EXPRESSION AND BIOLOGICAL ROLE OF THE CHROMOSOME 19 microRNA CLUSTER (C19MC) IN STEM CELL PLURIPOTENCY

By

NGUYEN PHAN NGUYEN NHI

A project report 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
Doctor of Philosophy of Medical Sciences

May 2017

ABSTRACT

EXPRESSION AND BIOLOGICAL ROLE OF THE CHROMOSOME 19 microRNA CLUSTER (C19MC) IN STEM CELL PLURIPOTENCY

Nguyen Phan Nguyen Nhi

Introduction of the transcription factors, OCT4, SOX2, KLF4 and c-MYC (OSKM), is able to 'reprogramme' somatic cells to become induced pluripotent stem cells (iPSCs). Several microRNAs (miRNAs) are reported to enhance reprogramming efficiency when co-expressed with OSKM. A primate-specific chromosome 19 miRNA cluster (C19MC) is essential in primate reproduction, development and differentiation. In part (I) of this work, miRNA profiling by microarray analysis showed 261 differentially expressed miRNAs in the iPSCs relative to the adipose-derived mesenchymal stem cells (MSCs) and pre-adipose cells from which the iPSCs were derived. Of these, 40 pairs (80 miRNAs) co-existed, and were co-up- or co-down-regulated. *En bloc* C19MC miRNAs were found to be activated in pluripotent stem cells but only selectively expressed in MSCs. Selective C19MC miRNA expression was confirmed by miRNA copy number analysis, which also showed selective C19MC activation in cancer cells with similar expression patterns. Sixteen C19MC miRNAs share the "AAGUGC" seed sequence with the well-characterised reprogramming miR-302 family. Bioinformatics-predicted putative targets of the C19MC-AAGUGC-miRNAs are involved in induced

pluripotency by modulating apoptosis and pluripotency-associated signalling pathways.

In part (II) of this work, a C19MC miRNA, miR-524-5p, was chosen to elucidate C19MC contribution to events involved in reprogramming. MiR-524-5p, which is highly homologous to the reprogramming miR-520d-5p, was expressed only in iPSCs but not MSCs. Co-expressing miR-524 with OSKM in the human fibroblast HFF-1 resulted in two-fold significant increase in the number of alkaline phosphate- and NANOG-positive ESC-like colonies. Furthermore, the putative target, TP53INP1, showed an inverse relationship in mRNA and protein expression levels with miR-524-5p. Direct miR-524-5p targeting at the 3'-UTR of the TP53INP1 mRNA was confirmed in luciferase assays. Down-regulation of TP53INP1 by miR-524-5p over-expression enhanced cell proliferation, suppressed apoptosis and up-regulated expression of pluripotency genes, all of which are critical events of the initial phase of the reprogramming process. MiR-524-5p also suppressed the epithelial-mesenchymal transition-related genes, ZEB2 and SMAD4, to promote mesenchymal-epithelial transition, a critical initial event of reprogramming. In conclusion, specific C19MC miRNAs are important in regulating stem cell self-renewal and pluripotency, as functionally demonstrated by the analysis of miR-524-5p.

ACKNOWLEDGEMENTS

First and foremost, my sincere appreciation and gratefulness to my supervisor, Senior Professor Dr. Choo Kong Bung, for his continuous guidance, encouragement, patience and advice to teach me throughout the project. I would like to thank my co-supervisor, Emeritus Professor Dr. Cheong Soon Keng, for his support and advice, and to express my gratitude to my co-supervisor Professor Huang Chiu-Jung, Chinese Culture University, Taiwan, for her opinions and feedback throughout the work. I have learned a lot from my supervisors and to have broadened my perspective in stem cell research and regulation of gene expression, and also in what it means to be a good researcher.

I would also like to thank Dr. Shigeki Sugii from Singapore BioImaging Consortium, A*Star, Singapore, Professor Dr. Lim Yang Mooi (UTAR) and Cryocord Sdn. Bhd. for providing the cell lines used in this study. This work was supported by a HIR-MoE Grant (UM.C/625/1/HIR/MOHE /CHAN/03) and UTARRF grants (6200/C70, 6200/C90 and 6200/CB7). I thank UTAR and Faculty of Medicine and Health Sciences for providing me with laboratory instruments and under good circumstances.

Not to forget my fellow lab-mates in the UTAR Postgraduate Laboratories, especially Tai Lihui, Michele Hiew Sook Yuin, Chai Kit Man and Vimalan Rengganatan, and in later stage of my candidature, Cheng Hang

Ping. I sincerely appreciate their sharing, encouragements and supports throughout the project. Last but not least, I am deeply indebted to my family for their endless sacrifice and emotional support and concern, which had enabled me to complete my postgraduate studies far away from home.

APPROVAL SHEET

This thesis entitled **"EXPRESSION AND BIOLOGICAL ROLE OF THE CHROMOSOME 19 microRNA CLUSTER (C19MC) IN STEM CELL PLURIPOTENCY"** was prepared by NGUYEN PHAN NGUYEN NHI and submitted as partial fulfilment of the requirements for the degree of Doctor of Philosophy of Medical Sciences at Universiti Tunku Abdul Rahman.

Approved by:

_____

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

_____

(Emeritus Prof. Dr. Cheong Soon Keng) Date: ……
Emeritus 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 THESIS

It is hereby certified that **Nguyen Phan Nguyen Nhi** (ID no: **14UMD06063**) has completed this final year project entitled "EXPRESSION AND BIOLOGICAL ROLE OF THE CHROMOSOME 19 microRNA CLUSTER (C19MC) IN STEM CELL PLURIPOTENCY" 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 and Prof. Dr. Huang Chiu-Jung (Co-Supervisor) from Department of Animal Science & Graduate Institute of Biotechnology, Chinese Culture University, Taipei, Taiwan.

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

Yours truly,

_____

(Nguyen Phan Nguyen Nhi)

DECLARATION

I NGUYEN PHAN NGUYEN NHI hereby declare that the thesis 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.

_____

NGUYEN PHAN NGUYEN NHI

Date: _____

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

3’UTR	3’ untranslated region
3-D scaffolds	Three-dimensional scaffolds
AID	Activation induced cytidine deaminase
AKT3	AKT serine/threonine kinase 3
AOF1	Lysine demethylase 1B
AOF2	Lysine demethylase 1A
AP	Alkaline phosphatase
ASCs	Adipose-derived stem cells
BM	Bone marrow
BM-MSCs	Bone marrow-derived mesenchymal stem/stromal cells
BMP2	Bone morphogenetic protein 2
BMPs	Bone morphogenetic proteins
BrdU	5-bromo-2’-deoxyuridine
BSA	Bovine serum albumin
C/EBPs	CCAAT/enhancer binding proteins
C19MC	Chromosome 19 miRNA cluster
CDH1	Cadherin 1
CDKN1A/p21	Cyclin dependent kinase inhibitor 1A
CDS	Coding sequence
CFU-F	Colony-forming unit fibroblasts
DGCR8	Double-stranded RNA-binding domain protein
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethyl sulfoxide
DNMT1	DNA methyltransferase 1
DNMT3B	DNA methyltransferase 3 beta
DOT1L	Histone H3 lysine methyltransferase
Dox	Doxycycline
DPPA2	Developmental pluripotency associated 2

ECM	Extracellular matrix
EdU	5-ethynyl-2'-deoxyuridine
ELISA	Enzyme-Linked Immunosorbent Assay
EMT	Epithelial to mesenchymal transition
EPCAM	Epithelial cell adhesion molecule
ER-	Estrogen receptor negative
ESCs	Embryonic stem cells
ESRRB	Estrogen related receptor beta
ETMRs	Multilayered rosettes
Exp5	Exportin-5
FGF1/2	Fibroblast growth factor 1/2
GO	Gene Ontology
H3K79me2	Histone H3 lysine 79
HCC	Hepatocellular carcinoma
HGF	Hepatocyte growth factor
hNSCs	Neural stem cells
HRP	Horseradish peroxidase
HSCs	Hematopoietic stem cells
HWP	Human white preadipocyte
IGF1	Insulin-like growth factor 1
IL-17	Interleukin
Int	Introns
iPSCs	Induced pluripotent stem cells
KEGG	Kyoto Encyclopedia of Genes and Genomes
LOX	Lysyl oxidases
MEF	Mouse Embryonic Fibroblast
MEM	Minimum Essential Medium
MET	Mesenchymal to epithelial transition
miRNAs	MicroRNAs
MSCs	Mesenchymal stem cells

MTT	3-(4,5-Dimethylthiazol-2-yl)-2,5-Diphenyltetrazolium Bromide
$NaHCO_3$	Sodium bicarbonate
NC	Negative control
NF-κB	Nuclear factor κB
NR2F2	Nuclear receptor subfamily 2 group F member 2
Nr5a2	Nuclear receptor subfamily 5 group A member 2
NSCLC	Non-small cell lung cancer
nt	Nucleotide
OSKM	OCT4, SOX2, KLF4, c-MYC
p16/INK4A	Cyclin dependent kinase inhibitor 2A
p53	Tumor protein p53
PNETs	Primitive neuroectodermal tumour
PPARγ	Peroxisome proliferator-activated receptor gamma
Pre-miRNA	Precursor miRNA
Pri-miRNAs	Primary miRNAs
PTEN	Phosphatase and tensin homolog
RB	Retinoblastoma
RBL2	RB transcriptional corepressor like 2
RISC	RNA-induced silencing complex
RPMI	Roswell Park Memorial Institute
SALL4	Spalt like transcription factor 4
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
SMA	Spinal muscular atrophy
SSEA1	Stage-specific embryonic antigen 1
SVF	Stromal vascular fraction
TBX3	T-box 3
TEMED	Tetramethylethylenediamine
TGFβ	Transforming growth factor β
TGFβ2	Transforming growth factor beta 2

TGFβR2	Transforming growth factor β receptor II
TIMP2	TIMP metallopeptidase inhibitor 2
TNFs	Tumor necrosis factors
TRBP	Transactivation-response RNA-binding protein
TTYH1	Tweety family member 1
WAT	White adipose tissue
ZEB1/2	Zinc finger E-box binding homeobox 1/2

CHAPTER 1

INTRODUCTION

Forced over-expression of core transcription factors, including OCT4, SOX2, KLF4 and c-MYC, abbreviated as OSKM, is able to 'reprogramme' somatic cells to become undifferentiated pluripotent stem cells (Takahashi and Yamanaka, 2006; Takahashi et al., 2007). The reprogrammed cells, called induced pluripotent stem cells (iPSCs), exhibit the essential characteristics of embryonic stem cells (ESCs) in proliferation and differentiation capability (Takahashi and Yamanaka, 2006). The reprogramming process is thought to involve three phases, *viz.* initiation, maturation and stabilization, each of which is driven by a cascade of expression changes in specific sets of genes to give rise to fully or partially reprogrammed cells (Buganim et al., 2013; David and Polo, 2014). Some important features of the early stage of reprogramming include increased proliferation, inhibition of apoptosis, acquisition of epithelial characteristics and up-regulation or activation of pluripotency-related genes (David and Polo, 2014). Due to low reprogramming efficiencies, elucidating the molecular events that regulate each step of the reprogramming process has been challenging (Plath and Lowry, 2011).

MicroRNAs (miRNAs) are 17-22-nucleotide non-coding and single-stranded RNAs that negatively regulate gene expression by inhibiting protein translation or destabilising mRNAs of the target genes (Bartel, 2009). An

ever-expanding panel of miRNAs has been shown to play important roles in modulating somatic cell reprogramming at the early and late stages (Henzler et al., 2013; Li et al., 2014) to govern pluripotent properties (Anokye-Danso et al., 2012; Y. Wang et al., 2013). MiRNAs have also been reported to promote dedifferentiation of somatic or cancers cells to pluripotency (He et al., 2014; Tsuno et al., 2014). MiRNA genes are often physically clustered in the genome to permit co-regulation (Leonardo et al., 2012); clustered miRNAs are often transcribed as a single primary transcript (Stadler et al., 2010). One such human miRNA cluster is mapped on chromosome 19, and is called the chromosome 19 miRNA cluster, or C19MC (Bortolin-Cavaille et al., 2009), a subject of the present thesis work.

C19MC, one of the largest miRNA gene clusters in the human genome, contains 46 miRNA genes within a ~100-kb genomic region (Bortolin-Cavaille et al., 2009). C19MC is a product of late evolution, and is found only in primates (Zhang et al., 2008). Hence, C19MC has been predicted to play critical roles in primate reproduction, development and differentiation (Lin et al., 2010), as is reflected in its restrictive expression in only reproductive tissues and in pluripotent ESCs (Lin et al., 2010; Razak et al., 2013; Liang et al., 2007). Moreover, aberrant activation of C19MC expression has been associated with the pathogenesis and progression of various forms of cancer (Li et al., 2009; Kleinman et al., 2014; Spence, Sin-Chan, et al., 2014).

It was recently reported that miR-520d-5p, which is a member of C19MC, was involved in reprogramming (Tsuno et al., 2014). It is, however, unclear if other C19MC miRNAs also contribute to inducing and regulating pluripotency.

Aim and specific objectives of the present study

The major goal of this study was to extend our understanding of the expression profile and the biological role of C19MC miRNA in stem cell pluripotency. The goal was achieved through the following specific study aims:

1. To first investigate the genome-wide miRNA expression profiles in general in iPSC lines derived from mesenchymal stem cells (MSCs).
2. To determine the expression and possible biological functions of C19MC miRNAs in various stem cell types, and in cancer cells.
3. Using miR-524-5p as a representative member of the C19MC cluster, this thesis also aimed to elucidate possible contribution and mechanism of a C19MC miRNA in the reprogramming process.

Major findings of the present study

Microarray data revealed *en bloc* C19MC expression in pluripotent stem cells, but only expression of specific C19MC miRNAs in the multipotent

mesenchymal stem cells (MSCs). Interestingly, the C19MC expression profiles in MSCs are highly similar to those in cancer cells. A subset of C19MC miRNAs shares the same "AAGUGC" seed sequence with miR-302 miRNAs, which are known cellular reprogramming factors, supporting C19MC involvement in the reprogramming process. Further focussing on a member of C19MC, the data showed that miR-524-5p, by targeting TP53INP1 and the epithelial to mesenchymal transition (EMT)-related genes ZEB2 and SMAD4, contributes to the early stage of induction of cellular pluripotency via regulation of cell proliferation, apoptosis, expression of pluripotency genes and EMT.

Data present in this work advance our understanding of possible biological roles of C19MC miRNAs in induction of pluripotency and, possibly, in the tumorigenesis process.

CHAPTER 2

LITERATURE REVIEW

2.1 General Introduction to Stem Cells

Stem cells are undifferentiated cells that do not possess many of the phenotypic characteristics of known adult tissues including epithelial, immune, neural, connective and muscle. However, stem cells have the capability to self-renew and to give rise to multiple differentiated cell types found in the adult tissues. Stem cells are present in the postnatal and all adult stages of life. Hence, stem cells are one of most essential biological components required for growth and development during embryogenesis. Yet stem cells have also play important roles in cellular replenishment of terminally differentiated cell types as stem cells replenish blood, muscle, epithelia, nervous system and other tissues with fresh cells throughout life. The major characteristics of stem cells are: (i) to potentially proliferate for unlimited period of time or prolonged self-renewal, (ii) to arise from a single cell (clonality), (iii) to differentiate into at least one type of mature, specialised cell (Volarevic et al., 2014).

Stem cells can be classified based on differentiation potentials (Figure 2.1). A fertilized egg (zygote) may be considered as the ultimate stem cell from which all stem cells originate. Therefore, a fertilised egg is a totipotent

cell. During growth and development, the totipotent fertilised egg drives the formation of both embryonic and extraembryonic tissues, which subsequently form the embryo and the placenta, respectively (Rossant, 2001). Following totipotent cells, pluripotent stem cells are able to drive the formation of the three germ layers, ectoderm, mesoderm and endoderm, thereby allowing the development of all tissues and organs (De Miguel et al., 2010). As an example, embryonic stem cells (ESCs) derived from the inner cell mass of the blastocyst are pluripotent in being able to differentiate into almost all cell type arising from the three germ layers (Evans and Kaufman, 1981). Recently, induced pluripotent stem cells (iPSCs) generated from adult somatic cells have been shown to share similar characteristics with ESCs (Takahashi and Yamanaka, 2006). The next level of differentiation potency is found in multipotent stem cells that are found in and isolated from various tissues. Multipotent stem cells are able to more limited in differentiation into specific differentiated cell types from a single germ layer (Ratajczak et al., 2012). An example of the most recognised multipotent stem cells is the mesenchymal stem cells (MSCs). MSCs are capable of differentiation into diverse mesoderm-derived lineages including adipose tissue, bone, cartilage and muscle (Ong and Sugii, 2013; Bara et al., 2014; Goldberg et al., 2017; Pavyde et al., 2016), and into ectoderm-derived tissue such as cells with neuronal-like morphology (Takeda and Xu, 2015). Eventually in the mature adult multipotent stem cells differentiate into unipotent stem cells which are capable of further differentiating into only one specific cell type (Beck and Blanpain, 2012). Glial cells are example of unipotent cells that can only differentiate into their own mature lineage. Another example of

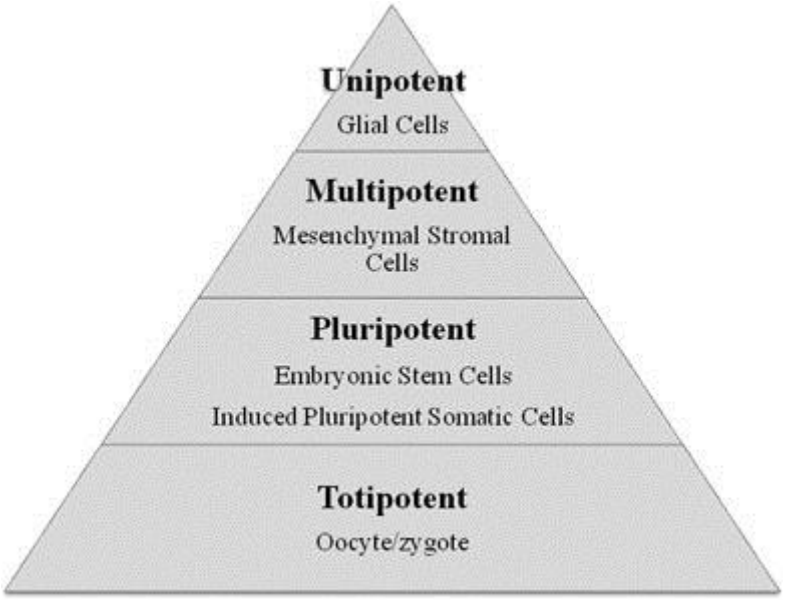


Figure 2.1 Specialisation of stem cells according to differentiation potentials. Stem cells have different degrees of differentiation potential. See text for details. [Figure was modified from (Spencer et al., 2011).]

unipotent cells is human white preadipocyte (HWP) which is only able to differentiate into mature adipocytes.

2.2 Mesenchymal Stem Cells (MSCs)

2.2.1 Overview of Mesenchymal Stem Cells (MSCs)

Adult mesenchymal stem cells are multipotent stem cells that have capability to self-renew and differentiate into both mesenchymal and nonmesenchymal lineages (Ong and Sugii, 2013; Takeda and Xu, 2015). The multilineage potentials have made MSCs the most attractive source for regenerative medicine in the recent decades. Historically, MSCs were first described in the study of Friedenstein and colleagues (Friedenstein et al., 1970). The study identified a minor subpopulation in the bone marrow (BM), which constituted about 0.01% to 0.001% of the mononuclear cells in BM, isolated based on their adherence to plastic. These plastic-adherent cells were able to form single-cell colonies, thus were given the name colony-forming unit fibroblasts (CFU-F) (Friedenstein et al., 1970). The plastic-adherent BM cells were later termed as "mesenchymal stem cell" because the cells could be differentiated into mesodermal cells *in vitro* and *in vivo* (Caplan, 1991). MSCs, today also known as multipotent stromal cells, are self-renewable and possess multipotent properties to allow the cells to proliferate and differentiate into mature cells of ectodermal, mesodermal and endodermal origins. MSCs are found in almost all adult tissues of diverse sources including bone marrow,

adipose tissue, umbilical cord, amniotic fluid, dental pulp, peripheral blood, synovial membranes and many more (X. Zhang et al., 2014; Choo, Tai, et al., 2014).

The hematopoietic microenvironment that is termed the "hematopoietic niche" is a local tissue microenvironment that maintains and regulates the development and differentiation of hematopoietic stem cells (HSCs) (Prockop, 1997). Bone marrow-derived mesenchymal stem/stromal cells (BM-MSCs) are one of most essential cells in the "hematopoietic niche" (Pittenger and Martin, 2004). Currently, no cell surface marker has been identified that is able to distinguish MSCs from HSCs. Therefore, the Mesenchymal and Tissue Stem Cell Committee of the International Society for Cellular Therapy has set minimal criteria to standardise human MSC isolated *in vitro*. Firstly, MSCs must have plastic adherent property under standard culture conditions. Secondly, MSCs must express positively for the CD105, CD73 and CD90 surface markers, and do not express the CD45, CD34, CD14 or CD11b, CD79α or CD19 and HLA-DR hematopoietic markers. Lastly, MSCs have the ability to differentiate *in vitro* into cells types of the three germ layers including osteoblasts, adipocytes and chrondroblasts (Dominici et al., 2006).

However, MSCs isolated from different adult tissue sources using identical culture conditions express significant differences in morphology, differentiation potential and gene expression profile (Nombela-Arrieta et al.,

2011). Hence, the MSCs derived from various sources may not be similar and display different degrees of self-renewal and multipotential capabilities.

2.2.2 Adipose-derived Stem Cells and Clinical Applications

The standard nomenclature recommended by the International Fat Applied Technology Society for plastic-adherent cells derived from the adipose tissue is designated as the adipose-derived stem cells (ASCs) (Gimble et al., 2007). ASCs are the MSCs isolated from the stromal vascular fraction (SVF) of infant adipose tissues (Rodriguez et al., 2005). ASCs display the characteristics of MSCs derived from the bone marrow (Friedenstein et al., 1970; Caplan, 1991), i.e., the self-renewable capacities that can be proliferated *in vitro* for more than 160 population doublings while maintaining the normal diploid karyotype. Currently, no unique cell surface markers for defining ASCs have been identified. Some differentially-expressed cell surface markers between ASCs and BM-MSCs have been reported despite the similarity between ASCs and BM-MSCs. CD49d is found to be expressed in ASCs but not in BM-MSCs while ASCs do not express CD49f, CD104 and CD106 the expression of which is detected in BM-MSCs (Lindroos et al., 2011; Pachón-Peña et al., 2011; Cawthorn et al., 2012).

ASCs are advantageous for clinical applications over the more commonly used BM-MSCs. ASCs abundantly reside in the white adipose tissue (WAT) localised throughout the body. In addition, ASCs can be isolated by using the minimally invasive liposuction (Ong and Sugii, 2013), and the

quantity of ASCs that can be obtained is relatively higher compared to only 0.01 - 0.001% BM-MSCs in bone marrow (Fraser et al., 2006). Furthermore, less implant migration and foreign-body reaction is observed when ASCs are transplanted to autologous or allogeneic body (Sterodimas et al., 2010). Hence, ASCs have become the most valuable source of adult stem cells with MSC potency for tissue engineering and regenerative medicine.

2.2.2.1 Plasticity of Human Adipose-derived Stem Cells

Numerous reports have demonstrated that ASCs possess the ability to differentiate into various cell types of multiple different lineages *in vitro* and *in vivo* (Figure 2.2) (Bunnell et al., 2008; Feisst et al., 2015). Given that ASCs originate from the mesoderm, it is not surprising that ASCs have been experimentally demonstrated in numerous studies to be able to differentiate into adipocytes, chondrocytes, osteoblasts and myocytes (Mizuno, 2009; Dai et al., 2016). However, the induction of ASCs *in vitro* to differentiate into these lineages requires the ASCs to be seeded onto suitable polymeric three-dimensional (3-D) scaffolds in culture media supplemented with specific growth factors under a well-defined external environment (Brayfield et al., 2010; Dai et al., 2016). For example, scaffold-derived bone morphogenetic protein 2 (BMP2) released at the site of transplantation was required for undifferentiated ASCs on 3-D scaffolds to form osteoid *in vivo* (Jeon et al., 2008).

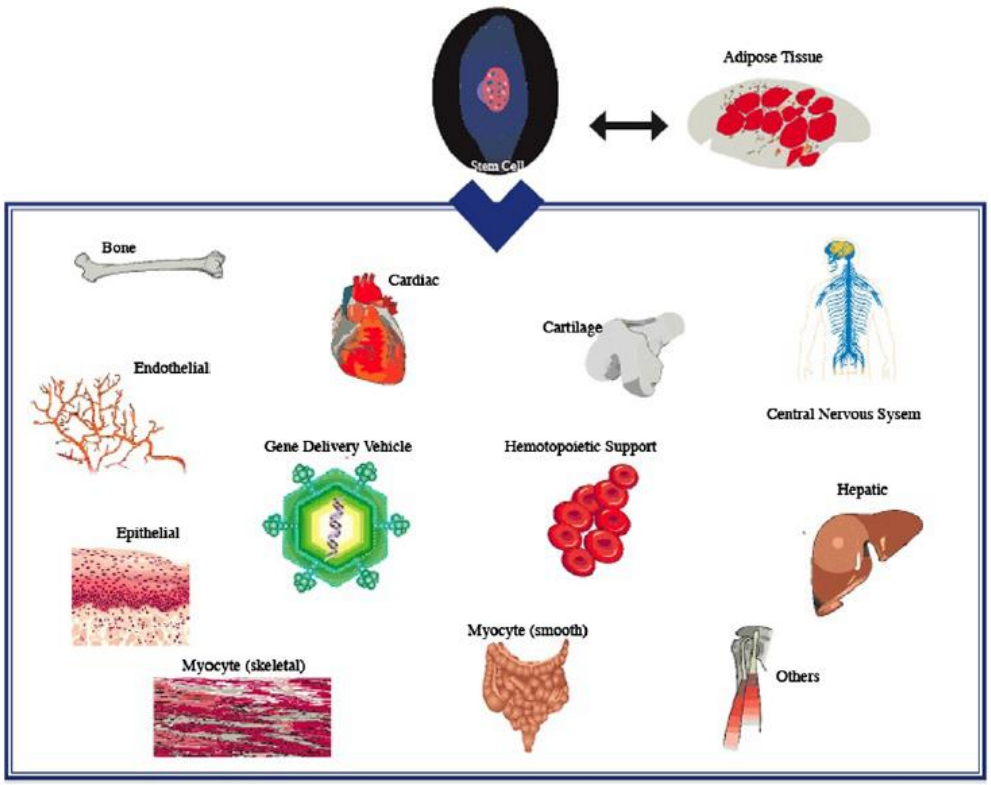


Figure 2.2 Multilineage differentiation potential of adipose-derived stem cells (ASCs). [Figure was modified from (Bunnell et al., 2008).]

Despite the mesodermal origin of ASCs, recent reports have also shown that ASCs are also able to transdifferentiate into non-mesodermal ecto- and endodermal lineage cells (Mizuno, 2009; Dai et al., 2016). Other studies have confirmed that ASCs could differentiate into ectodermal neuronal-like cells based on both morphology and function (Seo et al., 2005). Interestingly, the neuronal-like cells induced from ASCs are similar to those from BM-MSC (Ning et al., 2006) but with a higher proliferation capacity (Han et al., 2014). Intravenous administration of neuron and oligodendrocyte, which were differentiated from rat ASCs, was shown to significantly improve locomotor functions (Kang et al., 2006). The ability of ASCs to differentiate into endoderm-lineage cells was also reported. Several studies have demonstrated that ASCs could be induced into functional hepatocytes by the presence of cytokines hepatocyte growth factor (HGF) and fibroblast growth factor 1 (FGF1), or through co-culturing with other cell types (Aurich et al., 2009; Y. Zhang et al., 2014). Altogether, the plasticity of ASCs suggests that ASCs are a promising source for cell-based therapy.

2.2.2.2 Development of Adipose-derived Stem Cells into The Adipocyte Lineage

Not surprisingly, many studies have demonstrated that ASCs are able to promote adipogenesis (Bunnell et al., 2008; Mizuno, 2009; Dai et al., 2016). Adipogenesis is the process in which ASCs differentiate into mature adipocytes. The mature adipocytes are maintained by replacement of dead cells under a normal metabolic state whereas the number of adipocytes is

increased for hyperplastic WAT expansion in response to increased energy intake (Cawthorn et al., 2012; Ong and Sugii, 2013). The differentiation of ASCs is a two-step process. The first stage of adipogenesis involves the generation of the unipotent preadipocyte from ASCs, followed by terminal differentiation of preadipocytes into mature adipocytes. The mechanism that mediates the commitment of multipotent ASCs to the adipocyte lineage and the identity of the committed preadipocytes remain unclear. Bone morphogenic proteins (BMPs), insulin-like growth factor 1 (IGF1), interleukin 17 (IL-17), Wnt, transforming growth factor β (TGFβ), activin, FGF1 and FGF2 are identified as positive regulators for preadipocyte commitment (Tang and Lane, 2012). The regulatory adipocytic processes underlying terminal differentiation into adipocytes are well-defined by using 3T3-L1 as the preadipocyte cell model (Poulos et al., 2010). Generally, the adipogenic cascade is governed by the expression and activation of major transcriptional activators including peroxisome proliferator-activated receptor gamma (PPARγ), members of CCAAT/enhancer binding proteins (C/EBPs) and other transcription factors such as lysyl oxidase (LOX) (Lowe et al., 2011; Cawthorn et al., 2012).

2.3 Induced Pluripotent Stem Cells (iPSCs)

2.3.1 Induced Pluripotency in Somatic Cells

Human embryonic stem cell (ESC), derived from human blastocysts, is a unique cell type that is able to undergo indefinite self-renewal and possesses full developmental potentials to form all types of the somatic cells (Bradley et al., 1984). Such cells are also known as pluripotent cells, which have great clinical potentials to be used in cell-based therapies and regenerative medicine. However, besides ethical considerations, issues of immune rejection stemming from incompatibility between patient and donor cells have called for alternative approaches in generating pluripotent stem cells. Alternative sources of pluripotent stem cells have been investigated. Such pluripotent stem cells were first successfully generated by nuclear transfer, followed by the use of selected transcription factors (Wilmut et al., 1997; Takahashi and Yamanaka, 2006). Somatic cell nuclear transfer is able to reprogramme somatic cell nuclei to an undifferentiated state (Wilmut et al., 1997) while introduction of four transcription factors, OCT4, SOX2, KLF4 and c-MYC, or OSKM in short, by retrovirus-mediated transduction was also able to 'reprogramme' mouse and human fibroblast cells to undifferentiated pluripotent stem cells (Takahashi and Yamanaka, 2006). Similar to ESCs, the induced pluripotent stem cells (iPSCs) display ESC-like morphology, express pluripotent markers and have the potential to differentiate into the three germ layers of ectoderm, mesoderm, and endoderm. Remarkably, iPSCs could form teratoma *in vivo*, which is a noncancerous tumour containing differentiated progeny cells of the three germ

layers (Lee et al., 2013). Therefore, iPSCs could offer invaluable therapeutic implications in terms of *in vitro* disease modelling, pharmaceutical screening and cellular replacement therapies. Furthermore, the immune rejection issue can be easily overcome by iPSCs since they may be derived from the same patient.

2.3.2 Therapeutic Application of iPSCs

Similar to ESCs, iPSCs have unlimited self-renewable capacity and plasticity for trilineage differentiation. Furthermore, iPSCs could form differentiated cells that are functionally similar to those derived from ESCs, i.e. cardiomyocytes generated from human iPSCs and from human ESCs share similar properties (Zhang et al., 2009). Therefore, iPSCs are the ideal candidate to replace ESCs for autologous tissue regeneration and transplantation. It has been shown that iPSCs are able to form cardiac cells with different subtypes such as nodal-, atrial- and ventricular-like phenotypes, which are subsequently differentiated into cardiovascular lineages (Germanguz et al., 2011). Transplantation of cardiac progenitors induced from BM-MSC-derived iPSCs in a mouse model of acute myocardial infarction was shown to decrease infarct-size expansion and to enhance global cardiac functions with extensive survival and engraftment (Buccini et al., 2012). In that study, no cardiac tumorigenesis at the engraftment site was observed when compared to 21% in those transplanted with BM-MSC-derived iPSCs. The reasons for the avoidance of tumorigenesis are unclear.

Another important application of iPSCs is the possibility of correcting diseases caused by mutations by restoring normal physiological functions via gene targeting and correction technologies. A mouse model of sickle-cell anemia using iPSC-derived hematopoietic stem cells to correct the genetic defect was an early example of potential applications of iPSC technology in corrective medicine for genetic diseases (Hanna et al., 2007). In that study, the fibroblasts of mice with a hematopoietic genetic mutation were reprogrammed into iPSCs of which the defective gene was repaired by homologous recombination. The genetically corrected iPSCs were subsequently differentiated into hematopoietic progenitors for transplantation into the disease mice (Hanna et al., 2007). The results showed significant increased levels of β-globin proteins A and decreased levels of β-globin proteins B; there were also improvements of the symptoms of anemia and restoration of normal hemoglobin levels in the diseased model (Hanna et al., 2007).

The use of iPSCs generated from patient-specific cells allows recapitulation of pathologic processes *in vitro* for applications in disease modelling. One typical example of the application of iPSCs in disease modelling is demonstrated in the study of spinal muscular atrophy (SMA), a neurodegenerative genetic disorder which causes the loss of lower motor neurons (Ebert et al., 2009). The patient-specific iPSC-derived motor neurons have similar morphology and number comparing with those derived from wild-type iPSCs. However, after eight weeks in culture, the number and size of the patient iPSC-derived motor neurons selectively declined, suggesting that iPSC-derived motor neurons generated from patient have specific features

that are different from iPSCs from normal fibroblasts (Ebert et al., 2009). Such differences are likely reflection of the disease state that recapitulates disease progression to constitute a useful SMA disease model.

2.3.3 Molecular Insights of the Dynamics of Cellular Reprogramming

In order to create high-quality pluripotent cells suitable for clinical applications, a better understanding of the molecular mechanisms underlying the reprogramming process from somatic cell to pluripotency is necessary. Numerous studies have begun to interpret the events that occur during the reprogramming process (Hockemeyer et al., 2008; Buganim et al., 2013; David and Polo, 2014; Hussein et al., 2014; Smith et al., 2016). In general, cellular reprogramming seems to involve three major phases: initiation, maturation and stabilisation as discussed in more details below (Figure 2.3)

2.3.3.1 Initiation Phase

The initiation phase is marked by the changes from epithelial-to-mesenchymal transition (EMT) to mesenchymal-to-epithelial transition (MET). In molecular terms, the somatic transcriptional programme, which involves the down-regulated expression of the transcription factors SNAIL1/2 or zinc finger E-box binding homeobox 1/2 (ZEB1/2), is suppressed and is replaced by the gain of an epithelial signature defined by the up-regulation of

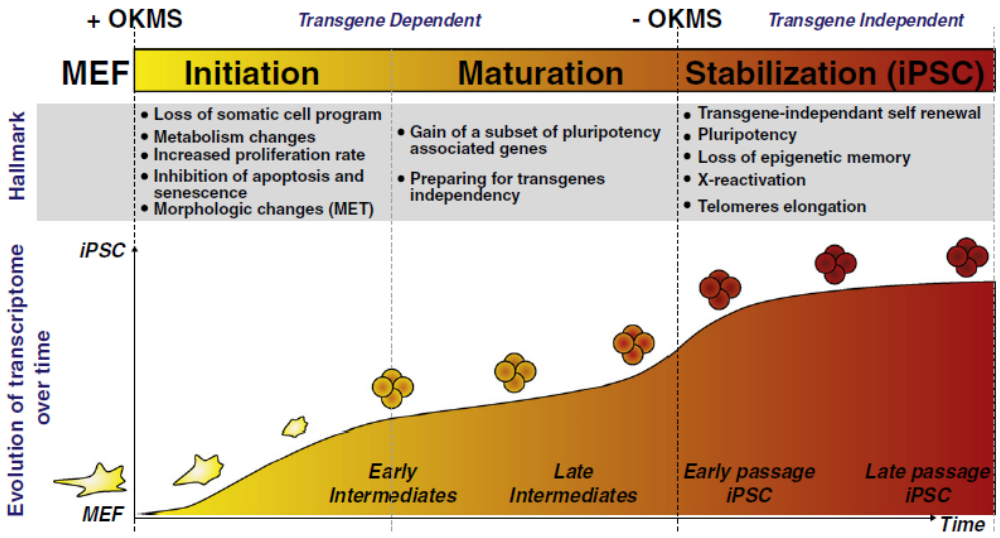


Figure 2.3 Sequential events-mediated reprogramming processes. [Figure modified from (David and Polo, 2014).]

cadherin 1 (CDH1), epithelial cell adhesion molecule (EPCAM) and other MET genes (Mikkelsen et al., 2008; Stadtfeld, Maherali, et al., 2008; David and Polo, 2014). Morphologically, the cells undergoing the initial phase of transition change from motile, multipolar or spindle-shape, all of which are mesenchymal features, to the typical epithelial characteristics of tight intracellular contacts with planar arrays of polarisation. Besides the MET-associated changes, enhancement of proliferation, inhibition of apoptosis and senescence cascades are crucial for the initiation phase of reprogramming (Mikkelsen et al., 2008; Buganim et al., 2013; David and Polo, 2014). Silencing of the tumour suppressor genes tumour protein p53 (p53), cyclin dependent kinase inhibitor 1A (CDKN1A/p21) and cyclin dependent kinase inhibitor 2A (p16/INK4A) was also observed leading to suppress cell cycle activities and enhance reprogramming (Ruiz et al., 2011). The initiation phase is considered complete when the first pluripotency-related genes are expressed. Experimentally, the initiation phase is marked by positive staining of alkaline phosphatase (AP) and stage-specific embryonic antigen 1 (SSEA1) (Buganim et al., 2013; David and Polo, 2014). The observed changes in gene and protein expression during the initiation phase also correlate with the epigenetic modifications of chromatin remodelling (Pasque et al., 2011; Plath and Lowry, 2011). The histone H3 lysine 79 (H3K79me2) methyltransferase DOT1L has an inhibitory effect on the reprogramming process in the early to middle phase by blocking KLF4 induction of MET (Onder et al., 2012). Silencing DOT1L results in enhancement of reprogramming efficiency by mediating the loss of H3K79me2 from EMT-related genes such as SNAIL1/2, ZEB1/2, and transforming growth factor beta 2 (TGFβ2).

2.3.3.2 Maturation Phase

Following initiation, the pluripotency-related genes are gradually activated during the maturation phase, which is the intermediate phase of reprogramming. In the maturation phase, early pluripotency markers nuclear receptor subfamily 5 group A member 2 (Nr5a2), NANOG, spalt like transcription factor 4 (SALL4), T-box 3 (TBX3), estrogen related receptor beta (ESRRB) and endogenous OCT4 are activated (Buganim et al., 2013; David and Polo, 2014). Some of these markers, including Nr5a2, TBX3, NANOG and ESRRB, have been reported to play pivotal roles in maintaining self-renewal of ESCs and are being used to replace some of the core reprogramming factors (Heng et al., 2010; Han et al., 2010; Hanna et al., 2009; Feng et al., 2009). However, it remains unclear how the maturation markers are activated and interconnected with exogenous OSKM. Some potential regulators have been hypothesised to be involved in the regulation of the maturation phase. The identified regulators are shown to suppress the maturation phase through BMP signalling, such as the suppression of the Mbd3 subunit of the NuRD complex and the recruitment of polycomb and Utx to genes involved in iPSC maturation (Chen et al., 2013; Luo et al., 2013; Welstead et al., 2012). It is noteworthy that in the late reprogramming event, expression of the exogenous transgenes OSKM is suppressed, and iPSCs are able to self-renew independent of the transgenes. The determinants of transgene independency were studied by using clonal analysis and the ability to survive transgene inhibition (Golipour et al., 2012). The study showed that termination of transgene expression in the late stage of the maturation phase

was required for the subsequent stabilisation phase and the eventual acquisition of pluripotency.

2.3.3.3 Stabilisation Phase

The reprogrammable cells can enter the stabilisation phase immediately after the suppression of transgene expression and the activation of the late pluripotency markers such as SOX2 and developmental pluripotency associated 2 (DPPA2) (Buganim et al., 2013; David and Polo, 2014). The activation of SOX2 triggers the expression of a core pluripotency circuitry by which the reprogrammable cells are able to be maintained independently from ectopic expression of the reprogramming factors and eventually be stabilised into the pluripotency state (Buganim et al., 2013). The endogenous pluripotency-associated genes OCT4, SOX2 and NANOG play essential roles in sustaining the pluripotency at this point (Boyer et al., 2005; Buganim et al., 2013). Numerous epigenetic changes are also in action while the iPSCs are at a pluripotency state. In a mouse model, the telomere is elongated to an ESC-like state and the inactive X chromosome in iPSCs derived from female mice is reactivated (Stadtfeld, Maherali, et al., 2008; David and Polo, 2014). Furthermore, the epigenetic memory found in both mouse and human iPSCs is reset (Buganim et al., 2013). The epigenetic reset is likely regulated by DNA methylation as treatment with 5-azacytidine, a DNA methylase inhibitor, was shown to reprogramme the epigenome to prevent differentiation biases (Kim et al., 2011; David and Polo, 2014).

Furthermore, regulators of DNA methylation such as activation induced cytidine deaminase (AID), TET family and DMNTs are reactivated during the late maturation/stabilisation phase (Polo et al., 2012). Among those regulators, AID has been reported to be actively enhancing the epigenetic reset (Kumar et al., 2013; David and Polo, 2014).

2.3.4 Pluripotent Stem Cells in Adipocyte Differentiation

Directed differentiation of pluripotent stem cells towards the adipocyte lineage provides a highly informative system to characterise the earliest steps underlying adipocyte development. The adipocyte lineage derived from direct differentiation of mouse ESCs was first demonstrated in 1997 (Dani et al., 1997). In that study, the mESC-derived adipocytes were found to express key adipocyte-specific genes, such as the C/EBP and the PPAR families, known to promote preadipocyte differentiation. Human ESCs and iPSCs have been shown to successfully differentiate adipocytes based on protocol obtained from mouse ESC studies (Xiong et al., 2005; Taura et al., 2009; Noguchi et al., 2013). Human iPSC-derived adipocytes possess a similar adipogenic potential comparable to those derived from ESCs (Taura et al., 2009; Noguchi et al., 2013). Notably, these studies showed that human iPSC-derived adipocytes could retain their adipogenic properties for six weeks post-transplantation into nude mice. Nevertheless, it was also noted that teratomas were formed several weeks after transplantation most likely due to the presence of other cell types such as immature neural cells and undifferentiated

iPSCs, which were also transplanted, besides the enrichment of adipocytes derived from the iPSCs (Taura et al., 2009; Noguchi et al., 2013). Furthermore, it was observed that the graft loss caused by transplantation of only mature adipocytes could be enhanced by introducing adipocyte progenitors (Noguchi et al., 2013). Therefore, a high adipogenic capacity obtained from purification of human iPSC-derived MSCs is necessary for an iPSC-based therapy.

2.3.5 Challenges in Reprogramming Cells to Pluripotency

As discussed above, iPSCs have become the promising tools used not only in basic research and disease modelling but also in a wide range of clinical applications in cell-based therapies and regenerative medicine. However, the induced reprogramming efficiency is currently low (~ 0.01 – 0.02%) (Takahashi and Yamanaka, 2006; Takahashi et al., 2007; Aasen et al., 2008). Furthermore, the delivery systems normally employ the integrative retroviral or lentiviral delivery systems (Sugii et al., 2011; Bazley et al., 2015), thereby raising the concerns in safety issues. Numerous nonintegrative methods including inducible lentivirus (Hockemeyer et al., 2008), sendai virus (Fusaki et al., 2009) and adenovirus (Stadtfeld, Nagaya, et al., 2008) have been developed to address the aforementioned safety concerns. Nevertheless, these techniques still do not comply with the safety standard requirements for clinical applications. In addition, iPSCs are also facing the challenge of retention of epigenetic memory of the parental cells. Although ESC and iPSC

are functionally similar, epigenetic characteristics between ESC and iPSC are notably different (Cahan and Daley, 2013). Inefficient silencing of genes of somatic cells to sufficiently erase the epigenetic memory, and the gaining of new methylation patterns, which results in epigenetic mutations, have been identified as the reasons for the reported differences in over a thousand methylated regions between ESC and iPSC (Lister et al., 2011). The results of gene expression profiling have further shown persistent expression of genes of the donor cells in iPSC, supporting retained epigenetic memory in iPSC (Marchetto et al., 2009).

A major difficulty in the practical applications of iPSCs is the high possibility of obtaining partially reprogrammed cells, an event associated with the continuous expression of exogenous reprogramming factors OSKM. These partially reprogrammed cells display some ESC-like characteristics including identical morphologies, high proliferation rate, activation of pluripotency-related genes, teratoma formation and chimeric contribution (Takahashi and Yamanaka, 2016). However, these characteristics in ESCs are more distinctive than those of partially reprogrammed cells. Additionally, partially reprogrammed cells still retain a greater degree of epigenetic memory of the somatic cell characteristics (Marchetto et al., 2009). Hence, partially reprogrammed cells can only undergo differentiation with flaws, and with gene expression profiles which could be highly dissimilar to those of ESC (Takahashi and Yamanaka, 2006). Improved reprogramming systems have been developed to select fully reprogrammed cells. In one reported system, expression of exogenous pluripotency genes is subjected to induction by

doxycycline (Hockemeyer et al., 2008). In the absence of doxycycline, only fully reprogrammed cells survive as these cells have already activated endogenous pluripotency genes, whereas partially reprogrammed cells that have failed to activate the endogenous genes do not survive. In summary, more studies are required to further develop safe and efficient reprogramming protocols for clinical translation of iPSCs.

2.4 MicroRNAs (miRNAs)

MicroRNAs (miRNAs) are short (~22 nucleotides) single-stranded RNAs (Bartel, 2009). Mammalian miRNAs function as guide molecules in gene silencing by either mRNA degradation or translational inhibition through imperfect pairing with the 3' untranslated region (3'UTR) of their target genes (Bartel, 2009). Since the discovery of the first miRNA (Lee et al., 1993), intense research has focused on the biological functions and mechanisms of action of miRNAs. The evolutionary conservation of miRNA has made miRNA one of most important regulators in various cellular and biological processes such as development, metabolism and homeostasis (Bartel, 2009). Due to the short length, each miRNA is able to target hundreds of different mRNA transcripts, and more than 50% of all human protein-coding transcripts may be regulated by collective expression of approximately 1,000 to 1,500 various miRNAs mapped on the human genome (Friedman et al., 2009). Aberrant miRNA expression is associated with many diseases such as cancer, neurodevelopmental disorders and chronic obstructive pulmonary diseases (Chang and Mendell, 2007; Li et al., 2009; De Smet et al., 2015). Furthermore,

distinct miRNA milieus are associated with different states of pluripotency (Clancy et al., 2014). Hence, it is important to investigate the transcriptional mechanisms as well as biological functions of miRNAs in cellular processes.

2.4.1 Biogenesis and Function of MicroRNAs

The biogenesis of miRNA is tightly mediated in a multistep process that begins at miRNA transcription in the nucleus of the cells and continues to processing to produce single-stranded mature miRNA in the cytoplasm where mature miRNA functions as the post-transcriptional regulator for RNA silencing.

The genomic locations where MiRNA sequences situated are used to define the different types of miRNAs. MiRNAs are classified as intergenic miRNAs when the miRNAs are located between annotated genes or regions of the human genome, and are transcribed via independent promoters (Lee et al., 2004). In contrast, intronic miRNAs are situated in the introns of protein-coding genes and, thus, are regulated and processed from the respective host promoters (Lee et al., 2004; Kim and Kim, 2007). MiRNAs are frequently located in clusters where the miRNA genes are physically close to each other (Z. Zhang et al., 2015; Rippe et al., 2012). Generally, miRNAs within a cluster are co-transcribed from a single master promoter, and the individual miRNAs are subsequently regulated by post-transcriptional processing (Rottiers and Näär, 2012). MiRNA transcription is performed by RNA polymerase II (Lee

et al., 2004) and is mediated by epigenetic controls such as DNA methylation and histone modification as well as by transcription factors such as p53 (Fornari et al., 2012; Rottiers and Näär, 2012). The primary transcripts (pri-miRNAs) are first formed after DNA transcription by RNA polymerase II. In the canonical miRNA biogenesis pathway, long pri-miRNAs are initially processed into a structure of 60- to 110-nucleotide (nt) hairpin precursor miRNA (pre-miRNA) by cellular RNase enzyme III, Drosha and the double-stranded RNA-binding domain protein, DGCR8 (Bushati and Cohen, 2007). In the non-canonical miRNA pathway, the pre-miRNAs are generated without involving Drosha by forming endogenous short hairpin RNAs (endo-shRNAs), or miRNAs are directly formed through splicing and then refolding into short-hairpin introns called mirtrons (Rottiers and Näär, 2012).

The pre-miRNAs generated from both the canonical and non-canonical pathways are transported to the cytoplasm by exportin-5 (Exp5) and a RAN-GTP-dependent process. In the cytoplasm, pre-miRNAs are then cleaved by another RNase III enzyme Dicer and the transactivation-response RNA-binding protein (TRBP) to generate a mature miRNA: miRNA* duplex (Lee et al., 2002). The functional guide strand of the duplex, which is given the nomenclature of miRNA, is incorporated into the Argonaute-containing RNA-induced silencing complex (RISC). The RISC-miRNA assembly is then guided to the target mRNA based on Watson–Crick base-pairing between the seed sequence in the mature miRNA and the 3’-UTR of the targeted transcripts (Maniataki and Mourelatos, 2005). If the base-pairing is perfectly complimentary, the RISC-bound mRNAs are degraded, whereas if the base-

paring is only partially complimentary, the miRNA suppresses mRNA translation. In contrast, the other strand, designated as miRNA* or called a passenger strand, is usually degraded.

In some cases, both the guiding and passenger strands of a pre-miRNAs survive the processing, and the two mature miRNAs excised from the 5'- and 3'- arms are the expression products, designated as the -5p and -3p species, respectively, of the miRNA gene (Choo, Soon, et al., 2014; Huang et al., 2014). Both the -5p and -3p miRNA species have been reported to be functional and target different mRNAs due to different sequences (Griffiths-Jones et al., 2006; Choo, Soon, et al., 2014). To avoid confusion, the human miRNA/miRNA* nomenclature has been retired by miRBase. Instead, the miRNA-5p and -3p nomenclature is now being applied widely according to 5'- or 3'-arms derivation of the miRNA species.

2.4.2 MiRNA-mediated Reprogramming of Somatic Cells to Pluripotency

A number of chromosomally clustered miRNAs has been found to be specifically up-regulated in pluripotent stem cells compared to mature differentiated cell types (Stadler et al., 2010; Leonardo et al., 2012). The better-characterised miRNA clusters include the miR-302/367, miR-17/92, miR-200 clusters, and the chromosome 19 miRNA cluster, abbreviated as C19MC. The miR-302/367 cluster comprises mature miR-302a-5p, 302a-3p,

302b-5p, 302b-3p, 302c-5p, 302c-3p, 302d, and -367 (Barroso-delJesus et al., 2008). Although miR-367 has slightly different seed sequence from miR-302a-d, they share a set of common mRNA targets. It has been reported that the miR-302/367 cluster alone is able to generate iPSCs (Anokye-Danso et al., 2011; Miyoshi et al., 2011); the reprogramming of the human foreskin fibroblasts to iPSCs was completely blocked when the miR-302/367 cluster was knocked out (Zhang et al., 2013). During the reprogramming process, the miR-302 family, which includes miR-302a through to miR-302d, was found to not only mediate the expression of the crucial pluripotency genes (Sandmaier and Telugu, 2015), but was also shown to promote proliferation and to accelerate G1 to S transition of the cell cycle. Cell cycle regulation was achieved by targeting the retinoblastoma (RB) family and CDK1NA/p21 to overcome the G1/S barrier of somatic cells and to adopt the pluripotent- like abbreviated G1 phase (Y. Wang et al., 2013). Moreover, the miR-302 miRNAs modify the epigenetic landscape of the reprogrammed cells by targeting the epigenetic regulators lysine demethylase 1B (AOF1) and lysine demethylase 1A (AOF2) to regulate DNA demethylation through down-regulation of DNA methyltransferase 1 (DNMT1) (Lin et al., 2011). Recently, the miR-302 family was demonstrated to facilitate MET process, an important event of the initiation phase of reprogramming, by blocking the expression of the transforming growth factor β receptor II (TGFβR2) (Subramanyam et al., 2011).

