The miR-200 Family is Controlled by Epigenetic-based Mechanisms and Mediates Transition between Non-stem and Stem-like Cell Phenotypes

Yat Yuen LIM (B.Sc)

Thesis submitted for the degree of

Doctor of Philosophy

The School of Medicine
Faculty of Health Sciences
The University of Adelaide
affiliated with
Cytokine Research Laboratory
Centre for Cancer Biology
IMVS, SA Pathology

October 2012

Contents

CHAPTER 1 – INTRODUCTION

CHAPTER 2 - MATERIALS and METHODS

CHAPTER 3 - Characterization of two cell lines - spontaneously converted stable normal sl-HMLE cells and SUM159PT breast cancer cell line as EMT/stem-like experimental models

CHAPTER 4 - The miR-200s are master regulators of epithelial plasticity between stem-like/EMT and non-stem/MET phenotypes

CHAPTER 5 - Identification of miR-200 targets in breast stem-like cells and bCSCs

CHAPTER 7 – DISCUSSION and CONCLUSION

Abstract

MicroRNAs (miRNAs) are ~22 nucleotide (nt) single-stranded non-coding RNAs which are important regulators of gene expression in many biological processes including controlling cellular phenotype. The epithelial to mesenchymal transition (EMT) and the reverse process termed mesenchymal to epithelial transition (MET) are key programs that control the transition of cells between stem-like and non-stem phenotypes which are collectively termed epithelial plasticity.

The miR-200 family is a key regulator of EMT however its role in controlling the transition between stem-like and non-stem phenotypes has not been well characterized. I utilized immortalized human mammary epithelial cells (HMLE) to investigate the function and regulation of the miR-200s during their conversion from a non-stem to a stem-like phenotype. HMLE cells were found to spontaneously convert from a non-stem to a stem-like phenotype. Isolation and comparison of the miR-200 expression between the spontaneously derived stem-like cells (sl-HMLE) and non-stem HMLE cells (nsl-HMLE) showed that the spontaneous conversion to a stem-like phenotype was accompanied by the loss of miR-200 expression. Likewise, miR-200 expression was also found to be down-regulated in prospective breast cancer stem cells (bCSCs) from metastatic pleural or ascites effusions and SUM159PT breast cancer cell line compared to non-CSC cells. This phenotypic change from a non-stem to a stem-like phenotype was directly controlled by the miR-200s as restoration of its expression partially converted the sl-HMLE cells to a non-stem phenotype with decrease stem-like properties and induction of an MET-like phenotype, although restoration of the miR-200 expression in SUM159PT prospective bCSCs did not have this effect.

Next, using bioinformatic approaches and cell-based assays, I aimed to identify new miR-200 targets that are responsible for regulating the stem-like properties in both sl-HMLE cells and SUM159PT prospective bCSCs. Although the predicted genes (WNT5A, PKCα and PKCε) were not direct miR-200 targets, preliminary data suggest those genes may be involve in the survival or anoikis-resistance of stem-like cells and bCSCs.

Investigation of the mechanism(s) controlling miR-200 expression revealed both DNA methylation and histone modifications were significantly altered in the stem-like and non-stem phenotypes. In particular, in the stem-like phenotype, the miR-200b/a/429 cluster was silenced primarily through polycomb group-mediated silencing whereas the miR-200c/141 cluster was repressed by DNA methylation. Furthermore, slight increase in EZH2 expression was observed in the stem-like phenotype and this might potentially contribute to the

polycomb group-mediated silencing of the miR-200b/a/429 cluster. Lastly, preliminary co-immunoprecipitation results suggest that the targeting of polycomb group proteins to the miR-200b/a/429 promoter is not dependent on the ZEB1 transcription factor which is a repressor of the miR-200 transcription.

Collectively, these results indicate that the miR-200 family plays a critical role in the transition between stem-like and non-stem phenotypes and that distinct epigenetic-based mechanisms regulate each miR-200 gene in this process. Therefore, combination of chemotherapy with therapies targeted against the miR-200 family members and epigenetic modifications would be beneficial towards treatment of breast cancer.

