ISOLATION AND CHARACTERISATION OF COLON CANCER STEM CELLS AND THE EFFECTS OF EPIGENETIC MODULATION ON PLURIPOTENT MARKERS

Brenda Lee Milner

A thesis submitted to the Faculty of Health Sciences, University of the Witwatersrand, in fulfilment of the requirements for the degree of

Doctor of Philosophy

Johannesburg, 2014

Declaration

I, Brenda Lee Milner declare that this thesis is my own work. It is being submitted for the degree of Doctor of Philosophy in the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at this or any other University.

…….

18th day of November, 2014

Dedication

This thesis is dedicated to my parents

Carol and Michael Lewin, my brother

Russell Milner and my late father Mervyn Milner.

In memory of my father

Mervyn Milner

1960-2009

Peer reviewed publications and presentations arising from this study

Oral Presentations

Milner, B.L., Penny, C.B., Gibbon, V. & Ruff, P. 2012. Isolation and characterisation of human colorectal cancer stem cells. *Microscopy Society of Southern Africa*, vol. 42, 43 p. Cape Town, South Africa, December.

Milner, B.L., Penny, C.B., Gibbon, V. & Ruff, P. 2010. Isolation and characterisation of colon cancer stem cells and the effects of epigenetic modulation on pluripotent markers. University of the Witwatersrand, Faculty of Health Sciences, Research Day, September.

Poster Presentations

Milner, B.L., Penny, C.B., Gibbon, V. & Ruff, P. 2012. Isolation and characterisation of colon cancer stem cells and the effects of epigenetic modulation on pluripotent markers. University of the Witwatersrand, Faculty of Science and Health Sciences, Molecular Biosciences Research Thrust Research Day, December.

Milner, B.L., Penny, C.B., Gibbon, V. & Ruff, P. 2012. Isolation and characterisation of colon cancer stem cells and the effects of epigenetic modulation on pluripotent markers. University of the Witwatersrand, Faculty of Health Sciences, Research Day, September.

Milner, B.L., Penny, C.B., Gibbon, V. & Ruff, P. 2012. Isolation and characterisation of colon cancer stem cells and the effects of epigenetic modulation on pluripotent markers. SASBMB FASBMB Congress, Champagne Sports Resort, Drakensburg, South Africa, January-February.

Milner, B.L., Penny, C.B., Gibbon, V. & Ruff, P. 2011. Isolation and characterisation of colon cancer stem cells and the effects of epigenetic modulation on pluripotent markers. SASCRO SASMO Congress, Sun City Resort, South Africa, August.

Milner, B.L., Penny, C.B., Gibbon, V. & Ruff, P. 2010. Isolation and characterisation of colon cancer stem cells and the effects of epigenetic modulation on pluripotent markers. The American Society for Cell Biology: 50th Annual Meeting, Philadelphia, USA, December.

Milner, B.L., Penny, C.B., Gibbon, V. & Ruff, P. 2010. Isolation and characterisation of colon cancer stem cells and the effects of epigenetic modulation on pluripotent markers. University of the Witwatersrand, Faculty of Science and Health Sciences, Molecular Biosciences Research Thrust: Health for Africa, November.

Abstract

Colorectal cancer has a 9.8% cumulative incidence rate, making it the third most common cancer in the Western world. Despite a 50-60% response rate in patients to current cancer therapies, drug resistance and tumour relapse remain a concern. While current therapies reduce the tumour mass, they possibly fail to eradicate a unique population of pluripotent tumour resident cells. These cells, known as cancer stem cells, may have similar properties of self-renewal and proliferation to embryonic and adult stem cells, as they also express a number of key pluripotent transcription factors, including amongst others, NANOG, OCT3/4 and SOX2. Furthermore, since discreet groups of such stem cells are proposed to essentially drive tumourigenesis, they present as potential novel targets for cancer therapy. This study aimed to isolate a putative CSC population from the advanced colon adenocarcinoma cell lines HT29 and DLD1 and to assess the therapeutic effects of the epigenetic drugs Valproic acid and Zebularine on pluripotent gene expression.

As cancer stem cells may be isolated from general tumour cell populations due to their expression of unique cell surface markers, this study began with an investigation of the existence of putative cancer stem cells in human colorectal adenocarcinoma cell lines. Using flow cytometry, the surface expression of the stem cell-specific biomarkers, CD133 and EpCAM were evaluated in each of the SW1116, HT29 and DLD1 cell lines; these being representative of early-, mid- and late-stage colon adenocarcinomas, respectively. In the SW1116 cell line, 88.70 percent of cells were $EpCAM^+$, in the HT29 cell line, 94.10 percent of cells were $EpCAM^+$, and in the DLD1 cell line, 89.50 percent of cells were $EpCAM^+$. Moreover, CD133 was expressed together with EpCAM within a subpopulation of each cell line: 2.33 percent of cells were $CD133^+/EpCAM^+$ within the SW1116 cell line; 5.11 percent of cells were $CD133^+/EpCAM^+$ within the HT29 cell line; while the DLD1 cell line contained some 10.30 percent of $CD133^+/EpCAM^+$ expressing cells. These results suggest that the raised frequency of CD133 expression may well be associated with increased tumour stage.

It was then of particular interest to evaluate the *in vitro* growth properties and clonogenic potential of $CD133^+/EpCAM^+$ cells derived from the more advanced HT29 and DLD1

cell lines. To accomplish this, magnetic cell separation was used to isolate a pure population of $CD133^+/EpCAM^+$ cells from each of these cell lines ($CD133^+/EpCAM^+$ cells were not isolated from the SW1116 cell line, as this represents an early stage, non-metastatic colon adenocarcinoma). A 98.60 and 98.00 percent purification of $CD133^+/EpCAM^+$ cells was obtained from the HT29 and DLD1 cell lines, respectively. CD133 was shown to be an important surface marker to isolate putative colorectal cancer stem cells, as an *in vitro* sphere formation assay showed that 21 percent of cells from each of the cell lines had clonogenic potential. Moreover, using indirect immunofluorescence, confocal microscopy localised CD133 and EpCAM expression to the cell membrane and within the cytoplasm of both HT29 and DLD1 derived cancer stem cells. While the cytoplasmic expression of CD133 may indicate cell invasiveness, it is nevertheless cell surface expression of both CD133 and EpCAM that are considered to be markers of putative cancer stem cells.

As epigenetic mechanisms such as methylation of gene promoters and histone activity are involved in the regulation of stem cell transcriptional programs and in the regulation of self-renewal and differentiation, it was of particular interest to evaluate gene and protein expression of the pluripotency associated transcription factors, NANOG, OCT3/4 and SOX2 in the $CD133^+/EpCAM^+$ HT29 and DLD1 derived cancer stem cells. These cells were treated with varying concentrations of a histone deacetylase inhibitor, Valproic acid and a DNA methyltransferase inhibitor, Zebularine, respectively; and their modulatory effects on protein expression of NANOG, OCT4 and SOX2 were assessed by confocal microscopy and on gene expression by quantitative real-time PCR analysis. Confocal microscopy showed that 2.5mM and 5mM Valproic acid up-regulated nuclear NANOG and OCT4, respectively in HT29 putative cancer stem cells; whereas 250µM and 500µM Zebularine up-regulated nuclear NANOG equally, and 500µM Zebularine up-regulated nuclear OCT4 in these cells ($p \leq 0.05$). The modulatory effects of these drugs on nuclear NANOG and OCT4 expression in DLD1 putative cancer stem cells were inconclusive ($p > 0.05$). Interestingly, nuclear SOX2 protein expression was unaffected by either drug in both HT29 and DLD1 derived cancer stem cells. Statistical analyses of *NANOG* and *OCT4* gene expression data indicated no significant change in mRNA transcript levels. Surprisingly, *SOX2* mRNA transcripts were undetectable in both HT29 and DLD1

putative cancer stem cells. A subsequent study showed that its absence was characteristic of these cells. Thus, a mechanistic model was proposed, wherein particular post-translational modifications of SOX2 regulate the negative feed-back loop in HT29 and DLD1 putative cancer stem cells.

In summary, Valproic acid and Zebularine modulate nuclear NANOG and OCT4 protein expression in HT29 putative cancer stem cells. As altered expression of either transcription factor is reported to induce differentiation in embryonic and adult stem cells, this similarly may induce differentiation in the HT29 derived cancer stem cells. Further research is needed to confirm whether Valproic acid and Zebularine induce differentiation, as these drugs may sensitise these cells to current anti-cancer therapies and possibly prevent tumour relapse.

Acknowledgements

I wish to thank my parents Carol and Michael Lewin, and my brother Russell Milner for their continuous support throughout my academic journey. I also wish to thank my father Mervyn Milner, who is no longer with me. Thank you all for providing me with the opportunity to follow my dreams, for your endless love and support, and for always believing in me.

I wish to thank my supervisors Dr Clem Penny, Dr Victoria Gibbon and Prof Paul Ruff for their valued support and guidance along this journey.

I wish to thank Prof Saraladevi Naicker, Dr Raquel Duarte, Therése Dix-Peek, Dr Caroline Dickens, Prof Margot Hosie, Marisha Meyer, Patti Kay, Jason Hemingway and Tanya Augustine for their valued assistance with my research. I wish to thank my best friend and colleague Ezio Fok for creating some of the stunning figures presented in this thesis, as well as for his valued support. I also wish to thank my other colleagues Natasia Kruger, Andrea Papadopoulos, Thandiwe Msibi, Nicole Crawford, Aadilah Omar, Kiashanee Moodley, Natanya Moodley, Drishna Govan, Lerato Mpye, Alex Kasembeli and Aleksandra Jovanovic for their friendship and support.

I wish to thank the following agencies for their financial support: Medical Research Council (MRC), National Research Foundation (NRF), Faculty Research Committee of Health Sciences (FRC), Wits Health Consortium, Cancer Association of South Africa (CANSA) and the University of the Witwatersrand for awarding me the Postgraduate Merit Award and Scholarship.

Table of Contents

List of Figures

List of Tables

Nomenclature

ACTB	β-Actin
ALDH1	Aldehyde dehydrogenase 1
ALCAM	Activated leukocyte cell adhesion molecule
AML	Acute myeloid leukaemia
APC	Adenomatous polyposis coli
ATCC	American Type Culture Collection
AXIN	Axis inhibition protein 1
β	Beta
b-FGF/FGF	Basic-fibroblast growth factor/fibroblast growth factor
β2M	Beta-2-microglobulin
β-TrCP	Beta-transducin repeat containing protein
BMP4	Bone morphogenetic protein-4
bp	Base pair
BRAF	V-raf murine sarcoma viral oncogene homolog B1
BSA	Bovine serum albumin
CD24, 34, 38, 44, 94f, 133, 166	Cluster of differentiation 24, 34, 38, 44, 94f, 133, 166
Cdx2	Caudal type homeobox 2
ChIP	Chromatin immunoprecipitation
CIMP	CpG methylator phenotype
CKI	Casein Kinase I
c-MYC	V-myc avian myelocytomatosis viral oncogene homolog
CO_2	Carbon dioxide
C_q	Quantification cycle

CSC	Cancer stem cell
CSI	Chromosomal instability
CK20	Cytokeratin 20
°C	Degrees celcius
DAPI	4', 6-diamidino-2-phenylindole
DKK-1, 2, 3, 4	Dickkopf-related protein-1, 2, 3, 4
DLL	Delta-like ligand
DMEM	Dulbecco's modified Eagles medium
DMSO	Dimethyl sulphoxide
DNA, cDNA, gDNA	Deoxyribonucleic acid, complementary deoxyribonucleic acid, genomic deoxyribonucleic acid
DNMT	DNA methyltransferase
ECM	Extracellular matrix
EDTA	Ethylenediaminetetraacetic acid
EGF	Epidermal growth factor
EGFR	Epidermal growth factor receptor
EMT	Epithelial-mesenchymal transition
ESA	Epithelial surface antigen
ESC	Embryonic stem cell
EpCAM	Epithelial cell adhesion molecule
EpICD	EpCAM intracellular domain
EZH2	Enhancer of zeste 2 polycomb repressive complex subunit
FBS	Foetal bovine serum
FcR	Fragment crystallizable region
FHL2	Four and a half LIM domains protein 2

FITC	Fluorescein isothiocyanate
FOLFIRI	Folinic acid, Fluorouracil, Irinotecan
FOLFOX	Folinic acid, Fluorouracil, Oxaliplatin
FSC-A	Forward scatter-area
FSC-H	Forward scatter-height
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
GDP	Guanosine-5'-diphosphate
GLI	Glioblastoma
GP130	Glycoprotein 130
g	Gram
GTP	Guanosine-5'-triphosphate
GSK-3	Glycogensynthase kinase-3
GUS	Glucuronidase
H2A, H2B, H3, H4	Histone 2A, 2B, 3, 4
H_2O	Water
HAT	Histone acetyltransferase
hEC	Human embryonal carcinoma derived
HEY	Hairy ears, Y-linked
HEYL	Hairy/enhancer-of-split related with YRPW motif-like
HES	Hairy and enhancer of split
HDAC	Histone deacetylase
HH	Hedgehog
HHIP	Hedgehog interacting protein
HMT	Histone methyl transferase
HPRT	Hypoxanthine phosphoribosyltransferase 1

Id	Inhibitor of differentiation
IPC	Inter-plate calibrator
iPSC	Induce pluripotent stem cell
IU	International unit
JAK-P	Janus kinase protein
K	Lysine
kDa	Kilodalton
KLF-4	Kruppel-like factor 4
LEF	Lymphoid enhancer-binding factor
LIF	Leukaemia inhibitory factor
LRP	LDL receptor-related protein
M	Molar
MACS	Magnetic activated cell sorting
MAPK	Mitogen-activated protein kinase
MEK	Methyl ethyl ketone
MET	Mesenchymal-epithelial transition
µg	Microgram
µl	Microlitre
µm	Mircon
mg	Milligram
MIQE	Minimum information for publication of quantitative real-time PCR experiments
ml	Millilitre
mM	Millimolar
MLH1	MutL homologue 1
MMR	Mismatch repair

MSH1, 2	MutS homologue 1, 2
MSI	Microsatellite Instability
ng	Nanogram
nM	Nanomolar
NT	No treatment
OCT4	Octamer-4
P110α	Phosphatidylinositide 3-kinase subunit alpha
PTCH	Patched
PBS	Phosphate buffered saline
PcG	Polycomb group proteins
PCR, qPCR	Polymerase chain reaction, quantitative real-time polymerase chain reaction
PDK1	Pyruvate dehydrogenase lipoamide kinase isozyme 1
PE	Phycoerythrin
pg	Picogram
PI3K	Phosphatidylinositide 3-kinase
PIK3CA	Phosphatidylinositol-4,5-bisphosphate 3-kinase
PIP2, 3	Phosphatidylinositol 4,5-bisphosphate-2, 3
PRC2	Polycomb repressive complex 2
p-value	Probability-value
RAF	Rapidly accelerated fibrosarcoma
RAS	Rat sarcoma
RBP-Jκ	Recombination signal binding protein J kappa
RIN	RNA integrity number
R-SMAD	Receptor-regulated mothers against decapentaplegic homolog

RNA, mRNA, miRNA, shRNA	Ribonucleic acid, messenger ribonucleic acid, micro ribonucleic acid, short hairpin ribonucleic acid
Rpm	Revolutions per minute
SC	Stem cell
SCID	Severe combined immunodeficient
SFRP2	Secreted frizzled-related protein 2
SHH	Sonic Hedgehog
SMAD1, 2, 3, 5	Mothers against decapentaplegic homolog 1, 2, 3, 5
SMO	Smoothened
SOX2, 21	Sex determining region Y (SRY)-related high mobility group (HMG) box 2, 21
SOX2OT	SOX2 overlapping transcript
SRY	Sex-determining region Y
STAT3	Signal transducer and activator of transcription 3
SUMO	Small ubiquitin-related modifier
TBE	Tris/Borate/Ethylenediaminetetraacetic acid
TBP	TATA box binding protein
TCF	T-cell factor
TGF-β	Transforming growth factor beta
T_M	Melting temperature
TP53	Tumour protein 53
TR	Transferrin receptor
TrxG	Trithorax group
TSA	Trichostatin A
TSG	Tumour suppressor gene
UBC	Ubiquitin C

VEGF	Vascular endothelial growth factor
VPA	Valproic acid
WIF-1	Wnt inhibitory factor 1
WNT	Wingless type
YWHAZ	Tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, zeta polypeptide
ZEB	Zebularine

CHAPTER 1

INTRODUCTION

Cancer is a worldwide burden in both the developed and developing world. In 2008, as many as 12.7 million cancer cases, and 7.6 million cancer deaths were estimated to have occurred. Colorectal cancer has a 9.8% cumulative incidence rate making it the third most common cancer in the Western world (Jemal *et al.*, 2009). It is estimated that colorectal cancer will cause as many as 26 270 male- and 24 040 female-deaths in the United States in 2014 (Cancer Facts and Figures, 2014). In 2007, The South African National Cancer registry had statistics on colorectal cancer prevalence in South Africans, for men it was 4.47% and women 3.59%. The age standardised incidence rate was 7.12 per 100 000 men and 4.34 per 100 000 women; whereas, the lifetime risk of developing colorectal cancer before the age of 74 was 1/115 men and 1/199 women. The lifetime risk was highest in Caucasian men (1/50) and women (1/85), followed by Asian men (1/59) and women (1/90) (Table 1.1). Although these figures may seem somewhat small compared to those of the Western world, they are indeed significant. To demonstrate their significance, women living in Uganda and Zimbabwe have a lower risk of developing cancer by the age of 65 compared with Western Europe; however, the risk of cancer related death is almost doubled (Ferlay *et al.*, 2004). This emphasises the need to study cancer prevention strategies and to design novel anti-cancer therapeutics to treat colorectal cancer in African populations.

Table 1.1: The lifetime risk (ages 0-74 years) of developing colorectal cancer per population group in South Africa.

Population	**Males**	**Females**
Caucasian	1/50	1/85
Asian	1/59	1/90
Coloured	1/63	1/100
Black	1/314	1/434
All	1/115	1/199

The standard treatment for patients with localised colorectal cancer requires surgery to remove the cancer; whereas, the treatment for patients with advanced metastatic disease involves chemotherapy and/or radiotherapy following surgery (DeSantis *et al.*, 2014). Adjuvant chemotherapy and/or radiotherapy may also be utilised to treat localised disease, but this is dependent on the patient and the grade of the tumour. The chemotherapeutic agents used include: 5-Fluorouracil, Leucovorin, Oxaliplatin, Irinotecan and combinations thereof, for example, FOLFOX (5-Fluorouracil, Leucovorin and Oxaliplatin) and FOLFIRI (5-Fluorouracil, Leucovorin and Irinotecan) (de Gramont *et al.*, 2000; Falcone *et al.*, 2007). In some instances, these drugs may be combined with the targeted therapeutic agents Bevacizumab or Cetuximab, these being monoclonal antibody derived inhibitors of vascular endothelial growth factor (VEGF) and epidermal growth factor receptor (EGFR), respectively (Hurwitz *et al.*, 2004; Giantonio *et al.*, 2007; Saltz *et al.*, 2008; Van Cutsem *et al.*, 2009; Dranitsaris *et al.*, 2010). However, despite a 50-60% response rate in patients to these chemotherapeutic agents, drug resistance and tumour relapse remains a concern (Paldino *et al.*, 2014).

It has been suggested that a subpopulation of tumour cells, the tumour initiating cells or cancer stem cells (CSCs), are highly resistant to these chemotherapeutic agents and are responsible for tumour relapse following therapy (Visvader *et al.*, 2008; Lin , 2014). In this regard, it is thought that CSCs may actively evade cytotoxic or radiotherapy, and further, that quiescent CSCs may possibly be more resistant to therapy. If correct, this implies that there may be a need to develop novel therapies that target these CSCs specifically (Visvader & Lindeman, 2012). Therefore, in this thesis, a putative colorectal CSC subpopulation will be isolated from two invasive colon adenocarcinoma cell lines to determine the potential regulatory effect(s) that the epigenetic drugs Valproic acid (VPA) and Zebularine (ZEB) have on pluripotent marker gene expression in these cells. Since CSCs are highly resistant to current cancer therapies, it is thought that altering the expression of key pluripotent genes that maintain "stemness" in these cells would induce cell differentiation and make them more susceptible to current therapy (Chiou *et al.*, 2010; Gudas & Wagner, 2011; Ibrahim *et al.*, 2012; Munoz *et al.*, 2012; Zhang *et al.*, 2012; Chen *et al.*, 2013; Shiozawa *et al.*, 2013).

1.1 The Adenoma-carcinoma Sequence

At a molecular level, colorectal cancer is known to arise from the sequential accumulation of multiple mutations in a single cell target. The sequence of genetic and epigenetic events leading to the initiation and progression of tumourigenesis in the colon is referred to as the adeno-carcinoma sequence, as shown in Figure 1.1 (Fearon & Vogelstein, 1990). The transformation of normal colonic epithelium to an adenoma occurs over a period of 5-20 years, and its progression to a carcinoma occurs anywhere from 5-15 years, as seen in Figure 1.1. There are two major pathways associated with the adenoma-carcinoma sequence, these being chromosomal instability (CSI) and microsatellite instability (MSI) (Grady & Carethers, 2008). These events are driven by the dysregulation of numerous signalling cascades implicated in colorectal cancer metastasis (see Figure 1.2).

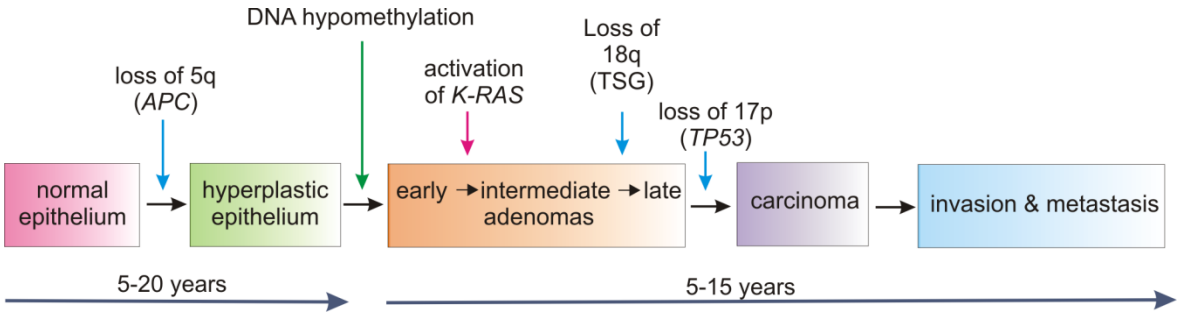

Figure 1.1. Colorectal tumour progression and the points where oncogene and tumour suppressor gene mutations occur. Initially, an epithelial cell acquires a mutation in the tumour suppressor gene *APC*, followed by DNA hypomethylation. This is followed by a sequential accumulation of mutations in the oncogenes *K-RAS*, *RAF* and *PIK3CA* and the tumour suppressor genes (TSG) *TP53* and Transforming growth factor beta (*TGF-β*). Additionally, there is a loss of chromosome 18q, where several other tumour suppressor genes such as *SMAD4* are located (Weinberg, 2007).

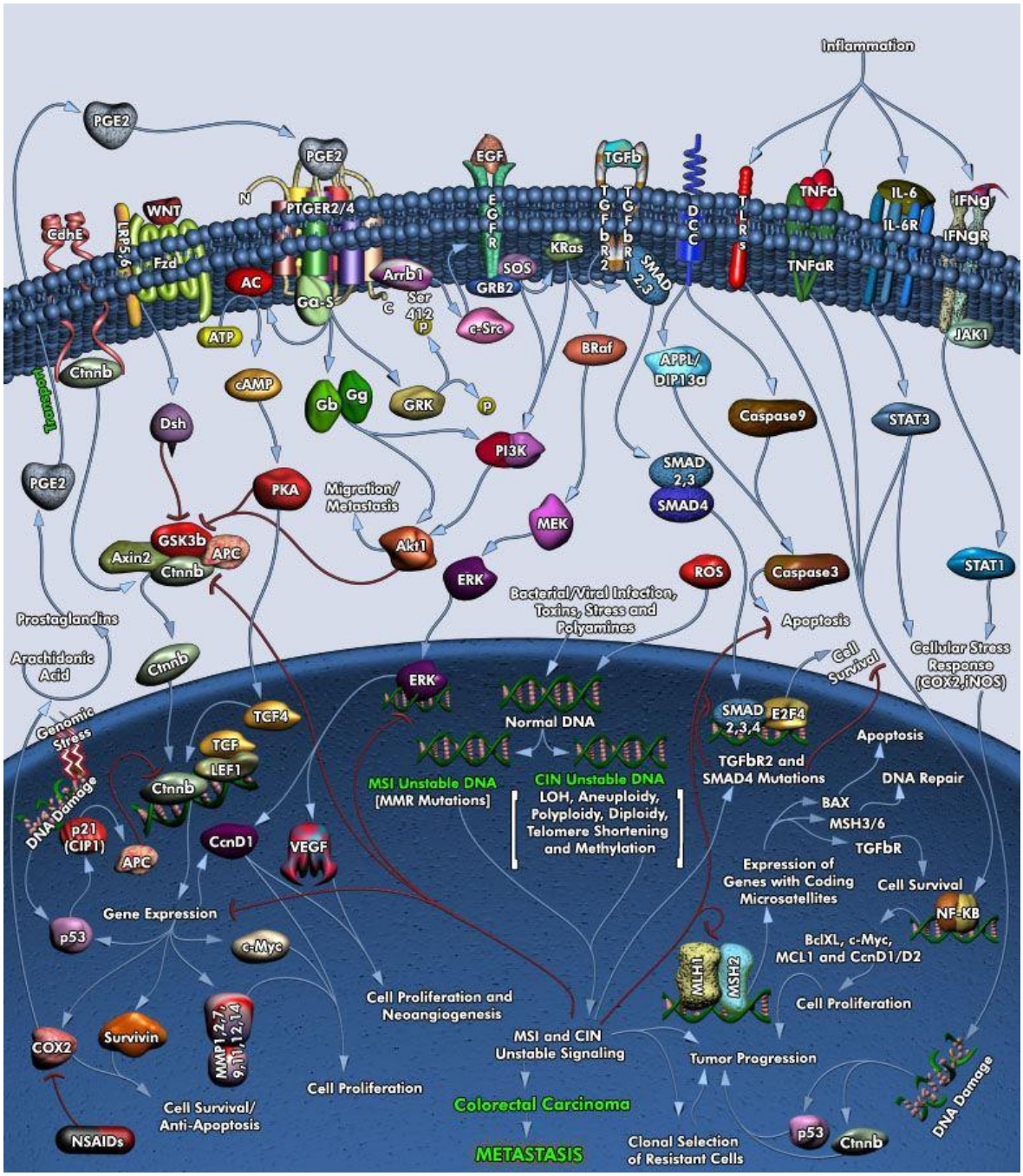

Figure 1.2. Cell signalling pathways involved in the tumourigenesis and metastasis of colorectal cancer. The dysregylation of these signalling cascades contribute to the chromosomal and microsatellite instability involved in the adenoma-carcinoma sequence (Qiagen).

1.1.1 Chromosomal Instability

CSI in the colon is characterised by the allelic loss of the chromosomal regions 5q (*Adenomatous polyposis coli*; *APC*), 18q (*Mothers against decapentaplegic homolog 4*, *SMAD4*) and 17p (*Tumour protein 53*; *TP53*) (Weinberg, 2007). Additionally, deoxyribonucleic acid (DNA) hypomethylation and alterations in the oncogenes *Rat sarcoma* (*RAS*), *Rapidly accelerated fibrosarcoma* (*RAF*) and *Phosphatidylinositol-4,5-bisphosphate 3-kinase* (*PIK3CA*) occur (Weinberg, 2007). The functional roles of these genes, and their loss-of-function, are briefly discussed below.

1.1.1.1 Loss of *APC*

Loss of *APC* function occurs in more than 80% of sporadic colorectal cancers (Morin *et al.*, 1997). *APC*, located on chromosome 5q, functions as a negative regulator of the canonical Wingless type (WNT) signalling pathway, and is additionally a regulator of apoptosis and cell-cycle progression (Fearon & Vogelstein, 1990; Goss & Groden, 2000). In the absence of WNT signalling, APC and Axis inhibition protein 1 (AXIN) bind to newly synthesised β-catenin. A destruction complex consisting of Casein kinase I (CKI) and Glycogensynthase kinase-3 (GSK-3) phosphorylate β-catenin at Serine and Threonine residues on its amino acid terminus. Phosphorylated β-catenin recruits ubiquitin ligase (Beta-transducin repeat containing protein; b-TrCP) resulting in its degradation (Reya & Clevers, 2005). Therefore, loss of APC function results in the accumulation of β-catenin in the nucleus where it activates the transcription of its target genes (Markowitz & Bertagnolli, 2009). Since *APC* plays a role in many cell regulatory functions, it is not surprising that the loss of *APC* function is considered to be a driver of colorectal tumourigenesis.

1.1.1.2 Loss of *SMAD4*

The tumour suppressor gene *SMAD4*, located on chromosome 18q, is a mediator in the Transforming growth factor beta (TGF-β) and Bone morphogenetic protein (BMP) signalling pathways (Dupont *et al.*, 2009). The binding of TGF-β ligands to type II receptors recruits type I receptors, which phosphorylate receptor-regulated SMAD2 and 3 (R-SMAD) (Massagué *et al.*, 2005; Roy & Majumdar, 2012). R-SMAD associates with

SMAD4 and together this complex translocates to the nucleus where it activates the transcription of its downstream target genes to regulate cell proliferation, differentiation and apoptosis (Heldin *et al.*, 1997; Duff & Clarke, 1998; Roy & Majumdar, 2012). In the BMP signalling pathway, SMAD4 associates with receptor-regulated SMAD1 and 5, where it too activates the transcription of genes involved in regulating cell proliferation, differentiation and apoptosis (Medvedev *et al.*, 2008). The loss of *SMAD4* is frequently found in 30-40% of colorectal cancers and its loss is also associated with liver metastases (Miyaki *et al.*, 1999; Kodach *et al.*, 2008). Furthermore, it promotes the "switching" of TGF-β tumour suppressor activity to tumour promoter activity (Zhang *et al.*, 2010). Voorneveld *et al.* (2014) also showed that in the absence of SMAD4, BMP signalling enhances both cell invasion and metastasis through its activation of SMAD4-independent BMP signalling pathways.

1.1.1.3 Loss of *TP53*

TP53 (also known as *p53*) is a tumour suppressor gene located on chromosome 17p. It is considered to be the "guardian of the genome" due to its functional role in regulating cell cycle progression, as well as its ability to stimulate DNA repair and promote cellular apoptosis (Lane, 1992). *TP53* mutations are highly abundant in colorectal cancer (Sameer, 2013). While mutations in this gene result in the loss of tumour suppressor activity, some mutations may however result in a gain-of-function phenotype (Leslie *et al.*, 2002). Gain-of-function attributes include increased genomic instability, enhanced cell migration and invasion, as well as resistance to pro-apoptotic signals (Oren & Rotter, 2010).

1.1.1.4 *RAS*, *RAF* and *PIK3CA* Mutations

RAS is a mediator in the Mitogen-activated protein kinase (MAPK) and Phosphatidylinositide 3-kinase (PI3K) signalling pathways that are implicated in controlling cell growth and survival (Courtney *et al.*, 2010; Nandan *et al.*, 2011). The MAPK signalling pathway is activated upon the binding of ligands to tyrosine kinase receptors resulting in their activation via phosphorylation. Activated receptors activate RAS proteins which have GTPase activity (Zenonos & Kyprianou, 2013). Activated RAS initiates a sequential phosphorylation cascade whereby it activates RAF which further activates Methyl ethyl ketone (MEK), and subsequently MAPK (Zenonos & Kyprianou,

2013). Phosphorylated MAPK is translocated to the nucleus where it activates transcription factors involved in maintaining cell proliferation (Ubeda *et al.*, 1999; Nandan *et al.*, 2011).

The PI3K signalling pathway becomes activated upon the binding of ligands to tyrosine kinase receptors (Zenonos & Kyprianou, 2013). PI3K is recruited to the phosphorylated receptors and its catalytic domain Phosphatidylinositide 3-kinase subunit alpha (P110α) phosphorylates Phosphatidylinositol 4,5-bisphosphate (PIP)-2 to PIP3 (Pacold *et al.*, 2000; Courtney *et al.*, 2010). PIP3 then recruits PDK1 and AKT resulting in AKT phosphorylation, which regulates cell proliferation, survival and migration (Zenonos & Kyprianou, 2013). RAS can also activate this pathway through its direct binding with P110α (Pacold *et al.*, 2000).

RAS mutations, specifically of the K-RAS isoform, are found in 40% of colorectal tumours. RAS mutations prevent the hydrolysis of RAS-GTP (active RAS) to RAS-GDP (inactive RAS) and therefore, RAS continuously signals through the MAPK pathway (Zenonos & Kyprianou, 2013). It has become of some importance clinically to identify the specific mutations in the *K-RAS* gene; these being mutations in codons 12 and 13 (about 95% of mutations) and codons 61 and 146 (between 35 to 40% of mutations), as patients with these mutations are unresponsive to anti-EGFR therapy (Cetuximab/Panitumumab) due to the constitutive up-regulation of the MAPK pathway (Tan & Du, 2012; Zenonos & Kyprianou, 2013). Furthermore, mutations of *BRAF* also leads to continuous MAPK signalling and is found in 5-10% of colorectal tumours (Zenonos & Kyprianou, 2013). Mutations of *PIK3CA*, which encodes the P110α catalytic domain of PI3K, often occur together with *RAS* mutations. It is found in 15-20% of colorectal tumours and results in increased PI3K activity (Fearon, 2011). Therefore, mutations of *RAS*, *RAF* and *PIK3CA* lead to enhanced cell proliferation, migration and survival.

1.1.2 Microsatellite Instability

MSI is characterised by a loss-of-function of DNA mismatch repair (MMR) genes (Grady & Carethers, 2008; Boland & Goel, 2010). Microsatellites are tandem repeats of oligonucleotides and these are highly prone to error during DNA replication (Watson *et al.*, 2014). Under normal circumstances mismatch errors are repaired by MMR proteins

however, these errors cannot be repaired when MMR function is lost (Fukui, 2010). As a result, microsatellite errors accumulate in key regulatory genes (Marra & Boland, 1996; Grady & Carethers, 2008). MSI occurs in 15-20% of sporadic colorectal cancers, and in more than 95% of patients with hereditary Lynch Syndrome (Grady & Carethers, 2008). In sporadic cases, MSI usually results due to epigenetic silencing of the MMR gene *MutL homologue 1* (*MLH1*) (Papadopoulos *et al.*, 1994). Patients with Lynch syndrome usually have a germline mutation in the MMR gene *MutS homologue 1* (*hMSH1*) or *MutS homologue 2* (*hMSH2*) (Leach *et al.*, 1993; Fishel *et al.*, 1993).

1.1.3 Epigenetic Instability

Although genomic instability is a well-established mechanism in colorectal tumour initiation and progression, epi-mutations also play a fundamental role and may occur more frequently than genetic mutations (Issa, 2004; Schnekenburger & Diederich, 2012). Epigenetics refers to heritable changes of the genome that alter gene expression without a change in the primary DNA sequence (Peedicayil, 2006; Goel & Boland, 2012; Migheli & Migliore, 2012; Schnekenburger & Diederich, 2012). These modifications include DNA methylation, histone modifications and non-coding ribonucleic acids (RNAs), and are briefly presented below (Migheli & Migliore, 2012).

Altered states of DNA methylation, involving both hyper- and hypomethylation, is highly prevalent in colorectal cancer. It involves the addition of methyl groups to CpG islands often located in the promoter regions of genes (Schnekenburger & Diederich, 2012). This modifies the three-dimensional conformation of DNA to prevent its interaction with transcription factors resulting in gene silencing (Graff *et al.*, 1997). Genome-wide methylation profiling studies have allowed researchers to identify several aberrantly methylated tumour suppressor genes associated with colorectal cancer. Cellular targets/functions often affected include WNT, RAS and MAPK signalling, cell cycle regulation, calcium and lipid metabolism, cell adhesion and migration, DNA repair and apoptosis (Goel & Boland, 2012; Schnekenburger & Diederich, 2012). Furthermore, colorectal tumours may be classified as having a CpG methylator phenotype (CIMP), should these methylated CpG islands occur extensively (Curtin *et al.*, 2011). Toyota *et al.* (1999) proposed a role of CIMP in colorectal tumour development and to date many

colorectal cancers are characterised as CIMP high (Migheli & Migliore, 2012). The number of effects that aberrant methylation has on cell function undoubtedly demonstrates its role in colorectal tumourigenesis and progression.

Histone modifications such as histone acetylation and deacetylation, and histone methylation and demethylation are best characterised in colorectal cancers, but their patterns are less understood compared to DNA methylation (van Engeland *et al.*, 2011; Goel & Boland, 2012). Histone modifications occur at histones H2A, H2B, H3 and H4, which associate with 150-200bp of DNA, termed the nucleosome (Kornberg, 1974; Luger & Richmond, 1998; Kornberg & Lorch, 1999). Histone tails protrude from these nucleosomes and are subject to post-translational modifications, including amongst others, acetylation, methylation, ubiquitylation, sumoylation and phosphorylation (Nakazawa *et al.*, 2012; Schnekenburger & Diederich, 2012). Of these modifications, the best understood ones include the acetylation of lysine residues, and methylation of lysine and arginine residues (Schnekenburger & Diederich, 2012). Histone mono-acetylation is catalysed by histone acetylases (HATs) and lessens the association of DNA with the core histone proteins, whereas histone mono-, di- and tri-methylation is catalysed by histone methyltransferases (HMTs) and strengthens their association with DNA (Bannister & Kouzarides, 2011). RNA polymerase can easily bind to DNA that is loosely associated with histones to allow gene expression (Figure 1.3) (Goel & Boland, 2012; Migheli & Migliore, 2012).

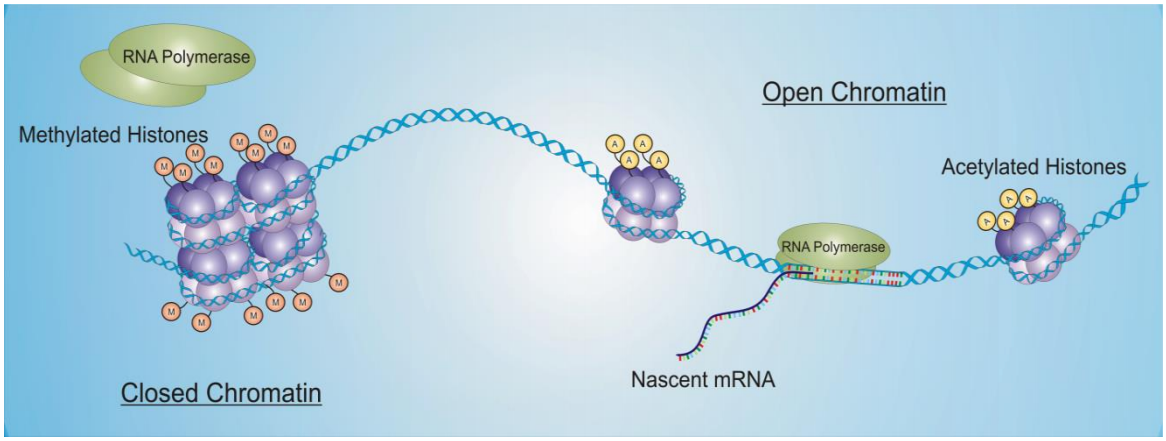

Figure 1.3. Histone methylation and acetylation alter gene expression. When histones are methylated (left) they tightly associate with DNA resulting in gene silencing. However, when histones are acetylated (right) they loosely associate with DNA to allow effective binding of RNA polymerase for gene transcription, and hence gene expression (adapted from Figure 1 in McClearly-Wheeler *et al.*, 2013, 213 p.).

The dysregulation of histone modifications, in addition to DNA methylation, contributes to colorectal carcinogenesis. The overexpression of histone deacetylases (HDACs) which catalyse the removal of acetyl groups are associated with enhanced cell proliferation and cell survival (Mariadason *et al.*, 2008; Weichert *et al.*, 2008). It was shown that class I HDACs (HDAC1, HDAC2 and HDAC3) are up-regulated in colorectal cancer, and that their overexpression is associated with poor patient survival (Weichert *et al.*, 2008; Ashktorab *et al.*, 2009). During colorectal tumourigenesis, the overexpression of HDAC2 in particular is followed by the acetylation of lysine residue 18 at histone 3 (H3K18) and lysine residue 12 at histone 4 (H4K12). As a result, these histone marks are good prognostic markers for predicting patient survival (Ashktorab *et al.*, 2009). Similarly, dysregulated methylation is observed in patients with colorectal cancer. Methylation of histones may result in either transcriptional activation or repression depending on the methylated site (Cloos *et al.*, 2008). Tri-methylation of lysine 4 at histone 3 (H3K4me3) is catalysed by the trithorax group (TrxG) of proteins and is associated with transcriptional activation; whereas, tri-methylation of lysine 27 at histone 3 (H3K27me3) is catalysed by the polycomb repressive complex 2 (PRC2), a class of the polycomb group (PcG) proteins, and is associated with gene silencing (Seenundun *et al.*, 2010). Genes located between

these “bivalent domain” chromatin marks are transcriptionally silent but have the capacity to be transcriptionally activated at a later stage (Felici, 2011). Enhancer of zeste 2 polycomb repressive complex (EZH2), the catalytic subunit of PRC2 of the polycomb complex, has been reported to be overexpressed in 87% of colorectal cancers and is associated with poor patient survival (Bracken & Helin, 2009; Crea *et al.*, 2011; Takawa *et al.*, 2011). Overall, the dysregulation of histone modifications contributes to colorectal tumourigenesis.

Micro RNAs (miRNAs) are an additional mechanism of epigenetic regulation. They are small non-coding RNAs generated from precursor RNAs that play a role in transcriptional and post transcriptional gene expression (Rane *et al.*, 2007). They anneal to complementary mRNA, often resulting in gene silencing. In colorectal cancer, miRNAs play a role in WNT signalling, cell migration and invasion, epithelial differentiation and cell cycle regulation (Lujambio *et al.*, 2007; Shi *et al.*, 2007; Huang *et al.*, 2008; Schnekenburger & Diederich, 2012). Table 1.2 lists some miRNAs involved in colorectal carcinogenesis, and their functional roles.

Altogether, epi-mutations involving the dysregulation of DNA methylation, histone modifications and miRNAs, contribute to colorectal tumourigenesis and therefore, the mechanisms involved are potential therapeutic targets, since epi-mutations are possibly reversible (Peltomäki, 2012). The therapeutic targeting of epigenetic mechanisms is further discussed in section 1.6.

Table 1.2: miRNAs and their role in colorectal cancer.

miRNA	**Function**
miRNA-145	WNT/β-catenin signalling (Shi *et al.*, 2007; Arndt *et al.*, 2009; Sarver *et al.*, 2009; Oberg *et al.*, 2011)
miRNA-135a	
miRNA-135b	
miRNA-139	
miRNA-145	
miRNA-134a	Apoptosis (Arndt *et al.*, 2009; Earle *et al.*, 2010)
miRNA-133b	
miRNA-195	
miRNA-34b/c	p53 signalling (Vogt *et al.*, 2011)
miRNA-18a	Proliferation (Arndt *et al.*, 2009; Motoyama *et al.*, 2009; Earle *et al.*, 2010)
miRNA-143	
miRNA-200c	
miRNA-373	Cell migration and invasion (Huang *et al.*, 2008)
miRNA-520c	
miRNA-141	Epithelial differentiation (Burk *et al.*, 2008)
miRNA-200c	
miRNA-124a	Cell cycle regulation (Lujambio *et al.*, 2007; Arndt *et al.*, 2009; Earle *et al.*, 2010; Tsang *et al.*, 2010)
miRNA-34a	
miRNA-192	
miRNA-215	
miRNA-675	
miRNA-126	Migration, invasion, metastasis (Arndt *et al.*, 2009; Earle *et al.*, 2010; Tsang *et al.*, 2010)
miRNA-200a/b/c	
miRNA-520c	

1.2 Colonic Stem Cells - a Target of Transformation

The colon and rectum as described by Ellis (2010) form the large intestine. The colon can be subdivided into ascending, transverse, descending and sigmoid sections. The colon was previously thought to play an unnecessary role in humans, but in fact the colonic epithelial lining is essential for nutrient absorption; also it serves as a barrier to the external environment and provides a microenvironment for gut microflora (Edwards 1997; Anderson *et al.*, 2011).

Colorectal cancer occurs in the epithelial lining of the colon and the rectum. The colonic epithelium is constantly being renewed by stem cells (SCs) located within the bottom of the villus crypts (Booth & Potten, 2000; Marshman *et al.*, 2002; Anderson *et al.*, 2011). Generally, 4-6 SCs reside in the crypt base, which continuously self-renew via symmetric cell divisions and generate transit-amplifying cells via asymmetric cell division (Potten *et al.*, 1997; Marshman *et al.*, 2002; Barker & Clevers, 2007). Transit-amplifying cells terminally differentiate to produce one of four cell types, each with a specialised biological function. These differentiated cell types include: colonocytes (absorptive cells), goblet cells (secrete mucin), enteroendocrine cells (produce hormones) and Paneth cells (secrete antimicrobial molecules) (Anderson *et al.*, 2011). They are located in the upper third of the colonic crypt and are shed into the lumen every 2-7 days (Willis *et al.*, 2008).

Since the colonic epithelium is constantly being renewed by long-lived SCs located within the crypt bottom, as opposed to transit-amplifying and differentiated cells that are short-lived, it is likely that SCs are the targets of transforming mutations (Booth & Potten, 2000; Marshman *et al.*, 2002; Vermeulen, *et al.*, 2008; Wang *et al.*, 2013a). This would imply a hierarchical organisation of cells within a tumour whereby transformed SCs, better known as CSCs or tumour initiating cells, are capable of producing the heterogeneous population of cells that comprise a tumour. This is best explained by the CSC model which is further discussed in section 1.2.

The exact origin of colorectal CSCs remains unknown as they may also arise from the de-differentiation of transit-amplifying and differentiated cells that have been subjected to transformation (Vermeulen, *et al.*, 2008; Anderson *et al.*, 2011; Wang *et al.*, 2013a). However, it is likely that they arise from adult SCs as they share many common features

with embryonic SCs. These common features include the ability to proliferate indefinitely whilst maintaining pluripotency, differentiation potential and migration capacity (Evans & Kaufman, 1981; Martin, 1981; Wong *et al*., 2008; Schoenhals *et al*., 2009; Kim *et al.*, 2010; Mizuno *et al*., 2010; Shats *et al*., 2011). Figure 1.4 shows a colonic crypt and the various cell types that comprise it. Also indicated in this figure are the likely points where transforming mutations, leading to the formation of CSCs, might occur.

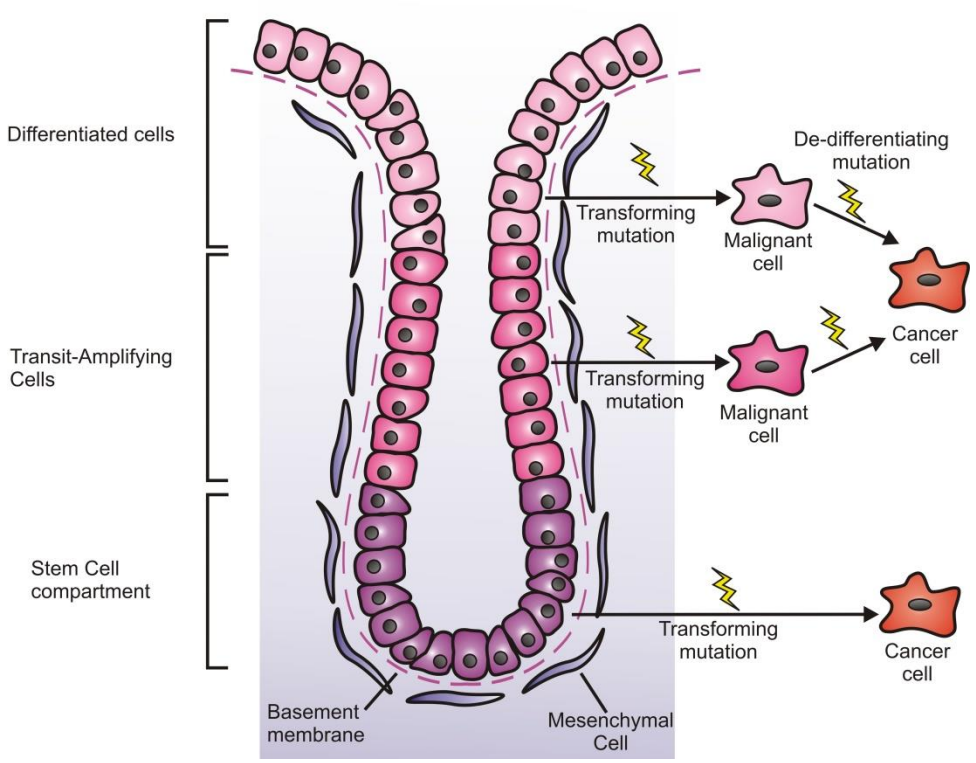

Figure 1.4. The colonic crypt showing the lower SC compartment (purple), the transit-amplifying cells (dark pink) and the differentiated cells (light pink) located in the upper third of the crypt. Indicated on the right are the points where transforming mutations may occur in colonic crypt cells (adapted from Figure 1 in Anderson *et al.*, 2011, 321 p.).

1.3 The Cancer Stem Cell Model

It has long been believed that tumours are comprised of a heterogeneous population of cells wherein a discrete subpopulation of CSCs are able to give rise to the remaining cells that comprise a tumour (Dalerba *et al.*, 2007a; O'Brien *et al.*, 2007; Ricci-Vitiani *et al.*,

2007 and Vermeulen, *et al.*, 2008). This key concept forms the basis of the CSC model. It differs to the conventional model that proposes tumour growth results from the self-renewal and differential potential of every cell type present within the tumour (Willis *et al.*, 2008). The hierarchical organisation of cells within a tumour, as proposed by the CSC model is best described by Vermeulen *et al*. (2008) and is illustrated in Figure 1.5. In this figure, it is shown that tumourigenic CSCs derived from adult SCs, or via the de-differentiation of somatic cells, can self-renew to maintain the SC pool. In addition, they are precursors to transit-amplifying cells that further produce non-tumourigenic, differentiated cells.

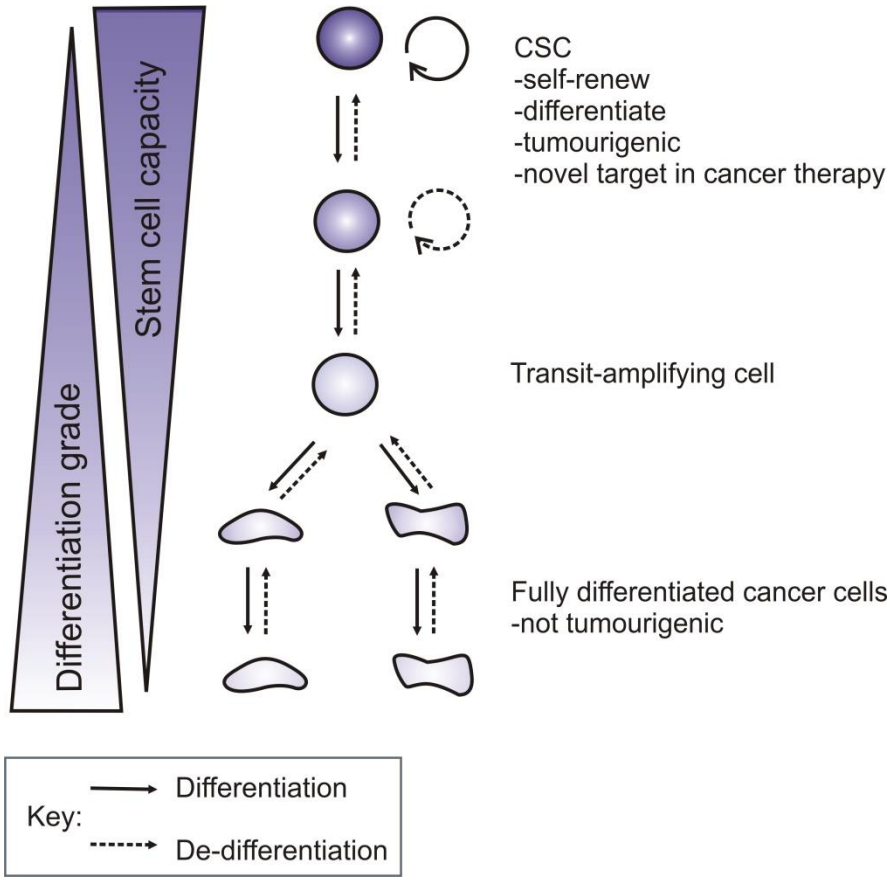

Figure 1.5. The cancer stem cell model and the hierarchical organisation of cells within a tumour (Vermeulen *et al*., 2008).

Evidence for the existence of a CSC subpopulation in tumours was first shown by Lapidot *et al.* (1994). They reported that a subset of acute myeloid leukaemia (AML) cells, characterised as cluster of differentiation (CD) $34^{+}/CD38^{-}$, could regenerate a phenocopy of the original leukaemia when these cells were transplanted into severe combined immunodeficient (SCID) mice. However, and in contrast to $CD34^{+}/CD38^{-}$ cells, AML cells characterised as $CD34^{+}/CD38^{+}$ or $CD34^{-}/CD38^{-}$ were unable to engraft tumours in immunodeficient mice. Similarly, evidence that a CSC population exists in malignant tumours was further extended to show that a CSC population exists in solid tumours. Al Hajj *et al.* (2003) identified and isolated breast CSCs characterised as epithelial cell adhesion molecule $(EpCAM)^{+}/CD44^{+}/CD24^{-/low}$ and demonstrated that these cells could produce tumours when serially transplanted into immunodeficient mice. Since then, studies have provided evidence for the existence of a CSC population in several solid tumours, including amongst others brain (Singh *et al.*, 2004), colon (Dalerba *et al.*, 2007b; O'Brien *et al.*, 2007) and lung (Eramo *et al.*, 2008).

1.4 Characterisation of Colorectal CSCs

A distinct profile of cell surface markers is used to characterise and isolate colorectal CSC populations (Horst *et al.*, 2008; Yeung *et al.*, 2010; Ren *et al.*, 2013). Currently, CD133 (also known as prominin-1) is one of the best-characterised markers used to identify and isolate colorectal CSCs through its binding with the antibodies AC133 that recognises the CD133/1 epitope, and AC141 and 293C that recognise the CD133/2 epitope (Zhou *et al.*, 2007; Kemper *et al.*, 2010). It is a promising marker since its increased expression has been correlated with increased tumour stage and reduced overall patient survival (Shimada *et al.*, 2011). Moreover, CD133 expression has been correlated with increased WNT activity in putative colorectal CSCs, a driver of colorectal tumourigenesis. Corbo *et al.* (2012) showed WNT signalling was activated in $CD133^{+}$ Caco2 and HCT116 cells. An increase in WNT activity resulted in the up-regulation of SRp20, a splicing factor whose activity has been linked to cell proliferation. In a study by Mak *et al.* (2012), they showed the association of CD133 with HDAC6 enhanced β-catenin stability in Caco2 colorectal cancer cells.

Although the biological function of CD133 is unknown, this five-transmembrane glycoprotein is however believed to be associated with cell-cell and cell-matrix interactions (Miraglia *et al.*, 1997; Yin et al., 1997; Ren *et al.*, 2013). This protein was originally discovered as a marker for neuroepithelial progenitor cells in mouse embryos, and later in humans as a marker of haematopoietic SCs (Weigmann *et al.*, 1997; Yin *et al.*, 1997). Figure 1.6 shows a structural model of CD133 as proposed by Miraglia *et al.* (1997). The transmembrane protein is characterised by an extracellular N-terminus, a cytoplasmic C-terminus and two extracellular loops that contain glycosylation sites that AC133 antibodies bind to.