On the other hand, members of the miR-302/367, miR-17/92 or miR-200 clusters, when co-expressed with the OSKM factors, are able to enhance

reprogramming efficiency (Hu et al., 2013; He et al., 2014; G. Wang et al., 2013). The miR-302 was reported to promote reprogramming efficiency via nuclear receptor subfamily 2 group F member 2 (NR2F2) repression whereas the miR-17/92 cluster mediates reprogramming by down-regulating phosphatase and tensin homolog (PTEN), a renowned tumour suppressor (Hu et al., 2013; He et al., 2014). Besides miR-302, the miR-200 cluster also plays crucial roles in the early stage of somatic cell reprogramming (G. Wang et al., 2013). Transcription of the miR-200 cluster is regulated by the promoter binding of OCT4 or SOX2 to mediate the initial phase of reprogramming by increasing the kinetics of MET through blocking the EMT-related genes ZEB1/ZEB2 (G. Wang et al., 2013).

p53 has been demonstrated as a key barrier that causes the low efficiency of iPSC generation: p53 deficiency has been shown to facilitate cellular reprogramming (Marion et al., 2009). Some miRNAs are mediators or barriers to iPSC reprogramming by involving in the regulation of p53 signalling pathway. MiR-138 acts as modulator to iPSC generation through suppression of p53 expression by directly targeting the 3′ UTR of p53 (Ye et al., 2012), while miR-34, which is transcriptionally activated by p53, in turn contributes to p53 repression at least in part through down-regulating the pluripotency genes NANOG, SOX2 and c-MYC, and hence repression of iPCS generation (Choi et al., 2011).

2.4.3 Roles of MiRNAs in Self-Renewal and Pluripotency

Besides the significant regulatory role of miRNAs in initiating reprogramming, miRNAs further play an important role in maintaining pluripotency, self-renewal and differentiation of ESCs and iPSCs. In this respect, miRNAs may be divided into two subgroups based on their functions, *viz.* the pluripotent and the pro-differentiation miRNAs. Pluripotent ESC-specific miRNAs, including miR-137, miR-184, miR-200, miR-290, miR-302, miR-9 and C19MC, function in the regulation of self-renewal and pluripotency. Previous studies have revealed that knockout of the miRNA processing enzymes in Dicer- and Dgcr8-deficient mouse ESCs markedly delayed cell cycle progression and differentiation (Wang et al., 2008; Y. Wang et al., 2013). However, the defective ES cell cycle in the Dicer- and -DGCR8 null mice could be rescued by over-expressing with ESC-specific miRNA-290 (Wang et al., 2008). Moreover, teratoma formation and chimerism, two pluripotency characteristics of fully-reprogrammed pluripotent stem cells, were not observed in the same group of knockout mice (Kanellopoulou et al., 2005; Mathieu and Ruohola-Baker, 2013). The results support the importance of ESC-specific miRNAs in regulating the pluripotency state. Indeed, ESC-specific miRNAs preserve the identity of pluripotent cells by balancing the expression levels of core ESC-specific transcription factors OCT4, SOX2 and NANOG (Anokye-Danso et al., 2012). A positive feed-back loop between the miR-302/367 cluster and OCT4, SOX2 and NANOG has, indeed, been identified (Barroso-delJesus et al., 2008; Lin et al., 2011; Anokye-Danso et al., 2012). MiR-302/367 is activated by the binding of pluripotency transcription factors to the miR-302/367 promoter (Barroso-delJesus et al., 2008). Upon

activation, miR-302/367 indirectly suppresses OCT4, SOX2 and NANOG expression by targeting the developmental genes RB transcriptional corepressor like 2 (RBL2), CDKN1A/p21 and many more other genes (Lin et al., 2011; Anokye-Danso et al., 2012).

In contrast, the pro-differentiation miRNAs, such as let-7, miR-296, miR-134 and miR-470, have been found to regulate the differentiation processes in pluripotent stem cells (Melton et al., 2010; Lüningschrör et al., 2013). Let-7 has been shown to facilitate the cell-fate decision between self-renewal and differentiation of stem cells. Let-7 was reportedly silenced in ESCs but was highly expressed in somatic cells (Lakshmipathy et al., 2007). The pluripotency factor LIN28 blocks the maturation of let-7 by binding to the pre-let-7 RNA resulting in inhibition of further miRNA processing by the Dicer ribonuclease (Rybak et al., 2008). During differentiation, LIN28 expression is down-regulated as expression of OCT4, SOX2 and NANOG is repressed. In the absence of LIN28, let-7 expression is up-regulated resulting in inhibition of genes engaged in promoting self-renewal (Lüningschrör et al., 2013).

2.5 Chromosome 19 MiRNA Cluster (C19MC)

2.5.1 Overview of C19MC

C19MC is a primate-specific miRNA cluster which was developed in a relatively short span of time during evolution. C19MC is the largest miRNA gene cluster in the human genome and accounts for approximately 8% of all known human miRNA genes. C19MC contains 46 highly-related miRNAs with 7 duplicates within a ~100 kb region of the genomic DNA (Bortolin-Cavaille et al., 2009). Notably, the C19MC cluster is interspersed among Alu repeats, which could have facilitated the evolutionary expansion of C19MC (Zhang et al., 2008) (Figure 2.4). C19MC expression is undetected in adult tissues (Landgraf et al., 2007; Liang et al., 2007) while high C19MC expression levels are found in the reproductive placenta system (Noguer-dance et al., 2010) and in human ESCs (Ren et al., 2009).

2.5.2 Transcriptional Mechanisms of C19MC in The Placenta

The transcriptional machinery that drives the expression of the C19MC miRNAs remains unclear. C19MC was previously shown to be transcribed by RNA polmerase III using the Alu repeats as promoters, which were verified by chromatin immunopreciptation (ChIP) and cell-free transcription assays (Borchert et al., 2006). However, in the placenta, it was later reported that the C19MC miRNAs are intron-encoded and are processed by the DGCR8-

Drosha complex, also known as the microprocessor, from the non-protein-coding RNA polymerase II transcript (Bortolin-Cavaille et al., 2009). It was shown in this study that, in the placenta, miRNAs in the C19MC cluster were first transcribed as a long primary transcript on the activation of the primary master promoter located at the 5'-end of the miRNA cluster. Subsequently, the primary transcript was proposed to be processed by splicing to generate the full complement of the C19MC mature miRNAs. Furthermore, the expression of C19MC was proposed to be driven by methylation at the upstream CpG-rich promoter region located 17.6 kb upstream of the first miRNA gene (Noguer-dance et al., 2010) as C19MC miRNA expression was activated upon treatment with DNA methylation inhibitors (Tsai et al., 2009; Flor et al., 2012). Of note, the transcriptional activity of C19MC is further regulated by genomic imprinting in which only the paternally-inherited allele is expressed in the placenta while the maternal allele is silenced by methylation. Yet C19MC is located adjacent to another imprinted ZFN331 gene (Rippe et al., 1999), which is mediated by maternal expression (Daelemans et al., 2010). Hence, C19MC may define a previously unrecognized large imprinted primate-specific chromosomal domain.

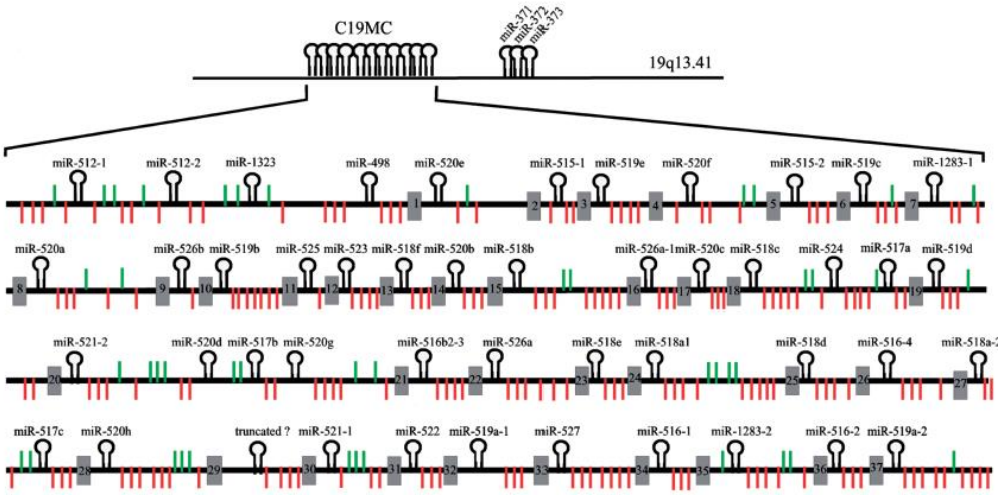


Figure 2.4 miRNA gene organisation at C19MC. Symbol of stem-loop structures are pre-miRNA genes. Grey boxes are repeated exons of C19MC. Repeated Alus are indicated by green and red bars which are corresponding to the sense and antisense orientations comparable to the pre-miRNA genes, respectively. [Figure was modified from (Bortolin-Cavaille et al., 2009).]

2.5.3 C19MC in Pregnancy

Methylation analysis has shown that both alleles of C19MC in normal adult tissues are methylated whereas the paternal allele of the placenta could escape epigenetic silencing to maintain active C19MC expression until birth before being shut down by hypermethylation (Tsai et al., 2009; Noguer-dance et al., 2010). The findings suggested crucial roles of C19MC in placenta-specific functions. The expression of twenty-one out of the forty-six C19MC pre-miRNAs was detected in the human placenta of a 5-week embryo (Zhang et al., 2008). Among the expressed pre-miRNAs, the expression of miR-498, miR-516-5p and miR-520e in normal-term placenta was preferentially detected in the cytoplasm of the syncytiotrophoblast. Similar results were also obtained for miR-517b, which was markedly increased in the syncytiotrophoblast situated in the trophoblast layer (Luo et al., 2009). Recently studies have shown that there is a relationship between changes in the expression levels of circulating miRNAs during pregnancy and in obstetrical and placental diseases (Hromadnikova et al., 2013; Miura et al., 2015). Circulating C19MC miRNAs, more notably miR-517-5p, miR-518b and miR-520h, are up-regulated in the early pregnancy of women at risk of preeclampsia, a disorder of pregnancy characterised by high blood pressure and signs of organ damage, in particular the liver and the kidneys (Hromadnikova et al., 2017). Interestingly, the use of serum level of miR-517-5p as a biomarker in the first trimester has led to the identification of a significant proportion of women with subsequent preeclampsia, suggesting

that the C19MC miR-517-5p is a potential predictive marker for preeclampsia (Hromadnikova et al., 2017).

Besides potential clinical applications as predictive markers in obstetrical and placental diseases, C19MC could also facilitate non-trophoblast cells to confer significant resistance to viral infection (Delorme-Axford et al., 2013). Members of C19MC are packaged within trophoblast-derived exosomes, which are subsequently delivered to recipient cells to induce viral resistance by the induction of autophagy (Delorme-Axford et al., 2013).

2.5.4 C19MC in Tumorigenesis

The hallmarks of human cancers are defined by activation of tumour-related developmental pathways that enable cells to maintain cell proliferation, escape growth suppressors, prevent cell death, allow replicative immortality, activate invasion and metastasis and to induce angiogenesis (Hanahan and Weinberg, 2011). It has been shown that dysregulated expression of C19MC miRNAs in cancers could disturb the cancer hallmarks for pathogenesis and progression of tumours (Rippe et al., 2012; Vaira et al., 2012; Kleinman et al., 2014). C19MC miRNAs could function as either oncogenes or tumour suppressors based on their target genes under specific circumstances.

2.5.4.1 C19MC as OncomiRs

In primitive neuroectodermal tumour, the activation of the C19MC locus is restricted to a sub-group of aggressive cells called CNS-PNETs with distinctly aggressive clinic-pathologic features (Li et al., 2009). The constitutive expression of miR-520g and -517c, two of the most highly expressed C19MC miRNAs in CNS-PNETs, inhibits differentiation of human neural stem cells (hNSCs), modulates cell survival, promotes *in vitro* cell proliferation and *in vivo* tumour xenograft growth (Li et al., 2009). The findings suggest that miR-520g and -517c are oncogenic miRNAs, or oncomirs, which can be used as potential biological markers for CNS-PNET. Moreover, the study on embryonal tumours with multilayered rosettes (ETMRs) of primitive neuroectodermal tumours has further demonstrated that the translocation-mediated fusion of the promoter of the highly active brain-specific tweety family member 1 (TTYH1) to C19MC results in high expression levels of the C19MC miRNAs in neural tissue (Kleinman et al., 2014). Following the amplification of the TTYH1-C19MC fusion gene, the activated C19MC miR-519a targets RBL2, leading to the suppression of DNA methyltransferase 3 beta (DNMT3B), thereby altering the cell cycle and the global epigenomic landscape to modulate tumourigenesis (Archer and Pomeroy, 2014). Likewise, the up-regulation of miR-517c, miR-519a, miR-521, miR-520c and miR-522 is correlated with aggressive breast cancer cells (Huang et al., 2008; Tan et al., 2014). Over-expression of miR-520c enhances breast cancer invasion and metastasis by directly targeting CD44 (Huang et al., 2008).

Controlling cell proliferation is an important feature in normal cells as abnormal cell-cycle progression, such as evasion of cell cycle-inhibited genes, leads to tumorigenesis. Once activated by p53 or the DNA methylation inhibitor 5-aza-2-deoxycytidine, miR-519d has been shown to inhibit apoptosis, enhance cell proliferation and invasion hepatocellular carcinoma cancer cells through targeting cell cycle-inhibited genes including CDKN1A, PTEN, AKT serine/threonine kinase 3 (AKT3) and TIMP metallopeptidase inhibitor 2 (TIMP2) (Fornari et al., 2012).

2.5.4.2 C19MC as Tumour Suppressor MiRNAs

In contrast to the oncogenic C19MC miRNAs described above, down-regulation of tumour suppressor C19MC miRNAs has been reported for the development of some tumours. Suppressed expression of tumour suppressive C19MC miRNAs is found in estrogen receptor negative (ER-) breast cancer, hepatocellular carcinoma, ovarian cancer and non-small cell lung cancer (NSCLC) (Keklikoglou et al., 2012; W. Zhang et al., 2012; Cong et al., 2015; Zhu et al., 2015). The down-regulated tumour suppressive C19MC miRNAs in cancers are normally responsible for the evasion of growth suppressors and maintenance of proliferative signalling in cancer cells.

E2F2 belongs to a subclass of E2F factors that act as transcriptional activators of cell cycle progression through the G1 to S transition (Wu et al.,

2001). Inhibition of E2F2 completely stops the capacity of mouse embryonic fibroblasts to enter the S phase of the cell cycle, abolishing proliferation. It is reported that tumour suppressor miR-520a targets E2F2 (Huang et al., 2016). In addition, miR-520a is negatively regulated by ANCCA/PRO2000, an important proliferation-associated protein (Huang et al., 2016), and miR-520a expression is reduced in human hepatocellular carcinoma (HCC) cells compared to normal hepatic cells (Dong et al., 2015). Hence, ANCCA/PRO2000 acts as a potent oncogene to enhance *in vitro* and *in vivo* growth of HCC cells by regulating E2F2 expression through suppressing miR-520a. Over-expression of miR-520b suppresses the *in vivo* growth of HCC cells through silencing cyclin D1 that is required for cell cycle progression in the G1 phase (Du et al., 2013).

C19MC miRNAs regulate cell proliferation not only by mediating cell cycle-associated components but also through modulating multiple signalling pathways. Nuclear factor κB (NF-κB) transcription factor regulates genes related to tumour growth, invasion and metastasis (Helbig et al., 2003; Park et al., 2007), while transforming growth factor-β (TGFβ) signalling is involved in a wide range of cellular processes, including cell proliferation, cell cycle, differentiation, EMT and metastasis (Yingling et al., 2004; Padua and Massague, 2009). Therefore, the modification of the NF-κB and TGFβ signalling pathways is central to the development of breast cancer (Park et al., 2007; Padua and Massague, 2009). MiRNA-520c acts as a tumour suppressor in ER- breast cancer through direct targeting RelA and TGFβR2, respective regulators of the NF-κB and TGFβ signalling pathways, to inhibit tumour

progression, metastasis and inflammation (Keklikoglou et al., 2012). Remarkably, the down-regulated miR-520c expression was also found to be associated with lymph node metastasis specifically in ER- breast cancer tumours (Keklikoglou et al., 2012). Hence, with the expression of miR-520c could interfere with the pathogenesis of ER- breast cancer.

2.5.5 C19MC in Stem Cells

The expression of C19MC is activated or up-regulated in pluripotent stem cells compared to mature differentiated cell types (Ren et al., 2009; Razak et al., 2013), suggesting an important role of C19MC in maintaining the properties of stem cells. In disease state, C19MC may be chromosomally amplified. The amplification of C19MC cluster in a subgroup of primitive neuroectodermal tumour (PNETs) is associated with high LIN28 expression (Spence, Perotti, et al., 2014). Notably, LIN28 has been shown to play important roles in somatic reprogramming, tissue development, organismal growth and metabolism, and in cancer (Shyh-Chang and Daley, 2013). Furthermore, miR-517c and -520g expressed in human neural stem cells (hNSCs) regulate and facilitate the non-canonical WNT/JNK signalling pathway that is associated with stem cell maintenance (Lien and Fuchs, 2014). More direct evidence for a role of C19MC in stem cell biology comes from the study of miR-520d-5p, a member of C19MC (Ishihara et al., 2014; Tsuno et al., 2014). Over-expression of miR-520d-5p was able to convert fibroblasts to form MSCs, which stained positively with CD105 as verified in both *in vitro*

experiments and in a *in vivo* xenografted model (Ishihara et al., 2014). Furthermore, introduction of miR-520d-5p caused the loss of malignant properties in hepatoma cells *in vivo* through the conversion of cancer cells to iPSC-like cells (Tsuno et al., 2014). Nevertheless, despites extensive studies of the involvement of C19MC in cancers, researches that focus on studying the role and function of C19MC in stem cells are currently still lacking. Hence, a major goal of the present thesis work was to elucidate the involvement of C19MC in stem cells especially the biological functions of C19MC in reprogramming.

CHAPTER 3

MATERIALS AND METHODS

3.1 Cell Culture and Maintenance of Cell Lines

Basal media including Dulbecco's Modified Eagle Medium (DMEM) high glucose, DMEM/F12, Roswell Park Memorial Institute (RPMI) 1640 and Minimum Essential Medium (MEM) (Gibco, California, USA) were prepared according to the manufacturer's instruction. To prepare 1 L basal medium, the powder of basal medium was dissolved in 900 mL nuclease-free ddH_2O. Subsequently, sodium bicarbonate ($NaHCO_3$) (EMD Millipore, Temecula, CA, USA) was added to the dissolved medium solution according to the recommended amount (per litre solution): DMEM high glucose, 3.7 g; DMEM/F12, 1.2 g; RPMI, 2 g and MEM, 2.2 g. The pH was adjusted to pH 7.2 – pH 7.4 prior to topping up with ddH_2O to one litre. The medium was processed immediately into sterile containers by membrane filtration with a 0.2-μm cellulose acetate filter (Techno Plastic Product, Trasadingen, Switzerland) using a vacuum pump system (GAST, MI, USA).

All non-pluripotent stem cell lines were cultured in the appropriate culture media (Table 3.1 and 3.2) and maintained in a 37 ^{o}C cell culture incubator with 5% CO_2. Culture medium was changed every alternative day. The cells were dissociated when the cells reached 70-80% confluency. Culture medium was discarded and the cells were washed once with 1X PBS at pH 7.4

Table 3.1 Formulation of media used in culture

Cell line	Formulation of media[1]
	Normal somatic cells
293FT	- DMEM high glucose - 10% fetal bovine serum (FBS) - 0.1 mM MEM non-essential amino acids (NEAA) - 6 mM L-glutamine - 1 mM MEM sodium pyruvate - 500 μg/ml geneticin
Fibroblast (HEF-1& MEF)	- DMEM high glucose (Gibco) - 15% FBS (Gibco) - 1% glutamax (Gibco)
Placenta (HS799.PI)	- Eagle's Minimum Essential Medium (EMEM)[2] - 10% FBS (Gibco)
Normal colon (CRL-1790)	- MEM (Gibco) - 10% FBS (Gibco)
	Stem cells
MSC (WJ0706)	- DMEM/F12 (Gibco) - 10% FBS (Gibco)
ESC/iPSC (HuES6, H9 & ESC-like colonies)	- DMEM/F12 (Gibco) - 20% knockout serum replacement (KOSR) - 1% glutamax - 1x NEAA (Gibco) - 0.1% β-mecaptoenthanol[3] - 10 ng/ml bFGF[4]
	Cancer cells
Placenta (JEG-3)	- Eagle's Minimum Essential Medium (EMEM) - 10% FBS
Colorectal/liver cancer (HCT-15, SKCO1 & HepG2)	- DMEM high glucose - 10% FBS
Breast cancer (MCF-7)	- RPMI 1640 - 10% FBS

[1]Supplier was Gibco or stated otherwise. [2]Supplier was ATCC. [3]Supplier was Calbiochem**.** [4]Supplier was ReproCELL

Table 3.2 Sources and maintenance of cell lines

Cell type	Cell line (designation)	Source	Culture medium
Normal somatic cells	293FT	Invitrogen	293 FT
	Foreskin fibroblast (HFF-1)	ATCC	Fibroblast
	Human normal placenta (HS799.PI)	ATCC	Placenta
	Human normal colon (ATCC CRL-1790)	ATCC	Normal colon
	Mouse embryonic fibroblast (MEF)	EMD Millipore	Fibroblast
Unipotent cells	Human white pre-adipocyte (HWP)	PromoCell	RNA provided by Dr. S.Sugii[1]
Multipotent MSC	Human adipose-derived MSCs: MSC-AT[1] ASC-Inv[1] ASC Lonza[1]	 PromoCell Invitrogen Lonza	 RNA provided by Dr. S.Sugii[1]
	Human umbilical cord Wharton's Jelly-derived MSC (WJ0706)	Cytopeutics Sdn. Bhd[2]	MSC
Pluripotent stem cells (PSCs)	Human embryonic stem cells (ESC) (HuES6 & H9)	A gift of Dr. S. Sugii[1]	ESC/iPSC
	Human induced PSCs: HWP-derived iPSC (HWP-iPSC) MSC-AT derived iPSC (MSC-iPSC) ASC-Inv-derived iPSC (ASC-iPSC) ASC Lonza derived iPSC (MH#1)	A gift of Dr. S. Sugii[1]	ESC/iPSC
ESC-like colonies	ESC-like colonies derived from HFF-1	Generated in this study	ESC/iPSC 2 µg/ml doxycycline
Human cancer cells	Placenta choriocarcinoma (JEG-3)	ATCC	Placenta
	Colorectal cancer (HCT-15 & SK-CO-1)	ATCC	Colorectal/ liver cancer
	Liver cancer (HepG2)	A Gift of Prof. Y.M. Lim[3]	Colorectal/ liver cancer
	Breast cancer (MCF-7)	A Gift of Prof. Y.M. Lim[3]	Breast cancer

[1]Gifts of Dr. S Sugii, Singapore BioImaging Consortium, A*STAR, Singapore; [2]Cytopeutics Sdn. Bhd, Selangor, Malaysia (http://www.cytopeutics.com); [3]Prof. Y.M. Lim, Cancer Research Center, Universiti Tunku Abdul Rahman, Selangor, Malaysia.

(Amresco, Ohio, USA). To detach the cells, appropriate amount of 0.25% Trypsin-EDTA solution (Gibco) was added to the cells and then incubated at 37 ^{o}C for 5 min. Complete culture medium (4 mL) was added to stop trypsin activity. The completely detached cells were gently transferred into a centrifuge tube prior to centrifugation at 1,500 rpm for 5 min. Finally, the supernatant was discarded and the pellet was re-suspended in complete culture medium. The cells were maintained in the incubator at 37 ^{o}C with 5% CO_2.

Pluripotent stem cells and ESC-like colonies derived from HFF-1 cell line which was established and deposited in ATCC in 2003 from normal human foreskin pooled from two individuals were cultured according to Table 3.1 and 3.2. The medium was replaced every day to suppress differentiation. Mitomycin C-treated mouse embryonic fibroblast (MEF) feeder layer was seeded one day before plating of the stem cell colonies. Stem cell colonies were cut into grids by using a blade (Swann-Morton, Sheffield, UK), and the colony grids were transferred to a mitomycin C-treated MEF feeder layer pre-seeded 24 h before use. The colonies were maintained in the incubator at 37 ^{o}C with 5% CO_2.

3.2 Cell Revival from Liquid Nitrogen Frozen Stock

Vials of frozen non-pluripotent stem cell lines, including normal somatic, unipotent, multipotent MSC and cancer cell lines that are listed in Table 3.1 and Table 3.2, were retrieved by placing in a 37-degree water bath

until cells were partially thawed. The thawed cells were transferred into a centrifuge tube containing 5 ml pre-warmed complete growth medium and centrifuged at 1,500 rpm for 5 min in a benchtop centrifuge (Allegra® X-15R, Beckman Coulter, CA, USA). The supernatant was aseptically decanted. The pellet was gently re-suspended in 1 mL complete culture medium. The cell suspension was then transferred into appropriate culture vessels and was maintained in the incubator at 37 ^{o}C with 5% CO_2.

3.3 Cryopreservation of Cell Lines

Following cell detachment as described in Section 3.1 above, the cell suspension was obtained by suspending cell pellets in a mixture containing 900 µl FBS and 100 µl dimethyl sulfoxide (DMSO) (Sigma-Aldrich, MO, USA). The cell suspension was aliquoted into sterile cryotubes (Corning, NY, USA) and frozen at 20 ^{o}C for 1 h at a rate of -1 °C per min using CoolCell alcohol-free cell freezing containers (BioCision, LLC, Mill Valley, CA, USA) before incubating at -80 ^{o}C for 24 h. For long-term storage, the cryotubes were kept into a liquid nitrogen container (Chart Industries, Ohio, USA) at approximately -180 ^{o}C.

3.4 Inactivated Mouse Embryonic Fibroblast Feeder Cells (MEFs)

Untreated MEFs (EMD Millipore) at passage 3 were thawed and expanded. When untreated cells reached 90 to 100% confluency, cells were

incubated with 10 µg/mL mitomycin C (EMD Millipore) at 37 °C, 5% CO_2 for 3 h to arrest mitotic growth. All traces of mitomycin C were removed by washing three times with 1X PBS (Amresco) before the inactivated MEF cells were trypsinised for cryopreservation at a desired cell concentration for future use.

3.5 Production of Lentiviral Vectors in 293FT Cells

Ten-cm tissue culture dishes were coated with 0.2% gelatin for at least 30 min before seeding of 293FT cells (Invitrogen, Carlsband, CA, USA). The day before transfection, 293FT cells were plated at a density of 6.5 x 10^6 per dish in 10 ml complete medium as described in Tables 3.1 and 3.2 without geneticin and penstrep. On the day of transfection, the cells were at 90-95% confluency. Twenty-four hours post-seeding, the medium was aspirated and 7 ml OptiMEM medium (Gibco), which contained 4 ml blank OptiMEM and 3 ml 5% FBS in OptiMEM medium, was replaced. Various lentiviral constructs listed in Table 3.3 were transfected into 293FT cells by using the Lipofectamine 2000 transfection reagent (Invitrogen). For each 10-cm dish, 10 µg lentivirus plasmid DNA, 2.5 µg pMD2.G and 7.5 µg psPAX2 were diluted in 1.5 ml OptiMEM medium without serum. The plasmid DNA mixture was then transferred to a sterile tube A. In a sterile tube B, 48 µL Lipofectamine 2000 (Invitrogen) was diluted in 1.5 ml OptiMEM medium without serum. The contents of tube A and tube B were mixed gently and the mixture was incubated at room temperature for 5 min, after which the diluted DNA in tube

A was added dropwise to the diluted Lipofectamine 2000 in tube B. The mixture was gently mixed and incubated at room temperature for 20 min. After incubation, the DNA-Lipofectamine 2000 complexes were added dropwise to the plated 293FT cells, and incubated at 37 °C, 5% CO_2 in a humidified incubator. After 24 h of incubation, the transfection reagent-containing medium was discarded and 10 ml 293 FT complete culture medium as described in Tables 3.1 and 3.2 without geneticin and penstrep was replaced. Viral supernatant collected at 48 and 72 h post-transfection was filtered through a 0.45-µm pore size PVDF filter (EMD Millipore). The filtered virus-containing supernatant was transferred into a sterile ultracentrifuge tube (Cat # 357002, Beckman Coulter) and concentrated by centrifugation at 25,000 rpm for 2 h at 4 ^{o}C using a JA-25.50 rotor (Beckman Coulter). The supernatant was removed to obtain the pellets which were then allowed to slowly dissolve into 200 µl sterile blank DMEM overnight at 4 ^{o}C in an up-right position. The next day, pellets were gently re-suspended and kept at -80 ^{o}C.

3.6 Reprogramming of HFF-1 to iPSCs

The HFF-1 cells were seeded in six-well plates 24 h prior to transduction. Individual lentiviruses carrying human m2rtTA, OCT4, SOX2, KLF4 or MYC at equal ratios supplemented with 8 µg/ml polybrene (EMD Millipore) were transduced together into HFF-1 cells with or without the lentivirus carrying the miR-524 precursor (System Biosciences,

Table 3.3 Lentiviral constructs used in this study

Vector	Description	Vector size (bp)	Cat. No.
FUW-m2rtTA	Expressing the reverse tetracycline transactivator	7,979	20342[1]
FUW-tetO-hOCT4	Expressing hOCT4	9,477	20726[1]
FUW-tetO-hSOX2	Expressing hSOX2	9,348	20724[1]
FUW-tetO-hKLF4	Expressing hKLF4	9,904	20725[1]
FUW-tetO-hMYC	Expressing hMYC	9,786	20723[1]
Mir-524	Expressing precursor mirna-524	8,125	PMIRH-524AA-1[2]
CD511	Scramble control hairpin in pCDH-CMV-MCS-EF1-copGFP	7,544	CD511B-1[2]
psPAX2	Second generation lentiviral packaging plasmid. Can be used with second or third generation lentiviral vectors and envelope expressing plasmid	10,703	12260[1]
pMD2.G	VSV-G envelope expressing plasmid	5,824	12259[1]

[1]Purchased from Addgene. [2]Purchased from System Biosciences.

Mountain View, CA, USA). The cell mixture was centrifuged at 300 xg for 2 h at 30 °C to enhance viral infection. On day 5 after transduction, 25,000 cells were transferred onto inactivated MEFs and incubated at 37 °C under a hypoxia condition at 5% O_2. On the next day, the culture medium was switched from HFF-1 growth medium to human iPSC (hiPSC) growth medium as described in Tables 3.1 and 3.2 containing 2 µg/ml doxycycline (Dox). Typically, human ESC-like colonies started to appear around day 7. Dox was withdrawn at day 25. ESC-like colonies that emerged were manually picked on day 30 and transferred to fresh inactivated MEF feeder layers and cultured with hiPSC medium until the next passage, which was defined as passage 1.

3.7 Live-Cell Immunofluorescence Staining of Putative ESC-like colonies

3.7.1 Alkaline Phosphatase (AP) Live Stain

ESC-like colonies were stained with alkaline phosphatase live stain (AP) (Life Technologies, Eugene, OR, USA) on day 14 following manufacturer's instructions. Briefly, culture medium was removed and the cells were washed twice with pre-warmed DMEM/F12 for 3 min per wash. The ESC-like colonies were incubated for 30 min with 1X AP live stain solution (Life Technologies). The AP live stain was removed and washed twice with DMEM/F12 for 5 min per wash. After the final wash, fresh

DMEM/F12 was added prior to the visualization of fluorescent-labelled colonies 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). Following visualization, the DMEM/F12 medium was removed and replaced with fresh growth medium.

3.7.2 NANOG Live Stain

The ESC-like colonies that stained positive with AP on day 14 were further stained with NANOG live stain SmartFlare™ (Cat # SF-875, EMD Millipore) on day 18 following manufacturer's instructions. Briefly, SmartFlare^{TM} probe was diluted with phosphate-buffered saline to create a 5.3 nM working solution, which was then directly added to the culture medium to yield a final concentration of 400 pM. Subsequently, the cells were incubated at 37 ^{o}C for 16 h prior to visualization by being directly imaged 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).

3.7.3 Calculation of Reprogramming Efficiency

Reprogramming efficiency was calculated based on the relative numbers of both AP and NANOG-positive colonies after infection of OSKM

+ CD511 or OSKM + mir-524 compared with colonies derived with transduction with OSKM alone, which was set as 1.0.

3.8 Cell Cycle Synchronization of MSC cells at the G2/M Phase Border

Synchronization of MSC WJ0706 cells were synchronized in the G2 phase by incubating the cells with the microtubule inhibitor nocodazole (Cat # M1404, Sigma-Aldrich) at 150 ng/ml final concentration for 18 h. Mitotic cells have round morphologies and loosely attach to the substratum; therefore, the cells can be easily dislodged by gentle mechanical agitation. Hence, mitotic shake-off was applied after nocodazole treatment to collect the suspended cells. The suspended cells were seeded in a new tissue culture plate and cultured with fresh complete medium. The cells were collected at different time points for immunofluorescence staining and flow cytometry analysis.

3.9 Cell Cycle Analysis by Flow Cytometry

Synchronized WJ0706 cells were trypsinised, washed with 1X PBS and fixed with 70% cold ethanol overnight at -20 °C. After fixation, cell pellets were obtained by centrifuging at 1,500 rpm for 5 min to discard ethanol solution. The cell pellets were treated with 500 µL 100 µg/mL RNase A (Thermo Fisher Scientific, San Jose, CA, USA) and stained with 10 µg/mL propidium iodide (Nacalai tesque, Kyoto, Japan) in the dark for 30 min at 37

^{o}C. The stained cells were filtered through a 70-µm cell strainer (BD Biosciences, Bedford, MA, USA) prior to analysis on a FACS Canto-II analyser (BD Biosciences). A total of 10,000 events were recorded at low rates for each sample. DNA distribution in different phases of the cell cycle was analysed using the ModFit LT™ software (BD Biosciences).

3.10 Determination of G1-to-S Transition by Immunoflorescence Staining of 5-Ethynyl-2'-Deoxyuridine (EdU)

The G1-to-S transition of synchronized WJ0706 cells was determined by Click-iT 5-ethynyl-2'-deoxyuridine (EdU) Imaging Kit (Cat # C10337, Invitrogen) according to the manufacturer's instructions. Briefly, the mitotic cells were incubated with 1X EdU solution for 3 h. After incubation, the cells were fixed with 3.7% formaldehyde for 15 min at room temperature, followed by permeabilising with 0.5% Triton X-100 for 20 min. The cells were incubated with the recommended Click-iT reaction cocktail for 30 min at room temperature. For nuclear staining, DAPI (Gibco) was used. Finally, the stained mitotic cells were imaged by Axio Observer A1 with ZEN 2011 Lite software fitted with Argon laser and filter sets BP 450/490 and BP 546/12 (Carl Zeiss).

3.11 RNA Isolation

Total RNA was isolated by using TRIzol reagent (Invitrogen) following the manufacturer's instructions. Cell pellets were homogenised in TRIzol reagent for 10 min. The homogenate was vigorously mixed for 15 s after adding 100 µL chloroform (Amresco) and incubated at room temperature for 15 min. The homogenate was centrifuged at 12,000 xg for 15 min at 4 ^{o}C. The upper transparent phase containing RNA was collected and mixed with 250 µL isopropanol (EMD Millipore), followed by incubation at room temperature for 10 min. The mixture was centrifuged at 12,000 xg for 15 min at 4 ^{o}C and the supernatant was discarded. The pellets were washed with 75% ethanol (EMD Millipore). The supernatant was removed and the pellet was air-dried for 15 min prior to elution of RNase-free water. The isolated RNA was quantified by NanoPhotometer (Implen, Munich, Germany). The isolated RNA samples that have A_{260}/A_{280} values within 1.8 – 2.0 were used for subsequent experiments. All the isolated RNAs were stored at -80 °C.

3.12 Genome-wide Analysis of MicroRNA (miRNA) Expression

3.12.1 MiRNA Profiling

Total RNA (1 µg) extracted from each cell line was analysed using a nanoscale miRNA real-time qRT-PCR array (SmartChip Human MicroRNA Panel v3; WaferGen Biosystems, Fremont, CA, USA) containing 1,036

miRNA-specific reactions in quadruplicates for a total of 5,184 reactions per sample. miRNA was first ligated to a pre-adenylated linker (3' adapter) by RNA ligase 2 at 22 °C for 60 min. The ligated RNA was subjected to one-step on-chip real-time qRT-PCR reaction at 52 °C for 5 min, 95 °C for 10 min, 95 °C for 1 min and 52 °C for 1 min, followed by 39 cycles of PCR amplification at 95 °C for 1 min and 60 °C for 1 min to synthesise cDNA and to amplify target on the SmartChip Cycler (Wafergen Biosystems). In this analysis, an additional 7 endogenous and 4 exogenous controls were included for data quality control.

3.12.2 Selection Criteria for Differentially Expressed miRNAs in Pairwise Comparison

The microarray data for iPSC and the parental cells were compared. For calculations of relative expression levels, the All-Mean normalisation method was employed, where mean C_t of all expressed genes were used (Pritchard et al., 2012). To compute the expression levels of the expressed miRNAs, the C_t values of each sample were compared to its average C_t (All-Mean) to obtain the ΔC_t values. The $\Delta\Delta C_t$ value was then calculated by the two ΔC_t values of the iPSC and its parental cell type. The $\log_2$(fold change) was $\log_2${Fold change (2^-[delta][delta]Ct)}. The selection criteria for differentially expression of miRNA was the $\log_2$(fold change) > 1.5 or < -1.5 with $p < 0.05$.

3.13 Quantification of the MiRNA Expression Levels

3.13.1 MicroRNA Quantitative Real-Time RT-PCR

Real-time qRT-PCR was performed using the NCode SYBR GreenER miRNA qRT-PCR kit (Invitrogen) following the supplier's instructions in a Rotor-Gene Q real-time PCR cycler (Qiagen, Hilden, Germany). Briefly, total RNA (1 µg) was tailed with polyadenylation (polyA) and incubated at 37 °C for 15 min. Following miRNA poly(A) tailing, first-strand cDNA was synthesised using the superscript III RT/RNaseOUT enzyme mix provided in the kit by incubating at 65 °C for 5 min, followed by incubating at 50 °C for 50 min, and 85 °C for 5 min in a 96-well Thermal Cycler (Takara, Shiga, Japan). Synthesised cDNA was amplified by real-time RT-PCR using SYBR Select master mix (Applied Biosystems, Foster City, CA, USA) in Rotor-Gene Q. Amplification was carried out for 40 cycles at 95 °C for 15 s and primer annealing at 60 °C for 1 min. Experiments were performed in triplicates and were normalised to the data of the small nuclear RNA (snRNA) U6. Primers used for miRNA quantification were as follows: forward primer (miR-9-F) 5'-TCTTTGGTTATCTAGCTGTATGA-3' and universal primer (provided in NCode SYBR GreenER miRNA qRT-PCR kit) as the reverse primer. The U6 oligonucleotide 5'-CACCACGTTTATACGCCGGTG-3' was used as the normalisation control. Relative miRNA levels were calculated using the comparative $\Delta\Delta C_t$ method.

3.13.2 MiRNA Stem-Loop RT-PCR

Primers for detection of mature miRNAs were designed according to (Chen et al., 2005). cDNAs were synthesised according to the manufacturer's manual (Invitrogen). Briefly, the annealing programme for the stem-loop primers (Table 3.4) was 5 min at 65 °C. Stem-loop products were then added to an RT reaction using Superscript III reverse transcriptase (Invitrogen) containing 4 μL first-strand buffer, 2 μL 0 .1 M DTT, 0.1 μL RNaseOUT and 0.25 μL (50 units) SuperScript III reverse transcriptase. The reaction was performed with the following incubation conditions: 16 °C for 30 min, followed by 60 cycles of 30 °C for 30 s, 42 °C for 30 s and 50 °C for 1 s. The enzyme was inactivated by incubation at 85 °C for 5 min. The cDNA was used at a dilution of 1:10 in ddH_2O in subsequent PCR reactions. The PCR reactions were performed by using Ex Taq DNA Polymerase (Hot Start version, Takara). The PCR was conducted at 98 °C for 5 min, followed by 40 cycles of 98 °C for 10 sec, 60 °C for 30 sec and 72 °C for 30 sec, and a final extension at 72 ^{o}C for 5 min using a Takara thermocycler. The PCR products were then detected by electrophoresis on 4% agarose gels, and cDNA loading controls were normalised with U6.

Table 3.4 Stem-loop RT-PCR primers

miRNAs	Stem-loop RT primers (5'-3')	microRNA-specific forward primers (5'-3')
hsa-miR-515-5p	GTC GTA TCC AGT GCA GGG TCC GAG GTA TTC GCA CTG GAT ACG ACC AGA AAG T	CGG CGG TTC TCC AAA AGA AAG CA
hsa-miR-515-3p	GTC GTA TCC AGT GCA GGG TCC GAG GTA TTC GCA CTG GAT ACG ACA ACG CTC C	CGG CGA GTG CCT TCT TTT GGA
hsa-mir-519e-5p	GTC GTA TCC AGT GCA GGG TCC GAG GTA TTC GCA CTG GAT ACG ACG AAA GTG C	CGG CGA TTC TCC AAA AGG GAG C
hsa-mir-519e-3p	GTC GTA TCC AGT GCA GGG TCC GAG GTA TTC GCA CTG GAT ACG ACA ACA CTC T	CGG CGC AAG TGC CTC CTT TTA G
hsa-miR-214-5p	GTC GTA TCC AGT GCA GGG TCC GAG GTA TTC GCA CTG GAT ACG ACG CAC AGC A	CGG CGT GCC TGT CTA CAC TTG
hsa-miR-214-3p	GTC GTA TCC AGT GCA GGG TCC GAG GTA TTC GCA CTG GAT ACG ACA CTG CCT G	CGG ACA GCA GGC ACA GAC A
hsa-miR-424-5p	GTC GTA TCC AGT GCA GGG TCC GAG GTA TTC GCA CTG GAT ACG ACT TCA AAA C	CGG GCA GCA GCA ATT CAT GT
hsa-miR-424-3p	GTC GTA TCC AGT GCA GGG TCC GAG GTA TTC GCA CTG GAT ACG ACA TAG CAG C	CGC AAA ACG TGA GGC GCT
U6-F	-	CTC GCT TCG GCA GCA CA
U6-R	-	AAC GCT TCA CGA ATT TGC GT
Universal reverse primer	-	CCA GTG CAG GGT CCG AGG TA

F: forward primer; R: reverse primer

3.13.3 Reverse Transcription and TaqMan miRNA Real-time PCR Assays

RNA was reverse-transcribed using the TaqMan MicroRNA Reverse Transcription Kit and miRNA-specific stem-loop primers (Applied Biosystems) listed in Table 3.5. The reverse transcription reaction (a final volume of 7.5 µl) was as follows: 0.075 µl 100 mM dNTP, 0.5 µl RT enzyme (50 U/µl), 0.75 µl 10X RT buffer, 0.094 µl RNase inhibitor (20 U/µl), 50 ng RNA. Reaction conditions were 16 °C for 30 min, 42 °C for 30 min, 85 °C for 5 min, and held at 4 °C. Real-time PCR assays of the transcribed cDNA were performed using the TaqMan MicroRNA assays (Applied Biosystems). The reaction mixture was as follows: 10 µl TaqMan 2X universal PCR master mix, 1 µl 20X TaqMan MicroRNA Assay, 1.33 µl cDNA and 7.67 µl nuclease-free water. Reaction conditions were: 95 °C for 10 min, followed by 40 cycles at 95 °C for 15 s and 60 °C for 1 min.

3.13.4 Determination of Absolute Copy Number of Mature miRNAs

Synthetic mature miRNAs (Integrated DNA Technologies IDT, Coralville, IA, USA) of the miRNAs under study were 10-fold serially diluted to final concentrations of 200-0.02 nM and 2-0.02 pM. The serially-diluted synthetic RNAs were reverse-transcribed and subjected to real-time PCR analysis concurrently with the sample RNAs. Standard curves were included on each plate of the miRNA TaqMan assays to convert the cycle threshold (C_t)

Table 3.5 TaqMan miRNAs used in this study

Taqman miRNAs[1]	Cat. No.
hsa-miR-512-3p	4427975
hsa-miR-520c-3p	4427975
hsa-miR-520d-5p	4427975
hsa-miR-519b-3p	4427975
hsa-miR-524-5p	4427975
hsa-miR-520f-3p	4427975
hsa-miR-517a-3p	4427975
hsa-miR-520g-3p	4427975

[1]Purchased from Applied Biosystems.

values of each sample into the corresponding number of miRNA copies in each cell, assuming that each cell contains 15 pg total RNA as previously described (Chen et al., 2005). C_t values ≥ 35 indicated that the miRNA expression levels were too low for accurate analysis, and were considered as "no detectable" expression. The cut-off threshold of miRNA expression was, therefore, standardised at $C_t < 35$.

3.14 Quantification of the mRNA Expression Levels

3.14.1 cDNA Synthesis by Reverse Transcription

Total RNA was converted into cDNA by using the Phusion RT-PCR kit (Thermo Fisher Scientific) according to the manufacturer's manual. Total RNA (1 µg) was mixed with 0.25 mM dNTP and 100 ng oligo(dT) primer and incubated at 65 °C for 5 min to pre-denature. Subsequently, the reaction mixture containing 1X RT buffer, RT enzyme mix and RNase-free water was added to the pre-denatured RNA reaction. The reaction tube was incubated at 25 °C for 10 min, 40 °C for 30 min and 85 °C for 5 min in a 96-well Thermal Cycler (Takara).

3.14.2 Determination of mRNA Expression by Direct RT-PCR

The expression of EMT-related genes at the mRNA level was determined by direct RT-PCR using Ex Taq DNA Polymerase (Hot Start

version, Takara). The PCR was conducted at 98 °C for 5 min, followed by 35 cycles of 98 °C for 10 sec, 60 °C for 30 sec and 72 °C for 30 sec, and a final extension at 72 ^{o}C for 5 min using a Takara thermocycler. The PCR products were then detected by electrophoresis on 1.5% agarose gels. The gel was pre-stained with 3X GelRed nucleic acid stain (Biotium, CA, USA) and visualised by exposure to 302 nm UV light under BioSpectrum Imaging System (Ultra-Violet Products Ltd, Cambridge, UK). The primers used are listed in Table 3.6.

3.14.3 Quantification of mRNA Expression by qRT-PCR

The SYBR Select Master Mix kit (Applied Biosystems) was used for real-time RT-PCR quantification of pluripotency genes according to the manufacture's protocol. Briefly, the cDNA was denatured at 95 °C for 10 min, followed by 40 amplification cycles of denaturing step at 95 °C for 15 s and primer annealing and extension step at 60 °C for 1 min. Melt-curve analysis was carried out to determine the specificity and quality of the amplified products by heating the PCR products from 60 °C to 95 °C. The real-time reaction was carried out in a Rotor-Gene Q. GAPDH was used as the normalisation control for all mRNA assays. mRNA expression was calculated based on the $\Delta\Delta C_t$ method. All experiments were done in triplicates, and three or more independent experiments were performed to obtain the results presented. The primers for mRNA quantification are shown in Table 3.7.

Table 3.6 Primers of EMT-related genes used for direct RT-PCR and qRT-PCR

Genes	Source	Primer Sequences (5'-3')	Amplicon size (bp)
TGFβR1	(He et al., 2016)	F: ACGGCGTTACAGTGTTTCTG R: GCACATACAAACGGCCTATCTC	167
SMAD2	(Subramanyam et al., 2011)	F: TTCAGTGCGTTGCTCAAGCATGTC R: AACAGTCCATAGGGACCACACACA	111
SMAD3	(Hao et al., 2016)	F: CCTCTCCAGCAATAATCCGAA R: TGCCCAATTTTCTTTACCAGT	240
SMAD4	(Hao et al., 2016)	F: ACTTGGCATCTCTACATTGTCC R: GCCACATCTATTTTGCTTGCT	226
ZEB1	(Subramanyam et al., 2011)	F: ATGCACAACCAAGTGCAGAAGAGC R: TTGCCTGGTTCAGGAGAAGATGGT	145
ZEB2	(Subramanyam et al., 2011)	F: ATATGGTGACACACAAGCCAGGGA R: GTTTCTTGCAGTTTGGGCACTCGT	177
TWIST1	(P. Zhang et al., 2015)	F: CTGTTGTTGCTGTGGCTGATA R: CCGTCCACAAGCAATGAGT	146

F: forward primer; R: reverse primer

Table 3.7 Primers of pluripotency genes used for direct RT-PCR and qRT-PCR

Genes	Source	Primer Sequences (5'-3')	Amplicon size (bp)
TP53INP1	(J. Zhang et al., 2014)	F: GCCCCACGTACAATGACTCTTCT R: GCCCTTCTT GGT TGG AGG AAG AAC	221
p53	(Ye et al., 2012)	F: AGCTGAATGAGGCCTTGGAACT R: AGGCCCTTCTGTCTTGAACAT	143
OCT4	(Takahashi et al., 2007)	F: GACAGGGGGAGGGGAGGAGCTAGG R: CTTCCCTCCAACCAGTTGCCCCAAAC	144
SOX2	(Takahashi et al., 2007)	F: GGGAAATGGGAGGGGTGCAAAAGAGG R: TTGCGTGAGTGTGGATGGGATTGGTG	151
NANOG	(Takahashi et al., 2007)	F: AGTCCCAAAGGCAAACAACCCACTTC R: TGCTGGAGGCTGAGGTATTTCTGTCTC	167
REX1	(Takahashi et al., 2007)	F: TCTGAGTACATGACAGGCAAGAA R: TCTGATAGGTCAATGCCAGGT	62
GAPDH	(P. Zhang et al., 2015)	F: GAAATCCCATCACCATCTTCCAGG R: GAGCCCCAGCCTTCTCCATG	120

F: forward primer; R: reverse primer

3.15 Determination of Protein Expression by Western Blot Analysis

3.15.1 Buffers and Reagents Preparation

The buffers and reagents used for western blot analysis are as described in Table 3.8.

3.15.2 Preparation of Cell Lysates and Quantification of Protein Lysates

Protein lysates were prepared by lysing the cell pellets in RIPA Lysis Buffer (Nacalai Tesque) following the manufacturer's instructions. Culture cells were first trypsinised and pelleted before adding RIPA Lysis. DNA fragmentation was done by passing the lysed suspension through a 21-G needle. The cell suspension was incubated on ice for 15 min, followed by centrifugation at 12,000 rpm for 15 min at 4 °C in a Herarus Fresco 21 refrigerated microfuge (Thermo Scientific). The supernatant was collected into a new microfuge tube and the cell pellet was discarded. All steps were performed on ice to prevent protein denaturation.

To standardise the protein amount loaded into the gel, protein quantification was carried out using Quick Start Bradford Protein Assay (Bio-Rad, CA, USA). A Protein Standard Curve was first constructed based on serial dilutions of concentrated bovine serum albumin (BSA), ranged from 125 to 2,000 µg/mL. The diluted protein standards (5 µL) and 5X diluted unknown

Table 3.8 Preparation of western blot buffers

Buffer	Methods of preparation
10% (w/v) sodium dodecyl sulphate (SDS) solution	- 10 g SDS (EMD Millipore) - 90 mL H_2O_2
10% (w/v) ammonium persulfate (APS)	- 0.1 g APS (EMD Millipore) - 1 mL H_2O_2
1.5 M Tris, pH 8.8	- 18.15 g Tris base (Norgen Biotek) - top up to 100 mL H_2O_2
0.5 M Tris, pH 6.8	- 6 g Tris base , pH 6.8 - top up to 100 mL H_2O_2
SDS-PAGE running buffer (0.05 M Tris, 0.384 M glycine, 0.1% SDS)	- 12 g Tris base - 57.6 g Glycine (EMD Millipore) - 2 g SDS - top up to 2 L H_2O_2
10X Tris-buffered saline (TBS) (0.5 M Tris, 1.5 M NaCl)	- 60.57 g Tris base, - 87.66 g NaCl - pH 7.5 - top up to 1 L H_2O_2
Washing buffer: 1X TBS with 0.05% Tween-20 (TBST)	- 100 mL of 10X TBS - 900 mL H_2O_2 - 500 µL Tween-20 (EMD Millipore)
Blocking buffer: TBST with 5% milk	- 2 g non-fat milk powder (Bio Basic Inc.) - 40 mL of 1X TBST
1X Transfer buffer (25 mM Tris, 192 mM glycine, 10% methanol)	- 3.03 g Tris base - 14.4 g glycine - 1 L mL H_2O_2 - 200 mL of methanol (EMD Millipore)
Mild stripping buffer	- 1.5 g of glycine - 0.1 g SDS - 1 mL Tween-20 - 100 mL H_2O_2 - pH 2.2

samples were pipetted into a 96-well microplate. Then, 250 µL 1X Dye Reagent (Bio-Rad) was added to each well, mixed and incubated in the dark for 15 min. The amount of proteins in the microplate was then measured at absorbance of 595 nm in an Infinite M200 PRO Microplate Reader (Tecan). All standards and samples were performed in triplicates. The concentration of the protein lysate was calculated based on average absorbance reading of the protein lysate and the BSA protein standard curve.

