, B.K.R., Sachidanandam, R., Fauvarque, M., McCarthy, A.A. & Pillai, R.S. 2016. PIWI slicing and EXD1 drive biogenesis of nuclear piRNAs from cytosolic targets of the mouse piRNA pathway. *Molecular cell*. 61 (1). pp. 138-152.
- Yang, B., O'Herrin, S.M., Wu, J., Reagan-Shaw, S., Ma, Y., Bhat, K.M., Gravekamp, C., Setaluri, V., Peters, N., Hoffmann, F.M., Peng, H., Ivanov, A.V., Simpson, A.J. & Longley, B.J. 2007. MAGE-A, mMage-b, and MAGE-C proteins form complexes with KAP1 and suppress p53-dependent apoptosis in MAGE-positive cell lines. *Cancer research*. 67 (20). pp. 9954-9962.
- Ye, S., Zhang, D., Cheng, F., Wilson, D., Mackay, J., He, K., Ban, Q., Lv, F., Huang, S., Liu, D. & Ying, Q.L. 2016. Wnt/beta-catenin and LIF-Stat3 signaling pathways converge on Sp5 to promote mouse embryonic stem cell self-renewal. *Journal of cell science*. 129 (2). pp. 269-276.
- Ying, M. & Chen, D. 2012. Tudor domain-containing proteins of *Drosophila melanogaster*. *Development, growth & differentiation*. 54 (1). pp. 32-43.
- Young, R.A. 2011. Control of the embryonic stem cell state. *Cell*. 144 (6). pp. 940-954.
- Yu, W., Briones, V., Lister, R., McIntosh, C., Han, Y., Lee, E.Y., Ren, J., Terashima, M., Leighty, R.M., Ecker, J.R. & Muegge, K. 2014. CG hypomethylation in Lsh^{-/-} mouse embryonic fibroblasts is associated with de novo H3K4me1 formation and altered cellular plasticity. *Proceedings of the National Academy of Sciences of the United States of America*. 111 (16). pp. 5890-5895.
- Zabala, M., Lobo, N., Qian, D., van Weele, L., Heiser, D., Clarke, M., Liu, H. & Lathia, J. 2016. *Overview: Cancer Stem Cell Self-Renewal*. Elsevier Cambridge, MA, USA: .

- Zafarana, G., Avery, S.R., Avery, K., Moore, H.D. & Andrews, P.W. 2009. Specific knockdown of OCT4 in human embryonic stem cells by inducible short hairpin RNA interference. *Stem cells*. 27 (4). pp. 776-782.
- Zeineddine, D., Hammoud, A.A., Mortada, M. & Boeuf, H. 2014. The Oct4 protein: more than a magic stemness marker. *American journal of stem cells*. 3 (2). pp. 74-82.
- Zeng, Y., He, Y., Yang, F., Mooney, S.M., Getzenberg, R.H., Orban, J. & Kulkarni, P. 2011. The cancer/testis antigen prostate-associated gene 4 (PAGE4) is a highly intrinsically disordered protein. *The Journal of biological chemistry*. 286 (16). pp. 13985-13994.
- Zeng, Y.X. & el-Deiry, W.S. 1996. Regulation of p21WAF1/CIP1 expression by p53-independent pathways. *Oncogene*. 12 (7). pp. 1557-1564.
- Zhang, Y., Fujita, N. & Tsuruo, T. 1999. p21Waf1/cip1 acts in synergy with bcl-2 to confer multidrug resistance in a camptothecin-selected human lung-cancer cell line. *International journal of cancer*. 83 (6). pp. 790-797.
- Zheng, Q.H., Ma, L.W., Zhu, W.G., Zhang, Z.Y. & Tong, T.J. 2006. p21Waf1/Cip1 plays a critical role in modulating senescence through changes of DNA methylation. *Journal of cellular biochemistry*. 98 (5). pp. 1230-1248.
- Zhou, J., Leu, N.A., Eckardt, S., McLaughlin, K.J. & Wang, P.J. 2014. STK31/TDRD8, a germ cell-specific factor, is dispensable for reproduction in mice. *PLoS One*. 9 (2). pp. e89471.
- Zou, C., Shen, J., Tang, Q., Yang, Z., Yin, J., Li, Z., Xie, X., Huang, G., Lev, D. & Wang, J. 2012. Cancer-testis antigens expressed in osteosarcoma identified by gene microarray correlate with a poor patient prognosis. *Cancer*. 118 (7). pp. 1845-1855.

Appendix A-1: TDRD12 variant sequences

Transcript variant 4,

XIVEKVEKFGLYGLAEKTLFHRVQVLEVNQKEDAWALDDILVEFIDEGRGTGLVTRDQLLHLPEHFHTLPPQAVEFIVCRVKPADNEIENWPKVTRYIHHKIVGKLHDAKVILALGNTVWIDPMVHITNLSSLKTSVIDYNVRAEILSMGMGIDNPEHIEQLKKLREDAKIPACEESLSQTPPRVTGTSPAQDQDHPSEEQGGQGTTPAEDAACLOSPQPEDTGAE
GGAESKTSSSENQKPE TLSGNTEGAFISRTAQPPPKSFHPQIKWFQKEDVVILKIRIRNVKDYKCQYLDRRVVFS
WVGDKFYLADELRGNIRKDDCQCVRNDEPVITLAKERREAWCHLLRQR