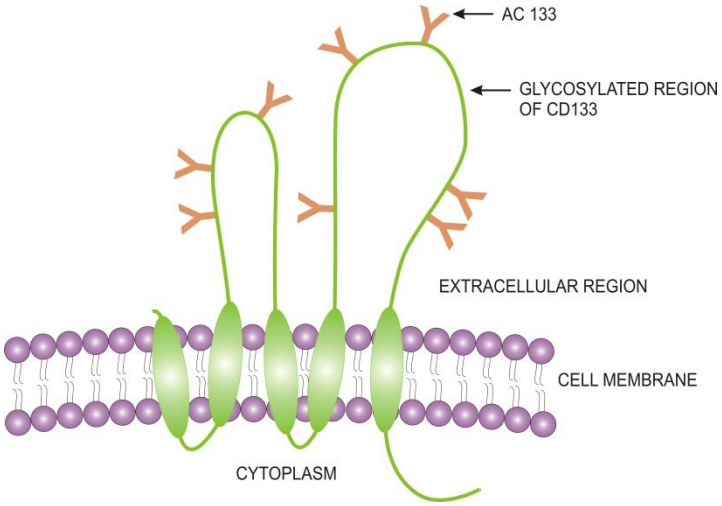

Figure 1.6. The five-transmembrane glycoprotein CD133: the glycosylated region of CD133, present in the extracellular region of the cell, can bind with AC133 antibodies used to isolate and characterise stem cells expressing the CD133 protein (adapted from Figure 1 in Yu *et al*. 2011, 255 p.).

Two independent studies by O'Brien *et al.* (2007) and Ricci-Vitiani *et al.* (2007) indicated that CD133 can be used as a biological marker to isolate and characterise colorectal CSCs. Both investigators isolated $CD133^+$ colon cancer cells from primary tumours and evaluated their tumourigenic potential *in vivo*. They reported that the transplantation of $CD133^+$ colon cancer cells, and not $CD133^-$ cells could initiate tumourigenesis in immunocompromised mice, whilst preserving the original tumour phenotype.

However, and in contrast to these studies, Shmelkov *et al.* (2008) observed that both CD133 positive and negative colon cancer cells isolated from a metastatic tumour, could elicit tumour formation in immunocompromised mice; although, the $CD133^{+}$ cells were found to be less aggressive, they did form larger tumours. Similarly, Kawamoto *et al.* (2010) observed that both CD133 colon cancer subsets were able to elicit tumour formation in immunocompromised mice. Shmelkov *et al.* (2008) took their study further and tracked the expression of CD133 in the colon. To do so, they generated knock-in lacZ reporter ($CD133^{lacZ/+}$) mice, and reported that CD133 was expressed in both undifferentiated and differentiated cells. They further demonstrated that CD133 expression was not restricted to CSCs of primary human and murine colorectal tumours. Based on these findings, it was concluded that CD133 expression is controversial and that it may not be a reliable marker when used alone to characterise and isolate colorectal CSCs.

It should however be taken into consideration that Shmelkov *et al.* (2008) isolated CD133 subsets from a metastatic tumour; whereas, both O'Brien *et al.* (2007) and Ricci-Vitiani *et al.* (2007) isolated these subsets from primary tumours (Ren *et al.*, 2013). During metastases, epithelial cells undergo epithelial-mesenchymal transition (EMT) which enables cells to migrate to and invade distant tissues (Tsai & Yang, 2013). During this phenotypic change, cells can be induced to gain SC characteristics which may explain why the $CD133^{-}$ subset was able to elicit tumour formation *in vivo* (Irollo & Pirozzi, 2013). Nonetheless, this does not explain Shmelkov's finding that both undifferentiated and differentiated cells express CD133 in murine colonic epithelial tissue, as well as in human and murine colorectal tumours.

Since the use of CD133 as a marker to identify putative colorectal CSCs is currently controversial, Kemper *et al.* (2010) investigated possible regulatory mechanisms of AC133. As a first approach, they investigated *CD133* promoter activity and mRNA expression in undifferentiated and differentiated cells of primary colorectal tumours. They reported that its promoter activity did not differ considerably between these cell types however, they did find that *in vitro AC133* expression decreased almost 10-fold in differentiated cells with unchanged *CD133* mRNA expression, when compared to undifferentiated CSCs. Since promoter activity could not explain differential AC133

recognition, they further investigated whether alternative splicing of *CD133* mRNA could influence the expression of the AC133 epitope. From their findings, they showed that alternative splicing did not contribute to differential recognition of AC133, since alternative splicing does not influence the extracellular loop where the AC133 epitope resides. Therefore, they further investigated the glycosylation status of CD133 upon CSC differentiation. Glycosylation is a translational and post-translational modification governed by the binding of glycan molecules to proteins (Irollo & Pirozzi, 2013). Surprisingly, they reported that differential glycosylation resulted in the differential folding of CD133 which may mask the AC133 epitope therefore, making it undetectable in differentiated cells. Based on their findings, it was concluded that both CSCs and differentiated cells express CD133 however, it is likely that AC133 is masked during cell differentiation. Furthermore, the consensus was that AC133 could be used to characterise and isolate CSCs when used with caution.

Therefore, other markers have been used in combination with CD133 to better identify colorectal CSCs. These markers include amongst others, CD44 (Dalerba *et al.*, 2007b; Vermeulen *et al.*, 2008; Du *et al.*, 2008; Chu *et al.*, 2009, Haraguchi *et al.*, 2008; Chen *et al.*, 2011), EpCAM (Langan *et al.*, 2012; Melin *et al.*, 2012) and CD166 (Vermeulen *et al.*, 2008; Fang *et al.*, 2010), and these are discussed below.

CD44 was initially used as a biological marker for the identification of breast CSCs, and it was later proposed to be a marker of colorectal CSCs (Al Hajj *et al.*, 2003). It is a surface adhesion molecule involved in cell-cell and cell-matrix interactions through its high affinity for binding hyaluronic acid and various components of the extracellular matrix (ECM), such as collagen and fibronectin (Naor *et al.*, 1997; Sahlberg *et al.*, 2014). The *CD44* gene contains twenty exons, ten of which remain unchanged. The remaining ten exons are subject to alternative splicing and are referred to as variable exons (v1-v10) (Naor *et al.*, 1997; Sahlberg *et al.*, 2014). Standard CD44 contains only the first- and last five- unchanged exons; however, the utilisation of the variable exons gives rise to many splice variants containing different combinations of v1-v10. The functional role of these splice variants is poorly understood; although, certain cancers including gastrointestinal cancers, express variant CD44 isoforms that are thought to mediate tumour metastasis (Du

et al., 2008; Banky *et al.*, 2012). Thus, the existence of multiple CD44 splice variants needs to be considered as this questions its use as a reliable CSC marker.

EpCAM, also known as ESA, is used in combination with CD133 to characterise putative colorectal CSCs (Fang *et al.*, 2010). EpCAM was discovered in 1979, and it is a 40kDa glycoprotein involved in cell-cell adhesion, cell-signalling, proliferation, migration and differentiation (Trzpis *et al.*, 2007; Schnell *et al.*, 2013). It is a promising marker used to characterise putative CSCs, as its overexpression has been correlated with reduced overall survival in patients with breast (Spizzo *et al.*, 2004), gall bladder (Varga *et al.*, 2004), urothelial (Brunner *et al.*, 2008) and colorectal (Patriarca *et al.*, 2012) cancers. It consists of three regions, an extracellular region containing three domains with several glycosylation sites, a transmembrane region and a cytoplasmic region known as EpCAM intracellular domain (EpICD) (Figure 1.7) (McLaughlin *et al.*, 1998; Dyer, 1999; Chong & Speicher, 2001; Sievers *et al.*, 2001; Dillman, 2002; Munz *et al.*, 2009). EpICD functions as an oncogenic signal transducer in the WNT signalling pathway which is activated following its cleavage by proteases and its release into the cytoplasm. EpICD then forms a complex with β-catenin, Lymphoid enhancer-binding factor (LEF) and Four and a half LIM domains protein 2 (FHL2) which is translocated to the nucleus where it activates the transcription of genes involved in cell adhesion, proliferation and migration (Maetzel *et al.*, 2009; Munz *et al.*, 2009).

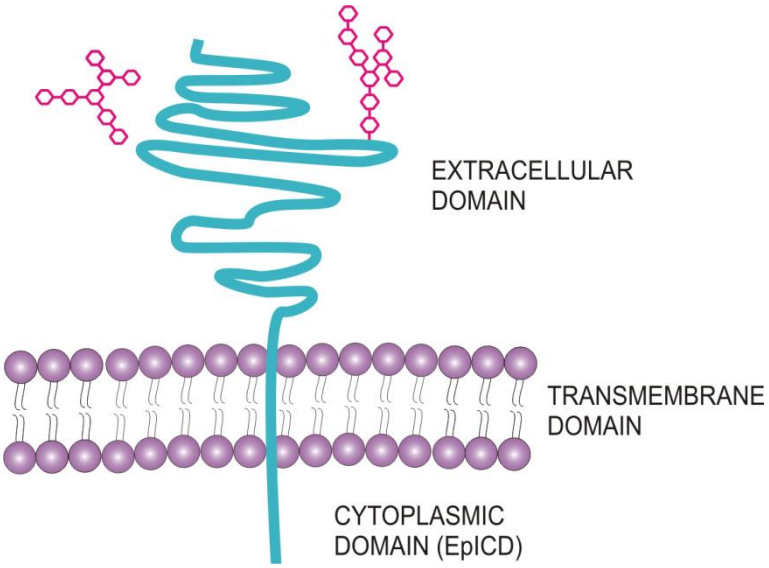

Figure 1.7. The glycoprotein EpCAM showing its extracellular, transmembrane and cytoplasmic domains (adapted from Figure 1 in Armstrong & Eck, 2003, 321 p.).

CD166, also known as activated leukocyte cell adhesion molecule (ALCAM) is a 105 kDa transmembrane protein, and as its alternative name suggests, it mediates cell-cell adhesion (Weidle *et al.*, 2010). CD166 is a member of the immunoglobulin superfamily of receptors. It consists of three domains, a short cytoplasmic domain, a transmembrane domain and an extracellular domain that contains five immunoglobulin-like domains referred to as VVC2C2C2 (Piazza *et al.*, 2005; Weidle *et al.*, 2010; Anderson *et al.*, 2011). This protein is expressed on a number of different cells including activated T cells and monocytes, epithelial cells, endothelial cells and fibroblasts although, its expression is usually restricted to cellular subsets involved in tissue maintenance, development and migration (Bowen *et al.*, 1995; Swart, 2002). Furthermore, it has been shown to mark the SC niche in the colonic crypt and therefore, it is used as a marker of colorectal CSCs (Levin *et al.*, 2010).

Chen *et al.* (2011) evaluated the expression of the surface markers CD44 and CD133 in the colorectal carcinoma cell line HCT116. They hypothesised that the use of both markers would further enrich for a tumour-initiating cell population. Four subpopulations were identified, these being: $CD44^{+}/CD133^{+}$, $CD44^{+}/CD133^{-}$, $CD44^{-}/CD133^{+}$ and $CD44^{-}/CD133^{-}$. They evaluated the self-renewal potential of these subsets using a colonosphere

assay and reported that the $CD44^+/CD133^+$ subset was the most efficient at generating colonospheres.

Similarly, Wang *et al.* (2012) enriched a tumour initiating subpopulation of cells using antibodies to the surface markers CD44 and CD133. They isolated four cellular subsets from the colorectal cancer cell line SW620 and showed that the subset characterised as $CD44^+/CD133^+$ formed colonospheres the most efficiently and with the highest proliferation rate; whereas, the $CD44^-/CD133^-$ subset was the least efficient with the lowest proliferation rate. Furthermore, the $CD44^+/CD133^+$ subset was more efficient at forming colonospheres with a higher proliferation rate when compared to the $CD44^+/CD133^-$ subset, and the $CD44^+/CD133^-$ subset was more efficient at forming colonospheres with a higher proliferation rate when compared to the $CD44^-/CD133^+$ subset. From the results of these two studies, it was suggested that CD44 contributes to colony formation and proliferation in the HCT116 and SW620 cell lines, and the addition of this marker further enriches for cells with a stem-like phenotype.

Dalerba *et al.* (2007b) showed that a subpopulation of colorectal cancer cells with stem-like features could be isolated based on CD44 and EpCAM surface marker expression. Flow cytometry analyses of these markers revealed that two distinct populations were present in tumours derived from six patients with primary colorectal cancer, these being $EpCAM^{high}/CD44^+$ and $EpCAM^{low}/CD44^-$. A subsequent tumourigenicity assay in immunodeficient mice showed that the cellular subset characterised as $EpCAM^{high}/CD44^+$ was able to elicit tumour formation whilst maintaining a differentiated phenotype. Furthermore, phenotypic heterogeneity resembling the primary tumour was maintained. Further screening of markers in the $EpCAM^{high}/CD44^+$ cellular subset showed that CD166 was co-expressed with these cells. To determine whether it could be used as an additional marker to identify cells with stem-like properties, a tumourigenicity assay was again utilised. The results showed tumourigenicity resided in the $CD44^+/CD166^+$ subset indicating that CD166 could be used as an additional marker to further enrich for colorectal CSCs.

In contrast to the above mentioned studies, Muraro *et al.* (2012) evaluated the stem-like features of putative CSC subpopulations isolated from ten established colorectal cancer

cell lines, and reported that neither of the putative CSC subpopulations showed stem-like characteristics. They isolated putative CSC subpopulations with the following expression phenotypes: $CD133^+$, $CD166^+/CD144^+$ and $CD44^+/CD24^+$ however, all three cell subpopulations failed to elicit tumour formation in immunodeficient mice and further lacked increased resistance to the chemotherapeutic drugs 5-Fluorouracil, Oxaliplatin and Irinotecan. It was suggested that putative CSCs isolated from established cell lines might not show "stemness" features as they lack the microenvironment found *in vivo*. This is important as signals from the microenvironment modulate marker expression and tumourigenicity; therefore, the expression profile of putative CSCs resident in solid tumours could differ to those present in established cell lines cultured *in vitro* (Visvader & Lindeman, 2012). To illustrate the effect of the tumour microenvironment on tumourigenicity, Rao *et al.* (2012) showed that murine tumour-associated macrophages enhanced the tumourigenicity potential of $CD44^+$ colorectal tumour cells, as these cells produced increased levels of osteopontin, a protein that has been found to be overexpressed in a number of cancers, including colorectal cancer (Li *et al.*, 2012a).

It should be noted that the detection of surface marker expression in both cell lines and tumours is dependent on the type of antibody used, as different antibodies recognise different epitopes. Moreover, the presence of different splice variants implies that some epitopes may be lost and would thus be undetectable. With regards to CD133 detection, it was demonstrated by Kemper *et al.* (2010) that glycosylation can alter the three-dimensional structure of CD133 thus masking the AC133 epitope. Therefore, cells positive for CD133 mRNA and protein expression might not be characterised as positive depending on the antibody used.

To further demonstrate the significance of using different antibodies, Sahlberg *et al.* (2014) utilised three different antibodies to determine the percentage of DLD1, HCT116 and HT29 cells that express CD24. It was shown in the DLD1, HCT116 and HT29 cell lines that 28%, 20% and 65% of these cells expressed CD24, respectively when the clone ML5 was used; whereas, these cells expressed 6%, 8% and 95% of CD24, respectively when the clone 32D12 was used. CD24 was not detected in either of these cell lines when the clone SN3 A5 2H10 was used. These data demonstrate the use of different antibodies result in ambiguous data, which makes it difficult to compare expression profiling studies.

In addition to the use of alternatively sourced antibodies resulting in disparate data, the use of cell lines and different techniques are also contributing factors. Chen *et al.* (2011) reported that more than 60% of HCT116 cells co-expressed CD44 and CD133; whereas, Wang *et al.* (2012) reported that only 14.8% of these cells were CD44 and CD133 double-positive. Again this may be attributed to the use of different antibodies, but may also be attributed to the use of different culturing techniques. In the study by Wang *et al.* (2012), Western blot analysis of CD133 protein expression in the DLD1 cell line showed the protein was absent however, flow cytometry analysis of surface marker expression showed that 12.00 ± 0.38% of these cells expressed CD133, which demonstrates that the use of different molecular techniques leads to different results. Together, these studies illustrate the difficulties involved when comparing studies amongst researchers. Therefore, there is an increasing need to identify universal markers to better characterise CSCs and to elucidate their cellular functions in both normal and cancerous intestinal epithelia.

1.5 Signalling Pathways and Transcription Factors Maintain Pluripotency

CSCs can be further characterised based on the activity of extrinsic signalling pathways and the expression of intrinsic pluripotent transcription factors. Pluripotency of embryonic stem cells (ESCs) is defined by the ability of these cells to produce cells of the three germinal layers (ectoderm, mesoderm and endoderm, excluding extra embryonic tissue) and gametes, through the process of differentiation (Medvedev *et al.*, 2008). Pluripotency of human embryonic and adult SCs such as CSCs is maintained by the signalling pathways WNT, BMP, TGFβ, NOTCH and HEDGEHOG, as well as by the transcription factors NANOG, Octamer-4 (OCT4), Sex determining region Y (SRY)-related high mobility group (HMG) box 2 (SOX2), Kruppel-like factor 4 (KLF-4) and V-myc avian myelocytomatosis viral oncogene homolog (c-MYC) (Figure 1.8) (Medvedev *et al.*, 2008). The expression of these signalling pathways and pluripotency associated transcription factors is modulated by epigenetic mechanisms that include DNA methylation, chromatin remodelling, PcG proteins and miRNAs (Orkin & Hochedlinger, 2011). Aberrant epigenetic modulation of these pluripotency associated signalling pathways and transcription factors drives colorectal tumourigenesis (Berdasco & Esteller, 2010; Munoz *et al.*, 2012).

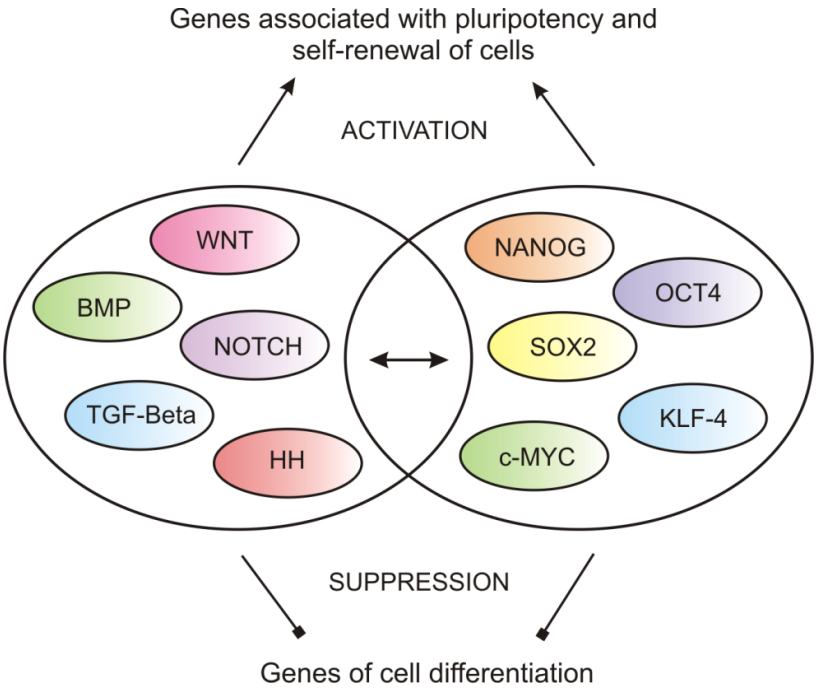

Figure 1.8. A representation of the cross-talk that occurs between key signalling pathways (left) and transcription factors (right) involved in maintaining pluripotency in human ESCs. These pathways and transcription factors activate the transcription of genes involved in the maintenance of ESC proliferation and self-renewal. In promoting pluripotency, the transcription of genes involved in cell differentiation are suppressed (adapted from Figure 1 in Medvedev *et al*., 2008, 1378 p.).

1.5.1 Extrinsic Signalling Pathways

WNT signalling maintains pluripotency in both mouse and human ESCs, and in adult SCs located in the colonic crypt, through the binding of WNT ligands to the Frizzled/LDL receptor-related protein (LRP) co-receptor (Roy & Majumdar, 2012). Activation of the Frizzled/LRP co-receptor leads to the recruitment of AXIN to the plasma membrane, which disrupts the destruction complex. Additionally, Dishevelled is activated via phosphorylation which further inhibits GSK-3 activity, a component of the destruction complex (Zeng *et al.*, 2008; Roy & Majumdar, 2012). β-catenin accumulates in the cell cytoplasm and is translocated to the nucleus where it associates with T-cell factor (TCF)/LEF, and together they activate the transcription of genes involved in maintaining pluripotency and cell proliferation (Figure 1.9) (Reya *et al.*, 2003; Shimizu *et al.*, 2008).

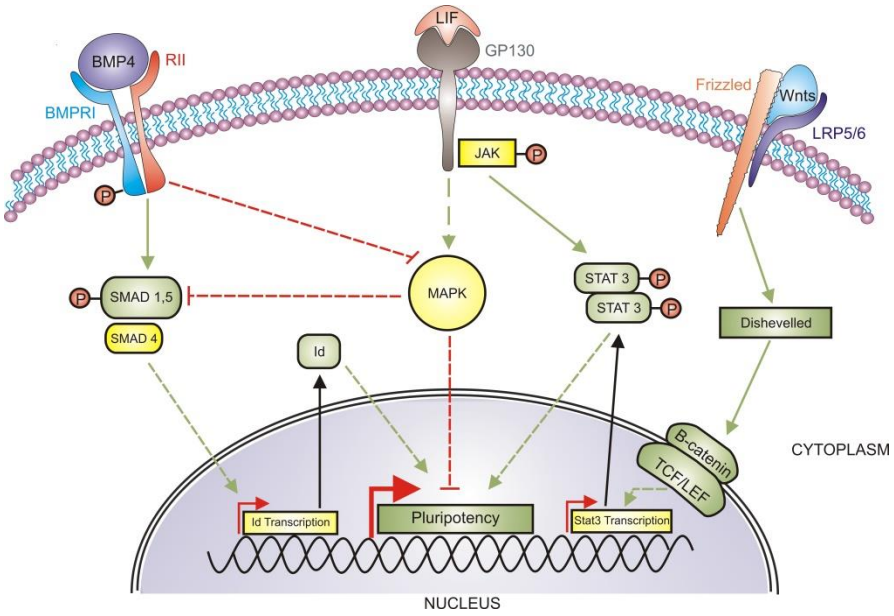

Figure 1.9. The synergistic interaction between BMP and LIF signalling, as well as WNT signalling, in maintaining pluripotency in mouse ESCs (adapted from Figure 6 in Hao *et al.*, 2006, 89 p.).

Within the colonic crypt there is a gradient of WNT activity along the axis required for compartmentalisation of the various cell types comprising it (Figure 1.10). WNT activity is highest in the SC compartment where it maintains pluripotency and SC proliferation, and it is gradually reduced along the axis towards the top of the crypt where the differentiated cells are located (Tesori *et al.*, 2013). The aberrant activation of WNT signalling in adult SCs located in the colonic crypt is associated with colorectal tumourigenesis (Roy & Majumdar, 2012). Gain-of-function mutations in the gene encoding *β-catenin* (*CTNNB1*), or loss-of-function mutations in *APC* are well characterised in colorectal tumourigenesis (Barker & Clevers, 2006; Fodde & Brabletz, 2007; Vermeulen *et al.*, 2010). The enhanced stabilisation of β-catenin results in constitutive WNT activity that designates the colorectal CSC population (Vermeulen *et al.*, 2010). Aberrant epigenetic modulation of the genes encoding WNT signalling players also drives colorectal tumourigenesis. DNA hypermethylation of the WNT inhibitors *SFRP2*, *WIF-1*, *DKK1*, *DKK2*, *DKK3*, *DKK4* and *SOX17*, as well as *APC* and *AXIN*, are silenced in colon cancer (Suzuki *et al.*, 2004; Taniguch *et al.*, 2005; Ying & Tao, 2009). Furthermore, aberrant gene silencing by

miRNA-145; -135a; -135b; -139 and -145 is also associated with colorectal tumourigenesis (Schnekenburger & Diedrich, 2012; Voorham *et al.*, 2013).

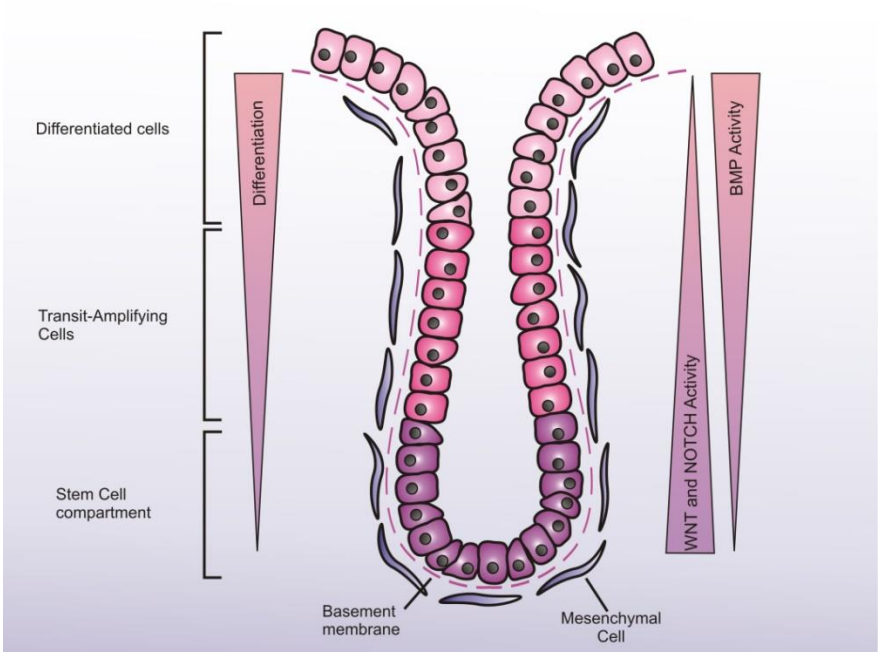

Figure 1.10. The colonic crypt showing the lower SC compartment (purple), the transit-amplifying cells (dark pink) and the differentiated cells (light pink) located in the upper third of the crypt. Indicated on the right are the WNT, BMP and NOTCH signalling gradients that are required for compartmentalisation (adapted from Figure 1 in Anderson *et al.*, 2011, 321 p.).

Like WNT, BMP signalling maintains pluripotency in both mouse and human ESCs however, with an opposite effect. In mouse ESCs, Bone morphogenetic protein-4 (BMP4) together with Leukaemia inhibitory factor (LIF) maintains pluripotency. BMP4 binds to type I (BMPR1A/B) and type II receptors that phosphorylate SMAD1 and 5 (R-SMAD) (Massagué *et al.*, 2005; Retting *et al.*, 2009). R-SMAD associates with SMAD4 and together this complex translocates to the nucleus where it activates a set of inhibitor of differentiation (*Id*) genes to prevent SC differentiation (Figure 1.9) (Medvedev *et al.*, 2008). Upon LIF binding to the GP130 receptor, Janus kinase protein (JAK-P) phosphorylates Signal transducer and activator of transcription 3 (STAT3), a transcription factor that enters the cell nucleus where it activates the transcription of genes required to

maintain pluripotency (Hao *et al.*, 2006). The binding of LIF to GP130 also activates MAPK, which is inhibited by BMP4 receptor binding (Figure 1.9). Therefore, in the absence of BMP4 signalling, activated MAPK activates SMAD1 and 5, which inhibit the activation of *Id* genes leading to cell differentiation (Hao *et al.*, 2006; Medvedev *et al.*, 2008).

Unlike mouse ESCs, LIF is not associated with human ESCs and the activation of BMP signalling leads to cell differentiation as opposed to the maintenance of pluripotency (Medvedev *et al.*, 2008). Rather, fibroblast growth factor (FGF) signalling together with a balance of BMP and TGF-β signalling is required to maintain pluripotency in these cells (James *et al.*, 2005; Wang *et al.*, 2005; Xu *et al.*, 2005). Upon FGF binding to its receptor, the PI3K pathway is activated which further activates AKT via phosphorylation. Activated AKT is able to maintain pluripotency as it promotes cell proliferation and survival (Hammerman *et al.*, 2005; Kim *et al.*, 2005). BMP signalling promotes differentiation upon receptor activation of SMAD1 and 5 whereas, TGF-β signalling maintains pluripotency upon receptor activation of SMAD2 and 3 (Chen *et al.*, 1998; Hao *et al.*, 2006).

Within the colonic crypt, a gradient of BMP activity is maintained by the expression of NOGGIN, an antagonist of BMP signalling. NOGGIN is expressed in the intestinal SC compartment where BMP activity is diminished to allow for cell proliferation and the maintenance of pluripotency (Figure 1.10) (Amit *et al.*, 2000; Xu *et al.*, 2005). Furthermore, the overexpression of NOGGIN enhances the translocation of cytoplasmic β-catenin into the nucleus suggesting there is cross-talk between WNT and BMP signalling (Rao & Wang, 2010).

In colorectal cancer, BMP signalling is impaired by the epigenetic silencing of the BMPRII receptor. The silencing of this receptor is mediated by PRC2 wherein its catalytic domain EZH2 tri-methylates H3 at lysine 27 (H3K27me3) resulting in the transcriptional repression of *BMPRII* and enhanced tumourigenicity (Parsons *et al.*, 1995; Krausova & Korinek., 2012). Furthermore, DNA methylation of the BMP2 ligand is associated with colorectal tumourigenesis (Wen *et al.*, 2006; Kodach *et al.*, 2008; Munoz *et al.*, 2012).

NOTCH signalling is important for maintaining pluripotency of adult SCs located in the colonic crypt where it also regulates the differentiation of crypt cells (Roy & Majumdar, 2012). Its signalling is mediated by the binding of Delta-like (DLL) and Jagged ligands to the extracellular domains of the NOTCH-1, -2, -3 and -4 transmembrane receptors (Krausova & Korinek, 2012). The intracellular region of NOTCH is cleaved by proteases to produce the NOTCH intracellular domain. This is translocated to the nucleus where it associates with recombination signal binding protein J kappa (RBP-Jκ) to activate the transcription of downstream target genes such as *Hairy and enhancer of split* (*HES*), *Hairy ears, Y-linked* (*HEY*) and *Hairy/enhancer-of-split related with YRPW motif-like* (*HEYL*) that repress genes involved in cell lineage differentiation (Katoh & Katoh, 2007; Roy & Majumdar, 2012).

Within the colonic crypt, a gradient of NOTCH signalling is maintained along the crypt axis. NOTCH signalling is highest in the SC compartment and is gradually reduced along the axis towards the top of the crypt where differentiated cells reside (Fig 1.9) (Roy & Majumdar, 2012). The dysregulation of NOTCH activity, as seen by an upregulation of *NOTCH1* and *HES1*, has been reported in colon adenocarcinomas (Reedjik *et al.*, 2008) however, the epigenetic mechanisms governing this pathway are poorly understood.

Hedgehog (HH) signalling is important for the maintenance and proliferation of embryonic and adult SCs. HH ligands bind to the 12-transmembrane receptor Patched (PTCH) resulting in the de-repression of the 7-transmembrane receptor Smoothened (SMO). SMO then mediates a cascade of downstream events leading to the activation of the Glioblastoma (GLI) transcription factors GLI1, GLI2 and GLI3 that further activate the transcription of HH target genes involved in maintaining pluripotency (Varjosalo & Taipale, 2008). Furthermore, there is cross-talk between this pathway and the BMP pathway in the colonic crypt (Bertrand *et al.*, 2012). Transit amplifying cells secrete HH ligands that associate with PTCH receptors expressed on mesenchymal cells. This induces the production of BMP ligands that mediate BMP signalling (Madison *et al.*, 2005).

While the role of HH signalling in colorectal carcinogenesis has remained quite controversial due to discordant data, Varnat *et al.* (2009) have nevertheless elucidated a functional role for HH-GLI signalling in early and advanced colorectal cancers. They

reported that *GLI1*, *PTCH1* and *Sonic Hedgehog* (*SHH*) are expressed in primary and metastatic colorectal cancer samples and that *GLI1* expression was enriched in the $CD133^+$ putative CSC population of metastatic cancers. Furthermore, they showed HH-GLI signalling is required for $CD133^+$ CSC self-renewal and survival *in vivo* and additionally that high HH-GLI is correlated with metastatic behaviour.

DNA hypermethylation of the *PTCH* and *Hedgehog interacting protein* (*HHIP*) promoters, the latter an antagonist of HH, both enhance HH-GLI signalling, and have been implicated in the development of various cancers including colorectal cancer (Taniguchi *et al.*, 2007). Also, the loss of miR-324-5p that normally functions to suppress *GLI1* expression, has also been shown to promote tumourigenesis (Ferretti *et al.*, 2008).

Altogether, these signalling pathways intricately maintain the pluripotency of adult SCs located in the colonic crypt and their dysregulation through epigenetic modulation drives colorectal carcinogenesis.

1.5.2 Intrinsic Transcription Factors

The core transcription factors NANOG, OCT4 and SOX2 are central to maintaining pluripotency in embryonic and adult SCs, although other transcription factors such as KLF-4 and c-MYC are also involved (Niwa *et al.*, 2000; Avilion *et al.*, 2003; Mitsui *et al.*, 2003; Rodda *et al.*, 2005). As they are widely expressed in many cancers, including colorectal cancer, this suggests that these "stemness" associated transcription factors play an essential role in driving colorectal carcinogenesis (Ben-Porath *et al.*, 2008; Liu *et al.*, 2013).

NANOG, often referred to as a "master gene", was first described by Wang *et al.* in 2003. It is a homeodomain transcription factor that is specifically expressed in embryonic and adult SCs where it maintains self-renewal capability (Tai *et al.*, 2005; Liu *et al.*, 2013). In ESCs, the down-regulation of *NANOG* is associated with cell differentiation whereas, in adult and cancer SCs, the up-regulation or down-regulation of *NANOG* is associated with differentiation (Wang *et al.*, 2003; Hyslop *et al.*, 2005; Zaehres *et al.*, 2005; Alvarez *et al.*, 2010; Kallas *et al.*, 2014). Alvarez *et al.* (2010) showed that the overexpression of *NANOG* in human mesenchymal SCs resulted in their differentiation into neural-like cells,

as seen by the expression of the neural cell lineage markers βIII-tubulin and glial fibrillary acidic protein. In a study by Kallas *et al.* (2014), they showed that NANOG protein expression was down-regulated in the human ESC line H9 following differentiation induced by sodium butyrate, an inhibitor of HDAC activity, whereas in the human embryonal carcinoma-derived (hEC) cell line, it remained high. Based on their findings, it was suggested that *NANOG* is regulated differently in "normal" ESCs compared to hEC cells.

In contrast to *NANOG*, small changes in *OCT4* expression, whether it be up-regulation or down-regulation, can induce ESC differentiation therefore, suggesting that its expression is tightly regulated (Niwa *et al.*, 2000; Matin *et al.*, 2004; Niwa *et al.*, 2005; Zaehres *et al.*, 2005; Rodriguez *et al.*, 2007; Kallas *et al.*, 2014). OCT4, also known as OCT3/4 or POU5F1, is a homeodomain transcription factor belonging to the POU family (Tai *et al*., 2005). OCT4 can heterodimerise with SOX2, a member of the SRY-related HMG box (SOX) family, and together they regulate the transcription of *NANOG* and other downstream target genes needed to maintain pluripotency (Boyer *et al*., 2005; Medvedev *et al.*, 2008). Altered expression of *SOX2* is also associated with cell differentiation (Chew *et al.*, 2005; Masui *et al.*, 2007; Kopp *et al.*, 2008; Adachi *et al.*, 2010; Kallas *et al.*, 2014). Similarly to *NANOG*, Kallas *et al.* (2014) reported that OCT4 protein expression was down-regulated upon H9 cell line differentiation, whereas it remained high in the hEC cell line. This again suggests that *OCT4* is also regulated differently in "normal" ESCs compared to hEC cells. Interestingly, they reported that SOX2 protein expression remained the same following ESC treatment with sodium butyrate. This led them to conclude that *NANOG* and *OCT4* are regulated at an epigenetic level whereas, other mechanisms such as post-translational modifications, regulate SOX2 expression in H9 human ESCs.

Since NANOG, OCT4 and SOX2 are widely expressed in colorectal tumours, they may play an essential role in maintaining CSC pluripotency to drive colorectal carcinogenesis (Shi *et al.*, 2012; Amini *et al.*, 2014). Shi *et al.* (2012) evaluated the expression of these transcription factors in the colorectal cancer cell lines SW480, SW620 and HT29. They also assessed their expression in putative CSC-enriched tumour spheres. Although all three transcription factors were expressed in all of the cell lines, the majority of their

expression was confined to their corresponding CSC-enriched tumour spheres. A subsequent knock-down study of *NANOG* in the SW620 cell line showed that these cells formed tumour spheres less efficiently, they proliferated less efficiently *in vitro* and had reduced tumourigenicity *in vivo*. Their findings emphasise the importance of these transcription factors, in particular NANOG, in maintaining CSC tumourigenicity in colorectal cancer.

KLF-4 and c-MYC are two additional transcription factors that maintain pluripotency in embryonic and adult SCs. KLF-4 (Kruppel-like factor 4) is a zinc finger transcription factor that functions to maintain cell proliferation, differentiation and apoptosis (Shi & Ai, 2013). KLF-4 has also been shown to be an upstream regulator of *NANOG* via its direct binding to its promoter (Chan *et al.*, 2009). Moreover, its involvement in maintaining pluripotency in breast (Yu *et al.*, 2011) and colon CSCs (Leng *et al.*, 2013) has been shown (Leng *et al.*, 2013). This is considered in relation to colon CSCs below.

Leng *et al.* (2013) investigated whether *KLF-4* expression was essential for maintaining self-renewal capability and tumourigenicity in DLD1 colorectal CSCs. Firstly, they cultured the DLD1 cells under special culture conditions to enrich for putative CSC spheroids and confirmed that these spheroids expressed the CSC associated markers *CD133*, *CD166, LGR5*, *ALDH1*, *NANOG*, *OCT4* and *SOX2*. They observed that the expression levels of these markers were higher in the putative CSC spheroids than in the DLD1 parental cell line. Next, they silenced *KLF-4* expression in these spheroids using a lentiviral vector bearing *KLF-4* short hairpin RNA (shRNA). Knock-down of *KLF-4* reduced the expression of *CD133, CD166, LGR5* and *ALDH1*. Importantly, the stem-like characteristics of these cells were altered, as they were less invasive and had reduced migration capability *in vitro*. Furthermore, they were less tumourigenic both *in vitro* and *in vivo* and had increased pharmacological susceptibility to 5- Fluorouracil.

c-MYC, a member of the MYC family of transcription factors, maintains pluripotency via the transcription of its many downstream target genes. Its role in maintaining pluripotency has been demonstrated by its ability to reprogram mouse embryonic fibroblasts into ES cell-like cells, termed induced pluripotent SCs (iPSCs) or *in vitro* engineered cells, together with *Oct4*, *Sox2* and *Klf-4* (Takahashi & Yamanaka, 2006). Its putative role in

maintaining pluripotency and tumourigenicity has also been shown by Wang *et al.* (2008). Wang *et al.* (2008) elucidated the functional role of *c-MYC* in glioma $CD133^{+}$ cells isolated from primary tumours by investigating the effect(s) of *c-MYC* knock-down in these cells. They reported that its knock-down impaired the ability of these cells to form neurospheres *in vitro* indicating that *c-MYC* is important for maintaining self-renewal capability. Additionally, *c-MYC* regulated cell proliferation and survival in $CD133^{+}$ glioma cells, as its knock-down reduced the number of cells that entered the S-phase of the cell cycle, and thus reduced their survival *in vitro*. Lastly, *c-MYC* knock-down impaired the ability of $CD133^{+}$ glioma cells to initiate tumourigenesis *in vivo*, as these cells failed to form tumours in immunodeficient mice. Altogether, their findings demonstrate the importance of *c-MYC* expression in maintaining CSC pluripotency and tumourigenicity.

1.5.2.1 Epigenetic Regulation of Pluripotent Transcription Factors

The epigenetic modulation of pluripotency associated transcription factors in cancer remains poorly understood. In embryonic and CSCs, the promoters of *NANOG* and *OCT4* have been shown to be methylated upon cell differentiation (Farthing *et al.*, 2008; Meissner, 2010; Wang *et al.*, 2013b). Wang *et al.* (2013b) explored the methylation status of the *NANOG* promoter in $CD133^{+high}$ putative colorectal CSCs isolated from the HCT116 cell line and reported that not only was its methylation reduced compared to the parental cell line, but that its expression was increased. Their results demonstrate the role that DNA methylation has on modulating *NANOG* expression in CSCs which may drive colorectal tumourigenesis.

The polycomb repressive complex 2 (PRC2) has also been shown to modulate NANOG expression in iPSCs (Villasante *et al.*, 2011). Villasante *et al.* (2011) reported that two NANOG expressing subpopulations were present in iPSCs, these being NANOG-high and NANOG-low. In the NANOG-high population it was observed that EZH2 expression was absent and was associated with low H3K27me3 at the promoter, whereas in the NANOG-low population, EZH2 expression was high and correlated with high H3K27me3. Chromatin immunoprecipitation (ChIP) showed the presence of the histone mark H3K4me3 which is associated with gene transcription; and H3K27me3 which is associated with gene silencing at the *NANOG* promoter. Based on these data, it was thought that the

NANOG promoter was bivalent; however, a further study using ChIP revealed that the NANOG-high population was associated with the H3K4me3 mark and the NANOG-low population was associated with the H3K27me3 mark. These findings provide good evidence for the regulation of *NANOG* by EZH2 in the maintenance of pluripotency.

In order to elucidate the mechanisms involved in the epigenetic modulation of pluripotency associated transcription factors in cancer, Wang *et al.* (2013b) evaluated the epigenetic status of *NANOG*, *OCT4*, *SOX2*, *KLF-4* and *c-MYC* in a variety of cancer cell lines. They characterised epigenetic signatures based on promoter methylation status and on the status of the histone marks H3K4me3 and H3K27me3. Some of these findings are presented below.

The *NANOG* and *OCT4* promoters, and exon 3 of *c-MYC*, were hypomethylated in the primary hepatocellular carcinoma (HCC) cell line PLC and in the metastatic HCC cell line 97L when compared to the normal hepatocyte cell line LO2 whereas, they were all found to be hypermethylated in the colon carcinoma cell line HCT116. The *SOX2* and *KLF-4* promoters were unmethylated in all of these cell lines. Altogether, this suggested a differential methylation profile of *NANOG*, *OCT4* and *SOX2* in HCC *versus* colorectal carcinoma cell lines. They further correlated the expression of *NANOG*, *OCT4*, *c-MYC* and *KLF-4* with methylation status and reported that the expression of all these genes was high in the 97L cell line but low in the HCT116 cell line, with the exception of *KLF-4*, when compared to their expression in the LO2 and PLC cell lines. Their results suggested that DNA methylation negatively regulates *NANOG*, *OCT4* and *c-MYC* expression.

Chromatin mark analyses using ChIP showed H3K4me3 was highly associated with the *NANOG* and *OCT4* promoters and exon 3 of *c-MYC* in 97L cells compared to LO2 and HCT116 cells, whereas the histone mark H3K27me3 was low in all of these cells. With regards to the *KLF-4* promoter, H3K4me3 was high in the 97L cell line and moderate in the HCT116 cell line; whereas H3K27me3 was lower in the HCT116 cell line than in the 97L cell line, which accounts for the high expression of *KLF-4* observed in both these cells. Together their results show that DNA methylation and histone modifications modulate pluripotent gene expression.

1.6 Implications of the CSC model and Therapeutics

Current anti-cancer therapies target the bulk of the tumour mass but they fail to eradicate the CSC population as these cells express multidrug resistance transporters that provide them with the ability to cope with toxic chemotherapeutics (Botchkina, 2013). Other resistant mechanisms used by these cells, include slow cycling, the presence of highly efficient DNA repair systems and the expression of anti-apoptotic proteins (Woodward & Bristow, 2009; Pajonk *et al.*, 2010). If the CSC model is correct, these CSCs would be able to recapitulate the tumour heterogeneity through self-renewal and differentiation, resulting in relapse after therapy (Banerji & Los 2006; Dalerba *et al.*, 2007a; Rashedi *et al.*, 2007; Vermeulen *et al.*, 2008; Meacham & Morrison, 2013; Puglisi *et al.*, 2013; Easwaran *et al.*, 2014). Therefore, it is critical to develop new therapies that target and eradicate the CSC population. Although CSC targeted therapies will not eliminate the tumour mass, they will prevent tumour growth, metastasis and relapse (Figure 1.11) (Reya *et al.*, 2001; Dalerba *et al.*, 2007b; Vermeulen *et al.*, 2008).

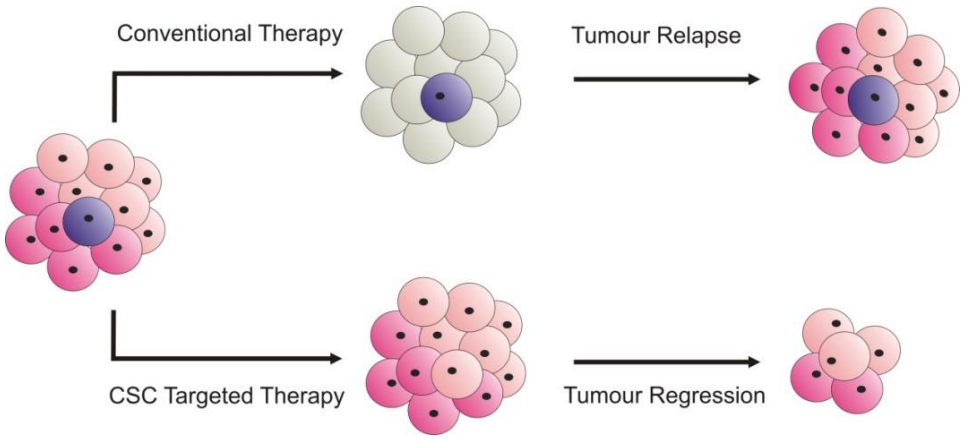

Figure 1.11. A representation of CSC targeted therapy *versus* conventional therapy. Conventional therapy targets the bulk of the tumour (pink) but is unable to eliminate CSCs (purple) leading to tumour relapse. CSC therapy aims to target the CSC population to result in tumour regression. An effective therapeutic would be to combine CSC targeted therapy with conventional therapy (adapted from Figure 3 in Dalerba *et al.*, 2007a, 279 p.).

Based on the roles of extrinsic signalling pathways and intrinsic transcription factors in maintaining pluripotency, and that epigenetic mechanisms modulate their expression in

embryonic and adult SCs, it is thought that targeting these pathways and transcription factors in CSCs could inhibit their activity and induce cell differentiation to ultimately sensitise these cells to current anti-cancer therapies (Gudas & Wagner, 2011; Cai & Zhu, 2012; Munoz *et al.*, 2012; Chen *et al.*, 2013). Therapeutic strategies include the use of small-molecules that either block or promote the activity of pluripotency associated signalling pathways; and the use of epigenetic drugs that target the enzymes involved in epigenetic modulation, namely histone deacetylases (HDACs) and DNA methyltransferases (DNMTs) (Zhang *et al*., 2012).

The non-steroidal anti-inflammatory drug Celecoxib has been shown to inhibit β-catenin-dependent transcription in colorectal cancer cells (Lepourcelet *et al.*, 2004; Tuynman *et al.*, 2008). Disruption of β-catenin activity results in the inhibition of cell proliferation and induces apoptosis (Anastas & Moon, 2013; Puglisi *et al.*, 2013). Differentiation therapy involving the use of BMP4 to stimulate BMP signalling activity has been shown to induce the differentiation of these cells *in vitro* and increase the sensitivity of $CD133^+$ colorectal adenocarcinoma CSCs to the chemotherapeutic drugs 5-Fluorouracil and Oxaliplatin *in vivo* (Lombardo *et al.*, 2011).

The use of anti-cancer epigenetic drugs, namely HDAC inhibitors and DNMT inhibitors, can be utilised in cancer differentiation therapy (Peedicayil, 2006). HDAC inhibitors are classified into four groups, namely the short chain fatty acids (sodium butyrate, Valproic acid), hydroxamic acids (Vorinostat, Panobinostat and Trichostatin A), cyclic tetrapeptides and benzamides (Yoo & Jones, 2006). The exact mechanism by which these inhibitors mediate their activity remains unknown; although, it is known that Vorinostat inhibits histone deacetylation by binding to the enzyme's catalytic domain. Therefore, it has been suggested that other hydroxamic acids inhibit histone deacetylation in a similar manner (Yoo & Jones, 2006).

DNMT inhibitors are classed as either nucleoside analogues (Zebularine, 5-azacytidine and 5-azadeoxycytidine) or non-nucleoside analogues, where DNA methylation is inhibited by the direct binding of non-nucleoside analogues to the catalytic domain of DNMT (Yoo & Jones, 2006). Antisense oligonucleotides, complementary to mRNA, comprise a third class of DNMT inhibitors (Peedicayil, 2006). Two potentially useful epigenetic

therapeutics, Valproic acid (HDAC inhibitor) and Zebularine (DNMT inhibitor) are described below, as the *in vitro* effects of both are assessed in the current study.

Valproic acid (VPA, 2-propylpentanoic acid) is a well-established short chain fatty acid drug used for the treatment of epilepsy and bipolar disorder (Löscher, 2002). The exact mechanism by which this drug mediates its activity is not well understood. It is however, believed to act by increasing gamma-aminobutyric acid levels in the brain, or alternatively, it may alter the properties of voltage gated ion channels (McLean & MacDonald, 1986; Owens & Memeroff, 2003). VPA has anti-tumour properties and is able to inhibit the growth of various cancer cells, promote cell differentiation, induce apoptosis, and inhibit angiogenesis (Blaheta & Cinatl, 2002; Kawagoe *et al.*, 2002; Phillips *et al.*, 2003; Michaelis *et al.*, 2004; Tang *et al.*, 2004; Zgouras *et al.*, 2004; Blaheta *et al.*, 2005; Kuendgen & Gattermann, 2007). It can be used for prolonged periods of time and may enhance the effect of other cancer therapies (Karagiannis *et al.*, 2006). The chemical structure of VPA is shown in Figure 1.12 (A).

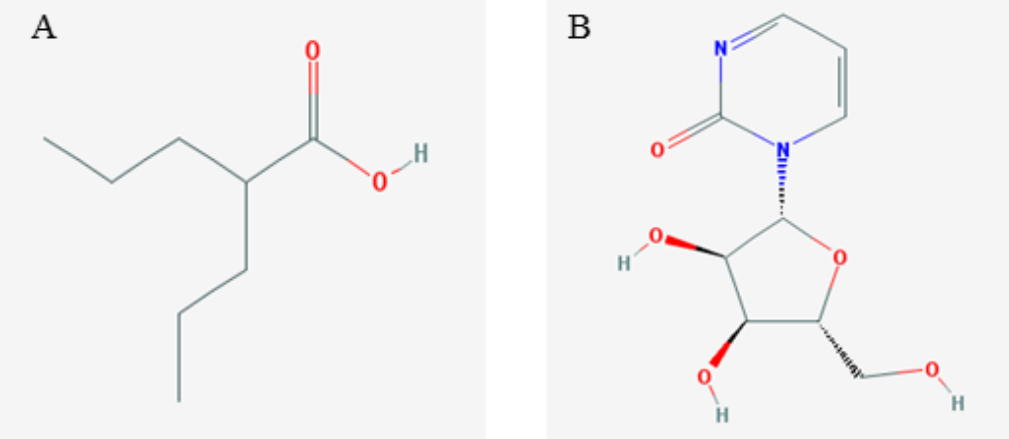

Figure 1.12. The chemical structure of (A) Valproic acid and (B) Zebularine (Bolton *et al.*, 2008).

Zebularine, a drug originally developed as a cytidine deaminase inhibitor (Kim *et al.*, 1986; Laliberté *et al.*, 1992), is an inhibitor of DNMTs. It prevents the transfer of methyl groups to the cytidines of CpG dinucleotides, resulting in the inhibition of DNMTs, as well as the re-activation of tumour suppressor genes previously silenced by DNMTs (Robertson & Jones, 2000; Egger *et al.*, 2004). It is a promising anti-cancer drug, because it is highly

stable in acidic and neutral solutions, is low on cytotoxicity both *in vivo* and *in vitro*, and can be administered orally. Furthermore, it has been shown to preferentially target cancer cells in a variety of cancer cell lines, including T24, SW480 and HT29 (Cheng *et al.*, 2004). The chemical structure of Zebularine is shown above in Figure 1.12 (B).

1.7 Research Aims

The role of SCs in the onset of haematopoietic cancers is well described. As there is compelling evidence for a role of SCs in the initiation and maintenance of solid tumours, and that these cells contribute to the development of therapeutic resistance, the aim of this study is to evaluate potential gene regulatory effect(s) of VPA and ZEB on pluripotency associated transcription factors in putative colorectal CSCs isolated from advanced colon adenocarcinoma cell lines. The research specific objectives for this study are presented below.

- Evaluate CD133 and EpCAM surface expression in the SW1116, HT29 and DLD1 cell lines which represent early-, mid- and late-stage colon adenocarcinomas, using flow cytometry. These markers are of particular interest here as they are both linked to the canonical WNT signalling pathway which is dysregulated in colorectal cancer.
- Isolate a putative colorectal CSC subpopulation from the two advanced colorectal adenocarcinoma cell lines HT29 and DLD1, using magnetic cell separation.
- Characterise the self-renewal capability of HT29 and DLD1 putative CSCs *in vitro* using a sphere formation assay.
- Culture and treat HT29 and DLD1 putative CSCs with the epigenetic drugs VPA and ZEB, to assess potential gene regulatory effects of these drugs on *NANOG*, *OCT4* and *SOX2* gene expression.
    - Evaluate the cellular expression of each of the pluripotency associated transcription factors NANOG, OCT4 and SOX2 following pharmacological treatments using confocal microscopy
    - Following on from the above, gene expression profiling will be used to assess epigenetic modulation of *NANOG*, *OCT4* and *SOX2* at the mRNA level.
- Propose a mechanism that potentially regulates SOX2 function in HT29 and DLD1 putative CSCs.

CHAPTER 2

CANCER STEM CELL ISOLATION AND CHARACTERISATION

The characterisation of colorectal cancer stem cells (CSCs) has important clinical relevance as they are suggested to be responsible for tumour initiation, metastasis and relapse, as proposed by the CSC model (Banerji & Los 2006; Dalerba *et al.*, 2007a; Rashedi *et al.*, 2007; Vermeulen *et al.*, 2008; Meacham & Morrison, 2013; Puglisi *et al.*, 2013; Easwaran *et al.*, 2014). Moreover, as they are also resistant to current chemo- and radiation-therapies, they have become a novel target in cancer therapy (Dalerba *et al.*, 2007a; O'Brien *et al.*, 2007; Ricci-Vitiani *et al.*, 2007; Vermeulen *et al.*, 2008; Visvader *et al.*, 2008; Saigusa *et al.*, 2010; Lin, 2014).

CSCs are potentially identified based on the expression of a unique panel of cell surface markers, which differ from differentiated, non-tumourigenic cells (Horst *et al.*, 2008; Yeung *et al.*, 2010; Ren *et al.*, 2013). A well characterised marker used to identify and isolate putative colorectal CSCs is the cluster of differentiation 133 (CD133). This marker holds particular promise as its increased expression has been correlated with enhanced tumour aggressiveness and reduced overall patient survival (Al-Hajj *et al.*, 2003; Maeda *et al.*, 2008; Zeppernick *et al.*, 2008; Shimada *et al.*, 2011; Patriarca *et al.*, 2012). However, it has yet to be established whether its mere presence identifies colorectal CSCs or whether its abundance holds greater value (Botchkina *et al.*, 2009).