3.15.3 Protein Separation by SDS-polyacrylamide Gel Electrophoresis (SDS-PAGE)

A Mini-PROTEAN Tetra Cell system (Bio-Rad) was used in SDS-polyacrylamide gel electrophoresis (SDS-PAGE). A short plate was placed on top of a spacer plate (1.0 mm thickness) and was secured on a casting stand to form a gel cassette assembly. A 10% resolving gel was prepared from 2.5 mL 1.5 M Tris (pH 8.8), 3.3 mL 30% (w/v) Bis/Acrylamide (Bio-Rad), 100 µL 10% SDS and 4.1 mL distilled water and mixed well. Tetramethylethylenediamine (TEMED) (Calbiochem, EMD Millipore) (5 µL) and 50 µL 10% APS were added last to the resolving gel solution. All the components were mixed well and pipetted in between the glass plates assembled on the gel casting stand. The resolving gel was gently overlaid with 50 µL isopropanol (EMD Millipore) to remove any bubble formation. After 45 min when the resolving gel was polymerised, isopropanol was rinsed off with distilled water and the residual was removed with Kimwipe tissue paper. A

5% stacking gel was prepared from 1 mL 0.5 M Tris (pH 6.8), 0.67 mL 30% (w/v) Bis/Acrylamide, 30 µL 10% SDS and 1.9 mL distilled water and mixed well. A total of 3.3 µL TEMED and 30 µL 10% APS were added last to the stacking gel solution and mixed well. The stacking gel solution was then layered on top of the polymerised resolving gel. A 10-well 1.0-mm comb was immediately inserted into the stacking gel layer without forming air bubbles.

When the stacking gel polymerised, the gel cassette was clamped onto an electrode assembly and placed into the Mini-PROTEAN Tatra Tank. The comb was removed and the formed wells were rinsed with 1X running buffer. The electrode assembly was filled with SDS-PAGE running buffer (Table 3.8) until full, and the tank was filled to the indicated level. Four volumes of protein lysate at a final amount of 50 µg were mixed with one volume of 5X Lane Marker Reducing Sample Buffer (Thermo Scientific), and the mixture was heated at 99 °C for 5 min in a block heater (Stuart Scientific, Staffordshire, UK). Aliquots of the denatured protein sample buffer mixture and 5 µL MagicMark™ XP Western Protein Standard (Thermo Fisher Scientific) were loaded into the wells accordingly. Electrophoresis was performed at a constant voltage of 120 V for approximately 75 min.

3.15.4 Semi-Dry Transfer of Protein from Gel to Nitrocellulose Membrane

While performing the gel electrophoresis, Hybong ECL nitrocellulose membrane (GE Healthcare, Little Chalfont, UK) and western blotting filter

paper (Thermo Scientific) were cut to the dimension of the gel, which was approximately 10 cm x 6 cm. When the dye front was approaching the bottom of the resolving gel, the power was turned off and the electrophoresis apparatus was disassembled. The gel was removed gently from the gel cassette and the stacking gel layer was removed, leaving only the resolving gel with separated proteins. The resolving gel, together with the nitrocellulose membrane and two pieces of blotting papers were soaked in 1X transfer buffer (Table 3.8) and equilibrated for 12 to 15 min.

The polyacrylamide gel was subjected to semi-dry electro-transfer by using the Trans-Blot SD Semi-Dry Electrophoretic Transfer Cell (Bio-Rad). A transfer sandwich was arranged accordingly. One piece of equilibrated blotting paper was placed at the bottom, followed by buffer-equilibrated nitrocellulose membrane, the buffer-equilibrated polyacrylamide gel and finally another piece of blotting paper. The transfer sandwich was placed on the platinum anode of the semi-dry transfer cell. A centrifuge tube was used as a roller to exclude all trapped bubbles before the upper cathode was placed onto the stack. Separated protein samples on the gel were blotted to membrane with a constant voltage of 15 V for 15 min. The nitrocellulose membrane was removed from the sandwich and the gel orientation was marked on the membrane with pencil. The membrane was washed with distilled water for 5 min with gentle agitation on a gyratory rocker. The membrane was stained with RedAlert Stain (Novagen, EMD Millipore) to verify the transfer of proteins before western blot analysis. The RedAlert stain was removed by

rocking the membrane in distilled water. The membrane was ready for blocking.

3.15.5 Membrane Blocking

The transferred nitrocellulose membrane was blocked with 5% milk to prevent non-specific binding of the antibodies onto the membrane. The membrane was incubated in 10 mL of blocking buffer (Table 3.8) for 1 h at room temperature with gentle agitation on a gyratory rocker. The blocking solution was discarded and the membrane was washed with 10 mL washing buffer (Table 3.8) three times for 10 min per each time with gentle agitation on a gyratory rocker, with fresh changes of washing buffer.

3.15.6 Antibody Staining

After the washing step, the membrane was blotted with polyclonal antibodies against TP53INP1 at 1:1,000 dilution (ab9777, Abcam, Cambridge, UK), SMAD4 (1:500, ab137861, Abcam), β-actin (1:1,000, ab8227, Abcam) or monoclonal antibody ZEB2 (1:500, sc-271984, Santa Cruz, CA, USA) and incubated overnight at 4 °C with slight agitation. The membrane was washed with 5 mL washing buffer with agitation on a gyratory rocker 3 times with 10 min per each time. The membrane was then incubated in 5 mL HRP-conjugated goat anti-rabbit IgG secondary antibody (ab97051, Abcam) for

TP53INP1, or HRP-conjugated goat anti-mouse IgG secondary antibody (ab6789, Abcam) for SMAD4 and β-actin. The secondary antibodies were diluted at 1:5,000 with the washing buffer and incubated with the membranes for 1 h at room temperature with gentle agitation. The membrane was again washed with 10 mL washing buffer three times 10 min each with agitation.

3.15.7 Chemiluminescence Detection

After the washing step, the excessive wash buffer was drained off from the blot. The membrane was placed on a sheet of transparent paper, with the protein side facing upward. Detection Solution 1 and 2, supplied in the Amersham Enhanced Chemiluminescent Western Blotting Detection Reagent (GE Healthcare), were mixed at a ratio of 1:1 and the mixture was added directly onto the membrane. Cumulative chemiluminescent signal emitted from the membrane after an exposure time of 15 min was captured by the LI-COR C-DiGit chemiluminescence western blot scanner (Li-Cor Biosciences; Lincoln, NE, USA). The band densitometric analysis was performed by Image Studio™ Lite software (Li-COR Biosciences) to measure the intensity of the protein bands developed in the western blot membrane.

3.15.8 Stripping and Reprobing

After chemiluminescence detection, the primary and secondary antibodies on the membrane were removed for further use of the membrane. The developed membrane was rinsed with the wash buffer followed by incubation in 7 mL mild stripping buffer (Table 3.8) for 10 min at room temperature twice, with gentle agitation. The membrane was washed with 7 mL 1X PBS twice for 10 min, and with 7 mL washing buffer twice for 5 min. Subsequently, the stripped membrane was again blocked in the blocking buffer, followed by staining with primary and secondary antibodies. Chemiluminescent signal detection was as described in Sections 3.15.5, 3.15.6 and 3.15.7.

3.16 Over-expression of miRNAs

3.16.1 Transient Transfection of Synthetic miRNAs

MirVana miR-524-5p mimic and negative control (NC) were designed and synthesised by Ambion (Foster City, CA, USA), whereas the ON-TARGETplus SMARTpool siRNA (Thermo Scientific) containing a mixture of four SMART selection-designed small interfering (siRNA) targeting the human TP53INP1 gene was used. The synthetic miRNAs were transfected into MSC WJ0706 cells by using Lipofectamine RNAiMAX Reagent (Invitrogen) according to the manufacturer's protocol. Briefly, on the day of transfection, the

WJ0706 cells were seeded at a density of 9.5 x 10^4 cells/well in a 6-well plate. Twenty-four hours post-seeding, 100 nM synthetic miRNAs was diluted in 250 µL OptiMEM medium (Gibco); 5 µL Lipofectamine RNAiMAX (Invitrogen) was also diluted in 250 µL OptiMEM medium (Gibco) and incubated for 5 min at room temperature. Both diluted synthetic miRNAs and Lipofectamine RNAiMAX were mixed and incubated at room temperature for 20 min. The synthetic miRNA-lipofectamine complex was added to the cells and incubated in a 37 °C cell culture incubator for 48 h. At 48 h post-transfection, the cells were trypsinised for further analysis.

3.16.2 Co-Transfection of Synthetic miRNAs and Plasmids Containing 3'UTR Regions of Predicted Target Genes

The synthetic miRNAs and plasmid constructs containing 3'UTR regions of predicted target genes were co-transfected into colorectal cancer cell line HCT-15 by using Lipofectamine 2000 Reagent (Invitrogen) to study the correlation between miRNAs and their predicted target genes. On the day of transfection, HCT-15 the cells were seeded onto 24-well plates at a density of 2.5 x 10^5 cells per well in a 24-well plate. Twenty-four hours post-seeding, 30 nM synthetic miRNAs and 250 ng/µL plasmids were diluted in 50 µL OptiMEM medium (Gibco); 1 µL Lipofectamine 2000 (Invitrogen) was also diluted in 50 µL OptiMEM medium (Gibco) and incubated for 5 min at room temperature. Both diluted synthetic miRNA/plasmid and Lipofectamine 2000 were mixed and incubated at room temperature for 20 min. The complex was added to the cells

and incubated in a 37 °C cell culture incubator for 24 h. At 24 h post-transfection, the medium was changed to fresh complete medium for HCT-15 (Table 3.1 and 3.2). At 72 h post-transfection, the cells were harvested for luciferase activity analysis.

3.17 Validation of miRNA-Targeted Transcripts by Luciferase Assays

3.17.1 Construction of pmirGLO Plasmids Containing 3'UTR of Putative Target Genes

To amplify 3'UTR sequences of putative target genes carrying the putative miRNA target sites, primers were designed by NCBI primer blast (Table 3.9). The cDNAs converted as described in Section 3.14.1 were used as template for amplification by using SeqAmp DNA Polymerase (Clontech, Palo Alto, CA, USA) according to manufacturer's manual. The PCR was conducted at 94 °C for 5 min, followed by 30 cycles of 94 °C for 10 sec, 60 °C for 30 sec and 72 °C for 1 min/kb, and a final extension at 72 ^{o}C for 5 min using a Takara thermocycler. The PCR products were purified by NucleoSpin Gel and PCR Clean-up (Macherey-Nagel, Duren, Germany) before subjecting to restriction enzymes double digestion at the SacI and XbaI restriction sites of the amplified products. The digested PCR fragments were ligated with linearised pmirGlo (GenBank accession FJ376737) (Promega, Madison, WI, USA) also digested at the SacI and XbaI restriction sites by using T4 DNA Ligase (New England Biolabs Inc., USA). The ligation products were then

Table 3.9 Primers for amplification and cloning of luciferase wild-type constructs

Primer designation	Sequence (5' > 3')
TP53INP1-F1	TAA GCA **GAG CTC** CAG TGT TTG GGG GTG TCT TT
TP53INP1-R1	CCG TGG **TCT AGA** AAT TGG CGG GAA GGA ATA GT
TP53INP1-F2	TAA GCA **GAG CTC** ACA CGG CGT CTC TTT TTC AT
TP53INP1-R2	CCG TGG **TCT AGA** AAT GCA TTT TGG CCA TGT TT
TP53INP1-F3&4	TAA GCA **GAG CTC** GGG AGG TTA GAT GTG TGT TT
TP53INP1-R3&4	CCG TAC **TCT AGA** GTA ACT CCA GGT AGT GCA AA
ZEB2-F	TAA GCA **GAG CTC** GCG GTT CAG CCA AGA CAG AT
ZEB2-R	CCG TGG **TCT AGA** ACT GAA GCT GGT GCA AAG GT
SMAD4-F	TAA GCA **GAG CTC** TCT CTT TGG AGC CAA GCC AC
SMAD4-R	CCT TGG **TCT AGA** GGC CTA GGA TGC CAC TTT GT

transformed into DH5α competent cells (Invitrogen) by heat-pulse at 42 °C for 45 sec. The transformed *E.coli* cells were incubated on ice for 2 min prior to add room-temperature recovery medium (250 µL) provided by the competent cell kit. The mixture was incubated in a shaking incubator at 250 rpm for 1 h at 37 °C. After the incubation, 80 µL transformed cells was plated on pre-warmed Luria Bertani (LB, Sigma-Aldrich) agar plate containing 100 µg ampicillin (Amresco). The plates were incubated overnight at 37 °C for 16-18 h for colony formation.

Mutations of the miRNA seed sequences were performed using the QuikChange Lightning Site-Directed Mutagenesis kit (Agilent Technologies, Santa Clara, CA, USA) as recommended by the supplier. Briefly, the recombinant wild-type plasmid constructs containing segments of 3'UTR covering predicted miRNA binding sites of target genes were used as template for amplification. The primer sequences for the mutant constructs are shown in Table 3.10. An amount of 100 ng wild-type plasmid constructs was added to the mixture of the provided reagents including 10X reaction buffer, dNTP mix, Quiksolution reagent, primers for the mutant constructs (Table 3.10), and QuikChange Lightning Enzyme. The mixture was then subjected to PCR amplification at 95 °C for 2 min, followed by 25 cycles of 95 °C for 20 s, 60 °C for 10s and 68 °C for 30 s/1 kb, and a final extension at 68 ^{o}C for 5 min using a Takara thermocycler. The amplification reactions were incubated with Dpn I restriction enzyme incubated at 37 °C for 5 min to digest the parental supercoiled dsDNA. Subsequently, 2 µl Dpn I-treated DNA was transformed into XL10-Gold ultracompetent cells by heat-pulse in a 42°C water bath for

Table 3.10 Primers for amplification and cloning of luciferase mutant constructs

Primer designation	Sequence (5' > 3')
INP1-mut1-24-F	TTCATTTTCATTTTATgaatTcTTACTTAATCTTTTAAGCAAGCA
INP1-mut1-24-R	TGCTTGCTTAAAAGATTAAGTAAgAattcATAAAATGAAAATGAA
INP1-mut2-24-F	GCCTTACCTGGGGCTAGTTTTTTATGCgaatTcCCTAGAAAAC
INP1-mut2-24-R	GTTTTCTAGGgaAttcGCATAAAAAACTAGCCCCAGGTAAGGC
INP1-mut3-24-F	CTGATTGGTTCATAGATGGTCAGTgaatTcCACAGACTGAAC
INP1-mut3-24-R	GTTCAGTCTGTGgAattcACTGACCATCTATGAACCAATCAG
INP1-mut4-24-F	TGTGTGTTAACACCTGTTCgaatTcATTGGGTTGTGGTGCAT
INP1-mut4-24-R	ATGCACCACAACCCAATgAattcGAACAGGTGTTAACACACA

Lower-case letters indicate the mutated sequences, which conveniently generated an EcoRI cleavage site for easy confirmation of the generated mutations.

30 s. The transformed E.coli cells were incubated on ice for 2 min prior to adding room-temperature recovery medium (250 µL). The mixture was incubated in a shaking incubator at 250 rpm for 1 h at 37 °C. After the incubation, 80 µL transformed cells was plated on pre-warmed Luria Bertani (LB, Sigma-Aldrich) agar plate containing 100 µg ampicillin (Amresco). The plates were incubated overnight at 37 °C for 16-18 h for colony formation.

3.17.2 Plasmid DNA Extraction and Selection of Transformed Colonies

Single colonies of transformed cells were picked from the LB agar plate with a pipette tip and dropped in pre-warmed LB broth (Sigma-Aldrich) containing 100 µg ampicillin (Amresco), and were incubated overnight at 37 °C with 250 rpm shaking. Plasmid DNA was purified using GeneJet Plasmid Miniprep Kit (Thermo Fisher Scientific). The bacterial culture was harvested by centrifugation at 8,000 rpm for 5 min in a Sorvall Legend centrifuge X1R machine (Thermo Fisher Scientific). The pelleted bacterial cells were re-suspended in 250 µL re-suspension solution which was then added with 250 µL lysis solution and immediately inverted four to six times until the solution became viscous and slightly clear. Next, 350 µL neutralization solution was added to the bacterial lysate and mixed thoroughly until the lysate turned cloudy, followed by centrifugation at 12,000 rpm for 5 min in a Sigma 1-14 benchtop centrifuge machine (Sartorius Corporation, NY, USA). The supernatant was pipetted to a new GeneJET spin column and centrifuged at 12,000 rpm for 1 min. The flow-through was discarded and the

column was returned to the same collection tube. Washing step was carried out twice by adding 500 µL wash solution to the spin column and centrifuged at 12,000 rpm for 60 sec. The flow-through was discarded and the column was centrifuged for an additional 1 min to remove residual wash solution. The spin buffer was added to the centre of the spin column membrane. The spin column was allowed to incubate for 5 min at room temperature before centrifugation at 12,000 rpm for 60 sec to elute the plasmid DNA. The purified plasmid DNA from different colonies were digested with SacI and XbaI restriction enzymes in the case of wild-type plasmid constructs, and with EcoRI for mutant plasmid constructs. The digested plasmids were run on 1% agarose and visualised by exposure to 302 nm UV light under BioSpectrum Imaging System (Ultra-Violet Products Ltd) in order to identify the recombinant plasmid constructs. The obtained recombinant plasmid constructs were confirmed by sequencing.

3.17.3 Luciferase Reporter Assays

The Dual-Luciferase Reporter Assay System (Promega, WI, USA) was used to identify miRNA direct targeting. Co-transfection of recombinant plasmid constructs and synthetic miRNAs into HCT-15 cells was performed by using Lipofectamine 2000 (Invitrogen) according to Section 3.16.2. Forty-eight hours post-transfection, the medium was discarded and the cells were washed twice with 1X PBS before addition of 100 µL 1X Passive Lysis Buffer to the transfected cells and rocking at room temperature for 15 min. Cell lysate

(20 µL) was then transferred to a 96-well plate pre-dispensed with 100 µL Luciferase Assay Reagent II to measure the Firefly activity in an Infinite M200 PRO Microplate Reader (Tecan). After quantifying the firefly luminescence, 100 µL Stop & Glo Reagent was added simultaneously to the same well to initiate Renilla activity. Transfection experiments were performed in two or more independent experiments with quadruple transfections each.

3.18 Prediction of miRNA Target Genes

The miRNA putative target genes were predicted based on previous studies with some modifications (Ritchie et al., 2009; Ritchie et al., 2013). Briefly, the 3'UTR sequences of putative target genes were retrieved from UCSC genome browser (http://genome.ucsc.edu). miRNA:mRNA interactions were predicted using the major miRNA databases TargetScan and microRNA.org. To identify genes and pathways specifically targeted by selected C19MC-AAGUGC-miRNAs, overlapping target gene sets of the selected miRNAs were used for the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway and Gene Ontology (GO) annotation analysis based on the web-based DAVID (Database for Annotation, Visualization and Integrated) algorithm. The criteria of analysis was EASE score ≤ 0.05, in which EASE score is a modified Fisher Exact P value in the DAVID system used for gene-enrichment analysis. An EASE score P value = 0 represents

perfect enrichment; $p \leq 0.05$ was considered as significant gene-enrichment in a specific annotation category.

3.19 Construction of Phylogeny Tree

Phylogenetic tree alignment of the 3'UTR of the TP53INP1 transcript sequences in different species was generated by the Clustal method using the DNAstar software (Madison, WI, USA). The mature sequences of 16 C19MC-AAGUGC-miRNAs and the known reprogramming miRNA families were downloaded from miRBase ver.21 and aligned using the Clustal W. Similarly, the stem-loop sequences based on precursor sequence for all C19MC miRNAs were also downloaded from miRBase ver. 21. A phylogenetic tree of the stem-loop sequences for C19MC was generated by multiple sequence alignment using the Clustal method of the Megalign project provided by DNAstar (USA).

3.20 Cell Proliferation Assays

3.20.1 Cell Growth Analysis

WJ0706 cells were transfected with a miRNA mimic, miRNA NC (negative control) or TP53INP1 siRNA (siTP53INP1) as described in Section 3.16.1. After 48 h incubation, cells were seeded in 6-well plates at a density of 1 x 104 cells/well. Two, 4 and 6 days post-transfection, the cells were

trypsinised and stained with trypan blue (Gibco Invitrogen). Total cells were counted every second day using a hemocytometer (Hirschmann). Data presented were from three independent experiments, and the results of the treated cells were normalised with the untreated control cells. The formula for determining the number of cells was as below:

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

3.20.2 5-bromo-2'-deoxyuridine (BrdU) Cell Proliferation Assay

WJ0706 cells were transfected with a miRNA mimic, miRNA NC or siTP53INP1 as described in Section 3.16.1. After 48 h incubation, cells were seeded in 96-well plates at a density of 5,000 cells/well for 24 h. Cell proliferation was measured by using the 5-bromo-2'-deoxyuridine (BrdU) cell proliferation assay kit (Cell Signaling Technology, Denvers, MA, USA) according to the manufacturer's instructions. Briefly, after 24 h, 1X BrdU solution was added to the medium and incubated at 37 ^{o}C for 2 h. Medium was then removed and 100 µL fixing/denaturing solution was added and incubated at room temperature for 30 min. Antibody solution (100 µL) was added and incubated for another hour. Washing step was carried out three times using 1X washing solution before incubation with 100 µL horseradish peroxidase (HRP)-conjugated secondary antibody solution for 30 min. The treated cells were again washed with 1X washing solution before addition of 100 µL tetra-methylbenzidine substrate. After incubation in the dark for 30 min, 100 µL STOP solution was added and the absorbance was measured at 450 nm in

Infinite M200 PRO Microplate Reader (Tecan). Data presented were from three independent experiments in triplicates, and the results of the treated cells were normalised with the untreated control cells.

3.21 MTT Assays for Cell Viability

MTT [3-(4,5-Dimethylthiazol-2-yl)-2,5-Diphenyltetrazolium Bromide] (Sigma-Aldrich) was used to quantify cell survival from H_2O_2-induced oxidative stress. Briefly, after 48 h post-transfection with miRNA or NC mimic or siTP53INP1 (Section 3.16.1), the transfected WJ-MSC cells were treated with 200 µM H_2O_2 for 2 h. Subsequently, the cells were trypsinised and seeded in 96-well plates at a density of 5,000 cells/well and cultured for 24 to 96 h, followed by addition of 10 µl 5 mg/m1 MTT to each well and incubation for 2.5 h. The reaction was stopped by adding 100 µl dimethyl sulfoxide (DMSO). Absorbance at 570 nm was determined using Infinite M200 PRO Microplate Reader (Tecan).

3.22 Histone/DNA Enzyme-Linked Immunosorbent Assay (ELISA) for Detection of Apoptosis

The Cell Death Detection ELISA plus kit (Roche Diagnostics, Penzberg, Germany) was employed to quantitatively detecting histone-associated DNA fragments in mono- and oligo-nucleosomes according to the manufacturer's protocol. Briefly, 48 h post-transfection (Section 3.16.1), cells

were trypsinised and seeded in 96-well plates at a density of 5,000 cells/well and cultured for 24 h. The cells were treated with 200 µM H_2O_2 for 6 h. After treatment, the cytoplasmic histone/DNA fragments from the cells were extracted and incubated in microtiter-plate modules coated with an anti-histone antibody. Subsequently, peroxidase-conjugated anti-DNA antibody was used for the detection of immobilised histone/DNA fragments followed by colour development with an ABTS substrate for peroxidase. The spectrophotometric absorbance of the samples was determined at 405 nm.

3.23 Statistical analysis

Data were analysed by paired Student's *t*-test (two-tailed distribution) comparing the differences of expression levels between treated and untreated cells. Statistical software Microsoft Excel was used. Statistical significance was accepted at $p < 0.05$.

CHAPTER 4

RESULTS AND DISCUSSION (PART I)

Selective Activation of miRNAs of The Primate-Specific Chromosome 19 miRNA Cluster (C19MC) in Stem Cells and Possible Biological Functions

[Published: Nguyen PNN *et al.* (2017) J. Biomed. Sci, 24:20, DOI: 10.1186/s12929-017-0326-z]

4.1 Background

MiRNAs are known to regulate the maintenance of stem cell self-renewal (Wang et al., 2008; Li et al., 2011), and specific miRNAs may be up-regulated in pluripotent stem cells population (Razak et al., 2013). Furthermore, miRNAs also act as tumour suppressors or oncogenes in the tumorigenesis process (Liang et al., 2007; Keklikoglou et al., 2012; S. Zhang et al., 2012; Kleinman et al., 2014; Lu et al., 2015).

For regulatory advantages, miRNAs, particular those from the same family, are often clustered in specific chromosomal locations (Wang et al., 2011). One such human miRNA cluster is mapped on chromosome 19, and is called the chromosome 19 miRNA cluster, or C19MC, which contains 46 highly homologous miRNA genes, including 7 duplicated pairs of the same genes, within a ~100-kb genomic region (Bortolin-Cavaille et al., 2009) (See

Chapter 2, subsection 2.5.1). In the human placenta, C19MC is expressed *en bloc* from the paternal allele thought to be regulated by a major promoter located 17.6 kb upstream of the cluster (Noguer-dance et al., 2010). The biological functions and expression patterns of C19MC members in other stem cell types and in cancer cells have not been systematically examined in a cluster-wide manner.

This chapter focuses on the study of the C19MC miRNA expression profile in pluripotent, multipotent and unipotent stem cells. As C19MC miRNAs are frequently selectively activated in cancer cells (Chapter 2, subsection 2.5.4), the expression pattern of C19MC miRNAs in cancer cell lines was also examined as a reference to expression in stem cells.

4.2 Study Design

To elucidate the expression profile of C19MC in different cell types, the study design of Part I was as shown in Figure 4.1.

Part I: The expression profile of C19MC in different stem cell types

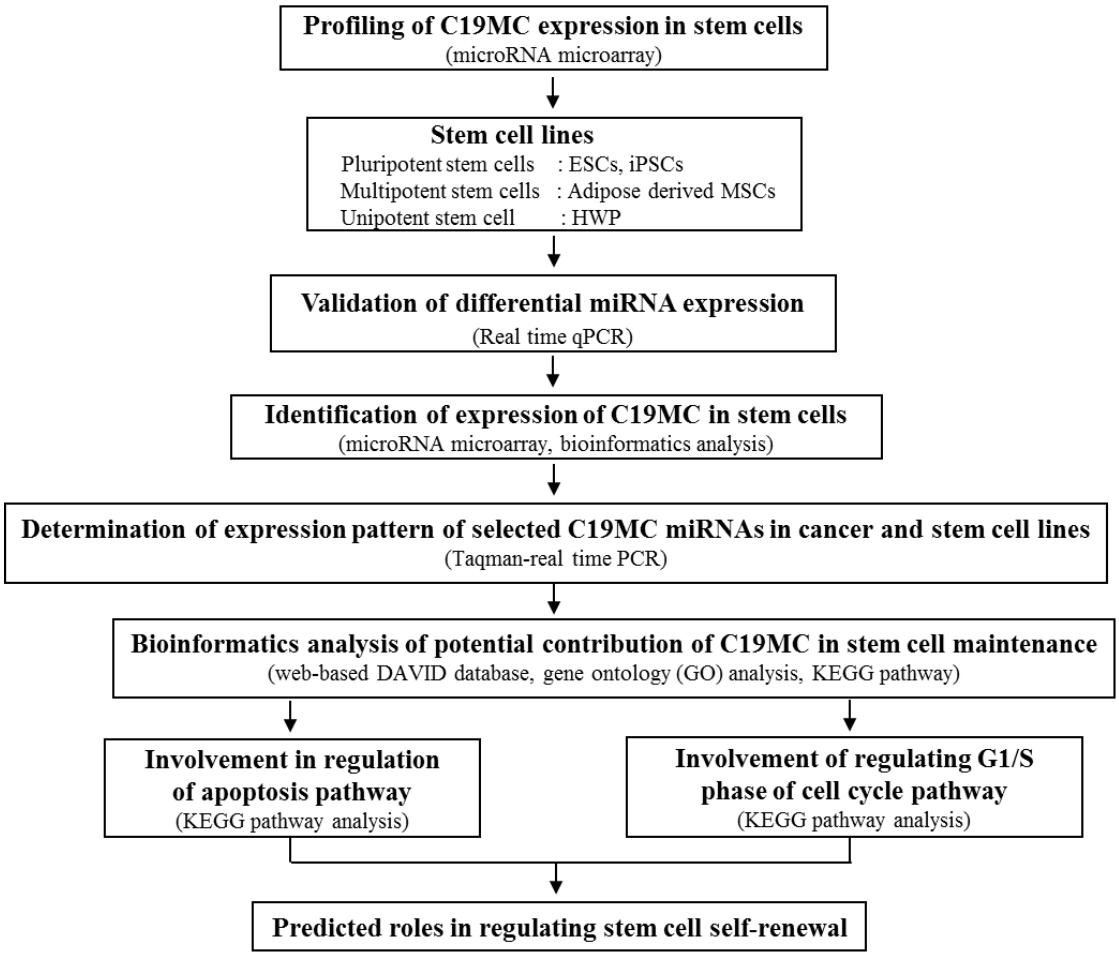


Figure 4.1 Study design of Part I. Elucidation of the expression profile of C19MC in different cell types**.** The experimental assays or procedures carried out are described in round brackets.

4.3 Results

4.3.1 Differential Expression of miRNAs in Different Stem Cell Types

To identify differentially expressed miRNAs in iPS cells relative to the parental cells from which they were derived via reprogramming, a HWP cell line, two human adipose-derived mesenchymal stem cell lines, designated as ASC and hMSC-AT, and the three induced pluripotent stem iPS cell lines derived from them were used (Table 4.1). Two well-characterized ESC lines, H6 and H9 were also included for comparison. Total RNAs prepared from these eight cell lines were subjected to quantitative miRNA profiling using a nanolitre-scale real-time RT-PCR microarray platform that included 1,036 miRNA species. On obtaining the microarray data, hierarchical clustering analysis of the miRNA profiles was performed between the two MSC and the three iPSC (Figure 4.2 and Appendix H). Since the HWP cells are unipotent and not multipotent cells (Ong and Sugii, 2013), HWP was omitted from the two multipotent MSC cell lines in the hierarchical clustering analysis. The clustergram showed that the miRNAs were clustered into two major (I & II) and one minor (III) clusters. Cluster I included miRNAs that were expressed in high levels in the MSC lines whereas cluster II included miRNAs highly expressed in iPSC. In each of these two clusters, there were also subclusters generated based on different miRNA levels observed. The data clearly showed that different miRNA signatures in the multipotent MSC relative to the pluripotent iPSC. Furthermore, miRNAs in cluster III were in high expression levels specifically in the HWP-derived iPSC in various lower expression

Table 4.1 Stem cell types used in microarray analysis

Cell type		Cell lines*	Abbreviation
Unipotent stem cell	HWP	- Human white preadipocyte	- HWP
Multipotent stem cell	MSC	- Adipose-derived MSC	- ASC - hMSC-AT
Pluripotent stem cells	ESC	- HuES6	- H6
		- H9	- H9
	iPSC	- Induced from ASC-derived MSCs	- ASC-iPSC - MSC-iPSC
		- Induced from human white preadipocyte (HWP)	- HWP-iPSC

*All cell lines were provided by Dr. Sugii's lab (Sugii et al., 2011)

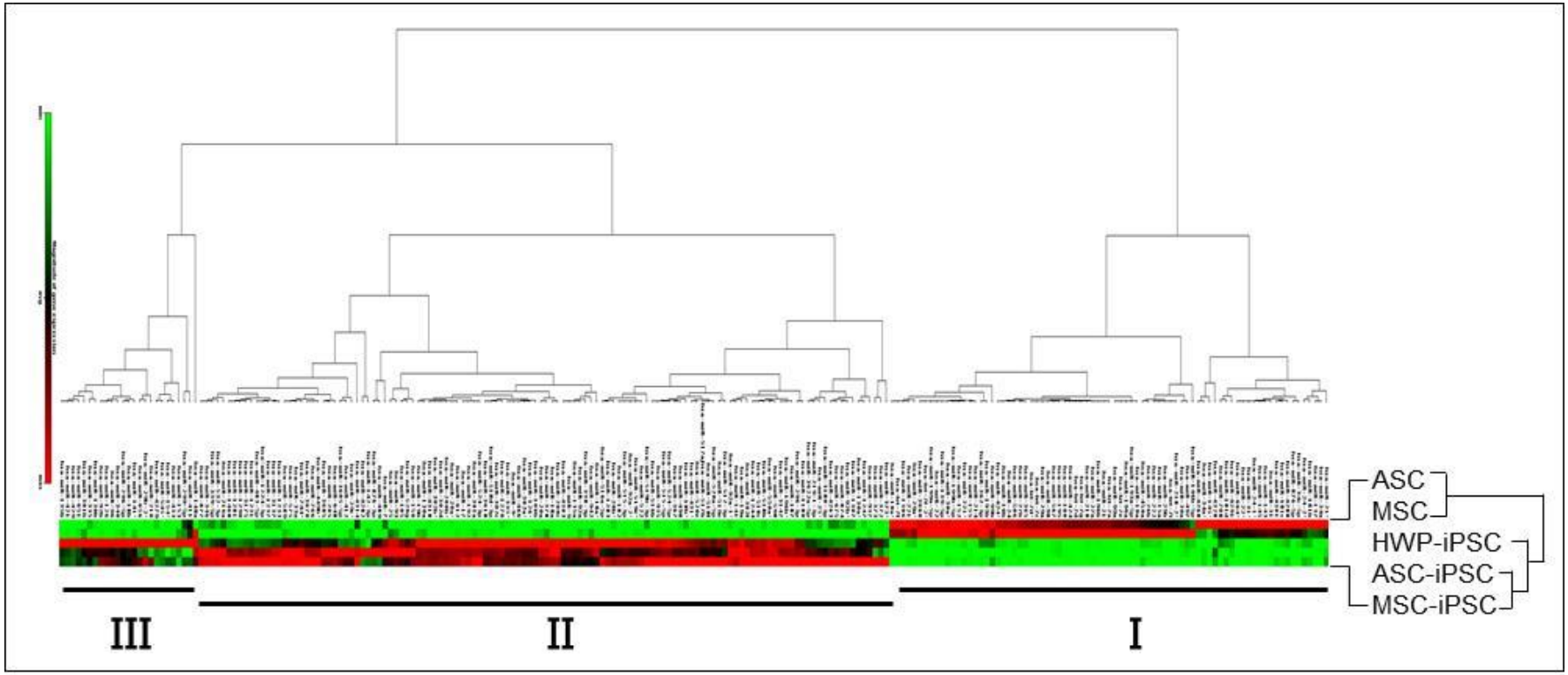


Figure 4.2. Hierarchical clustering analysis of miRNA profiles of MSC and iPSC. Level of gene expression is shown in the colour code shown at the bottom, ranging from minimal expression levels in green, average or weak miRNA expression in black and maximal levels in red. The analysis was performed using miScript miRNA PCR Array Data Analysis Web Portal.

levels in the other two MSC-derived iPSC lines, indicating iPSC derived from various sources are not entirely identical.

Pairwise comparisons between iPSC and ESC, and iPSC and MSC were next performed by using the cut-off threshold of $\log_2$(fold change) ≥ 1.5 or ≤ -1.5, and also the criteria that valid data were available for all the cell lines under consideration with statistical significance ($p < 0.05$). Furthermore, it was defined by the platform that a miRNA that had a threshold value of $C_t \geq 30$ was considered undetectable in expression level. Hence, in the iPSC-MSC comparison, a miRNA was considered activated in iPSC when this miRNA was in the detectable range in iPSC, but was undetectable ($C_t \geq 30$) in MSC. Likewise, a miRNA was considered shutdown in iPSC when the miRNA was detectable in MSC but was undetectable in iPSC. On the other hand, up- or down-regulated expression was used to describe increased or decreased detectable miRNA levels in pairwise comparison.

When the miRNA expression data of the two iPSC lines were first compared pairwise with their respective parental MSC lines, 441- 445 (42.6-43.0%) miRNAs were found to have altered in expression levels as defined (Table 4.2). On reprogramming of the unipotent HWP, 494 miRNAs (47.7%) were differentially expressed (Table 4.2). If the data were considered collectively under the criteria defined above, 261 miRNAs (25.2%) were found to be differentially expressed in the three iPSC relative to the two MSC lines (Tables 4.3; Appendix A) clearly indicating extensive changes in the

Table 4.2 miRNAs altered in expression levels in pairwise comparison between MSC/HWP and the derived iPS cells

Stem cell line			No. differentially expressed miRNA (% of miRNA analyzed)
iPSC		MSC/HWP	
hMSC-AT-iPSC	vs	hMSC-AT	441 (42.6%)
ASC-iPSC	vs	ASC	445 (43.0%)
HWP-iPSC	vs	HWP	494 (47.7%)

Table 4.3 Differentially expressed miRNAs in stem cells

	iPSC vs ESC[1]	iPSC vs MSC[2]
Activated	12 (37.5%)	66 (25.3%)
Up-regulated	13 (40.6%)	111 (42.5%)
Subtotal:	25 (78.1%)	177 (67.8%)
Shut-down	3 (9.4%)	17 (6.5%)
Down-regulated	4 (12.5%)	67 (25.7%)
Subtotal:	7 (21.9%)	84 (32.2%)
Total	32	261

[1]Data based on comparing 3 iPSC vs 2 ESC. [2]Data based on 3 iPSC vs 2 MSC; In both columns, data were $\log_2$(fold change) $\geq$1.5 or $\leq$-1.5.

miRNA profiles when MSC was reprogrammed to iPSC. On the other hand, when the miRNA expression data of the three iPSC lines were collectively compared with those of the two ESC lines, only 32 miRNAs (3.1%) were found to be differentially expressed (Table 4.3; Appendix A), consistent with ESC-like characteristics of iPSC.

In further iPSC-ESC pairwise comparison, 25 (78.1%) of the 32 differentially expressed miRNAs were found to be activated/up-regulated while 7 (21.9%) miRNAs were shutdown/down-regulated (Table 4.3). In MSC-iPSC comparison, 177 (67.8%) were activated/up-regulated and 84 (32.2%) miRNAs were shutdown/down-regulated (Table 4.3). The data indicated that in pluripotency, two-fold more miRNAs are activated/up-regulated than shutdown/down-regulated. The data indicated that on reprogramming to pluripotency, many more target genes are shutdown or down-regulated, which may be a significant event in rendering pluripotent stem cells the potential to differentiate into all cell types. The observation is consistent with the pluripotency of iPSC, and philosophically suggests that iPSC is life on hold, waiting for appropriate signals to release different sets of brakes to enter into differentiation into different tissues and organs.

4.3.2 *In silico* Validation of Differentially Expressed miRNA in iPSC on Reprogramming

The array of differentially expressed miRNAs when MSC/HWP was reprogrammed into iPSC was cross-checked with what was available in the literature (Table 4.4). In the first group of miRNAs targeting known reprogramming factors, miR-145, which was shown to modulate the Yamanaka factors, MYC, OCT4, KLF4 and SOX2 (Sachdeva et al., 2009; Xu et al., 2009), would be predicted to be down-regulated in iPSC on reprogramming. A down-regulated $\log_2$(fold change) of -6.2 was discerned on our analysis. Likewise, the let-7 family and miR-30a, which were shown to target LIN28A/B (Zhong et al., 2010), all 7 let-7 family members and miR-30a were down-regulated in our analysis. However, miR-9 was found to be up-regulated by 6.511 fold, which appeared to be inconsistent with the predicted down-regulation. To resolve the discrepancy, pairwise real-time RT-PCR was performed between the MSC/HWP and the derived iPSC lines (Table 4.5). In all the three pairs, up-regulated levels in iPSC were consistently obtained with a statistically significant mean up-regulated level of 3.14-fold, confirming the microarray data. It remains to be confirmed and investigated if miR-9 does modulate LIN28A/B to resolve the discrepancy. In the category of known reprogramming miRNAs (Table 4.4), all were found to be activated or up-regulated in our microarray datasets would be predicted, further supporting the validity of our results. There are also a group of miRNAs known to block reprogramming (Table 4.4, category III). With the

Table 4.4 *In silico* validation of WaferGen data on miRNAs that are differential expressed in iPSC relative to MSC/HWP

Reprogramming Factor / miRNA	Reference	WaferGen data (iPSC vs MSC)		
		Targeting miRNA	Expression	Log_2(fold change)
(I) <u>miRNAs targeting known reprogramming factors</u>				
MYC, OCT-4, KLF4, SOX2	Sachdeva et al. (2009) Xu et al. (2009)	miR-145	Down-regulated	-6.200
LIN28A/B	Zhong et al. (2010)	let-7 family (7 members)	All down-regulated	-9.539 to -7.492
		miR-9	Activated	6.511
		miR-30a	Down-regulated	-1.914
		miR-125	n.d.	--
(II) <u>Reprogramming miRNAs</u>				
miR-302 family (4 members)	Lin et al. (2008);	miR-302a, -302b	Activated	12.773, 14.438
miR-367	Kuo et al. (2012)	miR-302c, -302d, -367	All up-regulated	9.900 – 13.049
miR-302* species	This work	miR-302a*, -302b*, -302c*	All activated	8.611 – 11.068
miR-106a cluster	Li et al. (2011)	miR-106a, -18b, -19b, 92a	All up-regulated	3.286 – 7.711
miR-106b cluster		miR-106b, -93, -25	or activated	1.855 – 2.711
miR-17~92 cluster		miR-17, -18a, -19b, -92a		3.286 – 5.080
miR-106* species & miR-17~92* species	This work	8 star species	All up-regulated	1.876 – 3.618

Table 4.4 (Cont'd)

Reprogramming Factor / miRNA	Reference	WaferGen data (iPSC vs MSC)		
		Targeting miRNA	Expression	Log_2(fold change)
miR-200c	Miyoshi et al. (2011)	miR-200c	Activated	10.701
miR-369		miR-369	n.d.	--
miR-302 cluster		miR-302 cluster	Up-regulated	See above
	(III)	**Reprogramming barrier miRNAs**		
miR-34 family	Choi et al. (2011)	miR-34a, -34c-5p	Down-regulated	-3.594 to -1.804
miR-34a*	This work	miR-34a*	Down-regulated	-2.267
let-7 family	Viswanathan et al. (2008)	let-7a, -7c, -7d, -7e, 7f, -7f, -7i	All down-regulated	-9.539 to -7.492
let-7d*, let-7f-2*	This work	let-7d*	Shut-down	-6.404
		let-7f-2*	Down-regulated	-5.640
miR-143 & miR-145	Suzuki et al. (2009)	miR-143	Shut-down	-9.474
		miR-145	Down-regulated	-6.200
miR-145*	This work	miR-145*	Down-regulated	-2.449
miR-134, -296, -470	Tay et al., (2008)	miR-134	n.d.	--
		miR-296-5P	Up-regulated	3.081
		miR-296-3P	Activated	7.895
		miR-470	n.d.	--

n.d., not determined.

Table 4.5 Validation of miR-9 up-regulation in iPSC relative to MSC/HWP

MSC/HWP	vs	iPSC	Log_2(fold change)
ASC	vs	ASC-iPSC	2.73 ± 0.34
hMSC-AT	vs	AT-iPSC	2.96 ± 0.34
HWP	vs	HWP-iPSC	3.72 ± 0.79
Mean			3.14 ± 0.52 (p<0.01)

Data were obtained by real-time PCR analysis in three independent experiments.

exceptions of miR-134, -296 and -470, which were all suggested by Tay et al. (2008) to target the coding sequences of OCT4 and SOX2 transcripts, all other known reprogramming barrier miRNAs were down-regulated or shutdown on reprogramming of MSC/HWP to iPSC in this work. In summary, the microarray data were largely supported by published reports, supporting the robustness of the microarray platform used.

4.3.3 Co-expression of miRNA-5p/3p Pairs in Stem Cells

Recent miRNA studies are beginning to document frequent co-expression of both the miRNA and the miRNA* strands derived from the 5'- and 3' arms of the pre-miRNA duplex (Almeida et al., 2012; Shan et al., 2013; Choo, Soon, et al., 2014). When available in the microarray dataset in our analysis, the miRNA* species were identified and are included in this study (Appendix B). It is interesting to note that the identified miRNA* species were co-up- or co-down-regulated with their sister strand (Table 4.6 and Appendix B). miRBase has recently retired the human miRNA/miRNA* nomenclature but advised the use of miRNA-5p and -3p nomenclature based on derivation from the 5'- or 3'-arm of the pre-miRNA precursor. In the subsequent sections in this study, the 5p/3p nomenclature is used. Out of the 32 miRNAs that were differentially expressed in iPSC relative to ESC, there were only three (9%) 5p/3p pairs whereas 88 miRNAs out of 261 differentially expressed miRNAs (44 pairs, 33.7%) were co-expressed in 5p/3p when iPSC vs MSC data were compared (Table 4.6). On further examination, 26 (59%) 5p/3p miRNA pairs

Table 4.6 Number of co-regulated 5p-3p miRNA pairs in iPSC relative to ESC and MSC

	iPSC vs ESC[1]	**iPSC vs MSC[2]**
Co-up-regulated/-activated pairs	2 (66.7%)	26 (59%)
Co-down-regulated/-shutdown pairs	-	14 (32%)
Pairs with reverse regulation	1 (33.3%)	4 (9%)
Total	**3**	**44**

[1]Data based on comparing 3 iPSC vs 2 ESC. [2]Data based on 3 iPSC vs 2 MSC; In both columns, data were log2(fold change) ≥1.5 or ≤-1.5.

were co-up-regulated/co-activated and 14 (32%) pairs were co-down-regulated/co-shutdown. Interestingly, four pairs (9%) showed reverse directions of 5p/3p co-expression (Table 4.6), suggesting possible diversity in the biological functions of the 5p and 3p miRNA species. The co-regulation of both the 5p and 3p species of randomly selected pairs were confirmed in stem-loop RT-PCR (Figure 4.3 and Appendix I). Taken together, the data showed frequent (33.7%) co-expression of 5p/3p miRNAs in iPSC on reprogramming, and that the majority (91.9%) of the co-expressed 5p/3p pairs was co-up- or co-down- regulated in the same direction strongly suggesting concerted regulation of miRNA sister pairs in the reprogramming process.

4.3.4 Selective Activation of C19MC miRNAs in Mesenchymal Stem Cells

As shown above (Table 4.3), the microarray results indicated that more miRNAs were activated/up-regulated than down-regulated in pluripotent stem cells, suggesting essential roles for the activated/up-regulated miRNAs in regulating stem cells properties including self-renewal and pluripotency. Therefore, the activated/up-regulated miRNAs were further analysed. The up-regulated miRNAs were first grouped according to miRNA family and chromosomal location (Appendix C). Interestingly, beside up-regulation of the well-characterised reprogramming miR-302 family, numerous miRNAs of the miR-515 family or the chromosome 19 miRNA cluster (C19MC), which is located on chromosome 19q13.41, were also found to be expressed *en bloc* in

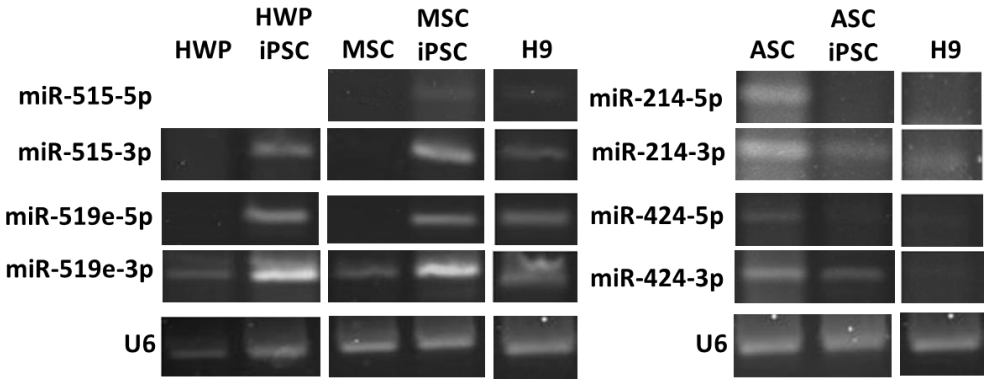


Figure 4.3 Verification of co-expression of miRNA 5p/3p pairs. The expression of co-expressed miRNA pairs was determined by stem-loop RT-PCR. PCR products were analysed in 4% agarose gels. The U6 snRNA was used as an internal control. HWP, human white pre-adipocyte; HWP-iPSC, HWP-derived induced pluripotent stem cell; ASC, adipose stem cell, ASC-iPSC: ASC derived iPSC; MSC, mesenchymal stem cell from adipose tissue; MSC-iPSC, MSC-derived iPSC.

pluripotent stem cells (Appendix C). Moreover, 12 out of the 26 co-up-regulated miRNA pairs were found to belong to C19MC cluster (Table 4.7). The observation suggested crucial roles of C19MC during reprogramming. Hence, expression of C19MC warranted further investigation.

All the forty-five C19MC miRNAs included in the microarray in either the 5p or 3p or in both 5p/3p configurations were expressed, albeit to different extents, in all the three pluripotent iPSC cell lines tested (Table 4.7), and in the hESC controls (data not shown). The miR-372 family that lies adjacent to the C19MC cluster (Figure 4.4A) was also included in the analysis since they have been reported to be expressed in pluripotent stem cells (Subramanyam et al., 2011). Of the forty-five C19MC miRNAs, thirty-nine were significantly ($p < 0.05$) expressed, as previously reported (Razak et al., 2013). Expression of the C19MC miRNAs in the iPSCs was generally two-fold or greater than that in the parental cell lines; the highest level of expression was 8.375 log2(fold change) in miR-520b (Table 4.7). Notably, both the 5p and 3p miRNA species were expressed in most cases; otherwise, the 3p species was the favoured precursor arm selected for the mature miRNAs, as opposed to frequent 5p arm expression in most other miRNA genes (Meijer et al., 2014).