Transcript variant 2, = 1177 bases = 141 kDa

MLQLLVLKIEDPGCFWVVIKGCSPFLDHDVDYQKLNSAMNDFYNSTCQDIEIKPLTLEEGQVCVVYCEELKCWCR
AIVKSITSSADQYLAECFLVDFAKNIPVKSKNIRVVESFMQLPYRAKKFSLYCTKPVTLHIDFCRDSTDIVPAK
KWDNAAIQYFQNLKATTQVEARLCAVEEDTFEVYLYVTIKDEKVCVNDDLVAKNYACYSPTKNKNLDYLEKPR
LNIKSAPSFNKNLPALTLPMPFLOGKDVQGMEDSHGVNFPAQSLQHTWCKGIVGDLRPTATAQDKAVKCNMDSLR
DSPKDKSEKKHHHCISLKDNTNKRVESSVYWPARGITITADPDVPEASALSQKSNEKPLRLTEKKEYDEKNSCVKL
LQFLNPDLRADGISDLQQLQKLKGLQPPVVVLRNKIKPCLTIDSSPLSADLKKALQRNKFPGPSHTESYSWPP
ARGCDVVVISHCESNPLLYLLPVLTVLQGTGACYKSLPSRNGPLAVIVCPGWKKAQFIFELLGEYSMSSRPLHPVL
LTIGLHKEEAKNTKLPRGCDVIVTTPYSLRLRLACQSLLFLRLCHLILDEVEVLFLEANEQMFALDNLKKNIEV
EERESAPHQIVAVGVHWNKHIEHLIKEFMNDPYIVITAMEEAAALYGNVQQVVHLCLECEKTSSLLQALDFIPSOQ
KTLIFTCSVAETEIVCKVVESSIFCLKMHKEMIFNLQNVLEQWKKKLSSGSQIILALTDDCVPLLAITDATCVI
HFSFPASPKVFGGRLYCMSDHFHAEQGSAPAEQGDKKAKSVLLLTEKDASHAVGLRYLERADAKVPAELYEFTAG
VLEAKEDKKAGRPLCPYLKAFGFCCKDKRICPDRHRINPETDLPRKLSSQALPSFGYIKIIPFYILNATNYFGRIV
DKHMDLYATLNAEMNEYFKDSNKTIVEKVEKFGLYGLAEKTLFHRVQVLEVNQKEDAWALDDILVEFIDEGRGTGL
VTRDQLLHLPEHFHTLPPQAVEFIVCRVKPADNEIENWPKVTRYIHHKIVGKLHDAKVILALGNTVWIDPMVHIT
NLSSLKTSVIDYNVRAEILSMGMGIDNPEHIEQLKKLREDAKIPACEESLSQTPPRVTGTSPAQDQDHPSEEQGG
QGTTPAEDAACLOSPQPEDTGAE
GGAESKTSSSENQKPGGYLVFKRWLSSNR

Transcript variant 1

MLQLLVLKIEDPGCFWVVIKGCSPFLDHDVDYQKLNSAMNDFYNSTCQDIEIKPLTLEEGQVCVVYCEELKCWCR
AIVKSITSSADQYLAECFLVDFAKNIPVKSKNIRVVESFMQLPYRAKKFSLYCTKPVTLHIDFCRDSTDIVPAK
KWDNAAIQYFQNLKATTQVEARLCAVEEDTFEVYLYVTIKDEKVCVNDDLVAKNYACYSPTKNKNLDYLEKPR
LNIKSAPSFNKNLPALTLPMPFLOGKDVQGMEDSHGVNFPAQSLQHTWCKGIVGDLRPTATAQDKAVKCNMDSLR
DSPKDKSEKKHHHCISLKDNTNKRVESSVYWPARGITITADPDVPEASALSQKSNEKPLRLTEKKEYDEKNSCVKL
LQFLNPDLRADGISDLQQT

Transcript variant 5

XTLSGNTEGAFISRTAQPPPKSFHPQIKWFQKEDVVILKIRIRNVKDYKCQYLDRRVVFS
AWVGDKFYLADELRGNIRKDDCQCVRNDEPVITLAKERREAWCHLLRQRNPNVAFDFDHWEDCEEDSHFPKVVNSKNLPYTVTEVVED
SSSTSEDDDESEREGE

Transcript variant 6 = 1344 bases = 162 kDa

MLQLLVLKIEDPGCFWVVIKGCSPFLDHDVDYQKLNSAMNDFYNSTCQDIEIKPLTLEEGQVCVVYCEELKCWCR
AIVKSITSSADQYLAECFLVDFAKNIPVKSKNIRVVESFMQLPYRAKKFSLYCTKPVTLHIDFCRDSTDIVPAK
KWDNAAIQYFQNLKATTQVEARLCAVEEDTFEVYLYVTIKDEKVCVNDDLVAKNYACYSPTKNKNLDYLEKPR
LNIKSAPSFNKNLPALTLPMPFLOGKDVQGMEDSHGVNFPAQSLQHTWCKGIVGDLRPTATAQDKAVKCNMDSLR
DSPKDKSEKKHHHCISLKDNTNKRVESSVYWPARGITITADPDVPEASALSQKSNEKPLRLTEKKEYDEKNSCVKL
LQFLNPDLRADGISDLQQLQKLKGLQPPVVVLRNKIKPCLTIDSSPLSADLKKALQRNKFPGPSHTESYSWPP
ARGCDVVVISHCESNPLLYLLPVLTVLQGTGACYKSLPSRNGPLAVIVCPGWKKAQFIFELLGEYSMSSRPLHPVL
LTIGLHKEEAKNTKLPRGCDVIVTTPYSLRLRLACQSLLFLRLCHLILDEVEVLFLEANEQMFALDNLKKNIEV
EERESAPHQIVAVGVHWNKHIEHLIKEFMNDPYIVITAMEEAAALYGNVQQVVHLCLECEKTSSLLQALDFIPSOQ
KTLIFTCSVAETEIVCKVVESSIFCLKMHKEMIFNLQNVLEQWKKKLSSGSQIILALTDDCVPLLAITDATCVI
HFSFPASPKVFGGRLYCMSDHFHAEQGSAPAEQGDKKAKSVLLLTEKDASHAVGLRYLERADAKVPAELYEFTAG
VLEAKEDKKAGRPLCPYLKAFGFCCKDKRICPDRHRINPETDLPRKLSSQALPSFGYIKIIPFYILNATNYFGRIV
DKHMDLYATLNAEMNEYFKDSNKTIVEKVEKFGLYGLAEKTLFHRVQVLEVNQKEDAWALDDILVEFIDEGRGTGL
VTRDQLLHLPEHFHTLPPQAVEFIVCRVKPADNEIENWPKVTRYIHHKIVGKLHDAKVILALGNTVWIDPMVHIT
NLSSLKTSVIDYNVRAEILSMGMGIDNPEHIEQLKKLREDAKIPACEESLSQTPPRVTGTSPAQDQDHPSEEQGG
QGTTPAEDAACLOSPQPEDTGAE
GGAESKTSSSENQKPGGYLVFKRWLSSNRXTLSGNTEGAFISRTAQPPPKSFH
PQIKWFQKEDVVILKIRIRNVKDYKCQYLDRRVVFS
AWVGDKFYLADELRGNIRKDDCQCVRNDEPVITLAK
ERREAWCHLLRQRNPNVAFDFDHWEDCEEDSHFPKVVNSKNLPYTVTEVVEDSSSTSEDDDESEREGE