Declaration

This work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution to Yat Yuen LIM and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text.

I give consent to this copy of my thesis, when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library catalogue, and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Yat Yuen LIM

Date:

Preface

Work presented in this thesis was generated through collaboration and assistance from the followings:

Dr Josephine Wright has contributed towards the collection and the study of miR-200 in the breast cancer patient samples. Dr Wright also designed the retroviral vectors used to stably over-express the miR-200s in sl-HMLE and SUM159PT cells. Fig. 3.5D and Fig. 4.6D were provided by Dr Wright.

DNA methylation assays for Fig. 6.1B-D were performed by Dr Eric Smith.

Dr Joanne Attema led the investigation into the potential regulation of the miR-200s by histone modifications in EMT. I joined her 2 to 3 months after this study had started to study the epigenetic regulation of miR-200 in stem-like cells. Therefore a substantial part of the antibody testing and optimization had been done by Dr Attema and her assistant Andrew Bert. ChIP assay for Fig. 6.6C was performed by Dr Attema while ChIP assays for Fig. 6.6A and Fig. 6.7A-B were performed by Andrew Bert. Real-time PCR for Fig. 6.11C was performed by Andrew Bert.

Acknowledgements

It is not possible to complete this PhD without the scientific mentorship, technical assistance, support and encouragement from many people. First and foremost, I would like to thank my primary supervisor, Associate Professor Gregory Goodall for initially accepting me as a student from Malaysia, providing all the necessary assistance I needed to come over to begin my PhD and more importantly, for your continual scientific insights and support in ideas and project strategies undertaken. All your efforts have moulded me to think critically and write proficiently. Furthermore, I am grateful for your help in making sure I can complete my PhD on time.

Thanks to my co-supervisor Dr Philip Gregory for your contribution during the establishment of this project, for teaching me the fundamental mRNA and miRNA techniques and for your continual support. I really enjoy having scientific discussion with you. I would also like to thank you for such a warm welcoming during my first day here in the Goodall laboratory. Your kindness has made me feel very comfortable and welcome in the Goodall Lab.

I would like to thank my second co-supervisor Dr Yeesim Khew-Goodall for all you have taught me on cell-based and microscopic work. Your expertise and lessons on cell biology especially on the complexity of cell membrane have corrected my naive perception of a cell and definitely broaden my knowledge on cell biology.

I am grateful to Dr Josephine Wright not just for your scientific mentorship but also for your enthusiastic support and encouragement which help build the confidence in me. I am inspired by your cheerfulness and positive attitude. Thank you for adapting to my strengths and weaknesses.

I am also grateful to Dr Joanne Attema for allowing me to join you to work on the epigenetic studies. The short 6 months working together had been a steep learning curve but thanks to you, I have equipped myself now with many valuable molecular techniques in the area of epigenetics. Thanks for sharing your ideas, insights and discussions as well as your support and encouragement.

I would also like to thank all the members of Goodall laboratory. Thanks to Andrew Bert, for teaching and helping me with all the real-time PCRs and the Taqman low density miRNA array. Thank you to Dr Cameron Bracken for your ingenious insights and discussions. Thanks to my fellow PhD student, Natasha Kolesnikoff for your kind assistance, for your technical support, for teaching me cloning and the luciferase gene reporter assay as well as being a

good friend. Thanks to Suraya Roslan for your assistance when I needed them, for the hilarious Friday pub drinks and for being a great friend. Thanks to Dr Emily Paterson, Matthew Anderson, Narrelle Mancini and Rosemary Sladic for all your help.

Thanks to the members from the Yeesim Khew-Goodall laboratory for all your help, insights and discussion during meetings.