On the other hand, only selected C19MC miRNAs were found to be expressed in MSC and HWP (Tables 4.7 and 4.8). Many of the expressed miRNAs share the "AAGUGC" seed sequence of the known reprogramming miR-302 miRNA family; these miRNAs are called the C19MC-AAGUGC-miRNAs in this study (see Figure 4.6A and depiction below). Twenty-two

Table 4.7 Expression of C19MC and the miR-372 family miRNAs in stem cells

miRNA[1]	Gene copy	Expression[2]			Log_2(fold change)
		HWP	MSC	iPSC	
miR-512-5p	2	-	+/-	+	5.265±0.58**
miR-512-3p		+	+	+	3.600±1.85*
miR-1323 (5p)	1	-	-	+	7.319±0.50**
miR-498 (5p)	1	+	+	+	3.857±0.99**
miR-520e (3p)	1	-	-	+	4.001±0.20**
miR-515-5p	2	-	-	+	7.053±0.63**
miR-515-3p		-	-	+	4.083±0.45**
miR-519e-5p	1	-	-	+	3.046±0.55**
miR-519e-3p		-	-	+	3.320±0.44**
miR-520f-3p	1	-	-	+	6.275±0.75**
miR-519c-3p	1	-	-	+	5.685±0.66**
miR-1283 (5p)	2	-	-	+	3.055±0.29**
miR-520a-5p	1	-	-	+	4.863±0.71**
miR-520a-3p		+	+	+	0.869±1.48
miR-526b-5p	1	-	+	+	4.233±0.13*
miR-526b-3p		-	+/-	+	6.906±0.81**
miR-519b-3p	1	-	+	+	2.620±1.36*
miR-518f-5p	1	-	+/-	+	3.901±0.50**
miR-518f-3p		-	+/-	+	6.971±0.59**
miR-520b (3p)	1	-	-	+	8.375±0.54**
miR-518b (3p)	1	+	+	+	3.706±0.98*
miR-526a (5p)	2	-	+	+	2.741±1.95
miR-520c-3p	1	-	-	+	8.285±0.70**
miR-518c-5p	1	n.d.	-	+	4.135±0.34
miR-518c-3p		-	-	+	6.282±0.76**
miR-524-5p	1	-	-	+	3.810±0.60**
miR-524-3p		-	-	+	4.202±0.63**
miR-517a-5p	1	-	+/-	+	2.328±0.65*
miR-517a-3p		-	-	+	8.262±0.62**
miR-519d-3p	1	-	+/-	+	4.819±2.56*

Table 4.7 (Cont'd)

miRNA[1]	Gene copy	Expression[2]			Log_2(fold change)
		HWP	MSC	iPSC	
miR-521 (3p)	2	+	+	+	1.500±0.30*
miR-520d-5p	1	-	-	+	3.933±0.34**
miR-520d-3p		n.d.	+	+	0.233±3.93
miR-517b-3p	1	-	+/-	+	6.474±0.73**
miR-520g-3p	1	+	+	+	6.014±1.27**
miR-516b-5p	2	-	-	+	3.243±0.52
miR-518e-5p	1	-	+	+	3.344±1.88
miR-518e-3p		+	+	+	1.386±0.49*
miR-518a-3p	2	+	+	+	2.013±1.04*
miR-518d-3p	1	-	-	+	3.613±0.47**
miR520h (3p)	1	-	+	+	5.618±0.82**
miR-522-3p	1	-	-	+	4.060±0.47**
miR-519a-3p	2	-	-	+	6.586±0.56**
miR-516a-5p	2	-	-	+	3.583±0.32**
miR-516a-3p		-	+/-	+	0.758±0.52*
miR-371a-5p	1	-	-	+	6.256±1.14**
miR-371a-3p		+	+/-	+	1.094±5.34
miR-372-3p	1	+	+	+	7.201±0.11**
miR-373-5p	1	-	-	+	2.500±1.27*
miR-373-3p		-	-	+	7.240±1.39**

[1]miRNA-5p and -3p designations are based on miRBase ver. 21; -5p and -3p designations in brackets are not annotated in miRBase, but are the presumptive precursor arms derived from sequence alignment. miRNAs are arranged in order of relative physical locations on chromosomae 19.13q.41; the neighbouring miR-371-3 cluster is also shown. [2]The two MSC cell lines were used in comparison with the three iPSC lines derived. "+" and "-" indicate detectable and undetectable expression of the miRNA, respectively, in both cell lines; "+/-" indicates that one of the two MSC was positive and the other one was negative. n.d., not done.

Table 4.8 Expression of C19MC miRNAs in different stem cell types

Stem cell type	Potency	miRNA[1]		. Number of miRNAs
		AAGUGC seed sequence[2]	Others	
iPSC/hESC	Pluripotent	16 miRNAs	29 miRNAs	45
MSC	Multipotent			
One cell line (ASC or MSC-AT)		miR-519d-3p, miR-526b-3p	miR-512-5p, miR-516a-3p, miR-517a-5p, miR-517b-3p, miR-518f-5p, miR-518f-3p	8
Both cell lines (ASC and MSC-AT)		miR-**512-3p**, **miR-519b-3p**, miR-520a-3p, miR-520d-3p, **miR-520g-3p**, miR520h	miR-498, miR-518a-3p, miR-518b, miR-518e-5p, miR-518e-3p, miR-521, miR-526a, miR-526b-5p,	14
HWP	Unipotent	miR-520a-3p, **miR-520g-3p**, **miR-512-3p**,	miR-498, miR-518a-3p, miR-518b, miR-518e-3p, miR-521	8
All cell lines	Pluri-/multi-/ unipotent	miR-520a-3p, **miR-520g-3p**	miR-498, **miR-512-3p**, miR-518a-3p, miR-518b, miR-518e-3p, miR-521	8

[1]miRNAs in bold letters were used for further quantification as depicted in Figure 4.4 and 4.5. [2]AAGUGC seed sequence-containing miRNAs are taken from Figure 4.6

(48.9%) of the forty-five C19MC miRNAs were activated in one or both MSC cell lines. Only eight miRNAs were expressed in HWP, which were, interestingly, also all expressed in the two MSC and all pluripotent cells (Table 4.8). This may suggest that these eight miRNAs constitute the minimal miRNA set require for minimal potency in the unipotent HWP. Thus, the cluster-wide microarray results indicated selective activation of twenty-two C19MC miRNAs in multipotent mesenchymal stem cells.

It has been reported that C19MC miRNAs are not expressed in adult tissues except in tissues of the reproductive system (Liang et al., 2007). To obtain further supporting evidences on selective activation, expression of eight miRNAs spanning the C19MC cluster (Figure 4.4A), but with different genomic structures, was selected for further experimentally verification; amongst the selected miRNAs, miR-512-3p is transcribed by the two miR-512-1 and- 512-2 genes located at the 5'-end of the C19MC miRNA gene cluster; miR-520c-3p, -519b-3p and -520f-3p are single miRNA genes located between previously proposed exons; miR-524-5p and -517a-3p are two of three miRNA genes mapped on intron 18 and miR-520d-5p and -520g-3p are two of four miRNAs mapped on intron 20 (Figure 4.4A) (Bortolin-Cavaille et al., 2009). Verification was done in three other different MSC cell lines, namely the MSC cell line WJ0706 derived from the Wharton's Jelly (Choo, Tai, et al., 2014), and two other adipose-derived MSC cell lines, ASC-Inv and ASC Lonza (Figure 4.4B). In the experiments, two other adipose MSC-derived iPSCs, ASC-iPSC and MH#1, and two hESCs, H6 and H9, were

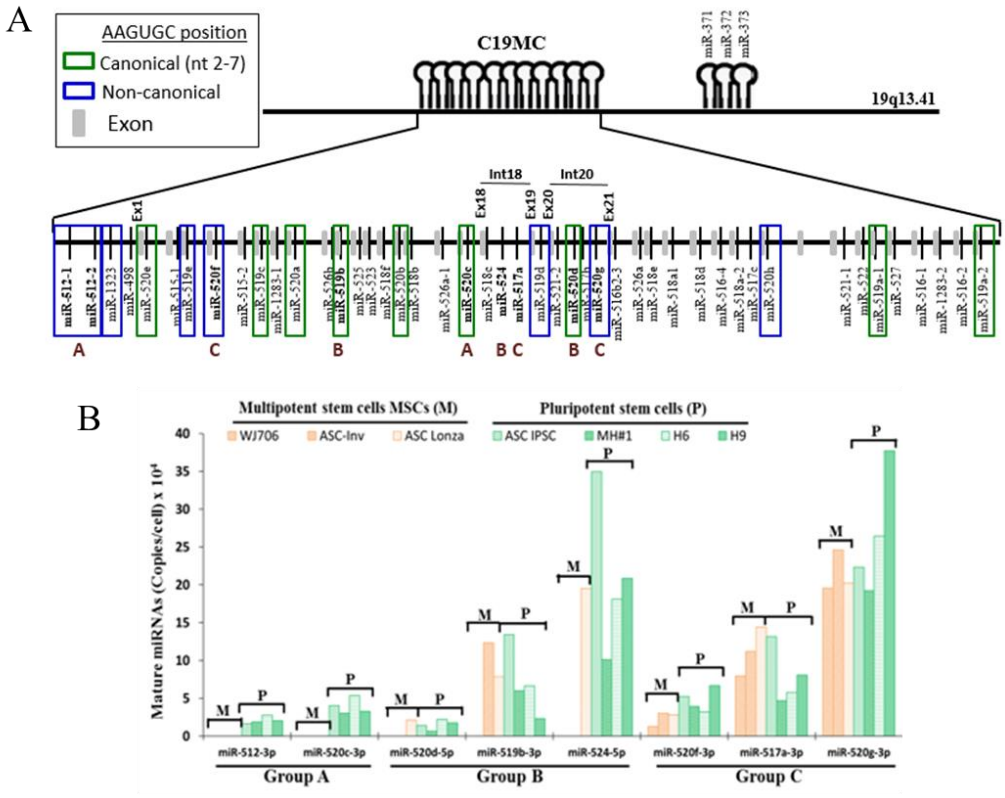


Figure 4.4 Expression of selected C19MC miRNAs in different stem cell lines. A. A scheme displaying relative genomic locations of the C19MC and the miR-372 family miRNAs on human chromosome 19q13.41. MiRNAs in green and blue boxes harbor the AAGUGC seed sequence in the canonical (nts 2-7) or non-canonical positions, respectively (see Figure. 4.6A). The proposed exon sequences (Ex) of the clusters (Bortolin-Cavaille et al., 2009) are shown in short grey bars between the miRNAs; introns (Int) 18 and 20 shown carry two of multiple miRNA genes analysed. The eight C19MC miRNAs selected for expression analysis in (B) below are shown in bold, with the expression A-C grouping designations established in (B) shown at the bottom of the miRNAs. **B**. Expression of selected C19MC miRNAs, determined based on copy number per cell, in mesenchymal stem cells (MSCs). The MSCs included are: WJ0706, ASC-INV and ASC. ASC IPSC and MH#1 are iPSCs derived from ASC-INV and ASC Lonza, respectively. H6 and H9 are human ESC cell lines. Ct values ≥ 35 was used as the cut-off threshold in the analysis.

included. The miRNA expression levels were determined as the absolute miRNA transcript copy number per cell, which ranged from 0 copy, at a real-time RT-PCR Ct value ≥ 35 (see materials and methods), to 377,200 copies per cell at a Ct value of 25.7 in miR-520g-3p in ESC H9 cells (Figure 4.4B).

Consistent with the miRNA microarray results, the selected miRNAs were all expressed to different levels in all four iPSCs and ESCs (Figure 4.4B). In contrast, the tested miRNAs were either not expressed, or expressed to different but lower levels in the MSCs tested. MSC expression of the eight C19MC miRNAs could be grouped in three expression patterns: group A, which included miR-512-3p and -520c-3p, showed very low or undetectable expression in the MSCs; expression of the group B miR-520d-5p, 519b-3p and -524-5p was detected in at least one or both MSC cell lines, whereas miR-520f-3p, -517a-3p and -520g-3p in group C were all expressed in all three MSC cell lines (Figure 4.4B and Table 4.8). The collective results obtained from the microarray and real-time RT-PCR experiments, therefore, confirmed selective C19MC expression in multipotent MSCs, and *en bloc* expression in pluripotent iPSCs. Furthermore, there seemed to be no correlation between the expression pattern and the physical location of the miRNA genes tested (Figure 4.4A). Notably, the miR-524-5p and -517a-3p and the miR-520d-5p and -520g-3p couples are flanked by two proposed exons but belong to different expression groups B and C (Figure 4.4A). The data suggest regulation by different promoters or transcriptional regulatory mechanism(s) other than simple splicing of the two flanking exon and co-processing of the spliced intron sequence as previously proposed for C19MC expression in a

choriocarcinoma JEG-3 cell line (Bortolin-Cavaille et al., 2009). The observation further suggests a critical biological role of the expressed C19MC miRNA in conferring different degrees of stemness to the stem cells, particularly in MSCs.

4.3.5 Selective Activation of C19MC miRNAs in Cancer Cells

Previous reports have indicated frequent activation of C19MC miRNAs in different cancer types, including colorectal cancer, breast cancer and primitive neuroectodermal brain tumour (Kleinman et al., 2014; Ma et al., 2016; Ren and Wang, 2016) (see below). To investigate C19MC expression in cancer cells, the expression of the same set of eight C19MC miRNAs was also quantified as gene copy number per cell in two colorectal cancer (HCT15 and SKCO1), one breast cancer (MCF-7) and one hepatocellular carcinoma (HepG2) cell lines; the choriocarcinoma (JEG-3) cell line, which was derived from the reproductive system, was included a positive control since JEG-3 cells have been shown to express all C19MC miRNAs in high levels (Noguer-dance et al., 2010) (Figure 4.5). Two cell lines CRL-1790 and HS799. PI, derived from normal colon and placenta tissues, respectively, were also included in the analysis. Despite *en bloc* and high-level C19MC expression in JEG-3 cells, only four of the eight miRNAs, namely miR-520d-5p of Group B as defined above for stem cell expression, and all three Group C miRNAs, miR-520f-3p, -517a-3p and -520g-3p, were shown to be expressed in the normal placenta cell line Hs799.PI. Furthermore, expression of the Group

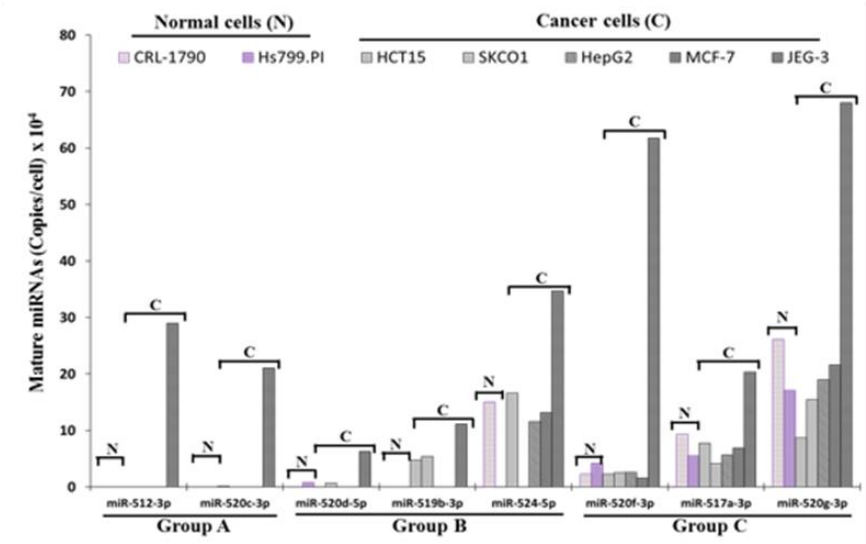


Figure 4.5 C19MC miRNA expression in cancer cell lines. The eight C19MC miRNAs selected for expression analysis in (Figure 4.4B) above were used. The expression A-C grouping designations established shown at the bottom of the miRNAs. The cell lines used are: CRL-1790: normal colon cells, Hs799. PI: normal placenta cells, HCT15 and SK-CO-1: colorectal cancer cells; HepG2: hepatocellular carcinoma cells, MCF7: breast cancer cells; JEG3: choriocarcinoma cells. Ct values ≥ 35 was used as the cut-off threshold in the analysis.

miR-524-5p, and all three Group C miRNAs was detected in CRL-1790, which was derived from normal fetal colon epithelium (Figure 4.5). The observed expression of selective C19MC in fetal colon epithelium and in the placenta is consistent with previous conclusions that C19MC is specifically expressed in reproduction and developmental process related tissues and is silenced in normal tissues (Razak et al., 2013; Liang et al., 2007; Lin et al., 2010). Interestingly, in the five cancer cell lines examined, the selective expression patterns of the eight miRNAs was similar to those shown in MSCs above (Figure 4.4B). Group A miRNAs also showed very low or undetectable expression in normal and cancer cells, except in JEG-3, whereas the Group B miRNAs were detected in one or more cancer cell lines; all three Group C miRNAs were expressed all four cancer cell lines (Figure 4.5). Taken together, quantitative expression analysis showed highly similar C19MC miRNA expression profiles found in MSCs and cancer cells, suggesting that the C19MC miRNAs may share some similar molecular and biological features in transcriptional regulation and in the etiological pathways in maintaining multipotency and cancer phenotype.

4.3.6 Identification of C19MC miRNAs Harbouring the "AAGUGC" Seed Sequence

miRNA-mRNA interactions involve the seed region at the 5' end of the miRNA; hence, seed sequences are important predictors for the identification of miRNA targeted transcripts (Bartel, 2009). MiRNAs that

share a common seed sequence also might share target specificity and possibly biological functions. On sequence alignment, sixteen C19MC miRNAs were found to share the same seed sequence, 5'-AAGUGC-3', with the reported reprogramming-able miR-302 and miR-372 miRNA families (Anokye-Danso et al., 2011; Subramanyam et al., 2011) (Figure 4.6A). These miRNAs are designated as "C19MC-AAGUGC-miRNAs". Furthermore, it is noted that the AAGUGC seed position at 5' end is variable among the C19MC-AAGUGC-miRNAs: subgroup I miRNAs, which includes eight miR-519 and -520 subfamilies, have the seed sequence located at the canonical and optimal 5'-nucleotide positions (nts) 2-7, as in the miR-302/-372 families; the seed sequence of the four subgroup IIa miRNAs is at location nts 1-6, and that of the remaining subgroup IIb miRNAs is at nts 3-8 and 4-9 (Figure 4.6A). Hence, despite the presence of the AAGUGC seed sequence, it is more likely that the nts 2-7 canonical subgroup of the C19MC-AAGUGC-miRNAs may target genes that share similar functions as the miR-302/-372 miRNAs.

While the 5p arm of a pre-miRNA precursor is normally selected for maturation (Meijer et al., 2014), it is noted that the C19MC-AAGUGC-miRNAs are predominantly derived from the 3p arm of the precursor miRNAs, hinting at an evolutionary bias in 3p selection with possible biological implications. Further supporting evidence of conservation of the C19MC-AAGUGC-miRNAs was derived from the construction of a phylogenetic tree of all precursor sequences of the C19MC miRNAs (Figure 4.6B). Most C19MC-AAGUGC-miRNAs are grouped into the same cluster in the top half of the phylogenetic tree. Four of the remaining C19MC-

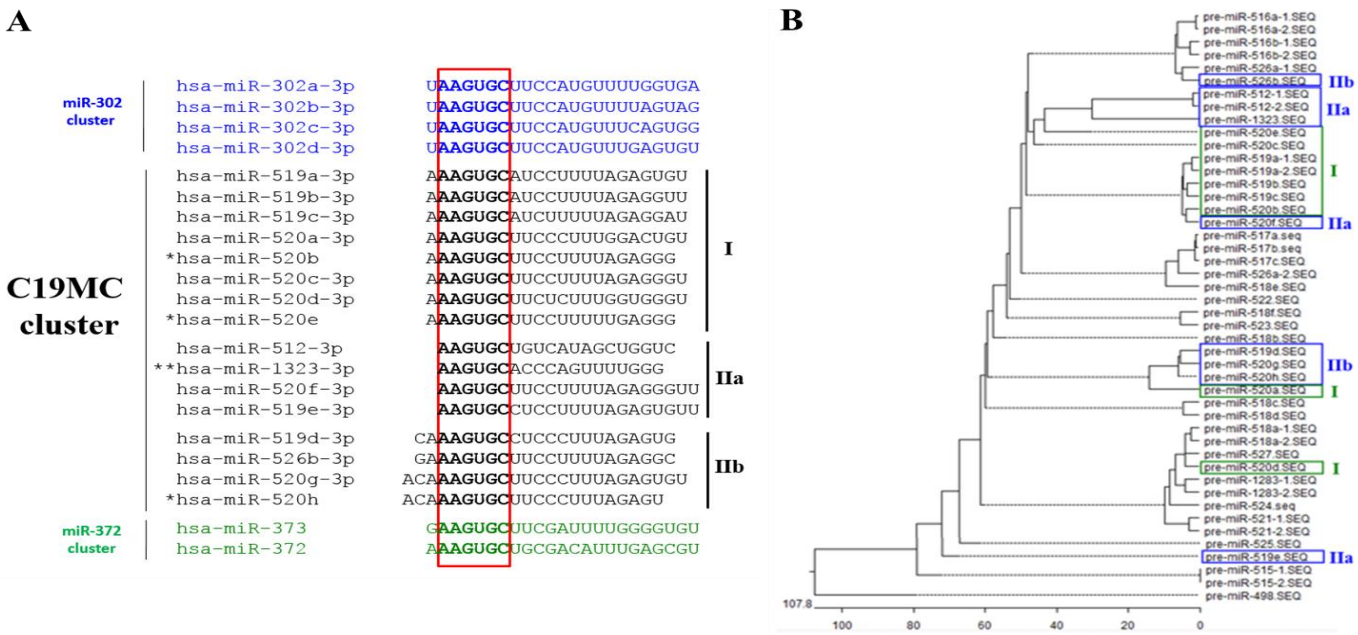


Figure 4.6. C19MC miRNAs harbouring the AAGUGC hexameric sequence. A. The sixteen C19MC miRNAs that share the AAGUGC hexameric seed sequence (in bold letters and boxed in red) with the miR-302 (in blue letters) and miR-372 (in green letters) families are shown. MiRNAs that have the AAGUGC seed sequence in the canonical nucleotides 2-7 position are called group I; other miRNAs in non-canonical position are called group II, with IIa and IIb subgroups as depicted. *miRNAs are in the 3p configuration as in the miRBase ver 21. **The existence of miR-1323-3p is based on computational prediction. **B.** Phylogenic tree of all C19MC miRNAs reconstructed with the precursors of the miRNAs. Groups I, IIa and IIb are AAGUGC-harbouring miRNAs as defined in (B) above (see text for further description).

AAGUGC-miRNAs form another cluster in the middle of the tree and the remaining two miRNAs are scattered in different branches in the lower half of the tree (Figure 4.6B).

4.3.7 Bioinformatics Predictions of Possible Biological Functions of Group I C19MC-AAGUGC-miRNAs

It is noted that the C19MC-AAGUGC-miRNAs with the canonical nts 2-7 seed position, defined here as Group I (Figure 4.6), contributed more significantly in gene targeting. Thus, in this study, we focused on analysis of potential biological functions of C19MC-AAGUGC-miRNAs in group I. Bioinformatics searches showed a total of 2058 putative target genes targeted by group I C19MC-AAGUGC-miRNAs (Figure 4.7 and Appendix D). However, construction of a Venn diagram showed that only 262 putative target genes are common between the miR-519 and miR-520 subfamilies in group I, indicating that the miR-519 and -520 subfamilies target different sets of genes. The overlapping gene sets among miR-302/372 and the miR-519 and miR-520 subfamilies in group I were further compared (Figure 4.7). The results showed that 1185 putative shared genes were obtained between the miR-520 and -302/372 families (Figure 4.7, blue box and appendix D, purple and yellow words), suggesting that the miR-520 subfamily might share similar biological functions with the miR-302/372 family. The group I miR-519 subfamily also shares 262 putative target genes with the miR-302/-372 families, far fewer than the miR-520 subfamily (Figure 4.7, red box).

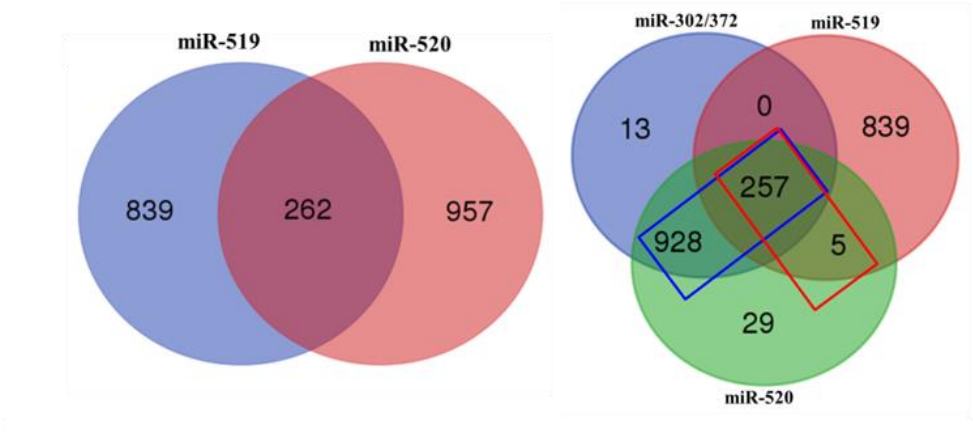


Figure 4.7 Venn diagrams of predicted target genes of the miR-302/372 families and group I of the C19MC-AAGUGC-miRNAs. (left panel) The miR-519 and -520 subfamilies share only a small number of target genes. (right panel) The miR-520 miRNAs share a significant number of target genes with the miR-302/372 families. Blue rectangle: putative genes shared between the miR-520 and -302/372 families. Red rectangle: putative genes shared between the miR-519 and -302/372 families.

Consistent with the bioinformatics prediction, a literature review showed that a number of validated targets have indeed been reported to be shared between the miR-302/372 and the group I C19MC-AAGCGU-miRNA families (Table 4.9).

The 2058 putative target genes were further subjected to GO analysis and KEGG pathway annotation (Figure 4.8A-C). Of the 828 predicted targets in the top 10 GO terms in biological functions, 616 (74.4%) putative genes are associated with transcriptional and translational regulation of gene expression (Figure 4.8A, GO terms 1, 2, 4, 5, 7 & 8). The remaining predicted targets regulate apoptosis, nervous system development, cellular response to DNA damage stimulus and cell cycle. The majority of the 2058 predicted genes in GO terms in molecular functions is likewise associated with transcriptional and translational regulation in some way (Figure 4.8B), and in epigenetic regulation (Figure 4.8B, GO term 9). Four hundred eleven genes (20.0%) are related to metal or zinc ion binding (Figure 4.8B, GO terms 2 & 4). which may also be components of signalling pathways. Taken together, the GO analysis data suggested that the group I C19MC-AAGUGC-miRNAs are mainly associated with the regulation of gene expression, cell proliferation and apoptosis via various signalling pathways.

The regulatory pathways were further annotated by interrogation of the KEGG database, which yielded 24 pathways, which included 568 genes in total (Appendix E); 14 of the 24 KEGG pathways which may be related to pluripotency and cancer are shown in Figure 4.8C. Ten of the 24 pathways,

Table 4.9 Common validated target genes shared between the C19MC-AAGUGC-miRNAs and the miR-302/-372 families

AAGUGC-miRNA		Seed position[a]	Target transcript	References
miR-302/-372	**C19MC**			
miR-302c	miR-520e	I	NIK	(S. Zhang et al., 2012; Gui et al., 2015)
miR-373	miR-520c	I	MT1-MMP, mTOR, SIRT1	(Lu et al., 2015; Liu and Wilson, 2012)
miR-372, -373	miR-520c, -520e	I	RelA	(Keklikoglou et al., 2012)
miR-302b, -372, -373	miR-520c, -520e	I	TGFβR2	(Subramanyam et al., 2011; Keklikoglou et al., 2012)
	miR-520b, -520e	I	CD46	(Cui et al., 2010)
miR-302c	miR-520c	I	MICA, MICB, ULBP2	(Min et al., 2013)
	miR-519a	I	RBL2	(Kleinman et al., 2014)
	miR-512	IIa		
	miR-519d, -520g	IIb	SMAD7	(Zhou et al., 2016; Kan et al., 2015)
	miR-520g, -520h	IIb	DAPK2	(Su et al., 2016; Zhang et al., 2016)
miR-302d, -372	miR-520b, -519b-3p, -520a-3p	I	CDKN1A	(Wang et al., 2008; Wu et al., 2010)
	miR-519e	IIa		
	miR-519d, -520h	IIb		

[a]Group I seed position is the canonical nts 2-7; IIa is nts 1-6 and IIb is other non-canonical position, as defined in Figure 4.6A.

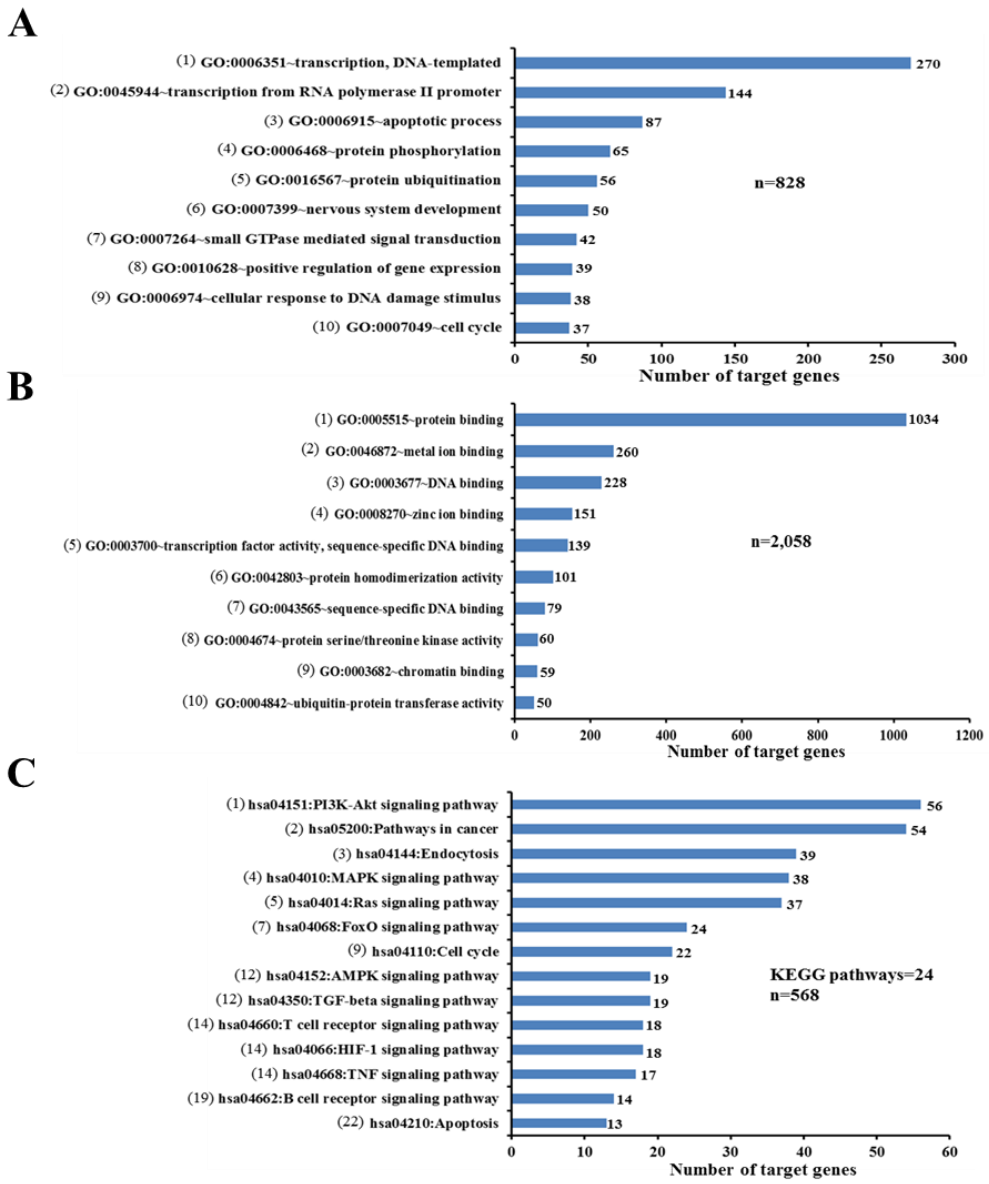


Figure 4.8 Bioinformatics analysis of predicted target genes of group I of the C19MC-AAGUGC-miRNAs. A & B. The top 10 highest scores and the most significantly enriched GO terms associated with biological process and molecular function, respectively. **C.** Fourteen of the 24 most enriched KEGG pathways displaying the ten signalling pathways identified; see text for explanation and Appendix E. The numerical in brackets shows the ranking of each pathway

which included 260 (45.8%) genes, are different signalling pathways that are known to be involved in the growth and development processes (Subramanyam et al., 2011; S. Zhang et al., 2012; Keklikoglou et al., 2012; Liu and Wilson, 2012; Zhang et al., 2011; Tsai et al., 2014; Brownlie and Zamoyska, 2013; Kurosaki et al., 2010; Danielsen et al., 2015). Notably, 129 (22.7%) genes are associated with pathways regulating apoptosis including PI3K-AKT, MAPK, HIF-1 and TNF (Figure 4.8C; see also Figure 4.9 and Discussion below). The highest enriched PI3K-Akt signalling pathway (56 genes) regulates cell survival by reducing apoptosis, stimulating cell growth and increasing proliferation (Danielsen et al., 2015). Furthermore, many of the genes are related to pathways that regulate the cell cycle (22 genes) and apoptosis (13 genes) (Figure 4.8C), which are important cellular events in the initiation and maintenance of stem cell pluripotency and tumorigenesis.

4.3.8 Possible Group I C19MC-AAGUGC-miRNAs Targeting of the Pro-apoptosis Functions in the Survival Pathway

Suppression of apoptosis is an important feature of the initiate phase of the reprogramming process (Smith et al., 2010). On the other hand, apoptosis dysregulation is associated with the different stages of tumorigenesis, including initiation, progression and metastasis (Fulda, 2009). A database search showed that the group I C19MC-AAGUGC-miRNAs target 179 apoptosis-associated genes (Appendix F). On the other hand, the KEGG pathway analysis above (Figure 4.8C) has also revealed that the highest

number of putative target genes of group I miRNAs are associated with PI3K-Akt, a survival pathway. Hence, we hypothesized that the group I miRNAs acted more specifically to inhibit apoptosis by targeting survival-related genes. Fifteen survival-related genes were predicted targets of the group I miRNAs (Table 4.10). Two out of the fifteen genes, viz. NIK and RelA, have been experimentally validated as direct targets miR-520e and miR-520c-3p (S. Zhang et al., 2012; Keklikoglou et al., 2012). Importantly, the group I miRNAs may promote apoptosis either by indirectly activating pro-apoptotic proteins BAK/BAX through suppression of the cell survival-related genes (Westphal et al., 2011), or by enhancing caspase-8 activation through targeting inhibitors of TRAIL-inducing apoptosis (Wang and El-Deiry, 2003; Garofalo and Croce, 2013). Taken together, the group I C19MC-AAGUGC-miRNAs were predicted by bioinformatics analysis to regulate apoptosis, which is important in the initial phase of cellular reprogramming, and in particular the cell survival pathways, which are directly relevant to tumorigenesis processes.

Table 4.10 Predicted group I C19MC-AAGUGC-miRNA target genes associated with cell survival pathways

Gene Symbol	Gene name
AKT1	AKT serine/threonine kinase 1
IGF1	Insulin-like growth factor 1 (somatomedin C)
IL2	Interleukin 2
KIT	KIT proto-oncogene receptor tyrosine kinase
MALT1	Mucosa associated lymphoid tissue lymphoma translocation gene 1
NIK/MAP3K14[1]	Mitogen-activated protein kinase kinase kinase 14
PIK3CA	Phosphoinositide-3-kinase, catalytic, alpha polypeptide
RELA[1]	V-rel reticuloendotheliosis viral oncogene homolog A (avian)
SOS1	SOS Ras/Rac guanine nucleotide exchange factor 1
TAK1/MAP3K7	Nuclear receptor subfamily 2 group C member 2
TLR4	Toll-like receptor 4
TNF/TNFα	Tumor necrosis factor (TNF superfamily, member 2)
TNFRSF10D/DcR2	Tumor necrosis factor receptor superfamily, member 10d, decoy with truncated death domain
TRAF6	TNF receptor-associated factor 6
TSP-1/THBS1	Thrombospondin 1

[1]Experimentally validated target genes (Keklikoglou et al., 2012; Zhang et al., 2012)

4.4 Discussion

4.4.1 Selective C19MC miRNA Expression in MSC and in Cancer Cells Suggests a Complex Transcriptional Regulatory Mechanism

In the present study and in the literature, data showed similar and disperse expression patterns of eight tested C19MC miRNAs in both mesenchymal stem and cancer cells (Figures 4.4 and 4.5), in contrast to the previous model of *en bloc* expression in the choriocarcinoma JEG-3 cell line regulated by a master promoter (Bortolin-Cavaille et al., 2009). Another studies have shown that the highly abundant Alu repetitive sequences embedded within the C19MC genomic region may function as independent RNA polymerase II promoters (Borchert et al., 2006; Saito et al., 2009). Our study clearly showed selective C19MC miRNA activation in MSCs and HWP, and in cancer cells, suggesting that C19MC transcripts are more likely regulated by multiple promoters, which may in turn be active by condition-specific transcription factors. Furthermore in cancer cells, chromosomal rearrangements, amplification and modification of the promoter(s) or specific transcription factors could further regulate the selective C19MC miRNA expression. Previous reports have, indeed, shown that translocation of chromosomal band 19q13.4 selectively activated C19MC miRNAs in thyroid adenomas, and that C19MC genomic amplifications in an aggressive primitive neuroectodermal brain tumours were associated with specific and abundant expression of miR-517c and -520g (Li et al., 2009). Moreover, epigenetic alterations in the C19MC genomic region may also play important role in

regulating C19MC expression, particularly in cancer and the dynamic stem cells. Promoter silencing of C19MC miRNAs by the DNA methylation inhibitor, 5-azacytidine, activated sixteen C19MC miRNAs (Tsai et al., 2009). Furthermore, placenta-derived mesenchymal stem cells were reported to escape epigenetic silencing of the paternal allele resulting in a number of the C19MC miRNAs being abundantly expressed (Inga and Bullerdiek, 2012). Specific activation of the C19MC miR-512-5p by histone deacetylase inhibitors was also reported in human gastric cancer cells (Saito et al., 2009). Transcription factors acting in trans are essential regulators of C19MC miRNA expression as shown by direct binding of p53 and the estrogen receptor α (ERα) to presumptive promoters of C19MC miR-519d and miR-515-5p, respectively, in chromatin immunoprecipitation assays (Fornari et al., 2012; Pinho et al., 2013). As a result of the specific DNA binding, miR-519d is up-regulated by p53, whereas ERα mediates both down- and up-regulated expression of miR-515-5p induced by estrogens and tamoxifen, respectively. Thus, transcription of specific C19MC miRNAs in MSC and cancer cells is likely highly complex, and may be dependent on the cellular and pathological state of the cells.

It was previously reported that C19MC is silenced in normal tissues (Liang et al., 2007; Lin et al., 2010) due to hypermethylation of both the paternal and maternal alleles (Noguer-dance et al., 2010). However, placenta is able to escape epigenetic silencing by maintaining paternal allele-expression (Noguer-dance et al., 2010; Tsai et al., 2009). Moreover, the expression of miR-498, a member of C19MC, was reported in the fetal brain (Zhang et al.,

2008), echoing our report of Group C miRNAs being expressed in a fetal colon epithelium-derived cell line, CRL-1790 and placental Hs799. PI (Figure 4.5), consistent with C19MC expression in reproductive and developmental process-related tissues, relevant to the primate specificity of the C19MC cluster.

In this study, we found that the 3p arms of the C19MC miRNA precursors were predominantly selected in ESCs and iPSCs (Table 4.8 and Figure 4.6). Several studies have previously demonstrated that preferred arm selection is temporal- and spatial-dependent (Jagadeeswaran et al., 2010; S. C. Li et al., 2012). Indeed, the 3p miRNA species have been shown to be more abundantly expressed in tumour tissues as opposed to preferred 5p selection in normal tissues (S. C. Li et al., 2012). Echoing these findings, the miR-302-like C19MC are also predominantly 3p-biased, possibly targeting genes which are biologically significant in regulating the stemness of stem cells and the tumour phenotype in cancers.

4.4.2 Structural and Function Significance of the Group I C19MC-AAGUGC-miRNAs

Our results also showed divergence in the positions of the AAGUGC seed sequence among C19MC miRNAs carrying the hexameric sequence. The canonical seed region situated at nucleotides 2-7 is a perfect seed match which markedly decreases the presence of false-positive bioinformatics predictions,

thus improving prediction reliability (Bartel, 2009). Furthermore, the canonical seed region is crucial and sufficient to trigger target silencing (Bartel, 2009). The hexamer of half of the C19MC-AAGUGC-miRNAs reported here are located at nts 2-7, designated as group I C19MC-AAGUGC-miRNAs in this report (Figure 4.6), suggesting high possibility that the predicted genes are the putative targets. Other non-canonical C19MC-AAGUGC-miRNAs are likely to have lower affinity and specificity and may be limited in mediating repression without the 3'-compensatory binding (Carroll et al., 2014).

The group I miRNAs is composed of the miR-519 and -520 subfamilies. Despite their similar seed location at nts 2-7, Venn diagram analysis shows that these two subfamilies share only a small number of putative target genes (Figure 4.7). Common prediction algorithms that use identical powerful prediction characteristics, such as the mandatory stringent seed base-pairing produce different prediction results properly due to usage of various UTR databases as well as different internal criteria (Akbari Moqadam et al., 2013). In this study, the putative target sets of the miR-519 and -520 subfamilies are overlapping gene sets predicted by two different prediction algorithms. Furthermore, one of the characteristics of target prediction, the sequence context surrounding the seed binding site of the target transcript (Bartel, 2009), between the miR-519 and -520 subfamilies are also dissimilar (Figure 4.6A). This may explain the different target gene sets of these two subfamilies. It has been reported that miR-96 and -182 that have identical seed region (UUGGCA, nucleotides 2-7) regulate different targets (Jalvy-Delvaille

et al., 2012). However, the miR-520 and miR-302/372 families share a significant number of target genes (Figure 4.7) suggesting common biological functions. Hence, it is highly likely that the group I miR-520 miRNAs may also contribute to reprogramming, as supported by the predicted involvement of miR-520 miRNAs in the reprogramming related apoptosis and cell proliferation pathways (see Figure 4.9 and Discussion below).

4.4.3 Regulation of C19MC miRNAs in Tumorigenesis and Stemness

Selective activation of C19MC miRNAs in MSCs and cancer cells reported here suggests functional involvement of the activated miRNAs in maintaining the stemness and promoting cancer development. Frequent aberrant C19MC miRNA expression in cancers has been reported (Kan et al., 2015; Li et al., 2009; Fornari et al., 2012). Activation of the C19MC miR-519d was shown to target CDKN1A/p21, PTEN, AKT3 and TIMP2, and is closely associated with the pathogenesis of hepatocellular carcinoma by promoting cell proliferation and invasion, and in inhibiting apoptosis (Fornari et al., 2012). In breast cancer, high expression levels of plasma miR-520g is correlated with patients with lymph node metastasis and mammary gland invasion, and suppressed p53 expression (Ren and Wang, 2016).

On the other hands, C19MC miRNAs have also been shown to play important role in cellular stemness state. In normal embryonic development, many C19MC miRNAs have been shown to be expressed only in

undifferentiated or germinal tissues, and C19MC expression inhibits differentiation of human embryonic stem cells (Razak et al., 2013; Noguer-dance et al., 2010; Ren et al., 2009). The observation that the cellular reprogramming-able transcription factors OCT4 and NANOG regulate C19MC miRNA expression in human embryonic stem cells further supports close association of C19MC with induced pluripotency (Bar et al., 2008). Moreover, the identification of sixteen miR-302-like C19MC miRNAs predicts functions in promoting "stemness" as the miR-302 and miR-372 families. Similarly, eight miR-302-like C19MC miRNAs were previously shown to promote cell proliferation and cell-cycle progression by targeting p21, an inhibitor of the G1/S transition, as for the miR-302 and -372 families (Wang et al., 2008; Wu et al., 2010).

4.4.4 Possible Involvement of Group I C19MC-AAGUGC-miRNAs in Regulating the Apoptosis Pathway Common to Stemness and Cancer Phenotype

Suppressed apoptosis is important to both the initial phase of acquiring pluripotency and in cancer progression (David and Polo, 2014; Portt et al., 2011). A combined expression profile and bioinformatics analysis reported in this work has, indeed, shown that the group I C19MC-AAGUGC-miRNAs, target genes related to the survival pathways (Table 4.10). Based on the predicted target genes, a scheme that correlates the group I C19MC-AAGUGC-miRNAs to stemness and cancer phenotype is proposed (Fig. 4.9).

In general, group I miRNAs may enhance apoptosis through the PIK3/ATK, TNFs/NF-κB and TRAIL pathways, as predicted by KEGG pathway analysis (Figure 4.9) (Wang and El-Deiry, 2003; Portt et al., 2011; Khandelwal et al., 2011; Schiaffino and Mammucari, 2011). The PIK3 pathway is activated by a wide range of extracellular signals, including cytokines, e.g. IL-2 (Fung et al., 2003), growth factors, e.g. IGF1 (Schiaffino and Mammucari, 2011) and components of the extracellular matrix (ECM) such as TSP-1 (Pallero et al., 2008), all of which are the predicted targets of the group I miRNAs (Figure. 4.9). It is proposed here that the miRNAs target and inactivate the PIK3/AKT3 pathway by inhibition of the PIK3-related upstream genes TSP-1, IL-2, IGF1, KIT, SOS1 and PIK3CA, and the downstream AKT1 gene. The second important mechanism of cell survival is tumour necrosis factors (TNFs) activation of anti-apoptotic proteins via the nuclear factors of kappa B (NF-κB) signalling cascade (Figure 4.9). Similar to the PIK3/ ATK pathway, group I C19MC-AAGUGC- miRNAs may enhance apoptosis by the predicted targeting of the TNFα, TLR4, TRAF6, TAK1, NIK, MALT1 and RelA genes. Thirdly, group I miRNAs are also predicted to silence genes, such as DcR2, that are inhibitory to the TRAIL-induced apoptosis pathway, resulting in pro-apoptosis (Wang and El-Deiry, 2003). The group I miRNAs-modulated pathways subsequently suppress the activation of downstream effector caspase-3, -6, and -7, thus inhibiting apoptosis and promoting proliferation (Indran et al., 2011).

4.5 Conclusions

In the present study, the data show selective expression of C19MC miRNAs in cancer and stem cells, offering insights into possible involvement of selective C19MC miRNAs in regulation of “stemness” and tumorigenesis, possibly via the cell survival pathways. More specifically, a subgroup of sixteen C19MC miRNAs has been identified that shares the same AAGUGC seed sequence as the reprogramming miR-302/372 family, predicting contribution of the C19MC-AAGUGC-miRNAs to the reprograming process. Further elucidation of the biological functions of C19MC miRNAs, particular the miRNA-302-like subclass, may lead to potential applications in more efficient cellular reprogramming and in cancer therapy.

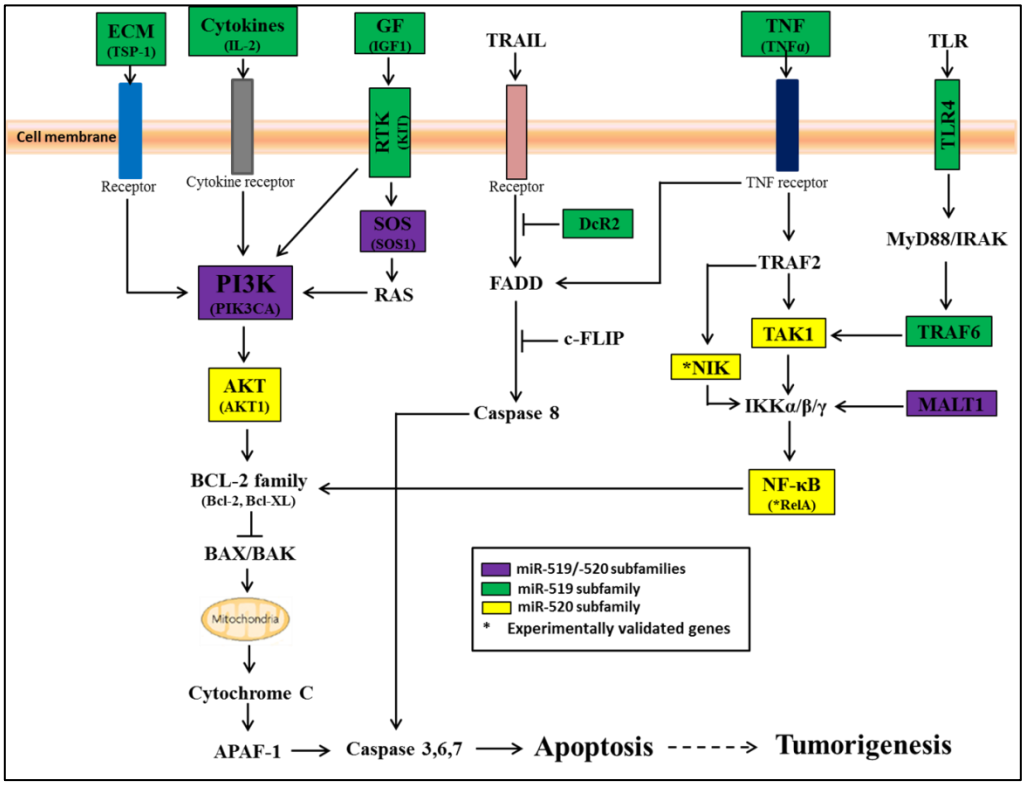


Figure 4.9 A proposed scheme that links the predicted group I C19MC-AAGUGC-miRNAs-targeted genes (in colour boxes) to cell survival functions and apoptosis pathways. Genes targeted by either or both the miR 519 or -520 subfamilies are shown in different colour boxes. See Discussion section for description of the proposed scheme

CHAPTER 5

RESULTS AND DISCUSSION (PART II)

MiR-524-5p of the Primate-Specific C19MC miRNA Cluster Targets TP53IPN1 and EMT-Associated Genes to Regulate Cellular Reprogramming [Manuscript submitted to Stem Cell Res Ther, Mar 2017]

5.1 Background

The cellular reprogramming process is thought to involve three phases, *viz.* initiation, maturation and stabilization, each of which is driven by a cascade of expression changes in specific set of genes to give rise to fully or partially reprogrammed cells (Buganim et al., 2013; David and Polo, 2014). Some important features of the early stage of reprogramming include increased proliferation, inhibition of apoptosis, acquisition of epithelial characteristics and up-regulation or activation of pluripotency-related genes (David and Polo, 2014). The molecular events that regulate each step of the reprogramming process are still being elucidated (Plath and Lowry, 2011).

Data in Chapter 4 have shown *en bloc* C19MC activation in pluripotent stem cells but expression of only selective C19MC miRNAs in multipotent

mesenchymal stem cells (MSCs) and a unipotent cell line (Nguyen et al., 2017). Bioinformatics analysis has further predicted that C19MC miRNAs may play a role in maintaining stem cell self-renewal and pluripotency by regulating the apoptosis and induced pluripotent- mediated signalling pathways (Nguyen et al., 2017). This chapter focuses on the study of the biological functions of a C19MC miRNA, viz. miR-524-5p, in the reprogramming process.

5.2 Study design

To elucidate possible contribution of a selected C19MC miRNA, viz. miR-524-5p, to the reprogramming process, the study design was as shown in Figure 5.1.

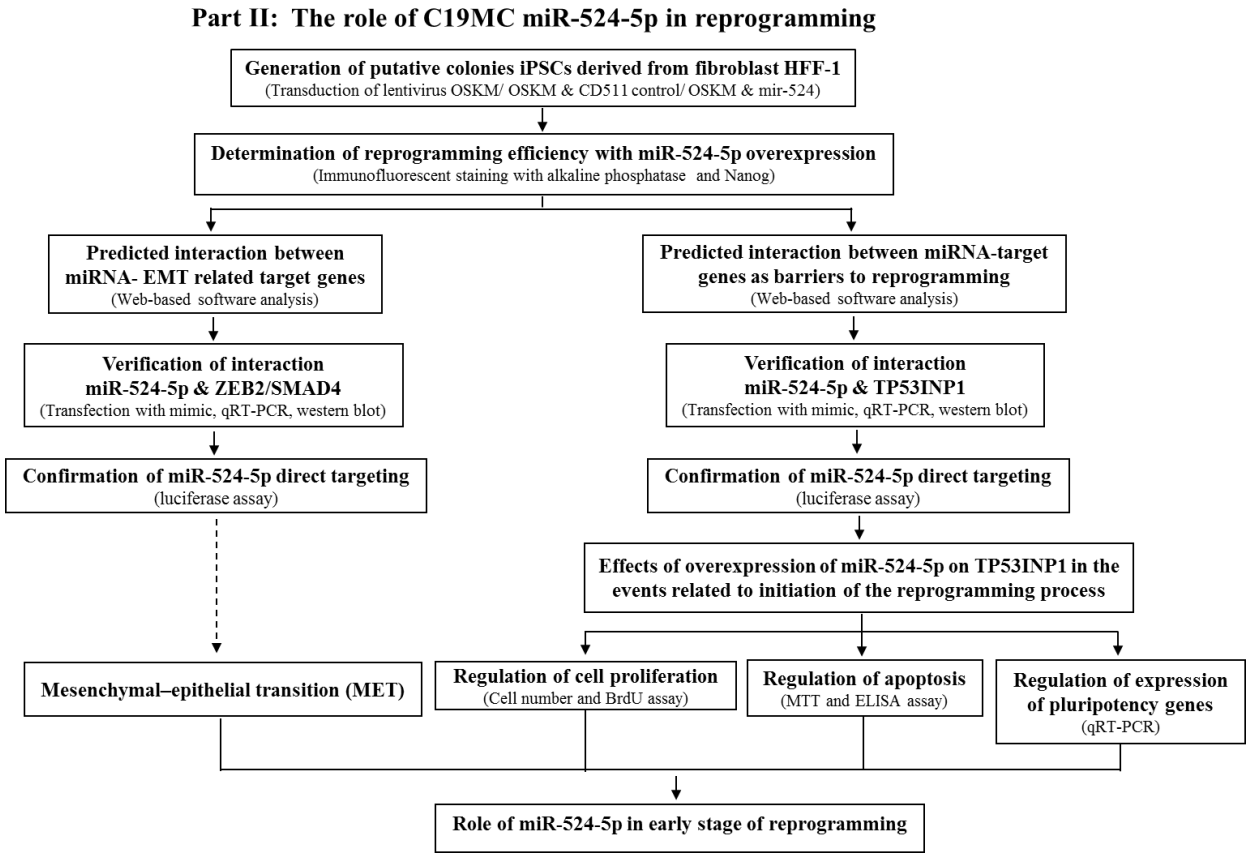


Figure 5.1 Study design in the elucidation of biological contribution of miR-524-5p to the reprogramming process. The assays carried out were described in round brackets.