Appendix A-2: the sequence of anti-TDRD12 antibody matches TV1 and TV2

Transcript variant 2, = 1177 bases = 141 kDa

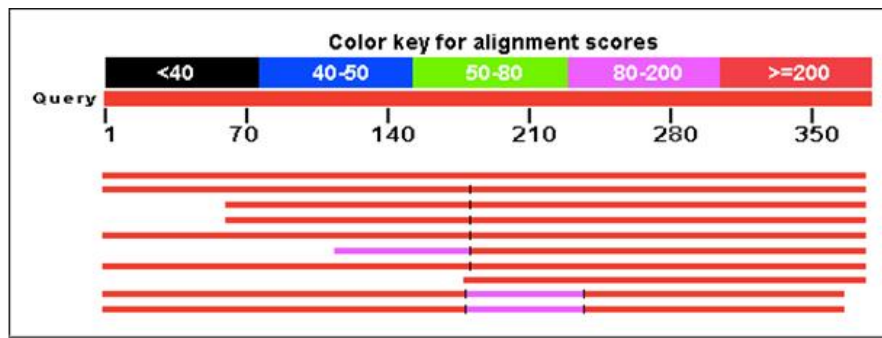
MLQLLVLKIEDPGCFWVVIKGCSPFLDHDVDYQKLSAMNDFYNSTCQDIEIKPLTLEEGQVCVVYCEELKCWCR
AIVKSITSSADQYLAECFLVDFAKNIPVKSKNIRVVVESFMQLPYRAKKFSLYCTKPVTLHIDFCRDSTDIVPAK
KWDNAAIQYFQNLLKATTQVEARLCAVEEDTFEVYLYVTIKDEKVCVNDDLVAKNYACYMSPTKNKNLDYLEKPR
LNIKSAPSFNKNLPALTLWPMFLQGKDVQGMEDSHGVNFPQSLQHTWCKGIVGDLRPTATAQDKAVKCNMDSLR
DSPKDKSEKKHHCTSLKDTNKRVESSVYWPARRGITYADPDVPEASALSQKSNEKPLRLTEKKEYDEKNSCVKL
LQFLNPDPLRADGISDLQQLQKLKGLQPPVVVLRNLIKPCLTIDSSPLSADLKKALQRNKFPGPSHTESYSWPPI
ARGCDVVVISHCESNPLLYLLPVLTVLQTGACYKSLPSRNGPLAVIVCPGWKKAQFIFELLGEYSMSSRPLHPVL
LTIGLHKEEAKNTKLPRGCDVIVTTPYSLLRLLACQSLLFLRLCHLILDEVEVLFLFLEANEQMFALDNFKKNIEV
EERESAPHQIVAVGVHWNKHIEHLIKEFMNDPYIVITAMEEAALYGNVQQVVHLCLECEKTSSLLQALDFIPSQQ
KTLIPTCSVAETEIVCKVVESSIFCLKMHKEMIFNLQNVLEQWKKKLSSGSQIILALTDDCVPLLAITDATCVI
HFSFPASPKVFGGRLYCMSDHFHAEQGSAPAEQGDKKAKSVLLLTEKDASHAVGVRLYLERADAKVPAELYEFTAG
VLEAKEDKKAGRPLCPYLKAFGFCKDKRICPDRHRINPETDLPRKLSSQALPSFGYIKIIPFYILNATNYFGRIV
DKHMDLYATLNAEMNEYFKDSNKTTEKVEKFGLYGLAEKTLFHRVQVLEVNQKEDAWALDDILVEFIDEGRITGL
VTRDQLLHLPEHFHTLPPQAVEFIVCRVKPADNEIENPNKVTRYIHHKIVGKLHDAKIVILALGNTVWIDPMVHIT
NLSSLKTSVIDYNVRAEILSMGMGIDNPEHIEQLKKLREDAKIPACEESLSQTPPRVTGTSPAQDQDHPSEEQGG
QGTTPAEDAACLQSPQPEDTGAEGGAESKTSSSENQKPGGYLVFKRWLSSNR

Transcript variant 1

MLQLLVLKIEDPGCFWVVIKGCSPFLDHDVDYQKLSAMNDFYNSTCQDIEIKPLTLEEGQVCVVYCEELKCWCR
AIVKSITSSADQYLAECFLVDFAKNIPVKSKNIRVVVESFMQLPYRAKKFSLYCTKPVTLHIDFCRDSTDIVPAK
KWDNAAIQYFQNLLKATTQVEARLCAVEEDTFEVYLYVTIKDEKVCVNDDLVAKNYACYMSPTKNKNLDYLEKPR
LNIKSAPSFNKNLPALTLWPMFLQGKDVQGMEDSHGVNFPQSLQHTWCKGIVGDLRPTATAQDKAVKCNMDSLR
DSPKDKSEKKHHCTSLKDTNKRVESSVYWPARRGITYADPDVPEASALSQKSNEKPLRLTEKKEYDEKNSCVKL
LQFLNPDPLRADGISDLQQT

Appendix A-3. EXDL1 sequencing results.