Thanks to Dr Daniel Thomas for initially helping to obtain the human ethics approval for patient sample collection and providing valuable advice and help on patient sample processing. Your research outlook which is patient-driven has really inspired me to think out of the box and how future basic researchers can work together with clinicians to translate basic research to the bed side.

A special thanks to Dr Eric Smith who taught and helped me with all the DNA methylation assays.

I would also like to acknowledge the staff from the Detmold family trust cell imaging centre for their tireless effort and time to teach and help me in the area of flow cytometry.

Last but not least, a personal thank you to my family for their endless support, understanding and patience. Thanks to all my friends here in Australia and Malaysia for their continual encouragement.

List of Figures

List of Tables

Abbreviations

°C	degrees Celsius
µg	Microgram
µl	Microlitre
µM	Micromolar
ABCC5	ATP-binding cassette, sub-family C (CFTR/MRP), member 5
ABCG2	The G (white) subfamily of ABC transporters
ACD	Acid Citrate Dextrose
ADP	Adenosine diphosphate
AE1/AE3	Pan-cytokeratin antibody
Ago	Argonaute
AKT	v-akt murine thymoma viral oncogene homolog
ALDH1	Aldehyde dehydrogenase 1
AML	Acute myeloid leukemia
ANRIL	Antisense noncoding RNA in the INK4 locus
APE1	Apurinic/apyrimidinic endonuclease 1
ARHGAP19	Rho GTPase activating protein 19
ASEL	Test receptor neuron cell fates ASE left
ASER	Test receptor neuron cell fates ASE right
ATM	Ataxia telangiectasia mutated
ATP	Adenosine triphosphate
BCA	Bicinchoninic acid
BcIX	B-cell lymphoma-extra
bCSCs	Breast cancer stem cells
bFGF	Basic fibroblast growth factor; FGF2
BLAST	Basic Local Alignment Search Tool
BMI1	B lymphoma Mo-MLV insertion region 1 homolog
BMP	Bone morphogenetic protein
bp	Base pair
BPE	Bovine pituitary extract
BRCA	Breast cancer susceptibility gene
BRCP	Breast cancer resistance protein
BRG1	Brahma/SWI2-related gene 1
BSA	Bovine serum albumin
CALLA	Common acute lymphoblastic leukaemia antigen, CD10
CBX	Chromobox
CCNE2	Cyclin E2
CDH1	E-cadherin
CDKN1C/p57	Cyclin-dependent kinase inhibitor 1C
cDNA	Complementary DNA
CEA	Carcinoembryonic antigen
CFL2	Cofilin 2
CHEK2	Checkpoint kinase 2
ChIP	Chromatin immunoprecipitation
CK	Cytokeratin
CK1	Casein kinase 1
cm	Centimeter
CMV	Cytomegalovirus
c-MYC	Cellular myelocytomatosis viral oncogene homolog
CO_2	Carbon dioxide
Co-IP	Co-Immunoprecipitation

CoREST	Corepressor for element-1-silencing transcription factor
Cq	Quantification cycle
CREB	cAMP responsive element binding protein
CSCs	Cancer stem cells
Ct	Cycle threshold
CtBP	C-terminal binding protein
CTNNB1	β-catenin
CXCL7	C-X-C motif chemokine 7
CYLD	Cylindromatosis (turban tumor syndrome)
DAPI	4´-6-Diamidino-2-phenylindole
dATP	Deoxyadenosine triphosphate
DCIS	Ductal carcinoma *in situ*
DCV	DyeCycle Violet
DGCR8	DiGeorge critical region 8
dGTP	Deoxyguanosine triphosphate
DKK-1	Dickkopf-1
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNMTs	DNA methyltransferases
DRFS	Distant relapse-free survival
dTTP	Deoxythymidine Triphosphate
dUTP	Deoxyuridine Triphosphate
EC50	Effective concentration of a drug that gives half-maximal response
ECM	Extracellular matrix
EDTA	Ethylenediaminetetraacetic acid
EED	Embryonic ectoderm development
EGF	Epidermal growth factor
EMT	Epithelial-to-mesenchymal transition
ENCODE	Encyclopedia Of DNA Elements
EpCAM	Epithelial cell adhesion molecule; ESA
ER	Estrogen receptor
ERRFI-1/MIG6	ERBB receptor feedback inhibitor 1
ES	Embryonic stem
ESA	Epithelial specific antigen
ETS1	v-ets erythroblastosis virus E26 oncogene homolog 1
EV	Empty vector
Exp5	Exportin-5
EZH1	Enhancer of zeste homolog 1
EZH2	Enhancer of zeste homolog 2
FACS	Fluorescence-activated cell sorting
FAP1	FAS-associated phosphatase 1
FasL	Fas ligand
FCS	Fetal calf serum
FGF	Fibroblast growth factor
FGFR	Fibroblast growth factor receptor
FHOD1	Formin homology domain containing protein 1
FN1	Fibronectin 1
FOG2	Friend of GATA 2
FOXC1	Forkhead box C1
FOXO	Forkhead box protein O

GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GATA3	GATA binding protein 3
GCDFP-15	Gross cystic disease fluid protein-15
GFP	Green fluorescent protein
GLI	Glioma-associated oncogene family zinc finger
GSC	Goosecoid
h	Hour
H2AK119ub	Mono-ubiquitylation of lysine 119 at histone H2A
H3	Histone 3
H3K9	Histone 3 at lysine 9
HA	Hyaluronan
HAS2	Hyaluronan synthase 2
HAT	Histone acetyltransferase
HBSS	Hank's Buffered Salt Solution
HBV	Hepatitis B virus
HCl	Hydrochloric acid
HDAC	Histone deacetylase
HDM	Histone demethyltransferase
HER2	Human Epidermal Growth Factor Receptor 2; *neu*; Erbb2
Hh	Hedgehog
HIF-1α	Hypoxia inducible factor 1, alpha subunit
HMECs	Primary human mammary epithelial cells
HMF	Human mammary fibroblasts
HMGA2	High-mobility group AT-hook 2
HMLE	Immortalized human mammary epithelial
HMLE+TGFβ1	HMLE cells treated with TGFβ1
HMLEN	HMLE cells transformed with an activated form of the HER2/*neu* oncogene
HMLER	HMLE cells transformed with a V12H-Ras oncogene
HMT	Histone methyltransferase
HOTAIR	HOX antisense intergenic RNA
HOXA10	Homeobox A10
H-RAS	v-Ha-ras Harvey rat sarcoma viral oncogene homolog
HRP	Horseradish peroxidase
HSA	Human serum albumin
HSF1	Heat shock transcription factor 1
hTERT	Human telomerase reverse transcriptase
HTwist	HMLE cells stably expressing TWIST
IAPs	Inhibitors of apoptosis proteins
IC50	Inhibition concentration of an inhibitor that gives half-maximal response
IDC	Infiltrating/invasive ductal carcinoma
IgG	Immunoglobulin G
IL-6	Interleukin-6
IL-8	Interleukin-8
IL8RA	Interleukin-8 receptor, alpha; CXCR1
ILC	Infiltrating/invasive lobular carcinoma
iPSC	Induced-pluripotent stem cells
IR	Infra-red
IRES	Internal ribosome entry site
ITGA6	Integrin, alpha 6; CD49f