5.3 Results

5.3.1 MiRNA-524-5p, but not MiR-512-3p, Promotes G1-S Transition

In the previous chapter, the C19MC miRNAs have been predicted to play important roles in the regulation of "stemness" and reprogramming process. In this chapter, the contribution of C19MC miRNAs in stem cells, particularly in iPSC induction, was further dissected. MiR-512-3p, the first miRNA in the C19MC cluster, belongs to the C19MC-AAGUGC-miRNAs that share the similar seed sequence with the miR-302 family (Figure 4.6A). Likewise, the C19MC miR-524-5p shares 19/20 nucleotides with miR-520d-5p (Figure 5.2), which was previously reported to convert cancer cells into iPSC-like cells (Ishihara et al., 2014; Tsuno et al., 2014). Hence, miR-512-3p and miR-524-5p may share biological roles in regulating self-renewal and pluripotency and are potentially good candidates to represent C19MC miRNAs for the study of the role of C19MC in mediating stemness.

An abbreviated G1 phase is an unique features of pluripotent stem cells (Wang et al., 2008; Ghule et al., 2011). MiRNAs have been shown to play essential roles in regulating the abbreviated G1 phase of stem cells by inhibiting several key regulators of the G1-S transition, resulting in enhancing the G1-to-S phase transition, thus, accelerating cell proliferation (Wang et al., 2008). Hence, in order to study the roles in reprogramming process, function of miRNA-512-3p and miR-524-5p in promoting G1-to-S phase transition was first investigated. Prior to examination of the effect of miR-512-3p and miR-

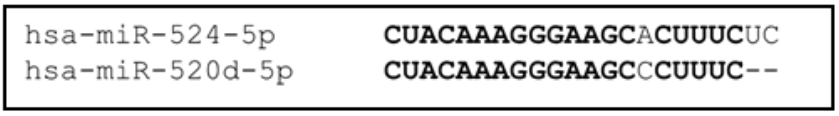


Figure 5.2 High degree of sequence homology (bold letters) between miR-524-5p and miR-520d-5p.

524-5p on G1-to-S phase transition, the conditions that mesenchymal stem cells WJ0706, derived from Wharton's Jelly, enter entirely the G2 phase were established (Figure 5.3). On treatment with nocodazole, 42% of WJ0706 cells were at G2 phase at 0 h (Figure 5.3A). The amount of G2-phase cells increased to 66% after 18 h of nocodazole (Figure 5.3B). The synchronized cells were then subjected to mitotic shake-off to obtain 100% cells accumulated at G2 phase (Figure 5.3C). It is observed that the synchronized cells started to enter the S phase 3 h after nocodazole withdrawal (Figure 5.4). In subsequent experiments in analysing the effects of miRNA, 3-h nocodazole withdrawal was used. To elucidate the involvement of miR-512-3p and miR-524-5p in regulating G1-S phase transition, WJ0706 was transfected with either miR-512-3p or miR-524-5p mimic to achieve miRNA over-expression. A negative control mimic (NC) was included in the experiment (Figure 5.5). Consistent to the mock, over-expression of miR-512-3p or miR-524-5p also resulted in increased number of cells stained positively with EdU 3 h after nocodazole withdrawal (Figure 5.5A). When miR-512-3p was over-expressed in the WJ0706, the number of EdU-positive cells observed remained unchanged compared with the mock or in the NC transfectant (Figure 5.5B). In contrast, introduction of miR-524-5p resulted in increased number of EdU-positive cells by approximately 1.5 fold compared with mock or NC (Figure 5.5B). The result was the first indication that miR-524-5p promoted the G1-to-S transition, suggesting that miR-524-5p plays an important role in mediating self-renewal and possibly pluripotency. Hence, in subsequent experiments, miR-524-5p was chosen as a representative C19MC miRNA to study the regulation during the reprogramming process.

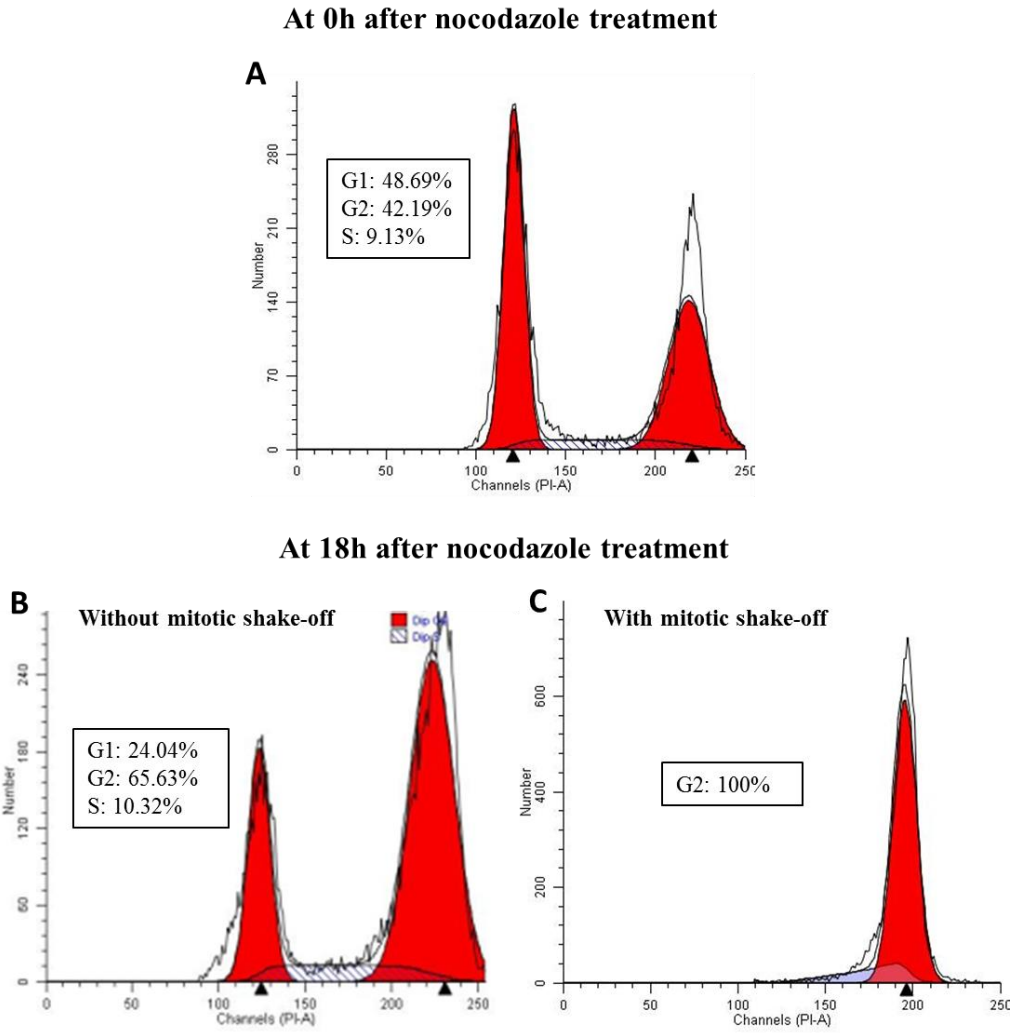


Figure 5.3 Analysis of cell cycle of MSC treated with nocodazole by flow cytometry. Mesenchymal stem cells WJ0706 were treated with nocodazole at a concentration of 150 ng/ml. The analysis of the cell cycle by flow cytometry of treated cells was performed at 0 h **(A)** or 18 h **(B & C)**. Synchronized cells that did not undergo mitotic shake-off **(B)** or mitotic shake-off **(C)** are shown.

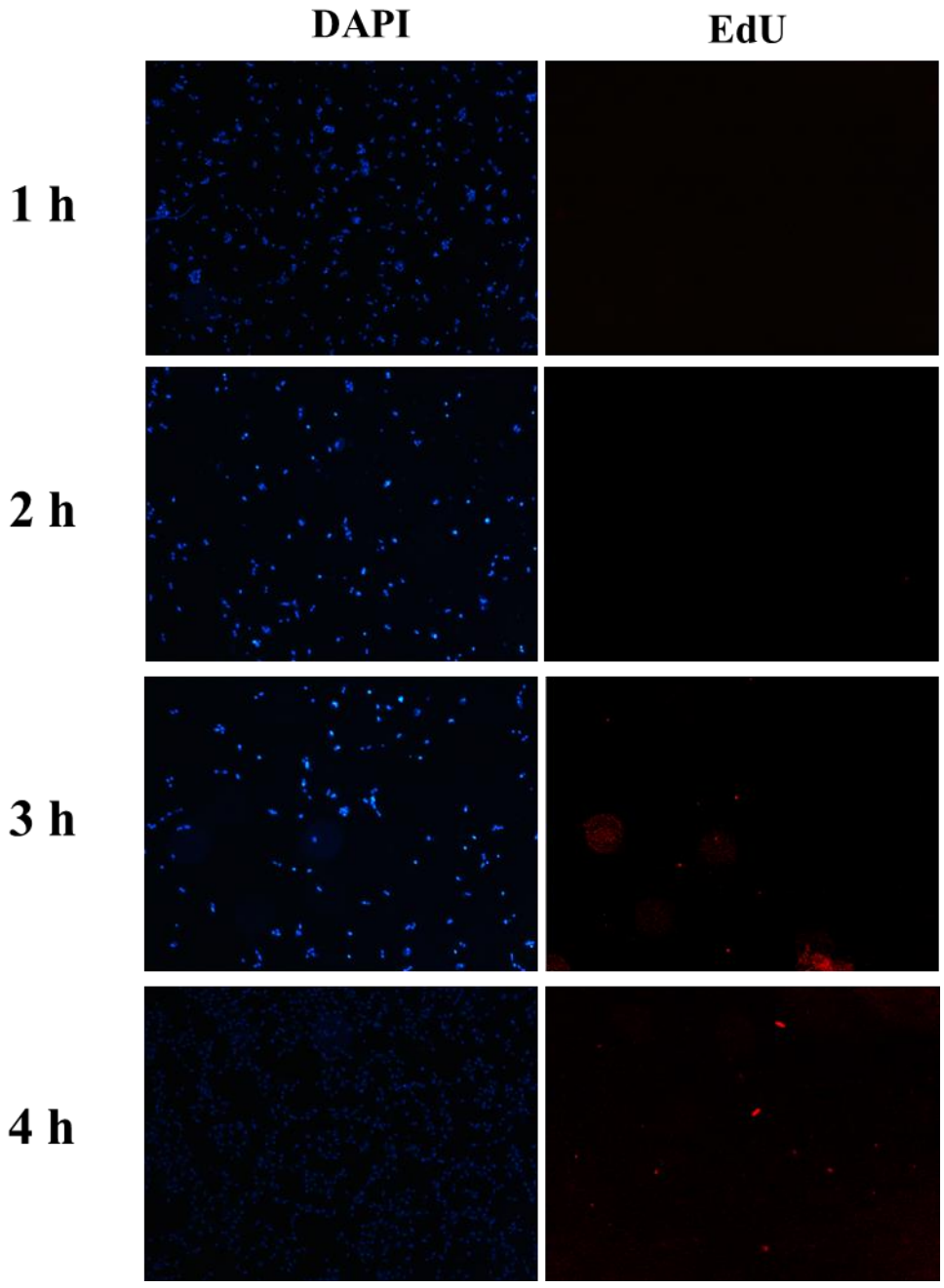


Figure 5.4 Duration of synchronized cells to enter the S phase. Immunofluorescence staining of 5-ethynyl-2´-deoxyuridine (EdU) uptake depicts nuclear staining DAPI (blue) and EdU (red) at 1 to 4 h after the synchronized cells re-entered cell cycle progression.

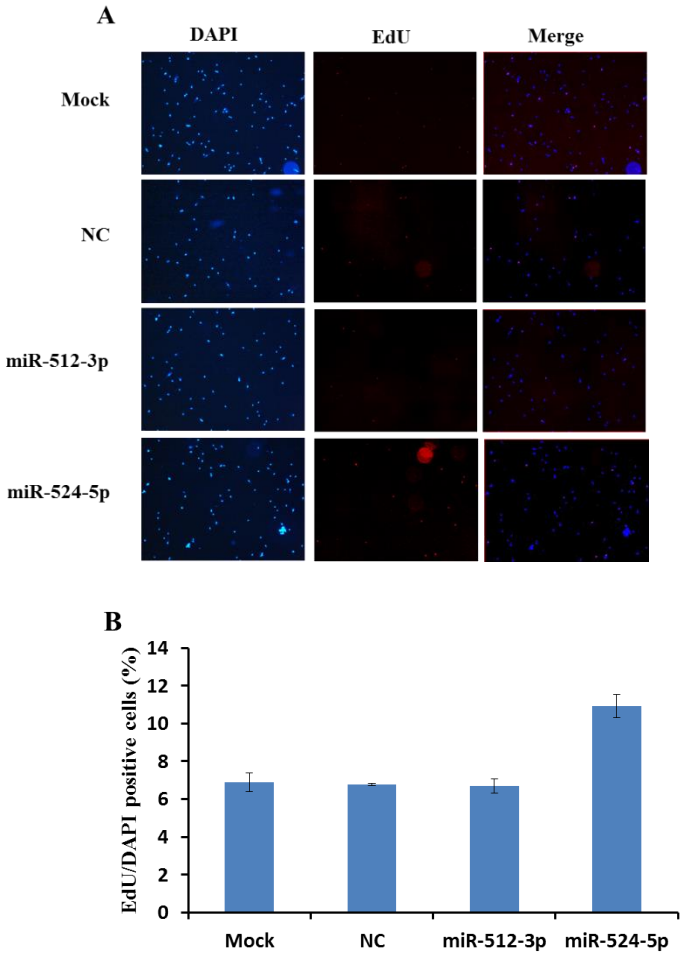


Figure 5.5 Effects of over-expression of miR-512-3p and miR-524-5p on the G1-to-S transition. A. Immunoflorescence staining of EdU uptake depicts nuclear staining DAPI (blue) and EdU (red) at 3 h after nocodazole withdrawal. MiR-512-3p and miR-524-5p mimics were transfected into WJ0706 cells for 48 h prior 18-h nocodazole treatment. As a control, a negative control mimic (NC) was also used in the transfection. **B.** Percentage of EdU/DAPI-positive cells. The percentage of the ratio of EdU/DAPI-labelled cells was calculated as percentage of the number of EdU-positive cells divided by the total number of DAPI-positive cells. The data shown were derived from two technical replicates of two independent experiments.

5.3.2 MiRNA-524-5p Enhances Reprogramming

In the previous chapter, miRNA microarray and miRNA copy number analyses have shown that miR-524-5p was expressed abundantly in pluripotent stem cells, including ESCs and iPSCs, whereas miR-524-5p expression was undetected or detected at low level in the MSC cell lines from which the iPSCs were derived (Nguyen et al., 2017). Furthermore, miR-524-5p shares 19/20 nucleotides with miR-520d-5p (Figure 5.2), suggesting identical biological functions for both miRNAs. Recent studies have indicated the ability of miR-520d-5p in converting cancer cells into iPSC-like cells (Tsuno et al., 2014; Ishihara et al., 2014). Moreover, the previous results have shown that over-expression of miR-524-5p enhances G1 to S phase transition. Hence, it is hypothesized that miR-524-5p may play an important role in reprogramming in iPSC.

To test if miR-524-5p also promotes reprogramming in the presence of other known reprogramming factors, the lentiviral vector-based pCDH-mir-524 construct encoding the mir-524 precursor was used to determine the effect of miR-524-5p on iPSC induction (Figures 5.6A to 5.6C). Human fibroblast HFF-1 cells were infected with pCDH-mir-524 together with the Dox-inducible lentiviral vectors each carrying the human cDNAs encoding the one of the four transcription factors OCT4, SOX2, c-MYC and KLF4 (OSKM), and a constitutively active lentivirus transducing the reverse tetracycline transactivator (FUW-m2rtTA) (see materials and methods). The blank vector pCDH-CD511 was included as a transduction control. After infection, the cells

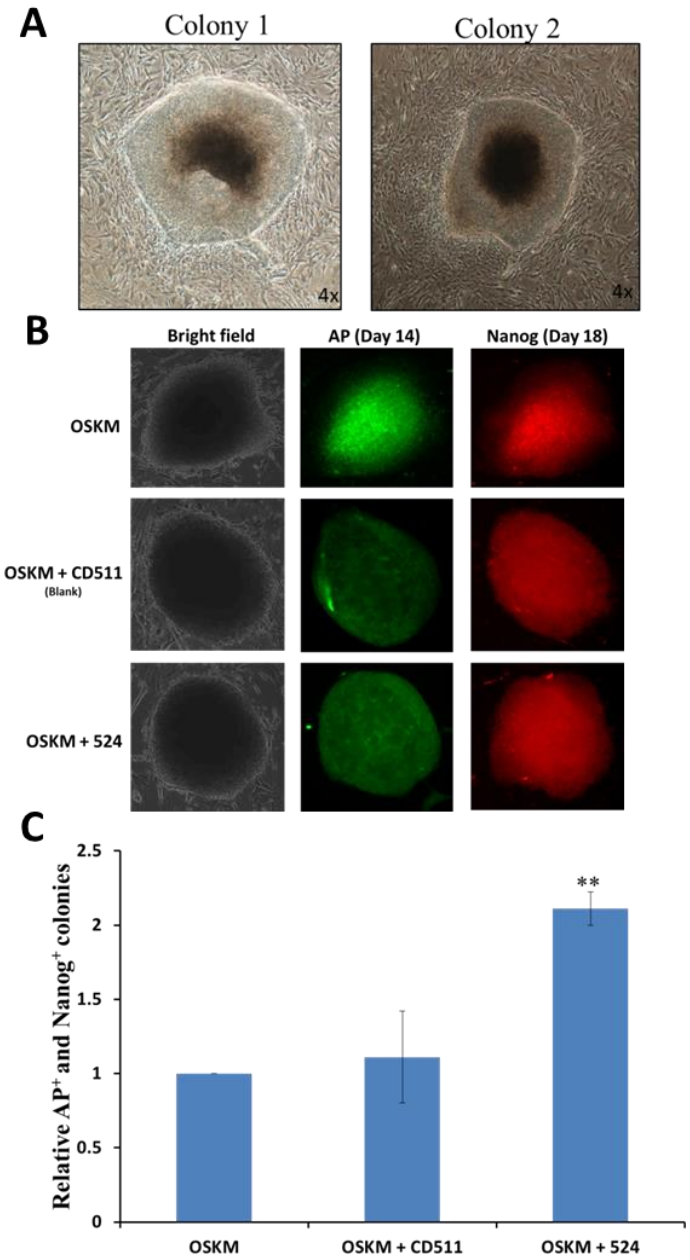


Figure 5.6 Over-expression of mir-524 precursor promotes OKSM-driven iPSC generation at the early stage of induction. A. ESC-like morphology of representative colonies formed at passage 1. **B.** Alkaline phosphatase-positive (AP^+) colonies at day 14 and NANOG-positive ($NANOG^+$) colonies at day 18 for transduction with OSKM only, OSKM in combination with either the blank vector CD511 or with a mir-524 precursor construct. **C.** Quantification of colonies stained both positive with AP at day 14 and NANOG at day 18 after infection of OSKM + CD511 or OSKM + mir-524 compared with colonies derived with transduction with OSKM alone, which was set as 1.0. The data shown were derived from three independent experiments. $^{**}p < 0.01$.

were cultured under standard conditions for human iPSC induction but in the presence of Dox, and the transduced cells were monitored daily for morphological changes. ESC-like colonies began to form on day 7, which when passaged on day 30 and cultured under Dox-free conditions, showed clear and round borders (Figure 5.6A). Furthermore, these colonies showed alkaline phosphatase (AP) staining on day 14 and NANOG at day 18 culturing under Dox-dependent medium (Figure 5.6B). Colonies that displayed ESC-like morphology and were AP- and NANOG- positive were considered *bona fide* iPSC colonies. When mir-524 was included in the iPSC induction with OSKM, the relative numbers of both AP and NANOG-positive colonies were significantly increased by approximately twofold compared with the induction of OSKM alone, or with OSKM in the presence of the blank vector CD511 (Figure 5.6C). The results indicated that the mir-524 precursor promoted iPSC induction in the presence of OSKM.

5.3.3 Bioinformatics Analysis of miRNA-524-5p and Predicted Target miRNA Interactions

To further understand the possible role of miR-524-5p in enhancing iPSC induction, the putative target genes of miR-524-5p were predicted by bioinformatics analysis. Since miR-524-5p was found to regulate the G1 to S transition which is also an important feature of regulation of stem cell self-renewal (Y. Wang et al., 2013; Kapinas et al., 2013), appropriate algorithms were interrogated to generate a repertoire of predicted G1 to S transition-

targets of miR-524-5p (Figure 5.7). Eight G1 to S transition-related genes, namely TGFβR1, SMAD2/3/4, RB1, PTEN, HIPK2 and TP53INP1, were predicted to be targeted by miR-524-5p. Repression of the PI3K/PKB/Akt/mTOR and TGFβ pathways, such as PTEN, p21/CDK1NA, TGFβR1 and SMAD2/3/4 has been reported to promote self-renewal and proliferation by blocking the G1 to S transition boundary of the cell cycle (Y. Wang et al., 2013; Jung et al., 2012; Itoh et al., 2014; Woltjen and Stanford, 2009). Furthermore, p53, a pro-apoptotic, anti-proliferative and anti-oxidant regulator, is indirectly regulated by miR-524-5p through targeting TP53INP1 and HIPK2 (Figure 5.7). Suppression of HIPK2 could inhibit p53 expression (Puca et al., 2010) whereas TP53INP1 is a major mediator of p53-driven responses to oxidative stress (Peuget et al., 2014). Besides p53, miR-524-5p was also predicted to target member of the RB family, RB1, an event that regulates the cell cycle by enhancing G1 to S transition and proliferation. TP53INP1 plays important roles not only in reprogramming by regulating p53 (Choi et al., 2011; He et al., 2007) but also in cancer stem cells in which TP53INP1 deficiency results in increased self-renewal and acquisition of cancer stem cell phenotype (Ma et al., 2010; Liu et al., 2015). Thus, the possible correlation between miR-524-5p and TP53INP1 was further investigated.

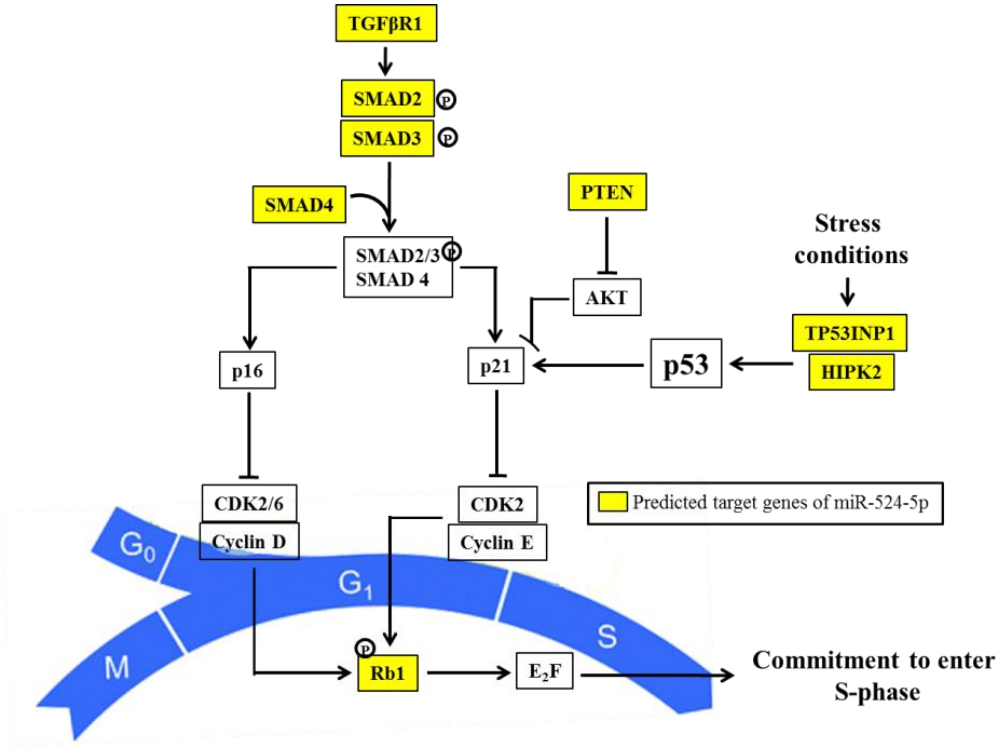


Figure 5.7 Predicted miR-524-5p-targeted genes regulate the G1 to S transition phase of the cell cycle. The predicted target genes were derived by interrogation of a variety of miRNA target prediction algorithms including the TargetScan, miRanda and DIANA-microT. Putative miR- 524-5p target genes are shown in yellow boxes.

5.3.4 TP53INP1 is a Direct Target of miR-524-5p

To investigate the relationship between miR-524-5p and TP53INP1, endogenous TP53INP1 expression in MSC and iPSC lines was first established by RT-PCR (Figure 5.8A). TP53INP1 was found to be expressed in both MSCs and iPSCs, albeit at higher levels in MSCs than in the derived iPSCs. Interestingly, our previous study has shown that miR-524-5p was expressed abundantly in iPSCs whereas miR-524-5p expression was undetected or detected at very low levels in MSC cell lines (Nguyen et al., 2017), suggesting an inverse correlation between the expression of miR-524-5p and TP53INP1 and negative regulation of TP53INP1 by miR-524-5p. ESC and placenta, which comprehensively express all the C19MC miRNAs, and, therefore, miR-524-5p (Nguyen et al., 2017), also expressed TP53INP1, and in higher levels in ESC. Surprisingly, the two cancer cell line controls, the colorectal HCT-15 and the breast cancer MCF-7, also expressed miR-524-5p to different levels (Figure 5.8A). Notably, all these cell lines, except for the placenta HS799.PI cells, also expressed various levels of miR-524-5p (Nguyen et al., 2017), but comparisons could not be made between the TP53INP1 and miR-524-5p expression levels.

When HCT-15 cells were transfected with a miR-524-5p mimic to achieve over-expression, TP53INP1 mRNA and protein levels were assayed in qRT-PCR and western blots (Figures 5.8B-D). As controls, a miRNA negative transfection control (NC) and a siRNA to knockdown TP53INP1 expression were also included. Transfection with the miR-524-5p mimic

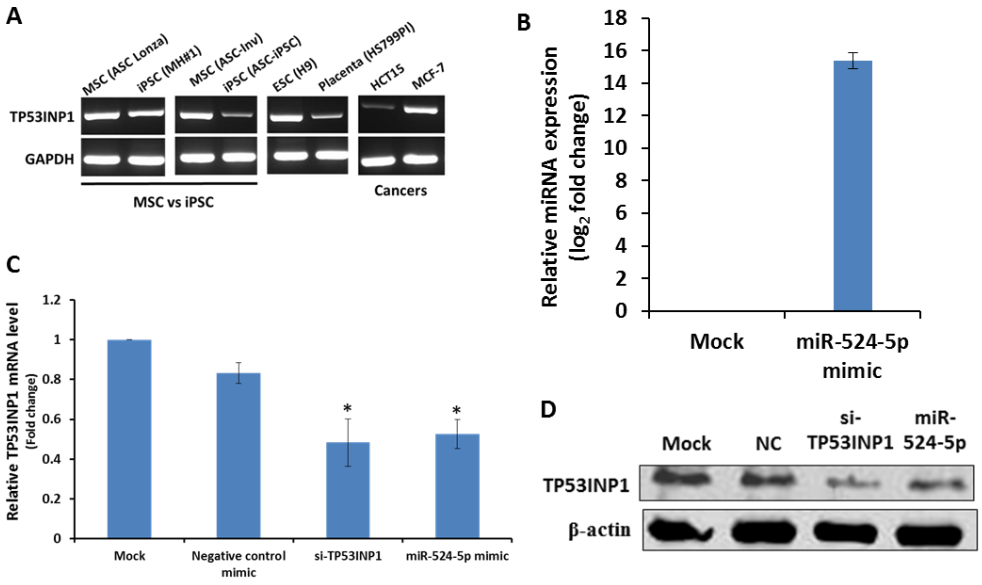


Figure 5.8 Inverse relationship between expression of miR-524-5p and TP53INP1. A. Expression of TP53INP1 in different cell lines determined by RT-PCR. iPSC (MH#1) and iPSC (ASC-iPSC) were derived from the adipose-derived MSC cell lines, ASC Lonza and ASC-Inv, respectively. ESC (H9), placenta (HS799.PI), colon cancer (HCT-15) and breast cancer (MCF-7) cell lines were included for comparison and as controls. **B**. Efficient transfection and over-expression of the miR-524-5p mimic in HCT-15 cells. **C, D** Effects of miR-524-5p over-expression on TP53INP1 expression. A miR-524-5p mimic was transfected into HCT-15 cells for 48 h before the cells were harvested for real-time RT-PCR (**C**) or western blot analysis (**D**) of TP53INP1 expression. As a control, a TP53INP1 siRNA (siTP53INP1) was also used in the transfection.

resulted in a 15-fold up-regulation of the miR-524-5p level 48 h post-transfection (Figure 5.8B). Similar to siRNA-mediated TP53INP1 suppression, forced over-expression of miR-524-5p significantly down-regulated TP53INP1 at both the mRNA and protein levels compared with the mock control (Figures 5.8C & 5.8D), further confirming a reverse relationship between miR-524-5p and TP53INP1 expression. The data further suggested that miR-524-5p regulated TP53INP1 expression probably via degradation of the TP53INP1 transcript.

Interestingly, when the TP53INP1 transcript sequences of various species were compared, TP53INP1 sequences of the primate, viz. human and chimpanzee, showed a tight clustering with a high sequence homology of 98.6% and late evolutionary emergence in comparison with other mammalian orthologs (Figure 5.9A). Keeping in mind that miR-524-5p belongs to the primate-specific C19MC cluster, the observation may hint at primate-specific co-evolution of the miR-524-5p and TP53INP1 gene sequences. The long 4,521-bp 3'-untranslated region (3'UTR) of the TP53INP1 transcript (NM_033285) encompasses four putative miR-524-5p-targeted sites, each with a 6- to 7-nucleotide seed sequence alignment with the TP53INP1 sequence, at nucleotide (nt) positions 1,461-1,466, nt 3,397-3,403, nt 5,450-5,455 and nt 5,530-5,536 of the transcript (Figure 5.9B).

Target sites 1, 3 and 4 did not show appreciable down-regulation of luciferase activities in luciferase assays using the pmirGLO vector (Appendix G). On the other hand, luciferase construct of target site 2, designated as WT2,

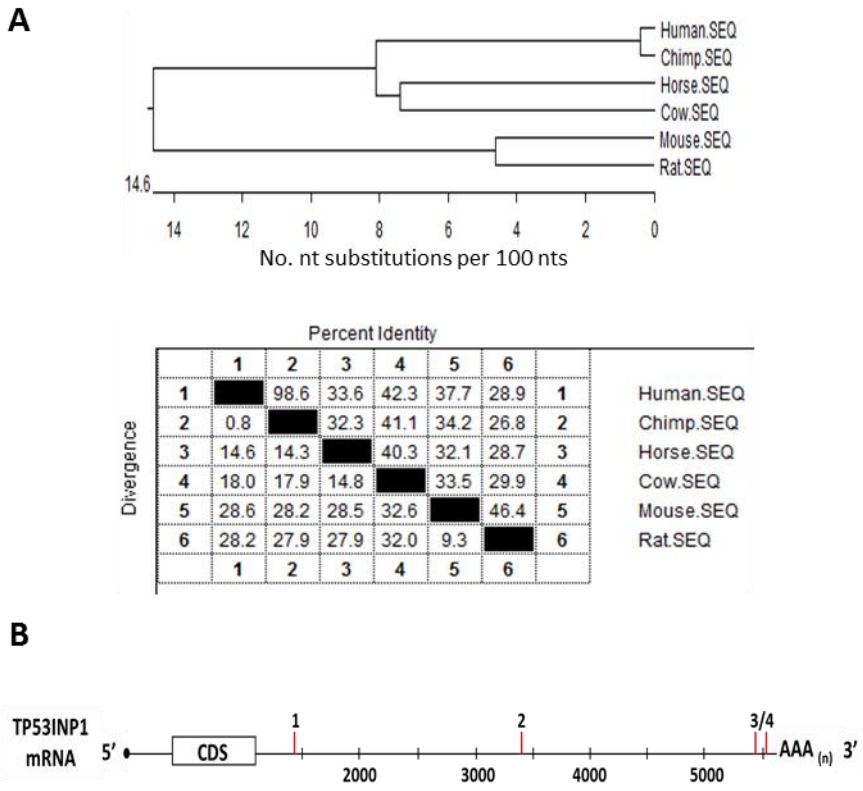


Figure 5.9 Possible co-evolution of miR-524-5p with TP53INP1. A. Phylogenetic tree alignment (top panel) and sequence comparison (bottom panel) of the 3'- UTR of TP53INP1 ortholog transcript sequences in different species. **B.** Identification of four putative miR-534-5p target sites (red vertical bars) in the 3'UTR of the human TP53INP1 transcript. CDS, coding sequence.

when co-transfected with the miR-524-5p mimic resulted in a reduction to ~40% of luciferase activity relative to that in the control cells transfected with the blank vector (Figure 5.10A, left panel). Specific targeting was confirmed with transfection of the mutated target site 2 in the Mut2 construct, which did not show appreciable effects on luciferase activities (Figure 5.10A, right panel). Echoing possible co-evolution of miR-524-5p and TP53INP1 suggested above (Figure 5.9), the active target site 2 also showed identical sequences between the two primate genes, but not with other mammalian orthologs (Figure 5.10B). Taken together, luciferase assays confirmed that miR-524-5p directly targets TP53INP1 to down-regulate TP53INP1 expression, and that the miRNA and target gene may have co-evolved late in the evolution of the primate.

5.3.5 MiR-524-5p Regulates Processes Relevant to Reprogramming

The cellular reprogramming process is a dynamic and tightly-controlled process driven by cascades of cellular and molecular events. Two such events that occur at the early stages of reprogramming are accelerated proliferation rates and suppression of apoptosis (David and Polo, 2014). To examine whether miR-524-5p expression influences cell proliferation rate, an MSC cell line was transfected with the miR-524-5p mimic and *in vitro* cell proliferation assays were performed. The data showed that transfection of the MSC with either the miR-524-5p mimic enhanced cell proliferation rates 3 to

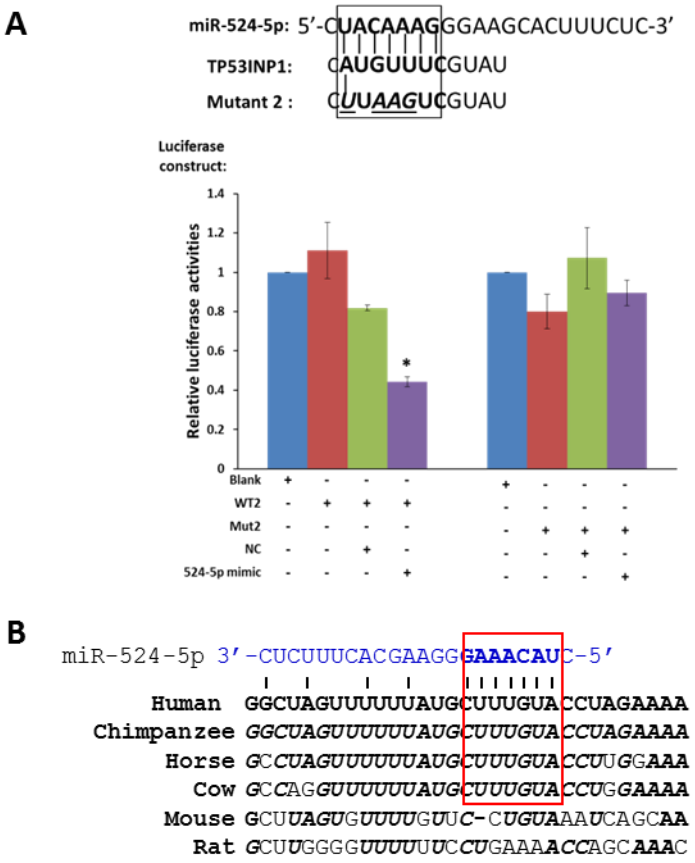


Figure 5.10 miR-524-5p direct targeting TP53INP1. A. Experimental validation of miR-524-5p targeting TP53INP1 in luciferase assays. In the top panel, alignment of miR-524-5p with the putative target site 2 (boxed and in bold letters) in the 3'-UTR of the TP53INP1 transcript (see text) is shown, so are the mutations (in italics and underlined) in the luciferase construct Mut2. HCT-15 transfected with the wild-type (WT2) or the mutant (Mut2) luciferase constructs alone, or in the presence of the miR-524-5p mimic or a validated negative control (NC) was performed before luciferase assays. The data shown were derived from two independent experiments in triplicates. *p< 0.05. **B.** Alignment of sequences around the active target site 2 of TP53INP1 in different species. The miR-524-5p miRNA sequence is shown in blue above the TP53IPN1 sequences. The targeted core sequences are underlined and shown in bold. Similar nucleotides relative to the human gene are shown in italics and in bold letters.

4 days post-transfection, correlating with results of the TP53INP1 knockdown (Figure 5.11A). The miR-524-5p-transfected cells also showed ~50% significantly higher BrdU incorporation compared to the mock control as in the experiment with TP53INP1 knock-down by siRNA transfection, which also significantly increased BrdU incorporation by 40% (Figure 5.11B). Hence, miR-524-5p enhances cell proliferation via TP53INP1 down-regulation.

Reprogramming is a stressful process that increases cellular levels of reactive oxygen species (ROS), resulting in the activation of apoptosis (Mah et al., 2011). To analyse the effects of miR-524-5p on cell viability in response to oxidative stress, miR-524-5p-transfected cells were exposed to 200 μM H_2O_2 for 2 h and the metabolic activities in the treated cells were measured by MTT assays. The results indicated that transfection of either the miR-524-5p mimic suppressed the damaging effects of H_2O_2-induced oxidative stress on cell viability, and the protection was likely via TP53INP1 since TP53INP1 knockdown also resulted in similar protection effects (Figure 5.12A). The effects of miR-524-5p on oxidative stress-induced apoptosis was further determined by ELISA quantification of the histone-associated DNA fragments in mono- and oligo-nucleosomes produced during nuclear DNA denaturation of apoptotic cells (Figure 5.12B). The result showed ~40% significant reduction of nucleosomes production in the miR-524-5p- or siTP53INP1-transfected cells compared to the mock control, further indicative of miR-524-5p suppression of apoptosis via TP53INP1.

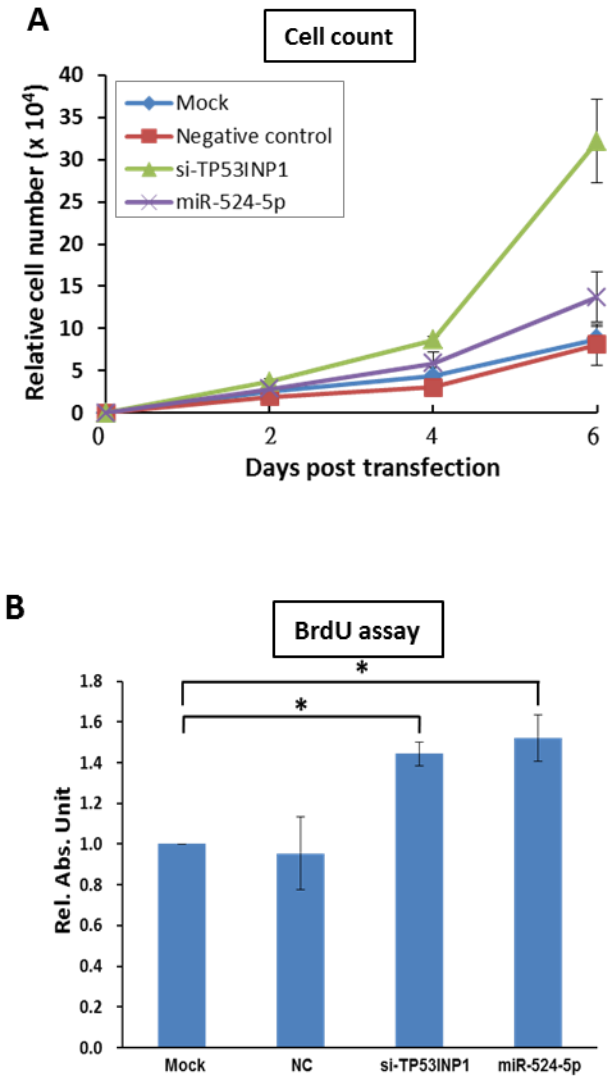


Figure 5.11 miR-524-5p enhances cell proliferation via targeting TP53INP1. In the experiments, a miR-524-5p mimic was transfected into the WJ0706 cell line, followed by biological assays to determine cell proliferation (**A & B**). All the experiments also included mock transfection with a negative control (NC) miRNA mimic, or a TP53INP1 siRNA (siTP53INP1) included as controls. For effects on cell proliferation, cell counts at different days post-transfection (A), or by BrdU ELISA measurements (**B**) were performed. *p<0.05.

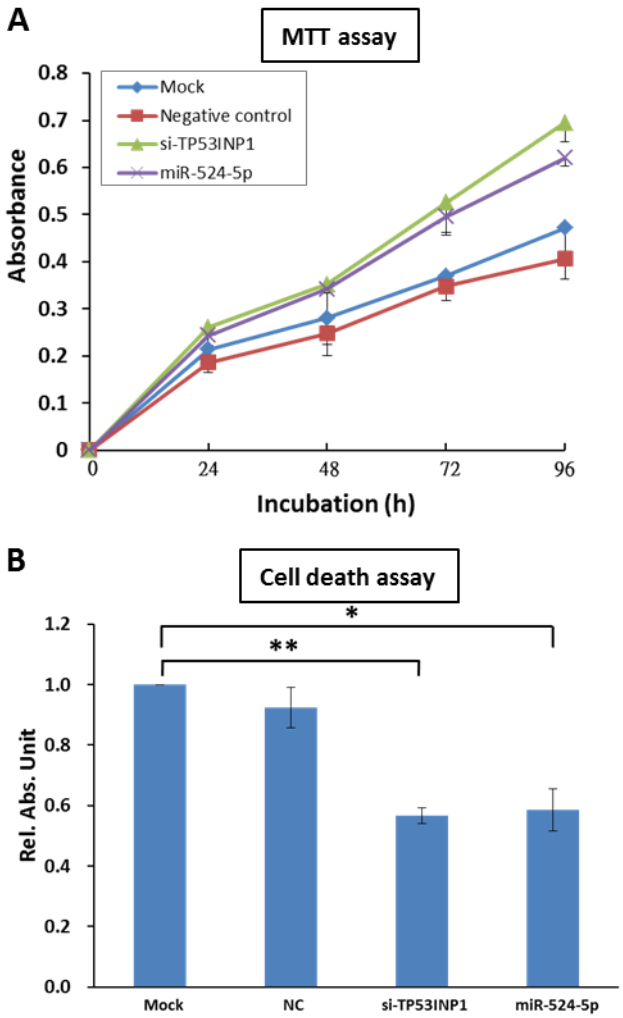


Figure 5.12 miR-524-5p inhibits oxidative stress-induced apoptosis via targeting TP53INP1. In the experiments, a miR-524-5p mimic was transfected into the WJ0706 cell line, followed by biological assays to determine cell viability (**A**) and apoptosis after oxidative stress induced with 200 μM H_2O_2 (**B**). All the experiments also included mock transfection with a negative control (NC) miRNA mimic, or a TP53INP1 siRNA (siTP53INP1) included as controls. Cell viability 2 h after incubation with H2O2 was determined by the MTT assay (**A**), or the histone-associated DNA fragments of apoptotic cells were quantified by ELISA assay 6 h after H2O2 treatment (**B**). Relative absorbance unit was determined as the relative absorbance units to the value of the mock experiment which was set as 1.0. *p<0.05 and **p<0.01.

The transition from the initiation to the maturation phase of the reprogramming process is also highlighted by up-regulation of pluripotency-associated genes, including OCT4, NANOG, SOX2 and REX1 (David and Polo, 2014; Plath and Lowry, 2011), which were assayed in miR-524-5p mimic- and siTP53INP1-transfected cells (Figure 5.13). Over-expression of miR-524-5p down-regulated TP53INP1 mRNA levels as anticipated, as did the transfection of siTP53INP1. Further, expression of the p53 gene, which is downstream of TP53INP1 (Figure 5.7), was concurrently down-regulated, also as anticipated (Peuget et al., 2014). The data clearly showed that both over-expression of miR-524-5p and TP53INP1 knockdown up-regulated the expression of all four pluripotency genes tested in the transfected cells (Figure 5.13). Hence, the results support that miR-524-5p negatively regulates TP53INP1, which in turn regulates p53, to up-regulate expression of pluripotency genes. Taken together, the results demonstrated involvement of miR-524-5p in events relevant to reprogramming, namely promotion of cell proliferation, suppression of oxidant-induced apoptosis and up-regulation of pluripotency-associated genes via targeting and degradation of the TP53INP1 transcript in the p53 signalling pathway.

5.3.6 MiR-524-5p Promotes MET Required for Initiating Reprogramming by Down-Regulating EMT-Related Genes

On reprogramming, besides promotion of cell proliferation, suppression of oxidant-induced apoptosis and up-regulation of pluripotency-

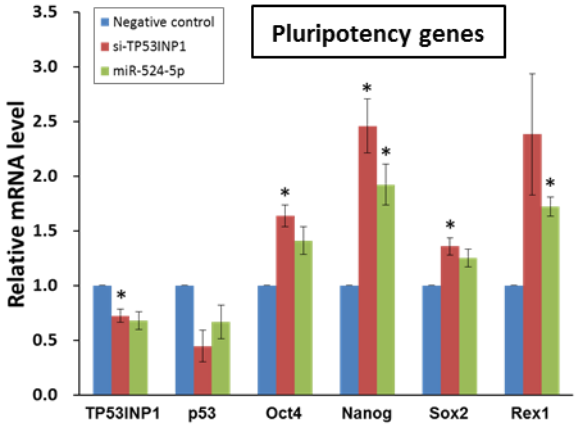


Figure 5.13 miR-524-5p up-regulates expression of pluripotency genes through targeting TP53INP1. In the experiments, a miR-524-5p mimic was transfected into the WJ0706 cell line, followed by biological assays to determine expression of pluripotency genes. All the experiments also included mock transfection with a negative control (NC) miRNA mimic, or a TP53INP1 siRNA (siTP53INP1) included as controls. Expression of pluripotency genes in the transfected cells was analysed by real-time RT-PCR 48 h post-transfection. MRNA level was determined as expression to the value of the negative control experiment which was set as 1.0. *p<0.05.

associated genes, MET is also an essential initiation step for progression towards pluripotency (David and Polo, 2014). The bioinformatics analysis revealed identical seed sequence and high degree of sequence homology between miR-524-5p and miR-520d-5p (Figure 5.2). Importantly, miR-520-5p has previously been reported to inhibit expression of the EMT-related gene TWIST1 (Tsukerman et al., 2014). Hence, we hypothesized that miR-524-5p may also enhance MET by targeting the EMT-associated genes that were predicted to be targeted by miR-524-5p. To test this hypothesis, a miR-524-5p mimic was introduced into Wharton's Jelly MSC 0706 (WJ0706) cell line and the expression of the predicted EMT-related target genes, including TGFβR1, SMAD2, SMAD3, SMAD4, ZEB1, ZEB2 and TWIST1, was first determined by RT-PCR. The results showed that ZEB2 and SMAD4 expression was obviously down-regulated 48 h post-transfection with the miR-524-5p mimic (Figure 5.14). MiR-524-5p down-regulation of endogenous expression of ZEB2 and SMAD4 was confirmed in quantitative real-time RT-PCR and western blot analysis (Figure 5.15A & 5.15B). Compared to transfection with the negative control mimic, forced expression of miR-524-5p significantly inhibited endogenous mRNA expression of ZEB2 and SMAD4 by almost 40% and 30%, respectively (Figure 5.15A). The protein levels of ZEB2 and SMAD4 were also diminished by miR-524-5p mimic (Figure 5.15B). Thus, these data suggested negative regulation of expression of ZEB2 and SMAD4 by miR-524-5p, which was confirmed by demonstration of direct miRNA targeting in luciferase assays (Figures 5.16A and 5.16B). 3'UTR construct of ZEB2 carried the two 5' predicted target sites whereas the SMAD4 construct carried a cluster of four putative target sites (Figure 5.16A). The ZEB2 or

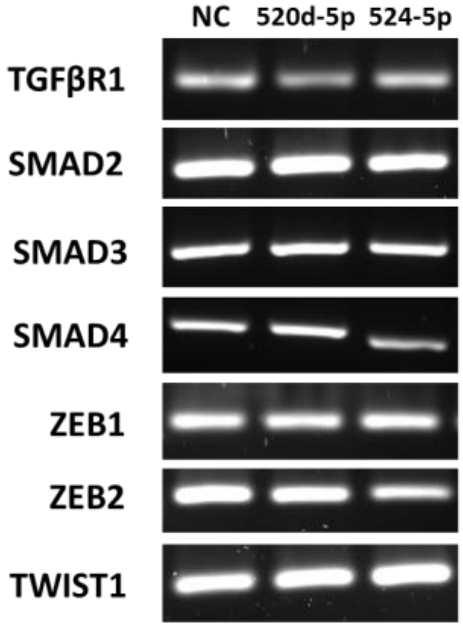


Figure 5.14 Effects of miR-524-5p over-expression on the expression of EMT-related genes. A miR-524-5p mimic was transfected to WJ0706 cells for 48 h before the cells were harvested. The expression of EMT-related genes TGFβR1, SMAD2, SMAD3, SMAD4, ZEB1, ZEB2 and TWIST1 was determined by RT-PCR. Negative control mimic (NC) and miR-520d-5p were included as controls.

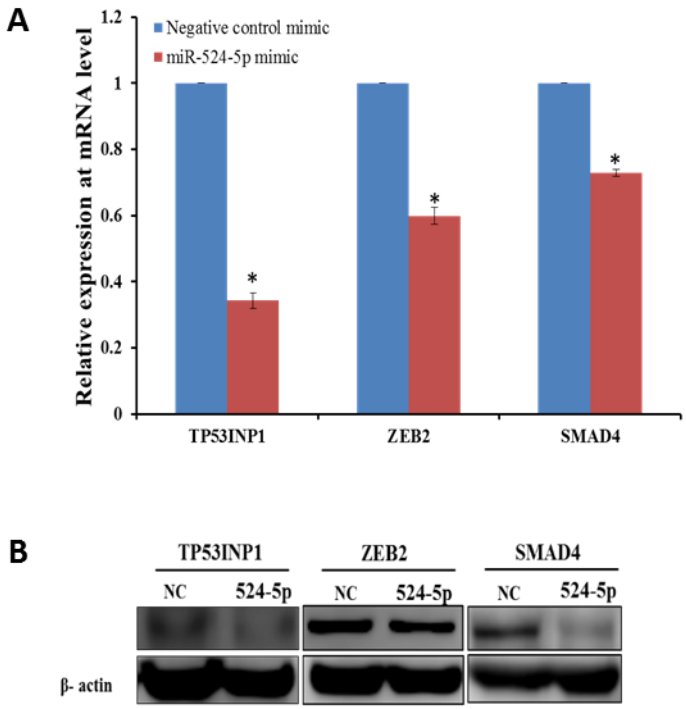


Figure 5.15 Effects of miR-524-5p over-expression on the expression of ZEB2 and SMAD4. A miR-524-5p mimic or negative control mimic (NC) were transfected to WJ0706 cells for 48 h before the cells were harvested for analysis. The analysis of the expression of ZEB2 and SMAD4 by real-time RT-PCR (**A**) and western blot analysis (**B**). In both experiments, TP53INP1 was included as a control. *$p < 0.05$.

