GENERATION AND MICRORNA EXPRESSION PROFILING OF COLORECTAL CANCER CELL-DERIVED INDUCED PLURIPOTENT CANCER CELLS

By

MICHELE HIEW SOOK YUIN

A dissertation submitted to the Department of Pre-Clinical Science,

Faculty of Medicine and Health Sciences,

Universiti Tunku Abdul Rahman,

in partial fulfilment of the requirements for the degree of

Master of Medical Sciences

October 2016

ABSTRACT

GENERATION AND MICRORNA EXPRESSION PROFILING OF COLORECTAL CANCER CELL-DERIVED INDUCED PLURIPOTENT CANCER CELLS

Michele Hiew Sook Yuin

Colorectal cancer (CRC) is the third most common malignancies worldwide and in Malaysia. Suitable models for elucidating the multi-step CRC progression are lacking. Induced pluripotent stem (iPS) cell technology via ectopic expression of OCT4, SOX2, KLF4 and c-MYC (OSKM) enables generation of induced pluripotent cancer (iPC) models. MicroRNAs (miRNAs) are also known to facilitate such reprogramming events. In this study, we generated iPC cells derived from two CRC cell lines via retroviral transduction of OSKM. The four CRC-iPC clones analysed showed ESC-like cellular morphology, expressed multiple pluripotency markers and differentiated into the three germ layers demonstrating pluripotency. The spontaneously-differentiated post-induced pluripotent cancer (post-iPC) cells showed up-regulation of germ-layer markers, also demonstrating ability to differentiate into cell lineages of the three germ layers. Expression of eight selected pluripotency genes was shown by RT-PCR in the CRC-iPC and parental CRC cells, sharing similar expression patterns. However, quantitative real-time RT-PCR results showed down-regulation of the pluripotency genes

OCT4, *SOX2*, *KLF4*, *c-MYC*, *NANOG* and *REX1* in CRC-iPC cells, hinting limited pluripotency. Pluripotency gene expression was partially restored in post-iPC cells, indicating reversibility of pluripotency gene expression, and possible epigenetic regulation. Furthermore, down-regulation of the mesenchymal-epithelial-transition (MET) genes, *CDH1* and *OCLN*, and up-regulation of the epithelial-mesenchymal-transition (EMT) genes, *SNAI1* and *VIM*, in the CRC-iPC cells indicated an inclination to a more mesenchyme-like state in the iPC cells. MET/EMT gene expression was also generally reversible in post-iPCs. Taken together, the molecular phenotypes of CRC-iPCs suggest partial reprogramming in cancer cells. MicroRNA microarray analysis revealed 52 up-regulated and 50 down-regulated miRNAs in CRC-iPCs, targeting 403 putative genes; these genes were predicted to be involved in cell mobility and migration, possibly through targeting the TGF-β and PI3K-Akt signalling pathways. The established CRC-iPC cells may have only achieved a pre-iPSC state. The limited pluripotency of CRC-iPC cells may possibly be due to partial reprogramming. MicroRNA profile of the CRC-iPC cells also suggested a novel role of miRNA in promoting induced pluripotency by modulating EMT activation. Schemes are presented to link the predicted signalling pathways that enhance stem cell features and promote EMT transition in support of the proposed a mesenchyme-like state in CRC-iPC cells. The CRC-iPC clones established may be developed into cancer disease models for elucidation of CRC progression, and for use in novel drug discovery.

ACKNOWLEDGEMENTS

First and foremost, I wish to express my heartfelt gratitude towards my supervisor, Senior Professor Dr. Choo Kong Bung, for selflessly providing guidance and valuable advices throughout the course of my study. I would also like to thank my co-supervisor, Emeritus Professor Dr. Cheong Soon Keng for his support and valuable opinions on the study.

Furthermore, I wish to extend my sincere appreciation to Professor Dr. Huang Chiu-Jung (Chinese Culture University, Taiwan), Dr. Chong Kowit-Yu (Chang Gung University, Taiwan) and Dr. Shen Chia-Ning (Academia Sinica, Taiwan) for their honest comments and thorough discussion on the issues encountered in this study. A special thank you to Dr. Shigeki Sugii from A* STAR, Singapore Bioimaging Consortium (SBIC) for the generosity in providing the pMXs vectors required to conduct this research.

Faculty of Medicine and Health Sciences has been of great assistance throughout my study, and that I wish to express my appreciation. More importantly, this work is financially supported by University of Malaya HIR-MoE Grant Initiative, (UM.C/625/1/HIR/MOHE/CHAN/03, account number - A000003 – 50001), and UTARRF research grant (no. 6200/C83).

Lastly, supports and encouragements from my fellow labmates, Nguyen Phan Nguyen Nhi, Tai Lihui, Chai Kit Man, Cheng Han Ping and Vimalan Rengganaten, have motivated me throughout the study. Needless to mention, my family has supported me throughout the study and has motivated me to achieve my goal.

APPROVAL SHEET

This dissertation entitled **"GENERATION AND MICRORNA EXPRESSION PROFILING OF COLORECTAL CANCER CELL-DERIVED INDUCED PLURIPOTENT CANCER CELLS"** was prepared by MICHELE HIEW SOOK YUIN and submitted as partial fulfillment of the requirements for the degree of of Medical Sciences at Universiti Tunku Abdul Rahman.

Approved by:

_____

(Senior Prof. Dr. Choo Kong Bung) Date: ……
Senior Professor/Supervisor
Department of Pre-clinical Sciences
Faculty of Medicine and Health Sciences
Universiti Tunku Abdul Rahman

_____

(Emeritus Prof. Dr. Cheong Soon Keng) Date: ……
Senior Professor/Co-supervisor
Department of Medicine
Faculty of Medicine and Health Sciences
Universiti Tunku Abdul Rahman

FACULTY OF MEDICINE AND HEALTH SCIENCES

UNIVERSITY TUNKU ABDUL RAHMAN

Date: _____

SUBMISSION OF DISSERTATION

It is hereby certified that **Michele Hiew Sook Yuin** (ID no: **14UMM05797**) has completed this dissertation entitled "GENERATION AND MICRORNA EXPRESSION PROFILING OF COLORECTAL CANCER CELL-DERIVED INDUCED PLURIPOTENT CANCER CELLS" under the supervision of Senior Prof. Dr. Choo Kong Bung (Supervisor) from the Department of Pre-clinical Sciences, Faculty of Medicine and Health Sciences, and Emeritus Prof. Dr. Cheong Soon Keng (Co-Supervisor) from the Department of Medicine, Faculty of Medicine and Health Sciences.

I understand that University will upload softcopy of my dissertation in pdf format into UTAR Institutional Repository, which may be made accessible to UTAR community and public.

Yours truly,

_____

(Michele Hiew Sook Yuin)

DECLARATION

I MICHELE HIEW SOOK YUIN hereby declare that the dissertation is based on my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously or concurrently submitted for any other degree at UTAR or other institutions.

_____

MICHELE HIEW SOOK YUIN

Date: _____

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

15-PGDH	15-prostaglandin dehydrogenase
5-FU	5-fluorouracil
A260	Absorbance at 260nm wavelength
A280	Absorbance at 280nm wavelength
ABC	ATP-binding cassette
AFP	Alpha-fetoprotein
AML	Acute myeloid leukaemia
AP	Alkaline phosphatase
APC	Adenomatous polyposis coli
ATCC	American type culture collection
BAX	BCL2-associated X protein
BER	Base-excision repair
BMP	Bone morphogenetic protein
bp	Base pair
BSA	Bovine serum albumin
CD133	Epitope of Prominin-1
CDH1	E-cadherin
cDNA	Complementary DNA
CDX2	Coudal type homeobox protein 2
CK	Cytokeratin
c-MYC	V-myc avian myelocytomatosis viral oncogene homolog
CO_2	Carbon dioxide
COX2	Cyclooxygenase 2
CRC	Colorectal cancer
CRC-iPC	Colorectal cancer-derived induced pluripotent cancer cell
CSC	Cancer stem cell
Ct	Cycle threshold
DAPI	4',6-diamidino-2-phenylindole
DAVID	Database for Annotation, Visualization, and Integrated Discovery
ddH2O	Double distilled water
DEPC	Diethylpyrocarbonate
DGCR8	Di George syndrome critical region gene 8
DMEM	Dulbecco's Modified Eagle Medium
DMEM/F12	Dulbecco's Modified Eagle Eedium/Ham F-12
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucliec acid
DNMT	DNA methyltransferase
DNMT3B	DNA (cytosine-5-)-methyltransferase 3 beta
dNTP	Deoxynucleotide

EBs	Embryoid bodies
EDTA	Ethylenediaminetetraacetic acid
EGFR	Epidermal growth factor receptor
EMT	Epithelial-mesenchymal transition
EpCAM	Epithelial cell adhesion molecule
ESCC-miRNA	ESC-specific cell cycle-regulating miRNA
FA	Familial dysautonomia
FAP	Familial adenomatous polyposis
FBS	Fetal bovine serum
FBXW7	F-Box and WD repeat domain containing 7
FC	Fold change
FDR	False discovery rate
FGF	Fibroblast growth factor
FIT	Faecal immunochemical test
FITC	Fluoresceine isothiocynate
FOBT	Faecal occult blood tests
FS	Flexible sigmoidoscopy
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GATA6	GATA binding protein 6
GDF3	Growth differentiation factor 3
GFP	Green fluorescence protein
GO	Gene ontology
GSK3	Glycogen synthase kinase 3
HEK293T	Human embryonic kidney 293 cell
hESCs	Human embryonic stem cells
HNPCC	Hereditary nonpolyposis colorectal cancer
iPC	Induced pluripotent cancer cell
iPSC	Induced pluripotent stem cell
KEGG	Kyoto Encyclopedia of Genes and Genomes
KLF4	Kruppel-like factor 4
KOSR	Knockout serum replacement
LEFTY1	Left-right determination factor 1
L-OHP	Oxaliplatin
MAP2	Microtubule-associated protein 2
MEFs	Mouse embryonic fibroblasts
MEK	Mitogen-activated protein kinase
MEM	Minimum Essential Medium
MET	Mesenchymal-epithelial transition
miR-	Mature form of the miRNA
miRNA	MicroRNA
MLH1	Mutl homolog 1
MMR	Mismatch-repair
mRNA	Messenger ribonucleic acid
MRE11	homolog A, double strand break repair nuclease

MSH2	Muts homolog 2
MSX1	Msh homeobox 1
NANOG	Nanog homeobox
NCBI	National Center for Biotechnology Information
NF-κB	Nuclear factor-κb
O_2	Oxygen
OCLN	Occludin
OCT4	Pou class 5 homeobox 1
OSKM	OCT4, SOX2, KLF4, and c-MYC
p53	Tumour protein p53
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PDAC	Pancreatic ductal adenocarcinoma
PI3K-Akt	Phosphatidylinositol 3-kinase-Protein Kinase B
Post-iPC	Post-induced pluripotent cancer cell
pre-miRNA	Precursor microRNA
pri-miRNA	Primary microRNA
PTEN	Phosphatase and tensin homolog
qRT-PCR	Quantitative reverse transcription polymerase chain reaction
REX1	Zinc finger protein 42
RISC	RNA-induced silencing complex
RNA	Ribonucleic acid
RNU6	Ribonucleic acid, U6 small nuclear 6
ROCK inhibitor	Rho-associated-coiled-containing protein kinase inhibitor
rpm	Resolution per minute
rRNA	Ribosomal ribonucleic acid
RT	Reverse transcription
RT-PCR	Reverse transcription polymerase chain reaction
SD	Standard deviation
SMA	Spinal muscular atrophy
SMAD4	SMAD family member 4
SMN1	Survival motor neuron 1
SNAI1	Snail homolog 1
SOX2	Sry (sex determining region y)-box 2
SSEA4	Stage specific embryonic antigens 4
TBE	Tris /borate/ ethylenediaminetetraacetic acid (EDTA)
TET	Ten eleven translocation
TGFBR2	Transforming growth factor-β receptor type II
TGFβ	Transforming growth factor beta
UDG	Uracil-DNA glycosylase
UTR	Untranslated region
UV	Ultraviolet

VEGF	Vascular endothelial growth factor
VIM	Vimentin
WJ	Wharton's jelly
WJ-MSC	Wharton's Jelly-derived mesenchymal stem cell
ΔΔCt	Delta delta ct

CHAPTER 1

INTRODUCTION

Background

Colorectal cancer (CRC) is one of the commonest malignancies globally (Pourhoseingholi, 2012). In Malaysia, CRC is ranked the second highest cancer incidence (Zainal and Nor Saleha, 2011; Lim, 2014). Besides hereditary germline mutations, extrinsic risk factors of CRC include lifestyle and dietary factors (Lim, 2014). CRC patients receive adjuvant or palliative chemotherapy, while some undergo surgery, or combined simultaneously (Centelles, 2012). Owing to the multifactorial nature of CRC, chemotherapic treatments are often inefficient. Moreover, cancer stem cells (CSCs), a subset of the cancer cell population which shows stem cell-like features including self-renewal and ability to generate a heterogenous population, are believed to be the culprit of cancer relapses (Marotta and Polyak, 2009). CSCs that survive chemotherapy proliferate and acquire the ability to metastasis, forming tumours at secondary sites (Spillane and Henderson, 2007). Hence, a study model which reflects stem cell-like characteristics is needed for CRC disease modeling. The cancer stem cell model is now achievable via induction of pluripotency in cancer cells (Chen et al., 2012).

Induced pluripotency is achieved via a de-differentiation process, which enables terminally differentiated cells to re-acquire pluripotency; such cells are called induced pluripotent stem cells (iPSCs) (Takahashi and Yamanaka, 2006) Induced pluripotency is first established in somatic cells by forced expression of defined factors, OCT4, SOX2, KLF4, and c-MYC (OSKM), by Yamanaka's group (Takahashi and Yamanaka, 2006; Takahashi et al., 2007). The somatic cell-derived iPSC highly mimic embryonic stem cells (ESC) in cellular morphology, gene expression profile and the ability to differentiate into cell lineages of three germ layers (Huang, 2010). On the other hand, iPSCs of cancer origin, called induced pluripotent cancer (iPC) cells, also exhibit distinctive characteristics of pluripotent stem cells, including self-renewal and the ability to differentiate into the three germ layers (Ramos-Mejia et al., 2012). Distinctive from somatic cell-derived iPSCs, iPCs also simultaneously express cancer surface markers (Sun and Liu, 2011). Furthermore, diminished tumorigenicity of iPCs has been suggested for reprogrammed cancer cells, which has led to the postulation that iPCs with diminished tumourigenicity simulate early stages of cancer, (Kim and Zaret, 2015). Therefore, reprogrammed cancers may be further developed to serve as a tumourigenesis model, allowing recapitulation of important steps in cancer development for mechanistic studies, discovery of the early-stage markers and development of novel therapeutic approaches (Kim and Zaret, 2015).

MicroRNAs (miRNAs) have been implicated in maintaining self-renewal, pluripotency and regulating cell fate in pluripotent stem cells (Lüningschrör et al., 2013). MicroRNAs have also been successfully used to

reprogramme somatic and cancer cells (Subramanyam et al., 2011; Pourrajab et al., 2014). In order to further understand the role of miRNAs in cellular reprogramming, miRNA expression profiles of various pluripotent stem cells, including embryonic stem cells and somatic cell-derived iPSCs have been established (Huang et al., 2014; Zhao et al., 2014). The distinctive miRNA expression profiles in reprogrammed somatic cells have been functionally annotated (Lüningschrör et al., 2013). Many such miRNAs have also been mapped to signalling pathways that orchestrated the modulation of cellular reprogramming (Hawkins et al., 2014).

Problem statement

However, there are fewer studies of miRNA expression in reprogrammed cancer cells, and limited understanding of miRNA-regulated mechanisms underlying the cancer cell-reprogramming process in cancers. A fully elucidated mechanistic of cancer reprogramming would also help to improve reprogramming efficiency, by eliminating cancer-specific reprogramming barriers (Ebrahimi, 2015) and for developing iPCs into an ideal disease model.

Research questions

This study was aimed to answer the research questions as follows:

(1) Is colorectal cancer reprogrammable to a pluripotent state via OSKM-cellular reprogramming?

(2) What are the phenotypic and molecular features of reprogrammed CRC?

(3) What are the changes in genome-wide miRNA profiles in reprogrammed cancer cells?

Hypothesis of study

In this study, we hypothesised that the OSKM-transduced CRC-iPC cells were able to demonstrate stem cell-like characteristics and tri-lineage differentiation potentiality, as previously demonstrated by other studies with normal somatic cells (Lowry et al., 2008; Loh et al., 2009; Miyoshi et al., 2010). Upon reprogramming, differential expression of miRNAs would be observed, which were associated with reprogramming and pluripotency via regulating multiple signalling pathways.

Research objectives

The specific objectives of this study were:

1) To generate and characterise features of induced pluripotent cancer (iPC) cells derived from colorectal cancer (CRC) cell lines.

2) To establish the genome-wide miRNA expression profile of colorectal cancer-derived iPC (CRC-iPC) cell lines and parental CRC.
3) To identify and predict possible functional roles of differentially-expressed miRNAs in relation to reprogramming in the established CRC-iPCs relative to parental CRCs.

Significance of study

It is anticipated at the conclusion of this study that a number of induced pluripotent colorectal cancer cell (CRC-iPC) clones would have been established. The CRC-iPC clones would be useful for further development of disease models for elucidating progression of pathogenesis of colorectal cancer, and in the discovery of novel biomarkers and chemotherapeutic approaches. The miRNA and gene expression profiles obtained would shed new light on understanding the molecular mechanism(s) that govern cancer cell pluripotency.

CHAPTER 2

LITERATURE REVIEW

2.1 Colorectal Cancer

Colorectal cancer (CRC), the cancer of colon and rectum, is one of the commonest malignancies globally (Haggar and Boushey, 2009). CRC is the third most common cancer type and also the third highest cancer death in both genders (Siegel et al., 2016). High incidence of CRC is reported in most industrialised countries, including United States, Canada and Northwestern Europe (Center et al., 2010). The recent increased incidences among the Asian community are thought to reflect the adaptation of westernised sedentary lifestyle and dietary, and are also associated with smoking and obesity (Center et al., 2010). Interestingly, Chinese living in Southeast Asia, particularly in Malaysia and Singapore, have higher incidence rates of CRC than other ethnic groups (Pourhoseingholi, 2014). The observation suggested that ethnicity may be an important aetiological factor in CRC occurrence in Asia (Pourhoseingholi, 2012). The overall mortality of CRC has also increased in most Asian countries in the past decade (Pourhoseingholi, 2014). The survival of CRC is generally is stage-dependent; higher survival rate is observed in patients who receive treatments at early stages of CRC (Haggar and Boushey, 2009).

CRC is a multi-factorial disease attributed to inherited germline mutations and various extrinsic factors (Centelles, 2012). The extrinsic risk factors are age, dietary intake and lifestyle. CRC is more frequently diagnosed in patients after the age of forty, and drastically increases after age fifty (Haggar and Boushey, 2009). In addition, fats- and meat-rich diets produced carcinogenic compounds that have been associated with CRC (Johnson and Lund, 2007). The fats- and meat-rich diets, along with low physical activities have led to obesity, consequently increases the risk of CRC (Campbell et al., 2007). Heavy cigarette smoking and alcohol consumption introduce carcinogens that enhance formation of adenomatous polyps, and are clearly associated with early onset of CRC (Tsong et al., 2007).

Familial adenomatous polyposis (FAP) and hereditary nonpolyposis colorectal cancer (HNPCC) are the most commonly reported hereditary conditions (Jackson-Thompson et al., 2006). The mutated genes adenomatous polyposis coli (*APC*), mutL homolog 1 (*MLH1*) and mutS homolog 2 (*MSH2*) are inherited in an autosomal dominant manner (Lynch et al., 2009; Poulogiannis et al., 2010). Patients with the mutated *APC*, *MLH1* and *MSH2* carry a lifetime risk, 80% of whom develop CRC in middle age (Rogers et al., 2012).

Routine screening is believed to effectively detect early symptoms of CRC, and hence prevents the onset of the disease. Currently, there are three CRC routine screening approaches namely, flexible sigmoidoscopy (FS), total colonoscopy and faecal occult blood tests (FOBT) (Cummings and Cooper,

2011). Colonoscopy has been used as the gold standard in CRC screening due to convenient and immediate removal of polyps on detection in the same procedure (Brenner et al., 2010). However, colonoscopy may miss nonpolypoid colorectal neoplasms that is indistinguishable from the surrounding mucosa (Soetikno et al., 2008). FOBT is a less invasive method which detects peroxidase activity of heme in the stool. The FOBT screening method is less sensitive and may miss mild cases (Guittet et al. 2007). In an improved FOBT procedure, the faecal immunochemical test (FIT) utilises antibodies specific to globin, which increases sensitivity towards high-risk adenomas (Smith et al., 2006).

Once detected, common therapeutic and management approaches in CRC include surgery, radiotherapy and chemotherapy (Centelles 2012). Among the drugs used in CRC, 5-fluorouracil (5-FU) and oxaliplatin (L-OHP) are commonly used in CRC management (Jonker et al., 2006). 5-FU is an anti-metabolite drug which inhibits thymidylate synthase, resulting in the decrease of pyrimidine thymidine required for DNA replication (Centelles, 2012). On the other hand, L-OHP is a platinum compound which forms DNA adducts and thus blocks DNA replication and inhibits RNA synthesis (Alcindor and Beauger, 2011). In addition, adjuvant chemotherapy combining the two drugs is reported to increase the survival rate of CRC (Jonker et al., 2006). Development of novel drugs targeting vascular endothelial growth factor (VEGF) and epidermal growth factor receptor (EGFR) is promising to improve the survival rate of patients with metastatic CRC (Centelles, 2012).

However, the current treatment modalities are often ineffective on preventing CRC relapses due to the presence of colorectal cancer stem cells (CSC). The role of CSC in colorectal cancer is elaborated in the Literature Review Section 2.2.

2.1.1 Colorectal cancer in Malaysia

In Malaysia, CRC is ranked the second highest cancer incidence (Zainal and Nor Saleha, 2011; Lim, 2014). From 2008 to 2013, there were 4, 501 CRC cases reported, with incidence rate in male 1.33 times higher than in female (Hassan et al., 2016). The same report also stated that the CRC incidence and mortality rates are highest among Chinese compared to Malays and Indians (Hassan et al., 2016).

A possible explanation of the high incidence and mortality rate of CRC in Malaysia is the low awareness of CRC risk factors and poor recognition of the early symptoms of the disease (Lim, 2014) due to lack of education on the screening methods (Hilmi et al., 2010; Yusoff et al., 2012). Consequently, delayed diagnosis in CRC cases has led to late discovery of the disease, often at advanced stage (Yusoff et al., 2012).

2.1.2 Molecular Basis of Colorectal Cancer

Cancer progression is a multistep process starting from initiation, promotion and followed by progression (Volgelstein and Kinzler, 2015). Briefly, CRC begins as a benign adenomatous polyp, and later develops into advanced adenoma, which subsequently transforms into an invasive carcinoma (Figure 2.1) (Markowitz and Bertagnolli, 2010).

DNA damage-induced mutations are frequently the first strike to initiate cancer development (De Gruijl et al., 2001). Chromosomal instability is often observed in CRC due to rare inactivating mutations of genes involved in the maintenance of chromosomal stability, including MRE11 homolog A, double-strand-break repair nuclease (*MRE11A*) and F-Box and WD repeat domain containing 7 (*FBXW7*) (Barber et al., 2008). The chromosomal instability subsequently leads to loss-of-function of tumour suppressor genes *APC*, *p53* and the SMAD family member 4 (*SMAD4*) (Fearon, 2011). Besides that, defects in DNA repair frequently occur through mutations in the mismatch-repair (MMR) genes *MLH1* and *MSH2*. The mutated alleles of mismatch-repair genes could be inherited; additionally, somatic inactivation of the MMR genes accelerates the development of CRC (Boland et al., 2008). Gene silencing is also mediated by DNA methylation in CRC (Kim et al., 2010). In sporadic CRC, *MLH1* expression is suppressed by aberrant DNA hypermethylation (Kim et al., 2010). Tumour suppressor genes, including transforming growth factor-β receptor type II (*TGFBR2*) and BCL2-associated X protein (*BAX*), are

inactivated due to inability to repair strand slippage within repetitive DNA sequences (Markowitz and Bertagnolli, 2009; Fearon, 2011).

On the other hand, the loss of APC protein in CRC promotes Wnt signalling as a result of insufficient degradation of β-catenin oncoprotein (Goss and Groden, 2000). Subsequently, Wnt signalling leads to excessive cell proliferation (Goss and Groden, 2000). Besides, inactivation of *TP53* by missense mutations and chromosomal deletions also allows evasion from cell-cycle arrest and p53-mediated apoptosis; hence, the loss of *TP53* promotes transition of adenomas to invasive carcinomas (Fearon, 2011).

As mentioned earlier, TGF-β signalling normally functions as a tumour suppressor pathway (Wood et al., 2007). Mutations or deletions of *TGFBR2, SMAD4* and *SMAD2* subsequently shut down TGF-β signalling, and thus, enable transition from adenoma to high-grade dysplasia (Markowitz and Bertagnolli, 2009; Fearon, 2011). Growth-factor pathways involving prostaglandin was also identified in CRC (La Vecchia et al., 1997). Elevated expression of cyclooxygenase 2 (*COX2*) enhances the production of the carcinogen, prostaglandin E2 (Cha and DuBois, 2007). Degradation of prostaglandin E2 is limited due to the loss of 15-prostaglandin dehydrogenase (15-PGDH), which is observed in 80% of colorectal adenomas. Epidermal growth factor receptor (EGFR) signalling merges with MAPK signalling leading to downstream phosphorylation of c-MYC and, hence, accelerates the cell cycle (Markowitz and Bertagnolli, 2009).

The understanding of genetic alterations and pathways modulating CRC progression is important in the development of novel drugs or gene therapy to provide better prognosis in CRC treatment. (O'Brien et al., 2007).

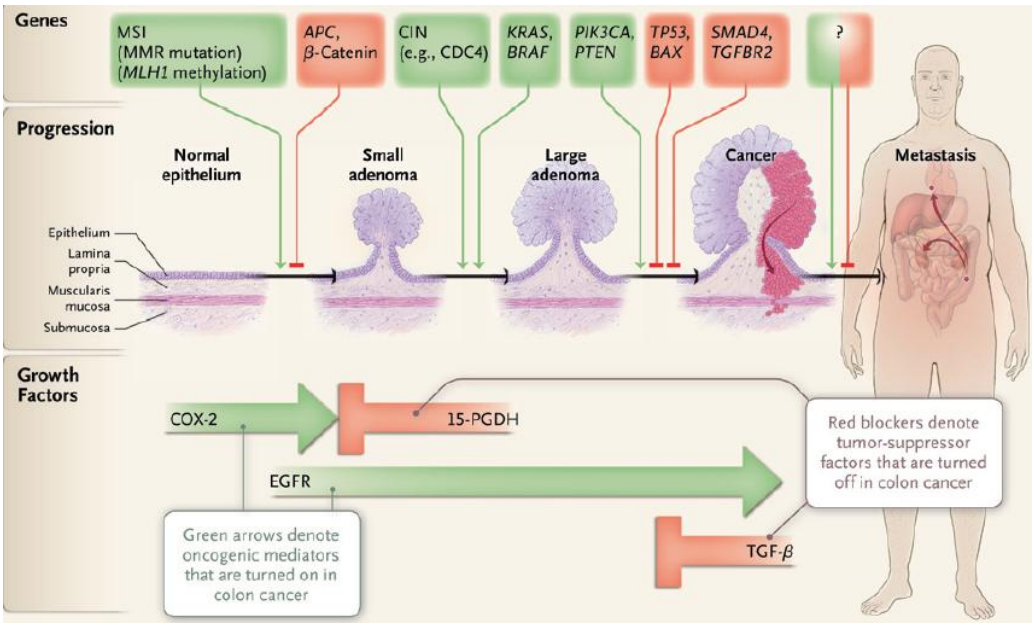


Figure 2.1: Molecular basis of colorectal cancer progression (Markowitz and Bertagnolli, 2009). An overview of genetic alterations and signalling pathways mediating progression of CRC is shown. The details are further elaborated in Literature Review Section 2.1.2.

2.2 Cancer Stem Cells (CSCs)

Cancer cells are reported to share some characteristics of pluripotent stem cells (Mooney et al., 2013). For instance, pluripotency-regulating pathways, including the Wnt/β-catenin, PI3K/Akt/mTOR and Hedgehog signalling pathways, are often activated and dysregulated in cancer cells (Mooney et al., 2013). In stem cells, these cell signalling pathways contribute to the maintenance of self-renewal, pluripotency and regular differentiation during development (Mooney et al., 2013). On the contrary, the effects of these pathways are amplified in cancer cells, leading to evasion of apoptosis, uncontrolled cell growth and enhanced motility (Mooney et al. 2013). Furthermore, common epigenetic regulating machinery mediated by DNMT (DNA methyltransferases) and TET (ten eleven translocation, catalyses DNA demethylation) are also commonly observed in cancer and stem cells (Hadjimichael et al., 2015). The molecular resemblances between cancer and stem cells described above have led to the presentation of stem cell-like characteristics in cancer (Semi et al., 2013), all of which are reflected in cancer stem cells (CSC), making CSC an ideal study models linking cancer and stem cell biology.

Cancer stem cells (CSCs) are a subset of cancer cells that reflects stem cell-like phenotypes (Marotta and Polyak, 2009). CSCs are well-characterised by unlimited self-renewal and multi-lineage differentiation potentials (Varghese et al., 2012). Self-renewal indicates asymmetrical cell division to generate an identical daughter stem cell with intact potential of cell

proliferation and differentiation while the other daughter cell may undergo differentiation (Seita and Weismann, 2010). Hence, CSCs are also known as tumour initiating cells, with the ability to form new tumours (Puglisi et al., 2013). Besides, heterogenous cell populations has been reported in colorectal cancer tumours, where some the cell populations are more relatively de-differentiated than others (Ashley et al., 2013). The less differentiated cells are believed to replenish the cancer population (Ashley et al., 2013). Previous studies have shown that CSCs enriched from gastrointestinal cancer express pluripotency genes and signalling pathways that are also found in embryonic stem cells (Müller et al., 2016).

2.2.1 Cancer Stem Cells in Colorectal Cancer

Presence of colorectal CSC was first demonstrated in *in vivo* tumour formation generated by the $CD133^-$ subpopulation of CRC (O'Brien et al., 2007). Furthermore, a subset of $CD133^+$/cytokeratin (CK) 20^- cells was found in CRC tumour samples (Ricci-Vitiani et al., 2007). The absence of epithelial differentiation marker CK20 suggested a less differentiated state of the cells. In the same report, the authors proposed that $CD133^+$ cells may be generated from an originally $CD133^-$ population (Ricci-Vitiani et al., 2007). In stem cells, ALDH1 protects the cell from oxidative stress and, hence, allows continuous cell proliferation. Expression of ALDH1 has been previously associated with $CD133^+$ cells, and the expression level is directly proportional to CRC

progression to carcinoma (Lugli et al., 2010). Hence, ALDH1 may serve as a potential marker in detecting colorectal CSCs.

The colorectal CSCs may be the culprit of cancer relapse due to incomplete eradication of cancer cells during chemotherapy (Reya et al., 2001). The failure is attributed to several drug-resistant mechanisms which are often observed in normal stem cells, including up-regulated expression of anti-apoptotic genes and ATP-binding cassette (ABC) transporters, which prevent cytotoxicity by limiting anti-cancer drug concentrations within cells (Reya et al., 2001). Hence, CSCs reservoir within the tumour bulk remains intact despite drug treatment. The CSCs that have survived from chemotherapy proliferate and acquire ability to metastasise, forming a new tumour at secondary sites (Spillane and Henderson, 2007).

Therefore, cancer stem cell model for CRC is essential to enhance the understanding of CRC pathology. Besides spheroidal-enrichment from *in vitro* cell culture (Wang, 2013), CSC-like cells have been generated through iPSC technology with increased sphere formation ability and enhanced chemoresistance (Oshima et al., 2014). The iPS technology will be further explained in Literature Review Sections 2.3-2.4.

2.3 Induced Pluripotency via Cellular Reprogramming

2.3.1 Embryonic Stem Cells (ESCs)

Embryonic stem cells (ESCs) are isolated from the inner cell mass of a pre-implantation blastocyst (Evans and Kaufman, 1981), and are pluripotent cells (Rippon and Bishop, 2004). In 1981, murine ESCs were first reported cultured and expanded *in vitro*. Since then, many researchers have used murine ESC as a model to study early embryo development and cellular differentiation. Subsequently, human ESCs were successfully derived in 1998 (Thomson et al. 1998).

Generally, ESCs are characterised by their self-renewal properties and the ability to differentiate into the three germ layers, and subsequently give rise to all somatic cell types (Rippon and Bishop, 2004). ESCs require specific culture conditions, such as the support of a feeder layer and cytokines to maintain an undifferentiated state. However, if induced with appropriate signals, ESCs are able to undergo lineage-specific differentiation to generate specific cell types of the three germ layers (Rippon and Bishop, 2004; Biswas and Hutchins, 2007).

Owing to the limitless propagative potential and flexibility in generating somatic cell types, ESC was widely used in early phases of regenerative research (Biswas and Hutchins, 2007). Furthermore, ESCs are also used in experimental analysis of gene regulation and functions in development and differentiation (Rippon and Bishop, 2004). However, the use of ESCs has

sparked ethical concerns due to the embryonic origin. Furthermore, there are several technical inconveniences that are hindering advances, including the availability and efficiency (Biswas and Hutchins, 2007).

2.3.2 Induced Pluripotency in Somatic Cells

In order to circumvent the ethical concerns and the aforementioned technical issues, other sources of pluripotent stem cells have been explored. Somatic-cell nuclear transfer utilises an enucleated donor egg to act as the recipient for the nuclear content of a somatic cell. The blastocyst formed thus develops into ESC-like cells (Nelson et al. 2010). Wilmut demonstrated that somatic cell nuclear transfer enabled *trans*-acting factors to reprogramme somatic cell nuclei to an undifferentiated state (Wilmut et al., 1997).

On the other hand, Yamanaka and his team first introduced pluripotency in murine somatic fibroblast cells by introducing a combination of four reprogramming factors, OCT4, SOX2, KLF4, and c-MYC (OSKM) (Takahashi and Yamanaka, 2006). The reprogrammed somatic cells are morphologically similar to ESC, express pluripotency markers and have acquired the ability to give rise to cell lineages of the three germ layers. Such cells are named as induced pluripotent stem cells (iPSCs). More importantly, the iPSCs are able to form teratoma *in vivo*, a benign tumour consisting of somatic cell types of the three germ layers (Huang, 2010). In the subsequent year, the human fibroblast-derived iPSC was generated by Yamanaka's group using a similar approach

(Takahashi et al., 2007). In the same year, another human fibroblast-derived iPSC was generated by replacing KLF4 and c-MYC with NANOG and LIN28 (Yu et al., 2007).

Potential applications of iPSC in disease modelling, regenerative therapies and drug discoveries have been actively researched (Huang, 2010). However, prior to clinical applications, it is important to investigate and map out the underlying mechanisms of iPSC generation. A fully elucidated mechanistic of reprogramming is expected to improve the efficiency and safety of cellular reprogramming (Ebrahimi, 2015).

2.3.3 Current Understanding on Underlying Mechanisms of Pluripotency Reprogramming

Genetic and epigenetic profile of iPSCs have been extensively studied (Buganim et al., 2013) to map the complex network of cell signalling and other pathways that govern the reprogramming process (Tanabe, 2015). To date, several pathways have been identified to participate in the three phases of reprogramming: initiation, maturation and stabilisation (David and Polo, 2014).

Loss of somatic genetic programmes by suppressing development-related genes was first observed in transfected cells (David and Polo, 2014). In addition, increased cell proliferation was observed during early reprogramming, which was shown to be attributed to involvement of fibroblast growth factor

(FGF) (Jiao et al., 2013) and the phosphatidylinositol-3-kinase (PI3K)-Akt (Chen et al., 2012) signalling. Furthermore, overexpression of miR-302 family enables the reprogrammed cells to overcome reprogramming-induced senescence (Banito et al., 2009). The PI3K-Akt signalling also aids the metabolic switch from oxidative phosphorylation to glycolysis in the initiation phase of reprogramming (Yoshida, 2015). Activators of PI3K-Akt signalling is reported to promote expression of glycolytic genes, and consequently, facilitates the reprogramming process (Zhu et al., 2010). Reports on hypoxic culture promoting reprogramming coincides with the metabolic shift to anaerobic glycolysis (Yoshida, 2015).