The PCR product for the EXDL1 gene was purified and sequenced. The sequence was aligned using Basic Local Aliment Search Tools (BLAST; NCBI <http://blast.ncbi.nlm.nih.gov/Blast.cgi>).



Homo sapiens exonuclease 3'-5' domain containing 1 (EXD1), transcript variant 2, mRNA
Sequence ID: [refNM_152596.3](#) Length: 3009 Number of Matches: 1

Range 1: 102 to 475 [GenBank](#) [Graphics](#)

▼ Next Match ▲ Previous Mat

Score	Expect	Identities	Gaps	Strand
680 bits(368)	0.0	373/375(99%)	1/375(0%)	Plus/Plus
Query 1	GCCGCCGGGTGGCGTGGAGCGTCCCGAGACTTCTTCCTTCATTGCAGCCCCCTCTTCTTTT	60		
Sbjct 102	GCCGCCGGGTGGCGTGGAGCGTCCCGAGACTTCTTCCTTCATTGCAGCCCCCTCTTCTTTT	161		
Query 61	AGTTCTAGACTGAAGTCGAGGAACGCACCACTTGGCTGAAATTAGGATCCAGATTTCCT	120		
Sbjct 162	AGTTCTAGACTGAAGTCGAGGAACGCACCACTTGGCTGAAATTAGGATCCAGATTTCCT	221		
Query 121	AAACTTCCTTCGGAGGAAAAAGCCCTAACGATGGAGGACAGTGAATTCCTAGCTTATGTG	180		
Sbjct 222	AAACTTCCTTCGGAGGAAAAAGCCCTAACGATGGAGGACAGTGAATTCCTAGCTTATGTG	281		
Query 181	GAAGTACTAGATGAAGTGAACAAGGCTCAGTGAGAGCAAAAGCATCTTCTGTTAGTCTA	240		
Sbjct 282	GAAGTACTAGATGAAGTGAACAAGGCTCAGTGAGAGCAAAAGCATCTTCTGTTAGTCTA	341		
Query 241	CATGCAGAAAGAACCTGGATGGAGAAAAATGAAAGTTGAAGACCTAAATGTATGTGAGCCT	300		
Sbjct 342	CATGCAGAAAGAACCTGGATGGAGAAAAATGAAAGTTGAAGACCTAAATGTATGTGAGCCT	401		
Query 301	GCTTCTCCTGCCCCCTGAAGCACCAGCTACCTCTCTGCTGAATGACCTCAAGTACAGCCCA	360		
Sbjct 402	GCTTCTCCTGCCCCCTGAAGCACCAGCTACCTCTCTGCTGAATGACCTCAAGTACAGCCCA	461		
Query 361	TCAGAGGATGGAGGA	375		
Sbjct 462	TCAGAGGAAG-AGGA	475		

Appendix B.

Appendix B-1. Data sheet information for normal tissues microarray samples

Cooperative Human Tissue Network (CHTN) Normal Tissue Microarray: Version CHTN_Norm2



Purpose:

To provide researchers with a tissue microarray of formalin fixed paraffin embedded samples that includes most of the cell types present in the human body.

Tissue samples:

Most samples are of normal, non-neoplastic adult tissue obtained from surgical resection specimens, fixed within one hour of removal from the donors. The fixative used is buffered zinc formaldehyde (3.7% formaldehyde) (Z-fix, Anatech, LTD., Battle Creek, MI).

Exceptions:

Parathyroid - tissue is from a hyperplastic parathyroid gland due to the limiting size and limited availability of normal parathyroid tissue.

Central nervous system tissue (cerebral cortex and white matter) - all obtained from autopsy specimens within 36 hours of death.

Pancreas, Islet Cell - tissue is from a pancreatic islet cell tumor because islet cells are not readily samples. Pancreatic islet cell tumor may serve as a substitute for this tissue.

TMA design:

Each tissue type was sampled multiple times with 1 mm needle cores in the original array design. The CHTN_Norm2 TMA series was constructed as 3 replicate TMA blocks, designated CHTN_Norm2A, CHTN_Norm2B, CHTN_Norm2C. Each array block is serially sectioned at 4 micron thickness. The histologic sections are placed on charged glass slides (Fisher Plus). At intervals, sections are stained and examined by a pathologist for quality assurance (QA) purposes. The desired tissue type must be present on at least one tissue spot to be scored as adequate. The number of tissue spots in which the desired tissue types resides will vary from section to section.

Quality assurance procedures:

Due to the variability inherent to tissue samples and histologic methodology, most of the tissue microarray sections will not contain all of the tissue spots in the original design. The number of tissue spots in which the desired tissue types resides will vary from section to section.

THE TISSUE SAMPLES HAVE BEEN ANONYMIZED, NO FURTHER DATA ON THE DONORS IS AVAILABLE OTHER THAN THAT FOUND IN THE ACCOMPANYING GUIDE SHEETS.

Guide sheets in Microsoft Excel format and H&E stained whole-slide images representative of the array are available on the CHTN TMA website: <http://chtn.sites.virginia.edu/>.

If despite our efforts you are missing a target tissue type on your array sections that you desire, please contact the CHTN Mid-Atlantic Division at (434) 924-9879.

Appendix B-2.

The properties of the samples included in the tissue microarrays employed in this study are shown in Table B-1, Table B-2, Table B-3 and Table B-4.