In the current study a dual approach is used to identify putative CSCs from colorectal cancer cell lines representative of tumour stages and then to further characterise CSCs derived from two advanced colorectal cancer cell lines, using a sphere formation assay. As the use of CD133 as a sole marker to identify putative CSCs is currently controversial, a second CSC marker, epithelial cell adhesion marker (EpCAM), will be used in conjunction with CD133. EpCAM is a promising marker of putative colorectal CSCs as it has been used to identify CSCs in breast (Al-Hajj *et al.*, 2003) pancreatic (Li *et al.*, 2007) and hepatocellular (Yamashita *et al.*, 2009; Chen *et al.*, 2012a) carcinomas. Furthermore, its overexpression has been correlated with reduced overall survival in patients with a variety

of cancers including breast (Spizzo *et al.*, 2004), gall bladder (Varga *et al.*, 2004), urothelial (Brunner *et al.*, 2008) and colorectal (Patriarca *et al.*, 2012) carcinomas.

In this study, it is of particular interest to evaluate the potential co-expression of CD133 and EpCAM in the advanced colorectal adenocarcinoma cell lines HT29 and DLD1, as this may increase the probability of identifying putative CSCs. Moreover, these markers are of particular significance here as their expression have been associated with WNT activity in SCs (Munz *et al.*, 2009; Mak *et al.*, 2012), WNT being a key signalling pathway dysregulated in colorectal cancer. In this regard, mutations in the transcription factors APC and/or β-catenin lead to enhanced WNT activity (Morin *et al.*, 1997). Since the HT29 (Dukes' stage B) and DLD1 (Dukes' stage C) cell lines both have mutations in *APC* (Sahlberg *et al.*, 2014), these markers make good candidates to identify, isolate and characterise putative CSCs in these cell lines. Additionally, as a comparison, CD133 and EpCAM expression will be assessed in the early Dukes' stage A colon adenocarcinoma cell line, SW1116, since previous studies have correlated the overexpression of these markers with poor overall survival in patients with advanced stages of disease (Shimada *et al.*, 2011; Patriarca *et al.*, 2012; Ying *et al.*, 2013). It is of interest here to determine whether there is a correlation between the proportion of cells that express these markers and tumour stage.

Next, after assessing surface expression of cells potentially co-expressing CD133 and EpCAM, this particular cell subpopulation will be isolated from the HT29 and DLD1 cell lines using magnetic cell separation. Since this study focuses on epigenetic modulation in putative CSCs isolated from advanced adenocarcinoma cell lines, $CD133^{+}/EpCAM^{+}$ cells will not be isolated from the SW1116 cell line as it represents an early-stage colon adenocarcinoma.

Once isolated, CSCs can be further characterised in relation to their growth properties. Such SCs have an inherent capacity to maintain long-term self-renewal and proliferation *in vitro* and/or *in vivo*. One such assay, the xenotransplantation assay is used to assess "stemness" *in vivo*. In this test, CSCs are serially transplanted into immunodeficient mice and their "stemness" is demonstrated by their capacity to form clonal tumours that resemble the primary tumour from which the cells were isolated (Sampieri & Fodde, 2012;

Dalerba *et al.*, 2007a; Vermeulen *et al.*, 2008). However, as such facilities were unavailable for this study, the classic *in vitro* sphere formation assay will be used to evaluate the growth properties of the isolated cells. This assay utilises a three-dimensional cell culture system together with a serum-free CSC enrichment medium (Singh *et al.*, 2004; Ponti *et al.*, 2005) to assess "stemness" *in vitro*. In this test, the self-renewal capability of single cells is evaluated by their ability to grow as free floating-spheroids in non-adherent culture conditions, or as adherent spheres when cultured on a three-dimensional substrate, for example, Matrigel™ (BD Biosciences). Furthermore, in subsequent work presented here, the expression of a number of essential SC associated transcription factors are determined in HT29 and DLD1 putative CSCs using both confocal microscopy and quantitative real-time polymerase chain reaction (PCR) (Chapter 3).

2.1 Materials and Methods

2.1.1 Cell Lines

The cell lines SW1116 (donated by Dr J. Magré, INSERM), HT29 (ATCC-American Type Culture Collection) and DLD1 (HSRRB-Health Science Research Resources Bank, Japan) are human colorectal adenocarcinomas. The SW1116 cell line was isolated from a 73 year old male Caucasian and represents a Dukes' stage A colon cancer (ATCC). The HT29 cell line, isolated in 1964 from a 44 year old female Caucasian, is a Dukes' stage B colon cancer (Fogh & Trempe, 1975). The DLD1 cell line, originally isolated from a male adult between 1977 and 1979, is representative of a Dukes' stage C colon cancer (ATCC), which has metastasised beyond the intestinal wall of the colon into the lymph nodes (Dukes, 1932).

2.1.2 Cell Culture

An ethics waiver for the culture of commercial cell lines was granted by the Human Research Ethics Committee, University of the Witwatersrand, Reference: W-CJ-090317-3 (Appendix E).

The SW1116 (passage number 10), HT29 (passage number 30) and DLD1 (passage number 13) cell lines were cultured in Dulbecco's modified Eagles medium: Hams F12 (DMEM:F12) (Lonza), supplemented with 10% heat inactivated foetal bovine serum

(FBS) (Invitrogen and Lonza) and 10000 IU penicillin and streptomycin (Lonza) in 25cm^2 flasks (Greiner BioOne). Each of the cell lines were routinely maintained at 37°C with an atmosphere of 5% CO_2 and 90% humidity. The cells were cultured until reaching 70-85% confluency, following which they were subcultured; the culture medium was discarded and the cells were rinsed with 3ml phosphate buffered saline (PBS) (Sigma Aldrich). Next, the PBS was discarded, and 2.5ml of trypsin-ethylenediaminetetraacetic acid (EDTA) (Lonza) was added to the flask. The cells were incubated at 37°C, 5% CO_2 and 90% humidity for 3-5 minutes until they had detached from the culture surface. To neutralise the trypsin, a 2.5ml solution of 5% FBS/DMEM:F12, was added to the cell suspension. This suspension was then centrifuged for 3 minutes at 800rpm in 15ml centrifuge tubes (Greiner) and the supernatant was discarded. The cell pellet was resuspended in complete DMEM supplemented with 10% FBS and 1% penicillin/streptomycin (Lonza) and aliquoted into flasks for further culture and/or experimentation.

2.1.3 Trypan Blue Exclusion Assay

The cells from each cell line were counted using the Trypan Blue Exclusion Assay. To 40µl of Trypan blue (Molecular Probes), 10µl of resuspended cells were added and pipette mixed to make a 1:5 dilution. Then 10µl of cells were transferred to one of the chambers of a fast-read haemocytometer slide (Davies Diagnostics). Each chamber is marked by a grid of ten squares, each square containing 16 smaller squares. Using a microscope (Zeiss Axiovert), the cells were counted in at least two of the larger squares (excluding the cells that had stained blue and cells that touched the left and bottom borders of each square). The following equation was used to determine the number of cells present in the cell suspension:

(Total number of cells/number of squares counted)(10^4 x dilution factor) = cells/ml

Alternatively, 10µl of cell suspension mixed with Trypan Blue (1:1) was transferred to one of the chambers of a Bio-Rad dual chamber counting slide, and the cells were counted using the Bio-Rad TC10 automated cell counter.

2.1.4 Characterisation of CD133 and EpCAM using Flow Cytometry

Up to 10^7 SW1116, HT29 and DLD1 cells were re-suspended in 80µl of a PBS solution (Sigma Aldrich) containing 2mM EDTA (Sigma Aldrich) and 0.5% bovine serum albumin (BSA) (Sigma Aldrich), 20µl fragment crystallizable region (FcR) blocking reagent (Miltenyi Biotec), 10µl Anti-CD133/2 (293C3-phycoerythrin (PE), 50µg/ml) from Miltenyi Biotec and 20µl anti-EpCAM (fluorescein isothiocyanate (FITC), 3µg/ml) from BD Biosciences respectively. Unstained cells, which served as a gating control, were re-suspended as above, and were incubated with 20µl FcR blocking reagent only. SW1116, HT29 and DLD1 cells were also stained with either CD133-PE or EpCAM-FITC. These single stained cells served as compensation controls, to ensure that there was no "spill over" of FITC into the PE channel. The cells were washed in 2ml of stain/wash buffer (Appendix A.4), and centrifuged for 10 minutes at 300g. Next, the cells were re-suspended in 500µl of stain/wash buffer for analysis on a BD LSRFortessa flow cytometer, together with the "DIVA software".

The unstained cell lines were analysed in 'setup mode' on a forward scatter-area (FSC-A) *versus* forward scatter-height (FSC-H) graph to differentiate doublets from singlets. The singlets were gated for each cell line respectively. Data was then acquired for both the unstained and stained samples respectively, and the data was then evaluated using "FlowJo" flow cytometry analysis software (version 9.4.11).

2.1.5 CSC Isolation using Magnetic Cell Separation

The MACS cell separator system (Miltenyi Biotech) was used to magnetically isolate $CD133^+$ cells from cultured HT29 and DLD1 cells. HT29 and DLD1 cells were trypsinised and centrifuged at 300rpm for 10 minutes. The supernatant was aspirated off and the cell pellets were re-suspended in 300µl of stain/wash buffer (Appendix A.4) respectively, per 10^8 cells. FcR blocking reagent (100µl per 10^8 cells) was added to the re-suspended pellets, followed by the addition of 100µl CD133/1 (clone AC133) MicroBeads per 10^8 cells (Miltenyi Biotech). The cell suspensions were pipette mixed and incubated at 4°C for 30 minutes. To each cell suspension, 10µl per 10^7 cells of conjugated antibody CD133/2 (293C3-PE, Miltenyi Biotech) was added, pipette mixed and incubated at 4°C for 5 minutes. A CD133/2 control antibody (Miltenyi Biotech) was used together with flow

cytometry to evaluate the success of the magnetic cell separation and to determine the purity of CD133 cells obtained using the MACS cell separator system. As CD133 is co-expressed with EpCAM in the HT29 and DLD1 cell lines (Figure 2.2 B and Figure 2.2 C), it follows that magnetically sorted $CD133^+$ cells will express EpCAM.

The cells were washed in 2ml of stain/wash buffer respectively, and were centrifuged at 300rpm for 10 minutes. The supernatant was aspirated off and the cells were re-suspended in 500µl of stain/wash buffer. Two MACS MS columns were placed in the magnetic field of the MACS separator (Miltenyi Biotech). The columns were washed with 500µl of stain/wash buffer followed by the addition of the HT29 and DLD1 cell suspensions, respectively.

The flow-through (negative fraction) was collected in a flow cytometry tube. The columns were washed three times with 500µl of stain/wash buffer and the flow-through was again collected in a second flow cytometry tube. The columns were removed from the magnetic field of the MACS separator, placed over a new flow cytometry tube, whereupon one ml of stain/wash buffer was added. The cells, representing the positive fraction, were then flushed into a clean tube. In order to increase the cell purity, the positive cell fraction was applied to a second MACS MS column.

The cells were analysed on a BD FACSCalibur (Beckton Dickinson), calibrated using the FACSComp software (Version 5.2.1), together with BD Calibrite beads (BD Biosciences). The data was acquired for each of the unlabelled, labelled but unseparated, negative and positive (separated) samples, respectively. The data was then evaluated using "FlowJo" (version 9.4.11). The forward *versus* side scatter profiles for the CD133 positive fraction of HT29 and DLD1 cells were gated and these gates were used to back-gate the forward *versus* side scatter profiles of the unstained cells (Figure 2.1). The purification of $CD133^+$ cells obtained using magnetic cell separation was then determined.

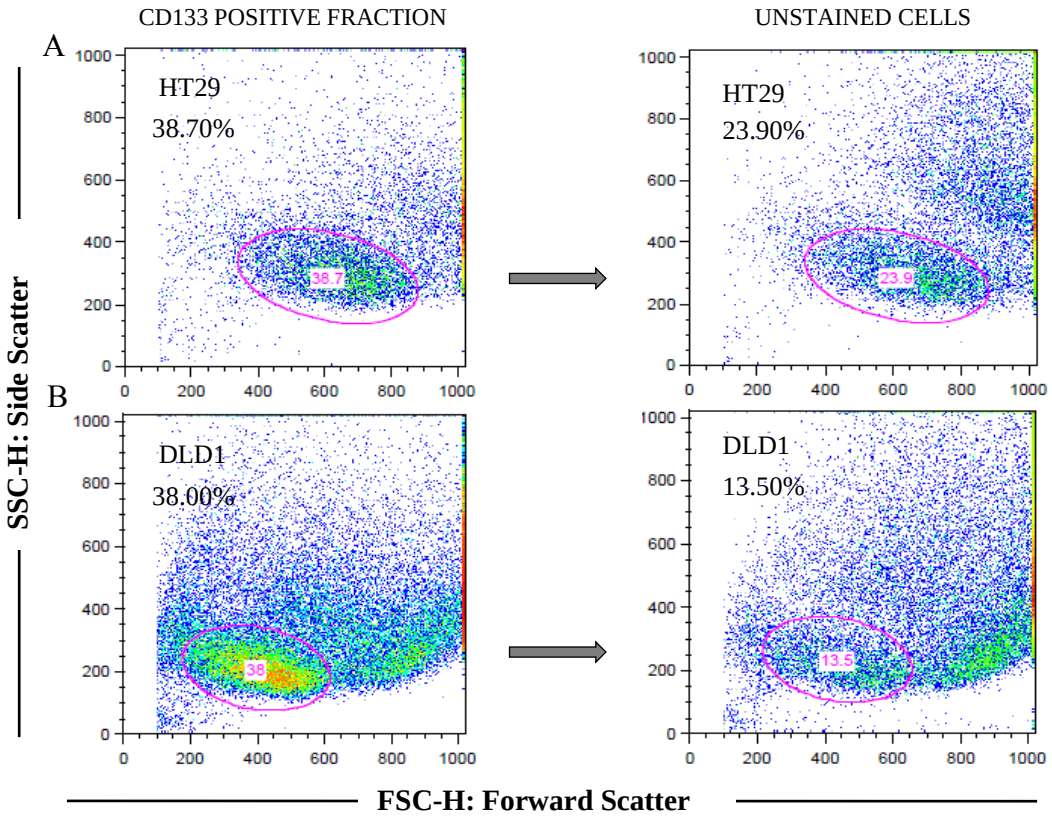

Figure 2.1. Flow cytometry dot plots showing forward scatter height *versus* side scatter height profiles for HT29 (A) and DLD1 (B) cells expressing CD133 following magnetic cell separation (left) and unstained cells (right). The forward *versus* side scatter profiles corresponding to the CD133 positive cell fractions were gated and these gates were used to back-gate the forward *versus* side scatter profiles of the unstained cells.

2.1.6 Three-Dimensional Cell Culture

The purified $CD133^+$ HT29 and DLD1 cell fractions were seeded into petri-dishes and/or flasks pre-coated with 2% agar (Appendix A.1). As agar prevents cell attachment, the cells grow in suspension, forming free-floating spheroids in a defined serum-free cancer stem cell enrichment medium supplemented with epidermal growth factor (EGF; 20ng/ml) and basic fibroblast growth factor (bFGF; 10ng/ml) (Mather, 2008) (Appendix A.1). Alternatively, these cells were seeded into petri-dishes and/or flasks pre-coated with growth factor reduced Matrigel™ (BD Biosciences) or Geltrex™ (Gibco) and were cultured as adherent spheres. Matrigel™ and Geltrex™ are soluble forms of basement

membrane extracted from Engelbreth-Holm-Swarm (EHS) mouse sarcoma cells. These matrices contain laminin, collagen IV, entactin and heparin sulphate proteoglycans that mimic the *in vivo* tumour microenvironment (BD Biosciences; Life Technologies). The thin gel (non-gelling) method was used to coat the petri-dishes and/or flasks and this method can be found in Appendix B.1.

2.1.7 Sphere Formation Assay

To evaluate the clonogenic potential (self-renewal capacity) of $CD133^+$ HT29 and DLD1 cells, 10^2 cells were seeded into the wells of a 6 well plate that had been pre-coated with Geltrex^{TM} (Gibco) (Appendix B.1). The cells were seeded at a low density to prevent the fusion of nascent spheres. The cells were cultured for 14 days in a CSC enrichment medium (Appendix A.1). The resulting spheres were counted under a light microscope (Zeiss Axiovert 25) and the clonogenic efficiency was calculated as the ratio of the number of spheres produced relative to the number of cells plated per well (Rafehi *et al.*, 2011). The assay was repeated a further two times (N=3).

2.1.8 Statistical Analyses

Statistical analyses were performed using Microsoft Office Excel, 2010. The results are presented as median expression and median counts, where appropriate.

2.2 Results

Anti-CD133/2 (PE) and anti-EpCAM (FITC) were used to characterise the surface marker expression of CD133 and EpCAM in the cell lines, SW1116, HT29 and DLD1, which are representative of early-, mid- and late-stage colon adenocarcinomas, respectively. The percentage of cells expressing $CD133^-/EpCAM^-$, $CD133^-/EpCAM^+$, and $CD133^+/EpCAM^+$ in each cell line are shown in Table C1 in Appendix C. For unstained cells, 99.50% of the SW1116 cells were negative for CD133 and EpCAM, and 100% of the HT29 and DLD1 cells were negative for these markers (Figure 2.1). For cells stained with both CD133 and EpCAM, it was found that EpCAM was co-expressed with CD133. Note that a $CD133^+/EpCAM^-$ cell population was not identified within any one of these cell lines. Within the SW1116 cell line, three distinct isotypic cell populations were identified including, $CD133^-/EpCAM^-$, $CD133^-/EpCAM^+$, and $CD133^+/EpCAM^+$, as shown in Figure

2.2 A. In contrast, only two distinct populations were identified in each of the HT29 and DLD1 cell lines, these being, $CD133^-/EpCAM^+$ and $CD133^+/EpCAM^+$, as shown in Figure 2.2 B and Figure 2.2 C.

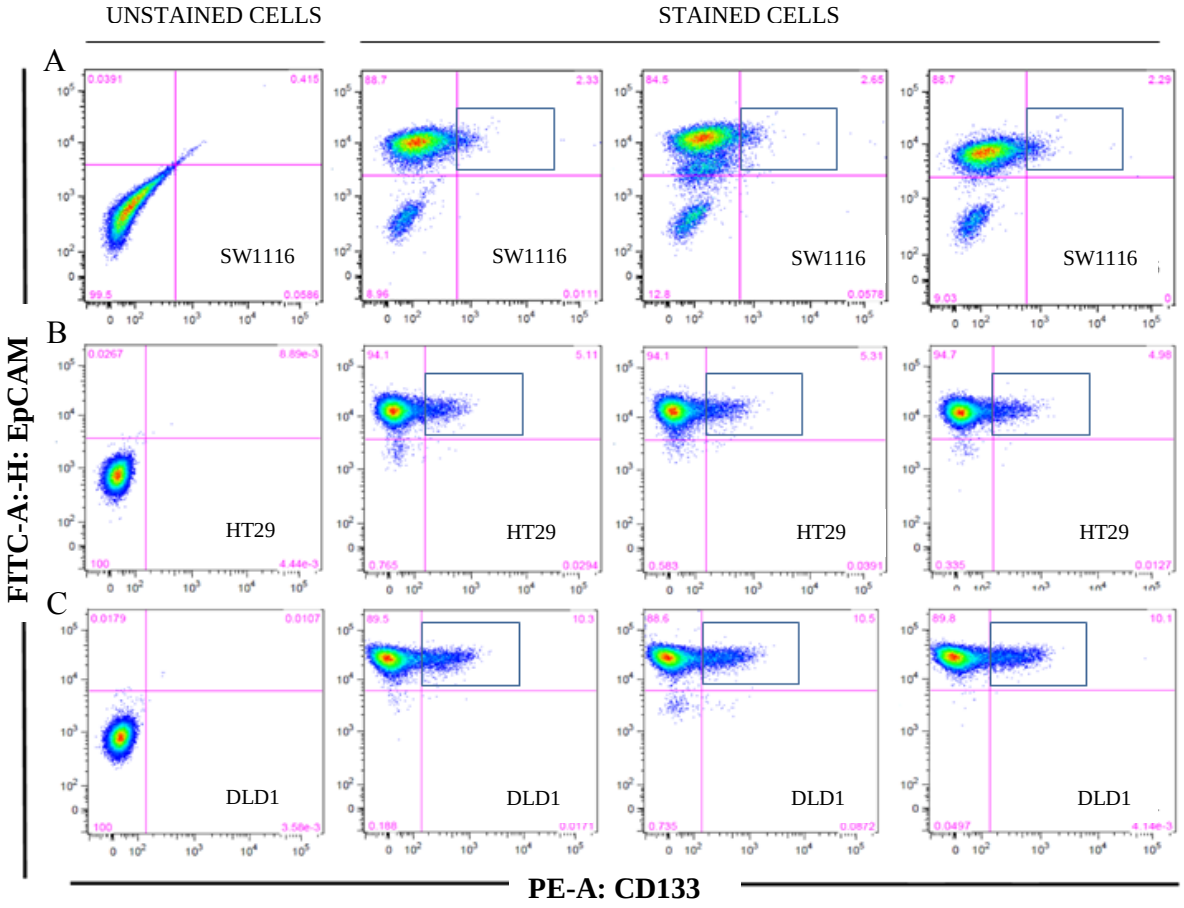

Figure 2.2. Flow cytometry detection of CD133 (PE) and EpCAM (FITC) in SW1116 (A), HT29 (B) and DLD1 (C) cell lines. Left to right these figures show unstained cells followed by triplicate cell staining of each cell line with CD133-PE and EpCAM-FITC antibodies. The unstained SW1116 cells show some degree of auto-fluorescence as 0.42% of these cells fluoresce in the PE and FITC channels (upper right quadrant). The gated cells co-express CD133 and EpCAM, whereas the un-gated cells do not express CD133, but may express EpCAM.

In the SW1116 cell line, 9.03% of these cells are $CD133^-/EpCAM^-$, 88.70% of these cells are $CD133^-/EpCAM^+$ and 2.33% of these cells are $CD133^+/EpCAM^+$ (Figure 2.3 A-C). In the HT29 cell line, 0.58% of these cells are $CD133^-/EpCAM^-$, 94.10% of these cells are

$CD133^-/EpCAM^+$ and 5.11% of these cells are $CD133^+/EpCAM^+$ (Figure 2.3 A-C). In the later stage DLD1 cell line, 0.19% of these cells are $CD133^-/EpCAM^-$, 89.50% of these cells are $CD133^-/EpCAM^+$ and 10.30% of these cells are $CD133^+/EpCAM^+$ (Figure 2.3 A-C). These results indicate that there is an increase in the proportion of $CD133^+$ expressing cells with advancement of tumour stage.

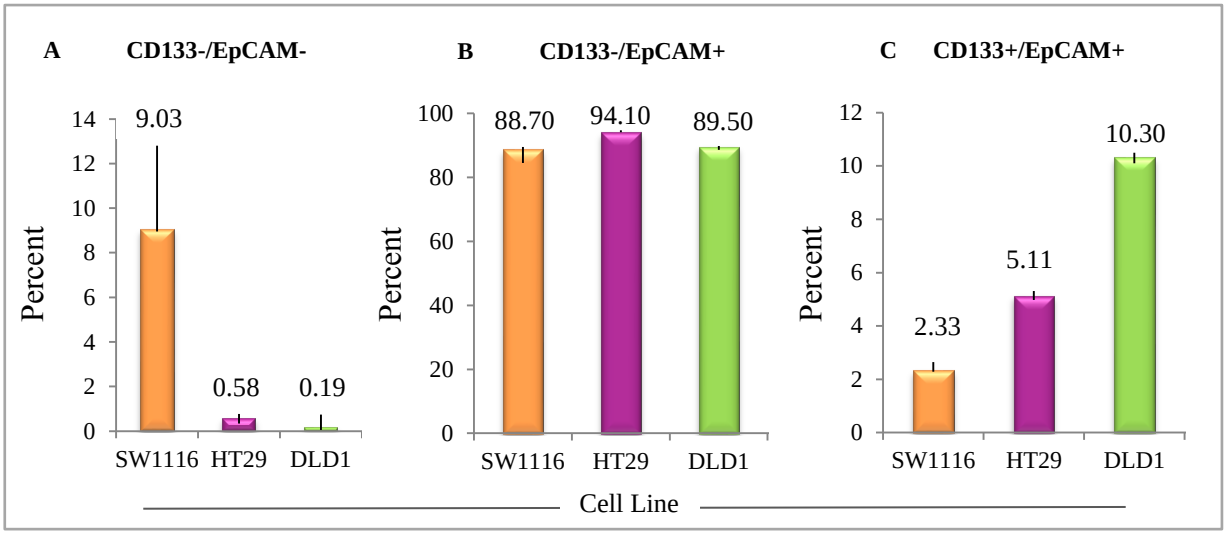

Figure 2.3. Flow cytometry analyses of $CD133^-/EpCAM^-$ (A), $CD133^-/EpCAM^+$ (B) and $CD133^+/EpCAM^+$ (C) expression in the SW1116, HT29 and DLD1 cell lines. The columns represent the median expression of each cellular subset and the lines show the maximum and minimum ranges of the data.

As the focus of this study centres on epigenetic modulation of $CD133^+/EpCAM^+$ cells isolated from advanced cell lines, MACS magnetic cell separation was used to isolate $CD133^+$ cells from the HT29 and DLD1 cell lines, but not from the early-stage SW1116 cell line. Furthermore, as CD133 is co-expressed with EpCAM (Figure 2.2), magnetically sorted $CD133^+$ cells will express EpCAM. A 98.60% and 98.00% purification of $CD133^+/EpCAM^+$ HT29 and DLD1 cells were obtained respectively, as confirmed by flow cytometry (Figure 2.4).

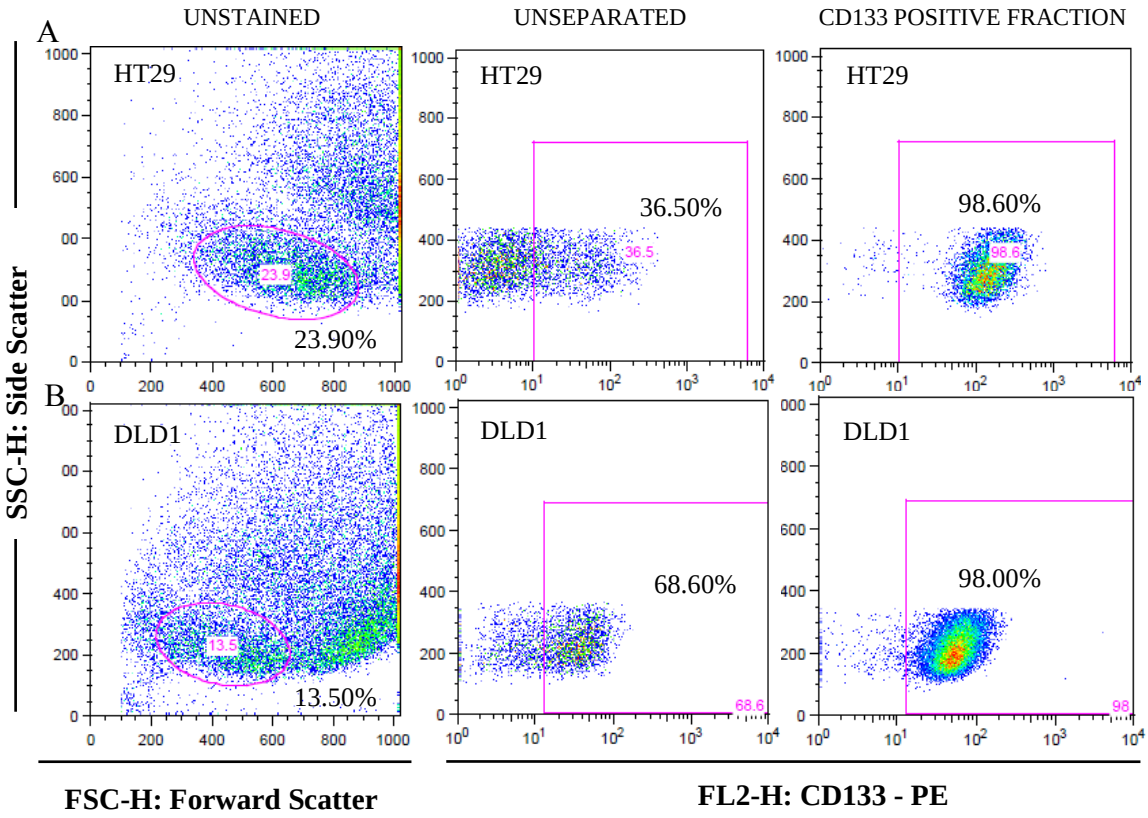

Figure 2.4. Flow cytometry analyses showing the efficiency of CD133 magnetic cell separation in the HT29 (A) and DLD1 (B) cell lines. Left to right these figures show gated, unstained cells followed by the percentage of cells expressing CD133, prior to and following separation.

Next, the $CD133^{+}/EpCAM^{+}$ cells were cultured as three-dimensional free-floating spheroids and/or as adherent spheres in a defined CSC enrichment medium supplemented with bFGF and EGF. $CD133^{+}/EpCAM^{+}$ HT29 and DLD1 cells produced free-floating spheroids that ranged in size from less than 2µm to more than 5µm in diameter when cultured over a period of two weeks (Figure 2.5). Three-dimensional cultures are favoured over two-dimensional monolayer cultures as they better reflect the physiology and microenvironment of an *in vivo* tumour (Pampaloni *et al.*, 2007; Friedrich *et al.*, 2009; Vinci *et al.*, 2012). Furthermore, these cells doubled approximately every 48-72 hours whereas their corresponding parental cell lines doubled approximately every 24 hours. These findings indicate that $CD133^{+}/EpCAM^{+}$ HT29 and DLD1 cells cycle much slower than their parental cell lines. This phenomenon of longer cycling times could possibly be

used in addition to both *in vitro* sphere formation assays and *in vivo* tumourigenic assays, to further characterise putative CSCs.

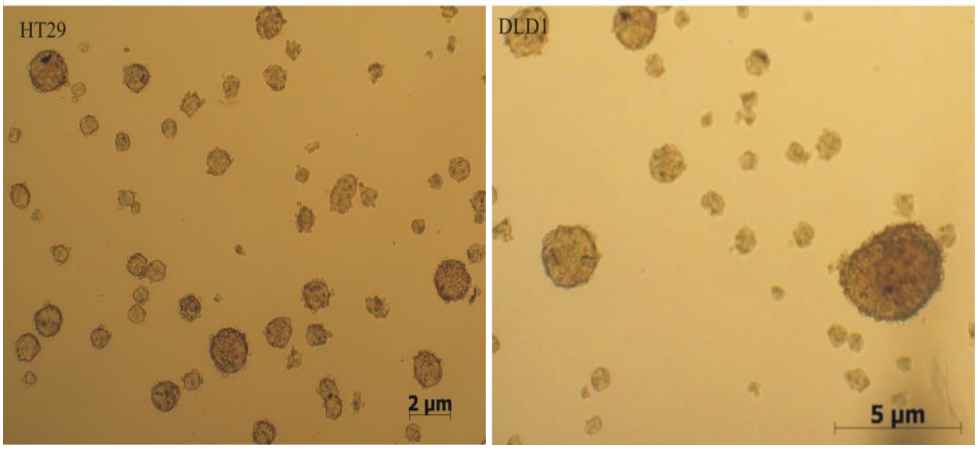

Figure 2.5. Brightfield images of HT29 (left) and DLD1 (right) free-floating putative CSC spheroids (Zeiss Axiovert 25, Original magnification 5X).

Subsequently, a sphere formation assay was carried out to determine the self-renewal capacity of these cells. This assay essentially recapitulates the *in vivo* xenotransplantation assay, *in vitro*. The number of spheres produced were counted (Table C2 in Appendix C) and the clonogenic efficiency was calculated as the ratio of the number of spheres produced to the number of cells plated per well. The clonogenic efficiency of both the $CD133^+/EpCAM^+$ HT29 and DLD1 cells was 21%. Therefore, approximately 1 in 5 of each of these cells produced spheres. Figure 2.6 shows the number of spheres produced per 10^2 cells plated per well.

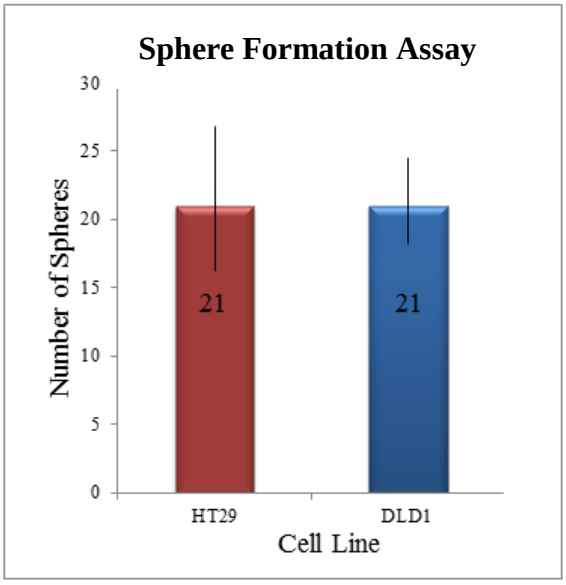

Figure 2.6. Sphere-forming capacity of $CD133^+/EpCAM^+$ HT29 and DLD1 cells. Equal numbers of $CD133^+/EpCAM^+$ HT29 and DLD1 cells (10^2 cells/well) each produced twenty-one spheres. The columns represent the median results, and the lines show the upper and lower interquartile ranges.

2.3 Discussion

In this aspect of the study, it was determined whether CD133 and EpCAM were co-expressed in three established colorectal adenocarcinoma cell lines that represent the three progressive stages of colorectal cancer; it was queried whether an increase in their expression could be correlated with tumour stage, and whether they could be used to isolate and characterise putative CSCs from advanced colorectal adenocarcinomas. These markers were of interest here as their increased expression has been correlated with enhanced WNT activity, a driver of colorectal tumourigenesis.

Flow cytometry characterisation of CD133 and EpCAM surface expression revealed the presence of three distinct populations in the SW1116 cell line ($CD133^-/EpCAM^-$, $CD133^-/EpCAM^+$ and $CD133^+/EpCAM^+$), and two in the HT29 and DLD1 cell lines ($CD133^-/EpCAM^+$ and $CD133^+/EpCAM^+$) (Figure 2.2). The expression of EpCAM was high in all of these cell lines and ranged from 88.70-94.10% (Figure 2.3 B). However, and in contrast to EpCAM expression, CD133 expression was found to vary from 2.33% in the SW1116

cell line to 5.11% and 10.30% in the HT29 and DLD1 cell lines, respectively (Figure 2.3 C). These data clearly indicate that there is a correlation between the number of cells expressing CD133 and tumour stage. More specifically, there was a two-fold increase in the number of HT29 cells expressing CD133, when compared to the number of SW1116 cells expressing this marker. Similarly, there was a two-fold increase in the number of DLD1 cells expressing CD133 compared to the HT29 cell line and a four-fold increase, relative to the SW1116 cell line. CD133 was also shown to be co-expressed with EpCAM. It has similarly been reported by Ricci-Vitiani *et al.* (2007) and Horst *et al.* (2008) that $CD133^{+}$ colorectal cancer cells also express EpCAM.

While the data reported here differs to other studies, it must however be emphasised that published data relating to CSCs is quite varied. This relates to the use of various cell lines, differing culture conditions, prolonged cell culture, cells sourced from patient tumours, different antibody clones and binding epitopes used to detect the same antigens (Dittfeld *et al.*, 2009; Hermansen *et al.*, 2011; Sahlberg *et al.*, 2014), combined with very different platforms for analysis. Some examples relating to colon CSCs are discussed below.

Ieta *et al.* (2008) using flow cytometry, evaluated the expression of CD133 in twelve colorectal cancer cell lines, including HT29 and DLD1; and reported that while 47.50% of the HT29 cells expressed CD133 only 0.40% of the DLD1 cells expressed this marker. In contrast, Elsaba *et al.* (2010) found that the HT29 cell line had in excess of 95% CD133 expression, whereas, the DLD1 cell line lacked CD133 expression. More recently, Sahlberg *et al.* (2014) reported that 63% of the HT29 cells and 14% of the DLD1 cells expressed CD133, respectively. It should be noted that Ieta *et al.* (2008) used an anti-CD133 antibody (clone not mentioned), Elsaba *et al.* (2010) used an anti-CD133/1 antibody and Sahlberg *et al.* (2014) used an anti-CD133/2 antibody which may contribute to the inconsistencies found amongst their data.

In contrast to the use of established cell lines to evaluate CD133 expression, Ricci-Vitiani *et al.* (2007) investigated the expression of CD133 in novel cell lines derived from many colonic tumours, ranging in Dukes' stages (A-D). Although they reported that 2.5 ± 1.4% of cells expressed CD133 in the majority of the colonic tumours, there was however, no evident trend associating CD133 expression with tumour stage (Table C3 in Appendix C).

In Table C3, the expression of CD133 varies from 0.7-2.4% in Dukes' stage A tumours, from 0.7-6.1% in Dukes' stage B tumours, from 1.4-2.7% in Dukes' stage C tumours, and in one Dukes' stage D tumour, the expression of CD133 was 2.5%. Similarly, O'Brien *et al.* (2007) investigated CD133 expression in 17 colonic tumours (six were primary tumours). Of these 17 colonic tumours, eleven were stage IV, three were stage IIIC, one was stage IIB and two were stage I. Overall, CD133 expression ranged from 1.8-24.5% (Table C4 in Appendix C). The expression of CD133 in stage I tumours ranged from 9.3-12.0%, in a stage IB tumour the expression of CD133 was 14.0%, in stage IIIC tumours the expression ranged from 7.5-15.9% and in stage IV tumours the expression ranged from 18.0-24.5%. Furthermore, O'Brien and colleagues (2007) investigated the expression of CD133 in the normal colonic tissues of six patients. CD133 was expressed in these normal tissues, however, CD133 expression was greater in the corresponding tumour tissues (Table C4). Overall, the findings reported by Ricci-Vitiani *et al.* (2007) and O'Brien *et al.*, (2007) suggest that there is no evident trend associating CD133 expression with tumour stage.

It is evident from the above that proportions of CSCs vary greatly even within samples derived from patients. Furthermore, it is difficult to compare the data obtained from established cell lines with clinical specimens, as established cell lines lack a tumour microenvironment that modulates surface marker expression (Platet *et al.*, 2007; Dittfeld *et al.*, 2009; Rao *et al.*, 2012; Visvader & Lindeman, 2012; Grosse-Gehling *et al.*, 2013; Hsu *et al.*, 2013). *In vitro* approaches at best attempt to mimic the *in vivo* environment with defined growth factors, and extracellular matrices to maintain stem growth.

Despite these differences, CD133 and EpCAM are associated with poor prognosis in colorectal cancer patients. Horst *et al.* (2008) evaluated CD133 expression in patient-derived colorectal tumours using immunohistochemistry and reported that a high expression of CD133 was associated with worse overall survival. Similarly, in a meta-analysis study by Chen *et al.* (2013), they too reported that a high expression of CD133 was associated with a worse prognosis. High EpCAM expression has also been shown to be associated with advanced tumour stage and with a worse overall survival (Patriarca *et al.*, 2012). However, and contrary to these findings, Lugli *et al.* (2010) reported that low

levels of membranous EpCAM, as opposed to high levels of EpCAM, is prognostic. Therefore, the use of EpCAM as a prognostic marker may need further validation.

Having confirmed the surface co-expression of CD133 and EpCAM in the SW1116, HT29 and DLD1 cell lines, magnetic cell separation was used to isolate a $CD133^+$ cell population from the advanced colorectal adenocarcinoma cell lines HT29 and DLD1. MACS magnetic cell separation is a technology used to isolate frequent or rare occurring cell populations. It uses paramagnetic microbeads, as small as 50nm in size, which have been conjugated with an antibody specific to the researcher's needs (Miltenyi Biotec). The rationale for subsequently isolating putative CSCs based on CD133 expression emanates from two studies by Ricci-Vitiani *et al.* (2007) and O'Brien *et al.* (2007). Ricci-Vitiani *et al.* (2007) evaluated the tumourigenic potential of $CD133^+$ colon cancer cells isolated from primary tumours and found that these cells could initiate a tumour when they were transplanted into SCID mice. However, this was not true for the $CD133^-$ colon cancer cells, as the transplantation of these cells into SCID mice failed to elicit tumour formation. Similarly, O'Brien *et al.* (2007) showed that colorectal cancer initiating cells were $CD133^+$ whereas, $CD133^-$ cells were unable to initiate tumourigenesis when transplanted into immunodeficient mice.

Prior to magnetic cell separation, it was found that 36.50% and 68.60% of the HT29 and DLD1 cells expressed CD133, respectively (Figure 2.4). These values differ considerably to those previously reported here, that is 36.50% *versus* 5.11% in the HT29 cell line, and 68.60% *versus* 10.30% in the DLD1 cell line. These discrepancies most likely result from the use of different flow cytometry gating strategies. Furthermore, these findings highlight the importance of standardising methods as the use of different methods can lead to varied data, making it difficult to compare data amongst studies.

The gating strategy used for analysing CD133 and EpCAM surface expression in the SW1116, HT29 and DLD1 cell lines involved gating on a singlet population. This is the preferred strategy (Figure 2.2) whereas, the gating strategy used for analysing the effectiveness of magnetic cell separation involved back-gating on the forward *versus* side scatter profile of the $CD133^+$ cell fractions (Figure 2.1). The unstained HT29 and DLD1 cells that were gated using the back-gating strategy, have a low forward scatter *versus* a

low side scatter profile. As CSCs are known to be smaller in size and less complex compared to the remainder of the population, it is likely that putative CSCs reside in a population of cells with a low forward and side scatter (De Paiva *et al.*, 2006; Allen *et al.*, 2009; Allahverdiyev *et al.*, 2012). Ideally these gating strategies should have been the same however, the parameters required to analyse a singlet population post-data acquisition were not set when the magnetic sorting data was captured. The inability to change the parameters post-data acquisition is one of the disadvantages of using an analogue instrument such as the BD FACS Calibur, whereas, this could have been done had the data been captured on a digital instrument such as the BD Fortessa. Although not ideal, the gating strategy used to analyse the pre- and post-magnetic cell separation data was not a limiting factor in this study as flow cytometry was merely used to determine the efficiency of magnetic cell separation. A 98.60% and a 98.00% pure population of $CD133^+$ cells were obtained from the HT29 and DLD1 cell lines, respectively. Therefore, magnetic cell separation was shown to be a valuable tool that can successfully be used to isolate and enrich for $CD133^+$ cells.

Since not all $CD133^+$ cells are putative CSCs, a sphere formation assay was used to determine the sphere-forming capacity of these $CD133^+/EpCAM^+$ HT29 and DLD1 cells *in vitro*. This assay evaluates the self-renewal capability of these cells. In the present study, approximately 1 in 5 (21%) $CD133^+/EpCAM^+$ HT29 and DLD1 cells produced spheres when cultured in a defined CSC medium that serves to maintain "stemness". A similar percentage of such cells could be predicted to initiate tumours upon transplantation into immunodeficient mice. A xenotransplantation assay should be used to compare the *in vivo* properties of these cells with their *in vitro* growth patterns, however, as previously mentioned (page 41), such facilities were unavailable for this study. In contrast, O'Brien *et al.* (2007) reported a far lower value, wherein only 1 in 262 $CD133^+$ cells were capable of initiating colorectal tumourigenesis when transplanted into immunodeficient mice. The relatively high percentage of spheres obtained here in the present study, may well reflect the use of the additional CSC enrichment step using magnetic cell separation. The use of additional markers may well be important to better define putative CSC populations and some of these are discussed below.

Other markers used to identify putative CSC include amongst others CD44, CD166 and CD24 (Dalerba *et al.*, 2007b; Chen *et al.*, 2011; Wang *et al.*, 2012; Sahlberg *et al.*, 2014). Although, the use of these markers may further enrich for colorectal CSCs, they may not be entirely reliable and their expression may not always be restricted to putative CSCs (Medema, 2013). For example, Chen *et al.* (2011) and Wang *et al.* (2012) showed that CD44 could be used to further enrich for putative CSCs that express CD133. However, since multiple CD44 splice variants exist whose functional roles remain poorly understood, this challenges its use as a marker of putative CSCs (Naor *et al.*, 1997; Medema, 2013; Sahlberg *et al.*, 2014). Similarly, Dalerba *et al.* (2007b) showed that CD44, when used in combination with $EpCAM^{high}$, marks colorectal CSCs and that CD166 could further enrich for $CD44^{+}$ putative CSCs. However, and in contrast to these studies, Muraro *et al.* (2012) showed that neither of these markers could be used to identify and characterise putative CSCs in any of the ten colorectal cancer cell lines used in their study. The differences in these findings may again be attributed to the lack of a tumour microenvironment when using established cell lines, which has been shown to modulate marker expression and tumourigenicity *in vivo* (Rao *et al.*, 2012; Visvader & Lindeman, 2012), as well as due to the use of different antibodies and techniques (Grosse-Gehling *et al.*, 2013; Sahlberg *et al.*, 2014). Nonetheless, these differences highlight the importance of identifying novel and universal markers that can be used to identify and characterise colorectal CSCs, as well as the importance to standardise the use of antibodies and techniques.

Although it was shown here that approximately 1 in 5 $CD133^{+}/EpCAM^{+}$ HT29 and DLD1 cells were capable of producing spheres *in vitro*, other studies have shown that $CD133^{-}$ cells are also capable of forming spheres *in vitro* (Botchkina *et al.*, 2009; Chen *et al.*, 2012a; Hsu *et al.*, 2013) and initiating tumourigenesis *in vivo* (Dalerba *et al.*, 2007b; Shmelkov *et al.*, 2008; Chu *et al.*, 2009; Kawamoto *et al.*, 2010). Chen *et al.* (2012) identified and isolated four cellular subsets characterised as $CD133^{+}/EpCAM^{+}$, $CD133^{+}/EpCAM^{-}$, $CD133^{-}/EpCAM^{+}$ and $CD133^{-}/EpCAM^{-}$ from the hepatocellular carcinoma cell line Huh7. Although, all the cellular subsets showed some degree of "stemness", the subpopulation characterised as $CD133^{+}/EpCAM^{+}$ produced more colonies that were larger in size compared to the other cellular subsets. Also, they possessed more biological features of CSCs including enhanced differentiation potential and enhanced

resistance to the chemotherapeutic agents, 5-Fluorouracil and Doxorubicin *in vitro*, and enhanced tumourigenicity *in vivo*.

Shmelkov *et al.* (2008) and Kawamoto *et al.* (2010) showed both $CD133^+$ and $CD133^-$ colorectal cancer cells could elicit tumour formation in immunocompromised mice. Shmelkov *et al.* (2008) isolated $CD133^+$ and $CD133^-$ cells from a metastatic tumour whereas, in contrast to the studies by O'Brien *et al.* (2007) and Ricci-Vitiani *et al.* (2007), they isolated $CD133^+$ and $CD133^-$ cells from primary tumours. Since epithelial cells undergo EMT during metastasis, these cells may acquire SC characteristics and this might explain why the $CD133^-$ cells showed tumourigenicity *in vivo* (Tsai & Yang, 2013). Another explanation is based on the findings by Kemper *et al.* (2010) who reported that CD133 expression is not confined to the putative CSC population, but rather the AC133 epitope is masked due to differential folding of the CD133 protein and therefore, CD133 would not be detected in differentiated cells when using AC133 antibodies. In the present study, the antibodies CD133/1 (linked to MACS magnetic beads) and CD133/2 (conjugated with PE) were used. Since both these antibodies recognise the AC133 epitope, their use was not a limiting factor with regards to putative CSC identification, isolation and characterisation in the HT29 and DLD1 cell lines as both of these cell lines were shown to express AC133. In addition, their $CD133^+$ cellular fractions were shown to have clonogenic potential *in vitro*.

Hsu *et al.* (2013) also showed that $CD133^+$ and $CD133^-$ cells isolated from the metastatic colorectal cancer cell line SW620 had clonogenic capability *in vitro* and tumourigenic potential *in vivo*. Both the cellular subsets showed similar characteristics with regards to their morphology, proliferative potential and resistance to the anti-tumour drugs Taxol, Cisplatin, Actinomycin D and Camptothecin when cultured under standard conditions. However, the $CD133^-$ cells showed enhanced growth potential when exposed to 3D Matrigel culture conditions and enhanced tumourigenicity when injected into immunocompromised mice suggesting that the tumour microenvironment may modulate metastatic progression. Therefore, they exposed the $CD133^+$ and $CD133^-$ cells to hypoxic conditions, cell adhesion free conditions and to ECM, the three environmental stressors faced by metastasising cells. Exposure to these conditions was found to promote cell-type switching, a phenomenon wherein $CD133^+$ cells switch to $CD133^-$ cells and *vice versa*.

Exposure to hypoxic and cell adhesion free conditions enhanced the switching of $CD133^-$ cells to $CD133^+$ cells whereas, exposure to ECM enhanced the switching of $CD133^+$ cells to $CD133^-$ cells. Based on their findings, they proposed a model on metastasis progression that might explain why $CD133^-$ cells, and not just $CD133^+$ cells, show tumourigenic potential *in vitro* and *in vivo*. In this model, both $CD133^+$ and $CD133^-$ cells are released into the bloodstream and migrate to the metastatic site where they are faced with environmental stressors. Exposure to ECM promotes the switching of $CD133^+$ cells to $CD133^-$ cells thereby enhancing their proliferation ability. Once colonisation has been established these cells are exposed to hypoxic and adhesion free conditions, which promotes the switching of $CD133^-$ cells to $CD133^+$ cells to enhance their "stemness" and to initiate tumour growth.

Altogether, the findings reported in the present study, as well as those of other studies, indicate that not all $CD133^+$ cells are necessarily putative CSCs. In this regard, the stable retention of SC properties are highly likely to be influenced by "factors", including growth factors and ECM molecules, within the *in vivo* and/or *in vitro* culture environments.

2.4 Conclusion

In this chapter, the SW1116, HT29 and DLD1 cell lines were shown to co-express CD133 and EpCAM. The percentage of cells expressing EpCAM was high in each of these cell lines whereas, the percentage of cells expressing CD133 increased with tumour stage. Furthermore, using magnetic cell separation, a pure population of $CD133^+$ cells were isolated from the advanced colorectal adenocarcinoma cell lines, HT29 and DLD1. In evaluating their *in vitro* growth properties, some 21% of these cells showed clonogenic potential. The use of other markers, in addition to CD133 and EpCAM, would be useful for further enrichment of putative CSCs in the HT29 and DLD1 cell lines.

The following phase of this study employs $CD133^+/EpCAM^+$ cells for subsequent epigenetic studies that focus on epigenetic modulation of the pluripotent-associated transcription factors NANOG, OCT4 and SOX2 in these cells.

CHAPTER 3

THE EFFECT(S) OF EPIGENETIC MODULATION ON PLURIPOTENT MARKERS

Colorectal cancer is known to arise from an accumulation of mutations in both oncogenes and tumour-suppressor genes, as proposed by Vogelstein *et al.* (1988), as well as due to epigenetic modifications (Issa *et al.*, 2004; Barker & Clevers, 2007; Vermeulen *et al.*, 2008; Markowitz & Bertognolli, 2009; Goel & Boland, 2012; Schnekenburger & Diederich, 2012). Epigenetic modifications include alterations in patterns of DNA methylation, histone modifications and the influence of non-coding RNAs on post-translational protein regulation, all of these being important in colorectal cancer initiation and progression (Migheli *et al.*, 2012). Of the three modifications, DNA methylation is the best understood and it involves the addition of methyl groups to CpG islands often located in the promoter regions of genes, thereby resulting in gene silencing (Graff *et al.*, 1997; Migheli *et al.*, 2012). In colorectal cancer, several genes are known to be methylated that affect cellular processes, including amongst others DNA repair, apoptosis, WNT signalling, cell cycle arrest, invasion and metastasis (Migheli *et al.*, 2012; Schnekenburger & Diederich, 2012).

Histone modifications, although less understood than DNA methylation, involve the addition and removal of acetyl and methyl groups from histone tails that protrude from nucleosomes, and these may either enhance or repress gene transcription (van Engeland *et al.*, 2011; Goel & Boland, 2012). In addition to histone acetylation and methylation, other histone modifications include phosphorylation, ubiquitylation and sumoylation although, acetylation and methylation are best characterised in colorectal cancers (van Engeland *et al.*, 2011; Goel & Boland, 2012; Nakazawa *et al.*, 2012; Schnekenburger & Diederich, 2012). Non-coding miRNAs further regulate gene expression by annealing to complementary mRNA resulting in mRNA degradation and gene silencing (Rane *et al.*, 2007; Fabian *et al.*, 2010). Cellular processes affected by miRNA regulation in colorectal cancer include WNT signalling, cell migration and invasion, epithelial differentiation and

cell cycle regulation (Lujambio *et al*., 2007; Shi *et al.*, 2007; Burk *et al*., 2008; Huang *et al.*, 2008; Schnekenburger & Diederich, 2012).

Since epigenetic defects are potentially reversible, and CSCs are thought to trigger tumour initiation, progression and metastasis, it is of interest here to evaluate the cellular expression of each of the SC associated pluripotent factors NANOG, OCT4 and SOX2 following epigenetic treatments in putative CSCs derived from the HT29 and DLD1 cell lines. These key pluripotency associated transcription factors are widely expressed in a variety of cancers, including colorectal cancer, suggesting that they might also play an essential role in maintaining the "stemness", and thus, the continued proliferation of colorectal CSCs (Ben-Porath *et al.*, 2008; Liu *et al.*, 2013). Therefore, each of these transcription factors are potential targets for the treatment of colorectal cancer in patients with advanced stages of disease. The expression of *NANOG*, *OCT4* and *SOX2* are tightly modulated by epigenetic mechanisms to allow them to maintain pluripotency in embryonic and adult SCs, however, these mechanisms are poorly understood (Farthing *et al.*, 2008; Meissner, 2010; Villasante *et al.*, 2011; Wang *et al.*, 2013b). Additionally, they further modulate their own expression via an auto-regulatory feed-back network (Figure 3.1 A) and regulate the transcription of *NANOG* and downstream target genes via a feed-forward loop (Figure 3.1 B; Boyer *et al.*, 2005).

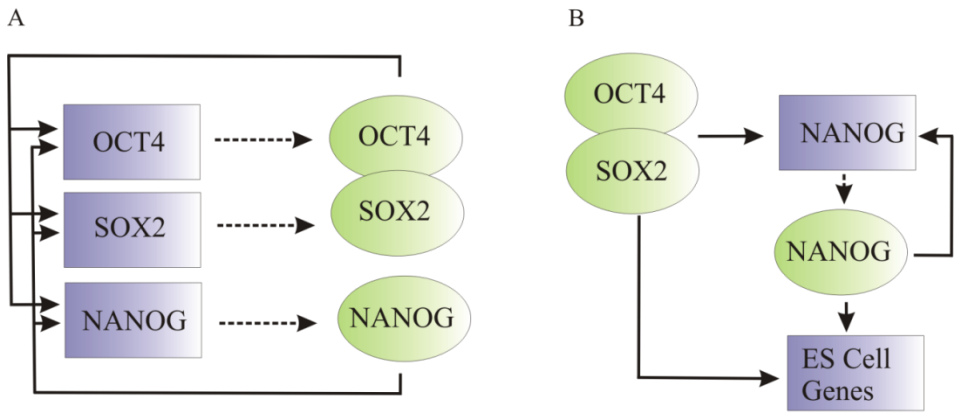

Figure 3.1. Transcriptional regulation of *NANOG*, *OCT4* and *SOX2* via an auto-regulatory feed-back loop (A); and transcriptional regulation of *NANOG* and ESC genes via a feed-forward loop (B). The promoters of genes are indicated as blue blocks and regulators are indicated as green ovals (adapted from Figure 4 in Boyer *et al.*, 2005, 952 p.).