More importantly, the initiation phase of cellular reprogramming is characterised by the activation of mesenchymal-epithelial transition (MET) (Sanges and Cosma, 2011). The OSKM reprogramming factors are reported to reverse EMT via suppression of SNAI1, whereas E-CADHERIN expression is promoted by KLF4 (Li et al., 2010; Redmer et al., 2011). MET activation is also aided by cell signalling. BMP signalling promotes expression of MET-associated genes, *E-CADHERIN* (*CDH1*), *OCCLUDIN* (*OCLN*) and epithelial cell adhesion molecule (*EPCAM*), which in turn enhance MET activation (Samavarchi-tehrani et al., 2010). Conversely, the inhibition of EMT-inducing TGF-β pathway also facilitates initiation of reprogramming (Maherali and Hochedlinger, 2009).

In the maturation phase of reprogramming, epigenetic changes occur, allowing the activation of endogenous pluripotency genes. Furthermore, the

expression of NANOG is crucial for cells to proceed to the stabilisation phase (Silva et al., 2009). Withdrawal of ectopic expression of pluripotency genes in the maturation phase was unable to halt the transition to a full iPSC state, as the reprogrammed cells rely on the endogenous expression of pluripotency genes (Samavarchi-tehrani et al., 2010). Wnt signalling was also observed to promote the maturation phase in mouse somatic cell reprogramming (Ho et al., 2013). Somatic cell reprogramming in a mouse model also demonstrated the significance of LIF/STAT signalling during the maturation phase (Tang and Tian, 2013). Expression of pluripotency genes are elevated by suppressing DNA methyltransferase DNMT1 and histone deacetylases 2, 3 and 8 via LIF/STAT3 signalling (Tang and Tian, 2013).

Subsequently, the stabilised iPSC is distinguishable by suppression of transgene expression (Golipour et al., 2012). Active epigenetic rearrangements are observed during stabilisation in order to reset the epigenetic memory of parental cells (Kim et al., 2011). As an example, DNA methylation changes that occur throughout this phase of stabilisation are thought to aid the resetting of the epigenetic programme (Kumar et al., 2013). The stabilisation of mouse iPSC also requires reactivating the X chromosomes (Stadtfeld and Hochedlinger, 2010), an event which is not necessary in human iPSCs, indicating molecular differences between the human and murine cells in pluripotent reprogramming.

2.3.4 Application of Induced Pluripotent Stem Cells (iPSCs)

The iPSCs are given special attention as they are able to propagate indefinitely in culture as the ESCs, and the cells have the potency to give rise to all cell lineages. In regenerative medicine, donor-derived iPSCs may be directed into specific lineage or cell type, and then clinically returned to the injured site of the donor for tissue regeneration and repair (Huang, 2010). Theoretically, the iPSC-mediated regenerative therapy allows minimal use of immunosuppressant after tissue transplantation (Scott et al., 2013).

Researchers have since been using iPSCs in disease modelling and to treat genetic disorders. Hanna et al. (2007) first demonstrated that sickle cell anaemia-derived mouse iPSC is able to generate mutation-free haemotopoietic progenitors. Similarly, the principle may be applied to other human diseases with known mutations. On the other hand, iPSC derived from a patient suffering from spinal muscular atrophy (SMA) was found to retain the disease phenotype while being able to differentiate into functional motor neurons (Ebert et al., 2009). However, the motor neurons differentiated from SMA-iPSC degenerated in long-term culture, indicating that disease-specific iPSC is able to recapitulate disease progression. In addition, genetic defects are found to regain the wild-type phenotype upon reprogramming. Mis-splicing of the *IKBKAP* gene was corrected in neural crest precursor formed from familial dysautonomia (FA)-derived iPSC (Lee et al., 2009).

On the other hand, disease-specific iPSCs are also used in drug discovery and development (Wang, 2014). The SMA-iPSC described above was reported to be responsive to drugs, valproic acid and tobramycin, and subsequently enhanced production of survival motor neuron 1 (SMN1), which was initially expressed in low levels in the diseased cells. Furthermore, iPSC may also be used to screen potential drug candidates by evaluating disease phenotypes after exposure to drugs (Stadtfeld and Hochedlinger, 2010). Jung et al. (2012) demonstrated that dantrolene was able to recover Ca^{2+} spark properties to normal levels, and thus rescued the arrhythmogenic phenotype of catecholaminergic polymorphic ventricular tachycardia-derived iPSC (Jung et al., 2012).

Nonetheless, full translation of iPSC into clinical applications faces several challenges. First and foremost, iPSCs have not been proven clinically safe in human (Vojnits and Bremer, 2010). The use of oncogenes *OCT4*, *KLF4* and *c-MYC* in reprogramming bestows high tumorigenic potentials in the resulting iPSCs (Palomo et al., 2015). Furthermore, inhibition of the tumour suppressor p53 gene was reported to highly enhance reprogramming efficiency (Hong et al., 2009). Suppressed p53 in iPSC may pose additional risk of tumourigenesis. Chromosomal integration of the transduced transcription factors leads to genomic instability and increases the risk of mutations (Chen et al., 2014). Hence, alternative reprogramming approaches have been developed. Transgenes-free approaches have been improved by the use of small compound mimicking the effects of transcription factors (Shi et al., 2008), or the use of alternate vectors to avoid insertional mutagenesis (Zhou and Freed, 2009). The

use of recombinant proteins and non-coding RNA molecules are also utilised to increase reprogramming efficiency (Ye et al., 2009; Vitale et al., 2011).

2.4 Induced Pluripotency in Cancers

2.4.1 Induced Pluripotent Cancer Cells (iPCs)

Induced pluripotent cancer (iPC) cells are iPSC of cancer origin. To date, iPCs have been generated from various cancer types including chronic myloid leukaemia (Carette et al., 2010; Kumano et al., 2012), gastrointestinal cancer (Miyoshi et al., 2010), sarcoma and osteosarcoma (Choong et al., 2014; Zhang et al. 2013), pancreatic ductal adenocarcinoma (Kim et al., 2013), glioblastoma (Stricker and Pollard, 2014), hepatocellular carcinoma (Koga et al., 2014), colorectal cancer (Miyazaki et al., 2015) and squamous cell carcinoma (Takaishi et al., 2016). The reprogrammed cancer cells showed distinct stem cell-like features of limitless self-renewal and the ability to differentiate into cell lineages of the three germ layers. Moreover, some of the reprogrammed cancer cells are reported to be able to form *in vivo* teratoma (Sun and Liu, 2011), supporting full pluripotency.

Early studies of reprogrammed cancer cells have focused on the pluripotency features of iPCs (Yilmazer et al., 2015). Interestingly, diminished tumorigenicity is often reported in reprogrammed cancer cells. The differentiated iPCs showed increased sensitivity to chemotherapeutic agents and to differentiation-inducing treatments when compared to natural cancer

cells (Miyoshi et al. 2010). Similarly, enhanced drug sensitivity was also reported in reprogrammed hepatocellular carcinoma cells and colorectal cancer cells (Koga et al., 2014; Miyazaki et al., 2015). Furthermore, the reprogrammed hepatocellular carcinoma cells (Koga et al., 2014) and colorectal cancer cells (Miyazaki et al., 2015) also showed decreased proliferation rate, and were less invasive when compared to the pa rental cancer. Both iPCs and differentiated iPCs, which are also called post-iPC cells, showed reduced *in vivo* tumorigenicity when xenografted into immunodeficient mice (Miyoshi et al., 2010; Zhang et al., 2013).

Notably, epigenetic remodelling of tumour suppressor genes and oncogenes may also contribute to the reduced tumorigenicity. Zhang et al. (2013) reported modified histone landmark at the *c-MYC* promoter in reprogrammed sarcoma cells, allowing suppression of *c-MYC* expression. In addition, global DNA methylation status analysis also showed hypomethylation of tumour suppressor genes, as opposed to that found in somatic cell-derived iPSCs (Allegrucci et al., 2011; Zhang et al., 2013; Miyazaki et al., 2015). Intriguingly, partially-reprogrammed squamous cell carcinoma cells showed MET activation and exemplified with concurrent down-regulated expression of EMT-associated genes (Takaishi et al., 2016).

These findings point out a common observation that cellular reprogramming may partially attenuates cancer malignancies via extensive epigenetic remodelling (Ramos-Mejia et al., 2012). Hence, it has been

proposed that the cellular reprogramming of cancers is a potential epigenetic therapy approach in cancer treatment.

2.4.2 Important Factors Regulating Cancer Cell Reprogramming

Though various types of cancers have been successfully reprogrammed into pluripotent or near-pluripotent state where the reprogrammed cells were capable of terminal differentiation but lacked identical ESC-like gene expression profile (Zhang et al., 2013), the process and outcome of cancer reprogramming is highly influenced by several factors.

Overexpression of the key reprogramming OKSM factors in cancer cells is frequently reported, and has been associated with poorer clinical outcome (Amini et al., 2014; Müller et al., 2016). Hence, additional ectopic expression of OSKM or other reprogramming-associated genes may impose higher tumourigenesis potentials in the resulting iPCs (Mooney et al., 2013). Therefore, basal expression levels of reprogramming factors should be determined prior to reprogramming, and the least-expressed genes are ideally used in reprogramming (Rao and Malik, 2012). Alternatively, cell lines derived from a cancer type with the least expression of reprogramming factors are best selected for reprogramming (Yilmazer et al., 2015).

In addition, epigenetic alterations in cancer cells, including hypermethylation of tumour suppressor genes and hypomethylation of

oncogenes, may pose as a variable in cancer cell reprogramming. Global DNA demethylation of tumour suppressor genes and oncogenes is previously reported in reprogrammed sarcoma cell lines (Zhang et al., 2013). In the same report, chromatic modification was also observed to occur in a permissive manner, which allows repression of *c-MYC*. Hence, epigenetic characteristics of reprogrammed cancer cells may serve as a clue in understanding the underlying mechanisms of cancer cell reprogramming. The variables elaborated above should be sorted out in order to enhance reproducibility and to ensure safe and effective clinical applications in the future.

2.4.3 Induced Pluripotent Cancer Cells As a Model for Monitoring Cancer Progression

Current human cancer models are mostly based on tumour cell lines that are frequently derived from primary tumour of advanced stage. In such cell lines, the early events of oncogenic transformation are not amenable to analysis (Sun and Liu, 2011). Hence, iPCs are believed to recapitulate the early progression of cancer, and allow discovery of novel biomarkers of early disease stages (Figure 2.2) (Kim and Zaret, 2015).

As such, reprogrammed chronic myeloid leukaemia (CML) cells showed enhanced resistance to imatinib, suggesting that the chemotherapeutic agent may only target cells with a differentiated cell state (Carette et al., 2010). In a separate study, the reprogrammed CML-derived hematopoietic cells re-

acquired drug sensitivity to imatinib (Kumano et al., 2012). In addition, reprogrammed pancreatic ductal adenocarcinoma (PDAC) generated pancreatic intraepithelial neoplastic lesion when injected into immunodeficient mice, and showed an invasive phenotype of the parental PDAC. The invasive phenotype was later found to be associated with signalling pathways, indicating that the reprogrammed PDAC is able to recapitulate the early PDAC progression (Kim et al., 2013).

The findings of reprogrammed cancers have suggested that iPCs may ideally serve as a disease model that allows new insights on cancer biology. The network underlying early progression of cancers may be targeted in alternative cancer therapeutic approaches (Ramos-Mejia et al., 2012).

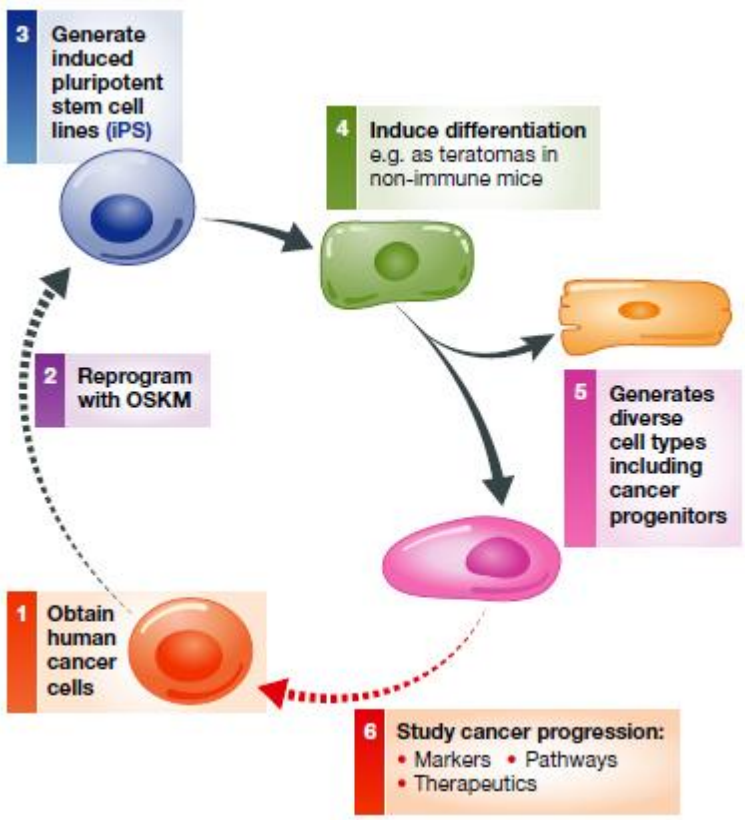


Figure 2.2: Reprogramming of cancer cells and the potential applications (Kim and Zaret, 2015). Cancer cells are first reprogrammed into induced pluripotent cancer cells (Steps 1–3), followed by differentiation to generate lineage-specific cancer progenitors (Steps 4–5). The cancer progenitors are used to recapitulate cancer progression, and allow discovery of novel biomarkers, pathways underlying cancer pathology and development of therapeutic approaches (Step 6).

2.5 MicroRNAs (miRNAs)

A group of small regulatory RNA, later known as microRNA (miRNA), was first discovered in 1993 (Lee et al., 1993). MiRNAs are a subset of non-coding RNAs with lengths of 18-22 nucleotides, and regulate gene expression post-transcriptionally (Bartel, 2004). MiRNAs anneal to 3'-untranslated regions (3'-UTRs) of target mRNAs, where base-pairing does not require complete homology (Bartel, 2004). This imperfect base pairing enables a miRNA to target multiple genes and lead to mRNA degradation or translational suppression, which in turn, regulates numerous cellular processes, including cell proliferation, tumourigenesis, pluripotency and cellular reprogramming (Wahid et al., 2010).

Most miRNAs are located in the intergenic regions of the genome, while some are found in intronic regions. Interestingly, miRNAs that reside in close proximity share the same promoter, and are co-expressed in poly-cistronic primary transcripts (Baskerville and Bartel, 2005). MiRNAs consist of a 6- to 8- nucleotide seed sequence at the 5' terminal end, which is responsible for target mRNA recognition at the 3'-UTR. The seed sequence is highly conserved among species and miRNA members of the same family (Laurent et al., 2008), suggesting that members from a miRNA family may evolve simultaneously and possibly target the same set of mRNAs. On the other hand, miRNAs are also available in mono-cistronic transcripts that are transcribed from specific promoter (Cai et al., 2004).

Rapid development of next-generation sequencing enables discovery of many novel miRNAs (Motameny et al., 2010). Furthermore, computational-based bioinformatics analysis has conveniently aid the prediction of the regulatory roles of miRNAs (Banwait and Bastola, 2015). These current technologies have highly promoted the studies of miRNA-mRNA interactions in cellular processes (Wahid et al., 2010).

2.5.1 Biogenesis and Mode of Function of MicroRNAs

The biogenesis of mature miRNAs is a multistep process, which requires processing in both the nucleus and cytoplasm of the cell (Figure 2.3) (Mallanna and Rizzino, 2010). MiRNAs are first transcribed into long primary miRNAs (pri-miRNAs) by DNA Polymerase II (Cai et al., 2004). A microprocessor-complex consist of RNase type III endonuclease Drosha, Di George syndrome critical region gene 8 (DGCR8) and other cofactors converts pri-miRNAs into precursor miRNAs (pre-miRNAs) with a hairpin-loop structure (Denli et al., 2004). Subsequently, the pre-miRNAs are transported to the cell cytoplasm via the Exportin-5/Ran-GTP complex (Bohnsack et al., 2004). A RNase-type III enzyme, Dicer, then recognises the 3'-overhangs of pre-miRNAs, and cleaves the pre-miRNA within the stem-loop structure (Hutvágner et al., 2001) to generate the mature miRNA duplexes.

The guiding strand of the miRNA duplex later incorporated into the Argonaute-containing RNA-induced silencing complex (RISC), directs the

RISC complex to the target mRNA and binds to the complimentary sequence at the 3'-UTR (Maniataki and Mourelatos, 2005). Consequently, the RISC-bound mRNAs are degraded if the base-pairing is perfectly complimentary (Figure 2.3) (Mallanna and Rizzino, 2010). Alternatively, if the base pairing between the guiding miRNA and the mRNA is only partially complimentary, translation of the mRNA into functional polypeptide protein is repressed (Mallanna and Rizzino, 2010). On the contrary, the other miRNA strand, the passenger miRNA is released and degraded (Winter et al., 2009).

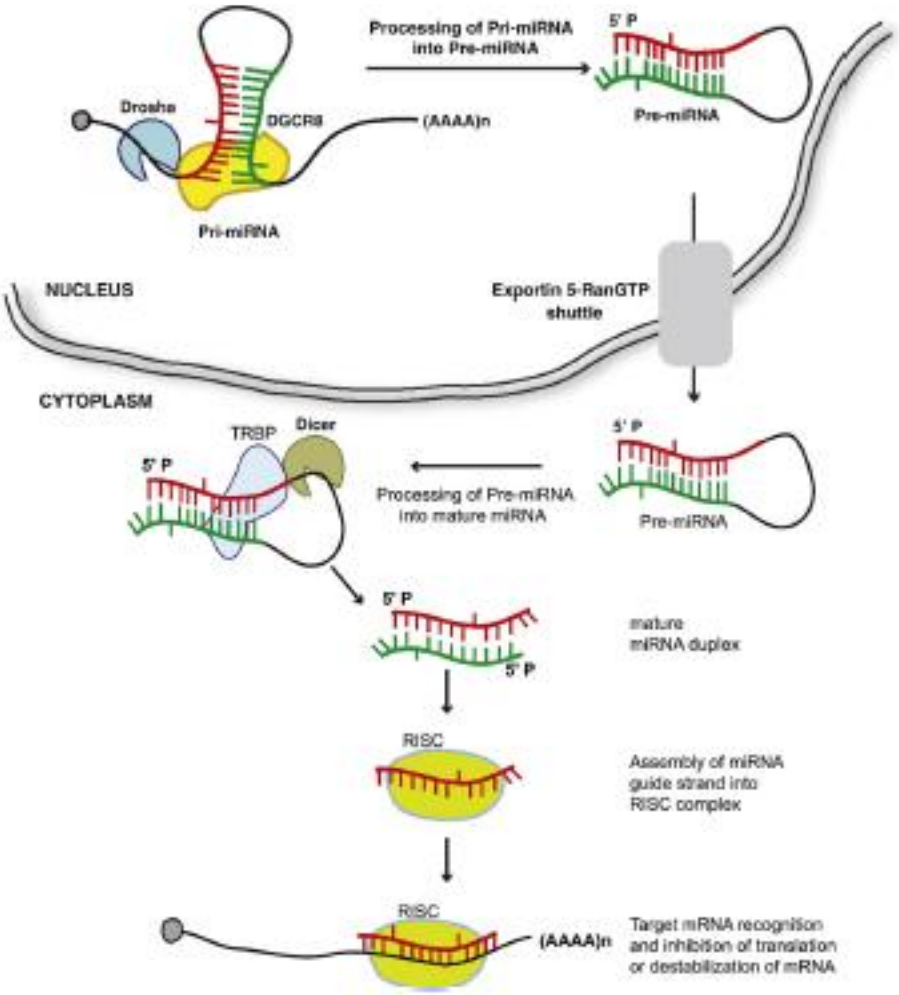


Figure 2.3: Biogenesis pathway of microRNAs (Mallanna and Rizzino, 2010). Processing of miRNAs is a multistep process, which occur in the nucleus and cytoplasm. The mature form of miRNA induces gene silencing post-transcriptionally by mRNA degradation or repression of mRNA translation into protein.

2.5.2 Role of MicroRNAs in Tumourigenesis

Dysregulated expression of miRNAs is frequently reported in cancers (Mendell and Olson, 2012). In general, miRNAs function as both tumour suppressors and promoters in cancer progression (Hammond, 2006). Oncogenic miRNAs, known as oncomiRs, target tumour suppressor genes and resulting in the loss of tumour suppressor functions in cells. In contrast, down-regulated expression of tumour suppressor miRNAs efficiently promotes translation of oncogenes into oncoproteins (Jansson and Lund, 2012). The combined effects synergistically drive tumour development by allowing active cell proliferation, evasion from apoptosis, invasion and metastasis (Jansson and Lund, 2012). Overexpression of miR-200 was shown to inhibit on sphere forming ability of breast cancer cells, whereas miR-34 induced p53-mediated cell cycle arrest, limiting cancer cell proliferation (He et al., 2007). On the contrary, miRNAs from the miR-155 and miR-181 families serve as oncomiRs, promoting tumour growth (Jiang et al., 2010).

Studies of global miRNA expression have demonstrated aberrant expression of miRNAs in CRC (Luo et al., 2011), suggesting that miRNA is indeed a critical regulator in CRC initiation and progression (Sassen et al., 2008). The expression of miR-143 and miR-145 are down-regulated compared to normal colorectal tissues, and are later recognised as tumour suppressor miRNAs (Michael et al., 2003). MiR-143 and -145 regulate cell proliferation by targeting cell-cycle inhibitors (Akao et al., 2007; Shi et al., 2007; Chen et al., 2009). Hence, suppressed expression of miR-143 and -145 allows excessive

cell proliferation in CRC. Additionally, OSKM is individually targeted by miR-145, which coincides with the OSKM overexpression in CRC (Xu et al., 2009; Müller et al., 2016). Down-regulation of miR-101 (Strillacci et al., 2009) in CRC results in up-regulated expression of the oncoprotein COX2, and, hence, promotes the production of carcinogenic metabolite prostaglandin E2 (Cha and DuBois, 2007). Other differentially-expressed miRNAs including miR-17-92 cluster, miR-106a, miR-31, miR-181b and miR-183, have also been identified in CRC (Luo et al., 2011).

The differentially-expressed miRNAs in various types of cancers may also serve as clinical diagnostic biomarkers. Overexpression of miR-21 is reported in glioblastoma (Chan et al., 2005) and in breast cancer (Yan et al., 2008); miR-21 may, therefore, be potentially used to detect the aforementioned cancers. Elevated expression of miR-92 was also detected in blood plasma of CRC patients (Ng et al., 2009), suggesting that levels of circulating miRNAs may conveniently be used in diagnosis of CRC. Besides, colonocytes isolated from stool samples of CRC patients also demonstrated higher expression levels of miR-21 and miR-106a (Link et al., 2010).

Given the regulatory role of miRNA in promoting and suppressing tumour growth, miRNA-based therapy targeting dysregulated miRNAs or genes may restore the miRNA expression to normal levels (Schetter et al., 2013). Notably, inhibition of oncomiRs may be achieved via delivery of antisense onligonucleotides or antagomiRs to cancer cells (Krützfeldt et al.

2005). The differentially-expressed miRNAs identified in this study may also provide novel miRNA candidates in miRNA-based cancer therapy.

2.5.3 Role of MicroRNAs in Maintaining Pluripotency

MiRNAs have long been shown to maintain self-renewal, pluripotency and regulation of cell fate in pluripotent stem cells (Lüningschrör et al., 2013). A subset of miRNAs, known as ESC-specific cell cycle-regulating miRNAs (ESCC-miRNAs), including the miR-302 family and miR-367, is highly enriched in ESC. The ESCC-miRNAs form an intertwined network with the ESC core transcription factors, OCT4, SOX2 and NANOG (Figure 2.4) (Anokye-Danso et al., 2012). Mechanistically, the core transcription factors first bind to the promoters of the ESCC-miRNAs, including the miR-302 and miR-367 clusters, activating miRNA transcription (Barroso-del Jesus et al., 2009). The ESCC-miRNAs are responsible to inhibit the expression developmental-associated genes. Besides, the ESCC-miRNAs also enhance the expression of other pluripotency genes, including *SOX2*, *LIN28* and *TBX3*, which in turn promote and maintain the expression of *OCT4*, *SOX2* and *NANOG* (Anokye-Danso et al., 2012).

Additionally, the core ESC transcription factors bind to the promoter region of genes and miRNAs associated with lineage commitment. Let-7 has been shown to inhibit ESC self-renewal (Melton et al., 2010). Elevated LIN28

expression impairs the maturation of let-7, and hence, maintains self-renewal in ESC (Shyh-Chang and Daley, 2013).

In short, the positive feed-forward regulatory loop of core ESC transcription factors and ESCC-miRNAs enhances the pluripotency programmes, and simultaneously suppresses differentiation. The combined effects maintain self-renewal and sustain pluripotency in ESC (Barroso-del Jesus et al., 2009; Anokye-Danso et al., 2012).

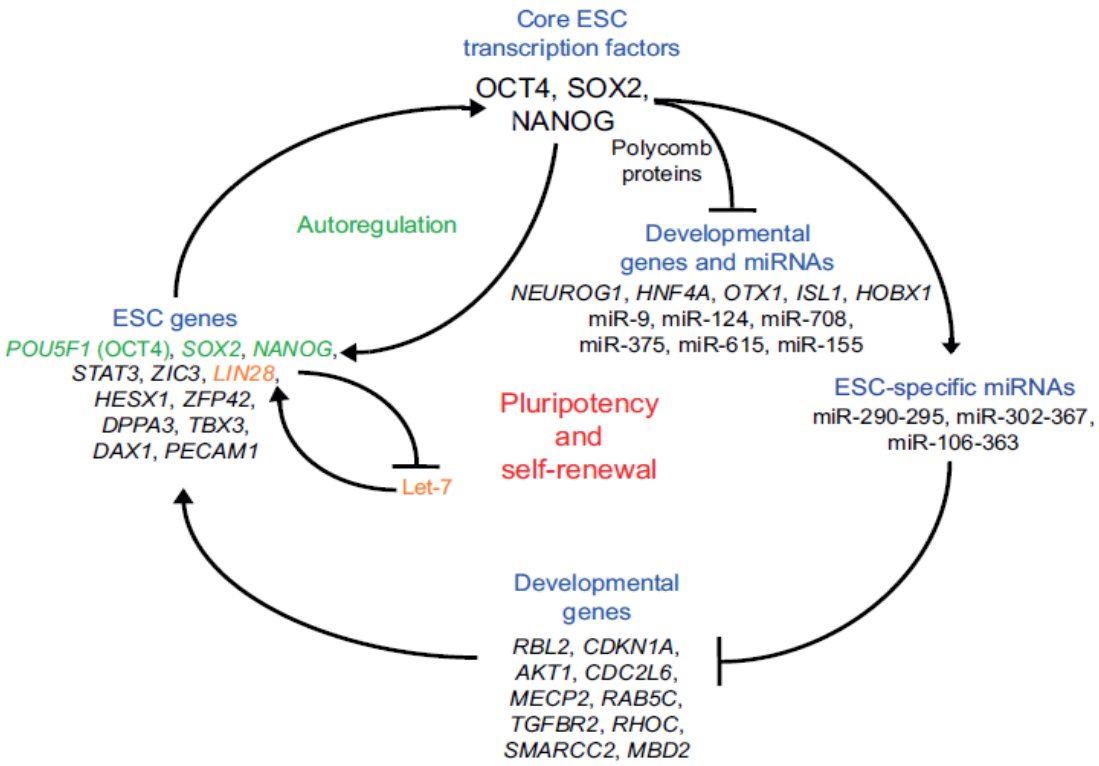


Figure 2.4: Regulatory circuit maintaining pluripotency and self-renewal in embryonic stem cell (Anokye-danso et al., 2012). Core ESC transcription factors form an intertwined network with ESC-specific miRNAs. The regulatory circuit promotes the expression of pluripotency genes and miRNAs. Expression of developmental-related genes and miRNAs are inhibited.

2.5.4 MiRNA-Mediated Reprogramming in Somatic and Cancer Cells

Due to the significance of miRNAs in pluripotency regulation, miRNA-induced reprogramming has been commonly used in iPSC and iPC studies. To date, miRNAs have been used to successfully generate iPSCs and iPCs from somatic and cancer cells (Subramanyam et al., 2011; Pourrajab et al., 2014).

Judson et al., (2009) reported that generation of mouse iPSC may be enhanced with transient transfection of miR-291-3p, -294, -295 and -302d in the presence of only three Yamanaka factors, OCT4, SOX2 and KLF4, thus, avoiding the use of the oncogenic c-MYC (Judson et al,. 2009). Other reprogramming experiments utilised miRNA groups including miR-302, -367 (Anokye-danso et al., 2011), miR-200c, -302s and -369s (Miyoshi et al., 2011) to induce ESC-like cells from mouse and human fibroblast cells, respectively. The miRNA-generated iPSCs are reported to show similar ESC-like features as the OSKM-mediated iPSCs, including expression of pluripotency genes and the ability to differentiate into cell lineages of the three germ layers (Anokye-danso et al., 2011; Miyoshi et al., 2011).

Similarly, miRNAs have also been used to reprogramme cancer cells. Lin et al. (2008) successfully generated iPC from human skin cancer cells using the miR-302 family (Lin et al., 2008). Other combinations of miR-200c, miR-302a-d and miR-369-3p and -5p have been used to reprogramme hepatocellular carcinoma and colorectal cancer cells (Koga et al., 2014; Miyazaki et al., 2015; Ogawa et al., 2015). These miRNA-induced iPC cells have shown suppressed

cell proliferation and invasion, and up-regulation of MET-associated genes, and thus raises the possibility that miRNA-mediated reprogramming attenuates malignancies of parental cancer cells (Koga et al., 2014; Miyazaki et al., 2015; Ogawa et al., 2015).

Cellular reprogramming using only miRNAs or in the presence of OSKM, was reported to have higher reprogramming efficiencies (Qi et al., 2009). Furthermore, miRNA-mediated reprogramming is highly dependent on the activation of endogenous pluripotency genes and relevant pathways. Therefore, epigenetic changes, such as chromatin modification and DNA methylation, in such iPSCs and iPCs may occur naturally as a result of altered expression level of endogenous genes (Wang et al., 2013). Hence, epigenetic scars are prevented in other reprogramming protocols that rely on transgene integration (Kim et al., 2010).

2.5.5 MicroRNAs Promote Reprogramming to Pluripotency

Efforts have been made to establish and compare miRNA expression profiles of pluripotent stem cells, including embryonic stem cells and somatic cell-derived iPSCs (Wilson et al., 2009; Huang et al., 2014; Lipchina et al., 2011). The distinctive expression profiles of miRNAs in reprogrammed cells have been functionally elucidated (Lüningschrör et al., 2013), and were found to be associated with OSKM-mediated transcriptional networks and numerous signalling pathways (Bao et al., 2013; Tanabe, 2015).

MiRNAs may play a role in suppressing OSKM inhibitors. NR2F2 which suppress the OCT4 promoter, is a target of the miR-302 family (Kuo et al., 2012). Orthologs of ESCC-miRNAs such as miRNA-302 and -372, improve reprogramming efficiency through targeting cell-cycle inhibitors and epigenetic modifiers (Subramanyam et al., 2011). Furthermore, miR-302 family also targets bone morphogenetic protein (BMP) inhibitors, resulting in BMP-mediated MET in the reprogrammed cells (Lipchina et al., 2011). Similarly, miR-200 and -205 have been reported to suppress transcriptional repressors of MET (Samavarchi-tehrani et al. 2010). As such, miRNAs play a crucial role in the activation of MET in the initiation phase of reprogramming (Bao et al., 2013; David and Polo, 2014). Conversely, inhibition of expression of miRNAs and proteins that antagonises the effects of ESCC-miRNAs also further enhances reprogramming efficiency (Melton et al., 2010).

MiRNAs also play a crucial role in regulating the p53 pathway during reprogramming (Lin et al., 2012). The *p53* gene has long been known to initiate cell cycle arrest, apoptosis and cellular senescence in response to cytotoxic stress and oncogene activation (Qian and Chen, 2010). The p53-induced senescence in reprogramming remains a major obstacle for the reprogrammed cells to achieve full pluripotency (Spike and Wahl, 2011). Hence, down-regulated p53 levels by overexpressing miR-138 or by suppressing miR-21 and -29a highly enhances reprogramming (Yang et al., 2011). Furthermore, p21-mediated cell-cycle arrest regulated by p53 is impaired by overexpression of miR-25, which in turn, promotes cellular reprogramming (Lu et al., 2012).

However, there are few systematic studies on miRNA expression in cancer-derived iPC cells. Although miRNAs may be thought to play similar roles in reprogrammed cancer cells as in reprogrammed somatic cells, knowledge of miRNA-regulated mechanisms modulating pluripotency reprogramming in cancer cells is currently insufficient. Genome-wide miRNA profiles of reprogrammed cancers may fill in the knowledge gap by identification of differentially-expressed miRNAs. The differentially-expressed miRNAs are expected to have biological significance in cancer reprogramming.

CHAPTER 3

MATERIALS AND METHODS

3.1 Cell lines

3.1.1 Colorectal Cancer Cell Lines

In this study, CRCs cell lines, HCT-15 and SK-CO-1, were obtained from American Type Culture Collection (ATCC, Manassas, VA, USA). HCT-15 was derived from a colorectal adenocarcinoma of Dukes' type C; and SK-CO-1 was derived from a metastatic site (ascites) of a colorectal adenocarcinoma patient. Both cell lines were adherent type of epithelial cells. The CRC cell lines were used to generate iPC cells, and acted as a control for the respective iPCs derived, in the subsequent experiments and gene expression analysis.

3.1.2 Embryonic Stem Cell Line

Embryonic stem cell line, H9 ESC, representing pluripotent stem cells, was a kind gift from Assistant Professor Dr. Shigeki Sugii (A*STAR Singapore Bioimaging Consortium, Singapore). In this study, H9 ESC was used as the positive control in expression analysis of pluripotency markers and genes, and in lineage-directed differentiation.

3.1.3 Virus Packaging 293FT Cell Line

293F-derived 293FT cells (Thermo Fisher Scientific, Illinois, USA) were originally established from primary embryonal human kidney cells and were transformed with sheared human adenovirus type 5 DNA (Graham et al., 1977; Harrison et al., 1977). The transformed cells were highly susceptible for virus production. 293FT was used as the retrovirus packaging cells for OSKM (Sugii et al., 2011).

3.1.4 Cell Lines Established in this Study

CRC-derived induced pluripotent cancer cells (CRC-iPCs), designated as iHCT-15 and iSK-CO-1, were generated from HCT-15 and SK-CO-1, respectively, via retroviral transduction of OSKM (See Materials and Methods Section 3.4). Pluripotency of the established CRC-iPC cell lines were firstly characterised, and subsequently, the cells were used in gene- and miRNA-expression analyses.

Post-iPC cells referred to the spontaneously differentiated cells from iPC clones via embryoid body formation (Materials and Methods Section 3.5.3). The post-iPC cell lines were used in pluripotency, germ-layer, and MET/EMT gene expression analysis.

3.1.5 Feeder Layer (Mouse Embryonic Fibroblasts)

Mouse embryonic fibroblast (MEF) derived from strain CF-1 (Merck Millipore, Massachusetts, USA) was used as the feeder layer for the growth of pluripotent stem cell lines. The non-mitomycin C-treated MEF was obtained at passage 3, and was further expanded prior to mitomycin C treatment. Mitomycin C treatment leads to cell cycle inhibition at the G2/M phase (Kang et al. 2001; An et al. 2015), and thus inhibited proliferation of the MEF cells. The non-proliferating feeder layer secretes important growth factors which help maintain pluripotency of stem cells, and also acted as a cellular matrix for stem cell attachment (Llames et al., 2015).

3.2 Cell Culture

Cell culture was performed in an Airstream Class II Biological Safety Cabinet (ESCO, Singapore) under aseptic environment. All cells were maintained in a 37 °C humidified cell culture incubator (ESCO) supplied with 5% CO_2. Disposable 3.5-, 6- and 10-cm cell culture dishes (BD Biosciences, New Jersey, USA), 6- and 24-well plates (SPL Life Sciences, Gyeonggi-do, South Korea), 5- and 10-mL serological pipettes (SPL Life Sciences) and 15- and 50-mL conical centrifuge tubes (BD Biosciences) were used for cell culture work. Media bottles and pipette tips were sterilised at 121 ºC for 15 min under a pressure of 975 kPa before use. An inverted phase contrast light

microscope (Nikon, Tokyo, Japan) was used to examine the cultured cells and to capture cell images.