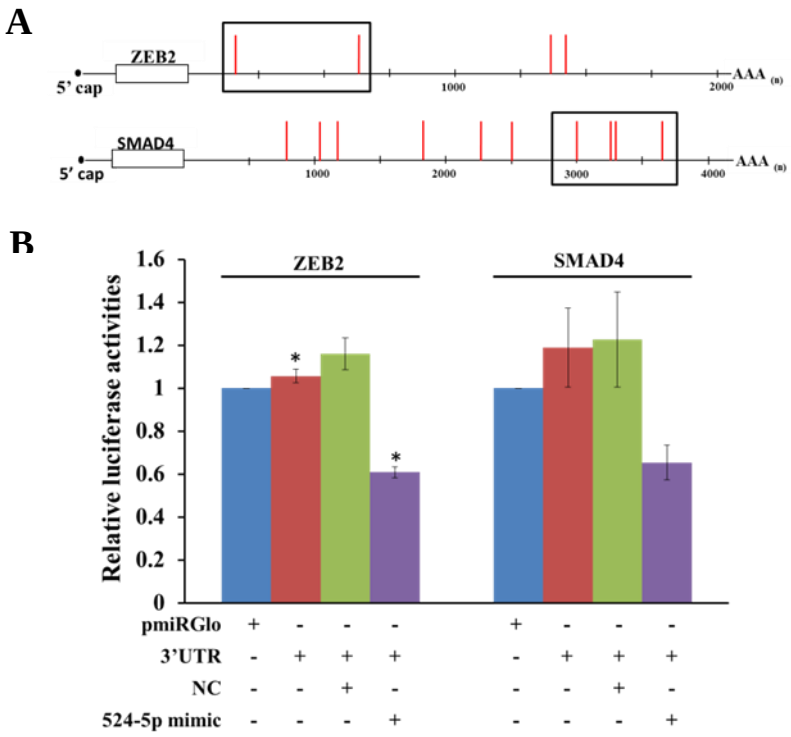


Figure 5.16 Experimental validation of miR-524-5p targeting of ZEB2 and SMAD4 in luciferase assays. Based on prediction of possible binding sites of miR-524-5p in 3'UTR of ZEB2 and SMAD4, a 3'-UTR luciferase construct of each gene was generated (boxed) (**A**). The blank pmiRGlo and 3'-UTR luciferase constructs were transfected alone, or in the presence of the miR-524-5p mimic or a validated negative control (NC) in colon cancer cell line (HCT-15) prior luciferase assays (**B**).*p< 0.05.

SMAD4 3'UTR constructs was individually transfected into HCT-15 cells alone, or in the presence of miR-524-5p mimic or a NC, which resulted in 60% and 65% reduction of luciferase activity, respectively, when co-transfected with the miR-524-5p mimic (Figure 5.16B). Taken together, the results confirmed that miR-524-5p targeted and repressed the expression of the EMT-associated genes ZEB2 and SMAD4, which may have direct bearing on the initial phase of establishing pluripotency.

5.4 Discussion

5.4.1 MiR-524-5p Enhances Reprogramming Efficiency by Targeting TP53INP1, ZEB2 and SMAD4

Emerging evidences have indicated that miRNAs play some crucial roles in somatic cell reprogramming, self-renewal and differentiation (Leonardo et al., 2012). Over-expression of miR-520d-5p alone has been reported to successfully convert hepatoma cells into iPSC-like cells (Tsuno et al., 2014). Since both C19MC members including miR-520d-5p and miR-524-5p are highly homologous in sequences and share the same seed sequence (Figure 5.2), the two miRNAs may share similar biological functions. It was first noted that miR-524-5p alone was unable to reprogram the normal fibroblast cells tested (data not shown). However, miR-524-5p was effective in enhancing the OSKM factor-driven reprogramming process. By targeting TP53INP1, miR-524-5p was shown to enhance proliferation and suppress

apoptosis (Figures 5.11 to 5.12), both of which are early crucial events for reprogrammable cells to enter subsequent phases of activation or up-regulation of pluripotency genes in the progression of the reprogramming process (David and Polo, 2014; Buganim et al., 2013). Furthermore, miR-524-5p was shown to target and down-regulate expression of the EMT-related genes, ZEB2 and SMAD4, and, hence, promoting MET, which is also an essential initial event of reprogramming (David and Polo, 2014; Buganim et al., 2013). Others have reported that introduction of multiple members of the miR-302/367 family was able to rapidly and efficiently reprogram fibroblasts into iPSCs with or without other reprogramming factors (Anokye-Danso et al., 2011; Hu et al., 2013). The miR-302/367 cluster enhances reprogramming efficiency by increasing cell division rate (Y. Wang et al., 2013), suppressing apoptosis (Z. Zhang et al., 2015) as well as promoting epigenetic reactivation of pluripotency genes (Lin et al., 2011), as shown here for miR-524-5p. In addition, miR-138 suppresses expression of p53 and its downstream genes, and significant enhances iPSC generation (Ye et al., 2012). Moreover, the miR-17-92, miR-106b-25, and miR-106a-363 clusters are highly expressed in the early phases of somatic cell reprogramming and directly target PTEN, p21 and TGFβR2, resulting in promoted iPSC induction by accelerating MET, cell cycle transitions, and regulation of epigenetic factors (He et al., 2014; Li et al., 2011). In the reprogramming of somatic cells, miRNAs are more likely to act as co-factors by enhancing the reprogramming process, as shown with miR-524-5p in this work, rather than acting in solo to exert their effects. It also seems likely that different miRNA-driven regulatory mechanisms and pathways may be involved in reprogramming normal somatic cells as opposed

to cancer cells, as in the case of miR-520d-5p and hepatoma (Tsuno et al., 2014).

5.4.2 MiR-524-5p Regulates Early Events of Reprogramming Process by Indirectly Mediating p53 Through Down-regulating TP53INP1

TP53INP1 and p53 are involved in many cellular processes, including apoptosis and regulation of cell cycle and radical oxygen species (ROS)-induced stress (Peuget et al., 2014; Seillier et al., 2012). On induction by p53, TP53INP1 is SUMOylated and, in turn, regulates p53 transcriptional activity by targeting anti-proliferative and pro-apoptotic genes, such as p21, BAX and PUMA, leading to cell-cycle arrest at the G1 phase, or apoptotic cell death (Peuget et al., 2014). In another study, TP53INP1 was also shown to regulate p53 activity on genes related to cell cycle regulation (Mdm2 and p21) and apoptosis (PIG3 and BAX) (Tomasini et al., 2003). Hence, TP53INP1 and p53 form a positive feedback loop in their action. Furthermore, ectopic expression of miR-504, miR-33 and miR-1285 has been shown to induce phenotypic changes associated with the loss of p53, including reduced apoptosis and increased stemness (Hermeking, 2012). Data from this and other works, therefore, strongly indicate that miRNA modulating the expression of both the TP53INP1 and p53 genes is important in fine-tuning the regulation of cell proliferation and apoptosis in the induction of pluripotency in iPSCs.

MiR-524-5p was also shown in this work to up-regulate the expression of pluripotency genes, OCT4, NANOG, SOX2 and REX1 (Figure 5.13). Expression of OCT4, NANOG and SOX2 is known to be negatively regulated by p53 (Takahashi et al., 2007; M. Li et al., 2012; Feng et al., 2009) and REX1 expression is, in turn, regulated by NANOG, OCT4 and SOX2 (Hosler et al., 1993; Shi et al., 2006). Hence, it may be speculated that the miR-524-5p/TP53INP1-induced up-regulated expression of pluripotency genes observed in this study may be a consequence of TP53INP1-induced p53 repression.

5.4.3 MiR-524-5p Promotes MET, an Essential Initial Event of Reprogramming, by Targeting ZEB2 and SMAD4

To achieve successful iPSC induction, exogenous factors are needed to initiate the MET program at the early stage of process by inhibiting EMT signals and activating the epithelial program (David and Polo, 2014; Lamouille et al., 2014). In this study, miR-524-5p was found to promote MET by inhibiting the expression of EMT-related genes, SMAD4 and ZEB2 (Figure 5.15 to 5.16), which may thereby be associated with enhancing the reprogramming process. More specifically, reprogramming has been reported to be associated with the loss of the somatic cell signatures, such as expression of the transcription factors SNAIL1/2 or ZEB1/2, and the gain of epithelial signatures, including expression of E-cadherin (David and Polo, 2014). A SNAIL1-SMAD3/4 complex has previously been shown to promote the

TGFβ-mediated down-regulation of E-cadherin while ZEB2 regulates repression by binding to the E-box motif of the regulatory sequence of the E-cadherin gene (Lamouille et al., 2014). Similarly, miR-302/367 and miR-200 play a crucial role in iPSC generation by targeting EMT-related gene TGFβR2 and ZEB1/ZEB2, respectively (Subramanyam et al., 2011; G. Wang et al., 2013), echoing our finding of miR-524-5p regulation of ZEB2.

5.4.4 A proposed Scheme of miR-524-5p Regulation Early Stage of Reprogramming

A scheme is proposed here to summarize the involvement of miR-524-5p in the reprogramming process via interactions with TP53INP1, ZEB2 and SMAD4, and the subsequent regulation of the p53 circuitry (Figure 5.17). In this scheme, miR-524-5p suppresses SMAD4 and ZEB2 resulting in up-regulation of the MET marker E-cadherin via the TGFβ pathway or by direct suppression of E-cadherin, respectively. On the other hand, direct suppression of TP53INP1 expression by miR-524-5p also leads to the p53 ablation, which in turn causes down-regulation of a cascade of p53-targeted genes involved in the cell cycle arrest and apoptosis, but up-regulates expression of pluripotency genes. Included in the scheme is also the previously reported ROS-induced p53 activation to form a feedback loop in the activation of TP53INP1 (Mah et al., 2011; Peuget et al., 2014).

5.5 Conclusions

In this work, we have provided experimental evidence to support that the C19MC miR-524-5p targets TP53INP1 to enhance cell proliferation and to suppress apoptosis, which are critical events in the early phase of cellular reprogramming. Our data also show that miR-524-5p targets the EMT-associated SMAD4 and ZEB2 gene to suppress MET, which is also a crucial step in initiating reprogramming. Other C19MC miRNAs, particularly those that share the same seed sequence with the known reprogramming miR-302/-372 families (Nguyen et al., 2017), may also be shown to contribute to cellular reprogramming in future studies.

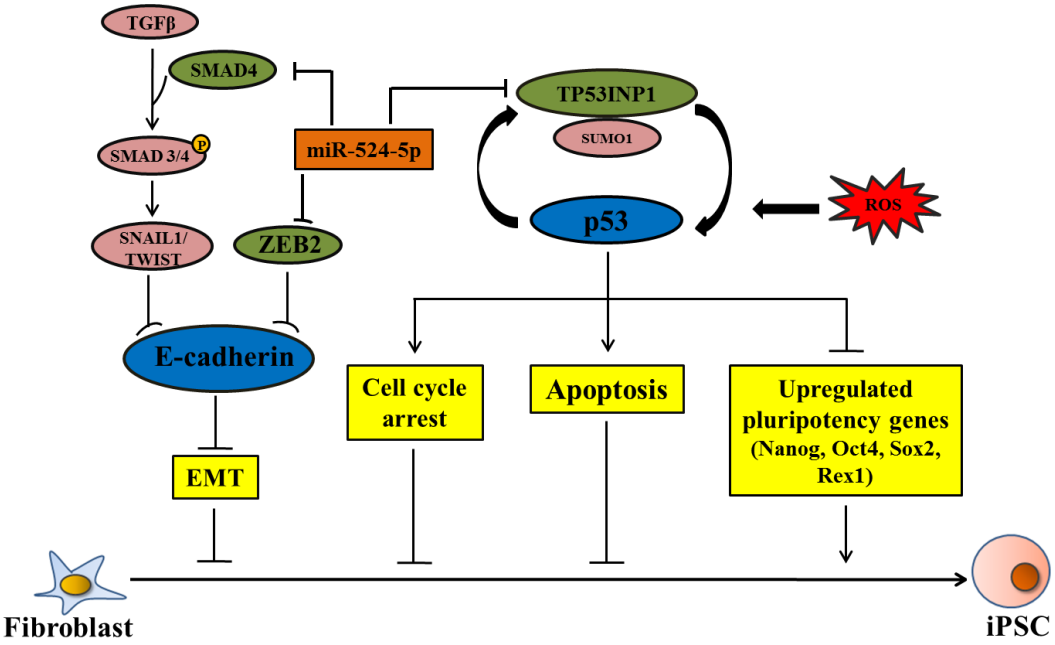


Figure 5.17 A proposed scheme of miR-524-5p regulation of the early stage of the reprogramming process (Buganim et al., 2013; David and Polo, 2014). In the scheme, miR-524-5p promotes reprogramming by down-regulating TP53INP1 to mediate processes associated with cell cycle, apoptosis and expression of pluripotency genes, which are essential for early stage of reprogramming. Furthermore, miR-524-5p also enhances MET, a required process for initial reprogramming, by targeting the EMT-related genes, ZEB2 and SMAD4. See text for further description of the proposed scheme.

CHAPTER 6

CONCLUSIONS AND FUTURE STUDIES

6.1 Conclusions

An overview of the major findings of the present study is presented in Figure 6.1.

In the miRNA microarray analysis of iPSCs, 261 miRNAs were found to be differentially expressed when compared with the parental AD-MSCs and pre-adipose cells. About a third of the differentially expressed miRNAs existed in both -5p and -3p forms, extending the range of target genes regulated. The 5p/3p miRNAs were co-up- or co-down-regulated indicating concerted 5p/3p regulation. MiRNAs of the C19MC cluster were found to be entirely activated in ESCs and iPSCs. However, in multipotent AD-MSCs, and in the unipotent HWP, only selected C19MC miRNAs were expressed. The C19MC expression profiles in MSCs were highly similar to those of the cancer cells analysed, suggesting that cancer and stem cells share miRNA-mediated gene regulatory mechanisms. Sixteen C19MC miRNAs share the same "AAGUGC" seed sequence with the miR-302 family, which are known cellular reprogramming factors, predicting that these C19MC-AAGUGC-miRNAs may be involved in induced pluripotency. Bioinformatics analysis of the putative targets of the C19MC-AAGUGC-miRNAs predicted significant

involvement of signalling pathways in reprogramming, many of which contribute to promoting apoptosis by indirect activation of the pro-apoptotic proteins BAK/BAX via suppression of genes of the cell survival pathways, or by enhancing caspase-8 activation through targeting inhibitors of TRAIL-inducing apoptosis.

To obtain experimental evidences to support possible involvement of C19MC miRNAs in reprogramming, the biological role of miR-524-5p was further explored. Co-expressing the miR-524 precursor with OSKM enhanced the OSKM-driven reprogramming efficiency. The putative target of miR-524-5p, TP53INP1, was confirmed in luciferase assays, and showed an inverse expression relationship with miR-524-5p. Functionally, miR-524-5p induced TP53INP1 down-regulation enhanced cell proliferation, suppressed apoptosis and up-regulated expression of pluripotency genes, all of which are critical early events of the reprogramming process. MiR-524-5p directly targeted the EMT-related genes, ZEB2 and SMAD4, and promoting MET. Hence, via targeting TP53INP1, ZEB2 and SMAD4, this work shows that miR-524-5p contributes to the early stage of inducing pluripotency by promoting cell proliferation, inhibiting apoptosis, up-regulating expression of pluripotency genes and enhancing MET.

In conclusion, data presented in this work indicate that specific C19MC miRNAs are important in regulating stem cell self-renewal and pluripotency, as experimentally demonstrated by the analysis of miR-524-5p.

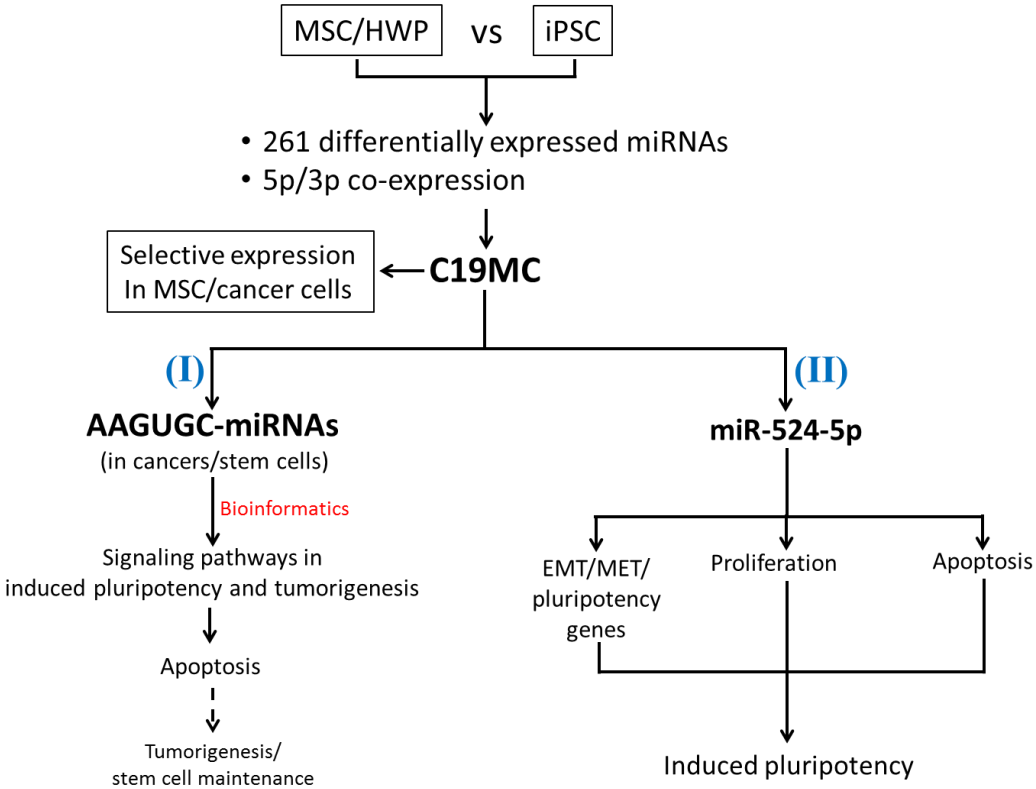


Figure 6.1 Overview of major findings in the present study. The findings were divided into two separated parts: (I) and (II). Part (I) focused on study the possible biological functions of C19MC-AAGUGC-miRNAs in mediating tumorigenesis and stem cell maintenance. Part (II) described contribution of C19MC miR-524-5p in induced pluripotency.

6.2 Future Studies

The bioinformatics analysis predicted possible involvement of selective C19MC miRNAs in regulating "stemness" and tumorigenesis, possibly via the cell survival pathways (subsection 4.3.3; Figure 4.9). The details of the pathways and mechanisms involved in C19MC in regulating stem cell and cancer properties remain to be elucidated.

In this study, it has been demonstrated that ectopic expression of C19MC miR-524-5p together with OSKM showed to obtain higher number of ESC-like colonies stained positively with alkaline phosphatase and NANOG comparing with those transduced OSKM alone or with blank vector CD511. However, the differentiation capacity of ESC-like colonies obtained from introduction of miR-524-5p has not been studied, therefore it is important to further verify that these ESC-like colonies could be pluripotency which are able to differentiate into derivatives of the three embryonic germ layers and into teratoma formation. Furthermore, this study only provided the indirect evidences for the mechanism that miR-524-5p enhances reprogramming efficiency through suppressing target genes including TP53INP1, ZEB2 and SMAD4. The experiments on silencing target genes during reprogramming are required for further indicating that the contribution of miR-524-5p in reprogramming process are in fact by direct targeting TP53INP1, ZEB2 and SMAD4.

Lastly, since numerous C19MC miRNAs share AAGUGC seed sequence with reprogramming miR-302 cluster, therefore it is worthy to expand the knowledge on regulatory mechanism of these miRNAs beside miR-524-5p in mediating and maintaining pluripotency.

REFERENCES

Aasen, T. et al., 2008. Efficient and rapid generation of induced pluripotent stem cells from human keratinocytes. *Nat Biotechnol*, 26(11), pp.1276–1284.

Akbari Moqadam, F., Pieters, R. and den Boer, M.L., 2013. The hunting of targets: challenge in miRNA research. *Leukemia*, 27(1), pp.16–23.

Almeida, M.I. et al., 2012. Strand-specific miR-28-5p and miR-28-3p have distinct effects in colorectal cancer cells. *Gastroenterology*, 142(4), p.886–896.e9.

Anokye-Danso, F. et al., 2011. Highly efficient miRNA-mediated reprogramming of mouse and human somatic cells to pluripotency. *Cell Stem Cell*, 8(4), pp.376–388.

Anokye-Danso, F., Snitow, M. and Morrisey, E.E., 2012. How microRNAs facilitate reprogramming to pluripotency. *J Cell Sci*, 125(18), pp.4179–4787.

Archer, T.C. and Pomeroy, S.L., 2014. A developmental program drives aggressive embryonal brain tumors. *Nat Genet*, 46(1), pp.2–3.

Aurich, H. et al., 2009. Hepatocyte differentiation of mesenchymal stem cells from human adipose tissue in vitro promotes hepatic integration in vivo. *Gut*, 58(4), pp.570–581.

Bar, M. et al., 2008. MicroRNA discovery and profiling in human embryonic stem cells by deep sequencing of small RNA libraries. *Stem Cells*, 26(10), pp.2496–2505.

Bara, J.J., Richards, R.G., Alini, M. and Stoddart, M.J., 2014. Concise review: Bone marrow-derived mesenchymal stem cells change phenotype following in vitro culture: implications for basic research and the clinic. *Stem Cells*, 32(7), pp.1713–1723.

Barroso-delJesus, A. et al., 2008. Embryonic stem cell-specific miR302-367 cluster: human gene structure and functional characterization of its core promoter. *Mol Cell Biol*, 28(21), pp.6609–6619.

Bartel, D.P., 2009. MicroRNAs: target recognition and regulatory functions. *Cell*, 136(2), pp.215–233.

Bazley, F.A. et al., 2015. Direct Reprogramming of Human Primordial Germ Cells into Induced Pluripotent Stem Cells: Efficient Generation of Genetically Engineered Germ Cells. *Stem Cells Dev*, 24(22), pp.2634–2648.

Beck, B. and Blanpain, C., 2012. Mechanisms regulating epidermal stem cells. *EMBO J*, 31(9), pp.2067–2075.

Borchert, G.M., Lanier, W. and Davidson, B.L., 2006. RNA polymerase III transcribes human microRNAs. *Nat Struct Mol Biol*, 13(12), pp.1097–1101.

Bortolin-Cavaille, M.L., Dance, M., Weber, M. and Cavaille, J., 2009. C19MC microRNAs are processed from introns of large Pol-II, non-protein-coding transcripts. *Nucleic Acids Res*, 37(10), pp.3464–3473.

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

Bradley, A., Evans, M., Kaufman, M.H. and Robertson, E., 1984. Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. *Nature*, 309(5965), pp.255–256.

Brayfield, C.A., Marra, K.G. and Rubin, J.P., 2010. Adipose tissue regeneration. *Curr Stem Cell Res Ther*, 5(2), pp.116–121.

Brownlie, R.J. and Zamoyska, R., 2013. T cell receptor signalling networks: branched, diversified and bounded. *Nat Rev Immunol*, 13(4), pp.257–269.

Buccini, S. et al., 2012. Cardiac progenitors derived from reprogrammed mesenchymal stem cells contribute to angiomyogenic repair of the infarcted heart. *Basic Res Cardiol*, 107(6), p.301.

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

Bunnell, B.A. et al., 2008. Adipose-derived stem cells: isolation, expansion and differentiation. *Methods*, 45(2), pp.115–120.

Bushati, N. and Cohen, S.M., 2007. microRNA functions. *Annu Rev Cell Dev Biol*, 23, pp.175–205.

Cahan, P. and Daley, G.Q., 2013. Origins and implications of pluripotent stem cell variability and heterogeneity. *Nat Rev Mol Cell Biol*, 14(6), pp.357–368.

Caplan, A.I., 1991. Mesenchymal stem cells. *J Orthop Res*, 9(5), pp.641–650.

Carroll, A.P., Goodall, G.J. and Liu, B., 2014. Understanding principles of miRNA target recognition and function through integrated biological and bioinformatics approaches. *WIREs RNA*, 5, pp.361–379.

Cawthorn, W.P., Scheller, E.L. and MacDougald, O.A., 2012. Adipose tissue stem cells meet preadipocyte commitment: going back to the future. *J Lipid Res*, 53(2), pp.227–246.

Chang, T.C. and Mendell, J.T., 2007. microRNAs in vertebrate physiology and human disease. *Annu Rev Genomics Hum Genet*, 8, pp.215–239.

Chen, C. et al., 2005. Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids Res*, 33(20), p.e179.

Chen, J. et al., 2013. H3K9 methylation is a barrier during somatic cell reprogramming into iPSCs. *Nat Genet*, 45, pp.34–42.

Choi, Y.J. et al., 2011. miR-34 miRNAs provide a barrier for somatic cell reprogramming. *Nat Cell Biol*, 13(11), pp.1353–1360.

Choo, K.B., Soon, Y.L., et al., 2014. MicroRNA-5p and -3p co-expression and cross-targeting in colon cancer cells. *J Biomed Sci*, 21(1), p.95.

Choo, K.B., Tai, L., et al., 2014. Oxidative stress-induced premature senescence in Wharton's jelly-derived mesenchymal stem cells. *Int J Med Sci*, 11(11), pp.1201–1207.

Clancy, J.L. et al., 2014. Small RNA changes en route to distinct cellular states of induced pluripotency. *Nature Communications*, 5, p.5522.

Cong, J. et al., 2015. Low miR-498 expression levels are associated with poor prognosis in ovarian cancer. *Eur Rev Med Pharmacol Sci*, 19(24), pp.4762–4765.

Cui, W. et al., 2010. miRNA-520b and miR-520e sensitize breast cancer cells to complement attack via directly targeting 3'UTR of CD46. *Cancer Biol Ther*, 10(3), pp.232–241.

Daelemans, C. et al., 2010. High-throughput analysis of candidate imprinted genes and allele-specific gene expression in the human term placenta. *BMC Genet*, 11, p.25.

Dai, R. et al., 2016. Adipose-Derived Stem Cells for Tissue Engineering and Regenerative Medicine Applications. *Stem Cells Int*, 2016, p.6737345.

Dani, C. et al., 1997. Differentiation of embryonic stem cells into adipocytes in vitro. *J Cell Sci*, 110 (Pt 1, pp.1279–1285.

Danielsen, S.A. et al., 2015. Portrait of the PI3K/AKT pathway in colorectal cancer. *Biochim Biophys Acta*, 1855(1), pp.104–121.

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

Delorme-Axford, E. et al., 2013. Human placental trophoblasts confer viral resistance to recipient cells. *Proc Natl Acad Sci U S A*, 110(29), pp.12048–12053.

Dominici, M. et al., 2006. Minimal criteria for defining multipotent mesenchymal stromal cells. The International Society for Cellular Therapy position statement. *Cytotherapy*, 8(4), pp.315–317.

Dong, Y. et al., 2015. MicroRNA-218 and microRNA-520a inhibit cell proliferation by downregulating E2F2 in hepatocellular carcinoma. *Mol Med Rep*, 12(1), pp.1016–1022.

Du, Z., Tong, X. and Ye, X., 2013. Cyclin D1 promotes cell cycle progression through enhancing NDR1/2 kinase activity independent of cyclin-dependent kinase 4. *J Biol Chem*, 288(37), pp.26678–26687.

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

Evans, M.J. and Kaufman, M.H., 1981. Establishment in culture of

pluripotential cells from mouse embryos. *Nature*, 292(5819), pp.154–156.

Feisst, V., Meidinger, S. and Locke, M.B., 2015. From bench to bedside: use of human adipose-derived stem cells. *Stem Cells Cloning*, 8, pp.149–162.

Feng, B. et al., 2009. Reprogramming of fibroblasts into induced pluripotent stem cells with orphan nuclear receptor Esrrb. *Nat Cell Biol*, 11, pp.197–203.

Flor, I. et al., 2012. Abundant expression and hemimethylation of C19MC in cell cultures from placenta-derived stromal cells. *Biochem Biophys Res Commun*, 422(3), pp.411–416.

Fornari, F. et al., 2012. In hepatocellular carcinoma miR-519d is up-regulated by p53 and DNA hypomethylation and targets CDKN1A/p21, PTEN, AKT3 and TIMP2. *J Pathol*, 227(3), pp.275–285.

Fraser, J.K., Wulur, I., Alfonso, Z. and Hedrick, M.H., 2006. Fat tissue: an underappreciated source of stem cells for biotechnology. *Trends Biotechnol*, 24(4), pp.150–154.

Friedenstein, A.J., Chailakhjan, R.K. and Lalykina, K.S., 1970. The development of fibroblast colonies in monolayer cultures of guinea-pig bone marrow and spleen cells. *Cell Tissue Kinet*, 3(4), pp.393–403.

Friedman, R.C., Farh, K.K.-H., Burge, C.B. and Bartel, D.P., 2009. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res*, 19(1), pp.92–105.

Fulda, S., 2009. Tumor resistance to apoptosis. *Int J Cancer*, 124(3), pp.511–515.

Fung, M.M., Rohwer, F. and Mcguire, K.L., 2003. IL-2 activation of a PI3K-dependent STAT3 serine phosphorylation pathway in primary human T cells. *Cell Signal*, 15(6), pp.625 – 636.

Fusaki, N. et al., 2009. Efficient induction of transgene-free human pluripotent stem cells using a vector based on Sendai virus, an RNA virus that does not integrate into the host genome. *Proc Jpn Acad Ser B Phys Biol Sci*, 85(8), pp.348–362.

Garofalo, M. and Croce, C.M., 2013. MicroRNAs as therapeutic targets in chemoresistance. *Drug Resist Updat*, 16(3–5), pp.47–59.

Germanguz, I. et al., 2011. Molecular characterization and functional properties of cardiomyocytes derived from human inducible pluripotent stem cells. *J Cell Mol Med*, 15(1), pp.38–51.

Ghule, P.N. et al., 2011. Reprogramming the pluripotent cell cycle: restoration of an abbreviated G1 phase in human induced pluripotent stem (iPS) cells. *J Cell Physiol*, 226(5), pp.1149–1156.

Gimble, J.M., Katz, A.J. and Bunnell, B.A., 2007. Adipose-derived stem cells for regenerative medicine. *Circ Res*, 100(9), pp.1249–1260.

Goldberg, A. et al., 2017. The use of mesenchymal stem cells for cartilage repair and regeneration: a systematic review. *J Orthop Surg Res*, 12, p.39.

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.

Griffiths-Jones, S. et al., 2006. miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res*, 34(Database issue), pp.D140–D144.

Gui, S. et al., 2015. Mir-302c mediates influenza A virus-induced IFNβ expression by targeting NF-κB inducing kinase. *FEBS Lett*, 589(24 Pt B), pp.4112–4118.

Han, C. et al., 2014. Human adipose-derived mesenchymal stem cells: a better cell source for nervous system regeneration. *Chin Med J (Engl)*, 127(2), pp.329–337.

Han, J. et al., 2010. Tbx3 improves the germ-line competency of induced pluripotent stem cells. *Nature*, 463(7284), pp.1096–1100.

Hanahan, D. and Weinberg, R.A., 2011. Hallmarks of cancer: the next generation. *Cell*, 144(5), pp.646–674.

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

Hanna, J. et al., 2007. Treatment of sickle cell anemia mouse model with iPS cells generated from autologous skin. *Science*, 318(5858), pp.1920–1923.

Hao, C. et al., 2016. MiR-708 promotes steroid-induced osteonecrosis of femoral head, suppresses osteogenic differentiation by targeting SMAD3. *Sci Rep*, 6, p.22599.

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

He, X. et al., 2014. Human fibroblast reprogramming to pluripotent stem cells regulated by the miR19a/b-PTEN axis. *PLoS One*, 9(4), p.e95213.

He, X., Liu, Z., Peng, Y. and Yu, C., 2016. MicroRNA-181c inhibits glioblastoma cell invasion, migration and mesenchymal transition by targeting TGF-β pathway. *Biochem Biophys Res Commun*, 469(4), pp.1041–1048.

Helbig, G. et al., 2003. NF-kappaB promotes breast cancer cell migration and metastasis by inducing the expression of the chemokine receptor CXCR4. *J Biol Chem*, 278(24), pp.21631–21638.

Heng, J.C.D. et al., 2010. The nuclear receptor Nr5a2 can replace Oct4 in the reprogramming of murine somatic cells to pluripotent cells. *Cell Stem Cell*, 6(2), pp.167–174.

Henzler, C.M. et al., 2013. Staged miRNA re-regulation patterns during reprogramming. *Genome Biol*, 14(12), p.R149.

Hermeking, H., 2012. MicroRNAs in the p53 network: micromanagement of tumour suppression. *Nat Rev Cancer*, 12(9), pp.613–626.

Hockemeyer, D. et al., 2008. A drug-inducible system for direct reprogramming of human somatic cells to pluripotency. *Cell Stem Cell*, 3(3), pp.346–353.

Hosler, B.A., Rogers, M.B., Kozak, C.A. and Gudas, L.J., 1993. An octamer motif contributes to the expression of the retinoic acid-regulated zinc finger gene Rex-1 (Zfp-42) in F9 teratocarcinoma cells. *Mol Cell Biol*, 13(5), pp.2919–2928.

Hromadnikova, I. et al., 2013. Circulating C19MC microRNAs in preeclampsia, gestational hypertension, and fetal growth restriction. *Mediators Inflamm*, 2013, p.186041.

Hromadnikova, I., Kotlabova, K., Ivankova, K. and Krofta, L., 2017. First trimester screening of circulating C19MC microRNAs and the evaluation of their potential to predict the onset of preeclampsia and IUGR. *PLoS One*, 12(2), p.e0171756.

Hu, S. et al., 2013. MicroRNA-302 increases reprogramming efficiency via repression of NR2F2. *Stem Cells*, 31(2), pp.259–268.

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

Huang, J. et al., 2016. An ANCCA/PRO2000-miR-520a-E2F2 regulatory loop as a driving force for the development of hepatocellular carcinoma. *Oncogenesis*, 5(5), p.e229.

Huang, Q. et al., 2008. The microRNAs miR-373 and miR-520c promote tumour invasion and metastasis. *Nat Cell Biol*, 10(2), pp.202–210.

Hussein, S.M.I. et al., 2014. Genome-wide characterization of the routes to pluripotency. *Nature*, 516, pp.198–206.

Indran, I.R., Tufo, G., Pervaiz, S. and Brenner, C., 2011. Recent advances in apoptosis, mitochondria and drug resistance in cancer cells. *Biochim Biophys Acta*, 1807(6), pp.735–745.

Inga, F. and Bullerdiek, J., 2012. The dark side of a success story: microRNAs of the C19MC cluster in human tumours. *J Pathol*, 227(3), pp.270–274.

Ishihara, Y. et al., 2014. Hsa-miR-520d converts fibroblasts into CD105+ populations. *Drugs R D*, 14(4), pp.253–264.

Itoh, F., Watabe, T. and Miyazono, K., 2014. Roles of TGF-β family signals in the fate determination of pluripotent stem cells. *Semin Cell Dev Biol*, 32, pp.98–106.

Jagadeeswaran, G. et al., 2010. Deep sequencing of small RNA libraries reveals dynamic regulation of conserved and novel microRNAs and

microRNA-stars during silkworm development. *BMC Genomics*, 11(1), pp.1–18.

Jalvy-Delvaille, S. et al., 2012. Molecular basis of differential target regulation by miR-96 and miR-182: the Glypican-3 as a model. *Nucleic Acids Res*, 40(3), pp.1356–1365.

Jeon, O. et al., 2008. In vivo bone formation following transplantation of human adipose-derived stromal cells that are not differentiated osteogenically. *Tissue Eng Part A*, 14(8), pp.1285–1294.

Jung, C.J. et al., 2012. Human ESC self-renewal promoting microRNAs induce epithelial-mesenchymal transition in hepatocytes by controlling the PTEN and TGFβ tumor suppressor signaling pathways. *Mol Cancer Res*, 10(7), pp.979–991.

Kan, H., Guo, W., Huang, Y. and Liu, D., 2015. MicroRNA-520g induces epithelial-mesenchymal transition and promotes metastasis of hepatocellular carcinoma by targeting SMAD7. *FEBS Lett*, 589(1), pp.102–109.

Kanellopoulou, C. et al., 2005. Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. *Genes Dev*, 19(4), pp.489–501.

Kang, S.K. et al., 2006. Autologous adipose tissue-derived stromal cells for treatment of spinal cord injury. *Stem Cells Dev*, 15(4), pp.583–594.

Kapinas, K. et al., 2013. The Abbreviated Pluripotent Cell Cycle. *J Cell Physiol*, 228(1), pp.9–20.

Keklikoglou, I. et al., 2012. MicroRNA-520/373 family functions as a tumor suppressor in estrogen receptor negative breast cancer by targeting NF-κB and TGF-β signaling pathways. *Oncogene*, 31(37), pp.4150–4163.

Khandelwal, N. et al., 2011. Nucleolar NF-κB/RelA mediates apoptosis by causing cytoplasmic relocalization of nucleophosmin. *Cell Death Differ*, 18(12), pp.1889–1903.

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

Kim, Y.K. and Kim, V.N., 2007. Processing of intronic microRNAs. *EMBO J*, 26(3), pp.775–783.

Kleinman, C.L. et al., 2014. Fusion of TTYH1 with the C19MC microRNA cluster drives expression of a brain-specific DNMT3B isoform in the embryonal brain tumor ETMR. *Nat Genet*, 46(1), pp.39–44.

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. *Biochem Biophys Res Commun*, 417(1), pp.11–16.

Kurosaki, T., Shinohara, H. and Baba, Y., 2010. B cell signaling and fate decision. *Annu Rev Immunol*, 28, pp.21–55.

Lakshmipathy, U. et al., 2007. MicroRNA expression pattern of undifferentiated and differentiated human embryonic stem cells. *Stem Cells Dev*, 16(6), pp.1003–1016.

Lamouille, S., Xu, J. and Derynck, R., 2014. Molecular mechanisms of epithelial–mesenchymal transition. *Nat Rev Mol Cell Biol*, 15(3), pp.178–196.

Landgraf, P. et al., 2007. A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell*, 129(7), pp.1401–1414.

Lee, M.O. et al., 2013. Inhibition of pluripotent stem cell-derived teratoma formation by small molecules. *Proc Natl Acad Sci U S A*, 110(35), pp.E3281-90.

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.

Lee, Y. et al., 2004. MicroRNA genes are transcribed by RNA polymerase II. *EMBO J*, 23(20), pp.4051–4060.

Lee, Y. et al., 2002. MicroRNA maturation: stepwise processing and subcellular localization. *EMBO J*, 21(17), pp.4663–4670.

Leonardo, T.R., Schultheisz, H.L., Loring, J.F. and Laurent, L.C., 2012. The functions of microRNAs in pluripotency and reprogramming. *Nat Cell Biol*, 14(11), pp.1114–1121.

Li, M. et al., 2012. Distinct regulatory mechanisms and functions for p53-activated and p53-repressed DNA damage response genes in embryonic stem cells. *Mol Cell*, 46(1), pp.30–42.

Li, M. et al., 2009. Frequent amplification of a chr19q13.41 microRNA polycistron in aggressive primitive neuroectodermal brain tumors. *Cancer Cell*, 16(6), pp.533–546.

Li, S.C. et al., 2012. miRNA arm selection and isomiR distribution in gastric cancer. *BMC Genomics*, 13 Suppl 1, p.S13.

Li, Z., Dang, J., Chang, K.-Y. and Rana, T.M., 2014. MicroRNA-mediated regulation of extracellular matrix formation modulates somatic cell reprogramming. *RNA*, 20(12), pp.1900–1915.

Li, Z., Yang, C.-S., Nakashima, K. and Rana, T.M., 2011. Small RNA-mediated regulation of iPS cell generation. *EMBO J*, 30(5), pp.823–834.

Liang, Y. et al., 2007. Characterization of microRNA expression profiles in normal human tissues. *BMC Genomics*, 8, p.166.

Lien, W.-H. and Fuchs, E., 2014. Wnt some lose some: transcriptional governance of stem cells by Wnt/β-catenin signaling. *Genes Dev*, 28(14),

pp.1517–1532.

Lin, S. et al., 2010. Computational identification and characterization of primate-specific microRNAs in human genome. *Comput Biol Chem*, 34(4), pp.232–241.

Lin, S.L. et al., 2008. Mir-302 reprograms human skin cancer cells into a pluripotent ES-cell-like state. *RNA*, 14(10), pp.2115–2124.

Lin, S.L. et al., 2011. Regulation of somatic cell reprogramming through inducible mir-302 expression. *Nucleic Acids Res*, 39(3), pp.1054–1065.

Lindroos, B., Suuronen, R. and Miettinen, S., 2011. The potential of adipose stem cells in regenerative medicine. *Stem Cell Rev*, 7(2), pp.269–291.

Lister, R. et al., 2011. Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature*, 471(7336), pp.68–73.

Liu, F., Kong, X., Lv, L. and Gao, J., 2015. TGF-β1 acts through miR-155 to down-regulate TP53INP1 in promoting epithelial-mesenchymal transition and cancer stem cell phenotypes. *Cancer Lett*, 359(2), pp.288–298.

Liu, P. and Wilson, M.J., 2012. miR-520c and miR-373 upregulate MMP9 expression by targeting mTOR and SIRT1, and activate the Ras/Raf/MEK/Erk signaling pathway and NF-κB factor in human fibrosarcoma cells. *J Cell Physiol*, 227(2), pp.867–876.

Lowe, C.E., O'Rahilly, S. and Rochford, J.J., 2011. Adipogenesis at a glance. *J Cell Sci*, 124(Pt 16), pp.2681–2686.

Lu, S. et al., 2015. Dual-Functions of miR-373 and miR-520c by Differently Regulating the Activities of MMP2 and MMP9. *J Cell Physiol*, 230(8), pp.1862–1870.

Lüningschrör, P., Hauser, S., Kaltschmidt, B. and Kaltschmidt, C., 2013. MicroRNAs in pluripotency, reprogramming and cell fate induction. *Biochim Biophys Acta*, 1833(8), pp.1894–1903.

Luo, M. et al., 2013. NuRD blocks reprogramming of mouse somatic cells into pluripotent stem cells. *Stem Cells*, 31(7), pp.1278–1286.

Luo, S.S. et al., 2009. Human villous trophoblasts express and secrete placenta-specific microRNAs into maternal circulation via exosomes. *Biol Reprod*, 81(4), pp.717–729.

Ma, S. et al., 2010. miR-130b Promotes CD133(+) liver tumor-initiating cell growth and self-renewal via tumor protein 53-induced nuclear protein 1. *Cell Stem Cell*, 7(6), pp.694–707.

Ma, W. et al., 2016. miR-517a is an independent prognostic marker and contributes to cell migration and invasion in human colorectal cancer. *Oncol Lett*, 11(4), pp.2583–2589.

Mah, N. et al., 2011. Molecular insights into reprogramming-initiation events

mediated by the OSKM gene regulatory network. *PLoS One*, 6(8), p.e24351.

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

Marchetto, M.C.N. et al., 2009. Transcriptional signature and memory retention of human-induced pluripotent stem cells. *PLoS One*, 4(9), p.e7076.

Marion, R.M. et al., 2009. A p53-mediated DNA damage response limits reprogramming to ensure iPS cell genomic integrity. *Nature*, 460(7259), pp.1149–1153.

Mathieu, J. and Ruohola-Baker, H., 2013. Regulation of stem cell populations by microRNAs. *Adv Exp Med Biol*, 786, pp.329–351.

Meijer, H.A., Smith, E.M. and Bushell, M., 2014. Regulation of miRNA strand selection: follow the leader? *Biochem Soc Trans*, 42(4), pp.1135–1140.

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.

De Miguel, M.P., Fuentes-Julian, S. and Alcaina, Y., 2010. Pluripotent stem cells: origin, maintenance and induction. *Stem Cell Rev*, 6(4), pp.633–649.

Mikkelsen, T.S. et al., 2008. Dissecting direct reprogramming through integrative genomic analysis. *Nature*, 454(7200), pp.49–55.

Min, D. et al., 2013. Downregulation of miR-302c and miR-520c by 1,25(OH)2D3 treatment enhances the susceptibility of tumour cells to natural killer cell-mediated cytotoxicity. *Br J Cancer*, 109(3), pp.723–730.

Miura, K. et al., 2015. Circulating chromosome 19 miRNA cluster microRNAs in pregnant women with severe pre-eclampsia. *J Obstet Gynaecol Res*, 41(10), pp.1526–1532.

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

Mizuno, H., 2009. Adipose-derived stem cells for tissue repair and regeneration: ten years of research and a literature review. *J Nippon Med Sch*, 76(2), pp.56–66.

Nguyen, P.N.N. et al., 2017. Selective activation of miRNAs of the primate-specific chromosome 19 miRNA cluster (C19MC) in cancer and stem cells and possible contribution to regulation of apoptosis. *J Biomed Sci*, 24, p.20.

Ning, H., Lin, G., Lue, T.F. and Lin, C.-S., 2006. Neuron-like differentiation of adipose tissue-derived stromal cells and vascular smooth muscle cells. *Differentiation*, 74(9–10), pp.510–518.

Noguchi, M. et al., 2013. In vitro characterization and engraftment of adipocytes derived from human induced pluripotent stem cells and embryonic stem cells. *Stem Cells Dev*, 22(21), pp.2895–2905.

Noguer-dance, M. et al., 2010. The primate-specific microRNA gene cluster (C19MC) is imprinted in the placenta. *Hum Mol Genet*, 19(18), pp.3566–3582.

Nombela-Arrieta, C., Ritz, J. and Silberstein, L.E., 2011. The elusive nature and function of mesenchymal stem cells. *Nat Rev Mol Cell Biol*, 12(2), pp.126–131.

Onder, T.T. et al., 2012. Chromatin-modifying enzymes as modulators of reprogramming. *Nature*, 483(7391), pp.598–602.

Ong, W.K. and Sugii, S., 2013. Adipose-derived stem cells: fatty potentials for therapy. *Int J Biochem Cell Biol*, 45(6), pp.1083–1086.

Pachón-Peña, G. et al., 2011. Stromal stem cells from adipose tissue and bone marrow of age-matched female donors display distinct immunophenotypic profiles. *J Cell Physiol*, 226(3), pp.843–851.

Padua, D. and Massague, J., 2009. Roles of TGFbeta in metastasis. *Cell Res*, 19(1), pp.89–102.

Pallero, M.A. et al., 2008. Thrombospondin 1 binding to calreticulin-LRP1 signals resistance to anoikis. *FASEB J*, 22(11), pp.3968–3979.

Park, B.K. et al., 2007. NF-kappaB in breast cancer cells promotes osteolytic bone metastasis by inducing osteoclastogenesis via GM-CSF. *Nat Med*, 13(1), pp.62–69.

Pasque, V. et al., 2011. Epigenetic factors influencing resistance to nuclear reprogramming. *Trends Genet*, 27(12), pp.516–525.

Pavyde, E. et al., 2016. Skeletal Muscle-Derived Stem/Progenitor Cells: A Potential Strategy for the Treatment of Acute Kidney Injury. *Stem Cells Int*, 2016, p.9618480.

Peuget, S. et al., 2014. Oxidative stress-induced p53 activity is enhanced by a redox-sensitive TP53INP1 SUMOylation. *Cell Death Differ*, 21(7), pp.1107–1118.

Pinho, F.G., Frampton, A.E. and Nunes, J., 2013. Downregulation of microRNA-515-5p by the estrogen receptor modulates sphingosine kinase 1 and breast cancer cell proliferation. *Cancer Res*, 73(19), pp.5936–5948.

Pittenger, M.F. and Martin, B.J., 2004. Mesenchymal stem cells and their potential as cardiac therapeutics. *Circ Res*, 95(1), pp.9–20.

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

Polo, J.M. et al., 2012. A molecular roadmap of reprogramming somatic cells into iPS cells. *Cell*, 151(7), pp.1617–1632.

Portt, L. et al., 2011. Anti-apoptosis and cell survival: a review. *Biochim Biophys Acta*, 1813(1), pp.238–259.

Poulos, S.P., Dodson, M. V and Hausman, G.J., 2010. Cell line models for differentiation: preadipocytes and adipocytes. *Exp Biol Med (Maywood)*, 235(10), pp.1185–1193.

Pritchard, C.C., Cheng, H.H. and Tewari, M., 2012. MicroRNA profiling: approaches and considerations. *Nat Rev Genet*, 13(5), pp.358–369.

Prockop, D.J., 1997. Marrow stromal cells as stem cells for nonhematopoietic tissues. *Science*, 276(5309), pp.71–74.

Puca, R., Nardinocchi, L., Givol, D. and D'Orazi, G., 2010. Regulation of p53 activity by HIPK2: molecular mechanisms and therapeutical implications in human cancer cells. *Oncogene*, 29(31), pp.4378–4387.

Ratajczak, M.Z. et al., 2012. Pluripotent and multipotent stem cells in adult tissues. *Adv Med Sci*, 57(1), pp.1–17.

Razak, S.R.A. et al., 2013. Profiling of microRNA in human and mouse ES and iPS cells reveals overlapping but distinct microRNA expression patterns. *PLoS One*, 8(9), p.e73532.

Ren, G. and Wang, L., 2016. Study on the relationship between miR-520g and the development of breast cancer. *Eur Rev Med Pharmacol Sci*, 20(4), pp.657–663.

Ren, J. et al., 2009. MicroRNA and gene expression patterns in the differentiation of human embryonic stem cells. *J Transl Med*, 7(20).

Rippe, V. et al., 1999. A KRAB zinc finger protein gene is the potential target of 19q13 translocation in benign thyroid tumors. *Genes Chromosomes Cancer*, 26(3), pp.229–236.

Rippe, V. et al., 2012. Activation of the two microRNA clusters C19MC and miR-371-3 does not play prominent role in thyroid cancer. *Mol Cytogenet*, 5(1), p.40.

Ritchie, W., Rajasekhar, M., Flamant, S. and Rasko, J.E.J., 2009. Conserved Expression Patterns Predict microRNA Targets. *PLoS Comput Biol*, 5(9), p.e1000513.

Ritchie, W., Rasko, J. and Flamant, S., 2013. MicroRNA target prediction and validation. Advances in experimental medicine and biology. *Adv Exp Med Biol*, 774, pp.39–53.

Rodriguez, A.-M. et al., 2005. Transplantation of a multipotent cell population from human adipose tissue induces dystrophin expression in the immunocompetent mdx mouse. *J Exp Med*, 201(9), pp.1397–1405.

Rossant, J., 2001. Stem cells from the Mammalian blastocyst. *Stem Cells*, 19(6), pp.477–482.

Rottiers, V. and Näär, A.M., 2012. MicroRNAs in metabolism and metabolic disorders. *Nat Rev Mol Cell Biol*, 13(4), pp.239–250.

Ruiz, S. et al., 2011. A high proliferation rate is required for cell reprogramming and maintenance of human embryonic stem cell identity. *Curr Biol*, 21(1), pp.45–52.

Rybak, A. et al., 2008. A feedback loop comprising lin-28 and let-7 controls pre-let-7 maturation during neural stem-cell commitment. *Nat Cell Biol*, 10(8), pp.987–993.

Sachdeva, M. et al., 2009. p53 represses c-Myc through induction of the tumor suppressor miR-145. *Proc Natl Acad Sci U S A*, 106(9), pp.3207–3212.

Saito, Y. et al., 2009. Chromatin remodeling at Alu repeats by epigenetic treatment activates silenced microRNA-512-5p with downregulation of Mcl-1 in human gastric cancer cells. *Oncogene*, 28(30), pp.2738–2744.

Sandmaier, S.E.S. and Telugu, B.P.V.L., 2015. MicroRNA-Mediated Reprogramming of Somatic Cells into Induced Pluripotent Stem Cells. *Methods Mol Biol*, 1330, pp.29–36.

Schiaffino, S. and Mammucari, C., 2011. Regulation of skeletal muscle growth by the IGF1-Akt/PKB pathway: insights from genetic models. *Skelet Muscle*, 1(1), p.4.