ITGB3	Integrin, beta 3; CD61
JAG1	Jagged 1
JAK2	Janus kinase 2
JARID2	Jumonji/ARID domain-containing protein 2
JNK	c-Jun N-terminal kinase
JUN	Jun proto-oncogene
K	Lysine
k	Thousand
kb	Kilobase
KCNQ1OT1	KCNQ1 overlapping transcript 1
KEAP1	Kelch-like ECH-associated protein 1
KLF4	Kruppel-like factor 4
LCIS	Lobular carcinoma *in situ*
LEPR	Leptin receptor
Lin	Lineage
LRP1	Low-density lipoprotein receptor-related protein 1
LSD1	Lysine-specific demethylase 1
LT	Large T
M	Molar
MAML	Mastermind-like
MAP2K4	Mitogen-activated protein kinase kinase 4
MAPK/ERK	Mitogen-activated protein kinase/extracellular signal-regulated kinase
MAPK14	Mitogen-activated protein kinase 14
MARCKS	Myristoylated alanine-rich protein kinase C substrate
MaSCs	Mammary stem cells
MBDs	Methyl-CpG binding domain proteins
MCP-1	Monocyte chemotactic protein 1
mCSCs	Migrating cancer stem cells
MDCK	Madin Darby canine kidney
MDCK+TGFβ1	MDCK cells treated with TGFβ1
MDCK-Pez	MDCK stably expressing the tyrosine phosphatase Pez
MDR1	Multidrug resistance 1 gene
MeCP2	Methylcytosine binding protein 2
MEGM	Mammary Epithelial Cell Growth Medium
MET	Mesenchymal to epithelial transition
MGG	May Grumwald Giemsa
Mi2/NuRD	Mi2/nucleosome remodeling and deacetylase
min	Minute
miR	MicroRNA
miRISC	MiRNA-containing RNA-induced silencing complex
miRNA*	Passenger strand of miRNA
miRNAs	MicroRNAs
miRNP	MicroRNA ribonucleoprotein complex
mM	Millimolar
ml	Millilitre
mg	Milligram
M-ref	Methylation reference
MMTV	Mouse mammary tumour virus
mRNA	messenger RNA
MRUs	Mammary repopulating units
MSCV	Murine Stem Cell Virus

MSN	Moesin
MSP	Stable spontaneously arising stem-like/mesenchymal HMLE subpopulation
MTS	(3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium)
MUC1	Mucin 1
MYOD	Myoblast determination protein 1; myogenic differentiation 1
MYT1	Myelin transcription factor 1
N.D	Not determine
N.V	Not validated by luciferase gene reporter assay
n/a	Not available
Na_3VO_4	Sodium vanadate
NaCl	Sodium chloride
NaF	Sodium fluoride
NANOG	Nanog homeobox
ncRNA	Non-coding RNA
nCSCs	Non-CSCs
neg	Negative
NF-kB	Nuclear factor kappa-light-chain-enhancer of activated B cells
ng	Nanogram
nM	Nanomolar
NOD/SCID	Nonobese diabetic/severe combined immune deficiency
nsl-HMLE	$CD44^{low}/CD24^{hi}$/epithelial-like cells
nt	Nucleotide
NTRK2	Neurotrophic tyrosine receptor kinase type 2
OD	Optical density
p21/WAF1/Cip1	CDK-interacting protein 1
p53	Tumour protein 53
p63	Tumour protein 63
p73	Tumour protein 73
p65/RELA	V-rel reticuloendotheliosis viral oncogene homolog A
PAX5	Paired box gene 5
PBS	Phosphate-buffered saline
PcG	Polycomb group
PDCD4	Program cell death 4
PFA	Paraformaldehyde
PGK	Phosphoglycerate kinase
PI3K	Phosphatidylinositol 3-kinase
PIC	Protease inhibitor cocktail
PKC	Protein kinase C
PLC-γ1	Phospholipase C gamma 1
PMS	Phenazine methosulfate
Pol	Polymerase
pos	Positive
PPARA	Peroxisome proliferator-activated receptor alpha
PPM1F	Mg2+/Mn2+ dependent protein phosphatase 1F
PR	Progesterone receptor
pRB	Retinoblastoma protein
PRC	Polycomb repressor complex
pre-miRNA	Precursor of miRNA
pri-miRNA	Primary miRNA transcripts