Table B-1 Normal tissue microarray samples

Tissue	Code	Age	Sex
BREAST			
breast, epithelium	BE	47	F
CARDIOVASCULAR			
aorta, smooth muscle	CASM	44	M
heart, myocardium	CHM	51	M
ENDOCRINE			
adrenal gland, cortex	EAGC	51	M
adrenal gland, medulla	EAGM	51	F
parathyroid (hyperplastic)	EPAD	45	F
thyroid	ET	38	F
GASTROINTESTINAL TRACT			
esophagus, squamous mucosa	GIE	58	M
gastric mucosa, antral	GIGA	59	F
gastric mucosa, oxyntic	GIGO	53	F
small intestine, mucosa	GISI	85	F
colon, mucosa	GIC	66	M
anus, mucosa	GIA	39	F
GENITAL TRACT, FEMALE			
ectocervix	GFEC	43	F
endocervix	GFEN	43	F
endometrium, secretory	GFES	45	F
fallopian tube	GFFT	49	F
ovary, 1° oocytes (Block A)	GFOO	54	F
ovary, 1° oocytes (Block B)	GFOO	54	F
ovary, 1° oocytes (Block C)	GFOO	42	F
ovary, corpus luteum	GFOCL	49	F
ovary, mesothelium (Block A)	GFOE	44	F
ovary, mesothelium (Block B)	GFOE	49	F
ovary, mesothelium (Block C)	GFOE	49	F
GENITAL TRACT, MALE			
seminiferous tubules	GMST	36	M
epididymis	GME	66	M
seminal vessicle	GMSV	44	M
prostate	GMP	52	M
HEPATIC & PANCREATOBILIARY			
gallbladder	HPG	56	M
pancreas, islet cell	HPI	70	M
liver	HPL	76	M
pancreas	HPP	67	F
LYMPHOID			
lymph node	LLN	6	F
mucosa assoc. lymphoid tissue, appendix	LMALT	66	F

spleen	LS	80	M
thymus	LT	48	F
NERVOUS SYSTEM			
cerebral cortex	NCCC	64	M
white matter (subcortical)	NCWM	33	F
peripheral nerve	NPPN	50	F
ORAL, SALIVARY & NASAL			
salivary gland (parotid)	OSSG	40	M
tonsil, squamous epithelium	OSTSE	24	M
PLACENTA			
amniotic membrane	PAM	21	F
placenta, villi	PV	24	F
RESPIRATORY TRACT			
alveoli	RA	43	M
bronchus, epithelium	RBE	52	M
SKIN			
skin, squamous epithelium	SSE	61	F
SOFT TISSUE			
cartilage, articular	STCA	43	M
skeletal muscle	STSKM	48	F
smooth muscle, intestine	STSMI	82	F
smooth muscle, uterus	STSMU	56	F
synovium	STS	30	F
UROLOGICAL TRACT			
kidney, cortex	UKC	47	F
kidney, medulla	UKM	61	M
bladder, transitional epithelium	UBTE	49	M