In this chapter, the potential regulatory effects of Valproic acid (VPA) and Zebularine (ZEB) on *NANOG*, *OCT4* and *SOX2* expression in HT29 and DLD1 putative CSCs are investigated. To date, no studies have explored the effects of these two drugs on the expression of these genes in putative CSCs derived from colorectal adenocarcinomas. In the first section of this chapter, confocal microscopy is utilised to qualitatively evaluate the expression of each of these pluripotency factors following epigenetic treatments *in vitro*. In the second section, and following on from the above, gene expression profiling (qPCR) will be used to quantitatively assess epigenetic modulation of *NANOG*, *OCT4* and *SOX2*.

VPA, an inhibitor of class I and IIa histone deacetylase (HDAC) activity, relaxes the association of core histones with DNA to allow transcription and hence, gene expression. Furthermore, its biochemical actions have been shown to inhibit cancer cell proliferation, promote differentiation, induce apoptosis, and inhibit angiogenesis in a number of cancer cell lines, including breast and colon adenocarcinomas (Göttlicher *et al*., 2001; Phiel *et al.*, 2001; Blaheta & Cinatl, 2002; Kawagoe *et al.*, 2002; Phillips *et al.*, 2003; Gurvich *et al.*, 2004; Michaelis *et al.*, 2004; Tang *et al.*, 2004; Thelen *et al.*, 2004; Zgouras *et al.*, 2004; Blaheta *et al.*, 2005; Friedmann *et al.*, 2006; Kuendgen & Gattermann, 2007).

ZEB, a nucleoside analogue of cytidine, inhibits DNA methyltransferase (DNMT) activity by inhibiting the transfer of methyl groups to the cytidines of CpG dinucleotides, serving to prevent the maintenance of DNA methylation (Yoo & Jones, 2006). This allows transcription factors to bind to the promoters of their target genes resulting in gene expression. Compared to similar DNA methylation inhibitors, 5-aza-2'deoxycytidine and 5-azacytidine, ZEB has low cytotoxicity both *in vitro* and *in vivo*, is highly stable in acidic and neutral solutions; and can be administered orally (Barchi *et al.*, 1992; Cheng *et al.*, 2003).

Based on the inhibitory activity that VPA and ZEB have on HDACs and DNMTs, it is proposed here that these epigenetic drugs will promote the up-regulation of *NANOG*, *OCT4* and *SOX2* expression in HT29 and DLD1 putative CSCs. Since their up-regulation in embryonic and adult SCs has been shown to promote cell differentiation (Niwa *et al.*, 2000; Rodriguez *et al.*, 2007; Kopp *et al.*, 2008; Alvarez *et al.*, 2010), epigenetic modulation of these transcription factors in CSCs may make these cells more susceptible to

current anti-cancer therapies. Ultimately, this could result in tumour regression and may prevent tumour relapse.

(i) Confocal Microscopy

3.1 Materials and Methods

3.1.1 CD133 and EpCAM

Prior to drug treatments, confocal microscopy was used to confirm the expression of CD133 and EpCAM in HT29 and DLD1 putative CSCs, isolated using magnetic cell separation (Chapter 2.1.5). HT29 and DLD1 putative CSCs were seeded onto sterile glass coverslips, placed in 30mm Petri dishes, and cultured in a serum free CSC enrichment medium at 37˚C in 5% CO_2 until they reached 60-70% confluency. Next, the cells were rinsed with PBS and fixed for 15 minutes in 3% formaldehyde diluted in PBS, followed by three PBS washes (refer to Appendix A.2). The cells were then permeabilised for 5 minutes in a PBS solution containing 0.5% bovine serum albumin (BSA) and 0.25% Triton X-100 (refer to Appendix A.2), followed by three PBS washes. The coverslips were incubated overnight at 4°C, in a moist sealed container with 1:100 of rabbit polyclonal anti-CD133 (H-284, Santa Cruz Biotechnology) and goat polyclonal anti-EpCAM (A-20, Santa Cruz Biotechnology). The coverslips were washed three times with a 0.5% PBS/BSA blocking solution (refer to Appendix A.2); and excess liquid was blotted on tissue paper. The coverslips were then incubated with 1:200 of Alexa Fluor 594, donkey anti-rabbit IgG (H+L) (Invitrogen), to detect CD133; and Alexa Fluor 568 donkey anti-goat IgG (Invitrogen), to detect EpCAM. Incubations were carried out in a moist sealed container in the dark at room temperature for one hour. The coverslips were washed three times in 0.5% PBS/BSA and blotted. This was followed by one hour incubation in the dark, in a moist sealed container with 1:20 Phalloidin (Alexa Fluor 488-green) (Lonza). Again the coverslips were washed three times in 0.5% PBS/BSA and blotted, and then incubated with 1:10000 of 4', 6-diamidino-2-phenylindole (DAPI) (Cambrex) in PBS for 5 minutes in the dark at room temperature, in a moist sealed container. The coverslips were washed a further three times in 0.5% PBS/BSA, followed by a PBS wash. The coverslips were blotted and mounted onto slides using Gel Mount Aqueous Mounting Media (Sigma-Aldrich) and allowed to set. The slides were viewed on the Zeiss LSM-780 confocal

microscope. Negative controls were carried out, wherein either primary or secondary antibodies were omitted, respectively. In such cases, cells were incubated with PBS/BSA for equivalent time periods to the experimental conditions.

3.1.2 Drug Treatments

Having confirmed the expression of CD133 and EpCAM in HT29 and DLD1 putative CSCs, these cells were seeded onto sterile coverslips and were serum starved for 12 hours to synchronise the cell cycle. The cells were treated for 24 hours, in triplicate, with the epigenetic drugs VPA (Sigma) and ZEB (Santa Cruz Biotechnology) respectively, to assess their possible effects on NANOG, OCT4 and SOX2 expression. A range of drug concentrations were tested, including 1mM, 2.5mM and 5mM VPA; and 100µM, 250µM and 500µM ZEB, respectively. Controls included untreated cells, and cells treated with dimethyl sulphoxide (DMSO) (Sigma) (ZEB carrier). The range of drug concentrations were carefully selected based on the following literature: Göttlicher *et al.* (2001), Phiel *et al.* (2001), Cheng *et al.* (2004), Gurvich *et al.* (2004), Scott *et al.* (2007), Billam *et al.* (2010), Venkataramani *et al.* (2010), Marquardt *et al.* (2011), and Witt *et al.* (2013). Next, the coverslips with attached cells were processed for indirect immunofluorescence staining as previously described in Section 3.1.1. The coverslips were incubated overnight at 4°C, in a moist sealed container with 1:100 of goat polyclonal anti-NANOG (W-18, Santa Cruz Biotechnology), 1:100 of goat polyclonal anti-OCT3/4 (N-19, Santa Cruz Biotechnology) and mouse monoclonal anti-SOX2 (Abcam). The coverslips were washed three times with a 0.5% PBS/BSA blocking solution (refer to Appendix A.2); and excess liquid was blotted on tissue paper. The coverslips were then incubated with 1:200 of Alexa Fluor 568, donkey anti-goat IgG (Invitrogen) to detect NANOG and OCT4; and Alexa Fluor 488 donkey anti-mouse IgG (Invitrogen) to detect SOX2. Incubations were carried out in a moist sealed container in the dark at room temperature for one hour. The coverslips were washed three times in 0.5% PBS/BSA and blotted, and then incubated with 1:10000 of DAPI (Cambrex) in PBS for 5 minutes in the dark at room temperature, in a moist sealed container. The coverslips were washed and mounted onto slides as previously described in Section 3.1.1. The slides were viewed on the Zeiss LSM-780 confocal microscope. Negative staining controls were performed by omitting either the primary or secondary

antibodies in each case and replacing these with an equivalent amount of PBS/BSA solution.

3.1.3 Scoring of Pluripotent Marker Expression

The drug treated HT29 and DLD1 putative CSC confocal images were analysed and scored according to zero, low, intermediate and high pluripotent marker expression in the nucleus. The scoring system used in this study is based on the system used by Yu *et al.* (2010). Six to eight cells per field of view were assessed. A score of zero was assigned to cells with no marker expression, one was assigned to cells with low marker expression, two was assigned to cells with intermediate marker expression and three was assigned to cells with high marker expression. The median fluorescence scores were plotted onto bar graphs. Cytoplasmic expression of NANOG, OCT4 and SOX2 was of less interest here as the exclusion of these transcription factors from the nucleus prevents them from binding to the promoters of their target genes (Baltus *et al.*, 2009).

3.1.4 Statistical Analyses

The Kruskal-Wallis and Mann-Whitney tests were used to assess whether there were any significant changes in pluripotent marker expression, with respect to drug treatment. The Kruskal-Wallis test examines variance in more than two independent samples and being non-parametric does not assume a normal distribution, whereas the Mann-Whitney test examines variance between two independent samples (Swinscow & Campbell, 2001). In analysing the confocal microscopy scored data, where the Kruskal-Wallis test indicated significant changes in nuclear expression, a Mann-Whitney *post hoc* analysis was subsequently performed. A two sample Mann-Whitney test was also used to determine whether there were significant differences between the untreated/no treatment (NT) and DMSO treatment groups. These statistical analyses were carried out using the "PAST" software, version 2.17c (Hammer *et al.*, 2001). Statistical significance was considered at a p-value ≤ 0.05.

The sign test was also considered to examine the data for trends. This test is less powerful than the Kruskal-Wallis and Mann-Whitney tests as it only considers the medians of variables whereas, the Kruskal-Wallis and Mann-Whitney tests consider the spread of

unmatched and unpaired variables, respectively (Hart, 2001; Pappas & DePuy, 2004). However, the smallest *p-value* that can be obtained when using the sign test on a sample size of three is 0.25 (one-sided) or 0.50 (two-sided), assuming that two "positives" (or "negatives") occur consecutively. Consequently, it is recommended that a minimal sample size of seven should be used when evaluating data using the sign test (Hammer *et al.*, 2001).

3.2 Results and Discussion

Immunofluorescence staining confirmed the expression of CD133 and EpCAM in HT29 and DLD1 derived putative CSCs following their isolation. In this aspect of the study, confocal microscopy shows that CD133 and EpCAM are localised on the cell membrane and in the cytoplasm in both HT29 and DLD1 putative CSCs (Figures 3.2-3.3). The primary and secondary antibody staining controls are shown in Figure D.1 in Appendix D and indicate a complete lack of non-specific antibody binding or fluorescence.

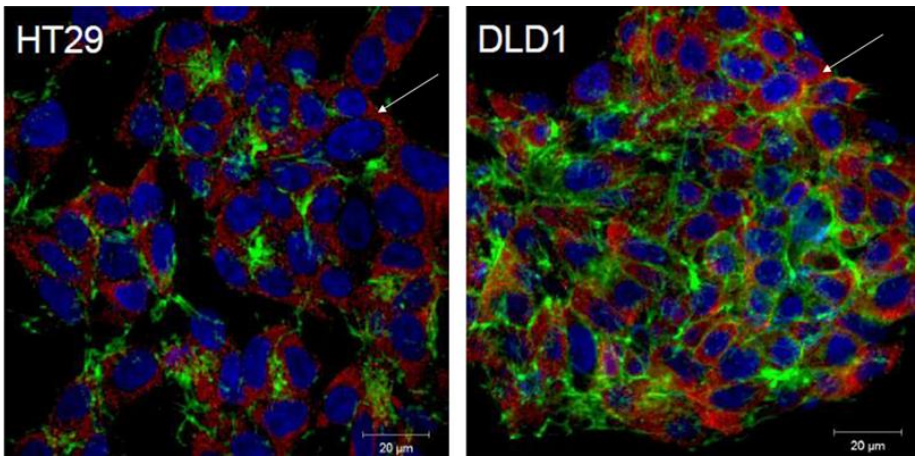

Figure 3.2. Immunofluorescence images showing CD133 expression in HT29 and DLD1 putative CSCs. Cells were stained with Phalloidin-Alexa Fluor 488 conjugate (F-actin) indicated in green, DAPI (nuclei) indicated in blue and CD133 detected with Alexa Fluor 594, donkey anti-rabbit IgG (H+L) indicated in red, as shown by the arrows. (Original magnification 63X).

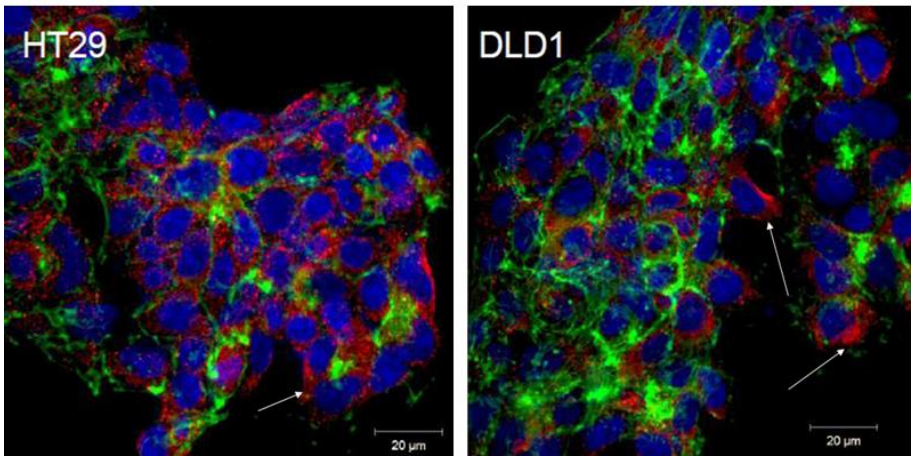

Figure 3.3. Immunofluorescence images showing EpCAM expression in HT29 and DLD1 putative CSCs. Cells were stained with Phalloidin-Alexa Fluor 488 conjugate (F-actin) indicated in green, DAPI (nuclei) indicated in blue and EpCAM detected with Alexa Fluor 568 donkey anti-goat IgG indicated in red, as shown by the arrows. (Original magnification 63X).

Similar to the findings reported here, some studies have described that CD133 is expressed exclusively on the membrane (Kojima *et al.*, 2008; Li *et al.*, 2009; Wang *et al.*, 2009; Garcia *et al.*, 2011), whilst others have reported that it is expressed on the cell membrane and in the cytoplasm of colorectal cancer cells (Takahashi *et al.*, 2010; Xi & Zhao, 2011; Coco *et al.*, 2012; Li *et al.*, 2012b; Reggiani *et al.*, 2012). Although it is cell surface expression of CD133 that is prognostic, the shift in its expression from the cytoplasm to the surface on the cell membrane is believed to be a characteristic of cell invasiveness (Takahashi *et al.*, 2010; Chen *et al.*, 2013). Therefore, the cytoplasmic expression of CD133 in HT29 and DLD1 putative CSCs may indicate the invasive nature of these cells. Similarly, cytoplasmic expression of EpCAM has been described in colorectal carcinoma cells (Gosens *et al.*, 2007; Yanamoto *et al.*, 2007). These findings support the observations that CD133 and EpCAM are expressed on the cell membrane and in the cytoplasm in HT29 and DLD1 putative CSCs.

Having confirmed CD133 and EpCAM expression in HT29 and DLD1 putative CSCs, these cells were treated with varying concentrations of VPA and ZEB to evaluate the epigenetic modulatory effects that these drugs have on NANOG, OCT4 and SOX2

expression in these cells. Cell proliferation and morphology were not affected by these epigenetic drugs. The confocal images of drug treated HT29 and DLD1 putative CSCs are shown in Figures D2-D13 in Appendix D (N=1). The primary and secondary antibody negative staining controls indicate a complete lack of non-specific fluorescence (refer to Figure D14 in Appendix D).

Confocal microscopy shows that these proteins are localised in the nucleus and cytoplasm in HT29 and DLD1 putative CSCs. Figure 3.4 shows an example of nuclear and cytoplasmic expression of these transcription factors in untreated HT29 putative CSCs. Although these proteins were shown to localise in the cell nucleus and cytoplasm, it is nuclear expression that is of particular interest here as the exclusion of these transcription factors from the nucleus prevents them from maintaining pluripotency as they are no longer able to bind to the promoters of their downstream target genes.

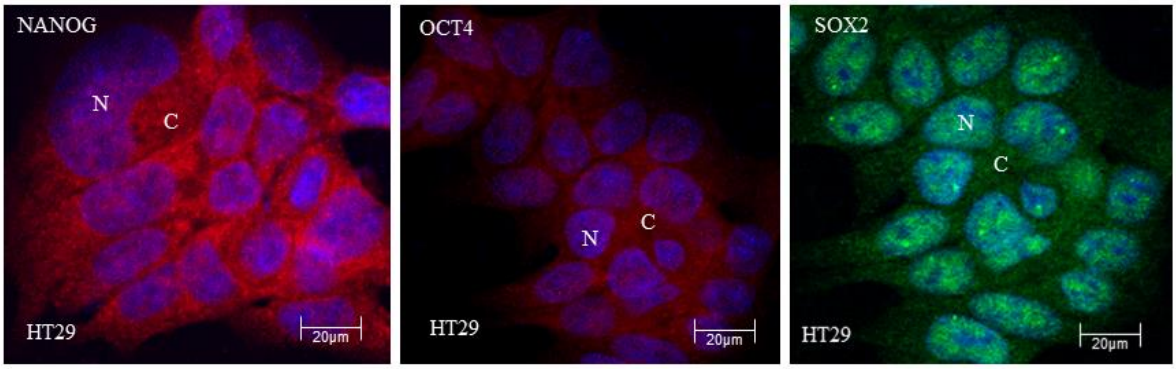

Figure 3.4. Immunofluorescence images showing nuclear (N) and cytoplasmic (C) expression of NANOG, OCT4 and SOX2 in untreated HT29 putative CSCs. Cells were stained with DAPI (nuclei) indicated in blue, NANOG detected with Alexa Fluor 568 donkey anti-goat IgG indicated in red, OCT4 detected with Alexa Fluor 568 donkey anti-goat IgG indicated in red and SOX2 detected with Alexa Fluor 488 donkey anti-mouse IgG indicated in green. (Original magnification 63X).

Similar to the findings reported here, Guo et al. (2011) reported that NANOG and OCT4 localised in the cell nucleus and cytoplasm in human glioma cells whereas, SOX2 localised exclusively in the nucleus. Despite the difference in SOX2 localisation, SOX2 has been

reported in other studies to shuttle from the nucleus into the cytoplasm in SCs, and in colorectal cancer cells (Lin *et al.*, 2010). In a study by Ishiguro *et al.* (2012), they reported that NANOG predominantly localised in the cytoplasm in HT29 and SW620 colorectal cancer cells whereas in Tera-2 cells, it predominantly localised in the nucleus. Alternative splicing of *NANOG* and *OCT4* mRNA transcripts, and post-translational modifications of NANOG and OCT4 proteins, could explain their nuclear and cytoplasmic localities (Lee *et al.*, 2006; Guo *et al.*, 2011). For example, the alternatively spliced and translated protein OCT4A is localised in the cell nucleus of pluripotent cells where it maintains pluripotency whereas, the alternatively spliced and translated protein OCT4B is localised in the cytoplasm of tumour and pluripotent cells (Cauffman *et al.*, 2005; Lee *et al.*, 2006). The functional role of OCT4B remains unknown; however, it may play a role in regulating the cell-cycle, and in regulating cell proliferation and survival (Cortes-Dericks *et al.*, 2013). Cortes-Dericks *et al.* (2013) showed that the knock-down of *OCT4B* in the lung adenocarcinoma cell line A549 increased the sensitivity of these cells to the chemotherapeutic drug Cisplatin by promoting cell-cycle progression, cell proliferation and apoptosis. Altogether, these findings support the observations that NANOG, OCT4 and SOX2 are expressed in the nucleus and cytoplasm in HT29 and DLD1 putative CSCs, which as stated above may indicate that *NANOG* and *OCT4* are alternatively transcribed and translated in these cells. Furthermore, the differential localisation of these pluripotency associated transcription factors in HT29 and DLD1 putative CSCs suggests that apart from their functional role in maintaining pluripotency, they might also participate in regulating other cellular functions.

Next, the nuclear expression of NANOG, OCT4 and SOX2 in HT29 and DLD1 putative CSCs were scored according to zero, low, intermediate and high expression (N=3). The data was then analysed using the Kruskal-Wallis and Mann-Whitney tests to determine whether VPA and ZEB significantly modulate the expression of these markers at a protein level. The medians of the scored expression data are shown in Tables C5-C6 in Appendix C and the results of the statistical analyses are shown in Tables C7-C8 in Appendix C. The median expression of nuclear NANOG, OCT4 and SOX2 in drug treated HT29 and DLD1 putative CSCs is graphically represented using bar graphs, and these are shown in Figures 3.5-3.10. The effects of VPA and ZEB on the nuclear expression of these transcription

factors was analysed relative to untreated and DMSO treated cells, respectively. Statistical significance where p≤0.05 is also presented in these figures.

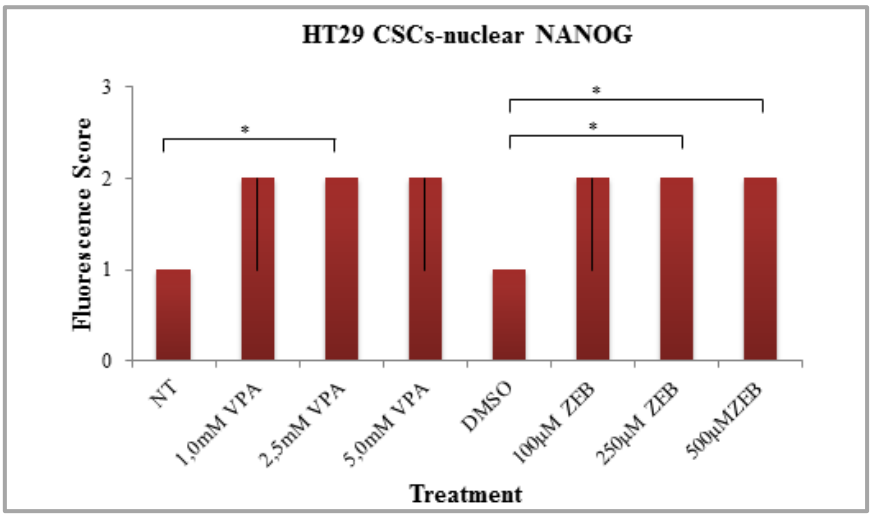

Figure 3.5. Confocal microscopy analyses of nuclear NANOG in HT29 putative CSCs treated with VPA and ZEB. The columns represent the median fluorescence and the lines show the maximum and minimum ranges of the data (*p≤0.05; N=3).

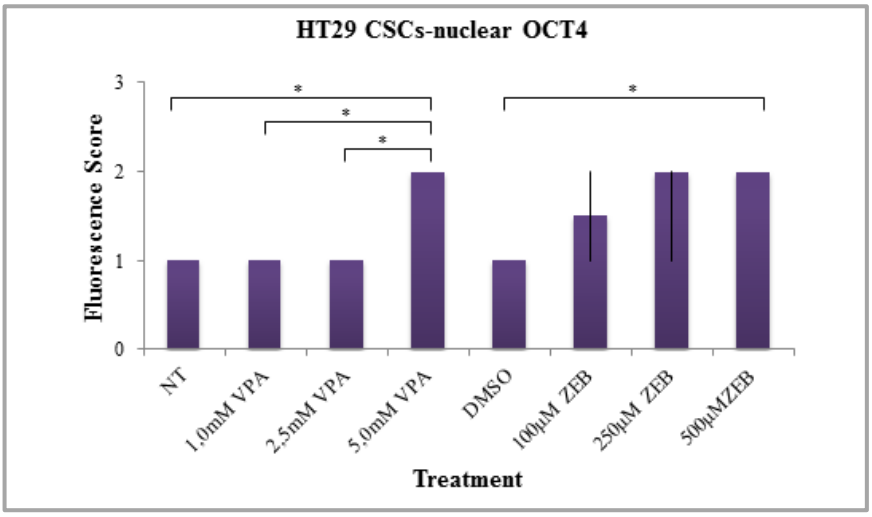

Figure 3.6. Confocal microscopy analyses of nuclear OCT4 in HT29 putative CSCs treated with VPA and ZEB. The columns represent the median fluorescence and the lines show the maximum and minimum ranges (*p≤0.05; N=3).

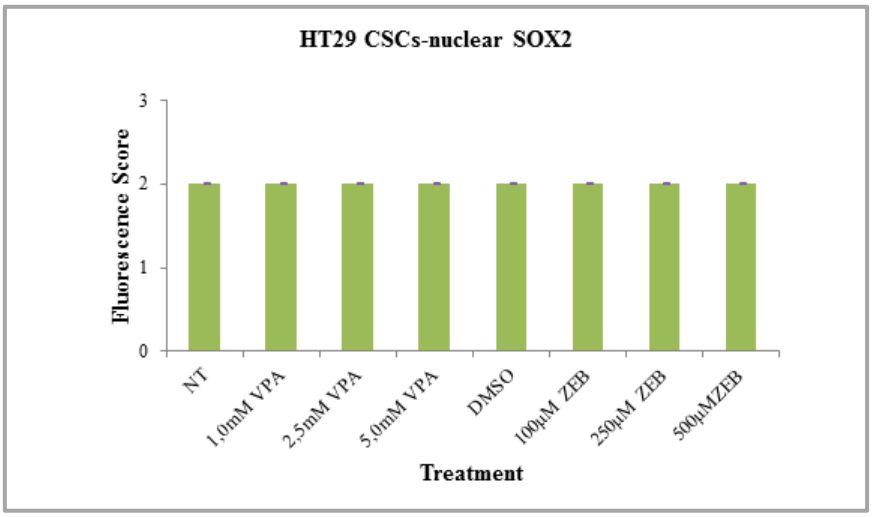

Figure 3.7. Confocal microscopy analyses of nuclear SOX2 in HT29 putative CSCs treated with VPA and ZEB. The columns represent the median fluorescence and the lines show the maximum and minimum ranges (*p≤0.05; N=3).

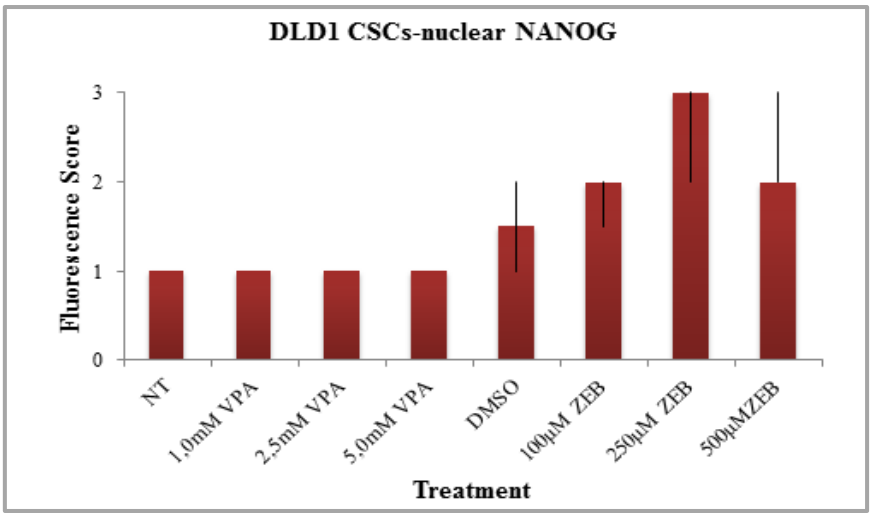

Figure 3.8. Confocal microscopy analyses of nuclear NANOG in DLD1 putative CSCs treated with VPA and ZEB. The columns represent the median fluorescence and the lines show the maximum and minimum ranges (*p≤0.05; N=3).

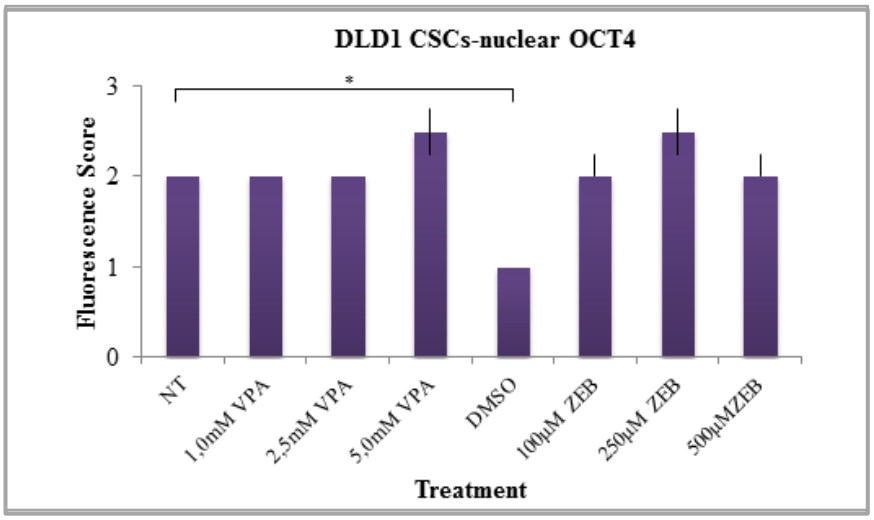

Figure 3.9. Confocal microscopy analyses of nuclear OCT4 in DLD1 putative CSCs treated with VPA and ZEB. The columns represent the median fluorescence and the lines show the maximum and minimum ranges (*p≤0.05; N=3).

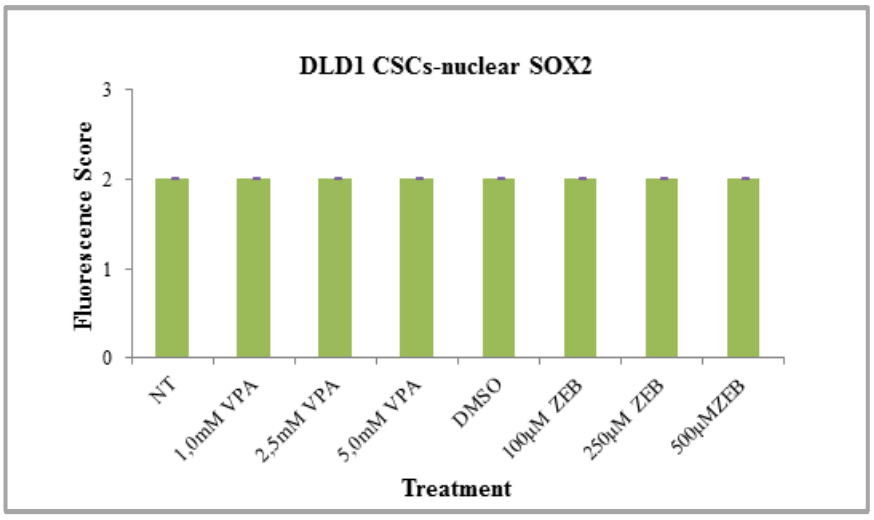

Figure 3.10. Confocal microscopy analyses of nuclear SOX2 in DLD1 putative CSCs treated with VPA and ZEB. The columns represent the median fluorescence and the lines show the maximum and minimum ranges (*p≤0.05; N=3).

In HT29 putative CSCs, 2.5mM and 5mM VPA up-regulate nuclear NANOG and OCT4, respectively (Figures 3.5 and 3.6). When these cells are treated with ZEB, concentrations of 250µM and 500µM up-regulate nuclear NANOG equally whereas, ZEB at 500µM up-regulates OCT4 in these cells (Figures 3.5 and 3.6). Neither VPA nor ZEB modulate SOX2 expression in these cells, as its expression remained unchanged (Figure 3.7).

In DLD1 putative CSCs, neither VPA nor ZEB modulate NANOG, OCT4 and SOX2 expression (Figures 3.8-3.10) as sample size may have been a limiting factor. Due to financial and time constraints experiments were carried out in triplicate, which is common amongst molecular and cellular biologists (Vaux, 2012). However, treatment of these cells with DMSO, a ZEB carrier, down-regulates OCT4 relative to untreated cells (Figure 3.9). Although DMSO was used as a drug carrier, it is also a differentiation inducing agent that has been shown to induce the differentiation of embryonic SCs (Torres *et al.*, 2012; Chetty *et al.*, 2013). Torres *et al.* (2012) reported that the treatment of E14Tg2a embryonic SCs with DMSO induced their differentiation. Moreover, they reported that DMSO reduced the expression of *NANOG* and *OCT4* by 0.18±0.07- and 0.70±0.02-fold, respectively. Although DMSO did not reduce the expression of NANOG in HT29 and DLD1 putative CSCs, it was shown in the present study to reduce the expression of OCT4 in DLD1 CSCs.

Taken together, specific concentrations of VPA and ZEB were shown to modulate nuclear NANOG and OCT4 expression in HT29 putative CSCs. Whether these drugs modulate NANOG and OCT4 expression in DLD1 putative CSCs remains inconclusive as a larger sample size is required to verify results. Moreover, SOX2 protein expression remained unchanged in both HT29 and DLD1 putative CSCs following their treatment with VPA and ZEB. This finding is worthy of note as SOX2 and OCT4 cooperatively regulate *NANOG* transcription. Consequently, it is thought that VPA and ZEB would up-regulate SOX2 protein expression, especially in HT29 putative CSCs, as these drugs were shown to up-regulate NANOG and OCT4 expression in these cells. Similar to the finding reported here, Kallas *et al.* (2014) reported that SOX2 protein expression remained unchanged in the human ESC line H9 following its differentiation induced by sodium butyrate, whereas NANOG and OCT4 were down-regulated upon differentiation. Based on their findings, they concluded that *NANOG* and *OCT4* are regulated at an epigenetic level whereas, other mechanisms such as post-translational modifications, regulate SOX2 expression in H9

ESCs. Although it is yet to be determined whether specific concentrations of VPA and ZEB induce the differentiation of HT29 and DLD1 putative CSCs, the findings reported here suggest that additional mechanism(s), other than epigenetic mechanisms, might modulate SOX2 expression in these cells.

In Part 2 of this chapter, gene expression profiling (qPCR) is used to assess the epigenetic modulatory effects that VPA and ZEB have on *NANOG*, *OCT4* and *SOX2* expression in HT29 and DLD1 putative CSCs. Surprisingly, *SOX2* mRNA transcripts were undetectable, or even absent, in HT29 and DLD1 putative CSCs. This finding might provide an explanation as to why SOX2 protein expression remained unchanged in HT29 and DLD1 putative CSCs following epigenetic drug treatments.

(ii) Gene Expression Profiling (qPCR)

3.3 Materials and Methods

3.3.1 Primer design

Gene specific DNA primers to *NANOG*, *OCT4* and *SOX2* were designed using various bioinformatic tools. The bioinformatic programs used for the design of these primers included: UCSC Genome Bioinformatics (UCSC), Ensembl Genome Browser (Ensembl) and Primer Blast NCBI (NCBI). Ensembl was used to obtain candidate gene sequences. UCSC Genome Bioinformatics (Blat) and Primer Blast NCBI were used to ensure that the DNA primers bound specifically to the candidate genes, and in using *in silico* PCR to ensure an optimal product size of 120-170 base pairs. The primers were specifically designed to span the exons of the candidate genes. All of these primers ranged from 20-30 base pairs in length and contained a GC content of 40-60%. Furthermore, they were designed to have melting temperatures (T_M) that ranged between 55°C and 65°C [calculated using the formula: $T_M = 4(G+C) + 2(A+T)$]. The primers were synthesised either by the Department of Molecular and Cell Biology, University of Cape Town or by Integrated DNA Technologies (Iowa). However, as the designed primers failed to amplify their target regions, possibly due to the existence of pseudogenes (Suo *et al.*, 2005; Zhang *et al.*, 2006), published primers were used. The published primer sequences are shown in

Table 3.1. The coding sequences, accession numbers and mapped primers to these genes are included in Figures D15-D17 (Appendix D).

Table 3.1: Primer pair sequences for *NANOG*, *OCT4* and *SOX2*.

Primer Pair	**Sequence (forward and reverse, respectively)**
NANOG (Freberg *et al.*, 2007)	5'- CAA AGG CAA ACA ACC CAC TT -3' 5'- TCT GCT GGA GGC TGA GGT AT -3'
OCT4 (Freberg *et al.*, 2007)	5'- ACA TCA AAG CTC TGC AGA AAG AAC -3' 5'- CTG AAT ACC TTC CCA AAT AGA ACC C -3'
SOX2 (Chen *et al.*, 2012b)	5'- AAA ACA GCC CGG ACC GCG TC -3' 5'- CTC GTC GAT GAA CGG CCG CT -3'

3.3.2 Drug Treatments

HT29 and DLD1 putative CSCs were cultured in a CSC enrichment medium (Chapter 2.1.6). The culture medium was aspirated off once the cells had reached 60% confluency, and were then serum starved for 12 hours in DMEM:F12 supplemented with 10000 IU penicillin and streptomycin. Following this, the serum-free medium was aspirated off and the cells were pharmacologically treated in triplicate with VPA and ZEB (see Section 3.1.2) for 24 hours. The cells were harvested, centrifuged at 800rpm for 5 minutes and their pellets re-suspended in 600µl RLT lysis buffer (Qiagen). The cell lysates were stored at -70˚C for up to six months.

3.3.3 RNA Extraction

For the extraction and purification of total RNA, the Qiagen RNA extraction kit was used and the protocol modified to accommodate for the use of both on-column DNase treatment (Qiagen) and off-column DNase treatment (New England Biolabs) (refer to Appendix B.2 for protocol).

The concentration as well as the purity of RNA (260/280 ratio) was measured on a NanoDrop-1000 spectrophotometer (Nanodrop Technologies) and the samples standardised to a concentration of 50ng/µl.

3.3.3.1 Assessment of RNA Integrity

The Agilent RNA 6000 Pico Kit (Agilent Technologies) was used to determine RNA integrity numbers (RINs). This technique is superior to gel electrophoresis, as it is standardised. A software algorithm is applied to RNA following microfluidic electrophoresis, to estimate total RNA integrity, this being classified from a range of one to ten (Schroeder *et al.*, 2006). A RIN of one is representative of completely degraded RNA and a RIN of ten is representative of completely intact RNA (Schroeder *et al.*, 2006).

RNA samples were diluted to 5000pg/µl, in RNase-free water. The RNA ladder and gel-dye mix was prepared according to the manufacturer's protocol, and were loaded into the wells of a RNA chip, together with the RNA conditioning solution and the RNA marker. Next, the RNA samples were loaded into the wells of the RNA chip, and the chip was vortexed for 1 minute at 2400rpm in an IKA vortex. The RNA chip was run on the Agilent 2100 Bioanalyzer and RINs were acquired. The output data is viewed as an electropherogram (Schroeder *et al.*, 2006). Figure 3.11 shows an example of an electropherogram for the RNA sample DLD1 NT1 with a RIN of 8.60. In this study, RINs ranged between 6.30 and 9.60 indicating that all the samples were intact and were of similar integrity (Table C9 in Appendix C).

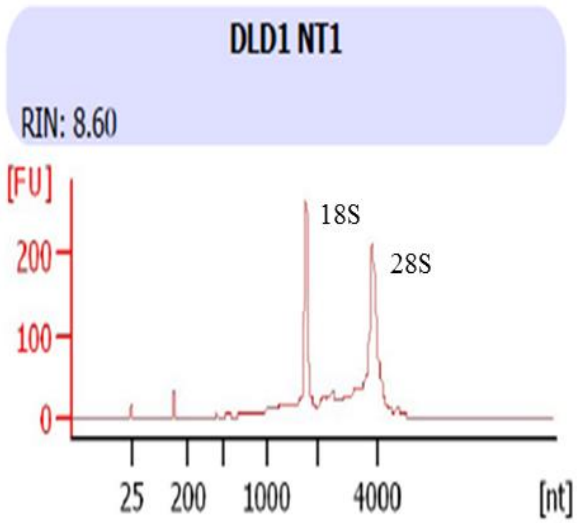

Figure 3.11. An electropherogram of RNA isolated from untreated (NT1) DLD1 putative CSCs showing the 18S (left) and 28S (right) rRNA peaks, respectively.

3.3.4 Reverse Transcription

The TaqMan Reverse Transcription reagents kit (Applied Biosystems) was used to reverse transcribe total RNA to complementary deoxyribonucleic acid (cDNA), using random hexamers, according to the manufacturer's protocol (Table C10 in Appendix C). Two reverse transcription reactions were prepared for each sample; one containing and one excluding, reverse transcriptase (the latter referred to as the negative/minus RT). The latter was performed as a control to rule out genomic DNA contamination. The reverse transcription thermal cycling parameters used are shown in Table C11 (Appendix C), and include an initial incubation at 25°C for 10 minutes, followed by one cycle of reverse transcription at 48°C for 30 minutes and Taq inactivation at 95°C for 5 minutes.

3.3.5 Quantitative Real-Time PCR (qPCR) Optimisation

qPCR reactions were prepared on ice, in duplicate, using 50ng cDNA from the untreated HT29 and DLD1 putative CSCs, with primers to *β-Actin* (*ACTB*) (a reference gene used as a positive control), *NANOG*, *OCT4* and *SOX2*. The qPCR reactions were prepared with a SYBR Green master mix in a total volume of 10µl, according to the manufacturer's protocol (Applied Biosystems) (Table C12 in Appendix C). The reactions were run on the 7500 real-time PCR system (Applied Biosystems) with the following thermal cycling parameters: an initial incubation step at 50°C for 2 minutes, a hot start step at 95°C for 10 minutes, followed by 40 cycles of denaturation at 95°C for 15 seconds and 40 cycles of primer annealing/extension at 55/60/65°C for 1 minute. This was followed by a dissociation/melt curve analysis of the PCR products. Annealing/extension temperatures for primer pairs *ACTB*, *NANOG* and *OCT4* were optimised to 60°C. Primers pairs to *SOX2* could not be optimised as *SOX2* expression was undetectable in untreated HT29 and DLD1 putative CSCs. Based on this latter finding, *SOX2* is further investigated in HT29 and DLD1 putative CSCs, and in their parental cell lines, in Chapter 4.

3.3.6 Selection of Reference Genes using geNORM

geNORM is an algorithm that determines the most stably expressed reference/house-keeping genes tested under certain experimental conditions, to ensure an accurate measure of candidate gene expression (Vandesompele *et al.*, 2002). Reference genes although

constitutively expressed, may differ in expression in various tissues and under different experimental conditions (Spanakis, 1993; Thellin *et al.*, 1999; Bustin, 2000; Suzuki *et al.*, 2000; Warrington *et al.*, 2000). Therefore, it is important to determine and validate the expression of a number of reference genes to ensure reliable normalisation of candidate gene expression.

Ten reference genes were selected based on literature review and availability; these included: *ACTB*, *GUS*, *β2M*, *GAPDH*, *HPRT*, *TBP*, *UBC*, *YWHAZ* and *TR* (Integrated DNA Technologies). Primers to these reference genes were obtained from the Department of Internal Medicine, University of the Witwatersrand, and their sequences are shown Table C13 (Appendix C).

qPCR reactions were prepared, in duplicate, using 50ng cDNA from HT29 and DLD1 CSCs that had been treated in the following ways: untreated, 5mM VPA, 500µM ZEB and DMSO, respectively. qPCR reactions were prepared in a total volume of 10µl as previously described (Section 3.3.5). The plates were run on the 7500 real-time PCR system (Applied Biosystems) with the aforementioned thermal cycling parameters (annealing/extension at 60˚C). The data were exported and analysed using geNORM (qBasePlus). The reference genes were ranked according to their expression stability. A minimum of three reference genes were required for reliable normalisation, these being *β-Actin* (*ACTB*), *Hypoxanthine phosphoribosyltransferase 1* (*HPRT*) and *Tyrosine 3-monooxygenase/tryptophan 5- monooxygenase activation protein, zeta polypeptide* (*YWHAZ*). The coding sequences, accession numbers and mapped primers to these reference genes are included in Figures D18-D20 (Appendix D).

3.3.7 qPCR of *NANOG* and *OCT4*

Following geNORM analysis, qPCR reactions were prepared in duplicate as above (Section 3.3.5) using 50ng cDNA from all of the treatment groups with primers to: *ACTB*, *HPRT*, *YWHAZ* (reference genes), and with primers to *NANOG* and *OCT4* (genes of interest). Controls included a minus RT control (genomic deoxyribonucleic acid (gDNA) contamination), a H_2O control (PCR product contamination) and an inter-plate calibrator (IPC) control. The IPC control consisted of a mix of cDNA from the HT29 and DLD1

treated putative CSCs to correct for inter-run variation when samples were distributed across plates.

The following equation was obtained from the TATAA Interplate Calibrator SYBR protocol (TATAA Biocenter) and was used to manually adjust quantification cycle (C_q) values to correct for inter-plate variation:

$$Cq_i^{\text{corrected}} = Cq_i^{\text{uncorrected}} - Cq_i^{\text{IPC}} + \frac{1}{\text{no.plates}} \sum_{i=1}^{no.plates} Cq_i^{\text{IPC}}$$

Having adjusted the C_q values, the data was analysed using qBasePlus (Biogazelle). qBasePlus uses the following equation as a quantification model:

$$NRQ = \frac{E_{goi}^{\Delta Ct,goi}}{\sqrt[n]{\prod_{i}^{n} E_{ref_i}^{\Delta Ct,ref_i}}}$$

This model allows for the use of several reference genes and is compliant with the minimum information for publication of quantitative real-time PCR experiments (MIQE) guidelines (Bustin *et al.*, 2009) and therefore it was the preferred method. The $\log_{10} 2^{-\Delta\Delta Cq}$ data, which is gene expression data that has been calibrated and normalised using the $2^{-\Delta\Delta Cq}$ method (Livak & Schmittgen, 2011), are shown in Tables C14 and C16 in Appendix C.

3.3.8 Statistical Analyses

The Kruskal-Wallis and Mann-Whitney tests were used to assess statistical significance in pluripotent marker expression, with respect to drug treatment (see Section 3.1.4 for a description of these statistical tests). Statistical significance was considered at a *p*-value $\leq$ 0.05. The sign test was also considered to examine the data for trends; however, as previously mentioned (page 65), a minimal sample size of seven is required to evaluate data using this test.

3.4 Results and Discussion

In this aspect of the study, qPCR was used to quantify any changes that VPA and ZEB have on *NANOG*, *OCT4* and *SOX2* gene expression in HT29 and DLD1 putative CSCs. As *SOX2* expression was absent, or at least undetectable, in HT29 and DLD1 putative CSCs, only *NANOG* and *OCT4* gene expression was quantified. The raw data obtained from the qPCR assay showed *NANOG* and *OCT4* expression. The C_q values of the raw data were manually adjusted to correct for inter-plate variation, and the data was analysed using qBasePlus. The gene expression data ($\log_{10}{}^{2-\Delta\Delta Cq}$) from qBasePlus is quantitative data that has been calibrated and normalised relative to the reference genes *ACTB*, *HPRT* and *YWHAZ*, and to the global mean expression of *NANOG* and *OCT4* (Hellemans *et al.*, 2007). The gene expression data are shown in Tables C14 and C16 (Appendix C). The median expression of *NANOG* and *OCT4* prior to and following drug treatments were plotted onto bar graphs (Figures 3.12 and 3.13) and the effects of VPA and ZEB on their expression were statistically analysed relative to untreated and DMSO treated cells, respectively. The results of the statistical analyses are shown in Table C18 (Appendix C).

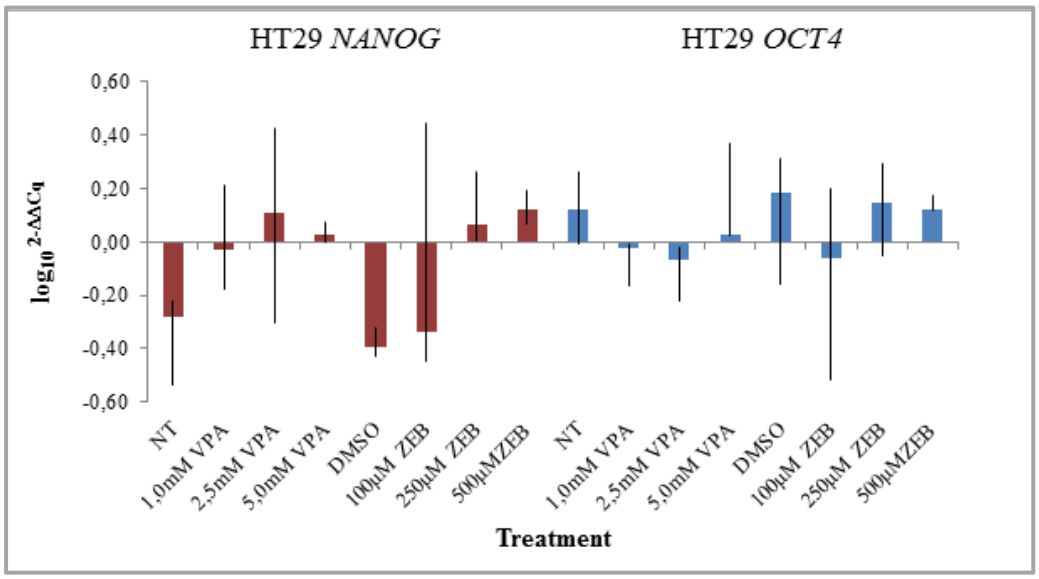

Figure 3.12. Relative gene expression of *NANOG* and *OCT4* in VPA and ZEB treated HT29 putative CSCs. The columns represent the sample median and the vertical lines show the maximum and minimum ranges (*p≤0.05; N=3).

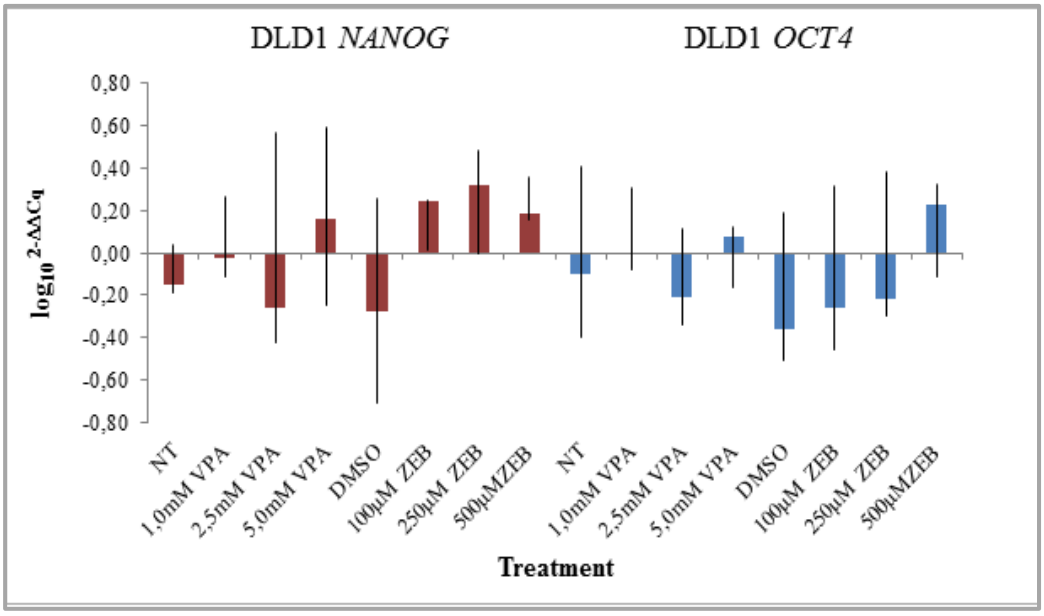

Figure 3.13. Relative gene expression of *NANOG* and *OCT4* in VPA and ZEB treated DLD1 putative CSCs. The columns represent the sample median and the vertical lines show the maximum and minimum ranges (*p≤0.05; N=3).

Statistical analyses show that VPA and ZEB do not modulate *NANOG* and *OCT4* gene expression in both HT29 and DLD1 putative CSCs. Statistical significance may have been obscured due to the large range in gene expression, as seen in Figures 3.12-3.13. Despite best efforts to control for inter-plate variation, experimental error is always a concern. This is apparent in the data from the third run of the replicates, as the gene expression values are consistently higher than those from the first and second runs. The RNA used for the first and second runs was extracted from HT29 and DLD1 putative CSCs that were drug treated simultaneously, whereas the RNA used for third run was extracted from CSCs that were drug treated approximately three months later. Although the cells were cultured under the same conditions, these cultures may change over time; which may account for some of the variation and potential experimental error observed in the third run.

Overall, the modulatory effects that VPA and ZEB have on *NANOG* and *OCT4* gene expression could not be determined in this aspect of the study, with sample size being a limiting factor. Even though these results are currently inconclusive, these drugs were shown to up-regulate nuclear NANOG and OCT4 protein expression in HT29 putative CSCs. Previous studies have shown that altering the expression of these pluripotent

associated transcription factors above and below endogenous levels in embryonic and adult SCs stimulates cell differentiation (Niwa *et al.*, 2000; Wang *et al.*, 2003; Matin *et al.*, 2004; Hyslop *et al*., 2005; Niwa *et al.*, 2005; Zaehres *et al.*, 2005; Rodriguez *et al.*, 2007; Alvarez *et al.*, 2010; Kallas *et al.*, 2014). Therefore, it would be of future interest to determine whether the up-regulation of NANOG and OCT4 in HT29 putative CSCs is associated with cell differentiation, as this may sensitise these cells to current anti-cancer therapies. Differentiation therapy holds promise as the use of the differentiation inducing agent Retinoic acid has been shown to have high cure rates in patients with acute promyelocytic leukaemia (Huang *et al.*, 1988). Moreover, differentiation therapy involving the use of BMP4 molecules has been shown to stimulate the differentiation of putative colorectal adenocarcinoma CSCs *in vitro* and enhance their sensitivity to the chemotherapeutic agents, 5-Fluorouracil and Oxaliplatin *in vivo* (Lombardo *et al.*, 2011). Therefore, stimulating putative CSC differentiation with epigenetic drugs that target the mechanisms involved in maintaining pluripotency may enhance their sensitivity to current chemo- and radiation-therapies. Ultimately, this approach could lead to tumour regression and prevent tumour relapse.

3.5 Conclusion

In this chapter, the effects of VPA and ZEB on the expression of the SC associated pluripotent transcription factors NANOG, OCT4 and SOX2 were investigated in HT29 and DLD1 putative CSCs. Confocal microscopy showed that CD133 and EpCAM were expressed on the cell membrane and in the cytoplasm; and that NANOG, OCT4 and SOX2 were expressed in the nucleus and cytoplasm in HT29 and DLD1 putative CSCs. Specific concentrations of VPA and ZEB were shown to up-regulate the expression of nuclear NANOG and OCT4 in HT29 putative CSCs whereas, neither drug modulated the expression of these transcription factors in DLD1 putative CSCs. Moreover, SOX2 expression was not affected by either of these drugs in both HT29 and DLD1 putative CSCs. Next, gene expression profiling was used to quantitatively assess the pharmacological effects of VPA and ZEB on *NANOG*, *OCT4* and *SOX2* gene expression. Interestingly, *SOX2* expression was either undetectable or possibly absent, in HT29 and DLD1 putative CSCs. The results obtained from the gene expression profiling assay of *NANOG* and *OCT4* were inconclusive possibly due to the amount of variability in gene

expression. Bearing in mind that statistical power is affected by sample size, expanding the sample size of this study would improve statistical robustness.

In Chapter 4 of this study, *SOX2* is further considered in HT29 and DLD1 putative CSCs and in their parental cell lines. Moreover, a mechanism that potentially regulates SOX2 function in HT29 and DLD1 putative CSCs is proposed.

CHAPTER 4

POST-TRANSLATIONAL MODIFICATIONS ALTER SOX2 FUNCTION

SOX2 is a member of the SRY-related HMG box (*SOX*) family. It is an intron-less gene located within the *SOX2* overlapping transcript gene sequence (Pruitt *et al.*, 2013). There are two *SOX2* splice variants, these being *SOX2-001* and *SOX2-201*, each has a transcript length of 2500bp and 1318bp, respectively (Flicek *et al.*, 2014). The protein coding sequence of these two variants codes for the same 317 amino acid long protein (Flicek *et al.*, 2014). The accession numbers for these transcripts and proteins are shown in Table C19 (Appendix C).