3.2.1 Preparation of Cell Culture Media

Basal media for various cell lines, including Dulbecco's Modified Eagle Medium (DMEM)/F12, DMEM/high glucose and Minimum Essential Medium (MEM) (Gibco, California, USA) were prepared according to the manufacturer's instruction. DMEM/F12, DMEM/high glucose and MEM powder were dissolved in 900 mL nuclease-free double-distilled water (ddH_2O) by constant stirring. Appropriate amount of sodium bicarbonate (Merck KGaA, Darmstadt, Germany) was added into the dissolved medium solution (DMEMF/12: 1.2 g; DMEM/high glucose: 3.7 g; MEM: 2.2 g). The pH of the medium was then adjusted to pH 7.2 – pH 7.4, and topped up with ddH_2O to 1 L. The medium was filter-sterilised with a 0.2-μm cellulose acetate membrane filter unit (Techno Plastic Product, Trasadingen, Switzerland) using a vacuum pump system (GAST, Michigan, USA). For preparation of the complete culture medium for non-pluripotent stem cell lines, 10% (v/v) fetal bovine serum (FBS) (Gibco) and 1% of 100 U/mL Penicillin (Gibco) were added to the filtered medium.

For H9 ESC and CRC-iPC cell lines, standard hESC medium was used. Standard hESC medium composed of DMEM/F12 media supplemented with 20% KnockOut™ Serum Replacement (Gibco), 0.1 mM MEM Non-Essential

Amino Acids solution (Gibco), 1 mM L-glutamine (Gibco), 100 μM β-mercaptoethanol (Calbiochem, USA), and 10 ng/ml human fibroblast growth factor (Miltenyi Biotec, Bergisch Gladbach, German). For 293FT culture, DMEM/high glucose was supplemented with 10% FBS (Gibco), 1 mM sodium pyruvate (Gibco), 1 mM L-glutamine (Gibco), MEM Non-essential Amino Acids solution (Gibco), and 500 μg/mL Geneticin® (Gibco). All prepared media were kept at 4 °C for storage.

3.2.2 Cell Revival from Liquid Nitrogen Frozen Stock

For revival of 293FT, MEF and CRC cells, a vial of cryopreserved cells was retrieved from liquid nitrogen tank and placed in a 37 °C water bath until partially thawed. Cell suspension was swiftly transferred to a sterile 15-mL centrifuge tube containing 9 mL complete culture medium that is pre-warmed to 37 °C, and centrifuged at 1,000 rpm for 5 min in a benchtop centrifuge machine (Allegra® X-15R, Beckman Coulter, California, USA). The supernatant was discarded and the pellet was re-suspended in 1 mL complete culture medium. The cell suspension was then transferred into the appropriate culture vessels and was maintained in a 37 °C cell culture incubator with 5% CO_2. The next day, the culture medium was discarded and replaced with fresh complete medium.

For pluripotent stem cells, thawed cell suspension was first transferred to a sterile 15-mL centrifuge tube prior to dropwise addition of four mL complete

hESC medium that is pre-warmed to 37 ºC. The cell suspension was centrifuged at 1,000 rpm for 5 min in a benchtop centrifuge. Supernatant containing the cryopreservation medium was discarded. The cell pellet was gently re-suspended twice with 1 mL complete hESC media and added onto pre-seeded feeder layer. ROCK-inhibitor Y27632 (Merck Millipore, Darmstadt, German) was added at a final concentration of 10 µM to enhance cell viability after revival.

3.2.3 Cell Culture Maintenance and Sub-culturing

For colorectal cancer cell lines, 293FT, WJ0706 and post-iPC cell lines, each cell line were cultured in appropriate culture media and maintained in a 37 ºC cell culture incubator with 5% CO_2. Culture medium was replaced every two to three days. Sub-culturing of cells was performed when the cells reached 70-80% confluency. Initial culture medium was discarded and the cells were washed twice with 1X PBS at pH 7.4 (Amresco, Ohio, USA) to remove all traces of serum which could inhibit trypsin activity. One mL 0.25% Trypsin-EDTA solution (Gibco) was added to the cells and incubated at 37 ºC for 5 min to enhance detachment of cells from the culture vessels. Upon complete cell detachment, two mL serum-containing culture medium was added to inactivate trypsin activity. The cell suspension was gently transferred into a sterile 15-mL centrifuge tube and centrifuged at 1,000 rpm for 5 min. Following centrifugation, the supernatant was discarded and the pellet was re-suspended in 1 mL culture medium.

If cell seeding with a fixed density was required, the cell suspension was diluted with 0.4% (w/v) trypan blue solution (MP Biomedicals, California, USA) at a 1:1 ratio. Cell count was performed using a haemocytometer (Hirschmann, Eberstadt, Germany) and the cell density was determined by the formula followed:

Cell concentration per mL = (the average counted cells) x (dilution factor) x (10^4)

Further dilutions or calculations were carried out to obtain desired cell seeding density prior to plating the cells onto appropriate culture vessels.

Pluripotent stem cells, H9 ESC and CRC-iPC cell lines, were cultured with complete hESC media. Medium change was performed every day to maintain an undifferentiated state of the stem cell lines. Prior to subculture, mitomycin c-treated MEF feeder layer was pre-seeded 24 h before plated with stem cell colonies. Stem cell colonies were cut into grids with a surgical blade (Swann-Morton, Sheffield, UK), and gently scraped with a 1-mL pipette tip to detach the cell grids from the dish surface. Cell suspension of colony grids were transferred to a fresh feeder layer, and labelled as the next passage.

Stem cell lines were also cultured under feeder-free conditions for normal maintenance and during differentiation assays. Culture vessels were pre-coated with Matrigel® hESC-Qualified Matrix, LDEV-Free (Corning, New York, USA) for one hour at 37 ºC. Similarly, media changes were performed

daily using mTeSR™1 medium (STEMCELL™ Technologies, Vancouver, Canada). To enhance cell detachment on Matrigel® during subculture, stem cell dissociation reagent (ATCC, USA) was added at a final concentration of 0.5 units/mL, and left to react for 15 min before being carefully discarded. Cell colonies were collected in 2 mL mTeSR™1 and transferred to a 15-mL tube, followed by the centrifugation step as performed above. The cell pellet was then gently re-suspended with 1 mL mTeSR™1, and seeded on new Matrigel®-coated culture vessels.

3.2.4 Cryopreservation of Cultured Cells

Following cell trypsinisation as described in Materials and Methods Section 3.2.3 above, the cell pellets obtained were re-suspended in a mixture containing 90% (v/v) FBS and 10% (v/v) dimethyl sulfoxide (Sigma-Aldrich, Missouri, USA). Aliquots of the cell suspension were transferred into sterile cryovials (Corning, New York, USA) and stored overnight at -80 ^{o}C. The next day, the cryovials were transferred into a liquid nitrogen container (Chart Industries, Ohio, USA) at approximately -180 ^{o}C for long-term storage.

Instead of FBS/DMSO mixture, stem cell pellets were re-suspended with CryoStem™ freezing medium (Stemgent, Massachusetts, USA), and aliquoted to sterile cryovials (Corning, USA). The cryopreserved cells were stored at -80 ^{o}C overnight and then transferred to liquid nitrogen for long-term storage.

3.2.5 Inactivation of Mouse Embryonic Feeder Cells

Primary MEF cells (Merck Millipore, German) was first expanded and cultured till ~100% confluency prior to mitomycin C treatment. The MEF was mitotically arrested by treating with final concentration of 10 µg/mL mitomycin C (Merck Millipore, German) for three hours incubating in a humidified atmosphere at 37 °C with 5% CO_2. Medium containing mitomycin C were aspirated, and followed by three washes with 1X PBS (Amresco, USA) to remove all traces of mitomycin C. The inactivated MEF cells were trypsinised and cryopreserved at a desired cell concentration for future use.

3.3 Transfection of 293FT Cells with Retroviral Vectors OSKM

Different aliquots of 18 x 10^5 293FT cells were plated in 6-cm culture dish to ensure 90% confluency the next day prior to transfection with OSKM retroviral constructs (pMX–retroviral constructs were kind gifts from Dr. Shigeki Sugii, Singapore), with Lipofectamine 2000 transfection reagent (Invitrogen, California, USA). Twenty-four hours post-seeding, 6 µg pMX retroviral constructs harbouring OCT4, SOX2, KLF4, c-MYC, or OSKM, and the green fluorescence protein (GFP) genes were individually added to 250 µL blank DMEM/high glucose, along with 4 µg retroviral gag-pol packaging plasmid and 2 µg Vesicular somatitis virus G glycoprotein (VSV-G; envelope protein) expression plasmid. The plasmid mixture was incubated at room temperature for 5 min. The recommended volume of Lipofectamine 2000

(Invitrogen) was simultaneously added to 250 µL of blank DMEM/high glucose, and incubated for 5 min at room temperature. The plasmid mixture and Lipofectamine 2000 were combined in a 1:1 ratio, and subsequently incubated at room temperature for 30 min. After incubation, the mixture was added dropwise to the plated 293FT cells, and incubated at humidified atmosphere at 37 °C with 5% CO_2. The medium was replaced with antibiotic-free complete medium the next day.

Forty-eight hours post-transfection, the culture medium was collected as the first virus-containing supernatant, and was replaced with fresh medium for further harvest of the virus the next day as the second virus-containing supernatant. The virus-containing supernatant was filtered through a 0.45 µm pore-size PVDF filter (Merck Millipore). Filtered OSKM retroviral supernatant prior use was stored at -80 ^{o}C.

3.4 Retroviral Transduction of OSKM into Colorectal Cancer Cells

The parental CRC cell lines, HCT-15 and SK-CO-1, were plated in a 6-well plate at 8 x 10^5 cells and 12 x 10^5 cells per well, respectively. Retroviral supernatant of OCT4, SOX2, KLF4, and c-MYC was pooled in a ratio of 1:1:1:1, and added to the CRC cells, supplemented with 5 µg/ml polybrene (Merck Millipore). Another set of parental cells was infected with retroviral GFP to monitor the transduction efficiency. The cell-mixture was centrifuged at 2,500 rpm for 90 min at 32 °C to enhance viral infection. The cells were

subsequently incubated at 37 °C under a hypoxia condition at 5% O_2 for 24 h (Yoshida et al., 2009). The medium was changed the next day.

Upon reaching confluency, OSKM-transduced colorectal cancer cells were passaged to fresh feeder layers. The next day, the medium was replaced with standard hESC medium and continuously cultured until the emergence of ESC-like colonies (Figure 4.1), which normally took 18-21 days. Emerged ESC-like colonies were manually picked and transferred to fresh MEF feeder layers, cultured with hESC medium until the next passage (Sugii et al., 2011). This stage was defined as passage 1 (Takahashi et al., 2007).

Reprogramming efficiency of each OSKM-retroviral transduction experiment was calculated as followed,

Reprogramming efficiency = (total number of colonies) / (total number of cells seeded) x 100%

3.5 Characterisation of Colorectal Cancer-derived Induced Pluripotent Cancer Cells (CRC-iPCs)

3.5.1 Immunofluorescence Staining of Pluripotency Markers

Cells were fixed with 4% paraformaldehyde (Sigma Aldrich, St. Louis, USA), incubated at room temperature for 30 min, followed by blocking for 2 h using 1% bovine serum albumin (Nacalai Tesque, Kyoto, Japan). Washing

with PBS was performed twice every time before adding the subsequent reagent. Primary antibodies of pluripotency markers, mouse anti-human-TRA-1-60, -TRA-1-81, -SSEA-4 and -OCT4 (STEMCELL Technologies, Vancouver, Canada), were added at a dilution factor of 1:100, and incubated overnight at 4 ^{o}C. For intracellular marker OCT4, the cells were washed and permeabilised with 0.2% Triton X-100 (Sigma Aldrich). Secondary antibody, FITC-conjugated rabbit anti-mouse antibody (Merck Millipore), was added and left for one hour-incubation at room temperature. The nuclei were counterstained with DAPI (Gibco). All phase contrast and fluorescent images were captured using Axio Observer A1 with ZEN 2011 Lite software fitted with Argon laser and filter sets BP 450/490 and BP 546/12 (Carl Zeiss, USA). All steps described above were performed in the dark or without direct light exposure.

3.5.2 Lineage-directed Differentiation

In both ectoderm- and endoderm-directed differentiation, hESC was used as a positive control. For ectoderm-directed differentiation, hESC and CRC-iPC clones were first passaged to 24-well plates pre-coated with Matrigel®. The next day, culture medium was replaced with DMEM/F12 medium, 10% FBS supplemented with 100 ng/ml Noggin (R&D Systems, Minneapolis, USA). Noggin was used to direct differentiation towards the ectodermal lineage (Zhang et al., 2013). The culture medium was changed every alternate day for one week. For differentiation towards endoderm lineages, CRC-iPCs were

cultured with DMEM/F12 medium, 10% FBS supplemented with 100 ng/ml Activin A (R&D Systems). Activin A has been shown to promote differentiation into endoderm precursor cells (Wang et al., 2015). Mouse anti-human antibodies specific to ectoderm (MAP2) (Merck Millipore) and endoderm (AFP) (Merck Millipore) markers were used to indicate the differentiation status of differentiated CRC-iPC cells by immunofluorescence staining (Material and Methods Section 3.5.1 above). Phase contrast and fluorescent images were captured using Axio Observer A1 with ZEN 2011 Lite software fitted with Argon laser and filter set BP 450/490 (Carl Zeiss, USA).

For mesoderm-directed differentiation, putative CRC-iPC clones were differentiated into osteocytes and adipocytes. Mesenchymal stem cell, WJ0706, was used as a control for the differentiation assay. Putative CRC-iPC colonies were passaged to 24-well plates pre-coated with hESC-qualified Matrigel®, and continuously cultured with osteogenic and adipogenic medium (Choong et al., 2014). Osteogenic medium, DMEM/F12 supplemented with 10% FBS, 50 µg/ml ascobate-2-phosphate, 10 mM β-gylcerophosphate, 100 nM dexamethasone (Sigma), 1% penicillin and streptomycin. Adipogenic medium, DMEM/F12 was supplemented with 10% FBS, 1.0 µM dexamethasone, 0.2 mM indomethacin, 0.01 mg/ml insulin, 0.5 mM 3-isobutyl-1-methyl-xanthine (Sigma), and 1% penicillin and streptomycin. Differentiation medium was changed every alternate day for 21-23 days. For validation of osteocyte-directed differentiation, Alizarin Red S (Merck Millipore) staining was used to indicate the presence of calcium deposits (Sidney et al., 2014). Cells were fixed in ice-cold 70% ethanol for one hour, and subsequently washing with 1X PBS

was performed twice. Alizarin Red S (Merck Millipore) stain was added to well and left for incubation at room temperature for 30 min. Traces of Alizarin Red S was removed by washing with 1X PBS twice. Appropriate amount of water was added to prevent drying.

Lipid droplets produced by adipocyte-like differentiated iPC cells were visualised by Oil Red O (Merck Millipore) staining (Qian et al., 2010). Briefly, the cells were fixed in 10% formaldehyde (Sigma Aldrich) at room temperature for 30 min. The fixative was removed and washed with 1X PBS twice, followed by incubation with 60% isopropanol (Merck KGaA) at room temperature for 5 min. The isopropanol was later discarded and Oil Red O dye was added to the well. The cells were left to stain for 5 min, and subsequently washed with 1X PBS to remove excess dye.

Both staining were visualised under an inverted microscope and phase contrast images were captured using Ecplise Ts 100 with NIS-Elements D software (Nikon, Japan).

3.5.3 Spontaneous Differentiation into Post-induced Pluripotent Cancer Cells

Putative CRC-iPC clones were first allowed to form embryoid body (EB) in liquid suspension culture. During subculture, cell grids were transferred to 3.5-cm bacterial culture dishes (BD Biosciences) to prevent attachment of cells

to the dish surface. The floating cells were cultured with standard hESC medium for 7 days. At day 8, the embryoid bodies formed were transferred to 0.1% gelatin coated-culture dish for attachment, and, thus, encouraged further differentiation. During this period, the attached embryoid bodies were fed with hESC medium (without FGF-2). At day 14, the attached embryoid bodies were subcultured (Materials and Methods Section 3.2.3) a passage further to form Post-iPC cells, which were fed with DMEM/F12, 10% FBS and 1% penicillin-streptomycin. Post-iPC cells were later used in gene expression analysis or cryopreserved for future experiments.

3.6 RNA Preparation

3.6.1 RNA Isolation

In order to examine the expression of germ layer-markers, pluripotency and MET/EMT genes, total RNAs were isolated from the parental CRCs, CRC-iPC clones and post-iPC cells. Total RNA was isolated from the cells using the TRIzol® reagent (Invitrogen, California, USA) according to manufacturer's guidelines. Following trypsinisation (Materials and Methods Section 3.2.3), cell pellets were harvested and homogenised in 500 µL TRIzol® reagent. The homogenate was left to react at room temperature for 10 min. One-hundred µL chloroform (Amresco, USA) was added to the homogenate and vortexed vigorously for 15 s, followed by incubation at room temperature for 15 min. By centrifugation at 12,000 x *g* at 4 ^{o}C for 15 min, the homogenate was separated into three phases. The upper transparent phase containing RNA

was transferred into a new microcentrifuge tube, and mixed with 250 μL isopropanol (Merck KGaA). The RNA/isopropanol mixture was left at room temperature for 10 min. The mixture was centrifuged at 12,000 x *g* for 15 min at 4 ºC and the supernatant was discarded. Ethanol (Merck KGaA) (75%, 500 μL) was added and subsequently centrifuged at 12,000 x *g* for 5 min at 4 ºC. The supernatant was discarded and the pellet was air-dried for 5 min before eluted with 30 μL RNase-free water. The total RNAs extracted were subsequently treated with DNase (Promega, Winconsin, USA) to eliminate DNA contamination in the samples. Briefly, the reaction mix contains 15 μL of RNA samples, DNase 10X Reaction buffer and the DNase. The reaction mix was topped up to 20 μL with RNase-free water, and incubated at 37 ºC for 2 h. 1 μL of DNase Stop Solution was added and resumed incubation at 65 ºC for 10 min to inactivate the DNase.

All centrifuge steps were carried out in a Herarus Merck Millipore 21 refrigerated microcentrifuge (Thermo Scientific, Illinois, USA). RNA was quantified and stored at -80 °C for future use.

3.6.2 RNA quantification and integrity assessment

The concentration of total RNAs isolated was determined by a NanoPhotometer (Implen, Munich, Germany). The ratio of the readings at 260 nm and 280 nm (A260/A280) provides an estimation of RNA purity. The concentration and purity of the extracted total RNA were determined (Table

3.1). The A260/A280 ratio of all RNA samples was within the ideal range of 1.8-2.1; A260/A230 ratio of the RNA samples had also achieved a minimum value of 2.0, indicating minimum organic contaminations or residual reagents (EDTA and guanidine isothiocyanate) carried over from the RNA preparations.

RNA integrity was accessed by agarose gel electrophoresis. An aliquot of isolated RNAs (200 ng) was mixed with 2 μL 6X loading dye [60% glycerol, 0.03% (w/v) bromophenol blue, 0.03% (w/v) xylene cyanol, 60 mM EDTA] (Thermo Fisher Scientific) and topped up to a final volume of 10 μL with distilled water. The RNA mixtures were subjected to 1% agarose gel electrophoresis at 90 V for 45 min using 1X TBE running buffer (89 mM Tris base, 89 mM boric acid, 2 mM EDTA, pH 8.0) (Vivantis, California, USA). The gel was pre-stained with 3X GelRed nucleic acid stain (Biotium, California, USA) and visualised by exposure to 302 nm UV light under BioSpectrum Imaging System (Ultra-Violet Products Ltd, California, USA). Agarose gel electrophoresis analysis (Appendix A) showed that the 28S and 18S ribosomal RNA (rRNA) bands were sharp and intense, and the intensity of the 28S band was approximately twice that of the 18S band. The RNA samples were considered intact, and indicated high quality RNA samples. Hence, the RNA samples were used for downstream gene expression analyses in reverse transcription PCR (RT-PCR) and quantitative real-time RT-PCR (qRT-PCR). The total RNA isolated was also qualified for genome-wide miRNA profiling.

Table 3.1: Purity and concentration of RNA samples from CRC, CRC-iPCs, and post-iPCs for miRNA microarray and qRT-PCR analysis

Cell lines	A_{260}/A_{280}	A_{260}/A_{230}	Concentration (ng/µL)
Colorectal cancer cell lines (CRCs)			
HCT-15	1.928	2.001	371
SK-CO-1	1.962	2.082	1, 516
CRC-derived induced pluripotent cancer cell (CRC-iPC) lines			
iHCT-15 Clone 1	2.054	2.043	1, 520
iHCT-15 Clone 5	1.796	1.898	934
iSK-CO-1 Clone 1	1.944	1.619	1, 182
iSK-CO-1 Clone 2	1.928	1.387	1, 040
Post-CRC-iPC cell lines			
Post- iHCT-15 Clone 1	1.934	1.891	2, 352
Post- iHCT-15 Clone 5	1.996	2.012	2, 134
Post-iSK-CO-1 Clone 1	2.046	1.988	571
Post-iSK-CO-1 Clone 2	2.000	2.124	312

3.7 Determination of Messenger RNA (mRNA) Levels

3.7.1 Primers for mRNA RT-PCR and qRT-PCR

The primer sequence for exogenous OSKM factors were adapted from Takahashi and Yamanaka, (2006). Specific primer pairs for GAPDH, germ layer-markers, pluripotency genes and MET- and EMT-related genes were designed using the National Center for Biotechnology Information (NCBI) Primer-BLAST (Ye et al., 2012). All gene primers were tested using Basic Local Alignment Search Tool (BLAST) and global alignment algorithm to ensure that the primer sequences were unique and specific and did not non-specifically match with other gene sequences published in the NCBI database. All primers (Table 3.2 and 3.3) were custom synthesised by First BASE Oligos (Singapore).

Table 3.2: Primers pairs for transgene expression (adapted from Takahashi and Yamanaka, 2006)

Genes	Primer Sequences (5'-3')	Applications
OCT4	CCCCAGGGCCCCATTTTGGTACC	OCT4 transgene RT-PCR
SOX2	GGCACCCCTGGCATGGCTCTTGGCTC	SOX2 transgene RT-PCR
KLF4	ACGATCGTGGCCCCGGAAAAGGACC	KLF4 transgene RT-PCR
c-MYC	CAACAACCGAAAATGCACCAGCCCCAG	c-MYC transgene RT-PCR
pMX-AS3200	TTATCGTCGACCACTGTGCTGCTG	Transgene amplification for OCT4, SOX2 and KLF4
pMX-L3205	CCCTTTTTCTGGAGACTAAATAAA	Transgene amplification for c-MYC

Table 3.3: Primers used in RT-PCR and qRT-PCR of mRNAs

Genes	Accession[1]	Primer Sequences (5'-3')	Amplicon size (bp)
		Pluripotency markers	
OCT4	NM_001285987.1	F: GACAGGGGGAGGGGAGGAGCTAGG	144
		R: CTTCCCTCCAACCAGTTGCCCCAAAC	
SOX2	NM_003106.3	F: GGGAAATGGGAGGGGTGCAAAAGAGG	151
		R: TTGCGTGAGTGTGGATGGGATTGGTG	
KLF4	NM_001314052.1	F: ACGATCGTGGCCCCGGAAAAGGACC	397
		R: TGATTGTAGTGCTTTCTGGCTGGGCTCC	
c-MYC	NM_002467.4	F: GCGTCCTGGGAAGGGAGATCCGGAGC	327
		R: TTGAGGGGCATCGTCGCGGGAGGCTG	
NANOG	NM_001297698.1	F: AGTCCCAAAGGCAAACAACCCACTTC	161
		R: TGCTGGAGGCTGAGGTATTTCTGTCTC	
DNMT3B	NM_175850.2	F: TGCTGCTCACAGGGCCCGATACTTC	241
		R: TCCTTTCGAGCTCAGTGCACCACAAAAC	
LEFTY1	NM_020997.3	F: CTTGGGGACTATGGAGCTCAGGGCGAC	255
		R: CATGGGCAGCGAGTCAGTCTCCGAGG	
REX1	NM_001304358.1	F: TCTGAGTACATGACAGGCAAGAA	62
		R: TCTGATAGGTCAATGCCAGGT	

Table 3.3 continued

Genes	Accession[1]	Primer Sequences (5'-3')	Amplicon size (bp)
		Germ layer-markers	
MSX1	NM_002448.3	F: GAGTTCTCCAGCTCGCTCAG	100
		R: TCTCCAGCTCTGCCTCTTGT	
GATA6	NM_005257.5	F: TCTACAGCAAGATGAACGGCCTCA	125
		R: TCTGCGCCATAAGGTGGTAGTTGT	
CDX2	NM_001265.4	F: AAAGGCTTGGCTGGTGTATG	66
		R: GTCAGGCCTGGAGTCCAATA	
		MET/EMT markers	
CDH1	NM_001317186.1	F: TCATGAGTGTCCCCCGGTATCTTC	117
		R: GAGAATCATAAGGCGGGGCTGTGG	
OCLN	NM_001205255.1	F: CTGCTGCTGATGAATACAATAGAC	149
		R: AGCCTTCTATGTTTTCTGTCTATC	
SNAI1	NM_005985.3	F: AAGATGCACATCCGAAGCCACACG	102
		R: AGTGTGGGTCCGGACATGGCCTTG	
VIM	NM_003380.3	F: ACTCCCTCTGGTTGATACCCACTC	131
		R: CACTGAGTGTGTGCAATTTTTATTC	
		Quantification control gene	
GAPDH	NM_001256799	F: GAAATCCCATCACCATCTTCCAGG	120
		R: GAGCCCCAGCCTTCTCCATG	

[1]NCBI was used in data derivation. F: forward primer; R: reverse primer

3.7.2 cDNA Synthesis by Reverse Transcription

Total RNA was reversed transcribed into cDNA by using the Phusion RT-PCR kit (Thermo Fisher Scientific) according to manufacturer's protocols. One microgram total RNA was pre-denatured in a mixture of 0.25 mM dNTP and 100 ng oligo(dT) primer at 65 °C for 5 min. Then, 1X RT buffer, RT enzyme mix and RNase-free water were added to the reaction to a final volume of 20 µL. The reaction tube was briefly spun down and incubated in a 96-well Thermal Cycler (Takara, Shiga, Japan) at 25 °C for 10 min for primer extension, 40 °C for 30 min for reverse transcription into cDNA and 85 °C for 5 min for termination of reaction.

3.7.3 mRNA Direct RT-PCR

The presence of pluripotency genes in CRC-iPC clones and parental CRC was determined by direct RT-PCR using Ex*Taq* DNA Polymerase (Hot Start version, Takara). The reaction mixture containing cDNA, Ex*Taq* buffer, dNTP mix, forward and reverse primers and TaKaRa Ex*Taq* HS polymerase was set up according to Table 3.4, and lastly topped up to 20 µl with DEPC-treated water. RT-PCR was carried out with the following conditions: pre-denaturation at 98 ^{o}C for 10 s, 30 amplification cycles of denaturation at 98 ^{o}C for 10 s, annealing at 60 ^{o}C for 30 s, elongation at 72 ^{o}C for 1 min, followed by a final extension at 72 ^{o}C for 5 min.

Presence of the amplified gene was assessed by agarose gel electrophoresis. Five–microlitre amplicon product was mixed with 1 μL 6X loading dye (Thermo Fisher Scientific), and was subjected to 1.5% agarose gel electrophoresis at 90 V for 45 min using 1X TBE running buffer (Vivantis, California, USA). The gel was pre-stained with 3X GelRed nucleic acid stain (Biotium, California, USA) and visualised by exposure to 302 nm UV light under BioSpectrum Imaging System (Ultra-Violet Products Ltd).

Table 3.4: mRNA direct RT-PCR reaction mixture

PCR components	Volume per reaction	Final concentration/amount
10X Ex*Taq* Buffer	5.0 µL	1X
dNTP mix (2.5 mM)	1.6 µL	0.2 mM
Forward primer (10 µM)	0.8 µL	0.4 µM
Reverse primer (10 µM)	0.8 µL	0.4 µM
Ex*Taq* (5 U/µL)	0.1 µL	0.5 U
cDNA template (10 ng/µL)	5.0 µL	50 ng
DEPC-treated water	6.7 µL	
Total volume	20 µL	

3.7.4 mRNA real-time quantitative RT-PCR (qRT-PCR)

A final volume of 20 µl PCR reaction containing Ampli*Taq*® DNA Polymerase, SYBR® GreenER™ fluorescent dye, heat-labile uracil-DNA glycosylase (UDG), dNTPs with dUTP/dTTP blend and optimised buffer components along with cDNA template was set up according to Table 3.5. The reaction mix was subjected to quantitative PCR analysis using a Rotor-Gene Q (Qiagen) thermal cycler. A negative control was included in each assay by substituting cDNA with DEPC-treated water to ensure no contamination in the PCR reaction. Quantitative-PCR was carried out with the following conditions: 50 °C for 2 min for UDG activation, 95 °C for 2 min for UDG inactivation and hot-start DNA polymerase activation, followed by 40 amplification cycles of denaturing step at 95 °C for 15 seconds and primer annealing and extension step at 60 °C for 1 min. Data was analysed by using the comparative $\Delta\Delta C_t$ method in order to obtain the relative $\log_2$ (fold change) of the mRNA expression (Livak and Schmittgen, 2001; Schmittgen and Livak, 2008). The mRNA expression was normalised to the expression level of GAPDH. Reactions were performed in triplicates, and data were analysed by Student's *t* test. $P<0.05$ were considered statistically significant.

Table 3.5: mRNA qRT-PCR reaction mixture

PCR components	Volume per reaction	Final concentration/amount
2X SYBR GreenER qPCR SuperMix	10 µL	1X
Forward primer (10 µM)	0.2 µL	100 nM
Reverse primer (10 µM)	0.2 µL	100 nM
cDNA template (10 ng/µL)	1.0 µL	10 ng
DEPC-treated water	8.6 µL	-
Total volume	20 µL	

3.8 Global microRNA (miRNA) expression profiling analysis

Genome-wide miRNA expression profiling was performed using the Agilent Human MiRNA Microarray (Release 21.0) (Agilent Technologies, California, USA). The array was designed to include human mature miRNAs as annotated in miRBase 21, which consisted of 2, 549 miRNAs. The array was also designed to include 48 negative controls to measure the intensity and variance of background fluorescence signal (Wang et al. 2007). The microarray analysis experiments were conducted by Science Valley Sdn Bhd (Malaysia). The quality of the total RNA was first verified by an Agilent 2100 Bioanalyzer profile. Total RNAs (100 ng) was first labelled with Cyanine-3 (Cy3) using miRNA Complete Labeling and Hybridization Kit (Agilent) according to the manufacturer's instruction. Spike-in solutions were also included in the experiment to help distinguish significant biological data from processing issues. During the labelling step, phosphatase treatment was performed on the total RNA samples; the dephosphorylated RNA was then denatured before proceeding to labelling with Cy-3. The labelled RNA was desalted by drying the samples in a vacuum concentrator for 30 minutes until fully dried.

The dried samples were re-suspended with nuclease-free water before the addition of the hybridisation mix. For the hybridisation mix, spike-in solution was also added in. Samples were then hybridised on SurePrint Human MiRNA Microarray Release 21.0 (Agilent), G4872A-070156, at 55 ºC for 20 h. After incubation, slides were washed using Agilent Wash Buffers 1 and 2, and placed in a slide holder before proceed for scanning.

The microarray slide was scanned using scan protocol Agilent G3_miRNA on Agilent SureScan Microarray Scanner, G4900DA (Agilent Technologies, USA). The scan images in .tiff format were then extracted using the Agilent Feature Extraction Software version 11.5.1.1.

3.8.1 Microarray Data Analysis

The miRNA microarray dataset obtained were analysed and normalised using the GeneSpring GX software version 13.0 (Agilent Technologies). Comparative analysis between the iPC and CRC group samples was carried out using the *t*-test (*p*-values) and the Benjamini-Hochberg False Discovery Rate (FDR) correction (adjusted *p*-values) to remove false positive miRNAs. The criteria set for selection of differentially-expressed miRNAs was $\log_2$ (fold change) ≥ 2 or ≤ -2 with adjusted *p*-values < 0.05. Hierarchical clustering was performed to identify and visualise patterns of miRNAs expression between samples.

3.8.2 Identification of target genes and functional prediction analysis

The differentially-expressed miRNAs were used to predict target genes using the microRNA.org (http://microrna.sanger.ac.uk) (Betel et al., 2008) and TargetScan (http://www.targetscan.org) algorithms. The predicted target genes were then uploaded to the web-based programme Database for Annotation,

Visualization, and Integrated (DAVID) (http://david.abcc.ncifcrf.gov/tools.jsp) (Huang et al., 2009a; Huang et al., 2009b) for gene ontology (GO) annotation and KEGG pathway analyses.

3.9 MicroRNA (miRNA) Real-time Quantitative RT-PCR (qRT-PCR)

3.9.1 MicroRNA (miRNA) Primers for qRT-PCR

For miRNA qRT-PCR, mature sequences of selected miRNAs were obtained from miRBase ver. 21 (www.mirbase.org) and synthesised for use as specific forward primers; the Universal qPCR Primer (Invitrogen) was used as the reverse primer. All miRNA primers (Table 3.6) were synthesised by First BASE Oligos (Singapore).

Table 3.6: Primer sequences of miRNAs analysed in qRT-PCR

miRNA	Accession[1]	Forward primer sequence (5' – 3')
miR-125b-5p	MIMAT0000423	GGTCCCTGAGACCCTAACTTGTGA
miR-150-3p	MIMAT0004610	CTGGTACAGGCCTGGGGGACAG
miR-199a-3p	MIMAT0000232	GCGACAGTAGTCTGCACATTGGTTA
miR-362-5p	MIMAT0000705	GGAATCCTTGGAACCTAGGTGTGAGT
miR-500-3p	MIMAT0002871	ATGCACCTGGGCAAGGATTCTG
miR-532-3p	MIMAT0004780	CCTCCCACACCCAAGGCTTGCA
RNU6	X07425	F: CTCGCTTCGGCAGCACA R: AACGCTTCACGAATTTGCGT

[1]miRBase version 21 was used in data derivation. F: forward primer; R: reverse primer.

3.9.2 Polyadenylation and Reverse Transcription of MicroRNA (miRNA)

For miRNA expression analysis, total RNA was first reverse transcribed into cDNA using NCode™ miRNA First-Strand cDNA Synthesis Kit (Invitrogen) according to manufacturer's instructions, as outlined in Figure 3.1. Briefly, a reaction mixture composed of 20 µL containing 1 µg total RNA, polyA polymerase, 100 µM ATP and 2.5 mM $MnCl_2$ and 1X miRNA Reaction Buffer was first subjected to polyadenylation. The reaction mixture was incubated at 37 ˚C for 15 min in a 96-well Thermal Cycler (Takara). Subsequently, the PolyA-tailed RNA (4 µL) was added to the Universal RT Primer and Annealing Buffer. Reverse transcription of polyA-tailed RNA was performed at 65 ºC for 5 min, followed by addition of 1X First-Strand Reaction Mix and SuperScript III Reverse Transcriptase/RNaseOUT Enzyme Mix. The final mixture was incubated at 50 ºC for 50 min, and 85 °C for 5 min in 96-well Thermal Cycler (Takara).

3.9.3 MicroRNA (miRNA) qRT-PCR

Quantitative real-time PCR of miRNA was carried out using the protocol as described in Materials and Methods Section 3.7.4. Briefly, the synthesised cDNA and the Universal qPCR primer (Invitrogen) were added to the SYBR® Select Master Mix (Thermo Fisher Scientific) reagents according to Table 3.7. Negative controls were included in the assay by substituting cDNA with DEPC-treated water to ensure no contaminant was detected in the PCR

reactions. All experiments were performed in triplicates. MiRNA expression were normalised to the expression level of the small nuclear RNA U6, and data were analysed using the comparative C_t ($\Delta\Delta C_t$) method. SnRNA U6 was commonly used as an endogenous control for miRNA expression normalisation (Choong et al., 2007; Corney et al., 2007).