Appendix B-3. Tumour microarray samples

Position	No.	Age	Sex	Organ/Anatomic Site	Pathology diagnosis	TNM	Stage	Type	Tissue ID.
A1	1	21	F	Uterus	Normal endometrium tissue	-	-	normal	Fur06N024
A2	2	21	F	Uterus	Normal endometrium tissue	-	-	normal	Fur06N023
A3	3	18	F	Uterus	Normal endometrium tissue	-	-	normal	Fur06N010
A4	4	35	F	Uterine cervix	Cancer adjacent normal cervical canals tissue	-	-	NAT	Fdu050245
A5	5	72	F	Uterine cervix	Cancer adjacent normal cervix tissue	-	-	NAT	Fdu031770
A6	6	65	F	Uterine cervix	Cancer adjacent normal cervix tissue (smooth muscle and blood vessel)	-	-	NAT	Fdu031772
A7	7	35	M	Skeletal muscle	Normal skeletal muscle tissue	-	-	normal	Doc11N001
A8	8	49	F	Skeletal muscle	Cancer adjacent normal skeletal muscle tissue	-	-	NAT	Fmg100041
A9	9	47	M	Lingual gland	Normal lingual gland tissue	-	-	normal	Doc07N027
B1	10	50	F	Skin	Cancer adjacent normal skin tissue	-	-	NAT	Kin040180
B2	11	37	M	Skin	Normal skin tissue of foot	-	-	normal	Kin08N012
B3	12	18	F	Skin	Normal skin tissue of head	-	-	normal	Kin06N010
B4	13	33	M	Nerve	Normal peripheral nerve tissue	-	-	normal	Scf07N026
B5	14	35	M	Nerve	Normal peripheral nerve tissue	-	-	normal	Scf08N001
B6	15	35	M	Nerve	Normal peripheral nerve tissue	-	-	normal	Scf08N006
B7	16	22	M	Lung	Normal mesothelial tissue	-	-	normal	Rln05N009
B8	17	27	F	Cardiac pericardium	Normal mesothelial tissue	-	-	normal	Apr07N025
B9	18	47	M	Lung	Normal lung and mesothelium tissue	-	-	normal	Rln06N004
C1	19	41	F	Cerebrum	Glioblastoma	-	-	malignant	Nct060300
C2	20	65	F	Cerebrum	Atypical meningioma	-	-	malignant	Nct020243
C3	21	15	F	Cerebrum	Malignant ependymoma	-	-	malignant	Nct030285
C4	22	39	F	Cerebrum	Oligodendroglioma	-	-	malignant	Nct090046
C5	23	63	F	Ovary	Serous adenocarcinoma	T2N0M0	II	malignant	Fov010785
C6	24	29	F	Ovary	Adenocarcinoma	T3N0M0	III	malignant	Fov020520
C7	25	12	F	Pancreas	Islet cell tumor	-	-	malignant	Dpa010125
C8	26	64	M	Pancreas	Adenocarcinoma	T3N0M0	II	malignant	Dpa020601
C9	27	32	M	Testis	Seminoma	T1N0M0	I	malignant	Mtt020007
D1	28	30	M	Testis	Embryonal carcinoma	T2N0M0	I	malignant	Mtt030413
D2	29	33	F	Thyroid	Medullary carcinoma	T3N0M0	II	malignant	Etg030492
D3	30	46	F	Thyroid	Papillary carcinoma	T4N0M0	III	malignant	Etg010040
D4	31	43	F	Breast	Invasive ductal carcinoma	T2N1M0	IIA	malignant	Fmg030341
D5	32	48	F	Breast	Intraductal carcinoma	TisN0M0	0	malignant	Fmg020962
D6	33	58	F	Breast	Invasive ductal carcinoma	T2N1M0	IIB	malignant	Fmg010766
D7	34	21	M	Spleen	Diffuse B-cell lymphoma	-	-	malignant	Is p030767
D8	35	61	M	Lung	Small cell undifferentiated carcinoma	T2N0M0	I	malignant	Rln020010
D9	36	66	M	Lung	Squamous cell carcinoma	T2N0M0	I	malignant	Rln020752
E1	37	42	M	Lung	Adenocarcinoma	T2N0M0	IB	malignant	Rln010049
E2	38	67	M	Esophagus	Neuroendocrine carcinoma	T2N1M0	IIA	malignant	Des040713
E3	39	63	M	Esophagus	Adenocarcinoma	T3N0M0	IIA	malignant	Des010320
E4	40	73	F	Stomach	Signet-ring cell carcinoma	T2N1M0	II	malignant	Dst010223
E5	41	64	F	Intestine	Adenocarcinoma	T4N0M0	II	malignant	Din010529
E6	42	71	F	Intestine	Stromal sarcoma	T2N0M0	IIB G2	malignant	Din050375
E7	43	72	F	Colon	Adenocarcinoma	T4N0M0	IIB	malignant	Dco020006
E8	44	54	M	Colon	Interstitialoma	T2N0M0	IIB G2	malignant	Aac050109
E9	45	36	M	Rectum	Adenocarcinoma	T2N0M0	I	malignant	Dre020645
F1	46	63	F	Rectum	Moderate malignant interstitialoma	T2N0M0	IIB G2	malignant	Dre030597
F2	47	55	F	Liver	Hepatocellular carcinoma	T3N0M0	IIIA	malignant	Dlv010154
F3	48	17	F	Liver	Hepatoblastoma	-	-	malignant	Dlv040278
F4	49	60	M	Kidney	Clear cell carcinoma	T2N0M0	II	malignant	Ukn060084
F5	50	80	M	Prostate	Adenocarcinoma (Gleason grade:4 Gleason score:4+5)	T4N1M1c	IV	malignant	Mpr020382
F6	51	77	M	Prostate	Adenocarcinoma (Gleason grade:3 Gleason score:3+4)	T2N0M0	II	malignant	Mpr040215
F7	52	33	F	Uterus	Leiomyoma	T2N0M0	IB G1	malignant	Fur010182
F8	53	50	F	Uterus	Adenocarcinoma	T1bN0M0	IB	malignant	Fur010190
F9	54	50	F	Uterus	Clear cell carcinoma of endometrium	T2bN0M0	IIB	malignant	Fur040227
G1	55	36	F	Uterine cervix	Squamous cell carcinoma	T1bN0M0	IB	malignant	Fdu020494
G2	56	49	F	Uterine cervix	Squamous cell carcinoma	T2N0M0	II	malignant	Fdu031255
G3	57	20	F	Striated muscle	Embryonal rhabdomyosarcom of left leg	T1aN0M0	IA	malignant	Srm050111
G4	58	70	F	Rectum	Malignant melanoma	T4N0M0	IIIB	malignant	Dre030434
G5	59	61	M	Skin	Basal cell carcinoma of left face	T2N0M0	II	malignant	Kin100009
G6	60	64	M	Skin	Squamous cell carcinoma of chest wall	T3N0M0	II	malignant	Kin020141
G7	61	59	M	Back	Neurofibroma	-	-	malignant	Sst040384
G8	62	1	F	Retroperitoneum	Neuroblastoma	T2bN0M0	IIIB	malignant	Apo050228
G9	63	60	M	Abdominal cavity	Malignant mesothelioma	T2N0M0	IIB G2	malignant	Aac060028
H1	64	49	M	Mediastinum	Diffuse B-cell lymphoma of lymph node	-	-	malignant	Amd030215
H2	65	50	F	Lymph node	Diffuse B cell lymphoma of right thigh	-	-	malignant	Ily020164
H4	67	19	F	Pelvic cavity	Anaplastic large cell lymphoma	-	-	malignant	Apc030192
H5	68	76	M	Bladder	Transitional cell carcinoma	T3aN0M0	III	malignant	Ubd010014
H6	69	62	M	Bladder	Low grade malignant leiomyosarcoma	T2bN0M0	IB G1	malignant	Ubd020317
H7	70	7	F	Bone	Osteosarcoma of right femur	T2N0M0	IIIB	malignant	Lbn020013
H8	71	48	F	Retroperitoneum	Spindle cell rhabdomyosarcoma	T2N0M0	IIB	malignant	Apo040221
H9	72	60	F	Smooth muscle	Moderate malignant leiomyosarcoma of left buttock	T2bN0M0	IIIB G3	malignant	Ssm040020

FDA808ci-2: Multiple organ tumor tissue array with cancer adjacent normal tissue as contral, including TNM, clinical stage and pathology grade, 72 cases/ 72 cores.