SOX2 primarily functions to maintain pluripotency in embryonic and adult SCs (Masui *et al.*, 2007). The putative role of *SOX2* in cellular pluripotency involves the dimerisation of SOX2 and OCT4 to regulate the transcription of *NANOG*, and other key pluripotent factors (Boyer *et al*., 2005; Medvedev *et al.*, 2008). Since *SOX2* imbues cells with pluripotent qualities by controlling the expression of a variety of "stemness" factors, its expression is tightly controlled via positive and negative feed-back loops, and by numerous feed-forward loops that respond to *SOX2* regulated gene products (Boyer *et al.*, 2005; Ormsbee *et al.*, 2013). However, the exact mechanism by which *SOX2* is regulated in pluripotent cells is not well understood (Boyer *et al.*, 2005; Ormsbee *et al.*, 2013).

To elucidate a mechanism that modulates *SOX2* gene expression in human embryonic SCs, Ormsbee *et al.* (2013) hypothesised that an increase in endogenous SOX2 inhibits *SOX2* transcription via a negative feed-back loop. They tested this hypothesis using genetically engineered embryonic SCs that inducibly express *OCT4*, *SOX2*, *KLF-4* and *c-MYC* (i-OSKM-ESC) when treated with Doxycycline. They reported that elevating the expression of these transcription factors resulted in the phosphorylation of AKT, a player in the PI3K signalling pathway, which activates a negative feed-back loop that specifically modulates *SOX2* expression in these cells. In a study by Kuzmichev *et al.* (2012), they showed that SOX2 regulates *Sox21* transcription to repress *Cdx2* in mouse colorectal cancer cells. As

Cdx2 is a repressor of *Sox2* transcription, its repression by *Sox21* allows these cells to maintain pluripotency (Kuzmichev *et al.*, 2012). Together, these studies illustrate that the mechanisms involved in modulating SOX2 expression are complex.

CSCs are thought to trigger tumour initiation, progression and metastasis. Since *SOX2* functions to maintain SC identity in embryonic and adult SCs, its dysregulation in CSCs is likely to play an important role in tumourigenesis. In this regard, *SOX2* dysregulation has been reported in cancers of the breast (Ben-Porath *et al.*, 2008), gastrointestinal tract (Otsubo *et al*., 2008), lung (Bass *et al.*, 2009; Maier *et al.*, 2011), oesophagus (Bass *et al.*, 2009; Maier *et al.*, 2011) and colon (Neumann *et al.*, 2011). A *SOX2* knock-down study by Han *et al.* (2012) showed reduced WNT signalling in the colorectal cell line SW620, resulting in mesenchymal-epithelial transition (MET). As a consequence, these cells lost their ability to metastasise *in vitro* and invade surrounding tissue *in vivo* (Han *et al.*, 2011). In another *SOX2* knock-down study by Alonso *et al.* (2011), brain tumour SCs cultured as neurospheres lost their ability to self-renew. These studies demonstrate the importance of *SOX2* in maintaining self-renewal and pluripotency, and that its aberrant expression may possibly drive cancer pathogenesis.

In the present study, *SOX2* mRNA transcripts were surprisingly undetectable by qPCR assay in both HT29 and DLD1 putative CSCs. Even so, the transcription factor was clearly localised subcellularly in these cells, as previously shown by confocal microscopy. The lack of mRNA transcripts in the presence of expressed proteins possibly suggests unusual post-translational processing of SOX2. In an attempt to determine whether this phenomenon is restricted to putative CSCs, the *SOX2* gene and its expression is further explored in HT29 and DLD1 putative CSCs, and in their parental cell lines. The results obtained in this aspect of the study are interpreted by a conceptual model.

4.1 Materials and Methods

As *SOX2* failed to amplify at the mRNA level in both HT29 and DLD1 putative CSCs, PCR amplification was carried out using genomic DNA (gDNA) to confirm the presence of this gene in HT29 and DLD1 putative CSCs, and in their parental cell lines. gDNA was extracted using the QIAamp DNA Mini Kit (Qiagen), according to the manufacturer's

protocol (Appendix B.5). The concentration, as well as the purity, of gDNA was measured on a NanoDrop-1000 spectrophotometer (Nanodrop Technologies).

Next, PCR reactions with primers that span the *SOX2* coding sequence were prepared, as in section 3.3.5, in a total volume of 10µl (refer to Table 3.1 for primer pair sequences). The reactions were run on the 7500 real-time PCR system (Applied Biosystems) with the following thermal cycling parameters: an initial incubation step at 50°C for 2 minutes, a hot start step at 95°C for 10 minutes, followed by 40 cycles of denaturation at 95°C for 15 seconds and 40 cycles of primer annealing/extension at 65°C for 1 minute. A H_2O control was included to rule out reagent contamination. For gel electrophoresis, 5µl of PCR product was mixed with 2µl loading dye, and was run on a 1.5% agarose gel for approximately one hour at 100V in 1X TBE buffer (Appendix A.3). The gel was analysed on the Bio-Rad Gel Doc EZ Imager system, together with the Image Lab software (version 5). A dissociation/melt curve was generated on the 7500 real-time PCR system (Applied Biosystems) with the remaining 5µl of PCR products.

Having confirmed the presence of the *SOX2* gene in HT29 and DLD1 putative CSCs, and in their parental cell lines, a subsequent PCR assay was carried out to evaluate its expression in these cells. RNA was extracted from the HT29 and DLD1 putative CSCs, and from their parental cell lines, as previously described in Chapter 3.3.3. The RNA from each of the samples was standardised to 50ng/µl and reverse transcribed (see Chapter 3.3.4). Next, PCR reactions with primers to *SOX2* and *ACTB* (positive control) were prepared, as in Chapter 3.3.5, in a total volume of 10µl. The reactions were run on the 7500 real-time PCR system (Applied Biosystems) with the same thermal cycling parameters described above. Negative/minus RT controls and H_2O controls were included to rule out gDNA contamination and reagent contamination, respectively. For gel electrophoresis, 5µl of PCR product was mixed with 2µl loading dye, and run on a 1.5% agarose gel for approximately one hour at 100V in 1X TBE buffer (Appendix A.3). The gel was analysed on the Bio-Rad Gel Doc EZ Imager system, together with the Image Lab software (version 5). As previously, a dissociation curve was generated on the 7500 real-time PCR system (Applied Biosystems) with the remaining 5µl of PCR products.

4.2 Results

As there was no amplification in the minus RT and H_2O controls, cDNA samples were neither compromised by gDNA/reagent contamination, nor by primer dimer formation. The PCR analysis of the gDNA confirmed the presence of the *SOX2* gene in both the HT29 and DLD1 putative CSC lines, and in their parental cell lines, with an amplicon of 176bp (Figure 4.1 A). The dissociation curve analysis produced a reproducible melting temperature of 87.12°C for each sample analysed (Figure 4.1 B). A subsequent PCR study using cDNA as a template, together with gel electrophoresis and melt curve analysis, shows that *SOX2* was not expressed at detectable levels in either HT29 or DLD1 putative CSC lines. Whereas, in the HT29 and DLD1 parental cell lines it was expressed at low levels when compared to *ACTB* (Figure 4.2 A).

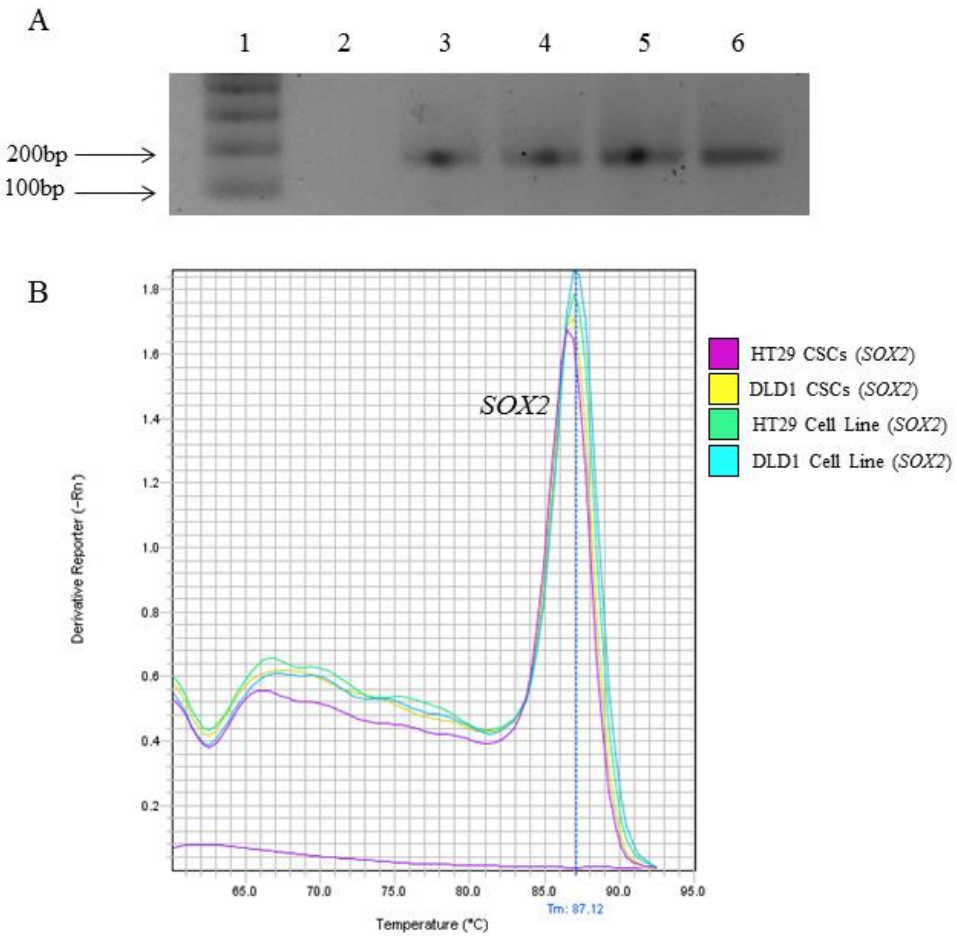

Figure 4.1. A 1.5% agarose gel (A), and its corresponding dissociation/melt curve (B) showing the presence of the *SOX2* amplicon (176bp) in gDNA isolated from DLD1 and HT29 cells, and from their associated putative CSCs. In Figure 4.1 A, lane 1 is a 100bp DNA ladder (Fermentas); lane 2 is a *SOX2* H_2O control; lanes 3, 4 are *SOX2* amplicons in DLD1 and HT29 cells respectively; lanes 5, 6 are *SOX2* amplicons in DLD1 and HT29 putative CSCs, respectively. In Figure 4.1 B, the melt curve shows that *SOX2* amplicons melt consistently at 87.12°C.

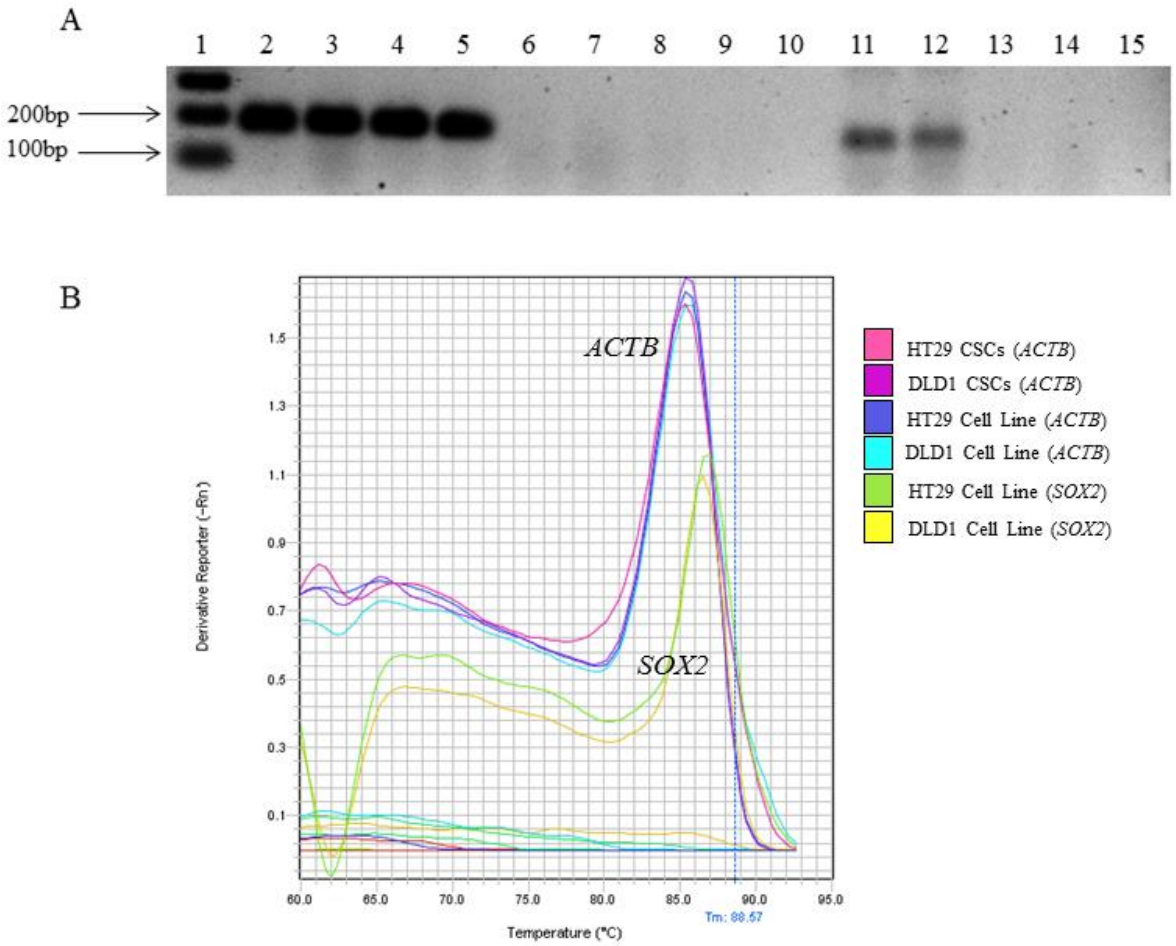

Figure 4.2. A 1.5% agarose gel (A), and its corresponding dissociation/melt curve (B) showing the presence of the *ACTB* (192bp) and *SOX2* amplicons (176bp) in DLD1 and HT29 cells, and the absence of the *SOX2* amplicon in DLD1 and HT29 putative CSCs. In Figure 4.2 A, lane 1 is a 100bp DNA ladder (Fermentas); lanes 2-5 are *ACTB* amplicons in DLD1 cells, HT29 cells, DLD1 putative CSCs and HT29 putative CSCs respectively; lanes 6-9 are *ACTB* minus RTs for all cell types; lane 10 is an *ACTB* H_2O control; lanes 11-12 are *SOX2* amplicons from DLD1 and HT29 cells, respectively; lanes 13-14 show *SOX2* amplicons from DLD1 and HT29 putative CSCs; and lane 15 is a *SOX2* H_2O control. In Figure 4.2 B, the melt curve shows *ACTB* amplicons melt at ~85.3°C in DLD1 and HT29 putative CSCs, and in their parental cells lines; whereas *SOX2* amplicons melt at ~87.12°C in DLD1 and HT29 cells only.

4.3 Discussion

The present findings indicate that the apparent lack (or at least, undetectable levels) of *SOX2* mRNA is characteristic of the HT29 and DLD1 putative CSCs, and not of their corresponding parental cell lines. The absence of *SOX2* mRNA transcripts in putative CSCs was unusual as the SOX2 protein was previously shown to localise subcellularly in these cells. This deficiency cannot be accounted for by the genotype of these cells, as the *SOX2* gene sequence was confirmed to be present in both the HT29 and DLD1 putative CSC lines, and in their parental cell lines. Moreover, the absence of *SOX2* mRNA does not result from alternative splicing as the protein coding sequences of its two variants are the same (Flicek *et al.*, 2014). In this aspect of the study, the *SOX2* primers used for the PCR assay spanned the protein coding sequence. As *SOX2* is known to regulate its own transcription via a negative feed-back loop (Ormsbee *et al.*, 2013), it is possible that post-translational modifications of SOX2 alter the negative regulatory feed-back mechanism in putative CSCs.

In embryonic and adult SCs, post-translational modifications such as sumoylation, phosphorylation and acetylation are known to modulate SOX2 function in these cells (Tsuruzoe *et al.*, 2006; Baltus *et al.*, 2009; Jeong *et al.*, 2010). Post-translational modifications alter protein function by affecting DNA binding capability, localisation and stability (Jeong *et al.*, 2012). Sumoylation involves the conjugation of small ubiquitin-related modifier (SUMO) proteins to cellular proteins and affects their ability to bind to target DNA. Such alterations to the DNA-binding properties of proteins may either enhance or prohibit this interaction (Johnson, 2004; Baltus *et al.*, 2009). Sumoylation alters SOX2 function by reducing its binding affinity to DNA, thus inactivating its transcriptional activity and pluripotent function (Tsuruzoe *et al.*, 2006). Phosphorylation of SOX2 enhances its nuclear stability by blocking the ubiquitination process that would otherwise result in proteolysis (Jeong *et al.*, 2010) whereas, acetylation of SOX2 results in its nuclear export to the cytoplasm where it can no longer maintain pluripotency (Baltus *et al.*, 2009).

As these post-translational modifications are known to modulate SOX2 function in embryonic and adult SCs, it is possible that these modifications are tailored to alter SOX2

function in putative CSCs. The regulation of *SOX2* gene expression and SOX2 function in HT29 and DLD1 putative CSCs, and in their parental cell lines, may be explained by the model developed through this study (Figure 4.3). It is proposed here that in HT29 and DLD1 putative CSCs, post-translational modifications of SOX2 result in high levels of active/stable SOX2 in the nucleus. These modifications might include SOX2 phosphorylation (Jeong *et al.*, 2010). It is here within the nucleus, that active SOX2 associates with OCT4, where together they up-regulate the transcription of SC associated pluripotent genes to imbue cells with pluripotency potential (Boyer *et al*., 2005; Medvedev *et al.*, 2008). Moreover, a high level of active SOX2 in the nucleus prevents further *SOX2* transcription via a negative feed-back loop (Ormsbee *et al.*, 2013). Eventually, *SOX2* mRNA is degraded and this would account for its absence in these cells. Alternatively, *SOX2* transcription might occur at levels low enough to maintain a SOX2 threshold level however, these levels may be too low to detect with PCR assay.

Conversely, in the HT29 and DLD1 cell lines, it is proposed here that post-translational modifications of nuclear SOX2 reduces its binding affinity to DNA (referred to as inactive SOX2) and/or results in its export to the cytoplasm. These post-translational modifications might include sumoylation and acetylation, respectively (Tsuruzoe *et al.*, 2006; Baltus *et al.*, 2009). Such *SOX2* status would result in the reduced expression of pluripotency associated transcription factors followed by cell differentiation. Moreover, a constant low level of active SOX2 in the nucleus would allow for continuous SOX2 transcription and this would account for its detection by PCR assay in the HT29 and DLD1 cell lines.

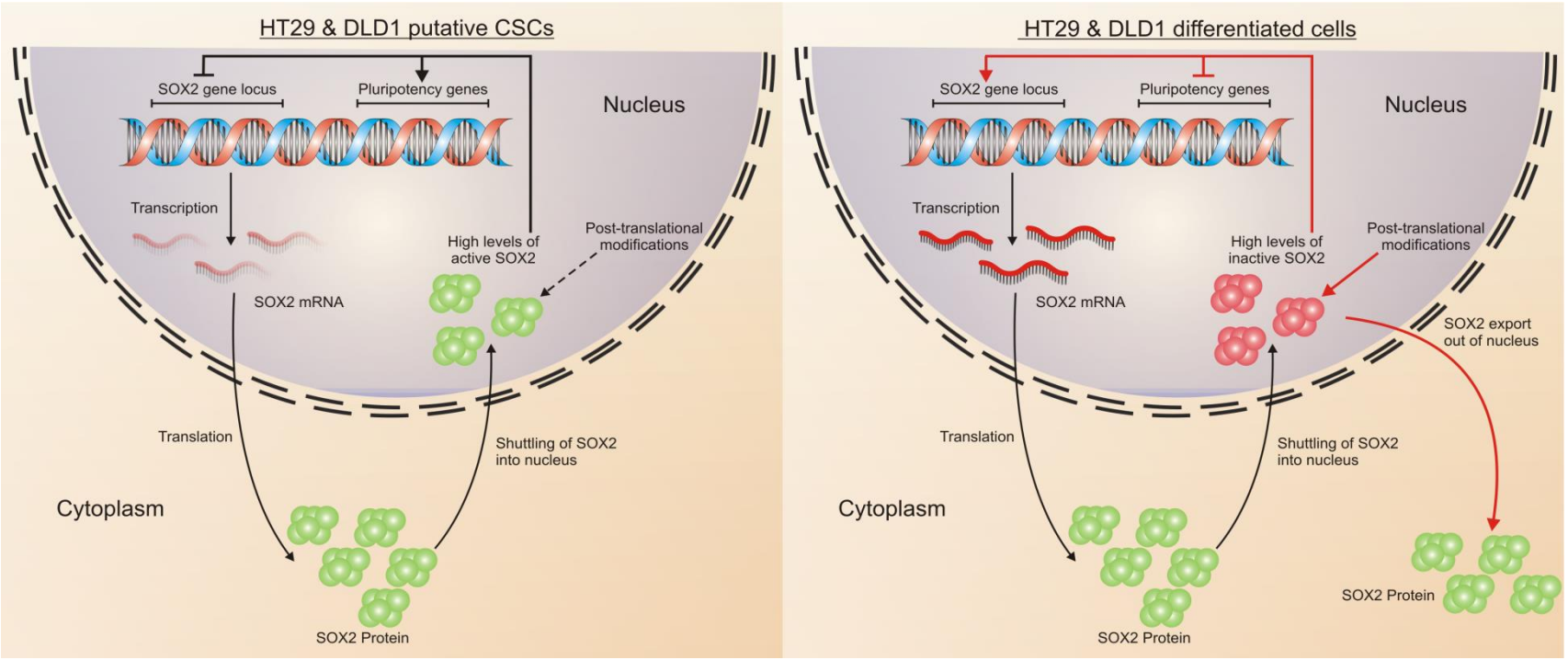

Figure 4.3. A conceptual model showing the regulation of *SOX2* gene expression in HT29 and DLD1 putative CSCs *versus* its regulation in HT29 and DLD1 differentiated cells. In both putative CSCs and differentiated cells, *SOX2* is transcribed and its mRNA is shuttled from the nucleus to the endoplasmic reticulum where it is translated. SOX2 protein is then shuttled into the nucleus where it is subjected to post-translational modifications. In HT29 and DLD1 putative CSCs, post-translational modifications result in high levels of active SOX2 in the nucleus that functions to maintain pluripotency in these cells. As well, SOX2 transcription is further prevented via a negative feed-back loop and its mRNA is degraded. In contrast, post-translational modifications of SOX2 in differentiated HT29 and DLD1 cells results in low levels of active SOX2 in the nucleus. This allows for SOX2 transcription and reduces the expression of pluripotency associated genes.

4.4 Conclusion

In this chapter, the *SOX2* gene and its expression at the mRNA level were further investigated in HT29 and DLD1 putative CSCs, and in their respective parental cell lines. PCR assay analyses showed that the *SOX2* gene was present in all of these cell types although, its expression was restricted to the HT29 and DLD1 cell lines. As the genotype of these cells did not account for the absence of *SOX2* expression in HT29 and DLD1 putative CSCs, and that alternative splicing was not responsible for its absence, a model involving post-translational modulation of SOX2 function in these cells was proposed. In HT29 and DLD1 putative CSCs, it was hypothesised that selective post-translational modifications of SOX2, such as phosphorylation, result in high levels of active SOX2 in the nucleus. Active SOX2 maintains pluripotency by increasing the transcription of pluripotent genes and by repressing its own transcription via negative feed-back. However, within the general population of HT29 and DLD1 cells, selective post-translational modifications, including sumoylation and acetylation, may result in low levels of active nuclear SOX2. This leads to reduced pluripotent gene expression and allows for increased *SOX2* transcription. Thus, post-translational modifications of SOX2 may possibly act as a bi-directional switch to modulate *SOX2* expression.

CHAPTER 5

CONCLUDING REMARKS AND FUTURE RESEARCH

The possible generation of cancer stem cells (CSCs) from adult stem cells (SCs) during early carcinogenesis may involve epigenetic reprogramming leading to altered gene expression, wherein genes associated with cell differentiation may be silenced and SC specific genes up-regulated. In cancer, the continued existence of treatment resistant tumour associated SCs poses a therapeutic challenge. A better understanding of epigenetic events in cancer SC biology may lead to the development of new therapies, particularly since epigenetic changes are potentially reversible.

The present study assessed the epigenetic modulation of key pluripotency genes in putative colorectal CSCs at both a protein and gene expression (mRNA) level. The cells were selectively isolated from colorectal adenocarcinoma cell lines on the basis of their expression of the cell surface markers CD133 and EpCAM. CD133 was found to be expressed at lower levels in the early stage colorectal adenocarcinoma cell line SW1116 (Dukes' stage A) and more abundantly in the later stage HT29 and DLD1 cell lines (Dukes' stage B and C, respectively); whereas, EpCAM was expressed at high levels in all the cell lines representative of Dukes' stages.

The later stage putative CSCs derived from the HT29 and DLD1 cell lines showed clonality, a reproducible feature of SCs. The effects of two epigenetic drugs, Valproic acid (VPA), a histone deacetylase inhibitor, and Zebularine (ZEB), a DNA methyltransferase inhibitor, were assessed on gene and protein expression of *NANOG*, *OCT4* and *SOX2*, these being universal genes necessary for the maintenance of "stemness" and thus continued cell proliferation. VPA and ZEB up-regulated nuclear NANOG and OCT4 expression in the HT29 derived CSCs, while the effects of these drugs on nuclear NANOG and OCT4 expression in the DLD1 derived CSCs were unclear. Nuclear SOX2 protein expression remained unchanged in either of the late stage derived CSCs. The gene expression profiling studies indicated a lack of modulation of both *NANOG* and *OCT4* in

both the HT29 and DLD1 sourced CSCs, while *SOX2* mRNA was undetectable. The lack of detectable *SOX2* mRNA transcripts was a feature that was unusual yet unique to these cells and in this regard, a regulatory model was proposed to account for this observation. These findings however suggest a potential post-translational response of *NANOG* and *OCT4* to epigenetic treatments in the HT29 derived CSCs, while *SOX2* would appear to be unresponsive to either epigenetic drug in CSCs derived from both the HT29 and DLD1 cell lines. Altering the expression of either transcription factor may possibly induce cell differentiation; however, further research would be needed to confirm this. Should VPA and ZEB induce cell differentiation, these drugs when combined with current anti-cancer therapies, may result in tumour regression and prevent tumour relapse. A model that captures the concept of combining CSC targeted therapy with current therapy is shown in Figure 5.1.

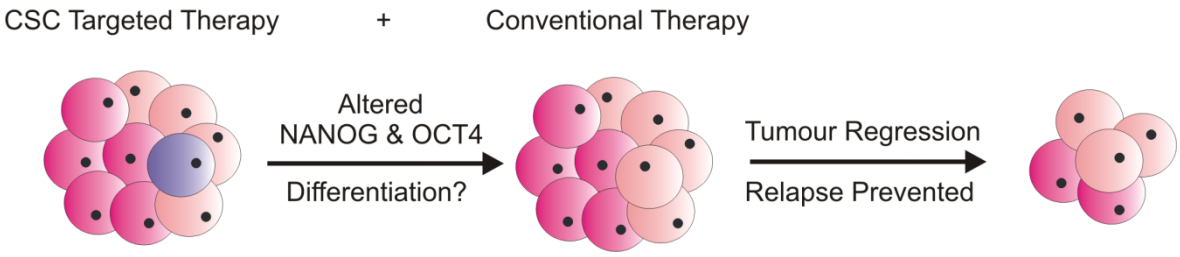

Figure 5.1. A conceptual model showing how combined CSC targeted therapy and conventional therapy results in tumour regression and prevents relapse (adapted from Figure 3 in Dalerba *et al.*, 2007a, 279 p.).

5.1 Future Directions

In this study, CD133 was shown as a valuable marker to isolate and characterise putative CSCs from the HT29 and DLD1 colon adenocarcinoma cell lines. However, as the sphere formation assay indicated that only 21% of these cells had clonogenic potential *in vitro*, it would therefore be of future interest to determine whether additional markers, such as CD44, CD166 and CD24 could further enrich putative CSC isolation from these cell lines. It would also be relevant to further characterise $CD133^{+}$ HT29 and DLD1 cells based on their invasive potential and sensitivity to current chemotherapeutic drugs *in vitro*; and to evaluate the tumourigenic potential of these cells using a xenotransplantation assay, as this method is the gold standard for demonstrating "stemness" *in vivo* (Wolf *et al.*, 2014).

Moreover, as some studies have shown that $CD133^-$ cells show clonogenic potential *in vitro* and tumourigenic potential *in vivo* (Shmelkov *et al.*, 2008; Kawamoto *et al.*, 2010; Hsu *et al.*, 2013), it would also be important to determine whether $CD133^-$ HT29 and DLD1 cells also show features of "stemness".

The treatment of HT29 putative CSCs with VPA and ZEB were shown to up-regulate nuclear NANOG and OCT4 in these cells but their effects on transcription factor modulation in DLD1 putative CSCs were unclear. Moreover, there was no clear evidence showing that these drugs modulate *NANOG* and *OCT4* gene expression at the mRNA level in HT29 and DLD1 putative CSCs. As sample size may have been a limiting factor in this study, it would be of future interest to increase the sample size and to verify the pharmacological effects of VPA and ZEB on *NANOG* and *OCT4* gene expression in these cells. It would be important to determine whether VPA and ZEB induce cell differentiation following NANOG and OCT4 modulation, in the HT29 and DLD1 derived CSCs, as these drugs may potentially prevent tumour relapse when combined with current anti-cancer therapies. An experimental approach would be to use flow cytometry/immunofluorescence microscopy together with antibodies to differentiation-specific markers, such as CK20, a marker specific to differentiated colorectal cells (Wildi *et al.*, 1999; Chen & Wang, 2004).

Since pluripotency of embryonic and adult SCs is also maintained by the pluripotent genes, *KLF-4* (Leng *et al.*, 2013) and *c-MYC* (Takahashi & Yamanaka, 2006; Wang *et al.*, 2008), as well as by a network of signalling pathways, including WNT and BMP4 (Medvedev *et al.*, 2008; Roy & Majumdar, 2012), it would be relevant to investigate the potential effects of VPA and ZEB on these markers, and on signal transduction, in HT29 and DLD1 putative CSCs. The epigenetic status of these pluripotency associated genes, pre-and post-treatment would also be of interest as epi-mutations are known to drive tumour initiation in the colon.

Lastly, in the context of the previously proposed model, the regulation of SOX2 function in HT29 and DLD1 putative CSCs may be of some importance. It would therefore be of future interest to elucidate the mechanism(s) that modulate SOX2 function in these cells. A relevant approach here would be to assess the interaction of SOX2 with partner binding

proteins, using chromatin immunoprecipitation (ChIP) assays. It is suggested that elucidating the underlying cell biological mechanism(s) involved in *SOX2* regulation, may well lead to the identification of potential targets for anti-cancer therapies.

REFERENCES

Adachi, K., Suemori, H. & Yasuda, S.Y., *et al*. 2010. Role of SOX2 in maintaining pluripotency of human embryonic stem cells. *Genes Cells*, vol. 15, no. 5, pp. 455-470.

Allahverdiyev, A.M., Bagirova, M. & Oztel, O.N., *et al*. 2012. *Aldehyde Dehydrogenase: Cancer and Stem Cells, Dehydrogenases*. Rijeka: In Tech, doi: 10.5772/48591, 18 p.

Al-Hajj, M., Wicha, M.S. & Benito-Hernandez, A., *et al.* 2003. Prospective identification of tumorigenic breast cancer cells. *Proc Natl Acad Sci USA*, vol. 100, no. 7, pp. 3983-3988.

Allen, J.E., Hart, L.S. & Dicker, D.T., *et al*. 2009. Visualization and enrichment of live putative cancer stem cell populations following p53 inactivation or Bax deletion using non-toxic fluorescent dyes. *Cancer Biol Ther*, vol. 8, no. 22, pp. 2194-2205.

Alonso, M.M., Diez-Valle, R. & Manterola, L., *et al.* 2011. Genetic and epigenetic modifications of Sox2 contribute to the invasive phenotype of malignant gliomas. *PLoS One*, vol. 6, no. 11, article ID e26740.

Alvarez, A., Hossain, M. & Dantuma, E., *et al.* 2010. Nanog overexpression allows human mesenchymal stem cells to differentiate into neural cells-Nanog transdifferentiates mesenchymal stem cells. *Scientific Res*, vol. 1, no. 1, pp.1-13.

Amini, S., Fathi, F. & Mobalegi, J., *et al.*, 2014. The expressions of stem cell markers: Oct4, Nanog, Sox2, nucleostemin, Bmi, Zfx, Tcl1, Tbx3, Dppa4, and Esrrb in bladder, colon, and prostate cancer, and certain cancer cell lines. *Anat Cell Biol*, vol. 47, no. 1, pp. 1-11.

Amit, M., Carpenter, M.K. & Inokuma, M.S., *et al*. 2000. Clonally derived human embryonic stem cell lines maintain pluripotency and proliferative potential for prolonged periods of culture. *Dev Biol*, vol. 227, no. 2, pp. 271-278.

Anderson, E.C., Hessman, C. & Levin, T.G., *et al.* 2011. The role of colorectal cancer stem cells in metastatic disease and therapeutic response. *Cancers (Basel)*, vol. 3, no. 1, pp. 319-339.

Armstrong, A. & Eck, S.L. 2003. EpCAM: a new therapeutic target for an old cancer antigen. *Cancer Biol Ther*, vol. 2, no. 4, pp. 320-326.

Arndt, G.M., Dossey, L. & Cullen, L.M., *et al.* 2009. Characterization of global microRNA expression reveals oncogenic potential of miR-145 in metastatic colorectal cancer. *BMC Cancer*, doi: 10.1186/1471-2407-9-374.

Ashktorab, H., Belgrave, K. & Hosseinkhah, F., *et al*. 2009. Global histone H4 acetylation and HDAC2 expression in colon adenoma and carcinoma. *Dig Dis Sci*, vol. 54, no. 10, pp. 2109-2117.

Avilion, A.A., Nicolis, S.K. & Pevny, L.H., *et al*. 2003. Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev*, vol. 17, no. 1, pp. 126-140.

Baltus, G.A., Kowalski, M.P. & Zhai, H., *et al.* 2009. Acetylation of sox2 induces its nuclear export in embryonic stem cells. *Stem cells*, vol. 27, pp. 2175-2184.

Banerji, S. & Los, M. 2006. Important differences between topoisomerase-I and -II targeting agents. *Cancer Biol Ther*, vol. 5, pp. 965-966.

Bánky, B., Rásó-Barnett, L. & Barbai, T., *et al*. 2012. Characteristics of CD44 alternative splice pattern in the course of human colorectal adenocarcinoma progression. *Mol Cancer*, doi: 10.1186/1476-4598-11-83.

Bannister, A.J. & Kouzarides, T. 2011. Regulation of chromatin by histone modifications. *Cell Res*, vol. 21, no. 3, pp. 381-395.

Barchi, J.J., Musser, S. & Marquez, V.E. 1992. The decomposition of 1-(beta-D-ribofuranosyl)-1,2-dihydropyrimidin-2-one (Zebularine) in alkali: mechanism and products. *J Org Chem*, vol. 57, pp. 536-541.

Barker, N. & Clevers, H. 2007. Tracking down the stem cells of the intestine: strategies to identify adult stem cells. *Gastroenterology*, vol. 133, no. 6, pp. 1755-1760.

Bass, A.J., Watanabe, H. & Mermel, C.H., *et al*. 2009. SOX2 is an amplified lineage-survival oncogene in lung and esophageal squamous cell carcinomas. *Nat Genet*, vol 41, pp. 1238-1242.

Ben-Porath, I., Thomson, M.W. & Carey, V.J., *et al*. 2008. An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. *Nat. Genet*, vol. 40, pp. 499-507.

Berdasco, M. & Esteller, M. 2010. Aberrant epigenetic landscape in cancer: how cellular identity goes awry. *Dev Cell*, vol. 19, no. 5, pp. 698-711.

Bertrand, F.E., Angus, C.W. & Partis, W.J., *et al*. 2012. Developmental pathways in colon cancer: crosstalk between WNT, BMP, Hedgehog and Notch. *Cell Cycle*, vol. 11, no. 23, pp. 4344-4351.

Billam, M., Sobolewski, M.D. & Davidson, N.E. 2010. Effects of a novel DNA methyltransferase inhibitor Zebularine on human breast cancer cells. *Breast Cancer Res Treat*, vol. 120, no. 3, pp. 581-592.

Blaheta, R.A. & Cinatl, J. 2002. Anti-tumor mechanisms of Valproate: a novel role for an old drug. *Med Res Rev*, vol. 22, no. 5, pp. 492-511.

Blaheta, R.A., Michaelis, M. & Driever, P.H., *et al.* 2005. Evolving anticancer drug Valproic acid: insights into the mechanism and clinical studies. *Med Res Rev*, vol. 25, no. 4, pp. 383-397.

Boland, C.R. & Goel, A. 2010. Microsatellite instability in colorectal cancer. *Gastroenterology*, vol. 138, no. 6, pp. 2073-2087.

Bolton, E., Wang, Y. & Thiessen, P.A., *et al.* 2008. *PubChem: Integrated Platform of Small Molecules and Biological Activities*. Oxford: Elsevier, vol. 4, pp. 217-240.

Bonnet, D. & Dick, J.E. 1997. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nat Med*, vol. 2, no. 7, pp. 730-737.

Booth, C. & Potten, C.S. 2000. Gut instincts: thoughts on intestinal epithelial stem cells. *J Clin Invest*, vol. 105, no. 11, pp. 1493-1499.

Botchkina, I.L., Rowehl, R.A. & Rivadeneira, D.E., *et al*. 2009. Phenotypic subpopulations of metastatic colon cancer stem cells: genomic analysis. *Cancer Genomics Proteomics*, vol. 6, no. 1, pp. 19-29.

Botchkina, G. 2013. Colon cancer stem cells-from basic to clinical application. *Cancer Lett*, vol. 338, no. 1, pp. 127-140.

Bowen, M.A., Patel, D.D. & Li, X., *et al*. 1995. Cloning, mapping, and characterization of activated leukocyte-cell adhesion molecule (ALCAM), a CD6 ligand. *J Exp Med*, vol. 181, no. 6, pp. 2213-2220.

Boyer, L.A., Lee, T.I. & Cole, M.F., *et al.* 2005. Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell*, vol. 122, no. 6, pp. 947-956.

Bracken, A.P. & Helin, K. 2009. Polycomb group proteins: navigators of lineage pathways led astray in cancer. *Nat Rev Cancer*, vol. 9, no. 11, pp. 773-784.

Brunner, A., Prelog, M. & Verdorfer, I., *et al*. 2008. EpCAM is predominantly expressed in high grade and advanced stage urothelial carcinoma of the bladder. *J Clin Pathol*, vol. 61, pp 307-310.

Brunner, A., Schaefer, G. & Veits, L., *et al*. 2008. EpCAM overexpression is associated with high-grade urothelial carcinoma in the renal pelvis. *Anticancer Res*, vol. 28, no. 1A, pp. 125-128.

Burk, U., Schubert, J. & Wellner, U., *et al.* 2008. A reciprocal repression between ZEB1 and members of the miRNA-200 family promotes EMT and invasion in cancer cells. *EMBO Rep*, vol. 9, no. 6, pp. 582-589.

Bustin, S.A., Benes, V. & Garson, J.A., *et al.* 2009. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem*, vol. 55, no. 4, pp. 611-622.

Cauffman, G., Van de Velde, H. & Liebaers, I., *et al.* 2005. Oct-4 mRNA and protein expression during human preimplantation development. *Mol Hum Reprod*, vol. 11, no. 3, pp. 173-181.

Chan, K.K., Zhang, J. & Chia, N.Y., *et al*. 2009. KLF4 and PBX1 directly regulate NANOG expression in human embryonic stem cells. *Stem Cells*, vol. 27, no. 9, pp. 2114-2125.

Chen, K.L., Pan, F. & Jiang, H., *et al*. 2011. Highly enriched CD133(+)CD44(+) stem-like cells with CD133(+)CD44(high) metastatic subset in HCT116 coloncancer cells. *Clin Exp Metastasis*, vol. 28, no. 8, pp. 751-763.

Chen, S., Song, X. & Chen, Z., *et al.* 2013. CD133 expression and the prognosis of colorectal cancer: a systematic review and meta-analysis. *PLoS One*, vol. 8, no. 2, article ID e56380.

Chen, S., Xu, Y. & Chen, Y., *et al*. 2012b. SOX2 gene regulates the transcriptional network of oncogenes and affects tumorigenesis of human lung cancer cells. *PLoS One*, vol. 7, no. 5, article ID e36326.

Chen, Y., Yu, D. & Zhang, H., *et al*. 2012a. CD133(+)EpCAM(+) phenotype possesses more characteristics of tumor initiating cells in hepatocellular carcinoma Huh7 cells. *Int J Biol Sci*, vol. 8, no. 7, pp. 1992-2004

Chen, Z.M. & Wang, H.L. 2004. Alteration of cytokeratin 7 and cytokeratin 20 expression profile is uniquely associated with tumorigenesis of primary adenocarcinoma of the small intestine. *Am J Surg Pathol*, vol. 28, no. 10, pp. 1352-1359.

Cheng, J.C., Matsen, C.B. & Gonzales, F.A., *et al.* 2003. Inhibition of DNA methylation and reactivation of silenced genes by Zebularine. *J Natl Cancer Inst*, vol. 95, no. 5, pp. 339-409.

Cheng, J.C., Weisenberger D.J. & Gonzales F.A., *et al.* 2004. Continuous Zebularine treatment effectively sustains demethylation in human bladder cancer cells. *Mol Cell Biol*, vol. 24, no. 3, pp. 1270-1278.

Chew, J.L., Loh, Y.H. & Zhang, W., *et al*. 2005. Reciprocal transcriptional regulation of Pou5f1 and Sox2 via the Oct4/Sox2 complex in embryonic stem cells. *Mol Cell Biol*, vol. 25, no. 14, pp. 6031-6046.

Chiou, S.H., Wang, M.L. & Chou, Y.T., *et al*. 2010. Coexpression of Oct4 and Nanog enhances malignancy in lung adenocarcinoma by inducing cancer stem cell-like properties and epithelial-mesenchymal transdifferentiation. *Cancer Res*, vol. 15, no. 70, pp. 10433-10444.

Chong, J.M. & Speicher, D.W. 2001. Determination of disulfide bond assignments and N-glycosylation sites of the human gastrointestinal carcinoma antigen GA733-2 (CO17-1A, EGP, KS1-4, KSA, and Ep-CAM). *J Biol Chem*, vol. 276, pp. 5804-5813.

Chu, P., Clanton, D.J. & Snipas, T.S., *et al*. 2009. Characterization of a subpopulation of colon cancer cells with stem cell-like properties. *Int J Cancer*, vol. 124, no. 6, pp. 1312-1321.

Cloos, P., Christensen, J. & Agger, A., *et al.*, 2008. Erasing the methyl mark: histone demethylases at the center of cellular differentiation and disease. *Genes Dev*, vol. 22, no. 9, pp. 1115–1140.

Coco, C., Zannoni, G.F. & Caredda, E., *et al*. 2012. Increased expression of CD133 and reduced dystroglycan expression are strong predictors of poor outcome in colon cancer patients. *J Exp Clin Cancer Res*, vol. 31, no. 71, doi: 10.1186/1756-9966-31-71.

Corbeil, D., Roper, K. & Hellwig, A., *et al.* 2000. The human AC133 hematopoietic stem cell antigen is also expressed in epithelial cells and targeted to plasma membrane protrusions. *J Biol Chem*, vol. 275, pp. 5512-5520.

Corbo, C., Orrù, S. & Gemei, M., *et al.* 2012. Protein cross-talk in CD133+ colon cancer cells indicates activation of the Wnt pathway and upregulation of SRp20 that is potentially involved in tumorigenicity. *Proteomics*, vol. 12, no. 12, pp. 2045-2059.

Cortes-Dericks, L., Yazd, E.F. & Mowla, S.J., *et al*. 2013. Suppression of OCT4B enhances sensitivity of lung adenocarcinoma A549 cells to cisplatin via increased apoptosis. *Anticancer Res*, vol. 33, no. 12, pp. 5365-5373.

Courtney, K.D., Corcoran, R.B. & Engelman, J.A. 2010. The PI3K pathway as drug target in human cancer. *J Clin Oncol*, vol. 28, no. 6, pp. 1075-1083.

Crea, F., Hurt, E.M. & Mathews, L.A., *et al*. 2011. Pharmacologic disruption of Polycomb Repressive Complex 2 inhibits tumorigenicity and tumor progression in prostate cancer. *Mol Cancer*, doi: 10.1186/1476-4598-10-40.

Curtin, K., Slattery, M.L. & Samowitz, W.S. 2011. CpG island methylation in colorectal cancer: past, present and future. *Patholog Res Int*, vol. 2011, doi: 10.4061/2011/902674.

Dalerba, P., Cho, R.W. & Clarke, M.F. 2007a. Cancer stem cells: models and concepts. *Annu Rev Med*, vol. 58, pp. 267-284.

Dalerba, P., Dylla, S.J. & Park, I.K., *et al.* 2007b. Phenotypic characterization of human colorectal cancer stem cells. *Proc Natl Acad Sci USA*, vol. 104, no. 24, pp. 10158-10163.

Debeb, B.G., Xu, W. & Mok, H., *et al.* 2010. Differential radiosensitizing effect of Valproic acid in differentiation versus self-renewal promoting culture conditions. *Int J Radiat Oncol Biol Phys*, vol. 76, no. 3, pp. 889-895.

de Gramont, A., Figer, A. & Seymour, M., *et al*. 2000. Leucovorin and fluorouracil with or without oxaliplatin as first-line treatment in advanced colorectal cancer. *J Clin Oncol*, vol. 18, no. 16, pp. 2938-2347.

De Paiva, C.S., Pflugfelder, S.C. & Li, D-Q. 2006. Cell size correlates with phenotype and proliferative capacity in human corneal epithelial cells. *Stem Cells*, vol. 24, no. 2, pp. 368-375.

DeSantis, C. E., Lin, C. C. & Mariotto, A. B., *et al.* 2014. Cancer treatment and survivorship statistics, 2014. *CA: A Cancer Journal for Clinicians*, vol. 64, pp. 252-271.

de Sousa, A.R., Penalva, L.O. & Marcotte, E.M., *et al.* 2009. Global signatures of protein and mRNA expression levels. *Mol Biosyst*, vol. 5, no. 12, pp. 1512-1526.

Dillman, R.O. 2002. Radiolabeled anti-CD20 monoclonal antibodies for the treatment of B-cell lymphoma. *J Clin Oncol*, vol. 20, pp. 3545-3557.

Dittfeld, C., Dietrich, A. & Peickert, S., *et al.* 2009. CD133 expression is not selective for tumor-initiating or radioresistant cell populations in the CRC cell lines HCT-116. *Radiother Oncol*, vol. 92, no. 3, pp. 353-361.

Dranitsaris, G., Edwards, S. & Edwards, J., *et al.* 2010. Bevacizumab in combination with folfiri chemotherapy in patients with metastatic colorectal cancer: an assessment of safety and efficacy in the province of Newfoundland and Labrador. *Curr Oncol*, vol. 17, no. 5, pp. 12-16.

Du, L., Wang, H. & He, L., *et al*. 2008. CD44 is of functional importance for colorectal cancer stem cells. *Clin Cancer Res*, vol. 14, no. 21, pp. 6751-6760.

Duff, E.K. & Clarke, A.R. 1998. Smad4 (DPC4)-a potent tumour suppressor? *Br J Cancer*, vol. 78, no. 12, pp. 1615-1619.

Dukes, C. E. 1932. The classification of cancer of the rectum. *The Journal of Pathology and Bacteriology*, vol. 35, pp. 323-332.

Dupont, S., Mamidi, A. & Cordenonsi, M., *et al.* 2009. FAM/USP9x, a deubiquitinating enzyme essential for TGFbeta signaling, controls Smad4 monoubiquitination. *Cell*, vol. 136, no. 1, pp. 123-135.

Dyer, M.J. 1999. The role of CAMPATH-1 antibodies in the treatment of lymphoid malignancies. *Semin Oncol*, vol. 26, pp. 52-57.

Earle, J.S., Luthra, R. & Romans, A., *et al*. 2010. Association of microRNA expression with microsatellite instability status in colorectal adenocarcinoma. *J Mol Diagn*, vol. 12, pp. 433-440.

Easwaran, H., Tsai, H.C. & Baylin, S.B., *et al*. 2014. Cancer epigenetics: tumor heterogeneity, plasticity of stem-like states, and drug resistance. *Mol Cell*, vol. 54, no. 5, pp. 716-727.

Edwards, C. 1997. Physiology of the colorectal barrier. *Adv Drug Deliver Rev,* vol. 28, no. 2, pp. 173-190.

Egger, G., Liang, G. & Aparicio, A., *et al.* 2004. Epigenetics in human disease and prospects for epigenetic therapy. *Nature*, vol. 429, no. 6990, pp. 457-463.

Ellis, H. 2010. Anatomy of the caecum, appendix and colon. *Surgery*, vol. 29, no.1, pp. 1-4.

Elsaba, T.M.A, Martinez-Pomares, L. & Robins, A.R., *et al.* 2010. The stem cell marker CD133 associates with enhanced colony formation and cell motility in colorectal cancer. *PLoS One*, vol. 5, no. 5, article ID e10714.

Eramo, A., Lotti, F. & Sette, G., *et al.* 2008. Identification and expansion of the tumorigenic lung cancer stem cell population. *Cell Death Differ*, vol. 15, no. 3, pp. 504-514.

Evans, M.J. & Kaufman, M.H. 1981. Establishment in culture of pluripotential cells from mouse embryos. *Nature*, vol. 292, no. 5819, pp. 154-156.

Fabian, M.R., Sonenberg, N. & Filipowicz, W. 2010. Regulation of mRNA translation and stability by microRNAs. *Annu Rev Biochem*, vol. 79, pp. 351-379.

Falcone, A., Ricci, S. & Brunetti, I., *et al*. 2007. Phase III trial of infusional fluorouracil, leucovorin, oxaliplatin, and irinotecan (FOLFOXIRI) compared with infusional fluorouracil, leucovorin, and irinotecan (FOLFIRI) as first-line treatment for metastatic colorectal cancer: the Gruppo Oncologico Nord Ovest. *J Clin Oncol*, vol. 25, no. 13, pp. 1670-1676.

Farthing, C.R., Ficz, G. & Ng, R.K., *et al*. 2008. Global mapping of DNA methylation in mouse promoters reveals epigenetic reprogramming of pluripotency genes. *PLoS One*, doi: 10.1371/journal.pgen.1000116.

Fearon, E.R. 2011. Molecular genetics of colorectal cancer. *Annu Rev Pathol*, vol. 6, pp. 479-507.

Fearon, E.R. & Vogelstein, B. 1990. A genetic model for colorectal tumorigenesis. *Cell*, vol. 61, no. 5, pp. 759-767.

Felici, M.D. 2011. Nuclear reprogramming in mouse primordial germ cells: epigenetic contribution. *Stem Cells International*, vol. 2011, article ID 425863.

Ferlay, J., Bray, F. & Pisani, P., *et al.* 2004. *GLOBOCAN 2002: Cancer Incidence, Mortality and Prevalence Worldwide*. Lyon: International Agency for Research on Cancer. France: IARCPress. no. 5, version 2.

Ferretti, E., De Smaele, E. & Miele, E., *et al*. 2008. Concerted microRNA control of Hedgehog signalling in cerebellar neuronal progenitor and tumour cells. *EMBO*, vol. 27, no. 19, pp. 2616–2627.

Fishel, R., Lescoe, M.K. & Rao, M.R., *et al*. 1993. The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. *Cell*, vol. 75, no. 5, pp. 1027-1038.

Flicek, P., Amode, M.R. & Barrell, D., *et al.* 2014. Ensembl 2014. *Nucleic Acids Res*, vol. 42 (database issue), D749-55.

Florek, M., Haase, M. & Marzesco, A.M., *et al.* 2005. Prominin-1/CD133, a neural and hematopoietic stem cell marker, is expressed in adult human differentiated cells and certain types of kidney cancer. *Cell Tissue Res*, vol. 319, pp. 15-26.

Fodde, R. & Brabletz, T. 2007. Wnt/beta-catenin signaling in cancer stemness and malignant behavior. *Curr Opin Cell Biol*, vol. 19, no. 2, pp. 150-158.

Fogh, J. & Trempe, G. 1975. *Human Tumor Cells in vitro*. New York: Plenum Publishing Corp, pp. 115-141.

Fong, D., Steurer, M. & Obrist, P., *et al*. 2008. Ep-CAM expression in pancreatic and ampullary carcinomas: frequency and prognostic relevance. *J Clin Pathol*, vol. 61, no. 1, pp. 31-35.

Freberg, C.T., Dahl, J.A. & Timoskainen, S., *et al.* 2007. Epigenetic reprogramming of *OCT4* and *NANOG* regulatory regions by embryonal carcinoma cell extract. *Mol Biol Cell*, vol. 18, pp. 1543-1553.

Friedmann, I., Atmaca, A. & Chow, K.U., *et al.* 2006. Synergistic effects of Valproic acid and mitomycin C in adenocarcinoma cell lines and fresh tumor cells of patients with colon cancer. *J Chemother*, vol. 18, no. 4, pp. 415-420.

Friedrich, J., Seidel, C. & Ebner, R., *et al*. 2009. Spheroid-based drug screen: considerations and practical approach. *Nat. Protoc*, vol. 4, no. 3, pp. 309-324.

Fukui, K. 2010. DNA Mismatch Repair in Eukaryotes and Bacteria. *J Nucleic Acids*, vol. 2010, article ID 260512.

Garcia, V.M., Batlle, J.F. & Casado, E., *et al*. 2011. Immunohistochemical analysis of tumour regression grade for rectal cancer after neoadjuvant chemoradiotherapy. *Colorectal Dis*, vol. 13, pp. 989-998.

Giantonio, B.J., Catalano, P.J. & Meropol, N.J., *et al.* 2007. Bevacizumab in combination with oxaliplatin, fluorouracil, and leucovorin (FOLFOX4) for previously treated metastatic colorectal cancer: results from the Eastern Cooperative Oncology Group Study E3200. *J Clin Oncol*, vol. 25, pp. 1539-1544.

Goel, A. & Boland, C.R. 2012. Epigenetics of colorectal cancer. *Gastroenterology*, vol. 443, no. 6, pp. 1442-1460.

Goldberg, R.M., Sargent, D.J. & Morton, R.F., *et al.* 2004. A randomized controlled trial of fluorouracil plus leucovorin, irinotecan, and oxaliplatin combinations in patients with previously untreated metastatic colorectal cancer. *J Clin Oncol*, vol. 22, no. 1, pp. 23-30.

Gosens, M.J., van Kempen, L.C. & van de Velde, C.J., *et al.* 2007. Loss of membranous Ep-CAM in budding colorectal carcinoma cells. *Mod Pathol*, vol. 20, no. 2, pp. 221-232.

Goss, K. H. & Groden, J. 2000. Biology of the adenomatous polyposis coli tumor suppressor. *J Clin Oncol*, vol. 18, pp. 1967-1979.

Göttlicher, M., Minucci, S. & Zhu, P., *et al.* 2001. Valproic acid defines a novel class of HDAC inhibitors inducing differentiation of transformed cells. *EMBO J*, vol. 20, pp. 6969-6978.