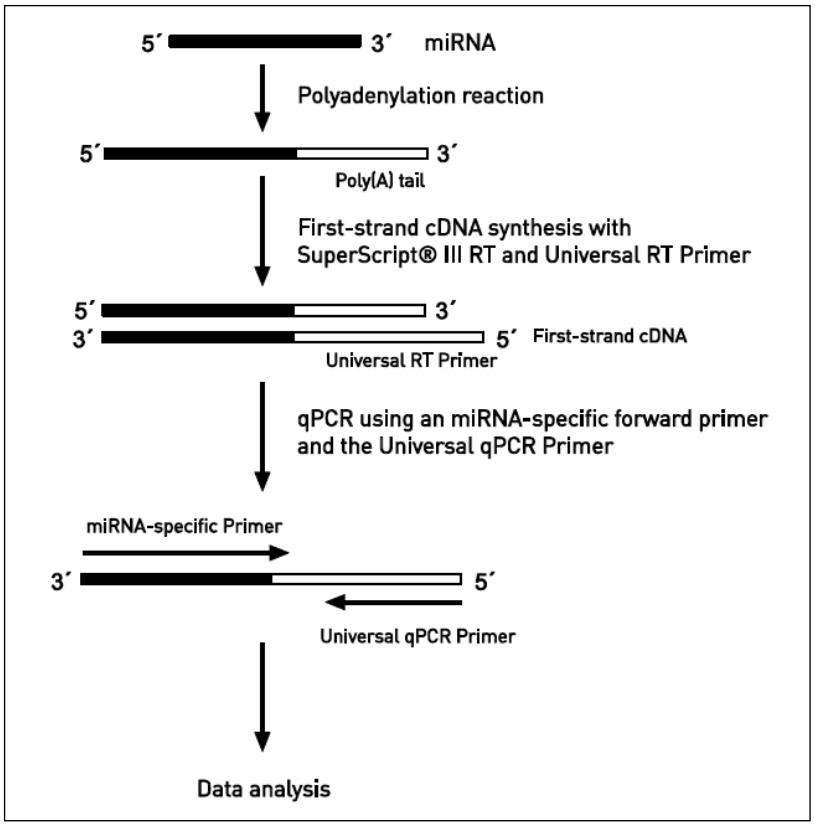


Figure 3.1: Workflow diagram of miRNA Real-Time qRT-PCR (Invitrogen, USA)

Table 3.7: miRNA qRT-PCR reaction mixture

PCR components	Volume per reaction	Final concentration/amount
2X SYBR GreenER qPCR SuperMix	10 µL	1X
Forward primer (10 µM)	0.2 µL	100 nM
Universal qPCR primer (10 µM)	0.2 µL	100 nM
cDNA template (10 ng/µL)	1.0 µL	10 ng
DEPC-treated water	8.6 µL	-
Total volume	20 µL	

3.10 Statistical analysis

Statistical analysis was performed on data tabulated in Microsoft Excel 2010 using the statistical formula to obtain mean values, deviation and variance, and to perform probability tests. Real-time RT-PCR results were reported as average of $\log_2$ (fold change) ± standard error. Data were analysed by Student's *t*-test (two-tailed distribution) by comparing the differences of miRNA or mRNA expression levels between the CRC-iPC clones and the respective parental CRC. $P<0.05$ values were considered statistically significant.

CHAPTER 4

RESULTS

4.1 Overview of Study Design

This study was divided into two major parts. Part 1 consisted of generation and characterisation of CRC-iPC clones, including expression of the pluripotency, MET- and EMT-associated genes. Part 2 of the study was aimed to establish the genome-wide miRNA profiles and to identify the differentially-expressed miRNAs in CRC-iPCs relative to the parental CRC cells. Prediction of biological functions, in particular in relation to pluripotency and MET/EMT, was done by bioinformatics analysis.

The experimental approaches to achieve the set aims of the two parts are shown in Figure 4.1.

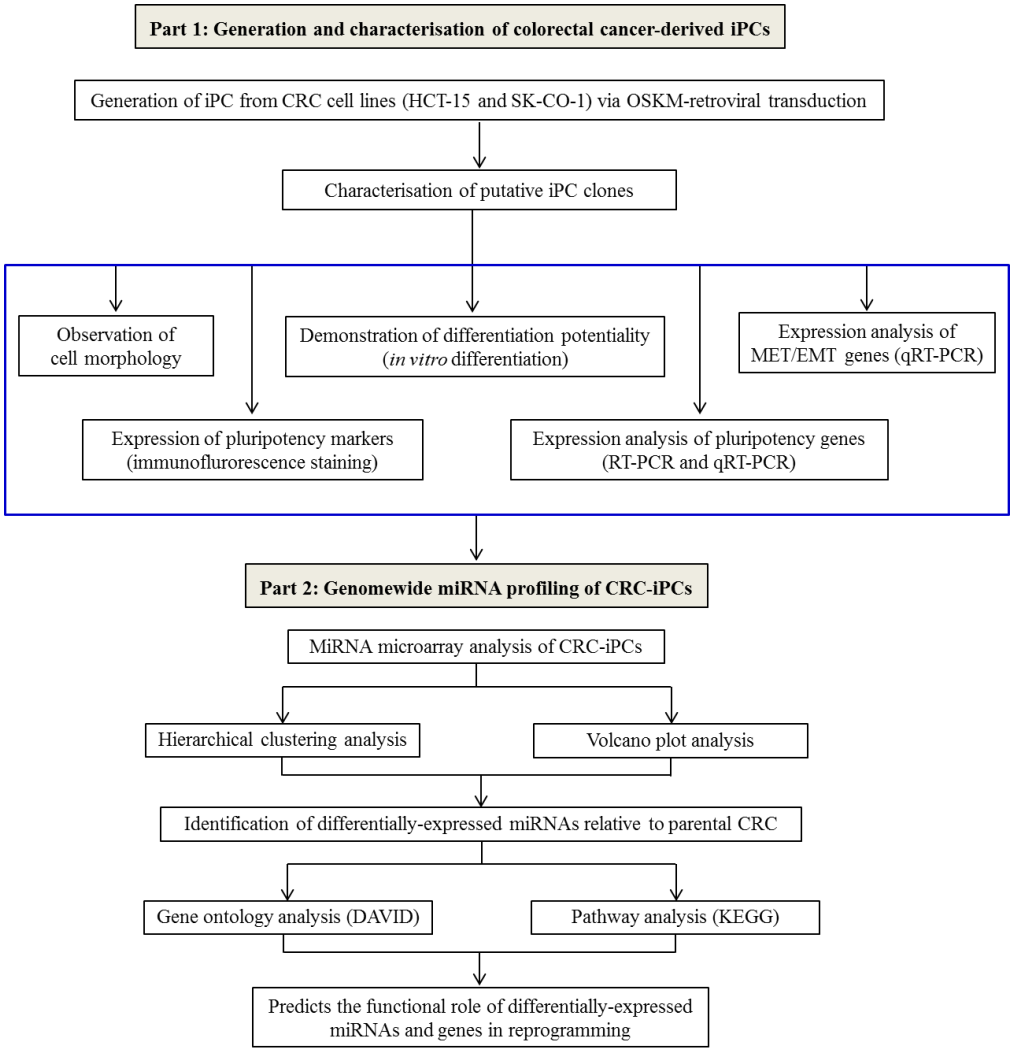


Figure 4.1: Overview of the study design. The study consisted of two parts. Part 1 included generation and characterisation of CRC-iPC clones, and expression analysis of the pluripotency and MET/EMT genes. Characterisation approaches of CRC-iPC clones are boxed in blue. Part 2 focused on genome-wide miRNA profiling of the characterised CRC-iPC clones and bioinformatics prediction of biological functions of the differentially-expressed miRNAs in reprogramming.

4.2 Retroviral OSKM Transduction of Colorectal Cancer Cells Generated Embryonic Stem Cell-like Colonies

OSKM-reprogramming is a well-established protocol to induce stem cell properties in both somatic and cancer cells (Rao and Malik, 2012). In this study, two colorectal cancer cell lines, HCT-15 (ATCC® CCL-225™), derived from a colorectal adenocarcinoma of Dukes' type C, and SK-CO-1 (ATCC® HTB39™), derived from ascites of a metastatic site colorectal adenocarcinoma, were subjected to cellular reprogramming. The two CRC cell lines were reprogrammed via retroviral transduction of the defined reprogramming factors, OCT4, SOX2, KLF4 and c-MYC (OSKM) (Takahashi et al., 2007) using the optimised protocol as described in Material and Method Section 3.4 (Sugii et al., 2011). The OSKM-transduced CRC cells formed rounded colonies at 18-21 days after OSKM-transduction. The colonies were picked and designated as iHCT-15 and iSK-CO-1 for iPCs derived from HCT-15 and SK-CO-1, respectively (Figure 4.2); different clones were designated with the prefix "C". The reprogramming efficiencies of OSKM transduction of HCT-15 and SK-CO-1 were 0.0043% for HCT-15 and 0.0125% for SK-CO-1, which were relatively lower than published data on other somatic cell-derived iPSCs (Takahashi et al., 2007) and cancer-derived iPCs (Koga et al., 2014). Nonetheless, two representative clones derived from each CRC cell line, Clone 1 and Clone 5 from iHCT-15, and Clone 1 and Clone 2 from iSK-CO-1 were selected for subsequent characterisation and gene expression experiments (Figure 4.2).

The iPC clones from iHCT-15 and iSK-CO-1 colonies showed cellular morphology similar to that of ESC, in that the cells proliferated in the form of colony and the cells were tightly packed within (Figure 4.2). Moreover, multilayer morphology was observed in the centre of the putative colonies, as the cells piled up due to rapid cell proliferation. Furthermore, the borders of the iPC colonies were not well-defined (Figure 4.2), as is found in ESC. Hence, the cellular morphology of CRC-iPC cells showed dissimilarities with ESCs and iPSCs.

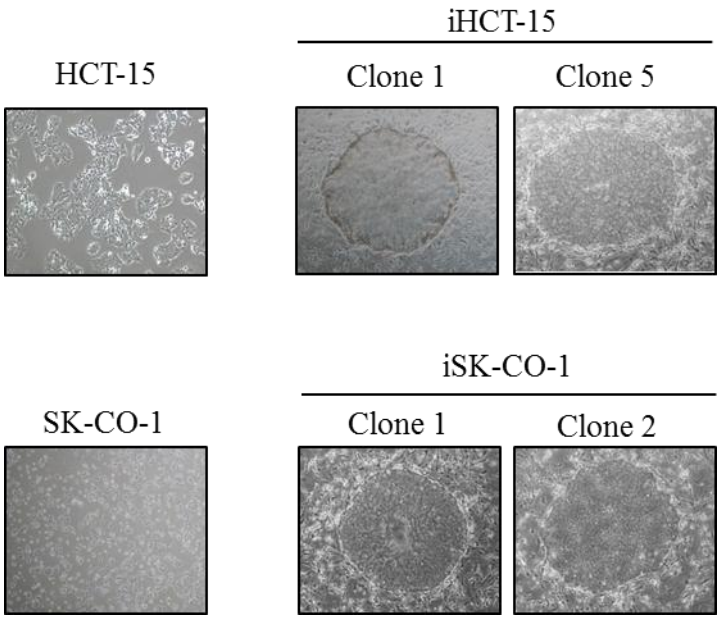


Figure 4.2: Cellular morphology of CRC and derived CRC-iPCs. Parental colorectal cancer cell lines, HCT-15 and SK-CO-1, were reprogrammed via OSKM-retroviral transduction. The OSKM-transduced CRC cell lines were designated as iHCT-15 and iSK-CO-1, respectively. Two representative clones of each reprogrammed CRC cell lines, Clone 1 and Clone 5 from iHCT-15, and Clone 1 and Clone 2 from iSK-CO-1 are shown. All images were captured at 10x magnification.

4.3 Expression of Pluripotency Markers in Colorectal Cancer-derived Induced Pluripotent Cancer Clones

Alkaline phosphatase (AP) is a phenotypic marker of pluripotent stem cells, and is commonly used as an early indicator of acquired stemness in reprogrammed somatic cells (Singh et al., 2012). The fluorescent images of AP live stain clearly indicated positive AP presence in the reprogrammed CRC cell lines (Figure 4.3). Interestingly, scattered positive fluorescent signals were also detected in both the parental HCT-15 and SK-CO-1 cell line in lower intensities.

Besides AP, the CRC-iPC clones were further examined for other common pluripotency markers, including TRA-1-60, TRA-1-81, SSEA-4, and OCT4. These pluripotency markers have previously been shown to be specifically present in pluripotent stem cells such as the ESC, but are absent in terminally-differentiated somatic cells (Zhao et al., 2012). The expression of the pluripotency markers in all four iPC clones was determined by immunofluorescence staining. The H9 embryonic stem cell line was used as the positive control. Pseudo-green green signal of TRA-1-60 showed that the CRC-iPC clones were positively stained, indicating the presence of TRA-1-60 (Figure 4.4). Similarly, TRA-1-81, SSEA-4 and OCT4 expression was positively detected in all iHCT-15 and iSK-CO-1 clones, indicated by the pseudo-green fluorescence signals (Figures 4.5 - 4.7).

It is also noted that the parental CRC cell lines also showed positive expression of AP, TRA-1-60, TRA-1-81, SSEA-4 and OCT4 (Figures 4.4 - 4.7), with various extents of fluorescent signal intensities. The results indicated that the stem cell-specific proteins were already expressed in the parental CRC cell lines. Previous studies have reported that ESC and cancer stem cells shared expression of factors that regulate self-renewal and cell proliferation (Zhao et al., 2012). Particularly, OCT4 was found to be abundantly expressed in various cancer types, as also observed in the two CRC cell lines in this study (Figure 4.7) and was intimately involved in promoting carcinogenesis (Samardzija et al., 2012). Taken together, the iHCT-15 and iSK-CO-1 clones showed the basic stem cell characteristics in cellular morphology and expression of the common pluripotency markers.

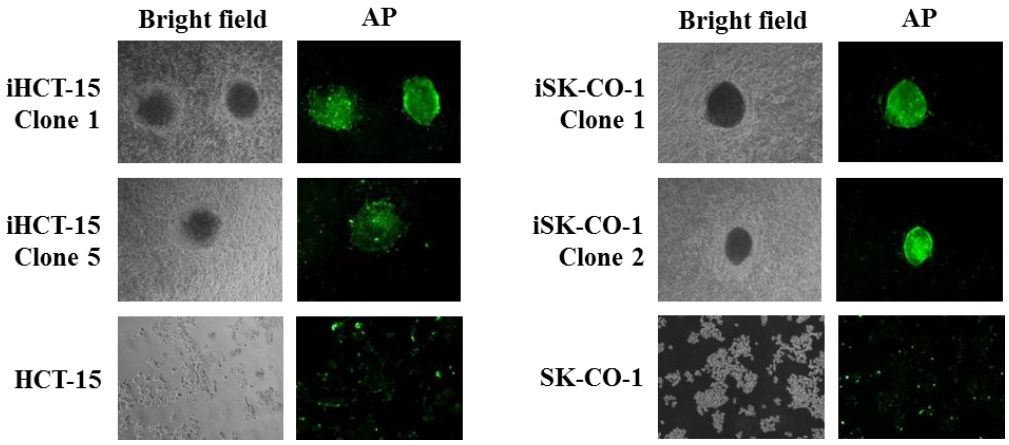


Figure 4.3: Live-staining of alkaline phosphatase (AP) in CRC-iPC clones. Presence of AP in CRC-iPC clones was indicated by pseudo-green fluorescence signal. Phase contrast (bright field) of each fluorescence image is included. All images were captured at 10x magnification.

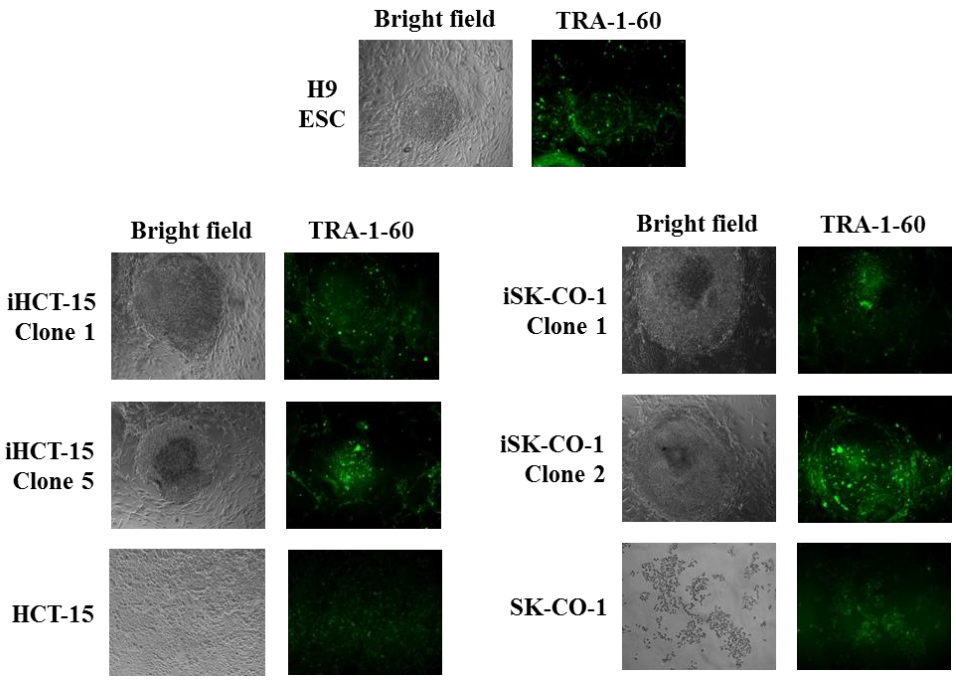


Figure 4.4: Expression of TRA-1-60 in H9 ESC, CRC-iPC clones and the parental CRCs. Presence of TRA-1-60 in each cell line was detected by immunofluorescence staining. Phase contrast (bright field) of each fluorescence image is included. All images were captured at 10x magnification.

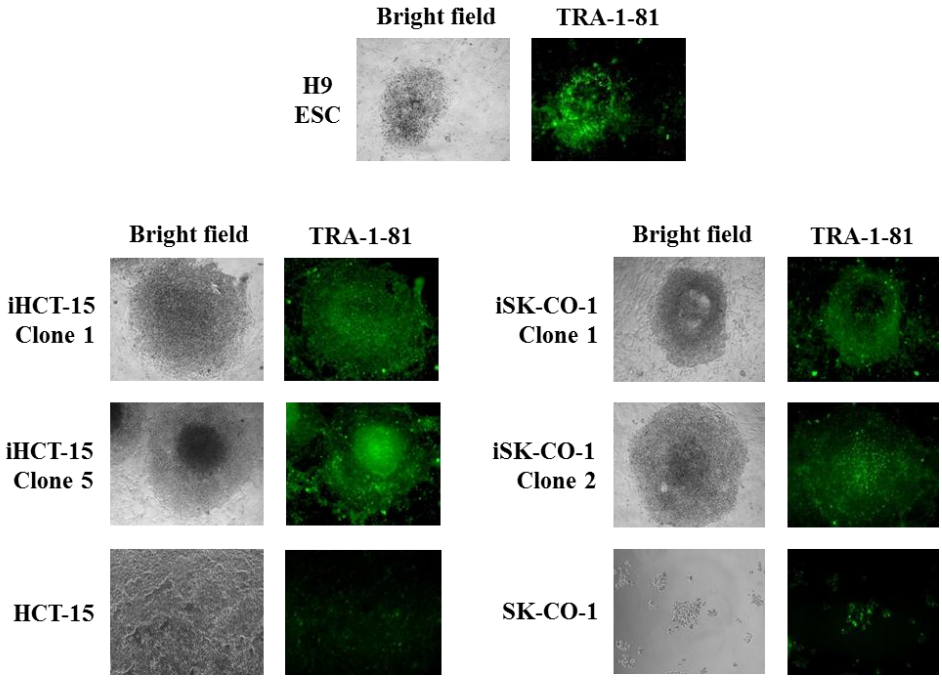


Figure 4.5: Expression of TRA-1-81 in H9 ESC, CRC-iPC clones and parental CRCs. Positive pseudo-green fluorescent signals indicated the presence of TRA-1-81 in the pluripotent stem cells and CRC cell lines. Phase contrast (bright field) of each fluorescence image is included. All images were captured at 10x magnification.

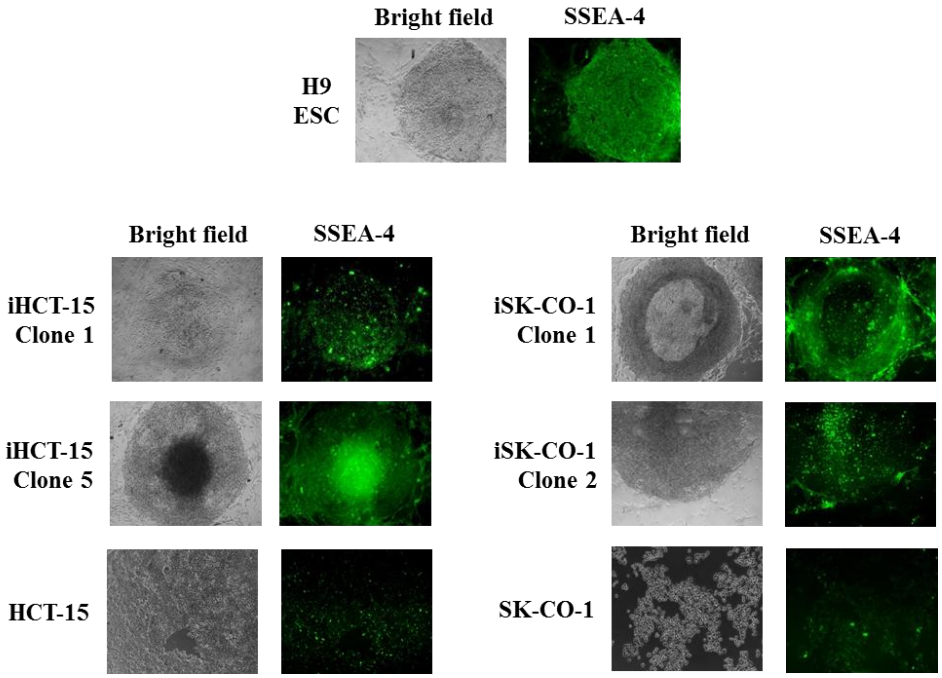


Figure 4.6: Expression of SSEA-4 in H9 ESC, CRC-iPC clones and parental CRCs. Expression of SSEA-4 was detected via immunofluorescence staining (green). Phase contrast (bright field) of each fluorescence image is included. All images were captured at 10x magnification.

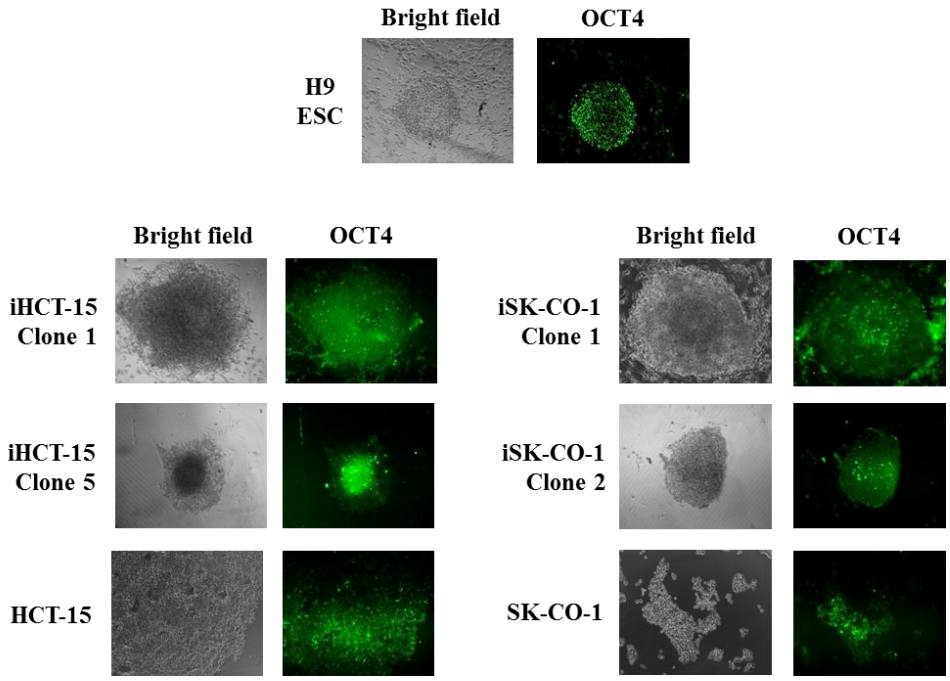


Figure 4.7: Expression of OCT4 in H9 ESC, CRC-iPC clones and parental CRCs. Expression of OCT4 was detected via immunofluorescence staining (green). Phase contrast (bright field) of each fluorescence image is included. All images were captured at 10x magnification.

4.4 *In vitro* Lineage-directed Differentiation of Colorectal Cancer-derived Induced Pluripotent Cancer Clones

The expression of common pluripotency markers is only the preliminary indication of acquired pluripotency in the CRC-derived iPC clones. In order to further demonstrate the pluripotency of the CRC-iPC clones, the cells were subjected to *in vitro* lineage-directed differentiation. The growth factors, noggin and activin A, were used to direct the differentiation towards the ectoderm (Zhang et al., 2013) and endoderm lineages (Teo et al., 2012), respectively. H9 ESC was used as the positive control. CRC-iPCs cultured with normal hESC medium was used as negative control. Expression of germ layer-markers, microtubule-associated protein 2 (MAP2, ectoderm) and alpha-fetoprotein (AFP, endoderm), in the differentiated CRC-iPC clones was positively detected by immunofluorescence staining (Figure 4.8), indicating that the CRC-iPC clones, indeed, differentiated into the ectoderm and endoderm lineages.

For mesoderm differentiation, the Wharton's Jelly-derived MSC cell line WJ0706, was used the positive control. Respective parental cells were served as negative control for mesoderm-directed differentiation. In osteogenic media, the WJ0706 differentiated into osteocyte-like cells of mesoderm lineage as indicated by the presence of extracellular calcium deposits stained bright orange-red by Alizarin Red S (Figure 4.9). Similarly, bright orange-red precipitates were also observed in the differentiated CRC-iPC cultures (Figure 4.9), indicated that the CRC-iPCs had differentiated into osteocyte-like cells,

where calcium salts were deposited around the cells due to mineralisation activities. Likewise, WJ0706 and CRC-iPC cells cultured in adipogenic media showed the presence of accumulated oil droplets, as visualised by Oil Red O staining (Figure 4.9).

In summary, the CRC-iPC cell lines, under *in vitro* induction conditions were able to differentiate into lineages or early progenitors of the three germ layers under the stimulation of appropriate growth factors, supportive of acquired pluripotency in the CRC-iPC cell lines.

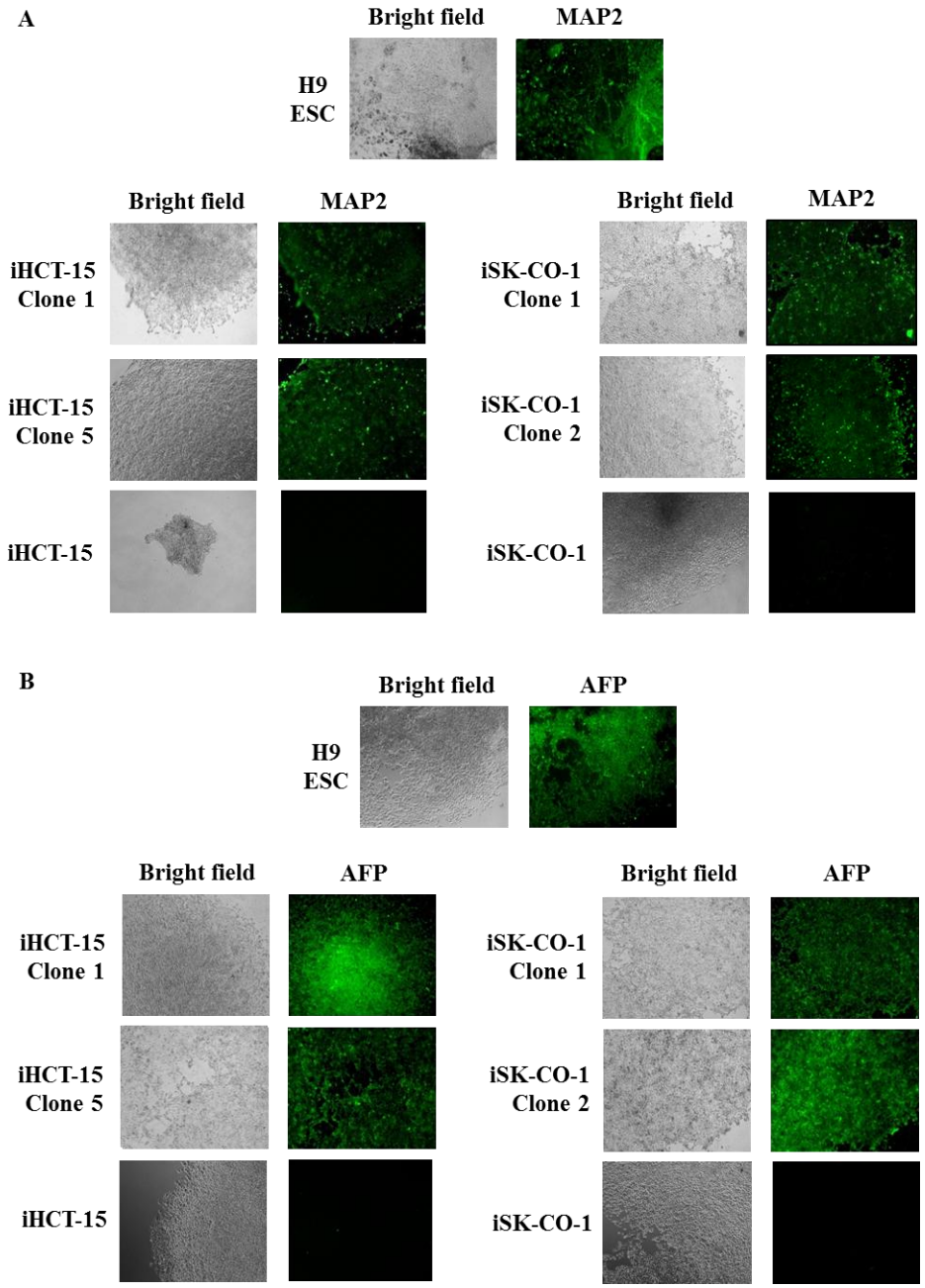


Figure 4.8: Expression of ectoderm (MAP2) and endoderm (AFP) markers in differentiated CRC-iPC clones. Pseudo-green fluorescence signal indicated expression of germ layer-markers MAP2 (A) and AFP (B) in differentiated CRC-iPC clones. Phase contrast (bright field) of each fluorescent image is included. All images were captured at 10x magnification.

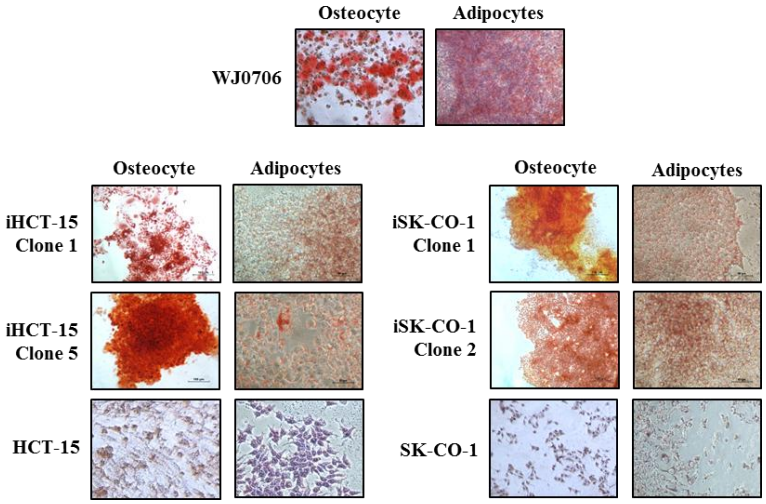


Figure 4.9: Mesoderm lineage-directed differentiation of CRC-iPC clones into osteocyte- and adipocyte-like cells. Differentiated CRC-iPC clones were stained with Alizarin Red S staining to detect the presence of calcium deposits; images were taken at 10x magnification. Lipid droplets were detected via Oil Red O staining; images were taken at 40x magnification.

4.5 *In vitro* Spontaneous Differentiation of Colorectal Cancer-derived Induced Pluripotent Cancer Clones

Further evidences of acquired pluripotency of the CRC-iPC cells were obtained by spontaneous differentiation into post-iPC cells, via embryoid body formation, following the protocol of Miyoshi et al. (2010). All iPC clones from both iHCT-15 and iSK-CO-1 were able to form spherical embryoid bodies in liquid suspension culture (Figure 4.10A), which is a typical differentiation feature of stem cells (Kurosawa, 2007). The floating embryoid bodies were then transferred to a primary culture condition, but supplemented with 10% FBS to encourage further differentiation; the cells thus obtained are called post-iPC cells (Miyoshi et al., 2010). The post-iPCs grew in monolayer instead of colonies, which was more similar to the parental CRCs (Figure 4.10A). The differentiation status of the post-iPCs was further assessed by the expression of the germ-layer markers, coudal type homeobox protein 2 (*CDX2*; ectoderm), Msh homeobox 1 (*MSX1*; mesoderm), and GATA binding protein 6 (*GATA6*; endoderm), relative to the CRC, and concurrently compared to the CRC-iPCs (Figure 4.10B).

On reprogramming, expression levels of the germ-layer markers were down-regulated in the CRC-iPC clones (Figure 4.10B), indicating that the reprogrammed cells may have achieved a more primitive cell state. On the contrary, expression of the germ-layer markers in the post-iPC cells was generally restored. restored to a level comparable to the respective parental CRC cell lines (Figure 4.10B). Notably, the expression levels of *MSX1* in post-

iHCT-15 Clone 1 and both post-iSK-CO-1 clones were up-regulated to a level higher than the respective parental cells. Similarly, *CDX2* expression in post-iSK-CO-1 Clone 2 was also up-regulated to a level higher than parental SK-CO-1.

In general, the OSKM-reprogramming seems to have reversed the parental CRC cancer cells into a more ESC-like pluripotent state, as reflected by, the down-regulation pattern of the germ-layer markers. On re-differentiation, the CRC-iPC clones were spontaneously differentiated into early progenitor cells, as indicated by the restoration of the expression levels of *CDX2*, *MSX1* and *GATA6*. The spontaneous differentiation process may have resulted in random commitment into different cell lineages, and thus may explain the various expression levels of germ-layer makers observed in CRC-iPCs and post-iPC cells.

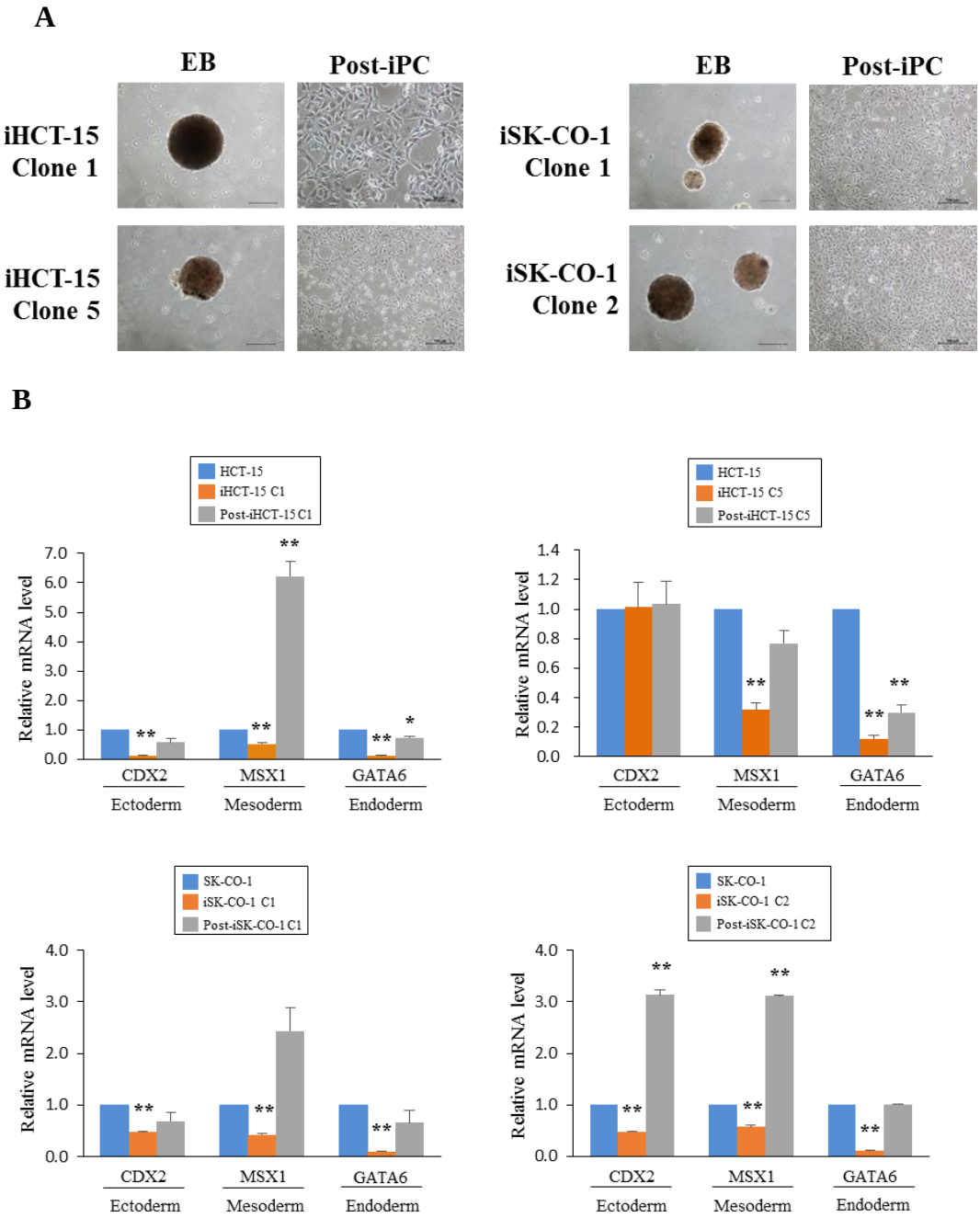


Figure 4.10: Spontaneous differentiation of CRC-iPC clones. (A) Cellular morphology of embryoid bodies (EB) and post-iPC cells derived from the CRC-iPC clones. All images were captured at 10x magnification. (B) Relative expression levels of germ-layer markers in CRC, CRC-iPC and post-iPC cell lines. Real-time RT-PCR data presented were from three independent experiments in triplicates. Expression levels of germ-layer markers were normalised to that of the respective parental CRC cell lines, which was arbitrarily set as 1.0. *p<0.05 and **p<0.01.