For more information, visit: https://www.biomax.us/tissue-arrays/Multiple_Organ/FDA808c-2

Appendix B-4. Ovarian carcinoma samples

Code	Age	Stage	Grade	Code	Age	Stage	Grade
SPC-1	62	pT3c [IIIC]	high	MAC-6	43	pT1a	low
SPC-2	64	pT3b [IIIB]	high	MAC-7	29	pT2b	high
SPC-3	64	pT2a [IIA]	high	MAC-8	62	pT3c	moderate
SPC-4	62	pT3c [IIIC]	high	MAC-9	63	pT1a	low
SPC-5	65	pT3a [IIIC]	high	MAC-10	66	n/a	moderate
SPC-6	74	pT3c [IIIC]	high	SBT-1	57	pT1c	low
SPC-7	42	pT2a [IIA]	high	SBT-2	72	pT1a	low
SPC-8	45	pT3c [IIIC]	high	SBT-3	46	pT1c	low
SPC-9	52	pT3c [IIIC]	high	SBT-4	46	pT3a	low
SPC-10	66	pT3c [IIIC]	high	SBT-5	49	pT1c	low
SPC-11	64	pT3b [IIIB]	high	SBT-6	39	pT1C	low
SPC-12	71	pT3c [IIIC]	high	MBT-1	70	pT1C	low
EAC-1	58	pT1c [IC]	moderate	MBT-2	50	pT1c	low
EAC-2	59	pT2c [IIC]	moderate	MBT-3	13	n/a	low
EAC-3	67	pT1a [IA]	low	MBT-4	67	pT1c	low
EAC-4	52	pT1c [IC]	low	MBT-5	54	pT1a	low
EAC-5	51	pT1c [IC]	low	MBT-6	49	pT1a	low
EAC-6	52	pT2b [IIB]	high	NPE-1	57	N/A	N/A
EAC-7	65	pT1a [IA]	high	NPE-2	43	N/A	N/A
EAC-8	42	pT3c [IIIC]	high	NPE-3	43	N/A	N/A
EAC-9	49	pT1a	high	NPE-4	46	N/A	N/A
EAC-10	76	pT1c	moderate	NPE-5	41	N/A	N/A
EAC-11	70	pT1a	moderate	NPE-6	57	N/A	N/A
EAC-12	83	pT1c	low	NFT-1	67	N/A	N/A
CCC-1	62	pT1c [IC]	high	NFT-2	61	N/A	N/A
CCC-2	77	pT3a [IIIA]	high	NFT-3	47	N/A	N/A
CCC-3	73	pT2C [IIC]	high	NFT-4	61	N/A	N/A
CCC-4	72	pT1a [IA]	high	NFT-5	75	N/A	N/A
CCC-5	42	pT3a [IIIA]	high	NFT-6	64	N/A	N/A
CCC-6	27	pT3c [IIIC]	high	SCA-1	50	N/A	N/A
CCC-7	46	pT3c [IIIC]	high	SCA-2	71	N/A	N/A
CCC-8	68	pT3a [IIIA]	high	SCA-3	76	N/A	N/A
CCC-9	88	pT1c [IC]	high	SCA-4	53	N/A	N/A
CCC-10	39	pT1a [IA]	high	SCA-5	46	N/A	N/A
CCC-11	76	pT2c [IIC]	high	SCA-6	74	N/A	N/A
CCC-12	50	pT1c	high	MCA-1	51	N/A	N/A
MAC-1	20	pT1a	low	MCA-2	42	N/A	N/A
MAC-2	63	pT1a	high	MCA-3	60	N/A	N/A
MAC-3	57	pT1a	high	MCA-4	76	N/A	N/A
MAC-4	40	pT1a	high	MCA-5	42	N/A	N/A
MAC-5	69	pT3a	not assigned	MCA-6	25	N/A	N/A

SP: serous papillary carcinoma, well to moderately differentiate.

CC: clear cell carcinoma.

EC: endometrioid adenocarcinoma.

MC: mucinous adenocarcinoma.

PD: poorly differentiated serous papillary carcinoma, or undifferentiated carcinoma.

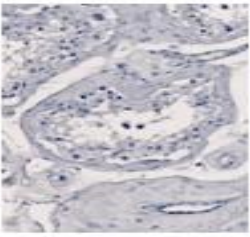
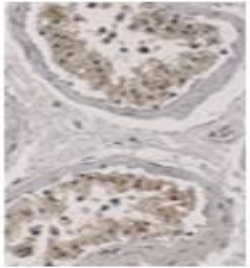
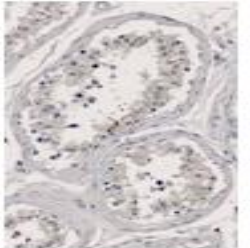
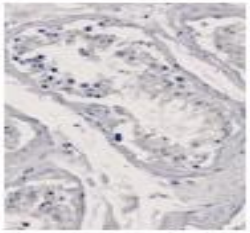
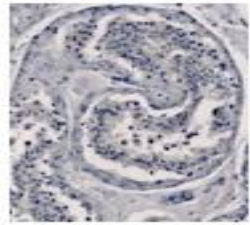
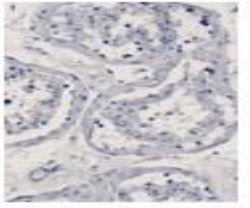

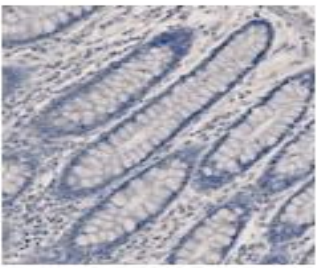
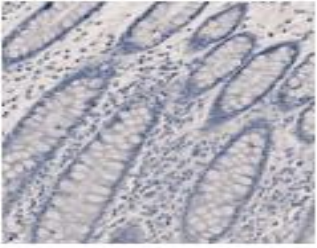
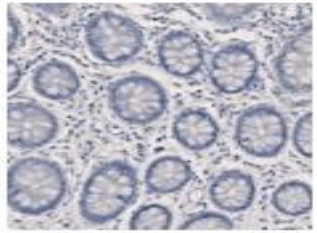
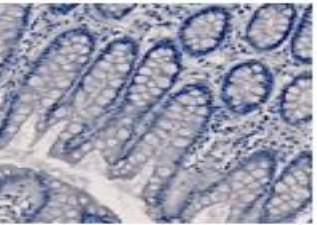