Seillier, M. et al., 2012. TP53INP1, a tumor suppressor, interacts with LC3 and ATG8-family proteins through the LC3-interacting region (LIR) and promotes autophagy-dependent cell death. *Cell Death Differ*, 19(9), pp.152515–152535.

Seo, M.J., Suh, S.Y., Bae, Y.C. and Jung, J.S., 2005. Differentiation of human adipose stromal cells into hepatic lineage in vitro and in vivo. *Biochem Biophys Res Commun*, 328(1), pp.258–264.

Shan, S.W. et al., 2013. Mature miR-17-5p and passenger miR-17-3p induce hepatocellular carcinoma by targeting PTEN, GalNT7 and vimentin in different signal pathways. *J Cell Sci*, 126(Pt 6), pp.1517–1530.

Shi, W. et al., 2006. Regulation of the pluripotency marker Rex-1 by Nanog and Sox2. *J Biol Chem*, 281(33), pp.23319–23325.

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.

De Smet, E.G. et al., 2015. Non-coding RNAs in the pathogenesis of COPD. *Thorax*, 70(8), pp.782–791.

Smith, Z.D., Nachman, I., Regev, A. and Meissner, A., 2010. Dynamic single-cell imaging of direct reprogramming reveals an early specifying event. *Nat Biotechnol*, 28(5), pp.521–526.

Smith, Z.D., Sindhu, C. and Meissner, A., 2016. Molecular features of cellular reprogramming and development. *Nat Rev Mol Cell Biol*, 17(3), pp.139–154.

Spence, T., Perotti, C., et al., 2014. A novel C19MC amplified cell line links Lin28/let-7 to mTOR signaling in embryonal tumor with multilayered rosettes.

Neuro Oncol, 16(1), pp.62–71.

Spence, T., Sin-Chan, P., et al., 2014. CNS-PNETs with C19MC amplification and/or LIN28 expression comprise a distinct histogenetic diagnostic and therapeutic entity. *Acta Neuropathol*, 128(2), pp.291–303.

Spencer, N.D., Gimble, J.M. and Lopez, M.J., 2011. Mesenchymal stromal cells: past, present, and future. *Vet Surg*, 40(2), pp.129–139.

Stadler, B. et al., 2010. Characterization of microRNAs involved in embryonic stem cell states. *Stem Cells Dev*, 19(7), pp.935–950.

Stadtfeld, M., Nagaya, M., et al., 2008. Induced pluripotent stem cells generated without viral integration. *Science*, 322(5903), pp.945–949.

Stadtfeld, M., Maherali, N., Breault, D.T. and Hochedlinger, K., 2008. Defining molecular cornerstones during fibroblast to iPS cell reprogramming in mouse. *Cell Stem Cell*, 2(3), pp.230–240.

Sterodimas, A., de Faria, J., Nicaretta, B. and Pitanguy, I., 2010. Tissue engineering with adipose-derived stem cells (ADSCs): current and future applications. *J Plast Reconstr Aesthet Surg*, 63(11), pp.1886–1892.

Su, C.M. et al., 2016. miR-520h is crucial for DAPK2 regulation and breast cancer progression. *Oncogene*, 35(9), pp.1134–1142.

Subramanyam, D. et al., 2011. Multiple targets of miR-302 and miR-372 promote reprogramming of human fibroblasts to induced pluripotent stem cells. *Nat Biotechnol*, 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. *Nat Protoc*, 6(3), pp.346–358.

Suzuki, H.I. et al., 2009. Modulation of microRNA processing by p53. *Nature*, 460(7254), pp.529–533.

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., 2016. A decade of transcription factor-mediated reprogramming to pluripotency. *Nat Rev Mol Cell Biol*, 17(3), pp.183–193.

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.

Takeda, Y.S. and Xu, Q., 2015. Neuronal Differentiation of Human Mesenchymal Stem Cells Using Exosomes Derived from Differentiating Neuronal Cells. *PLoS One*, 10(8), p.e0135111.

Tan, S.M. et al., 2014. Sequencing of captive target transcripts identifies the network of regulated genes and functions of primate-specific miR-522. *Cell*

Rep, 8(4), pp.1225–1239.

Tang, Q.Q. and Lane, M.D., 2012. Adipogenesis: from stem cell to adipocyte. *Annu Rev Biochem*, 81, pp.715–736.

Taura, D. et al., 2009. Adipogenic differentiation of human induced pluripotent stem cells: comparison with that of human embryonic stem cells. *FEBS Lett*, 583(6), pp.1029–1033.

Tay, Y.M. et al., 2008. MicroRNA-134 modulates the differentiation of mouse embryonic stem cells, where it causes post-transcriptional attenuation of Nanog and LRH1. *Stem Cells*, 26(1), pp.17–29.

Tomasini, R. et al., 2003. TP53INP1s and homeodomain-interacting protein kinase-2 (HIPK2) are partners in regulating p53 activity. *J Biol Chem*, 278(39), pp.37722–37729.

Tsai, C.H. et al., 2014. Resistin promotes tumor metastasis by down-regulation of miR-519d through the AMPK/p38 signaling pathway in human chondrosarcoma cells. *Oncotarget*, 6(1), pp.258–270.

Tsai, K. et al., 2009. Epigenetic control of the expression of a primate-specific microRNA cluster in human cancer cells. *Epigenetics*, 4(8), pp.587–592.

Tsukerman, P. et al., 2014. MiR-520d-5p directly targets TWIST1 and downregulates the metastamiR miR-10b. *Oncotarget*, 5(23), pp.12141–12150.

Tsuno, S. et al., 2014. Hsa-miR-520d induces hepatoma cells to form normal liver tissues via a stemness-mediated process. *Sci Rep*, 4, p.3852.

Vaira, V. et al., 2012. The microRNA cluster C19MC is deregulated in parathyroid tumours. *J Mol Endocrinol*, 49(2), pp.115–124.

Viswanathan, S.R., Daley, G.Q. and Gregory, R.I., 2008. Selective blockade of microRNA processing by Lin28. *Science*, 320(5872), pp.97–100.

Volarevic, V. et al., 2014. Stem cells as new agents for the treatment of infertility: current and future perspectives and challenges. *Biomed Res Int*, 2014, p.507234.

Wang, G. et al., 2013. Critical regulation of miR-200/ZEB2 pathway in Oct4/Sox2-induced mesenchymal-to-epithelial transition and induced pluripotent stem cell generation. *Proc Natl Acad Sci U S A*, 110(8), pp.2858–2863.

Wang, J. et al., 2011. Regulatory coordination of clustered microRNAs based on microRNA-transcription factor regulatory network. *BMC Syst Biol*, 5, p.199.

Wang, S. and El-Deiry, W.S., 2003. TRAIL and apoptosis induction by TNF-family death receptors. *Oncogene*, 22(53), pp.8628–8633.

Wang, Y. et al., 2008. Embryonic stem cell-specific microRNAs regulate the G1-S transition and promote rapid proliferation. *Nat Genet*, 40(12), pp.1478–

1483.

Wang, Y. et al., 2013. miR-294/miR-302 promotes proliferation, suppresses G1-S restriction point, and inhibits ESC differentiation through separable mechanisms. *Cell Rep*, 4(1), pp.99–109.

Welstead, G.G. et al., 2012. X-linked H3K27me3 demethylase Utx is required for embryonic development in a sex-specific manner. *Proc Natl Acad Sci U S A*, 109(32), pp.13004–13009.

Westphal, D., Dewson, G., Czabotar, P.E. and Kluck, R.M., 2011. Molecular biology of Bax and Bak activation and action. *Biochim Biophys Acta*, 1813(4), pp.521–531.

Wilmut, I. et al., 1997. Viable offspring derived from fetal and adult mammalian cells. *Nature*, 385(6619), pp.810–813.

Woltjen, K. and Stanford, W.L., 2009. Inhibition of Tgf-beta signaling improves mouse fibroblast reprogramming. *Cell Stem Cell*, 5(5), pp.457–458.

Wu, L. et al., 2001. The E2F1-3 transcription factors are essential for cellular proliferation. *Nature*, 414, pp.457–462.

Wu, S. et al., 2010. Multiple microRNAs modulate p21Cip1/Waf1 expression by directly targeting its 3' untranslated region. *Oncogene*, 29(15), pp.2302–2308.

Xiong, C. et al., 2005. Derivation of adipocytes from human embryonic stem cells. *Stem Cells Dev*, 14(6), pp.671–675.

Xu, N. et al., 2009. MicroRNA-145 regulates OCT4, SOX2, and KLF4 and represses pluripotency in human embryonic stem cells. *Cell*, 137(4), pp.647–658.

Ye, D. et al., 2012. MiR-138 promotes induced pluripotent stem cell generation through the regulation of the p53 signaling. *Stem Cells*, 30(8), pp.1645–1654.

Yingling, J.M., Blanchard, K.L. and Sawyer, J.S., 2004. Development of TGF-beta signalling inhibitors for cancer therapy. *Nat Rev Drug Discov*, 3(12), pp.1011–1022.

Zhang, J. et al., 2009. Functional cardiomyocytes derived from human induced pluripotent stem cells. *Circ Res*, 104(4), pp.e30-41.

Zhang, J. et al., 2014. microRNA-155 acts as an oncogene by targeting the tumor protein 53-induced nuclear protein 1 in esophageal squamous cell carcinoma. *Int J Clin Exp Pathol*, 7(2), pp.602–610.

Zhang, J. et al., 2016. MicroRNA-520g promotes epithelial ovarian cancer progression and chemoresistance via DAPK2 repression. *Oncotarget*, 7(18), pp.26516–26534.

Zhang, P. et al., 2015. Cordycepin (3'-deoxyadenosine) suppressed HMGA2,

Twist1 and ZEB1-dependent melanoma invasion and metastasis by targeting miR-33b. *Oncotarget*, 6(12), pp.9834–9853.

Zhang, R., Wang, Y.Q. and Su, B., 2008. Molecular evolution of a primate-specific microRNA family. *Mol Biol Evol*, 25(7), pp.1493–1502.

Zhang, S. et al., 2012. MicroRNA-520e suppresses growth of hepatoma cells by targeting the NF-κB-inducing kinase (NIK). *Oncogene*, 31(31), pp.3607–3620.

Zhang, W. et al., 2012. MicroRNA-520b inhibits growth of hepatoma cells by targeting MEKK2 and cyclin D1. *PLoS One*, 7(2), p.e31450.

Zhang, X. et al., 2014. Transplantation of Autologous Adipose Stem Cells Lacks Therapeutic Efficacy in the Experimental Autoimmune Encephalomyelitis Model. *PLoS One*, 9(1), p.e85007.

Zhang, X., Tang, N., Hadden, T.J. and Rishi, A.K., 2011. Akt, FoxO and regulation of apoptosis. *Biochim Biophys Acta*, 1813(11), pp.1978–1986.

Zhang, Y. et al., 2014. Concise Review: Differentiation of Human Adult Stem Cells Into Hepatocyte-like Cells In vitro. *Int J Stem Cells*, 7(2), pp.49–54.

Zhang, Z. et al., 2013. Dissecting the roles of miR-302/367 cluster in cellular reprogramming using TALE-based repressor and TALEN. *Stem Cell Reports*, 1(3), pp.218–225.

Zhang, Z. et al., 2015. MicroRNA-302/367 cluster governs hESC self-renewal by dually regulating cell cycle and apoptosis pathways. *Stem Cell Reports*, 4(4), pp.645–657.

Zhong, X. et al., 2010. Identification of microRNAs regulating reprogramming factor LIN28 in embryonic stem cells and cancer cells. *J Biol Chem*, 285(53), pp.41961–41971.

Zhou, J.Y. et al., 2016. MiR-519d facilitates the progression and metastasis of cervical cancer through direct targeting Smad7. *Cancer Cell Int*, 16, p.21.

Zhu, X. et al., 2015. Inhibition of RAC1-GEF DOCK3 by miR-512-3p contributes to suppression of metastasis in non-small cell lung cancer. *Int J Biochem Cell Biol*, 61, pp.103–114.

APPENDICES

APPENDIX A. Differentially expressed miRNAs in stem cells determined by microarray analysis

A. miRNAs altered in iPSC relative to ESC

No.	Activated		Up-regulated		Shutdown		Down-regulated	
	miRNA	**Log_2(fold change)**	**miRNA**	**Log_2(fold change)**	**miRNA**	**Log_2(fold change)**	**miRNA**	**Log_2(fold change)**
1	HSA-MIR-502-3P	9.05477652	HSA-MIR-199A-5P	3.289988274	HSA-MIR-636	-4.973121579	HSA-MIR-3615	-1.680011726
2	HSA-MIR-656	6.91477652	HSA-MIR-199A-3P	2.913321607	HSA-MIR-664*	-3.518121579	HSA-MIR-548D-3P	-1.575011726
3	HSA-MIR-130A*	6.758109853	HSA-MIR-154*	2.771654941	HSA-MIR-585	-2.218121579	HSA-LET-7F-2*	-1.560011726
4	HSA-MIR-378*	6.371443186	HSA-MIR-199B-5P	2.589988274			HSA-MIR-4279	-1.520011726
5	HSA-MIR-648	5.41477652	HSA-MIR-411	2.213321607				
6	HSA-MIR-589	3.91477652	HSA-MIR-432	2.129988274				
7	HSA-MIR-196B	2.51477652	HSA-MIR-181C	2.081654941				
8	HSA-MIR-4328	2.311443186	HSA-MIR-369-3P	1.998321607				
9	HSA-MIR-451	2.108109853	HSA-MIR-651	1.973321607				
10	HSA-MIR-1297	2.03477652	HSA-MIR-487B	1.898321607				
11	HSA-MIR-548V	1.56477652	HSA-MIR-410	1.798321607				
12	HSA-MIR-181C*	1.52477652	HSA-MIR-181B	1.584988274				
13			HSA-MIR-329	1.506654941				
	12		**13**		**3**		**4**	
	Total = 32							

APPENDIX A (Cont'd)

B. miRNAs altered in iPSC relative to MSC

No.	Activated		Up-regulated		Shutdown		Down-regulated	
	miRNA	Log_2(fold change)	miRNA	Log_2(fold change)	miRNA	Log_2(fold change)	miRNA	Log_2(fold change)
1	HSA-MIR-302F	13.15477652	HSA-MIR-302D	13.04792715	HSA-MIR-10A	-12.45851603	HSA-LET-7I	-10.99207285
2	HSA-MIR-1231	11.40810985	HSA-MIR-367	11.66459382	HSA-MIR-139-5P	-10.19351603	HSA-LET-7G	-9.538739512
3	HSA-MIR-302A*	11.06810985	HSA-MIR-302C	9.899593821	HSA-MIR-143	-9.473516032	HSA-MIR-199A-5P	-8.530406179
4	HSA-MIR-1290	10.89810985	HSA-MIR-135A	8.932927154	HSA-LET-7D	-9.083516032	HSA-MIR-199A-3P	-8.457072846
5	HSA-MIR-200C	10.70144319	HSA-MIR-518F	7.177927154	HSA-MIR-22*	-7.923516032	HSA-LET-7F	-8.120406179
6	HSA-MIR-3665	10.39144319	HSA-MIR-517B	6.777927154	HSA-MIR-1306	-6.408516032	HSA-LET-7A	-7.933739512
7	HSA-MIR-302C*	9.658109853	HSA-MIR-135B	6.499593821	HSA-LET-7D*	-6.403516032	HSA-MIR-199B-5P	-7.690406179
8	HSA-MIR-520C-3P	8.621443186	HSA-MIR-141	6.446260488	HSA-MIR-3660	-5.818516032	HSA-LET-7C	-7.603739512
9	HSA-MIR-302B*	8.611443186	HSA-MIR-335*	6.177927154	HSA-MIR-98	-5.808516032	HSA-LET-7E	-7.492072846
10	HSA-MIR-517A	8.598109853	HSA-MIR-372	5.261260488	HSA-MIR-876-5P	-4.248516032	HSA-MIR-137	-7.203739512
11	HSA-MIR-3670	8.26477652	HSA-MIR-96	5.239593821	HSA-MIR-4268	-3.883516032	HSA-MIR-29A	-6.937072846
12	HSA-MIR-187	8.101443186	HSA-MIR-429	5.204593821	HSA-MIR-664*	-3.033516032	HSA-MIR-145	-6.200406179
13	HSA-MIR-182	7.91477652	HSA-MIR-520H	5.201260488	HSA-MIR-10A*	-2.973516032	HSA-MIR-99A	-6.082072846
14	HSA-MIR-296-3P	7.89477652	HSA-MIR-520G	5.174593821	HSA-MIR-506	-2.778516032	HSA-MIR-100	-6.072072846
15	HSA-MIR-550A*	7.828109853	HSA-MIR-18A	5.079593821	HSA-MIR-567	-2.428516032	HSA-MIR-196A	-5.812072846
16	HSA-MIR-18B	7.711443186	HSA-MIR-3175	4.871260488	HSA-MIR-24-1*	-2.193516032	HSA-MIR-196B	-5.688739512

APPENDIX A (Cont'd)

No.	Activated		Up-regulated		Shutdown		Down-regulated	
	miRNA	Log_2(fold change)	miRNA	Log_2(fold change)	miRNA	Log_2(fold change)	miRNA	Log_2(fold change)
17	HSA-MIR-373	7.57477652	HSA-MIR-335	4.171260488	HSA-MIR-218-2*	-1.613516032	HSA-LET-7F-2*	-5.640406179
18	HSA-MIR-515-5P	7.388109853	HSA-MIR-205	4.134593821			HSA-MIR-29B	-5.588739512
19	HSA-MIR-150	7.348109853	HSA-MIR-935	4.122927154			HSA-MIR-31	-5.380406179
20	HSA-MIR-3180-3P	7.11477652	HSA-MIR-598	4.121260488			HSA-MIR-22	-5.193739512
21	HSA-MIR-519A	6.921443186	HSA-MIR-106A	4.116260488			HSA-MIR-31*	-5.182072846
22	HSA-MIR-3648	6.918109853	HSA-MIR-526B	4.049593821			HSA-MIR-27A	-4.660406179
23	HSA-MIR-518C	6.618109853	HSA-MIR-29B-2*	3.831260488			HSA-MIR-193A-5P	-4.615406179
24	HSA-MIR-371-5P	6.591443186	HSA-MIR-766	3.831260488			HSA-MIR-214	-4.365406179
25	HSA-MIR-9	6.511443186	HSA-MIR-421	3.827927154			HSA-MIR-155	-4.342072846
26	HSA-MIR-4303	6.498109853	HSA-MIR-520E	3.821260488			HSA-MIR-424*	-4.215406179
27	HSA-MIR-3147	6.391443186	HSA-MIR-17	3.816260488			HSA-MIR-21	-4.025406179
28	HSA-MIR-1180	6.32477652	HSA-MIR-20A	3.697927154			HSA-MIR-24	-3.940406179
29	HSA-MIR-4282	6.05477652	HSA-MIR-18A*	3.617927154			HSA-MIR-424	-3.883739512
30	HSA-MIR-449B*	6.038109853	HSA-MIR-126	3.566260488			HSA-MIR-181C	-3.833739512
31	HSA-MIR-656	5.74477652	HSA-MIR-301A	3.539593821			HSA-MIR-34A	-3.593739512
32	HSA-MIR-3605-5P	5.47477652	HSA-MIR-489	3.514593821			HSA-MIR-152	-3.495406179
33	HSA-MIR-4266	5.458109853	HSA-MIR-19B	3.436260488			HSA-MIR-218	-3.442072846
34	HSA-MIR-33A	5.21477652	HSA-MIR-532-3P	3.429593821			HSA-MIR-221*	-3.382072846
35	HSA-MIR-520A-5P	5.198109853	HSA-MIR-20A*	3.411260488			HSA-MIR-221	-3.355406179
36	HSA-MIR-412	4.92477652	HSA-MIR-498	3.382927154			HSA-MIR-10B	-3.118739512
37	HSA-MIR-3685	4.878109853	HSA-MIR-1281	3.329593821			HSA-MIR-222	-2.923739512

APPENDIX A (Cont'd)

No.	Activated		Up-regulated		Shutdown		Down-regulated	
	miRNA	Log_2(fold change)	miRNA	Log_2(fold change)	miRNA	Log_2(fold change)	miRNA	Log_2(fold change)
38	HSA-MIR-433	4.788109853	HSA-MIR-92A	3.286260488			HSA-MIR-29A*	-2.775406179
39	HSA-MIR-192	4.611443186	HSA-MIR-19A	3.224593821			HSA-MIR-146B -5P	-2.755406179
40	HSA-MIR-524-3P	4.538109853	HSA-MIR-651	3.117927154			HSA-MIR-4288	-2.755406179
41	HSA-MIR-515-3P	4.418109853	HSA-MIR-518B	3.102927154			HSA-MIR-365	-2.648739512
42	HSA-MIR-522	4.39477652	HSA-MIR-296 -5P	3.081260488			HSA-MIR-21*	-2.475406179
43	HSA-MIR-105	4.371443186	HSA-MIR-572	3.039593821			HSA-MIR-145*	-2.448739512
44	HSA-MIR-3150B	4.301443186	HSA-MIR-92B	2.919593821			HSA-MIR-708	-2.393739512
45	HSA-MIR-612	4.27477652	HSA-MIR-338 -3P	2.912927154			HSA-MIR-181B	-2.300406179
46	HSA-MIR-520D-5P	4.268109853	HSA-MIR-548F	2.864593821			HSA-MIR-34A*	-2.267072846
47	HSA-MIR-648	4.24477652	HSA-MIR-126*	2.861260488			HSA-MIR-15A	-2.163739512
48	HSA-MIR-524-5P	4.14477652	HSA-MIR-340	2.831260488			HSA-MIR-30A*	-2.105406179
49	HSA-MIR-518D-3P	3.948109853	HSA-MIR-219-1 -3P	2.806260488			HSA-MIR-181D	-1.990406179
50	HSA-MIR-516A-5P	3.918109853	HSA-MIR-877	2.792927154			HSA-MIR-30A	-1.913739512
51	HSA-MIR-744*	3.728109853	HSA-MIR-20B*	2.774593821			HSA-MIR-450A	-1.893739512
52	HSA-MIR-9*	3.721443186	HSA-MIR-106B	2.771260488			HSA-MIR-411	-1.862072846
53	HSA-MIR-575	3.691443186	HSA-MIR-1274A	2.766260488			HSA-MIR-27B	-1.852072846
54	HSA-MIR-200B*	3.678109853	HSA-MIR-130A	2.721260488			HSA-MIR-497	-1.847072846
55	HSA-MIR-519E*	3.65477652	HSA-MIR-940	2.721260488			HSA-MIR-34C -5P	-1.803739512

APPENDIX A (Cont'd)

No.	Activated		Up-regulated		Shutdown		Down-regulated	
	miRNA	Log_2(fold change)	miRNA	Log_2(fold change)	miRNA	Log_2(fold change)	miRNA	Log_2(fold change)
56	HSA-MIR-516B	3.578109853	HSA-MIR-602	2.626260488			HSA-MIR-26B	-1.788739512
57	HSA-MIR-92A-2*	3.558109853	HSA-MIR-512-3P	2.574593821			HSA-MIR-193B	-1.782072846
58	HSA-MIR-1283	3.391443186	HSA-MIR-92A-1*	2.567927154			HSA-MIR-4301	-1.763739512
59	HSA-MIR-519E	3.381443186	HSA-MIR-346	2.557927154			HSA-MIR-4328	-1.707072846
60	HSA-MIR-373*	2.83477652	HSA-MIR-3621	2.542927154			HSA-MIR-4291	-1.697072846
61	HSA-MIR-466	2.79477652	HSA-MIR-2277 -3P	2.531260488			HSA-MIR-625	-1.673739512
62	HSA-MIR-589	2.74477652	HSA-MIR-210	2.504593821			HSA-MIR-299-5P	-1.670406179
63	HSA-MIR-141*	2.44477652	HSA-MIR-3618	2.496260488			HSA-MIR-195	-1.647072846
64	HSA-MIR-1273C	1.938109853	HSA-MIR-874	2.491260488			HSA-MIR-24-2*	-1.627072846
65	HSA-MIR-876-3P	1.84477652	HSA-MIR-142 -3P	2.439593821			HSA-MIR-362-3P	-1.540406179
66	HSA-MIR-196B*	1.83477652	HSA-MIR-483 -5P	2.396260488			HSA-MIR-99B	-1.523739512
67			HSA-MIR-1468	2.379593821			HSA-MIR-3158	-1.500406179
68			HSA-MIR-503	2.297927154				
69			HSA-MIR-1247	2.262927154				
70			HSA-MIR-3181	2.221260488				
71			HSA-MIR-1233	2.217927154				
72			HSA-MIR-3622A-5P	2.194593821				
73			HSA-MIR-942	2.167927154				
74			HSA-MIR-4314	2.097927154				

APPENDIX A (Cont'd)

No.	Activated		Up-regulated		Shutdown		Down-regulated	
	miRNA	Log_2(fold change)	miRNA	Log_2(fold change)	miRNA	Log_2(fold change)	miRNA	Log_2(fold change)
75			HSA-MIR-518E*	2.089593821				
76			HSA-MIR-519B-3P	2.086260488				
77			HSA-MIR-1244	2.057927154				
78			HSA-MIR-93	2.022927154				
79			HSA-MIR-339-5P	2.011260488				
80			HSA-MIR-25*	1.999593821				
81			HSA-MIR-675	1.987927154				
82			HSA-MIR-153	1.934593821				
83			HSA-MIR-545*	1.911260488				
84			HSA-MIR-3656	1.901260488				
85			HSA-MIR-718	1.879593821				
86			HSA-MIR-18B*	1.876260488				
87			HSA-MIR-216A	1.874593821				
88			HSA-MIR-1224-3P	1.871260488				
89			HSA-MIR-374B	1.869593821				
90			HSA-MIR-25	1.854593821				
91			HSA-MIR-130B	1.847927154				
92			HSA-MIR-92B*	1.834593821				
93			HSA-MIR-1237	1.819593821				
94			HSA-MIR-526A	1.811260488				

APPENDIX A (Cont'd)

No.	Activated		Up-regulated		Shutdown		Down-regulated	
	miRNA	Log_2(fold change)	miRNA	Log_2(fold change)	miRNA	Log_2(fold change)	miRNA	Log_2(fold change)
95			HSA-MIR-4321	1.726260488				
96			HSA-MIR-937	1.719593821				
97			HSA-MIR-7	1.692927154				
98			HSA-MIR-4279	1.674593821				
99			HSA-MIR-378	1.667927154				
100			HSA-MIR-154*	1.661260488				
101			HSA-MIR-3615	1.654593821				
102			HSA-MIR-3137	1.636260488				
103			HSA-MIR-148A	1.619593821				
104			HSA-MIR-301B	1.611260488				
105			HSA-MIR-431*	1.604593821				
106			HSA-MIR-1270	1.601260488				
107			HSA-MIR-29C*	1.571260488				
108			HSA-MIR-130A*	1.539593821				
109			HSA-MIR-4318	1.537927154				
110			HSA-MIR-4263	1.534593821				
111			HSA-MIR-1307	1.527927154				
	66		**111**		**17**		**67**	
				Total = 261				

APPENDIX B. Mir families and expression levels in 5p/3p pairwise comparison for iPSC vs MSC (n=82)

I. Up-regulated in iPSC (both -5p and -3p species log_2(FC) >1.5) (n=14 pairs, 28 miRNAs)

No	miRNA family	Chromos' l site	miRNA-5p	Other name(s)	Expression status	Log2(FC)	miRNA-3p	Other name(s)	Expression status	Log2(FC)
1	mir-17	13q31.3	miR-18a-5p	miR-18A	Up	5.823	miR-18a-3p	miR-18A*	Up	4.133
2		13q31.3	miR-20a-5p	miR-20A	Up	4.226	miR-20a-3p	miR-20*	Up	4.096
3		Xq26.2	miR-20b-5p	miR-20b	Up	5.171	miR-20b-3p	miR-20B*	Up	3.023
4		Xq26.2	miR-106a-5p	miR-106A	Up	4.640	miR-106a-3p	miR-106A*	Up	2.739
5	mir-8	12p13.31	miR-141-5p	miR-141*	Act	2.110	miR-141-3p	miR-141	Up	6.446
6	mir-25	1q22	miR-92b-5p	miR-92B*	Up	1.835	miR-92b-3p	miR-92B	Up	2.920
7		7q22.1	miR-25-5p	miR-25*	Up	2.000	miR-25-3p	miR-25	Up	1.983
8		13q31.3	miR-92a-1-5p	miR-92A-1*	Up	2.568	miR-92a-3p	miR-92A	Up	2.923
9	mir-126	9q34.3	miR-126-5p	miR-126*	Up	2.133	miR-126-3p	miR-126	Up	3.186
10	mir-130	11q12.1	miR-130a-5p	miR-130A*	Up	1.610	miR-130a-3p	miR-130a	Up	3.770
11	mir-135	3p21.2	miR-135a-5p	miR-135A	Up	8.933	miR-135a-3p	miR-135A*	Act	5.265
12	mir-335	7q32.2	miR-335-5p	miR-335	Up	4.416	miR-335-3p	miR-335*	Up	5.066
13	mir-340	5q35.3	miR-340-5p	miR-340	Up	3.283	miR-340-3p	miR-340*	Up	1.922
14	mir-515	19q13.42	miR-518f-5p	miR-518F*	Up	3.556	miR-518-3p	miR-518F	Up	7.178

APPENDIX B (Cont'd)

II. Down-regulated in iPSC (both -5p and -3p species Log2(FC)<-1.5) (n=10 pairs, 20 miRNAs)

No	miRNA family	Chromos' l site	miRNA-5p	Other name(s)	Expression Status	Log2(FC)	miRNA-3p	Other name(s)	Expression Status	Log2(FC)
1	mir-29	7q32.3	miR-29a-5p	miR-29A*	Down	-1.807	miR-29a-3p	miR-29A	Down	-6.487
2		7q32.3	miR-29b-1-5p	miR-29B-1*	Down	-5.000	miR-29b-3p	miR-29B	Down	-4.553
3	mir-31	9p21.3	miR-31-5p	miR-31	Down	-4.737	miR-31-3p	miR-31*	Down	-4.447
4	mir-145	5q32	miR-145-5p	miR-145	Down	-4.8407	miR-145-3p	miR-145*	Down	-2.0969
5	mir-193	17q11.2	miR-193a-5p		Down	-4.6307	miR-193a-3p	miR-193A	Down	-2.026
6	mir-199	19p13.2	miR-199a-5p	miR-199A	Down	-7.677	miR-199a-3P	miR-199a*	Down	-8.274
7	mir-214	1q24.3	miR-214-5p	miR-214*	Down	-2.869	miR-214-3p	miR-214	Down	-3.834
8	mir-218	5q34	miR-218-5p	miR-218	Down	-3.257	miR-218-2-3p	miR-218-2*	Shut	-1.540
19	mir-221	Xp11.3	miR-221-5p	miR-221*	Down	-3.277	miR-221-3p	miR-221	Down	-3.080
10	mir-322	Xq26.3	miR-424-5p	miR-424	Down	-3.657	miR-424-3p	miR-424*	Down	-3.934

APPENDIX B (Cont'd)

III. Activated in iPSC (both 5p/3p Log2(FC)>1.5) (n=12 pairs, 24 miRNA)

No	miRNA family	Chromos' l site	miRNA-5p	Other name(s)	Expression status	Log2(FC)	miRNA-3p	Other name(s)	Expression status	Log2(FC)
1	mir-8	12p13.31	miR-200c-5p	miR-200C*	Up	1.583	miR-200c-3p	miR-200C	Act	10.366
2	mir-17	Xq26.2	miR-18b-5p	miR-18B	Act	7.376	miR-18b-3p	miR-18B*	Up	2.453
3	mir-373	19q13.42	miR-373-5p	miR-373*	Act	2.500	miR-373-3p	miR-373	Act	7.240
4	mir-515	19q13.42	miR-515-5p		Act	7.053	miR-515-3p		Act	4.083
5	mir-290	19q13.42	miR-371a-5p	miR-371-5P	Act	6.256	miR-371a-3p	miR-371-3P	Up	4.024
6	mir-296	20q13.32	miR-296-5p	miR-296	Up	3.083	miR-296-3p		Act	7.560
7	mir-515	19q13.42	miR-519e-5p	miR-519E*	Act	3.320	miR-519e-3p	miR-519E	Act	3.046
8	mir-339	7p22.3	miR-339-5p	miR-339	Up	2.690	miR-339-3p		Act	5.490
9	mir-589	7p22.1	miR-589-5p	miR-589	Act	2.410	miR-589-3p	miR-589*	Up	1.734
10	mir-3180		miR-3180-5p		Up	2.189	miR-3180-3p		Act	6.780
11	mir-148	7p15.2	miR-148a-5p	miR-148A*	Act	4.625	miR-148a-3p	miR-148A	Up	1.650
12	mir-744	7p12	miR-744-5p	miR-744	Up	1.536	miR-744-3p	miR-744*	Act	3.393

APPENDIX B (Cont'd)

IV. Shutdown in iPSC (both <1.5) (n=4 pairs, 8 miRNAs)

No	miRNA family	Chromos' l site	miRNA-5p	Other name(s)	Expression status	Log2(FC)	miRNA-3p	Other name(s)	Expression status	Log2(FC)
1	mir-10	7q21.32	miR-10a-5p	miR-10A	Shut	-12.420	miR-10a-3p	miR-10A*	Shut	-2.790
2	let-7	9q22.32	let-7d-5p	let-7D	Shut	-9.120	let-7d-3p	let-7D*	Shut	-6.090
3	mir-22	17p13.3	miR-22-5p	miR-22*	Shut	-7.880	miR-22-3p	miR-22	Down	-4.527
4	mir-24	9q22.32	miR-24-1-5p	miR-24-1*	Shut	-1.927	miR-24-3p	miR-24	Down	-3.167

V. Reverse direction in iPSC (both -5p and -3p species Log2(FC)<-1.5) (n=4 pairs, 8 miRNAs)

No	miRNA family	Chromos' l site	miRNA-5p	Other name(s)	Expression Status	Log2(FC)	miRNA-3p	Other name(s)	Expression Status	Log2(FC)
1	mir-146	10q24.32	miR-146b-5p	miR-146B	Down	-4.367	**miR-146b-3p**		**Up**	**1.500**
2	mir-196	7p15.2	miR-196b-5p	miR-196B	Down	-5.587	**miR-196b-3p**	**miR-196B***	**Act**	**1.500**
3	**mir-139**	**11q13.4**	**miR-139-5p**	**miR-139**	**Shut**	**-9.810**	**miR-139-3p**		**Up**	**1.943**
4	**mir-876**	**9p21.1**	**miR-876-5p**		**Shut**	**-3.797**	**miR-876-3p**		**Act**	**1.510**

APPENDIX C. Activated/up-regulated miRNAs arranged according to miRNA family and chromosomal location

Family	miRNA	-5p/3p name	Chromosome location
mir-8	hsa-miR-141*	hsa-miR-141-5p	12p13.31
	hsa-miR-200c	hsa-miR-200c-3p	
	hsa-miR-200b*	hsa-miR-200b-5p	1p36.33
mir-9	hsa-miR-9	hsa-miR-9-5p	1q22
	hsa-miR-9*	hsa-miR-9-3p	
mir-17	hsa-miR-18b	hsa-miR-18b-5p	Xq26.2
mir-25	hsa-miR-92a-2*	hsa-miR-92a-2-5p	Xq26.2
mir-33	hsa-miR-33a	hsa-miR-33a-5p	22q13.2
mir-105	hsa-miR-105	hsa-miR-105-5p	Xq28
mir-150	hsa-miR-150	hsa-miR-150-5p	19q13.33
mir-154	hsa-miR-656	-	14q32.31
mir-182	hsa-miR-182	hsa-miR-182-5p	7q32.2
mir-187	hsa-miR-187	hsa-miR-187-3p	18q12.2
mir-192	hsa-miR-192	hsa-miR-192-5p	11q13.1
mir-196	hsa-miR-196b*	hsa-miR-196b-3p	7p15.2
mir-290	hsa-miR-371-5p	hsa-miR-371a-5p	19q13.41
mir-296	-	hsa-miR-296-3p	20q13.32
mir-302	hsa-miR-302a*	hsa-miR-302a-5p	4q25 (known reprogramming miRNAs)
	hsa-miR-302b*	hsa-miR-302b-5p	
	hsa-miR-302c*	hsa-miR-302c-5p	
	hsa-miR-302f	-	
mir-373	hsa-miR-373	hsa-miR-373-3p	19q13.41
	hsa-miR-373*	hsa-miR-373-5p	
mir-412	hsa-miR-412	-	14q32.31
mir-433	hsa-miR-433	-	14q32.2
mir-449	hsa-miR-449b*	hsa-miR-449b-3p	5q11.2
mir-467	hsa-miR-466	-	3

APPENDIX C (Cont'd)

Family	miR	-5p/3p name	Chromasome location
mir-515 (C19MC) (n=17)	-	hsa-miR-515-5p	
	-	hsa-miR-515-3p	
		hsa-miR-516a-5p	
	hsa-miR-516b	hsa-miR-516b-5p	
	hsa-miR-517a	hsa-miR-517a-3p	
	hsa-miR-518c	hsa-miR-518c-3p	
	-	hsa-miR-518d-5p	
	hsa-miR-518d	hsa-miR-518d-3p	
	hsa-miR-519a	hsa-miR-519a-3p	19q13.42
	hsa-miR-519e	hsa-miR-519e-3p	
	hsa-miR-519e*	hsa-miR-519e-5p	
	hsa-miR-520a*	hsa-miR-520a-5p	
	hsa-miR-520c	hsa-miR-520c-3p	
	hsa-miR-520d*	hsa-miR-520d-5p	
	hsa-miR-522	hsa-miR-522-3p	
	hsa-miR-524*	hsa-miR-524-5p	
	hsa-miR-524	hsa-miR-524-3p	
	hsa-miR-1283	-	19
mir-550	hsa-miR-550a*	hsa-miR-550a-3p	7p14.3
mir-575	hsa-miR-575	-	4q21.22
mir-589	hsa-miR-589	hsa-miR-589-5p	7p22.1
mir-612	hsa-miR-612	-	11q13.1
mir-648	hsa-miR-648	-	22q11.21
mir-744	hsa-miR-744*	hsa-miR-744-3p	17p12
mir-876	-	hsa-miR-876-3p	9p21.1
mir-1180	hsa-miR-1180	-	17
mir-1231	hsa-miR-1231	-	1
mir-1273c	hsa-miR-1273c	-	6
mir-1290	hsa-miR-1290	-	1

APPENDIX D. Putative target genes of group I C19MC-AAGUGC-miRNAs

No.	Gene ID	No.	Gene ID	No.	Gene ID	No.	Gene ID	No.	Gene ID
1	ANKRD45	61	HIPK3	121	PDCD1LG2	181	FAM196B	241	CR2
2	KLHL2	62	BCL6	122	SOBP	182	SLC45A2	242	RSAD2
3	CXCR4	63	MFAP3L	123	TESK2	183	SETD5	243	TGFBR2
4	GALNT3	64	UNKL	124	DNM1L	184	CYB5D2	244	JAK1
5	POLK	65	LMX1A	125	ZNF2	185	STK33	245	ZNF77
6	RTN1	66	FCMR	126	DEK	186	HNRNPUL1	246	ST8SIA2
7	SLC15A2	67	NELL2	127	PLVAP	187	DAPK2	247	LRP4
8	NFIA	68	KLF9	128	FOXF2	188	SFRP4	248	TAPT1
9	EPS8	69	BROX	129	R3HDM1	189	PPP1R10	249	HPS5
10	DAD1	70	TRIM6	130	CCP110	190	JOSD1	250	SHC4
11	ROCK2	71	PARP8	131	RBM25	191	CUX2	251	PLAC8
12	LGI2	72	BNC1	132	TMTC1	192	POLQ	252	BECN1
13	C6orf15	73	TRIM24	133	FLT4	193	CCDC62	253	ZNFX1
14	DMRT3	74	HFM1	134	RNASE4	194	ACTN2	254	ANKRD55
15	LAMP3	75	VPS35	135	MAP3K9	195	TXNIP	255	TRPS1
16	UGT2B10	76	COL4A2	136	PIGM	196	RSBN1L	256	MIB1
17	PLIN2	77	LUM	137	TNIK	197	SMYD2	257	STK17B
18	RFX3	78	TCF7L1	138	SOS1	198	PTGDR2	258	SKA2
19	ATP11AUN	79	HOXD8	139	FRMD3	199	TMEM154	259	SAR1B
20	ARL13B	80	FOXJ3	140	MALT1	200	PARP1	260	TCEAL1
21	RACGAP1	81	SASH1	141	FAM151B	201	MAGEH1	261	RPS6KA3
22	C1orf168	82	ERCC4	142	IRF2BP2	202	SEC23A	262	SLC16A10
23	ERBB2IP	83	IRAK2	143	PHLPP2	203	ZNF333	263	RAB8A
24	TOX3	84	LCLAT1	144	ACTL6A	204	PRDM8	264	CD44
25	RAD18	85	DCAF6	145	SNIP1	205	EZH1	265	CEP250
26	PIK3AP1	86	CT55	146	ZNF148	206	TVP23B	266	CIPC
27	HABP4	87	KLHL15	147	FYCO1	207	ACAD9	267	ABRA
28	MFAP5	88	SNX5	148	TIPARP	208	HIVEP2	268	CMTR2
29	CNOT6	89	MYLK	149	SLC40A1	209	ATAD2	269	ATF6B
30	ITPRIPL2	90	SOD1	150	UHRF2	210	SUV39H1	270	MUM1L1
31	OSR2	91	ARL4D	151	HSD17B11	211	MED17	271	PFKP
32	PKIB	92	ZNF317	152	MED8	212	LPA	272	ZNF799
33	SLAIN1	93	NEO1	153	HOXB3	213	MYO15A	273	DMP1
34	SERTAD2	94	STAG1	154	MYNN	214	PIK3CA	274	CUBN
35	RBBP6	95	SLITRK3	155	SUN1	215	MTURN	275	NAGPA
36	DYRK4	96	BLCAP	156	CROT	216	MAP3K2	276	CRTAM
37	BMP6	97	SLAIN2	157	MPC1	217	SDC1	277	MSL1
38	PLSCR4	98	TUSC3	158	FGF19	218	IFIT5	278	DCAF5
39	FSTL5	99	GLIS3	159	GOLGA1	219	JUNB	279	C12orf66
40	ANKRD45	100	TMEM50B	160	USP9X	220	GRM5	280	SP140L
41	C17orf100	101	MUC17	161	CYP4V2	221	CA10	281	FAT4
42	G6PC2	102	CPA3	162	NPAP1	222	CAPRIN2	282	FXYD6
43	GRHL1	103	ACVR1	163	LEFTY1	223	ABHD3	283	TOB2
44	C5	104	FAM175B	164	LRRC66	224	USP34	284	TET3
45	NHLH2	105	CYBB	165	LIMA1	225	PIGS	285	NINL
46	TANC2	106	SNRK	166	ZNF236	226	ANKRD29	286	MED13L
47	NTN4	107	FBXL17	167	KLLN	227	CXCL1	287	C10orf12
48	FAM120A	108	BAHD1	168	MRRF	228	MXD3	288	FAHD1
49	SYDE2	109	SRPX	169	AKTIP	229	TNFRSF10B	289	GEN1
50	TLX1	110	KIF13A	170	PLEKHA3	230	GABRR1	290	TMEM19
51	GNS	111	FILIP1L	171	JAKMIP1	231	RYR2	291	CLCC1
52	C5orf56	112	ESCO2	172	PDE8A	232	TSEN34	292	NOL8
53	TEX2	113	BRMS1L	173	VLDLR	233	GSTM3	293	ACSBG2
54	CEP164	114	GUCA1C	174	PSMB2	234	DKK1	294	CNN1
55	ZFP90	115	PURB	175	CLUL1	235	PHF14	295	LRGUK
56	ESR1	116	ZNF512B	176	PHKA1	236	DPP8	296	KPNA2
57	CMKLR1	117	CPOX	177	SGTB	237	ZDHHC14	297	ASB13
58	PDE4D	118	FAM219B	178	PCDHB4	238	OTUD7B	298	TXNDC8
59	PHF12	119	ZIM3	179	XRRA1	239	C3orf18	299	C1QTNF3
60	RASGEF1A	120	NEK9	180	MCCD1	240	HFE	300	MANEA

APPENDIX D (Cont'd)

No.	Gene ID	No.	Gene ID	No.	Gene ID	No.	Gene ID	No.	Gene ID
301	PDCD4	366	HP1BP3	431	ZNF362	496	TRIP11	562	ZNRF3
302	RHOC	367	AGGF1	432	GABRE	497	ECT2	563	C1QTNF6
303	PRR16	368	HLF	433	EFCAB5	498	SLC22A5	564	ZBTB21
304	WDR45	369	IDI2	434	CHN2	500	PTDSS1	565	CEP55
305	BCLAF1	370	TDRKH	435	PPCDC	501	GRIN3A	566	OPCML
306	BORCS7	371	IGDCC3	436	THAP5	502	GALNT15	567	ENPEP
307	VWA5A	372	CDHR5	437	ZSCAN12	503	SAXO1	568	FBXL22
308	GLE1	373	GNB4	438	TOX	504	HORMAD2	569	SZT2
309	C19orf43	374	NR2C2	439	UBE2Q2	505	RAB24	570	DDIAS
310	PGBD5	375	ALDH1L2	440	DPY19L3	506	INPP5F	571	VENTX
311	NPC2	376	FAM218A	441	CHMP3	507	MTUS2	572	FAM134C
312	BHMT	377	RBM44	442	FMOD	508	IGHMBP2	573	HCFC2
313	ISM1	378	CCDC172	443	ALG9	509	FAM3C	574	RBBP7
314	GYPC	379	ASAH1	444	NOL4L	510	PAM	575	SH2D7
315	FBXO11	380	NFYA	445	SRRT	511	GLO1	576	PRDM10
316	LMOD3	381	PBLD	446	DMTF1	512	RBM34	577	DNAJC27
317	SYT13	382	RPP14	447	CALCOCO2	513	CLEC4D	578	PLAC1
318	POLE3	383	TIAM1	448	ZKSCAN1	514	MKNK2	579	KLHL18
319	PTK2	384	AHNAK	449	ITGB8	515	F2RL2	580	SMPD3
320	BTN3A1	385	KIF14	450	HSD17B7	516	SLC35F5	581	CHRFAM7A
321	OPHN1	386	FAM175A	451	ITGB2	517	URGCP	582	SART1
322	NELFE	387	DUSP18	452	CLLU1	518	CCR6	583	SCIMP
323	ZNF25	388	PPP1R37	453	FCRLB	519	KRTAP1-3	584	PIWIL2
324	HS2ST1	389	BCAP29	454	HN1	520	PNPLA3	585	MYCN
325	LRIT1	390	SLC35E2	455	KIF9	521	MYL12B	586	RSBN1
326	YWHAQ	391	HNRNPUL2	456	SETBP1	522	SHE	587	MOBP
327	RBFOX1	392	PRRX1	457	DERL2	523	GIPC2	588	GPR6
328	TMEM156	393	UNC5CL	458	ZXDC	524	SLC22A23	589	GDF11
329	ZNF239	394	TAGAP	459	CEP68	525	KDM7A	590	MCM2
330	HPCA	395	MMP23B	460	KANK4	526	RTN4IP1	591	MPP5
331	WDFY2	396	BCL2L15	461	FHAD1	527	FNDC9	592	TARDBP
332	GPD1	397	TMEM140	462	WDR26	528	UBN1	593	RAB34
333	C2orf44	398	TMEM25	463	BLOC1S5	529	ABCA10	594	ARSD
334	NPM1	399	STRIP1	464	FRMD4A	530	TBCEL	595	FAM50B
335	CRHBP	400	CHAF1A	465	DTWD2	531	HDAC4	596	ILDR1
336	BARX2	401	PLA2G6	466	GGT6	532	ST7L	597	ERLIN1
337	USP42	402	SLC46A2	467	ZC3H12C	533	GHITM	598	DGCR14
338	UQCRC1	403	HRH2	468	C19orf44	534	CCDC3	599	RNF24
339	TMEM202	404	CASC4	469	GP2	535	LIAS	600	IFNLR1
340	DDX18	405	SYTL4	470	ZFP1	536	IRX4	601	GLRX
341	LUC7L2	406	TNRC18	471	LAMA3	537	HSD17B3	602	NR0B2
342	SERF1A	407	ADRB1	472	SLC11A1	538	WDR92	603	IRF2
343	MFN2	408	DNAJC9	473	GPD1L	539	NHLRC2	604	ADAM18
344	RANBP6	409	TNKS2	474	MBD2	540	ADHFE1	605	SOX1
345	ARL16	410	TIMM17A	475	KIR2DL1	541	FAM117A	606	ZNF474
346	DNAJB3	411	BVES	476	ABHD14B	542	FAM57A	607	FBXL7
347	BAMBI	412	PI15	477	CLIC1	543	MYSM1	608	ADORA2B
348	UPF3A	413	FAM177A1	478	WDR82	544	MS4A4A	609	ZNF532
349	APPBP2	414	PAX1	479	AMPD3	545	TMEM38A	610	LRRC20
350	FAAP24	415	POLR3G	480	CA12	546	CAPNS1	611	SYDE1
351	SUV39H2	416	KLHL28	481	ZNF3	547	EID2B	612	MED12L
352	EPX	417	PROSER3	482	TMEM62	548	SPOP	613	TSLP
353	PALM2	418	CYB5R4	483	ALDH5A1	549	IPPK	614	C18orf32
354	HINT3	419	TRAPPC2	484	CDC40	550	BLVRA	615	RNH1
355	DAPP1	420	SERP1	485	SLC2A1	551	ZSCAN5B	616	MNT
356	RPL32	421	TACR2	486	PTGER4	552	PIP4K2A	617	ARL6IP4
357	POU6F1	422	EIF2S1	487	METAP1	553	MEX3A	618	RAPGEF5
358	LRRC10	423	TMPRSS4	488	MMP24	554	LDHAL6B	619	C7orf73
359	GNG12	424	ACADVL	489	CD109	555	RAB6A	620	CDC23
360	UFL1	425	GRIA1	490	NR2F2	556	TFAP4	621	PNLIPRP3
361	HPS4	426	ZNF627	491	RHO	557	GPR161	622	PPA2
362	VSIG10L	427	BDP1	492	CCDC47	558	TRIM66	623	MAP3K7
363	AGPS	428	LPPR4	493	SLC22A3	559	SNX9	624	RAB11FIP5
364	SAC3D1	429	TAOK2	494	P2RY14	560	NECAP1	625	NEUROD6
365	DRAM2	430	MCM10	495	WFS1	561	BNIP1	626	ADAMTSL5

APPENDIX D (Cont'd)