4.6 Ectopic expression of Reprogramming Factors OSKM in Colorectal Cancer-derived Induced Pluripotent Cancer Clones

The molecular features of the reprogrammed CRC cells are important to provide insights on the reprogramming status of the resulted iPC cells. During the stabilisation phase of reprogramming normal somatic cells, endogenous expression of pluripotency genes were activated, while the transgenes were silent, and hence, marking the maturation of an authentic iPSC (Buganim et al., 2013). To determine ectopic expression of the reprogramming factors in the established iPC clones, primers were designed that included the vector sequence (Table 3.2). Early- (passage 10 or earlier) and late-passage (passage >25) CRC- iPC cells were used in the analysis. Both early and late passage CRC-iPC clones were subjected to transgene amplication by direct RT-PCR.

Exogenous expression of *OCT4*, *SOX2*, *KLF4* and *c-MYC* was not expressed in both early and late passages of iHCT-15 clones and iSK-CO-1 Clone 1 (Figure 4.11), indicating that the CRC-iPC clones were transgene-independent from the early passage, and that the ectopic expression was not re-activated in the pro-longed culture. On the other hand, exogenous *OCT4* and *KLF4* were absent in both early and late passages of iSK-CO-1 Clone 2, and remained silent throughout the pro-longed culture. In contrast, strong ectopic expression of *SOX2* and weak *c-MYC* expression were observed in the early passage of iSK-CO-1 Clone 2, and ectopic expression of these genes was still detected at late passage. The expression of the *SOX2* and *c-MYC* transgenes

was indicative of partial reprogramming, and, thus, the clone possibly remained at a pre-iPSC stage (Chen et al., 2012) and were not identical in genetic and epigenetic profiles of ESC (Chen et al., 2012). Nonetheless, iSK-CO-1 Clone 2 demonstrated multiple pluripotency characteristics as described in the previous sections, including the ability of tri-lineage differentiation, and was included in subsequent analyses.

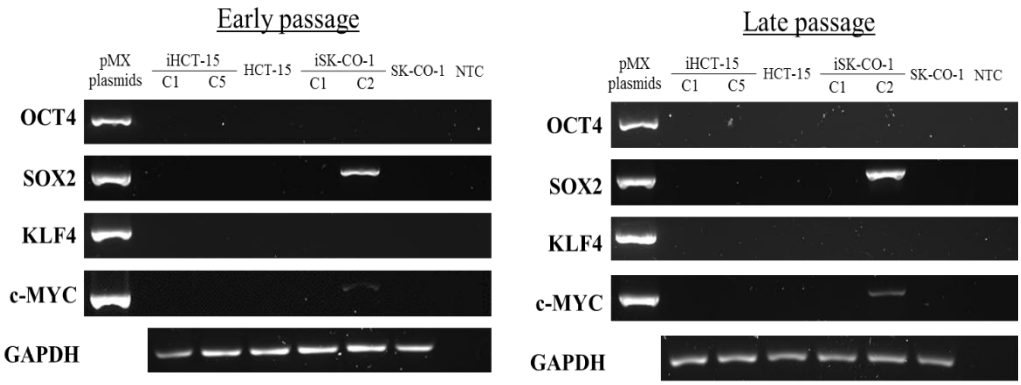


Figure 4.11: Ectopic expression of reprogramming factors OSKM in early and late passages of CRC-iPCs. Expression of the exogenous OSKM factors was examined in CRC-iPC cell lines of early (passage 10 or earlier) and late (passage >25) passage by direct RT-PCR. Individual pMX plasmid was used as the positive control for each gene expression. The respective parental cell lines served as non-OSKM-transduced negative controls. Non-template control (NTC) was used as negative control and GAPDH as the reference gene in RT-PCR.

4.7 Down-regulated Expression of Pluripotency Genes on Reprogramming

To obtain further confirmation on the reprogramming status of the CRC-iPC clones, endogenous expression of pluripotency genes in the CRC-iPCs clones and parental CRC cell lines was next investigated using the H9 ESC as a reference. Direct RT-PCR results showed that while all the genes were expressed in H9 ESC cells, the parental CRC cells expressed the pluripotency genes, *OCT4*, *SOX2*, *KLF4*, *c-MYC*, *NANOG*, *DNMT3B*, *LEFTY1* and *REX1*, but not *GDF3*; *DNMT3B* was expressed at very low levels (Figure 4.12). The observation is consistent with other reports on the expression of pluripotency genes in cell lines derived from various types of cancers (Schoenhals et al., 2009; Amini et al., 2014). The CRC-iPC clones were observed to share a similar expression pattern of the pluripotency genes as the respective parental CRC cell lines (Figure 4.12).

To further quantify the expression changes of *OCT4*, *SOX2*, *KLF4*, *c-MYC*, *NANOG* and *REX1* on reprogramming, the relative expression levels between the parental CRC and CRC-iPCs were determined by real-time quantitative RT-PCR. Post-iPC cell lines were also included in the pluripotency gene expression analysis. For clarity, a summary of pairwise comparison of the relative expression levels between the CRC, CRC-iPC and post-iPC cells is shown in Table 4.1. The results showed that, the expression of aforementioned pluripotency genes was generally down-regulated in all four CRC-iPC clones (Figure 4.13). The only exception was that *SOX2* expression

in the two iHCT-15 clones, and that *REX1* was slightly up-regulated in iHCT-15 Clone 5 (Figure 4.13).

On the other hand, expression of the pluripotency genes was up-regulated in all the post-iPC cell lines. Notably, *OCT4* and *SOX2* expression in iHCT-15 Clone 1 and Clone 5 was restored to an expression level comparable to that of the parental HCT-15 (Figure 4.13). Interestingly, *KLF4* expression was further down-regulated in the post-iHCT-15 cells. Similarly, up-regulation of *OSKM*, *NANOG* and *REX1* expression was also observed in the post-iSK-CO-1 cells (Figure 4.12).

In summary, the pluripotency genes were found to be already expressed in the CRC cell lines used, but expression of the pluripotency genes were generally down-regulated on reprogramming, but not totally shutdown. The collective data of exogenous and endogenous expression of pluripotency genes in iSK-CO-1 Clone 2 are consistent with that iSK-CO-1 Clone 2 was reprogrammed to a pre-iPSC state and raises the possibility of partial reprogramming (Plath & Lowry, 2011).

As mentioned in Materials and Methods Section 3.1.1, HCT-15 and SK-CO-1 cell lines were established from different origins of CRC, which may result in clonal heterogeneity of the iPC clones derived (Mills et al., 2013; Choong et al., 2014), and thus the different gene expression profiles observed in the iPC and post-iPC cells. However, the CRC-iPC clones thus derived were maintained up to 60 passages (data not shown), implying that the maintenance

of pluripotency and self-renewal properties of the reprogrammed colorectal cancer cells may not be entirely dependent on the expression levels of the pluripotency genes.

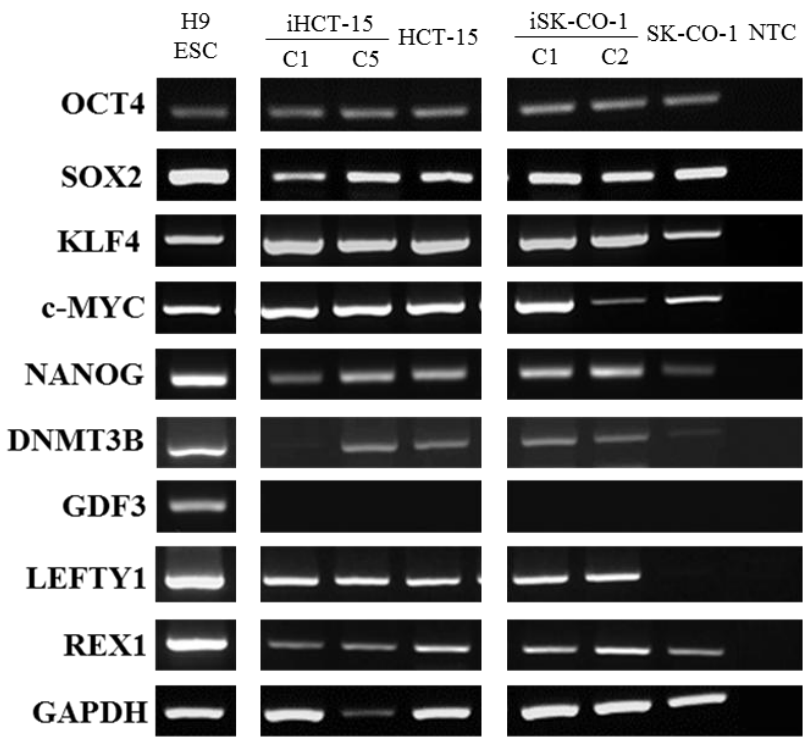


Figure 4.12: Expression of pluripotency genes in H9 ESC, CRCs and CRC-iPCs. Expression of the pluripotency genes in each cell line was examined through direct RT-PCR. Two representative clones from each iPC cell line were included. H9 ESC was used as a positive control for gene expression. Non-template control (NTC) was used as negative control and GAPDH as the reference gene in RT-PCR.

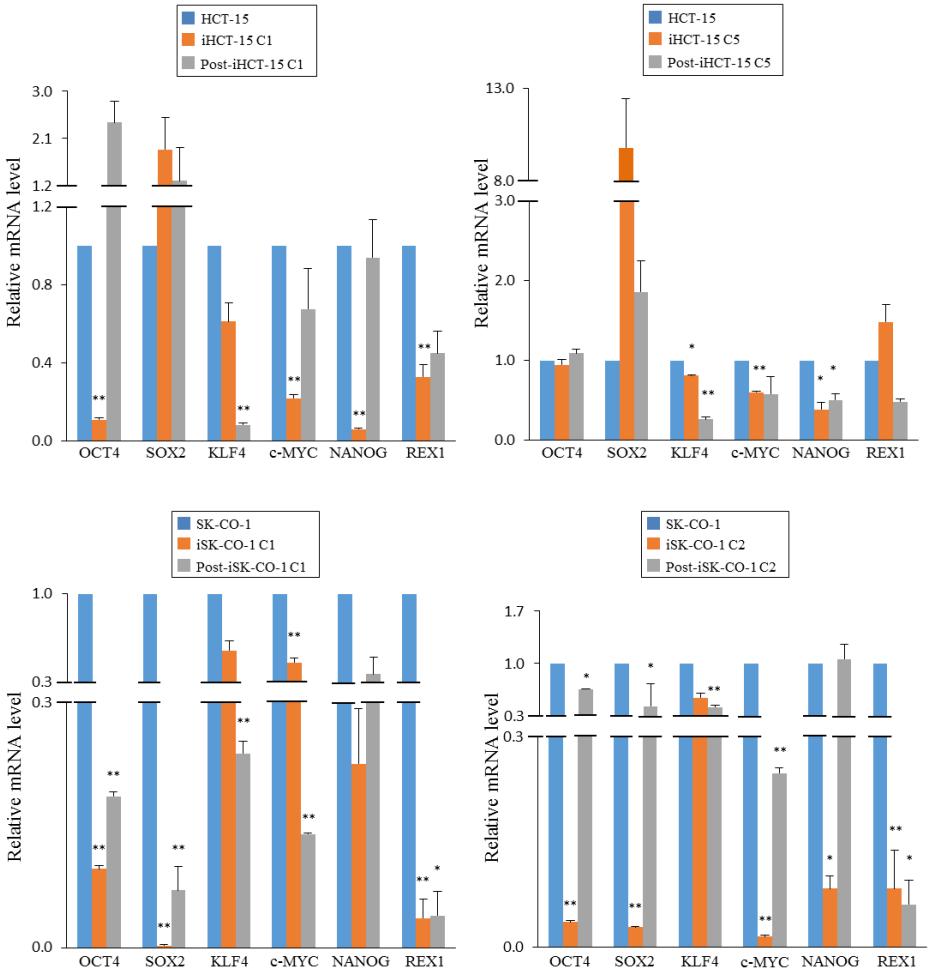


Figure 4.13: Relative expression levels of pluripotency genes in CRC-iPC clones and the derived post-iPCs. Real-time RT-PCR data presented were from three independent experiments in triplicates. Expression levels of pluripotency genes were normalised to that of the respective parental CRC, which was arbitrarily set as 1.0. *p<0.05 and **p<0.01.

Table 4.1: Relative expression of pluripotency genes between CRC, iPC and post-iPC

Gene	HCT-15		SK-CO-1		General expression trend
	C1	C5	C1	C2	
			A. Expression in iPC relative to CRC		
OCT4	2-	Unchanged	2-	4-	Down-regulated
SOX2	+	2+	4-	4-	Up-regulated in iHCT-15 clones; down-regulated in iSK-CO-1 clones
KLF1	-	-	-	-	Down-regulated
c-MYC	2-	-	2-	4-	Down-regulated
NANOG	3-	2-	2-	3-	Down-regulated
REX1	2-	+	4-	3-	Down-regulated
			B. Expression in post-iPC relative to iPC		
OCT4	4+	Unchanged	+	3+	Up-regulated
SOX2	-	2-	4+	3+	Down-regulated in iHCT-15 clones; up-regulated in iSK-CO-1 clones
KLF1	2-	2-	2-	-	Down-regulated
c-MYC	2+	Unchanged	2-	3+	Up-regulated
NANOG	3+	+	+	3+	Up-regulated
REX1	+	2-	Unchanged	-	Up/down-regulated

"+" to "4+"' and "-" to "4-" indicate > 20-, 11-19.9-, 2-10- and < 2.0-fold up or down-regulation, respectively between the pair in comparison. "Unchanged" is up- or down-regulated between +1.2- to -1.2-fold change.

4.8 Expression of MET/EMT-related Genes in Colorectal Cancer-derived Induced Pluripotent Cancer Clones

Cellular morphology of human ESCs and iPSCs are recognised by the form of cell colonies with tightly packed cells within (Courtot et al., 2014). Such cellular morphology indicates high degree of cell-cell contact attributable to the mesenchymal-epithelial transition (MET) process (Pieters and van Roy, 2014). In line with this, MET is promoted during somatic cell-cellular reprogramming, and is exemplified by up-regulation of epithelial-associated genes, including *E-CADHERIN* (*CDH1*) and *OCCLUDIN* (*OCLN*) (Chen et al., 2012). Hence, the expressions of genes of MET and epithelial-to-mesenchymal transition (EMT) serve as important indicators of cell state. In this work, the EMT/MET gene expression in the CRC-iPC cells was determined by real-time RT-PCR. Intriguingly, real-time RT-PCR results showed significant down-regulation of the MET-associated genes, *CDH1* and *OCLN*, in all four CRC-iPC clones, when compared to that of the parental CRC cell lines (Figure 4.14). Among the four CRC-iPC clones, only iHCT-15 Clone 5 showed minimal down-regulation of *CDH1* and *OCLN* expression. On re-differentiation, the expression of *CDH1* and *OCLN* was up-regulated again in three post-iPC cell lines, except for post-iSK-CO-1 Clone 1, which showed minimum or no significant changes in the expression levels. However, the expression levels of *CDH1* and *OCLN* in post-iPCs was not restored to that of the respective parental cells.

On the contrary, expression of the EMT-related genes, SNAI1 (*SNAI1*) and VIMENTIN (*VIM*), was generally up-regulated in the CRC-iPC clones and the derived post-iPC cell lines. Expression of *SNAI1* showed significant up-regulation only in iSK-CO-1 Clone 1 (Figure 4.15). Interestingly, the expression was still highly up-regulated relative to the parental CRC and iPC clones on re-differentiation in all four post-iPC cell lines (Figure 4.15). In contrast, *VIM* expression was significantly and extensively up-regulated in all four CRC-iPC clones. The expression level of *VIM* was down-regulated in all four of the post-iPC cells on re-differentiation. However, the expression levels of *SNAI1* and *VIM* in the post-iPC cells were still higher when compared to the corresponding parental cells, except for *VIM* expression in post-HCT-15 Clone 5.

Taken together, expression of the MET-associated genes was surprisingly down-regulated in the reprogrammed CRC cells, unlike the somatic cell-derived iPSCs (Li et al., 2010). On the other hand, up-regulation of *SNAI1* and *VIM* was observed in CRC-iPC cells as anticipated in response to the suppression of epithelial-related genes. Analysis of the molecular phenotype of the MET/EMT gene has collectively suggested an inclination towards a mesenchyme-like cell state in the CRC-iPC cells on reprogramming.

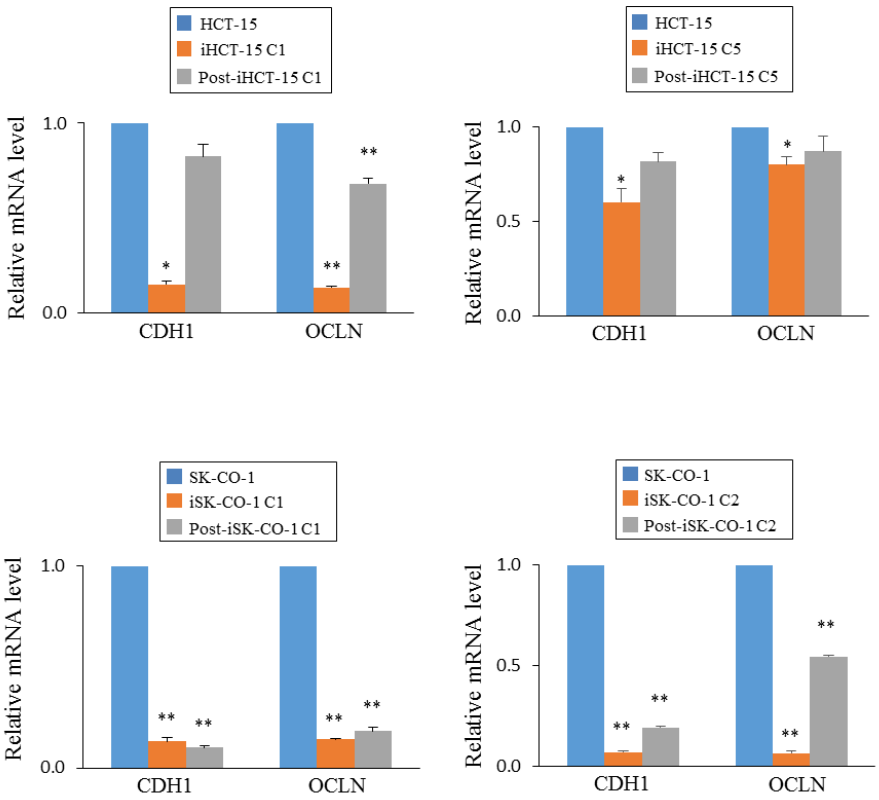


Figure 4.14: Relative expression of MET genes in CRC-iPC clones and the derived post-iPC. Real-time RT-PCR data presented were from three independent experiments in triplicates. Expression levels of MET genes were normalised to that of the respective parental CRC, which was arbitrarily set as 1.0. *p<0.05 and **p<0.01.

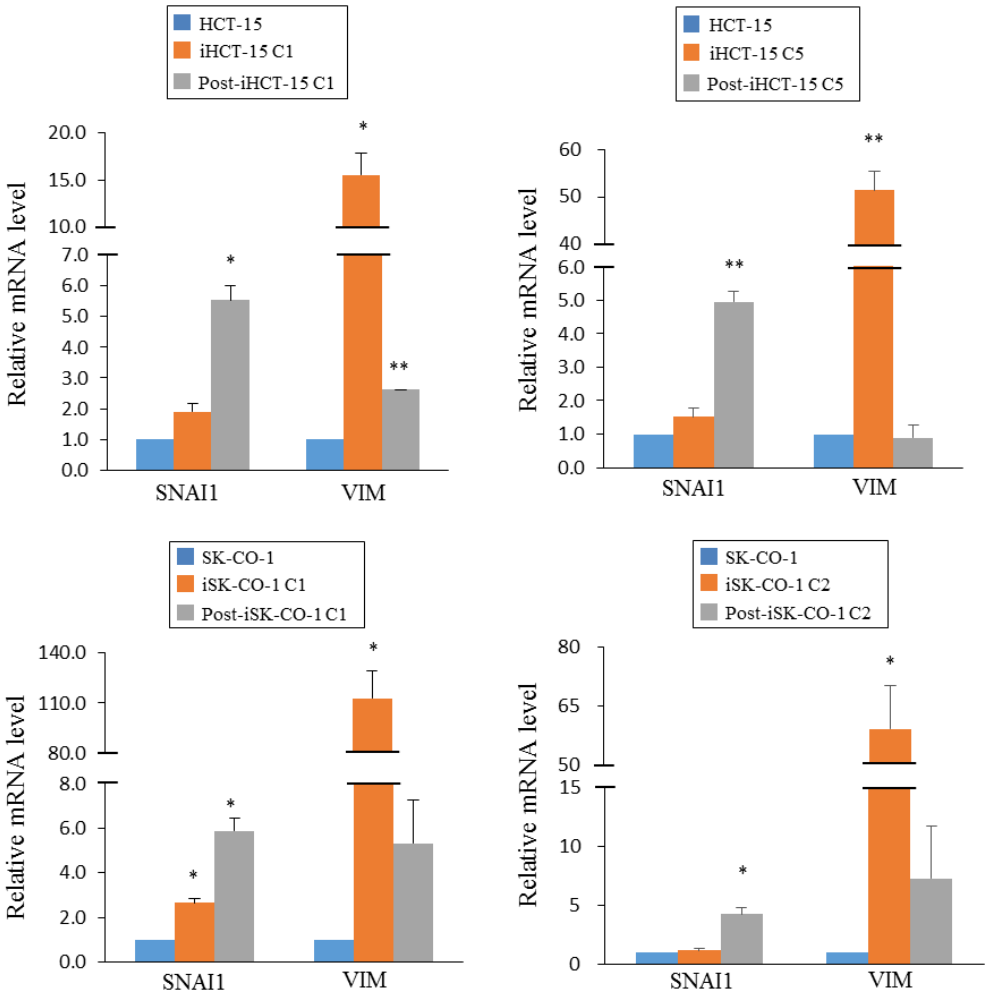


Figure 4.15: Relative expression of EMT genes in CRC-iPC clones and the derived post-iPC. Real-time RT-PCR data presented were from three independent experiments in triplicates. Expression levels of EMT genes were normalised to that of the respective parental CRC, which was arbitrarily set as 1.0. *p<0.05 and **p<0.01.

4.9 Genome-wide MicroRNA (miRNA) Profiling of in Colorectal Cancer-derived Induced Pluripotent Cancer Clones and Colorectal Cancer Cell Lines

Previous reports have shown altered miRNA expression in iPSCs and iPCs (Lüningschrör et al., 2013; Koga et al., 2014) compared to the parental cells. In addition, miRNA-mediated cellular reprogramming has been proven to be effective in somatic cells and cancer cells (Pourrajab et al., 2014; Koga et al., 2014). These reports indicate the importance of miRNAs in regulating pluripotency and cell fate. However, studies on the miRNA expression profile in reprogrammed cancers are lacking. In order to establish a genome-wide miRNA expression profile of iPCs, the four CRC-iPC clones described above and the respective parental CRC cell lines were subjected to miRNA microarray analysis using the Agilent SurePrint Human MiRNA Microarray Release 21.0. The iPC miRNA profile was compared to that of the parental CRC cell lines.

Hierarchical clustering was first performed on the microarray data to display the miRNA expression patterns between the iPCs and CRCs. From the hierarchical clustering diagram (Figure 4.16), parental CRC cell lines were clustered in a group (red vertical bar), whereas the four CRC-iPC cell lines were clustered in another group (Figure 4.16; blue vertical bar on the left). The clustering results indicated distinct miRNA expression patterns between the CRC and iPC groups, which is attributed to OSKM reprogramming. Interestingly, SK-CO-1 showed distinct miRNA expression pattern in the

clustergram (Figure 4.16) when compared to the other CRC cell line HCT-15. The possible explanation of this phenomenon may be that SK-CO-1 was derived from a metastatic site while HCT-1 is adenocarcinoma *in situ.* The distinctive miRNA expression patterns of HCT-15 and SK-CO-1 may suggest tumour stage-specific miRNA profiles (Choo et al., 2014).

In order to identify the differentially-expressed miRNAs, volcano plot analysis was conducted using the selection criteria of $\log_2$ (FC) $\geq$ 2.0 or $\leq$ -2.0, and a *p*-value < 0.05. The identified miRNAs are presented (blue cubes) above the green line in the volcano plot (Figure 4.17). A total of 102 differentially-expressed miRNAs were identified, 52 of which were up-regulated and 50 were down-regulated (Figure 4.17). The full list of the 102 miRNAs is shown in Appendix A and B. The miRNA microarray data have been deposited in the NCBI GEO archive with the accession number GSE87280.

Ten miRNAs with the highest $\log_2$ (fold change) in up- and down-regulation are shown in Table 4.2. Among the top 10 up-regulated miRNAs, five miRNAs, miR-6789-5p, miR-1181, miR-199a-3p, miR-150-3p and miR-125a-3p, were found on chromosome 19 (Table 4.2). Interestingly, among the 20 top up- and down-regulated miRNAs, five are clustered on chromosome 1, four miRNAs belongs to chromosome 17 and five are mapped on chromosome 19, in particular in the 19q13.2 – 19q13.4 region. The data appear to indicate OSKM-reprogramming may have resulted in major expression changes of miRNAs mapping in chromosomes 1, 17 and 19.

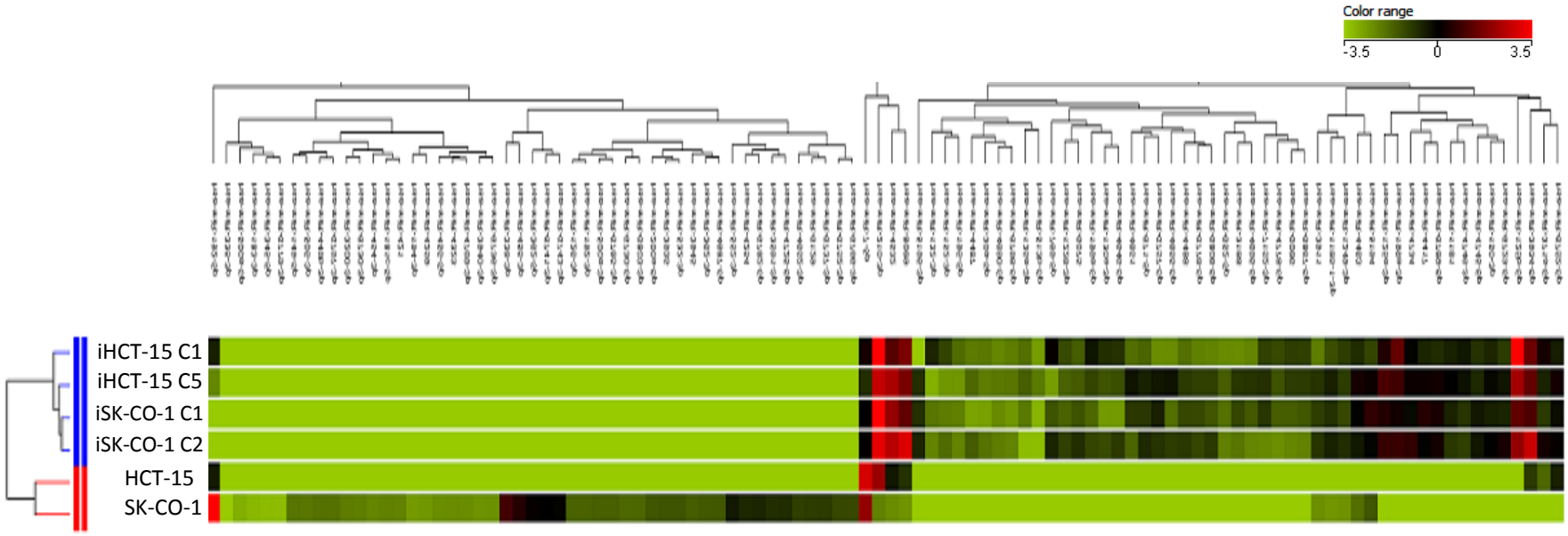


Figure 4.16: Hierarchical clustering analysis of miRNA profiles of parental CRC and CRC-iPC clones. Colour-code bar indicates the relative expression level of a miRNA: maximum expression is in red, minimum expression in green. The analysis was performed and generated by GeneSpring GX software version 13.0.

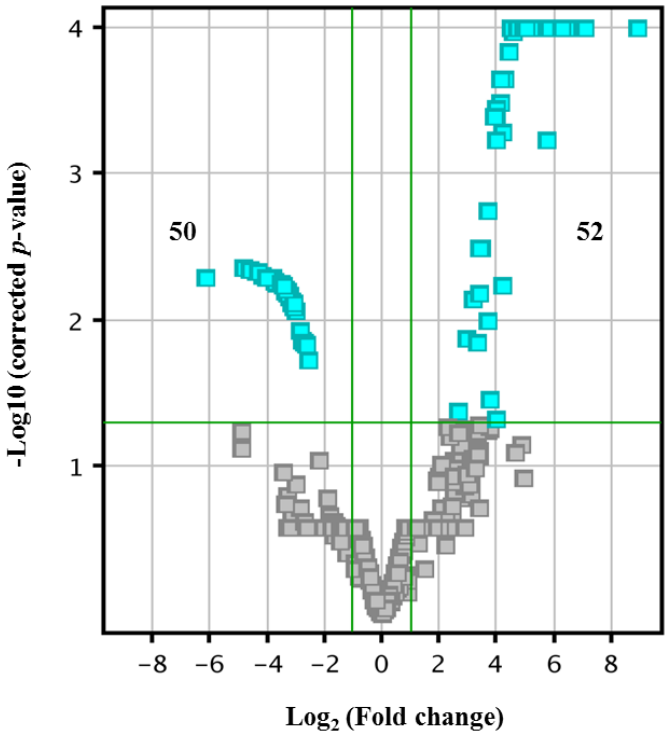


Figure 4.17: Volcano plot analysis of miRNA profiles of parental CRC and CRC-iPC clones. The analysis was performed and generated by GeneSpring GX software version 13.0. MiRNAs that were significantly and differentially-expressed with $\log_2$ (fold change) $\geq$ 2.0 or $\leq$ -2, and *p*-value <0.05 (light blue-colored cubes), were identified.

Table 4.2: Top 10 differentially-expressed miRNAs in reprogrammed CRC-iPCs

miRNA	Chromosome	miRNA family[1]	Log_2 FC[2]
	A. Up-regulated		
miR-4417	1	NA	6.21
miR-6723-5p	1	NA	5.77
miR-3934-5p	6	mir-3934	5.71
miR-125b-5p	11q24.1	mir-10	8.88
miR-4734	17	NA	6.36
miR-6789-5p	19	NA	6.24
miR-1181	19	mir-1181	5.67
miR-199a-3p	19p13.2	mir-199	7.04
miR-150-3p	19q13.33	mir-150	5.86
miR-125a-3p	19q13.41	mir-10	6.89
	B. Down-regulated		
miR-6741-3p	1	NA	-4.37
miR-4254	1	NA	-4.02
miR-552-3p	1p34.3	mir-552	-4.15
miR-455-3p	9q32	mir-455	-4.63
miR-6743-3p	11	NA	-4.34
miR-192-5p	11q13.1	mir-192	-6.17
miR-4725-5p	17	NA	-3.99
miR-6782-5p	17	NA	-4.05
miR-338-3p	17q25.3	mir-338	-4.84
miR-362-5p	Xp11.23	mir-362	-4.42

[1]miRBase version 21 was used in data derivation.[2] Microarray data of Log_2 fold change (FC) relative to the parental CRC cells with $p < 0.05$. NA, not annotated

4.10 Validation of MicroRNA (miRNA) Expression in Colorectal Cancer-derived Induced Pluripotent Cancer Clones

To validate the microarray data, three miRNAs each from the up- and down-regulated miRNA groups were randomly selected for data validation using miRNA real-time qRT-PCR. To more accurately compare the expression levels of the miRNAs, the mean of the $\log_2$ (FC) of four CRC-iPC clones relative to the respective parental CRC cell lines was taken (Figure 4.18). The real-time RT-PCR results showed that miR-500a-3p, miR-362-5p and miR-532-3p were down-regulated, consistent with the microarray data. Similarly, expression of miR-125b-5p, miR-199a-3p and miR-150-3p was up-regulated, as in the microarray data (Figure 4.18). Thus, the qRT-PCR data supported the validity and reliability of the microarray-generated profiling data.

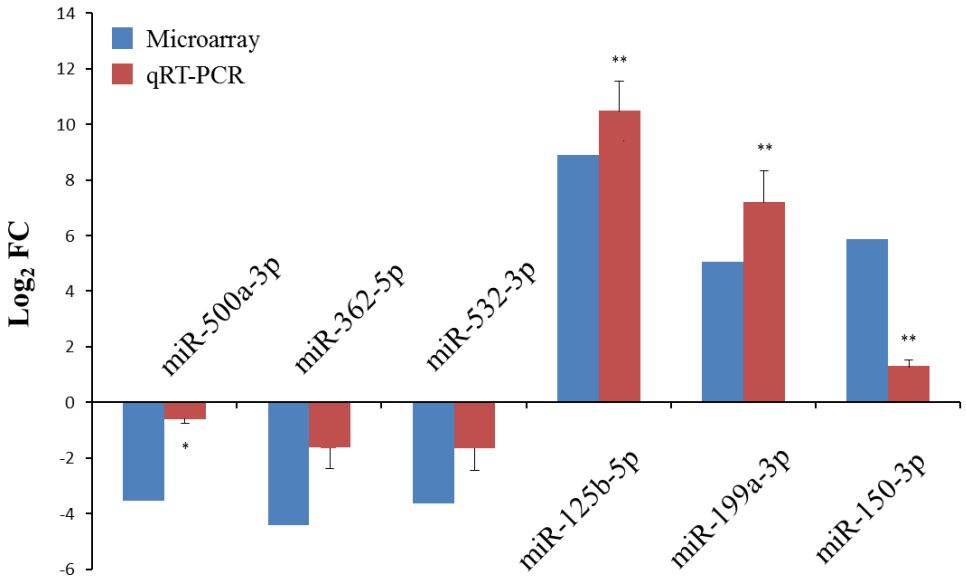


Figure 4.18: Validation of differentially-expressed miRNAs by quantitative real-time PCR. Expression validation of six randomly-selected miRNAs in CRC-iPC clones by real-time RT-PCR. The data shown were the mean of data obtained from all four iPC clones used in the microarray analysis. Real-time RT-PCR data presented were from three independent experiments in triplicates. Expression levels of miRNAs were normalised to that of the respective parental CRC. *p<0.05 and **p<0.01.

4.11 Gene Ontology Analysis of Predicted Target Genes

The differentially-expressed miRNAs were used to predict the targeted transcripts by interrogating the TargetScan and MicroRNA.org databases. The putative target genes were further analyzed using DAVID to determine the Gene Ontology (GO) annotation dataset based on biological processes, cellular components and molecular functions (Figure 4.19) (Huang et al., 2009a; Huang et al., 2009b). The *p*-value for each GO term indicated the probability that observed numbers of total gene counts could have been resulted from inaccurate distribution to each GO term, hence, the most specific GO term is expected to have the lowest *p*-values.

The top 10 biological processes assigned from the analysed gene list were reproductive developmental process, ectoderm development, regulation of cell motion, regulation of locomotion, epidermis development, regulation of cell migration, rhythmic process, regulation of cell morphogenesis, regulation of lipid metabolic process, positive regulation of cell motion (Figure 4.19A). Notably, five out of the ten assigned GO terms were involved in cell motility, indicating that the MET/EMT processes are crucial in cancer cell reprogramming. In fact, cell migration during early reprogramming was reported by Megyola et al. (2013), indicating involvement of EMT in the early phase of reprogramming (Megyola et al., 2013). Cell locomotion, cell migration and cell morphogenesis were previously associated with EMT during gastrulation and differentiation during early embryo development (Ferrer-Vaquer et al., 2010).