Grady, W.M. & Carethers, J.M. 2008. Genomic and epigenetic instability in colorectal cancer pathogenesis. *Gastroenterology*, vol. 135, pp. 1079-1099.

Graff, J.R., Herman, J.G. & Myöhänen, S., *et al.* 1997. Mapping patterns of CpG island methylation in normal and neoplastic cells implicates both upstream and downstream regions in de novo methylation. *J Biol Chem*, vol. 272, no. 35, pp. 22322-22329.

Grosse-Gehling, P., Fargeas, C.A. & Dittfeld, C., *et al*. 2013. CD133 as a biomarker for putative cancer stem cells in solid tumours: limitations, problems and challenges. *J Pathol*, vol. 229, no. 3, pp. 355-378.

Gudas, L.J. & Wagner, J.A. 2011. Retinoids regulate stem cell differentiation. *J Cell Physiol*, vol. 226, no. 2, pp. 322-330.

Guo, Y., Liu, S. & Wang, P., *et al.* 2011. Expression profile of embryonic stem cell-associated genes Oct4, Sox2 and Nanog in human gliomas. *Histopathology*, vol. 59, no. 4, pp. 763-775.

Guo, Y., Mantel, C. & Hromas, R.A., *et al.* 2008. Oct 4 is critical for survival/antiapoptosis of murine embryonic stem cells subjected to stress: effects associated with STAT3/Survivin. *Stem Cells*, vol. 26, no. 1, article ID 30.

Gurvich, N., Tsygankova, O.M. & Meinkoth, J.L., *et al*. 2004. Histone deacetylase is a target of Valproic acid-mediated cellular differentiation. *Cancer Res*, vol. 64, no. 3, pp.1079-1086.

Hammer, Ø., Harper, D.A.T. & Ryan, P.D. 2001. PAST: Paleontological statistics software package for education and data analysis. *Palaeontologia Electronica*, vol. 4, no. 1, pp. 1-9.

Hammerman, P.S., Fox, C.J. & Birnbaum, M.J., *et al*. 2005. Pim and Akt oncogenes are independent regulators of hematopoietic cell growth and survival. *Blood*, vol. 105, no. 11, pp. 4477-4483.

Han, X., Fang, X. & Lou, X. 2012. Silencing SOX2 induced mesenchymal-epithelial transition and its expression predicts liver and lymph node metastasis of CRC patients. *PLoS One*, vol. 7, no. 8, doi: 10.1371/journal.pone.0041335.

Hanahan, D. & Weinberg, R.A. 2000. The hallmarks of cancer. *Cell*, vol. 100, no. 1, pp. 57-70.

Hao, J., Li, T-G. & Qi, X., *et al*. 2006. WNT/β-catenin pathway up-regulates Stat3 and converges on LIF to prevent differentiation of mouse embryonic stem cells. *Dev Biol*, vol. 290, pp. 81-91.

Haraguchi, N., Ohkuma, M. & Sakashita, H., *et al*. 2008. CD133+CD44+ population efficiently enriches colon cancer initiating cells. *Ann Surg Oncol*, vol. 15, no. 10, pp. 2927-2933.

Hart, A. 2001. Mann-Whitney test is not just a test of medians: differences in spread can be important. *BMJ*, vol. 323, no. 7309, pp. 391-393.

Hay, D.C., Sutherland, L. & Clark, J., *et al.* 2004. Oct-4 knockdown induces similar patterns of endoderm and trophoblast differentiation markers in human and mouse embryonic stem cells. *Stem Cells*, vol. 22, no. 2, pp. 225-235.

Heldin, C.H., Miyazono, K. & ten Dijke, P. 1997. TGF-beta signalling from cell membrane to nucleus through SMAD proteins. *Nature*, vol. 390, no. 6659, pp. 465-71.

Hellemans, J., Mortier, G. & De Paepe, A., *et al*. 2007. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biol*, vol. 8, no.2, R19.

Hermansen, S.K., Christensen, K.G. & Jensen, S.S., *et al*. 2011. Inconsistent immunohistochemical expression patterns of four different CD133 antibody clones in glioblastoma. *J Histochem Cytochem*, vol. 59, no. 4, pp. 391-407.

Horst, D., Kriegl, L. & Engel, J., *et al.* 2008. CD133 expression is an independent prognostic marker for low survival in colorectal cancer. *Br J Cancer*, vol. 99, no. 8, pp. 1285-1289.

Hsu, C-S., Tung, C-Y. & Yang, C-Y., *et al*. 2013. Response to stress in early tumor colonization modulates switching of CD133-positive and CD133-negative subpopulations in a human metastatic colon cancer cell line, SW620. *PLoS One*, vol. 8, no. 4, doi: 10.1371/journal.pone.0061133.0.

Huang, E.H., Hynes, M.J. & Zhang, T., *et al.* 2009. Aldehyde dehydrogenase 1 is a marker for normal and malignant human colonic stem cells (SC) and tracks SC population during colon tumorigenesis. *Cancer Res*, vol. 69, no. 8, pp. 3382-3389.

Huang, Q., Gumireddy, K. & Schrier, M., *et al.* 2008. The microRNAs miR-373 and miR-520c promote tumour invasion and metastasis. *Nat Cell Biol*, vol. 10, pp. 202-210.

Hurwitz, H., Fehrenbacher, L. & Novotny, W., *et al*. 2004. Bevacizumab plus irinotecan, fluorouracil, and leucovorin for metastatic colorectal cancer. *N Engl J Med*, vol. 350, pp. 2335-2342.

Hyslop, L., Stojkovic, M. & Armstrong, L., *et al*. 2005. Downregulation of NANOG induces differentiation of human embryonic stem cells to extraembryonic lineages. *Stem Cells*, vol. 23, no. 8, pp. 1035-1043.

Ibrahim, E.E., Babaei-Jadidi, R. & Saadeddin, A., *et al*. 2012. Embryonic NANOG activity defines colorectal cancer stem cells and modulates through AP1- and TCF-dependent mechanisms. *Stem Cells*, vol. 30, no. 10, pp. 2076-2087.

Ieta, K., Tanaka, F. & Haraguchi, N., *et al.* 2008. Biological and genetic characteristics of tumor-initiating cells in colon cancer. *Ann Surg Oncol*, vol. 15, pp. 638-648.

Irollo, E. & Pirozzi, G. 2013. CD133: to be or not to be, is this the real question? *Am J Transl Res*, vol. 5, no. 6, pp. 563–581.

Ishiguro, T., Sato, A. & Ohata, H., *et al.* 2012. Differential expression of Nanog1 and Nanogp8 in colon cancer cells. *Biochem Biophys Res Commun*, vol. 418, no. 2, pp. 199-204.

Issa, J-P. 2004. CpG island methylator phenotype in cancer. *Nat Rev Cancer*, vol. 4, pp. 988-993.

James, D., Levine, A.J. & Besser, D., *et al*. 2005. TGFbeta/activin/nodal signaling is necessary for the maintenance of pluripotency in human embryonic stem cells. *Development*, vol. 132, no. 6, pp. 1273-1282.

Jemal, A., Siegel, R. & Ward, E., *et al.* 2009. Cancer statistics, 2009. *CA Cancer J Clin*, vol. 59, pp. 225-249.

Jeong, C.H., Cho, Y.Y. & Kim, M.O., *et al*. 2010. Phosphorylation of Sox2 cooperates in reprogramming to pluripotent stem cells. *Stem Cells*, vol. 28, no. 12, pp. 2141-2150.

Johnson, E.S. 2004. Protein modification by SUMO. *Annu Rev Biochem*, vol. 73, pp. 355-382.

Jones, S., Zhang, X. & Parsons, D.W., *et al.* 2008. Core signalling pathways in human pancreatic cancers revealed by global genomic analyses. *Science*, vol. 32, pp. 1801-1806.

Kallas, A., Pook, M. & Trei, A., *et al*. 2014. SOX2 is regulated differently from NANOG and OCT4 in human embryonic stem cells during early differentiation initiated with sodium butyrate. *Stem Cells International*, vol. 2014, article ID 298163.

Karagiannis, T.C., Kn, H. & El-Osta, A. 2006. The epigenetic modifier, Valproic acid, enhances radiation sensitivity. *Epigenetics*, vol. 1, pp. 131-137.

Katoh, M. & Katoh, M. 2007. Notch signaling in gastrointestinal tract (review). *Int J Oncol*, vol. 30, no. 1, pp. 247-251.

Kawagoe, R., Kawagoe, H. & Sano, K. 2002. Valproic acid induces apoptosis in human leukemia cells by stimulating both caspase-dependent and -independent apoptotic signalling pathways. *Leuk Res*, vol. 26, no. 5, pp. 495-502.

Kawamoto, H., Yuasa, T. & Kubota, Y., *et al.* 2010. Characteristics of CD133(+) human colon cancer SW620 cells. *Cell Transplant*, vol. 19, pp. 857–864.

Kemper, K., Sprick, M.R. & de Bree, M., *et al*. 2010. The AC133 epitope, but not the CD133 protein, is lost upon cancer stem cell differentiation. *Cancer Res*, vol. 70, no. 2, pp. 719-729.

Kim, C.H., Marquez, V.E. & Mao, D.T., *et al*. 1986. Synthesis of pyrimidin-2-one nucleosides as acid-stable inhibitors of cytidine deaminase. *J Med Chem*, vol. 29, pp. 1374-1380.

Kim, C.S., Vasko, V.V. & Kato, Y., *et al*. 2005. AKT activation promotes metastasis in a mouse model of follicular thyroid carcinoma. *Endocrinology*, vol. 146, no. 10, pp. 4456-4463.

Kim, J., Woo, A.J. & Chu, J., *et al*. 2010. A myc network accounts for similarities between embryonic stem and cancer cell transcription programs. *Cell*, vol. 143, no. 2, pp. 313-324.

Klonisch, T., Wiechec, E. & Hombach-Klonisch, S., *et al.* 2008. Cancer stem cell markers in common cancers - therapeutic implications. *Trends Mol Med*, vol. 14, no. 10, pp. 450-460.

Kodach, L.L., Wiercinska, E. & de Miranda, N.F., *et al*. 2008. The bone morphogenetic protein pathway is inactivated in the majority of sporadic colorectal cancers. *Gastroenterology*, vol. 134, no. 5, pp. 1332-1341.

Kojima, M., Ishii, G. & Atsumi, N., *et al*. 2008. Immunohistochemical detection of CD133 expression in colorectal cancer: a clinicopathological study. *Cancer Sci*, vol. 99, pp. 1578-1583.

Kopp, J.L., Ormsbee, B.D. & Desler, M., *et al.* 2008. Small increases in the level of Sox2 trigger the differentiation of mouse embryonic stem cells. *Stem Cell*, vol. 24, no. 4, pp. 903-911.

Kornberg, R.D. 1974. Chromatin structure: a repeating unit of histones and DNA. *Science*, vol. 184, no. 4139, pp. 868-871.

Kornberg, R.D. & Lorch, Y. 1999. Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. *Cell*, vol. 98, no. 3, pp. 285-294.

Krausova, M. & Korinek, V. 2012. Signal transduction pathways participating in homeostasis and malignant transformation of the intestinal tissue. *Neoplasma*, vol. 59, no. 6, pp. 708-718.

Kuendgen, A. & Gattermann, N. 2007. Valproic acid for the treatment of myeloid malignancies. *Cancer*, vol. 110, no. 5, pp. 943-954.

Kuzmichev, A.N., Kim, S.K. & D'Alessio, A.C., *et al.* 2012. Sox2 acts through Sox21 to regulate transcription in pluripotent and differentiated cells. *Curr Biol*, vol. 22, no. 18, pp. 1705-1710.

Laliberté, J., Marquez, V.E. & Momparler, R.L. 1992. Potent inhibitors for the deamination of cytosine arabinoside and 5-aza-2'-deoxycytidine by human cytidine deaminase. *Cancer Chemother Pharmacol*, vol. 30, no. 1, pp. 7-11.

Lane, D.P. 1992. Cancer. p53, guardian of the genome. *Nature*, vol. 358, no. 6381, pp. 15-16.

Langan, R.C., Mullinax, J.E. & Ray, S., *et al*. 2012. A pilot study assessing the potential role of non-CD133 colorectal cancer stem cells as biomarkers. *J Cancer*, vol. 3, pp. 231-240.

Lapidot, T., Sirard, C. & Vormoor, J., *et al.* 1994. A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. *Nature*, vol. 367, no. 6464, pp. 645-648.

Leach, F.S., Nicolaides, N.C. & Papadopoulos, N., *et al*. 1993. Mutations of a mutS homolog in hereditary nonpolyposis colorectal cancer. *Cell*, vol. 75, no. 6, pp. 1215-1225.

Lee, J., Kim, H.K. & Rho, J.Y., *et al.* 2006. The human Oct-4 isoforms differ in their ability to confer self-renewal. *J Biol Chem*, vol. 281, no. 44, pp. 33554-33565.

Leng, Z., Tao, K. & Xia, Q., *et al*. 2013. Krüppel-like factor 4 acts as an oncogene in colon cancer stem cell-enriched spheroid cells. *PLoS One*, vol. 8, no. 2, doi: 10.1371/journal.pone.0056082.

Leslie, A., Carey, F.A. & Pratt, N.R., *et al*. 2002. The colorectal adenoma-carcinoma sequence. *Br J Surg*, vol. 89, no. 7, pp. 845-860.

Levin, T.G., Powell, A.E. & Davies, P.S., *et al*. 2010. Characterization of the intestinal cancer stem cell marker CD166 in the human and mouse gastrointestinal tract. *Gastroenterology*, vol. 139, no. 6, pp. 2072-2082.

Li, C., Heidt, D.G. & Dalerba, P., *et al*. 2007. Identification of pancreatic cancer stem cells. *Cancer Res*, vol. 67, no. 3, pp. 1030-1037.

Li, C.Y., Li, B.X. & Liang, Y., *et al*. 2009. Higher percentage of CD133+ cells is associated with poor prognosis in colon carcinoma patients with stage IIIB. *J Transl Med*, vol. 7, no. 56, doi: 10.1186/1479-5876-7-56.

Li, H., Zhao, P. & Lu, Y., *et al*. 2012b. Correlation of aberrant expression of CD133 with FHIT and malignant phenotype of colorectal adenocarcinoma. *Int J Colorectal Dis*, vol 27, pp. 1015-1020.

Li, J., Yang, G.Z. & Zhu, Z.M., *et al*. 2012a. Osteopontin is overexpressed in colorectal carcinoma and is correlated with P53 by immunohistochemistry. *Exp Ther Med*, vol. 3, no. 4, pp. 621-624.

Lin, C.Y. 2014. Differentiation Therapy with Cancer Stem Cells. *J Stem Cell Res Ther*, vol. 4, issue. 1, article ID 1000e117.

Lin, S.C., Wani, M.A. & Whitsett, J.A., *et al.* 2010. Klf5 regulates lineage formation in the pre-implantation mouse embryo. *Development*, vol. 137, no. 23, pp. 3953-3963.

Litvinov, S.V., Balzar, M. & Winter, M. J., *et al.* 1997. Epithelial adhesion molecule Ep-CAM modulates cell–cell interactions mediated by classic cadherins. *J Cell Biol*, vol. 139, no. 55, pp. 1337-1348.

Liu, A., Yu, X. & Liu, S. 2013. Pluripotency transcription factors and cancer stem cells: small genes make a big difference. *Chin J Cancer*, vol. 32, no. 9, pp. 483–487.

Livak, K.J. & Schmittgen, T.D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. *Methods*, vol. 25, no. 4, pp. 402-408.

Lombardo, Y., Scopelliti, A. & Cammareri, P., *et al.* 2011. Bone morphogenetic protein 4 induces differentiation of colorectal cancer stem cells and increases their response to chemotherapy in mice. *Gastroenterology*, vol. 140, no. 1, pp. 297-309.

Löscher, W. 2002. Basic pharmacology of Valproate: a review after 35 years of clinical use for the treatment of epilepsy. *CNS Drugs*, vol. 16, no. 10, pp. 669-694.

Luger, K. & Richmond, T.J. 1998. The histone tails of the nucleosome. *Curr Opin Genet Dev*, vol. 8, no. 2, pp. 140-146.

Lugli, A., Iezzi, G. & Hostettler, I., *et al.* 2010. Prognostic impact of the expression of putative cancer stem cell markers CD133, CD166, CD144s, EpCAM and ALDH1 in colorectal cancer. *Br J Cancer*, vol. 103, no. 3, pp. 382-390.

Lujambio, A., Ropero, S. & Ballestar, E., *et al.* 2007. Genetic unmasking of an epigenetically silenced microRNA in human cancer cells. *Cancer Res*, vol. 67, pp. 1424-1429.

Madison, B.B., Braunstein, K. & Kuizon, E., *et al*. 2005. Epithelial hedgehog signals pattern the intestinal crypt-villus axis. *Development*, vol. 132, no. 2, pp. 279-289.

Maeda, S., Shinchi, H. & Kurahara, H., *et al*. 2008. CD133 expression is correlated with lymph node metastasis and vascular endothelial growth factor-C expression in pancreatic cancer. *Br J Cancer*, vol. 98, no. 8, pp. 1389‑1397.

Maetzel, D., Denzel, S. & Mack, B., *et al*. 2009. Nuclear signalling by tumour-associated antigen EpCAM. *Nat Cell Biol*, vol. 11, no. 2, pp. 162-171.

Maier, S., Wilbertz, T. & Braun, M., *et al*. 2011. SOX2 amplification is a common event in squamous cell carcinomas of different organ sites. *Hum Pathol*, vol. 42, pp. 1078-1088.

Maier, T., Güell, M. & Serrano, L. 2009. Correlation of mRNA and protein in complex biological samples. *FEBS Lett*, vol. 583, pp. 3966–3973.

Mak, A.B., Nixon, A.M. & Kittanakom, S., *et al*. 2012. Regulation of CD133 by HDAC6 promotes β-catenin signaling to suppress cancer cell differentiation. *Cell Rep*, vol. 2, no. 4, 951-963.

Mariadason, J.M. 2008. HDACs and HDAC inhibitors in colon cancer. *Epigenetics*, vol. 3, no. 1, pp. 28-37.

Markowitz, S.D. & Bertagnolli, M.M. 2009. Molecular origins of cancer: molecular basis of colorectal cancer. *N Engl J Med*, vol. 361, no. 25, pp. 2449-2460.

Marquardt, J.U., Raggi, C. & Andersen, J.B., *et al.* 2011. Human hepatic cancer stem cells are characterized by common stemness traits and diverse oncogenic pathways. *Hepatology*, vol. 54, no. 3, pp. 1031-1042.

Marra, G & Boland, C.R. 1996. DNA repair and colorectal cancer. *Gastroenterol Clin North Am*, vol. 25, no. 4, pp. 755-772.

Marshman, E., Booth, C. & Potten, C.S. 2002. The intestinal epithelial stem cell. *Bioessays*, vol. 24, no. 1, pp. 91-98.

Martin, G. R. 1981. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc Natl Acad Sci USA*, vol. 78, no. 12, pp. 7634-7638.

Massagué, J., Seoane, J. & Wotton, D. 2005. Smad transcription factors. *Genes Dev*, vol. 19, no. 23, pp. 2783-2810.

Masui, S., Nakatake, Y. & Toyooka, Y., *et al*. 2007. Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. *Nat. Cell Biol*, vol. 9, pp. 625-635.

Mather, J.P. 2008. *Methods in Cell Biology*, vol. 86. California: Raven Biotechnologies, Inc. 317 p.

Matin, M.M., Walsh, J.R. & Gokhale, P.J., *et al.* 2004. Specific knockdown of Oct4 and Beta2-microglobulin expression by RNA interference in human embryonic stem cells and embryonic carcinoma cells. *Stem Cells*, vol. 22, pp. 659-668.

McCleary-Wheeler, A.L., Lomberk, G.A. & Weiss, F.U., *et al.* 2013. Insights into the epigenetic mechanisms controlling pancreatic carcinogenesis. *Cancer Lett*, vol. 328, no. 2, pp. 212-221.

McLaughlin, P., Grillo-Lopez, A.J. & Link, B.K., *et al.* 1998. Rituximab chimeric anti-CD20 monoclonal antibody therapy for relapsed indolent lymphoma: half of patients respond to a four-dose treatment program. *J Clin Oncol*, vol. 16, pp. 2825-2833.

McLean, M.J. & Macdonald, R.L. 1986. Sodium Valproate, but not ethosuximide, produces use- and voltage-dependent limitation of high frequency repetitive firing of action potentials of mouse central neurons in cell culture. *J Pharmacol Exp Ther*, vol. 237, no. 3, pp. 1001-1011.

Meacham, C.E. & Morrison, S.J. 2013. Tumour heterogeneity and cancer cell plasticity. *Nature*, vol. 501, no. 7467, pp. 328-337.

Medema, J.P. 2013. Cancer stem cells: the challenges ahead. *Nat Cell Biol*, vol. 15, no. 4, pp. 338-344.

Medvedev, S.P., Shevchenko, A.I. & Mazurok, N.A., *et al.* 2008. Oct4 and Nanog are the key genes in the system of pluripotency maintenance in mammalian cells. *Genetika*, vol. 44, no. 12, pp. 1589-1608.

Meissner, A. 2010. Epigenetic modifications in pluripotent and differentiated cells. *Nat Biotechnol*, vol. 28, no. 10, pp. 1079-1088.

Mélin, C., Perraud, A. & Akil, H., *et al*. 2012. Cancer stem cell sorting from colorectal cancer cell lines by sedimentation field flow fractionation. *Anal Chem*, vol. 84, no. 3, pp. 1549-1556.

Michaelis, M., Michaelis, U.R. & Fleming, I., *et al.* 2004. Valproic acid inhibits angiogenesis in vitro and in vivo. *Mol Pharmacol*, vol. 65, no. 3, pp. 520-527.

Migheli, F. & Migliore, L. 2012. Epigenetics of colorectal cancer. *Clin Genet*, vol. 81, no. 4, pp. 312-318.

Miraglia, S., Godfrey, W. & Yin, A.H., *et al*. 1997. A novel five-transmembrane hematopoietic stem cell antigen: isolation, characterization, and molecular cloning. *Blood*, vol. 90, no. 12, pp. 5013-5021.

Mitsui, K., Tokuzawa, Y. & Itoh, H., *et al*. 2003. The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell*, vol. 113, no. 5, pp. 631-642.

Miyaki, M., Iijima, T. & Konishi, M., *et al*. 1999. Higher frequency of Smad4 gene mutation in human colorectal cancer with distant metastasis. *Oncogene*, vol. 18, no. 20, pp. 3098-3103.

Mizrak, D., Brittan, M. & Alison, M.R. 2008. CD133: molecule of the moment. *J Pathol*, vol. 214, pp. 3-9.

Mizuno, H., Spike, B.T. & Wahl, G.M., *et al.* 2010. Inactivation of p53 in breast cancers correlates with stem cell transcriptional signatures. *Proc Natl Acad Sci USA*, vol. 107, pp. 22745-22750.

Momburg, F., Moldenhauer, G. & Hammerling, G. J., *et al.* 1987. Immunohistochemical study of the expression of a Mr 34,000 human epithelium-specific surface glycoprotein in normal and malignant tissues. *Cancer Res*, vol. 47, pp. 2883-2891.

Morin, P.J., Sparks, A.B. & Korinek, V., *et al*. 1997. Activation of beta-catenin-Tcf signaling in colon cancer by mutations in beta-catenin or APC. *Science*, vol. 275, no. 5307, pp. 1787-1790.

Motoyama, K., Inoue, H. & Takatsuno, Y., *et al*. 2009. Over- and under-expressed microRNAs in human colorectal cancer. *Int J Oncol*, vol. 34, pp. 1069-1075.

Muñoz, P., Iliou, M.S. & Esteller, M. 2012. Epigenetic alterations involved in cancer stem cell reprogramming. *Mol Oncol*, vol. 6, no. 6, pp. 620-636.

Munz, M., Baeuerle, P.A. & Gires, O. 2009. The emerging role of EpCAM in cancer and stem cell signalling. *Cancer Res*, vol. 69, no. 14, pp. 5627-5629.

Muraro, M.G., Mele, V. & Däster, S., *et al*. 2012. CD133+, CD166+CD44+, and CD24+CD44+ phenotypes fail to reliably identify cell populations with cancer stem cell functional features in established human colorectal cancer cell lines. *Stem Cells Transl Med*, vol. 1, no. 8, pp. 592-603.

Nakazawa, T., Kondo, T. & Ma, D., *et al.* 2012. Global histone modification of histone H3 in colorectal cancer and its precursor lesions. *Hum Pathol*, vol. 43, no. 6, pp. 834-842.

Nandan, M.O. & Yang, V.W. 2011. An update on the biology of RAS/RAF mutations in colorectal cancer. *Curr Colorectal Cancer Rep*, vol. 7, no. 2, pp. 113-120.

Naor, D., Sionov, R.V. & Ish-Shalom, D. 1997. CD44: structure, function, and association with the malignant process. *Adv Cancer Res*, vol. 71, pp. 241-319.

Neumann, J., Bahr, F. & Horst, D., *et al*. 2011. SOX2 expression correlates with lymph-node metastases and distant spread in right-sided colon cancer. *BMC Cancer*, vol. 11, no. 518, doi: 10.1186/1471-2407-11-518.

Niwa, H., Miyazaki, J. & Smith, A.G. 2000. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat Genet*, vol. 24, no. 4, pp. 372-376.

Niwa, H., Toyooka, Y. & Shimosato, D., *et al*. 2005. Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation. *Cell*, vol. 123, no. 5, pp. 917-929.

Oberg, A.L., French, A.J. & Sarver, A.L., *et al*. 2011. miRNA expression in colon polyps provides evidence for a multihit model of colon cancer. *PLoS One*, vol. 6, no. 6, article ID e20465.

O'Brien, C.A., Pollett, A. & Gallinger, S., *et al.* 2007. A human colon cancer cell capable of initiating tumour growth in immunodeficient mice. *Nature*, vol. 445, no. 7123, pp. 106-110.

Oren, M. & Rotter, V. 2010. Mutant p53 gain-of-function in cancer. *Cold Spring Harb Perspect Biol*, vol. 2, no. 2, article ID a001107.

Orkin, S.H. & Hochedlinger, K. 2011. Chromatin connections to pluripotency and cellular reprogramming. *Cell*, vol. 145, no. 6, pp. 835-850.

Ormsbee, G.B.D., Wuebben, E.L. & Rizzino, A. 2013. Sox2 expression is regulated by a negative feedback loop in embryonic stem cells that involves AKTsignaling and FoxO1. *PLoS One*, vol. 8, no. 10, doi: 10.1371/journal.pone.0076345.

Otsubo, T., Akiyama, Y. & Yanagihara, K. 2008. SOX2 is frequently downregulated in gastric cancers and inhibits cell growth through cell-cycle arrest and apoptosis. *Br J Cancer*, vol. 98, no. 4, pp. 824-831.

Owens, M.J. & Nemeroff, C.B. 2003. Pharmacology of Valproate. *Psychopharmacol Bull*, vol. 37, suppl. 2, pp. 17-24.

Pacold, M.E., Suire, S. & Perisic, O., *et al.* 2000. Crystal structure and functional analysis of Ras binding to its effector phosphoinositide 3-kinase gamma. *Cell*, vol. 103, no. 6, pp. 931-943.

Pajonk, F., Vlashi, E. & McBride, W.H. 2010. Radiation resistance of cancer stem cells: the 4 R's of radiobiology revisited. *Stem Cells*, vol. 28, no. 4, pp. 639-648.

Paldino, E., Tesori, V. & Casalbore, P., *et al.* 2014. Tumor initiating cells and chemoresistance: which is the best strategy to target colon cancer stem cells? *Biomed Res Int*, article ID 859871.

Pampaloni, F., Reynaud, E.G. & Stelzer, E.H. 2007. The third dimension bridges the gap between cell culture and live tissue. *Nat Rev Mol Cell Biol*, vol. 8, no. 10, pp. 839-845.

Papadopoulos, N., Nicolaides, N.C. & Wei, Y.F., *et al*. 1994. Mutation of a mutL homolog in hereditary colon cancer. *Science*, vol. 263, no. 5153, pp. 1625-1629.

Pappas, P.A. & DePuy, V. 2004. *An overview of non-parametric tests in SAS: when, why, and how*. Duke Clinical Research Institute, Durham, pp. 1-5.

Parsons, R., Myeroff, L.L. & Liu, B., *et al*. 1995. Microsatellite instability and mutations of the transforming growth factor beta type II receptor gene in colorectal cancer. *Cancer Res*, vol. 55, no. 23, pp. 5548-5550.

Patriarca, C., Macchi, R.M. & Marschner, A.K., *et al*. 2012. Epithelial cell adhesion molecule expression (CD326) in cancer: a short review. *Cancer Treat Rev*, vol. 38, no. 1, pp. 68-75.

Peedicayil, J. 2006. Epigenetic therapy: a new development in pharmacology. *Indian J Med Res*, vol. 123, no. 1, pp. 17-24.

Peltomäki, P. 2012. Mutations and epimutations in the origin of cancer. *Exp Cell Res*, vol. 318, no. 4, pp. 299-310.

Phiel, C.J., Zhang, F. & Huang, E.Y., *et al.* 2001. Histone deacetylase is a direct target of Valproic acid, a potent anticonvulsant, mood stabilizer, and teratogen. *J Biol Chem*, vol. 39, pp. 36734-36741.

Phillips, A., Bullock, T. & Plant, N. 2003. Sodium Valproate induces apoptosis in the rat hepatoma cell line, FaO. *Toxicology*, vol. 192, no. 2-3, pp. 219-227.

Piazza, T., Cha, E. & Bongarzone, I., *et al*. 2005. Internalization and recycling of ALCAM/CD166 detected by a fully human single-chain recombinant antibody. *J Cell Sci*, vol. 118, pp. 1515-1525.

Platet, N. Liu, S.Y. & Atifi, M.E., *et al*. 2007. Influence of oxygen tension on CD133 phenotype in human glioma cell cultures. *Cancer Lett*, vol. 258, pp. 286–290.

Ponti, D., Costa, A. & Zaffaroni, N., *et al.* 2005. Isolation and *in vitro* propagation of tumorigenic breast cancer cells with stem/progenitor cell properties. *Cancer Res*, vol. 65, no. 13, pp. 5506–5511.

Potten, C.S., Booth, C. & Pritchard, D.M. 1997. The intestinal epithelial stem cell: the mucosal governor. *Int J Exp Pathol*, vol. 78, pp. 219-243.

Pruitt, K.D., Brown, G.R. & Hiatt, S.M., *et al.* 2013. RefSeq: an update on mammalian reference sequences. *Nucleic Acids Res*, 42 (database issue), D756-763.

Puglisi, M.A., Tesori, V. & Lattanzi, W., *et al*. 2013. Colon cancer stem cells: Controversies and perspectives. *World J Gastroenterol*, vol. 19, no. 20, pp. 2997–3006.

Rane, S., Sayed, D. & Abdellatif, M. 2007. MicroRNA with a MacroFunction. *Cell Cycle*, vol. 6, no. 15, pp. 1850-1855.

Rao, A.N. & Wang, J-Y. 2010. *Regulation of Gastrointestinal Mucosal Growth.* San Rafael (CA): Morgan & Claypool Life Sciences. Bookshelf ID: NBK54095.

Rao, G., Wang, H. & Li, B., *et al*. 2012. Reciprocal interactions between tumor-associated macrophages and CD44-positive cancer cells via osteopontin/CD44 promote tumorigenicity in colorectal cancer. *Clin Cancer Res*, vol. 19, no. 4, pp. 785-797.

Rashedi, I., Panigrahi, S. & Ezzati, P., *et al*. 2007. Autoimmunity and apoptosis-therapeutic implications. *Curr Med Chem*, vol. 14, pp. 3139–3159.

Rasheed, Z.A., Yang, J. & Wang, Q., *et al*. 2010. Prognostic significance of tumorigenic cells with mesenchymal features in pancreatic adenocarcinoma. *J Natl Cancer Inst*, vol. 102, no. 5, pp. 340-351.

Reggiani, B.L., Migaldi, M. & Caredda, E., *et al*. 2012. Increased expression of CD133 is a strong predictor of poor outcome in stage I colorectal cancer patients. *Scand J Gastroenterol*, vol. 47, pp. 1211-1217.

Ren, F., Sheng, W.Q. & Du, X. 2013. CD133: a cancer stem cells marker, is used in colorectal cancers. *World J Gastroenterol*, vol. 19, no. 17, pp. 2603-2611.

Retting, K.N., Song, B. & Yoon, B.S., *et al*. 2009. BMP canonical Smad signaling through Smad1 and Smad5 is required for endochondral bone formation. *Development*, vol. 136, no. 7, pp. 1093-1104.

Reya, T. & Clevers, H. 2005. Wnt signalling in stem cells and cancer. *Nature*, vol. 434, no. 7035, pp. 843-850.

Reya, T., Duncan, A.W. & Ailles, L., *et al*. 2003. A role for Wnt signalling in self-renewal of haematopoietic stem cells. *Nature*, vol. 423, no. 6938, pp. 409-414.

Reya, T., Morrison, S.J. & Clarke, M.F., *et al*. 2001. Stem cells, cancer, and cancer stem cells. *Nature*, vol. 414, pp. 105-111.

Ricci-Vitiani, L., Lombardi, D.G. & Pilozzi, E., *et al*. 2007. Identification and expansion of human colon-cancer-initiating cells. *Nature*, vol. 445, no. 7123, pp. 111-115.

Robertson, K.D. & Jones, P.A. 2000. DNA methylation: past, present and future directions. *Carcinogenesis*, vol. 21, pp. 461-467.

Rodda, D.J., Chew, J.L. & Lim, L.H., *et* al. 2005. Transcriptional regulation of NANOG by OCT4 and SOX2. *J Biol Chem*, vol. 280, no. 26, pp. 24731-24737.

Rodriguez, R.T., Velkey, J.M. & Lutzko, C., *et al.* 2007. Manipulation of OCT4 levels in human embryonic stem cells results in induction of differential cell types. *Exp Biol Med (Maywood)*, vol. 232, no. 10, pp. 1368-1380.

Roy, S. & Majumdar, A. 2012. Signaling in colon cancer stem cells. *J Mol Signal*, vol. 7, no. 11, doi: 10.1186/1750-2187-7-11.

Sahlberg, S., Spiegelberg, D. & Glimelius, B., *et al*. 2014. Evaluation of cancer stem cell markers CD133, CD44, CD24: association with AKT isoforms and radiation resistance in colon cancer cells. *PLoS One*, doi: 10.1371/journal.pone.0094621.

Saltz, L.B., Clarke, S. & Díaz-Rubio, E., *et al.* 2008. Bevacizumab in combination with oxaliplatin-based chemotherapy as first-line therapy in metastatic colorectal cancer: a randomized phase III study. *J Clin Oncol*, vol. 26, pp. 2013-2019.

Sampieri, K. & Fodde, R. 2012. Cancer stem cells and metastasis. *Semin Cancer Biol*, vol. 22, no. 3, pp. 187-193.

Sarver, A.L., French, A.J. & Borralho, P.M., *et al.* 2009. Human colon cancer profiles show differential microRNA expression depending on mismatch repair status and are characteristic of undifferentiated proliferative states. *BMC Cancer*, doi: 10.1186/1471-2407-9-401.

Schnekenburger, M. & Diederich, M. 2012. Epigenetics offer new horizons for colorectal cancer prevention. *Curr Colorectal Cancer Rep*, vol. 8, no. 1, pp. 66-81.

Schnell, U., Cirulli, V. & Giepmans, B.N. 2013. EpCAM: structure and function in health and disease. *Biochim Biophys Acta*, vol. 1828, no. 8, pp. 1989-2001.

Schoenhals, M., Kassambara, A. & De Vos, J., *et al*. 2009. Embryonic stem cell markers expression in cancers. *Biochem Biophys Res Commun*, vol. 383, pp. 157-162.

Schroeder, A., Mueller, O. & Stocker, S., *et al.* 2006. The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Mol Biol*, vol. 7, no. 3, doi:10.1186/1471-2199-7-3.

Schwanhäusser, B., Busse, D. & Li, N., *et al.* 2011. Global quantification of mammalian gene expression control. *Nature*, vol. 473, no. 7347, pp. 337-342.

Scott, S.A., Lakshimikuttysamma, A. & Sheridan, D.P., *et al.* 2007. Zebularine inhibits human acute myeloid leukemia cell growth in vitro in association with p15INK4B demethylation and reexpression. *Exp Hematol*, vol. 35, no. 2, pp. 263-273.

Seenundun, S., Rampalli, S. & Liu, Q.C., *et al*. 2010. UTX mediates demethylation of H3K27me3 at muscle-specific genes during myogenesis. *EMBO J*, vol. 29, no. 8, pp. 1401-1411.

Shats, I., Gatza, M.L. & Chang, J.T., *et al*. 2011. Using a stem cell-based signature to guide therapeutic selection in cancer. *Cancer Res*, vol. 71, no. 5, pp. 1772-1780.

Shi, B., Sepp-Lorenzino, L. & Prisco, M., *et al.* 2007. Micro RNA 145 targets the insulin receptor substrate-1 and inhibits the growth of colon cancer cells. *J Biol Chem*, vol. 282, pp. 32582-32590.

Shi, Y. & Ai, W. 2013. *Pluripotent Stem Cells*. Rijeka: In Tech, doi: 10.5772/54370, 317 p.

Shi, Z., Bai, R. & Fu, Z., *et al*. 2012. Induced pluripotent stem cell-related genes influence biological behavior and 5-fluorouracil sensitivity of colorectal cancer cells. *J Zhejiang Univ Sci B*, vol. 13, no. 1, pp. 11-19.

Shimada, S.R., Iinuma, H. & Akahane, T., *et al*. 2011. Prognostic significance of ctcs and cscs of tumor drainage vein blood in Dukes' stage B and C colorectal cancer patients. *Oncology Reports*, vol. 27, pp. 947-953.

Shimizu, T., Kagawa, T. & Inoue, T., *et al*. 2008. Stabilized beta-catenin functions through TCF/LEF proteins and the Notch/RBP-Jkappa complex to promote proliferation and suppress differentiation of neural precursor cells. *Mol Cell Biol*., vol. 28, no. 24, pp. 7427-7441.

Shiozawa, Y., Nie, B. & Pienta, K.J., *et al*. 2013. Cancer stem cells and their role in metastasis. *Pharmacol Ther*, vol. 138, no. 2, pp. 285-293.

Shmelkov, S.V., Butler, J.M. & Hooper, A.T., *et al.* 2008. CD133 expression is not restricted to stem cells, and both CD133+ and CD133- metastatic colon cancer cells initiate tumors. *J Clin Invest*, vol. 118, no. 6, pp. 2111-2120.

Sievers, E.L., Larson, R.A. & Stadtmauer, E.A., *et al.* 2001. Efficacy and safety of gemtuzumab ozogamicin in patients with CD33-positive acute myeloid leukemia in first relapse. *J Clin Oncol*, vol. 19, pp. 3244-3254.

Singh, S.K., Hawkins, C. & Clarke, I.D., *et al.* 2004. Identification of human brain tumour initiating cells. *Nature*, vol. 432, pp. 396-401.

Spanakis, E. 1993. Problems related to the interpretation of autoradiographic data on gene expression using common constitutive transcripts as controls. *Nucleic Acids Res*, vol. 21, pp. 3809-3819.

Spangrude, G.J., Heimfeld, S. & Weissman, I.L. 1988. Purification and characterization of mouse hematopoietic stem cells. *Science*, vol. 241, no. 4861, pp. 58-62.

Spizzo, G., Went, P. & Dirnhofer, S., *et al.* 2004. High Ep-CAM expression is associated with poor prognosis in node-positive breast cancer. *Breast Cancer Research and Treatment*, vol. 86, no. 3, pp. 207-213.

Suo, G., Han, J. & Wang, X., *et al.* 2005. Oct4 pseudogenes are transcribed in cancers. *Biochem Biophys Res Commun*, vol. 337, no. 4, pp. 1047-1051.

Suzuki, H., Watkins, D.N. & Jair, K.W., *et al*. 2004. Epigenetic inactivation of SFRP genes allows constitutive WNT signaling in colorectal cancer. *Nat Genet*, vol. 36, no. 4, pp. 417-422.

Suzuki, T., Higgins, P.J. & Crawford, D.R. 2000. Control selection for RNA quantitation. *Biotechniques*, vol. 29, no. 2, pp. 332-337.

Swart, G.W. 2002. Activated leukocyte cell adhesion molecule (CD166/ALCAM): developmental and mechanistic aspects of cell clustering and cell migration. *Eur J Cell Biol*, vol. 81, no. 6, pp. 313-321.

Swinscow, T.D.V. & Campbell, M.J. 2001 *Statistics at Square One*. London: Blackwell BMJ Books. pp. 195-197.

Tai, M.H., Chang, C.C. & Kiupel, M., *et al.* 2005. Oct4 expression in adult human stem cells: evidence in support of the stem cell theory of carcinogenesis. *Carcinogenesis*, vol. 26, no. 7, pp. 495-502.

Takahashi, K. & Yamanaka, S. 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*, vol. 126, no. 4, pp. 663-676.

Takahashi, S., Kamiyama, T. & Tomaru, U., *et al.*, 2010. Frequency and pattern of expression of the stem cell marker CD133 have strong prognostic effect on the surgical outcome of colorectal cancer patients. *Oncol Rep*, vol. 24, pp. 1201-1212.

Takawa M., Masuda, K. & Kunizaki, M., *et al*. 2011. Validation of the histone methyltransferase EZH2 as a therapeutic target for various types of human cancer and as a prognostic marker. *Cancer Sci*, vol. 102, no. 7, pp. 1298-1305.

Tan, C. & Du, X. 2012. *KRAS* mutation testing in metastatic colorectal cancer. *World J Gastroenterol*, vol. 18, no. 37, pp. 5171-5180.

Tang, R., Faussat, A.M. & Majdak, P., *et al.* 2004. Valproic acid inhibits proliferation and induces apoptosis in acute myeloid leukemia cells expressing P-gp and MRP1. *Leukemia*, vol. 18, no. 7, pp. 1246-1251.

Taniguchi, H., Yamamoto, H. & Akutsu, N., *et al*. 2007. Transcriptional silencing of hedgehog-interacting protein by CpG hypermethylation and chromatic structure in human gastrointestinal cancer. *J Pathol*, vol. 213, no. 2, pp. 131-139.

Taniguchi, H., Yamamoto, H. & Hirata, T., *et al*, 2005. Frequent epigenetic inactivation of Wnt inhibitory factor-1 in human gastrointestinal cancers. *Oncogene*, vol. 24, no. 53, pp. 7946-7952.

Tesori, V., Puglisi, M.A. & Lattanzi, W., *et al*. 2013. Update on small intestinal stem cells. *World J Gastroenterol*, vol. 19, no. 29, pp. 4671–4678.

Thelen, P., Schweyer, S. & Hemmerlein, B., *et al.* 2004. Expressional changes after histone deacetylase inhibition by Valproic acid in LNCaP human prostate cancer cells. *Int J Oncol*, vol. 24, no. 1, pp. 25-31.

Thellin, O., Zorzi, W. & Lakaye, B., *et al.* 1999. Housekeeping genes as internal standards: use and limits. *J Biotechnol*, vol. 75, no. 2-3, pp. 291-295.

Tournigand, C., André, T. & Achille, E., *et al.* 2004. FOLFIRI followed by FOLFOX6 or the reverse sequence in advanced colorectal cancer: a randomized GERCOR study. *J Clin Oncol*, vol. 22, no. 2, pp. 229-237.

Toyota, M., Ahuja, N. & Ohe-Toyota. M., *et al*. 1999. CpG island methylator phenotype in colorectal cancer. *Proc Natl Acad Sci USA*, vol. 96, no. 15, pp. 8681-8686.

Trzpis, M., McLaughlin, P.M. & de Leij, L.M., *et al*. 2007. Epithelial cell adhesion molecule: more than a carcinoma marker and adhesion molecule. *Am J Pathol*, vol. 171, no. 2, pp. 386-395.

Tsai, J.H. & Yang, J. 2013. Epithelial-mesenchymal plasticity in carcinoma metastasis. *Genes Dev*, vol. 27, no. 20, pp. 2192-2206.

Tsang, W.P., Ng, E.K. & Ng, S.S., *et al*. 2010. Oncofetal H19-derived miR-675 regulates tumor suppressor RB in human colorectal cancer. *Carcinogenesis*, vol. 31, no. 3, pp. 350-358.

Tsuruzoe, S., Ishihara, K. & Uchimura, Y., *et al*. 2006. Inhibition of DNA binding of Sox2 by the SUMO conjugation. *Biochem Biophys Res Commun*, vol. 351, no. 4, pp. 920-926.

Ubeda, M., Vallejo, M. & Habener, J.F. 1999. CHOP enhancement of gene transcription by interactions with Jun/Fos AP-1 complex proteins. *Mol Cell Biol*, vol. 19, no. 11, pp. 7589-7599.

Van Cutsem, E., Köhne, C.H. & Hitre, E., *et al.* 2009. Cetuximab and chemotherapy as initial treatment for metastatic colorectal cancer. *N Engl J Med,* vol. 360, pp. 1408-1417.

van Engeland, M., Derks, S. & Smits, K.M., *et al*. 2011. Colorectal cancer epigenetics: complex simplicity. *J Clin Oncol*, vol. 29, no. 10, pp. 1382-1391.

Vandesompele, J., De Preter, K. & Pattyn, F., *et al.* 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol*, vol. 3, no. 7, research 0034.1-0034.11.

Varga, M., Obrist, P. & Schneeberger, S., *et al*. 2004. Overexpression of epithelial cell adhesion molecule antigen in gallbladder carcinoma is an independent marker for poor survival. *Clin Cancer Res*, vol. 10, no. 9, pp. 3131-3136.

Varjosalo, M. & Taipale, J. 2008. Hedgehog: functions and mechanisms. *Genes Dev*, vol. 22, no. 18, pp. 2454-2472.

Varnat, F., Duquet, A. & Malerba, M., *et al*. 2009. Human colon cancer epithelial cells harbour active HEDGEHOG-GLI signalling that is essential for tumour growth, recurrence, metastasis and stem cell survival and expansion. *EMBO Mol Med*, vol. 1, no. 6-7, pp. 338-351.

Vaux, D.L. 2012. Research methods: Know when your numbers are significant. *Nature*, vol. 492, no. 7428, pp. 180-181.

Venkataramani, V., Rossner, C. & Iffland, L., *et al*. 2010. Histone deacetylase inhibitor Valproic acid inhibits cancer cell proliferation via down-regulation of the alzheimer amyloid precursor protein. *J Biol Chem*, vol. 285, no. 14, pp. 10678-10689.

Vermeulen, L., Todaro, M. & de Sousa M.F., *et al.* 2008. Single-cell cloning of colon cancer stem cells reveals a multi-lineage differentiation capacity. *Proc Natl Acad Sci USA*, vol. 105, no. 36, pp. 13427-13432.

Villasante, A., Piazzolla, D. & Li, H., *et al*. 2011. Epigenetic regulation of Nanog expression by Ezh2 in pluripotent stem cells. *Cell Cycle*, vol. 10, no. 9, pp. 1488-1498.

Vinci, M., Gowan, S. & Boxall, F., *et al*. (2012). Advances in establishment and analysis of three-dimensional tumor spheroid-based functional assays for target validation and drug evaluation. *BMC Biology*, vol. 10, doi: 10.1186/1741-7007-10-29.

Visvader, J.E. & Lindeman, G.J. 2008. Cancer stem cells in solid tumours: accumulating evidence and unresolved questions. *Nature*, vol. 8, pp. 755-768.

Visvader, J.E. & Lindeman, G.J. 2012. Cancer stem cells: current status and evolving complexities. *Cell Stem Cell*, vol. 10, no. 6, pp. 717-728.

Vogelstein, B., Fearon, E.R. & Hamilton, S.R., *et al.* 1988. Genetic alterations during colorectal-tumor development. *N Engl J Med*, vol. 319, no. 9, pp. 525-532.

Vogt, M., Munding, J. & Grüner, M., *et al*. 2011. Frequent concomitant inactivation of miR-34a and miR-34b/c by CpG methylation in colorectal, pancreatic, mammary, ovarian, urothelial, and renal cell carcinomas and soft tissue sarcomas. *Virchows Arch*, vol. 458, no. 3, pp. 313-322.

Voorham, Q.J., Janssen, J. & Tijssen, M., *et al*. 2013. Promoter methylation of Wnt-antagonists in polypoid and nonpolypoid colorectal adenomas. *BMC Cancer*, doi: 10.1186/1471-2407-13-603.

Voorneveld, P.W., Stache, V. & Jacobs, R.J., *et al*. 2013. Reduced expression of bone morphogenetic protein receptor IA in pancreatic cancer is associated with a poor prognosis. *Br J Cancer*, vol. 109, no. 7, pp. 1805-1812.

Wang, C., Xie, J. & Guo, J., *et al*. 2012. Evaluation of CD44 and CD133 as cancer stem cell markers for colorectal cancer. *Oncol Rep*, vol. 28, no. 4, pp. 1301-1308.

Wang, J., Wang, J. & Li, Z., *et al*. 2008. c-Myc is required for maintenance of glioma cancer stem cells. *PLoS ONE*, vol. 3, no. 11, doi: 10.1371/journal.pone.0003769.

Wang, L., Li, L. & Menendez, P., *et al*. 2005. Human embryonic stem cells maintained in the absence of mouse embryonic fibroblasts or conditioned media are capable of hematopoietic development. *Blood*, vol. 105, no. 12, pp. 4598-4603.

Wang, P., Gao, Q. & Suo, Z., *et al*. 2013a. Identification and charcaterization of cells with cancer stem cell properties in human primary lung cancer cell lines. *PLoS One*, doi: 10.1371/journal.pone.0057020.

Wang, Q., Chen, Z.G. & Du, C.Z., *et al*. 2009. Cancer stem cell marker CD133+ tumour cells and clinical outcome in rectal cancer. *Histopathology*, vol. 55, pp. 284-293.

Wang, S.H., Tsai, M.S. & Chiang, M.F., *et al.* 2003. A novel NK-type homeobox gene, ENK (early embryo specific NK), preferentially expressed in embryonic stem cells. *Gene Expr Patterns*, vol. 3, pp. 99-103.

Wang, X.Q., Ng, R.K. & Ming, X., *et al.* 2013b. Epigenetic regulation of pluripotent genes mediates stem cell features in human hepatocellular carcinoma and cancer cell lines. *PLoS One*, doi: 10.1371/journal.pone.0072435.

Warrington, J.A., Nair, A. & Mahadevappa, M., *et al*. 2000. Comparison of human adult and fetal expression and identification of 535 housekeeping/maintenance genes. *Physiol. Genomics*, vol. 2, pp. 143-147.

Watson, M., Berg, M. & Søreide, K. 2014. Prevalence and implications of elevated microsatellite alterations at selected tetranucleotides in cancer. *Br J Cancer*, vol. 111, pp. 823-827.

Weichert, W., Denkert, C. & Noske, A., *et al*. 2008. Expression of class I histone deacetylases indicates poor prognosis in endometrioid subtypes of ovarian and endometrial carcinomas. *Neoplasia*, vol. 10, no. 9, pp. 1021-1027.

Weidle, U.H., Eggle, D. & Klostermann, S., *et al*. 2010. ALCAM/CD166: cancer-related issues. *Cancer Genomics Proteomics*, vol. 7, no. 5, pp. 231-243.

Weigmann, A., Corbeil, D. & Hellwig, A., *et al*. 1997. Prominin, a novel microvilli-specific polytopic membrane protein of the apical surface of epithelial cells, is targeted to plasmalemmal protrusions of non-epithelial cells. *Proc Natl Acad Sci USA*, vol. 94, pp. 12425-12430.

Weinberg, R.A. 2007. *The Biology of Cancer, 1st edition*. New York: Taylor & Francis Group, LLC. 409 p.

Wen, X.Z., Akiyama, Y. & Baylin, S.B., *et al*. 2006. Frequent epigenetic silencing of the bone morphogenetic protein 2 gene through methylation in gastric carcinomas. *Oncogene,* vol. 25, no. 18, pp. 2666-2673.

Went, P., Vasei, M. & Bubendorf, L., *et al*. 2006. Frequent high-level expression of the immunotherapeutic target Ep-CAM in colon, stomach, prostate and lung cancers. *Br J Cancer*, vol. 94, pp. 128-135.

Wildi, S., Kleeff, J. & Maruyama, H., *et al.* 1999. Characterization of cytokeratin 20 expression in pancreatic and colorectal cancer. *Clin Cancer Res*, vol. 5, no. 10, pp. 2840-2847.

Willis, N.D., Przyborski, S.A. & Hutchison, C.J., *et al*. 2008. Colonic and colorectal cancer stem cells: progress in the search for putative biomarkers. *J Anat*, vol. 213, no. 1, pp. 59-65.

Witt, D., Burfeind, P. & von Hardenberg, S., *et al.* 2013. Valproic acid inhibits the proliferation of cancer cells by re-expressing cyclin D2. *Carcinogenesis*, vol. 34, no. 5, pp. 1115-1124.

Wolf, J., Müller-Decker, K. & Flechtenmacher, C., *et al*. 2014. An *in vivo* RNAi screen identifies *SALL1* as a tumor suppressor in human breast cancer with a role in *CDH1* regulation. *Oncogene,* vol. 33, pp. 4273-4278.

Wong, D.J., Liu, H. & Ridky, T.W., *et al.* 2008. Module map of stem cell genes guides creation of epithelial cancer stem cells. *Cell Stem Cell*, vol. 2, pp. 333–344.

Woodward, W.A. & Bristow, R.G. 2009. Radiosensitivity of cancer-initiating cells and normal stem cells (or what the Heisenberg uncertainly principle has to do with biology). *Semin Radiat Oncol*, vol. 19, pp. 87-95.

Xi, H.Q. & Zhao, P. 2011. Clinicopathological significance and prognostic value of EphA3 and CD133 expression in colorectal carcinoma. *J Clin Pathol*, vol. 64, pp. 498-503.

Xu, R.H., Peck, R.M. & Li, D.S., *et al.* 2005. Basic FGF and suppression of BMP signalling sustain undifferentiated proliferation of human ES cells. *Nat Methods*, vol. 2, no. 3, pp. 185-190.

Yamashita, T., Ji, J. & Budhu, A., *et* al. 2009. EpCAM-positive hepatocellular carcinoma cells are tumor-initiating cells with stem/progenitor cell features. *Gastroenterology*, vol. 136, no. 3, pp. 1012-1024.

Yanamoto, S., Kawasaki, G. & Yoshitomi, I., *et al.* 2007. Clinicopathologic significance of EpCAM expression in squamous cell carcinoma of the tongue and its possibility as a potential target for tongue cancer gene therapy. *Oral Oncol*, vol. 43, no. 9, pp. 869-877.

Yeung, T.M., Gandhi, S.C. & Wilding, J.L., *et al*. 2010. Cancer stem cells from colorectal cancer-derived cell lines. *Proc Natl Acad Sci USA*, vol. 107, no. 8, pp. 3722-3727.

Yin, A.H., Miraglia, S. & Zanjani, E.D., *et al.* 1997. AC133, a novel marker for human hematopoietic stem and progenitor cells. *Blood*, vol. 90, no. 12, pp. 5002-5012.

Ying, Y. & Tao, Q. 2009. Epigenetic disruption of the WNT/beta-catenin signaling pathway in human cancers. *Epigenetics*, vol. 4, pp. 307-312.

Yoo, C.B. & Jones, P.A. 2006. Epigenetic therapy of cancer: past, present and future. *Nat Rev Drug Discov*, vol. 5, no. 1, pp. 37-50.

Yu, K.S., Lee, Y. & Kim, C.M. 2010. The protease inhibitor, elafin, induces p53-dependent apoptosis in human melanoma cells. *Int J Cancer*, vol. 127, no. 6, doi: 10.1002/ijc.25125.