No.	Gene ID	No.	Gene ID	No.	Gene ID	No.	Gene ID	No.	Gene ID
627	RPS6KA1	692	MTCH2	757	ZNF230	822	CNGB3	887	ZFP30
628	UNC5D	693	FBXL4	758	GIMAP8	823	SDHAF3	888	TMUB2
629	KDM1B	694	MYO1D	759	NLRP3	824	RAI14	889	LGSN
630	ZNF830	695	TECPR1	760	PKD1	825	ZSCAN26	890	MAP6D1
631	TMEM186	696	DEPDC4	761	DIRC2	826	FGD4	891	EBF3
632	CPNE1	697	FAM13B	762	RAB3GAP1	827	TCHHL1	892	SPRED1
633	GCNT3	698	PLAUR	763	KIAA1522	828	COA5	893	SERF1B
634	SLC2A2	699	MAP3K11	764	ZNF500	829	PRAF2	894	CCSAP
635	BNIP3L	700	AGTR2	765	DIAPH2	830	ARL4C	895	RABGAP1
636	CFLAR	701	KAZN	766	TGM2	831	ITSN2	896	DRAM1
637	AJUBA	702	MSANTD4	767	NFATC4	832	TENM2	897	HK1
638	PTPMT1	703	JRKL	768	BMPR1A	833	PRRG4	898	WDR4
639	FAM210B	704	TMEM26	769	ARHGAP30	834	HOXB13	899	LHX8
640	STXBP1	705	SIK1	770	TIMP3	835	WDR5B	900	CD300LG
641	SLC46A3	706	FBXL13	771	TMEM86A	836	PQLC2L	901	TCEB3
642	TET1	707	KCNJ8	772	ABCG1	837	ORMDL3	902	ANKRD54
643	HILPDA	708	ABTB1	773	EPHA2	838	CORO2B	903	CRTC2
644	HOOK1	709	ZNF697	774	SQSTM1	839	PRDM4	904	ZBTB7A
645	PTGR2	710	ZNF776	775	MAN2A2	840	RAPGEFL1	905	ARHGEF18
646	ZNF737	711	KMT2C	776	FBXO27	841	DHX40	906	RBM46
647	NFYB	712	SI	777	YOD1	842	TP73	907	OSTM1
648	C9orf40	713	SERPINA10	778	PHKA2	843	LPXN	908	DCAF7
649	MXD1	714	TEX15	779	PTGDR	844	RBMS2	909	AGTR1
650	WDR37	715	REPIN1	780	SLC16A7	845	ENAM	910	WBSCR17
651	ATG12	716	ARL8B	781	CARD18	846	TRIM65	911	AGFG2
652	MTCP1	717	WDR73	782	NEDD4	847	MYO10	912	CDCA7
653	WHAMM	718	C14orf177	783	PPP3R2	848	TUSC2	913	MAP7D2
654	CLDN1	719	NXPH2	784	RBBP9	849	TMEM74	914	DDA1
655	SWI5	720	YPEL2	785	BTG1	850	CXCL8	915	SPECC1
656	AREL1	721	LETMD1	786	C5orf28	851	KHK	916	DENND5B
657	SPATA31D3	722	SYAP1	787	CFAP61	852	OTX1	917	RND3
658	MUC13	723	SRSF12	788	FBLIM1	853	ABHD15	918	TNS1
659	BCL6B	724	CHRNA7	789	SLC16A12	854	C10orf76	919	MAT2B
660	PLEKHG1	725	FANCD2	790	UBQLN4	855	GPR12	920	RBL2
661	SLC25A45	726	TMTC2	791	TRPV6	856	PBRM1	921	TRIB1
662	CTSA	727	MEF2C	792	TMEM223	857	PPP4R3B	922	CD46
663	UBFD1	728	DRD1	793	HLA-E	858	CC2D1A	923	RALGDS
664	HOXA11	729	KLF3	794	MRPL24	859	NR2E1	924	PRMT6
665	NLRP12	730	B3GNT5	795	BTN3A2	860	ZBTB43	925	ARFGAP2
666	ST3GAL5	731	SMNDC1	796	ADAT2	861	MTF1	926	ZKSCAN4
667	ORC4	732	USP30	797	NEUROG3	862	FAM169A	927	TMEM106A
668	TSPAN4	733	AMER2	798	REEP3	863	FBXO40	928	E2F5
669	AGBL2	734	TMPRSS11B	799	ZNF596	864	FNDC7	929	RBM12B
670	JMJD7	735	CCDC25	800	RBM4	865	MCM3	930	PRAP1
671	GPR68	736	KIF26B	801	ACBD5	866	ZNF502	931	KY
672	HTR2A	737	OTUD1	802	CLEC2B	867	LEF1	932	PPP1R26
673	MLH1	738	CCNA1	803	GMCL1	868	SNAPC4	933	MRPL17
674	POLD3	739	PANX1	804	IPO8	869	CTSS	934	CAMK2N1
675	AMPD2	740	FGL2	805	ULK1	870	KLHDC8B	935	AK2
676	RSF1	741	ADAMTS18	806	MAP3K14	871	LYST	936	THRSP
677	KRT12	742	AIFM1	807	C3AR1	872	CNRIP1	937	FRMPD2
678	NCOA7	743	CXCL14	808	FAM102B	873	SYNRG	938	DCDC2
679	CEP128	744	SLC17A1	809	CD36	874	HIST1H2BB	939	HAUS8
680	ZSWIM3	745	APOBEC4	810	USP24	875	GOLGA8A	940	IGSF5
681	RNF222	746	SENP1	811	KIF5B	876	LPCAT2	941	LONP2
682	PTCHD1	747	CILP	812	DCAF12	877	GDAP1	942	TXLNA
683	ANKMY2	748	ZNF473	813	TMEM233	878	PDPR	943	PBK
684	LRAT	749	C8A	814	DUSP2	879	UNK	944	DIS3
685	TTLL5	750	ZBED3	815	KCTD18	880	B4GALT6	945	HNRNPH3
686	BCL10	751	SMARCC2	816	CAPN14	881	ZRANB1	946	ZBTB6
687	INO80	752	GPR180	817	ZNF665	882	EDNRB	947	MCL1
688	FBXO39	753	SYNE2	818	DCLK1	883	RAD52	948	RGS9BP
689	LPPR5	754	LDHD	819	ASF1B	884	EGLN1	949	PRR11
690	ASB9	755	SQRDL	820	USP16	885	CACUL1	950	FAF2
691	IRF9	756	IRF1	821	CAMK2D	886	C11orf95	951	ZNF260

APPENDIX D (Cont'd)

No.	Gene ID	No.	Gene ID	No.	Gene ID	No.	Gene ID	No.	Gene ID
952	ZNF629	1017	WIPF2	1082	C9orf131	1147	POFUT1	1213	FRMD4B
953	KIF3B	1018	IL12RB2	1083	GTDC1	1148	GIMAP5	1214	NDN
954	SNAPIN	1019	KIR2DL3	1084	OR51E1	1149	RGMA	1215	KLHL24
955	MCCC2	1020	NFIB	1085	PAX8	1150	PGBD2	1216	RAB7A
956	PDCD2	1021	SASS6	1086	RELA	1151	RNF149	1217	GOLGA8B
957	DMRTA2	1022	AADACL3	1087	TMEM231	1152	RSRC2	1218	GPLD1
958	PSTPIP2	1023	C1GALT1C1L	1088	SLC2A4	1153	TBC1D2	1219	PIK3IP1
959	SUV420H2	1024	GTPBP4	1089	VSX1	1154	CSF2RA	1220	SLC39A6
960	TMEM100	1025	LHX3	1090	LYPD5	1155	PON2	1221	A4GNT
961	ARHGEF40	1026	SLC35D3	1091	HIF1AN	1156	NOV	1222	ABCD2
962	ZCCHC14	1027	SPTLC2	1092	METTL7A	1157	SSX2IP	1223	ABCE1
963	FAM120C	1028	EXPH5	1093	FGD5	1158	POMT2	1224	ABHD10
964	SEL1L	1029	C22orf29	1094	KIAA1468	1159	SS18L1	1225	ABHD5
965	LAMP5	1030	GALK2	1095	ZNF180	1160	ITGB3	1226	ABT1
966	ENTPD5	1031	MYO19	1096	C2CD2	1161	F2RL3	1227	ACP2
967	SLC7A2	1032	CLINT1	1097	EIF2AK4	1162	QRSL1	1228	ACSS2
968	MRPL43	1033	REEP5	1098	ABHD11	1163	HIST2H2BF	1229	ACTR8
969	TNFAIP1	1034	ZKSCAN8	1099	C7orf43	1164	BRF2	1230	ACVR1C
970	PKN2	1035	PLCL1	1100	SIRT3	1165	BLOC1S6	1231	ADAMTS8
971	GOLGA7	1036	CYLC1	1101	CYP26B1	1166	SPARC	1232	ADGRD1
972	VDR	1037	FAM177B	1102	RAB6C	1167	PAF1	1233	AFF1
973	SRCIN1	1038	ZNF766	1103	SUGP2	1168	SFT2D1	1234	AGO1
974	ZC3H13	1039	PTPN21	1104	BSCL2	1169	IL20RB	1235	AGTPBP1
975	FASLG	1040	SPATA31D4	1105	GNB5	1170	ALDOB	1236	AHCTF1
976	PHRF1	1041	SAP18	1106	IL16	1171	FBXO18	1237	AHI1
977	C9orf64	1042	PAFAH2	1107	ITFG1	1172	ZNF385A	1238	AIDA
978	KIR2DS4	1043	TSHZ3	1108	HAUS5	1173	STX16	1239	AIG1
979	KCNK2	1044	FOXL2	1109	ZC3H11A	1174	FEZ2	1240	AKAP8
980	KRT14	1045	ZBTB9	1110	UBTD2	1175	ZNF12	1241	AKAP9
981	ABCG2	1046	APOBEC3H	1111	LMO3	1176	GNGT2	1242	ALG6
982	PLA2G3	1047	PHYHIPL	1112	XKR6	1177	UHRF1	1243	ALKBH4
983	ZNF394	1048	IRF5	1113	AKAP5	1178	PLA2G4B	1244	AMIGO2
984	FAM13C	1049	GPR34	1114	BACH1	1179	HIP1	1245	ANKFY1
985	CLSTN1	1050	TRPM1	1115	PRPF38A	1180	C18orf54	1246	ANKRD10
986	DPF3	1051	ENDOD1	1116	C6orf201	1181	OLA1	1247	ANKRD12
987	ODF2	1052	ACTR1B	1117	REL	1182	MFHAS1	1248	ANTXR2
988	AARS2	1053	AUNIP	1118	ATP6V1C2	1183	SPRYD3	1249	ANXA6
989	H2AFJ	1054	SPACA4	1119	CD320	1184	METTL3	1250	AP2B1
990	CCDC36	1055	TRAF4	1120	INSR	1185	HS3ST4	1251	APCDD1
991	TMEM45B	1056	MICB	1121	CYTH3	1186	CAPN7	1252	APH1A
992	SHCBP1	1057	GLTSCR1L	1122	INTS4	1187	SLC14A1	1253	API5
993	PRPF4	1058	RASSF2	1123	ZFYVE21	1188	IRAK4	1254	APOBEC3F
994	DPP3	1059	KLHL8	1124	METTL4	1189	PRRT2	1255	APPL1
995	BTN1A1	1060	C16orf89	1125	TSSK1B	1190	PNN	1256	AQP1
996	FAM101A	1061	KLHL22	1126	HAAO	1191	TXNDC17	1257	ARAP2
997	SH2D3A	1062	ANKRD50	1127	RPL13A	1192	ARHGEF10	1258	ARHGAP1
998	ZDHHC8	1063	IPO7	1128	RAB11A	1193	SH2D5	1259	ARHGAP12
999	LEFTY2	1064	GNPDA2	1129	ARHGEF17	1194	C20orf197	1260	ARHGAP24
1000	ATP6V1B2	1065	ZNF674	1130	KIAA0319	1195	PPP4R4	1261	ARHGAP29
1001	CAPZA1	1066	LHX6	1131	SOWAHC	1196	C21orf58	1262	ARHGAP31
1002	ERP29	1067	SLC20A2	1132	PLEKHS1	1197	VAV3	1263	ARHGAP5
1003	SCARA5	1068	IQSEC1	1133	SETD7	1198	ALDH9A1	1264	ARHGAP8
1004	CAPN13	1069	KIAA1549	1134	SCUBE2	1199	EVPLL	1265	ARHGEF11
1005	PGRMC1	1070	STK4	1135	QSER1	1200	SUSD1	1266	ARL1
1006	ZNF597	1071	ISM2	1136	PIGO	1201	PITPNB	1267	ARL6IP1
1007	GDAP2	1072	PLAGL2	1137	P2RX4	1203	TRIM8	1268	ARX
1008	RPS20	1073	ALX4	1138	CCND2	1204	ATMIN	1269	ASAP2
1009	NUDT11	1074	CORIN	1139	ZNF443	1205	SEMA3C	1270	ASXL2
1010	ZNF229	1075	TMCC1	1140	AKT1	1206	RNF6	1271	ATAD3C
1011	CHRNA1	1076	SRPRB	1141	MCM6	1207	SCN5A	1272	ATF7IP2
1012	MLLT6	1077	NCAPD2	1142	RNMTL1	1208	RBL1	1273	ATG101
1013	NOD1	1078	TNN	1143	SLC6A9	1209	GRIA2	1274	ATG16L1
1014	PCGF5	1079	PPM1L	1144	CATSPER2	1210	APP	1275	ATG2B
1015	GRAPL	1080	TEC	1145	CISD1	1211	SNX21	1276	ATL3
1016	P2RY6	1081	APCS	1146	C9orf131	1212	C12orf73	1277	ATM

APPENDIX D (Cont'd)

No.	Gene ID	No.	Gene ID	No.	Gene ID	No.	Gene ID	No.	Gene ID
1278	ATP12A	1343	CDK2	1408	DAPL1	1473	FAM26E	1538	HEY2
1279	ATP13A3	1344	CDKL2	1409	DCAF8	1474	FAM45A	1539	HIF1A
1280	ATP6V0A2	1345	CDKN1A	1410	DCTPP1	1475	FAM60A	1540	HIF3A
1281	ATP6V0E1	1346	CDS1	1411	DDX26B	1476	FAM63B	1541	HIGD1A
1282	ATP6V1C1	1347	CECR1	1412	DDX5	1477	FAM69C	1542	HIP1R
1283	ATP6V1D	1348	CEMIP	1413	DEDD	1478	FAM73B	1543	HIST1H2BD
1284	ATP7B	1349	CENPN	1414	DEDD2	1479	FAM78A	1544	HLA-F
1285	ATPAF2	1350	CENPO	1415	DERL1	1480	FANCM	1545	HMGN3
1286	ATXN1L	1351	CENPQ	1416	DGKG	1481	FASTK	1546	HOXA3
1287	B3GALT2	1352	CEP120	1417	DHDDS	1482	FBXL5	1547	HPGD
1288	BAG5	1353	CEP57	1418	DHRS12	1483	FBXO10	1548	HS3ST5
1289	BANK1	1354	CEP70	1419	DIRC1	1484	FBXO31	1549	HSDL1
1290	BBX	1355	CEP85L	1420	DKC1	1485	FBXO48	1550	HSPA13
1291	BCO1	1356	CEP97	1421	DLEC1	1486	FBXW11	1551	HSPA4L
1292	BCO2	1357	CERCAM	1422	DLG5	1487	FGF5	1552	HSPA8
1293	BDH1	1358	CHD5	1423	DNAH12	1488	FGF9	1553	HTR1F
1294	BEST3	1359	CHP2	1424	DNAJB9	1489	FIBIN	1554	ICMT
1295	BHLHE41	1360	CHRM2	1425	DNAJC16	1490	FIGNL1	1555	IER3IP1
1296	BICD2	1361	CHST11	1426	DNAJC3	1491	FJX1	1556	IFNAR2
1297	BIRC5	1362	CHTOP	1427	DNAJC30	1492	FOXA1	1557	IGF1
1298	BIRC6	1363	CIDEA	1428	DOCK7	1493	FOXG1	1558	IL2
1299	BMP8B	1364	CIT	1429	DOK5	1494	FOXP1	1559	ING1
1300	BMPR2	1365	CLTC	1430	DPYSL2	1495	FOXQ1	1560	INHBC
1301	BOC	1366	CMPK2	1431	DTNBP1	1496	FRMPD4	1561	IQSEC2
1302	BTBD10	1367	CNIH1	1432	DUSP8	1497	FUCA2	1562	IRF8
1303	BTG3	1368	CNNM3	1433	DUT	1498	FZD3	1563	ITGA11
1304	C11orf58	1369	CNP	1434	E2F1	1499	GAB1	1564	ITGB4
1305	C12orf74	1370	CNPY2	1435	E2F2	1500	GABBR1	1565	ITK
1306	C15orf57	1371	CNTN4	1436	EFCAB1	1501	GADD45B	1566	ITPR1
1307	C18orf25	1372	CNTNAP3	1437	EFCAB14	1502	GALNT13	1567	JPH1
1308	C1orf52	1373	COBLL1	1438	EFHC1	1503	GALNT16	1568	KANSL1L
1309	C3orf35	1374	COL17A1	1439	EFNA5	1504	GAPVD1	1569	KATNAL1
1310	C3orf38	1375	COL4A4	1440	EFR3A	1505	GBF1	1570	KBTBD2
1311	C5AR2	1376	COL5A1	1441	EGLN3	1506	GBP3	1571	KBTBD6
1312	C6orf141	1377	COL8A2	1442	EGR2	1507	GCA	1572	KCNC2
1313	C7orf60	1378	COMMD6	1443	EIF4G2	1508	GCC2	1573	KCND2
1314	CAAP1	1379	COX10	1444	EIF5A2	1509	GJA1	1574	KCNN2
1315	CAB39L	1380	COX7A2L	1445	EIF5AL1	1510	GLDN	1575	KCTD10
1316	CAMSAP1	1381	COX8C	1446	ELK3	1511	GLS	1576	KDM2A
1317	CANT1	1382	CPEB1	1447	ELK4	1512	GMNC	1577	KDM5B
1318	CASD1	1383	CPEB2	1448	EMC3	1513	GNPTAB	1578	KIAA0141
1319	CASP7	1384	CPEB4	1449	EMX2	1514	GOLM1	1579	KIAA1191
1320	CASP8	1385	CRCT1	1450	ENPP4	1515	GORASP2	1580	KIAA1715
1321	CBLL1	1386	CREB5	1451	ENPP5	1516	GOSR1	1581	KIAA1919
1322	CBX1	1387	CREG2	1452	EP400	1517	GPATCH2	1582	KIF16B
1323	CCDC121	1388	CRIPT	1453	EPB41L5	1518	GPR176	1583	KIF23
1324	CCDC129	1389	CRY2	1454	EPC2	1519	GPR45	1584	KIF5A
1325	CCDC137	1390	CSDE1	1455	EPHB4	1520	GPT2	1585	KIT
1326	CCDC142	1391	CSF1	1456	ERBB4	1521	GRAMD1A	1586	KLF10
1327	CCDC176	1392	CSGALNACT1	1457	EREG	1522	GRB10	1587	KLHL20
1328	CCDC71L	1393	CSNK1G1	1458	ERVFRD-1	1523	GRPEL2	1588	KLK7
1329	CCL1	1394	CSRNP3	1459	ETNK1	1524	GSTA1	1589	KLRD1
1330	CCT5	1395	CSTB	1460	F3	1525	GUCY1A3	1590	KMO
1331	CCT7	1396	CTAGE4	1461	FAM104B	1526	GZMK	1591	KMT2B
1332	CD164L2	1397	CTAGE6	1462	FAM109B	1527	HADHA	1592	KREMEN1
1333	CD177	1398	CTAGE9	1463	FAM117B	1528	HADHB	1593	KRT23
1334	CD69	1399	CTSK	1464	FAM129A	1529	HAND2	1594	KRT38
1335	CD83	1400	CXCL6	1465	FAM129C	1530	HARS	1595	KRT76
1336	CD8A	1401	CXorf57	1466	FAM13A	1531	HAS2	1596	KRT81
1337	CDADC1	1402	CYP19A1	1467	FAM155A	1532	HAUS6	1597	KRTAP10-4
1338	CDC37L1	1403	CYP27B1	1468	FAM160B2	1533	HBP1	1598	KRTAP4-7
1339	CDC42BPA	1404	CYP4A11	1469	FAM179B	1534	HDHD2	1599	L3MBTL3
1340	CDCA4	1405	CYS1	1470	FAM19A1	1535	HECA	1600	LACE1
1341	CDH13	1406	CYTH4	1471	FAM210A	1536	HEG1	1601	LAGE3
1342	CDK12	1407	DAB2	1472	FAM216B	1537	HERC3	1602	LAP3

APPENDIX D (Cont'd)

No.	Gene ID	No.	Gene ID	No.	Gene ID	No.	Gene ID	No.	Gene ID
1603	LAPTM4A	1668	MOG	1733	PATE2	1798	RAB32	1863	SLC25A48
1604	LASP1	1669	MON1B	1734	PAX6	1799	RAC2	1864	SLC29A2
1605	LCE2B	1670	MON2	1735	PCDH20	1800	RAG1	1865	SLC35D1
1606	LDLR	1671	MPDU1	1736	PCNP	1801	RAPGEF4	1866	SLC35F3
1607	LDLRAD3	1672	MPV17L	1737	PCYOX1	1802	RASD1	1867	SLC41A1
1608	LIMK1	1673	MREG	1738	PDGFD	1803	RASGRP4	1868	SLC43A3
1609	LINC01588	1674	MRGPRX2	1739	PDRG1	1804	RASL11B	1869	SLC45A4
1610	LIPI	1675	MRGPRX3	1740	PDZD11	1805	RASL12	1870	SLC4A5
1611	LNX2	1676	MS4A10	1741	PEAR1	1806	RAX	1871	SLC4A9
1612	LPGAT1	1677	MSL2	1742	PER1	1807	RBAK	1872	SLC6A2
1613	LRIG1	1678	MSMO1	1743	PEX13	1808	RBM20	1873	SLTM
1614	LRPPRC	1679	MSTO1	1744	PEX19	1809	RCCD1	1874	SMAD5
1615	LRRC45	1680	MTMR11	1745	PF4V1	1810	RCHY1	1875	SMAD7
1616	LRRC55	1681	MVK	1746	PFKFB3	1811	RD3	1876	SMCO1
1617	LRRC57	1682	MYB	1747	PGM2L1	1812	RDH11	1877	SMCR8
1618	LRRC61	1683	MYBL1	1748	PGM5	1813	REEP1	1879	SMIM5
1619	LSM14B	1684	MYH10	1749	PGP	1814	REPS2	1880	SMIM8
1620	LTBP2	1685	MYO5C	1750	PHF23	1815	RFX5	1881	SMOC1
1621	LYN	1686	MYOZ2	1751	PID1	1816	RFXAP	1882	SMOC2
1622	LYPD6B	1687	MYPN	1752	PIP4K2C	1817	RGPD4	1883	SNCA
1623	LYPD8	1688	MYRF	1753	PITPNA	1818	RHOXF1	1884	SNX31
1624	LYRM1	1689	MYT1	1754	PKIA	1819	RILPL1	1885	SOD2
1625	LYRM9	1690	N4BP2L2	1755	PKMYT1	1820	RNASEH2B	1886	SORL1
1626	M6PR	1691	NAA50	1756	PKNOX1	1821	RNF114	1887	SOX4
1627	MAATS1	1692	NAB1	1757	PLA1A	1822	RNF121	1888	SP3
1628	MAB21L1	1693	NABP1	1758	PLA2G4C	1823	RNF128	1889	SP4
1629	MACROD2	1694	NAGK	1759	PLA2G5	1824	RNF145	1890	SPARCL1
1630	MAML3	1695	NCF2	1760	PLEKHA2	1825	RNF150	1891	SPG7
1631	MAN1C1	1696	NDEL1	1761	PNKD	1826	RNF216	1892	SRGAP1
1632	MAP10	1697	NDUFB9	1762	POC1A	1827	RORC	1893	SRGAP3
1633	MAP1B	1698	NEDD4L	1763	POGZ	1828	RPA2	1894	SRSF7
1634	MAP2	1699	NETO2	1764	POLE4	1829	RPRD2	1895	SSH1
1635	MAP3K12	1700	NEUROD4	1765	PPP1R15B	1830	RRN3	1896	SSH2
1636	MAP3K5	1701	NEUROG1	1766	PPP1R21	1831	RSRP1	1897	STARD7
1637	MAP3K8	1702	NEUROG2	1767	PPP1R3B	1832	RUNDC1	1898	STAT3
1638	MAP6	1703	NFATC2IP	1768	PPP1R9A	1833	RUNDC3A	1899	STC2
1639	MAP7	1704	NFE2L1	1769	PPP2R1B	1834	RYR3	1900	STK11IP
1640	MAPK6	1705	NIPA1	1770	PPP3CA	1835	SACS	1901	STK38
1641	MAPRE1	1706	NKIRAS1	1771	PPP6R2	1836	SALL1	1902	STK38L
1642	MAPRE3	1707	NLRC5	1772	PPTC7	1837	SATL1	1903	STOM
1643	MARK2	1708	NMUR2	1773	PRDM16	1838	SCAMP4	1904	STX12
1644	MASTL	1709	NNT	1774	PRDM6	1839	SCAMP5	1905	STX3
1645	MBD4	1710	NOL4	1775	PRIMA1	1840	SCCPDH	1906	STXBP5L
1646	MBOAT2	1711	NOX4	1776	PRKAA1	1841	SCLT1	1907	STYX
1647	MCC	1712	NPAS2	1777	PRNP	1842	SCN2B	1908	SUMF1
1648	MCHR2	1713	NPAT	1778	PRR14L	1843	SCP2	1909	SVIL
1649	MCM7	1714	NPFFR2	1779	PRR15	1844	SCRN3	1910	SWSAP1
1650	MCMDC2	1715	NPHP3	1780	PRR18	1845	SDHAF2	1911	SYBU
1651	MCPH1	1716	NPLOC4	1781	PRR5	1846	SECISBP2L	1912	SYT10
1652	MCTS1	1717	NPY2R	1782	PSD	1847	SEMA3G	1913	SYT16
1653	MDFIC	1718	NR2C2AP	1783	PSG1	1848	SEMA4B	1914	SYT6
1654	MDM4	1719	NR2E3	1784	PSG3	1849	SEMA7A	1915	TACC1
1655	MED14	1720	NRBP1	1785	PTGES3	1850	SERPINB13	1916	TAOK3
1656	METTL14	1721	NRBP2	1786	PTGFRN	1851	SERPINB8	1917	TAX1BP1
1657	MEX3D	1722	NRSN1	1787	PTH	1852	SERPINB9	1918	TBC1D1
1658	MFF	1723	NUTM2G	1788	PTHLH	1853	SFMBT1	1919	TBC1D12
1659	MFSD6	1724	NXF1	1789	PTPN4	1854	SGIP1	1920	TBC1D15
1660	MIDN	1725	OGT	1790	PTPN9	1855	SGMS1	1921	TBC1D2B
1661	MIOS	1726	OPRL1	1791	PTPRE	1856	SH3PXD2A	1922	TBC1D9
1662	MKNK1	1727	ORAI1	1792	PUDP	1857	SHTN1	1923	TBX19
1663	MKRN2	1728	OSM	1793	PUF60	1858	SLC12A3	1924	TBXAS1
1664	MKRN3	1729	PAG1	1794	PXDN	1859	SLC17A7	1925	TCEAL7
1665	MLC1	1730	PAIP2B	1795	PXYLP1	1860	SLC25A12	1926	TES
1666	MMAA	1731	PAK6	1796	QDPR	1861	SLC25A2	1927	TESC
1667	MMP2	1732	PAN3	1797	RAB10	1862	SLC25A23	1928	TFAP2C

APPENDIX D (Cont'd)

No.	Gene ID	No.	Gene ID	No.	Gene ID	No.	Gene ID	No.	Gene ID
1929	THAP1	1954	TNFRSF10D	1979	TUB	2014	WDFY1	2039	ZNF280C
1930	THBS1	1955	TNFRSF19	1980	TXLNB	2015	WDR1	2040	ZNF304
1931	THBS2	1956	TNFRSF21	1981	TXNDC9	2016	WEE1	2041	ZNF441
1932	THRA	1957	TNFRSF8	1982	U2SURP	2017	XRN1	2042	ZNF460
1933	TIGAR	1958	TNFSF11	1983	UBA6	2018	YTHDF2	2043	ZNF514
1934	TIMP2	1959	TNKS1BP1	1984	UBAP1	2019	YWHAG	2044	ZNF543
1935	TLR4	1960	TNRC6C	1985	UBC	2020	ZBTB38	2045	ZNF583
1936	TLR7	1961	TOM1L1	1986	UBE2E2	2021	ZBTB4	2046	ZNF671
1937	TM2D2	1962	TOM1L2	1987	UBE3A	2022	ZC3H14	2047	ZNF684
1938	TMBIM6	1963	TOPORS	1988	UBXN2A	2023	ZC3H7B	2048	ZNF704
1939	TMEM127	1964	TP53RK	1999	UEVLD	2024	ZDHHC9	2049	ZNF71
1940	TMEM133	1965	TPRG1L	2000	UFD1L	2025	ZFAND4	2050	ZNF74
1941	TMEM136	1966	TRAF6	2001	ULBP2	2026	ZFC3H1	2051	ZNF791
1942	TMEM159	1967	TRAPPC4	2002	UNC5C	2027	ZFP91	2052	ZNF80
1943	TMEM167B	1968	TRERF1	2003	USP15	2028	ZFYVE9	2053	ZNF831
1944	TMEM168	1969	TRIM3	2004	UXS1	2029	ZHX2	2054	ZNF879
1945	TMEM170B	1970	TRIM37	2005	VAMP3	2030	ZNF112	2055	ZNHIT6
1946	TMEM182	1971	TRIP10	2006	VAPA	2031	ZNF134	2056	ZPLD1
1947	TMEM192	1972	TROVE2	2007	VCP	2032	ZNF140	2057	ZSCAN2
1948	TMEM50A	1973	TRPC1	2008	VCPKMT	2033	ZNF202	2058	ZSCAN20
1949	TMEM55B	1974	TSC22D1	2009	VPS26A	2034	ZNF217		
1950	TMOD2	1975	TSG101	2010	VPS50	2035	ZNF250		
1951	TMPPE	1976	TSPAN9	2011	VWA8	2036	ZNF266		
1952	TMPRSS11A	1977	TTC17	2012	WARS2	2037	ZNF276		
1953	TNF	1978	TTC38	2013	WBSCR22	2038	ZNF277		

*Genes in purple and yellow: putative genes of the miR-519/-520/-302 and miR-520/-302 families, respectively.

APPENDIX E. KEGG pathways of the group I C19MC-AAGUGC-miRNAs

No	KEGG designation	Pathway	No. target genes
1	**hsa04151**	**PI3K-Akt signaling pathway**	**56**
2	hsa05200	Pathways in cancer	54
3	hsa04144	Endocytosis	39
4	**hsa04010**	**MAPK signaling pathway**	**38**
5	**hsa04014**	**Ras signaling pathway**	**37**
6	hsa05205	Proteoglycans in cancer	30
7	hsa05162	Measles	24
	hsa04068	**FoxO signaling pathway**	**24**
9	hsa04110	Cell cycle	22
	hsa04360	Axon guidance	22
	hsa04380	Osteoclast differentiation	22
12	**hsa04350**	**TGF-beta signaling pathway**	**19**
	hsa04152	**AMPK signaling pathway**	**19**
14	**hsa04066**	**HIF-1 signaling pathway**	**18**
	hsa04660	**T cell receptor signaling pathway**	**18**
	hsa04931	Insulin resistance	18
17	**hsa04668**	**TNF signaling pathway**	**17**
18	hsa00564	Glycerophospholipid metabolism	16
19	**hsa04662**	**B cell receptor signaling pathway**	**14**
	hsa05133	Pertussis	14
	hsa04612	Antigen processing and presentation	14
22	hsa04210	Apoptosis	13
23	hsa05134	Legionellosis	11
24	hsa03030	DNA replication	9
		Total	**568**

The ten signaling pathways are shown in bold letters.

APPENDIX F. Predicted target genes of group I C19MC-AAGUGC-miRNAs related to apoptosis

Gene ID	Gene name
ACTN2	Actinin, alpha 2
AHI1	Abelson helper integration site 1
AIFM1	Apoptosis-inducing factor, mitochondrion-associated, 1
AKT1	AKT serine/threonine kinase 1
AKTIP	AKT interacting protein
ALX4	ALX homeobox 4
AMIGO2	Adhesion molecule with Ig-like domain 2
APH1A	Anterior pharynx defective 1 homolog A (C. elegans)
API5	API5-like 1; apoptosis inhibitor 5
AQP1	Aquaporin 1 (Colton blood group)
AREL1	Apoptosis resistant E3 ubiquitin protein ligase 1
ARHGEF11	Rho guanine nucleotide exchange factor (GEF) 11
ARHGEF17	Rho guanine nucleotide exchange factor (GEF) 17
ARHGEF18	Rho/Rac guanine nucleotide exchange factor (GEF) 18
ARL6IP1	ADP-ribosylation factor-like 6 interacting protein 1
ATM	ATM serine/threonine kinase
BCAP29	B-cell receptor-associated protein 29
BCL10	B-cell CLL/lymphoma 10
BCL2L15	BCL2-like 15
BCL6	B-cell CLL/lymphoma 6
BCLAF1	Similar to Bcl-2-associated transcription Factor 1 (Btf); BCL2-associated transcription factor 1
BECN1	Beclin 1, autophagy related
BIRC5	baculoviral IAP repeat containing 5
BIRC6	Baculoviral IAP repeat-containing 6
BMP6	Bone morphogenetic protein 6
BMP8B	Bone morphogenetic protein 8b
BNIP1	BCL2/adenovirus E1B 19kDa interacting protein 1
BNIP3L	BCL2/adenovirus E1B 19kDa interacting protein 3-like
C3ORF38	Chromosome 3 open reading frame 38
CAAP1	Caspase activity and apoptosis inhibitor 1
CARD18	Caspase recruitment domain family, member 18
CASP7	Caspase 7, apoptosis-related cysteine peptidase
CASP8	CASP8 and FADD-like apoptosis regulator
CCND2	Cyclin D2
CD44	CD44 molecule (Indian blood group)
CDCA7	Cell division cycle associated 7
CDKN1A	Cyclin-dependent kinase inhibitor 1A (p21, Cip1)
CFLAR	CASP8 and FADD like apoptosis regulator
CHMP3	Charged multivesicular body protein 3
CHST11	Carbohydrate (chondroitin 4) sulfotransferase 11
CIDEA	Cell death-inducing DFFA-like effector a
CSRNP3	Cysteine-serine-rich nuclear protein 3
CXCR4	Chemokine (C-X-C motif) receptor 4
DAB2	DAB2, clathrin adaptor protein
DAD1	Defender against cell death 1
DAPK2	Death-associated protein kinase 2

APPENDIX D (Cont'd)

Gene ID	Gene name
DcR2/TNFRSF10D	Tumor necrosis factor receptor superfamily, member 10d, decoy with truncated death domain
DDIAS	DNA damage-induced apoptosis suppressor
DEDD2	Death effector domain containing 2
DLG5	Discs large MAGUK scaffold protein 5
DNAJC3	DnaJ (Hsp40) homolog, subfamily C, member 3
DNM1L	Dynamin 1-like
DRAM1	DNA-damage regulated autophagy modulator 1
DRAM2	DNA-damage regulated autophagy modulator 2
ECT2	Epithelial cell transforming 2
EDNRB	Endothelin receptor type B
EFNA5	Ephrin-A5
EGLN3	Egl-9 family hypoxia-inducible factor 3
ERBB4	Erb-b2 receptor tyrosine kinase 4
ESR1	Estrogen receptor 1
FASLG	Fas ligand (TNF superfamily, member 6)
FBXO10	F-box protein 10
FCMR	Fc fragment of IgM receptor
FGD4	FYVE, RhoGEF and PH domain containing 4
FIGNL1	Fidgetin-like 1
FLT4	Fms-related tyrosine kinase 4
FOXL2	Forkhead box L2
GADD45B	Growth arrest and DNA-damage-inducible, beta
GDF11	Growth differentiation factor 11
GHITM	Growth hormone inducible transmembrane protein
GJA1	Gap junction protein, alpha 1, 43kDa
GLO1	Glyoxalase I
GPLD1	Glycosylphosphatidylinositol specific phospholipase D1
HIF3A	Hypoxia inducible factor 3, alpha subunit
HIGD1A	HIG1 hypoxia inducible domain family, member 1A
HIP1	Huntingtin interacting protein 1 related
HIP1R	Huntingtin interacting protein 1 related
HIPK3	Homeodomain interacting protein kinase 3
IGF1	Insulin-like growth factor 1 (somatomedin C)
IL2	Interleukin 2
INHBC	Inhibin, beta C
IRF1	Interferon regulatory factor 1
IRF5	Interferon regulatory factor 5
ITGB2	Integrin, beta 2
KIF14	Kinesin family member 14
KIT	KIT proto-oncogene receptor tyrosine kinase
KLHL20	Kelch-like family member 20
KLLN	Killin, p53-regulated DNA replication inhibitor
LEF1	Lymphoid enhancer-binding factor 1
LEFTY1	Left-right determination factor 1
LEFTY2	Left-right determination factor 2
LHX3	LIM homeobox 3
MAGEH1	Melanoma antigen family H1
MALT1	Mucosa associated lymphoid tissue lymphoma translocation gene 1
MAP3K5	Mitogen-activated protein kinase kinase kinase 5
MAP3K8	Mitogen-activated protein kinase kinase kinase 8
MAP3K9	Mitogen-activated protein kinase kinase kinase 9
MCL1	Myeloid cell leukemia sequence 1 (BCL2-related)

APPENDIX D (Cont'd)

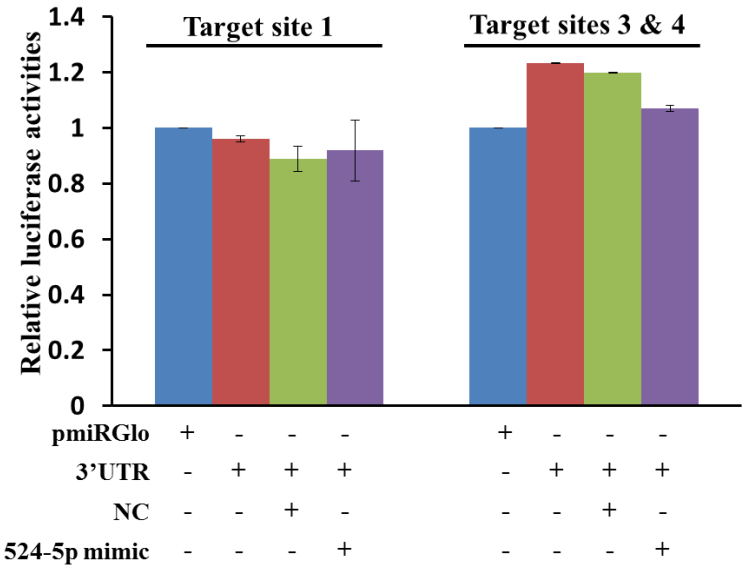
Gene ID	Gene name
MCM2	Minichromosome maintenance complex component 2
MDM4	MDM4, p53 regulator
MEF2C	Myocyte enhancer factor 2C
MFN2	Mitofusin 2
MTCH2	Mitochondrial carrier 2
NIK/MAP3K14	Mitogen-activated protein kinase kinase kinase 14
NLRP3	NLR family, pyrin domain containing 3
NOD1	Nucleotide-binding oligomerization domain containing 1
NOX4	NADPH oxidase 4
NPM1	Nucleophosmin
NR2E1	Nuclear receptor subfamily 2, group E, member 1
OGT	O-linked N-acetylglucosamine (GlcNAc) transferase
PAFAH2	Platelet-activating factor acetylhydrolase 2
PAK6	P21 protein (Cdc42/Rac)-activated kinase 6
PAX8	Paired box 8
PDCD2	Programmed cell death 2
PDCD4	Programmed cell death 4
PIK3CA	Phosphoinositide-3-kinase, catalytic, alpha polypeptide
PKN2	Protein kinase N2
PLAC8	Placenta-specific 8
PLAUR	Plasminogen activator, urokinase receptor
PRKAA1	Protein kinase, AMP-activated, alpha 1 catalytic subunit
PRNP	Prion protein
PTK2	Protein tyrosine kinase 2
PUF60	Poly-U binding splicing factor 60KDa
RBM25	RNA binding motif protein 25
RelA	V-rel reticuloendotheliosis viral oncogene homolog A (avian)
RNF216	Ring finger protein 216
RPS6KA1	Ribosomal protein S6 kinase, 90kDa, polypeptide 1
RPS6KA3	Ribosomal protein S6 kinase, 90kDa, polypeptide 3
SAP18	Sin3A-associated protein, 18kDa
SERPINB9	Serpin peptidase inhibitor, clade B (ovalbumin), member 9
SFRP4	Secreted frizzled-related protein 4
SGMS1	Sphingomyelin synthase 1
SHC4	SHC (Src homology 2 domain containing) family, member 4
SLC40A1	Solute carrier family 40 (iron-regulated transporter), member 1
SLTM	SAFB-like, transcription modulator
SMNDC1	Survival motor neuron domain containing 1
SNCA	Synuclein, alpha
SOD1	Superoxide dismutase 1, soluble
SOD2	Superoxide dismutase 2, mitochondrial
SOS1	SOS Ras/Rac guanine nucleotide exchange factor 1
SOX4	SRY (sex determining region Y)-box 4
SQSTM1	Sequestosome 1
STAT3	signal transducer and activator of transcription 3
STK17B	Serine/threonine kinase 17b
STK4	Serine/threonine kinase 4
TAK1/MAP3K7	Mitogen-activated protein kinase kinase kinase 7
TAOK2	TAO kinase 2
TAX1BP1	Tax1 (human T-cell leukemia virus type I) binding protein 1

APPENDIX D (Cont'd)

Gene ID	Gene name
TFAP4	Transcription factor AP-4 (activating enhancer binding protein 4)
TGFBR2	Transforming growth factor, beta receptor II
TGM2	Transglutaminase 2
TIAM1	T-cell lymphoma invasion and metastasis 1
TIGAR	TP53 induced glycolysis regulatory phosphatase
TLR4	Toll-like receptor 4
TMBIM6	Transmembrane BAX inhibitor motif containing 6
TNF/TNFα	Tumor necrosis factor (TNF superfamily, member 2)
TNFAIP1	Tumor necrosis factor, alpha-induced protein 1 (endothelial)
TNFRSF10B	Tumor necrosis factor receptor superfamily, member 10b
TNFRSF19	Tumor necrosis factor receptor superfamily, member 19
TNFRSF21	Tumor necrosis factor receptor superfamily, member 21
TNFRSF8	Tumor necrosis factor receptor superfamily, member 8
TOPORS	Topoisomerase I binding, arginine/serine-rich, E3 ubiquitin protein ligase
TOX3	TOX high mobility group box family member 3
TP73	Tumor protein p73
TRAF4	TNF receptor-associated factor 4
TRAF6	TNF receptor-associated factor 6
TRIM24	Tripartite motif containing 24
TSP-1/THBS1	Thrombospondin 1
TXNIP	Thioredoxin interacting protein
UBC	Ubiquitin C
UNC5C	Unc-5 netrin receptor C
UNC5D	Unc-5 netrin receptor D
USP47	Ubiquitin specific peptidase 47
VAV3	Vav 3 guanine nucleotide exchange factor
VCP	Valosin containing protein
WDR92	WD repeat domain 92
ZNF385A	Zinc finger protein 385A
ZNF443	Zinc finger protein 443
ZNF830	Zinc finger protein 830

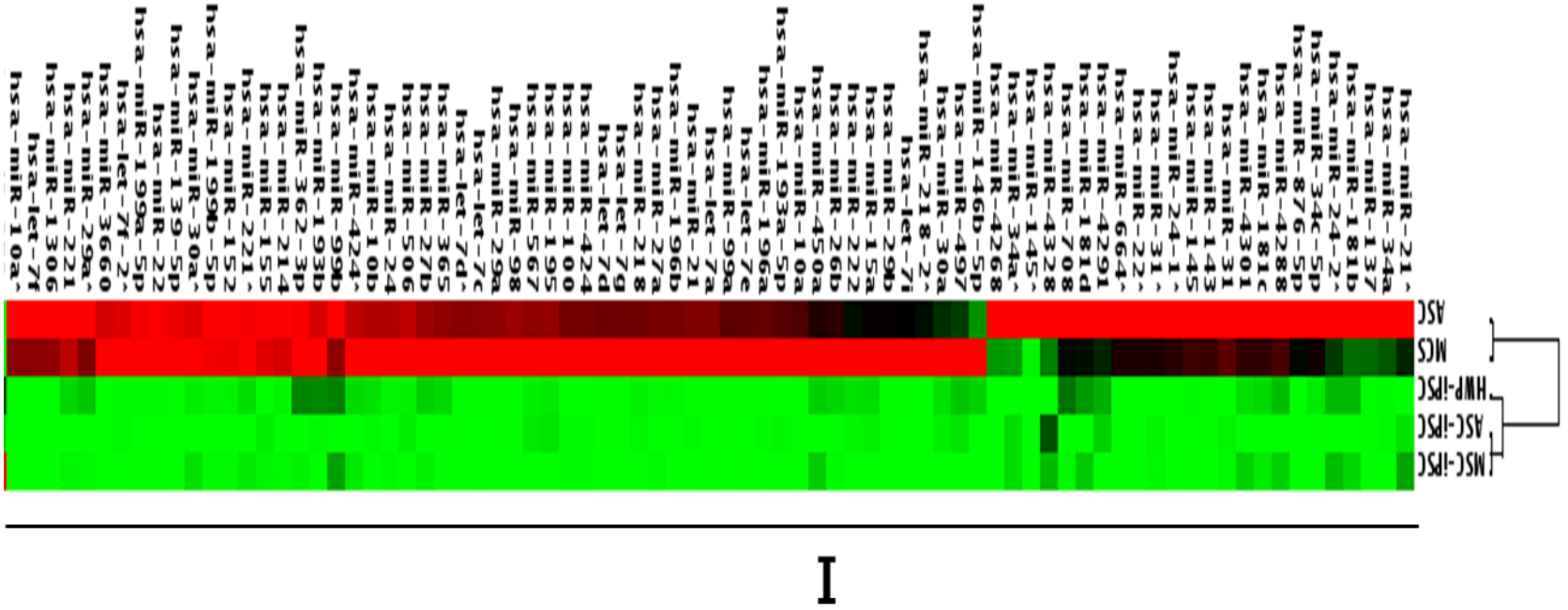
*Bold genes: Genes related to survival pathway

APPENDIX G



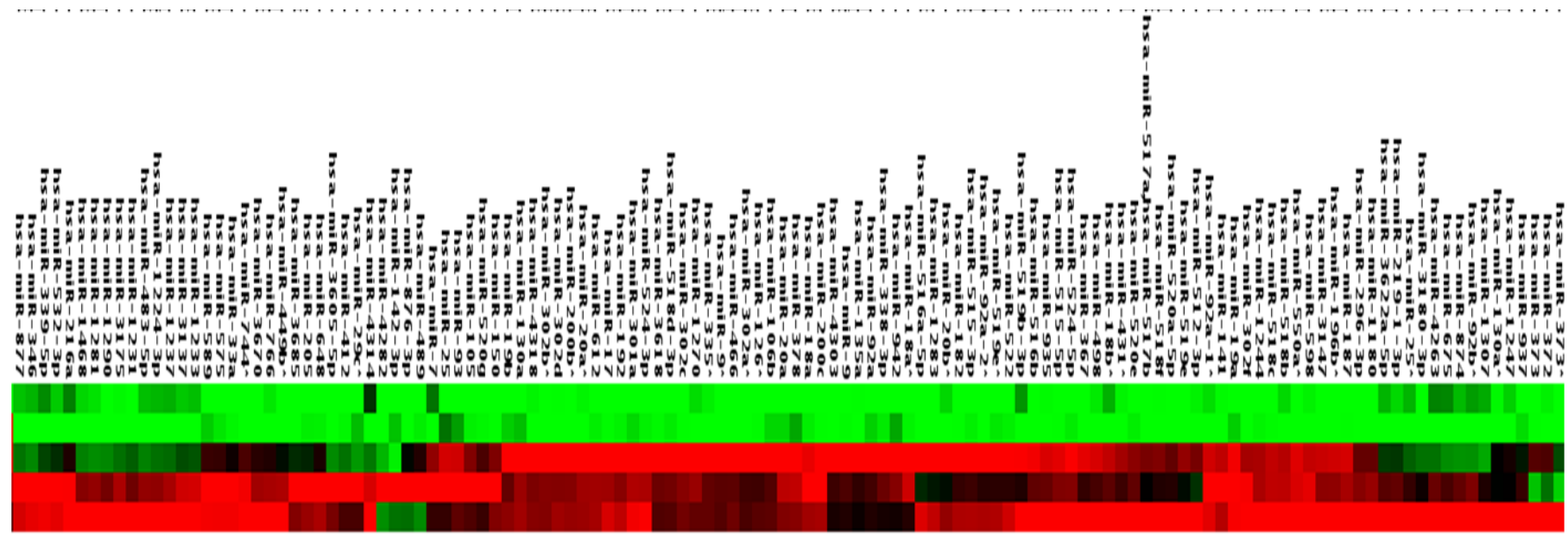
Luciferase assays to experimentally validate miR-524-5p targeting TP53INP1 at the putative target sites 1, 3 and 4. HCT-15 transfected with the luciferase constructs containing target sites 1, 3 and 4 or in the presence of the miR-524-5p mimic or a validated negative control (NC) was performed before luciferase assays. The data shown were derived from two independent experiments in duplicates.

APPENDIX H



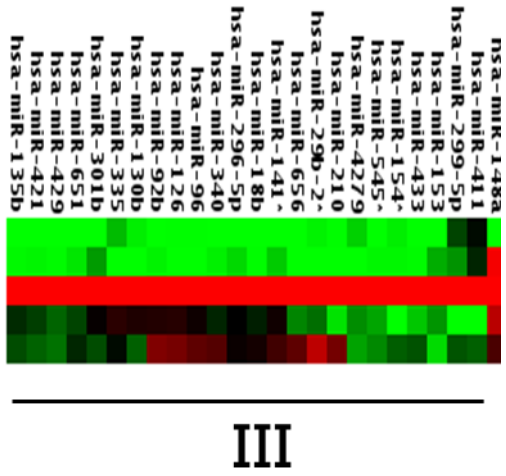
Enlarge Cluster I of Figure 4.2 from the text

APPENDIX H (Cont'd)



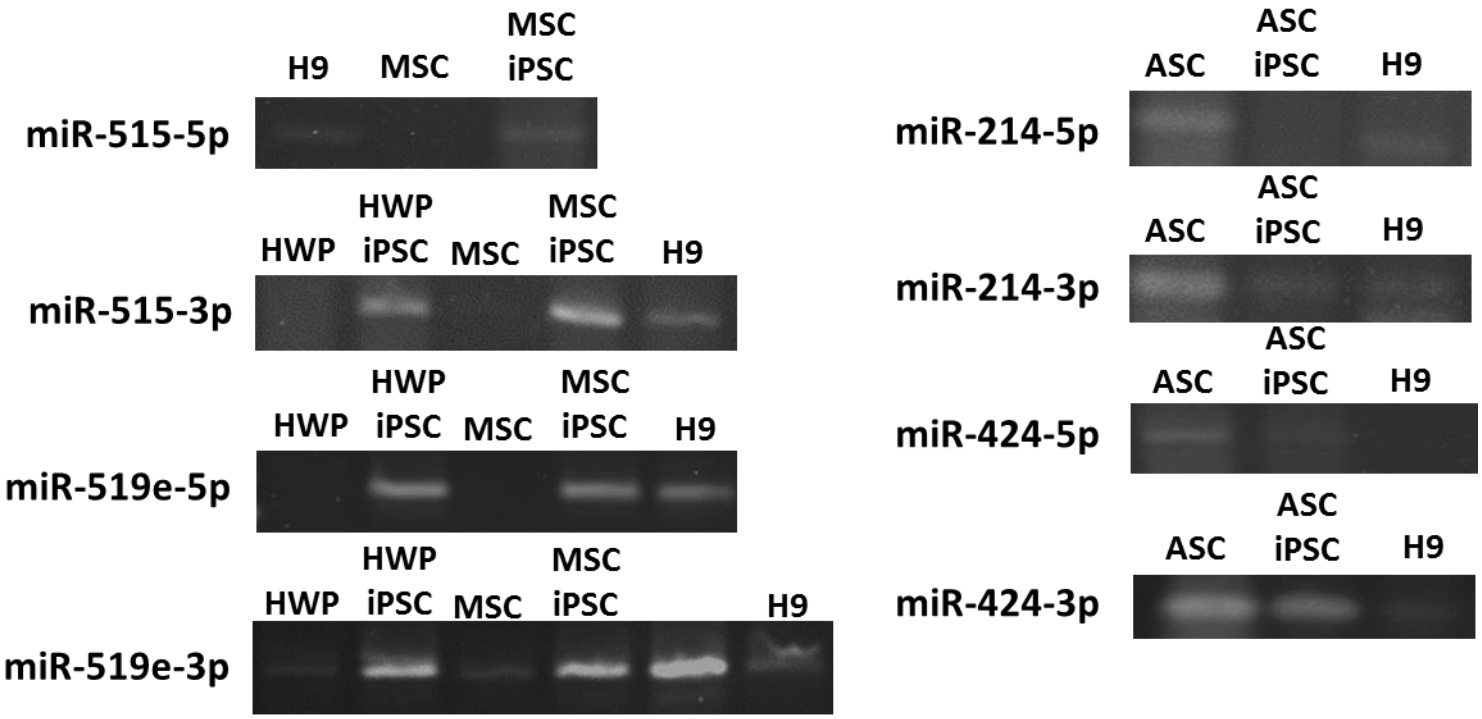
Enlarge Cluster II of Figure 4.2 from the text

APPENDIX H (Cont'd)



Enlarge Cluster III of Figure 4.2 from the text

APPENDIX I



Extended length of RT-PCR agarose gels of Figure 4.3

(Tay et al., 2008)

(Lin et al., 2008)

(Kuo et al., 2012)

(Viswanathan et al., 2008)

(Suzuki et al., 2009)