The top 10 cellular locations of the gene products (cellular components) indicated involvement in regulation of cellular activities in the mitochondrion, golgi apparatus, mitochondrial part, cell junction, synapse, synapse part, mitochondrial matrix, mitochondrial lumen, ribosome, and ribosomal subunit (Figure 4.19B). The assigned GO in cellular components showed a major involvement of mitochondria. One possible explanation would be that mitochondrial activities were previously shown to induce TGF-β signalling, and thus up-regulating mesenchymal markers (Naito et al., 2008). Moreover, most gene products were also found to regulate cellular activities at cell junctions, including cell-cell adhesion, which may affect the cell state (Pieters and van Roy, 2014). Therefore, in relation to this study, the assigned GO terms indicated crucial MET/EMT regulation in the reprogramming of CRC cells.

The top 10 molecular functions of the predicted target genes of the differentially-expressed miRNAs were involved in protein kinase activity, enzyme inhibitor activity, structural constituent of ribosome, nuclease activity, extracellular ligand-gated ion channel activity, protein phosphatase regulator activity, phosphatase inhibitor activity, carboxylase activity, phosphoprotein phosphatase inhibitor activity and phosphopyruvate hydratase activity (Figure 4.19C). Protein kinases and phosphatase activities are crucial in signal transduction (Bononi et al., 2011). Complex networks involving FGF-2 induce phosphorylation of phosphatidylinositol-3 kinase (PI3K)-AKT, mitogen-activated protein kinase-ERK-1/2 kinase (MEK), and PKC/ERK-1/2 kinase, all of which serve to enhance long-term self-renewal of pluripotent stem cells

(Kinehara et al., 2013). Role of phosphatase PTEN is also well-characterised in stem cells, and is involved in regulation of cell-cycle progression, cell proliferation, DNA damage response (Bononi et al., 2011).

Taken together, the gene ontology analysis suggests that the miRNA-regulating putative target genes may be involved in cell motility, a process governed by MET/EMT in line with the findings in Results Section 4.7.

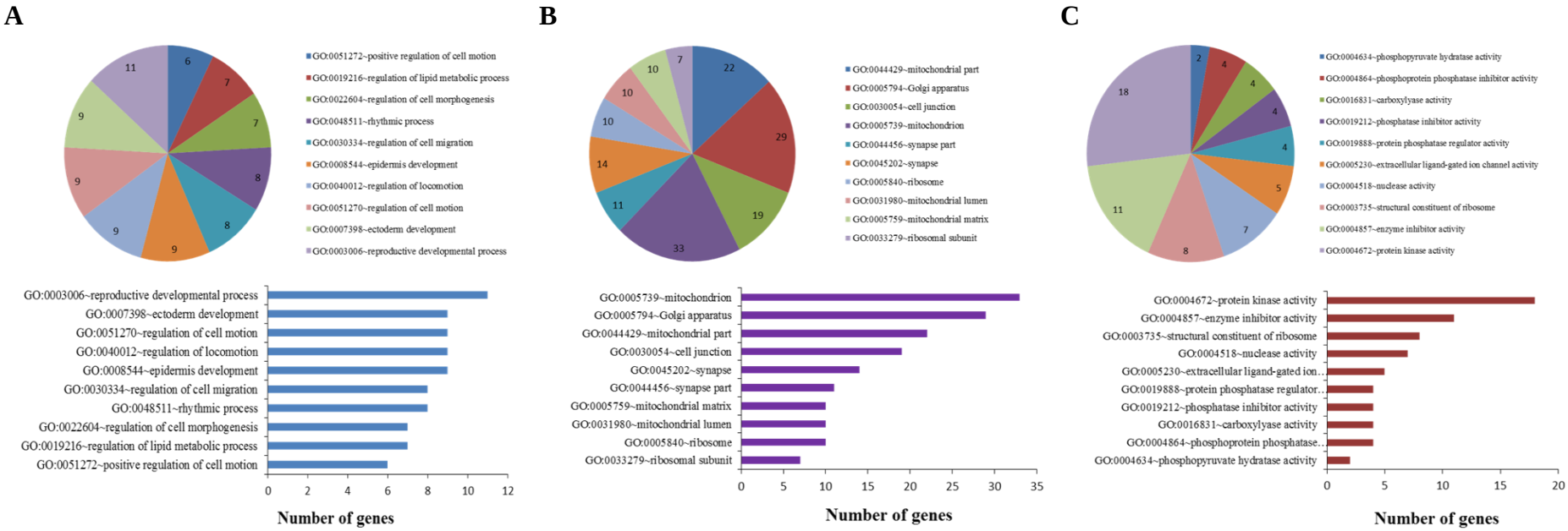


Figure 4.19: Gene ontology analysis of the predicted target transcripts of 102 differentially-expressed miRNAs. Top 10 enriched gene ontology (GO) terms in (A) biological processes (B) cellular components and (C) molecular functions are shown. The data were generated by the Database for Annotation, Visualization, and Integrated Discovery (DAVID) algorithm.

4.12 Predicted Pathways Regulated by Differentially-expressed MicroRNAs (miRNAs)

The KEGG pathways analysis based on the predicted target genes of the differentially-expressed miRNAs showed possible involvement in cytokine-cytokine receptor interaction (hsa04060), B-cell receptor signalling pathway (hsa04662), transforming growth factor-beta (TGF-β) signalling (hsa04350), Fc gamma r-mediated phagocytosis (hsa04666), Fc epsilon RI signalling pathway (hsa04664), chronic myeloid leukaemia (hsa05220), pancreatic cancer (hsa05212), ABC transporters (hsa02010), and base excision repair (hsa03410) (Figure 4.20). The predicted genes involved in each pathway are shown in Table 4.3.

Cytokine-cytokine receptor interaction was ranked the first in the KEGG pathway, with 12 predicted genes involved (Figure 4.20). Cytokines are crucial intercellular regulators engaged in numerous cellular processes, including maintaining pluripotency and promoting reprogramming (Kristensen et al., 2005). Cytokine-cytokine receptor interaction was first identified as one of the novel developmental signalling pathways in mouse ESC, and the pathway was regulated by the endogenous Yamanaka factors (Liu et al., 2008; Huang et al., 2009). On the other hand, B-cell receptor signalling pathway was ranked second in the pathway prediction with eight genes matched. Activation of B-cell receptor has been shown to lead to subsequent modulation of other downstream signalling cascades, including PI3K-AKT, MAPK, NF-κB signalling pathways (Woyach et al., 2012). These pathways were previously

associated with maintenance of pluripotency and the undifferentiated state of ESC and iPSC (Takase et al., 2013). The results collectively may indicate activation of pluripotency-related pathways in the reprogrammed CRC cells.

Interestingly, seven of the predict target transcripts were enriched in the TGF-β signalling pathway (Table 4.4). Previous reports have shown that inhibition of the TGF-β pathway enhances reprogramming by activation of the MET process. However, the EMT-inducing TGF-β pathway (Xu et al., 2009) coincides with the up-regulated EMT gene expression (Figure 4.15). Notably, AKT3 is enriched in five out of the nine predicted pathways, implicating involvement of PI3K-AKT in cellular reprogramming, as has also previously been reported (Tang et al., 2014). The possible involvement of TGF-β and PI3K-AKT signalling pathways are further discussed in Results Section 4.12.

Despite correlation of the predicted pathways with pluripotency or cellular reprogramming described above, the relationship between the predicted Fc gamma R-mediated phagocytosis and Fc epsilon RI signalling pathways with pluripotency and reprogramming is currently unclear. However, Fc gamma R-mediated phagocytosis is responsible in regulating actin cytoskeletons via RAF1 (Wang et al., 2013), VAV2 (Liu and Burridge, 2000), and AKT3 (Xue and Hemmings, 2013), which contribute to cell motility (Shankar and Nabi, 2015), and in agreement with the data of GO analysis (Figure 4.19). Fc epsilon RI signalling pathway stimulates the production of cytokines (Hernandez-Hansen et al., 2005), which may be involved in cytokine-cytokine receptor interactions. Additionally, chronic leukemia and

pancreatic cancer have been shown to have the same set of overlapping predicted genes, including *TGFβ2*, *SMAD4* from TGF-β signalling, and *AKT3* from the PI3K-AKT signalling pathway. The data indicate convergence of signalling pathways in regulating pluripotency and tumourigenesis (Mooney et al., 2013).

ATP-binding cassettes transporters (ABC transporters) mediate drug-resistant mechanisms in cancer, and are also known to protect stem cells from toxic substances (Erdei et al., 2014). Particularly, ABCG2-expressing stem cells showed higher toleration to stress, such as physical stress, drugs and UV light exposure (Erdei et al., 2013). Furthermore, previous reports have also stated that the ABC transport protein families are selectively expressed in pluripotent stem cells (Hirata et al., 2014; Kuo et al., 2014). Hence, with further validation, the predicted target genes, *ABCG5*, *ABCC3*, *ABCB7* and *ABCA13*, may serve as specific markers in distinguishing between various types of stem cells.

On the other hand, base-excision repair (BER) is known to safeguard the genome integrity and protects the cells from DNA damages (Rocha et al., 2012). Previous reports have shown that BER-related genes are highly expressed in pluripotent stem cells, when compared to differentiated cells, to accommodate continuous mutation-free self-renewal and differentiation (Luo et al., 2012). Hence, the predicted activated BER pathway in CRC-iPC clones may have contributed to the maintenance of genome stability during reprogramming (Momcilovic et al., 2010). Furthermore, BER-related proteins

are also involved in DNA methylation and histone modification in epigenetic regulation of gene expression (Li et al., 2013). Based on the observed reversibility of gene expression, the expression changes in the pluripotency and MET/EMT genes may be due to epigenetic regulation involving BER-related proteins (Results in Section 4.7-4.8).

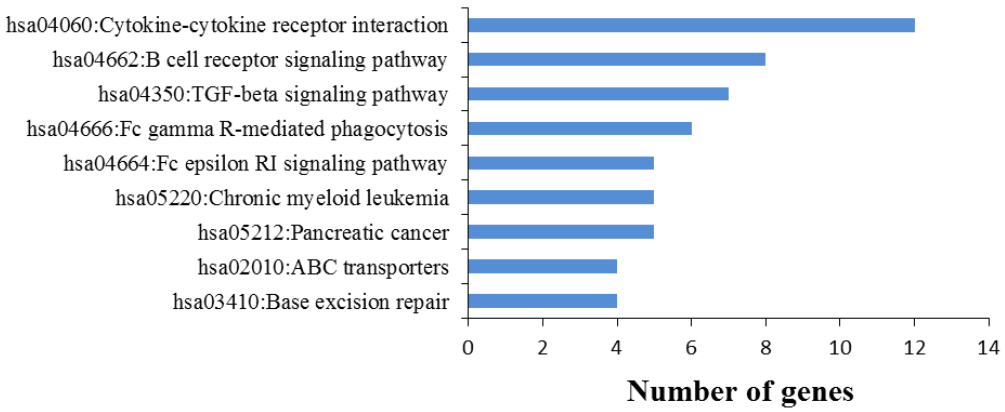


Figure 4.20: KEGG pathway analysis of predicted target genes. Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis was conducted via Database for Annotation, Visualization, and Integrated Discovery (DAVID). Predicted targets genes involved in the enriched KEGG pathways are also shown in Table 4.3.

Table 4.3: Predicted KEGG pathways and target genes of the CRC-iPC differentially-expressed miRNAs[1]

KEGG ID	KEGG pathway	Predicted target genes involved	Fold Enrichment	*p*-value
hsa04060	Cytokine-cytokine receptor interaction	CCL1, INHBB, TNFRSF6B, ACVR2A, TNFRSF1B, ACVRL1, TNFRSF19, IL15, MPL, EDA, KDR, TGFB2	2.06	0.028
hsa04662	B cell receptor signalling pathway	CARD11, MAP2K1, LYN, CD22, RAF1, CD72, VAV2, AKT3	4.8	0.001
hsa04350	TGF-beta signalling pathway	INHBB, ACVR2A, ACVRL1, SMAD4, RPS6KB1, BMP8B, TGFB2	3.62	0.011
hsa04666	Fc gamma R-mediated phagocytosis	MAP2K1, LYN, RAF1, RPS6KB1, VAV2, AKT3	2.84	0.057
hsa04664	Fc epsilon RI signalling pathway	MAP2K1, LYN, RAF1, VAV2, AKT3	2.88	0.092
hsa05220	Chronic myeloid leukaemia	MAP2K1, SMAD4, RAF1, AKT3, TGFB2	3.00	0.082
hsa05212	Pancreatic cancer	MAP2K1, SMAD4, RAF1, AKT3, TGFB2	3.12	0.073
hsa02010	ABC transporters	ABCG5, ABCC3, ABCB7, ABCA13	4.09	0.071
hsa03410	Base excision repair	POLD3, POLE4, MBD4, NTHL1	5.14	0.040

[1]Arranged in decreasing total number of predicted genes involved.

4.13 Predicted miRNA-target mRNA Interactions in the TGF-β and PI3K-AKT Signalling Pathways

Based on the bioinformatics and KEGG analyses (Figures 4.18 - 4.19), the predicted target genes of the differentially-expressed miRNAs were found to be highly involved in the regulation of cell motility, and enriched in TGF-β and PI3K-Akt signalling pathways. The pathways were previously associated with maintenance of pluripotency, reprogramming and tumourigenesis (Mooney et al., 2013). Possible roles of TGF-β and PI3K-Akt pathways in relation to reprogramming using the schemes created, with depiction of miRNA-mRNA interactions are proposed in scheme presented in Figures 4.20 - 4.21.

In the TGF-β pathway, TGFβ2, which is a ligand of the TGFβ family, was predicted to be targeted by 10 miRNAs, seven of which, miR-7-5p, miR-149-5p, miR-345-5p, miR-362-5p, miR-371a-5p, miR-552-3p, and miR-500a-5p, were down-regulated; miR-132-3p, miR-152-3p, and miR-199a-5p were up-regulated. (Figure 4.21). Similarly, expression of another predicted target gene SMAD4 was regulated by seven differentially-expressed miRNAs, among which miR-362-5p, miR-421, miR-4259, miR-552-3p were down-regulated, while miR-130a-3p, miR-132-3p and miR-125a-3p were up-regulated (Figure 4.21). The down-regulation of these miRNAs may promote TGFβ2 and SMAD4 translation into functional proteins. Hence, the signalling cascade was promoted via the upstream TGFβ2 ligand binding, and/or downstream binding of the SMAD4/R-SMAD complex, to the promoters or cofactor regions of the

genes, resulting in transcription of the EMT inducers, *SNAI1*, *SLUG* and *ZEB1* (Figure 4.21). Additionally, activation of TGF-β signalling also leads to repression of epithelial genes, which also reverses the MET process. Notably, the down-regulated miRNAs, miR-362-5p and miR-552-3p, were concurrently predicted to target TGFβ2 and SMAD4 (Figure 4.21), indicating that the two miRNAs may play a vital role in reprogramming-induced TGF-β signalling pathway.

In PI3K-AKT signalling, an isoform of the protein kinase B AKT, namely AKT3, was predicted to be regulated by eight miRNAs, five of which were down-regulated (miR-149-5p, miR-181c-5p, miR-362-3p, miR-455-3p, and miR-4254) (Figure 4.22). On the other hand, three of the predicted AKT-regulating miRNAs were up-regulated, including miR-195-5p, miR-671-5p and miR-769-5p (Figure 4.22). Similarly, up-regulation of AKT is essential to promote cell proliferation and survival of reprogrammed cells, and thus, enhances the reprogramming efficiency (Tang et al., 2014). Furthermore, previous studies have reported the role of AKT-induced EMT, which subsequently promotes the mobility and invasiveness of various cancers (Grille et al., 2003). Phosphorylation of AKT has been shown to facilitate degradation of GSKβ3 (Xu et al., 2016), allowing accumulation of β-catenin and SNAI1, which subsequently bind to TCF/LEF, promoting gene transcription of EMT-associated genes (Savagner, 2010).

Both the TGF-β and PI3K-AKT signalling pathways may individually or act synergistically to promote the transition towards to mesenchyme-like state

(EMT), contributing to the de-differentiation process. Taken together, the miRNA-regulated TGF-β and PI3K-AKT signalling pathways may provide insights of the underlying mechanisms of cancer reprogramming, although further validation is required.

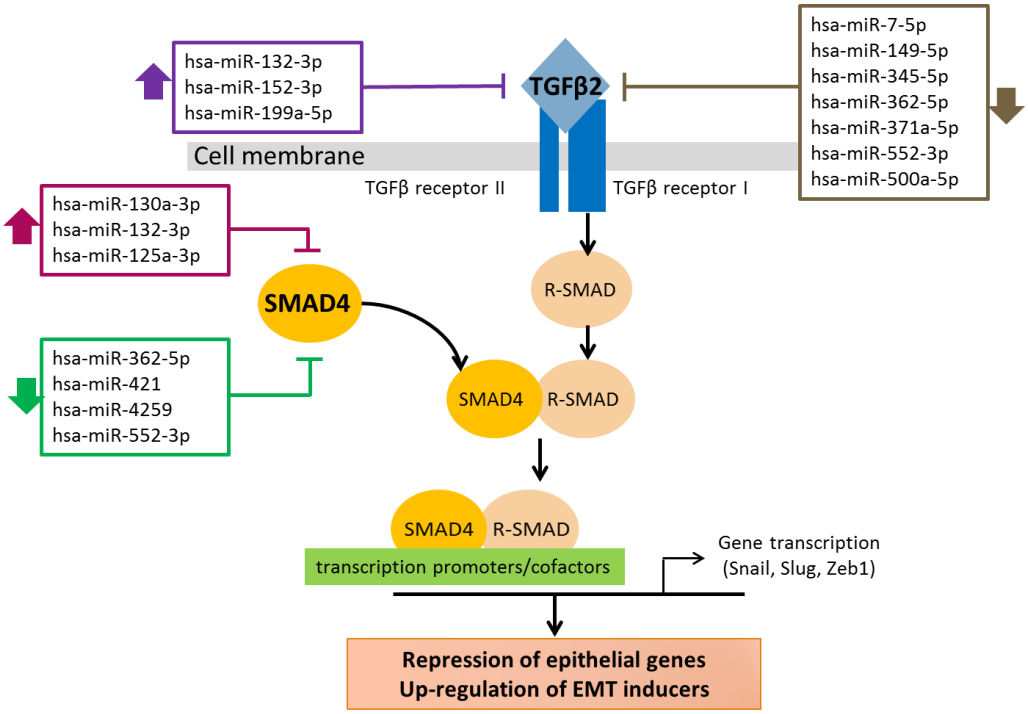


Figure 4.21: Predicted miRNA-mRNA interactions in the TGF-β signalling pathway. The predicted transcripts targeted by miRNAs are shown in boxes; up- and downward arrows indicate the up or down-regulated miRNAs. The TGF-β pathway scheme was modified from Miyazono et al., 2009.

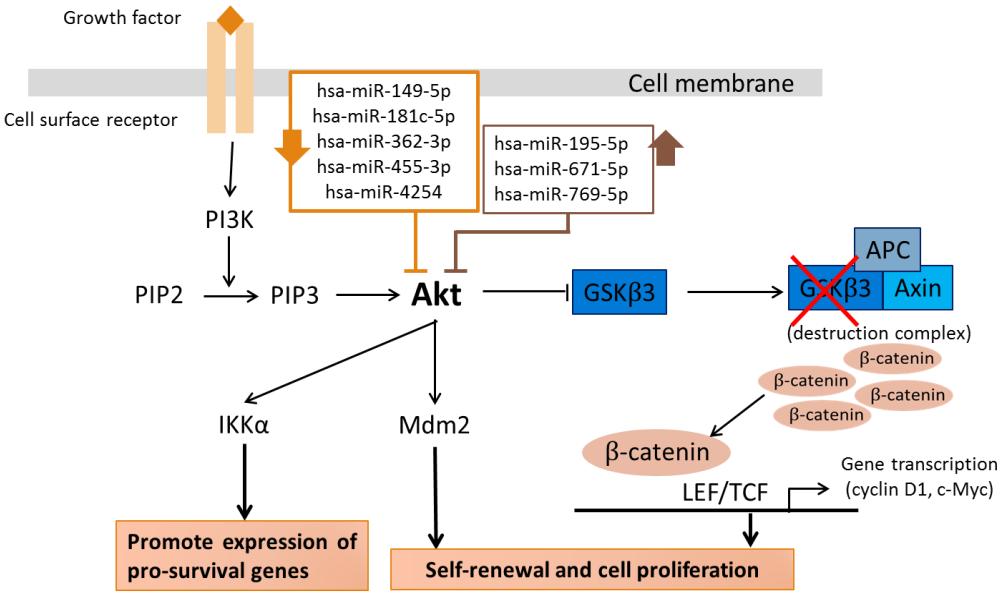


Figure 4.22: Predicted miRNA-mRNA interactions in the PI3K-AKT signalling pathway. The predicted target transcripts targeted by miRNAs are shown in boxes; up- and downward arrows indicate the up or down-regulated miRNAs. The PI3K-AKT signalling scheme was modified from Matsuoka et al., 2014.

4.14 Prediction of Differentially-expressed MicroRNAs (miRNAs) Targeting MET/EMT genes

The prediction of miRNA involvement in EMT-inducing TGF-β and PI3K-AKT signalling pathways was consistent with the observation of down-regulation of MET genes, and up-regulation of EMT genes, in the four CRC-iPC clones (Figures 4.13 - 4.14). The data collectively indicated that the differentially-expressed miRNAs may also regulate MET/EMT processes by directly targeting at the MET genes, *CDH1* and *OCLN*, and the EMT genes, *SNAI1* and *VIM*, analysed in Result Section 4.7. The genes and the corresponding up- or down- regulating miRNAs are shown in Table 4.6. Through interrogation of online database, including TargetScan 7.0, miRWalk 2.0 and Diana tools (microT-CDS), *CDH1* was targeted by 10 miRNAs; *OCLN* was predicted to be regulated by 11 miRNAs. Similarly, the EMT genes, *SNAI1* and *VIM* were putatively targeted by four and six miRNAs, respectively. In short, the four MET/EMT genes analysed were found to be putatively regulated by one or more miRNA from the 102 differentially-expressed miRNA groups, suggesting that miRNA is an additional level of regulation in the MET/EMT processes in reprogrammed CRC cells. However, further functional studies are required to more directly demonstrate and validate the relationship of miRNA-targeted pathways and MET/EMT in modulation of cancer cell reprogramming.

Table 4.4: Differentially-expressed miRNAs targeting at the MET/EMT genes

Gene	miRNAs
	A. Up-regulated
	MET
CDH1	miR-195-5p, miR-671-5p, miR-769-5p, miR-4695-5p
OCLN	miR-132-3p, miR-513b-5p, miR-1228-3p, miR-4463
	EMT
SNAI1	miR-30a-5p, miR-199a-5p, miR-4695-5p
VIM	miR-30a-5p, miR-3188,miR-4745-5p
	B. Down-regulated
	MET
CDH1	miR-149-5p, miR-335-3p, miR-362-3p, miR-532-3p, miR-660-5p, miR-4725-5p,
OCLN	miR-335-3p, miR-362-3p, miR-362-5p, miR-449b-3p, miR-500a-5p, miR-660-5p, miR-3591-3p
	EMT
SNAI1	miR-194-3p
VIM	miR-7-5p, miR-335-3p, miR-500a-5p

Target genes were predicted by TargetScan 7.0, miRWalk2.0 and Diana tools (microT-CDS).

CHAPTER 5

DISCUSSION

5.1 Establishment of Colorectal Cancer-derived Induced Pluripotent Cancer Clones via OSKM-retroviral Transduction

In this study, the OSKM cocktail was used to induce pluripotency in CRC cell lines. OSKM-reprogramming is a well-established reprogramming protocol commonly used in somatic and cancer cell reprogramming (Hussein and Nagy, 2012). However, OSKM have been individually associated with cancer development by enhancing cell survival, invasion and EMT (Bose and Sudheer Shenoy, 2014). Overexpression of OSKM potentially implies high tumourigenesis potentials in the reprogrammed cells. OCT4 and SOX2 are further reported to form a core transcriptional regulatory circuitry with NANOG that regulates expression of pluripotency and somatic genes by binding to the respective promoters (Boyer et al., 2005). Furthermore, c-MYC is responsible for facilitating the expression of proliferation-related genes (Dang, 2012), and for amplifying transcription activities of active promoters (Lin et al., 2012; Nie et al., 2012). The OSKM cocktail also initiates chromatin remodelling to allow access to the pluripotency loci, and conversely, somatic loci are restricted (Xu et al., 2016).

Ectopic expression of OSKM in this study was delivered via retroviral vectors harbouring OCT4, SOX2, KLF4 and c-MYC. The retroviral-based delivery approach has successfully generated iPSC and iPC (Lim et al., 2016). One advantage of the retroviral vector approach is that termination of the transgene transcription occurs when the transduced cells have achieved pluripotent state (Medvedev et al., 2010). Hence, genuine iPSC and iPC cells are able to be identified as the cells solely rely on the endogenous expression of OSKM. However, retroviral DNA is randomly integrated into the host genome, hampering genomic stability (Takahashi et al., 2007). Despite the concerns of genomic instability in the resulting reprogrammed cells, retroviral-based delivery remains as one of the most efficient reprogramming approaches due to its high transduction efficiency (Nigro and Verfaillie, 2013).

In present study, retroviral transduction of OSKM generated CRC-iPC clones with features of pluripotent stem cells, including positive expression of pluripotency markers (Figures 4.3 - 4.7) and ability to undergo in vitro tri-lineage differentiation, be it lineage-directed or via embryoid body formation. Intriguingly, the CRC-iPC clones established in this study exhibited some dissimilarities in cellular morphology, gene expression and miRNA profiles when compared to the features reported for ESC and other somatic cell-derived iPSCs. The reprogrammed CRC colonies showed multilayer cellular morphology with ill-defined border (Figure 4.2), in contrast to the well-defined border of somatic cell-derived iPSCs (Amit and Itskovitz-Eldor, 2012). The expression of pluripotency genes was unexpectedly down-regulated on reprogramming of CRC (Figure 4.12). The molecular phenotypes of

MET/EMT genes (Figures 4.13 - 4.14) and the miRNA profiles (Figures 4.15 - 4.21) collectively indicated that the CRC-iPCs inclined towards a mesenchyme-like state, as opposed to the reported epithelia-state of ESC and iPSC (Hawkins et al., 2014). The dissimilarities raise two possible explanations. One, the CRC-iPC may have only undergone partial reprogramming, leading to the acquired limited pluripotency. The inferences of limited pluripotency and mesenchyme-like state of CRC-iPC clones will be discussed further in Discussion Sections 5.4 and 5.8, below.

Alternatively, the underlying mechanism of cancer reprogramming is distinct from that of the somatic-cell reprogramming. The reprogramming of cancer cells may encounter different barriers that are not found in somatic-cell reprogramming, thus the need of specific reprogramming enhancers. Consequently, the phenotypes observed in reprogrammed cancer cells may also differ from those of iPSC. Thus, the current understanding of somatic cell-derived iPSC may not be appropriately applied on iPC, or applied with great caution. Instead, new perspectives on the underlying mechanisms may be need to aid further understanding and the development of iPC studies.

On the other hand, re-differentiation of CRC-iPC cells into intestinal lineages is indeed crucial to provide insights on the early developmental stages of CRCs. Kim et al. (2013) reported that reprogrammed pancreatic ductal adenocarcinoma (PDAC) generated pancreatic intraepithelial neoplastic lesion when injected into immunodeficient mice, and showed an invasive phenotype of the parental PDAC. The invasive phenotype was later found to be associated

with signalling pathways, indicating that the reprogrammed PDAC is able to recapitulate the early PDAC progression (Kim et al., 2013). In fact, *in vivo* tumorigenicity of CRC-iPC cells and post-iPC cells were relatively compared to the parental CRCs (data not shown), and the results indicated diminished tumourigenicity.

5.2 Down-regulation of Pluripotency Genes

The pluripotency genes, *OSKM*, *NANOG* and *REX1*, tested in this work were found to be already steadily expressed in the parental CRC cell lines (Figure 4.11), suggesting overlapped functional roles of the pluripotency genes in cancer and stem cells. Overexpression of OSKM and NANOG is repeatedly observed in various human malignancies (Ben-Porath et al., 2008; Schoenhals et al., 2009). Amini et al. (2014) also demonstrated that the expression levels of the pluripotency-regulating factors were directly proportional to tumour grades. Given the major role of pluripotency genes in self-renewal, cell proliferation and survival, it is not surprising that the same set of pluripotency genes is dysregulated in cancer cells. In fact, *OCT4* and *NANOG* are proven to enhance tumour initiating ability and drug resistance, and to promote EMT in lung adenocarcinoma (Chiou et al., 2010). *SOX2* enhances migration and invasion abilities, thus, aiding cancer metastasis in nasopharyngeal carcinoma (Luo et al., 2013; Weina and Utikal, 2014).

On reprogramming, the pluripotency genes were down-regulated in the reprogrammed CRC cells (Figure 4.12). The pluripotency gene expression patterns were in disagreement with the general up-regulation trend observed in other reprogrammed cancers (Miyoshi et al., 2010; Miyazaki et al., 2015; Koga et al., 2014). Despite down-regulation of the pluripotency genes, which mainly contributes to self-renewal and proliferation (Buganim et al., 2013), the iPC lines were able to be continuously cultured up to the 60th passage (data not shown), indicating that the lower expression levels of pluripotency genes observed were sufficient to sustain self-renewal and proliferation of the reprogrammed cancer cells.

The aforementioned pluripotency-maintaining genes have been reported to be also highly expressed in various cancer types (Schoenhals et al., 2009; Amini et al., 2014) and some of the pluripotency genes work in concert with each other to aid cancer development (Hadjimichael et al., 2015). It is tempting to interpret the down-regulation of pluripotency genes to predict diminished tumorigenicity, as has been shown in other iPC reports. Reprogrammed colorectal cancer (Miyazaki et al., 2015) and hepatocellular carcinoma cells (Koga et al., 2014) showed enhanced drug sensitivity, low proliferative ability and down-regulated *c-MYC* expression. Furthermore, the proposed reduced malignancy is consistent with the hypothesis that cellular reprogramming reverses an advanced cancer stage to an earlier disease stage (Kim et al., 2013). Hence, recapitulation of cancer progression is possible for clinical monitoring (Kim and Zaret, 2015).

5.3 Limited Pluripotency of Reprogrammed Colorectal Cancer Cells

The phenomenon of partial reprogramming has been reported in other studies of reprogrammed somatic cells. The partially-reprogrammed somatic cells are known as pre-iPSC, and are often characterised by the low levels of endogenous pluripotency genes while the exogenous reprogramming factors may still be sustained in the pre-iPSC (Plath and Lowry, 2011), as observed in iSK-CO-1 Clone 2 failed to attenuate ectopic expression of *SOX2* and *c-MYC* (Figure 4.11), indicating incomplete reprogramming (Buganim et al., 2013). Pre-iPSCs do not exactly mimic the genetic and epigenetic characteristics of the fully pluripotent ESC (Kim et al., 2010), including expression of pluripotency genes (Zhao et al., 2014). However, the pre-iPSCs are still able to differentiate into tri-lineages and form *in vivo* teratoma (Gonzalez et al., 2009). In other reports, constant expression of reprogramming factors in neonatal fibroblast- and cord blood-derived pre-iPSC may result in inefficient differentiation into haemotapoietic and neural lineages (Ramos-Mejía et al., 2012). Similar observation was also reported by Toivonen et al. (2013). An iPSC cell line with exogenous *KLF4* expression suffered reduced differentiation efficiency (Toivonen et al., 2013)

In another human fibroblast-derived iPSC generated via CAG-driven polycistronic OSKM and GFP showed expression of transgenes in some derived clones (Montserrat et al., 2011). Therefore, residual transgene expression is not specific to retroviral transduction method (Montserrat et al., 2011). In fact, OSKM-retroviral transduced sarcoma (Zhang et al., 2013) and

colorectal cancer cells (Miyoshi et al., 2010) frequently showed efficient silencing of the OSKM transgenes.

Though iHCT-15 clones and iSK-CO-1 Clone 1 showed silent OSKM transgene, the down-regulation of pluripotency genes may also indicate limited activation of endogenous pluripotency genes (David and Polo, 2014). In addition, the miR-302 and miR-367 families, which were previously used to generate iPSCs and iPCs (Kuo and Ying, 2012), were not found among the 102 differentially-expressed miRNAs in the CRC-iPCs generated (Appendices B and C). On the contrary, the miRNA expression profiles indicated miRNA involvement in the regulation of EMT, which raises the possibility of an alternate route in cancer-cell reprogramming.

Given that cancer cells are heterogeneous and have accumulated multiple mutations and acquired an altered epigenetic programme, complete reprogramming of cancer cells faces greater challenges (Ramos-Mejia et al., 2012). Indeed, reprogrammed cancer cell clones frequently fail to form teratoma *in vivo* (Nishikawa et al., 2012) suggesting that limited pluripotency may be a common feature in reprogrammed cancer cells, as may also have resulted in the low reprogramming efficiency observed in this study.

Our data and other reports thus support the elite model of cancer reprogramming in that only a subpopulation of cancer cells may be fully reprogrammed (Lai et al., 2013). In relation to limited pluripotency, Zhang et al. (2013) proposed a pluripotency hierarchy model in which pluripotent stem

cells such as ESC are assigned the highest differentiation potentials, followed by somatic cell-derived iPSCs. On the contrary, in terms of differentiation potentials, the cancer cell-derived iPCs are induced to a pre-iPSC state, just a level above multipotent stem cells, such as mesenchymal stem cells. The pre-iPSC state of reprogrammed cancer cells may be reflected in the down-regulation of pluripotency genes (Zhang et al., 2013).

5.4 Differences Between Reprogramming of Somatic and Cancer Cells

Based on the observations and inferences discussed above, it is noted that reprogramming of somatic and cancer cells are distinctively different. Reprogramming of somatic cells is reported to follow a continuous stochastic model, where all transduced cells are at equal probability to transform into a pluripotent state (Hanna et al., 2009; Yamanaka, 2009). In contrast, Lai et al. (2013) has shown evidences that the elite model fits the context of cancer reprogramming better. The elite model proposes that the reprogramming of cancer cells is a biased process, in which only selected subset of cells in the heterogenous parental population is reprogrammed into iPC (Lai et al., 2013). The discrimination of cells during reprogramming has indicated a major difference between somatic and cancer cell reprogramming.

Cellular reprogramming of cancer cells faces many challenges due to the obstacles discussed above. In somatic cells, the pluripotency genes are initially dormant (Park et al., 2008), and are only re-activated upon reprogramming.

Therefore, the expression of pluripotency genes serves as a key indicator of ESC-mimicking gene expression profile in iPSC. On the contrary, the pluripotency genes are abundantly expressed in many cancer cell lines (Amini et al., 2014) (See Discussion Section 5.3 above). Hence, examining only the gene expression profile of the pluripotency genes may not be significant in the context of iPC. Consequently, differentiation ability of the reprogrammed cells has become an important clue in validating acquired pluripotency (Nishikawa et al., 2012). Owing to the latent subpopulation of multipotent cancer stem cells in the parental cancer cell population (Varghese et al., 2012), the parental CRC cells may also show various extents of differentiation ability. Hence, further support of the *in vitro* differentiation assays may be derived from teratoma formation to authenticate the pluripotency of the reprogrammed cancers (Zhang et al., 2012).

5.5 Differentially-expressed MicorRNAs (miRNAs) Possibly Mediate the EMT Process in Reprogramming

Upon cellular reprogramming, cells undergo epigenetic changes, including DNA methylation and chromatic modification, to achieve high degrees of similarity in the epigenetic state of ESC (Mattout et al., 2011). Given the regulatory role of epigenetic-related miRNA in the maintenance of pluripotency and cellular reprogramming (Anokye-Danso et al., 2012) (see Literature Review Section 2.4.3-2.4.5), differential expression of miRNA upon cellular reprogramming is anticipated. Wilson et al. (2009) first highlighted up-

regulated expression of the miR-302, miR-17 and miR-92 clusters in human fibroblast-derived iPSCs, and conversely, down-regulation of the let-7 family (Porciuncula et al., 2013). Similarly, the present study has shown differential expression of 102 miRNAs in CRC-iPC cells.

MiR-302 and -367 have been shown to facilitate reprogramming by enhancing self-renewal and cell cycle progression (Kuo et al., 2012). The reprogrammed CRC cells may not have functional contributions from miR-302 and -367 in self-renewal and cell cycle progression since these miRNAs were not activated in the CRC-iPC cells (Appendices B and C). However, there are other miRNA groups which may have supported cell proliferation in colorectal cancer (Cekaite et al., 2012). Up-regulation of miR-31 is reported in CRC and has been shown to inhibit the cell cycle suppressor E2F2, which in turn promotes CRC proliferation (Li et al., 2015). Furthermore, other miRNAs, including miR-106a/b, miR-20a/b and -17 (Yang et al., 2016) and miR-96 (Gao and Wang, 2015) have also been shown to contribute to the proliferation of CRC, possibly substituting the role of miR-302 and -367.