Yu, X., Lin, Y. & Yan, X., *et al.* 2011. CD133, stem cells, and cancer stem cells: myth or reality? *Curr Colorectal Cancer Rep*, vol. 7, no. 4, pp. 253-259.

Zaehres, H., Lensch, M.W. & Daheron, L., *et al.* 2005. High-efficiency RNA interference in human embryonic stem cells. *Stem Cells*, vol. 23, no. 3, pp. 299-305.

Zeng, X., Huang, H. & Tamai, K., *et al*. 2008. Initiation of Wnt signaling: control of Wnt coreceptor Lrp6 phosphorylation/activation via frizzled, dishevelled and axin functions. *Development*, vol. 135, no. 2, pp. 367-375.

Zenonos, K. & Kyprianou. K. 2013. RAS signaling pathways, mutations and their role in colorectal cancer. *World J Gastrointest Oncol*, vol 5, no. 5, pp. 97-101.

Zeppernick, F., Ahmadi, R. & Campos, B., *et al*. 2008. Stem cell marker CD133 affects clinical outcome in glioma patients. *Clin Cancer Res*, vol. 14, no. 1, pp. 123-129.

Zgouras, D., Becker, U. & Loitsch, S., *et al.*, 2004. Modulation of angiogenesis-related protein synthesis by Valproic acid. *Biochem Biophys Res Commun*, vol. 316, no. 3, pp. 693-697.

Zhang, B., Halder, S.K. & Kashikar, N.D., *et al*. 2010. Antimetastatic role of Smad4 signaling in colorectal cancer. *Gastroenterology*, vol. 138, no. 3, pp. 969-980.

Zhang, J., Wang, X. & Li, M., *et al*. 2006. NANOGP8 is a retrogene expressed in cancers. *FBS J*, vol. 273, no. 8, pp. 1723-1730.

Zhang, W., Kai, K. & Choi, D.S., *et al.* 2012. Microfluidics separation reveals the stem-cell–like deformability of tumor-initiating cells. *PNAS*, vol. 109, no. 46, pp. 18707-18712.

Zhou, L., Wei, X. & Cheng, L., *et al.* 2007. CD133, one of the markers of cancer stem cells in Hep-2 cell line. *Laryngoscope*, vol. 117, pp. 455-460.

E-REFERENCES

1. American Type Culture Collection. DLD-1 characteristics. USA. <http://www.atcc.org/> (Accessed 3.3.2013).
2. American Type Culture Collection. SW1116 characteristics. 2013. USA. <http://www.atcc.org/> (Accessed 3.3.2013).
3. American Cancer Society. Cancer Facts & Figures 2014. Atlanta. <http://www.cancer.org/acs/groups/content/@research/documents/webcontent/acspc-042151.pdf> (Accessed 28.3.14).
4. BD Biosciences. Matrigel™. USA. <http://www.bdbiosciences.com/documents/ecm_cancer_stem_cell.pdf> (Accessed 5.10.14).
5. Ensembl Genome Browser <http://www.ensembl.org/index.html> (Accessed 11.6.12).
6. Life Technologies. Geltrex™. USA. <http://www.lifetechnologies.com/order/catalog/product/A1413202> (Accessed 5.10.14).
7. National Centre for Biotechnological Information, BLAST. <http://www.blast.ncbi.nlm.nih.gov/Blast.cgi> (Accessed 14.6.12).
8. MACS Magnetic Cell Separation. <http://www.miltenyibiotec.com> (Accessed 6.11.13).
9. Qiagen. <http://www.qiagen.com/us/products/genes%20and%20pathways/pathway%20details.aspx> (Accessed 17.11.14).
10. TATAA Interplate Calibrator SYBR protocol. <http://www.tataa.com/wp-contenct/uploads/2012/10/TATAA-Manual_Interplate_Calibrator_SYBR_v01_.pdf> (Accessed 2.7.13).
11. The South African National Cancer Registry. 2007. < http://www.nioh.ac.za/?page=cancer_statistics&id=163> (Accessed 10.8.2014).

12. UCSC Genome Bioinformatics. Human BLAT search. <http://www.genome.ucsc.edu/cgi-bin/hgBlat?command=start> (Accessed 14.6.12).

Appendix A

Procedures

All the chemicals used in this section of this study were sourced from Sigma Aldrich unless otherwise stated.

A.1 Cell Culture

Serum Free Cancer Stem Cell Selection Medium (Mather, 2008)

The serum-free cancer stem cell selection medium (DMEM/F12) was supplemented with the following cell culture grade ingredients:

6mg/ml glucose
1mg/ml sodium bicarbonate
2mM L-glutamine
4µg/ml heparin
4mg/ml BSA
10ng/ml bFGF
20ng/ml EGF
100µg/ml apotransferrin
25µg/ml insulin
9.6µg/ml putrescine
30nM sodium selenite anhydrous
20nM progesterone

Add 100µl 10000 IU penicillin and streptomycin (Lonza) per 50ml media

Filter sterilise

Only add bFGF and EGF, once needed

2% Agar

2g Agar (Difco)

100ml distilled water

Autoclave

Once cooled, add enough agar to cover the area of the growth surface of a dish (plate wells or flasks)

A.2 Immunofluorescence Microscopy

Phosphate Buffered Saline (PBS)

5 PBS tablets in one litre distilled water (Sigma Aldrich)
Autoclave

Blocking Solution/Buffer

PBS 0.5% BSA

50ml PBS

0.25g BSA

Dissolve

Fixing Solution/Buffer

PBS 3% Formaldehyde

16.5ml PBS

1.5ml Formaldehyde (37%, Merck)

Mix

Permeabilisation Solution/Buffer

PBS 0.5% BSA 0.25% Triton X-100

20ml PBS

0.1g BSA

50µl Triton X

A.3 Gel Electrophoresis

10X TBE (Tris, borate, ethylenediaminetetraacetic acid) buffer

108g Tris (Merck)

55g Boric acid (Merck)

9.3g EDTA (Merck)

Dissolve and make up to a total volume of 1l, adjust pH to 8.3 and store at room temperature.

To make a one litre working solution of 1X TBE buffer, mix 100ml 10X TBE buffer with 900ml distilled water.

1.5% agarose gel

1.5g agarose (Bio-Rad)

100ml TBE buffer

In a microwave, dissolve agarose in 1X TBE buffer, once slightly cooled add 5µl ethidium bromide (Promega), swirl to mix. Pour into a casting tray with a comb and allow to set.

A.4 Flow Cytometry

Wash Buffer

PBS, 2mM EDTA and 0.5% BSA

80ml PBS

14.89mg EDTA (Merck)

Adjust pH to 8.0 using 10N NaOH

Add 0.5g BSA and dissolve

Make up to 100ml

Appendix B

Methods

B.1 Thin Gel (non-gelling) Method for Three-dimensional Cell Culture (Matrigel/Geltrex)

1. Thaw Matrigel/Geltrex (BD Biosciences/Gibco) on ice
2. Gently pipette-mix the Matrigel/Geltrex and dilute 1ml into 29ml DMEM/F12
3. Add enough diluted Matrigel/Geltrex to cover the area of the growth surface of a dish (plate wells or flasks)
4. Incubate the coated plastics at 37°C for one hour, and remove the DMEM/F12 medium
5. Immediately plate cells, or alternatively seal the plastics with parafilm and store at 2-8°C for up to 2 weeks.

B.2 RNA Extraction and DNase Digestion

Qiagen Easy Plus RNA Extraction and On-column DNase Treatment (Qiagen)

This method of RNA extraction utilises a column wherein a silica-gel-membrane with selective binding properties, is placed. The sample is lysed and homogenised. Ethanol is added to provide appropriate binding conditions, and the sample is added to the column. RNA binds to the membrane and unbound sample is passed through by centrifugal force. The sample is washed and treated with DNase to ensure complete removal of DNA from the sample. An appropriate elution buffer is applied to the column and purified RNA is eluted.

1. Collect and disrupt cells in 600µl RLT Buffer containing 10µl/ml β-mercaptoethanol (Sigma-Aldrich). The lysates are stable for up to 6 months at -70°C.

2. Thaw cell lysates and homogenise in QIAshredder spin columns centrifuged for 2 minutes at 14 000rpm.
3. Transfer the homogenised lysates to genomic DNA eliminator spin columns and centrifuge for 30 seconds at 10 900rpm.
4. Add 600µl ice cold ethanol (70%) to the flow-through and pipette mix.
5. Transfer up to 700µl of the sample to an RNeasy spin column and centrifuge for 15 seconds at 10 900rpm. Add the remainder of the flow-through and repeat.
6. Discard the flow-through, add 350µl Buffer RW1 and centrifuge for 15 seconds at 10 900rpm and discard the flow- through.
7. Add 10µl of re-constituted DNase mixed in 70µl Buffer RDD onto the RNeasy spin column membrane and incubate at room temperature for 15 minutes followed by the addition of 350µl Buffer RW1. Centrifuge the columns for 15 seconds at 10 900rpm and discard the flow-through. Add 500µl Buffer RPE to the column and centrifuge for 15 seconds at 10 900rpm followed by an additional wash with 500µl Buffer RPE, centrifuged at 10 900rpm for 2 minutes.
8. Place RNeasy spin columns in new 2ml collection tubes and centrifuge for 1 minute at 14 000rpm.
9. Place spin columns in new 1.5ml collection tubes, add 30µl RNase-free water to the column membrane and incubate for 5 minutes followed by centrifugation at 10 900rpm for 1 minute to elute RNA.
10. Add 5µl eluted RNA from step 9 to a 1.5ml eppendorf tube. Add 85µl RNase-free water, 10µl 10X DNase I Reaction Buffer (New England Biolabs) and 2µl DNase I (RNase-free, 2000U/ml) (New England Biolabs).
11. Incubate for 10 minutes at 37°C followed by the addition of 1µl 0.5M EDTA (filter sterilised). Incubate for 10 minutes at 75°C and briefly centrifuge to remove droplets from the lid.
12. Add 350µl RLT Buffer, mix and add 250µl ice cold ethanol (96-100%). Pipette mix and apply 750µl of sample to a RNeasy column placed in a 2ml collection tube, followed by centrifugation at 10 900rpm for 15 seconds.

13. Transfer the column to a new 2ml collection tube and add 500µl Buffer RPE followed by centrifugation at 10 900rpm for 15 seconds.
14. Discard the flow-through and add 500µl Buffer RPE followed by centrifugation at 10 900rpm for 2 minutes.
15. Repeat steps 8 and 9 to elute RNA.

B.3 Genomic DNA Extraction (Qiagen)

Qiagen QIAamp DNA Mini Kit

This method of DNA extraction utilises a column wherein a silica-gel-membrane with selective binding properties, is placed. The sample is lysed and ethanol is added to provide appropriate binding conditions. The sample is added to the column and DNA is adsorbed onto the membrane during centrifugation. The DNA is washed and purified DNA is eluted in an appropriate elution buffer.

1. Collect and wash cells with 2ml PBS. Centrifuge cells for 3 minutes at 800rpm, and discard the supernatant. Repeat the wash and re-suspend the cell pellet in 200µl PBS, in a 1.5ml microcentrifuge tube.
2. Add 20µl Proteinase K and 200µl Buffer AL. Pulse vortex for 15 seconds and incubate at 56°C for 10 minutes.
3. Add 200µl ice cold ethanol (96-100%) and pulse vortex for 15 seconds.
4. Apply the suspension to a QIAamp spin column placed in a 2ml collection tube, and centrifuge for 1 minute at 8000rpm.
5. Place the column in a new 2ml collection tube and add 500µl Buffer AW1. Centrifuge for 1 minute at 8000rpm.
6. Place the column in a new 2ml collection tube and add 500µl Buffer AW2. Centrifuge for 3 minutes at 14000rpm.
7. Place the column in a new 2ml collection tube and centrifuge for 1 minute at 14000rpm.

8. Place the column in a 1.5ml microcentrifuge tube and add 200µl Buffer AE to the column. Incubate for 1 minute at room temperature, and centrifuge for 1 minute at 8000rpm.

Appendix C

Tables

C.1 Flow Cytometric Characterisation of CD133 and EpCAM

Table C1: $CD133^-/EpCAM^-$, $CD133^-/EpCAM^+$ and $CD133^+/EpCAM^+$ expression in the SW1116, HT29 and DLD1 cell lines (N=3).

Cell Line	Marker Expression (%)		
	$CD133^-/EpCAM^-$	**$CD133^-/EpCAM^+$**	**$CD133^+/EpCAM^+$**
SW1116	8.96 12.80 9.03	88.70 84.50 88.70	2.33 2.65 2.29
Median	9.03	88.70	2.33
HT29	0.77 0.58 0.34	94.10 94.10 94.70	5.11 5.31 4.98
Median	0.58	94.10	5.11
DLD1	0.19 0.74 0.05	89.50 88.60 89.80	10.30 10.50 10.10
Median	0.19	89.50	10.30

C.2 Sphere Formation Assay

Table C2: Sphere-forming capacity of $CD133^{+}/EpCAM^{+}$ HT29 and DLD1 cells (N=3).

Replicate	**Number of Spheres/well**	
	HT29	DLD1
Replicate 1	23 13 18 19 17 14	29 25 16 18 21 23
Replicate 2	14 29 24 29 16 16	13 19 23 28 21 19
Replicate 3	24 28 29 26 17 27	25 26 18 22 21 15
Median	21	21
Lower Quartile	16	18
Upper Quartile	27	25

C.3 CD133 Expression in Colonic Tumours

Table C3: CD133 expression in colonic tumours with varying Dukes' stages (adapted from Ricci-Vitiani *et al.*, 2007).

Case	Dukes' Stage	CD133 Expression (%)
1	C	2.40
2	C	2.00
3	D	2.50
4	A	1.70
5	C	1.40
6	C	2.40
7	C	1.80
8	A	0.70
9	A	2.40
10	B	0.70
11	A	-
12	C	-
13	B	2.60
14	B	4.10
15	C	2.70
16	B	6.10
17	A	1.70
18	B	4.60
19	D	-

C.4 CD133 Expression in Colonic Tumours and Normal Tissues

Table C4: CD133 expression in 17 colonic tumours of varying stages, and in their corresponding normal tissues (adapted from O'Brien *et al.*, 2007).

Patient Number	Tumour Stage	$CD133^+$ in tumour (%)	$CD133^+$ in normal (%)
1	IB	14.00	2.10
2	IV	5.20	-
3	IV	14.70	-
4	IV	1.80	-
5	IV	24.50	-
6	IV	6.30	-
7	I	9.30	1.20
8	IIIC	7.50	0.40
9	IV	12.10	-
10	IIIC	8.90	1.30
11	IV	19.00	-
12	IIIC	15.90	0.85
13	I	12.00	1.90
14	IV	17.60	-
15	IV	18.20	-
16	IV	10.40	-
17	IV	3.20	-

C.5 Pluripotent Marker Expression in HT29 putative CSCs

Table C5: Median pluripotent marker scores in HT29 putative CSCs (N=3).

Treatment	**Median**		
HT29 NANOG	Replicate 1	Replicate 2	Replicate 3
NT	1.0	1.0	1.0
1mM VPA	1.0	2.0	2.0
2.5mM VPA	2.0	2.0	2.0
5mM VPA	1.0	2.0	2.0
DMSO	1.0	1.0	1.0
100µM ZEB	1.0	2.0	2.0
250µM ZEB	2.0	2.0	2.0
500µM ZEB	2.0	2.0	2.0
HT29 OCT4	Replicate 1	Replicate 2	Replicate 3
NT	1.0	1.0	1.0
1mM VPA	1.0	1.0	1.0
2.5mM VPA	1.0	1.0	1.0
5mM VPA	2.0	2.0	2.0
DMSO	1.0	1.0	1.0
100µM ZEB	2.0	1.5	1.0
250µM ZEB	1.0	2.0	2.0
500µM ZEB	2.0	2.0	2.0
HT29 SOX2	Replicate 1	Replicate 2	Replicate 3
NT	2.0	2.0	2.0
1mM VPA	2.0	2.0	2.0
2.5mM VPA	2.0	2.0	2.0
5mM VPA	2.0	2.0	2.0
DMSO	2.0	2.0	2.0
100µM ZEB	2.0	2.0	2.0
250µM ZEB	2.0	2.0	2.0
500µM ZEB	2.0	2.0	2.0

Key: NT=No treatment, VPA=Valproic acid, ZEB=Zebularine, DMSO=Dimethyl sulphoxide.

C.6 Pluripotent Marker Expression in DLD1 putative CSCs

Table C6: Median pluripotent marker scores in DLD1 putative CSCs (N=3).

Treatment	**Median**		
DLD1 NANOG	Replicate 1	Replicate 2	Replicate 3
NT	1.0	1.0	1.0
1mM VPA	1.0	1.0	1.0
2.5mM VPA	1.0	1.0	1.0
5mM VPA	1.0	1.0	1.0
DMSO	1.5	2.0	1.0
100µM ZEB	2.0	2.0	1.5
250µM ZEB	2.0	3.0	3.0
500µM ZEB	2.0	2.0	3.0
DLD1 OCT4	Replicate 1	Replicate 2	Replicate 3
NT	2.0	2.0	1.0
1mM VPA	2.0	2.0	1.0
2.5mM VPA	2.0	2.0	1.0
5mM VPA	2.5	3.0	2.0
DMSO	1.0	1.0	1.0
100µM ZEB	2.0	2.5	1.0
250µM ZEB	2.0	2.5	2.0
500µM ZEB	2.5	2.0	2.0
DLD1 SOX2	Replicate 1	Replicate 2	Replicate 3
NT	2.0	2.0	2.0
1mM VPA	2.0	2.0	2.0
2.5mM VPA	2.0	2.0	2.0
5mM VPA	2.0	2.0	2.0
DMSO	2.0	2.0	2.0
100µM ZEB	2.0	2.0	2.0
250µM ZEB	2.0	2.0	2.0
500µM ZEB	2.0	2.0	2.0

Key: NT=No treatment, VPA=Valproic acid, ZEB=Zebularine, DMSO=Dimethyl sulphoxide.

C.7 Statistical Analyses of Pluripotent Marker Expression in HT29 Putative CSCs

Table C7: Kruskal-Wallis and Mann-Whitney tests of nuclear NANOG, OCT4 and SOX2 expression in HT29 putative CSCs (N=3).

HT29 putative CSCs				
Treatment	**Statistical Test**	**NANOG**	**OCT4**	**SOX2**
All treatments	Kruskal-Wallis	*p=0.038	*p=0.011	p=1.000
NT *versus* 1mM VPA	Mann-Whitney	p=0.187	p=1.000	-
NT *versus* 2.5mM VPA	Mann-Whitney	*p=0.046	p=1.000	-
NT *versus* 5mM VPA	Mann-Whitney	p=0.187	*p=0.046	-
1mM *versus* 2.5mM VPA	Mann-Whitney	p=0.505	p=1.000	-
1mM *versus* 5mM VPA	Mann-Whitney	p=0.792	*p=0.046	-
2.5mM *versus* 5mM VPA	Mann-Whitney	p=0.505	*p=0.046	-
NT *versus* DMSO	Mann-Whitney	p=1.000	p=1.000	-
DMSO *versus* 100µM ZEB	Mann-Whitney	p=0.187	p=0.196	-
DMSO *versus* 250µM ZEB	Mann-Whitney	*p=0.046	p=0.187	-
DMSO *versus* 500µM ZEB	Mann-Whitney	*p=0.046	*p=0.046	-
100µM *versus* 250µM ZEB	Mann-Whitney	p=0.505	p=0.813	-
100µM *versus* 500µM ZEB	Mann-Whitney	p=0.505	p=0.196	-
250µM *versus* 500µM ZEB	Mann-Whitney	p=1.000	p=0.505	-

*$p \leq 0.05$ is statistically significant; Key: NT=No treatment, VPA=Valproic acid, ZEB=Zebularine, DMSO=Dimethyl sulphoxide.

C.8 Statistical Analyses of Pluripotent Marker Expression in DLD1 Putative CSCs

Table C8: Kruskal-Wallis and Mann-Whitney tests of nuclear NANOG, OCT4 and SOX2 expression in DLD1 putative CSCs (N=3).

DLD1 putative CSCs				
Treatment	**Statistical Test**	**NANOG**	**OCT4**	**SOX2**
All treatments	Kruskal-Wallis	*p=0.004	*p=0.032	p=1.000
NT *versus* 1mM VPA	Mann-Whitney	p=1.000	p=1.000	-
NT *versus* 2.5mM VPA	Mann-Whitney	p=1.000	p=1.000	-
NT *versus* 5mM VPA	Mann-Whitney	p=1.000	p=0.196	-
1mM *versus* 2.5mM VPA	Mann-Whitney	p=1.000	p=1.000	-
1mM *versus* 5mM VPA	Mann-Whitney	p=1.000	p=0.196	-
2.5mM *versus* 5mM VPA	Mann-Whitney	p=1.000	p=0.196	-
NT *versus* DMSO	Mann-Whitney	p=0.196	*p=0.046	-
DMSO *versus* 100µM ZEB	Mann-Whitney	p=0.479	p=0.059	-
DMSO *versus* 250µM ZEB	Mann-Whitney	p=0.115	p=0.063	-
DMSO *versus* 500µM ZEB	Mann-Whitney	p=0.164	p=0.059	-
100µM *versus* 250µM ZEB	Mann-Whitney	p=0.157	p=0.479	-
100µM *versus* 500µM ZEB	Mann-Whitney	p=0.301	p=0.792	-
250µM *versus* 500µM ZEB	Mann-Whitney	p=0.619	p=0.479	-

*$p \leq 0.05$ is statistically significant; Key: NT=No treatment, VPA=Valproic acid, ZEB=Zebularine, DMSO=Dimethyl sulphoxide.

C.9 RNA Integrity Numbers (RINs)

Table C9: RNA integrity numbers of RNA extracted from HT29 and DLD1 putative CSCs.

Treatment (N=3)	RINs-HT29 CSCs	RINs-DLD1 CSCs
NT1	7.30	8.60
NT2	8.10	8.10
NT3	9.30	7.70
1mM VPA1	9.00	9.20
1mM VPA2	7.20	8.80
1mM VPA3	8.70	8.50
2.5mM VPA1	8.60	9.20
2.5mM VPA2	8.00	9.00
2.5mM VPA3	6.60	8.80
5mM VPA1	7.20	9.50
5mM VPA2	8.80	9.50
5mM VPA3	8.60	7.40
DMSO1	7.30	9.50
DMSO2	8.90	9.00
DMSO3	8.20	8.10
100µM ZEB1	7.90	9.70
100µM ZEB2	8.30	9.60
100µM ZEB3	9.00	8.20
250µM ZEB1	6.30	9.50
250µM ZEB2	8.30	8.60
250µM ZEB3	8.40	7.60
500µM ZEB1	7.70	8.90
500µM ZEB2	6.70	9.30
500µM ZEB3	7.80	7.80

Key: NT=No treatment, VPA=Valproic acid, ZEB=Zebularine, DMSO=Dimethyl sulphoxide.

C.10 Reverse Transcription Master Mix (Applied Biosystems)

Table C10: Stock concentrations and volumes of reagents used per 10µl reverse transcription reaction.

Reagent	**Volume and stock concentration**
Taq Man Buffer	1.00µl (10X)
$MgCl_2$	2.20µl (25mM)
dNTPs	2.00µl (10mM)
Random Hexamers/Oligo d(T)16s	0.50µl (2.5µM)
RNase Inhibitor	0.20µl (20µM)
Reverse Transcriptase	0.25µl (5U/µl)
RNA + RNase free water	3.85µl (700-1000ng/µl)
Total volume	**10.00µl**

C.11 Reverse Transcription Thermal Cycling Parameters

Table C11: Reverse transcription thermal cycling parameters (number of cycles=1).

Step	**Temperature (°C)**	**Time (minutes)**	**Number of cycles**
Incubation	25	10	1
Reverse Transcription	48	30	1
Taq Inactivation	95	5	1

C.12 SYBR Green PCR Master Mix

Table C12: Stock concentrations and volumes of reagents used per 10µl real-time PCR reaction.

Reagent	**Volume and Stock concentration**
SYBR green master mix	5.00µl (5X)
Forward primer	1.00µl (2µM)
Reverse primer	1.00µl (2µM)
cDNA	3.00µl
Total volume	**10.00µl**

C.13 Reference Gene Primers

Table C13: Primer sequences of reference genes.

Primer	Sequence (forward and reverse, respectively)
ACTB	5'- ACC AAC TGG GAC GAC ATG GAG AAA -3' 5'- TAG CAC AGC CTG GAT AGC AAC GTA -3'
GUS	5'- ACG AAC GCC CTG CCT ATC TGT ATT -3' 5'- ATG AGG AAC TGG CTC TTG GTG ACA -3'
β2M	5'- TGC TGT CTC CAT GTT TGA TGT ATC T -3' 5'- TCT CTG CTC CCC ACC TCT AAG T -3'
GAPDH	5'- TGC ACC ACC ACC TGC TTA GC -3' 5'- GGC ATG GAC TGT GGT CAT GAG -3'
HPRT	5'- TGA CAC TGG CAA AAC AAT GCA -3' 5'- GGT CCT TTT CAC CAG CAA GCT -3'
TBP	5'- TGA TGC CTT ATG GCA CTG GAC TGA -3' 5'- CTG CTG CCT TTG TTG CTC TTC CAA -3'
UBC	5'- ATT TGG GTC GCG GTT CTT -3' 5'- TGC CTT GAC ATT CTC GAT GGT -3'
YWHAZ	5'- ACT TTT GGT ACA TTG TGG CTT CAA -3' 5'- CCG CCA GGA CAA ACC AGT AT -3'
TR	5'- GGC ACC ATC AAG CTG CTG AAT GAA -3' 5'- GTT GAT CAC GCC AGA CTT TGC TGA -3'

C.14 Gene Expression ($\log_{10}{}^{2-\Delta\Delta Cq}$) in HT29 putative CSCs

Table C14: *ACTB*, *HPRT*, *NANOG*, *OCT4* and *YWHAZ* gene expression in HT29 putative CSCs.

Sample	***ACTB***	***HPRT***	***NANOG***	***OCT4***	***YWHAZ***
HT29 CSC NT1	-0,01	0,06	-0,28	0,26	-0,05
HT29 CSC NT2	-0,15	0,06	-0,53	0,00	0,09
HT29 CSC NT3	0,10	-0,08	-0,22	0,12	-0,02
HT29 CSC 1mM VPA1	-0,10	0,07	-0,18	-0,16	0,03
HT29 CSC 1mM VPA2	-0,02	0,04	-0,03	-0,02	-0,03
HT29 CSC 1mM VPA3	0,17	-0,14	0,22	-0,01	-0,03
HT29 CSC 2.5mM VPA1	-0,11	-0,03	-0,30	-0,22	0,14
HT29 CSC 2.5mM VPA2	-0,01	-0,04	0,11	-0,02	0,05
HT29 CSC 2.5mM VPA3	0,02	-0,04	0,42	-0,07	0,02
HT29 CSC 5mM VPA1	-0,13	0,06	0,00	0,37	0,07
HT29 CSC 5mM VPA2	-0,20	0,10	0,07	0,02	0,10
HT29 CSC 5mM VPA3	0,04	-0,02	0,03	0,02	-0,03
HT29 CSC DMSO1	0,19	-0,08	-0,43	0,31	-0,12
HT29 CSC DMSO2	-0,06	0,03	-0,32	-0,16	0,03
HT29 CSC DMSO3	0,10	-0,01	-0,40	0,18	-0,09
HT29 CSC 100uM ZEB1	0,13	-0,15	-0,45	-0,06	0,02
HT29 CSC 100uM ZEB2	-0,03	0,02	-0,34	-0,51	0,01
HT29 CSC 100uM ZEB3	0,07	0,05	0,45	0,20	-0,12
HT29 CSC 250uM ZEB1	-0,11	0,14	0,07	0,14	-0,03
HT29 CSC 250uM ZEB2	-0,13	0,14	0,26	-0,05	-0,02
HT29 CSC 250uM ZEB3	0,00	0,07	0,01	0,30	-0,07
HT29 CSC 500uM ZEB1	-0,08	0,14	0,12	0,11	-0,06
HT29 CSC 500uM ZEB2	-0,02	0,03	0,07	0,12	-0,01
HT29 CSC 500uM ZEB3	0,18	-0,02	0,19	0,18	-0,16

Key: NT=No treatment, VPA=Valproic acid, ZEB=Zebularine, DMSO=Dimethyl sulphoxide.

C.15 Standard Error in Gene Expression (log_{10} $^{2-\Delta\Delta Cq}$) in HT29 putative CSCs

Table C15: Standard Error in *ACTB*, *HPRT*, *NANOG*, *OCT4* and *YWHAZ* gene expression in HT29 putative CSCs.

Sample	***ACTB***	***HPRT***	***NANOG***	***OCT4***	***YWHAZ***
HT29 CSC NT1	6,28	0,87	0,09	0,34	1,23
HT29 CSC NT2	0,08	0,26	0,01	6,30	0,44
HT29 CSC NT3	1,52	0,48	0,12	0,52	2,99
HT29 CSC 1mM VPA1	0,22	0,35	0,06	0,19	1,32
HT29 CSC 1mM VPA2	1,64	1,93	2,98	1,35	2,36
HT29 CSC 1mM VPA3	0,79	0,24	0,28	5,50	0,83
HT29 CSC 2.5mM VPA1	0,27	1,45	0,12	0,14	0,92
HT29 CSC 2.5mM VPA2	8,47	2,60	1,12	5,84	6,09
HT29 CSC 2.5mM VPA3	0,58	0,39	0,49	0,56	1,31
HT29 CSC 5mM VPA1	0,18	1,79	19,85	0,19	0,50
HT29 CSC 5mM VPA2	0,27	0,77	0,65	1,79	1,16
HT29 CSC 5mM VPA3	0,80	3,59	2,48	1,13	0,72
HT29 CSC DMSO1	0,59	0,64	0,03	0,40	0,24
HT29 CSC DMSO2	1,01	0,86	0,07	0,28	1,47
HT29 CSC DMSO3	0,19	6,97	0,01	0,10	0,18
HT29 CSC 100uM ZEB1	0,76	0,11	0,07	0,63	1,27
HT29 CSC 100uM ZEB2	0,78	1,20	0,07	0,05	10,42
HT29 CSC 100uM ZEB3	0,49	0,61	0,13	0,28	0,30
HT29 CSC 250uM ZEB1	0,50	0,27	0,42	0,69	0,94
HT29 CSC 250uM ZEB2	0,24	1,01	0,28	1,21	5,40
HT29 CSC 250uM ZEB3	4,61	0,26	1,57	0,10	0,22
HT29 CSC 500uM ZEB1	0,23	0,24	0,94	1,08	0,75
HT29 CSC 500uM ZEB2	2,47	2,13	0,46	0,37	2,97
HT29 CSC 500uM ZEB3	0,44	1,18	0,15	0,33	0,12

Key: NT=No treatment, VPA=Valproic acid, ZEB=Zebularine, DMSO=Dimethyl sulphoxide.

C.16 Gene Expression ($\log_{10}{}^{2-\Delta\Delta Cq}$) in DLD1 putative CSCs

Table C16: *ACTB*, *HPRT*, *NANOG*, *OCT4* and *YWHAZ* gene expression in DLD1 putative CSCs.

Sample	***BA***	***HPRT***	***NANOG***	***OCT4***	***YWHAZ***
DLD1 CSC NT1	0,00	0,00	-0,19	-0,10	0,00
DLD1 CSC NT2	0,07	-0,16	-0,15	-0,40	0,09
DLD1 CSC NT3	-0,05	-0,04	0,04	0,41	0,10
DLD1 CSC 1mM VPA1	0,06	0,15	-0,11	-0,01	-0,21
DLD1 CSC 1mM VPA2	0,07	-0,11	-0,02	-0,08	0,04
DLD1 CSC 1mM VPA3	0,07	-0,10	0,27	0,31	0,03
DLD1 CSC 2.5mM VPA	0,19	-0,11	-0,26	-0,21	-0,08
DLD1 CSC 2.5mM VPA2	0,07	-0,12	-0,42	-0,34	0,05
DLD1 CSC 2.5mM VPA3	0,02	-0,17	0,56	0,11	0,14
DLD1 CSC 5mM VPA1	0,12	-0,13	-0,25	-0,16	0,01
DLD1 CSC 5mM VPA2	0,04	-0,13	0,16	0,13	0,08
DLD1 CSC 5mM VPA3	-0,02	-0,10	0,59	0,07	0,12
DLD1 CSC DMSO1	-0,06	0,10	-0,27	-0,36	-0,04
DLD1 CSC DMSO2	-0,37	0,16	-0,71	-0,51	0,22
DLD1 CSC DMSO3	-0,14	0,06	0,26	0,19	0,08
DLD1 CSC 100uM ZEB1	-0,06	0,15	0,02	-0,26	-0,09
DLD1 CSC 100uM ZEB2	0,06	0,15	0,24	-0,46	-0,22
DLD1 CSC 100uM ZEB3	0,03	-0,01	0,25	0,31	-0,01
DLD1 CSC 250uM ZEB1	0,00	0,03	0,32	-0,22	-0,04
DLD1 CSC 250uM ZEB2	-0,06	0,18	-0,01	-0,30	-0,12
DLD1 CSC 250uM ZEB3	-0,03	0,00	0,48	0,38	0,03
DLD1 CSC 500uM ZEB1	0,12	-0,02	0,15	0,32	-0,10
DLD1 CSC 500uM ZEB2	0,01	-0,17	0,18	-0,11	0,16
DLD1 CSC 500uM ZEB3	0,03	-0,03	0,36	0,23	0,00

Key: NT=No treatment, VPA=Valproic acid, ZEB=Zebularine, DMSO=Dimethyl sulphoxide.

C.17 Standard Error in Gene Expression ($log_{10}{}^{2-\Delta\Delta Cq}$) in DLD1 putative CSCs

Table C17: Standard Error in *ACTB*, *HPRT*, *NANOG*, *OCT4* and *YWHAZ* gene expression in DLD1 putative CSCs.

Sample	***ACTB***	***HPRT***	***NANOG***	***OCT4***	***YWHAZ***
DLD1 CSC NT1	7,12	1,97	0,16	0,06	4,94
DLD1 CSC NT2	0,54	0,40	0,22	0,05	0,60
DLD1 CSC NT3	0,47	1,39	2,68	0,15	0,51
DLD1 CSC 1mM VPA1	0,45	0,33	0,47	26,12	0,11
DLD1 CSC 1mM VPA2	0,32	0,06	1,92	0,09	0,53
DLD1 CSC 1mM VPA3	1,81	0,57	0,65	0,30	1,57
DLD1 CSC 2.5mM VPA1	0,53	1,27	0,15	0,18	0,79
DLD1 CSC 2.5mM VPA2	0,17	0,02	0,04	0,04	0,07
DLD1 CSC 2.5mM VPA3	3,30	0,10	0,15	0,38	0,24
DLD1 CSC 5mM VPA1	0,35	0,10	0,04	0,08	5,36
DLD1 CSC 5mM VPA2	1,02	0,15	0,87	0,68	1,05
DLD1 CSC 5mM VPA3	2,64	0,37	1,33	1,86	0,39
DLD1 CSC DMSO1	1,09	0,34	0,19	0,05	1,05
DLD1 CSC DMSO2	0,06	0,20	0,01	0,06	0,35
DLD1 CSC DMSO3	0,30	0,99	0,48	0,42	1,28
DLD1 CSC 100uM ZEB1	0,06	0,06	1,06	0,04	0,12
DLD1 CSC 100uM ZEB2	1,08	0,27	0,62	0,03	0,21
DLD1 CSC 100uM ZEB3	3,51	1,84	0,29	0,52	2,40
DLD1 CSC 250uM ZEB1	39,31	0,96	0,36	0,09	1,11
DLD1 CSC 250uM ZEB2	0,33	0,52	9,18	0,04	0,22
DLD1 CSC 250uM ZEB3	2,59	13,85	0,50	0,49	2,13
DLD1 CSC 500uM ZEB1	0,42	1,62	0,84	0,20	0,72
DLD1 CSC 500uM ZEB2	4,49	0,43	0,37	0,70	0,65
DLD1 CSC 500uM ZEB3	2,92	1,27	0,21	0,37	22,40

Key: NT=No treatment, VPA=Valproic acid, ZEB=Zebularine, DMSO=Dimethyl sulphoxide.

C.18 Statistical Analyses of Gene Expression Profiling

Table C18: Kruskal-Wallis and Mann-Whitney tests of *NANOG* and *OCT4* mRNA expression in HT29 and DLD1 putative CSCs.

	Statistical Test	**HT29 CSCs**	**DLD1 CSCs**
NANOG			
NT *versus* VPA	Kruskal-Wallis	p=0.204	p=0.621
NT *versus* 1mM VPA	Mann-Whitney	p=0.080	p=0.382
NT *versus* 2.5mM VPA	Mann-Whitney	p=0.382	p=0.662
NT *versus* 5mM VPA	Mann-Whitney	p=0.080	p=0.662
NT *versus* DMSO	Mann-Whitney	p=0.662	p=0.662
DMSO *versus* ZEB	Kruskal-Wallis	p=0.218	p=0.496
DMSO *versus* 100µM ZEB	Mann-Whitney	p=1.000	p=0.662
DMSO *versus* 250µM ZEB	Mann-Whitney	p=0.080	p=0.190
DMSO *versus* 500µM ZEB	Mann-Whitney	p=0.080	p=0.382
OCT4			
NT *versus* VPA	Kruskal-Wallis	*p=0.037	p=0.679
NT *versus* 1mM VPA	Mann-Whitney	p=0.080	p=0.662
NT *versus* 2.5mM VPA	Mann-Whitney	p=0.080	p=1.000
NT *versus* 5mM VPA	Mann-Whitney	p=1.000	p=1.000
NT *versus* DMSO	Mann-Whitney	p=1.000	p=0.662
DMSO *versus* ZEB	Kruskal-Wallis	p=0.739	p=0.361
DMSO *versus* 100µM ZEB	Mann-Whitney	p=0.662	p=0.662
DMSO *versus* 250µM ZEB	Mann-Whitney	p=1.000	p=0.382
DMSO *versus* 500µM ZEB	Kruskal-Wallis	p=0.662	p=0.190

*$p \leq 0.05$ is statistically significant; Key: NT=No treatment, VPA=Valproic acid, ZEB=Zebularine, DMSO=Dimethyl sulphoxide.

C.19 SOX2 Splice Variants

Table C19: SOX2 transcript and protein accession numbers (Flicek *et al.*, 2014).

Name	Transcript ID	Length	Protein ID	Length	CCDS
SOX2-001	ENST00000325404	2500bp	ENSP00000323588	317aa	CCDS3239
SOX2-201	ENST00000431565	1318bp	ENSP00000439111	317aa	CCDS3239

Appendix D

Figures

D.1 Immunofluorescence Microscopy-CD133 and EpCAM Negative Staining Controls

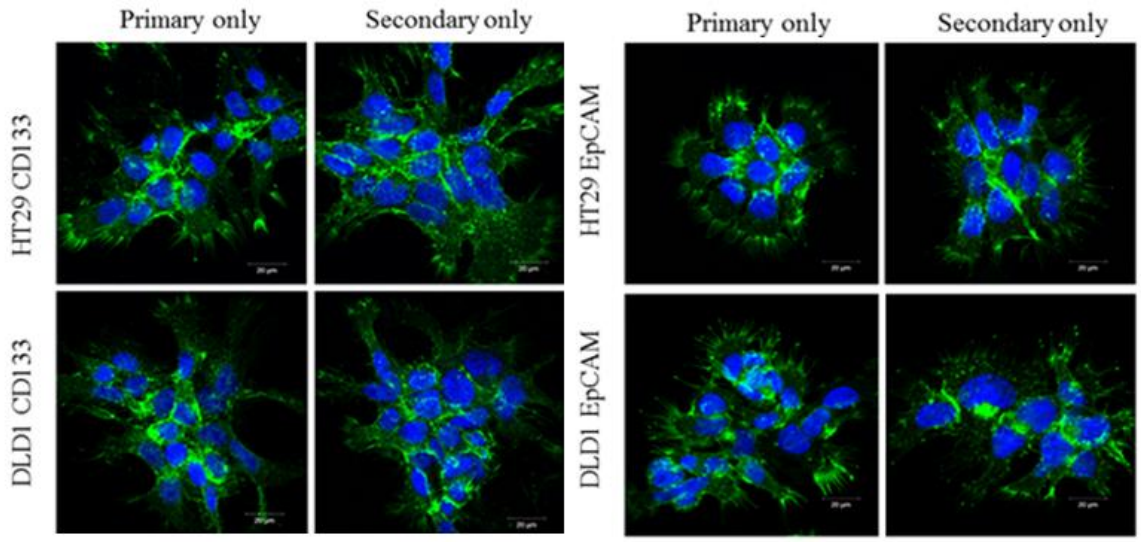

Figure D.1. Immunofluorescence images showing primary and secondary antibody negative control staining for CD133 and EpCAM in HT29 and DLD1 putative CSCs. Cells were stained with Phalloidin-Alexa Fluor 488 conjugate (F-actin) indicated in green and DAPI (nuclei) indicated in blue. The cells are also stained with CD133 and EpCAM primary antibodies only, and with Alexa Fluor 594, donkey anti-rabbit IgG (H+L) and with Alexa Fluor 568 donkey anti-goat IgG, respectively. (Original magnification 63X).

D.2 Immunofluorescence Microscopy-Drug Treatments

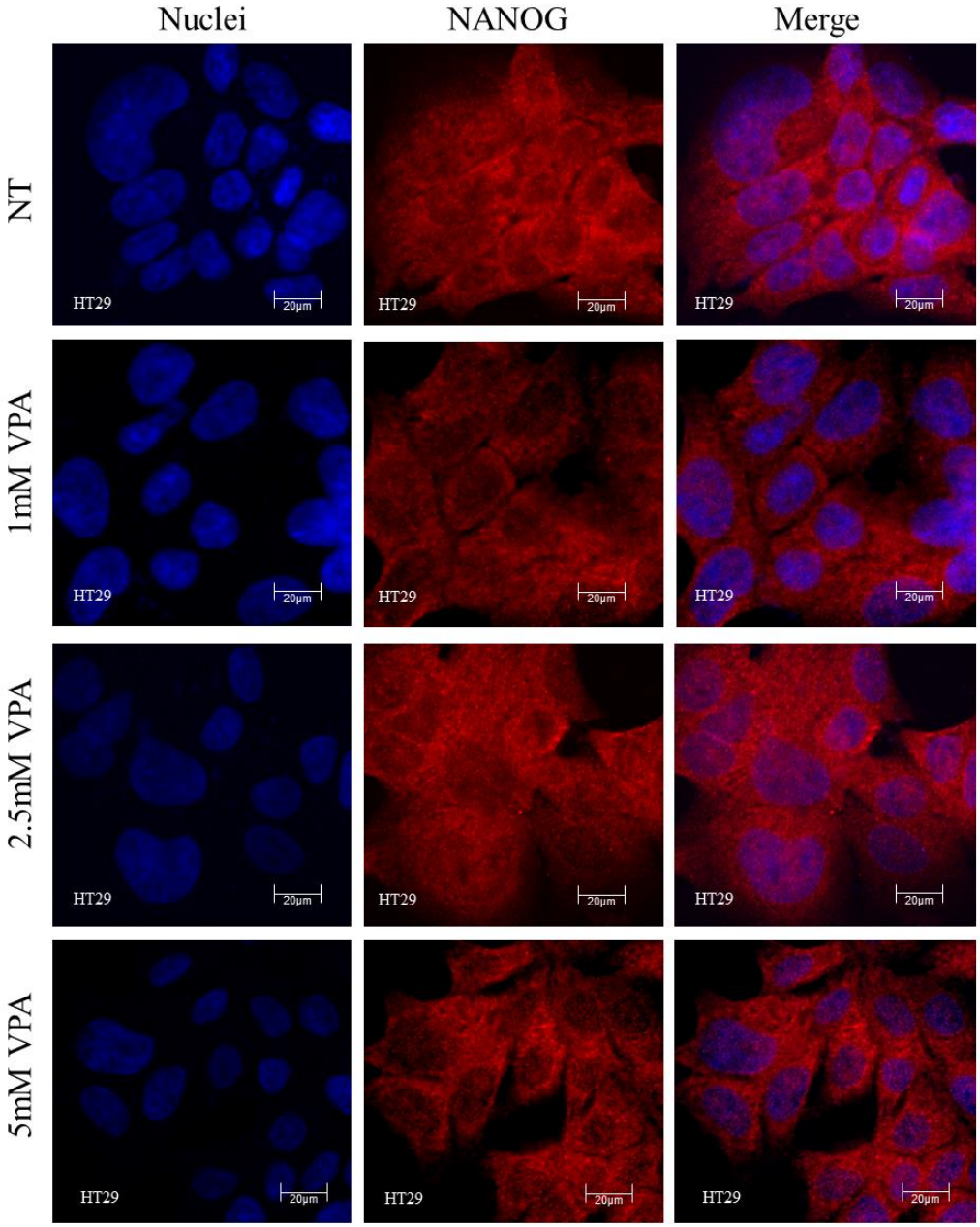

Figure D.2. Immunofluorescence images showing NANOG expression in HT29 putative CSCs treated with varying concentrations of VPA. Cells were stained with DAPI (nuclei) indicated in blue and NANOG detected with Alexa Fluor 568 donkey anti-goat IgG indicated in red. (Original magnification 63X) (N=1).

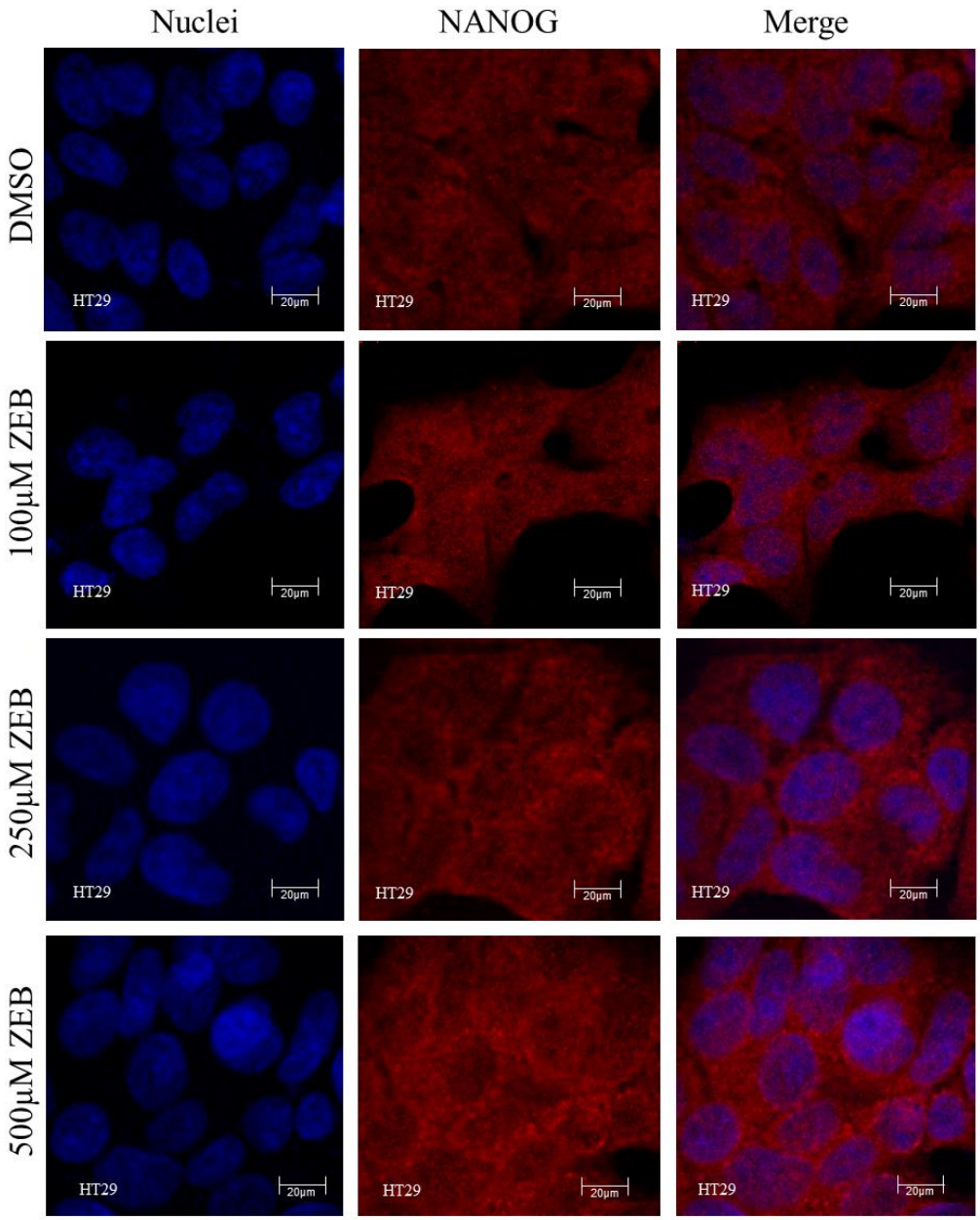

Figure D.3. Immunofluorescence images showing NANOG expression in HT29 putative CSCs treated with varying concentrations of ZEB. Cells were stained with DAPI (nuclei) indicated in blue and NANOG detected with Alexa Fluor 568 donkey anti-goat IgG indicated in red. (Original magnification 63X) (N=1).

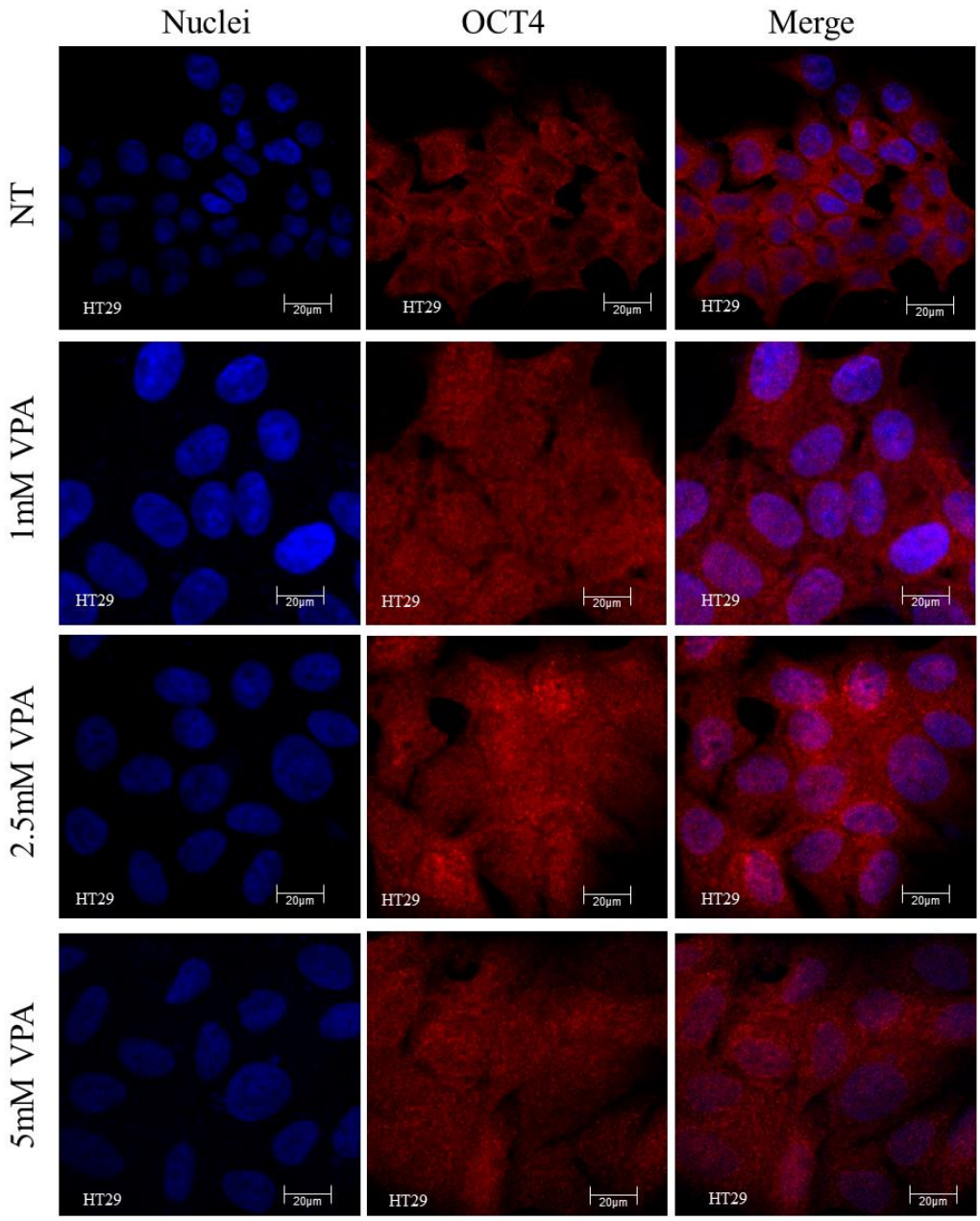

Figure D.4. Immunofluorescence images showing OCT4 expression in HT29 putative CSCs treated with varying concentrations of VPA. Cells were stained with DAPI (nuclei) indicated in blue and OCT4 detected with Alexa Fluor 568 donkey anti-goat IgG indicated in red. (Original magnification 63X) (N=1).

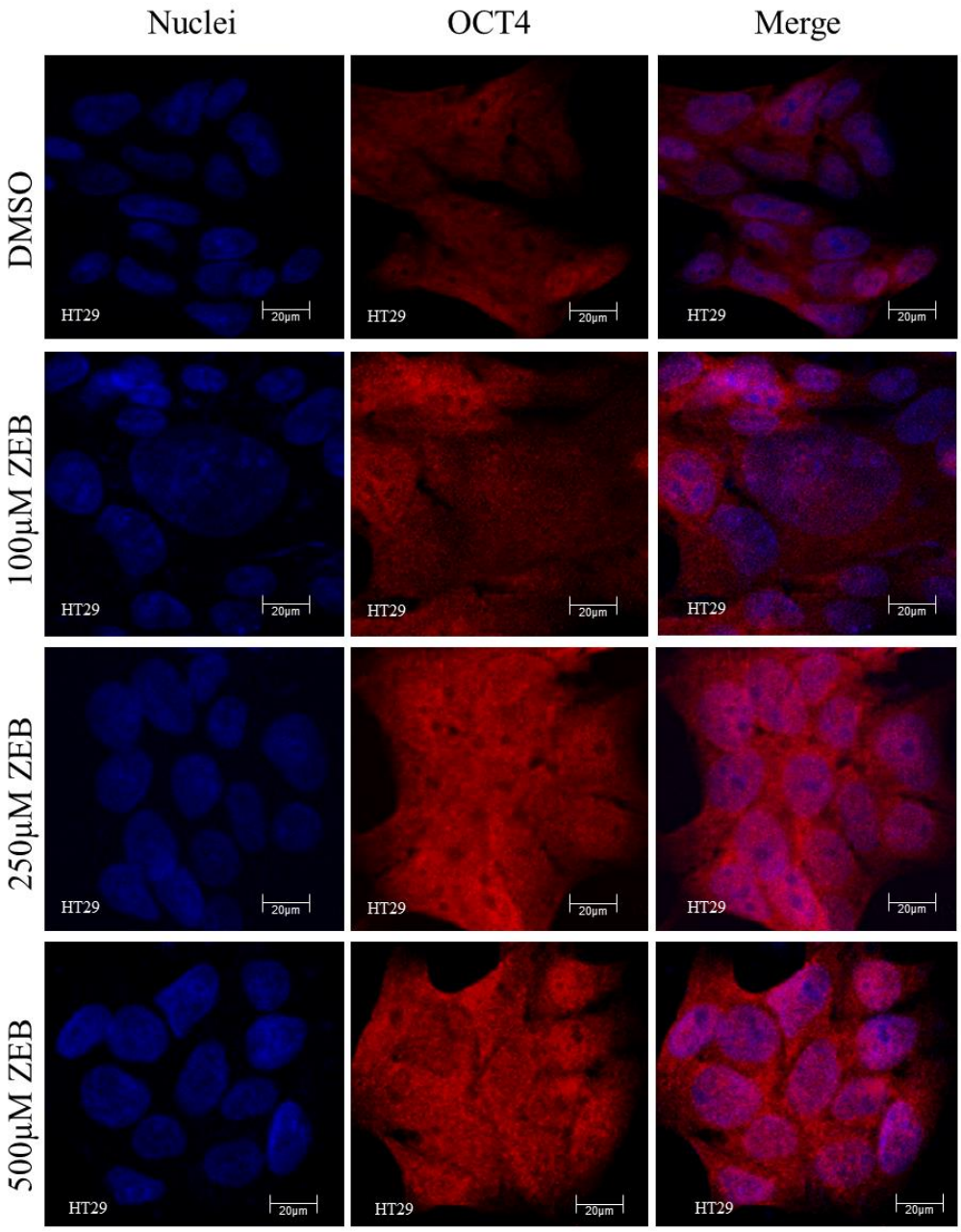

Figure D.5. Immunofluorescence images showing OCT4 expression in HT29 putative CSCs treated with varying concentrations of ZEB. Cells were stained with DAPI (nuclei) indicated in blue and OCT4 detected with Alexa Fluor 568 donkey anti-goat IgG indicated in red. (Original magnification 63X) (N=1).