To see more details; <https://chtn.sites.virginia.edu/tissue-microarrays#OvCa2>

Appendix B-5. CHTN_CRC2 - Colorectal Carcinoma Progression

TMA code	Diagnosis	Location in Colon†	Gender	Age	pT stage*	pN stage*	Grade*	Metastatic Site
Bnc1	diverticulosis	D	M	68	NA	NA		
Bnc2	diverticulosis	D	M	66	NA	NA		
Bnc3	diverticulosis	D	F	75	NA	NA		
Bnc4	diverticulosis	D	F	47	NA	NA		
Bnc5	diverticulosis	D	F	48	NA	NA		
Bnc6	diverticulosis	D	M	39	NA	NA		
Bnc7	diverticulosis	D	M	44	NA	NA		
Bc1	mucosa adjacent to adenocarcinoma	D	F	68	NA	NA		
Bc2	mucosa adjacent to adenocarcinoma	D	M	68	NA	NA		
Bc3	mucosa adjacent to adenocarcinoma	D	M	56	NA	NA		
Bc4	mucosa adjacent to adenocarcinoma	P	M	61	NA	NA		
Bc5	mucosa adjacent to adenocarcinoma	P	M	72	NA	NA		
Bc6	mucosa adjacent to adenocarcinoma	D	F	60	NA	NA		
Bc7	mucosa adjacent to adenocarcinoma	D	M	71	NA	NA		
Bi1	ulcerative colitis	D	M	71	NA	NA		
Bi2	ulcerative colitis	D	M	10	NA	NA		
Bi3	ulcerative colitis	D	F	54	NA	NA		
Bi4	ulcerative colitis	D	F	19	NA	NA		
Bi5	ulcerative colitis	D	M	40	NA	NA		
Bi6	ulcerative colitis	P	F	53	NA	NA		
Bi7	ulcerative colitis	P	F	25	NA	NA		
AS1	tubulovillous adenoma	D	F	47	NA	NA		
AS2	tubulovillous adenoma	U	M	71	NA	NA		
AS3	tubular adenoma	D	F	57	NA	NA		
AS4	tubular adenoma	P	F	60	NA	NA		
AS5	tubular adenoma	U	M	50	NA	NA		
AS6	tubulovillous adenoma	U	F	70	NA	NA		
AS7	tubulovillous adenoma	D	M	71	NA	NA		
AL1	tubulovillous adenoma	P	M	63	NA	NA		
AL2	tubulovillous adenoma	D	F	71	NA	NA		
AL3	tubulovillous adenoma	D	M	50	NA	NA		
AL4	tubulovillous adenoma	D	F	83	NA	NA		
AL5	tubulovillous adenoma	D	M	54	NA	NA		
AL6	tubulovillous adenoma	D	F	78	NA	NA		
AL7	tubulovillous adenoma	U	F	45	NA	NA		
CE1	adenocarcinoma	P	F	60	pT1	pN0	G3	
CE2	adenocarcinoma	D	F	29	pT2	pN0	G2	
CE3	adenocarcinoma	D	M	68	pT2	pN1a	G2	
CE4	adenocarcinoma	D	M	70	pT2	pN0	G2	
CE5	adenocarcinoma	U	M	51	pT1	pN0	G1	
CE6	adenocarcinoma	P	M	62	pT2	pNx	G2	
CE7	adenocarcinoma	P	M	66	pT2	pN0	G2	
CL1	adenocarcinoma	D	F	68	pT4b	pN0	G2	
CL2	adenocarcinoma	P	M	71	pT3	pN2	G2	
CL3	adenocarcinoma	D	M	56	T4b	pN1a	G2	
CL4	adenocarcinoma	D	M	61	pT3	pN2b	G2	
CL5	adenocarcinoma	D	M	72	pT3	pN0	G2	
CL6	adenocarcinoma	P	F	70	pT3	pN1b	G3	
CL7	adenocarcinoma	P	F	69	pT3	pN1a	G2	
CL8	adenocarcinoma	D	M	61	pT4b	pN1a	G3	
LN1	metastatic adenocarcinoma	NA	M	71	pT3	pN2b	G2	lymph node
LN2	metastatic adenocarcinoma	NA	M	61	pT3	pN2b	G2	lymph node
LN3	metastatic adenocarcinoma	NA	M	68	pT2	pN1a	G2	lymph node
LN4	metastatic adenocarcinoma	NA	M	56	T4b	pN1a	G2	lymph node
LN5	metastatic adenocarcinoma	NA	M	48	pT3	pN2b	G2	lymph node
LN6	metastatic adenocarcinoma	NA	M	51	pT4b	pN1a	G3	lymph node
LN7	metastatic adenocarcinoma	NA	M	58	pT4b	pN1b	G2	lymph node
M1	metastatic adenocarcinoma	NA	M	59	pT4b	pN2a	G3	liver
M2	metastatic adenocarcinoma	NA	F	61	pT4a	pN1a	G2	pelvic soft tissue
M3	metastatic adenocarcinoma	NA	F	71	pT4b	pN1b	G3	liver
M4	metastatic adenocarcinoma	NA	M	52	pT4b	pNX	G2	liver
M5	metastatic adenocarcinoma	NA	M	58	pT4b	pNx	G2	liver
M6	metastatic adenocarcinoma	NA	M	48	pT3	pN2b	G2	liver
M7	metastatic adenocarcinoma	NA	M	52	pT3	pN1	G2	liver

†P = proximal (cecum, ascending, transverse colon), D = distal (descending, sigmoid colon & rectum), U = unknown, NA = not applicable. (<https://chtn.sites.virginia.edu/chtn-crc2>).

Appendix C.

	Negative	1/250	1/500	1/750	1/1000	1/2000
Testis						
Colon						

9.1 IHC was conducted to optimise anti-LKAAE/AR1 antibody (ab 108142) using different dilutions from negative, 1/250, 1/500, 1/750, 1/1000 and 1/2000.