The miR-302 and miR-372 clusters have also been reported to inhibit EMT-inducing TGF-β signalling pathway during reprogramming, resulting in the acceleration of MET process (Subramanyam et al., 2011). However, the gene ontology and pathway analyses presented in this study collectively suggested EMT activation via TGF-β and PI3K-Akt signalling pathways (Figures 4.20 - 4.21), indicating a novel regulatory role of miRNA in mediating the cancer cells reprogramming.

5.6 Colorectal Cancer-derived Induced Pluripotent Cancer Cells Incline Towards a Mesenchyme-like State

During cellular reprogramming of somatic cells, MET is re-activated (David and Polo, 2014), as exemplified by up-regulation of epithelial proteins, E-CADHERIN (CDH1) and OCCLUDIN (OCLN) (Chen et al., 2012), resembling the epithelia-like state of ESC (Hawkins et al., 2014). In a study of partially-reprogrammed squamous cell carcinoma (SCC), the OSKM/LIN28-transduced SCC cells regained epithelia-like phenotypes despite down-regulated expression of pluripotency genes (Takaishi et al., 2016). Interestingly, the epithelia-associated gene, CDH1 has been shown to act as an alternative for OCT4 of the Yamanaka factors (Redmer et al., 2011). In miRNA-mediated reprogramming, the miR-302/367 clusters have been shown to accelerate reprogramming by modulating the MET process (Liao et al., 2011). These reports indicate that MET favours reprogramming. On the contrary, the reverse EMT process antagonises the de-differentiation process. Inhibition of the EMT-inducing TGF-β pathway greatly enhances reprogramming and, thus, inhibitors of the TGF-β pathway may replace SOX2 and c-MYC in reprogramming (Maherali and Hochedlinger, 2009). The analysed miRNA profile of the CRC-iPCs also showed possible activation or enhancement of two EMT-inducing pathways, TGF-β and PI3K-Akt signalling (Figure 4.21 and Figure 4.22), supporting the deduction of mesenchyme-like state of CRC-iPCs.

The findings that disagree with MET re-activation in somatic cell-derived iPSCs (Li et al., 2010; Chen et al., 2012) may suggest an alternative role of EMT in reprogramming. Liu et al. (2013) demonstrated that sequential introduction of the OSKM factors induced an early response of EMT prior to a delayed MET. Despite the fact that inhibition of TGF-β signalling, has positive effects during reprogramming, inhibition of TGF-β signalling is found to antagonise the initial phase of de-differentiation process, in agreement with the findings of Liu et al. (2013) (Hochedlinger and Plath, 2009; Maherali and Hochedlinger, 2009). The EMT-associated genes, *SNAI1* (Unternaehrer et al., 2014) and *VIM* (Kong et al., 2014), have been shown to successfully reprogramme both mouse and human cells by targeting well-known reprogramming barriers, p53 and let-7. Taken together, dynamic transition of EMT-MET is crucial during cellular reprogramming. In line with that, a hypothesis is proposed that temporary EMT allows reprogrammed cells to reach the optimal epithelial state, which subsequently converges with the delayed MET state (Liu et al., 2013) to enhance reprogramming.

5.7 Down-regulation of Pluripotency Genes Possibly Enhances EMT Activation

A scheme is proposed to reconcile the down-regulation of pluripotency genes with the possible mesenchyme-like state of CRC-iPCs (Figure 5.1). OCT4 and SOX2, have been previously shown to inhibit SNAI1, which works synergistically with KLF4 to induce epithelial phenotypes by up-regulating

CDH1 (Li et al., 2010). However, KLF4 is, in turn, negatively regulated by SNAI1 (De Craene et al., 2005). Thus, the up-regulation of SNAI1 may have inhibited KLF4, leading to down-regulation of CDH1 and hence, the suppression of MET. However, it is unclear how down-regulation of NANOG promotes EMT. In addition, the core pluripotency genes OSKM have previously been shown to inhibit various factors of the TGF-β signalling pathway (Li et al., 2010), and that OCT4 and SOX2 also suppress the EMT-inducing factor, SNAI1 (Li et al., 2010). Hence, OSKM down-regulation is likely to have contributed to the TGF-β signalling pathway to up-regulate SNAI1 expression and, therefore, promoting EMT (Figure 5.1). Transition between MET and EMT, linked by expression of pluripotency genes, may play a vital role in inducing pluripotency in cells of cancer origin.

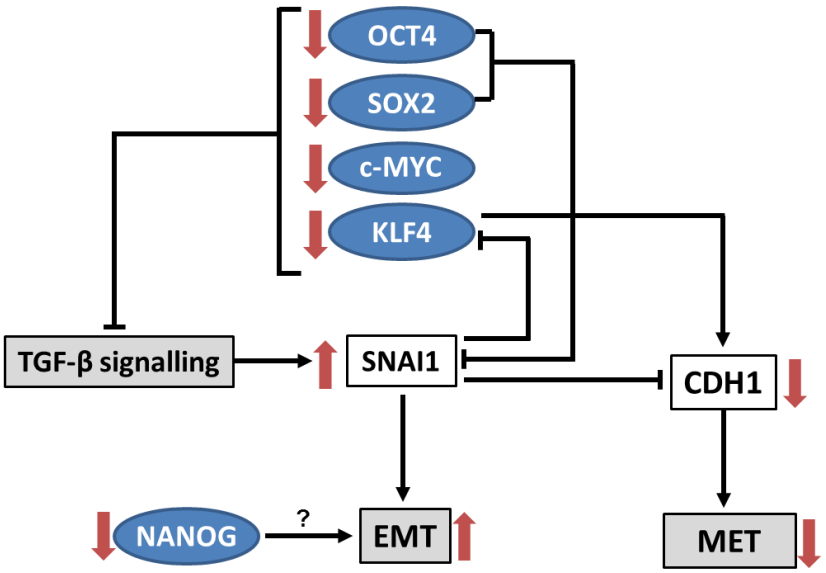


Figure 5.1: A proposed model of down-regulated pluripotency genes leading to EMT in CRC reprogramming. Red up- and down-ward arrows indicating gene expression observed in this study. Blue oval indicates pluripotency genes. White rectangle indicate MET/EMT genes. Grey rectangle indicates cellular processes. Down-regulation of pluripotency genes, collectively lead to promotion of EMT and suppression of MET.

CHAPTER 6

CONCLUSION

6.1 Conclusion

The current study demonstrated the generation of iPC cells from CRC cell lines via retroviral transduction of OSKM. The established CRC-iPC cells were in a near-pluripotent state with acquired tri-lineage differentiation ability. Interestingly, distinct molecular phenotypes in expression of pluripotency and MET/EMT genes were also observed in the CRC-iPC cells. The observed down-regulated expression of pluripotency genes may reflect partial reprogramming, suggesting limited pluripotency of the CRC-iPC cells. Furthermore, the down-regulated expression levels of the MET genes, exemplified by increased EMT gene expression, proposed a more mesenchyme-like state of the CRC-iPC cells. Genome-wide miRNA expression profiling revealed that 102 miRNAs were dysregulated in expression in the CRC-iPC cells upon reprogramming. Based on bioinformatics and gene ontology analyses, the predicted target genes of the differentially-expressed miRNAs were highly involved in regulation of cell motility. In addition, KEGG pathway analysis predicted possible activation of the TGF-β and PI3K-Akt cell signalling pathways in the CRC-iPC cells. The possible activation of TGF-β and PI3K-Akt signalling was consistent with

induction of the EMT process, likely resulting in the mesenchyme-like state of the CRC-iPC cells.

6.2 Limitations of current study

Despite the establishment of CRC-iPC cells presented in this study, low reprogramming efficiency was observed. The low reprogramming efficiency could possibly be linked to the limited pluripotency of the CRC-iPC cells. Hence, in order to enhance reprogramming efficiency, additional reprogramming factors, including LIN28 and NANOG, may be tested in the reprogramming cocktail in future reprogramming work (Zhang et al., 2013). Furthermore, inclusion of miR-302 and -371 and other reprogramming-able miRNAs may further enhance the reprogramming efficiency (Anokye-danso et al., 2011). In addition, more research needs to be done to improve the selection and identification of authentic iPC colonies, such as the doxycycline-induced transgene expression system that allows selection of fully reprogrammed colonies (Brambrink et al., 2008). In this system, the reprogrammed colonies retain ESC-like morphology upon removal of doxycycline as the cells have already entered a self-sustaining pluripotent state. On the other hand, the parental CRC share similar expression of pluripotency genes and surface markers as the CRC-iPC cells, complicating data interpretation, but does draw our attention to the intrinsic molecular differences between reprogramming normal somatic cells and cancer cells.

6.3 Future studies

The present study demonstrated possible activation of EMT-inducing signalling pathways via miRNA regulation in CRC reprogramming. As opposed to the antagonising role of EMT in somatic cell reprogramming, EMT activation in CRC reprogramming may indicate an alternative mechanism modulating cancer reprogramming. Therefore, validation of predicted miRNA-mRNA interactions in TGF-β and PI3K-Akt signalling is crucial to obtain direct experimental evidences, and to provide insights on the miRNA-regulated EMT process in cancer reprogramming. Such studies would further confirm novel roles of miRNAs in promoting pluripotency via EMT/MET regulation.

The iPC and post-iPC cells obtained in this study were mainly subjected to analysis of expression of the pluripotency and MET/EMT genes. However, future work should test the predicted diminished tumorigenicity of the reprogrammed cancer cells in animal studies (Zhang et al., 2013). Besides, sensitivity to chemotherapeutic drugs, such as 5-FU and L-OHP, of the iPC and post-iPC cells should be determined (Miyoshi et al., 2010; Koga et al., 2014). The CRC-iPC cells may also be further developed into three-dimensional organoid model by lineage-specific differentiation, and hence, serve as a disease model for CRC progression (Kim and Zaret, 2015; Fatehullah et al., 2016)

REFERENCES

Akao, Y., Nakagawa, Y. and Naoe, T., 2007. MicroRNA-143 and -145 in colon cancer. *DNA and Cell Biology*, 26(5), pp.311–320.

Alcindor, T. and Beauger, N., 2011. Oxaliplatin: A review in the era of molecularly targeted therapy. *Current Oncology*, 18(1), pp.18–25.

Allegrucci, C. et al., 2011. Epigenetic reprogramming of breast cancer cells with oocyte extracts. *Molecular Cancer*, 10(1), p.7.

Alvarez Palomo, A.B. et al., 2015. Prospects for clinical use of reprogrammed cells for autologous treatment of macular degeneration. *Fibrogenesis and Tissue Repair*, 8 (1), p.9.

Amini, S., Fathi, F., Mobalegi, J., Sofimajidpour, H. and Ghadimi, T., 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. *Anatomy and Cell Biology*, 47(1), pp.1–11.

Amit, M. and Itskovitz-Eldor, J., 2012. Atlas of human pluripotent stem cells. *Stem Cell Biology and Regenerative Medicine*, pp.15–39.

An, Q. et al., 2015. In vitro effects of mitomycin C on the proliferation of the non-small-cell lung cancer line A549. *International Journal of Clinical and Experimental Medicine*, 8(11), pp.20516–20523

Anokye-danso, F. et al., 2011. Highly efficient miRNA-mediated reprogramming of mouse and human somatic cells to pluripotency. *Cell Stem Cell*, 8(6), p.853.

Anokye-Danso, F., Snitow, M. and Morrisey, E.E., 2012. How microRNAs facilitate reprogramming to pluripotency. *Journal of Cell Science*, 125(18), pp.4179–87.

Ashley, N., Yeung, T.M. and Bodmer, W.F., 2013. Stem cell differentiation and lumen formation in colorectal cancer cell lines and primary tumors. *Cancer Research*, 73(18), pp.5798–5809.

Banito, A. et al., 2009. Senescence impairs successful reprogramming to pluripotent stem cells. *Genes and Development*, 23(18), pp.2134–2139.

Banwait, J.K. and Bastola, D.R., 2015. Contribution of bioinformatics prediction in microRNA-based cancer therapeutics. *Advanced Drug Delivery Reviews*, 81, pp.94–103.

Bao, X. et al., 2013. MicroRNAs in somatic cell reprogramming. *Current Opinion in Cell Biology*, 25(2), pp.208–214.

Barber, T.D. et al., 2008. Chromatid cohesion defects may underlie chromosome instability in human colorectal cancers. *Proceedings of the National Academy of Sciences of the United States of America*, 105(9), pp.3443–3448.

Barroso-del Jesus, A., Lucena-Aguilar, G. and Menendez, P., 2009. The miR-302-367 cluster as a potential stemness regulator in ESCs. *Cell Cycle*, 8(3), pp.394–398.

Bartel, D.P., 2004. MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. *Cell*, 116(2), pp.281–297.

Baskerville, S. and Bartel, D.P., 2005. Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. *RNA*, 11(3), pp.241–247.

Ben-Porath, I. et al., 2008. An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. *Nature Genetics*, 40(5), pp.499–507.

Betel, D., Wilson, M., Gabow, A., Marks, D.S. and Sander, C., 2008. The microRNA.org resource: Targets and expression. *Nucleic Acids Research*, 36, pp.149–153.

Biswas, A. and Hutchins, R., 2007. Consice Review: Embryonic Stem Cells. *Handbook of Stem Cells*, 740, pp.275–286.

Bohnsack, M.T., Czaplinski, K. and Gorlich, D., 2004. Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *RNA*, 10(2), pp.185–91.

Boland, C.R. et al., 2008. The biochemical basis of microsatellite instability and abnormal immunohistochemistry and clinical behavior in Lynch Syndrome: From bench to bedside. *Familial Cancer*, 7(1), pp.41–52.

Bononi, A. et al., 2011. Protein kinases and phosphatases in the control of cell fate. *Enzyme Research*, 2011, pp.1-26

Bose, B. and Sudheer Shenoy, P., 2014. Stem cell versus cancer and cancer stem cell : intricate balance decides their respective usefulness or harmfulness in the biological system. *Journal of Stem Cell Research & Therapy* , 4(2).

Boyer, L.A., et al., 2005. Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell*, 122(6), pp.947–956.

Brambrink, T. et al., 2008. Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells. *Cell Stem Cell*, 2(2), pp.151–159.

Brenner, H. et al., 2010. Protection from right-and left-sided colorectal neoplasms after colonoscopy: population-based study. *Journal of the National Cancer Institute*, 102(2), pp.89–95.

Buganim, Y., Faddah, D.A. and Jaenisch, R., 2013. Mechanisms and models of somatic cell reprogramming. *Nature Reviews Genetics*, 14(6), pp.427–439.

Cai, X., Hagedorn, C.H. and Cullen, B.R., 2004. Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA*, 10(12), pp.1957–1966.

Campbell, P.T. et al., 2007. Excess body weight and colorectal cancer risk in Canada: Associations in subgroups of clinically defined familial risk of cancer. *Cancer Epidemiology Biomarkers and Prevention*, 16(9), pp.1735–1744.

Carette, J.E. et al., 2010. Generation of iPSCs from cultured human malignant cells. *Blood*, 115(20), pp.4039–4042.

Cekaite, L. et al., 2012. MiR-9, -31, and -182 deregulation promote proliferation and tumor cell survival in colon cancer 1,2. *Neoplasia*, 14(9), pp.868–IN21.

Centelles, J.J., 2012. General aspects of colorectal cancer. *ISRN Oncology*, 2012, p.139268.

Center, M.M., Jemal, A., Smith, R.A. and Ward, E., 2009. Worldwide variations in colorectal cancer. *CA: A Cancer Journal for Clinicians*, *59*(6), pp.366-378

Cha, Y.I. and DuBois, R.N., 2007. NSAIDs and cancer prevention: targets downstream of COX-2. *Annual Review of Medicine*, 58, pp.239–252.

Chan, J.A., Krichevsky, A.M. and Kosik, K.S., 2005. MicroRNA-21 is an antiapoptotic factor in human glioblastoma cells. *Cancer Research*,*65*(14), pp.6029-6033

Chen, J., Han, Q. and Pei, D., 2012. EMT and MET as paradigms for cell fate switching. *Journal of Molecular Cell Biology*, 4(2), pp.66–69.

Chen, L. et al., 2012. A model of cancer stem cells derived from mouse induced pluripotent stem cells. *PloS one*, 7(4), p.e33544.

Chen, M. et al., 2012. Promotion of the induction of cell pluripotency through metabolic remodeling by thyroid hormone triiodothyronine-activated PI3K/AKT signal pathway. *Biomaterials*, 33(22), pp.5514–5523.

Chen, X. et al., 2009. Role of miR-143 targeting KRAS in colorectal tumorigenesis. *Oncogene*, 28(10), pp.1385–1392.

Chen, Y., Williams, V., Filippova, M., Filippov, V. and Duerksen-Hughes, P., 2014. Viral carcinogenesis: factors inducing DNA damage and virus integration. *Cancers*, 6(4), pp.2155–2186.

Chiou, S.H. 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 Research*, 70(24), pp.10433–10444.

Choo, K.B., Soon, Y.L., Nguyen, P.N.N., Hiew, M.S.Y. and Huang, C.J., 2014. MicroRNA-5p and-3p co-expression and cross-targeting in colon cancer cells. *Journal of Biomedical Science*, *21*(1), p.1-14.

Choong, M.L., Yang, H.H. & McNiece, I., 2007. MicroRNA expression profiling during human cord blood-derived CD34 cell erythropoiesis. Experimental Hematology, 35(4), pp.551–564.

Choong, P.F. et al., 2014. Heterogeneity of osteosarcoma cell lines led to variable responses in reprogramming. *International Journal of Medical Sciences*, 11(11), pp.1154–1160.

Corney, D.C. et al., 2007. MicroRNA-34b and MicroRNA-34c are targets of p53 and cooperate in control of cell proliferation and adhesion-independent growth. *Cancer Research*, 67(18), pp.8433–8438

Courtot, A.-M. et al., 2014. Morphological analysis of human induced pluripotent stem cells during induced differentiation and reverse programming. *BioResearch open access*, 3(5), pp.206–216.

De Craene, B. et al., 2005. The transcription factor snail induces tumor cell invasion through modulation of the epithelial cell differentiation program. *Cancer Research*, 65(14), pp.6237–6244.

Cummings, L.C. and Cooper, G.S., 2011. Colorectal cancer screening: Update for 2011. *Seminars in Oncology*, 38(4), pp.483–489.

Dang, C. V., 2012. MYC on the path to cancer. *Cell*, 149(1), pp.22–35.

David, L. and Polo, J.M., 2014. Phases of reprogramming. *Stem Cell Research*, 12(3), pp.754–761.

Denli, A.M. et al., 2004. Processing of primary microRNAs by the Microprocessor complex. *Nature*, 432(7014), pp.231–5.

Ebert, A.D. et al., 2009. Induced pluripotent stem cells from a spinal muscular atrophy patient. *Nature*, 457(7227), pp.277–280.

Ebrahimi, B., 2015. Engineering cell fate: the roles of ipsc transcription factors, chemicals, barriers and enhancing factors in reprogramming and transdifferentiation. *BioRxiv*, p.019455.

Erdei, Z. et al., 2013. Dynamic ABCG2 expression in human embryonic stem cells provides the basis for stress response. *European Biophysics Journal*, 42(2–3), pp.169–179.

Erdei, Z. et al., 2014. Expression pattern of the human ABC transporters in pluripotent embryonic stem cells and in their derivatives. *Cytometry Part B: Clinical Cytometry*, 86(5), pp.299–310.

Evans, M.J. and Kaufman, M.H., 1981. Establishment in culture of pluripotential cells from mouse embryos. *Nature*, 292, pp.154–156.

Fatehullah, A., Hui Tan, S. and Barker, N., 2016. Organoids as an in vitro model of human development and disease. *Nature Cell Biology*, 18(3), pp.246–254.

Fearon, E.R., 2011. Molecular genetics of colorectal cancer. *Annual Review of Pathology: Mechanisms of Disease*, 6, pp.479–507.

Ferrer-Vaquer, A., Viotti, M. and Hadjantonakis, A.K., 2010. Transitions between epithelial and mesenchymal states and the morphogenesis of the early mouse embryo. *Cell Adhesion and Migration*, 4(3), pp.447–457.

Gao, F. and Wang, W., 2015. MicroRNA-96 promotes the proliferation of colorectal cancer cells and targets tumor protein p53 inducible nuclear protein 1, forkhead box protein O1 (FOXO1) and FOXO3a. *Molecular Medicine Reports*, 11(2), pp.1200–1206.

Golipour, A. et al., 2012. A late transition in somatic cell reprogramming requires regulators distinct from the pluripotency network. *Cell Stem Cell*, 11(6), pp.769–782.

Goss, K.H. and Groden, J., 2000. Biology of the adenomatous polyposis coli tumor suppressor. *Journal of Clinical Oncology*, 18(9), pp.1967–1979.

Graham, F.L. et al., 1977. Characteristics of a Human Cell Line Transformed by D N A from Human Adenovirus Type 5. *Journal of General Virology*, 36(1), pp.59–72.

Grille, S.J. et al., 2003. The protein kinase akt induces epithelial mesenchymal transition and promotes enhanced motility and invasiveness of squamous cell carcinoma lines *Cancer research*, 63(9), pp.2172-2178.

De Gruijl, F.R., Van Kranen, H.J. and Mullenders, L.H.F., 2001. UV-induced DNA damage, repair, mutations and oncogenic pathways in skin cancer. *Journal of Photochemistry and Photobiology B: Biology*, 63(1–3), pp.19–27.

Guittet, L. et al., 2007. Comparison of a guaiac based and an immunochemical faecal occult blood test in screening for colorectal cancer in a general average risk population. *Gut*, 56(2), pp.210–214.

Hadjimichael, C. et al., 2015. Common stemness regulators of embryonic and cancer stem cells. *World Journal of Stem Cells*, 7(9), pp.1150–1184.

Haggar, F.A. and Boushey, R.P., 2009. Colorectal cancer epidemiology : incidence , mortality , survival , and risk factors. *Clinics in Colon and Rectal Surgery*, 22(04), pp.191–197.

Hammond, S.M., 2006. MicroRNAs as oncogenes. *Current Opinion in Genetics and Development*, 16(1), pp.4–9.

Hanna, J. et al., 2009. Direct cell reprogramming is a stochastic process amenable to acceleration. *Nature*, 462(7273), pp.595–601.

Harrison, T., Graham, F. and Williams, J., 1977. Host-range mutants of adenovirus type 5 defective for growth in HeLa cells. *Virology*, 77(1), pp.319–329.

Hassan, A.M.R. et al., 2016. Incidence and mortality rate of colorectal cancer in Malaysia. *Epidemiology and Health*, 38, p.e2016007.

Hawkins, K., Joy, S. and McKay, T., 2014. Cell signalling pathways underlying induced pluripotent stem cell reprogramming. *World Journal of Stem Cells*, 6(5), pp.620–628.

He, L. et al., 2007. A microRNA component of the p53 tumour suppressor network. *Nature*, 447(7148), pp.1130–1134.

Hernandez-Hansen, V. et al., 2005. Increased expression of genes linked to FcepsilonRI Signaling and to cytokine and chemokine production in Lyn-deficient mast cells. *Journal of Immunology*, 175(12), pp.7880–7888.

Hilmi, I., Hartono, J.L. and Goh, K.L., 2010. Negative perception in those at highest risk - Potential challenges in colorectal cancer screening in an urban Asian population. *Asian Pacific Journal of Cancer Prevention*, 11(3), pp.815–822.

Hirata, N. et al., 2014. A chemical probe that labels human pluripotent stem cells. *Cell Reports*, 6(6), pp.1165–1174.

Ho, R., Papp, B., Hoffman, J.A., Merrill, B.J. and Plath, K., 2013. Stage-specific regulation of reprogramming to induced pluripotent stem cells by wnt signaling and t cell factor proteins. *Cell Reports*, 3(6), pp.2113–2126.

Hochedlinger, K. and Plath, K., 2009. Epigenetic reprogramming and induced pluripotency. *Development*, 136(4), pp.509–523.

Hong, H. et al., 2009. Suppression of induced pluripotent stem cell generation by the p53-p21 pathway. *Nature*, 460(7259), pp.1132–1135.

Huang, C.J. et al., 2014. Frequent co-expression of miRNA-5p and -3p species and cross-targeting in induced pluripotent stem cells. *International Journal of Medical Sciences*, 11(8), pp.824–833.

Huang, D.W., Sherman, B.T. and Lempicki, R.A., 2009. Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Research*, 37(1), pp.1–13.

Huang, D.W., Sherman, B.T. and Lempicki, R.A., 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*, 4(1), pp.44–57.

Huang, G.T.J., 2010. Induced pluripotent stem cells-a new foundation in medicine. *Journal of Experimental and Clinical Medicine*, 2(5), pp.202–217.

Huang, J. et al., 2009. More synergetic cooperation of Yamanaka factors in induced pluripotent stem cells than in embryonic stem cells. *Cell Research*, 19(10), pp.1127–1138.

Hussein, S.M.I. and Nagy, A.A., 2012. Progress made in the reprogramming field: new factors, new strategies and a new outlook. *Current Opinion in Genetics and Development*, 22(5), pp.435–443..

Hutvágner, G. et al., 2001. A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science*, 293(5531), pp.834–838.

Jackson-Thompson, J., Ahmed, F., German, R.R., Lai, S.M. and Friedman, C., 2006. Descriptive epidemiology of colorectal cancer in the United States, 1998-2001. *Cancer*, 107(S5), pp.1103–1111.

Jansson, M.D. and Lund, A.H., 2012. MicroRNA and cancer. *Molecular Oncology*, 6(6), pp.590–610.

Jiang, S. et al., 2010. MicroRNA-155 functions as an oncomiR in breast cancer by targeting the suppressor of cytokine signaling 1 gene. *Cancer Research*, 70(8), pp.3119–3127.

Jiao, J. et al., 2013. Promoting reprogramming by FGF2 reveals that the extracellular matrix is a barrier for reprogramming fibroblasts to pluripotency. *Stem Cells*, 31(4), pp.729–740.

Johnson, I.T. and Lund, E.K., 2007. Review article: nutrition, obesity and colorectal cancer. *Alimentary Pharmacology and Therapeutics*, 26(2), pp.161–181.

Jonker, D., Rumble, R.B. and Maroun, J., 2006. Role of oxaliplatin combined with 5-fluorouracil and folinic acid in the first- and second-line treatment of advanced colorectal cancer. *Current Oncology*, 13(5), pp.173–184.

Judson, R.L., Babiarz, J.E., Venere, M. and Blelloch, R., 2009. Embryonic stem cell-specific microRNAs promote induced pluripotency. *Nature Biotechnology*, 27(5), pp.459–61.

Jung, C.B. et al., 2012. Dantrolene rescues arrhythmogenic RYR2 defect in a patient-specific stem cell model of catecholaminergic polymorphic ventricular tachycardia. *EMBO Molecular Medicine*, 4(3), pp.180–191.

Kang, S.G. et al., 2001. Mechanism of growth inhibitory effect of Mitomycin-C on cultured human retinal pigment epithelial cells: apoptosis and cell cycle arrest. Current eye research, 22(3), pp.174–181.

Kim, J. et al., 2013. An iPSC Line from Human Pancreatic Ductal Adenocarcinoma Undergoes Early to Invasive Stages of Pancreatic Cancer Progression. *Cell Reports*, 3(6), pp.2088–2099.

Kim, J. and Zaret, K.S., 2015. Reprogramming of human cancer cells to pluripotency for models of cancer progression. *EMBO Journal*, 34(6), pp.739–747.

Kim, K. et al., 2011. Donor cell type can influence the epigenome and differentiation potential of human induced pluripotent stem cells. *Nature Biotechnology*, 29(12), pp.1117–1119.

Kim, K. et al., 2010. Epigenetic memory in induced pluripotent stem cells. *Nature*, 467(7313), pp.285–290.

Kim, M.S., Lee, J. and Sidransky, D., 2010. DNA methylation markers in colorectal cancer. *Cancer and Metastasis Reviews*, 29(1), pp.181–206.

Kinehara, M. et al., 2013. Protein kinase c regulates human pluripotent stem cell self-renewal. *PLoS ONE*, 8(1), p.e54122.

Koga, C. et al., 2014. Reprogramming using microRNA-302 improves drug sensitivity in hepatocellular carcinoma cells. *Annals of Surgical Oncology*, 21(4), pp.591–600.

Kong, Q. et al., 2014. Identification and characterization of an oocyte factor required for porcine nuclear reprogramming. *Journal of Biological Chemistry*, 289(10), pp.6960–6968.

Kristensen, D.M., Kalisz, M. and Nielsen, J.H., 2005. Cytokine signalling in embryonic stem cells. *Apmis*, 113(11–12), pp.756–772.

Krützfeldt, J. et al., 2005. Silencing of microRNAs in vivo with "antagomirs". *Nature*, 438(7068), pp.685–689.

Kumano, K. et al., 2012. Generation of induced pluripotent stem cells from primary chronic myelogenous leukemia patient samples. *Blood* , 119(26), pp.6234–6242.

Kumar, R. et al., 2013. AID stabilizes stem-cell phenotype by removing epigenetic memory of pluripotency genes. *Nature*, 500(7460), pp.89–92.

Kuo, C.H., Deng, J.H., Deng, Q. and Ying, S.Y., 2012. A novel role of miR-302/367 in reprogramming. *Biochemical and Biophysical Research Communications*, 417(1), pp.11–16.

Kuo, C.H. and Ying, S.Y., 2012. Advances in microRNA-mediated reprogramming technology. *Stem Cells International*.

Kuo, T.F. et al., 2014. Selective elimination of human pluripotent stem cells by a marine natural product derivative. *Journal of the American Chemical Society*, 136(28), pp.9798–801.

Kurosawa, H., 2007. Methods for inducing embryoid body formation: in vitro differentiation system of embryonic stem cells. *Journal of Bioscience and Bioengineering*, 103(5), pp.389–398.

Lai, J., Kong, C.M., Mahalingam, D., Xie, X. and Wang, X., 2013. Elite Model for the Generation of Induced Pluripotent Cancer Cells (iPCs). *PLoS ONE*, 8(2), p.e56702.

Laurent, L.C. et al., 2008. Comprehensive microRNA profiling reveals a unique human embryonic stem cell signature dominated by a single seed sequence. *Stem Cells*, 26(6), pp.1506–1516.

Lee, G. et al., 2009. Modelling pathogenesis and treatment of familial dysautonomia using patient-specific iPSCs. *Nature*, 461(7262), pp.402–406.

Lee, R.C., Feinbaum, R.L. and Ambros, V., 1993. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell*, 75(5), pp.843–854.

Li, J., Braganza, A. and Sobol, R.W., 2013. Base excision repair facilitates a functional relationship between Guanine oxidation and histone demethylation. *Antioxid Redox Signal*, 18(18), pp.2429–2443.

Li, R. et al., 2010. A mesenchymal-to-epithelial transition initiates and is required for the nuclear reprogramming of mouse fibroblasts. *Cell Stem Cell*, 7(1), pp.51–63.

Li, T., Luo, W., Liu, K., Lv, X. and Xi, T., 2015. MiR-31 promotes proliferation of colon cancer cells by targeting E2F2. *Biotechnology Letters*, 37(3), pp.523–532.

Liao, B. et al., 2011. MicroRNA cluster 302-367 enhances somatic cell reprogramming by accelerating a mesenchymal-to-epithelial transition. *Journal of Biological Chemistry*, 286(19), pp.17359–17364.

Lim, K.G., 2014. A review of colorectal cancer research in Malaysia. *Medical Journal of Malaysia*, 69, pp.23–32.

Lim, K.L. et al., 2016. Expert opinion on biological therapy reprogramming cancer cells: overview and current progress. *Expert Opinion on Biological Therapy*, pp.1-11.

Lin, C.P., Choi, Y.J., Hicks, G.G. and He, L., 2012. The emerging functions of the p53-miRNA network in stem cell biology. *Cell Cycle*, 11(11), pp.2063–2072.

Lin, C.Y. et al., 2012. Transcriptional amplification in tumor cells with elevated c-Myc. *Cell*, 151(1), pp.56–67.

Lin, S.L. et al., 2008. Mir-302 reprograms human skin cancer cells into a pluripotent ES-cell-like state. *RNA*, 14(10), pp.2115–24.

Link, A. et al., 2010. Fecal microRNAs as novel biomarkers for colon cancer screening. *Cancer Epidemiology Biomarkers and Prevention*, 19(7), pp.1766–1774.

Lipchina, I. et al., 2011. Genome-wide identification of microRNA targets in human ES cells reveals a role for miR-302 in modulating BMP response. *Genes and Development*, 25(20), pp.2173–2186.

Liu, B.P. and Burridge, K., 2000. Vav2 activates Rac1, Cdc42, and RhoA downstream from growth factor receptors but not beta1 integrins. *Molecular and Cellular Biology*, 20(19), pp.7160–7169.

Liu, X. et al., 2013. Sequential introduction of reprogramming factors reveals a time-sensitive requirement for individual factors and a sequential EMT-MET mechanism for optimal reprogramming. *Nature Cell Biology*, 15(7), pp.829–838.

Liu, X. et al., 2008. Yamanaka factors critically regulate the developmental signaling network in mouse embryonic stem cells. *Cell Research*, 18(12), pp.1177–1189.

Livak, K.J. and Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2^{-} ΔΔCT method. *Methods*, 25(4), pp.402-408.

Llames, S., García-Pérez, E., Meana, A., Larcher, F. and del Río, M., 2015. Feeder layer cell actions and applications. *Tissue engineering Part B: Reviews*, 21(4), pp.345–353.

Loh, Y.-H. et al., 2009. Generation of induced pluripotent stem cells from human blood. *Blood*, 113(22), pp.5476–5479.

Lowry, W.E. et al., 2008. Generation of human induced pluripotent stem cells from dermal fibroblasts. *Proceedings of the National Academy of Sciences*, 105(8), pp.2883–2888.

Lu, D. et al., 2012. MiR-25 regulates Wwp2 and Fbxw7 and promotes reprogramming of mouse fibroblast cells to iPSCs. *PLoS ONE*, 7(8), p.e40938.

Lugli, A. et al., 2010. Prognostic impact of the expression of putative cancer stem cell markers CD133, CD166, CD44s, EpCAM, and ALDH1 in colorectal cancer. *British journal of cancer*, 103(3), pp.382–390.

Lüningschrör, P., Hauser, S., Kaltschmidt, B. and Kaltschmidt, C., 2013. MicroRNAs in pluripotency, reprogramming and cell fate induction. *Biochimica et Biophysica Acta - Molecular Cell Research*, 1833(8), pp.1894–1903.

Luo, L.Z. et al., 2012. DNA repair in human pluripotent stem cells is distinct from that in non-pluripotent human cells. *PLoS ONE*, 7(3), p.e30541.

Luo, W. et al., 2013. Embryonic stem cells markers Sox2, Oct4 and Nanog expression and their correlations with epithelial-mesenchymal transition in nasopharyngeal carcinoma. *PLoS ONE*, 8(2), p.e56324.

Luo, X., Burwinkel, B., Tao, S. and Brenner, H., 2011. MicroRNA signatures: novel biomarker for colorectal cancer? *Cancer Epidemiology Biomarkers and Prevention*, 20(7), pp.1272–1286.

Lynch, H.T. et al., 2009. Review of the Lynch syndrome: history, molecular genetics, screening, differential diagnosis, and medicolegal ramifications. *Clinical Genetics*, 76(1), pp.1–18.

Maherali, N. and Hochedlinger, K., 2009. Tgfβ signal inhibition cooperates in the induction of iPSCs and replaces Sox2 and cMyc. *Current Biology*, 19(20), pp.1718–1723.

Mallanna, S.K. and Rizzino, A., 2010. Emerging roles of MicroRNAs in the control of embryonic stem cells and the generation of induced pluripotent stem cells. *Developmental Biology*, 344(1), pp.16–25.

Maniataki, E. and Mourelatos, Z., 2005. A human, ATP-independent, RISC assembly machine fueled by pre-miRNA. *Genes and Development*, 19(24), pp.2979–2990.

Markowitz, S. and Bertagnolli, M., 2009. Molecular basis of colorectal cancer. *New England Journal of Medicine*, *361*(25), pp.2449-2460.

Marotta, L.L.C. and Polyak, K., 2009. Cancer stem cells: a model in the making. *Current Opinion in Genetics and Development*, 19(1), pp.44–50.

Mattout, A., Biran, A. and Meshorer, E., 2011. Global epigenetic changes during somatic cell reprogramming to iPS cells. *Journal of Molecular Cell Biology*, 3(6), pp.341–350.

Medvedev, S.P., Shevchenko, A.I. and Zakian, S.M., 2010. Induced pluripotent stem cells : problems and advantages when applying them in regenerative medicine. *Acta Naturae*, 2(5), pp.18–27.

Megyola, C.M. et al., 2013. Dynamic migration and cell-cell interactions of early reprogramming revealed by high-resolution time-lapse imaging. *Stem Cells*, 31(5), pp.895–905.