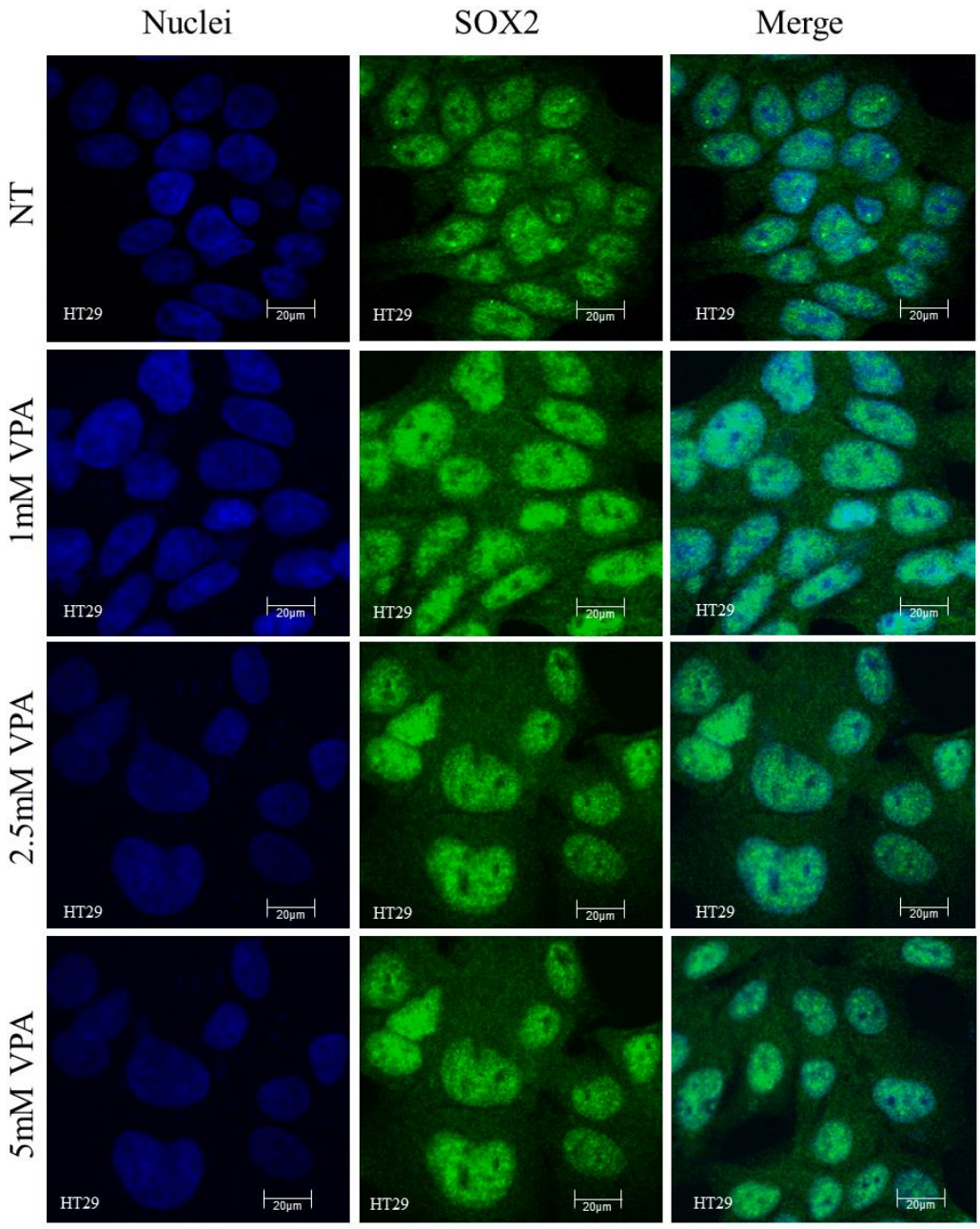

Figure D.6. Immunofluorescence images showing SOX2 expression in HT29 putative CSCs treated with varying concentrations of VPA. Cells were stained with DAPI (nuclei) indicated in blue and SOX2 detected with Alexa Fluor 488 donkey anti-mouse IgG indicated in green. (Original magnification 63X) (N=1).

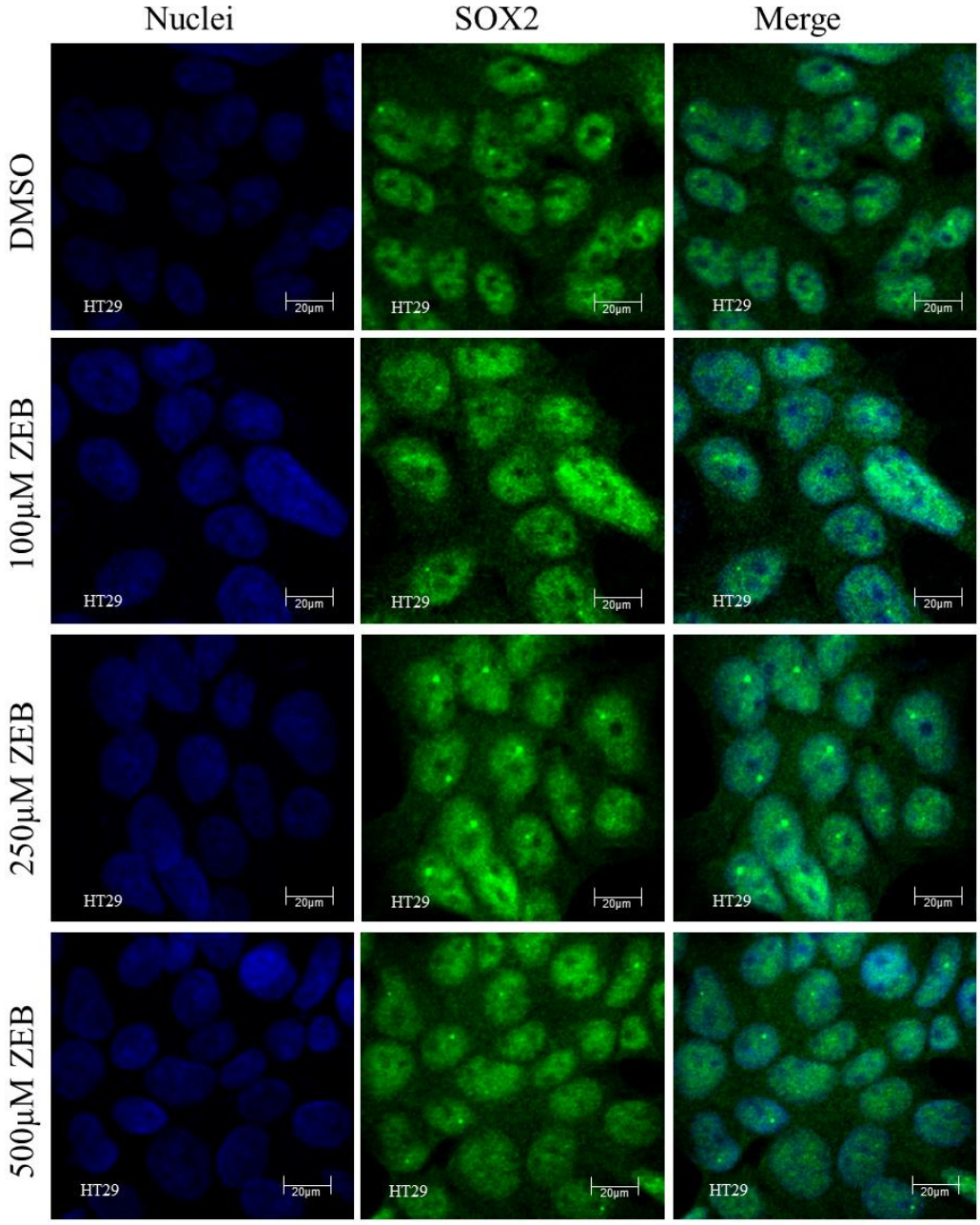

Figure D.7. Immunofluorescence images showing SOX2 expression in HT29 putative CSCs treated with varying concentrations of ZEB. Cells were stained with DAPI (nuclei) indicated in blue and SOX2 detected with Alexa Fluor 488 donkey anti-mouse IgG indicated in green. (Original magnification 63X) (N=1).

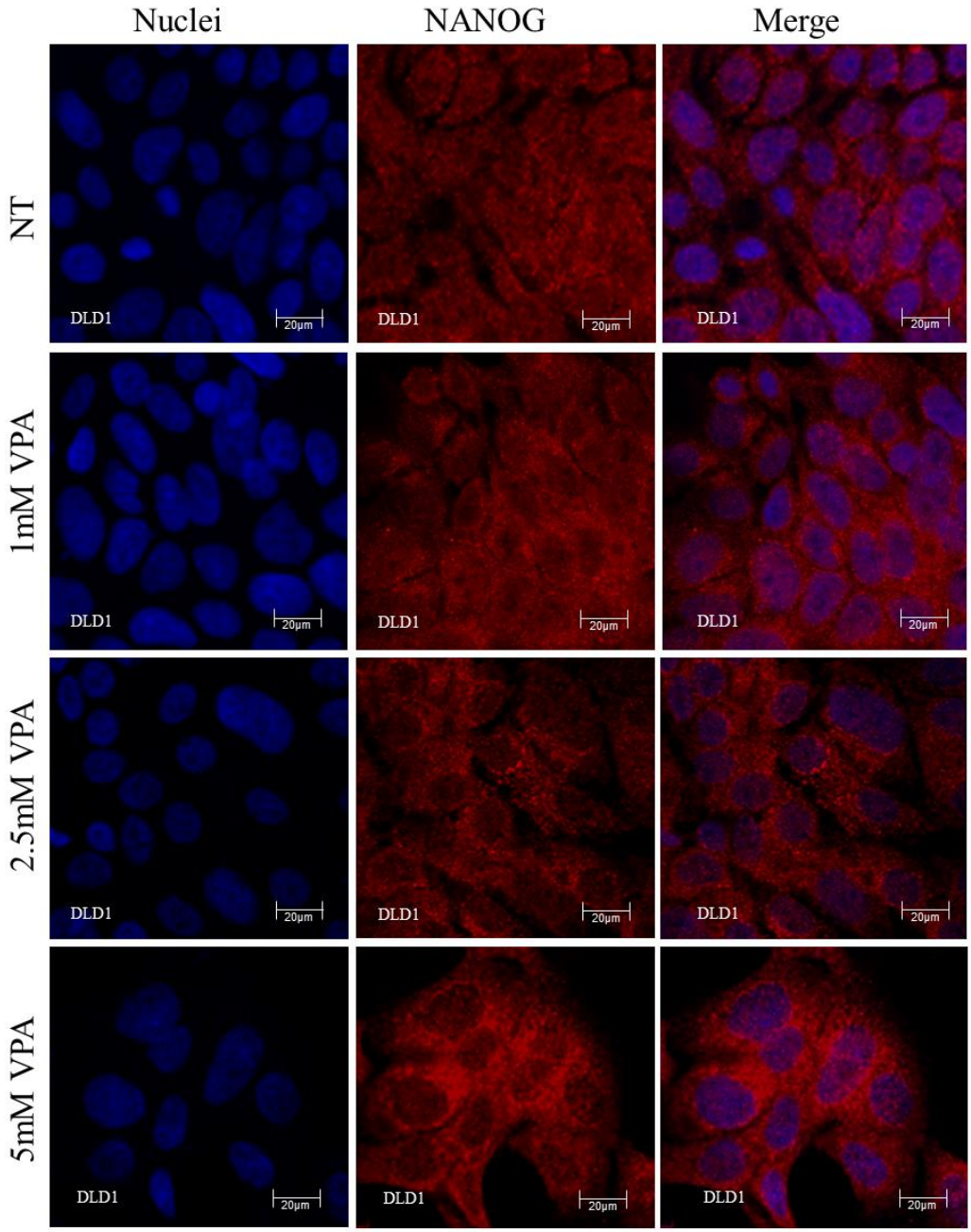

Figure D.8. Immunofluorescence images showing NANOG expression in DLD1 putative CSCs treated with varying concentrations of VPA. Cells were stained with DAPI (nuclei) indicated in blue and NANOG detected with Alexa Fluor 568 donkey anti-goat IgG indicated in red. (Original magnification 63X) (N=1).

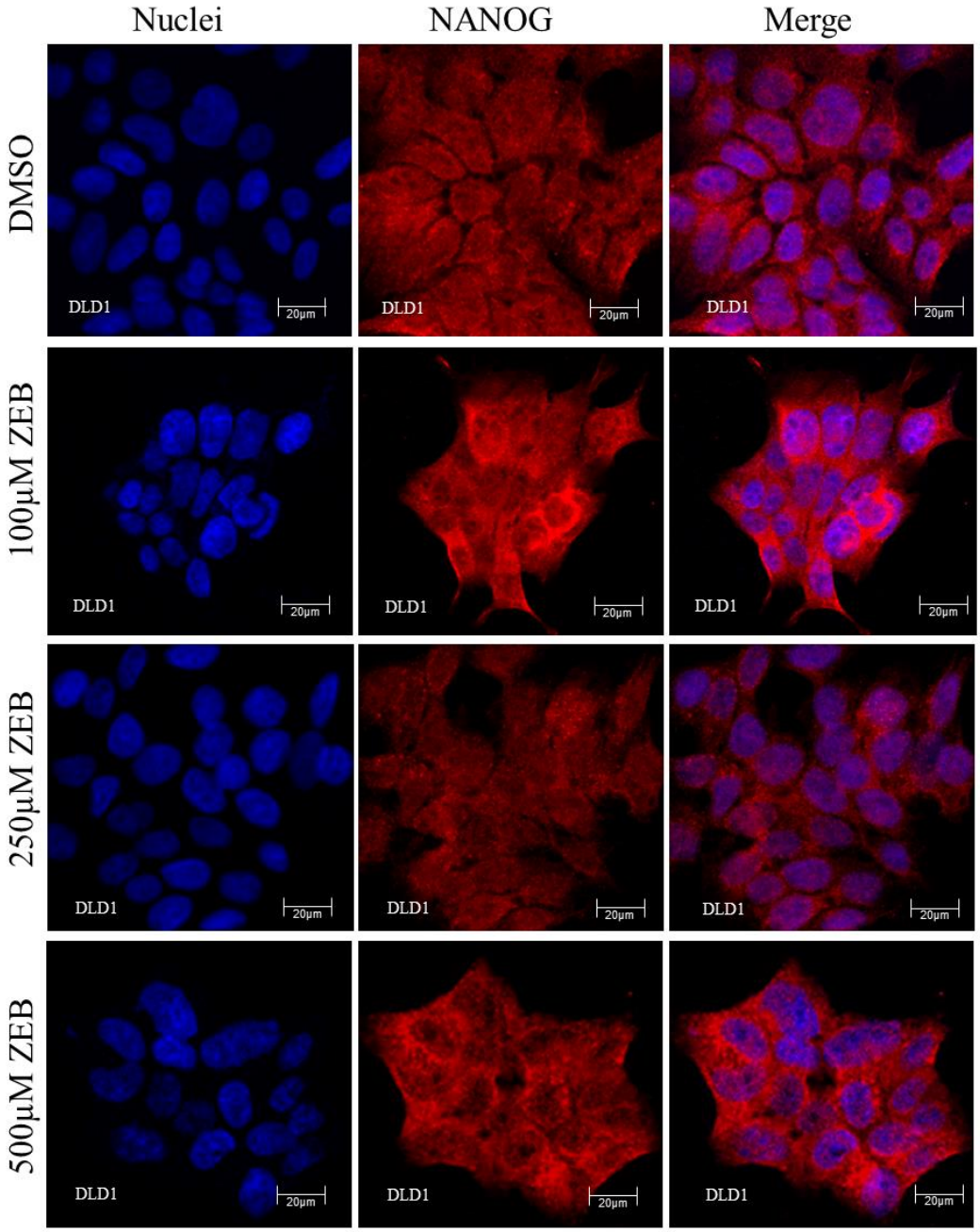

Figure D.9. Immunofluorescence images showing NANOG expression in DLD1 putative CSCs treated with varying concentrations of ZEB. Cells were stained with DAPI (nuclei) indicated in blue and NANOG detected with Alexa Fluor 568 donkey anti-goat IgG indicated in red. (Original magnification 63X) (N=1).

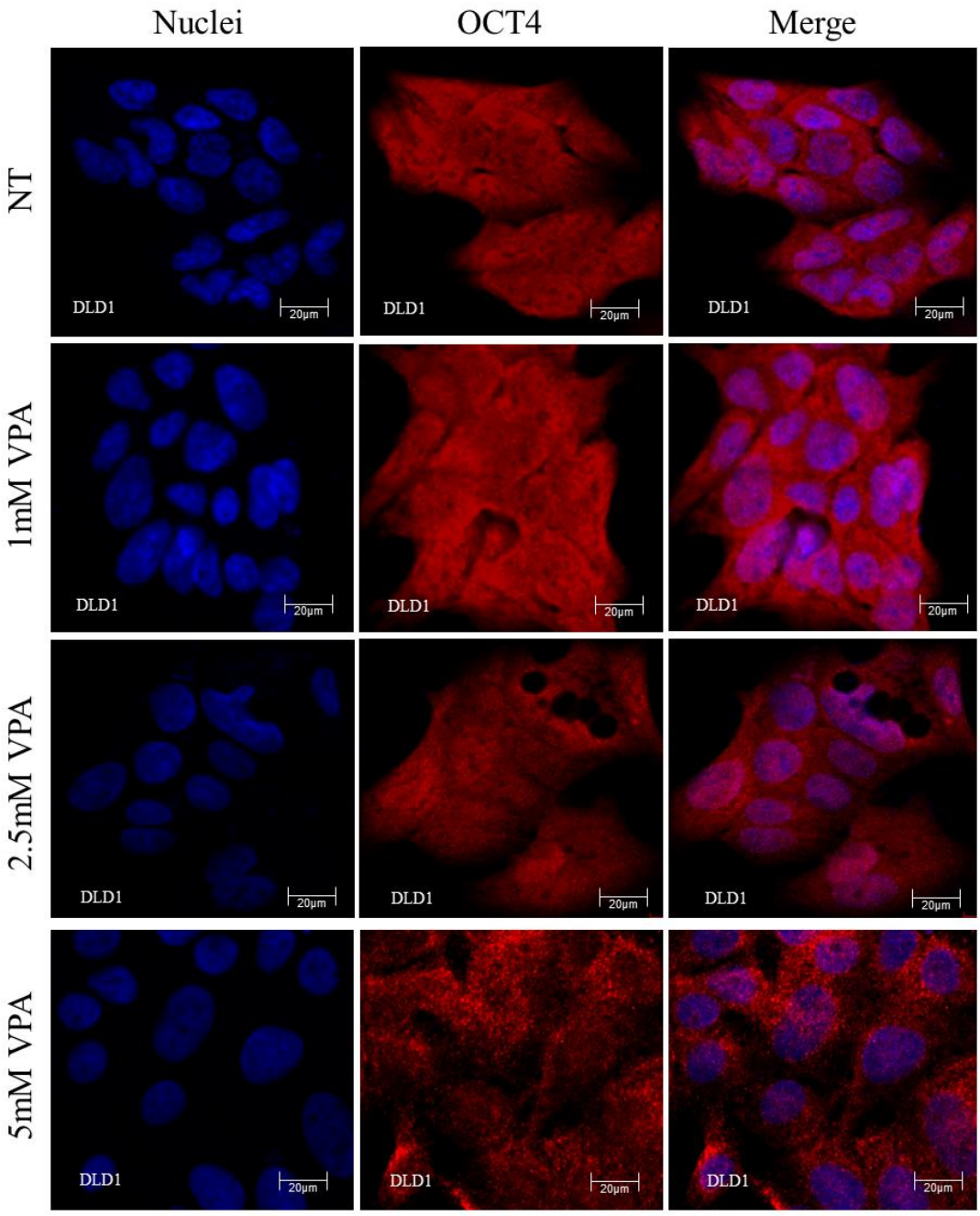

Figure D.10. Immunofluorescence images showing OCT4 expression in DLD1 putative CSCs treated with varying concentrations of VPA. Cells were stained with DAPI (nuclei) indicated in blue and OCT4 detected with Alexa Fluor 568 donkey anti-goat IgG indicated in red. (Original magnification 63X) (N=1).

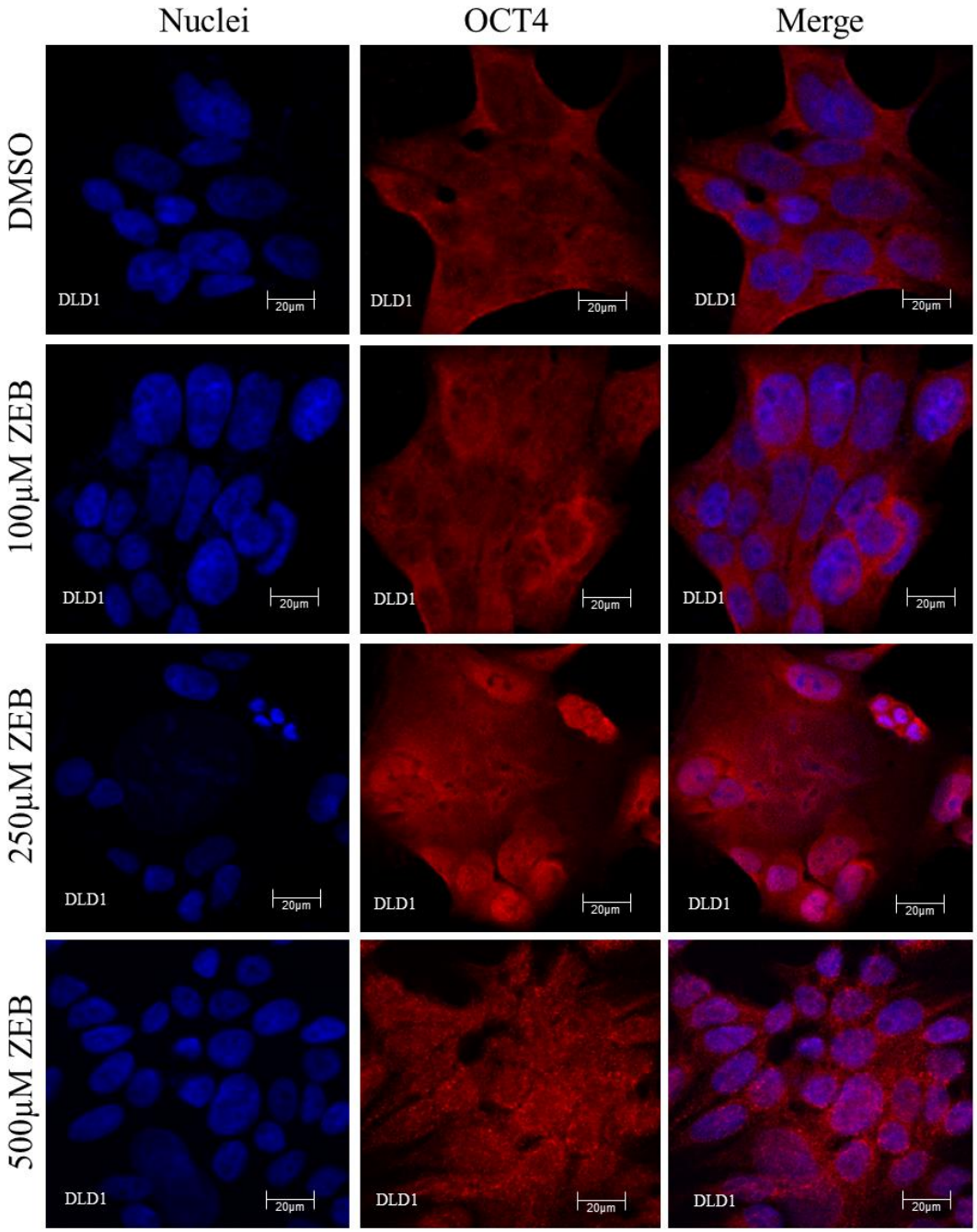

Figure D.11. Immunofluorescence images showing OCT4 expression in DLD1 putative CSCs treated with varying concentrations of ZEB. Cells were stained with DAPI (nuclei) indicated in blue and OCT4 detected with Alexa Fluor 568 donkey anti-goat IgG indicated in red. (Original magnification 63X) (N=1).

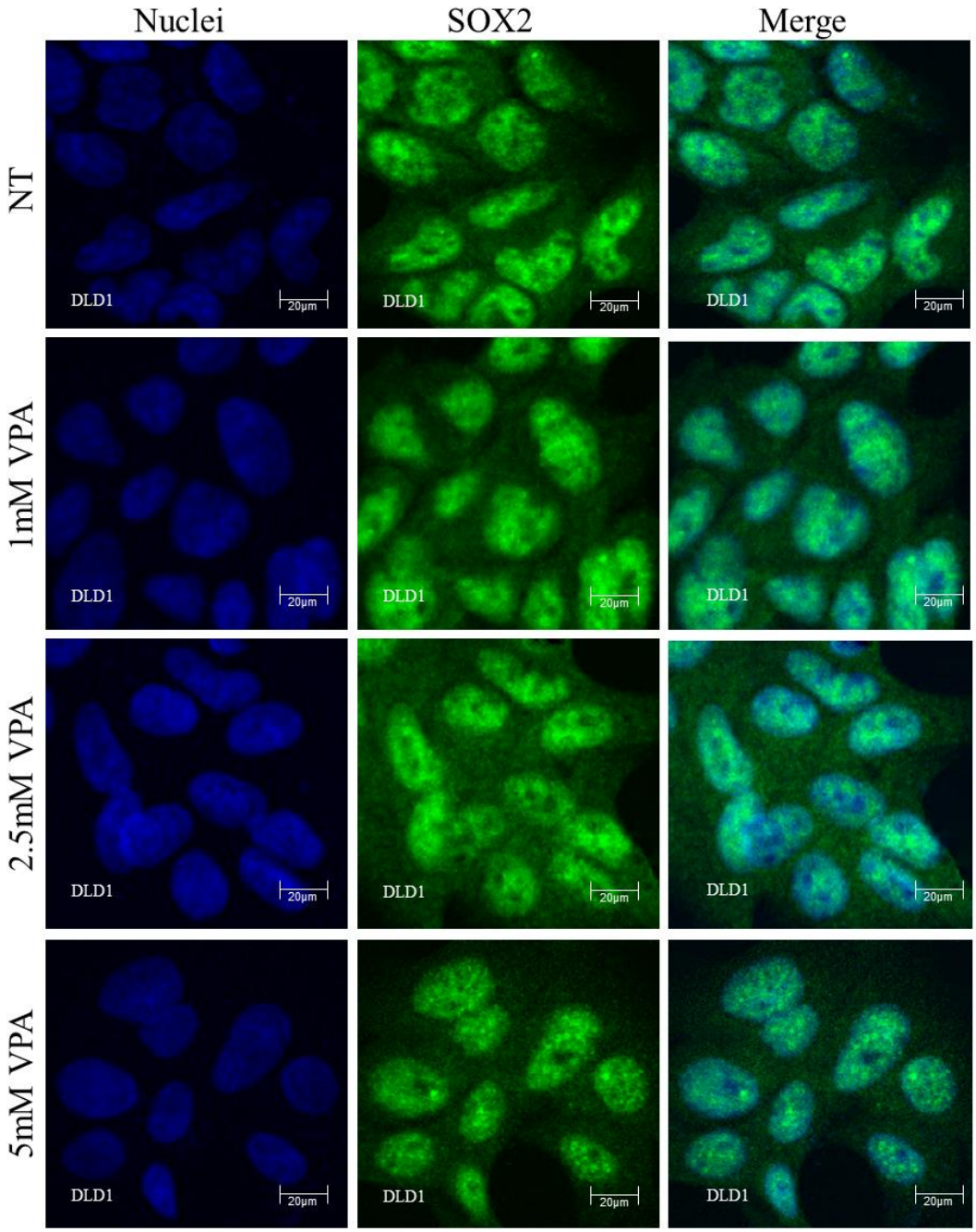

Figure D.12. Immunofluorescence images showing SOX2 expression in DLD1 putative CSCs treated with varying concentrations of VPA. Cells were stained with DAPI (nuclei) indicated in blue and SOX2 detected with Alexa Fluor 488 donkey anti-mouse IgG indicated in green. (Original magnification 63X) (N=1).

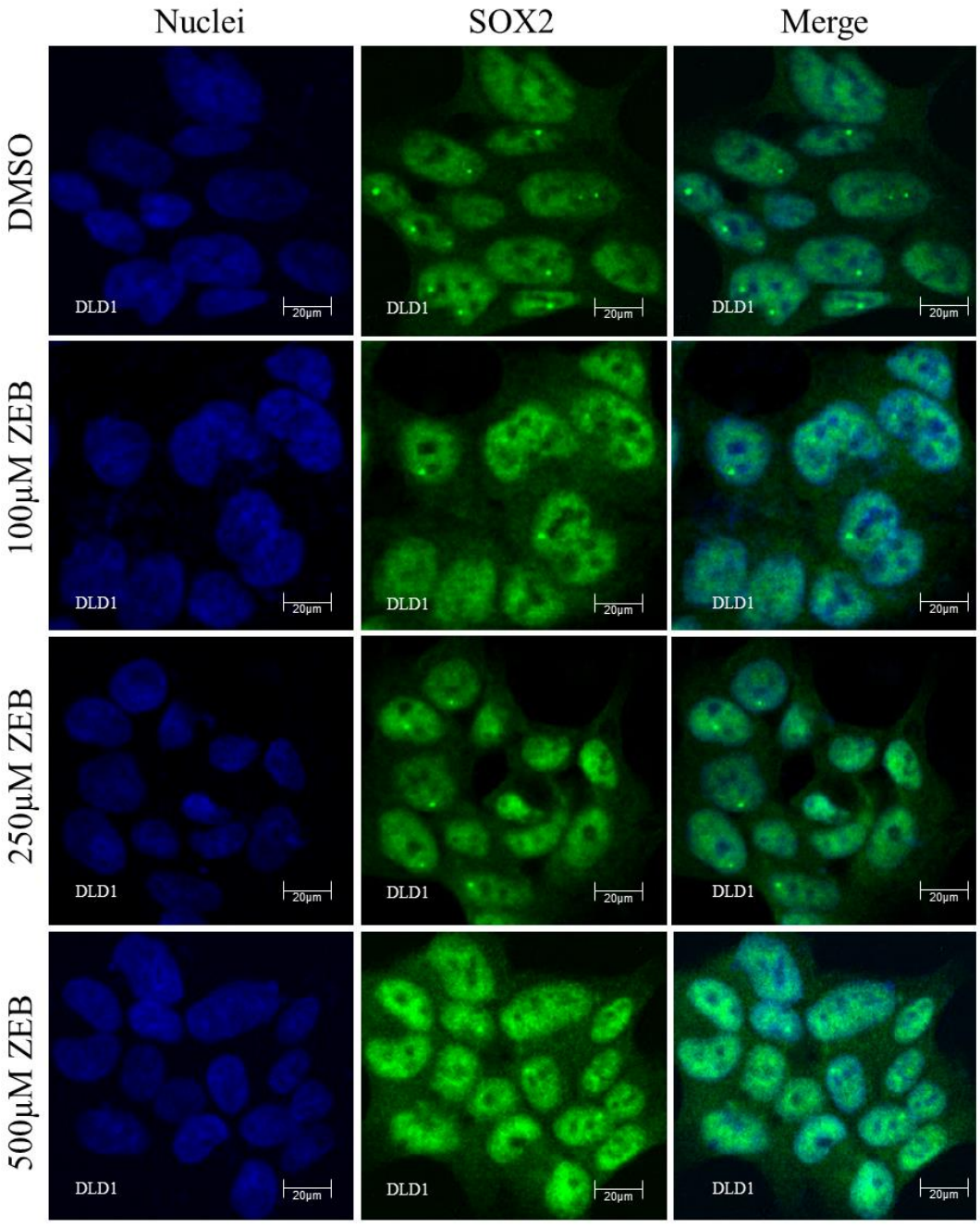

Figure D.13. Immunofluorescence images showing SOX2 expression in DLD1 putative CSCs treated with varying concentrations of ZEB. Cells were stained with DAPI (nuclei) indicated in blue and SOX2 detected with Alexa Fluor 488 donkey anti-mouse IgG indicated in green. (Original magnification 63X) (N=1).

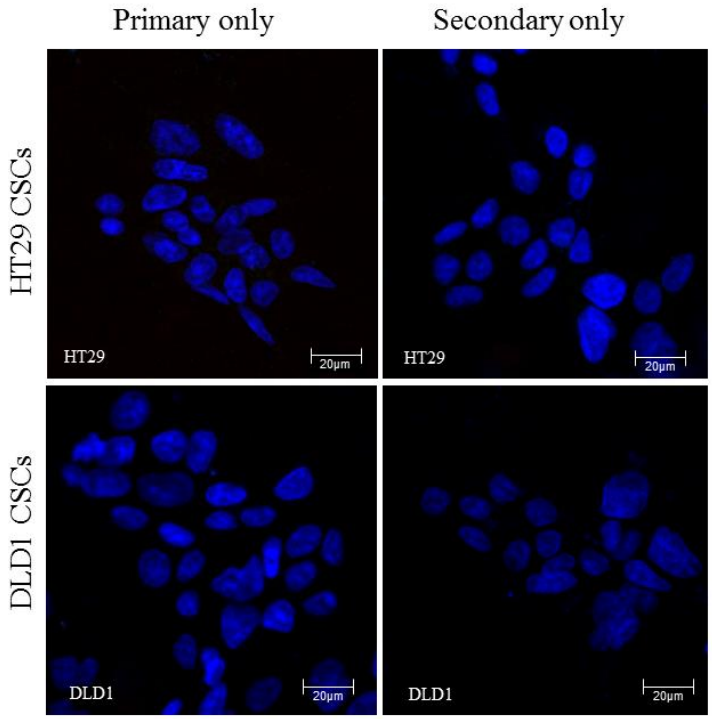

Figure D.14. Immunofluorescence images showing primary and secondary antibody control staining for NANOG, OCT4 and SOX2 in HT29 and DLD1 putative CSCs. Cells are stained with DAPI (nuclei) indicated in blue. The cells are also stained with a combination of NANOG, OCT4 and SOX2 primary antibodies only, and with Alexa Fluor 568, donkey anti-goat IgG and with Alexa Fluor 488 donkey anti-mouse IgG. (Original magnification 63X).

D.3 Mapped NANOG Primers

```
1    ATGAGTGTGGATCCAGCTTGTCCCCAAAGCTTGCCTTGCTTTGAAGCATCCGACTGTAAAGAATCTTCAC
71   CTATGCCTGTGATTTGTGGGCCTGAAGAAAACTATCCATCCTTGCAAATGTCTTCTGCTGAGATGCCTCA
141  CACGGAGACTGTCTCTCCTCTTCCTTCCTCCATGGATCTGCTTATTCAGGACAGCCCTGATTCTTCCACC
              F primer
211  AGTCCCAAAGGCAAACAACCCACTTCTGCAGAGAAGAGTGTCGCAAAAAAGGAAGACAAGGTCCCGGTCA
281  AGAAACAGAAGACCAGAACTGTGTTCTCTTCCACCCAGCTGTGTGTACTCAATGATAGATTTCAGAGACA
              R primer
351  GAAATACCTCAGCCTCCAGCAGATGCAAGAACTCTCCAACATCCTGAACCTCAGCTACAAACAGGTGAAG
421  ACCTGGTTCCAGAACCAGAGAATGAAATCTAAGAGGTGGCAGAAAAACAACTGGCCGAAGAATAGCAATG
491  GTGTGACGCAGAAGGCCTCAGCACCTACCTACCCCAGCCTTTACTCTTCCTACCACCAGGGATGCCTGGT
561  GAACCCGACTGGGAACCTTCCAATGTGGAGCAACCAGACCTGGAACAATTCAACCTGGAGCAACCAGACC
631  CAGAACATCCAGTCCTGGAGCAACCACTCCTGGAACACTCAGACCTGGTGCACCCAATCCTGGAACAATC
701  AGGCCTGGAACAGTCCCTTCTATAACTGTGGAGAGGAATCTCTGCAGTCCTGCATGCAGTTCCAGCCAAA
771  TTCTCCTGCCAGTGACTTGGAGGCTGCCTTGGAAGCTGCTGGGGAAGGCCTTAATGTAATACAGCAGACC
841  ACTAGGTATTTTAGTACTCCACAAACCATGGATTTATTCCTAAACTACTCCATGAACATGCAACCTGAAG
911  ACGTGTGA
```

Figure D.15. *NANOG* mRNA sequence showing the forward (F) and reverse (R) primer binding sites, amplicon size: 158bp (NCBI accession number: NM_024865.2).

D.4 Mapped OCT4 Primers

```
1    ATGGCGGGACACCTGGCTTCGGATTTCGCCTTCTCGCCCCCTCCAGGTGGTGGAGGTGATGGGCCAGGGG
71   GGCCGGAGCCGGGCTGGGTTGATCCTCGGACCTGGCTAAGCTTCCAAGGCCCTCCTGGAGGGCCAGGAAT
141  CGGGCCGGGGGTTGGGCCAGGCTCTGAGGTGTGGGGGATTCCCCCATGCCCCCCGCCGTATGAGTTCTGT
211  GGGGGGATGGCGTACTGTGGGCCCCAGGTTGGAGTGGGGCTAGTGCCCCAAGGCGGCTTGGAGACCTCTC
281  AGCCTGAGGGCGAAGCAGGAGTCGGGGTGGAGAGCAACTCCGATGGGGCCTCCCCGGAGCCCTGCACCGT
                                                          F primer
351  CACCCCTGGTGCCGTGAAGCTGGAGAAGGAGAAGCTGGAGCAAAACCCGGAGGAGTCCCAGGACATCAAA
421  GCTCTGCAGAAAGAACTCGAGCAATTTGCCAAGCTCCTGAAGCAGAAGAGGATCACCCTGGGATATACAC
                                    R primer
491  AGGCCGATGTGGGGCTCACCCTGGGGGTTCTATTTGGGAAGGTATTCAGCCAAACGACCATCTGCCGCTT
561  TGAGGCTCTGCAGCTTAGCTTCAAGAACATGTGTAAGCTGCGGCCCTTGCTGCAGAAGTGGGTGGAGGAA
631  GCTGACAACAATGAAAATCTTCAGGAGATATGCAAAGCAGAAACCCTCGTGCAGGCCCGAAAGAGAAAGC
701  GAACCAGTATCGAGAACCGAGTGAGAGGCAACCTGGAGAATTTGTTCCTGCAGTGCCCGAAACCCACACT
771  GCAGCAGATCAGCCACATCGCCCAGCAGCTTGGGCTCGAGAAGGATGTGGTCCGAGTGTGGTTCTGTAAC
841  CGGCGCCAGAAGGGCAAGCGATCAAGCAGCGACTATGCACAACGAGAGGATTTTGAGGCTGCTGGGTCTC
911  CTTTCTCAGGGGGACCAGTGTCCTTTCCTCTGGCCCCAGGGCCCCATTTTGGTACCCCAGGCTATGGGAG
981  CCCTCACTTCACTGCACTGTACTCCTCGGTCCCTTTCCCTGAGGGGGAAGCCTTTCCCCCTGTCTCCGTC
1051 ACCACTCTGGGCTCTCCCATGCATTCAAACTGA
```

Figure D.16 *OCT3/4* (*POU5F1*) mRNA sequence showing the forward (F) and reverse (R) primer binding sites, amplicon size: 127bp (NCBI accession number: NM_002701.5).

D.5 Mapped SOX2 Primers

```
1     ATGTACAACATGATGGAGACGGAGCTGAAGCCGCCGGGCCCGCAGCAAACTTCGGGGGGCGGCGGCGGCA
                                                F primer
71    ACTCCACCGCGGCGGCGGCCGGCGGCAACCAGAAAAACAGCCCGGACCGCGTCAAGCGGCCCATGAATGC
141   CTTCATGGTGTGGTCCCGCGGGCAGCGGCGCAAGATGGCCCAGGAGAACCCCAAGATGCACAACTCGGAG
                                                                  R primer
211   ATCAGCAAGCGCCTGGGCGCCGAGTGGAAACTTTTGTCGGAGACGGAGAAGCGGCCGTTCATCGACGAGG
281   CTAAGCGGCTGCGAGCGCTGCACATGAAGGAGCACCCGGATTATAAATACCGGCCCCGGCGGAAAACCAA
351   GACGCTCATGAAGAAGGATAAGTACACGCTGCCCGGCGGGCTGCTGGCCCCCGGCGGCAATAGCATGGCG
421   AGCGGGGTCGGGGTGGGCGCCGGCCTGGGCGCGGGCGTGAACCAGCGCATGGACAGTTACGCGCACATGA
491   ACGGCTGGAGCAACGGCAGCTACAGCATGATGCAGGACCAGCTGGGCTACCCGCAGCACCCGGGCCTCAA
561   TGCGCACGGCGCAGCGCAGATGCAGCCCATGCACCGCTACGACGTGAGCGCCCTGCAGTACAACTCCATG
631   ACCAGCTCGCAGACCTACATGAACGGCTCGCCCACCTACAGCATGTCCTACTCGCAGCAGGGCACCCCTG
701   GCATGGCTCTTGGCTCCATGGGTTCGGTGGTCAAGTCCGAGGCCAGCTCCAGCCCCCCTGTGGTTACCTC
771   TTCCTCCCACTCCAGGGCGCCCTGCCAGGCCGGGGACCTCCGGGACATGATCAGCATGTATCTCCCCGGC
841   GCCGAGGTGCCGGAACCCGCCGCCCCCAGCAGACTTCACATGTCCCAGCACTACCAGAGCGGCCCGGTGC
911   CCGGCACGGCCATTAACGGCACACTGCCCCTCTCACACATGTGA
```

Figure D.17. *SOX2* mRNA sequence showing the forward (F) and reverse (R) primer binding sites, amplicon size: 176bp (NCBI accession number: NM_003106.3).

D.6 Mapped ACTB Primers

```
1     ATGGATGATGATATCGCCGCGCTCGTCGTCGACAACGGCTCCGGCATGTGCAAGGCCGGCTTCGCGGGCG
71    ACGATGCCCCCCGGGCCGTCTTCCCCTCCATCGTGGGGCGCCCCAGGCACCAGGGCGTGATGGTGGGCAT
141   GGGTCAGAAGGATTCCTATGTGGGCGACGAGGCCCAGAGCAAGAGAGGCATCCTCACCCTGAAGTACCCC
                              F primer
211   ATCGAGCACGGCATCGTCACCAACTGGGACGACATGGAGAAAATCTGGCACCACACCTTCTACAATGAGC
281   TGCGTGTGGCTCCCGAGGAGCACCCCGTGCTGCTGACCGAGGCCCCCCTGAACCCCAAGGCCAACCGCGA
                                                                  R primer
351   GAAGATGACCCAGATCATGTTTGAGACCTTCAACACCCCAGCCATGTACGTTGCTATCCAGGCTGTGCTA
421   TCCCTGTACGCCTCTGGCCGTACCACTGGCATCGTGATGGACTCCGGTGACGGGGTCACCCACACTGTGC
491   CCATCTACGAGGGGTATGCCCTCCCCCATGCCATCCTGCGTCTGGACCTGGCTGGCCGGGACCTGACTGA
561   CTACCTCATGAAGATCCTCACCGAGCGCGGCTACAGCTTCACCACCACGGCCGAGCGGGAAATCGTGCGT
631   GACATTAAGGAGAAGCTGTGCTACGTCGCCCTGGACTTCGAGCAAGAGATGGCCACGGCTGCTTCCAGCT
701   CCTCCCTGGAGAAGAGCTACGAGCTGCCTGACGGCCAGGTCATCACCATTGGCAATGAGCGGTTCCGCTG
771   CCCTGAGGCACTCTTCCAGCCTTCCTTCCTGGGCATGGAGTCCTGTGGCATCCACGAAACTACCTTCAAC
841   TCCATCATGAAGTGTGACGTGGACATCCGCAAAGACCTGTACGCCAACACAGTGCTGTCTGGCGGCACCA
911   CCATGTACCCTGGCATTGCCGACAGGATGCAGAAGGAGATCACTGCCCTGGCACCCAGCACAATGAAGAT
981   CAAGATCATTGCTCCTCCTGAGCGCAAGTACTCCGTGTGGATCGGCGGCTCCATCCTGGCCTCGCTGTCC
1051  ACCTTCCAGCAGATGTGGATCAGCAAGCAGGAGTATGACGAGTCCGGCCCCTCCATCGTCCACCGCAAAT
1121  GCTTCTAG
```

Figure D.18. *ACTB* mRNA sequence showing the forward (F) and reverse (R) primer binding sites, amplicon size: 192bp (NCBI accession number: NM_001101.3).

D.7 Mapped HPRT Primers

```
1    ATGGCGACCCGCAGCCCTGGCGTCGTGATTAGTGATGATGAACCAGGTTATGACCTTGATTTATTTTGCA
71   TACCTAATCATTATGCTGAGGATTTGGAAAGGGTGTTTATTCCTCATGGACTAATTATGGACAGGACTGA
141  ACGTCTTGCTCGAGATGTGATGAAGGAGATGGGAGGCCATCACATTGTAGCCCTCTGTGTGCTCAAGGGG
211  GGCTATAAATTCTTTGCTGACCTGCTGGATTACATCAAAGCACTGAATAGAAATAGTGATAGATCCATTC
281  CTATGACTGTAGATTTTATCAGACTGAAGAGCTATTGTAATGACCAGTCAACAGGGGACATAAAAGTAAT
                                                                F primer
351  TGGTGGAGATGATCTCTCAACTTTAACTGGAAAGAATGTCTTGATTGTGGAAGATATAATTGACACTGGC
421  AAAACAATGCAGACTTTGCTTTCCTTGGTCAGGCAGTATAATCCAAAGATGGTCAAGGTCGCAAGCTTGC
          R primer
491  AGCTTGCTGGTGAAAAGGACC TTTGTTGGATTTGAAATTCCAGACAAGTT
561  TGTTGTAGGATATGCCCTTGACTATAATGAATACTTCAGGGATTTGAATCATGTTTGTGTCATTAGTGAA
631  ACTGGAAAAGCAAAATACAAAGCCTAA
```

Figure D.19. *HPRT* mRNA sequence showing the forward (F) and reverse (R) primer binding sites, amplicon size: 130bp (NCBI accession number: NM_000194.2).

D.8 Mapped YWHAZ Primers

```
1    CTTTCTCCTTCCCCTTCTTCCGGGCTCCCGTCCCGGCTCATCACCCGGCCTGTGGCCCACTCCCACCGCC
71   AGCTGGAACCCTGGGGACTACGACGTCCCTCAAACCTTGCTTCTAGGAGATAAAAAGAACATCCAGTCAT
141  GGATAAAAATGAGCTGGTTCAGAAGGCCAAACTGGCCGAGCAGGCTGAGCGATATGATGACATGGCAGCC
211  TGCATGAAGTCTGTAACTGAGCAAGGAGCTGAATTATCCAATGAGGAGAGGAATCTTCTCTCAGTTGCTT
281  ATAAAAATGTTGTAGGAGCCCGTAGGTCATCTTGGAGGGTCGTCTCAAGTATTGAACAAAAGACGGAAGG
351  TGCTGAGAAAAAACAGCAGATGGCTCGAGAATACAGAGAGAAAATTGAGACGGAGCTAAGAGATATCTGC
421  AATGATGTACTGTCTCTTTTGGAAAAGTTCTTGATCCCCAATGCTTCACAAGCAGAGAGCAAAGTCTTCT
491  ATTTGAAAATGAAAGGAGATTACTACCGTTACTTGGCTGAGGTTGCCGCTGGTGATGACAAGAAAGGGAT
561  TGTCGATCAGTCACAACAAGCATACCAAGAAGCTTTTGAAATCAGCAAAAAGGAAATGCAACCAACACAT
631  CCTATCAGACTGGGTCTGGCCCTTAACTTCTCTGTGTTCTATTATGAGATTCTGAACTCCCCAGAGAAAG
701  CCTGCTCTCTTGCAAAGACAGCTTTTGATGAAGCCATTGCTGAACTTGATACATTAAGTGAAGAGTCATA
771  CAAAGACAGCACGCTAATAATGCAATTACTGAGAGACAACTTGACATTGTGGACATCGGATACCCAAGGA
841  GACGAAGCTGAAGCAGGAGAAGGAGGGGAAAATTAACCGGCCTTCCAACTTTTGTCTGCCTCATTCTAAA
911  ATTTACACAGTAGACCATTTGTCATCCATGCTGTCCCACAAATAGTTTTTTGTTTACGATTTATGACAGG
981  TTTATGTTACTTCTATTTGAATTTCTATATTTCCCATGTGGTTTTTATGTTTAATATTAGGGGAGTAGAG
1051 CCAGTTAACATTTAGGGAGTTATCTGTTTTCATCTTGAGGTGGCCAATATGGGGATGTGGAATTTTTATA
                                    F primer
1121 CAAGTTATAAGTGTTTGGCATAGTACTTTTGGTACATTGTGGCTTCAAAAGGGCCAGTGTAAAACTGCTT
                                    R primer
1191 CCATGTCTAAGCAAAGAAAACTGCCTACATACTGGTTTGTCCTGGCGGGGAATAAAAGGGATCATTGGTT
1261 CCAGTCACAGGTGTAGTAATTGTGGGTACTTTAAGGTTTGGAGCACTTACAAGGCTGTGGTAGAATCATA
1331 CCCCATGGATACCACATATTAAACCATGTATATCTGTGGAATACTCAATGTGTACACCTTTGACTACAGC
1401 TGCAGAAGTGTTCCTTTAGACAAAGTTGTGACCCATTTTACTCTGGATAAGGGCAGAAACGGTTCACATT
1471 CCATTATTTGTAAAGTTACCTGCTGTTAGCTTTCATTATTTTTGCTACACTCATTTTATTTGTATTTAAA
1541 TGTTTTAGGCAACCTAAGAACAAATGTAAAAGTAAAGATGCAGGAAAAATGAATTGCTTGGTATTCATTA
1611 CTTCATGTATATCAAGCACAGCAGTAAAACAAAAACCCATGTATTTAACTTTTTTTTAGGATTTTTGCTT
1681 TTGTGATTTTTTTTTTTTTGATACTTGCCTAACATGCATGTGCTGTAAAAATAGTTAACAGGGAAATAAC
1751 TTGAGATGATGGCTAGCTTTGTTTAATGTCTTATGAAATTTTCATGAACAATCCAAGCATAATTGTTAAG
1821 AACACGTGTATTAAATTCATGTAAGTGGAATAAAAGTTTTATGAATGGACTTTTCAACTACTTTCTCTAC
1891 AGCTTTTCATGTAAATTAGTCTTGGTTCTGAAACTTCTCTAAAGGAAATTGTACATTTTTTGAAATTTAT
1961 TCCTTATTCCCTCTTGGCAGCTAATGGGCTCTTACCAAGTTTAAACACAAAATTTATCATAACAAAAATA
2031 CTACTAATATAACTACTGTTTCCATGTCCCATGATCCCCTCTCTTCCTCCCCACCCTGAAAAAAATGAGT
2101 TCCTATTTTTTCTGGGAGAGGGGGGGATTGATTAGAAAAAAATGTAGTGTGTTCCATTTAAAATTTTGGC
2171 ATATGGCATTTTCTAACTTAGGAAGCCACAATGTTCTTGGCCCATCATGACATTGGGTAGCATTAACTGT
2241 AAGTTTTGTGCTTCCAAATCACTTTTTGGTTTTTAAGAATTTCTTGATACTCTTATAGCCTGCCTTCAAT
2311 TTTGATCCTTTATTCTTTCTATTTGTCAGGTGCACAAGATTACCTTCCTGTTTTAGCCTTCTGTCTTGTC
2381 ACCAACCATTCTTACTTGGTGGCCATGTACTTGGAAAAAGGCCGCATGATCTTTCTGGCTCCACTCAGTG
2451 TCTAAGGCACCCTGCTTCCTTTGCTTGCATCCCACAGACTATTTCCCTCATCCTATTTACTGCAGCAAAT
2521 CTCTCCTTAGTTGATGAGACTGTGTTTATCTCCCTTTAAAACCCTACCTATCCTGAATGGTCTGTCATTG
2591 TCTGCCTTTAAAATCCTTCCTCTTTCTTCCTCCTCTATTCTCTAAATAATGATGGGGCTAAGTTATACCC
2661 AAAGCTCACTTTACAAAATATTTCCTCAGTACTTTGCAGAAAACACCAAACAAAAATGCCATTTTAAAAA
2731 AGGTGTATTTTTTCTTTTAGAATGTAAGCTCCTCAAGAGCAGGGACAATGTTTTCTGTATGTTCTATTGT
2801 GCCTAGTACACTGTAAATGCTCAATAAATATTGATGATGGGAGGCAGTGAGTCTTGATGATAAGGGTGAG
```

```
2871 AAACTGAAATCCCAAACACTGTTTTGTTGCTTGTTTTATTATGACCTCAGATTAAATTGGGAAATATTGG
2941 CCCTTTTGAATAATTGTCCCAAATATTACATTCAAATAAAAGTGCAATGGAGAAAAAAAAAAA
```

Figure D.20. *YWHAZ* mRNA sequence showing the forward (F) and reverse (R) primer binding sites, amplicon size: 94bp (NCBI accession number: NM_003406.3).

Appendix E

Ethics Waiver

University
of the Witwatersrand,
Johannesburg

Human Research Ethics Committee (Medical)
(formerly Committee for Research on Human Subjects (Medical))

Secretariat: Research Office, Room SH10005, 10th floor, Senate House • Telephone: +27 11 717-1234 • Fax: +27 11 339-5708
Private Bag 3, Wits 2050, South Africa

Ref: W-CJ-090317-3
17/03/2009

TO WHOM IT MAY CONCERN:

Waiver: This certifies that the following research does not require clearance from the Human Research Ethics Committee (Medical).

Investigator: **Ms B Milner and Dr C B Penny**

Project title: **Isolation and characterization of colon cancer stem cells.**

Reason: This is a wholly laboratory study using commercial cell lines from the American Culture Type Collection. There are no humans involved.

UNIVERSITY OF THE WITWATERSRAND
PROF PE CLEATON-JONES
HREC (MEDICAL)
2009 -03- 17
JOHANNESBURG

Professor Peter Cleaton-Jones
Chair: Human Research Ethics Committee (Medical)

copy: Anisa Keshav, Research Office, Senate House, Wits