PRMT5	Protein Arginine Methyltransferase 5
PROCR	Protein C receptor, endothelial
PTEN	Phosphatase and tensin homolog
qPCR	Quantitative PCR
qRT-PCRs	Quantitative real-time PCR
RAD51	Rad51 homolog protein
RAH	Royal Adelaide Hospital
RANK	Receptor activator of NF-KB
RANKL	Receptor activator of nuclear factor kappa-B ligand
RBAP46/48	Retinoblastoma protein associated protein 46/48
RECK	Reversion-inducing-cysteine-rich protein with kazal motifs
REDD1	also known as DDIT4, DNA-damage-inducible transcript 4
RepA	Repeat A
RIN	RNA integrity number
RING1B	RING finger protein 1B
RIPA	Radioimmunoprecipitation assay
RL	Renilla luciferase
RMA	Robust multichip averaging
RNA	Ribonucleic acid
RNAi	RNA interference
RNAse	Ribonuclease
RND3	Rho family GTPase 3
RNU6B	RNA, U6B small nuclear
ROR	Receptor tyrosine kinase-like orphan receptor
ROS	Reactive oxygen species
rpm	Revolutions per minute
RPMI	Roswell Park Memorial Institute medium
RT	Reverse transcriptase
s	Second
SAGE	Serial Analysis of Gene Expression
Sca-1	Stem cell antigen 1
SD	Standard deviation
SDS	Sodium dodecyl sulfate
SEC23A	Sec23 homolog A
SEM	Standard error of the mean
SFE	Sphere-forming efficiency
SFRP-1	Secreted frizzled-related protein 1
SIP1	Smad-Interacting Protein 1
siRNA	Short interfering RNA
SIRT1	Sirtuin 1
sl-HMLE	*De novo* spontaneously derived $CD44^{hi}/CD24^{low}$/mesenchymal-like HMLE
SLUG	also known as SNAI2
SMAD3	SMAD family member 3
SNAI1	Snail homolog 1
snRNA	Small nucleolar RNA
SOX2	SRY-related HMG-box 2
SP	Side population
SRC	V-src sarcoma (Schmidt-Ruppin A-2) viral oncogene homolog
STAT3	Signal transducer and activator of transcription 3
STK11	Serine/threonine kinase 11

SUV39H	Suppressor of variegation 3-9 homolog 1
SUZ12	Suppressor of zeste 12
TAF	Tumour associated fibroblast
TAK1	TGFβ1–activated kinase 1
TAM	Tumour associated macrophages
TBST	Tris Buffered Saline with Tween 20
TBX3	T-box 3
TCF3	Transcription factor 3
TDLUs	Terminal ductal lobular units
TDT	Terminal deoxynucleotidyl transferase
TE	Tris EDTA
TEBs	Terminal end buds
TGFβ	Transforming growth factor beta
TGFβR	Transforming growth factor beta receptor
Thy-1	Thy-1 cell surface antigen
TICs	Tumour-initiating cells
TNBC	Triple negative breast cancer
TNF-α	Tumor Necrosis Factor alpha
TNT	Tris-NaCl-Tween buffer
TRAIL	TNF-related apoptosis-inducing ligand
tRNAs	Transfer RNAs
TSSs	Transcription start sites
TTF-1	Thyroid transcription factor 1
TUBB3	Class III β-tubulin
U	Units
UBC9	Ubiquitin carrier protein 9
UCSC GBD	University of California, Santa Cruz Genome Browser Database
UDG	DNA glycosylase
uPA	Urokinase type plasminogen activator
UTR	Untranslated region
VEGFR	Vascular endothelial growth factor receptor
VIM	Vimentin
vol	Volume
WAVE3	WAS protein family member 3
Wnt	Wingless-type MMTV integration site family
Xist	X-inactive specific transcript
YB1	Y box binding protein 1
YY1	Ying-Yang 1
ZEB	Zinc finger E-box binding homeobox
ZO-1	Zona occludens 1
α-SMA	Alpha smooth muscle actin
ΔNp63	p63 inhibitory isoform