Melton, C., Judson, R.L. and Blelloch, R., 2010. Opposing microRNA families regulate self-renewal in mouse embryonic stem cells. *Nature*, 463(7281), pp.621–626.

Mendell, J.T. and Olson, E.N., 2012. MicroRNAs in stress signaling and human disease. *Cell*, 148(6), pp.1172–1187.

Michael, M.Z. et al., 2003. Reduced accumulation of specific microRNAs in colorectal neoplasia. *Molecular cancer research*, 1(12), pp.882–891.

Mills, J. a et al., 2013. Clonal genetic and hematopoietic heterogeneity among human induced pluripotent stem cell lines. *Blood*, 122(12), pp.2047–2052.

Miyazaki, S. et al., 2015. A cancer reprogramming method using micrornas as a novel therapeutic approach against colon cancer. *Annals of Surgical Oncology*, 22(3), pp.1394-1401

Miyoshi, K. et al., 2010. Generation of human induced pluripotent stem cells from oral mucosa. *Journal of Bioscience and Bioengineering*, 110(3), pp.345–350.

Miyoshi, N. et al., 2010. Defined factors induce reprogramming of gastrointestinal cancer cells. *Proceedings of the National Academy of Sciences*, 107(1), pp.40–45.

Miyoshi, N. et al., 2011. Reprogramming of mouse and human cells to pluripotency using mature microRNAs. *Cell Stem Cell*, 8(6), pp.633–638.

Momcilovic, O. et al., 2010. DNA damage responses in human induced pluripotent stem cells and embryonic stem cells. *PLoS ONE*, 5(10), p.e13410.

Montserrat, N., Garreta, E., González, F., Gutiérrez, J., Eguizábal, C., Ramos, V., Borrós, S. and Belmonte, J.C.I., 2011. Simple generation of human induced pluripotent stem cells using poly-β-amino esters as the non-viral gene delivery system. *Journal of Biological Chemistry*, 286(14), pp.12417-12428.

Mooney, B.M., Raof, N.A., Li, Y. and Xie, Y., 2013. Convergent mechanisms in pluripotent stem cells and cancer: Implications for stem cell engineering. *Biotechnology Journal*, 8(4), pp.408–419.

Motameny, S., Wolters, S., Nürnberg, P. and Schumacher, B., 2010. Next generation sequencing of miRNAs - Strategies, resources and methods. *Genes*, 1(1), pp.70–84.

Müller, M. et al., 2016. The role of pluripotency factors to drive stemness in gastrointestinal cancer. *Stem Cell Research*, 16(2), pp.349–357.

Naito, A. et al., 2008. Progressive tumor features accompany epithelial-mesenchymal transition induced in mitochondrial DNA-depleted cells. *Cancer Science*, 99(8), pp.1584–1588.

Nelson, T.J. et al., 2010. Induced pluripotent stem cells: advances to applications. *Stem cells and cloning: advances and applications*, 3, pp.29–37.

Ng, E.K. et al., 2009. Differential expression of microRNAs in plasma of patients with colorectal cancer: a potential marker for colorectal cancer screening. *Gut*, 58(10), pp.1375–1381.

Ng, E.K.O. et al., 2009. MicroRNA-143 targets DNA methyltransferases 3A in colorectal cancer. *British Journal of Cancer*, 101(4), pp.699–706.

Nie, Z. et al., 2012. c-Myc is a universal amplifier of expressed genes in lymphocytes and embryonic stem cells. *Cell*, 151(1), pp.68–79.

Nigro, A. Lo and Verfaillie, C.M., 2013. Multiple Paths to Reprogramming.

Nishikawa, S. et al., 2012. MicroRNA-based cancer cell reprogramming technology. *Experimental and Therapeutic Medicine*, 4(1), pp.8-14.

O'Brien, C.A., Pollett, A., Gallinger, S. and Dick, J.E., 2007. A human colon cancer cell capable of initiating tumour growth in immunodeficient mice. *Nature*, 445(7123), pp.106–110.

Ogawa, H. et al., 2015. MicroRNAs induce epigenetic reprogramming and suppress malignant phenotypes of human colon cancer cells. *PLoS ONE*, 10(5), p.e0127119.

Oshima, N. et al., 2014. Induction of cancer stem cell properties in colon cancer cells by defined factors. *PLoS ONE*, 9(7), p.e101735.

Park, I.H. et al., 2008. Reprogramming of human somatic cells to pluripotency with defined factors. *Nature*, 451(7175), pp.141–146.

Pieters, T. and van Roy, F., 2014. Role of cell-cell adhesion complexes in embryonic stem cell biology. *Journal of Cell Science*, 127(12), pp.2603–2613.

Plath, K. and Lowry, W.E., 2011. Progress in understanding reprogramming to the induced pluripotent state. *Nature reviews. Genetics*, 12(4), pp.253–265.

Porciuncula, A. et al., 2013. MicroRNA signatures of iPSCs and endoderm-derived tissues. *Gene Expression Patterns*, 13(1–2), pp.12–20.

Poulogiannis, G., Frayling, I.M. and Arends, M.J., 2010. DNA mismatch repair deficiency in sporadic colorectal cancer and Lynch syndrome. *Histopathology*, 56(2), pp.167–179.

Pourhoseingholi, M.A., 2014. Epidemiology and burden of colorectal cancer in Asia-Pacific region: what shall we do now? *Translational Gastrointestinal Cancer*, 3(4), pp.169–173.

Pourhoseingholi, M.A., 2012. Increased burden of colorectal cancer in Asia. *World Journal of Gastrointestinal Oncology*, 4(4), p.68.

Pourrajab, F. et al., 2014. MicroRNA-based system in stem cell reprogramming; differentiation/dedifferentiation. *International Journal of Biochemistry and Cell Biology*, 55, pp.318–328.

Puglisi, M.A., Tesori, V., Lattanzi, W., Gasbarrini, G.B. and Gasbarrini, A., 2013. Colon cancer stem cells: controversies and perspectives. *World Journal of Gastroenterology*, 19(20), pp.2997–3006.

Qi, J. et al., 2009. MicroRNAs regulate human embryonic stem cell division. *Cell cycle*, 8(22), pp.3729–3741.

Qian, S.W. et al., 2010. Characterization of adipocyte differentiation from human mesenchymal stem cells in bone marrow. *BMC Developmental Biology*, 10(1), p.47.

Qian, Y. and Chen, X., 2010. Tumor suppression by p53: making cells senescent. *Histology and Histopathology*, 25(4), pp.515–526.

Ramos-Mejia, V., Fraga, M.F. and Menendez, P., 2012. IPSCs from cancer cells: challenges and opportunities. *Trends in Molecular Medicine*, 18(5), pp.245–247.

Ramos-Mejía, V. et al., 2012. Residual expression of the reprogramming factors prevents differentiation of iPSC generated from human fibroblasts and cord blood CD34+ progenitors. *PLoS ONE*, 7(4), p.e35824.

Rao, M.S. and Malik, N., 2012. Assessing iPSC reprogramming methods for their suitability in translational medicine. *Journal of Cellular Biochemistry*, 113(10), pp.3061–3068.

Redmer, T. et al., 2011. E-cadherin is crucial for embryonic stem cell pluripotency and can replace OCT4 during somatic cell reprogramming. *EMBO reports*, 12(7), pp.720–726.

Reya, T., Morrison, S.J., Clarke, M.F. and Weissman, I.L., 2001. Stem cells, cancer, and cancer stem cells. *Nature*, 414(6859), pp.105–111.

Ricci-Vitiani, L. et al., 2007. Identification and expansion of human colon-cancer-initiating cells. *Nature*, 445(7123), pp.111–115.

Rippon, H.J. and Bishop, A.E., 2004. Embryonic stem cells. *Cell Proliferation*, 37(1), pp.23–34.

Rocha, C.R.R., Lerner, L.K., Okamoto, O.K., Marchetto, M.C. and Menck, C.F.M., 2013. The role of DNA repair in the pluripotency and differentiation of human stem cells. *Mutation Research*, 752(1), pp.25–35.

Rogers, A.C., Hanly, A.M., Collins, D., Baird, A.W. and Winter, D.C., 2012. Review article: loss of the calcium-sensing receptor in colonic epithelium is a key event in the pathogenesis of colon cancer. *Clinical Colorectal Cancer*, 11(1), pp.24–30.

Samardzija, C., Quinn, M., Findlay, J.K. and Ahmed, N., 2012. Attributes of Oct4 in stem cell biology: perspectives on cancer stem cells of the ovary. *Journal of Ovarian Research*, 5(1), p.37.

Samavarchi-tehrani, P. et al., 2010. Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. *Cell Stem Cell*, 7(1), pp.64-77.

Sanges, D. and Cosma, M., 2011. Reprogramming cell fate to pluripotency: the decision-making signalling pathways. I*nternational Journal of Developmental Biology*, 54(11-12), pp.1575-1587.

Sassen, S., Miska, E.A. and Caldas, C., 2008. MicroRNA - Implications for cancer. *Virchows Archiv*, 452(1), pp.1–10.

Savagner, P., 2010. The epithelial-mesenchymal transition (EMT) phenomenon. *Annals of Oncology*, 21, pp.89–92.

Schetter, A.J., Okayama, H. and Harris, C.C., 2013. Role of microRNA in colorectal cancer. *Cancer Journal*, 18(3), pp.244–252.

Schmittgen, T.D. and Livak, K.J., 2008. Analyzing real-time PCR data by the comparative CT method. *Nature protocols*, 3(6), pp.1101-1108.

Schoenhals, M. et al., 2009. Embryonic stem cell markers expression in cancers. *Biochemical and Biophysical Research Communications*, 383(2), pp.157–162.

Scott, C.W., Peters, M.F. and Dragan, Y.P., 2013. Human induced pluripotent stem cells and their use in drug discovery for toxicity testing. *Toxicology Letters*, 219(1), pp.49–58.

Seita, J. and Weismann, I.L., 2010. Hematopoietic stem cell: self-renewal versus differentiation. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, 2(6), pp.640-653

Semi, K. et al., 2013. Cellular reprogramming and cancer development. *International Journal of Cancer*, 132(6), pp.1240–1248.

Shankar, J. and Nabi, I.R., 2015. Actin cytoskeleton regulation of epithelial mesenchymal transition in metastatic cancer cells. *PLoS ONE*, 10(3), pp.1–12.

Shi, B., Sepp-Lorenzino, L., Prisco, M., Linsley, P. and Baserga, R., 2007. MicroRNA-145 targets the insulin receptor substrate-1 and inhibits the growth of colon cancer cells. *Journal of Biological Chemistry*, 282(45), pp.32582–32590.

Shi, Y. et al., 2008. Induction of pluripotent stem cells from mouse embryonic fibroblasts by oct4 and klf4 with small-molecule compounds. *Cell Stem Cell*, 3(5), pp.568–574.

Shyh-Chang, N. and Daley, G.Q., 2013. Lin28: primal regulator of growth and metabolism in stem cells. *Cell Stem Cell*, 12(4), pp.395–406.

Sidney, L.E., Kirkham, G.R. and Buttery, L.D., 2014. Comparison of osteogenic differentiation of embryonic stem cells and primary osteoblasts revealed by responses to IL-1β, TNF-α, and IFN-γ. *Stem Cells and Development*, 23(6), pp.605–617.

Siegel, R.L., Miller, K.D. and Jemal, A., 2016. Cancer statistics, 2016. *CA: Cancer Journal for Clinicians*, 66(1), pp.7–30.

Silva, J. et al., 2009. Nanog is the gateway to the pluripotent ground state. *Cell*, 138(4), pp.722–737.

Singh, U. et al., 2012. Novel live alkaline phosphatase substrate for identification of pluripotent stem cells. *Stem Cell Reviews and Reports*, 8(3), pp.1021–1029.

Smith, A., Young, G.P., Cole, S.R. and Bampton, P., 2006. Comparison of a brush-sampling fecal immunochemical test for hemoglobin with a sensitive guaiac-based fecal occult blood test in detection of colorectal neoplasia. *Cancer*, 107(9), pp.2152–2159.

Soetikno, R.M. et al., 2008. Prevalence of nonpolypoid (flat and depressed) colorectal neoplasms in asymptomatic and symptomatic adults. *JAMA: The Journal of The American Medical Association*, 299(9), pp.1027–1035.

Spike, B.T. and Wahl, G.M., 2011. p53, stem cells, and reprogramming: tumor suppression beyond guarding the genome. *Genes and Cancer*, 2(4), pp.404–419.

Spillane, J.B. and Henderson, M.A., 2007. Cancer stem cells: a review. *ANZ Journal of Surgery*, 77(6), pp.464–468.

Stadtfeld, M. and Hochedlinger, K., 2010. Induced pluripotency: history, mechanisms, and applications. *Genes and Development*, 24(20), pp.2239–2263.

Stricker, S.H. and Pollard, S.M., 2014. Reprogramming cancer cells to pluripotency: an experimental tool for exploring cancer epigenetics. *Epigenetics*, 9(6), pp.798–802.

Strillacci, A. et al., 2009. MiR-101 downregulation is involved in cyclooxygenase-2 overexpression in human colon cancer cells. *Experimental Cell Research*, 315(8), pp.1439–1447.

Subramanyam, D. et al., 2011. Multiple targets of miR-302 and miR-372 promote reprogramming of human fibroblasts to induced pluripotent stem cells. *Nature Biotechnology*, 29(5), pp.443–448.

Sugii, S., Kida, Y., Berggren, W.T. and Evans, R.M., 2011. Feeder-dependent and feeder-independent iPS cell derivation from human and mouse adipose stem cells. *Nature Protocols*, 6(3), pp.346–358.

Sun, C. and Liu, Y.K., 2011. Induced pluripotent cancer cells: progress and application. *Journal of Cancer Research and Clinical Oncology*, 137(1), pp.1–8.

Takahashi, K. et al., 2007. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell*, 131(5), pp.861–872.

Takahashi, K. and Yamanaka, S., 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*, 126(4), pp.663–676.

Takaishi, M. et al., 2016. Mesenchymal to epithelial transition induced by reprogramming factors attenuates the malignancy of cancer cells. *PLoS ONE*, 11(6), pp.1–15.

Takase, O. et al., 2013. The role of nf- k b signaling in the maintenance of pluripotency of human induced pluripotent stem cells. *PLoS ONE*, 8(2), pp.1–9.

Tanabe, S., 2015. Signaling involved in stem cell reprogramming and differentiation. *World Journal of Stem Cells*, 7(7), pp.992–998.

Tang, Y. et al., 2014. Differential effects of Akt isoforms on somatic cell reprogramming. *Journal of Cell Science*, 127(18), pp.3998–4008.

Tang, Y. and Tian, X., 2013. JAK-STAT3 and somatic cell reprogramming. *Jak-Stat*, 2(4), p.e24935

Teo, A.K.K. et al., 2012. Activin and BMP4 synergistically promote formation of definitive endoderm in human embryonic stem cells. *Stem Cells*, 30(4), pp.631–642.

Thomson, J.A. et al., 1998. Embryonic stem cell lines derived from human blastocysts. *Science*, 282(5391), pp.1145–1147.

Tsong, W.H. et al., 2007. Cigarettes and alcohol in relation to colorectal cancer: the Singapore Chinese Health Study. *British Journal of Cancer*, 96(5), pp.821–827.

Toivonen, S. et al., 2013. Comparative analysis of targeted differentiation of human induced pluripotent stem cells (hiPSCs) and human embryonic stem cells reveals variability associated with incomplete transgene silencing in retrovirally derived hiPSC lines. *Stem cells translational medicine*, 2(2), pp.409–419.

Unternaehrer, J.J. et al., 2014. The epithelial-mesenchymal transition factor SNAIL paradoxically enhances reprogramming. *Stem Cell Reports*, 3(5), pp.691–698.

Varghese, S., Whipple, R., Martin, S.S. and Alexander, H.R., 2012. Multipotent cancer stem cells derived from human malignant peritoneal mesothelioma promote tumorigenesis. *PLoS ONE*, 7(12), pp.1–10.

La Vecchia, C. et al., 1997. Aspirin and colorectal cancer. *British Journal of Cancer*, 76(5), pp.675–677.

Vitale, A.M., Wolvetang, E. and Mackay-Sim, A., 2011. Induced pluripotent stem cells: a new technology to study human diseases. *The International Journal of Biochemistry and Cell Biology*, 43(6), pp.843–846.

Vojnits, K. and Bremer, S., 2010. Challenges of using pluripotent stem cells for safety assessments of substances. *Toxicology*, 270(1), pp.10–17.

Volgelstein, B. and Kinzler, K.W., 2015. The Path to cancer -- three strikes and you ’re out. *The New England Journal of Medicine*, 373(20), pp.1895–1898.

Wahid, F., Shehzad, A., Khan, T. and Kim, Y.Y., 2010. MicroRNAs: Synthesis, mechanism, function, and recent clinical trials. *Biochimica et Biophysica Acta - Molecular Cell Research*, 1803(11), pp.1231–1243.

Wang, R., Mercaitis, O.P., Jia, L., Panettieri, R.A. and Tang, D.D., 2013. Raf-1, actin dynamics, and abelson tyrosine kinase in human airway smooth muscle cells. *American Journal of Respiratory Cell and Molecular Biology*, 48(2), pp.172–178.

Wang, S., 2013. Enrichment and selective targeting of cancer stem cells in colorectal cancer cell lines. *Human Genetics and Embryology*, 1(S2).

Wang, T., Shi, S. and Sha, H., 2013. MicroRNAs in regulation of pluripotency and somatic cell reprogramming: small molecule with big impact. *RNA Biology*, 10(8), pp.1255–1261.

Wang, Y., 2014. Myocardial reprogramming medicine: the development, application , and challenge of induced pluripotent stem cells. *New Journal of Science*, 2014, pp.1–22.

Wang, Z. et al., 2015. Activin A can induce definitive endoderm differentiation from human parthenogenetic embryonic stem cells. Biotechnology Letters, 37(8), pp.1711–1717.

Weina, K. and Utikal, J., 2014. SOX2 and cancer: current research and its implications in the clinic. *Clinical and Translational Medicine*, *3*(1), pp.1–10.

Wilmut, I., Schnieke, A.E., Mchwir, J., Kind, A.J. and Campbell, K.H.S., 1997. Viable offspring derived from fetal and adult mammalian cells. *Letters to Nature*, 385, pp.810-813

Wilson, K.D. et al., 2009. MicroRNA profiling of human-induced pluripotent stem cells. *Stem cells and Development*, 18(5), pp.749–757.

Winter, J., Jung, S., Keller, S., Gregory, R.I. and Diederichs, S., 2009. Many roads to maturity: microRNA biogenesis pathways and their regulation. *Nature Cell Biology*, 11(3), pp.228-234.

Wood, L.D. et al., 2007. The genomic landscapes of human breast and colorectal cancers. *Science*, 318(5853), pp.1108–1113.

Woyach, J.A., Johnson, A.J. and Byrd, J.C., 2012. The B-cell receptor signaling pathway as a therapeutic target in CLL. *Blood*, 120(6), pp.1175–1184.

Xu, J., Lamouille, S. and Derynck, R., 2009. TGF-β-induced epithelial to mesenchymal transition. *Cell Research*, 19(2), pp.156–172.

Xu, N., Papagiannakopoulos, T., Pan, G., Thomson, J.A. and Kosik, K.S., 2009. MicroRNA-145 Regulates OCT4, SOX2, and KLF4 and represses pluripotency in human embryonic stem cells. *Cell*, 137(4), pp.647–658.

Xu, W., Yang, Z. and Lu, N., 2015. A new role for the PI3K/Akt signaling pathway in the epithelial-mesenchymal transition. *Cell Adhesion & Migration*, 9(4), pp.317-324

Xu, Y. et al., 2016. Transcriptional Control of Somatic Cell Reprogramming. *Trends in Cell Biology*, 26(4), pp.272–288.

Xue, G. and Hemmings, B.A., 2013. PKB/akt-dependent regulation of cell motility. *Journal of the National Cancer Institute*, 105(6), pp.393–404.

Yamanaka, S., 2009. Elite and stochastic models for induced pluripotent stem cell generation. *Nature*, 460(7251), pp.49–52.

Yan, L. et al., 2008. MicroRNA miR-21 overexpression in human breast cancer is associated with advanced clinical stage, lymph node metastasis and patient poor prognosis. *RNA*, pp.2348–2360.

Yang, C., Li, Z. and Rana, T.M., 2011. MicroRNAs modulate iPS cell generation. *RNA*, 17(8), pp.1451–1460.

Yang, L., Luo, P., Fei, X., and Zhang, P., 2016. A miRNAs panel promotes the proliferation and invasion of colorectal cancer cells by targeting GABBR1. *Cancer Medicine*, 5(8), pp.2022–2031.

Ye, J. et al., 2012. Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics*, 13(1), p.134.

Ye, Z. et al., 2009. Human-induced pluripotent stem cells from blood cells of healthy donors and patients with acquired blood disorders. *Blood*, 114(27), pp.5473–5480.

Yeo, J.C. and Ng, H.-H., 2013. The transcriptional regulation of pluripotency. *Cell Research*, 23(1), pp.20–32.

Yilmazer, A., Lázaro, I. De and Taheri, H., 2015. Reprogramming cancer cells: a novel approach for cancer therapy or a tool for disease-modeling? *Cancer Letters*, 369(1), pp.1–8.

Yoshida, G.J., 2015. Metabolic reprogramming: the emerging concept and associated therapeutic strategies. *Journal of Experimental and Clinical Cancer Research*, 34(1), p.111.

Yoshida, Y., Takahashi, K., Okita, K., Ichisaka, T. and Yamanaka, S., 2009. Hypoxia enhances the generation of induced pluripotent stem cells. *Cell Stem Cell*, 5(3), pp.237–241.

Yu, J. et al., 2007. Induced pluripotent stem cell lines derived from human somatic cells. *Science*, 318(5858), pp.1917–1920.

Yusoff, H.M., Daud, N., Noor, N.M. and Rahim, A.A., 2012. Participation and barriers to colorectal cancer screening in Malaysia. *Asian Pacific Journal of Cancer Prevention*, 13(8), pp.3983–3987.

Zainal Ariffin, O. and Nor Saleha, I.T., 2011. National cancer registry report 2007. *Malaysia: Ministry of Health*.

Zhang, W.Y., De Almeida, P.E. and Wu, J.C., 2012. Teratoma formation: a tool for monitoring pluripotency in stem cell research. *Current Protocols in Stem Cell Biology*, p.4a.8.1-4a.8.17.

Zhang, X., Cruz, F.D., Terry, M., Remotti, F. and Matushansky, I., 2013. Terminal differentiation and loss of tumorigenicity of human cancers via pluripotency-based reprogramming. *Oncogene*, 32(18), pp.2249–2260.

Zhang, Y., Zhou, J., Fang, Z., Jiang, M. and Chen, X., 2013. Noggin versus basic fibroblast growth factor on the differentiation of human embryonic stem cells. *Neural Regeneration Research*, 8(23), pp.2171–2177.

Zhao, B. et al., 2014. Genome-wide mapping of miRNAs expressed in embryonic stem cells and pluripotent stem cells generated by different reprogramming strategies. *BMC Genomics*, 15(1), pp.1–11.

Zhao, W., Ji, X., Zhang, F., Li, L. and Ma, L., 2012. Embryonic stem cell markers. *Molecules*, 17(6), pp.6196-6236

Zhao, X. et al., 2014. Expression level of pluripotent genes in incomplete reprogramming. *Asian Pacific Journal of Tropical Medicine*, 7(8), pp.639–644.

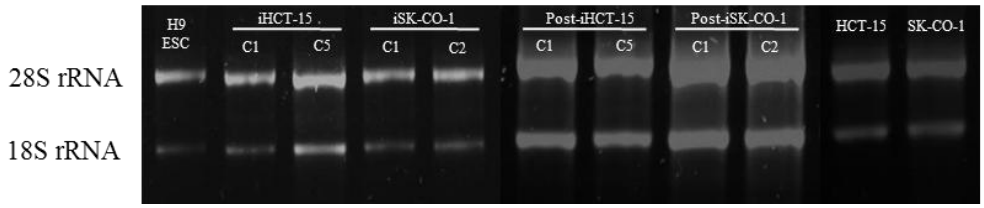
Zhou, W. and Freed, C.R., 2009. Adenoviral gene delivery can reprogram human fibroblasts to induced pluripotent stem cells. *Stem Cells*, 27(11), pp.2667–2674.

Zhu, S. et al., 2010. Brief report reprogramming of human primary somatic cells by OCT4 and chemical compounds. *Cell Stem Cell*, 7(6), pp.651–655.

APPENDICES

APPENDIX A

Integrity of RNA prepared from CRC, CRC-iPC and post-iPC cell lines



APPENDIX B

Supplementary Table 1: Fifty-two up-regulated miRNAs in CRC-iPCs

miRNA	Family[1]	Accession no[1]	Chrom. Location[1]	Sequence (5' → 3') [1]	Log_2 (FC)[2]
miR-125b-5p	mir-10	MIMAT0000423	11q24.1	tcacaagttagggtctc	8.881**
miR-199a-3p	mir-199	MIMAT0000232	19p13.2	taaccaatgtgcagactact	7.041**
miR-125a-3p	mir-10	MIMAT0004602	19q13.41	ggctcccaagaacctca	6.889**
miR-4734	NA	MIMAT0019859	17	cgccctgaccgc	6.360**
miR-6789-5p	NA	MIMAT0027478	19	cccgcgcgcc	6.237**
miR-4417	NA	MIMAT0018929	1	ccctccgggaagc	6.214**
miR-150-3p	mir-150	MIMAT0004610	19q13.33	ctgtcccccaggc	5.856**
miR-6723-5p	NA	MIMAT0025855	1	gccccgacgttact	5.773**
miR-3934-5p	mir-3934	MIMAT0018349	6	ctgcctcagtttcca	5.710**
miR-1181	mir-1181	MIMAT0005826	19	cggctcgggtgg	5.670**
miR-4745-5p	NA	MIMAT0019878	19	cgccgtcccgg	5.610**

Supplementary Table 1 (continued)

miRNA	Family[1]	Accession no[1]	Chrom. Location[1]	Sequence (5' → 3') [1]	Log_2 (FC)[2]
miR-4746-3p	NA	MIMAT0019881	19	tcggcccgcagg	5.522**
miR-671-5p	mir-671	MIMAT0003880	7q36.1	ctccagcccct	5.202**
miR-6757-5p	NA	MIMAT0027414	12	tcatcctggcctccc	5.136**
miR-199a-5p	mir-199	MIMAT0000231	19p13.2	gaacaggtagtctgaacac	5.062**
miR-769-5p	mir-769	MIMAT0003886	19q13.32	agctcagaacccagaggtc	5.019**
miR-4651	NA	MIMAT0019715	7	gcccgacctcacc	4.955**
miR-4488	mir-4488	MIMAT0019022	11	cgccggagccc	4.910**
miR-7152-3p	NA	MIMAT0028215	10	gcctcctgtccagg	4.889**
miR-6075	NA	MIMAT0023700	5	caccaatgccgcc	4.846**
miR-4655-5p	NA	MIMAT0019721	7	cgaccctctgcca	4.775**
miR-6778-5p	NA	MIMAT0027456	17	acctgcctcctgtcc	4.747**
miR-1228-3p	mir-1228	MIMAT0005583	12	gggggggcgagg	4.745**

Supplementary Table 1 (continued)

miRNA	Family[1]	Accession no[1]	Chrom. Location[1]	Sequence (5' → 3') [1]	Log_2 (FC)[2]
miR-6808-5p	NA	MIMAT0027516	1	catggtcccacctcc	4.737**
miR-6867-5p	NA	MIMAT0027634	17	tcccttcttcctctaca	4.729**
miR-4646-5p	NA	MIMAT0019707	6	tccctcagctcctct	4.723**
miR-6086	NA	MIMAT0023711	X	ctctgcccttccca	4.684**
miR-130a-3p	mir-130	MIMAT0000425	11q12.1	atgccctttaacattgca	4.646**
miR-3188	mir-3188	MIMAT0015070	19	ccccgtatccgca	4.550**
miR-4778-5p	NA	MIMAT0019936	2	cctcttcttcctttacag	4.496**
miR-4695-5p	NA	MIMAT0019788	1	cctgctcgcccac	4.437**
miR-152-3p	mir-148	MIMAT0000438	17q21.32	ccaagttctgtcatgc	4.246**
miR-132-3p	mir-132	MIMAT0000426	17p13.3	cgaccatggctgtaga	4.199**
miR-371a-5p	mir-290	MIMAT0004687	19q13.42	agtgcccccacag	4.177**
miR-652-5p	mir-652	MIMAT0022709	Xq23	tgaatggcaccctctc	4.148**

Supplementary Table 1 (continued)

miRNA	Family[1]	Accession no[1]	Chrom. Location[1]	Sequence (5' → 3') [1]	Log_2 (FC)[2]
miR-4487	NA	MIMAT0019021	11	ctgcccttcagcca	4.106**
miR-135a-3p	mir-135	MIMAT0004595	3p21.1	cgccacggctcca	3.963**
miR-195-5p	mir-15	MIMAT0000461	17p13.1	gccaatatttctgtgctgc	3.956**
miR-4690-5p	NA	MIMAT0019779	11	ttcagcccagcctcg	3.948**
miR-30a-5p	mir-30	MIMAT0000087	6q13	cttccagtcgaggatg	3.942**
miR-5195-3p	NA	MIMAT0021127	14	agccccctcagagaa	3.940*
miR-6798-5p	NA	MIMAT0027496	19	cccaagctcgccc	3.919**
miR-6752-5p	NA	MIMAT0027404	11	ccccccctggctcc	3.722*
miR-8089	NA	MIMAT0031016	5q35.3	ctgccccaatcccc	3.654**
miR-210-3p	mir-210	MIMAT0000267	11p15.5	tcagccgctgtcacac	3.647*
miR-4532	NA	MIMAT0019071	20	cgccgggctcc	3.468**
miR-4463	NA	MIMAT0018987	6	cgccccacccca	3.399**

Supplementary Table 1 (continued)

miRNA	Family[1]	Accession no[1]	Chrom. Location[1]	Sequence (5' → 3') [1]	Log_2 (FC)[2]
miR-513b-5p	mir-506	MIMAT0005788	Xq27.3	ataaatgacacctccttgt	3.380**
miR-5194	NA	MIMAT0021125	8	ccatcccattccaaac	3.308*
miR-1185-1-3p	mir-154	MIMAT0022838	14	ataagagtctccccctg	3.144**
miR-1249-3p	mir-1249	MIMAT0005901	22q13.31	tgaagaaggggggga	2.911*
miR-3911	NA	MIMAT0018185	9	tgcctcctccagga	2.641*

[1]miRBase version 21 was used in data derivation.[2] Microarray data of Log_2 Fold change (FC) relative to the parental CRC cells with*p<0.05 or **p<0.01. Only log_2(FC)> 2.0 or < -2.0 are shown. NA, not annotated

Appendix C

Supplementary Table 2: Fifty down-regulated miRNAs in CRC-iPCs

miRNA	Family[1]	Accession no[1]	Chrom. Location[1]	Sequence (5' → 3') [1]	Log_2 (FC)[2]
miR-192-5p	mir-192	MIMAT0000222	11q13.1	ctgtgacctatggaattgg	-6.167**
miR-338-3p	mir-338	MIMAT0000763	17q25.3	caacaaaatcactgatgctgg	-4.837**
miR-455-3p	mir-455	MIMAT0004784	9q32	gtgtatatgcccatgga	-4.629**
miR-362-5p	mir-362	MIMAT0000705	Xp11.23	actcacacctaggttcc	-4.415**
miR-6741-3p	NA	MIMAT0027384	1	ctagggtgagggaga	-4.366**
miR-6743-3p	NA	MIMAT0027388	11	tgtgggcagggaga	-4.336**
miR-552-3p	mir-552	MIMAT0003215	1p34.3	ttgtctaaccagtcacctgt	-4.154**
miR-6782-5p	NA	MIMAT0027464	17	acacccctgaattccc	-4.052**
miR-4254	NA	MIMAT0016884	1	gagatggtggagtagc	-4.023**
miR-4725-5p	NA	MIMAT0019843	17	ggtgggaaggctgc	-3.989**

Supplementary Table 2 (continued)

miRNA	Family[1]	Accession no[1]	Chrom. Location[1]	Sequence (5' → 3') [1]	Log_2 (FC)[2]
miR-3591-3p	mir-122	MIMAT0019877	18	gtggagtgtgacaatgg	-3.969**
miR-6737-3p	NA	MIMAT0027376	1	ctgggtaggggtga	-3.859**
miR-6129	mir-6129	MIMAT0024613	17	tatacacccaactccct	-3.850**
miR-4652-3p	NA	MIMAT0019717	7	tgaggggatgggttaa	-3.824**
miR-6752-3p	NA	MIMAT0027405	11	ctgggagtatggggg	-3.742**
miR-6766-3p	NA	MIMAT0027433	15	tgagggtgggggaa	-3.737**
miR-3935	NA	MIMAT0018350	16	gtggctggtgctcg	-3.639**
miR-200a-5p	mir-8	MIMAT0001620	1p36.33	tccagcactgtccggt	-3.625**
miR-4687-5p	NA	MIMAT0019774	11	tttgggtgcgggag	-3.608**
miR-362-3p	mir-362	MIMAT0004683	Xp11.23	tgaatccttgaataggtgtg	-3.597**
miR-3945	NA	MIMAT0018361	4	atatcaaccctctcctatgc	-3.566**
miR-6785-3p	NA	MIMAT0027471	17	ctggggaaggtggg	-3.540**

Supplementary Table 2 (continued)

miRNA	Family[1]	Accession no[1]	Chrom. Location[1]	Sequence (5' → 3') [1]	Log_2 (FC)[2]
miR-500a-3p	mir-500	MIMAT0002871	Xp11.23	cagaatccttgcccaggt	-3.535**
miR-192-3p	mir-192	MIMAT0004543	11q13.1	ctgtgacctatggaattgg	-3.515**
miR-126-3p	mir-126	MIMAT0000445	9q34.3	cgcattattactcacggt	-3.515**
miR-6803-3p	NA	MIMAT0027507	19	ctgagggtgagaaggc	-3.477**
miR-6730-5p	NA	MIMAT0027361	1	tctgacaacccctcc	-3.467**
miR-6797-3p	NA	MIMAT0027495	19	gtggggagggaagg	-3.383**
miR-149-5p	mir-149	MIMAT0000450	2q37.3	gggagtgaagacacggag	-3.359**
miR-505-3p	mir-505	MIMAT0002876	Xq27.1	aggaaaccagcaagtgttg	-3.336**
miR-3200-3p	mir-3200	MIMAT0015085	22	cagacctgagtagcgc	-3.294**
miR-6730-3p	NA	MIMAT0027362	1	tgagggcagatgggg	-3.280**
miR-181c-5p	mir-181	MIMAT0000258	19p13.13	actcaccgacaggttgaat	-3.247**
miR-421	mir-95	MIMAT0003339	Xq13.2	gcgcccaattaatgtctg	-3.241**

Supplementary Table 2 (continued)

miRNA	Family[1]	Accession no[1]	Chrom. Location[1]	Sequence (5' → 3') [1]	Log_2 (FC)[2]
miR-454-3p	mir-454	MIMAT0003885	17q22	accctataagcaatattgcac	-3.216**
miR-4323	NA	MIMAT0016875	19	tctgaggctgtgggg	-3.154**
miR-4769-3p	NA	MIMAT0019923	X	gtaggggagggagg	-3.147**
miR-6736-3p	NA	MIMAT0027374	1	ctgtgggtagagagga	-3.116**
miR-3940-3p	mir-3940	MIMAT0018356	19	aagtgggctgggatc	-3.103**
miR-4259	NA	MIMAT0016880	1	tcctgacccctagacc	-3.062**
miR-194-3p	mir-194	MIMAT0004671	11q13.1	cagataacagcagcccc	-3.033**
miR-500a-5p	mir-500	MIMAT0004773	Xp11.23	tctcacccaggtag	-2.849*
miR-183-3p	mir-183	MIMAT0004560	7q32.2	ttatggcccttcggt	-2.770*
miR-345-5p	mir-345	MIMAT0000772	14q32.2	gagccctggactag	-2.745*
miR-6779-3p	NA	MIMAT0027459	17	agatgggaggagacagg	-2.736*
miR-7-5p	mir-7	MIMAT0000252	15q26.1	acaacaaaatcactagtcttcc	-2.610*

Supplementary Table 2 (continued)

miRNA	Family[1]	Accession no[1]	Chrom. Location[1]	Sequence (5' → 3') [1]	Log_2 (FC)[2]
miR-335-3p	mir-335	MIMAT0004703	7q32.2	ggtcaggagcaataatgaaaaa	-2.605*

[1]miRBase version 21 was used in data derivation.[2] Microarray data of Log_2 Fold change (FC) relative to the parental CRC cells with *p<0.05 or **p<0.01. Only log_2 (FC)> 2.0 or < -2.0 are shown. NA, not annotated.